Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 17 (2015), Issue 3   ISSN 1438-8871   Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Viewpoint

Original Papers

Reviews

Corrigenda and Addendas

Viewpoint

Systems Medicine 2.0: Potential Benefits of Combining Electronic Health Care Records With Systems Science Models

Taavi Tillmann[1,2], BSc, MBChB, MPH, MFPH; Alexander R Gibson[2], BSc, MBChB, DiMM; Gregory Scott[3], BEng, MBBS, MSc; Oliver Harrison[4,5], MA, MBBS, MPH; Anna Dominiczak[2], MD, MRCP, FRCP, FMedSci; Phil Hanlon[2], BSc, MD, MPH

[1]Department of Epidemiology & Public Health, University College London, London, United Kingdom
[2]College of Medical, Veterinary and Life Sciences, University of Glasgow, United Kingdom
[3]Division of Brain Sciences, Imperial College London, United Kingdom
[4]Healthways International, London, United Kingdom
[5]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Corresponding Author:
Taavi Tillmann, BSc, MBChB, MPH, MFPH
Department of Epidemiology & Public Health
University College London
1-19 Torrington Place
London, WC1E 7HB
United Kingdom
Phone: 44 20 7679 8252
Fax: 44 20 7813 0242
Email: taavi.tillmann@gmail.com

Abstract

Background: The global burden of disease is increasingly dominated by non-communicable diseases.These diseases are less amenable to curative and preventative interventions than communicable disease. This presents a challenge to medical practice and medical research, both of which are experiencing diminishing returns from increasing investment.

Objective: Our aim was to (1) review how medical knowledge is generated, and its limitations, (2) assess the potential for emerging technologies and ideas to improve medical research, and (3) suggest solutions and recommendations to increase medical research efficiency on non-communicable diseases.

Methods: We undertook an unsystematic review of peer-reviewed literature and technology websites.

Results: Our review generated the following conclusions and recommendations. (1) Medical knowledge continues to be generated in a reductionist paradigm. This oversimplifies our models of disease, rendering them ineffective to sufficiently understand the complex nature of non-communicable diseases. (2) Some of these failings may be overcome by adopting a “Systems Medicine” paradigm, where the human body is modeled as a complex adaptive system. That is, a system with multiple components and levels interacting in complex ways, wherein disease emerges from slow changes to the system set-up. Pursuing systems medicine research will require larger datasets. (3) Increased data sharing between researchers, patients, and clinicians could provide this unmet need for data. The recent emergence of electronic health care records (EHR) could potentially facilitate this in real-time and at a global level. (4) Efforts should continue to aggregate anonymous EHR data into large interoperable data silos and release this to researchers. However, international collaboration, data linkage, and obtaining additional information from patients will remain challenging. (5) Efforts should also continue towards “Medicine 2.0”. Patients should be given access to their personal EHR data. Subsequently, online communities can give researchers the opportunity to ask patients for direct access to the patient’s EHR data and request additional study-specific information. However, selection bias towards patients who use Web 2.0 technology may be difficult to overcome.

Conclusions: Systems medicine, when combined with large-scale data sharing, has the potential to raise our understanding of non-communicable diseases, foster personalized medicine, and make substantial progress towards halting, curing, and preventing non-communicable diseases. Large-scale data amalgamation remains a core challenge and needs to be supported. A synthesis of “Medicine 2.0” and “Systems Science” concepts into “Systems Medicine 2.0” could take decades to materialize but holds much promise.

(J Med Internet Res 2015;17(3):e64) doi:10.2196/jmir.3082

KEYWORDS

gene-environment interaction; systems theory; electronic health records; epidemiology; online social networks; crowd-sourcing; Web 2.0

Current Limitations in the Study and Management of Chronic Disease

Medical science has brought clear and dramatic improvements to health over the past 150 years. This has included effective cures and preventions as well as important public health measures. As a consequence of many of these advancements, the leading causes of death and disability have shifted from infectious diseases to more complex non-communicable diseases. In the past 50 years, expenditure on health research and health care has increased dramatically to meet the new challenges that treating and preventing these multifaceted diseases presents. However, combating chronic disease has proven significantly more difficult and costly than infectious disease, with it becoming increasingly difficult to continue raising life expectancy and healthy life expectancy [1]. Many spheres of academia and clinical practice have shifted away from an intent to cure, to an attempt to slow down pathological processes and prevent complications. This is in part due to the inability of science to determine which individuals will suffer the most from any given risk factor, or who will benefit most from specific interventions.

Current medical science is largely conducted using the reductionist paradigm, which assumes that complex entities are best understood by breaking them down into smaller, simpler components. Detailed analysis of the weaknesses of this assumption has been done elsewhere [2], but principally, reductionism limits our ability to understand how multiple variables interact with one another to create emergent effects.

The reductionist approach does offer a useful first step for the understanding of a complex system because it helps identify key components. However, a strong and enduring emphasis on the reductionist approach risks over-simplification (focusing only on a handful of major factors with the biggest effect, while the sum of minor factors may be considerable) and generalization (assuming that a common cause-effect relationship applies equally in all cases). Moreover, an excessive focus on a limited number of pathways may impede our ability to understand both the behavior of the system as a whole, as well as system variability between individuals. Our etiological models of seemingly disparate chronic diseases include a striking number of common pathways. For example, alterations of the tumor necrosis factor alpha (TNF-alpha) gene have been implicated in 88 clinically distinct diseases [3]. However, it remains unclear under what circumstances increased TNF-alpha levels cause an individual to develop rheumatoid arthritis, atherosclerosis, or a septic cytokine cascade. It is assumed that many such pathways interact to produce disease outcomes. Rarely can we adequately describe why a disease develops in an individual, their prognosis, the effect of risk factor modification on an individual, or the likelihood of family being affected by a similar disease.

The type of drugs dispensed to patients further reflects our limited understanding of the true cause of disease. While some drugs can reverse the original disease process to provide a near cure (eg, antibiotics, chemotherapy), many more provide only temporary relief to the current physiological imbalance and fail to cure (eg, thyroxin, antidepressants, anticonvulsants, steroids, diuretics). Others still simply dampen the body’s capacities to exacerbate the condition in a symptomatic or palliative [4] fashion (eg, beta blockers, warfarin) with significant side effects. Another weakness in our management of chronic diseases stems from the way research studies and clinical guidelines consider each disease in isolation, while in reality most patients have at least one comorbidity. This leaves patients with an expanding burden of treatment, polypharmacy, increased side effects, unintended drug interactions, and reduced adherence.

This pattern of diminishing returns necessitates a reevaluation of the approach that medical science has taken in the study and management of chronic diseases. As chronic diseases are substantially more complex than infectious diseases, additional approaches are probably required to overcome the limitations of reductionism. In this essay, we aim to (1) document medical science’s first steps in moving away from reductionism towards more complex models, (2) assess the potential benefit of introducing Systems Science into medical science and evaluate the relative strengths and weaknesses of various technologies in facilitating this (such as Medicine 2.0), and (3) present our suggestions of how to best increase medical research efficiency. To achieve this, we combine an unsystematic review of peer-reviewed literature, with a review of two websites known for the dissemination of novel technologies and ideas (Wired, TED). We then try to bring together these disparate lines of thought, across a range of disciplines and industries, into a subjective but hopefully thought provoking synthesis.

Supplementing the Reductionist Paradigm

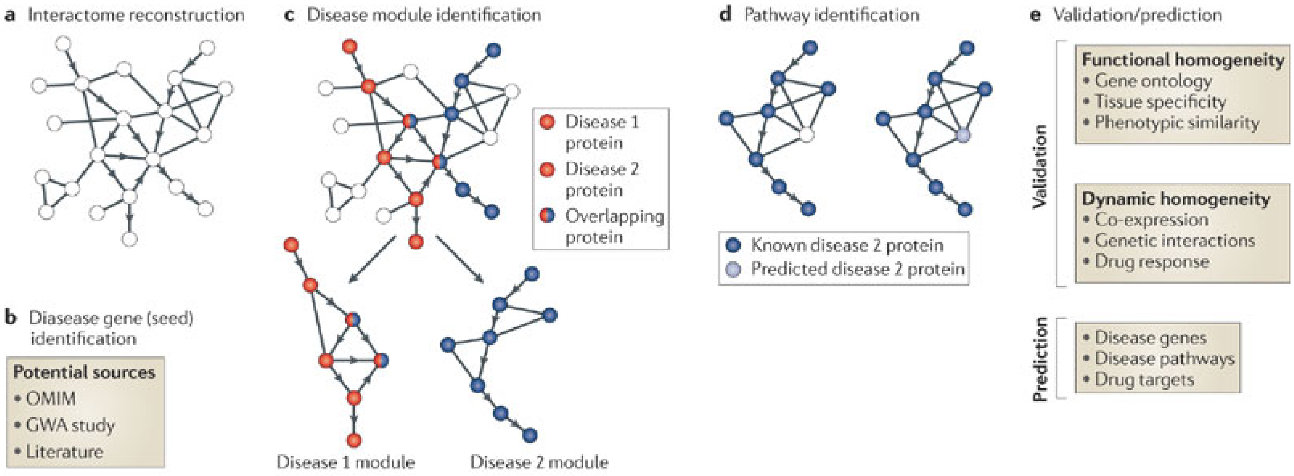
Should the prevalent theoretical model, framework, or paradigm find it increasingly difficult to account for experimental data, alternatives to the prevalent model should be considered [5]. This can be done by replacing the old model or by adding supplements to the existing model [6]. We argue that three such supplements have recently been added to the core of reductionism. First, select chemotherapy agents were found to work particularly well for subtypes of breast cancers and leukemia, which corresponded to genetic subtypes of the disease. This reveals that diseases may look similar on the outside but can function very differently on the inside. As our knowledge of this etiome [3] (the precise etiological pathways by which genetic and environmental agents cause a disease) grows, it becomes useful to name each diagnostic subcategory with appropriate subdivisions. The term “intermediate pathophenotype” [7] has been suggested to capture these subdivisions. Linguistically, this denotes a focus on end-state pathology, the likely object of interest for pathologists and systems biologists. Clinicians, epidemiologists, and biomedical scientists are, however, more interested in upstream etiology and the narration of an individual’s past and likely future. Perhaps adopting the term “etiphenotype” instead would make Systems Medicine more accessible to clinicians (or alternatively, letting the term “etignosis” complement “diagnosis”).

The second supplement to reductionism came from the observation that most patients have more than one disease at any given time and that some diseases tend to cluster together. For example, people with diabetes have a greater risk of developing certain other diseases, such as stroke. Recent evidence has identified numerous early commonalities, such as “the metabolic syndrome” or “diabesity”, but our scientific language and modes of thought struggle to describe the interconnectedness of these tightly intertwined etiological processes [8]. Reductionism is ill-suited to deal with the question of comorbidity, prompting a search for alternative models [9]. Third, in many cases it appears better to treat complex diseases with a cocktail of drugs administered simultaneously. This has been applied in the treatment of infections (eg, tuberculosis and human immunodeficiency virus), as well as for chronic conditions (eg, acute coronary syndrome, exacerbations of chronic obstructive pulmonary disease) [10]. Advances in such complex, combined interventions have prompted the Medical Research Council to issue guidance on how they should be monitored and evaluated [11]. We believe that this will give only modest improvements, as intervention development will still remain somewhat blind and ignorant of the intricacies of etiological processes.

Modifying the Reductionist Paradigm With Network Medicine

A transitionary model of medical research is emerging that attempts to reconcile some of the discrepancies outlined above. There is much hope that network medicine could be better suited to understand the basic biological processes that culminate in health and disease [12,13]. This framework suggests that complex diseases are the emergent result of perturbations in multiple genes that are interconnected to one another to create a disease module (Figure 1). A hypothetical study of diabetes could begin by mapping the system of proteins that are responsible for healthy function like glucose regulation (Figure 1, panel a), also known as the interactome of proteins. Next, existing literature is used to identify a candidate gene or protein critical in etiology (Figure 1, panel b). This protein is mapped onto the interactome to identify a smaller set of proteins that directly influence the candidate protein (Figure 1, panel c), to identify a suspected diabetes disease module (colored red). Concurrent mapping of comorbid diseases (such as stroke in blue) can identify structural reasons for comorbidity as well as common premorbid states (such as the metabolic syndrome, marked by the red-blue proteins). Models like this can account for the effectiveness of combined interventions (Figure 1, panel d, dark blue proteins) and identify novel drug targets (Figure 1, panel d, light blue). Such models have been used to identify novel genes across a range of cancers [14-16]. It has also allowed drugs to be re-positioned for other diseases, such as using the anti-ulcer drug cimetidine in the treatment of lung cancer [17].

Figure 1. The methodological steps by which network medicine identifies disease modules and predicts novel protein targets for intervention.



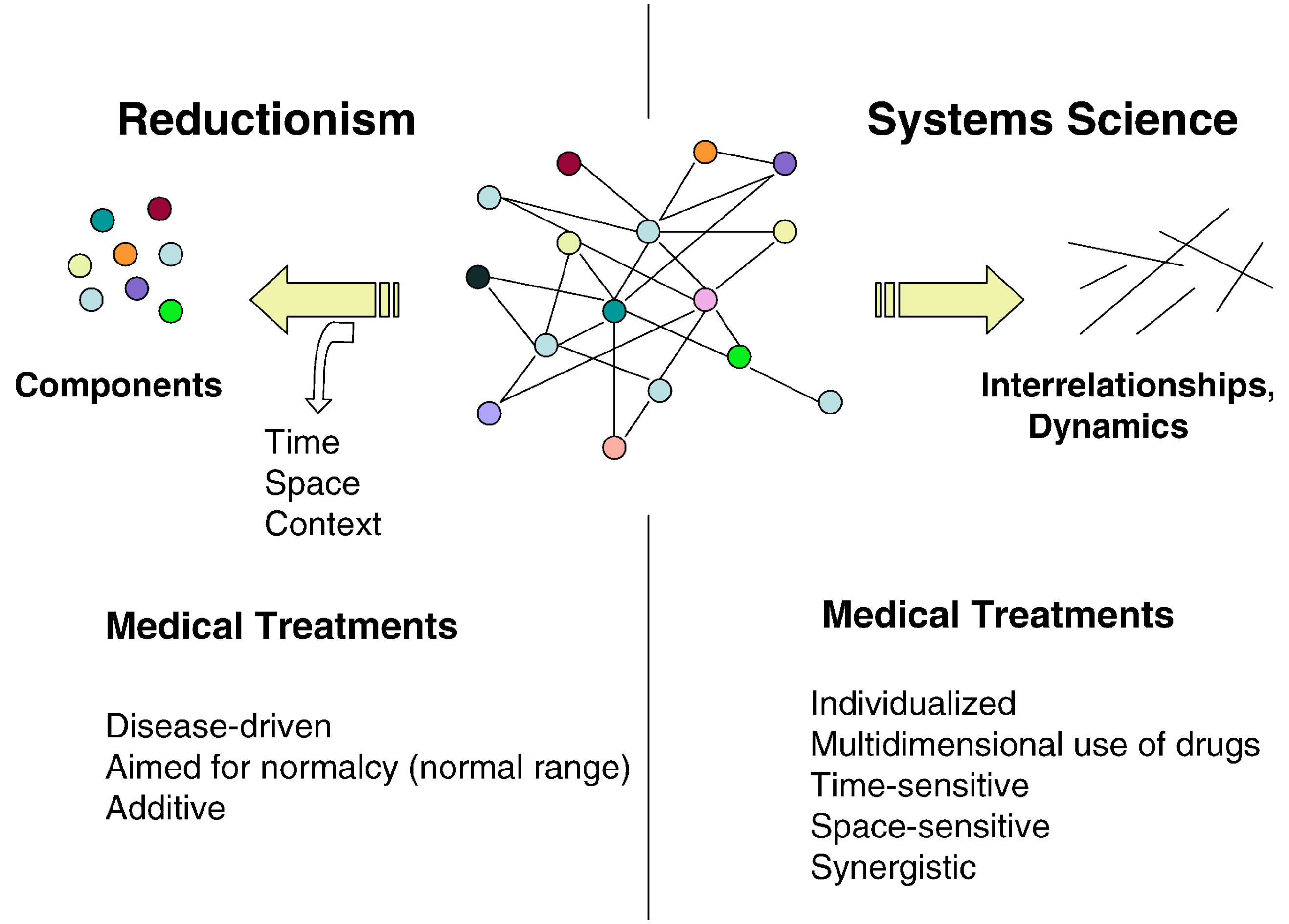
Nature Reviews | Genetics

We believe that the network medicine approach can achieve half the paradigm shift required toward systems medicine. However, such models remain incomplete because of the following weaknesses [18]. (1) When considering disease, network medicine looks only at snapshots of end-stage illness. The model cannot be applied to longitudinal data to describe the gradual shift that takes place when healthy states slowly transform into diseased states. This weakness stems from the model’s implicit assumption that disease processes (such as glucose dysregulation) are merely broken flipsides of healthy processes (glucose regulation). The cascade of events that caused the shift from healthy state to diseased state are irrelevant and not investigated [19]. (2) Network medicine focuses exclusively on the intracellular level of proteins and genes. Researchers interested in higher order risk factors (such as how weight gain is influenced by childhood nutrition, family upbringing, socioeconomic status, or the physical environment of green spaces and fast food outlets) struggle to utilize network medicine models [20]. (3) Simplistic causal modeling, wherein diseased genes and proteins are thought to exert their effects uniformly across a range of interindividual and interenvironmental variation. Thus network medicine struggles to model how genetic risk interacts with environmental risk to create disease. We feel that these three weaknesses will keep our models of disease largely incomplete, thus perpetuating many of the shortcomings outlined in our introduction.

Replacing the Reductionist Paradigm With Systems Medicine

It helps to begin by clarifying our core terms. *A System* can be defined as a set of components that are related to one another in a meaningful way (centerpiece, Figure 2). *Systems Science* is the formal study of systems [21]. *Systems Thinking* describes modes of thought that focus on the connectedness and interrelationships of components, rather than focusing on the components themselves. *Systems Theory* is a set of theories that try to derive generalizable organizing principles that apply to all systems. For example, complex systems can maintain robust performance during external perturbations, thanks to system-wide properties such as *modularity* and *built-in redundancy* (ie, having two kidneys), features that are seen in biological and non-biological systems alike [19]. *Systems Biology* is the study of how biological functions emerge from the interactions between the components of living systems and how these emergent properties in turn influence the behavior of lower-level components [22]. The field developed quickly during the past decade, due to technological advances in generating large amounts of high-throughput data very quickly, along with the interdisciplinary ability to compute, model, and make sense of this data.

Figure 2. Schematic illustration of the core differences between reductionism and systems science, when analyzing the properties of a system.



To date, little of the above paradigms have been incorporated into medical research. In part, this may be due to difficulties in amalgamating large-scale clinical datasets. However, as such technical barriers begin to fall away, increasing attribution can be placed on lack of knowledge about Systems Science among the medical research community. Indeed, those looking for a working definition of Systems Medicine may be baffled by the diversity of opinions out there [19,23,24]. Historically, Systems Medicine has been defined as the clinical application of Systems Biology approaches to medicine, where traditional model-driven experiments are informed by data-driven models in an iterative manner [25]. We see Systems Medicine as the long-term objective of a wider paradigm shift in medical science, at the end of which a range of different models and approaches will coexist under the Systems Medicine umbrella. All of these models will be substantially more complex than the models used in Reductionism or Network Medicine. We suggest that Systems Medicine models should include two or more of the following organizing principles of the human body: non-linearity, multi-agency, multi-levelness, or adaptivity [26].

Non-linearity means that the independent variables interact with one another and modify each other's effects on the dependent variable. This makes the dependent variable exhibit emergent properties that can be understood only when *all* of the independent variables are assessed concurrently. Such interactions have been documented between behavioral risk factors, between two genes [27], between transcribed mRNA and regulatory microRNA [28], between single nucleotide polymorphisms (SNP) and expression quantitative trait loci [29], as well as between transcription factors [30]. Perhaps most important of all are gene-environment interactions, which have begun to emerge for obesity [31], coronary artery disease [32], asthma [33], colorectal cancer [34], depression [35], eczema [36], Alzheimer's disease [37], and multiple sclerosis [38] to name a few. Models of systems medicine need to be built and empirical data assembled so that the detection of interactive effects and non-linear dynamics is facilitated. Contrast this with the stance taken in most introductory epidemiology and medical statistics courses, where interactive effects are seen as "nuisance" phenomena and students are discouraged against opening such cans of worms because of seemingly unmanageable type 1 errors. Regrettably, the study of interaction phenomena remains a niche field. This stifles progress at understanding any non-linear mechanisms of the human body.

Non-linearity can be present in relatively simple systems [39]. Consider the example of a system where smoking and drinking alcohol leads to increased stroke risk, through the upregulation of a hypothetical inflammatory cytokine called "smokdrink". These two risk factors can interact with one another and produce non-linear responses that are quite complex in nature (such as a dramatic increase in smokdrink, but only if you drink more than 3 units a day *and* have accumulated more than 20 years of smoking damage). Nonetheless, our model remains simply non-linear (Multimedia Appendix 1, frame 1). Next, we can expand the model to account for how stroke is further influenced by another non-linear system, namely the cholesterol system (frame 2). If we view smokdrink and cholesterol as our two agents of stroke, then we have built a multi-agent model (frame 3). This concept is useful to bear in mind as the outcome of complex systems is rarely determined by one agent, but rather is the interaction between multiple agents (Multimedia Appendix 1, red arrows at the top of frame 3). Many of these agents remain directly unmeasurable, but their parameters can be estimated. For example, even before we discover the inflammatory cytokine smokdrink, we can calculate that 20 cigarettes + 3 pints a day has a similar effect on stroke, as does 10 cigarettes + 5 pints a day. Constructing intermediary agents *in silico* (such as smokdrink and cholesterol) facilitates our conceptual and mathematical understanding of the nature of dynamic disease processes. Such constructs could also include the inflammatome [40], the Metabolic Syndrome scale, the Hypothalamus-Pituitary-Axis dysregulation scale, or the allostatic load scale [41].

Multi-level models account for how people often form natural groups, which in turn influence individual behavior and disease outcomes. A grouping variable, such as "living in a deprived neighborhood or not", can exert its effect on the disease outcome directly just like any other risk factor (Multimedia Appendix 1, frame 4, dotted blue line), or indirectly by modifying the effect of a lower order risk factor (such as drinking alcohol; Multimedia Appendix 1, red dotted lines). Although it may be tempting to enter "living in a deprived neighborhood" into the model alongside the other variables, this violates certain statistical assumptions. Accordingly, models known interchangeably as nested, hierarchical, or multiscale models, are useful to take account of the multi-levelness of such real-life phenomena.

Finally, all the models presented so far compare states of full health against states of full illness. The sequence of events over which full health deteriorates towards full illness remains unknown. To understand this process, we must build dynamic models of the human body. The core principle here is *adaptivity*, wherein after a certain stimulus, the system modifies its own response pattern in anticipation of similar stimuli in the future. It is through processes such as down-regulation, long-term potentiation, habituation, and synaptic learning, that the body modifies its definitions of "an ideal state" or "an ideal response". Thus the rules that explain the behavior of homeostatic systems are nested in larger systems, called homeodynamic systems [41]. Imagine an individual whose homeostatic system is trying to keep its blood glucose level around 6, and Body Mass Index (BMI) around 22 (also known as an attractor zone). The system tolerates and shows resilience in the face of a range of perturbations in the short-term (such as skipping a meal, met by glycogenolysis), medium-term (entering Ramadan, met by lipolysis), or permanent (increased energy demands, met by increasing appetite). Whatever happens, the system will try to return to the initial attractor zone. This is easy to do if the person's state meanders only slightly in the immediate phase space around the central attractor zone, also known as the system's *basin* (eg, glucose 4-8, BMI 20-24). However, if the person's state moves far away enough from the basin (eg, glucose 2, BMI 10), the integrity of the system is at stake and death from starvation may follow. In other diseases, the attractor zone may shift to BMI 26 (diagnosed as overweight) or to glucose 12 (diagnosed with diabetes) over a few decades. Other diseases still may show less gradual but more sudden shifts with bifurcation into a far away attractor zone (as seen in an acute epileptic seizure [42], many other acute presentations, but also in agent-based models of community interventions [43]). Other diseases can show cyclical flip-flop behavior between two attractor zones (such as is seen in bipolar disorder [44]). Dynamical Systems Theory focuses on understanding how and why these attractor zones change over time. Chaos Theory describes a type of system change that is particularly sensitive to initial conditions (such as seen in the Barker and Hygiene hypotheses). Similar models have been used to understand cell fates [45], endothelial function [46], cytokine function [47], heart rate variability [48], atrial fibrillation [49], septic shock [50], and multiple organ dysfunction syndrome [51]. Understanding these dynamic processes may lead to targeted interventions of preventative [52] or curative nature. For example, animal models suggest that infection with helmliths can reprogram autoimmune disease states towards healthier ones in colitis [53], gastric atrophy [54], multiple sclerosis [55], and diabetes [56].

Analytical Approaches to Systems Medicine

Conducting systems medicine research will require numerous changes to how medical scientists and epidemiologists conduct their daily work. Such changes have been poorly addressed in the existing literature. Fundamentally, an entire new approach to data collection and analyses is required. If most risk factors interact with one another to create small interactive effects, many of which are nested in intricate multi-levels of hierarchy, then the detection of such intricacies will require large numbers of patients with large numbers of variables [57].

Once assembled, such high-fidelity datasets could be explored for “hot spots”, or loci of association, between risk factor data (including behavioral, environmental, and sociological risks), molecular data (including genomic, trascriptomic, and proteomic data), and clinical outcomes, in order to identify and treat etiphenotypical subgroups. Some of this exploration will take the narrow shape of testing a priori hypotheses. Other exploration will use high-throughput techniques to scan more broadly and identify novel hypotheses for subsequent scrutiny [58,59]. This may be facilitated by machine learning algorithms [60]. For example, the world’s most advanced supercomputer, IBM Watson, has recently been given access to data on over a million cancer patients, as well as the emerging oncological literature. It is hoped that Watson will support and enhance clinical decision making in real time [61]. This may shatter the traditional view of humans generating hypotheses on which computers calculate the test statistic, as computers like Watson may also become indispensible in hypothesis generation. Our vision of Systems Medicine becomes impossible without seamless collaboration with mathematicians, modelers, and data scientists, so that the right modeling tools can be used and combined to appropriately balance sufficient complexity with practical utility.

A Roadmap to Systems Medicine

Although multiple challenges impede the integration of systems approaches within medical research, we prioritize five that deserve attention first. (1) Existing models that view the human body as complex adaptive systems are limited and need specialist development. Our review of the literature suggests that while textbooks exist on modeling for Systems Science and Systems Biology, little consensus exists on modeling for Systems Medicine. (2) While large-scale research projects (such as the Human Genome Project, Wellcome Trust Case Control Consortium, Virtual Physiological Human, ECell, Cancer Genome Atlas, Human Brain Project) are developing at a promising rate, the majority of research continues to be conducted in an almost proprietary-like mindset of non-sharing, with individual teams collecting and holding data in small silos. This makes it difficult to create the large datasets required for systems analysis. (3) While some organizations (like IBM, Sage Bionetworks, and Farr Institute) can foster multidisciplinary teams between medical researchers and information specialists, this collaboration gap remains large in most areas of public research. (4) Although research funders are beginning to recognize and invest in long-term and larger research endeavors, particularly in database curation, the vast majority of funds continue to be spent on low-complexity projects with short-term deliverables and seemingly low risk. Many such projects are underpowered in breadth, depth, and complexity to sufficiently address the problems they seek to address, with sometimes negligible subsequent improvement in our knowledge base. (5) The obstacles above will yield with sufficient energy and leadership; however, progress at each front continues to be slow due to the prevalent reductionist culture in the medical research community and general risk aversion to engage with new technology (particularly those that may alter existing power relations). Therefore, greater discussion, awareness, and education of differences between systems medicine and reductionism are required across the board, in order to promote and facilitate interest and activity in this important transition.

Sharing Data

The medical research community itself is a complex adaptive system. Shifting this toward a more efficient, systems-science configuration will require time and effort to be applied at multiple levels. We believe that the second obstacle mentioned above, creating large datasets, is most pressing so we discuss this in detail. There are two possible non-competing solutions: (1) sharing data between researchers or (2) using data from electronic health records (EHR) for research purposes. Although data sharing appears to be increasing, most research data are not shared or recycled outside the original project team. A range of factors discourage data sharing. Project-specific data are often collected using context-specific priorities, definitions, and measurement tools that are rarely compatible with other peer-researchers, let alone researchers looking at the picture from a higher or lower level of order [62]. Data access requests are often cumbersome and slow, and some researchers may be wary of giving up their competitive edge [63]. Ultimately, even if all research groups pool their data successfully, a rich trove of clinical process and outcome data will continue to be held by health care providers who are not involved in research.

Health care data are never assembled for research purposes, which substantially hampers its transferability. For example, patients with borderline disease states are often “upcoded” into more severe disease states, as this brings financial rewards to the health care provider. Some EHR data may always be absent or of insufficient quality to be useful in medical research (eg, tracking the daily improvement of inpatients with cellulitis or fluid overload, which is difficult to quantify). However, other EHR data lend themselves very well to scientific analyses, particularly as drug prescription data are of exceptional quality. Furthermore, as health care is carried out on real patients, studying the real-life effects of drugs is preferable to unrepresentative clinical trials, particularly in understanding how one drug can accidentally influence two disparate diseases [64,65]. Future work could clarify which health care data are transferrable to medical research, and develop some tools to facilitate this. For example, should some severe diagnoses, when seen in conjunction with other hospital laboratory parameters, be downcoded back to a more truthful level, to correct for financial upcoding in EHR? The integration of health care and research data will create its own technical, logistical, and legal challenges, and at present we cannot tell if this will be worthwhile. If attempted, we suggest that groups of patients could be subdivided empirically based on their differential risk factor profiles, disease trajectories, or molecular data [66], and these subgroups subsequently explored in more detail for evidence of etiphenotypical subgroups.

Unlocking Health Care Data

There is growing support for the collation of EHRs for secondary research purposes [67]. Transmission of data is typically thought to proceed with researchers asking health care providers for access to anonymized data [68]. Added value can be gained by integrating data from multiple providers [69]. Overall, this approach offers substantial productivity gains. However, this benefit must be considered against the drawback of reduced data accuracy.

As an alternative, we suggest an approach where researchers ask patients directly for their data, who in turn have to ask hospitals for their data. The advantage here is that routine medical data could be augmented by study-specific scientific tests (such as genotyping). The disadvantage is that samples will not be representative of the wider patient group, but biased towards those comfortable with sharing their personal data with scientists over the Internet, a concept that probably frightens most patients today. It also requires EHRs to be personally accessible and personally controlled—a feature that is still under development [70], but with some proof of concept from private settings [71].

Online Communities

Web 2.0 technology is creating a social Web, where users create content, share useful information with interested peers, and moderate each other’s activities (eg, Facebook, eBay, ReseacherGate) [72]. Web 2.0 has been combined with the power of crowdsourcing (the force behind projects like Wikipedia) to foster Science 2.0 (also known as Open Science or Cyberscience 2.0) [73]. Here, online platforms facilitate large-scale data sharing between researchers who can merge or re-analyze each other’s data (eg, HapMap project, Sage Bionetworks). Some of these projects (eg, GalaxyZoo) are also enlisting the power of citizens in the scientific endeavor. A seminal article by Eysenbach in 2008 [74] detailed how the application of these Web 2.0 technologies to personally controlled health care records will create “Medicine 2.0”. This will be a world where the collaboration gap between researcher, clinician, and patient will narrow. Online patient communities such as PatientsLikeMe are not only useful to facilitate peer transfer of knowledge [75] but also in scientifically testing new drug applications [76].

To the best of our knowledge, a major gap in the current literature is the fact that systems medicine researchers have never suggested that their vision could be achieved quicker using Medicine 2.0 tools. Similarly, although Medicine 2.0 enthusiasts have begun to associate their ideas to Big Data, Web Science, Health Web Science, and the Semantic Web [77], none have spoken of the potential for Medicine 2.0 to further our understanding of the human body as a complex adaptive system. We urge these two important communities to begin collaboration on what we suggest could be called “Systems Medicine 2.0” (Figure 3). Understandably, few patients will trust their most sensitive secrets with an amorphous, faceless research community. An online community with detailed researcher profiles, faces, and credentials (eg, ReseacherGate) could be augmented by allowing participants and researchers to interact with one another. Reliable feedback systems (such as those seen in Amazon and eBay) will be central to creating sufficient trust. To date, most sites have focused on fostering researcher-researcher collaboration [78,79], or patient-patient collaboration. The next step would be to foster bi- or tri-directional links (also to clinicians), with the added challenge of increased knowledge asymmetries.

Another challenge for Systems Medicine 2.0 is how health care data has very poor interoperability, both nationally and internationally. We welcome the work of the Joint Initiative on SDO Global Health Informatics Standardization to facilitate the rapid adoption of common health informatics standards. A related challenge is how various Systems Medicine research teams will initially create their own datasets, analytical programs, and models. Reporting these via traditional journals will hamper progress. Datasets, programs, and models should be shared, scrutinized, and developed, perhaps by making use of the efficiency tools perfected by the Open Source movement [80]. Open access [81], open standards, open source software [82], and open competitions [83] can foster faster innovation and drive efficiency. Ideally, the entire Systems Medicine community could see themselves as working on the same project, akin to the global communities who brought us Linux, Mozilla Firefox, Wikipedia, HUGO, and Systems Biology.

To start the discussion of how the research community can move towards Systems Medicine, we have compiled a list of suggestions (Textbox 1). This is neither exhaustive nor comprehensive and is designed merely to stimulate discussion, development, and implementation of Systems Medicine.

Textbox 1. Initial suggestions for the advancement of systems medicine.

Governments and key research funders:

- training systems scientists, big data engineers, and medical researchers
- incentivizing and normalizing multidisciplinary collaboration for systems medicine
- incentivizing and normalizing the sharing of laboratory data as well as health care data
- earmarking research funds for systems medicine
- national leadership on systems medicine strategy and implementation

Legislation for the Web 2.0 era:

- EHRs to be recognized as patient property
- EHR systems must facilitate the simple and free export/import of data by abiding to global standards and/or open standards
- anonymized research on EHRs is permitted and automatically facilitated, unless patients opt out
- define if and how data can be accessed for commercial purposes and how any proceeds are divided

Scientific community:

- opinion leaders to promote the benefits of Systems Science approaches in medicine
- disease-specific expert bodies (eg, colleges) to facilitate common standards of data collection in their fields
- early Systems Medicine 2.0 academics to focus their papers at one flagship journal (eg, *Journal of Medical Internet Research*)

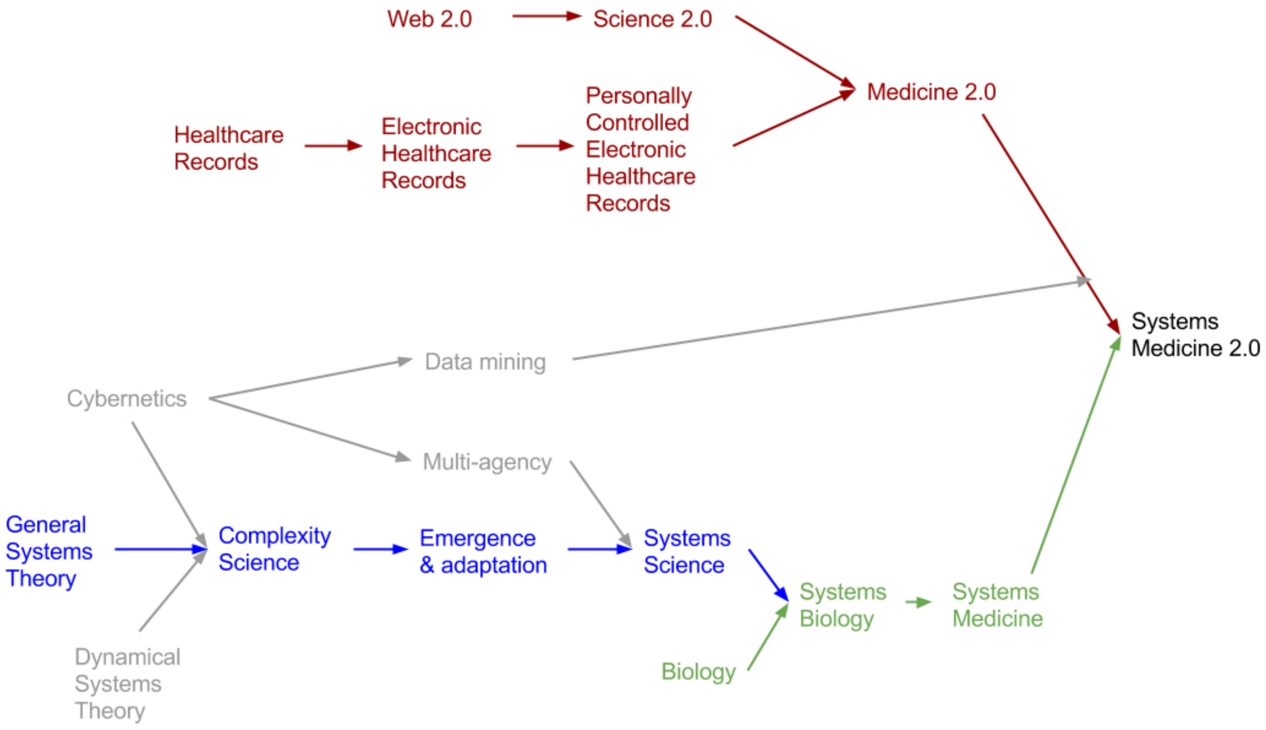
Software developers:

- develop an online community that links together systems scientists and modelers, big data engineers, medical researchers, and subsequently patients

International co-ordination:

- international leadership and co-ordination (eg, World Health Organization) to promote all of the above
- an online data-sharing directory listing all large datasets pertaining to Systems Medicine

Figure 3. The historical origins of “Systems Science” and “Medicine 2.0”, and the potential for their combination into “Systems Medicine 2.0”.



Conclusion

The prevalent paradigm in medical science is reductionism, whose limitations have become increasingly apparent, resulting in diminishing returns. This paradigm has been supplemented and modified with the framework of network medicine. If further developed with Systems Science approaches, this has the potential to evolve into fully fledged Systems Medicine paradigm. This neatly complements advances in Internet-powered medicine (Medicine 2.0), but as yet the two fields have yet to take advantage of each other's nascent existence. We hope that our paper can bridge this conceptual gap and advance mutual interest and collaboration between these two, to foster Systems Medicine 2.0. This could lead to significant advances in the prevention and treatment of non-communicable diseases.

Acknowledgments

We wish to thank Michael Baum and Jayant Vaidya for their contributions on how systems science has been applied to breast cancer. We also thank anonymous reviewers at JAMIA and JMIR, as their comments substantially developed the paper. This work was supported by the Wellcome Trust (grant number 097821/Z/11/A), and the lead author has received fellowships from the National Institute of Health Research and the Wellcome Trust.

Authors' Contributions

TT wrote the first concept paper (in 2009), and the final draft. AG and TT wrote the first draft. GS contributed information about the use of EHRs for secondary research in the United Kingdom. OH contributed information about the use of EHRs in United Arab Emirates and wrote the third draft. AD contributed information about proteomics and metabolomics. PH supervised the research throughout, developed the systems theory concept, and wrote the second draft.

Conflicts of Interest

OH is a Senior Vice President with Healthways International, a global population health and well-being improvement company. As such, he seeks to promote holistic and effective approaches to improving health.

Multimedia Appendix 1

A PDF illustrating various aspects of a Systems Medicine model of disease: non-linearity (frame 1), multi-agency (frame 3), multilevelness (frame 4).

[PDF File (Adobe PDF File), 121KB - jmir_v17i3e64_app1.pdf]

References

1. Crimmins EM, Saito Y. Trends in healthy life expectancy in the United States, 1970-1990: gender, racial, and educational differences. Soc Sci Med 2001 Jun;52(11):1629-1641. [Medline: 11327137]
2. Ahn AC, Tewari M, Poon CS, Phillips RS. The limits of reductionism in medicine: could systems biology offer an alternative? PLoS Med 2006 May;3(6):e208 [FREE Full text] [doi: 10.1371/journal.pmed.0030208] [Medline: 16681415]
3. Liu YI, Wise PH, Butte AJ. The "etiome": identification and clustering of human disease etiological factors. BMC Bioinformatics 2009;10 Suppl 2:S14 [FREE Full text] [doi: 10.1186/1471-2105-10-S2-S14] [Medline: 19208189]
4. Yildirim MA, Goh KI, Cusick ME, Barabási AL, Vidal M. Drug-target network. Nat Biotechnol 2007 Oct;25(10):1119-1126. [doi: 10.1038/nbt1338] [Medline: 17921997]
5. Kuhn TS. The Structure of Scientific Revolutions. Chicago: University Of Chicago Press; 1962.
6. Lakatos I. Falsification and the methodology of scientific research programmes. In: Can Theories be Refuted? . Springer: Netherlands; 1975:205.
7. Loscalzo J, Kohane I, Barabasi AL. Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. Mol Syst Biol 2007;3:124 [FREE Full text] [doi: 10.1038/msb4100163] [Medline: 17625512]
8. Hidalgo CA, Blumm N, Barabási AL, Christakis NA. A dynamic network approach for the study of human phenotypes. PLoS Comput Biol 2009 Apr;5(4):e1000353 [FREE Full text] [doi: 10.1371/journal.pcbi.1000353] [Medline: 19360091]
9. May C, Montori VM, Mair FS. We need minimally disruptive medicine. BMJ 2009;339:b2803. [Medline: 19671932]
10. Csermely P, Agoston V, Pongor S. The efficiency of multi-target drugs: the network approach might help drug design. Trends Pharmacol Sci 2005 Apr;26(4):178-182. [doi: 10.1016/j.tips.2005.02.007] [Medline: 15808341]
11. Medical Research Council. Developing and evaluating complex interventions: new guidance. 2008. URL: http://www.mrc.ac.uk/complexinterventionsguidance/ [WebCite Cache ID 6KC76KPKI]
12. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet 2011 Jan;12(1):56-68 [FREE Full text] [doi: 10.1038/nrg2918] [Medline: 21164525]
13. Loscalzo J, Barabasi AL. Systems biology and the future of medicine. Wiley Interdiscip Rev Syst Biol Med 2011;3(6):619-627 [FREE Full text] [doi: 10.1002/wsbm.144] [Medline: 21928407]

14. Chuang HY, Lee E, Liu YT, Lee D, Ideker T. Network-based classification of breast cancer metastasis. Mol Syst Biol 2007;3:140 [FREE Full text] [doi: 10.1038/msb4100180] [Medline: 17940530]
15. Ergün A, Lawrence CA, Kohanski MA, Brennan TA, Collins JJ. A network biology approach to prostate cancer. Mol Syst Biol 2007;3:82 [FREE Full text] [doi: 10.1038/msb4100125] [Medline: 17299418]
16. Chang W, Ma L, Lin L, Gu L, Liu X, Cai H, et al. Identification of novel hub genes associated with liver metastasis of gastric cancer. Int J Cancer 2009 Dec 15;125(12):2844-2853. [doi: 10.1002/ijc.24699] [Medline: 19569046]
17. Sirota M, Dudley JT, Kim J, Chiang AP, Morgan AA, Sweet-Cordero A, et al. Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci Transl Med 2011 Aug 17;3(96):96ra77 [FREE Full text] [doi: 10.1126/scitranslmed.3001318] [Medline: 21849665]
18. Mesarovic M, Sreenath SN. Beyond the Flat Earth Perspective in Systems Biology. Biological Theory 2006 Jan;1(1):33-34. [doi: 10.1162/biot.2006.1.1.33]
19. Gross F. What systems biology can tell us about disease. Hist Philos Life Sci 2011;33(4):477-496. [Medline: 22662506]
20. Scholes S, Bajekal M, Norman P, O'Flaherty M, Hawkins N, Kivimäki M, et al. Quantifying policy options for reducing future coronary heart disease mortality in England: a modelling study. PLoS One 2013;8(7):e69935 [FREE Full text] [doi: 10.1371/journal.pone.0069935] [Medline: 23936122]
21. Klir G. Facets of systems science. New York: Kluwer Academic/Plenum Publishers; 2001.
22. Wolkenhauer O. Why model? Front Physiol 2014;5:21 [FREE Full text] [doi: 10.3389/fphys.2014.00021] [Medline: 24478728]
23. Zhang B, Gaiteri C, Bodea LG, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell 2013 Apr 25;153(3):707-720 [FREE Full text] [doi: 10.1016/j.cell.2013.03.030] [Medline: 23622250]
24. Wolkenhauer O, Shibata DK, Mesarović MD. A stem cell niche dominance theorem. BMC Syst Biol 2011;5:4 [FREE Full text] [doi: 10.1186/1752-0509-5-4] [Medline: 21214945]
25. Bousquet J, Anto JM, Sterk PJ, Adcock IM, Chung KF, Roca J, et al. Systems medicine and integrated care to combat chronic noncommunicable diseases. Genome Med 2011;3(7):43 [FREE Full text] [doi: 10.1186/gm259] [Medline: 21745417]
26. Capra F. The web of life: a new synthesis of mind and matter. London: Flamingo; 1996.
27. Cordell HJ. Detecting gene-gene interactions that underlie human diseases. Nat Rev Genet 2009 Jun;10(6):392-404 [FREE Full text] [doi: 10.1038/nrg2579] [Medline: 19434077]
28. Ziebarth JD, Bhattacharya A, Chen A, Cui Y. PolymiRTS Database 2.0: linking polymorphisms in microRNA target sites with human diseases and complex traits. Nucleic Acids Res 2012 Jan;40(Database issue):D216-D221 [FREE Full text] [doi: 10.1093/nar/gkr1026] [Medline: 22080514]
29. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 2010 Apr;6(4):e1000888 [FREE Full text] [doi: 10.1371/journal.pgen.1000888] [Medline: 20369019]
30. Gerke J, Lorenz K, Cohen B. Genetic interactions between transcription factors cause natural variation in yeast. Science 2009 Jan 23;323(5913):498-501 [FREE Full text] [doi: 10.1126/science.1166426] [Medline: 19164747]
31. Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ, et al. Variations in DNA elucidate molecular networks that cause disease. Nature 2008 Mar 27;452(7186):429-435 [FREE Full text] [doi: 10.1038/nature06757] [Medline: 18344982]
32. Zak I, Niemiec P, Balcerzyk A, Krauze J. Combined "pro-atherosclerotic" variants of the ACE and APOE genes increase the risk of the coronary artery disease associated with the presence of cigarette smoking. Acta Cardiol 2008 Dec;63(6):741-747. [Medline: 19157170]
33. Choudhry S, Avila PC, Nazario S, Ung N, Kho J, Rodriguez-Santana JR, et al. CD14 tobacco gene-environment interaction modifies asthma severity and immunoglobulin E levels in Latinos with asthma. Am J Respir Crit Care Med 2005 Jul 15;172(2):173-182. [doi: 10.1164/rccm.200409-1232OC] [Medline: 15879416]
34. Liu AY, Scherer D, Poole E, Potter JD, Curtin K, Makar K, et al. Gene-diet-interactions in folate-mediated one-carbon metabolism modify colon cancer risk. Mol Nutr Food Res 2013 Apr;57(4):721-734 [FREE Full text] [doi: 10.1002/mnfr.201200180] [Medline: 22961839]
35. Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 2003 Jul 18;301(5631):386-389 [FREE Full text] [doi: 10.1126/science.1083968] [Medline: 12869766]
36. Bisgaard H, Simpson A, Palmer CNA, Bønnelykke K, McLean I, Mukhopadhyay S, et al. Gene-environment interaction in the onset of eczema in infancy: filaggrin loss-of-function mutations enhanced by neonatal cat exposure. PLoS Med 2008 Jun 24;5(6):e131 [FREE Full text] [doi: 10.1371/journal.pmed.0050131] [Medline: 18578563]
37. Itzhaki RF, Lin WR, Shang D, Wilcock GK, Faragher B, Jamieson GA. Herpes simplex virus type 1 in brain and risk of Alzheimer's disease. Lancet 1997 Jan 25;349(9047):241-244. [doi: 10.1016/S0140-6736(96)10149-5] [Medline: 9014911]
38. Ramagopalan SV, Maugeri NJ, Handunnetthi L, Lincoln MR, Orton SM, Dyment DA, et al. Expression of the multiple sclerosis-associated MHC class II Allele HLA-DRB1*1501 is regulated by vitamin D. PLoS Genet 2009 Feb;5(2):e1000369 [FREE Full text] [doi: 10.1371/journal.pgen.1000369] [Medline: 19197344]
39. Higgins JP. Nonlinear systems in medicine. Yale J Biol Med 2002;75(5-6):247-260 [FREE Full text] [Medline: 14580107]

40. Wang IM, Zhang B, Yang X, Zhu J, Stepaniants S, Zhang C, et al. Systems analysis of eleven rodent disease models reveals an inflammatome signature and key drivers. Mol Syst Biol 2012;8:594 [FREE Full text] [doi: 10.1038/msb.2012.24] [Medline: 22806142]
41. McEwen BS, Wingfield JC. The concept of allostasis in biology and biomedicine. Horm Behav 2003 Jan;43(1):2-15. [Medline: 12614627]
42. Stam CJ. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. Clin Neurophysiol 2005 Oct;116(10):2266-2301. [doi: 10.1016/j.clinph.2005.06.011] [Medline: 16115797]
43. Chalabi Z, Lorenc T. Using agent-based models to inform evaluation of complex interventions: examples from the built environment. Prev Med 2013 Nov;57(5):434-435. [doi: 10.1016/j.ypmed.2013.07.013] [Medline: 23872173]
44. Gottschalk A, Bauer MS, Whybrow PC. Evidence of chaotic mood variation in bipolar disorder. Arch Gen Psychiatry 1995 Nov;52(11):947-959. [Medline: 7487343]
45. Huang S, Eichler G, Bar-Yam Y, Ingber DE. Cell fates as high-dimensional attractor states of a complex gene regulatory network. Phys Rev Lett 2005 Apr 1;94(12):128701. [Medline: 15903968]
46. Aird WC. Endothelial cell dynamics and complexity theory. Crit Care Med 2002 May;30(5 Suppl):S180-S185. [Medline: 12004233]
47. Callard R, George AJ, Stark J. Cytokines, chaos, and complexity. Immunity 1999 Nov;11(5):507-513 [FREE Full text] [Medline: 10591175]
48. Lombardi F. Chaos theory, heart rate variability, and arrhythmic mortality. Circulation 2000;101(1):8-10 [FREE Full text] [Medline: 10618296]
49. Garfinkel A, Chen PS, Walter DO, Karagueuzian HS, Kogan B, Evans SJ, et al. Quasiperiodicity and chaos in cardiac fibrillation. J Clin Invest 1997 Jan 15;99(2):305-314 [FREE Full text] [doi: 10.1172/JCI119159] [Medline: 9005999]
50. Buchman TG. The community of the self. Nature 2002 Nov 14;420(6912):246-251. [doi: 10.1038/nature01260] [Medline: 12432410]
51. Seely AJ, Christou NV. Multiple organ dysfunction syndrome: exploring the paradigm of complex nonlinear systems. Crit Care Med 2000 Jul;28(7):2193-2200. [Medline: 10921540]
52. Sherman RC, Langley-Evans SC. Early administration of angiotensin-converting enzyme inhibitor captopril, prevents the development of hypertension programmed by intrauterine exposure to a maternal low-protein diet in the rat. Clin Sci (Lond) 1998 Apr;94(4):373-381. [Medline: 9640343]
53. Khan WI, Blennerhasset PA, Varghese AK, Chowdhury SK, Omsted P, Deng Y, et al. Intestinal nematode infection ameliorates experimental colitis in mice. Infect Immun 2002 Nov;70(11):5931-5937 [FREE Full text] [Medline: 12379667]
54. Fox JG, Beck P, Dangler CA, Whary MT, Wang TC, Shi HN, et al. Concurrent enteric helminth infection modulates inflammation and gastric immune responses and reduces helicobacter-induced gastric atrophy. Nat Med 2000 May;6(5):536-542. [doi: 10.1038/75015] [Medline: 10802709]
55. Sewell D, Qing Z, Reinke E, Elliot D, Weinstock J, Sandor M, et al. Immunomodulation of experimental autoimmune encephalomyelitis by helminth ova immunization. Int Immunol 2003 Jan;15(1):59-69 [FREE Full text] [Medline: 12502726]
56. Weinstock JV. Autoimmunity: The worm returns. Nature 2012 Nov 8;491(7423):183-185 [FREE Full text] [doi: 10.1038/491183a] [Medline: 23135449]
57. Palla L, Higgins JPT, Wareham NJ, Sharp SJ. Challenges in the use of literature-based meta-analysis to examine gene-environment interactions. Am J Epidemiol 2010 Jun 1;171(11):1225-1232 [FREE Full text] [doi: 10.1093/aje/kwq051] [Medline: 20406760]
58. Hubner N, Wallace CA, Zimdahl H, Petretto E, Schulz H, Maciver F, et al. Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. Nat Genet 2005 Mar;37(3):243-253. [doi: 10.1038/ng1522] [Medline: 15711544]
59. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet 2012 Jun;13(6):395-405. [doi: 10.1038/nrg3208] [Medline: 22549152]
60. Bilal E, Dutkowski J, Guinney J, Jang IS, Logsdon BA, Pandey G, et al. Improving breast cancer survival analysis through competition-based multidimensional modeling. PLoS Comput Biol 2013;9(5):e1003047 [FREE Full text] [doi: 10.1371/journal.pcbi.1003047] [Medline: 23671412]
61. Forbes. IBM's Watson Gets Its First Piece Of Business In Healthcare. 2013. URL: http://www.forbes.com/sites/bruceupbin/2013/02/08/ibms-watson-gets-its-first-piece-of-business-in-healthcare/ [accessed 2013-10-09] [WebCite Cache ID 6KEn7iwcP]
62. Leonelli S. Global data for local science: Assessing the scale of data infrastructures in biological and biomedical research. BioSocieties 2013 Nov 11;8(4):449-465. [doi: 10.1057/biosoc.2013.23]
63. Research Information Network. Stewardship of digital research data: a framework of principles and guidelines. 2008. URL: http://www.rin.ac.uk/system/files/attachments/Stewardship-data-guidelines.pdf [accessed 2013-10-09] [WebCite Cache ID 6KEnWLAYq]
64. Brauer R, Smeeth L, Anaya-Izquierdo K, Timmis A, Denaxas SC, Farrington CP, et al. Antipsychotic drugs and risks of myocardial infarction: a self-controlled case series study. Eur Heart J 2014 Jul 8:-. [doi: 10.1093/eurheartj/ehu263] [Medline: 25005706]

65. Quint JK, Herrett E, Bhaskaran K, Timmis A, Hemingway H, Wedzicha JA, et al. Effect of β blockers on mortality after myocardial infarction in adults with COPD: population based cohort study of UK electronic healthcare records. BMJ 2013;347:f6650 [FREE Full text] [Medline: 24270505]
66. Clermont G, Auffray C, Moreau Y, Rocke DM, Dalevi D, Dubhashi D, et al. Bridging the gap between systems biology and medicine. Genome Med 2009;1(9):88 [FREE Full text] [doi: 10.1186/gm88] [Medline: 19754960]
67. Souhami R. Personal data for public good: using health information in medical research. London: Academy of Medical Sciences; 2006.
68. Mandl KD, Kohane IS. Tectonic shifts in the health information economy. N Engl J Med 2008 Apr 17;358(16):1732-1737. [doi: 10.1056/NEJMsb0800220] [Medline: 18420506]
69. Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, van ST, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. BMJ 2013;346:f2350 [FREE Full text] [Medline: 23692896]
70. Tiik M, Ross P. Patient opportunities in the Estonian Electronic Health Record System. Stud Health Technol Inform 2010;156:171-177. [Medline: 20543352]
71. Bourgeois FT, Simons WW, Olson K, Brownstein JS, Mandl KD. Evaluation of influenza prevention in the workplace using a personally controlled health record: randomized controlled trial. J Med Internet Res 2008;10(1):e5 [FREE Full text] [doi: 10.2196/jmir.984] [Medline: 18343794]
72. Kidd MR. Personal electronic health records: MySpace or HealthSpace? BMJ 2008 May 10;336(7652):1029-1030 [FREE Full text] [doi: 10.1136/bmj.39567.550301.80] [Medline: 18460554]
73. Lunshof JE, Chadwick R, Vorhaus DB, Church GM. From genetic privacy to open consent. Nat Rev Genet 2008 May;9(5):406-411. [doi: 10.1038/nrg2360] [Medline: 18379574]
74. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
75. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. J Med Internet Res 2010;12(2):e19 [FREE Full text] [doi: 10.2196/jmir.1549] [Medline: 20542858]
76. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. Nat Biotechnol 2011 May;29(5):411-414. [doi: 10.1038/nbt.1837] [Medline: 21516084]
77. Luciano JS, Cumming GP, Wilkinson MD, Kahana E. The emergent discipline of health web science. J Med Internet Res 2013;15(8):e166 [FREE Full text] [doi: 10.2196/jmir.2499] [Medline: 23968998]
78. Schleyer T, Spallek H, Butler BS, Subramanian S, Weiss D, Poythress ML, et al. Facebook for scientists: requirements and services for optimizing how scientific collaborations are established. J Med Internet Res 2008;10(3):e24 [FREE Full text] [doi: 10.2196/jmir.1047] [Medline: 18701421]
79. Kahlon M, Yuan L, Daigre J, Meeks E, Nelson K, Piontkowski C, et al. The use and significance of a research networking system. J Med Internet Res 2014;16(2):e46 [FREE Full text] [doi: 10.2196/jmir.3137] [Medline: 24509520]
80. Bot BM, Burdick D, Kellen M, Huang ES. clearScience: Infrastructure for Communicating Data-Intensive Science. AMIA Jt Summits Transl Sci Proc 2013;2013:27. [Medline: 24303291]
81. Leonelli S. Why the Current Insistence on Open Access to Scientific Data? Big Data, Knowledge Production, and the Political Economy of Contemporary Biology. Bulletin of Science, Technology & Society 2013 Aug 20;33(1-2):6-11. [doi: 10.1177/0270467613496768]
82. Reynolds CJ, Wyatt JC. Open source, open standards, and health care information systems. J Med Internet Res 2011;13(1):e24 [FREE Full text] [doi: 10.2196/jmir.1521] [Medline: 21447469]
83. Plenge RM, Greenberg JD, Mangravite LM, Derry JMJ, Stahl EA, Coenen MJH, International Rheumatoid Arthritis Consortium (INTERACT), et al. Crowdsourcing genetic prediction of clinical utility in the Rheumatoid Arthritis Responder Challenge. Nat Genet 2013 May;45(5):468-469 [FREE Full text] [doi: 10.1038/ng.2623] [Medline: 23619782]

Abbreviations

BMI: Body Mass Index
EHR: Electronic Health Record

Edited by G Eysenbach; submitted 05.11.13; peer-reviewed by V Gavgani, S Green; comments to author 01.03.14; revised version received 05.09.14; accepted 12.09.14; published 23.03.15.

Please cite as:
Tillmann T, Gibson AR, Scott G, Harrison O, Dominiczak A, Hanlon P
Systems Medicine 2.0: Potential Benefits of Combining Electronic Health Care Records With Systems Science Models
J Med Internet Res 2015;17(3):e64
URL: http://www.jmir.org/2015/3/e64/
doi:10.2196/jmir.3082
PMID:25831125

Original Paper

The Impact of a Web-Based App (eBalance) in Promoting Healthy Lifestyles: Randomized Controlled Trial

Jenny Safran Naimark[1], PhD; Zecharia Madar[1], PhD; Danit R Shahar[2], PhD

[1]Robert H. Smith Faculty of Agriculture, Food and Environment, Hebrew University of Jerusalem, Rehovot, Israel

[2]S. Daniel Abraham International Center for Health and Nutrition,Department of Public Health, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel

Corresponding Author:
Jenny Safran Naimark, PhD
Robert H. Smith Faculty of Agriculture, Food and Environment
Hebrew University of Jerusalem
POBox 12
Rehovot, 76100
Israel
Phone: 972 524788559
Fax: 972 37320370
Email: jennynaimark@gmail.com

Abstract

Background: The use of Web-based apps to promote a healthy lifestyle is increasing, although most of these programs were not assessed using suitable epidemiological methods. We evaluated the effectiveness of a newly developed Web-based app in promoting a healthy lifestyle and educating adults on such lifestyles. We also analyzed predictors for success in acquiring and maintaining a healthy lifestyle.

Objective: Our aim was to compare people receiving a new Web-based app with people who got an introductory lecture alone on healthy lifestyle, weight change, nutritional knowledge, and physical activity, and to identify predictors of success for maintaining a healthy lifestyle.

Methods: Subjects were recruited from the community and were randomized into intervention and control groups. The intervention subjects received access to the app without any face-to-face support; the control subjects continued their standard lifestyle. Measurements were taken by the researcher at baseline and after 14 weeks and included weight and waist circumference. Nutritional knowledge, diet quality, and physical activity duration were obtained using online questionnaires. The new Web-based app was developed based on current US Department of Agriculture and Israel Ministry of Health recommendations for healthy lifestyle. The app provides tools for monitoring diet and physical activity while instructing and encouraging healthy diet and physical activity.

Results: Out of 99 subjects who were randomized into app and control groups, 85 participants (86%) completed the study, 56 in the intervention and 29 in the control group. The mean age was 47.9 (SD 12.3) years, and mean Body Mass Index was 26.2 (SD 3.9). Among the intervention group only, frequency of app use was 2.7 (SD 1.9) days/week. The mean change in physical activity was 63 (SD 20.8) minutes in the app group and -30 (SD 27.5) minutes in the control group (*P*=.02). The mean weight change was -1.44 (SD 0.4) kg in the app group and -0.128 (SD 0.36) kg in the control group (*P*=.03). Knowledge score increased significantly in the app group, 76 (SD 7.5) to 79 (SD 8.7) at the end of the study (*P*=.04) compared with the control group. Diet quality score also increased significantly at the end of the study, from 67 (SD 9.8) to 71 (SD 7.6; *P*<.001) in contrast to the control group. Success score (represents the success in maintaining healthy lifestyle) was higher among the app group (68%) compared with 36% in the control group (*P*<.001). The app frequency of use was significantly related to a higher success score (*P*<.001).

Conclusions: We showed a positive impact of a newly developed Web-based app on lifestyle indicators during an intervention of 14 weeks. These results are promising in the app's potential to promote a healthy lifestyle, although larger and longer duration studies are needed to achieve more definitive conclusions.

Trial Registration: Clinicaltrial.gov number: NCT01913496; http://www.clinicaltrials.gov/ct2/show/NCT01913496 (Archived by WebCite at http://www.webcitation.org/6WSTUEPuJ).

(J Med Internet Res 2015;17(3):e56) doi:10.2196/jmir.3682

KEYWORDS

Web-based; healthy lifestyle; mobile apps; mhealth

Introduction

Although chronic diseases have reached epidemic proportions, they could be significantly reduced through prevention or reduction of their risk factors, early detection, and timely treatments [1]. Most chronic diseases are strongly associated and causally linked with four particular behaviors: tobacco use, physical inactivity, unhealthy diet, and excessive consumption of alcohol. These behaviors lead to four key metabolic/physiological changes: hypertension, overweight/obesity, hyperglycemia, and hyperlipidemia. In terms of attributable deaths, the leading chronic disease risk factors globally are hypertension (attributed with 13% of global deaths), tobacco use (9%), hyperglycemia (6%), physical inactivity (6%), and overweight and obesity (5%) [2].

A healthy lifestyle was proven to lower the risk for chronic diseases. In the European Prospective Investigation into Cancer and Nutrition (EPIC) study, 23,153 German participants aged 35-65 were followed for a mean of 7.8 years. Adherence to four health behaviors not smoking, exercising 3.5 hours per week, eating a healthy diet (high intake of fruits, vegetables, and whole-grain bread, and low meat consumption), and having a Body Mass Index (BMI) <30 kg/m^2 at baseline was associated with 78% lower risk for developing chronic disease (diabetes 93%, myocardial infarction 81%, stroke 50%, and cancer 36%) than participants without these protective factors [3,4].

Despite the importance of following healthy lifestyle recommendations for preventing or lowering the incidence of chronic diseases, multiple population studies showed that only a minority of individuals adheres to healthy lifestyle behaviors. In a comparative analysis of middle-aged adults aged 40-74 years participating in the National Health and Nutrition Examination Study (NHANES III) 1988-1994 and 2001-2006 surveys, the proportion of adults who adhered to all five healthy habits (at least five fruits and vegetables/day, regular exercise >12 times/month, maintaining a BMI between 18.5 and 29.9 kg/m^2, moderate alcohol consumption, and not smoking) decreased from 15% to 8% [5,6]. Additional available data suggest that 31% of the world's population is not meeting the minimum recommendations for physical activity [7].

The number of behavior change interventions that have become available through the Internet has increased steadily over the last decade [8]. Evidence exists to support the effectiveness of Web-based interventions in changing behaviors. Research has shown these interventions to be effective in different areas of health care [9-14]. Among them are promoting physical activity [15], nonsmoking behavior [10], and weight loss [16]. However, the effects are usually small [17]; the actual reach of Internet-delivered interventions seems lower than expected and attrition rates are generally high [10]. To the best of our knowledge, there have not been any randomized controlled trials (RCTs) of a Web/mobile phone app as a healthy lifestyle intervention in itself that focuses on education and self-monitoring of diet and physical activity for the healthy population.

A trial of this type is important because Web/mobile phone apps are readily available to the public and provide an opportunity to develop interventions with lower costs, less burden, and a greater reach, particularly as mobile phone use rates continue to rise. Adding Web/mobile phone technology offers a significant potential for improving the user's self-monitoring experience and adherence to health promotion programs and thus to enjoy their benefits.

In this study, we hypothesized that using a theory-guided, technology-supported, healthy lifestyle–promoting Web-based app will result in high use, improved nutritional knowledge, increased duration of physical activity, and greater weight loss compared to a control approach.

The objective of the study was to compare people receiving a new Web-based app with people who got an introductory lecture on healthy lifestyle alone, on resulting weight change, nutritional knowledge, and physical activity, and to identify predictors of success for maintaining a healthy lifestyle.

Methods

Eligibility and Study Design

People from the community (high-tech companies, kibbutzim, large organizations) were recruited between the end of 2012 and January 2013 from the south and center of Israel, via an email invitation. The invitation described the research goals and informed the participants that their participation was voluntary and involved no incentive payments or gifts.

The study lasted 14 weeks and had two arms: intervention (Web-based app) and control. Eligible participants were healthy people, aged ≥18 years with Web and email experience, interested in a healthy lifestyle, and willing to commit the necessary time and effort to the study. Exclusion criteria were participating in a program for weight reduction, pregnancy, and lactation.

The first face-to-face meeting of the study was the introductory meeting. In this meeting, all participants received a presentation on a healthy lifestyle, recommendations for healthy nutrition, and the clinical benefits of physical activity for the recommended weekly duration. The participants signed an informed consent form and were randomized to intervention and control groups.

The members of the app group received a password for the app and were instructed to use it as much as they wanted to. The members of the control group were instructed to continue living a healthy lifestyle as they understood it and based on the presentation in which they participated. The control group did not receive any other intervention.

Anthropometric measurements of weight and waist were obtained by the researcher in the introductory meeting. Nutrition knowledge, diet quality, and physical activity type and duration questionnaires were completed by participants through the Internet (Google docs). Anthropometric measurements and questionnaires were repeated after 14 weeks (the last face-to-face meeting). The intervention participants also completed a satisfaction questionnaire at the end of the study. In total there were only two face-to-face meetings between the researcher and the study participants during the study period.

Randomization

Following the introductory meeting, and after signing an informed consent form, participants were randomized to the intervention and control groups. The randomization was performed in clusters according to the place of recruitment: kibbutz, company, organization. Participants were numbered using a random number table where the numbers were randomly assigned to intervention and control groups using a 2:1 ratio.

This ratio was used in order to enhance recruitment by increasing the chances of being in the intervention group. This ratio enabled us to recruit more people and to achieve our calculated sample size.

Ethics and Informed Consent

Ethics approval was received by the Hebrew University, Faculty of Agriculture Food and Environment (AGHS-07.12). Each participant provided written informed consent at the beginning of the study. This study is a registered clinical trial (NCT01913496).

Intervention

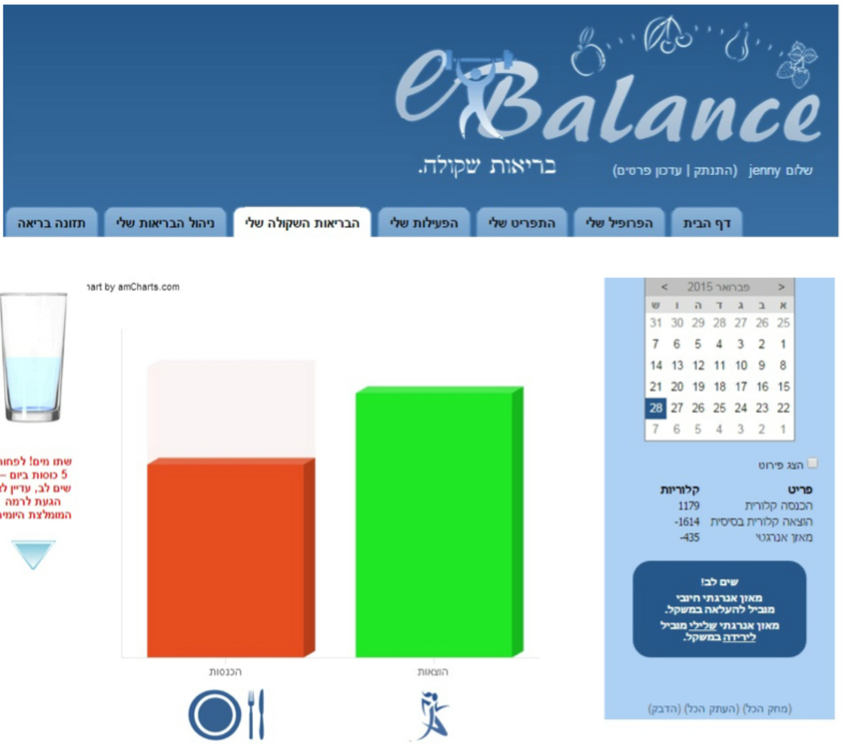
The new Web-based eBalance app was designed for a healthy non-professional audience interested in self-management and achievement of a healthy lifestyle. It was developed over a period of 1 year by one of the authors (JSN) and was programmed by a software expert (Figures 1 and 2). The app is based on guidelines published by the Israel Ministry of Health [18] and the Dietary Guidelines for Americans 2010 [19] recommendations for a Healthy Lifestyle. The theoretical basis of the app and the elements required to keep users engaged was given much consideration. The app was designed to be user-friendly for all types of populations, and specifically for people with only basic computer and Internet skills. If additional explanations are necessary, the user can watch videos with instructions on how to use the app (located in the app itself). eBalance was designed to be a helpful tool for facilitating self-management of a healthy lifestyle. Much thought was invested in enabling the easiest and fastest possible entry of data (food intake and physical activity).

The app enables the users to monitor their dietary intake and physical activity by receiving real-time feedback. Based on the Control Systems Theory (CST) of self-regulation [20], we postulated that the experience of receiving consistent and immediate feedback from eBalance would reinforce adherence to self-monitoring and healthy behavior change.

The essence of the app is based on the flexibility of choosing your own preferred diet within the Dietary Reference Intake (DRI) recommendations that are related to the user. This was achieved through feedback from the app. This feedback includes information on nutrient intake compared with the recommended level of DRI and potential food sources for the nutrients, in addition to monitoring the calorie intake and expenditure.

For example, Figure 1 shows the diet of a user who did not achieve the daily recommendations of several nutrients. In this case, the app can suggest which types of foods the user can add to their menu in order to reach the recommended level.

Figure 1. Screenshot of the eBalance Web-based app.



שתו מים! לפחות
5 כוסות ביום –
שים לב, עדיין לא
הגעת לרמה
המומלצת היומית

פריט	קלוריות
הכנסה קלורית	1179
הוצאה קלורית בסיסית	-1614
מאזן אנרגטי	-435

שים לב!
מאזן אנרגתי חיובי
מוביל להעלאה במשקל.
מאזן אנרגתי שלילי מוביל
לירידה במשקל.

Figure 2. Monitoring key nutrients by the app. Nutrients are divided into three groups: (1) carbohydrates, proteins, fats, (2) minerals, and (3) vitamins (green bar – intake according to the recommendations, red bar – intake below the recommendations).

חלבונים פחמימות ושומנים

חלבונים	פחמימות	אומגה 3	סיבים תזונתיים	שומנים	כולסטרול	שומן רווי
16.30	19.69	0.32	5.04	8.73	20.10	4.12
(מומלץ: 133.16)	(מומלץ: 177.54)	(מומלץ: 1.10)	(מומלץ: 25.00)	(מומלץ: 59.18)	(מומלץ: 300.00)	(מומלץ: 177.54)

מינרלים

אבץ	ברזל	מגנזיום	נתרן	סידן
1.59	1.71	65.49	467.58	273.90
(מומלץ: 8.00)	(מומלץ: 18.00)	(מומלץ: 320.00)	(מומלץ: 1600.00)	(מומלץ: 1000.00)

ויטמינים

ויטמין A	ויטמין B12	ויטמין C	ויטמין E
231.09	0.72	207.11	2.61
(מומלץ: 700.00)	(מומלץ: 2.40)	(מומלץ: 75.00)	(מומלץ: 15.00)

Outcomes Assessment

App Usability

App usability was determined by frequency and convenience of use. Frequency of use was evaluated from the app database according to the reports of the users. Google Analytics was also used as an objective tool for evaluating visit duration and number of pages per visit. Convenience of use was reported by the app users via a satisfaction questionnaire. The satisfaction questionnaire was developed for eBalance based on a questionnaire developed by Shahar et al [21].

Lifestyle Questionnaires

The lifestyle questionnaires were adapted for use in the current study.

Nutrition Knowledge

There are two ways to improve knowledge using the app. The first is by entering the “Healthy Nutrition” page. This page contains information on the importance of the key nutrients for a healthy lifestyle and displays the food sources for these key nutrients. The second is through routine use of the app, which leads to learning about the caloric value of food intake or physical activity expenditure and also about the food’s key components.

Knowledge was assessed using a 43-item online self-report questionnaire based on Parmenter’s general nutrition knowledge questionnaire for adults [22].

Diet Quality

The diet quality questionnaire is a 16-item questionnaire based on Parmenter’s general nutrition knowledge questionnaire for adults [22].

Physical Activity

This questionnaire included 28 questions on the type, frequency, and duration of time each participant spent in physical activity per week. This questionnaire is based on an international physical activity questionnaire [23,24].

Anthropometric Measurements

Weight was measured in kilograms with a portable digital scale (Beurer GmbH & Co. KG, Germany). Weight measurement was performed in light clothes and without shoes to the nearest 0.1 kg. Waist circumference was measured on the navel.

Sample Size Calculation

The sample size for the RCT was calculated based on the findings from a previous pilot study. Cholesterol intake was used for the calculation; the sample size was calculated using P=.05, power=80%. A sample size of 50 participants in the app users group was found to be sufficient for detecting a 10% reduction in level of cholesterol intake at the end of study period compared with the baseline. We added 20% to the number of participants to compensate for potential dropouts.

Statistical Analysis

Analyses were carried out using SPSS 21.0. Descriptive statistics were used to describe and compare baseline characteristics of participants in the app and control groups. Univariate analyses were performed first to compare pre- and post- values using the paired t test. The change from baseline was compared using the t test for independent samples. Categorical measurements were compared using the chi-square test. An analysis of variance (ANOVA) model followed by Scheffe post-hoc tests were used to compare differences between three groups (control and the two app groups, light users and heavy users). This analysis was performed post hoc and is presented separately from the main outcomes in each table and figure.

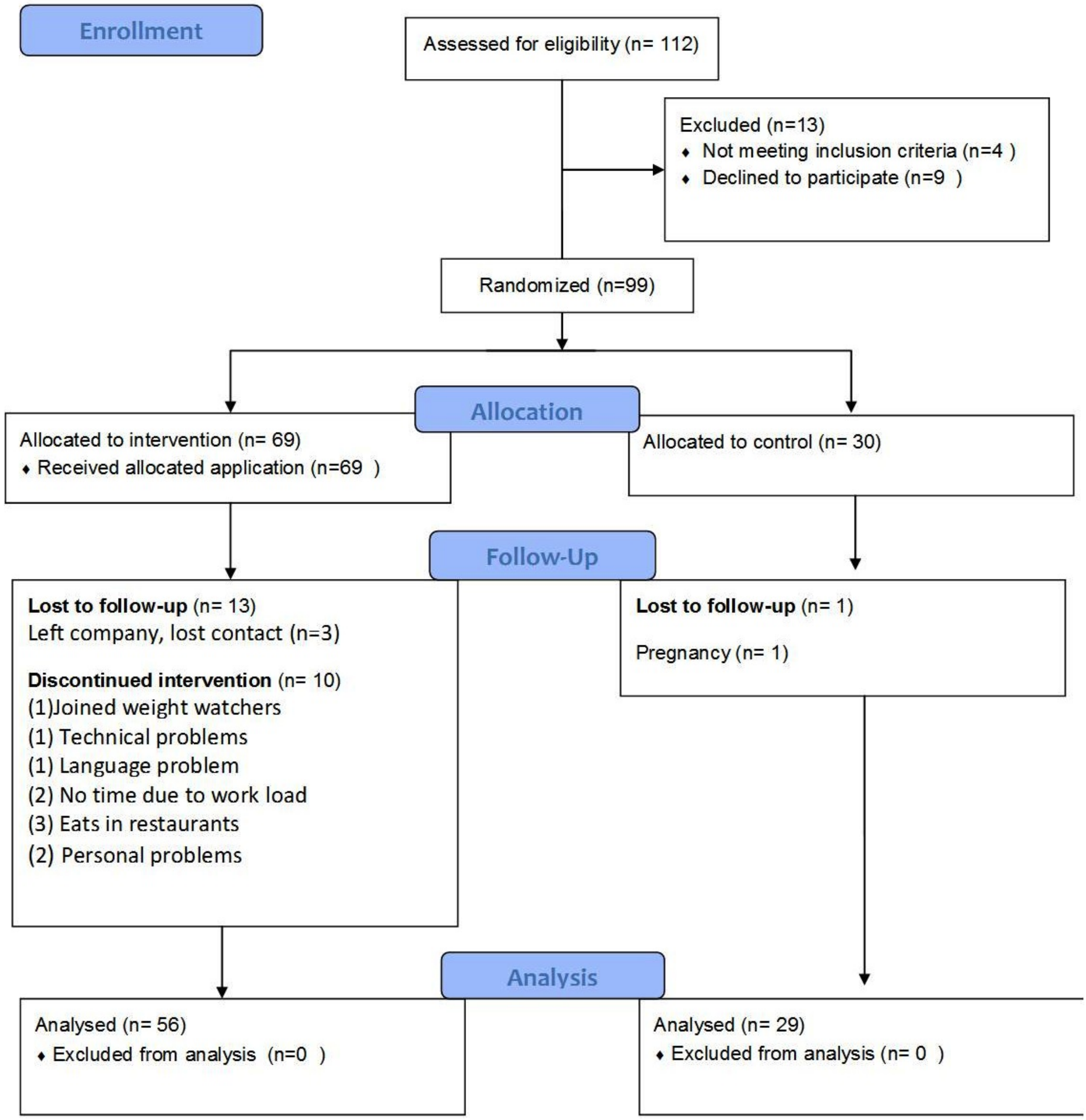
In order to estimate the effect of the intervention on lifestyle, we defined an integrated score (success score) based on the following parameters: decrease or maintain body weight, duration of physical activity equal to or above 150 minutes per week, and scores of at least 70 in knowledge and diet quality questionnaires. Each parameter contributed one point to the score, summing the 4 points for all users. An analysis of covariance (ANCOVA) model was applied to identify the factors that affect the score. Age and gender were included as confounders, frequency of use was entered as the tested factor with three levels: control (no use at all), light users (less than three usage days per week), and heavy users (at least three usage days per week). The model was also tested for app users versus the control group. The statistical significance level was defined as alpha=.05.

Results

Recruitment

Of the 112 who agreed to participate in the study, 99 participants were found eligible and were randomized into two groups (Figure 3). A total of 85 participants completed the follow-up visit (86% retention rate); 56 of the 69 participants who started the study in the app group completed the follow-up (81% retention rate) and 29 of 30 participants who started the study in the control group completed the follow-up (97% retention rate). Three participants left their company and their contact was lost. One joined WeightWatchers, one had technical problems, one had a language problem, two had no time due to work load, three frequently ate in restaurants and did not always know the components of the food, and two had personal problems. In the control group, one participant did not complete the follow-up due to pregnancy. There were no differences in the demographic characteristics at baseline of the app non-completers group and the app users group.

Figure 3. Flow of the participants through the RCT.



Baseline Data

Characteristics of the study participants are presented in Table 1. Our study included 85 adults: 54 (64%) women, 31 (36%) men, with a mean age of 47.9 years (SD 12.3). No differences in age, weight, BMI, height, and waist circumference existed between the app and control groups at baseline.

Table 1. Baseline characteristics of the RCT participants.

Characteristics	All (n=85)	Web-based app (n=56)	Control (n=29)	*P*
Age in years, mean (SD)	47.9 (12.3)	48.5 (11.3)	46.7 (14.2)	.53
Females, n (%)	54 (64)	33 (59)	21 (72)	.25
Weight in kg, mean (SD)	74.0 (15.4)	76.3 (15.1)	69.7 (15.2)	.07
Height in cm, mean (SD)	1.69 (0.1)	1.70 (0.1)	1.67 (0.1)	.09
BMI, mean (SD)	25.8 (4.0)	26.2 (3.9)	25.0 (4.4)	.19
Waist circumference in cm, mean (SD)	93.5 (12.4)	94.3 (11.7)	91.7 (14)	.39

Frequency of Use

The mean frequency of use among the intervention group was 2.7 (SD 1.9) days a week (95% CI 2.2-3.2). The average period of use was 7.8 (SD 4.3) weeks; this is the period that a user stayed connected to the app (the period between the first and the last visit to the app).

Adherence rate of 56% (7.8/14 weeks) was calculated for all app users including light and heavy users. The average frequency of use during the 7.8 weeks was 4.9 days (SD 1.8).

Self-monitoring declined over the study period. At the end of 14 weeks, 27% of the users were still active on the app. According to Google Analytics, the average duration of visits was 7.5 minutes (SD 0.9). The average number of visits per user per day was 1.7 (SD 0.25), and the average number of pages per visit was 6.2 (SD 0.6).

We used the use days as a measure of compliance and divided the intervention group into light and heavy users. Light users were defined as using the app less than 3 days a week (1-40 use days), and heavy users were defined as using the app 3 or more days a week (41-98 use days). The cut-off was set at 40 days since at 40 days (almost half of the study duration) 50% of the participants were still using the app. The light and heavy users as part of the intervention group completed the 98 days (14 weeks) of the study.

The light users group included 30 participants (53% of the app group). The heavy users included 26 users (47% of the app group). At baseline, the mean weight of the light users was 77.9 kg (SD 14.8) with a BMI of 26.3 kg/m^2 (SD 3.7); the mean weight of the heavy users was 74.4 kg (SD 15.5) with a BMI of 26.1 kg/m^2 (SD 4.1). There were no differences at baseline between the light and heavy app users groups and the control group (*P*>.05).

The comparison between heavy and light users was a post-hoc analysis. It was not part of the main objective of the study. The main analysis depicts the comparison between intervention and control groups. The comparison between heavy and light users was added post hoc in order to illuminate the consistent trend of the beneficial effect of the app for heavy users.

Convenience of Use

In terms of convenience, 79% of the users reported that data entry was very easy; 93% reported that use of the app was very easy. Users' satisfaction with the app was high: 93% will make a "moderate" to "very high" recommendation to their friends to use the app. The median of the satisfaction from the app on a scale of 1-10 was 8 (25th and 75th percentiles were 6 and 8.5, respectively). The average score was 7.3 (SD 1.9). The app greatly affected 72.1% of the intervention group to maintain good nutrition, 69.7% reported that the app helped them lose weight, and 76.8% reported that the app encouraged them to increase their physical activity.

Nutrition Knowledge

Table 2 shows that the score on the knowledge questionnaire did not change significantly (*P*=.16) at the end of the study in the control group but did increase significantly in the app group, from 76 (SD 7.5) to 79 (SD 8.7) at the end of the study (*P*=.04). The heavy users improved their score in the questionnaire significantly from 76 (SD 6.4) at baseline to 79 (SD 7.6) at the end of the study (*P*=.03), contrary to the light users who did not improve their knowledge significantly (*P*=.47; Table 2).

Table 2. Knowledge score of control and app groups and light/heavy users.

	Mean score				*P*
	Baseline	SD	End of study	SD	
Group					
Control (n=29)	71	9.5	73	8.1	.16
Web-based app (n=56)	76	7.5	79	8.7	.04
Web-based app users					
Light users (n=30)	76	8.8	78	9.9	.47
Heavy users (n=26)	76	6.4	79	7.6	.03

Physical Activity

Figure 4 presents the results of the analysis of the physical activity questionnaire. The analysis revealed that the app users significantly increased their weekly duration of physical activity compared with the control group, who decreased their weekly duration of physical activity. The mean change in the weekly duration of physical activity was 63 (SE 20.8) minutes in the app group and -30 (SE 27.5) minutes in the control group (*P*=.02; Figure 4).

Table 3 shows the percentage of participants who reached the healthy threshold of 150 minutes as recommended by the 2008 Physical Activity Guidelines for Americans [25]. At baseline, no significant difference between the control and the app groups (*P*=.42) was shown in the percentage of participants engaging in physical activity of different weekly durations.

However, at the end of the study there was a significant difference between the control group and app users. The app motivates the users to significantly increase the weekly duration of physical activity. More users increased their weekly duration of physical activity to the healthy range of more than 150 minutes a week, which may afford substantial health benefits.

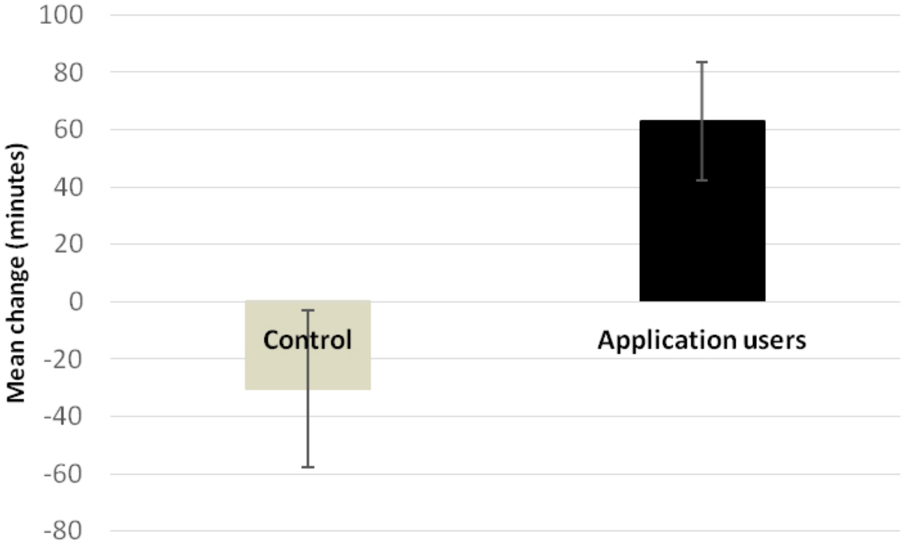
Table 3. Percentage of participants engaging physical activity at different weekly durations.

	Study groups[a]				Web-based app users[b]			
	Control (n=29)		App users (n=56)		Light users (n=30)		Heavy users (n=26)	
	Baseline	End of study	Baseline	End of study	Baseline	End of study	Baseline	End of study
Low activity (<150 min.)	45.0%	55.0%	28.8%	17.3%	26.9%	23.1%	30.8%	11.5%
Recommended activity (150-300 min.)	30.0%	25.0%	36.5%	32.7%	42.3%	38.5%	30.8%	26.9%
High activity (>300 min.)	25.0%	20.0%	34.7%	50.0%	30.8%	38.4%	38.4%	61.6%

[a] *P* value for difference between study groups. Baseline, *P*=.42; end of study, *P*=.01.

[b] *P* value for difference between light and heavy users. Baseline, *P*=.64; end of study, *P*=.009.

Figure 4. Mean change of physical activity duration in the control and app groups (P=.02).



Diet Quality

Table 4 shows that there was no difference in the score on the diet quality questionnaire in the control group (*P*=.41). In contradistinction to the control group, the app users improved their score significantly at the end of the study from 67 (SD 9.8) to 71 (SD 7.6; *P*<.001; Table 4). No significant difference was found between light and heavy users. Both subgroups improved their quality of diet score similarly. Based on analysis of the diet quality questionnaire, the app users improved their diet quality more than the control group did.

Table 4. Diet quality score at baseline and at the end of the RCT.

	Mean score				*P*
	Baseline	SD	End of study	SD	
Group					
Control (n=29)	61	8.9	62	10.1	.41
Web-based app (n=56)	67	9.8	71	7.6	<.001
Web-based app users					
Light users (n=30)	66	10.3	71	7.1	.017
Heavy users (n=26)	67	9.6	72	8.2	.001

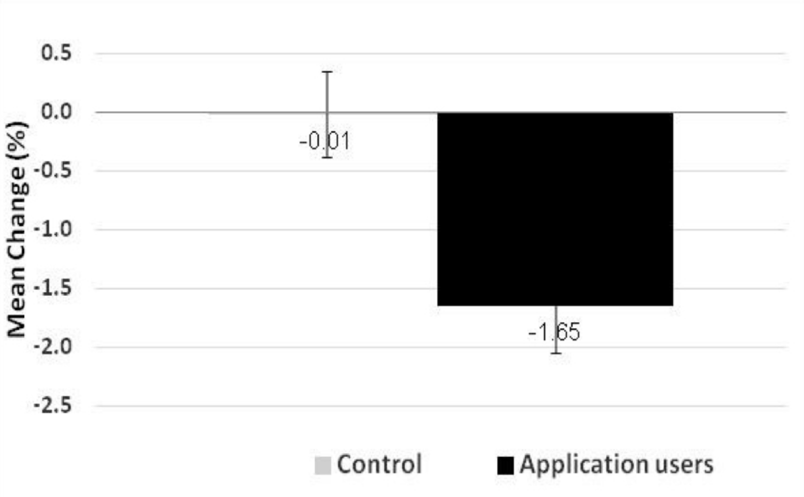
Weight and Waist Circumference

The app users lost more weight compared to the control group. At the end of the study, the mean weight and BMI change of the app group was significantly greater than that of the control group. The mean weight change of the app group was -1.44 kg (SE 0.40, 95% CI 0.1-2.5) versus -0.13 kg (SE 0.36) in the control group (*P*=.03). The mean BMI change was -0.48 kg/m^2 (SE 0.13, 95% CI 0.05-0.84) in the app group, but only -0.03

kg/m^2 (SE 0.12) in the control group (*P*=.03). The mean weight change presented in Figure 5 was -1.65% (SE 0.4) of the initial body weight in the app group, versus only 0.01% (SE 0.5) in the control group (*P*=.003). The mean change in weight and BMI in the heavy users was statistically significant, as opposed to the light users. Mean change in weight in heavy users was -2.1 kg (SE 0.64, *P*=.01) and in BMI -0.7 kg/m^2 (SE 0.21, *P*=.002). Mean change in weight in the light users group was -0.84 kg (SE 0.47, *P*=.08) and BMI -0.3 kg/m^2 (SE 0.14, *P*=.07).

No statistically significant difference was found in waist circumference between the app and control groups. At the end of study, the mean change in waist circumference was -2.31 cm (SE 0.56) in the app group and -0.9 cm (SE 0.58) in the control group (*P*=.09).

Figure 5. Mean change in weight (% of body weight) P=.01.



Success Predictors

Success was defined as follows: reducing or maintaining the original weight, duration of physical activity equal to or more than 150 minutes per week, a score of more than 70 points on the nutrition and quality of diet questionnaires. The highest mean success score normalized to a range of 100% was of the app group’s 68%; 36% was the success score of the control group (*P*<.001). The heavy users received a higher mean success score (82%) than the light users (61%; *P*=.02). We did not find a difference in gender (*P*=.42) between heavy and light users. Table 5 presents the results of the linear regression model. The explained variance was 41% for the total variable parameters, and includes age, gender, weight at baseline, and frequency of app use. The key parameter for predicting success was the frequency of app use. The app frequency of use was significantly related to a higher success score (*P*<.001). The weight at baseline did not predict success, whereas being a woman did predict success.

Table 5. Linear regression model presenting success predictors.

Dependent variable	R^2=explained variance, %	*P* value
Age	1	.28
Weight at baseline	1	.2
Frequency of use	32	.01
Gender (female)	4	.03
Total of model variability	41	

Discussion

Principal Findings

In this study, we describe the development of a Web-based app for promoting healthy lifestyle and its performance evaluation using a RCT.

The app was designed as a theory-guided, technology-supported lifestyle-promoting program integrating elements pointed out in studies as important for supporting lifestyle changes. Its strengths may stem from its individually tailored approach with personal goal setting and feedback. Understanding the psychological determinants of a self-change process of altering behaviors is needed for designing effective intervention programs. The app contains features to enhance usability, feedback tailored to the user according to the DRI recommendations, a large Ministry of Health–food database, and a physical activity database.

This is one of the first studies that tested the impact of a Web-based app on lifestyle using established epidemiological methodology. The findings of the RCT indicate that the app is acceptable, effective, and can serve as a supporting tool in promoting a healthy lifestyle. These findings are strengthened by the fact that heavy users achieved better results than light users, suggesting a dose-effect relationship.

Compliance

High compliance is a key component in the applicability of health promotion methods. The compliance to the intervention in our study was high as can be seen by the low attrition rate of 19%. This finding contrasts with an RCT by Carter et al [26], which estimated adherence to a smartphone app for weight loss compared with a website and paper diary with an attrition rate of 38%. Attrition was not distributed equally among the groups, with higher attrition in the diary and Web groups compared to the smartphone group. A review focusing specifically on Web-based interventions for weight loss found that most interventions had attrition rates greater than 20% [8]. In a systematic review of long-term weight loss trials in obese adults, attrition ranged from 30% to 60% [27]. In our study, various methods and strategies were used to support adherence to the intervention. These strategies included immediate user-tailored feedback as text messages and a graphic view of the user's nutrition and physical activity. It seems that these strategies were effective at least in the short term in improving compliance and reducing attrition rate.

Frequency of Use

The average frequency of the app use was found to be 2.7 (SD 1.9) days a week, and visit duration was evaluated as 8 minutes by Google Analytics. Duration of about 10 minutes per visit is sufficient to not take too much time for introducing food and physical activity, but still learn from the feedback received by the app. Brouwer [8] reported visit durations to be from less than 10-20 minutes.

In a recent literature review, Vandelanotte et al [10] reported that better outcome measures on improvement of physical activity were identified when participants visited the intervention website more than five times. However, other studies reported that only a minority of participants visited the intervention website more than once [8,28].

The frequency of self-monitoring required for successful outcomes was not clearly answered in the literature. In our study, frequency of use of 3 times per week had an added value in promoting healthy lifestyle as opposed to a lower frequency of use. Heavy users showed a trend of better outcomes and better orientation to a healthy lifestyle compared with light users.

Adherence

The percentage of participants who adhered to the app was 56%, which is higher than described in other studies. However, it is in line with the high frequency of app use.

In the review by Kelders [29], the adherence rate of 83 interventions was 50%, and in lifestyle Web-based interventions it was only 33%. The adherence of a Web-based family intervention for overweight children was reported to be 51.1% in a 4-week study [30]. As expected, a gradual decline in the use of the app was detected. The gradual decline in adherence that was seen during the study is typical for Web-based interventions. Studies reported that 25% of the sample continues to self-monitor at the end of study [31,32]. However, some of them added in-person support. In our study, which did not include any personal support, 27% (33) of users used the app for the entire duration of the study (14 weeks). In the study by Carter et al studying adherence to a smartphone intervention [26], 12% of Web users (5 users) and 19% of smartphone users (8 users) reported for 3 months. We can conclude at this point that the adherence to the app was reasonable in our study, although increased adherence would probably improve outcomes even more.

The high rate of adherence in this study is also attributable to convenience of use, among other factors. The design of the app was oriented to enhance convenience of use. For example, the app permits one to save frequently eaten meals and thus eliminate the need for repeated searching and entry. The high percentage of users evaluating the app as easy to use is very encouraging.

Weight and Physical Activity Change

We showed that the mean change in weekly duration of physical activity in the app group increased in our study, with more users progressing toward the recommended 150 minutes of moderate activity per week. We also found a decrease in the weekly duration of physical activity in the control group; this decrease in the control group is an interesting finding. Our expectation was that no change in any baseline measurements would occur in the control group. Some RCTs report a modest beneficial effect for the participants in control groups due to contamination. It may emphasize the importance of having a tool to encourage behavior change on a daily basis. Continuous monitoring by the app provides immediate feedback on the daily and weekly duration of physical activity and motivates for a behavior change. Without this tool (control group) physical activity duration is not consistent and depends on external factors or on the perception of the participant that the duration of physical activity is sufficient; however, reality indicates differently.

In a pilot study of smartphone-assisted behavioral obesity treatment [33], participants reported engaging in physical activity an average of 125.1 (SE 10.8) minutes per week during the first 12 weeks, and an average of 140.7 (SE 12.3) minutes per week during the second 12 weeks.

Participation in 150 minutes of moderate physical activity each week (or equivalent) is estimated to reduce the risk for ischemic heart disease by approximately 30%, the risk for diabetes by 27%, and the risk for breast and colon cancer by 21-25% [2]. Furthermore, physical activity lowers the risk for stroke, hypertension, and depression. It is a key determinant of energy expenditure and thus fundamental to energy balance and short- and long-term weight control [34]. An increased level of physical activity/exercise, regardless of body weight or weight loss, increases health [35]. Exercise offers a way to mitigate the health-damaging effects of obesity, even without weight loss [35,36]. Our findings show that users of the app increased the duration of physical activity, thus moving towards a healthier lifestyle. A larger percentage of heavy users increased the duration of physical activity to 150 minutes and above, than light users. Jing Wang [37] also reported that physical activity may increase if recording the number of minutes engaged in physical activity becomes an integral part of the self-monitoring regimen.

It is important to note that although the goal of the intervention was weight maintenance, the app users lost significantly more weight than the control group and also gained less weight, which highlights the impact of the app. Our results are consistent with studies that show a significant positive relationship between self-monitoring diet, physical activity or weight, and successful outcomes related to weight management [32].

Moreover, the same trend that more frequent app use leads to better outcomes that was seen in other comparisons was also seen regarding change in weight. Heavy users lost more weight than light users. These results are in line with other studies that claim that more frequent self-monitoring was consistently and significantly associated with weight loss compared to less frequent self-monitoring [38-40].

Nutrition Knowledge and Diet Quality

We found that the app group received a higher score in the nutrition knowledge and diet quality questionnaires compared to the control group. Although the difference in the score between baseline and follow-up was small, it captures an increase in knowledge regarding the nutritional recommendations. This relatively small change can also be attributed to selection bias due to the increased awareness of our participants of healthy lifestyle. However, this relatively small change in knowledge score is in line with a systematic review on the impact of computer-mediated nutrition education programs for adolescents who also showed a small change in nutrition knowledge [41].

However, even promotion of small changes in eating and healthful eating can be used as preventive strategies for adults [40]. According to the World Health Organization, adequate consumption of fruits and vegetables reduces the risk for cardiovascular diseases, stomach cancer, and colorectal cancer; 1.7 million deaths (2.8%) worldwide are attributable to low fruit and vegetable consumption. Changes in specific eating patterns, such as reducing food rich in saturated fat or trans fat, can lead to a healthier diet and weight. High consumption of saturated fats and trans fatty acids is linked to heart disease. Heavy users improved their knowledge in contrast to the light users.

As the participants of the study were already inclined to a healthy lifestyle, the app was the trigger to implement what they had already known. Therefore, no difference between heavy and light users regarding diet quality was seen. However, in order to increase knowledge, the more one uses the app, the more knowledge one gains.

Success Predictor

The success score estimates the effect of the intervention on lifestyle by integrating all measured parameters, weight at baseline, duration of physical activity, score of nutrition knowledge, and quality of diet questionnaires, into a combined score. Frequency of the app use was found to be the key factor affecting the success score. This finding supported our hypothesis that the frequency of the app use is the key factor for succeeding in maintaining or reducing weight, increasing the duration of physical activity, and improving knowledge on nutrition and diet quality.

Since studies on Web/mobile phone apps are very limited, this analysis was performed only on our app and we could not compare our findings to other studies.

Strengths and Limitations

The app was evaluated in a randomized controlled trial. Most of the study was performed online (ie, the questionnaires and the app), with little contact between the participants and the researcher, which was limited to two meetings at the beginning and end of the study. Our findings were achieved even though there was no payment or any incentives given to the participants and face-to-face contact was limited to two meetings. Objective (Google Analytics) tools were used to evaluate the usability of the app.

However, the sample comprised predominantly adult, well-educated, white females. Therefore, the results may not be generalizable to some ethnic minority populations, adolescents, males, or less educated persons. The two study groups were unbalanced in size. The number of participants in the study group was almost twice the number in the control group. This was because it was difficult to recruit participants for the control group. The total sample was small and included 85 participants. The duration of the study was only 14 weeks. A longer duration of 12 months or more and including various biomarkers along the way is recommended for future studies. Our findings can be used as an initial 3-month testing of our intervention strategy.

Conclusions

Analysis of the RCT results indicate that the newly developed app is effective, as reflected in high usability, increased knowledge in nutrition, increased duration of physical activity, and reduction in weight that were found in the app group only.

eBalance differs from commercially available apps in that it has taken an evidence-based approach and was guided by feedback from potential system users, contains features to enhance usability, feedback tailored to the user according to the DRI recommendations, a large Ministry of Health–food database, and physical activities database.

Testing the performance of the app in an RCT is unique and highly important for selecting the appropriate tool for public health. Evaluation of the app’s efficacy was performed by objective scientific tools. These results are promising in the app’s potential to promote healthy lifestyle, although larger and longer duration studies are needed to achieve more definitive conclusions.

Acknowledgments

The publication fee was partially supported by JMIR Publications.

Conflicts of Interest

None declared.

References

1. American College of Preventive Medicine. Lifestyle medicine: evidence review. 2009 Jun 30. URL: http://c.ymcdn.com/sites/www.acpm.org/resource/resmgr/lmi-files/lifestylemedicine-literature.pdf [accessed 2015-02-21] [WebCite Cache ID 6WVhbNTcm]
2. World Health Organization. Global health risks: mortality and burden of disease attributable to selected major risks. Geneva; 2009. URL: http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf [accessed 2015-02-21] [WebCite Cache ID 6WVhm5P2k]
3. Lianov L, Johnson M. Physician competencies for prescribing lifestyle medicine. JAMA 2010 Jul 14;304(2):202-203. [doi: 10.1001/jama.2010.903] [Medline: 20628134]
4. Ford ES, Bergmann MM, Kröger J, Schienkiewitz A, Weikert C, Boeing H. Healthy living is the best revenge: findings from the European Prospective Investigation Into Cancer and Nutrition-Potsdam study. Arch Intern Med 2009 Aug 10;169(15):1355-1362. [doi: 10.1001/archinternmed.2009.237] [Medline: 19667296]
5. Dall TM, Fulgoni VL, Zhang Y, Reimers KJ, Packard PT, Astwood JD. Potential health benefits and medical cost savings from calorie, sodium, and saturated fat reductions in the American diet. Am J Health Promot 2009;23(6):412-422. [doi: 10.4278/ajhp.080930-QUAN-226] [Medline: 19601481]
6. King DE, Mainous AG, Carnemolla M, Everett CJ. Adherence to healthy lifestyle habits in US adults, 1988–2006. American Journal of Medicine 2009 Jun;122(6):528-534. [doi: 10.1016/j.amjmed.2008.11.013] [Medline: http://www.ncbi.nlm.nih.gov/pubmed/?term=adherence+to+healthy+lifestyle+habits+in+us+adults%252c+1988%25e2%2580%25932006.+american+journal+of+medicine]
7. Kohl HW, Cora LC, Estelle VL, Shigeru I, Jasem RA, Grit L, for the Lancet Physical Activity Series Working Group. The pandemic of physical inactivity: Global action for public health. Lancet 2012;380(9838):294-305. [doi: 10.1016/S0140-6736(12)60898-8] [Medline: http://www.ncbi.nlm.nih.gov/pubmed/?term=the+pandemic+of+physical+inactivity%253a+global+action+for+public+health.+lancet%252c+2012]
8. Brouwer W, Kroeze W, Crutzen R, de NJ, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
9. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
10. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
11. 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]
12. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
13. Cuijpers P, van SA, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med 2008 Apr;31(2):169-177 [FREE Full text] [doi: 10.1007/s10865-007-9144-1] [Medline: 18165893]
14. Barak A, Hen L, Boniel-Nissim M, Shapira N. A Comprehensive Review and a Meta-Analysis of the Effectiveness of Internet-Based Psychotherapeutic Interventions. Journal of Technology in Human Services 2008;26(2/4):109-160 [FREE Full text] [Medline: http://construct.haifa.ac.il/~azy/B401-MetaAnalysisInterventionsBarak.pdf]
15. van den Berg MH, Schoones JW, Vliet Vlieland TPM. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
16. Weinstein PK. A review of weight loss programs delivered via the Internet. J Cardiovasc Nurs 2006;21(4):251-8; quiz 259. [Medline: 16823276]
17. Anderson-Bill ES, Winett RA, Wojcik JR, Winett SG. Web-based guide to health: relationship of theoretical variables to change in physical activity, nutrition and weight at 16-months. J Med Internet Res 2011;13(1):e27 [FREE Full text] [doi: 10.2196/jmir.1614] [Medline: 21447470]
18. Israeli Ministry of Health. URL: http://www.health.gov.il/UnitsOffice/HD/PH/nutrition/Pages/default.aspx [accessed 2015-02-21] [WebCite Cache ID 6WVn1FlPT]
19. U.S. Department of Agriculture. health.gov. 2010. Dietary Guidelines for Americans 2010 URL: http://www.health.gov/dietaryguidelines/dga2010/DietaryGuidelines2010.pdf [accessed 2015-02-22] [WebCite Cache ID 6WWggXJe9]

20. Carver C, Scheier MF. On the Self-Regulation of Behavior. Cambridge: Cambridge University Press; 1998.
21. Shahar DR, Henkin Y, Rozen GS, Adler D, Levy O, Safra C, et al. A controlled intervention study of changing health-providers' attitudes toward personal lifestyle habits and health-promotion skills. Nutrition 2009 May;25(5):532-539. [doi: 10.1016/j.nut.2008.11.020] [Medline: 19230614]
22. Parmenter K, Wardle J. Development of a general nutrition knowledge questionnaire for adults. Eur J Clin Nutr 1999 Apr;53(4):298-308. [Medline: 10334656]
23. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 2000 Sep;32(9 Suppl):S498-S504. [Medline: 10993420]
24. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
25. Physical Activity Guidelines for Americans. 2008. URL: http://www.health.gov/paguidelines/pdf/usersguide.pdf [accessed 2015-02-22] [WebCite Cache ID 6WWhtXffF]
26. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
27. Douketis JD, Macie C, Thabane L, Williamson DF. Systematic review of long-term weight loss studies in obese adults: clinical significance and applicability to clinical practice. Int J Obes (Lond) 2005 Oct;29(10):1153-1167. [doi: 10.1038/sj.ijo.0802982] [Medline: 15997250]
28. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
29. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen Julia E W C. 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]
30. Delamater AM, Pulgaron ER, Rarback S, Hernandez J, Carrillo A, Christiansen S, et al. Web-based family intervention for overweight children: a pilot study. Child Obes 2013 Feb;9(1):57-63 [FREE Full text] [doi: 10.1089/chi.2011.0126] [Medline: 23308372]
31. Acharya SD, Elci OU, Sereika SM, Music E, Styn MA, Turk MW, et al. Adherence to a behavioral weight loss treatment program enhances weight loss and improvements in biomarkers. Patient Prefer Adherence 2009;3:151-160 [FREE Full text] [Medline: 19936157]
32. Burke LE, Styn MA, Steenkiste AR, Music E, Warziski M, Choo J. A randomized clinical trial testing treatment preference and two dietary options in behavioral weight management: preliminary results of the impact of diet at 6 months--PREFER study. Obesity (Silver Spring) 2006 Nov;14(11):2007-2017. [doi: 10.1038/oby.2006.235] [Medline: 17135618]
33. Thomas JG, Wing RR. Health-e-call, a smartphone-assisted behavioral obesity treatment: pilot study. JMIR Mhealth Uhealth 2013;1(1):e3 [FREE Full text] [doi: 10.2196/mhealth.2164] [Medline: 25100672]
34. Riboli E, Norat T. Epidemiologic evidence of the protective effect of fruit and vegetables on cancer risk. Am J Clin Nutr 2003 Sep;78(3 Suppl):559S-569S [FREE Full text] [Medline: 12936950]
35. Carroll S, Dudfield M. What is the relationship between exercise and metabolic abnormalities? A review of the metabolic syndrome. Sports Med 2004;34(6):371-418. [Medline: 15157122]
36. Wu T, Gao X, Chen M, van Dam RM. Long-term effectiveness of diet-plus-exercise interventions vs. diet-only interventions for weight loss: a meta-analysis. Obes Rev 2009 May;10(3):313-323. [doi: 10.1111/j.1467-789X.2008.00547.x] [Medline: 19175510]
37. Wang J, Sereika SM, Chasens ER, Ewing LJ, Matthews JT, Burke LE. Effect of adherence to self-monitoring of diet and physical activity on weight loss in a technology-supported behavioral intervention. Patient Prefer Adherence 2012;6:221-226 [FREE Full text] [doi: 10.2147/PPA.S28889] [Medline: 22536058]
38. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-321. [doi: 10.1111/j.1467-789X.2009.00646.x] [Medline: 19754633]
39. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
40. Hill JO, Wyatt HR, Reed GW, Peters JC. Obesity and the environment: where do we go from here? Science 2003 Feb 7;299(5608):853-855. [doi: 10.1126/science.1079857] [Medline: 12574618]
41. Ajie WN, Chapman-Novakofski K. Impact of Computer-Mediated, Obesity-Related Nutrition Education Interventions for Adolescents: A Systematic Review. Journal of Adolescent Health 2014;54(6):631-645. [doi: 10.1016/j.jadohealth.2013.12.019]

Abbreviations

ANOVA: analysis of variance

ANCOVA: analysis of covariance
DRI: Dietary Reference Intake
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 07.07.14; peer-reviewed by R Pojednic, P Janwantanakul; comments to author 23.09.14; revised version received 19.11.14; accepted 06.12.14; published 02.03.15.

Please cite as:
Safran Naimark J, Madar Z, R Shahar D
The Impact of a Web-Based App (eBalance) in Promoting Healthy Lifestyles: Randomized Controlled Trial
J Med Internet Res 2015;17(3):e56
URL: http://www.jmir.org/2015/3/e56/
doi:10.2196/jmir.3682
PMID:25732936

Original Paper

Effectiveness of a Web-Based Tailored Interactive Health Communication Application for Patients With Type 2 Diabetes or Chronic Low Back Pain: Randomized Controlled Trial

Nina Weymann[1*], GradDipPsych; Jörg Dirmaier[1*], PhD; Alessa von Wolff[1], PhD; Levente Kriston[1], PhD; Martin Härter[1], PhD

Department of Medical Psychology, Center for Psychosocial Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[*]these authors contributed equally

Corresponding Author:
Jörg Dirmaier, PhD
Department of Medical Psychology
Center for Psychosocial Medicine
University Medical Center Hamburg-Eppendorf
Martinistr. 52
Hamburg, 20246
Germany
Phone: 49 40 7410 59137
Fax: 49 40 7410 54940
Email: dirmaier@uke.de

Abstract

Background: The prevalence of chronic diseases such as type 2 diabetes and chronic low back pain is rising. Patient empowerment is a key strategy in the management of chronic diseases. Patient empowerment can be fostered by Web-based interactive health communication applications (IHCAs) that combine health information with decision support, social support, and/or behavioral change support. Tailoring the content and tone of IHCAs to the needs of individual patients might improve their effectiveness.

Objective: The main objective was to test the effectiveness of a Web-based, tailored, fully automated IHCA for patients with type 2 diabetes or chronic low back pain against a standard website with identical content without tailoring (control condition) on patients' knowledge and empowerment.

Methods: We performed a blinded randomized trial with a parallel design. In the intervention group, the content was delivered in dialogue form, tailored to relevant patient characteristics. In the control group, the sections of the text were presented in a content tree without any tailoring. Participants were recruited online and offline and were blinded to their group assignments. Measurements were taken at baseline (t_0), directly after the first visit (t_1), and at 3-month follow-up (t_2). The primary hypothesis was that the tailored IHCA would have larger effects on knowledge and patient empowerment (primary outcomes) than the control website. The secondary outcomes were decisional conflict and preparation for decision making. All measurements were conducted by online self-report questionnaires. Intention-to-treat (ITT) and available cases (AC) analyses were performed for all outcomes.

Results: A total of 561 users agreed to participate in the study. Of these, 179 (31.9%) had type 2 diabetes and 382 (68.1%) had chronic low back pain. Usage was significantly higher in the tailored system (mean 51.2 minutes) than in the control system (mean 37.6 minutes; *P*<.001). Three months after system use, 52.4% of the sample was retained. There was no significant intervention effect in the ITT analysis. In the AC analysis, participants using the tailored system displayed significantly more knowledge at t_1 (*P*=.02) and more emotional well-being (subscale of empowerment) at t_2 (*P*=.009). The estimated mean difference between the groups was 3.9 (95% CI 0.5-7.3) points for knowledge and 25.4 (95% CI 6.3-44.5) points for emotional well-being on a 0-100 points scale.

Conclusions: The primary analysis did not support the study hypothesis. However, content tailoring and interactivity may increase knowledge and reduce health-related negative effects in persons who use IHCAs. There were no main effects of the intervention on other dimensions of patient empowerment or decision-related outcomes. This might be due to our tailored IHCA being, at its core, an educational intervention offering health information in a personalized, empathic fashion that merely additionally provides decision support. Tailoring and interactivity may not make a difference with regard to these outcomes.

Trial Registration: International Clinical Trials Registry: DRKS00003322; http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00003322 (Archived by WebCite at http://www.webcitation.org/6WPO0lJwE).

(J Med Internet Res 2015;17(3):e53) doi:10.2196/jmir.3904

KEYWORDS

Type 2 diabetes mellitus; back pain; randomized controlled trial; health communication; Internet

Introduction

Long-term conditions such as type 2 diabetes (T2D) and chronic low back pain (CLBP) are chronic diseases with high and still rising prevalence [1,2], which cause a significant burden on individuals as well as negative social and economic effects [3-8]. Thus, there is a strong need for cost-effective ways to improve the care of these long-term conditions.

To improve care of long-term conditions, patients, practitioners, scientists, and politicians have called for a greater empowerment of patients in the management of their chronic diseases [9]. Patient empowerment can be observed as a motivational construct reflecting the ability to positively influence self-management and health behavior. The main aspects of patient empowerment are knowledge of the disease, its course and treatment options, the ability to be involved in making medical decisions and relate to health care providers [10], and to manage one’s health behavior and treatment regimens [11,12]. Schulz and Nakamoto additionally stressed that these factors must be accompanied by a volitional component to better predict changes in individuals’ behavior [13]. The most popular definition of patient empowerment is probably that of Funnell et al [14] who defined patient empowerment as “the discovery and development of one’s inherent capacity to be responsible for one’s own life. People are empowered when they have sufficient knowledge to make rational decisions, sufficient control, and resources to implement their decisions, and sufficient experience to evaluate the effectiveness of their decisions”. Patient empowerment and health-related knowledge can be considered as predictors of improved self-management and health outcomes [15,16].

In times of rapidly growing Internet adoption, the Web holds the opportunity to deliver health information [17] and self-management support [18] to large numbers of participants at a comparatively low cost and at the preferred time, place, and learning speed of the individuals. Existing systematic reviews and meta-analyses of Internet interventions in somatic diseases aimed at improving lifestyles (smoking, alcohol consumption, diet, physical exercise) show promising effects on either health- or cost-related outcome measures [19]. More specifically, recent reviews and studies on Internet interventions for adults with T2D [20,21] and CLBP [22-24] also found effects on knowledge, self-efficacy, health behavioral changes, and clinical outcomes. Evidence for Internet interventions can also be found with regard to effects on more proximal outcomes such as patient empowerment [22,25,26] or specific antecedents and mediators of patient empowerment [27].

A specific application of Internet interventions combines health information with at least one other type of support, for example, social support, decision support, or behavior change support: interactive health communication applications (IHCAs). These Internet interventions are expected to improve the knowledge, involvement in decision making, motivation, and self-efficacy of users, resulting in enhanced patient empowerment [28]. This improved empowerment can then enable users to initiate changes in health behaviors, which might result in improved clinical outcomes [28,29]. A Cochrane review found that IHCAs could have positive effects on knowledge, self-efficacy, and behavioral and clinical outcomes. However, the authors demanded more evidence regarding the most suitable application and delivery approaches of IHCAs and the effects of IHCAs for different chronic diseases [28].

Still, the effectiveness of those online applications is limited by high attrition rates [30,31], and users often visit a health intervention website only once [32-34]. A major body of evidence suggests that the effect of online interventions increases with the dose (longer stays, repeated website visits, total contact hours) [35,36], and the effectiveness is maximized if patients intensively work with the information offered [37-39] and return for repeated visits [40,41].

Computer tailoring strategies such as the individualization and personalization of information, as well as an interactive presentation, have been found to effectively increase the exposure to [42] and effectiveness of Web-delivered interventions [43,44]. However, these previous studies predominantly focused on tailoring in health behavior change interventions, with great variability in how the tailoring was carried out. In addition to the question of which elements of the intervention work, one remaining challenge of research with regard to Internet interventions is finding out which delivery methods (interactivity, tailoring, individualization) are effective [45]. Therefore, evidence is especially needed with regard to disease-specific tailoring and individualization strategies in IHCAs for T2D and CLBP, focusing on more proximal outcomes such as health-related empowerment and knowledge.

In this randomized controlled trial, we compared a tailored IHCA presenting information on T2D and CLBP, self-management education, and decision support to a website presenting the same information in a content tree without tailoring. The primary hypothesis was that the tailored and individualized delivery format has a greater effect on knowledge and patient empowerment than the control website. The secondary hypothesis was that users, when facing a health decision, experience less decisional conflict and feel better prepared for the consultation after using the tailored rather than the control website. This paper reports on the trial using the two guidelines that were published in 2011 on designing and reporting Internet intervention research [18,46].

Methods

Study Design

We performed a blinded two-armed randomized controlled trial with a parallel design. Measurements were scheduled immediately before the first use of the system (t_1), immediately after use (t_2), and at 3-months follow-up (t_3). Knowledge (primary outcome) and decisional conflict and preparation for decision making (secondary outcomes) were assessed immediately after the first visit. Patient empowerment (primary outcome) was assessed 3 months after the first visit. All measurements were online self-assessment questionnaires. The study design and procedures have been published in two study protocols [47,48]. There were no important changes to the study design, methods, or trial outcomes after trial commencement. Data collection took place between August 2012 and April 2013.

Study Population

The eligibility criteria were age ≥18 years, access to the Internet, sufficient computer/Internet literacy, and a self-reported diagnosis of T2D or CLBP. CLBP was defined as pain in the lower back almost every day for more than 12 weeks [49].

Recruitment

In general, based on the Cochrane review by Murray et al [28], we expected a small effect (Cohen's *d*=0.2) of the IHCA. Based on the review by van Vugt et al [20] for diabetes, and based on a similar previous study [23] for patients with back pain, we did expect that the tailored intervention would perform better (*d*=0.2) than the control on the primary outcome knowledge for both patient groups. Based on the meta-analysis by Samoocha et al [25], we also expected a small effect with regard to the primary outcome empowerment for patients with T2D and CLBP. To detect a small effect with an alpha of .05 and a power of .80 (one-tailed *t* test), a sample size of 310 (155 per group) was required. Due to the experiences of other Internet trials [40] and the effect of incentives [50], we expected a dropout rate of 20% between registration and immediately after the first visit. Thus, we aimed to include a sample of 414 at baseline. Because we were not aware of differences in the dropout rates between T2D and CLBP patients, we calculated with the same expected dropout rate for both groups.

Recruitment took place using a number of pathways. Two pension funds and six health insurance companies were contacted to request whether they were interested in informing their insurants about the study (eg, via their website, magazine, or newsletter). Three outpatient treatment networks (in which mainly primary care and specialized practices are organized), 15 diabetology practices, 15 practices specialized in CLBP, 87 primary care practices, six rehabilitation centers and hospitals, seven patient associations, and 192 self-help groups were contacted and asked whether they were interested in displaying flyers. Additionally, information on the study and a link to it were disseminated via the mailing list of a population-representative online panel of the University of Münster. Information on the study was also available on the study website. Information and links were placed on the website of the University Medical Centre Hamburg-Eppendorf, as well as on websites that are structurally connected to the work group, one external private diabetes information website, and the website of a doctors' and therapists' CLBP network. An article was also published in a regional newspaper (Hamburger Abendblatt).

Study Procedures

In this purely Web-based trial without any face-to-face component, every person meeting the eligibility criteria could register for the study on the study website (open survey on a site created exclusively for the study) by providing a unique email address and choosing a password for login. After providing online informed consent and completing the pre-assessment (T2D: eligibility criteria, demographic data, time since diagnosis, treatment; CLBP: eligibility criteria, demographic data, chronic pain grade [51]), the participants were randomly assigned to the tailored system or the control system with the content tree. The informed consent was the first page entered after login. The participants were told the approximate length of time of the survey, where data were stored and for how long, who the investigators were, and the purpose of the study. Consent was provided via checkbox. Pre-assessments were completed after providing informed consent and before randomization. Only users who had filled in the pre-assessment were allowed to use the intervention (mandatory survey). In the control condition, tailoring variables (T2D: diabetes self-care [52], barriers to insulin treatment (BIT) [53], knowledge; CLBP: coping style [54], knowledge) were assessed immediately after randomization and before the intervention. In the tailored version, coping style (CLBP) was also assessed immediately after randomization and before the intervention, whereas knowledge, diabetes self-care, and barriers to insulin treatment were assessed throughout the intervention. The reason for this is that when tailoring to coping style, the user's coping type is determined in the beginning. At different places throughout the intervention, messages are tailored to this pre-assessed type. However, when tailoring to knowledge, diabetes self-care, and barriers to insulin treatment, there is no typology. Instead, individual items are assessed at different places throughout the intervention, and at that assessment point, one single message is tailored to the user's answer to the single item. Immediately after their first visit to the tailored IHCA or the control website, all participants were asked to fill in the post-assessment.

All participants received an email 3 months after their first visit asking them to fill in the online follow-up questionnaire. Participants were reminded by email twice, at 2 weeks and 4 weeks after the first email. Because non-monetary incentives have been shown to reduce attrition in online trials [50,55], participants who had answered all questionnaires received a €10 Amazon gift voucher. The voucher code was sent to them by email at the end of the study.

Participants were free to use the intervention as often and as long as they wished. Between the post and follow-up assessments, no prompts or reminders were used. No recommendations were provided regarding the duration or frequency of use, but the IHCA was designed to be used in one "go". Consequently, there were no prompts to use the intervention. No payment was required. Information on the frequency and duration of usage was gathered via server registrations. Usage data, data from the self-assessment questionnaires, and personal data such as name and email address were saved separately. Data were pseudonymized. After data collection, personal data were deleted. If participants withdrew their informed consent to study participation, their data were immediately erased. All data will be erased 5 years after the end of the study.

The study was approved by the Hamburg Medical Chamber ethics committee.

Treatment Allocation

The informed consent outlined that participants would be randomly assigned in consecutive order (50:50) to one of two presentation formats holding the same content. The random allocation (simple randomization) of the participants was automatically performed by the software program, which also provided the website and triggered automatic emails to participants. This centralized, software-driven, computerized, simple randomization procedure to the intervention or control group assured the concealment of allocation, so that randomization could not be subverted by the team of researchers. The two formats were not further elucidated, so participants did not know whether they were in the intervention or control group.

Description of the Intervention and Control Conditions

The tailored IHCA is designed as a stand-alone intervention that complements usual care. The T2D content of both the tailored IHCA and the control website covered basic information on diabetes (pathophysiology, epidemiology, subtypes, symptoms) and its sequelae (neuropathy, nephropathy, retinopathy, heart and vessel problems, sexual dysfunction, and depression), information on health behavior and lifestyle changes, and treatment options (see Table 1). The CLBP content covered essential information on CLBP (physiology of pain, acute vs chronic pain, chronification, epidemiology, psychological aspects, coping and pain management) and related psychological problems (depression, anxiety), diagnostic procedures, and treatment options (pharmacological and non-pharmacological; see Table 1). The look of the website (colors, font, figures, and pictures) was identical in both conditions. After registration, each participant received a password via email with which they could log onto the system as often as they wished.

Table 1. Overview over the IHCA contents.

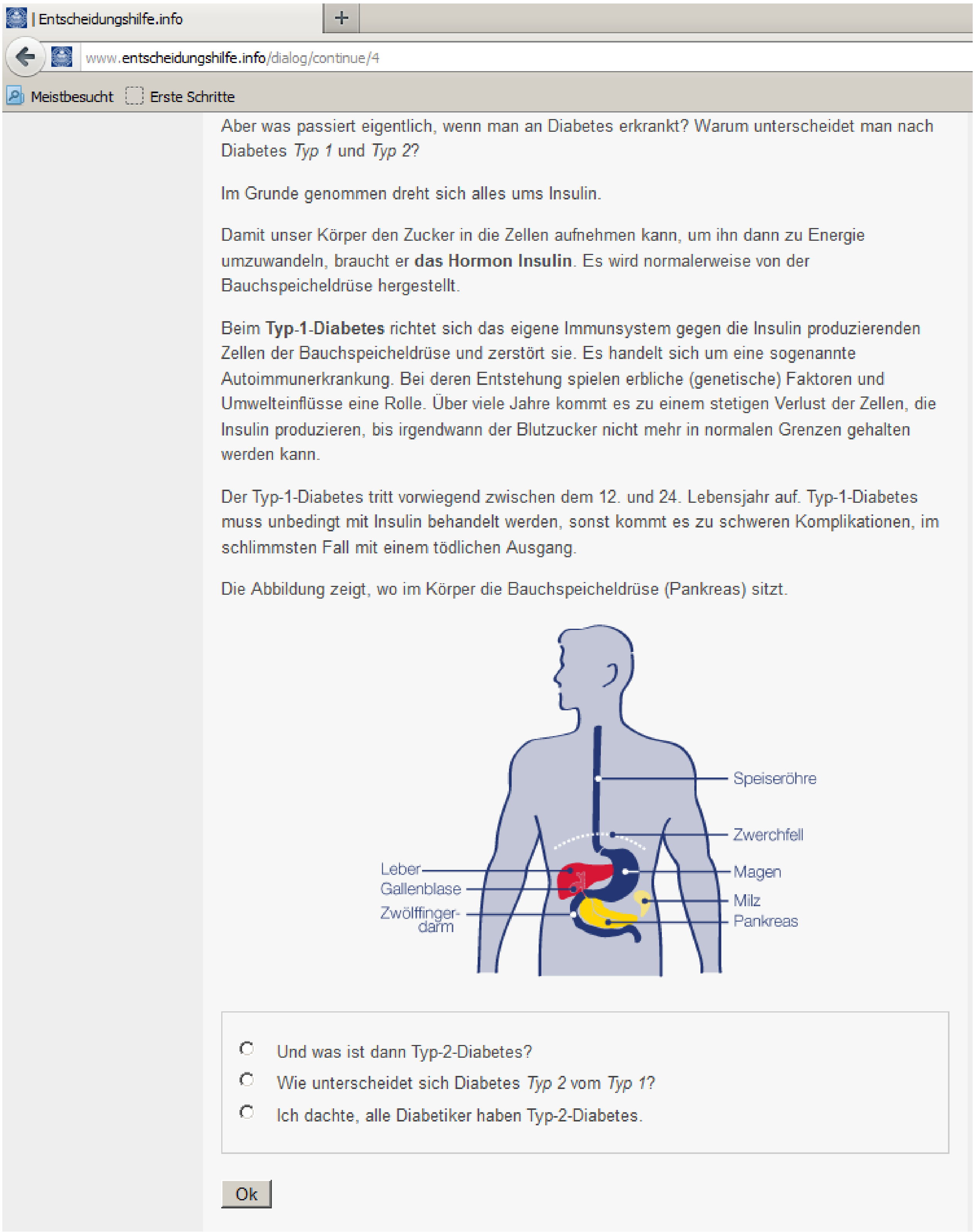
Type 2 diabetes	Chronic low back pain
1. Introduction: What is this website?	1. Introduction: What is this website?
1.1. Where does the information on this site come from?	1.1. Where does the information on this site come from?
2. Basics	2. CLBP Basics
2.1. Different diabetes types	2.1. Physiological basics: back, spine, and intervertebral discs
2.2. How do I know I have type 2 diabetes?	2.2. What exactly is pain?
2.3. What causes type 2 diabetes?	2.3. What is the difference between acute and chronic pain?
2.4. How many people live with type 2 diabetes?	2.4. Why does the pain stay when the physical injury heals?
2.5. How is type 2 diabetes diagnosed?	2.5. How many people live with CLBP?
2.6. Diabetes ABCs	2.6. Managing CLBP in everyday life
2.7. Blood sugar control	3. How is CLBP diagnosed?
3. How is type 2 diabetes treated?	3.1. How much diagnostics makes sense and at which point?
3.1. What are the goals of diabetes treatment?	3.2. Diagnostic options
3.2. What can you do to treat your diabetes?	4. How is CLBP treated?
3.3. When should you consider taking pills?	4.1. How much treatment makes sense and at which point?
3.4. Insulin treatment	4.2. What is the natural, untreated course of CLBP?
3.5. Summary and overview of the treatment options	5. Are there accompanying conditions or sequelae of CLBP?
4. Acute complications and sequelae	6. Treatment options
4.1. Which acute complications can occur?	6.1. How do I recognize good treatment?
4.2. Which sequelae can occur?	7. Summary
5. Additional information and literature	8. Additional information and literature
5.1. Associations and self-help	8.1. Associations and self-help
5.2. Websites	8.2. Websites
5.3. Journals	8.3. Journals
5.4. Books	8.4. Books
6. Glossary	9. Glossary
7. Legal notice	10. Legal notice
8. References	11. References

Tailored Condition

In the tailored condition, the delivery format was a dialogue-based, tunneled design tailoring the content and tone of the dialogue to relevant patient characteristics. It was developed based on two preliminary studies exploring the quality of existing websites [56] and assessing patient needs [57]. A tunneled design, in which the user is guided through the content, has been found to increase website use and knowledge gained from a website more than a design with more user control [58]. Still, it might annoy the user and evoke resistance [59]. Consequently, we decided to give the user some control over the path they take through the dialogue: at the end of each text passage, the user chose one of at least three reply options. These options always included at least one answer that expressed disagreement or doubt. The user then received a tailored answer that mirrored what the user had said, respected disagreement, conveyed esteem, and empathy and built an individualized bridge to the next content block. It was not possible to skip a whole content block (meaning the subheadings in Table 1), but it was possible to view the content in more or less detail.

Tailoring was performed using the following characteristics for diabetes patients: current T2D knowledge and preferred level of detail, attitudes toward self-care, and, if insulin treatment was a relevant topic, psychological barriers to it. The questionnaires that assessed patient characteristics were presented during the dialogue. At the beginning of the respective section (eg, diabetic foot), the participant was asked about their knowledge or attitude toward the topic, and the following section was then modified according to their answer. Figure 1 shows such a dialogue window.

Figure 1. Dialogue window.



Aber was passiert eigentlich, wenn man an Diabetes erkrankt? Warum unterscheidet man nach Diabetes *Typ 1* und *Typ 2*?

Im Grunde genommen dreht sich alles ums Insulin.

Damit unser Körper den Zucker in die Zellen aufnehmen kann, um ihn dann zu Energie umzuwandeln, braucht er **das Hormon Insulin**. Es wird normalerweise von der Bauchspeicheldrüse hergestellt.

Beim **Typ-1-Diabetes** richtet sich das eigene Immunsystem gegen die Insulin produzierenden Zellen der Bauchspeicheldrüse und zerstört sie. Es handelt sich um eine sogenannte Autoimmunerkrankung. Bei deren Entstehung spielen erbliche (genetische) Faktoren und Umwelteinflüsse eine Rolle. Über viele Jahre kommt es zu einem stetigen Verlust der Zellen, die Insulin produzieren, bis irgendwann der Blutzucker nicht mehr in normalen Grenzen gehalten werden kann.

Der Typ-1-Diabetes tritt vorwiegend zwischen dem 12. und 24. Lebensjahr auf. Typ-1-Diabetes muss unbedingt mit Insulin behandelt werden, sonst kommt es zu schweren Komplikationen, im schlimmsten Fall mit einem tödlichen Ausgang.

Die Abbildung zeigt, wo im Körper die Bauchspeicheldrüse (Pankreas) sitzt.

○ Und was ist dann Typ-2-Diabetes?
○ Wie unterscheidet sich Diabetes *Typ 2* vom *Typ 1*?
○ Ich dachte, alle Diabetiker haben Typ-2-Diabetes.

Diabetes Section and its Tailoring

Users' attitudes toward self-care were assessed with items that we adapted from the Summary of Diabetes Self-care Activities Measure (SDSCA) [52] to match the respective content section (see Table 2). Users were asked how important a certain self-care activity or piece of advice is for them. Every item had three reply options: "important or very important", "a little important", and "not important". The goal and techniques were inspired by Motivational Interviewing, a counseling method for addressing ambivalence about change [60].

For example, if a user attached great importance to the self-care behavior in question, this behavior was reinforced, positive consequences of the self-care behavior were stressed, and/or ideas were provided on how to keep up motivation. If a user found the self-care behavior in question “a little important”, an understanding of the users’ ambivalence was expressed, and the importance the user attached to the self-care behavior (little as it might be) was stressed and reinforced. Finally, if a user rated the self-care behavior as not important, the autonomy expressed in this answer was respected in order not to elicit resistance. Table 2 shows an example of self-care tailoring.

Table 2. Example of self-care tailoring: Response to “If you feel thirsty and urinate frequently, it usually means your blood sugar is…”.

Response options	Reply
High (correct answer)	That’s correct! If you want to learn more about what happens in the body and how you know that you have type 2 diabetes, you can go into more detail. Otherwise you can proceed to the next question. • I’d like to learn more about that topic. • I’d like to proceed to the next question.
Low (wrong answer)	No, that’s not correct. Actually, it’s the other way around: When you have type 2 diabetes, there is too much sugar in your blood. Unfortunately, you don’t realize it in the beginning. However, there are warning signs. The most important signs are […]
I don’t know	That’s ok, [name], that’s what we are here for: to learn, for example, what high blood sugar does to your body. When you have type 2 diabetes, there is too much sugar in your blood. Unfortunately, you don’t realize it in the beginning. However, there are warning signs. The most important signs are […]

Psychological barriers to insulin treatment were assessed using the Barriers to Insulin Treatment (BIT) questionnaire [53]. The BIT assesses the following expectations regarding insulin treatment: fear of injection and self-testing; expectations regarding positive insulin-related outcomes; expected hardships from insulin treatment; stigmatization by insulin injections; and fear of hypoglycemia. There are two or three items per subscale. In every item, a certain hope or fear with respect to insulin treatment is expressed (eg, “I am afraid of the pain when injecting insulin”). The user is asked to rate their agreement on a scale from 1 (completely disagree) to 10 (completely agree). A score from 1-10 can be calculated for each subscale. A validating and understanding (if a fear was expressed) or reinforcing (if a hope was expressed) answer was given. If there was a sign of fear (score >1), further information on the topic in question was provided.

Chronic Low Back Pain Section and its Tailoring

For CLBP, the concepts of coping style according to the avoidance endurance model (AEM) [37] and current CLBP knowledge and preferred level of detail were used for tailoring the provided information to the individual preferences of the users. The individual coping style was assessed using a questionnaire, which was presented before starting the dialogue. There are four AEM subtypes: the “depressed endurer”, which is high endurance coping (EC) and high depressiveness (D), the “happy endurer”, which is high EC and low D, the “depressed avoider”, which means low EC and high D, and the “adaptive coper”, which means low EC and low D (see Table 3). During the virtual conversation, the content, tone, and messages were tailored to the coping style of the individual user. The items that assess CLBP knowledge were presented during the dialogue. In the beginning of the respective section (eg, physiological basics), the user was asked about their level of knowledge on this subject. Depending on the response, the subsequent section was accordingly amended.

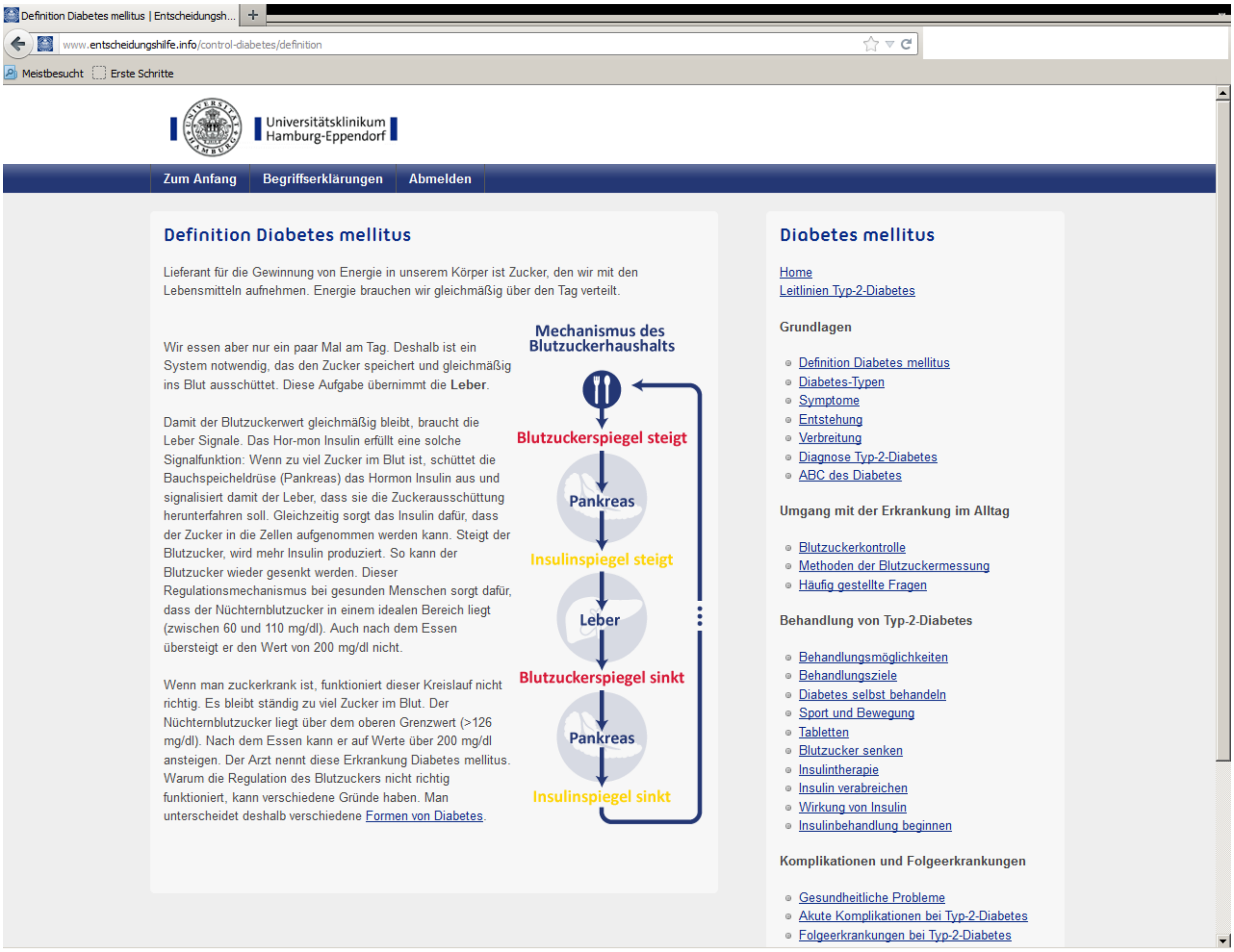
Table 3. Example of tailoring to coping style (CLBP).

Coping type	Adaptive coper	Happy endurer	Depressed endurer	Depressed avoider
Description of coping style	You go about your pain in a matter-of-fact manner. You know that on one hand, there is no serious disease behind it but that on the other hand, it can signal to you physical strain. You are good at taking short breaks at the right time to keep up your daily routine – maybe temporarily a little slower than usual.	You tend to keep going in your daily routine even if the pain is strong. This is, on one hand, a personal strength. However, at the same time, you run the risk of actually straining your muscles, ligaments, joints, and intervertebral discs.	You are a multi-tasker. Saying “No” to someone or not getting things done is hard on you. To meet requirements and get things done, you push yourself to your limits and beyond. Often, you don’t listen to your body before it is overstrained.	You are unsettled by your pain. You are worried that there might be a serious disease behind it, and / or you avoid activities that might increase the pain.
Take home message	Keep on like that! Make exercise part of your routine if you haven´t yet. Choose something fun and back-friendly. If you strengthen your muscles and stick to your relaxing breaks, the pain should soon vanish.	Even if it’s hard, try to pay more attention to your pain and take breaks early enough. Keep working, do things that are pleasant and fun, and keep moving – but remember to pause when you might need to!	Reconsider what you are asking from yourself: do you really have to demand so much? Maybe there are times when it is possible to leave something undone, to do it o.k. instead of perfectly, or to ask for assistance. These things are closely related to your pain.	Pain is unpleasant but not dangerous. Don’t let it suffocate you. Expand your limits step by step, and make pleasant activities a part of your everyday life.

Control Condition

On the control website, the content was not tailored and was not presented in a dialogue format. In contrast to the tailored, interactive version, the control website was not tunneled, and there was no guidance through the content. On the right side of each page, a content tree displayed a menu of all content sections that the participant could click on to get to the content of interest (see Figure 2). On both the intervention and control websites, the institutional affiliation of the University Medical Center Hamburg-Eppendorf was displayed at the top of each webpage.

Figure 2. Control window.



Definition Diabetes mellitus

Lieferant für die Gewinnung von Energie in unserem Körper ist Zucker, den wir mit den Lebensmitteln aufnehmen. Energie brauchen wir gleichmäßig über den Tag verteilt.

Wir essen aber nur ein paar Mal am Tag. Deshalb ist ein System notwendig, das den Zucker speichert und gleichmäßig ins Blut ausschüttet. Diese Aufgabe übernimmt die **Leber**.

Damit der Blutzuckerwert gleichmäßig bleibt, braucht die Leber Signale. Das Hor-mon Insulin erfüllt eine solche Signalfunktion: Wenn zu viel Zucker im Blut ist, schüttet die Bauchspeicheldrüse (Pankreas) das Hormon Insulin aus und signalisiert damit der Leber, dass sie die Zuckerausschüttung herunterfahren soll. Gleichzeitig sorgt das Insulin dafür, dass der Zucker in die Zellen aufgenommen werden kann. Steigt der Blutzucker, wird mehr Insulin produziert. So kann der Blutzucker wieder gesenkt werden. Dieser Regulationsmechanismus bei gesunden Menschen sorgt dafür, dass der Nüchternblutzucker in einem idealen Bereich liegt (zwischen 60 und 110 mg/dl). Auch nach dem Essen übersteigt er den Wert von 200 mg/dl nicht.

Wenn man zuckerkrank ist, funktioniert dieser Kreislauf nicht richtig. Es bleibt ständig zu viel Zucker im Blut. Der Nüchternblutzucker liegt über dem oberen Grenzwert (>126 mg/dl). Nach dem Essen kann er auf Werte über 200 mg/dl ansteigen. Der Arzt nennt diese Erkrankung Diabetes mellitus. Warum die Regulation des Blutzuckers nicht richtig funktioniert, kann verschiedene Gründe haben. Man unterscheidet deshalb verschiedene Formen von Diabetes.

Diabetes mellitus

Home
Leitlinien Typ-2-Diabetes

Grundlagen

- Definition Diabetes mellitus
- Diabetes-Typen
- Symptome
- Entstehung
- Verbreitung
- Diagnose Typ-2-Diabetes
- ABC des Diabetes

Umgang mit der Erkrankung im Alltag

- Blutzuckerkontrolle
- Methoden der Blutzuckermessung
- Häufig gestellte Fragen

Behandlung von Typ-2-Diabetes

- Behandlungsmöglichkeiten
- Behandlungsziele
- Diabetes selbst behandeln
- Sport und Bewegung
- Tabletten
- Blutzucker senken
- Insulintherapie
- Insulin verabreichen
- Wirkung von Insulin
- Insulinbehandlung beginnen

Komplikationen und Folgeerkrankungen

- Gesundheitliche Probleme
- Akute Komplikationen bei Typ-2-Diabetes
- Folgeerkrankungen bei Typ-2-Diabetes

Potential Risks for Participants

Research focusing on the negative effects of Internet interventions is scarce. One recent study on the side effects of Internet interventions for social anxiety disorder found that 14% of participants experienced negative effects, of which the most frequent was the emergence of new symptoms [61]. Concerning long-term conditions like T2D or CLBP, possible negative effects on cognitive or emotional variables such as self-efficacy or anxiety should be considered, because of high demands concerning self-management tasks or fear-inducing information. However, several recent reviews did not find any contraindications or negative side effects of IHCAs [21,28].

Intervention Development and Trial Design

The development process was user-oriented, evidence-based, and peer-reviewed. Two preliminary studies were conducted informing intervention development. To find out which topics are relevant to patients with T2D or CLBP, we performed a needs assessment with two steps. First, we conducted semistructured interviews with 12 physicians (T2D: 7 internists, 2 of whom were specialized in diabetology; CLBP: 5 physicians specialized in orthopedics) and 19 patients (10 with T2D, 9 with CLBP). In the second step, a self-assessment questionnaire was developed based on the main results of the interviews, and it was administered to a new and larger patient sample (T2D: N=178, CLBP: N=117). The needs assessment for T2D is described in more detail elsewhere [57]. We then conducted a cross-sectional study on the information and support available online, evaluating the formal quality, usability, and presence and quality of decision support of websites for CLBP or T2D. The results on T2D have been published elsewhere [56]. To ensure that the information is evidence-based, selected treatment guidelines were used as primary sources. Based on review articles [62,63] and up-to-dateness, the British [64] and the American [65] T2D guidelines were chosen. For CLBP, certain guidelines [49,66,67] and Cochrane reviews [68-73] were chosen. The theoretical foundations and the development of the T2D IHCA are described in more detail elsewhere [74]. Programming and graphic design were performed by the Gaia AG, a subcontractor specializing in Web-based health interventions. The intervention was not changed during the trial.

Outcomes Assessment

The primary outcomes were knowledge (assessed immediately after the first visit) and patient empowerment (assessed at 3-months follow-up).

T2D knowledge was assessed immediately after the first visit with 16 items, and CLBP knowledge was assessed with 29 items. The items were developed to map the content covered in the sections of the tailored IHCA and could be answered with true/false/I don’t know.

For the context of long-term conditions, patient empowerment was defined as a feeling of confidence and the ability to manage the challenges resulting from the chronic disease [10]. An empowered patient can better understand and participate in care processes, use resources, and measures to reduce negative emotions, and enhance strategies to cope with chronic disease. Consequently, patient empowerment includes intrapersonal and behavioral dimensions [75]. However, at the time this study started, we could not identify a generic, adequately validated questionnaire of empowerment for general use in long-term conditions as reported by a systematic review [76]. Patient empowerment was therefore measured with the Health Education Impact Questionnaire (heiQ) [77,78]. The heiQ includes 42 items and eight dimensions: Positive and Active Engagement in Life, Health Directed Behavior, Skill and Technique Acquisition, Constructive Attitudes and Approaches, Self-Monitoring and Insight, Health Service Navigation, Social Integration and Support, and Emotional Well-being. Schuler et al [79] translated the questionnaire into German and evaluated its psychometric properties (Raykov’s Composite Reliability Coefficient, factorial and concurrent validity). They were able to replicate the structure of the eight scales and found the questionnaire to be a reliable and valid measure. We removed Social Integration and Support from our testing battery because we did not expect an effect of our IHCA on that dimension. Although these 7 heiQ scales may not comprehensively measure the multidimensional construct of empowerment given, the selected scales do cover the intrapersonal and behavioral dimensions that are part of health-related empowerment. Patient empowerment was assessed only at 3-months follow-up because we expected changes on the heiQ to take more time.

The secondary outcomes were decisional conflict and preparation for decision making, assessed immediately after the first visit. Decisional conflict was assessed with the Decisional Conflict Scale (DCS) by O’Connor [80,81]. This questionnaire measures personal perceptions of uncertainty in choosing options, modifiable factors contributing to uncertainty such as feeling uninformed, unclear about personal values, and unsupported in decision making, and effective decision making such as feeling that the choice is informed, values-based, and likely to be implemented and expressing satisfaction with the choice. Reliability is good, with a Cronbach alpha between .78 and .92 [80]. The discriminant validity is acceptable. Preparation for decision making was measured with the Preparation for Decision Making Scale (PDMS).

Preparation for decision making was measured with the Preparation for Decision Making Scale (PDMS) [82,83]. This 11-item scale assesses a patient’s or participant’s perception of how useful a decision aid or decision support intervention was in preparing them to communicate with their practitioner in making a health decision. The reliability is very good, ranging from alpha=.92 to alpha=.94. Both questionnaires were offered only to those participants who had indicated that they were facing a health decision concerning their T2D or CLBP.

To avoid missing data, all questionnaires included validation checks that alerted participants when their answers were implausible or when items were skipped. Usage data were assessed via log files. Before going online, the usability and technical functionality of the electronic questionnaire was tested by members of the research team. All outcomes were self-assessed through online questionnaires. The questionnaires were not validated for online use.

Statistical Analysis

Baseline Data

Data on sample characteristics were analyzed using *t* tests (for metric data) and chi-square tests (for categorical data) to test for differences between treatment groups. A dropout analysis was performed to test for possible attrition bias. The effects of the intervention (tailored vs control condition), disease (T2D vs CLBP), gender, age, education, family status, and employment status on attrition were evaluated using *t* tests (for metric data) and chi-square tests (for categorical data).

Intention-to-Treat Analysis

To evaluate the effectiveness of the tailored IHCA, multiple linear regression analyses were performed using the intervention, the disease, and their interaction term as dummy-coded predictors. Intention-to-treat (ITT) and available cases (AC) analyses were performed for all outcomes. The ITT approach pooled 10 analyses, estimating missing values by a multiple regression approach using all outcomes, demographic data, and diseases but not intervention information for multiple data imputation (MI). In the primary ITT analysis, a corrected level of significance was used for testing the eight primary outcomes (Bonferroni adjustment); thus, the results with a type I error rate of $P<.001$ were considered statistically significant. For secondary outcomes, $P<.05$ was used.

Sensitivity Analysis (Available Cases)

The AC analysis included all of the available participants providing valid data on t_1 and/or t_2. In both analyses, estimated marginal means with standard errors for both the tailored and control conditions were calculated with analysis of variance (ANOVA). Additionally, these parameters were also retained for subgroups stratified by condition. In all AC analyses, results with a type I error rate of $P<.05$ were considered statistically significant. All analyses were performed using SPSS 18.0.

Results

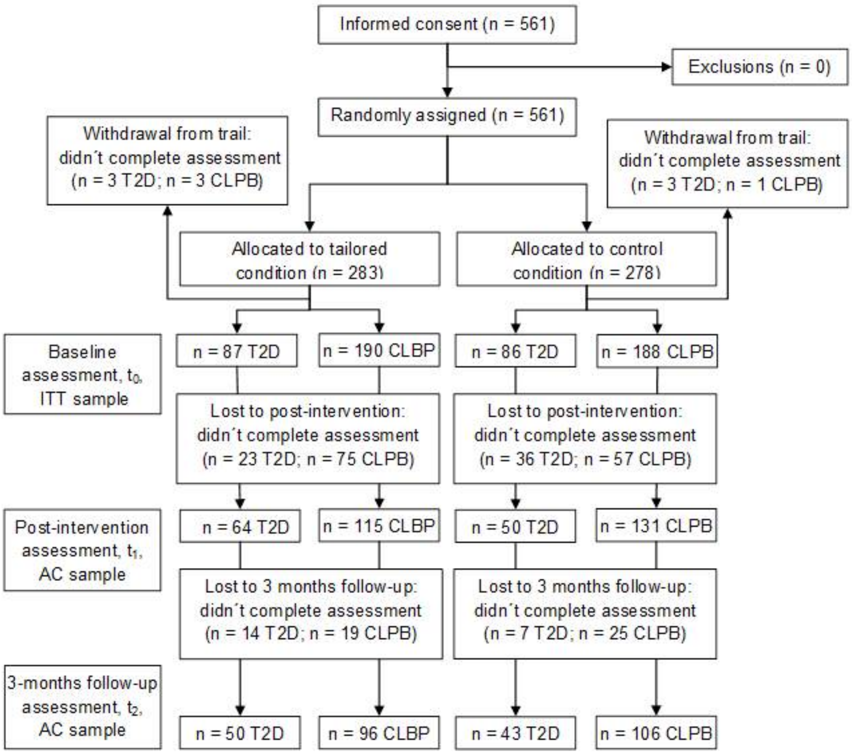
Participant Flow

A total of 561 users agreed to participate in the study. Of these, 179 (31.9%) had T2D, and 382 (68.1%) had CLBP. Analyzable data (availability of at least basic demographic information such as age and gender) at t_0 were available from 551 users. For data analysis at t_1, data for 360 participants was available (availability of data for at least one of the outcomes of t_1). Three months after system use, the questionnaires of 295 participants contained data on at least one of the three outcomes at t_2 and could thus be used for analyses (Figure 3). There were no significant differences with regard to gender, age, family status, educational level, or working status between those participants who provided all questionnaires and those who dropped out of the study after providing at least demographic data. Participants with T2D who were treated with oral anti-diabetics provided data at

$t_1$significantly more often than those who were treated with dietary changes or insulin. Those participants who provided data at t_1 spent significantly more time using the system, and participants in the tailored condition spent significantly more time in the IHCA than participants in the control condition spent in the control website (see Table 4).

There was also selective dropout between t_0 and $t_1$among participants with CLBP. At t_1, participants with CLBP were significantly (*P*=.015) younger in the tailored condition (mean 48.0; SD 12.9) than in the control condition (mean 52.0; SD 12.7). Additionally, there are significantly (*P*=.021) more participants with higher education in the tailored condition (62.6%) than in the control condition (48.9%). Among the participants with T2D, there was no selective dropout between t_0 and t_1. At t_2, there were no significant differences in either of the two diseases (T2D or CLBP).

Figure 3. Flow of participants after randomization (ITT=intention-to-treat, AC=available cases).



Baseline Data

The mean age was 52.2 years (SD 13.1) in the tailored condition and 52.7 years (SD 13.0) in the control condition. Of the participants using the IHCA, 58.5% (162/277) were female (control condition: 59.1%, 162/274). There were no statistically significant differences in further demographic variables such as marital status, educational level, and working status. Sample characteristics are shown in Table 4.

Table 4. Sample characteristics[a].

	Tailored condition t_0 (n=277)	Control condition t_0 (n=274)	Baseline differences (tailored vs control condition), *P* value	Total t_1 (n=360)	Dropout analysis (t_1 available vs t_1 not available), *P* value
Female, n (%)	162 (58.5)	162 (59.1)	.474	216 (60.0)	.467
Age in years, mean (SD)	52.2 (13.1)	52.7 (13.0)	.668	51.8 (13.1)	.116
Family status, n (%)					
Single	67 (24.2)	77 (28.1)	.742	100 (27.8)	.341
Married	162 (58.5)	150 (54.7)		194 (53.9)	
Divorced	39 (14.1)	37 (13.5)		52 (14.4)	
Widowed	9 (3.2)	10 (3.6)		14 (3.9)	
Educational level, high[b], n (%)	148 (53.4)	140 (51.1)	.322	198 (55.0)	.089
Working status, employed, n (%)	145 (55.6)	160 (58.4)	.282	207 (57.5)	.786
Years since diagnosis[c], mean (SD)	11.1 (7.6)	10.5 (8.0)	.649	10.7 (8.2)	.858
Current diabetes treatment [a]					
Dietary change	40 (44.4)	46 (51.7)	.371	57 (50.0)	.535
Insulin	35 (38.9)	25 (28.1)	.154	43 (37.7)	.139
Oral anti-diabetics	59 (65.6)	55 (61.8)	.643	80 (70.2)	.023
Disability score[d], mean (SD)	41.4 (22.5)	42.7 (22.8)	.573	42.2 (20.9)	.855
System usage in minutes, mean (SD)	51.16 (39.7)	37.6 (35.0)	<.001	49.7 (35.1)	<.001

[a] t_0 = demographic data available (ITT population); t_1= at least one outcome after intervention reported.

[b] more than 10 years of education.

[c] for patients with diabetes.

[d] for patients with back pain.

Intention-to-Treat Analysis

The following results were obtained using the ITT approach including all randomized participants. The results of the sensitivity analysis using the available cases approach are reported in a separate section. Table 5 shows all of the results in detail.

Knowledge Immediately After the First Visit (t_1)

With regard to knowledge of T2D or CLBP users in the tailored condition had a mean score of 77.9 (SE 1.2) compared with 76.3 (SE 1.3) in the control condition. There were no significant differences between groups (*P*=.53). There was, however, a significant difference between users with T2D and CLBP (*P*<.001), indicating higher knowledge scores in the T2D group. In addition, we observed a significant interaction effect between intervention and disease (*P*=.04), more strongly favoring the tailored condition over the control condition in CLBP (estimated mean difference of 4.6 [95% CI 1.0-8.2] points on a 0-100 points scale) than in T2D participants (estimated mean difference of -1.6 points [95% CI -7.4 to 4.2] on a 0-100 points scale).

Patient Empowerment at 3-Month Follow-Up (t_2)

The heiQ does not provide a total score for patient empowerment. Table 5 shows the results for the seven included dimensions. There was no significant intervention main effect or interaction. However, there was a significant disease main effect on the dimensions Skill and Technique Acquisition (*P*=.01) and Self-Monitoring and Insight (*P*=.04), both indicating higher scores for users with T2D.

Decisional Conflict Immediately After the First Visit (t_1)

There was a highly significant disease main effect. After the first use of the system, decisional conflict was lower in the CLBP group than in the T2D group (*P*<.001). There was no significant intervention main effect and no significant interaction.

Preparation for Decision Making Immediately After the First Visit (t_1)

There was no significant main effect or interaction.

Table 5. Results of ITT and AC analyses.

	N	Tailored condition			Control condition			Intervention main effect *P*	Disease main effect *P*	Intervention x disease *P*
		T2D,M (SE)	CLBP,M (SE)	Total, M (SE)	T2D, M (SE)	CLBP, M (SE)	Total, M (SE)			
Intention-to-treat analysis										
Primary outcomes										
Knowledge	551	81.3 (1.9)	74.4 (1.2)	77.9 (1.2)	82.9 (2.3)	69.8 (1.4)	76.3 (1.3)	.53	<.001	.04
Positive and active engagement in life	551	71.9 (2.5)	69.7 (1.8)	70.8 (1.4)	71.4 (2.3)	70.9 (1.8)	71.2 (1.4)	.88	.86	.43
Health directed behavior	551	63.5 (3.9)	68.7 (2.4)	66.1 (2.4)	63.7 (3.3)	68.3 (2.4)	66.0 (2.0)	.97	.28	.92
Emotional well-being	551	68.8 (3.9)	63.2 (2.8)	66.0 (2.6)	62.6 (3.7)	60.2 (2.8)	61.4 (2.3)	.28	.60	.66
Constructive attitudes and approaches	551	78.3 (2.9)	75.4 (2.1)	76.8 (1.9)	75.8 (2.5)	75.6 (1.9)	75.7 (1.6)	.498	.95	.59
Skill and technique acquisition	551	77.6 (2.6)	65.1 (1.7)	71. 4 (1.5)	75.8 (2.9)	67.6 (1.7)	71.7 (1.8)	.62	.01	.36
Self-monitoring and insight	551	80.1 (2.1)	70.8 (1.4)	75.4 (1.4)	79.5 (2.2)	73.4 (1.3)	76.5 (1.2)	.85	.04	.52
Health service navigation	551	77.9 (3.1)	70.0 (2.1)	73.9 (2.0)	74.0 (2.9)	69.7 (1.8)	71.8 (1.6)	.32	.24	.44
Secondary outcomes										
Decisional conflict	551	79.7 (2.3)	61.3 (1.6)	70.5 (1.5)	75.5 (2.3)	60.3 (1.7)	67.9 (1.4)	.15	<.001	.33
Preparation for decision making	551	60.5 (3.4)	53.8 (2.5)	56.7 (2.1)	57.6 (3.7)	51.2 (2.3)	54.4 (2.2)	.57	.14	.85
Available cases analysis										
Primary outcome										
Knowledge	330	81.1 (1.9)	77.1 (1.4)	79.1 (1.2)	81.8 (2.1)	68.7 (1.3)	75.2 (1.2)	.02	<.001	.008
Positive and active engagement in life	295	71.8 (2.6)	69.9 (1.8)	70.9 (1.6)	71.3 (2.8)	71.3 (1.8)	71.3 (1.6)	.86	.68	.68
Health directed behavior	295	63.0 (3.4)	69.4 (2.5)	66.2 (2.1)	64.9 (3.7)	68.7 (2.4)	66.8 (2.2)	.84	.10	.68
Emotional well-being	295	70.8 (3.7)	66.1 (2.6)	68.5 (2.3)	60.7 (3.9)	59.3 (2.5)	60.0 (2.3)	.009	.35	.60
Constructive attitudes and approaches	295	78.8 (2.8)	76.1 (2.0)	77.5 (1.7)	3.2 (0.09)	74.5 (3.0)	75.2 (1.9)	.30	.68	.51
Skill and technique acquisition	295	78.3 (2.4)	64.3 (1.7)	71.3 (1.5)	75.0 (2.6)	68.8 (1.6)	71.9 (1.5)	.78	<.001	.06
Self-monitoring and insight	295	80.3 (1.9)	70.0 (1.3)	75.2 (1.2)	79.3 (2.0)	74.7 (1.3)	77.0 (1.2)	.27	<.001	.09
Health service navigation	295	79.1 (2.7)	71.2 (1.9)	75.2 (1.6)	73.4 (2.9)	69.8 (1.8)	71.6 (1.7)	.13	.02	.37
Secondary outcomes										
Decisional conflict	324	79.9 (2.4)	61.9 (1.8)	70.9 (1.5)	74.8 (2.7)	60.4 (1.7)	67.6 (1.6)	.13	<.001	.47
Preparation for decision making	324	61.0 (3.3)	52.1 (2.4)	56.4 (2.0)	55.7 (3.6)	51.2 (2.2)	53.5 (2.1)	.29	.02	.47

Sensitivity Analysis (Available Cases)

In addition to the ITT approach, we performed all calculations following the AC approach, including only participants who filled in all of the questionnaires. The aim of this procedure was to determine the extent to which missing data impacted the results reported above (sensitivity analysis).

Knowledge Immediately After the First Visit (t_1)

The AC analysis showed a significant intervention main effect for knowledge (*P*=.02) indicating higher scores for the tailored condition (mean 79.1, SE 1.2) than for the control condition (mean 75.2, SE 1.2). The estimated mean difference between groups was 3.9 (95% CI 0.5-7.3) points on a 0-100 points scale. There was a significant disease x intervention interaction (*P*=.008) for knowledge, indicating the superiority of the tailored condition over the control condition in CLBP (estimated mean difference of 8.4 [95% CI 4.7-12.1] points on a 0-100 points scale) but not in T2D participants (estimated mean difference of -0.7 [95% CI -6.5 to 5.1] points on a 0-100 points scale). Additionally, there was a significant disease main effect for knowledge favoring the T2D group.

Patient Empowerment at 3-Month Follow-Up (t_2)

We found a significant intervention main effect for Emotional Well-being (meaning less health-related negative effects such as anxiety, anger, and depression [78]) (*P*=.009) favoring the tailored condition (mean 68.5, SE 2.3) over the control condition (mean 60.0, SE 2.3). The estimated mean difference between groups was 25.4 (95% CI 6.3-44.5) points on a 0-100 points scale. Finally, there were significant disease main effects for Skill and Technique Acquisition (*P*<.001), Self-Monitoring and Insight (*P*<.001), and Health Service Navigation (*P*=.02) favoring the T2D group.

Decisional Conflict Immediately After the First Visit (t_1)

We found a significant disease main effect (*P*<.001) showing more decisional conflict in the T2D group.

Preparation for Decision Making Immediately After the First Visit (t_1)

There was a significant disease main effect (*P*=.02) indicating higher scores for the T2D compared with the CLBP group.

Discussion

Principal Findings

In a randomized controlled trial, we compared a Web-based, tailored, dialogue-based information system containing information on T2D or CLBP (tailored condition) with a website providing identical information without dialogue structure, tailoring, or interactive elements (control condition). The primary outcomes of the trial were knowledge and patient empowerment. Secondary outcomes were decisional conflict and preparation for decision making.

We expected that the tailored IHCA would be more attractive than the control website, be used more, and would thus lead to more knowledge and more empowerment. Indeed, participants spent significantly more time with the tailored website than the control website. Still, this did not lead to more knowledge or empowerment in the primary ITT analysis. In the AC analysis, the participants in the tailored condition displayed more knowledge at t_1 and more Emotional Well-being at t_2. This indicates that the tailored IHCA was more effective on these two dimensions than the control website. This was not the case for all users included; this was only the case for those who remained in the study and thus spent more time using the system. Contrary to the hypothesis, the tailored IHCA did not result in higher scores on the other six heiQ scales. It is possible that the effect was limited to the emotional level and could not be transferred to the cognitive or behavior level. This is in line with the results of Pal et al, who found that positive effects on cognitive outcomes could not be converted into behavioral changes [21]. There was a significant intervention x disease interaction favoring the tailored condition over the control condition more strongly in CLBP than in T2D participants. This superiority might indicate that tailoring in the CLBP IHCA may be more effective than tailoring in the T2D IHCA.

Other recent studies aimed directly at behavioral changes found effects on behavioral outcomes [24], and a meta-analysis on Internet-based cognitive behavioral therapy for patients with chronic somatic diseases found effects on psychological and physical outcomes [84]. A Web-based intervention aimed at psychosocial well-being in older adults with diabetes found improvements in depression, quality of life, social support, and self-efficacy [85], and a Web-based depression treatment for people with diabetes was found to reduce diabetes-specific emotional distress but had no beneficial effect on glycemic control [86]. Taken together, these results suggest that interventions aimed specifically at certain outcomes reliably have effects on these outcomes but have fewer effects on related or more distal outcomes. Consequently, our IHCA, as an educational intervention providing health information and adding behavioral change and decision support, has more consistent effects on knowledge (in persons who actually use it) than on cognitive or behavioral outcomes.

There were no significant effects regarding decisional conflict or preparation for decision making. A recent Cochrane review found that decision aids have, among other outcomes, an impact on knowledge and decisional conflict [87]. Again, the fact that we did find an impact on knowledge in the AC analysis but not on decisional conflict or preparation for decision making might be due to our IHCA being more of an educational intervention, providing the information necessary for shared decision making, than a classical decision tool.

Users with T2D yielded significantly better results regarding knowledge, preparation for decision making (only AC), and three (ITT: two) dimensions of the heiQ than participants with CLBP. One possible explanation might be that education and empowerment are traditionally cornerstones of diabetes management [88], which is not as explicitly true for the treatment of CLBP. Still, this result should be interpreted cautiously, because the instruments used to measure knowledge were different in both groups.

Strengths and Limitations

The work presented is the first trial on a German language IHCA on T2D or CLBP. The intervention was designed carefully based on two preliminary studies. There are some limitations to the work. One limitation concerns the representativeness of the sample. Only people with Internet access could be included in the study. Of the German general population, 73% are online [89], but of the population over 50, only 47% use the Internet. Because the prevalence of both T2D [90] and CLBP increases with age [91,92], there might be a selection bias in our sample. The diagnosis was self-assessed. In addition, this presents a limitation regarding the implementation and reach of online support for these diseases. Still, attrition was comparatively low for an online trial [23]. At t_2, 52.4% of the sample was retained. The comparatively low attrition rate in the tailored and control conditions might be due to the incentive given for complete datasets. Because none of the outcome criteria were assessed at t_0, we cannot know whether the differences between conditions at t_1 were caused by the intervention or had been there from the beginning.

We did not include quantitative or qualitative feedback on user acceptance. We also did not assess potential confounders (eg, which other interventions the participants used while enrolled in the study). These variables might have added to our understanding of the IHCA effects. Going beyond the scope of our study, investigating the effectiveness of the tested intervention, further research should focus on the mechanisms of change and the role of context variables through analyzing potential mediators and moderators [93]. Although the participants were blinded to the group assignment, it might be possible that participants identified the intervention group due to the unusual dialogue-based delivery format used in the intervention group. However, the design and content of both groups were nearly identical.

Another limitation arises from the measures used. First, there are concerns regarding data quality and response rates in online questionnaires [94,95]. Psychometric properties have been found to be equivalent to or even better than data obtained from paper pencil questionnaires [96,97]. There are also advantages of online assessment: data quality can additionally be improved by validation checks that alert participants if their answers are implausible or if items are skipped [96]. Furthermore, online assessment seems to be less prone to social desirability [98]. Second, only some of the measures used in this trial are standardized (DCS, PDMS, BIT), whereas others are adapted (attitudes toward self-care) for our purposes. The measure to assess the primary outcome of T2D/CLBP knowledge was developed for the purpose of this study and has not been validated. Different versions of this outcome measure with different numbers of items for T2D and CLBP are used. None of the measures have been adapted for online use, which limits their comparability to results obtained from paper pencil tests [99].

Finally, the intervention had multiple components. We cannot know which component resulted in which effect. Future research should determine which components are effective and which are not.

Conclusions

The tailored IHCA enhanced knowledge and empowerment in persons who actually used it but failed to have effects in the total study population and on more distal outcomes. It might be concluded that tailoring and interactivity do not have effects with regard to these outcomes. Intervention components more specifically targeting cognitive and behavioral outcomes might enhance the effects. Pathways of change connecting intervention components and effects should be explored.

With regard to implementation, the IHCA could function outside of the study without major changes. Still, it would require some resources for updates and maintenance. Involving sponsors from the beginning might facilitate implementation. If our IHCA had made it to this stage, there would have been steps taken to extend its reach and effectiveness. In addition to being more specific, adaptability to tablets and mobile phones might have been an asset [21]. Another feature could be blended care to more explicitly integrate personal contacts, telephone, and online support [100]. The opportunity to share information and experiences with peers might be an especially attractive and important feature. The Pew Internet and American Life Project [101] found that people living with a chronic disease are more actively using the opportunities of Web 2.0: they generate and share content on their disease, use social media, blog, and chat more than people with no chronic conditions. Stepping into a multimedia dialogue with the users and letting expert-generated content and user-generated content spur each other might be the next step toward patient-centeredness in online support.

Acknowledgments

This study was funded by the German Federal Ministry of Education and Research (grant number: 01GX0710).

Authors' Contributions

NW participated in the conception and design of the study, interpreted the findings, and drafted the manuscript. JD participated in the conception and design of the study, interpreted the findings, and revised the manuscript. AW analyzed the data, interpreted the findings, and participated in drafting and revising the manuscript. LK participated in the conception and design of the study, analyzed the data, interpreted the findings, and participated in revising the manuscript. MH was awarded the grant, participated in the conception and design of the study and the interpretation of the findings, and revised the manuscript.
All authors read and approved the final manuscript.

Conflicts of Interest

NW, JD, and MH were among the developers of the intervention.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [46].

[PDF File (Adobe PDF File), 93KB - jmir_v17i3e53_app1.pdf]

References

1. Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, et al. A systematic review of the global prevalence of low back pain. Arthritis Rheum 2012 Jun;64(6):2028-2037 [FREE Full text] [doi: 10.1002/art.34347] [Medline: 22231424]
2. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Blood Glucose). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. Lancet 2011 Jul 2;378(9785):31-40. [doi: 10.1016/S0140-6736(11)60679-X] [Medline: 21705069]
3. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004 May;27(5):1047-1053. [Medline: 15111519]
4. Becker A, Held H, Redaelli M, Strauch K, Chenot JF, Leonhardt C, et al. Low back pain in primary care: costs of care and prediction of future health care utilization. Spine (Phila Pa 1976) 2010 Aug 15;35(18):1714-1720. [Medline: 21374895]
5. Freburger JK, Holmes GM, Agans RP, Jackman AM, Darter JD, Wallace AS, et al. The rising prevalence of chronic low back pain. Arch Intern Med 2009 Feb 9;169(3):251-258. [doi: 10.1001/archinternmed.2008.543] [Medline: 19204216]
6. Knauer SR, Freburger JK, Carey TS. Chronic low back pain among older adults: a population-based perspective. J Aging Health 2010 Dec;22(8):1213-1234. [doi: 10.1177/0898264310374111] [Medline: 20657007]
7. Centers for Disease Control and Prevention (CDC). Prevalence and most common causes of disability among adults--United States, 2005. MMWR Morb Mortal Wkly Rep 2009 May 1;58(16):421-426 [FREE Full text] [Medline: 19407734]
8. Agency for Healthcare Research and Quality (AHRQ): Medical Expenditure Panel Survey. Total expenses and percent distribution for selected conditions by type of service. 2014. URL: http://meps.ahrq.gov/data_stats/tables_compendia_hh_interactive.jsp?_SERVICE=MEPSSocket0&_PROGRAM=MEPSPGM.TC.SAS&File=HC2Y2011&Table=HC2Y2011_CNDXP_C&_Debug= [WebCite Cache ID 6QTNU8klE]
9. Segal L. The importance of patient empowerment in health system reform. Health Policy 1998 Apr;44(1):31-44. [Medline: 10180200]
10. Wang JM, Walter S, Mantovani A. Re-evaluation of the chemotactic activity of tumour necrosis factor for monocytes. Immunology 1990 Nov;71(3):364-367. [Medline: 2269474]
11. Aujoulat I, d'Hoore W, Deccache A. Patient empowerment in theory and practice: polysemy or cacophony? Patient Educ Couns 2007 Apr;66(1):13-20. [doi: 10.1016/j.pec.2006.09.008] [Medline: 17084059]
12. 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]
13. Schulz PJ, Nakamoto K. Health literacy and patient empowerment in health communication: the importance of separating conjoined twins. Patient Educ Couns 2013 Jan;90(1):4-11. [doi: 10.1016/j.pec.2012.09.006] [Medline: 23063359]
14. Funnell MM, Anderson RM, Arnold MS, Barr PA, Donnelly M, Johnson PD, et al. Empowerment: an idea whose time has come in diabetes education. Diabetes Educ 1991;17(1):37-41. [Medline: 1986902]
15. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]
16. Camerini L, Schulz PJ, Nakamoto K. Differential effects of health knowledge and health empowerment over patients' self-management and health outcomes: a cross-sectional evaluation. Patient Educ Couns 2012 Nov;89(2):337-344. [doi: 10.1016/j.pec.2012.08.005] [Medline: 22959333]
17. Lustria MLA, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
18. Proudfoot J, Klein B, Barak A, Carlbring P, Cuijpers P, Lange A, et al. Establishing guidelines for executing and reporting Internet intervention research. Cogn Behav Ther 2011;40(2):82-97. [doi: 10.1080/16506073.2011.573807] [Medline: 25155812]
19. Elbert NJ, van OMH, van RW, Ekeland AG, Hakkaart-van RL, Raat H, Nijsten Tamar E C, Pasmans Suzanne G M A. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]

20. van VM, de WM, Cleijne WHJ, Snoek FJ. Use of behavioral change techniques in web-based self-management programs for type 2 diabetes patients: systematic review. J Med Internet Res 2013;15(12):e279 [FREE Full text] [doi: 10.2196/jmir.2800] [Medline: 24334230]
21. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013;3:CD008776. [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
22. Riva S, Camerini AL, Allam A, Schulz PJ. Interactive sections of an Internet-based intervention increase empowerment of chronic back pain patients: randomized controlled trial. J Med Internet Res 2014;16(8):e180 [FREE Full text] [doi: 10.2196/jmir.3474] [Medline: 25119374]
23. Simon D, Kriston L, von WA, Buchholz A, Vietor C, Hecke T, et al. Effectiveness of a web-based, individually tailored decision aid for depression or acute low back pain: a randomized controlled trial. Patient Educ Couns 2012 Jun;87(3):360-368. [doi: 10.1016/j.pec.2011.10.009] [Medline: 22154867]
24. Schulz DN, Kremers SPJ, Vandelanotte C, van Adrichem MJG, Schneider F, Candel MJJM, et al. Effects of a web-based tailored multiple-lifestyle intervention for adults: a two-year randomized controlled trial comparing sequential and simultaneous delivery modes. J Med Internet Res 2014;16(1):e26 [FREE Full text] [doi: 10.2196/jmir.3094] [Medline: 24472854]
25. Samoocha D, Bruinvels DJ, Elbers NA, Anema JR, van der Beek AJ. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. J Med Internet Res 2010;12(2):e23 [FREE Full text] [doi: 10.2196/jmir.1286] [Medline: 20581001]
26. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res 2013;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]
27. Ciere Y, Cartwright M, Newman SP. A systematic review of the mediating role of knowledge, self-efficacy and self-care behaviour in telehealth patients with heart failure. J Telemed Telecare 2012 Oct;18(7):384-391. [doi: 10.1258/jtt.2012.111009] [Medline: 23019605]
28. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive Health Communication Applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
29. Camerini L, Schulz PJ. Effects of functional interactivity on patients' knowledge, empowerment, and health outcomes: an experimental model-driven evaluation of a web-based intervention. J Med Internet Res 2012;14(4):e105 [FREE Full text] [doi: 10.2196/jmir.1953] [Medline: 22810046]
30. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]
31. 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]
32. Brouwer W, Oenema A, Raat H, Crutzen R, de NJ, de Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010 Aug;25(4):585-595 [FREE Full text] [doi: 10.1093/her/cyp063] [Medline: 19897515]
33. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. J Med Internet Res 2010;12(4):e69 [FREE Full text] [doi: 10.2196/jmir.1640] [Medline: 21156470]
34. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res 2011;13(1):e9 [FREE Full text] [doi: 10.2196/jmir.1391] [Medline: 21371992]
35. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
36. Fan L, Sidani S. Effectiveness of Diabetes Self-management Education Intervention Elements: A Meta-analysis. Canadian Journal of Diabetes 2009 Jan;33(1):18-26. [doi: 10.1016/S1499-2671(09)31005-9]
37. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
38. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. J Med Internet Res 2006;8(3):e15 [FREE Full text] [doi: 10.2196/jmir.8.3.e15] [Medline: 16954125]
39. 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]

40. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
41. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
42. 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]
43. Lustria MLA, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
44. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
45. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
46. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
47. Dirmaier J, Härter M, Weymann N. A tailored, dialogue-based health communication application for patients with chronic low back pain: study protocol of a randomised controlled trial. BMC Med Inform Decis Mak 2013;13:66 [FREE Full text] [doi: 10.1186/1472-6947-13-66] [Medline: 23768119]
48. Weymann N, Härter M, Dirmaier J. A tailored, interactive health communication application for patients with type 2 diabetes: study protocol of a randomised controlled trial. BMC Med Inform Decis Mak 2013;13:24 [FREE Full text] [doi: 10.1186/1472-6947-13-24] [Medline: 23406466]
49. Becker A, Chenot JF, Niebling W, Kochen MM, Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin. [Guidelines for back pain]. Z Orthop Ihre Grenzgeb 2004;142(6):716-719. [doi: 10.1055/s-2004-832488] [Medline: 15614654]
50. Khadjesari Z, Murray E, Kalaitzaki E, White IR, McCambridge J, Thompson SG, et al. Impact and costs of incentives to reduce attrition in online trials: two randomized controlled trials. J Med Internet Res 2011;13(1):e26 [FREE Full text] [doi: 10.2196/jmir.1523] [Medline: 21371988]
51. Klasen BW, Hallner D, Schaub C, Willburger R, Hasenbring M. Validation and reliability of the German version of the Chronic Pain Grade questionnaire in primary care back pain patients. Psychosoc Med 2004;1:Doc07 [FREE Full text] [Medline: 19742049]
52. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000 Jul;23(7):943-950 [FREE Full text] [Medline: 10895844]
53. Petrak F, Stridde E, Leverkus F, Crispin AA, Forst T, Pfützner A. Development and validation of a new measure to evaluate psychological resistance to insulin treatment. Diabetes Care 2007 Sep;30(9):2199-2204. [doi: 10.2337/dc06-2042] [Medline: 17575092]
54. Hasenbring M. Hallner D: Telemedizinisches Patienten-Diagnose-System (TPDS) : Selbsterklärende PC-Lösung zur Analyse von Risikofaktoren der Chronifizierung von Rückenschmerzen. Dtsch Arztebl 1999;6:49-50.
55. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2009(3):MR000008. [doi: 10.1002/14651858.MR000008.pub4] [Medline: 19588449]
56. Weymann N, Härter M, Dirmaier J. Quality of online information on type 2 diabetes: a cross-sectional study. Health Promot Int 2014 Mar 30:epub ahead of print. [doi: 10.1093/heapro/dau019] [Medline: 24688114]
57. Weymann N, Härter M, Dirmaier J. Information and decision support needs in patients with type 2 diabetes. Health Informatics J 2014 Jun 10:-. [doi: 10.1177/1460458214534090] [Medline: 24916569]
58. Crutzen R, Cyr D, de Vries NK. The role of user control in adherence to and knowledge gained from a website: randomized comparison between a tunneled version and a freedom-of-choice version. J Med Internet Res 2012;14(2):e45 [FREE Full text] [doi: 10.2196/jmir.1922] [Medline: 22532074]
59. Danaher BG, McKay HG, Seeley JR. The information architecture of behavior change websites. J Med Internet Res 2005;7(2):e12 [FREE Full text] [doi: 10.2196/jmir.7.2.e12] [Medline: 15914459]
60. Miller WR, Rollnick S. Motivational Interviewing, Third Edition: Helping People Change (Applications of Motivational Interviewing). New York: The Guilford Press; 2012.
61. Boettcher J, Rozental A, Andersson G, Carlbring P. Side effects in Internet-based interventions for Social Anxiety Disorder. Internet Interventions 2014 Mar;1(1):3-11. [doi: 10.1016/j.invent.2014.02.002]
62. Burgers JS, Bailey JV, Klazinga NS, Van Der Bij AK, Grol R, Feder G. Inside guidelines: comparative analysis of recommendations and evidence in diabetes guidelines from 13 countries. Diabetes Care 2002 Nov;25(11):1933-1939. [Medline: 12401735]

63. Stone MA, Wilkinson JC, Charpentier G, Clochard N, Grassi G, Lindblad U, et al. Evaluation and comparison of guidelines for the management of people with type 2 diabetes from eight European countries. Diabetes Res Clin Pract 2010 Feb;87(2):252-260. [doi: 10.1016/j.diabres.2009.10.020] [Medline: 19932517]
64. National Institute for Health and Clinical Excellence (NICE): National Guideline C. The management of type 2 diabetes. Type 2 diabetes URL: http://www.guideline.gov/content.aspx?id=15202 [accessed 2014-06-20] [WebCite Cache ID 6QTSPsUbU]
65. American Diabetes Association. Standards of medical care in diabetes--2010. Diabetes Care 2010 Jan;33 Suppl 1:S11-S61 [FREE Full text] [doi: 10.2337/dc10-S011] [Medline: 20042772]
66. Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, COST B13 Working Group on Guidelines for Chronic Low Back Pain. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 2006 Mar;15 Suppl 2:S192-S300 [FREE Full text] [doi: 10.1007/s00586-006-1072-1] [Medline: 16550448]
67. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF): Nationale Versorgungsleitlinie Kreuzschmerz – Kurzfassung Version. 2010. URL: http://www.versorgungsleitlinien.de/themen/kreuzschmerz/pdf/nvl-kreuzschmerz-lang-4.pdf [accessed 2014-06-20] [WebCite Cache ID 6QTTXa0Az]
68. Deshpande A, Furlan A, Mailis-Gagnon A, Atlas S, Turk D. Opioids for chronic low-back pain. Cochrane Database Syst Rev 2007(3):CD004959. [doi: 10.1002/14651858.CD004959.pub3] [Medline: 17636781]
69. Furlan AD, Imamura M, Dryden T, Irvin E. Massage for low-back pain. Cochrane Database Syst Rev 2008(4):CD001929. [doi: 10.1002/14651858.CD001929.pub2] [Medline: 18843627]
70. Hayden JA, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. Cochrane Database Syst Rev 2005(3):CD000335. [doi: 10.1002/14651858.CD000335.pub2] [Medline: 16034851]
71. Heymans MW, van Tulder MW, Esmail R, Bombardier C, Koes BW. Back schools for non-specific low-back pain. Cochrane Database Syst Rev 2004(4):CD000261. [doi: 10.1002/14651858.CD000261.pub2] [Medline: 15494995]
72. Staal JB, de BR, de Vet HCW, Hildebrandt J, Nelemans P. Injection therapy for subacute and chronic low-back pain. Cochrane Database Syst Rev 2008(3):CD001824. [doi: 10.1002/14651858.CD001824.pub3] [Medline: 18646078]
73. Urquhart DM, Hoving JL, Assendelft WWJ, Roland M, van Tulder MW. Antidepressants for non-specific low back pain. Cochrane Database Syst Rev 2008(1):CD001703. [doi: 10.1002/14651858.CD001703.pub3] [Medline: 18253994]
74. Weymann N, Härter M, Petrak F, Dirmaier J. Health information, behavior change, and decision support for patients with type 2 diabetes: development of a tailored, preference-sensitive health communication application. Patient Prefer Adherence 2013;7:1091-1099 [FREE Full text] [doi: 10.2147/PPA.S46924] [Medline: 24174871]
75. Maunsell E, Lauzier S, Brunet J, Pelletier S, Osborne RH, Campbell HS. Health-related empowerment in cancer: validity of scales from the Health Education Impact Questionnaire. Cancer 2014 Oct 15;120(20):3228-3236. [doi: 10.1002/cncr.28847] [Medline: 24988944]
76. Herbert RJ, Gagnon AJ, Rennick JE, O'Loughlin JL. A systematic review of questionnaires measuring health-related empowerment. Res Theory Nurs Pract 2009;23(2):107-132. [Medline: 19558027]
77. Nolte S, Elsworth GR, Sinclair AJ, Osborne RH. The extent and breadth of benefits from participating in chronic disease self-management courses: a national patient-reported outcomes survey. Patient Educ Couns 2007 Mar;65(3):351-360. [doi: 10.1016/j.pec.2006.08.016] [Medline: 17027221]
78. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007 May;66(2):192-201. [doi: 10.1016/j.pec.2006.12.002] [Medline: 17320338]
79. Schuler M, Musekamp G, Faller H, Ehlebracht-König I, Gutenbrunner C, Kirchhof R, et al. Assessment of proximal outcomes of self-management programs: translation and psychometric evaluation of a German version of the Health Education Impact Questionnaire (heiQ™). Qual Life Res 2013 Aug;22(6):1391-1403. [doi: 10.1007/s11136-012-0268-6] [Medline: 22987145]
80. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 1995;15(1):25-30. [Medline: 7898294]
81. Buchholz A, Hölzel L, Kriston L, Simon D. Härter M: Die Decisional Conflict Scale in deutscher Sprache (DCS-D) - Dimensionale Struktur in einer Stichprobe von Hausarztpatienten. Klinische Diagnostik und Evaluation 2011;4:15-30.
82. Graham I, O'Connor A. Preparation for Decision Making Scale - User manual. 1996. l URL: http://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_PrepDM.pdf [accessed 2014-06-20] [WebCite Cache ID 6QTUmSFnj]
83. Buchholz A, Simon D. Härter M: Die Preparation for Decision-Making Scale (PDMS-G). Faktorielle und konkurrente Validität in einer Onlinestichprobe von Personen mit Kreuzschmerz und Depression. Klinische Diagnostik und Evaluation 2011;4:31-45.
84. van BS, Ferwerda M, Hoeve D, Rovers MM, Spillekom-van KS, van MH, et al. Internet-based cognitive behavioral therapy for patients with chronic somatic conditions: a meta-analytic review. J Med Internet Res 2014;16(3):e88 [FREE Full text] [doi: 10.2196/jmir.2777] [Medline: 24675372]
85. Bond GE, Burr RL, Wolf FM, Feldt K. The effects of a web-based intervention on psychosocial well-being among adults aged 60 and older with diabetes: a randomized trial. Diabetes Educ 2010;36(3):446-456. [doi: 10.1177/0145721710366758] [Medline: 20375351]

86. van Bastelaar KMP, Pouwer F, Cuijpers P, Riper H, Snoek FJ. Web-based depression treatment for type 1 and type 2 diabetic patients: a randomized, controlled trial. Diabetes Care 2011 Feb;34(2):320-325 [FREE Full text] [doi: 10.2337/dc10-1248] [Medline: 21216855]
87. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, Wu Julie H C. 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]
88. Krans H, Porta M, Keen H. Diabetes Care and Research in Europe: the St. Vincent Declaration Action Programme. Copenhagen: World Health Organization. Regional Office for Europe. (EUR/ICP/CLR 055/3); 1992. URL: http://www.codex.vr.se/texts/SVD.pdf [accessed 2014-06-20] [WebCite Cache ID 6QTUXjtY0]
89. van Eimeren B, Frees B. Ergebnisse der ARD/ZDF-Onlinestudie 2011: Drei von vier Deutschen im Netz -- ein Ende des digitalen Grabens in Sicht? media perspektiven 2011;7:334-349.
90. Centers for Disease Control and Prevention, US Department of Health and Human Services. Centers for Disease Control and Prevention (CDC): Diabetes Report Card 2012. Atlanta, GA; 2012. URL: http://www.cdc.gov/diabetes/pubs/pdf/diabetesreportcard.pdf [accessed 2014-06-20] [WebCite Cache ID 6QTW37ZfO]
91. Göbel H. [Epidemiology and costs of chronic pain syndromes exemplified by specific and unspecific low back pain]. Schmerz 2001 Apr;15(2):92-98. [Medline: 11810338]
92. Rückenschmerzen Kapitel 1.2.5.1. 2006. Gesundheitsberichterstattung des Bundes (GBE): Gesundheit in Deutschland URL: http://www.gbe-bund.de/gbe10/abrechnung.prc_abr_test_logon?p_uid=gastg&p_aid=&p_knoten=FID&p_sprache=D&p_suchstring=10410::Dorsopathien [WebCite Cache ID 6QTelF7jk]
93. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655 [FREE Full text] [Medline: 18824488]
94. Wyatt JC. When to use web-based surveys. J Am Med Inform Assoc 2000;7(4):426-429 [FREE Full text] [Medline: 10887170]
95. Best SJ, Krueger B, Hubbard C, Smith A. An Assessment of the Generalizability of Internet Surveys. Social Science Computer Review 2001 May 01;19(2):131-145. [doi: 10.1177/089443930101900201]
96. van Gelder MMHJ, 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]
97. Pouwer F, Snoek FJ, van der Ploeg HM, Heine RJ, Brand AN. A comparison of the standard and the computerized versions of the Well-being Questionnaire (WBQ) and the Diabetes Treatment Satisfaction Questionnaire (DTSQ). Qual Life Res 1998 Jan;7(1):33-38. [Medline: 9481149]
98. Booth-Kewley S, Larson GE, Miyoshi DK. Social desirability effects on computerized and paper-and-pencil questionnaires. Computers in Human Behavior 2007 Jan;23(1):463-477. [doi: 10.1016/j.chb.2004.10.020]
99. Buchanan T. Internet-based questionnaire assessment: appropriate use in clinical contexts. Cogn Behav Ther 2003;32(3):100-109. [doi: 10.1080/16506070310000957] [Medline: 16291542]
100. Brouwer W, Kroeze W, Crutzen R, de NJ, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
101. Fox S, Purcell K. California Health Care Foundation: Chronic Disease and the Internet. 2010. URL: http://www.pewinternet.org/2010/03/24/chronic-disease-and-the-internet/ [accessed 2014-06-20] [WebCite Cache ID 6QTeO3hTY]

Abbreviations

ADA: American Diabetes Association
AEM: avoidance endurance model
ANCOVA: analysis of covariance
DCS: Decisional Conflict Scale
heiQ: Health Education Impact Questionnaire
IHCA: Interactive Health Communication Application
PDMS: Preparation for Decision Making Scale

Edited by G Eysenbach; submitted 30.09.14; peer-reviewed by J Ross, S Wangberg; comments to author 06.11.14; revised version received 17.12.14; accepted 23.12.14; published 03.03.15.

Please cite as:
Weymann N, Dirmaier J, von Wolff A, Kriston L, Härter M
Effectiveness of a Web-Based Tailored Interactive Health Communication Application for Patients With Type 2 Diabetes or Chronic Low Back Pain: Randomized Controlled Trial
J Med Internet Res 2015;17(3):e53
URL: http://www.jmir.org/2015/3/e53/
doi:10.2196/jmir.3904
PMID:25736340

Original Paper

A Web-Based, Computer-Tailored Smoking Prevention Program to Prevent Children From Starting to Smoke After Transferring to Secondary School: Randomized Controlled Trial

Henricus-Paul Cremers[1], MSc; Liesbeth Mercken[1], PhD; Math Candel[2], PhD; Hein de Vries[1], PhD; Anke Oenema[1], PhD

[1]Maastricht University, Department of Health Promotion, School for Public Health and Primary Care (CAPHRI), Maastricht, Netherlands
[2]Maastricht University, Department of Methodology and Statistics, School for Public Health and Primary Care (CAPHRI), Maastricht, Netherlands

Corresponding Author:
Henricus-Paul Cremers, MSc
Maastricht University
Department of Health Promotion
School for Public Health and Primary Care (CAPHRI)
PO Box 616
Maastricht, 6200 MD Maastricht
Netherlands
Phone: 31 (0)43 388 2379
Fax: 31 (0)43 367 1032
Email: p.cremers@maastrichtuniversity.nl

Abstract

Background: Smoking prevalence rates among Dutch children increase rapidly after they transit to secondary school, in particular among children with a low socioeconomic status (SES). Web-based, computer-tailored programs supplemented with prompt messages may be able to empower children to prevent them from starting to smoke when they transit to secondary school.

Objective: The main aim of this study is to evaluate whether computer-tailored feedback messages, with and without prompt messages, are effective in decreasing children's smoking intentions and smoking behavior after 12 and 25 months of follow-up.

Methods: Data were gathered at baseline (T0), and after 12 months (T1) and 25 months (T2) of follow-up of a smoking prevention intervention program called Fun without Smokes. A total of 162 schools were randomly allocated to a no-intervention control group, an intervention prompt group, or an intervention no-prompt group. A total of 3213 children aged 10 to 12 years old participated in the study and completed a Web-based questionnaire assessing their smoking intention, smoking behavior, and sociocognitive factors, such as attitude, social influence, and self-efficacy, related to smoking. After completion, children in the intervention groups received computer-tailored feedback messages in their own email inbox and those messages could be accessed on the intervention website. Children in the prompt group received prompt messages, via email and short message service (SMS) text messaging, to stimulate them to reuse the intervention website with nonsmoking content. Multilevel logistic regression analyses were performed using multiple imputations to assess the program effects on smoking intention and smoking behavior at T1 and T2.

Results: A total of 3213 children participated in the Fun without Smokes study at T0. Between T0 and T1 a total of 1067 children out of the original 3213 (33.21%) dropped out of the study. Between T0 and T2 the number of children that did not participate in the final measurement was 1730 out of the original 3213 (53.84%). No significant program effects were observed for any of the intervention groups compared to the control group at T1 for the intention to engage in smoking—prompt, OR 0.67 (95% CI 0.30-1.50), no-prompt, OR 0.76 (95% CI 0.34-1.67)—or for smoking behavior—prompt, OR 1.13 (95% CI 0.13-9.98), no-prompt, OR 0.50 (95% CI 0.04-5.59). Similar nonsignificant program effects were found at T2 for the intention to start smoking—prompt, OR 0.78 (95% CI 0.26-2.32), no-prompt, OR 1.31 (95% CI 0.45-3.82)—and smoking behavior—prompt, OR 0.53 (95% CI 0.12-2.47), no-prompt, OR 1.01 (95% CI 0.24-4.21).

Conclusions: This study showed that the Web-based, computer-tailored feedback messages with and without prompt messages were not effective in modifying children's smoking intentions and smoking behavior as compared to no information. Future smoking prevention interventions are recommended to start closer to the age of actual smoking uptake. Furthermore, future studies on Web-based, computer-tailored smoking prevention programs should focus on assessing and controlling exposure to the educational content and the response to the prompt messages.

Trial Registration: Netherlands Trial Register NTR3116; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3116 (Archived by WebCite at http://www.webcitation.org/6O0wQYuPI).

(J Med Internet Res 2015;17(3):e59) doi:10.2196/jmir.3794

KEYWORDS

Web-based intervention; primary school children; smoking prevention; prompts; computer-tailoring

Introduction

Smoking among children and adolescents remains a public health problem [1-3], potentially leading to chronic diseases, cardiovascular diseases, or cancer at a later age [4,5]. Although smoking prevalence among Dutch primary school children at the age of 12 has decreased in the past decade from a range of 2% to 5%, down to 0% [6,7], prevalence still increases rapidly when children are in secondary school (15% of children are monthly smokers at age 14) [6]. One prevention strategy suggests starting smoking prevention programs at primary school, before positive beliefs toward smoking are formed [8]. Given the advantages of a Web-based computer-tailored approach (ie, reduced cost and an expanded reach of participants) [9-11] and the increasing use of the Internet among Dutch children (ie, 96% of 11- to 14-year-olds use the Internet) [12], Web-based computer-tailored smoking prevention programs may be helpful in decreasing smoking initiation rates among children.

Prior research already indicated that schools are an effective setting for reaching children and promoting a healthy lifestyle [13,14]. Hence, numerous smoking prevention programs have been developed to prevent the uptake or continuation of smoking among children and adolescents in school settings [15,16]. Most of the foregoing prevention programs were in-school interventions, however, schools are known to have limited time [14] and teachers are often not educated to perform health promotion activities [13,17]. Therefore, out-of-school interventions may be promising. The Octopus intervention, developed by Ausems et al, is an out-of-school, computer-tailored smoking prevention program (ie, students completed the questionnaire at school and received computer-tailored feedback letters at home sent by postal mail) for Dutch primary school children, aged 11 to 12 years. This program was reported to be more effective as compared to an in-school or combined (ie, in-school and out-of-school) program after 6 months of follow-up [18]. However, this program was provided via postal letters and did not use digital delivery channels, such as the Internet. Previous studies [9,10,19] reported Web-based, computer-tailored programs to be effective in changing unhealthy behaviors among adults, adolescents, and children. Positive program effects have also been observed in Web-based, computer-tailored smoking prevention and cessation studies among adolescents [20,21]. However, despite the positive findings of a pilot study [22], no conclusive evidence is currently available concerning the efficacy of a Web-based, computer-tailored smoking prevention program for primary school children. In this study, the paper-based Octopus intervention was translated to a Web-based, computer-tailored smoking prevention intervention and evaluated for effectiveness among children in the final grades of primary school and after their transition to secondary school. Although Web-based interventions have numerous benefits, a disadvantage is the consistently low reported adherence rates of participants [23,24]. This requires the use of optimal strategies to improve adherence. According to previous research [25-29] prompt messages may be effective in stimulating participants to reuse a Web-based intervention. However, using prompt messages in smoking prevention trials has not been studied among children before.

The aim of this study was to evaluate whether computer-tailored feedback messages, with and without prompt messages, are effective in decreasing the smoking intentions and smoking behavior of Dutch primary school children, aged 10 to 12 years, after 12 and 25 months of follow-up. Furthermore, it is known that children with a low socioeconomic status (SES) engage in smoking more often [30] and have a higher intention to start smoking [31], as compared to high SES children. Therefore, we will also assess whether SES moderates the effects of the two versions of the intervention.

Methods

Study Design, Participants, and Procedure

A cluster randomized controlled trial with three study arms was conducted to evaluate the effects of the Web-based, computer-tailored smoking prevention intervention, Fun without Smokes [32](Figure 1), on the intention to start smoking and smoking behavior. Children were followed for 2 years. During this period, children participated at three measurement sessions: baseline (T0, October to November 2011), follow-up at 12 months (T1, October to November 2012) and follow-up at 25 months (T2, November to December 2013). Primary schools, as the unit of randomization, were randomly assigned to one of the three study arms (ie, prompt, no-prompt, or control group) in a computer-determined sequence using a clustered randomization scheme.

The Fun without Smokes study was approved by the Medical Ethics Committee of the Atrium-Orbis-Zuyd Hospital (NL32093.096.11 / MEC 11-T-25) and registered in the Netherlands Trial Register (NTR3116).

Figure 1. Homepage of the Fun without Smokes website.

Leuk Zonder Peuk stopt... !

Het onderzoek Leuk Zonder Peuk is afgelopen. De afgelopen 3 jaar hebben jullie onze vragenlijsten ingevuld. We willen jullie daar hartelijk voor bedanken. De laatste prijswinnaars worden binnenkort bekend gemaakt. We wensen jullie veel plezier met jullie prijs en heel veel succes op school.

Based on a sample size calculation, 81 schools and 3240 children were needed at T0 in the Fun without Smokes study. This calculation predicted that 15% of the no-prompt group, 8% of the prompt group, and 24% of the control group would have smoked by T2. Since children had to complete the final Web-based questionnaire outside of school, it was taken into account that 60% of the participating children at T0 would have dropped out by T2. The Optimal Design program of Raudenbush [33] was used to calculate the sample size (ie, two-sided testing with Type I error rate=.05, power=.80, intraclass correlation=.04). For the present study, approximately 3500 primary schools were approached by seven Dutch Municipal Health Promotion Organizations and Maastricht University. A total of 162 primary schools participated at T0 in the intervention study with 3213 children. In Figure 2, a flowchart shows the number of participating children and schools at T0, T1, and T2. Children of all participating schools were included in the intervention trial at T0 unless they or their parents refused to be involved—of all participants, 1.7% refused to be involved at T0. The participating children in this study at T0 were Dutch primary school children in grade 7, aged 10 to 11 years. The students were followed when they entered grade 8 (T1) and when they transferred to secondary school (T2). Children had to complete a Web-based questionnaire concerning their smoking behavior, smoking intention, and sociocognitive factors related to smoking (ie, attitude, social influence, and self-efficacy expectations). After completion of the questionnaire, children in the prompt and no-prompt group received computer-tailored feedback messages via email and at the Fun without Smokes website.

At T0 and T1, children completed the Web-based questionnaire at their primary school under supervision by their teacher. At T2, children made the transition to secondary school and had to complete the Web-based questionnaire outside of school on their own initiative. During this period, all children who participated at T0 received an information letter sent by postal mail to their home address asking them to complete the Web-based questionnaire for the last time. In this information letter, children were also informed that they could win one of 500 incentives (eg, film voucher, gift card, or a subscription to a magazine) that would be raffled off among the children that filled out the final questionnaire completely. If children had provided their email address and/or mobile phone number at T0 or T1, they also received an email message and/or a short message service (SMS) text message to remind them to complete the final Web-based questionnaire.

Figure 2. Fun without Smokes study flowchart of baseline (T0), 12-month follow-up (T1), and 25-month follow-up (T2) measurement sessions.

Measurements

Overview

The primary outcome measures were intention to start smoking and the smoking behavior of the participating children. Both measures were assessed at T0, T1, and T2, and were based on self-reports using a previously used staging question [18].

Intention to Start Smoking

For intention to start smoking, children could indicate which one of seven statements would describe their intentions best. The statements were the following: “I am sure I will never start smoking,” “I think I will never start smoking,” “I think I will start smoking in the future,” “I think I will start smoking within 5 years,” “I think I will start smoking within 1 year,” “I think I will start smoking within 6 months,” and “I think I will start smoking within 1 month.” Children who indicated that they intended to start smoking anytime in the future were categorized as having the intention to smoke (scored as 1). Children who indicated they would never start smoking in the future (ie, “I am sure I will never start smoking” and “I think I will never start smoking”) were categorized as not having the intention to smoke (scored as 0).

Smoking Behavior

Smoking behavior of the children was also assessed by a staging algorithm. Children could indicate which one of nine statements described their smoking behavior best. The statements were the following: “I have never smoked, not even a puff,” “I have tried smoking, but do not smoke anymore,” “I stopped smoking, I used to smoke less than once a month,” “I stopped smoking after I smoked at least once a week,” “I try smoking occasionally,” “I smoke less than once a month,” “I do not smoke every week, but at least once a month,” “I do not smoke daily, but at least once a week,” and “I smoke at least once a day.” Children who indicated that they were smokers (ie, occasionally, monthly, weekly, or daily) were scored with a 1, otherwise a child was considered to be a nonsmoker (ie, never smoked or stopped smoking) and scored with a 0.

Background Variables

Background variables measured the age in years, gender (boy=1, girl=2), SES, and ethnicity of the participating children.

Socioeconomic Status

SES of the participating children was based on their postal code, which they had provided in the questionnaire. The Netherlands Institute for Social Research—a Dutch government agency that conducts research into the social aspects of all areas of government policy—calculated an index score for the 4-digit postal code of all Dutch inhabitants, based on their income, occupation, and education [34,35]. Based on this index score, children living in a low SES neighborhood were coded with a 0, and children living in a high SES neighborhood were coded with a 1.

Ethnicity

To assess ethnicity, children reported their parents’ and their own places of birth. In line with the guidelines of Statistics Netherlands, a child was considered to have a Western ethnic background (coded as 1) if he/she and both parents had been born in the Netherlands, another European country, North America, Oceania, Indonesia (a former Dutch colony), or Japan. Otherwise, the child was considered to have a non-Western ethnic background (coded as 2) [36].

Sociocognitive Variables

Sociocognitive variables were only measured at T0 and T1 and were derived from the I-Change Model [37].

Attitude

Attitude was measured toward both the positive and negative consequences of smoking and answers were scored on a 4-point scale. Advantageous attitudes (Cronbach alpha=.85) included nine questions concerning the benefits of smoking (eg, feeling mature, cool, or sociable), whereas disadvantageous attitudes (Cronbach alpha=.81) comprised ten questions concerning the drawbacks of smoking (eg, less physically fit, will become ill, or will get addicted).

Social Influence

Social influence was measured using perceived social norms (ie, norms about the smoking status of important people in the child’s environment) and modeling (ie, the smoking behavior of people in the child’s environment). The social norm measure included seven questions concerning the norms of the smoking status of the children’s father, mother, brother(s), sister(s), friends, best friend, and most people that are important to them. These questions were scored on a 5-point Likert scale (Cronbach alpha=.70). The modeling measure included eight questions about the smoking behavior of the children’s father, mother, brother(s), sister(s), and best friend. Furthermore, the number of friends, family members, and classmates who smoked was assessed. The modeling measure questions were also scored on a 5-point Likert scale.

Self-Efficacy

Self-efficacy expectations were measured with ten questions concerning the ability of the child to refuse cigarettes in different situations. Participants indicated how easy or difficult it was to refuse cigarettes using a 5-point Likert scale, where -2 was “very difficult” and +2 was “very easy” (Cronbach alpha=.94).

The Fun Without Smokes Intervention

Overview

All participating children received personalized log-in codes to access the Fun without Smokes website and to complete the measurements at T0, T1, and T2. The Fun without Smokes intervention consisted of computer-tailored feedback messages (ie, educational content) and prompt messages to stimulate reuse (ie, additional log-ins and views) of the Fun without Smokes website. In the coming paragraphs, the Fun without Smokes intervention will be explained for each study arm.

No-Prompt Group

After completion of the Web-based questionnaire, children randomized into the no-prompt group received three computer-tailored feedback messages on three consecutive days. The first feedback message provided advice on the children’s attitude toward smoking, the second provided advice on perceived social influence, and the third addressed children’s self-efficacy expectations concerning their ability to refuse cigarettes. The messages were sent to the children’s email addresses as a PDF file and were also available at the Fun without Smokes website. The computer-tailored feedback messages were tailored to the children’s personal (ie, name, age, and gender) and sociocognitive characteristics (ie, attitude, social influence, and self-efficacy expectations), which they had provided in the Web-based questionnaire.

Prompt Group

Similar to the children in the no-prompt group, children randomized to the prompt group received three computer-tailored feedback messages upon completion of the Web-based questionnaire. However, children in the prompt group also received six prompt messages via email and SMS every year encouraging them to reuse the Fun without Smokes website. At the Fun without Smokes website, children were able to read smoking and nonsmoking information, watch animated videos with nonsmoking content, play games concerning nonsmoking, fill out the Web-based questionnaire, or read the computer-tailored feedback messages. The aim of the website was to repeatedly expose children to nonsmoking information during the course of the year in addition to the tailored feedback messages. The content of the website changed regularly to include new information and interactive elements. The prompt messages were sent to announce a new topic related to smoking prevention that was addressed at the Fun without Smokes website (ie, games, animated videos, or new smoking and nonsmoking information). Children in the prompt and no-prompt group were able to reuse the Fun without Smokes website during the entire intervention period. However, children in the no-prompt group were not prompted to reuse the website.

Among the prompt and no-prompt group, significant differences were observed in the numbers of reuse actions of the Fun without Smokes website after the Web-based questionnaire was completed. Between T0 and T1, the mean number of reuse actions was 2.14 (SD 7.53) in the prompt group and 0.47 (SD 2.30) in the no-prompt group [38]. Between T1 and T2, the mean number of reuse actions was 0.67 (SD 2.79) in the prompt group and 0.06 (SD 0.63) in the no-prompt group.

Control Group

Children in the control group also completed the Web-based questionnaire at the Fun without Smokes website, but did not receive computer-tailored feedback or prompt messages. They were only able to use the intervention website during the three measurement sessions and not during the intervening periods. They also did not have access to the nonsmoking information or interactive elements of the website. More detailed information about the Fun without Smokes study is available elsewhere [32].

Statistical Analyses

Children in this study were nested in schools and, therefore, multilevel analyses were performed. Attrition analysis was done using multilevel logistic regression analysis to assess which factors (ie, age, gender, ethnicity, SES, advantageous or disadvantageous attitudes, social norms, modeling, self-efficacy, smoking behavior at T0, and intention to start smoking at T0) could explain the dropout between T0 and T1, and between T0 and T2. To describe the demographic characteristics of the children at T0 and potential differences concerning their primary outcome measures, general descriptive analyses were carried out (ie, means, standard deviations, and percentages) on the children that participated in the baseline measurement. Furthermore, analysis of variance (ANOVA) was used to assess whether attitude, social influence, and self-efficacy expectations differed between the study groups at T0. Additionally, to determine the number of children that changed their smoking intentions or smoking behaviors during the intervention period, basic analyses were performed to report the transition from a negative intention at T0 to a positive intention to engage in smoking at T1 and T2. The transition of nonsmoking at T0 to smoking at T1 and T2 was also analyzed using the data provided by the children that participated in the measurements at T0, T1, and T2.

Multilevel logistic regression analyses were done to assess the program effects on the prompt and no-prompt group as compared to the control group. Separate analyses were performed to assess the intervention effects on smoking intention and smoking behavior among children at T1 and T2. These analyses were adjusted for age, gender, ethnicity, SES, advantageous attitude, disadvantageous attitude, social norms, modeling, and self-efficacy. In the analyses concerning both the smoking behavior and intention to start smoking, children who smoked at T0 were excluded from the analyses. Because of the high dropout rate, multiple imputation of missing variables was applied. In the multiple imputation analyses, missing values were imputed using the intervention factor (ie, control-prompt and control-no-prompt), background variables (ie, age, gender, ethnicity, and SES), smoking intention, smoking behavior, and sociocognitive variables (ie, advantageous or disadvantageous attitudes, social norms, modeling, and self-efficacy) as predictor variables. Based on the percentage of missing data, a total of 50 datasets were imputed for T0 and T1 data, and 62 datasets were imputed for T0 and T2 data [39,40]. The program effects were analyzed by averaging the results of all the datasets (pooling). The prompt and no-prompt group were dummy coded with the control group as a reference. We set about to identify whether there were differential effects between the control, prompt, and no-prompt groups based on children's smoking intentions or smoking behavior at T1 and T2. To do this, we performed multilevel logistic regression analysis that included SES as the only interaction term to examine potential differences among children living in high- and low-SES neighborhoods. Those analyses were adjusted for age, gender, ethnicity, advantageous and disadvantageous attitude, social norms, modeling, and self-efficacy. If an interaction effect by SES was present, separate analyses were performed for the high- and low-SES groups. All analyses were performed using SPSS version 20.0 and MLwiN version 2.28 [41]. Differences were considered significant when *P*≤.05. Interaction effects were considered to be significant when *P*≤.10.

Results

Attrition Analyses

A total of 3213 children participated in the Fun without Smokes study at T0. Between T0 and T1 a total of 1067 children out of the original 3213 (33.21%) dropped out of the study. Between T0 and T2 the number of children that did not participate in the final measurement was 1730 out of the original 3213 (53.84%).

The attrition analysis showed that older children were more likely to drop out at T1 (odds ratio [OR] 1.30, 95% CI 1.01-1.67). At T2, children dropped out more frequently if they were boys (OR 0.64, 95% CI 0.55-0.75), older (OR 1.25, 95% CI 1.07-1.46), had a non-Western ethnic background (OR 1.51, 95% CI 1.17-1.94), were randomized into the prompt group as compared to the control group (OR 1.43, 95% CI 1.16-1.78), and had more smokers in their environment (OR 1.73, 95% CI 1.37-2.19).

Sample Characteristics

Table 1 shows the characteristics of the children at T0 that were randomized into the prompt, no-prompt, and control groups. Overall, slightly more girls (1625/3213, 50.58%) participated in the first measurement and the majority of the children had a Western ethnic background (2836/3213, 88.27%). At T0, 3.39% (109/3213) had a positive intention to start smoking and 1.15% (37/3213) indicated current smoking behavior. In the control group, significantly more children (*P*<.001) were of high SES, as compared to the prompt and no-prompt groups. No significant differences (*P*>.05) between study groups at T0 were observed for the other smoking-related factors (ie, attitude, social influence, and self-efficacy expectations).

Table 2 indicates that between T0 and T1, 48 children out of 2006 (2.39%) changed their negative intention to engage in smoking into a positive one. Additionally, a total of 9 children out of 2094 (0.43%) actually started to smoke after 12 months of follow-up. After 25 months, 23 children out of 1402 (1.64%) indicated they had a positive intention to start smoking and 13 children out of 1462 (0.89%) indicated current smoking behavior.

Table 1. Sample characteristics of children in each study group at T0.

Sample characteristic	Study group						
	Overall (n=3213)	Prompt (n=1207)	No-prompt (n=1003)	Control (n=1003)	*F* test (df)	*P* value	Missing values, n (%)
Age in years, mean (SD)	10.36 (0.55)	10.36 (0.55)	10.35 (0.54)	10.38 (0.55)	1.02 (2)	.36	107 (3.33)
Gender (female), n (%)	1625 (50.59)	618 (51.20)	495 (49.35)	512 (51.05)	0.44 (2)	.65	0 (0)
SES (high), n (%)	1354 (42.14)	440 (36.45)	431 (42.97)	483 (48.16)	13.21 (2)	<.001[a]	577 (17.96)
Ethnicity (Western), n (%)	2836 (88.27)	1072 (88.82)	875 (87.23)	889 (88.63)	0.75 (2)	.48	9 (0.28)
Intention to smoke at T0 (positive), n (%)	109 (3.39)	35 (2.90)	37 (3.69)	37 (3.69)	0.76 (2)	.47	75 (2.33)
Smoking behavior at T0 (smoker), n (%)	37 (1.15)	16 (1.33)	10 (1.00)	11 (1.10)	0.26 (2)	.77	34 (1.06)

[a] Significantly more children in the control group were of high SES, as compared to the prompt and no-prompt groups.

Table 2. Transition of intention to start smoking and smoking behavior between T0 and T1, and between T0 and T2.

Transition of smoking intention and behavior between measurement sessions	Study group				
	Overall	Prompt	No-prompt	Control	Missing values, n (%)
T0 to T1					
Negative to positive smoking intention, n (%)	48/2006 (2.39)	13/701 (1.85)	18/623 (2.89)	17/682 (2.49)	1023/3029 (33.77)
Nonsmoking to smoking behavior, n (%)	9/2094 (0.43)	4/727 (0.55)	2/649 (0.31)	3/718 (0.42)	1048/3142 (33.35)
T0 to T2					
Negative to positive smoking intention, n (%)	23/1402 (1.64)	7/491 (1.43)	10/446 (2.24)	6/465 (1.29)	1627/3029 (53.71)
Nonsmoking to smoking behavior, n (%)	13/1462 (0.89)	3/504 (0.59)	5/470 (1.06)	5/488 (1.02)	1680/3142 (53.47)

Intervention Effects on Smoking Intention and Smoking Behavior

The program effects on the intention to start smoking between T0 and T1, and between T0 and T2, were calculated. Multilevel logistic regression analyses indicated no significant differences at T1 between the control and prompt groups (OR 0.67, 95% CI 0.30-1.50) or between the control and no-prompt groups (OR 0.76, 95% CI 0.34-1.67). Similar nonsignificant effects concerning the intention to start smoking were observed at T2 between the control and prompt groups (OR 0.78, 95% CI 0.26-2.32) and between the control and no-prompt groups (OR 1.31, 95% CI 0.45-3.82).

The program effects on smoking behavior between T0 and T1, and between T0 and T2, were calculated. Multilevel logistic regression analyses indicated that at T1, no significant program effects were found between the control and prompt groups (OR 1.13, 95% CI 0.13-9.98). As well, no significant program effects were found between the control and no-prompt groups (OR 0.50, 95% CI 0.04-5.59). At T2, no significant differences in smoking behavior were observed between the control and prompt groups (OR 0.53, 95% CI 0.12-2.47) or between the control and no-prompt groups (OR 1.01, 95% CI 0.24-4.21).

Effect of Socioeconomic Status Within Study Groups

The results of the intervention by SES interaction showed that SES did not moderate the association between intention to start smoking and smoking behavior on the one hand and type of intervention at the other, not at T1 nor at T2 (*P*>.10). For that reason, no further subgroup analyses were performed.

Discussion

Principal Findings

The main aim of this study was to evaluate whether computer-tailored feedback messages, with and without prompt messages, are effective in decreasing smoking intention and smoking behavior of Dutch primary school children (aged 10 to 12 years) after 12 and 25 months of follow-up. Feedback messages were meant to stimulate reuse of the intervention website and increase exposure to the nonsmoking information. Since the smoking initiation and smoking intention rates were low among the study sample, findings of this study indicate that the two versions of the intervention—prompt and no-prompt messages—were not able to reduce smoking intention and smoking behavior at either of the time points.

The findings of this study are in line with the results reported by a recently published, home-based smoking prevention program [42] in which no preventive effects were found in smoking initiation among Dutch primary school children. In that study, the nonsignificant findings were attributed to the low smoking prevalence rates at this young age. This explanation is also valid for our study since the smoking intention and smoking initiation rates were low in all study arms, ranging from 1.29% to 2.89% and from 0.31% to 1.06%, respectively. The Fun without Smokes study was based on the out-of-school intervention of the Octopus study which was shown to be effective in preventing the initiation and continuation of smoking among primary school children (aged 11 to 12 years) [18]. However, the Octopus study started over a decade ago when monthly smoking prevalence rates among children were higher (2% to 5% smoking prevalence at 10 to 12 years of age), whereas nowadays those prevalence rates are 0% [7]. This decreasing trend may have been caused due to the implementation of policies to reduce smoking (ie, smoking bans in public places and workplaces, or tax increases) [43,44] or the changing norms regarding smoking [45]. Thereby, the smoking initiation rates have shifted to later ages among children and adolescents during the last ten years [7]. Hence, a smoking prevention program for 10 to 12 year olds may not be as relevant as it would be for adolescents attending secondary school when the actual uptake of smoking starts [6,7]. This suggestion has also been supported by other studies [16,20] that reported positive preventive effects for smoking prevention interventions among adolescents.

Besides the societal and smoking-climate changes during the last decade, other factors might have contributed to the nonsignificant findings of the Fun without Smokes study and should be considered in the development of future Web-based interventions. The Octopus study was a paper-based program, whereas the Fun without Smokes study was delivered via Internet. Although previous research already showed that Web-based interventions were effective in decreasing substance use [46] and improving dietary behavior [47] among children, no consistent evidence for effective Web-based smoking prevention interventions has been found yet. The educational content of both the Octopus study and the Fun without Smokes study was provided via computer-tailored feedback messages, but the content and extensiveness of the information provided was not the same. Tailored information via the Internet should be concise [48] and, therefore, the messages were edited down to brief messages that were suitable for provision through the Internet. A further dissimilarity between the studies was the difference in follow-up measurements. Both studies used multiple measurements and included both the transition periods from grade 7 to grade 8, and from primary to secondary school. In the Octopus study, there were follow-up measurements at 6, 9, 20, 30, and 36 months and in the Fun without Smokes study after 12 and 25 months. However, positive effects were observed in the Octopus study only at 6 months of follow-up [18]. This may indicate that the positive preventive effects in the Fun without Smokes study were missed because follow-up periods were too long. Moreover, in the Octopus study and the Fun without Smokes study, a similar staging question was used to assess children's smoking behavior, but the studies categorized smokers and nonsmokers differently—the Octopus study used the categories *never smokers*, *noncurrent smokers*, and *current smokers*, while the Fun without Smokes study used the categories *smokers* and *nonsmokers*. It is, however, unlikely that this difference explains the nonsignificant findings in the Fun without Smokes study, since the smoking prevalence rates of the children were low.

Despite the nonsignificant findings in the present study, it has been recommended that smoking prevention programs start before attitudes and beliefs toward smoking are formed [8], since the transition period of primary to secondary school is a crucial time when adolescents may engage in smoking [49]. Therefore, it is to be expected that smoking prevention programs may influence sociocognitive factors (ie, attitude, social influence, and self-efficacy expectations) even if no effects on smoking intentions and smoking behavior are indicated. A review by Hopfer et al [50] revealed that substance use (ie, alcohol, cigarette smoking, and other drugs) prevention programs were able to change children's attitudes, subjective norms, self-efficacy expectations, and knowledge. Similar results are observed among other studies [22,51] in which Web-based interventions demonstrated changes in behavioral determinants. However, at the final measurement session of this study, no sociocognitive factors were assessed. Therefore, no statements can be made about whether the computer-tailored feedback messages or the prompt messages were able to change sociocognitive factors.

To increase the likelihood of lower smoking intention and smoking behavior rates, this study used prompt messages to stimulate reuse of the intervention website. Although a prior study [38] showed that prompt messages were able to stimulate reuse of the Fun without Smokes website, the percentage of reuse remained low and, therefore, the exposure to the nonsmoking content was limited. This may have been caused by the content of the prompt messages. The content may not have been stimulating enough, or the children may have felt they had no reason to reuse the Fun without Smokes website. Another reason may have been the timing or frequency of the prompt messages. The first three prompt messages were sent at one-month time intervals, whereas the last three prompts were sent every two months. Previous studies [52,53] indicated that shorter time between prompts would be most effective. However, no unequivocal conclusion has been reported for the optimal prompt timing and frequency. Additionally, national reports [12,54,55] indicated that the majority of Dutch primary school children have an email address and mobile phone, though the actual use of those devices among this age group may be low. Therefore, future research should not only focus on optimal prompt timing, frequency, and content, but also on effective delivery channels to stimulate children to reuse an intervention website.

Strengths and Limitations

A strength of this study is the large sample of Dutch primary school children, since children from all regions in the Netherlands participated in the Fun without Smokes study. Another strength is the follow-up period of 25 months, which enables a long-term evaluation of the intervention effects. However, this study was also subject to some limitations. Since a large number of schools were approached for participation (n=3500) and only 162 schools were able to participate, the study sample should not be seen as representative of all schools in the Netherlands and outside the Netherlands. However, the smoking prevalence rates that were found in this study are comparable to those found in Dutch national reports [6], indicating that there were no large differences between our sample and the Dutch population of primary school children in terms of smoking prevalence. Due to differences in smoking prevalence between the Netherlands and other countries, the results may be less generalizable to countries with higher smoking prevalence rates among children aged 10 to 12 years. Another limitation may be the lack of a process evaluation of the Fun without Smokes intervention. Therefore, we did not receive in-depth information concerning the children’s opinions toward the intervention, which may explain the nonsignificant differences among the three study arms. However, children had to complete the final questionnaire at home on their own initiative. To increase the likelihood that children answered all of the questions, we chose to leave out a process evaluation. Nevertheless, it is advisable for future research to evaluate the process of the Web-based intervention, which may affect the direction of further research. A final limitation is the use of the SES index score that was based on the children’s postal code. This index score reflects the SES at a neighborhood level and not the individual SES of children. The SES measure that was used was based on the income, occupation, and education of the inhabitants living in that neighborhood and is known to correlate strongly with a more precise 6-digit postal code [35].

Conclusions

No program effects in the intervention groups (ie, prompt and no-prompt groups) were found in the Web-based, computer-tailored smoking prevention trial at 12 and 25 months of follow-up. This indicates that the Web-based, computer-tailored feedback messages were not able to change the smoking intentions and smoking behaviors of the participating children. Although the prompt messages were meant to stimulate reuse of the intervention website, resulting in an increased exposure to the nonsmoking information, no effects were observed concerning the smoking intentions and smoking behaviors of children in the prompt group. It is not completely clear why the Fun without Smokes intervention was not found to be effective. This may have been caused by the low smoking initiation rates among children and the lack of exposure to the intervention content. Future Web-based smoking prevention programs should, therefore, take place closer to the age of actual smoking uptake (ie, secondary school). Furthermore, future evaluations of smoking prevention interventions should focus on assessing and controlling exposure to the intervention content and the response to the prompt messages.

Acknowledgments

The authors would like to thank ZonMw, the Netherlands Organization for Health Research and Development (200110011) for supporting this study.

Conflicts of Interest

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

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [56].

[PDF File (Adobe PDF File), 80KB - jmir_v17i3e59_app1.pdf]

References

1. Lando HA, Hipple BJ, Muramoto M, Klein JD, Prokhorov AV, Ossip DJ, et al. Tobacco is a global paediatric concern. Bull World Health Organ 2010 Jan;88(1):2 [FREE Full text] [doi: 10.2471/BLT.09.069583] [Medline: 20428341]
2. Mackay J, Eriksen M. The Tobacco Atlas. Geneva, Switzerland: World Health Organization; 2002.
3. WHO Policy Series: Health Policy for Children and Adolescents Issue 1. International Report. Copenhagen, Denmark: WHO Regional Office for Europe; 2000. Health and health behaviour among young people URL: http://www.euro.who.int/__data/assets/pdf_file/0006/119571/E67880.pdf [accessed 2015-02-23] [WebCite Cache ID 6WZPCxPLi]
4. WHO. The World Health Report 1999. Geneva, Switzerland: World Health Organization; 1999. Combating the tobacco epidemic URL: http://www.who.int/whr/1999/en/whr99_ch5_en.pdf [accessed 2015-02-23] [WebCite Cache ID 6WZPddcNx]

5. Hipple B, Lando H, Klein J, Winickoff J. Global teens and tobacco: a review of the globalization of the tobacco epidemic. Curr Probl Pediatr Adolesc Health Care 2011 Sep;41(8):216-230. [doi: 10.1016/j.cppeds.2011.02.010] [Medline: 21821205]
6. STIVORO. The Hague, The Netherlands: Stichting Volksgezondheid en Roken [Dutch Foundation on Smoking and Health]; 2012. Smoking, the firm facts: Youth 2012 (Roken, de harde feiten: Jeugd 2012) URL: http://stivoro.nl/wp-content/uploads/2012/docs/factsheets/jongeren/Fact%20sheet%20RJM%202012.pdf [accessed 2015-02-23] [WebCite Cache ID 6WZPyE1Ju]
7. STIVORO. Vallen en opstaan (Trial and error). In: Jaarverslag 2012 (Annual Report 2012). The Hague, The Netherlands: Stichting Volksgezondheid en Roken [Dutch Foundation on Smoking and Health]; 2013.
8. NHS Centre for Reviews and Dissemination. Effective Health Care Bulletin. London, UK: Royal Society of Medicine Press; 1999 Oct. Preventing the uptake of smoking in young people URL: http://www.york.ac.uk/inst/crd/EHC/ehc55.pdf [accessed 2015-02-23] [WebCite Cache ID 6WZQmchdJ]
9. Neville LM, O'Hara B, Milat AJ. Computer-tailored dietary behaviour change interventions: a systematic review. Health Educ Res 2009 Aug;24(4):699-720 [FREE Full text] [doi: 10.1093/her/cyp006] [Medline: 19286893]
10. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
11. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
12. De Haan J. NL Kids Online. The Hague, The Netherlands: Sociaal en Cultureel Planbureau (The Netherlands Institute of Social Research); 2010:1-58.
13. Leurs MT, Bessems K, Schaalma HP, de Vries H. Focus points for school health promotion improvements in Dutch primary schools. Health Educ Res 2007 Feb;22(1):58-69 [FREE Full text] [doi: 10.1093/her/cyl043] [Medline: 16766607]
14. Boot NM, van Assema P, Hesdahl B, Leurs M, de Vries NK. Tijdschrift voor Gezondheidswetenschappen (TSG). Houten, The Netherlands: Bohn Stafleu van Loghum (BSL), Springer Media; 2010 Nov 12. Gezondheidsbevordering en voortgezet onderwijs: verstandshuwelijk of echte liefde? URL: http://www.gezondeschool.info/object_binary/o9457_artikel-TSG-Boot.pdf [accessed 2015-02-23] [WebCite Cache ID 6WZR6IZuR]
15. Thomas RE, McLellan J, Perera R. School-based programmes for preventing smoking. Cochrane Database Syst Rev 2013;4:CD001293. [doi: 10.1002/14651858.CD001293.pub3] [Medline: 23633306]
16. Crone MR, Reijneveld SA, Willemsen MC, van Leerdam FJ, Spruijt RD, Sing RA. Prevention of smoking in adolescents with lower education: a school based intervention study. J Epidemiol Community Health 2003 Sep;57(9):675-680 [FREE Full text] [Medline: 12933772]
17. Boot N, Assema PV, Hesdahl B, Vries ND. Professional assistance in implementing school health policies. Health Educ 2010;110(4):294-308. [doi: 10.1108/09654281011052646]
18. Ausems M, Mesters I, van Breukelen G, De Vries H. Short-term effects of a randomized computer-based out-of-school smoking prevention trial aimed at elementary schoolchildren. Prev Med 2002 Jun;34(6):581-589. [doi: 10.1006/pmed.2002.1021] [Medline: 12052017]
19. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
20. de Josselin de Jong S, Candel M, Segaar D, Cremers HP, de Vries H. Efficacy of a Web-based computer-tailored smoking prevention intervention for Dutch adolescents: randomized controlled trial. J Med Internet Res 2014 Mar;16(3):e82 [FREE Full text] [doi: 10.2196/jmir.2469] [Medline: 24657434]
21. Prokhorov AV, Kelder SH, Shegog R, Conroy JL, Murray N, Peters R, et al. Project ASPIRE: an Interactive, Multimedia Smoking Prevention and Cessation curriculum for culturally diverse high school students. Subst Use Misuse 2010 May;45(6):983-1006. [doi: 10.3109/10826080903038050] [Medline: 20397881]
22. Shegog R, McAlister AL, Hu S, Ford KC, Meshack AF, Peters RJ. Use of interactive health communication to affect smoking intentions in middle school students: a pilot test of the "Headbutt" risk assessment program. Am J Health Promot 2005;19(5):334-338. [Medline: 15895535]
23. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
24. Wangberg SC, Bergmo TS, Johnsen JA. Adherence in Internet-based interventions. Patient Prefer Adherence 2008 Feb;2:57-65 [FREE Full text] [Medline: 19920945]
25. Woodall WG, Buller DB, Saba L, Zimmerman D, Waters E, Hines JM, et al. Effect of emailed messages on return use of a nutrition education website and subsequent changes in dietary behavior. J Med Internet Res 2007;9(3):e27 [FREE Full text] [doi: 10.2196/jmir.9.3.e27] [Medline: 17942389]
26. Svensson M, Svensson T, Hansen AW, Trolle Lagerros Y. The effect of reminders in a web-based intervention study. Eur J Epidemiol 2012 May;27(5):333-340. [doi: 10.1007/s10654-012-9687-5] [Medline: 22531973]
27. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]

28. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]
29. Schneider F, van Osch L, Schulz DN, Kremers SP, de Vries H. The influence of user characteristics and a periodic email prompt on exposure to an internet-delivered computer-tailored lifestyle program. J Med Internet Res 2012;14(2):e40 [FREE Full text] [doi: 10.2196/jmir.1939] [Medline: 22382037]
30. Conrad KM, Flay BR, Hill D. Why children start smoking cigarettes: predictors of onset. Br J Addict 1992 Dec;87(12):1711-1724. [Medline: 1490085]
31. Cremers HP, Oenema A, Mercken L, Candel M, de Vries H. Explaining socio-economic differences in intention to smoke among primary school children. BMC Public Health 2014;14:191 [FREE Full text] [doi: 10.1186/1471-2458-14-191] [Medline: 24555819]
32. Cremers HP, Mercken L, Oenema A, de Vries H. A web-based computer-tailored smoking prevention programme for primary school children: intervention design and study protocol. BMC Public Health 2012;12:277 [FREE Full text] [doi: 10.1186/1471-2458-12-277] [Medline: 22490110]
33. Raudenbush SW. Statistical analysis and optimal design for cluster randomized trials. Psychol Methods 1997;2(2):173-185. [doi: 10.1037/1082-989X.2.2.173]
34. The Netherlands Institute for Social Research. Sociaal en Cultureel Planbureau. Statusscores URL: http://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores [accessed 2013-10-09] [WebCite Cache ID 6KF4XgLV9]
35. Knol FA, Boelhouwer J, Veldheer V. Statusontwikkeling van Wijken in Nederland 1998-2010 (Status Development of Districts in the Netherlands 1998-2010). The Hague, The Netherlands: The Netherlands Institute for Social Research; 2012:1-97.
36. Keij I. Standaarddefinitie Allochtonen (Standard Definition Immigrants). Voorburg, The Netherlands: Statistics Netherlands; 2000:24-25.
37. de Vries H, Kremers SP, Smeets T, Brug J, Eijmael K. The effectiveness of tailored feedback and action plans in an intervention addressing multiple health behaviors. Am J Health Promot 2008;22(6):417-425. [doi: 10.4278/ajhp.22.6.417] [Medline: 18677882]
38. Cremers HP, Mercken L, Crutzen R, Willems P, de Vries H, Oenema A. Do email and mobile phone prompts stimulate primary school children to reuse an Internet-delivered smoking prevention intervention? J Med Internet Res 2014 Mar;16(3):e86 [FREE Full text] [doi: 10.2196/jmir.3069] [Medline: 24642082]
39. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007 Sep;8(3):206-213. [doi: 10.1007/s11121-007-0070-9] [Medline: 17549635]
40. Bodner TE. What improves with increased missing data imputations? Structural Equation Modeling: A Multidisciplinary Journal 2008 Oct 22;15(4):651-675. [doi: 10.1080/10705510802339072]
41. Rasbash J, Charlton C, Browne WJ, Healy M, Cameron B. MLwiN version 2.28. In: University of Bristol. Bristol, UK: Centre for Multilevel Modelling, University of Bristol; 2013.
42. Hiemstra M, Ringlever L, Otten R, van Schayck OC, Jackson C, Engels RC. Long-term effects of a home-based smoking prevention program on smoking initiation: a cluster randomized controlled trial. Prev Med 2014 Mar;60:65-70. [doi: 10.1016/j.ypmed.2013.12.012] [Medline: 24355576]
43. Nagelhout GE, Willemsen MC, de Vries H. The population impact of smoke-free workplace and hospitality industry legislation on smoking behaviour. Findings from a national population survey. Addiction 2011 Apr;106(4):816-823. [doi: 10.1111/j.1360-0443.2010.03247.x] [Medline: 21182553]
44. Nagelhout GE, Levy DT, Blackman K, Currie L, Clancy L, Willemsen MC. The effect of tobacco control policies on smoking prevalence and smoking-attributable deaths. Findings from the Netherlands SimSmoke Tobacco Control Policy Simulation Model. Addiction 2012 Feb;107(2):407-416. [doi: 10.1111/j.1360-0443.2011.03642.x] [Medline: 21906197]
45. Nyborg K, Rege M. On social norms: the evolution of considerate smoking behavior. J Econ Behav Organ 2003 Nov;52(3):323-340. [doi: 10.1016/S0167-2681(03)00031-3]
46. Champion KE, Newton NC, Barrett EL, Teesson M. A systematic review of school-based alcohol and other drug prevention programs facilitated by computers or the internet. Drug Alcohol Rev 2013 Mar;32(2):115-123. [doi: 10.1111/j.1465-3362.2012.00517.x] [Medline: 23039085]
47. Hamel LM, Robbins LB. Computer- and web-based interventions to promote healthy eating among children and adolescents: a systematic review. J Adv Nurs 2013 Jan;69(1):16-30. [doi: 10.1111/j.1365-2648.2012.06086.x] [Medline: 22757605]
48. Prins RG. Environmetal influences on physical activity among adolescents: studies on determinants and intervention strategies. Rotterdam, The Netherlands: Erasmus University Rotterdam; 2012.
49. Crone MR, Spruijt R, Dijkstra NS, Willemsen MC, Paulussen TG. Does a smoking prevention program in elementary schools prepare children for secondary school? Prev Med 2011 Jan;52(1):53-59 [FREE Full text] [doi: 10.1016/j.ypmed.2010.11.003] [Medline: 21078340]
50. Hopfer S, Davis D, Kam JA, Shin Y, Elek E, Hecht ML. A review of elementary school-based substance use prevention programs: identifying program attributes. J Drug Educ 2010;40(1):11-36. [Medline: 21038761]

51. Oenema A, Brug J, Lechner L. Web-based tailored nutrition education: results of a randomized controlled trial. Health Educ Res 2001 Dec;16(6):647-660 [FREE Full text] [Medline: 11780705]
52. Schneider F, de Vries H, Candel M, van de Kar A, van Osch L. Periodic email prompts to re-use an internet-delivered computer-tailored lifestyle program: influence of prompt content and timing. J Med Internet Res 2013;15(1):e23 [FREE Full text] [doi: 10.2196/jmir.2151] [Medline: 23363466]
53. Skov-Ettrup LS, Ringgaard LW, Dalum P, Flensborg-Madsen T, Thygesen LC, Tolstrup JS. Comparing tailored and untailored text messages for smoking cessation: a randomized controlled trial among adolescent and young adult smokers. Health Educ Res 2014 Apr;29(2):195-205. [doi: 10.1093/her/cyt112] [Medline: 24399268]
54. De Graaf I, Naaborgh L. Uw Tiener en het Internet (Your Teenager and the Internet). Utrecht, The Netherlands: Trimbos Instituut Netherlands (Institute of Mental Health and Addiction); 2010:1-10.
55. Duimel M, Pijpers R, Borgdorff MW. Mijn kind online (My child online). Zoetermeer, The Netherlands: Mijn Kind Online Foundation; 2012 Mar. Hey, what's app? URL: http://mijnkindonline.nl/sites/default/files/uploads/MKO%20Mobiel%20Hey%2C%20what%27s%20app%20-%20DEF2.pdf [accessed 2015-02-23] [WebCite Cache ID 6WZUEtGOS]
56. Eysenbach G, Consort- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

ANOVA: analysis of variance
OR: odds ratio
SES: socioeconomic status
SMS: short message service

Edited by G Eysenbach; submitted 18.08.14; peer-reviewed by B Bock, L Akers; comments to author 25.09.14; revised version received 06.11.14; accepted 23.11.14; published 09.03.15.

Please cite as:
Cremers HP, Mercken L, Candel M, de Vries H, Oenema A
A Web-Based, Computer-Tailored Smoking Prevention Program to Prevent Children From Starting to Smoke After Transferring to Secondary School: Randomized Controlled Trial
J Med Internet Res 2015;17(3):e59
URL: http://www.jmir.org/2015/3/e59/
doi:10.2196/jmir.3794
PMID:25759248

Original Paper

Expectations Among Patients and Health Professionals Regarding Web-Based Interventions for Depression in Primary Care: A Qualitative Study

Jesús Montero-Marín[1,2,3,4], PhD; Javier Prado-Abril[5]; Cristina Botella[4,6], PhD; Fermin Mayoral-Cleries[7], PhD; Rosa Baños[4,8,9], PhD; Paola Herrera-Mercadal[2,3,4]; Pablo Romero-Sanchiz[7], PhD; Margalida Gili[2,10], PhD; Adoración Castro[2,10]; Raquel Nogueira[11], PhD; Javier García-Campayo[2,3,4,12], PhD

[1]Faculty of Health and Sports, University of Zaragoza, Huesca, Spain

[2]REDIAPP “Research Network on Preventative Activities and Health Promotion” (RD06/0018/0017), Zaragoza, Spain

[3]IIS Aragon: Health Research Institute, Zaragoza, Spain

[4]Red de Excelencia PROMOSAM (PSI2014-56303-REDT). MINECO, Madrid, Spain

[5]Psychosocial Rehabilitation and Recovery Center, Mental Health Service, Santa Maria Hospital, GSS “Gestió de Serveis Sanitaris", Lerida, Spain

[6]Department of Basic Psychology, Clinic and Psychobiology, University Jaume I, Castellon, Spain

[7]UGC Salud Mental, Hospital Regional Universitario, Instituto de Biomedicina, Málaga, Spain

[8]CIBER Fisiopatología Obesidad y Nutrición (CIBERObn), Instituto Salud Carlos III, Santiago de Compostela, Spain

[9]Department of Personality, Evaluation and Psychological Treatment, University of Valencia, Valencia, Spain

[10]Institut Universitari d’Investigació en Ciències de la Salut (IUNICS), University of Balearic Islands, Mallorca, Spain

[11]Department of Personality, Assesment and Psychological Treatment, University of Málaga, Málaga, Spain

[12]Psychiatry Service, Miguel Servet University Hospital, Zaragoza, Spain

Corresponding Author:
Javier García-Campayo, PhD
Psychiatry Service
Miguel Servet University Hospital
Avda Isabel La Católica 1
Zaragoza, 50009
Spain
Phone: 34 976253621
Fax: 34 976254006
Email: jgarcamp@gmail.com

Abstract

Background: One-quarter of the world’s population will suffer from depression symptoms at some point in their lives. Mental health services in developed countries are overburdened. Therefore, cost-effective interventions that provide mental health care solutions such as Web-based psychotherapy programs have been proposed.

Objective: The intent of the study was to identify expectations regarding Web-based psychotherapy for the treatment of depression in primary care among patients and health professionals that might facilitate or hinder its effects.

Methods: The expectations of untreated patients and health professionals were examined by means of interviews and focus groups. There were 43 participants (20 patients with mild and moderate levels of depression, 11 primary care physicians, and 12 managers; 22 of them for interviews and 21 for groups). A thematic content analysis from the grounded theory for interviews, and an analysis of the discursive positions of participants based on the sociological model for groups were performed. Interpretations were achieved by agreement between three independent analysts.

Results: All participants showed a good general acceptance of Web-based psychotherapy, appreciating possible advantages and improvements. Patients, physicians, and managers shared the same conceptualization of their expectations, although highlighting different aspects. Patients focused on the need for individualized and personalized interaction, while professionals highlighted the need for the standardization of the program. Physicians were concerned with extra workload, while managers were worried about optimizing cost-effectiveness.

Conclusions: Expectations of the different participants can conflict with each other. Finding a balanced position among them is needed if we are to harmoniously implement effective Web-based interventions for depression in routine clinical practice.

(J Med Internet Res 2015;17(3):e67) doi:10.2196/jmir.3985

KEYWORDS

depression; computer-delivered psychotherapy; qualitative methods; expectations

Introduction

Depression currently ranks fourth in disorders with the highest disease burden, and in the coming five years, it will cause the second-highest level of disability worldwide [1-3]. It has been said that 25% of the population will experience depression symptoms at some point in their lives [4]. Depression brings with it significant economic, personal, inter-personal, and societal costs, and is associated with important impairments in quality of life and increased mortality [5-8].

Antidepressant medication is the most common form of treatment for depressive patients in primary care [9], but adherence is low [10,11]. In fact, two-thirds of depressed primary care patients would prefer psychological treatments to pharmacotherapy [12-17]. Psychological treatments obtain better results than pharmacological ones in adherence, relapse, recovery rates, and chronicity, reducing the number of physician visits and hospital days [18,19]. Despite their effectiveness, face-to-face psychotherapy interventions face some serious limitations. These include difficulties in delivering interventions to the community due to strong barriers in providing psychotherapies in current models of service delivery; constraints in health care resources; limited availability of clinicians, especially in rural areas; and low participation rates, even if access to those interventions is at little or no cost [20-23]. Mental health services in developed countries are overburdened, and given the lack of intended resources, are in need of cost-effective alternatives [24].

Attempts to overcome barriers to access have been addressed recently through the evolution of a new understanding in mental health care that recognizes low-intensity services, such as bibliotherapy, psychoeducation, and Web-based interventions [25]. Low-intensity interventions signify treatments that limit specialist time or use this time in a cost-effective manner and are provided, when appropriate, as a first option before referral to high-intensity interventions in which the presence of the therapist is higher [26]. For example, when offered by an integrated health care provider using a stepped-care model, it has been suggested that the first choice should be the least intensive intervention that is appropriate for a person, enabling people to step up or down the pathway according to changing needs and in response to treatment [18]. The Internet offers a new way of providing psychological treatment for common mental health problems such as depression [27,28], and may attract people who do not make use of traditional mental health services [23]. Web-based treatment has been described as an intervention that is operationalized and transformed for delivery via the Internet [29]. It is usually highly structured and can be roughly divided into unguided and guided interventions, with an identified therapist [27,28,30]. Guided Internet interventions are considered to be more effective in reducing depression symptoms than unguided ones [31,32], with average levels of adherence estimated at 26% in unguided interventions and 72% in guided interventions [33]. Web-based interventions may also resemble traditional psychotherapy with scheduled sessions [34], but a general characteristic is that patients are reached from a distance and that therapist time is reduced compared with face-to-face treatments [35].

Several studies show that Web-based psychological treatments are effective for the treatment of depression, in both unguided and guided interventions, although the effects are more favorable in the latter case [33,36-38]. It has even been said that Web-based treatments with clinician assistance as brief as 1 hour per patient can work as well as face-to-face therapy [39]. Web-based psychotherapy could be viable not only in the context of specialist services, but also at the primary care level [40,41], and could be particularly recommended as a first treatment step for self-help in treating depression before visiting a psychiatrist or psychologist [42]. However, in spite of its efficiency, there are some limitations to its systematic use. One of the main ones is that attrition rates could be substantial [28,43-45]. In most cases, patients withdraw for personal reasons, not because of problems with the technology or the social environment. Interestingly, these kinds of interventions are generally acceptable for both patients and professionals, although the attitudes of professionals toward this type of psychotherapy seem to be more negative than the attitudes of patients themselves [38,46,47].

Therefore, if we are to develop and deploy this kind of program, maximizing acceptance and using it as a part of public health services, as in this case, we must be sensitive to the different views and attitudes of all participants in the therapeutic process. There may be different types of expectations or requirements according to the perspective of each type of participant. Given this background, the aim of the present study was to identify expectations among both patients and health professionals that may serve as barriers or facilitators to engagement with treatment when performing Web-based psychotherapy for depression in primary care, in order to increase adherence and effectiveness, improving the care offered to these types of patients.

Methods

A qualitative design was used to collect information from a wide range of purposefully and theoretically guided samples of depressed patients and health professionals. Both in-depth interviews and discussion groups were used to access the subjectivity and the processes involved in generating expectations [48]. In-depth interviews were carried out by a single interviewer and focus groups were moderated by an interviewer and an observer, both of them female psychologists and researchers, with previous experience in the field and no previous contact with participants.

Patients were recruited from the Spanish autonomous region of Aragon during their visits to primary care, and diagnosis was established using the MINI Neuropsychiatric Interview in its Spanish version [49]. Patients were selected when they reported mild or moderate depression symptoms according to cut-off points of the Beck Depression Inventory [50]. Family physicians and managers were chosen from the catchment area, based on whether their workplace was rural or urban, and contacted by telephone. Participants were selected with consideration given to the variables of age, gender, residential setting (“urban”, “rural”), and affinity for technology (by means of the question: what is your level of affinity for new technologies?, with “high”, “intermediate”, “low” as possible responses), in order to gather plentiful and varied information. Participants had no previous knowledge of the researchers or the project. No patients refused to participate in the interviews, although some of them did not attend the discussion groups. No physicians or managers refused to participate; however, it was difficult to establish meeting dates because of their work schedule. Table 1 outlines the main characteristics of the 43 participants (20 patients, 11 physicians, 12 managers).

A standardized protocol was designed to guide the interviews and groups, including the preparation of a topic list (Table 2) to be addressed, with previously tested, open suggestions that could be of interest.

The objectives of the study were indirectly raised and questions asked about the topics in an open and progressive way. The interviewer and/or moderator were introduced to the participants as research psychologists and assumed a minimally orientative role, limiting their interventions to addressing the topics in the script. The setting for data collection was a neutral room in the hospital where the project was conducted, without the presence of non-participants. No individual or group interviews lasted more than 90 minutes. They were digitally audio-recorded and transcribed to obtain the final set of qualitative data for the analysis, which were revised by participants and added to the field notes made after interviews and groups. Participants provided written informed consent to participate in the study. There were no repeat interviews. This study forms part of a mixed-method research project that includes a randomized controlled trial with different participants [51]. The project was approved by the Regional Ethics Committee of Aragon, Spain.

A thematic content analysis was performed from grounded theory in order to develop and to define until saturation the emergent categories of analysis derived from the interview data [52-54]. Secondly, a sociological approach that analyzed discursive positions of participants in focal groups was used to complement the first approximation [55,56]. All the analyses were developed using Maxqda-2007 software, in an iterative way by agreement between the three researchers (JMM, JPA, JGC), and discussed with those in charge of the interviews and groups (PR, PHM) [56]. Moreover, participants provided their agreement with the interpretations. This methodological triangulation was able to increase consistency and rigor by combining multiple techniques, maximizing the breadth and depth of perspectives. The complete details of the study protocol are reported elsewhere [57].

Table 1. Characteristics of participants (n=43).[a]

Stratification variables		Interviews			Groups		
		Patients n=12	Physicians n=5	Managers n=5	Patients n=8	Physicians n=6	Managers n=7
Age							
	20-40 years	3 (7%)	2 (5%)	0 (0%)	2 (5%)	3 (7%)	1 (2%)
	41-60 years	6 (14%)	2 (5%)	5 (12%)	5 (12%)	2 (5%)	5 (12%)
	>60 years	3 (7%)	1 (2%)	0 (0%)	1 (2%)	1 (2%)	1 (2%)
Sex							
	Male	5 (12%)	1 (2%)	4 (9%)	1 (2%)	2 (5%)	2 (5%)
	Female	7 (16%)	4 (9%)	1 (2%)	7 (16%)	4 (9%)	5 (12%)
Residential setting							
	Urban	9 (21%)	3 (7%)	4 (9%)	5 (12%)	4 (9%)	5 (12%)
	Rural	3 (7%)	2 (5%)	1 (2%)	3 (7%)	2 (5%)	2 (5%)
Level of education							
	Primary	4 (9%)	0 (0%)	0 (0%)	3 (7%)	0 (0%)	0 (0%)
	Secondary	5 (12%)	0 (0%)	0 (0%)	3 (7%)	0 (0%)	0 (0%)
	University	3 (7%)	5 (12%)	5 (12%)	2 (5%)	6 (14%)	7 (16%)
Affinity for technology							
	High	6 (14%)	2 (5%)	2 (5%)	3 (7%)	4 (9%)	5 (12%)
	Intermediate	3 (7%)	3 (7%)	3 (7%)	4 (9%)	2 (5%)	2 (5%)
	Low	3 (7%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)	0 (0%)

[a] Out of the total of participating patients, 40% (17/43) had university studies and 60% (26/43) had high school education. Half of the patients (10/20) presented with mild symptoms of depression, and the other half (10/20) presented with moderate symptoms.

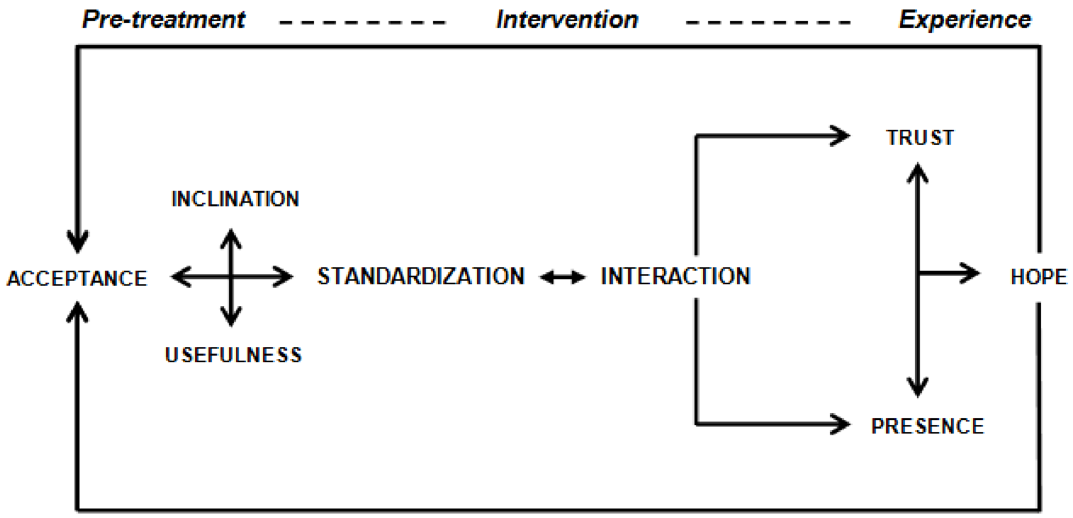
Table 2. Questions asked and topic list.

Questions asked	Topic list
What is your relationship with technology? Have you received any training?	
	Skills
	Resources
	Routine use
	Reasons
	Like / dislike
Do you think new technologies can help to improve people's lives?	
	Relationship machine/man
	Face-to-face vs Web-based
	Support tool vs replacement tool
What do you think about the use of Web-based therapy to support the actions of health professionals? And as further therapies?	
	Possibilities of Web-based therapy
	Virtual and anonymous relationship
	Health professional vs program
What about the possibility of offering Web-based therapy for depressive disorders? Do you see it as feasible?	
	What would you ask/expect?
	Information
	Difficulties
	Aspects that could be helpful
	Process from primary care
	Professional support

Results

The patients, primary care physicians, and managers showed a series of expectations regarding engagement with the use of Web-based psychotherapy for the treatment of depression, which can be divided into three main groups: aspects related to pre-treatment conditions, such as their initial acceptance, their general inclination to make use of the treatment, and its perceived usefulness; aspects related to the particular characteristics of this type of intervention, such as its standardization and the processes of interacting with the program; and the representations made of what could be seen as the expected therapeutic experience, such as experiences of trust, presence, and hope. These groupings can be arranged according to the moment in the therapeutic process to which they correspond (Figure 1).

Figure 1. Engagement expectations regarding online psychotherapy for depression.



Conceptualization of the Discourse

The general acceptance of Web-based psychotherapy for depression depended on prior conditions, such as the inclination to use new technologies, which in turn consisted of subtopics such as “skill level”, “reasons for use”, “available technical means”, and “dedication”. Acceptance also depended on its usefulness as an instrument for psychotherapy, and usefulness consisted of subtopics such as “facilitating communication” and “access to information”, which in the case of physicians took on the meaning of consulting the records or background of patients:

> *We have limited access to sources of information during a consultation. The computer is the tool that provides access to the most updated, complete, and verified information, depending on the platforms you consult. It’s an important tool for both receiving and conveying information. It also allows us to have the patient’s records online and on hand.* [physician, female, 50 years old, high affinity]

Whereas, for patients, it meant having the possibility of consulting information related to their conditions and its treatment:

> *The machine [sic] should explain the diagnosis of the disease or anomaly a patient has and quickly help to find solutions for daily life.* [patient, male, 65 years old, high affinity]

Other subtopics under usefulness were “adding convenience”, which for managers meant providing patient access to health care resources:

> *Maybe it will make professionals more accessible to users; it could become more convenient to avoid having to travel, and interfere less with people’s work because they can do it at home.* [manager, female, 51 years old, high affinity]

Whereas, patients understood it to mean the possibility of improving reminders for control of and compliance with the instructions given by the professional:

> *The program should have some mechanism to do online activities and revision; for example, what the doctor told me, to refresh the information so as not to forget it.* [patient, female, 59 years old, low affinity]

Finally, usefulness also included “enabling reflectiveness”, understood by patients to mean the opportunity to reflect on their dysfunctional situation:

> *Does the program have a written format? If that’s the case, writing means thinking and rethinking; it’s like a kind of meditation or awareness of the day, of how things are going, how they develop. It gives you perspective, and that’s positive.* [patient, female, 52 years old, intermediate affinity]

Physicians particularly focused on their awareness of diagnostic impressions resulting from explorations:

> *The advantage it may have is that you can prepare beforehand the information you’re going to introduce: the questions; results of a particular test; the answers; and you can leave a record of your recommendations, so that there are no misinterpretations. This encourages patients to elaborate on what took place during the consultation once outside, in a more relaxed atmosphere like that of their own homes.* [physician, female, 42 years old, high affinity]

Acceptance depended on the processes required to standardize the program, which consisted of the subtopics “learning process”, “definition of goals”, “complementary tool”, “prior dissemination”, “service offered” (particularly for managers), “professionals involved” (especially highlighted by physicians), "patients served", and "activities to carry out". The last was understood by patients according to their own needs:

> *I do yoga on Wednesday and I feel great. Maybe it could include relaxation techniques and different strategies for different states of mind; some days are better than others.* [patient, male, 46 years old, intermediate affinity]

Whereas physicians underscored the importance of being able to use their own paradigmatic coordinates:

> *In our specialty, we refer to those of a cognitive behavioral model. Other more introspective or psychodynamic models revolve around the doctor-patient relationship, which cannot be applied using technology.* [physician, male, 51 years old, intermediate affinity]

In turn, the necessary processes for standardization were associated with the idea of specific forms of interaction between the participants and the program. This idea of interaction combined subtopics such as "feeling of security", "program universality", "possibilities for expression", "supervision by a therapist", and "individualized attention for patients". The last of these had special relevance for patients, in the sense of personal and close contact:

> *Rapport. You have to be on the same wavelength and make a connection. You need warmth—they're personal matters—personal contact, personal proximity.* [patient, male, 64 years old, intermediate affinity]

For professionals, it was simply focused on the need to adapt to the individual characteristics of each patient:

> *If it were also a little... if the doctor could tailor the treatment a little and if it were based on a shared model of decision-making, with a certain proximity between the professional and patient.* [physician, female, 54 years old, intermediate affinity]

Finally, if these forms of individualized interaction, associated with the processes for standardization of the program, created expectations of a therapeutic experience based on trust and on presence, the discourse would pave the way to hope, which could lead to positive attitudes once again linked to acceptance as a necessary condition for engagement with the therapeutic process.

Table 3 shows the definitions of the topics and subtopics that comprise the conceptualization of the discourse presented above, and Multimedia Appendix 1 provides examples of participants' responses to all of them.

Table 3. Conceptualization of discourse and definitions.

Topic/subtopic	Definition
Acceptance	
Acceptance	Favorable attitude to therapeutic use of new technologies (NTs), with responsibility to accept the program with the necessary effort, recognizing its possible benefits.
Inclination	
Skill level	Level of knowledge and skill when using NTs.
Reasons for use	Reasons for the use of NTs, for work, leisure, etc.
Available technical means	Available technical means and possibilities of updating them.
Dedication	Time devoted NTs and form of use, taking into account the risks of patterns of abusive use in some individuals.
Usefulness as an instrument for therapy	
Facilitating communication	NTs as alternative means of communicating and forming relationships with others, of keeping in contact, alerting, and giving notification of results.
Access to information	Possibility of consulting background information and patient records (physicians) and important matters related to medical conditions and treatments (patients).
Adding convenience	Solutions to problems with daily clinical practice, such as facilitating access (managers) or reminders for, control of, and compliance with instructions (patients)
Enabling reflectiveness	Possibility of thinking/reconsidering through the use of writing, making possible a reflective attitude to their dysfunctional situation (patients), or of taking note of exploration and diagnostic impressions (physicians).
Program standardization	
Learning process	Ease of learning the program and ease of use.
Definition of goals	Structured definition of the goals to meet with the program and their timeframe.
Complementary tool	Combined and complementary use of Web-based and traditional therapies, without the intention of replacing the presence of professionals.
Prior dissemination	Dissemination of information about the program in order to overcome possible resistance to change based on old habits.
Service offered	Implementation in health care delivery system, with a structure focused on a new model of health care, and with support from health care providers, IT experts, and administrative staff that allow referrals from primary care in coordination with mental health services.
Professionals involved	Recruitment of professionals with a positive attitude toward NTs who have specific skills depending on the characteristics of this type of intervention.
Patients served	Recruitment of patients according to their willingness to use NTs, low or mild severity of their disorder, and their demand for greater information or anonymity.
Activities to carry out	Incorporation of modules that appeal to patients, arouse their interest, and meet their needs (patients), with real possibilities of implementation within paradigmatic coordinates (physicians).
Interaction processes	
Feeling of security	Combination of safeguarding anonymity, data protection, and program quality.
Program universality	Sharing experiences of illness within a group by means of supervised forums, owing to the general nature of dysfunction shared by all the patients.
Possibilities for expression	Possibility of manifesting or expressing feelings and emotions through NTs.
Supervision by a therapist	Appraisal of the evolution of the disorder throughout the therapeutic process (physicians).
Individualized attention for patients	Personal, empathetic, sincere, and close contact (patients), with flexibility to be adapted to the individual characteristics of each patient (physicians).
Trust	Peace of mind when using the program in an intimate and participative way.
Presence	Support from and availability of a professional who listens and pays attention (patients).
Hope	An optimistic perspective for the outcome of the treatment.

Participants' Perspectives

In general, the participants similarly conceptualized their expectations regarding Web-based psychotherapy for depression, although the focus of their discourse offered differing perspectives depending on their position (Table 4). For example, the health professionals—physicians and managers—placed emphasis on the intervention processes from the perspective of standardization of the program. More specifically, the physicians centered their interest on the professionals involved, with the concern of whether this type of intervention could improve the current delivery of health care, but at the cost of increasing their already excessive workload. However, the discourse given by the managers was focused on the idea of the service provided, with interest shown in the cost-effectiveness of this resource, by means of an organization of work that would create a new dynamic for distributing patients, and in a way that would solve practical problems, such as reducing waiting lists. Finally, patients highlighted the intervention processes from the perspective of interaction with the program, through which they expected to receive quality, personal, and individualized service, which would offer them the possibility of safely sharing experiences with people in the same situation as theirs, and with an accessible professional who would be in charge of supervising the treatment.

The different perspectives expressed by the participants—physicians, managers, and patients—reflected their different concerns based on their position. Could a patient be placed in the program without increasing my workload? Would this program solve the problem of waiting lists without added expense? Would this program sufficiently attend to my particular condition? Each position raised the need to satisfy different requirements, and any of these could cause resistance to the introduction of this type of instrument into the health service.

Table 4. Participants' discursive perspectives.

Participant	Focus of discourse	Quotes
Physicians	Professionals	"Maybe we could use a little help so as not to overload our schedules even more." (physician, male, 45 years old, high affinity)
		"The design is relatively simple to imagine, but it means a change to our routine practice. How long will it take to train us to use and implement this program? Would we be released from our clinical practice for that time?" (physician, female, 54 years old, intermediate affinity)
		"Not all of us have the same technological know-how or motivation. Should all doctors use the tool or only the most qualified? Those of us who volunteer to participate could be released from our routine workload to devote ourselves more specifically to this matter". (physician, female, 35 years old, high affinity)
		"Telemedicine platforms are already showing great improvement in places with dispersed populations, assisting with the organization of care and better communication between primary and specialist care services, which not only improves final care, but saves money." (manager, female, 51 years old, high affinity)
Managers	Service	"The tool should speed up care processes, especially waiting lists and user access to resources, providing maximum efficacy with a small investment that is quickly justified." (manager, male, 55 years old, intermediate affinity)
		"If we all save time with this tool, seeing more patients in distance mode and therefore freeing up our schedules, close, properly programmed, face-to-face monitoring could be carried out." (manager, male, 50 years old, high affinity)
		"Of course, every patient is different, so if you decide to help them, first you have to know their whole story." (patient, male, 46 years old, intermediate affinity)
Patients	Individualization	"Depending on the patient, on the state they're in, there can be interventions that are for everybody. But some patients, given their particular evolution, will need more personalized care." (patient, female, 52 years old, intermediate affinity)
		"Therapy has to be something that is very personalized, with the possibility of asking and speaking about other subjects..." (patient, female, 59 years old, low affinity)

Discussion

Principal Findings

Our results show how important it is for primary care physicians that the implementation of Web-based psychotherapy programs does not mean an added workload for them; rather, this should be reduced. They also show that the main concern for health care management professionals is to offer greater accessibility to services, and to improve their cost-effectiveness. Patients expect to receive individualized care, which requires investment in a certain level of resources, workload for professionals, and infrastructure in order to meet their needs. These particular viewpoints reflect different expectations, which taken individually, seem to respond to conflicting needs and interests. In order to overcome any possible resistance simultaneously, we have to find common ground. One starting point in this regard could be the optimization of the processes for management and patient flow in an integrated model, which has proven to be very useful for balancing the burdens of care delivery [58].

Depressive disorder is currently one of the greatest challenges for public health systems. In its mild and moderate forms, it responds satisfactorily to different types of psychotherapeutic approaches [59], such as Web-based cognitive behavioral therapy, including those with little or no support [60,61]. This type of treatment could become an alternative for public health systems to facilitate patients' acceptance and access to care [62], adding efficacy to early interventions [22,40,41]. Forecasts show that the dissemination and use of Web-based psychotherapy will grow exponentially in the coming decade [63]. In this context, it is necessary to establish procedures for the implementation of these type of programs that meet the many varied expectations placed on them by all participants in the therapeutic process, whether patients or health professionals.

A fundamental aspect in this regard would be to determine the level of intensity of the required supervision. The Spanish health care system treats all individuals equally, which means that the referral process for patients is not a fast and smooth one. In an attempt to optimize the therapeutic process, patients should be and feel cared for, without overloading professionals, but while reducing costs and waiting times [64]. In order to do this, it would be necessary to perform the specific task of disseminating the program, which would allow the patients most likely to benefit from this type of treatment to be identified and chosen, without losing sight of the factors that would enhance compliance. Those patients with symptoms of mild severity are seen as good candidates, as well as those who take responsibility for the treatment and who attribute success to themselves, focusing on action [65-67]. According to our results, all participants would show moderate severity at most, with an attitude of general acceptance toward Web-based therapy and a certain inclination for the use of new technologies, as has been seen in other works [68], but most of all, they have to understand their possible usefulness in the specific field of psychotherapy.

Both patients and health professionals have to share a certain willingness to use new technologies [68], must have a certain level of skill in their use, and need to have the necessary infrastructure [69], which is particularly important when we encounter difficult-to-access settings, or in cases with scarcity of training and resources [70,71]. All participants need to perceive the possibilities offered by Web-based psychotherapy as useful, with regard to convenience, time, and place of use, and value the reduction of transport costs, and the possibilities for communication, reflectiveness, memory, and information that it brings, while giving patients an active role and enabling its use whenever they want [72]. Nevertheless, all participants also highlighted the importance of not falling into patterns of abuse. In other words, the use of Web-based psychotherapy for the treatment of depression should provide everybody involved with a certain sense of efficacy and comfort [69]. This is how depression can become easier to handle, probably through patients' actual perception of their own self-efficacy.

We have seen that the discourse of professionals basically revolves around the difficulties of attempts at standardization, which is necessary in order to provide institutional support for Web-based therapy. This is a reference to the general culture in which health services are immersed [73]. This culture is present in questions such as learning processes, definition of goals to be met and activities to carry out, and in relation to the professionals involved and the services provided. That is to say, the group of health professionals—physicians and managers—are interested in the implementation of the program as a tool for clinical intervention, with the expectation that the needs of their users/patients have to be met, but also those of the professionals involved, implying possible strategies for promotion and improvement in quality and efficiency. We have thus seen how health care workers highlight the importance of standardization processes for Web-based treatments as part of an integrated service model, which is in agreement with the results of other studies [68-74].

In general, health professionals have positive attitudes to Web-based treatments [75,76], and these positive attitudes tend to be associated with an open mind and availability to new treatments, the convenience that computers offer, and easy access to technology [69], which is in line with our results. Nevertheless, therapists have also been said to express doubts as to whether new technologies produce substantial improvements, and although they are generally open to incorporating Web-based psychotherapy into their clinical practice, they show doubt depending on their skill, and out of the belief that this type of therapy may harm their relationship with patients [69]. In this regard, our study has pointed out, through the discourse of both professionals and patients, that therapeutic interaction by computer should not completely replace human contact, which underscores the limits to the application of indiscriminate Web-based psychotherapy.

Web-based therapeutic interventions are rarely conducted without face-to-face contact at some point in the process [27,77,78]. The presence of a therapist is key to guiding patients, given that this benefits the treatment, and supervision of the therapeutic process is associated with achieving positive results, while withdrawals from treatment increase when there is not sufficient contact [27,30,79]. On the contrary, adding contacts may improve efficacy, probably as a result of the trust created in the process [80]. When we asked patients about the possibilities of Web-based psychotherapy, we immediately saw the emergence of thematic content that was linked to the need to construct a therapeutic alliance as a vector that produces the hope of a positive outcome [81-83]. In fact, it is the base on which all psychotherapy is constructed, as a relationship of help [84,85]. Another important aspect to take into account is the necessary confidentiality of the entire process, for the security and anonymity that both patients and professionals require [68], and given that this could help to prevent stigmas [65], if clinical symptoms are able to be resolved at the primary health care level, providing solutions without the need for involving mental health units [86].

In general, the context for applying these results does not seem to be logically restricted to the particular characteristics of the place where the study was conducted. Instead, it points out the possible general characteristics and interests of the different participants. While all of the participants resided in only one region of Spain, it is true that Spanish primary care practices on the whole are currently a setting with a very high care delivery burden [87], and patient profiles also show a certain tendency toward personal contact in visits, as occurs in the majority of countries studied [88]. For this reason, our results seem to be fairly reliable. Nonetheless, it would be logical to propose future studies that make use of our procedures in other countries, in order to verify the extent to which our conclusions can be generalized in different populations, or whether they are in fact the specific product of the studied population. The fact that the interviewer and analysis team were mental health specialists, psychologists, and psychiatrists should not be overlooked, and they may therefore have focused the framework of discourse generation and the creation of emerging categories toward topics of special interest in this field of knowledge. We would add that the randomized controlled study conducted parallel to this work, although with different participants, may have exerted a certain influence on the processes of analysis and interpretation of the data.

Conclusions

This study shows how patients, physicians, and managers see the possibility of carrying out Web-based psychotherapy for the treatment of depression revolve around two attractors: the need for individualized interaction, which enables the establishment of a therapeutic alliance, and the processes for standardizing the program, which would prevent professional overloading while optimizing cost-effectiveness. In order to find a balance between both expectations, Web-based psychotherapy should be understood as a complementary intervention, which allows the amount of human contact to be reduced, but without completely eliminating it. This type of intervention provides a way of treating depression, which would permit more patients to be supervised with fewer visits, offering a solution to the problem of waiting lists, and without being an additional burden for health professionals. Research shows that Web-based psychotherapy as an approach to treating depression is in its infancy as a therapeutic option. In order to facilitate its implementation in primary care, we should be aware of the expectations of all participants by designing effective programs that are able to offer patients individualized treatment by means of standard interventions in the different contexts. Incorporating Web-based therapy into health systems harmoniously will show its value in improving delivery of care [89], by facilitating access to resources and improving working conditions for professionals.

Acknowledgments

This study was funded by the Institute of Health Carlos III (FIS PI10/01982) of the Ministry of Economy and Competitiveness and the Ministry of Health and Consumption (Spain), through the Network for Prevention and Health Promotion in Primary Care (redIAPP, RD12/0005), the CIBER (PSI2010-17563), and with European Union ERDF funds.

Authors' Contributions

JMM and JGC are the principal researchers and developed the original idea for the study and the study design. PRS and PHM were in charge of the interviews and groups. JMM, JPA, and JGC developed the data analysis and wrote the first manuscript. All the authors have read and corrected draft versions and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Participants’ responses to subtopics.

[PDF File (Adobe PDF File), 16KB - jmir_v17i3e67_app1.pdf]

References

1. World Health Organization. The Global Burden of Disease: 2004 Update. Switzerland: WHO; 2004. URL: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf?ua=1 [accessed 2015-03-06] [WebCite Cache ID 6WpsrqZA3]
2. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006 Nov;3(11):e442 [FREE Full text] [doi: 10.1371/journal.pmed.0030442] [Medline: 17132052]

3. Richards D. Prevalence and clinical course of depression: a review. Clin Psychol Rev 2011 Nov;31(7):1117-1125. [doi: 10.1016/j.cpr.2011.07.004] [Medline: 21820991]
4. Sartorius N. Psychological disorders in general medical settings. Toronto: Published on behalf of the World Heath Organization by Hogrefe & Huber; 1990.
5. Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000? J Clin Psychiatry 2003 Dec;64(12):1465-1475. [Medline: 14728109]
6. Sobocki P, Jönsson B, Angst J, Rehnberg C. Cost of depression in Europe. J Ment Health Policy Econ 2006 Jun;9(2):87-98. [Medline: 17007486]
7. Cuijpers P, Schoevers RA. Increased mortality in depressive disorders: a review. Curr Psychiatry Rep 2004 Dec;6(6):430-437. [Medline: 15538991]
8. Rapaport MH, Clary C, Fayyad R, Endicott J. Quality-of-life impairment in depressive and anxiety disorders. Am J Psychiatry 2005 Jun;162(6):1171-1178. [doi: 10.1176/appi.ajp.162.6.1171] [Medline: 15930066]
9. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]
10. McHugh RK, Whitton SW, Peckham AD, Welge JA, Otto MW. Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review. J Clin Psychiatry 2013 Jun;74(6):595-602 [FREE Full text] [doi: 10.4088/JCP.12r07757] [Medline: 23842011]
11. van Schaik DJ, Klijn AF, van Hout HP, van Marwijk HW, Beekman AT, de Haan M, et al. Patients' preferences in the treatment of depressive disorder in primary care. Gen Hosp Psychiatry 2004;26(3):184-189. [doi: 10.1016/j.genhosppsych.2003.12.001] [Medline: 15121346]
12. Bedi N, Chilvers C, Churchill R, Dewey M, Duggan C, Fielding K, et al. Assessing effectiveness of treatment of depression in primary care. Partially randomised preference trial. Br J Psychiatry 2000 Oct;177:312-318 [FREE Full text] [Medline: 11116771]
13. Brody DS, Khaliq AA, Thompson TL. Patients' perspectives on the management of emotional distress in primary care settings. J Gen Intern Med 1997 Jul;12(7):403-406 [FREE Full text] [Medline: 9229277]
14. Churchill R, Khaira M, Gretton V, Chilvers C, Dewey M, Duggan C, Nottingham Counselling and Antidepressants in Primary Care (CAPC) Study Group. Treating depression in general practice: factors affecting patients' treatment preferences. Br J Gen Pract 2000 Nov;50(460):905-906 [FREE Full text] [Medline: 11141877]
15. Dwight-Johnson M, Sherbourne CD, Liao D, Wells KB. Treatment preferences among depressed primary care patients. J Gen Intern Med 2000 Aug;15(8):527-534 [FREE Full text] [Medline: 10940143]
16. O'Mahen HA, Flynn HA. Preferences and perceived barriers to treatment for depression during the perinatal period. J Womens Health (Larchmt) 2008 Oct;17(8):1301-1309. [doi: 10.1089/jwh.2007.0631] [Medline: 18816202]
17. Priest RG, Vize C, Roberts A, Roberts M, Tylee A. Lay people's attitudes to treatment of depression: results of opinion poll for Defeat Depression Campaign just before its launch. BMJ 1996 Oct 5;313(7061):858-859 [FREE Full text] [Medline: 8870574]
18. Treating depression in adults.: National Institute for Health and Clinical Excellence (NICE); 2009. URL: https://www.nice.org.uk/guidance/cg90/resources/guidance-depression-in-adults-pdf [accessed 2014-10-27] [WebCite Cache ID 6TdRxt4CA]
19. Huhn M, Tardy M, Spineli LM, Kissling W, Förstl H, Pitschel-Walz G, et al. Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: a systematic overview of meta-analyses. JAMA Psychiatry 2014 Jun;71(6):706-715. [doi: 10.1001/jamapsychiatry.2014.112] [Medline: 24789675]
20. Morgan AJ, Jorm AF, Mackinnon AJ. Email-based promotion of self-help for subthreshold depression: Mood Memos randomised controlled trial. Br J Psychiatry 2012 May;200(5):412-418 [FREE Full text] [doi: 10.1192/bjp.bp.111.101394] [Medline: 22442102]
21. Christensen H, Griffiths KM. The prevention of depression using the Internet. Med J Aust 2002 Oct 7;177 Suppl:S122-S125. [Medline: 12358571]
22. Buntrock C, Ebert DD, 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 2014;14:25 [FREE Full text] [doi: 10.1186/1471-244X-14-25] [Medline: 24485283]
23. Cuijpers P, van Straten A, Warmerdam L, van Rooy MJ. 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]
24. Lehtinen V, Riikonen E, Latineen E. Promotion of mental health on the European agenda.: Finnish Ministry of Social Affairs and Health, Department for Prevention and Promotion; 2000. URL: https://www.julkari.fi/bitstream/handle/10024/75859/Promotion.pdf?sequence=1 [accessed 2015-03-03] [WebCite Cache ID 6WkdVWVtP]

25. Richards D, Timulak L, Doherty G, Sharry J, Colla A, Joyce C, et al. Internet-delivered treatment: its potential as a low-intensity community intervention for adults with symptoms of depression: protocol for a randomized controlled trial. BMC Psychiatry 2014;14:147 [FREE Full text] [doi: 10.1186/1471-244X-14-147] [Medline: 24886179]
26. Bennett-Levy J, Richards DA, Farrand P. Oxford guide to low intensity CBT interventions. Oxford: Oxford University Press; 2010:3-18.
27. Andersson G. Using the Internet to provide cognitive behaviour therapy. Behav Res Ther 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010] [Medline: 19230862]
28. van BW, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to Internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]
29. Ritterband LM, Andersson G, Christensen HM, Carlbring P, Cuijpers P. Directions for the International Society for Research on Internet Interventions (ISRII). J Med Internet Res 2006;8(3):e23 [FREE Full text] [doi: 10.2196/jmir.8.3.e23] [Medline: 17032639]
30. Patten SB. Prevention of depressive symptoms through the use of distance technologies. Psychiatr Serv 2003 Mar;54(3):396-398. [Medline: 12610251]
31. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-9; quiz 870. [doi: 10.1586/ern.12.63] [Medline: 22853793]
32. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
33. 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]
34. Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered Internet psychotherapy for depression in primary care: a randomised controlled trial. Lancet 2009 Aug 22;374(9690):628-634. [doi: 10.1016/S0140-6736(09)61257-5] [Medline: 19700005]
35. Bergström J, Andersson G, Ljótsson B, Rück C, Andréewitch S, Karlsson A, et al. Internet-versus group-administered cognitive behaviour therapy for panic disorder in a psychiatric setting: a randomised trial. BMC Psychiatry 2010;10:54 [FREE Full text] [doi: 10.1186/1471-244X-10-54] [Medline: 20598127]
36. So M, Yamaguchi S, Hashimoto S, Sado M, Furukawa TA, McCrone P. Is computerised CBT really helpful for adult depression?-A meta-analytic re-evaluation of CCBT for adult depression in terms of clinical implementation and methodological validity. BMC Psychiatry 2013;13:113 [FREE Full text] [doi: 10.1186/1471-244X-13-113] [Medline: 23587347]
37. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
38. Kaltenthaler E, Parry G, Beverley C, Ferriter M. Computerised cognitive-behavioural therapy for depression: systematic review. Br J Psychiatry 2008 Sep;193(3):181-184 [FREE Full text] [doi: 10.1192/bjp.bp.106.025981] [Medline: 18757972]
39. 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]
40. Clarke G, Reid E, Eubanks D, O'Connor E, DeBar LL, Kelleher C, et al. Overcoming depression on the Internet (ODIN): a randomized controlled trial of an Internet depression skills intervention program. J Med Internet Res 2002 Dec;4(3):E14 [FREE Full text] [doi: 10.2196/jmir.4.3.e14] [Medline: 12554545]
41. Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. J Med Internet Res 2005;7(2):e16 [FREE Full text] [doi: 10.2196/jmir.7.2.e16] [Medline: 15998607]
42. Learmonth D, Rai S. Taking computerized CBT beyond primary care. Br J Clin Psychol 2008 Mar;47(Pt 1):111-118. [doi: 10.1348/014466507X248599] [Medline: 17939879]
43. 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]
44. Gerhards SAH, de Graaf LE, Jacobs LE, Severens JL, Huibers MJH, Arntz A, et al. Economic evaluation of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. Br J Psychiatry 2010 Apr;196(4):310-318 [FREE Full text] [doi: 10.1192/bjp.bp.109.065748] [Medline: 20357309]
45. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
46. Gun SY, Titov N, Andrews G. Acceptability of Internet treatment of anxiety and depression. Australas Psychiatry 2011 Jun;19(3):259-264. [doi: 10.3109/10398562.2011.562295] [Medline: 21682626]

47. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
48. Mays N, Pope C. Qualitative research in health care. London: BMJ Books; 2000:1-168.
49. Ferrando L, Franco A, Soto M, Bobes J, Soto O, Franco L, et al. MINI International Neuropsychiatric Interview. Madrid:Instituto IAP 1998:1-25 Versión en español 5.0.0. DSM-IV [FREE Full text]
50. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, Texas: Psychological Corp; 1996.
51. López-del-Hoyo Y, Olivan B, Luciano JV, Mayoral F, Roca M, Gili M, et al. Low intensity vs. self-guided internet-delivered psychotherapy for major depression: a multicenter, controlled, randomized study. BMC Psychiatry 2013;13:21 [FREE Full text] [doi: 10.1186/1471-244X-13-21] [Medline: 23312003]
52. Corbin JM, Strauss AL. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. London: Sage Publications, Inc; 1998.
53. Strauss A. Qualitative analysis for social scientists. Cambridge: Cambridge University Press; 1987.
54. Glaser BG, Strauss AL. The discovery of grounded theory; strategies for qualitative research. Chicago: Aldine Publishing Co; 1967.
55. Alonso LE. La mirada cualitativa en sociología. Una aproximación interpretativa [A qualitative approach to sociology. An interpretative approach]. Madrid: Editorial Fundamentos; 1998.
56. Phillips N, Hardy C. Discourse analysis: investigating processes of social construction. Thousand Oaks, Calif: Sage Publications; 2002.
57. Montero-Marín J, Carrasco JM, Roca M, Serrano-Blanco A, Gili M, Mayoral F, et al. Expectations, experiences and attitudes of patients and primary care health professionals regarding online psychotherapeutic interventions for depression: protocol for a qualitative study. BMC Psychiatry 2013;13:64 [FREE Full text] [doi: 10.1186/1471-244X-13-64] [Medline: 23425435]
58. Davison GC. Stepped care: doing more with less? J Consult Clin Psychol 2000 Aug;68(4):580-585. [Medline: 10965633]
59. Cuijpers P, van Straten A, Andersson G, van Oppen P. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J Consult Clin Psychol 2008 Dec;76(6):909-922. [doi: 10.1037/a0013075] [Medline: 19045960]
60. Andersson G. Internet-based cognitive-behavioral self help for depression. Expert Rev Neurother 2006 Nov;6(11):1637-1642. [doi: 10.1586/14737175.6.11.1637] [Medline: 17144778]
61. Spek V, Nyklícek I, Smits N, Cuijpers P, Riper H, Keyzer J, et al. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. Psychol Med 2007 Dec;37(12):1797-1806. [doi: 10.1017/S0033291707000542] [Medline: 17466110]
62. Wootton BM, Titov N, Dear BF, Spence J, Kemp A. The acceptability of Internet-based treatment and characteristics of an adult sample with obsessive compulsive disorder: an Internet survey. PLoS One 2011;6(6):e20548 [FREE Full text] [doi: 10.1371/journal.pone.0020548] [Medline: 21673987]
63. Norcross JC, Pfund RA, Prochaska JO. Psychotherapy in 2022: A Delphi poll on its future. Professional Psychology: Research and Practice 2013;44(5):363-370. [doi: 10.1037/a0034633]
64. Reins JA, Ebert DD, Lehr D, Riper H, Cuijpers P, Berking M. Internet-based treatment of major depression for patients on a waiting list for inpatient psychotherapy: protocol for a multi-centre randomised controlled trial. BMC Psychiatry 2013;13:318 [FREE Full text] [doi: 10.1186/1471-244X-13-318] [Medline: 24279841]
65. Ellis LA, Collin P, Hurley PJ, Davenport TA, Burns JM, Hickie IB. Young men's attitudes and behaviour in relation to mental health and technology: implications for the development of online mental health services. BMC Psychiatry 2013;13:119 [FREE Full text] [doi: 10.1186/1471-244X-13-119] [Medline: 23601273]
66. Bendelin N, Hesser H, Dahl J, Carlbring P, Nelson KZ, Andersson G. Experiences of guided Internet-based cognitive-behavioural treatment for depression: a qualitative study. BMC Psychiatry 2011;11:107 [FREE Full text] [doi: 10.1186/1471-244X-11-107] [Medline: 21718523]
67. Titov N. Status of computerized cognitive behavioural therapy for adults. Aust N Z J Psychiatry 2007 Feb;41(2):95-114. [doi: 10.1080/00048670601109873] [Medline: 17464688]
68. Dünnebeil S, Sunyaev A, Blohm I, Leimeister JM, Krcmar H. Determinants of physicians' technology acceptance for e-health in ambulatory care. Int J Med Inform 2012 Nov;81(11):746-760. [doi: 10.1016/j.ijmedinf.2012.02.002] [Medline: 22397989]
69. Becker EM, Jensen-Doss A. Computer-assisted therapies: examination of therapist-level barriers to their use. Behav Ther 2013 Dec;44(4):614-624. [doi: 10.1016/j.beth.2013.05.002] [Medline: 24094787]
70. McCloskey MS. Training in empirically supported treatments using alternative learning modalities. Clinical Psychology:Science and Practice 2011;18(1):84-88. [doi: 10.1111/j.1468-2850.2011.01237.x]
71. Weingardt KR. The role of instructional design and technology in the dissemination of empirically supported manual-based therapies. Clinical Psychology: Science and Practice 2004;11(3):313-331. [doi: 10.1093/clipsy.bph087]
72. Klaw E, Humphreys K. Facilitating client involvement in self-help groups. In: Koocher G, Norcross J, Hill S, editors. Psychologists' Desk Reference, 2nd edition. New York: Oxford University Press; 2005:502-505.

73. Parand A, Dopson S, Renz A, Vincent C. The role of hospital managers in quality and patient safety: a systematic review. BMJ Open 2014;4(9):e005055 [FREE Full text] [doi: 10.1136/bmjopen-2014-005055] [Medline: 25192876]
74. Wangberg SC, Gammon D, Spitznogle K. In the eyes of the beholder: exploring psychologists' attitudes towards and use of e-therapy in Norway. Cyberpsychol Behav 2007 Jun;10(3):418-423. [doi: 10.1089/cpb.2006.9937] [Medline: 17594266]
75. Gund A, Lindecrantz K, Schaufelberger M, Patel H, Sjöqvist BA. Attitudes among healthcare professionals towards ICT and home follow-up in chronic heart failure care. BMC Med Inform Decis Mak 2012;12:138 [FREE Full text] [doi: 10.1186/1472-6947-12-138] [Medline: 23190602]
76. Perle JG, Langsam LC, Randel A, Lutchman S, Levine AB, Odland AP, et al. Attitudes toward psychological telehealth: current and future clinical psychologists' opinions of internet-based interventions. J Clin Psychol 2013 Jan;69(1):100-113. [doi: 10.1002/jclp.21912] [Medline: 22975897]
77. Palmqvist B, Carlbring P, Andersson G. Internet-delivered treatments with or without therapist input: does the therapist factor have implications for efficacy and cost? Expert Rev Pharmacoecon Outcomes Res 2007 Jun;7(3):291-297. [doi: 10.1586/14737167.7.3.291] [Medline: 20528315]
78. Marks IM, Cuijpers P, Cavanagh K, van Straten A, Gega L, Andersson G. Meta-analysis of computer-aided psychotherapy: problems and partial solutions. Cogn Behav Ther 2009;38(2):83-90. [doi: 10.1080/16506070802675239] [Medline: 20183689]
79. Christensen H, Griffiths KM, Mackinnon AJ, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. Psychol Med 2006 Dec;36(12):1737-1746. [doi: 10.1017/S0033291706008695] [Medline: 16938144]
80. Carlbring P, Bohman S, Brunt S, Buhrman M, Westling BE, Ekselius L, et al. Remote treatment of panic disorder: a randomized trial of internet-based cognitive behavior therapy supplemented with telephone calls. Am J Psychiatry 2006 Dec;163(12):2119-2125. [doi: 10.1176/appi.ajp.163.12.2119] [Medline: 17151163]
81. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy: Theory, Research & Practice 1979;16(3):252-260. [doi: 10.1037/h0085885]
82. Hill CE, Knox S. Processing the therapeutic relationship. Psychother Res 2009 Jan;19(1):13-29. [doi: 10.1080/10503300802621206] [Medline: 19206019]
83. Prado-Abril J, García-Campayo J, Sánchez-Reales S. Funcionamiento de la terapia cognitivo-interpersonal en los trastornos de la personalidad: Estudio de dos casos [Efficacy of interpersonal-cognitive therapy in personality disorders: Two cases]. RPPC 2014 Jul 21;18(2):139. [doi: 10.5944/rppc.vol.18.num.2.2013.12770]
84. Gelso CJ. The real relationship in a postmodern world: theoretical and empirical explorations. Psychother Res 2009 May;19(3):253-264. [doi: 10.1080/10503300802389242] [Medline: 19034731]
85. Addis ME, Wade WA, Hatgis C. Barriers to dissemination of evidence-based practice: Promises, obstacles and future directions. Clin Psychol-Sci Pr 1999;9(4):367-378. [doi: 10.1093/clipsy.9.4.367]
86. Farrer L, Christensen H, Griffiths KM, Mackinnon A. Web-based cognitive behavior therapy for depression with and without telephone tracking in a national helpline: secondary outcomes from a randomized controlled trial. J Med Internet Res 2012;14(3):e68 [FREE Full text] [doi: 10.2196/jmir.1859] [Medline: 22738715]
87. Fernández A, Saameño JAB, Pinto-Meza A, Luciano JV, Autonell J, Palao D, et al. Burden of chronic physical conditions and mental disorders in primary care. Br J Psychiatry 2010 Apr;196(4):302-309 [FREE Full text] [doi: 10.1192/bjp.bp.109.074211] [Medline: 20357307]
88. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
89. Ebert DD, Lehr D, Baumeister H, Boß L, Riper H, Cuijpers P, et al. GET.ON Mood Enhancer: efficacy of Internet-based guided self-help compared to psychoeducation for depression: an investigator-blinded randomised controlled trial. Trials 2014;15:39 [FREE Full text] [doi: 10.1186/1745-6215-15-39] [Medline: 24476555]

Edited by G Eysenbach; submitted 28.10.14; peer-reviewed by J Boggs, G Culjak; comments to author 30.12.14; revised version received 26.01.15; accepted 12.02.15; published 10.03.15.

Please cite as:
Montero-Marín J, Prado-Abril J, Botella C, Mayoral-Cleries F, Baños R, Herrera-Mercadal P, Romero-Sanchiz P, Gili M, Castro A, Nogueira R, García-Campayo J
Expectations Among Patients and Health Professionals Regarding Web-Based Interventions for Depression in Primary Care: A Qualitative Study
J Med Internet Res 2015;17(3):e67
URL: http://www.jmir.org/2015/3/e67/
doi:10.2196/jmir.3985
PMID:25757358

Review

The Effect of Technology-Based Interventions on Pain, Depression, and Quality of Life in Patients With Cancer: A Systematic Review of Randomized Controlled Trials

Stephen O Agboola[1,2,3], MD, MPH; Woong Ju[4], MD, PhD, MPH; Aymen Elfiky[3,5], MA, MD, MPH; Joseph C Kvedar[1,2,3], MD; Kamal Jethwani[1,2,3], MD, MPH

[1]Partners Healthcare Center for Connected Health, Boston, MA, United States
[2]Massachusetts General Hospital, Boston, MA, United States
[3]Harvard Medical School, Boston, MA, United States
[4]Medical Research Institute, College of Medicine, Department of Obstetrics and Gynecology, Ewha Womans University, Seoul, Republic Of Korea
[5]Dana-Farber Cancer Institute, Boston, MA, United States

Corresponding Author:
Stephen O Agboola, MD, MPH
Partners Healthcare Center for Connected Health
25 New Chardon St.
Suite 300
Boston, MA, 02114
United States
Phone: 1 617 643 0291
Fax: 1 617 228 4619
Email: sagboola@partners.org

Abstract

Background: The burden of cancer is increasing; projections over the next 2 decades suggest that the annual cases of cancer will rise from 14 million in 2012 to 22 million. However, cancer patients in the 21st century are living longer due to the availability of novel therapeutic regimens, which has prompted a growing focus on maintaining patients' health-related quality of life. Telehealth is increasingly being used to connect with patients outside of traditional clinical settings, and early work has shown its importance in improving quality of life and other clinical outcomes in cancer care.

Objective: The aim of this study was to systematically assess the literature for the effect of supportive telehealth interventions on pain, depression, and quality of life in cancer patients via a systematic review of clinical trials.

Methods: We searched PubMed, EMBASE, Google Scholar, CINAHL, and PsycINFO in July 2013 and updated the literature search again in January 2015 for prospective randomized trials evaluating the effect of telehealth interventions in cancer care with pain, depression, and quality of life as main outcomes. Two of the authors independently reviewed and extracted data from eligible randomized controlled trials, based on pre-determined selection criteria. Methodological quality of studies was assessed by the Cochrane Collaboration risk of bias tool.

Results: Of the 4929 articles retrieved from databases and relevant bibliographies, a total of 20 RCTs were included in the final review. The studies were largely heterogeneous in the type and duration of the intervention as well as in outcome assessments. A majority of the studies were telephone-based interventions that remotely connected patients with their health care provider or health coach. The intervention times ranged from 1 week to 12 months. In general, most of the studies had low risk of bias across the domains of the Cochrane Collaboration risk of bias tool, but most of the studies had insufficient information about the allocation concealment domain. Two of the three studies focused on pain control reported significant effects of the intervention; four of the nine studies focus on depression reported significant effects, while only the studies that were focused on quality of life reported significant effects.

Conclusions: This systematic review demonstrates the potential of telehealth interventions in improving outcomes in cancer care. However, more high-quality large-sized trials are needed to demonstrate cogent evidence of its effectiveness.

(J Med Internet Res 2015;17(3):e65) doi:10.2196/jmir.4009

KEYWORDS

telehealth; connected health; cancer; telephone; pain; depression; quality of life; systematic review; randomized controlled trials

Introduction

The burden of cancer is increasing globally; projections over the next two decades suggest that the annual cases of cancer will rise from 14 million in 2012 to 22 million [1]. Cancer is the leading cause of death worldwide and the second leading cause of death in the United States [1,2]. Encouragingly, cancer patients in the 21st century are living longer due to a combination of early detection, availability of novel therapeutic regimens, and improved supportive care. According to the National Cancer Institute, the 5-year survival rate for all cancers increased significantly from about 48.7% in 1975 to about 68.5% in 2006 in the United States [2]. Despite these notable improvements in cancer outcomes, many patients experience physical and/or emotional distress, resulting from complex interplays between the disease process and treatment modalities, which significantly impact quality of life [3,4]. In this context, extended longevity has necessarily prompted a growing focus on better defining, capturing, and maintaining health-related quality of life (HR-QOL).

An increasingly popular model for delivering supportive care for patients with cancer is telehealth or other terminologies including connected health, eHealth, mHealth, that are used to describe health care delivery that leverages technology. Telehealth offers patients the opportunity for long-term home monitoring, health education and coaching, behavioral modification, sharing health information with care providers, and timely feedback. It has been largely employed in the management of chronic disease such as diabetes, hypertension, and heart failure [5-10]. Nowadays, with many patients with cancer living longer, it is increasingly being used to engage patients with cancer. Over the last decade, a growing body of studies regarding the application of telehealth in cancer care has been published. Some of the common applications in cancer care include management of pain, cancer-related psychological effects, and overall use to improve quality of life. However, the evidence of its effectiveness in cancer care is still not solid due to difficulty in designing or implementing non-biased randomized controlled trials (RCT) exploring its true effect. For this reason, there is a dearth of published systematic reviews or meta-analyses that summarize this topic. In this study, we evaluate the effect of telehealth on pain, depression, and quality of life in cancer patients via a systematic review of RCTs.

Methods

Literature Search

We first searched PubMed, EMBASE, Google Scholar, CINAHL, and PsycINFO in July 2013 for prospective RCTs evaluating telehealth in cancer care regarding pain, depression, and quality of life. The search was updated in January 2015.The keywords were as follows: “neoplasms [MeSH]”, “cancer” and “Remote Consultation [Mesh]”, “mHealth”, “connected health”, “text messaging”, “telemedicine”, “telehealth”, “ehealth”, “telephone therapy”, “teleconsultation”, “mobile technology”, “telecare”, “Internet”, “digital health”, “mobile phone*”, “smartphone”, “apps”, and “mobile application”.

Selection Criteria

We included RCTs that met all of the following criteria: reported the effect of telehealth on pain, depression, or quality of life in cancer patients. If data were duplicated or shared in more than one study, the last published or more comprehensive study was included in the analysis.

Selection of Relevant Studies

Based on the pre-determined selection criteria, 2 authors (JW, SA) independently selected all trials retrieved from the databases and bibliographies. Disagreements between evaluators were resolved by discussion.

Assessment of Methodological Quality

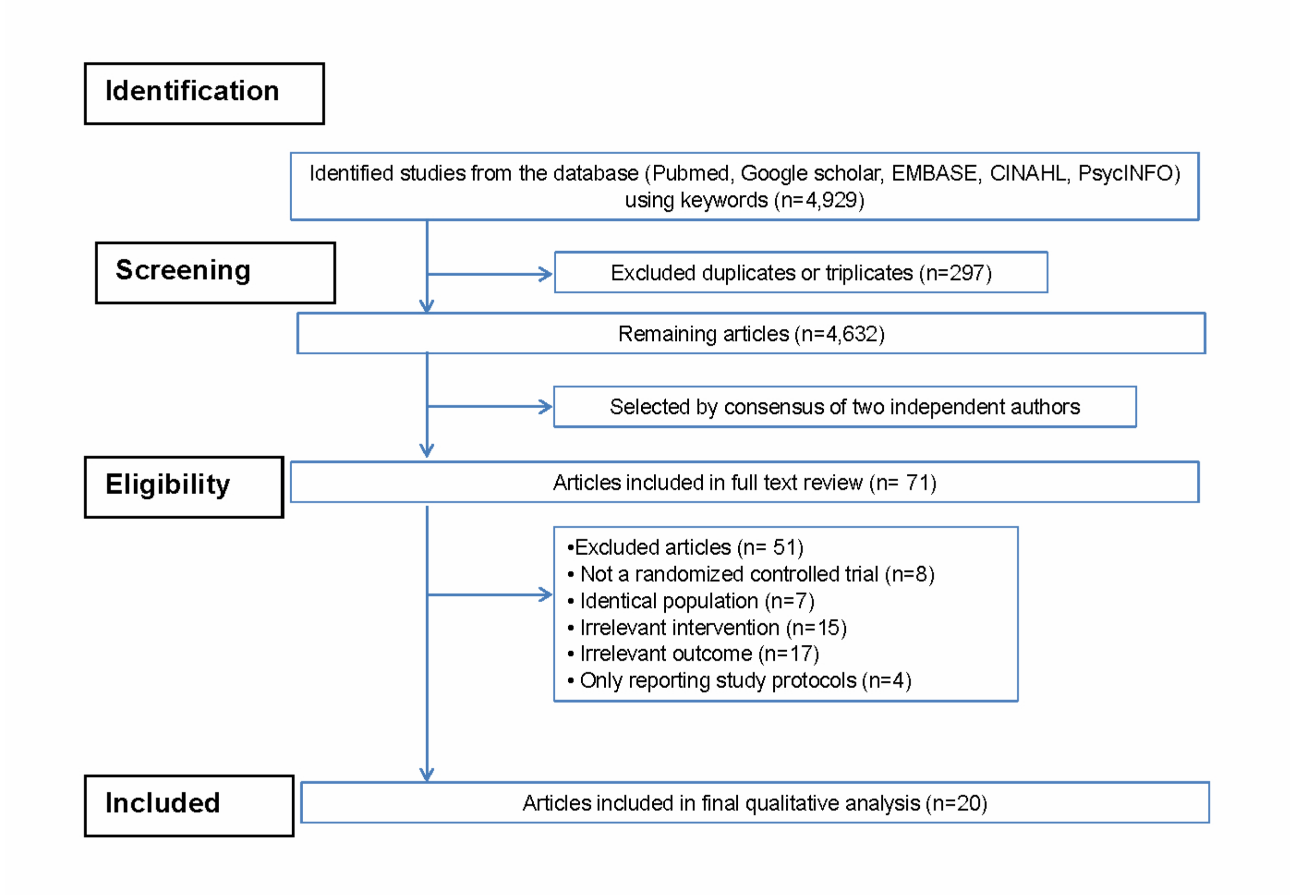
The methodological quality of included studies was assessed by the Cochrane Collaboration’s risk-of-bias tool [11], a commonly used tool to report the risk of bias in individual studies included in systematic reviews. The tool assesses several internal validity domains, which include sequence generation, allocation concealment, blinding of study participants, personnel and outcome assessors, incomplete outcome data, selective reporting, and other sources of bias. We classified each of these domains as being at high, low, or unclear risk of bias. Data were entered into Review Manager 5.3, and a risk of bias graph was generated for all included studies.

Results

Overview

Figure 1 depicts a flow diagram of how we identified relevant clinical trials included in this review. Using the above-mentioned keywords, a total of 4929 articles were identified from the literature search of five databases, that is, PubMed, EMBASE, Google Scholar, CINAHL, and PsycINFO. After excluding 297 duplicated articles, 2 authors independently reviewed and excluded an additional 4561 articles that did not satisfy the pre-determined selection criteria based on each article’s title and abstract. We reviewed the full texts of the remaining 71 articles and excluded 51 articles because of the following reasons: identical trials with the same population (n=7), nonrandomized studies (n=8), trials not related to the subject (intervention/outcome) of this study (n=32), and trials reporting only the study protocol (n=4). A total of 20 trials were included in the final analysis [12-31]. Since the studies included in this review are RCTs with comparator groups, we report only between-group effect estimates. Pre- and post measures within groups were not considered.

Figure 1. Flow diagram depicting the systematic review process.



Characteristics of Included Studies

Table 1 shows the general characteristics of the 20 trials. Eight (40%) of the 20 studies focused on improving quality of life [16,18,19,22,24,26,28,31], another nine (45%, 9/20) on improving depression outcome (Sherman et al actually focused on psychological well-being) [12,15,17,20,21,25,27,29,30], two (10%, 2/20) on improving pain control [13,14], and one (5%, 1/20) has both pain intensity and depression as main outcomes [23]. The Functional Assessment of Cancer Therapy (FACT) was the most (62.5 of all studies with HR-QOL as main outcome) commonly used measure of quality of life. The Center for Epidemiologic Studies Depression Scale (CES-D) and the Hospital Anxiety and Depression Scale (HADS) were the most commonly used measure for depression and the Brief Pain Inventory most commonly (66.7% of all studies with pain intensity as main outcome) used to evaluate pain outcomes. The studies were published over a period of 9 years from 2006-2014.

Table 1. Characteristics of randomized trials included in the systematic review on telehealth for cancer patients[a].

Author, year, country	Technology	Participants	Objectives	Intervention	Comparator	Intervention time
Badger, [12] 2012, USA	Telephone	70 breast cancer patients and their supportive partners (SPs)	To evaluate the efficacy of two telephone-delivered interventions in improving quality of life among Latinas with breast cancer and their family members or friends	Telephone interpersonal counseling delivered by trained interventionist	Telephone health education delivered by trained professionals	8 weeks: eight weekly sessions for patients and four sessions every other week for SPs
Borosund [27] 2014, Sweden	Internet	167 breast cancer patients	To evaluate the effect of the components of a Web-based support tool on symptom distress, anxiety and depression	Two intervention arms: (1) Internet-based patient-provider communication (IPPC) tool, (2) Web-choice + IPPC. Web-choice facilitates symptom monitoring, self-management and communication with other patients	Usual standard of care at the hospital of treatment	6 months
Duffecy, [] 2012, USA	Internet	31 patients with any cancer	To evaluate the feasibility of a Web-based intervention in increasing adherence to the intervention and efficacy in reducing symptoms of depression in post cancer treatment survivors	Individual Internet Intervention +Internet Support Group (ISG). ISG included a discussion board and features to enhance supportive accountability	Individual Internet Intervention is a self-management program, based on cognitive behavioral principles, for the treatment of depression	8 weeks
Freeman, [28] 2014, USA	Video-conference	118 breast cancer survivors	To evaluate the effect of an imagery-based group intervention on quality of life in breast cancer survivors	Two intervention groups with five 4-hr weekly group session delivered by trained professionals via live sessions or video-conferencing plus weekly telephone calls	Wait-list controls	3 months
Gotay, [25] 2007, USA	Telephone	305 breast cancer patients	To evaluate the effectiveness of a peer-delivered telephone support intervention on psychosocial outcomes in patients with a first recurrence of breast cancer	Telephone counseling/ information sessions delivered by trained peer counselors at a breast cancer advocacy organization	Standard care	4-8 sessions weekly with 1-2 calls per week for 1 month
Harrison, [19] 2011, Australia	Telephone	75 colorectal cancer (CRC) patients	To evaluate the effectiveness of a nurse-delivered telephone supportive intervention in reducing unmet supportive care needs, reducing health service utilization, and improving HR-QOL post- discharge from the hospital after surgery for CRC	CONNECT: post-surgery follow-up telephone calls delivered by an experienced colorectal cancer nurse who has undergone training in telephone communication	Usual care: follow-up appointment with a general practitioner and surgeon	5 calls over 6 months

Author, year, country	Technology	Participants	Objectives	Intervention	Comparator	Intervention time
Hawkins, [22] 2010, USA	Telephone and web	434 breast patients	To evaluate the mediating processes of two communication interventions to improve HR-QOL in patients with breast cancer	3 intervention groups: (1) Access to the Web-based comprehensive Health Enhancement Support System (CHESS), (2) Telephone-based Cancer information mentor, (3) CHESS + Cancer Information Mentor	Internet training and access	10 times over 6 months
Kim, [14] 2013, Korea	Telephone	108 patients with any solid-organ tumor	To evaluate the effectiveness of standardized education and telemonitoring in improving pain, distress, anxiety, depression, HR-QOL, and performance in outpatients with advanced cancers	Telemonitoring performed by an NP trained in pain management	Standardized pain education based on the WHO and NCCN pain control guidelines delivered by NP on the first visit	30 mins every day for 1 week
Kroenke, [23] 2010, USA	Telephone and Internet	405 cancer patients	To evaluate the effect of a telephone-based care management combined with automated symptom monitoring on depression and pain in patients with cancer	Telephonic care management by a nurse care manager combined with automated symptom monitoring (via interactive voice-recorded telephone calls or Web-based surveys)	Usual care provided by oncologists.	Follow-up calls and automated symptom monitoring staggered over 12 months
Lepore, [29] 2014, USA	Internet	184 breast cancer patients	To test the mental health benefits of two Internet support group (ISG) interventions in women with breast cancer	Pro-social Internet support group (ISG) which includes all features of the Standard-ISG plus tips on recognizing and responding to others' need for support and participation in a breast cancer awareness outreach activity	Standard-ISG with weekly live 90-minutes chats facilitated by PhD level interventionist plus discussion board for asynchronous text communication	6 weeks
Livingston, [20] 2009, Australia	Telephone	571 male colorectal (CRC) and prostate cancer patients	To evaluate the psychological impact of a referral and telephone intervention, involving information and support, among men with CRC and prostate cancer	Cancer Helpline: telephone calls from cancer nurses to help patients address issues they may experience during cancer care. 2 intervention groups: (1) Active Referral—4: four outcalls, (2) Active Referral—1: one outcall.	Passive Referral: usual care which involved a specialist referral to the Helpline but contact was at the participant's initiative	Active Referral—4: four outcalls staggered over 6 months post-diagnosis. Active Referral—1: outcall within 1 week of diagnosis.
Loprinzi, [18] 2011, USA	Telephone	25 breast cancer survivors	To evaluate the effect of a Stress Management and Resiliency Training (SMART) program for increasing resiliency and for decreasing stress and anxiety among breast cancer mentors who themselves were previously diagnosed with breast cancer	The SMART program: consisted of 3 parts: 2 small-group, 90-minute sessions teaching the SMART program; a brief individual follow-up session with a study investigator; and 3 follow-up telephone calls	Wait list group. Intervention delayed by 12 weeks.	12 weeks: telephone calls at 4-week intervals. Each call lasted approximately 15 minutes

Author, year, country	Technology	Participants	Objectives	Intervention	Comparator	Intervention time
Marcus, [21] 2009, USA	Telephone	304 breast cancer patients	To evaluate the effect of a telephone counseling program on psychosocial outcomes among breast cancer patients post-treatment	Usual care + Telephone Counseling program delivered by four Masters-level psychosocial oncology counselors	Usual care: booklet listing psychosocial and other social service and rehabilitation resources in their community for breast cancer	16 sessions delivered over a 12-month period. Each session lasted 45 mins
Nelson, [24] 2008, USA	Telephone	50 cervical cancer patients	To evaluate the feasibility of a psychosocial telephone counseling intervention designed for patients with cervical cancer on improving HR-QOL	Psychosocial telephone counseling intervention, delivered by a psychologist, designed to help women cope with the stressful events and feelings of distress associated with cervical cancer	Usual care	5 weeks: weekly session about 45 to 50 min in length + 1 month booster later
Park, [16] 2012, Korea	Telephone	48 breast cancer patients	To evaluate the effect of a psycho-educational support program on HR-QOL and symptom experience for women in the first year post-breast cancer treatment survivorship	Psychoeducation plus Standard care. The psychoeducational program consisted of individual face-to-face education using a participant handbook, telephone-delivered health-coaching sessions, and small-group meetings	Standard care from their medical team plus a short booklet on cancer care	12 weeks: 10-30 mins telephone coaching sessions every other week
Rustoen, [13] 2013, Norway	Telephone	179 cancer patients with bone metastasis	To evaluate the efficacy of PRO-SELF in decreasing pain intensity scores and increasing opioid intake in cancer patients.	PRO-SELF: Individualized pain management education delivered by oncology intervention nurses who visited patients in their homes at weeks 1, 3, and 6 and conducted telephone interviews at weeks 2, 4, and 5	Cancer pain management booklet plus home visits and nurse telephone interviews with the same frequency as patients in the intervention to monitor level of adherence with completing the pain diary	6 weeks
Ryhanen, [31] 2013, Finland	Internet	90 breast cancer patients	To evaluate the effect of the Breast Cancer Patient Pathway (BCPP) program on patients’ empowerment process. Specifically looking at quality of life, anxiety, and side-effects	Hospital standard of care plus the BCPP program - an Internet-based patient education tool to increase patients’ knowledge about breast cancer	Oral and written education materials according to hospital standards	Throughout the treatment period, average of 9 months
Sandgren, [26] 2006, USA	Telephone	218 breast cancer patients	To evaluate the effectiveness of two telephone-based interventions in improving mood and HR-QOL in patients with breast cancer	Telephone counseling including health education and emotional expression therapy delivered by oncology nurses	Standard care	5 weekly 30-minutes phone calls, with a 6th, follow-up call, made approx. 3 months later

Author, year, country	Technology	Participants	Objectives	Intervention	Comparator	Intervention time
Sherman, [17] 2012, USA	Telephone	249 breast cancer patients	To evaluate the effect of three technology-based interventions on physical, emotional, and social adjustment of women with early-stage breast cancer	3 intervention groups: (1) usual care + four phase-specific psychoeducational videos, (2) Usual care + four phase-specific telephone counseling sessions delivered by nurse interventionist, (3) usual care + phase-specific psycho-educational videos+ phase-specific telephone counseling sessions	Usual care was standardized across all sites according to national treatment protocols for the diagnosis and treatment of breast cancer.	Phase-specific: four phases of the breast cancer experience: diagnosis, post-surgery, adjuvant therapy and ongoing recovery
Stanton, [30] 2013, USA	Internet	88 breast cancer patients	To evaluate the effect of an Internet-based invention designed for chronicling the cancer experience and promoting communication	Project Connect Online: patients taught how to develop personalized website where they can journal their cancer experience and share content with their social networks	Waiting-list control	6 months

[a] HR-QOL: Health-related Quality of Life; CHESS: Comprehensive Health Enhancement Support System; WHO: World Health Organization; NP: nurse practitioner; SP: supportive partner; NCCN: National Comprehensive Cancer Network; CRC: colorectal carcinoma; SMART: Stress Management and Resiliency Training.

The majority (13/20, 65%) of the studies were conducted in the United States. The other countries represented include Australia (2/20, 10%), South Korea (2/20, 10%), Sweden (1/20, 5%), Finland (1/20, 5%), and Norway (1/20, 5%). The sample size in each of the studies ranged from 25-571 for a total of number 3789 subjects in all. The median follow-up time was 4 months with a range of 1 week to 18 months. Thirteen (65%) of the 20 trials, were conducted among patients with breast cancer [12,16-18,21,22,25-31], followed by four trials in patients with any solid cancer [13-15,23], and one trial each with focus on cervical [24], colorectal [19], and colorectal/prostate cancers [20].

Many of the included studies (14/20, 70%) were telephone-based interventions, although two of them were used in conjunction Web-based systems. Most of these telephone-based interventions (12/14, 85.7%) involved a professional interventionist (nurses, psychologists, or counselors) trained to provide counseling, while the remaining two studies [18,25] were delivered by peer counselors who are cancer survivors. Additionally, only one of these telephone-based studies utilized automated voice response [23], which was actually used in conjunction with life-support personnel. Five of the studies [15,27,29-31] used Web-based delivery systems for their interventions, and one study [28] utilized store-and-forward video-recorded sessions to deliver their intervention. The duration and frequency of the interventions varied and so also the total intervention time with a median of 12 weeks and range of 1 week to 12 months. Table 2 summarizes the main results from each of the studies showing effects of the intervention on primary outcomes.

Table 2. Results showing effects of the intervention on primary outcomes[a].

Author, year, country	Follow-up time	Outcome	Outcome measurement	Effect measure	Effect size	*P* value
Kim, [14] 2013, Korea	1 week	Pain	BPI	Mean pain score; proportion with pain score ≥4	-0.3; -16%	.24; .02
Rustoen, [13] 2013, Norway	6 weeks	Pain	Numerical rating scale	Mean change in pain intensity score	No effect	NS
Kroenke, [23] 2010, USA	12 months	Pain, depression	BPI, HSCL-20	Mean difference	-0.70; -0.26	<.001; <.001
Badger, [12] 2013, USA	16 weeks	Depression	CES-D	Mean difference	No effect	NS
Borosund, [27] 2014, Sweden	6 months	Depression	HADS	Mean difference compared with control	Webchoice: -0.79; IPPC: 0.69	.03; .03
Duffecy, [15] 2013, USA	8 weeks	Depression	HADS	Mean difference	0.26	--
Gotay, [25] 2007, USA	3 months	Depression	CES-D	Odds ratio of proportion with scores ≥16	1.38	.24
Lepore, [29] 2014, USA	1 month	Depression	HADS	Unstandardized regression coefficients (S-ISG=0, P-ISG=1)	1.11	.028
Livingston, [20] 2010, Australia	12 months	Depression	HADS	Mean difference	0.16; -0.19	.55; .57
Marcus, [21] 2010, USA	18 months	Depression	CES-D	Mean difference; Proportion with scores ≥16	No change in mean scores; 0.23	NS; .06
Stanton, [30] 2013, USA	6 months	Depression	CES-D	Adjusted group means	5.8	.009
Freeman, [28] 2014, USA	3 months	HR-QOL	SF-36; FACT-B	Adjusted group means	Comparing LD vs TD vs WL: SF-36 PCS: 48.32 vs 49.93 vs 46.81; SF-36 MCS: 48.77 vs 49.40 vs 44.30 FACT-B: 24.66 vs 26.03 vs 23.66	.15; .02; .08
Harrison, [19] 2011, Australia	6 months	HR-QOL	FACT-C	Mean difference	7.4	.19
Hawkins, [22] 2010, USA	6 weeks	HR-QOL	WHOQOL	Mean difference	0.26, 0.19, 0.24	All <.05
Loprinzi, [18] 2011, USA	12 weeks	HR-QOL	LASA QOL	Mean difference	2.3	–
Nelson, [24] 2008, USA	4 months	HR-QOL	FACT-Cx	Mean difference	11.57	.012
Park, [16] 2012, Korea	3 months	HR-QOL	FACT-B	Mean difference	-17.18	.002
Ryhanen, [31] 2013, Finland	Throughout treatment period, average 9 months	HR-QOL	Quality of life instrument - breast cancer patient version	Mean QOL scores (ANOVA)		.82
Sandgren, [26] 2007, USA	13 months	HR-QOL	FACT-G	Mean score	96.84 vs 95.50 vs 97.00	>.11

Author, year, country	Follow-up time	Outcome	Outcome measurement	Effect measure	Effect size	*P* value
Sherman, [17] 2012, USA	Phase-specific: within 14 days of completion of adjuvant chemotherapy or 6 months post-surgery	Psychological well-being	PAL-C	Mean change	No effect	NS

[a] HR-QOL: Health-related Quality of Life; CES-D: Center for Epidemiological Studies-Depression Scale; BPI: Brief Pain Inventory; HADS: Hospital Anxiety and Depression Scale; FACT-B: Functional Assessment of Cancer Therapy-Breast; PAL-C: Profile of Adaptation to Life Clinical Scale; LASA QOL: Linear Analog Self-Assessment Quality of Life; FACT-C: Functional Assessment of Cancer Therapy-Colorectal; WHOQOL: World Health Organization Quality of Life; HSCL-20: 20-item Hopkins Symptom Checklist; FACT-Cx: Functional Assessment of Cancer Therapy-Cervical; FACT-G: Functional Assessment of Cancer Therapy Scale-General; NS: non-significant.

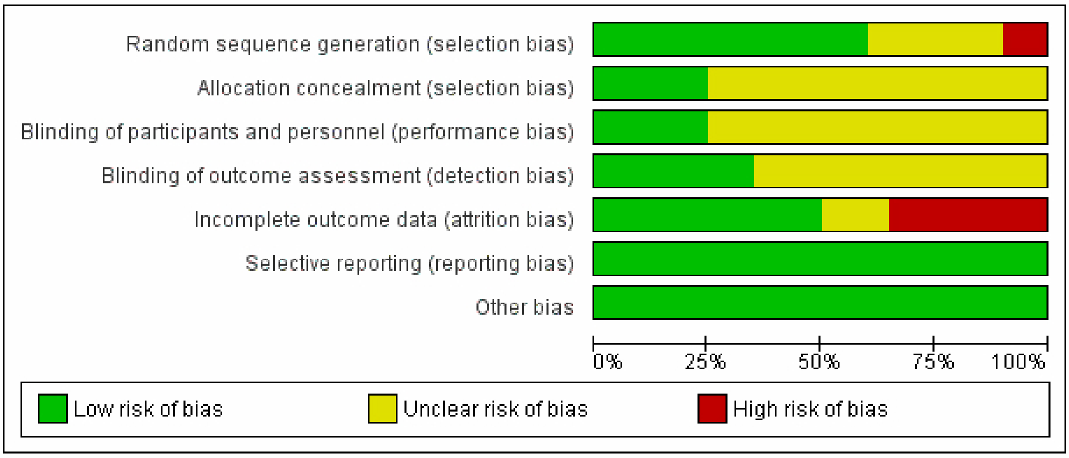
Figure 2 depicts the methodological quality of the studies included in this review. Most of the studies provided information about the method of generation of random sequence, while two studies [14,20] applied inappropriate methods of random sequence generation. Only a few studies (25%, 5/20) provided information about allocation concealment. Similarly, only a few studies (25%, 5/20) [15,17,23,29,31] had low risk of bias on the blinding of subjects and study personnel domain. Seven studies [13,15,18,21,22,24,27] were judged to have a risk of bias on the incomplete outcome reporting because of imbalance in dropout rates by group or insufficient accounting of all study participants.

Kim et al [14] compared the efficacy of pain education alone (control arm) and pain education plus telemonitoring (experimental arm) on pain and depression in a total of 108 patients with advanced solid tumors. In their trial, nursing specialists provided video-assisted educational material in both arms and daily telemonitoring for the first week in the experimental arm. There was significant improvement in pain and depression outcomes comparing baseline and final outcome in all study participants. They also reported significant reductions in number of intervention subjects with pain intensity scores ≥4 compared with control group (35%-19%, *P*=.02). Although average pain score over the past 24 hours (-1.2 vs -1.9) and worst pain scores (-0.7 vs -1.9) decreased compared to control group, these were not significant. Similarly, Harrison et al [19] evaluated the effectiveness of a nurse-delivered telephone supportive intervention (the "CONNECT" intervention) compared with usual care in 75 colorectal cancer patients. The CONNECT intervention consisted of five calls from a specialist nurse in the 6 months after initial discharge from the hospital [32]. They also found time-dependent improvement in HR-QOL within each arm but failed to reach a statistical significance comparing intervention and control groups.

Livingston et al [20] enrolled 571 newly diagnosed male CRC (n=182) and prostate (n=389) cancer patients and randomized them into three arms: two intervention arms and a passive referral arm. In the active referral arms, the specialist actively referred men to a Cancer Helpline. In Active Referral-4, patients received calls from the Helpline within 1 week of diagnosis, at 6 weeks, 3 months, and 6 months post diagnosis. In the Active Referral-1 arm, patients received only one call within 1 week of diagnosis. In the control arm, Passive Referral, patients were referred to contact the Helpline at their own initiative. The telephone helplines were developed by many cancer organizations in Australia to provide information tailored to the cancer patient's needs, support, and referral to supportive service [33]. The study included only male patients based on prior work that suggested that men were less likely to utilize supportive services [33,34]. However, they found no psychological impact of the telephone-based intervention; mean changes over time in cancer-specific depression outcomes were similar between study arms.

In 2010, Kroenke et al reported the results of the Indiana Cancer Pain and Depression (INCPAD) trial [23]. In this trial, 202 patients were randomly assigned to receive the intervention and 203 to receive usual care. Patients in the intervention group received centralized telecare management by a nurse-physician specialist team coupled with automated home-based symptom monitoring by interactive voice recording or online. They reported that the intervention resulted in improved pain and depression outcomes in cancer patients assigned to receive the intervention. The standardized effect size for between group differences at 3 and 12 months was 0.67 (95% CI 0.33-1.02) and 0.39 (95% CI 0.01-0.77) for pain, and 0.42 (95% CI 0.16-0.69) and 0.41 (95% CI 0.08-0.72) for depression.

Figure 2. Risk of bias graph for included studies.



Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

Discussion

Principal Findings

This systematic review of randomized controlled trials evaluating the effect of telehealth interventions on pain, depression, and HR-QOL outcomes in cancer care included 20 studies published over a 9-year period from 2006-2014. From this review, we make a number of observations. First, the application of telehealth in cancer care is in early stages, and all the studies were conducted in high-income countries. Second, the studies were largely heterogeneous in design and outcome assessments, making it difficult to pool effects in a meta-analysis. Third, the interventions are diverse in terms of type, content, intervention times, follow-up periods, and outcome measures.

Two of the three studies included in this review that focused on pain control reported a positive effect of telehealth on improving outcomes [14,23]. However, the study by Rustoen et al attributed inadequate dose of the psychoeducational intervention as one of the probable reasons for lack of efficacy. This is in contrast to the other two studies that relied heavily on collaborative care management between patients, their caregivers, and health care providers with extensive patient education. The INCPAD trial also included automated symptom monitoring as part of the intervention. Previous systematic reviews have identified patient education as a key component in improving cancer pain management. Lovell et al proposed four core principles that should guide the basis for patient education to successfully improve cancer pain management. These include the principles that education should be (1) patient-centered, (2) be an integral component of the patient-provider relationship, (3) aimed at patient empowerment for self-management, and (4) incorporated as part of ongoing care to counsel and support patients in the context of the severity of their pain, their needs and self-management plans. Therefore, incorporating collaborative patient-centered psychoeducational strategies in the design of technology-based interventions for pain management could improve efficacy of technology-based pain management interventions.

In contrast to the pain-focused studies, three of the eight studies evaluating the effect of telehealth on HR-QOL demonstrated improved outcomes. A recent review by Dickson et al evaluating the use of technology-based interventions for cancer follow-up surmised that the interventions did not decrease HR-QOL [35]. Similarly, four of the ten studies evaluating the effect of telehealth interventions on depression demonstrated significant improvement in depression outcomes. The SMART oncology trial, another collaborative care management approach that was delivered by a care manager under the supervision of a psychiatrist, demonstrated improvement in depression outcomes [36]. Also, a meta-analysis evaluating the effect of various interventions on depression in cancer patients showed that compared with controls, psychotherapy significantly improved outcomes and cognitive behavioral therapy was particularly associated with better outcomes [37].

It is noteworthy that the majority of interventions reported in this review were telephone-based. This is unsurprising because telephone systems are one of the oldest and most reliable information technologies available today, which makes them a very popular communication tool across different generations. They also enable personal live interactions between patients and providers, which can enhance the patient’s sense of being supported. While they are also very cheap, the cost of maintaining a professional to deliver care coupled with the fact that treatment effects could be provider-dependent may hinder scalability. While there are no doubts that the current dominance of telephone-based interventions will continue, current trends suggest that mobile phones will be the primary medium of delivery. This is evidenced by the near global ubiquity of cellular coverage and the increasing affordability, portability, and ease of use of smartphones [38]. Current estimates suggest that about 56% of US adults own a smartphone, and we envisage that this upward trend will continue [39]. While the telephone-based interventions in this review are largely voice communications, text messaging now appears to be a dominant function that is being used to engage patients across multiple disease groups.

Although not specific to connected health-related studies, previous studies have highlighted similar challenge [40].

Okuyama et al in their review of clinical trials evaluating psychosocial telephone interventions in patients with cancers and survivors also reported a similar finding that the majority of the studies reviewed lacked a standardization of outcome assessments and did not adhere adequately to reporting according to CONSORT guidelines. To standardize the reporting of technology-based interventions, the CONSORT-EHEALTH Group developed checklists to provide useful guidelines in reporting technology-based trials [41]. The time is ripe to capitalize on the current optimism of the potential of telehealth to transform care delivery. To realize this goal, we cannot overemphasize the need to design high-quality trials to comprehensively establish evidence of the effectiveness of telehealth-related studies.

Limitations

This study is not without limitations. The fact that we included only studies reported in English could have led to the exclusion of relevant studies, but we do not believe this will significantly impact our findings. We limited our search to five databases and also did not search the gray literature to find relevant studies nor did we include non-randomized, retrospective studies. We believe that evidence from prospective randomized trials will be sufficient to demonstrate effectiveness. In addition, there was a heterogeneity of outcomes, outcomes assessment measures, and comparators, which makes it difficult to estimate overall effects.

Conclusions

This is one of the first studies seeking to evaluate the effect of telehealth on pain, depression, and quality of life outcomes in patients at different stages of their cancer experience. While the studies evaluating cancer pain outcomes proved to be effective, the same could not be reported for those evaluating depression and quality of life outcomes. In total, our findings suggest that the application of telehealth in cancer care is still at a very early stage and is mostly utilized in developed nations. Evidence of its effectiveness demonstrates promise of improving pain, depression, and HR-QOL-related outcomes in cancer patients. There is a need to invest resources into developing rigorous larger-sized clinical trials, standardize outcome assessments, and improve reporting of clinical trials to demonstrate the effect of telehealth and realize the potential of transforming care delivery.

Conflicts of Interest

None declared.

References

1. World Health Organization. Cancer URL: http://www.who.int/mediacentre/factsheets/fs297/en/ [accessed 2014-11-05] [WebCite Cache ID 6TrOg0d6u]
2. National Cancer Institute. SEER Stat Fact Sheets: All Cancer Sites URL: http://seer.cancer.gov/statfacts/html/all.html [accessed 2014-11-05] [WebCite Cache ID 6TrOleNQK]
3. Zabora J, BrintzenhofeSzoc K, Curbow B, Hooker C, Piantadosi S. The prevalence of psychological distress by cancer site. Psychooncology 2001;10(1):19-28. [Medline: 11180574]
4. VanHoose L, Black LL, Doty K, Sabata D, Twumasi-Ankrah P, Taylor S, et al. An analysis of the distress thermometer problem list and distress in patients with cancer. Support Care Cancer 2014 Oct 16:-. [doi: 10.1007/s00520-014-2471-1] [Medline: 25315367]
5. Kvedar JC, Nesbitt T, Kvedar JG, Darkins A. E-patient connectivity and the near term future. J Gen Intern Med 2011 Nov;26 Suppl 2:636-638 [FREE Full text] [doi: 10.1007/s11606-011-1763-0] [Medline: 21989615]
6. Agboola S, Havasy R, Myint- U, Kvedar J, Jethwani K. The impact of using mobile-enabled devices on patient engagement in remote monitoring programs. J Diabetes Sci Technol 2013 May;7(3):623-629 [FREE Full text] [Medline: 23759394]
7. Jethwani K, Ling E, Mohammed M, Myint- U, Pelletier A, Kvedar JC. Diabetes connect: an evaluation of patient adoption and engagement in a web-based remote glucose monitoring program. J Diabetes Sci Technol 2012;6(6):1328-1336 [FREE Full text] [Medline: 23294777]
8. Pena-Robichaux V, Kvedar JC, Watson AJ. Text messages as a reminder aid and educational tool in adults and adolescents with atopic dermatitis: a pilot study. Dermatol Res Pract 2010;2010:- [FREE Full text] [doi: 10.1155/2010/894258] [Medline: 20885940]
9. Watson AJ, Kvedar JC, Rahman B, Pelletier AC, Salber G, Grant RW. Diabetes connected health: a pilot study of a patient- and provider-shared glucose monitoring web application. J Diabetes Sci Technol 2009 Mar;3(2):345-352 [FREE Full text] [Medline: 20144366]
10. Kulshreshtha A, Kvedar JC, Goyal A, Halpern EF, Watson AJ. Use of remote monitoring to improve outcomes in patients with heart failure: a pilot trial. Int J Telemed Appl 2010;2010:870959 [FREE Full text] [doi: 10.1155/2010/870959] [Medline: 20508741]
11. Higgins JPT, Green S. The Cochrane Collaboration. Version 5.0. 2011. Cochrane handbook for systematic reviews of interventions URL: http://community.cochrane.org/handbook [accessed 2015-03-06] [WebCite Cache ID 6WpMBUb1N]
12. Badger TA, Segrin C, Hepworth JT, Pasvogel A, Weihs K, Lopez AM. Telephone-delivered health education and interpersonal counseling improve quality of life for Latinas with breast cancer and their supportive partners. Psychooncology 2013 May;22(5):1035-1042. [doi: 10.1002/pon.3101] [Medline: 22573418]

13. Rustøen T, Valeberg BT, Kolstad E, Wist E, Paul S, Miaskowski C. A randomized clinical trial of the efficacy of a self-care intervention to improve cancer pain management. Cancer Nurs 2014;37(1):34-43. [doi: 10.1097/NCC.0b013e3182948418] [Medline: 23666269]
14. Kim HS, Shin SJ, Kim SC, An S, Rha SY, Ahn JB, et al. Randomized controlled trial of standardized education and telemonitoring for pain in outpatients with advanced solid tumors. Support Care Cancer 2013 Jun;21(6):1751-1759 [FREE Full text] [doi: 10.1007/s00520-013-1722-x] [Medline: 23338230]
15. 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]
16. Park JH, Bae SH, Jung YS, Kim KS. Quality of life and symptom experience in breast cancer survivors after participating in a psychoeducational support program: a pilot study. Cancer Nurs 2012;35(1):E34-E41. [doi: 10.1097/NCC.0b013e318218266a] [Medline: 21760491]
17. Sherman DW, Haber J, Hoskins CN, Budin WC, Maislin G, Shukla S, et al. The effects of psychoeducation and telephone counseling on the adjustment of women with early-stage breast cancer. Appl Nurs Res 2012 Feb;25(1):3-16. [doi: 10.1016/j.apnr.2009.10.003] [Medline: 20974079]
18. Loprinzi CE, Prasad K, Schroeder DR, Sood A. Stress Management and Resilience Training (SMART) program to decrease stress and enhance resilience among breast cancer survivors: a pilot randomized clinical trial. Clin Breast Cancer 2011 Dec;11(6):364-368. [doi: 10.1016/j.clbc.2011.06.008] [Medline: 21831722]
19. Harrison JD, Young JM, Solomon MJ, Butow PN, Secomb R, Masya L. Randomized pilot evaluation of the supportive care intervention "CONNECT" for people following surgery for colorectal cancer. Dis Colon Rectum 2011 May;54(5):622-631. [doi: 10.1007/DCR.0b013e31820bc152] [Medline: 21471765]
20. Livingston PM, White VM, Hayman J, Maunsell E, Dunn SM, Hill D. The psychological impact of a specialist referral and telephone intervention on male cancer patients: a randomised controlled trial. Psychooncology 2010 Jun;19(6):617-625. [doi: 10.1002/pon.1609] [Medline: 19673008]
21. Marcus AC, Garrett KM, Cella D, Wenzel L, Brady MJ, Fairclough D, et al. Can telephone counseling post-treatment improve psychosocial outcomes among early stage breast cancer survivors? Psychooncology 2010 Sep;19(9):923-932 [FREE Full text] [doi: 10.1002/pon.1653] [Medline: 19941285]
22. Hawkins RP, Pingree S, Shaw B, Serlin RC, Swoboda C, Han JY, et al. Mediating processes of two communication interventions for breast cancer patients. Patient Educ Couns 2010 Dec;81 Suppl:S48-S53 [FREE Full text] [doi: 10.1016/j.pec.2010.10.021] [Medline: 21081261]
23. Kroenke K, Theobald D, Wu J, Norton K, Morrison G, Carpenter J, et al. Effect of telecare management on pain and depression in patients with cancer: a randomized trial. JAMA 2010 Jul 14;304(2):163-171 [FREE Full text] [doi: 10.1001/jama.2010.944] [Medline: 20628129]
24. Nelson EL, Wenzel LB, Osann K, Dogan-Ates A, Chantana N, Reina-Patton A, et al. Stress, immunity, and cervical cancer: biobehavioral outcomes of a randomized clinical trial [corrected]. Clin Cancer Res 2008 Apr 1;14(7):2111-2118 [FREE Full text] [doi: 10.1158/1078-0432.CCR-07-1632] [Medline: 18381952]
25. Gotay CC, Moinpour CM, Unger JM, Jiang CS, Coleman D, Martino S, et al. Impact of a peer-delivered telephone intervention for women experiencing a breast cancer recurrence. J Clin Oncol 2007 May 20;25(15):2093-2099 [FREE Full text] [doi: 10.1200/JCO.2006.07.4674] [Medline: 17513815]
26. Sandgren AK, McCaul KD. Long-term telephone therapy outcomes for breast cancer patients. Psychooncology 2007 Jan;16(1):38-47. [doi: 10.1002/pon.1038] [Medline: 16862634]
27. Børøsund E, Cvancarova M, Moore SM, Ekstedt M, Ruland CM. Comparing effects in regular practice of e-communication and Web-based self-management support among breast cancer patients: preliminary results from a randomized controlled trial. J Med Internet Res 2014;16(12):e295 [FREE Full text] [doi: 10.2196/jmir.3348] [Medline: 25525672]
28. Freeman LW, White R, Ratcliff CG, Sutton S, Stewart M, Palmer JL, et al. A randomized trial comparing live and telemedicine deliveries of an imagery-based behavioral intervention for breast cancer survivors: reducing symptoms and barriers to care. Psychooncology 2014 Aug 22:-. [doi: 10.1002/pon.3656] [Medline: 25146413]
29. Lepore SJ, Buzaglo JS, Lieberman MA, Golant M, Greener JR, Davey A. Comparing standard versus prosocial internet support groups for patients with breast cancer: a randomized controlled trial of the helper therapy principle. J Clin Oncol 2014 Dec 20;32(36):4081-4086. [doi: 10.1200/JCO.2014.57.0093] [Medline: 25403218]
30. Stanton AL, Thompson EH, Crespi CM, Link JS, Waisman JR. Project connect online: randomized trial of an internet-based program to chronicle the cancer experience and facilitate communication. J Clin Oncol 2013 Sep 20;31(27):3411-3417 [FREE Full text] [doi: 10.1200/JCO.2012.46.9015] [Medline: 23940228]
31. Ryhänen AM, Rankinen S, Siekkinen M, Saarinen M, Korvenranta H, Leino-Kilpi H. The impact of an empowering Internet-based Breast Cancer Patient Pathway program on breast cancer patients' clinical outcomes: a randomised controlled trial. J Clin Nurs 2013 Apr;22(7-8):1016-1025. [doi: 10.1111/jocn.12007] [Medline: 23480498]
32. Young J, Harrison J, Solomon M, Butow P, Dennis R, Robson D, et al. Development and feasibility assessment of telephone-delivered supportive care to improve outcomes for patients with colorectal cancer: pilot study of the CONNECT intervention. Support Care Cancer 2010 Apr;18(4):461-470. [doi: 10.1007/s00520-009-0689-0] [Medline: 19626346]

33. Jefford M, Black C, Grogan S, Yeoman G, White V, Akkerman D. Information and support needs of callers to the Cancer Helpline, the Cancer Council Victoria. Eur J Cancer Care (Engl) 2005 May;14(2):113-123. [doi: 10.1111/j.1365-2354.2005.00505.x] [Medline: 15842458]
34. McCaughan E, McKenna H. Information-seeking behaviour of men newly diagnosed with cancer: a qualitative study. J Clin Nurs 2007 Nov;16(11):2105-2113. [doi: 10.1111/j.1365-2702.2006.01785.x] [Medline: 17331096]
35. Dickinson R, Hall S, Sinclair JE, Bond C, Murchie P. Using technology to deliver cancer follow-up: a systematic review. BMC Cancer 2014;14:311 [FREE Full text] [doi: 10.1186/1471-2407-14-311] [Medline: 24885758]
36. Sharpe M, Walker J, Holm HC, Martin P, Symeonides S, Gourley C, SMaRT (Symptom Management Research Trials) Oncology-2 Team. Integrated collaborative care for comorbid major depression in patients with cancer (SMaRT Oncology-2): a multicentre randomised controlled effectiveness trial. Lancet 2014 Sep 20;384(9948):1099-1108. [doi: 10.1016/S0140-6736(14)61231-9] [Medline: 25175478]
37. Hart SL, Hoyt MA, Diefenbach M, Anderson DR, Kilbourn KM, Craft LL, et al. Meta-analysis of efficacy of interventions for elevated depressive symptoms in adults diagnosed with cancer. J Natl Cancer Inst 2012 Jul 3;104(13):990-1004 [FREE Full text] [doi: 10.1093/jnci/djs256] [Medline: 22767203]
38. International Telecommunication Union. The world in 2014: ICT facts and figures URL: http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf [accessed 2014-11-05] [WebCite Cache ID 6TrOAJtwm]
39. Pew Research Internet Project. Smartphone Ownership 2013 URL: http://www.pewinternet.org/2013/06/05/smartphone-ownership-2013/ [accessed 2014-11-05] [WebCite Cache ID 6TrO209gg]
40. Okuyama S, Jones W, Ricklefs C, Tran ZV. Psychosocial telephone interventions for patients with cancer and survivors: a systematic review. Psychooncology 2014 Oct 18:-. [doi: 10.1002/pon.3704] [Medline: 25328103]
41. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

BPI: Brief Pain Inventory
CES-D: Center for Epidemiological Studies-Depression Scale
CHESS: Comprehensive Health Enhancement Support System
CRC: colorectal carcinoma
FACT-B: Functional Assessment of Cancer Therapy-Breast
FACT-C: Functional Assessment of Cancer Therapy-Colorectal
FACT-Cx: Functional Assessment of Cancer Therapy-Cervical
FACT-G: Functional Assessment of Cancer Therapy Scale-General
HADS: Hospital Anxiety and Depression Scale
HR-QOL: Health-Related Quality of Life
HSCL-20: 20-item Hopkins Symptom Checklist
INCPAD: Indiana Cancer Pain and Depression trial
LASA QOL: Linear Analog Self Assessment Quality of Life
NCCN: National Comprehensive Cancer Network
NCI: National Cancer Institute
NP: nurse practitioner
PAL-C: Profile of Adaptation to Life Clinical Scale
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial
SMART: Stress Management and Resiliency Training
WHO: World Health Organization
WHOQOL: World Health Organization Quality of Life

Edited by G Eysenbach; submitted 05.11.14; peer-reviewed by T Badger, E Nelson; comments to author 30.11.14; revised version received 18.02.15; accepted 18.02.15; published 13.03.15.

Please cite as:
Agboola SO, Ju W, Elfiky A, Kvedar JC, Jethwani K
The Effect of Technology-Based Interventions on Pain, Depression, and Quality of Life in Patients With Cancer: A Systematic Review of Randomized Controlled Trials
J Med Internet Res 2015;17(3):e65
URL: http://www.jmir.org/2015/3/e65/
doi:10.2196/jmir.4009
PMID:25793945

Original Paper

A Web-Based Health Promotion Program for Older Workers: Randomized Controlled Trial

Royer F Cook[1], PhD; Rebekah K Hersch[1], PhD; Dana Schlossberg[1], MPH; Samantha L Leaf[1], PhD

ISA Associates, Alexandria, VA, United States

Corresponding Author:
Royer F Cook, PhD
ISA Associates
201 North Union Street
Suite 330
Alexandria, VA, 22314
United States
Phone: 1 703 739 0880 ext 12
Fax: 1 703 739 0462
Email: rcook@isagroup.com

Abstract

Background: Recent evidence supports the efficacy of programs that promote improvements in the health practices of workers 50 years and older who are at higher risk for chronic diseases than younger workers are. Internet-based programs that promote healthy practices have also shown promise and, therefore, should be especially appropriate for workers aged 50 years and older.

Objective: The purpose of the research was to evaluate the effectiveness of HealthyPast50, a fully automated Web-based health promotion program based on social cognitive theory and aimed specifically at workers 50 years and older.

Methods: The randomized controlled trial was conducted across multiple US offices of a large global information technology company. The sample included 278 employees aged 50 to 68 who were recruited online and randomly assigned to the Web-based HealthyPast50 program or to a wait-list control condition. Self-report measures of diet, physical activity, stress, and tobacco use were collected online before and 3 months after the program group was given access to the program. Use data included number of log-ins and number of pages accessed. The primary analysis was multiple linear regression, following intent-to-treat principles with multiple imputation using the Markov chain Monte Carlo (MCMC) approach for nonmonotone missing data. Potential moderators from demographic characteristics and program dosage effects were assessed using multiple linear regression models. Additional analyses were conducted on complete (nonimputed) cases, excluding program participants who used the program for less than 30 minutes.

Results: Retention rates were good for both groups: 80.4% (111/138) for the program group and 94.3% (132/140) for the control group. Program group participants spent a mean of 102.26 minutes in the program (SD 148.32), logged in a mean of 4.33 times (SD 4.28), and viewed a mean of 11.04 pages (SD 20.08). In the analysis of the imputed dataset, the program group performed significantly better than the control group on diet behavioral change self-efficacy (estimated adjusted difference [Δ]=0.16, *P*=.048), planning healthy eating (Δ=0.17, *P*=.03), and mild exercise (Δ=1.03, *P*=.01). Moderator and dosage analyses of the dataset found no significant program effects. Analyses of the nonimputed dataset comparing program users with controls found additional significant program effects on eating practices (Δ=0.09, *P*=.03), exercise self-efficacy (Δ=0.12, *P*=.03), exercise planning (Δ=0.18, *P*=.03), and aging beliefs (Δ=0.17, *P*=.01). Moderator analysis of this dataset also found significant moderator effects of gender on multiple measures of exercise.

Conclusions: A Web-based health promotion program showed promise for making a significant contribution to the short-term dietary and exercise practices of older working adults. Gender effects suggest that the program effects on exercise are due mainly to improvements among women.

(J Med Internet Res 2015;17(3):e82) doi:10.2196/jmir.3399

KEYWORDS

Internet; health promotion; middle aged; nutritional requirements; exercise

Introduction

Decades of studies have confirmed the link between chronic illness and common modifiable risk factors, such as smoking, physical inactivity, poor diet, and high stress [1,2]. Moreover, there is ample evidence that the large and expanding group of older workers—a sizable group numbering more than 50 million people—are at higher risk for costly chronic diseases than their younger coworkers [3]. It has also been shown that if workers in midlife can reduce their modifiable health risks, they can forestall disability and reduce their utilization of health care services. By doing so, they can increase the likelihood of a healthful, enjoyable life in their later years while also contributing to a possible reduction in health care costs [3-5].

Consequently, providing older workers (50 and older) with effective tools that can help them improve their health practices could be beneficial to these workers, who are more likely to be affected by major diseases such as cancer, cardiovascular disease, diabetes, and musculoskeletal disease. Moreover, decreasing the health risks of older workers has special relevance for employers and health care companies because annual losses from chronic diseases exceed US $1 trillion (including productivity losses), and older workers account for a disproportionate share of the organization's health care costs [6]. In addition, from a public policy perspective, there is a particularly strong incentive for reducing the health risks of workers in this group because they will be moving onto the Medicare rolls at age 65, swelling the enrollee population to more than 70 million by 2030. There is ample evidence that well-constructed health promotion programs for the workplace can be excellent mechanisms for increasing worker health, decreasing health care costs, and improving productivity [7-9]. Moreover, there is accumulating evidence that health promotion programs developed specifically for older adults, with their distinctive set of health risks and age-appropriate health practices (eg, focusing on regular moderate activity rather than highly aerobic exercise), can significantly improve their health practices and lower their health risks [10-13].

Computer-based approaches to workplace health promotion and disease prevention strategies have become more numerous, including Web-based modes of delivery [14-22]. With the advent of broadband, high-speed connections (especially prevalent in workplaces), Web-based programs offer the advantages of tailored, media-rich psychoeducational experiences accessible at any time or place where an Internet connection is available. Research on Web-based programs indicates that such approaches can be an effective means of contributing to positive changes in diet, physical activity, stress management, and substance misuse [14-19]. However, with the exception of the study by Hughes and associates [11], none of the more promising studies focused on older working adults, but drew their samples through Internet recruiting [16,18] and church groups [15].

Because there are physical and psychological characteristics of older workers that distinguish them from younger workers, any workplace intervention directed at this older group must be tailored to their particular needs and characteristics. Most of the health promotion topics that comprise the typical workplace health promotion program (eg, promoting physical activity, healthy eating, stress management) should also be part of an older workers' program. However, both the content of the programs and the recruitment and motivational strategies would most appropriately be altered for this age group.

The purpose of the research was to evaluate the impact of a multimedia Web-based health promotion program on central health attitudes and practices of older workers. This program, called HealthyPast50, was designed as a stand-alone intervention to address a wide variety of health behavior topics, including physical activity, healthy eating, stress management, and tobacco cessation—the health behaviors that contribute to the prevention of major diseases.

Methods

Design

The impact of a multimedia Web-based program on the health practices of workers 50 years of age or older was assessed in a randomized controlled trial (RCT) in which participants were randomly assigned to the Web-based program condition or to a wait-list control condition. Participation was voluntary and all protocols and procedures were approved by the ISA Associates, Inc Institutional Review Board, Alexandria, VA. All participants, employees of a large global information technology company, were surveyed on multiple outcome measures before and 3 months after the program group received access to the Web-based program. The specific objectives of the study were to assess the extent to which the Web-based program produced significant positive changes, from baseline to posttest, on measures of stress, diet, physical activity, aging beliefs, and tobacco use. All outcome measures were self-reports, collected through an online survey. The full study protocol is shown in Multimedia Appendix 1.

Procedures

A recruitment flyer briefly describing the purpose of the study was emailed by company officials to all employees 50 years of age and older (approximately 2500 employees) located in multiple US offices of a global information technology company. The flyer stated that the study was being conducted by a research organization through a grant from the National Institutes of Health. The flyer also explained that participants would receive US $25 for completing the first questionnaire and US $25 for completing the second questionnaire, and that their name would be entered into a drawing in which 1 participant would receive US $500 during each questionnaire round. Interested employees who fit the inclusion criteria (age 50 and older) were instructed to contact the project staff directly by email or telephone. When interested employees contacted the study team, they were provided additional information about the study and their eligibility to participate (ie, age and employment at the company) was confirmed. Employees interested in participating after learning more about the project provided the study team with an email address to be used to send a personalized email and link to the online baseline survey, which also included the consent document. The full consent document is shown in Multimedia Appendix 2.

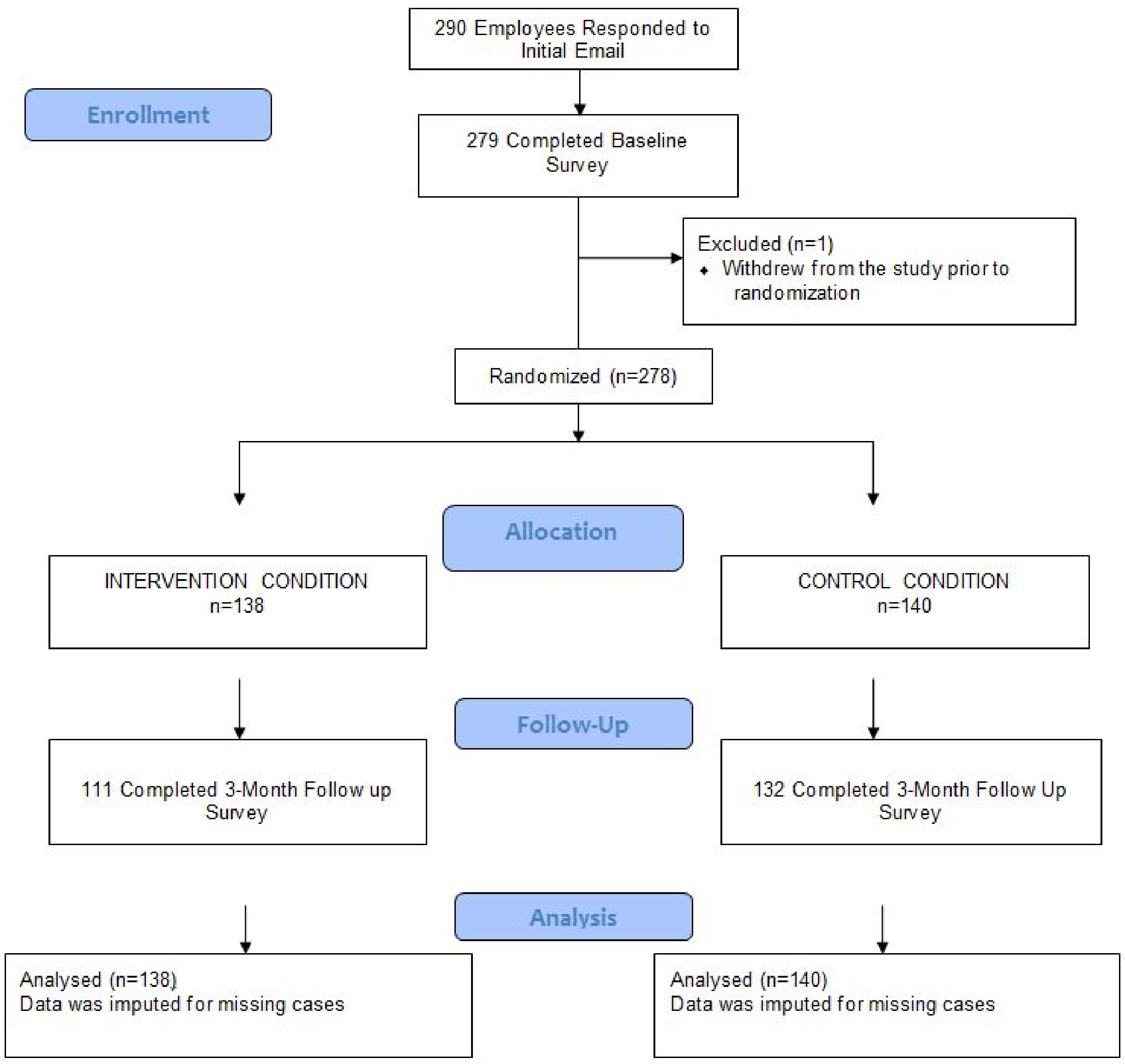
After reading the consent document, participants selected 1 of 2 responses indicating whether they consented or declined to participate. Participants were not able to continue with the survey until they acknowledged and indicated that they consented to participate. No participant declined to participate (although, as noted subsequently, some initially interested employees did not complete the baseline survey and 1 employee withdrew after consenting to participate and completing the survey). An electronic copy of the consent document was then emailed to participants after they completed the baseline survey.

A total of 290 employees contacted the study team in response to the initial email and were sent a link to the online survey, 279 of whom completed the baseline survey. One participant subsequently contacted the study team to withdraw from the study leaving a total sample size of 278 participants (Figure 1).

Randomization was conducted by the second author using a block-randomized design with blocks of 4 and 6. The 0 and 1 within each block were random and the order of the group of 4 and the group of 6 was random. Randomization occurred after each participant completed the baseline survey. The online survey program was checked every day to determine who completed the survey each day and individuals were assigned to the next condition on the randomization table as they completed the survey. Once randomization was complete, participants were notified of the condition to which they were assigned (no blinding procedures were employed) and were informed of next steps; the program group was given the program link and log-in information and the control group was told that their access to the program would be delayed until the end of the test period. Participants could complete the online questionnaires and (for those in the program condition) access the online program on work time or at home.

Participants in the program group could access the Web-based program at any time during the 3-month test period, both at work or outside work (eg, at home). The program operated alone and automatically; no human contact was involved. A “project update” email was sent to participants at 1 month and 2 months after randomization. For the program group, the emails included a reminder to use the program and information about the latest update to the program (a brief message on the home page). With the exception of the project update, the program was unchanged throughout the test period. For the control group, the email included information about when the second survey would be available. In addition, the project staff was always available to answer questions by telephone and email if participants had any difficulty accessing the program. The data collection started on October 10, 2012 (first participant enrolled) and ended on February 23, 2013 (last participant completed posttest). Individual access to program by participants was limited to the 3-month test period. No discernable secular events of note occurred during the test period. At the end of the 3-month test period and prior to administering the follow-up survey, access to the program was blocked. Three months after randomization, participants were sent an email with the link to the follow-up survey. After the posttest survey was complete, all participants received access to the program.

Figure 1. Participant flow diagram.



Sample

Study participants were 278 employees of a large global information technology company whose US offices are located primarily in Massachusetts and California. The study was powered for a sample size of 250 (after attrition) based on previous studies by the authors. The computer literacy of the workforce was estimated to be relatively high. Demographic characteristics of the sample are displayed in Table 1.

Participants ranged in age from 50 to 68 years, 67.3% (187/278) of whom were male and 89.6% (249/278) identified as Caucasian. Participants were relatively affluent and well educated: 64.7% (180/278) of participants had a Bachelor’s degree or higher and 70.1% (195/278) reported a household income of US $100,000 or more. A comparison of the program group with the control group on demographics and outcome variables at baseline revealed no significant differences between the groups indicating that randomization was successful.

Table 1. Participant characteristics at baseline.

Demographics	Control, n (%) n=140	Program, n (%) n=138	Total, n (%) N=278
Gender			
Male	89 (63.6)	98 (71.0)	187 (67.3)
Female	50 (35.7)	40 (29.0)	90 (32.4)
Prefer not to answer	1 (0.7)	0	1 (0.4)
Race			
Black or African American	6 (4.3)	3 (2.2)	9 (3.2)
Caucasian	123 (87.9)	126 (91.3)	249 (89.6)
American Indian/Alaska Native	1 (0.7)	0	1 (0.4)
Asian	5 (3.6)	2 (1.4)	7 (2.5)
Native Hawaiian or Other Pacific Islander	1 (0.7)	1 (0.7)	2 (0.7)
Multiracial	2 (1.4)	4 (2.9)	6 (2.2)
Prefer not to answer	2 (1.4)	1 (0.7)	3 (1.1)
Latino			
Yes	4 (2.9)	4 (2.9)	8 (2.9)
No	134 (95.7)	133 (96.4)	267 (96.0)
Prefer not to answer	2 (1.4)	1 (0.7)	3 (1.1)
Age			
50-54	74 (52.9)	64 (46.4)	138 (49.6)
55-59	36 (25.7)	47 (34.1)	83 (29.9)
60-64	22 (15.7)	22 (15.9)	44 (15.8)
65-69	8 (5.7)	5 (3.6)	13 (4.7)
Marital status			
Single	6 (4.3)	9 (6.5)	15 (5.4)
Married	114 (81.4)	101 (73.2)	215 (77.3)
Divorced	14 (10.0)	16 (11.6)	30 (10.8)
Separated	0	5 (3.6)	5 (1.8)
Widowed	2 (1.4)	3 (2.2)	5 (1.8)
Living with a partner	4 (2.9)	4 (2.9)	8 (2.9)
Income (US $)			
<$60,000	4 (2.9)	6 (4.3)	10 (3.6)
$60,000-$79,999	7 (5.0)	12 (8.7)	19 (6.8)
$80,000-$99,999	15 (10.7%)	11 (8.0)	26 (9.4)
$100,000 or more	102 (72.9)	93 (67.4)	195 (70.1)
Prefer not to answer	12 (8.6)	16 (11.6)	28 (10.1)
Education			
High school diploma or equivalent	8 (5.7)	6 (4.3)	14 (5.0)
Vocational/technical training after high school	5 (3.6)	12 (8.7)	17 (6.1)
Some college, but no degree	22 (15.7)	19 (13.8)	41 (14.7)
Associate’s degree (AA, AS)	12 (8.6)	14 (10.1)	26 (9.3)
Bachelor’s degree (BA, BS, AB)	45 (32.1)	43 (31.2)	88 (31.7)
Graduate or professional school, but no degree	10 (7.1)	13 (9.4)	23 (8.3)

Demographics	Control, n (%) n=140	Program, n (%) n=138	Total, n (%) N=278
Master’s degree (MA, MS)	34 (24.3)	28 (20.3)	62 (22.3)
Doctorate degree (PhD, EDD)	4 (2.9)	2 (1.4)	6 (2.2)
Professional degree beyond Bachelor’s degree (MD, DDS, JD)	0	1 (0.7)	1 (0.4)

Intervention

HealthyPast50 is a Web-based multimedia program containing information and guidance on the major health promotion topics of healthy aging, diet, physical activity, stress management, and tobacco use. In addition, a central module of the program contained assessments across the major health topics, providing recommendations on particular segments of HealthyPast50 that users should visit based on the results of the assessments. Sample screenshots are shown in Figure 2 and in Multimedia Appendix 3. An outline of program content is shown in Figure 3. On the home page, users were encouraged to complete the assessments before accessing the other modules. These recommendations were the only form of tailoring used in the program. With the exception of the module on healthy aging, all the major modules incorporated material from previously developed and tested programs from our group (for which evidence of efficacy was shown), modifying the content and approaches for the 50 and older audience [19,22-23].

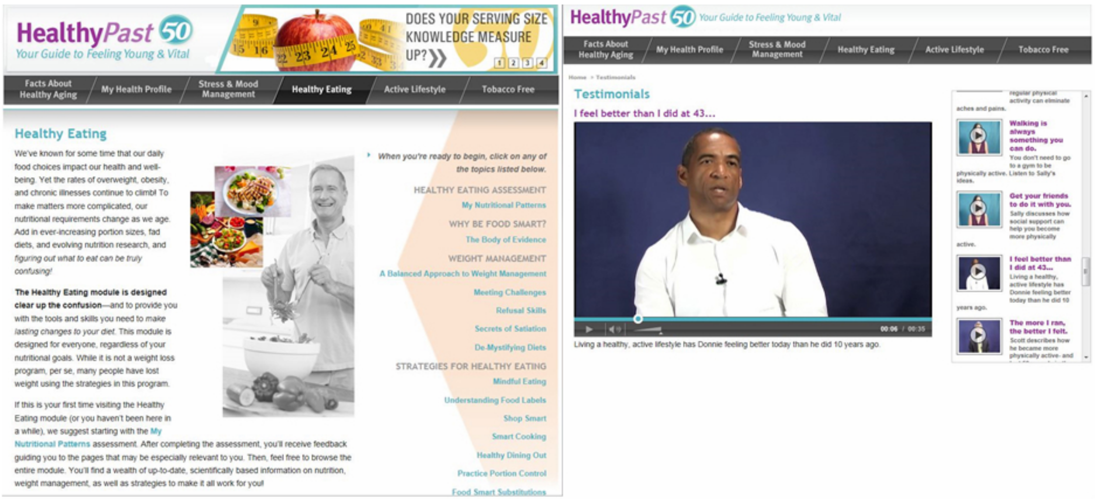
Program content was shaped specifically for older adults in many ways. Throughout the program, from the home page through all the modules, there are numerous depictions (eg, photos, graphics) of older adults along with dozens of video segments (testimonials) featuring people older than 50 years of age describing their health practices, struggles, and successes. An entire module is devoted to “Facts About Healthy Aging,” emphasizing the importance of making positive changes in health practices now, and containing a segment on “Aging Myths and Facts.” The assessment module (My Health Profile) is geared to older adults, tailored by age (eg, 50-59, 60-69) and including questions about current diseases and preventive health screenings. A central segment of the Stress and Mood Management module describes with text, video, graphics, and narration, the relationship between stress and aging, featuring detailed information on how stress accelerates the aging process at the cellular level by shortening the ends of chromosomes (telomeres). Dietary information in the Healthy Eating module emphasizes the role of a nutritious diet and a healthy weight in preventing chronic disease and increasing longevity. Much of the content of the Active Lifestyle module is oriented toward older users, including the use of the Physical Activity Readiness Questionnaire (PAR-Q), assessing one’s readiness for physical activity, the relationship of physical activity to mobility and energy in later years, and the need to moderate the intensity of exercise as one ages.

The program functions were monitored throughout the test period to ensure quality of operation. As with our group’s previous Web-based interventions, HealthyPast50 was shaped by a social-cognitive conceptual model based primarily on the work of Bandura [24-25], emphasizing the boosting of self-efficacy, self-regulation, and planning. A central premise of this model is that to achieve lasting improvements in health behavior, an intervention must do more than provide information about a given health topic; it must also provide the skills and motivation that are essential to making lasting improvements in one’s health practices.

The Web-based HealthyPast50 program was developed by our group over a 2-year period through multiple cycles of development and testing, beginning with focus groups of older workers (age 50 and older) providing feedback on specific features of the planned program, followed by ratings of prototype content, and culminating in the workplace-based RCT. The program was constructed using the ColdFusion-based Mura content management system, with the many interactive elements and assessments developed in Flash. The program also contains ample graphics, audio, and video. Many of these main elements are congruent with health behavior change theory and principles (eg, providing opportunities for observational learning, building self-efficacy, and self-tailoring of content and sequence).

For access to the full HealthyPast50 program for review or replication purposes, please contact the first author.

Figure 2. Screenshots from HealthyPast50.



Healthy Eating

We've known for some time that our daily food choices impact our health and well-being. Yet the rates of overweight, obesity, and chronic illnesses continue to climb! To make matters more complicated, our nutritional requirements change as we age. Add in ever-increasing portion sizes, fad diets, and evolving nutrition research, and *figuring out what to eat can be truly confusing!*

The Healthy Eating module is designed clear up the confusion—and to provide you with the tools and skills you need to *make lasting changes to your diet*. This module is designed for everyone, regardless of your nutritional goals. While it is not a weight loss program, per se, many people have lost weight using the strategies in this program.

If this is your first time visiting the Healthy Eating module (or you haven't been here in a while), we suggest starting with the My Nutritional Patterns assessment. After completing the assessment, you'll receive feedback guiding you to the pages that may be especially relevant to you. Then, feel free to browse the entire module. You'll find a wealth of up-to-date, scientifically based information on nutrition, weight management, as well as strategies to make it all work for you!

When you're ready to begin, click on any of the topics listed below.

Living a healthy, active lifestyle has Donnie feeling better today than he did 10 years ago.

regular physical activity can eliminate aches and pains.

Walking is always something you can do. You don't need to go to a gym to be physically active. Listen to Sally's ideas.

Get your friends to do it with you. Sally discusses how social support can help you become more physically active.

I feel better than I did at 43... Living a healthy, active lifestyle has Donnie feeling better today than he did 10 years ago.

The more I ran, the better I felt. Scott describes how he became more physically active- and

Figure 3. HealthyPast50 content outline.

- ***Welcome & Introduction***
  - About the Program
- ***Facts About Healthy Aging***
  - Why Healthy Past 50
  - Making the Right Choice
  - The Time is Now
  - Keys to Success
  - Myths and Facts
- ***My Health Profile***
  - Health Questionnaire and Recommendations
- ***Stress and Mood Management***
  - Assess Your Stress
  - The Importance of Managing Stress
    - Good and Bad Stress
    - Stress and Your Body
    - Stress and Age
  - Manage Stress
    - The Basics of Stress Management
    - Reacting vs Responding
    - Change Your View of Stressors
      - The Mental Lens
      - Learned Optimism/Authentic Happiness
      - Building Resilience
      - Mindfulness
      - Changing Negative Thinking
    - Change Your Body's Response to Stress
      - Physical Activity
      - Progressive Relaxation
      - Breathwork
      - Guided Imagery
      - Meditation
      - Social Support
    - Change the Stressful Situation
      - Assertiveness
      - Problem Solving
  - When It's More Than Just Stress
    - Depression, Anxiety, Getting Help
- **Healthy Eating**
  - Healthy Eating Assessment
  - Why Be Food Smart?
    - Body of Evidence
  - Weight Management
    - A Balanced Approach to Weight Management
    - Meeting Challenges
    - Refusal Skills
    - Secrets of Satiation
    - De-Mystifying Diets
  - Strategies for Healthy Eating
    - Mindful Eating
    - Understanding Food Labels
    - Shop Smart
    - Smart Cooking
    - Healthy Dining Out
    - Practice Portion Control
    - Food Smart Substitutions
    - Setting S.M.A.R.T Goals
  - Balancing the Food on Your Plate
    - Introducing the "Healthy Eating Plate"
    - Is This a Healthy Plate?
    - Get the Facts
- **Active Lifestyle**
  - Active Lifestyle Assessments
  - Why Be Physically Active?
    - Body of Evidence
    - What is Fit?
  - Starting and Staying Active
    - What's Holding You Back?
    - Setting S.M.A.R.T Goals
    - Top 15 Tips for Staying Motivated
    - Balanced Fitness
    - How Much, How Hard, How Often
    - Measuring Intensity
    - Adjusting Intensity
    - Structuring Your Physical Activity Session
    - Fueling Your Activity
    - Preventing Injury and Illness
    - If Injury Happens
    - Shoes and Gear
- **Tobacco Free**
  - Assess Your Tobacco Use
  - Why Be Tobacco Free?
    - Body of Evidence
    - Nicotine is a Drug
    - Improve Your Health Immediately
    - What Doctors Say
  - Preparing to Quit
    - Understanding My Smoking Habit
    - Setting S.M.A.R.T Goals
    - Plan Your Quit Program
    - Preparation Checklist
    - Getting Through Quit Day
  - Remaining Tobacco Free
    - Healthy Alternatives
    - The Importance of Social Support
    - Learning to Say "NO!"
    - Managing Triggers
    - Dealing with Cravings
    - Handling Slips
    - Concerns About Weigh Gain
    - Staying on Track

Measures

Overview

The 30-45 minute online self-report survey contained the measures described subsequently. In addition, program utilization data, including number of log-ins, minutes in the program, and pages accessed, were recorded for each member of the program group.

Demographics

Seven items assessing respondents' gender, age, race, ethnicity, marital status, education, and income.

Symptoms of Distress

A 15-item scale developed by Orioli et al [26] and used in multiple studies by our team [19,27]. Each item describes a physical or emotional symptom of distress (eg, muscle tension, nervousness) with a 4-point response scale indicating the frequency with which the symptom was felt in the past 30 days, ranging from 1 (nearly every day) to 4 (never); a higher score=less stress (α=.85).

Coping With Stress

Twelve items assessing the type of strategies one uses to cope with difficult situations and events [26]. Questions are answered on a 4-point scale ranging from 1 (never) to 4 (almost always); higher score=better coping. Typical questions included "I often put things aside for a while to get perspective on them" and "I decide certain problems are not worth worrying about" (α=.76).

Diet Outcome Expectancies

A 9-item scale developed and validated by Trenkner and associates [28] assessing perceived benefits to eating a healthy diet. The response scale is a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree); a higher score=higher perceived benefits. Typical items included "Eating a poor diet increases my chances of getting diseases like heart disease, cancer, and diabetes" and "Eating more fruits and vegetables will make me healthier" (α=.80).

Barriers to a Healthy Diet

An 8-item scale also developed and validated by Trenkner and associates [28], assessing perceived barriers to eating a healthy diet. The response scale is a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree); a higher score=lower perceived barriers. Typical items included "It's easy to buy healthy foods in a grocery store" and "A lot of things get in the way of my eating a more healthy diet" (α=.82).

Eating Practices

A 10-item subscale that is part of the Weight Control Assessment scale developed by O'Neil and Rhodes [29], the measure assesses the frequency with which respondents exercise control over their eating during the past 30 days. The response scale is a 4-point Likert scale ranging from 1 (almost always) to 4 (never); a higher score=better control over eating. It contained items such as "How often do you eat between meals" and "Do you have trouble controlling your eating when your favorite foods are around the house?" (α=.77).

Overeating Self-Efficacy

A 15-item scale assessing one's confidence in resisting overeating in different situations, with responses on a 7-point scale ranging from 1 (no difficulty controlling overeating) to 7 (most difficulty controlling overeating). Typical items included "overeating when depressed" and "overeating around holiday time." Developed by McCann et al [30], this scale is a shortened version of the 25-item Eating Self-Efficacy Scale [31]. A lower score indicates greater self-efficacy (α=.95).

Diet Change Self-Efficacy

A 5-item scale previously used by the study team [22] that asks respondents how confident they are that they can change their dietary practices. The response scale has 5 points, from 1 (not confident) to 5 (extremely confident); a higher score=higher self-efficacy. Typical items included "How confident are you that you have the skills to eat a healthy diet?" and "How confident are you that you have the skills to eat more fruits and vegetables?" (α=.88).

Planning Healthy Eating

A 2-item scale that asks respondents if they have a good plan for "maintaining a nutritious diet" and for "minimizing the amount of fats and sugars in my diet." The response scale for both items is a 4-point scale, ranging from 1 (not at all true) to 4 (definitely true); a higher score=better planning (item r=.77).

Weight and Body Mass Index

Two items ask respondents to report their height and weight, yielding both weight and body mass index (BMI) measures.

Exercise Habits

The Godin Leisure-Time Exercise Questionnaire [32] is a brief 4-item query of usual leisure-time exercise habits, generating 5 activity scores. The first 3 items ask respondents to indicate the times per week they engage in strenuous, moderate, and mild exercise; the 3 items are combined to yield a score for overall exercise. The fourth item (the "sweat" score) asks how often the respondent engages in an activity long enough to work up a sweat; the 3-point response scale ranges from 1 (often) to 3 (never/rarely). A lower sweat score is better. Reliability and concurrent validity of the measure were demonstrated by Godin and Shephard [32].

Exercise Self-Efficacy

An 8-item scale assessing the respondent's confidence in engaging in regular exercise. All items begin with "I am confident that..." and are answered on a 4-point response scale from 1 (not at all true) to 4 (always true); a higher score=higher exercise self-efficacy. Typical items included "I can accomplish the physical activity and exercise goals that I set" and "I can overcome barriers to and challenges to physical activity and exercise if I try hard enough." This scale was developed by Kroll et al [33] (α=.91).

Self-Efficacy for Overcoming Barriers to Exercise

A 13-item scale that asks respondents how confident they are that they can "exercise 3 times a week for the next 3 months" under potentially difficult circumstances, including bad weather, lack of interest, schedule conflicts, etc. The 10-point response scale ranges from 0 (not at all confident) to 100 (highly confident) in 10-point intervals; a higher score=higher self-efficacy to overcome barriers to regular exercise. This scale was developed by McAuley [34] (α=.94).

Exercise Planning

A 2-item scale that asks respondents if they have a good plan for "incorporating regular physical activity into my life" and "overcoming barriers to getting regular physical activity." The response scale for both items is a 4-point scale, ranging from 1 (not at all true) to 4 (definitely true); a higher score=a better plan (item *r*=.85).

Beliefs About Aging

A 5-item scale that asks respondents about the extent to which they hold healthful attitudes about aging. Typical items include "The slowing of metabolism with age makes it even more important to eat a nutritious and well-balanced diet" and "Physical and mental decline is a natural part of the aging process, and you really can't slow it down" (reverse scored). The response scale is a 5-point Likert scale ranging from 1 (strongly agree) to 5 (strongly disagree); a higher score=more healthful attitudes (α=.54).

Tobacco Use

Seven items assessed whether or not participants currently smoke cigarettes or use other tobacco products and, if so, how often they use tobacco products, how many times they use tobacco products per day, and whether and how often they have tried to quit using tobacco products.

Analysis

Multiple linear regression models were used to examine program effects on outcome measures and potential moderators and to assess dosage effects. Analyses followed intent-to-treat principles, including all participants irrespective of protocol violations and events arising postrandomization [35]. Multiple imputation was conducted for missing values of all outcome variables, using the Markov chain Monte Carlo (MCMC) approach for nonmonotone missing data [36]. Data used to construct the imputation models were baseline demographic characteristics (eg, gender, age, race, marital status, education, and income), group assignment, and the corresponding baseline (pretest) measures. The program effects, moderators, and program dosage effects for each outcome variable were assessed using multiple linear regression models after adjusting for the baseline measure, based on 20 imputed datasets. SAS Version 9.3 (SAS Inc, Cary, NC, USA) was used to perform each analysis (PROC GLM for the regression models; PROC MI and MIANALYZE for the multiple imputation).

On most outcome variables, missing data occurred for less than 15% of the participants. Before imputing missing data, attrition analysis was conducted to examine potential baseline differences between those who completed the follow-up survey and those who did not. These data indicated that responders and nonresponders were virtually equivalent. Chi-square tests showed no differences between the groups on age (*P*=.41), gender (*P*=.99), race (*P*=.38), income (*P*=.80), or education (*P*=.33). Comparing the 2 groups on all outcome measures at baseline revealed significant differences only on symptoms of distress (t_{269}=2.53, *P*=.01).

To assess program effects on program users, we also conducted the same multiple linear regression analysis on a nonimputed dataset that excluded participants in the program group who used the program for less than 30 minutes.

Results

Program Effects on Primary Outcomes

Table 2 presents the results of the estimated adjusted posttest difference between program and control groups on measures of diet, exercise, and stress.

The program group showed significantly greater improvement than the control group on diet behavioral change self-efficacy (estimated adjusted difference [Δ]=0.16, *P*=.048), and planning healthy eating (Δ=0.17, *P*=.03). The estimated adjusted difference [Δ] for eating practices was 0.07 and was not statistically significant (*P*=.08). In addition, there were no differences between the groups on diet outcome expectancies, healthy diet barriers, or overeating self-efficacy.

Compared to the control group, the program group showed significant improvement on mild exercise (Δ=1.03, *P*=.01). Two other measures of physical activity, including moderate exercise (Δ=0.47, *P*=.06) and overall exercise (Δ=4.98, *P*=.08) were not statistically significant. In addition, there were no differences between the groups on strenuous exercise, sweat, exercise self-efficacy, exercise planning, or self-efficacy for overcoming barriers to exercise. There were no differences between the program and control groups on symptoms of distress or coping with stress. However, the lack of a significant effect may be partly a function of a ceiling effect (ie, the mean scores of both groups at baseline indicated that participants in both groups entered the test period with relatively low stress and high coping skills).

There was also no significant difference between the groups on the measure of aging beliefs, although the trend was in the desired direction (Δ=0.09, *P*=.16).

There were insufficient numbers of smokers in the sample (only 16 smokers in the sample) to perform meaningful analysis on the measure of tobacco use.

Table 2. Adjusted program effects on dependent measures

Measure	Estimated adjusted posttest difference between program and control groups,[a] Δ (95% CI)	*P*
Eating and diet measures		
Diet outcome expectancies	0.02 (–0.08, 0.11)	.76
Healthy diet barriers	0.05 (–0.07, 0.17)	.43
Eating practices	0.07 (–0.01, 0.15)	.08
Overeating self-efficacy	–0.14 (–0.36, 0.08)	.20
Diet change self-efficacy	0.16 (0.00, 0.31)	.05
Planning health eating	0.17 (0.01, 0.33)	.03
BMI	0.07 (–0.28, 0.41)	.70
Exercise measures		
Godin: Strenuous exercise	–0.11 (–0.52, 0.31)	.61
Godin: Moderate exercise	0.47 (–0.01, 0.96)	.06
Godin: Mild exercise	1.03 (0.26, 1.81)	.01
Godin: Sweat	0.08 (–0.08, 0.23)	.33
Godin: Overall exercise	4.98 (–0.66, 10.62)	.08
Overcoming barriers to exercise self-efficacy	–0.68 (–5.55, 4.19)	.78
Exercise self-efficacy	0.08 (–0.02, 0.18)	.11
Exercise planning	0.11 (–0.04, 0.25)	.15
Stress and coping measures		
Symptoms of distress	0.05 (–0.03, 0.13)	.22
Coping with stress	0.01 (–0.05, 0.07)	.79
Aging beliefs	0.09 (–0.03, 0.21)	.16

[a] Adjusted for the corresponding baseline measure.

Moderator and Dosage Effects

To determine whether program effects on outcomes differed based on participant demographics, potential moderators were tested on all outcome measures. Regression analysis models included main effects as well as interactions between condition and potential moderators. No significant interactions were detected between condition and gender, age, marital status, education, or income on any outcome measures.

Data on number of log-ins, minutes in the program, and number of pages accessed were recorded for all members of the program group. An examination of the distribution of number of minutes in the program revealed one major outlier of 2053 minutes (more than 34 hours), more than twice as long as the next longest number of minutes. In calculating average use of the program, this participant's data were removed. The mean number of log-ins was 4.33 (SD 4.28, range 0-28), the mean number of minutes in the program was 102.26 minutes (SD 148.32), and the mean number of pages viewed was 11.04 (SD 20.08, range 0-120). An examination of the distribution of number of minutes in the program showed that 39 participants spent less than 30 minutes in the program and 99 participants—71.7% (99/138) of the program group—spent 30 minutes or more in the program.

To assess the extent to which the program effects were associated with the extent to which participants in the program group accessed the HealthyPast50 program, multiple regression analysis was conducted to examine the dosage effect of the number of pages viewed on all outcome variables. There were no significant associations between the number of pages viewed and any of the outcome measures after adjusting for the corresponding baseline measures.

Program Effects Excluding Nonusers

A total of 39 participants in the program group who used the program less than 30 total minutes were defined as "nonusers" and were excluded from the analysis of the nonimputed dataset. The results of these analyses are shown in Table 3. In addition to the significant program effects found in the analysis of the imputed dataset, the analysis of the nonimputed dataset with nonusers excluded found significant program effects on eating practices (Δ=0.09, *P*=.03), exercise self-efficacy (Δ=0.12, *P*=.03), exercise planning (Δ=0.18, *P*=.03), and aging beliefs (Δ=0.17, *P*=.01).

Table 3. Adjusted program effects on dependent measures comparing legitimate program users and controls using complete cases.

Measure	Estimated adjusted posttest difference between program and control groups,[a] Δ (95% CI)	*P*
Eating and diet measures		
Diet outcome expectancies	–0.01 (–0.12, 0.10)	.85
Healthy diet barriers	0.06 (–0.06, 0.19)	.34
Eating practices	0.09 (0.01, 0.18)	.03
Overeating self-efficacy	–0.24 (–0.47,–0.00)	.047
Diet change self-efficacy	0.18 (0.02, 0.34)	.03
Planning health eating	0.24 (0.07, 0.40)	.01
BMI	0.05 (–0.38, 0.47)	.83
Exercise measures		
Godin: Strenuous exercise	–0.09 (–0.55, 0.36)	.68
Godin: Moderate exercise	0.35 (–0.23, 0.92)	.24
Godin: Mild exercise	0.78 (–0.06, 1.62)	.07
Godin: Sweat	0.03 (–0.14, 0.21)	.70
Godin: Overall exercise	3.43 (–2.76, 9.63)	.28
Overcoming barriers to exercise self-efficacy	–0.44 (–5.62, 4.74)	.87
Exercise self-efficacy	0.12 (0.01, 0.23)	.03
Exercise planning	0.18 (0.02, 0.34)	.03
Stress and coping measures		
Symptoms of distress	0.08 (–0.01, 0.17)	.09
Coping with stress	0.03 (–0.04, 0.10)	.37
Aging beliefs	0.17 (0.04, 0.29)	.01

[a] Adjusted for the corresponding baseline measure.

Moderator and Dosage Effects Excluding Nonusers

Moderator analysis of the dataset excluding the 39 nonusers found that the interactions between gender × condition on overall exercise (*P*=.05), moderate exercise (*P*=.06), and sweat (*P*=.07) did not meet the threshold for statistical significance. As shown in Table 4, analysis of program effects conducted separately for males and females found significant program effects for females on overall exercise and exercise planning, with improvements in the desired direction on moderate exercise. No program effects were found for males on any of the exercise outcomes.

Multiple regression analysis of the dataset excluding the 39 nonusers examined the dosage effect of the number of pages viewed on all outcome variables. There were no significant associations between the number of pages viewed and any of the outcome measures after adjusting for the corresponding baseline measures.

Table 4. Adjusted program effects on dependent measures comparing program users and controls by gender.

Measure	Estimated adjusted posttest difference between program and control groups,[a] Δ (95% CI)	*P*
Godin: Moderate exercise		
Male	–0.06 (–0.71, 0.59)	.85
Female	1.09 (–0.08, 2.27)	.07
Godin: Sweat		
Male	0.15 (–0.06, 0.37)	.16
Female	–0.20 (–0.53, 0.12)	.22
Godin: Overall exercise		
Male	–0.37 (–7.63, 6.88)	.92
Female	12.47 (1.20, 23.75)	.03
Exercise planning		
Male	0.09 (–0.10, 0.27)	.34
Female	0.39 (0.06, 0.72)	.02

[a] Adjusted for the corresponding baseline measure.

Discussion

Principal Results

This randomized trial showed that working adults 50 years of age and older who were given access to the Web-based HealthyPast50 program showed significantly greater improvement on key health constructs over the 3-month test period than their 50 years and older counterparts in the control group. In the analysis of the imputed dataset, the program group performed significantly better than the control group on diet behavioral change self-efficacy, planning healthy eating, and mild exercise, and there were improvements on eating practices, moderate exercise, and overall exercise, but these did not meet the threshold for statistical significance. Moderator and dosage analyses of the imputed dataset found no significant effects. No significant program effects were found on measures of stress or aging beliefs. These results, following intent-to-treat principles and using multiple imputation methods, stand as the primary findings of the study and suggest that the HealthyPast50 program moved participants toward healthier eating and exercise practices. However, the program effects were selective because several measures of dietary and exercise outcomes showed no effects on participants.

Although the results from the analysis of the imputed dataset provide the most rigorous assessment of the program's efficacy, the results of the parallel analyses on the nonimputed dataset comparing program users with controls are also of interest because they indicate the extent to which HealthyPast50 improved health outcomes for the participants defined as program users—those who used the program for at least 30 minutes (71.7% of the program group). These analyses found additional significant program effects on eating practices, exercise self-efficacy, exercise planning, and aging beliefs. Interestingly, the moderator analyses of the program users showed that program effects on exercise were due largely to significant effects on females, not males. An examination of pretest and posttest means on exercise measures indicated that the lack of significant improvement by men was at least partly a function of the fact that men in both groups exhibited much higher mean scores than women at pretest.

The effects of the Web-based program on multiple measures of diet and exercise are noteworthy, as numerous studies have shown diet and exercise to be critical behaviors—perhaps the most critical health behaviors after tobacco use—in reducing the risk of major disease. By reducing, at least partially, these key modifiable health risks in midlife, working adults might be laying the foundation for improved health and vitality in their later years.

These findings indicate not only that HealthyPast50 was more effective for women than men, but also that the program did not attract as many men with poor health practices as desired. There has not been a focus on these particular types of gender effects in studies of Web-based interventions; however, the role of self-efficacy and planning among women in this trial has a parallel in a study of an in-person exercise intervention tested in Finland [37]. At 3 months, increases in self-efficacy and planning among women predicted increases in their exercise levels; the same relationship was not found among men.

Comparison With Prior Work

These findings are congruent in many respects to the findings from previous research by our group, as well as research by other investigators who have tested the efficacy of Internet-based interventions on the health practices of working adults. Across several randomized trials testing the effects of multiple Web-based programs on working adults, our group found significant effects on outcome measures of diet, exercise, stress, and substance misuse [17,19,22,38]. However, in this trial the effects of the HealthyPast50 program on dietary and exercise practices were somewhat stronger than in the trials of the other programs. The test of the Web-based Health Connection found effects on working adults' dietary attitudes, but not on dietary or exercise practices [22]. Similarly, the Web-based Heart Healthy program exhibited significant impact on dietary attitudes and self-efficacy, as well as exercise practices, but did not show significant effects on eating practices [38]. Like Health Connection and Heart Healthy, HealthyPast50 is a multimedia, interactive, theory-based program. The main difference between HealthyPast50 and the previously tested programs is that the content of HealthyPast50 was tailored to—and tested on—adults aged 50 years and older.

To our knowledge, HealthyPast50 is the only Web-based program specifically developed for—and shown to be efficacious with—working adults 50 years and older. A recent trial of a health promotion program targeting older workers compared an in-person program with a publicly available Web-based program and found significant effects on both diet and exercise behaviors for the in-person program, but few effects for the Web-based program [11].

Tests by other investigators of Web-based programs only have found results similar to the findings of our trial. The multiple trials of the Web-based Guide to Health (GTH) program have generated impressive results on both dietary and physical behaviors [15,18]. Of particular interest is their study involving older participants (mean age 58.11 years) which showed significant effects on physical activity and related social cognitive theory (SCT) constructs [18]. In a causal model, increases in self-efficacy at 7 months led to increased physical activity levels at 16 months, suggesting that interventions with aging adults that boost self-efficacy might help older participants become more active [18]. These findings would seem to lend support to the view that the self-efficacy increases found in the HealthyPast50 trial contributed to the improvements in dietary and exercise practices, especially among women. In their randomized trial of a Web-based program to promote physical activity, Carr and associates found that compared to a control group with access to public health websites (eg, Mayo Clinic), the program group showed significant improvements at 3 months in total minutes of physical exercise, although the difference was not maintained at 6 months [16]. They attributed the promising effects of their Web-based program to the use of formative focus groups and the targeted inclusion of SCT elements in their program. With their Web-based programs rooted in SCT and containing multiple elements designed to increase self-efficacy and related SCT constructs, these interventions appear to be quite similar to HealthyPast50—and have generated similar positive effects on health behaviors.

Limitations

Although this study exhibited a variety of strengths, including a randomized design, an advanced Web-based intervention, and a sizable workforce sample, the study also has some limitations, including—and perhaps foremost—the single posttest at 3 months and the reliance on self-reports. In addition, because of the particular characteristics of the sample, caution should be exercised in generalizing these findings to workforces that are less educated and affluent. Future research on HealthyPast50 should include longer-term posttests and the inclusion of physical measures (eg, weight, waist circumference, blood pressure). The program also needs to be tested on older workers who are less educated and affluent.

Conclusions

The Web-based HealthyPast50 program demonstrated significant effects on the short-term dietary and exercise practices of older working adults. Significant program effects were not shown on measures of stress or aging beliefs, and there were too few smokers in the sample for meaningful analysis. Analysis of the nonimputed data indicated that program effects were stronger for women than men. The findings suggest that a multimedia Web-based program could be a promising vehicle for delivering health promotion material to older working adults.

Acknowledgments

We would like to thank Dr Susan Hughes, University of Illinois, for her assistance throughout the project. We would also like to thank Dr Xin He for his helpful assistance with the data analysis. This research was funded by the National Institute on Aging, Grant Number R44AG033964.

Conflicts of Interest

The HealthyPast50 program tested in this study is owned by ISA Associates, Inc, a company of which RC and RH are owners.

Multimedia Appendix 1

Trial protocol.

[PDF File (Adobe PDF File), 95KB - jmir_v17i3e82_app1.pdf]

Multimedia Appendix 2

Informed consent document.

[PDF File (Adobe PDF File), 40KB - jmir_v17i3e82_app2.pdf]

Multimedia Appendix 3

Selected screenshots from HealthyPast50.

[PDF File (Adobe PDF File), 1MB - jmir_v17i3e82_app3.pdf]

Multimedia Appendix 4

CONSORT-EHEALTH checklist V1.6.2 [39].

[PDF File (Adobe PDF File), 77KB - jmir_v17i3e82_app4.pdf]

References

1. Breslow L, Breslow N. Health practices and disability: some evidence from Alameda County. Prev Med 1993 Jan;22(1):86-95. [doi: 10.1006/pmed.1993.1006] [Medline: 8475014]
2. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. JAMA 2004 Mar 10;291(10):1238-1245. [doi: 10.1001/jama.291.10.1238] [Medline: 15010446]
3. Wegman DH, McGee JP, editors. Health and Safety Needs of Older Workers. Washington, DC: The National Academies Press; 2004.
4. Fries JF. Aging, natural death, and the compression of morbidity. N Engl J Med 1980 Jul 17;303(3):130-135. [doi: 10.1056/NEJM198007173030304] [Medline: 7383070]
5. Fries JF, Koop CE, Sokolov J, Beadle CE, Wright D. Beyond health promotion: reducing need and demand for medical care. Health Aff (Millwood) 1998;17(2):70-84 [FREE Full text] [Medline: 9558786]
6. DeVol R, Bedrousian A. An Unhealthy American: The Economic Burden of Chronic Disease. Santa Monica, CA: Milken Institute; 2007.
7. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: update VI 2000-2004. J Occup Environ Med 2005 Oct;47(10):1051-1058. [Medline: 16217246]
8. Goetzel RZ, Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. Annu Rev Public Health 2008;29:303-323. [doi: 10.1146/annurev.publhealth.29.020907.090930] [Medline: 18173386]
9. Ozminkowski RJ, Goetzel RZ, Smith MW, Cantor RI, Shaughnessy A, Harrison M. The impact of the Citibank, NA, health management program on changes in employee health risks over time. J Occup Environ Med 2000 May;42(5):502-511. [Medline: 10824303]
10. Hughes SL, Seymour RB, Campbell RT, Whitelaw N, Bazzarre T. Best-practice physical activity programs for older adults: findings from the national impact study. Am J Public Health 2009 Feb;99(2):362-368. [doi: 10.2105/AJPH.2007.131466] [Medline: 19059858]
11. Hughes SL, Seymour RB, Campbell RT, Shaw JW, Fabiyi C, Sokas R. Comparison of two health-promotion programs for older workers. Am J Public Health 2011 May;101(5):883-890 [FREE Full text] [doi: 10.2105/AJPH.2010.300082] [Medline: 21421955]
12. Phelan EA, Williams B, Snyder SJ, Fitts SS, LoGerfo JP. A five state dissemination of a community-based disability prevention program for older adults. Clin Interv Aging 2006;1(3):267-274 [FREE Full text] [Medline: 18046880]
13. Leveille SG, Wagner EH, Davis C, Grothaus L, Wallace J, LoGerfo M, et al. Preventing disability and managing chronic illness in frail older adults: a randomized trial of a community-based partnership with primary care. J Am Geriat Soc 1998;46:1-8. [Medline: 9777899]
14. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
15. Winett RA, Anderson ES, Wojcik JR, Winett SG, Bowden T. Guide to health: nutrition and physical activity outcomes of a group-randomized trial of an Internet-based intervention in churches. Ann Behav Med 2007 Jun;33(3):251-261. [doi: 10.1080/08836610701358045] [Medline: 17600452]
16. Carr LJ, Dunsiger SI, Lewis B, Ciccolo JT, Hartman S, Bock B, et al. Randomized controlled trial testing an internet physical activity intervention for sedentary adults. Health Psychol 2013 Mar;32(3):328-336. [doi: 10.1037/a0028962] [Medline: 22823069]
17. Deitz DK, Cook RF, Hendrickson A. Preventing prescription drug misuse: field test of the SmartRx Web program. Subst Use Misuse 2011;46(5):678-686. [doi: 10.3109/10826084.2010.528124] [Medline: 21043788]
18. Anderson-Bill ES, Winett RA, Wojcik JR, Winett SG. Web-based guide to health: relationship of theoretical variables to change in physical activity, nutrition and weight at 16-months. J Med Internet Res 2011;13(1):e27 [FREE Full text] [doi: 10.2196/jmir.1614] [Medline: 21447470]
19. Billings DW, Cook RF, Hendrickson A, Dove DA. A web-based approach to managing stress and mood disorders in the workforce. J Occup Environ Med 2008 Aug;50(8):960-968. [doi: 10.1097/JOM.0b013e31816c435b] [Medline: 18695455]

20. Marcus BH, Ciccolo JT, Sciamanna CN. Using electronic/computer interventions to promote physical activity. Br J Sports Med 2009 Feb;43(2):102-105. [doi: 10.1136/bjsm.2008.053744] [Medline: 19052143]
21. Marcus BH, Lewis BA, Williams DM, Dunsiger S, Jakicic JM, Whiteley JA, et al. A comparison of Internet and print-based physical activity interventions. Arch Intern Med 2007 May 14;167(9):944-949. [doi: 10.1001/archinte.167.9.944] [Medline: 17502536]
22. Cook RF, Billings D, Hersch RK, Back AS, Hendrickson A. A field test of a web-based workplace health promotion program to improve dietary practices, reduce stress, and increase physical activity: randomized controlled trial. J Med Internet Res 2007;9(2):e17 [FREE Full text] [doi: 10.2196/jmir.9.2.e17] [Medline: 17581811]
23. Deitz D, ISA Associates. A field test of a web-based cardiac risk reduction program: Final report [submitted]. National Heart, Lung and Blood Institute 2010:1-1 (forthcoming).
24. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
25. Bandura A. Self Efficacy: The Exercise of Control. New York: WH Freeman and Company; 1997.
26. Orioli EM, Jaffe DT, Scott CD. StressMap: Personal Diary Edition. The Ultimate Stress Management, Self-Assessment, and Coping Guide. New York: Newmarket Press; 1991.
27. Cook RF, Back AS, Trudeau JV, McPherson TL. Integrating substance abuse prevention into health promotion programs in the workplace: A social cognitive intervention targeting the mainstream user. In: Bennett JB, Lehman WEK, editors. Preventing Workplace Substance Abuse: Beyond Drug Testing to Wellness. Washington, DC: American Psychological Association; 2003:97-133.
28. Trenkner LL, Rooney B, Viswanath K, Baxter J, Elmer P, Finnegan JR, et al. Development of a scale using nutrition attitudes for audience segmentation. Health Educ Res 1990;59(4):479-487.
29. O'Neil PM, Rhodes SK. Module 2: Assessment, screening and goal-setting. In: Wolfe BL, editor. The Lifestyle Counselor's Guide for Weight Control. Dallas, TX: American Health Publishing; 1996.
30. McCann BS, Bovbjerg VE, Brief DJ, Turner C, Follette WC, Fitzpatrick V, et al. Relationship of self-effecacy to cholesterol lowering and dietary change in hyperlipidemia. Ann Behav Med 1995 Sep;17(3):221-226. [doi: 10.1007/BF02903916] [Medline: 24203533]
31. Glynn SM, Ruderman AJ. The development and validation of an eating self-efficacy scale. Cognit Ther Res 1986;10(4):403-420.
32. Godin G, Shephard RJ. Godin leisure-time exercise questionnaire. Med Sci Sports Exerc 1997;29(June Supplement):S36-S38.
33. Kroll T, Kehn M, Ho PS, Groah S. The SCI Exercise Self-Efficacy Scale (ESES): development and psychometric properties. Int J Behav Nutr Phys Act 2007;4:34 [FREE Full text] [doi: 10.1186/1479-5868-4-34] [Medline: 17760999]
34. McAuley E. The role of efficacy cognitions in the prediction of exercise behavior in middle-aged adults. J Behav Med 1992 Feb;15(1):65-88. [Medline: 1583674]
35. Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials. New York: Springer; 2010.
36. Schafer JL. Analysis of Incomplete Multivariate Data. New York: Chapman and Hall; 1997.
37. Hankonen N, Absetz P, Ghisletta P, Renner B, Uutela A. Gender differences in social cognitive determinants of exercise adoption. Psychol Health 2010 Jan;25(1):55-69. [doi: 10.1080/08870440902736972] [Medline: 20391207]
38. Deitz D, Cook RF, Hersch RK, Leaf S. Heart healthy online: an innovative approach to risk reduction in the workplace. J Occup Environ Med 2014 May;56(5):547-553. [doi: 10.1097/JOM.0000000000000148] [Medline: 24806568]
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;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

BMI: body mass index
GTH: Guide to Health
MCMC: Markov chain Monte Carlo
RCT: randomized controlled trial
SCT: social cognitive theory

Edited by G Eysenbach; submitted 14.03.14; peer-reviewed by E Anderson-Bill, D Bowen; comments to author 23.09.14; revised version received 05.01.15; accepted 03.02.15; published 25.03.15.

Please cite as:
Cook RF, Hersch RK, Schlossberg D, Leaf SL
A Web-Based Health Promotion Program for Older Workers: Randomized Controlled Trial
J Med Internet Res 2015;17(3):e82
URL: http://www.jmir.org/2015/3/e82/
doi:10.2196/jmir.3399
PMID:25830503

Review

The Potential of Technology-Based Psychological Interventions for Anorexia and Bulimia Nervosa: A Systematic Review and Recommendations for Future Research

Sandra Schlegl[1], Dipl Psych, MSc, PhD; Carolina Bürger[1], MSc; Luise Schmidt[1], Dipl Psych; Nirmal Herbst[2], Dipl Psych; Ulrich Voderholzer[2,3], MD

[1]Department of Psychiatry and Psychotherapy, University of Munich, München, Germany
[2]Department of Psychiatry and Psychotherapy, Freiburg, Germany
[3]Schön Klinik Roseneck, Prien am Chiemsee, Germany

Corresponding Author:
Sandra Schlegl, Dipl Psych, MSc, PhD
Department of Psychiatry and Psychotherapy
University of Munich
Nussbaumstraße 7
München, 80336
Germany
Phone: 49 89440053369
Fax: 49 89440054572
Email: sandra.schlegl@med.uni-muenchen.de

Abstract

Background: Previous studies have shown an unmet need in the treatment of eating disorders. In the last decade, interest in technology-based interventions (TBIs) (including computer- and Internet-based interventions [CBIs] or mobile interventions) for providing evidence-based therapies to individuals with different mental disorders has increased.

Objective: The aim of this review was to systematically evaluate the potential of TBIs in the field of eating disorders, namely for anorexia nervosa (AN) and bulimia nervosa (BN), for both prevention and treatment, and also for carers of eating disorder patients.

Methods: A systematic literature search was conducted using Medline and PsycINFO. Bibliographies of retrieved articles were also reviewed without date or study type restrictions.

Results: Forty studies resulting in 45 publications reporting outcomes fulfilled the inclusion criteria: 22 randomized controlled trials, 2 controlled studies, and 16 uncontrolled studies. In total, 3646 patients were included. Overall, the studies provided evidence for the efficacy of guided CBIs, especially for BN patients and for compliant patients. Furthermore, videoconferencing also appeared to be a promising approach. Evaluation results of Internet-based prevention of eating disorders and Internet-based programs for carers of eating disorder patients were also encouraging. Finally, there was preliminary evidence for the efficacy of mobile interventions.

Conclusions: TBIs may be an additional way of delivering evidence-based treatments to eating disorder patients and their use is likely to increase in the near future. TBIs may also be considered for the prevention of eating disorders and to support carers of eating disorder patients. Areas of future research and important issues such as guidance, therapeutic alliance, and dissemination are discussed.

(J Med Internet Res 2015;17(3):e85) doi:10.2196/jmir.3554

KEYWORDS

anorexia nervosa; bulimia nervosa; computers; Internet; mobile phone; cognitive behavioral therapy

Introduction

Up to 4% of women have an eating disorder. Among these, the prevalence is 0.3% for anorexia nervosa (AN), 1% for bulimia nervosa (BN) [1], and 2.4% for eating disorders not otherwise specified (EDNOS) [2]. Evidence-based psychological treatments exist, especially for BN, and include cognitive behavioral therapy (CBT) and interpersonal psychotherapy (IPT) [3,4]. For adolescent AN patients, the efficacy of family-based therapy is apparent [3,4]. For adult AN patients, a combination of renourishment and psychotherapy (eg, CBT, supportive clinical management, IPT) is recommended [5]. There is still a lack of evidence-based treatments for adults with AN due to several methodological difficulties. These include difficulties in recruiting participants because this disorder is relatively rare, lack of treatment acceptance, and high dropout rates [6,7].

However, previous international studies on health service utilization showed an unmet need for the treatment of patients with eating disorders [8,9]. In their meta-analysis, Hart et al [10] found that only one-quarter of sufferers sought eating disorder-specific treatment. Barriers discussed were shame or fear of stigmatization when going to a psychologist. On the other hand, the number of patients seeking treatment in primary care has increased, especially for BN [11]. However, patients often undergo psychological treatment only after a considerable delay, and waiting times for treatment are long [12,13]. Furthermore, there is a general shortage of specialized therapists and institutions, and evidence-based methods are still lacking in routine care.

Technology-based interventions (TBIs), including computer- and Internet-based interventions (CBIs) or mobile interventions, have the potential to reach patients who otherwise may not access help, and to improve health care for those seeking treatment, by offering immediate access to evidence-based interventions. Communication in TBIs can take place synchronously (in real time, such as videoconference, chat rooms) or asynchronously (with lag between contacts, such as email, postings on a secure website, text messaging). Information exchange can occur in writing, just via audio communication, or via webcam. TBIs can be administered individually or in the form of group sessions.

There are various forms of CBIs that differ especially in their amount of therapist guidance:

1. Computer- and Internet-based unguided self-help (unguided CBI): this is the generic term for self-help interventions primarily delivered via computer technology. Available programs mainly are multimedia-based as well as CBT-based. Patients can either use them at home or in health care settings, and programs are designed to enable patients to work through them independent of a therapist.
2. Computer- and Internet-based guided self-help (guided CBI): support can range from screening for suitability, offering technical advice, monitoring progress and outcome, as well as giving emotional support [14].
3. Internet-based therapist-delivered treatments: these can be delivered using different methods, such as email, chat rooms, and videoconferencing, either solely Internet-delivered or in combination with face-to-face treatment. Patients have regular contact with a therapist.

In the last decade, a large body of empirical evidence on the acceptance and efficacy of CBIs for mental disorders has accrued. Several reviews and meta-analyses have shown that new CBI treatments hold great promise in the treatment of adults with depression [15-17], depression and anxiety [18,19], anxiety [20,21], obsessive-compulsive disorder [22,23], and traumatic stress [24,25]. CBIs have been found to be as effective as face-to-face treatment [18,26]. However, guided interventions seem to result in better outcomes [16,27]. Preliminary evidence also suggests that CBIs may be acceptable and effective interventions for adolescents with mental disorders [28,29].

Researchers in the field of eating disorders thus hypothesized that CBIs may also be a suitable approach for eating disorder patients because they mostly come from an age group that uses the computer and Internet frequently. This led them to examine the efficacy of CBIs for eating disorders. There are some reviews of CBIs in the field of eating disorders, but most of them do not meet the requirements of a systematic review [30-33]. To date, Myers et al [34] and Marks et al [35] have presented comprehensive reviews on the use of new technologies in the treatment of a broad range of eating disorders, but only 5 and 8 studies covered CBIs for AN and BN, respectively. The authors concluded that CBIs are a promising approach for the treatment of eating disorders that should be further explored. Dölemeyer et al [36] presented a meta-analysis including 8 controlled studies evaluating CBIs for eating disorders. Guided programs seemed promising as medium to large effect sizes were found. Bauer and Moessner [37] recently provided a review of randomized controlled trials (RCTs) of CBIs for eating disorders and very useful recommendations for future research.

Nevertheless, these reviews do not give a comprehensive overview of the current state of research in the field of TBIs in eating disorders due to their strict inclusion criteria. Our review addresses the fast changes in new technologies and the growing knowledge and increasing research on TBIs for eating disorders. The aims of this review are to provide a comprehensive and up-to-date picture on the developing field of TBIs for eating disorders, specifically for AN and BN, by including a wider range of study designs beyond just RCTs and to discuss the past findings with respect to both acceptance and efficacy of TBIs for AN and BN. We primarily focus on treatment studies, but also consider studies on prevention, motivation, and on programs for carers of eating disorder patients.

Methods

We searched Medline and PsycINFO for eligible studies published in English, German, Spanish, Italian, French, or Portuguese up to August 2014. The following search terms were used: “(online* OR internet* OR e-mail* OR email* OR web* OR media* OR computer* OR remote* OR tele* OR virtual* OR “interactive voice response*” OR www OR cd* OR dvd* OR flopp* OR audio* OR video* OR palmtop* OR e-health* OR technolog* OR chat* OR software* OR text-messag* OR “text messag*” OR “internet telephony” OR mobile* OR sms*).” Search terms were combined with (anore* OR bulimi*

OR “eating disorder*” OR “disordered eating” OR body*image*) in title. The bibliographies of the retrieved articles were also reviewed.

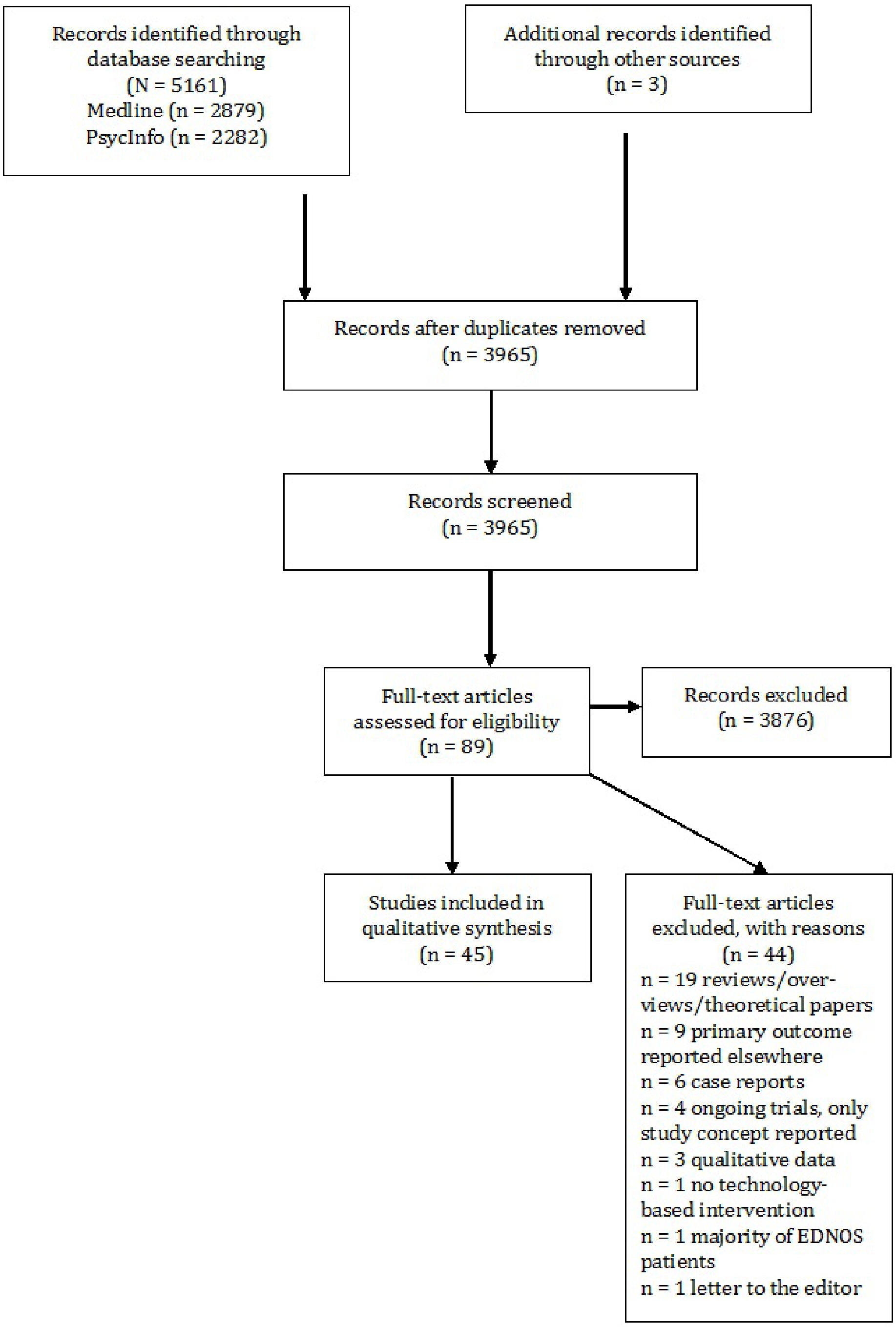
We included studies that met the following inclusion criteria: (1) technology-based psychological interventions, (2) samples including patients with AN and/or BN according to *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) criteria (*DSM-III* and *DSM-IV*), persons with body image concerns and disordered eating, subthreshold eating disorders, or carers of patients with eating disorders, (3) outcome data at postintervention and/or follow-up, and (4) sample size of at least N=10 per study. There were no age or study type restrictions.

We excluded psychoeducational or counseling interventions, online support groups, as well as computer-based assessment methods. Studies that evaluated virtual reality in the treatment of body image were excluded because a review covering this issue has already been published [38]. Furthermore, studies evaluating the Internet-based prevention program StudentBodies for eating disorders that generally met our inclusion criteria were excluded because a meta-analysis was recently published on this program [39,40]. Unpublished studies, abstracts of conference proceedings, dissertations, and letters were also not considered. No attempt was made to include unpublished data. Furthermore, inclusion and exclusion criteria were not documented in any official review protocol. Figure 1 shows the flowchart of the literature search.

We separately evaluated CBIs and mobile interventions and divided studies on CBIs into the following main categories: treatment of AN and BN, relapse prevention, prevention and early intervention, and interventions for carers. Treatment studies for AN and BN were further classified into 3 sections depending on the amount of guidance patients received. Studies in which patients were offered no contact with a coach or a clinician were classified as “unguided CBIs.” Studies in which patients worked through programs delivered on a computer or via the Internet and were guided by email, phone, or face-to-face contact with a professional were classified as “guided CBIs.” Finally, interventions that were completely delivered by a therapist (email therapy, videoconferencing) were classified as “therapist-delivered treatments.”

For studies for which relevant data were available, effect sizes (standardized mean differences) were calculated using pooled standard deviations [41]. Within-group effect sizes were calculated between baseline and the following assessment time points (posttreatment, follow-up 1, follow-up 2). Between-group effect sizes at posttreatment and/or follow-up 1 and/or follow-up 2 were calculated. Furthermore, lower and upper confidence intervals for effect sizes were determined [41].

Figure 1. PRISMA flow chart of the literature search.



Results

A total of 40 studies met the inclusion criteria with outcomes reported in 45 publications. There were 41 publications on the results of the efficacy of CBIs (n=21 CBIs for the treatment of eating disorder patients; n=12 programs for prevention and early intervention of eating disorders; n=3 relapse prevention interventions; n=5 interventions for carers of eating disorder patients). A total of 4 publications reported on the efficacy of mobile interventions. Overall, 5 studies focused on AN and 5 studies on adolescents. 22 of the included studies were RCTs, 2 were controlled studies, and the remaining 16 studies were uncontrolled. The studies included a total number of 3646 patients.

Table 1 provides a short overview of included studies. The table in Multimedia Appendix 1 gives detailed information on study characteristics, outcomes, and weaknesses of studies.

In the following, results in effect sizes are given for all categories. Effect sizes and confidence intervals for all individual studies are presented in Multimedia Appendix 2.

Table 1. Studies evaluating efficacy of computer- and Internet-based treatment for eating disorders.[a]

Study	Diagnosis (N)	Intervention (study design)	Duration
Computer- and Internet-based unguided self-help			
Bara-Carril et al [42]	BN (N=46); EDNOS (N=11)	"Overcoming Bulimia" (uncontrolled study)	8 modules, 4-8 weeks
Schmidt et al [43]	BN (N=60); EDNOS (N=37)	"Overcoming Bulimia" (RCT)	8 modules, 8-12 weeks
Johnston et al [44]	BN (N=94)	Therapeutic writing (RCT)	20 min on 3 consecutive days
Computer- and Internet-based guided self-help			
Huon [45]	BN (N=120)	Internet-guided self-help (RCT)	7 modules, 7 months
Graham & Walton [46]	BN (N=13); BED (N=27)	"Overcoming Bulimia" (uncontrolled study)	8 modules, 8 weeks
Murray et al [47]	BN (N=77); EDNOS (N=5)	"Overcoming Bulimia" (controlled study)	8 modules, 8-12 weeks
Sánchez-Ortiz et al [48]	BN (N=39); EDNOS (N=37)	"Overcoming Bulimia" (RCT)	8 modules, 8-12 weeks; continued access to the online sessions for 24 weeks
Ljotsson et al [49]	BN (N=33); BED (N=36)	"Overcoming Binge eating" (RCT)	6 modules, 12 weeks
Carrard et al [50]	BN (N=41); EDNOS (N=4)	"SALUT" (uncontrolled study)	7 modules, 4 months
Liwowsky et al [51]	BN (N=22)	"SALUT" (uncontrolled study)	7 modules, 4 months
Nevonen et al [52]	BN (N=27); EDNOS (N=11)	"SALUT" (uncontrolled study)	7 modules, 6 months
Fernández-Aranda et al [53]	BN (N=62)	"SALUT" (controlled study)	7 modules, 4 months
Carrard et al [54]	BN (N=100); EDNOS (N=27)	"SALUT" (uncontrolled study)	7 modules, 4 months
Wagner et al [55]	BN/EDNOS (N=155)	"SALUT" (RCT)	7 modules, 4-7 months
Leung et al [56]	ED (N=280)	"SMART EATING" (uncontrolled study)	6 components, open-end
Pretorius et al [57]	BN (N=61); EDNOS (N=40)	"Overcoming Bulimia" (adapted for adolescents) (uncontrolled study)	8 modules, 3 months
Wagner et al [58]	BN (N=126)	"SALUT" (RCT)	7 modules, 4-7 months
Internet-based therapist-delivered treatment			
Robinson & Serfaty [59]	BN (N=18); BED (N=4); EDNOS (N=1)	Email therapy (uncontrolled study)	3 months
Robinson & Serfaty [60]	BN (N=51); EDNOS (N=20); BED (N=26)	Email therapy (RCT)	3 months
Simpson et al [61]	BN (N=5); AN (N=1); EDNOS (N=6)	Videoconferencing (uncontrolled study)	12-20 sessions CBT; 6-8 optional sessions of nutritional education
Mitchell et al [62]	BN (N=71); EDNOS (N=57)	Videoconferencing (RCT)	20 sessions manual-based CBT, 4 months
Internet-based relapse prevention			
Fichter et al [63,64]	AN (N=258)	"VIA" (RCT)	9 modules, 9 months
Mezei et al [65]	BN/EDNOS/BED/ AN binge-purging subtype (N=39)	"EDINA" (uncontrolled study)	Online platform for peer support and professional consultation, 4 months
Body image and eating disorder prevention			
Gollings & Paxton [66]	Body dissatisfaction and disordered eating (N=40)	"Set Your Body Free" (pilot RCT)	Weekly 90-min group sessions, 8 weeks
Paxton et al [67]	Body dissatisfaction (N=116)	"Set Your Body Free" (RCT)	Weekly 90-min group sessions, 8 weeks

Study	Diagnosis (N)	Intervention (study design)	Duration
Stice et al [68,69]	Body dissatisfaction (N=107)	“eBody Project” (RCT)	6 modules, 3 weeks
Serdar et al [70]	Weight and/or shape concerns (N=333)	Dissonance-based eating disorder prevention (RCT)	3 sessions for 60 min
Zabinski et al [71]	Weight concerns (N=60)	Synchronous Internet relay chat (RCT)	Weekly (60 min) chat discussions, 8 weeks
Ohlmer et al [72]	Women at risk for AN (N=36)	“Student Bodies” for AN (uncontrolled study)	Weekly sessions for 45-90 min, 10 weeks
Heinicke et al [73]	Body image or eating problems (N=73)	“My Body, My Life” (RCT)	Weekly 90-min online sessions, 6 weeks
Eating disorder symptoms/subthreshold eating disorders			
Ruwaard et al [74]	Bulimic symptoms (N=105)	Online cognitive behavioral treatment (RCT)	20 weeks
Jacobi et al [75]	Subthreshold eating disorder (N=29); Eating disorder symptoms at lower level (N=97)	“Student Bodies+” (RCT)	8 sessions, 8 weeks
Motivation			
Hötzel et al [76]	Eating disorder symptoms (N=212)	“ESS-KIMO“ (RCT)	Weekly online sessions for 45 min, 6 weeks
Leung et al [77]	Eating disorder (N=185)	“SMART EATING” (uncontrolled study)	11 worksheets to enhance individuals’ motivation to change their eating behaviors
Carers/Parents			
Binford Hopf et al [78]	Parents of AN patients (N=13)	Internet-based chat support groups (uncontrolled study)	15 weekly online chat sessions for 90 min
Grover et al [79]	Carers of AN patients (N=27)	“Overcoming anorexia online” (uncontrolled study)	9 workbooks (encouraged 1 per week, but no formal time limit)
Grover et al [80]	Carers of AN patients (N=64)	“Overcoming anorexia online” (RCT)	8 modules, 4 months
Hoyle et al [81]	Carers of AN patients (N=37)	“Overcoming anorexia online” (RCT)	7 modules + 2 additional modules for carers, 7 weeks
Bruning Brown et al [82]	Parents of sophomore students (N=69)	“Student Bodies” parent intervention (RCT)	Unstructured web-based intervention, 4 weeks
Mobile/SMS text messaging			
Shapiro et al [83]	BN (N=31)	Text messaging (uncontrolled study)	Daily, 24 weeks + 12 CBT face-to-face group sessions for 90 min
Robinson et al [84]	BN (N=21)	SMS-based intervention (uncontrolled study)	Weekly, 6 months
Bauer et al [85]	BN (N=97); EDNOS (N=68)	Aftercare SMS-based intervention (RCT)	Weekly symptom report via SMS text message, 16 weeks
Cardi et al [86]	AN (N=18); BN (N=13)	Mp4 player or iPod with 10 video clips (vodcasts) (uncontrolled study)	Between 3 to 20 min each, workbook and daily monitoring forms, 3 weeks

[a] AN: anorexia nervosa, BN: bulimia nervosa, BED: binge eating disorder, CBT: cognitive behavioral therapy, ED: eating disorder, EDNOS: eating disorders not otherwise specified, RCT: randomized controlled trial, SMS: short message service.

Efficacy of Computer- and Internet-Based Interventions for Anorexia Nervosa and Bulimia Nervosa

Efficacy of Computer- and Internet-Based Unguided Self-Help

A total of 3 studies with 248 participants including 94 controls examined unguided CBIs for eating disorders. Unguided CBIs showed no effects [44] using the total score of the Bulimic Investigatory Test-Edinburgh (BITE) [87] or small effects [43] using the global score of the Eating Disorder Examination (EDE) [88] from pre- to posttreatment. Furthermore, from preintervention to follow-up, small (EDE global, BITE total) [43,44] to medium effects (binging, vomiting) [42] were found. There was no difference between control groups and unguided CBIs for EDE global or BITE total at posttreatment [43,44], and no (BITE total) [44] to small (EDE global) [43] between-group effects at follow-up.

Efficacy of Computer- and Internet-Based Guided Self-Help

Adults

Five different treatment approaches of guided CBIs for adults were evaluated in 12 publications. A total of 1051 participants were included. Overall, 135 participants were controls without intervention or on the waiting list; 56 were active controls. From pre- to posttreatment, small to large effects were found for binging and vomiting (in 1 study there was no effect for vomiting) and medium to large effects for eating disorder psychopathology (total score of the Eating Disorder Inventory-2 [EDI-2] [89], EDE global) [48,49,54-56]. From baseline to follow-up treatment, small to large effects were reported for binging, small to medium effects for vomiting, and medium to large effects for eating disorder psychopathology (EDI-2 total, EDE global) [48,52,55]. No differences were observed between guided bibliotherapy and guided CBI both at postintervention and at follow-up (binging, vomiting, EDI-2 total) [55]. Compared to a waiting list control, there were small to medium effects for binging, small to large effects for vomiting, and large effects for EDE global at postintervention [48,49]. At follow-up, there was a medium effect for binging, a small effect for vomiting, and a large effect on EDE global [48].

Adolescents

There was 1 study evaluating the potential of CBIs for adolescents and 1 additional study that provided data in the context of a subgroup analysis comparing adult and adolescent patients receiving guided CBI. A total of 130 adolescents were included. Adolescents were compared to 97 adults in 1 study. Small to large effects for binging and medium to large effects for vomiting at both postintervention and follow-up were found [57,58]. Furthermore, medium (EDI-2 total) to large (EDE global) effects were reported for eating disorder psychopathology [57,58] at both time points. In comparison to adults, small effects for binging and vomiting, but no effect for EDI-2 total was shown at postintervention. At follow-up, 1 study showed a medium effect for binging as well as small effects for vomiting and for EDI-2 total [58].

Efficacy of Internet-Based Therapist-Delivered Treatments

Four studies examined the efficacy of Internet-based therapist-delivered treatments: 2 investigated email therapy and 2 evaluated videoconferencing. In all, 260 patients participated of whom 27 were waiting list patients and 100 active controls. Internet-based therapist-delivered treatments resulted in medium (BITE severity and symptoms, binging) [59, 62] to large effects (vomiting) [62] from pre- to posttreatment. From preintervention to follow-up, medium (binging) to large (vomiting) effects were found [62]. In comparison to a face-to-face treatment, the effects for binging and vomiting at posttreatment were small and zero at follow-up [62].

Efficacy of Internet-Based Relapse Prevention

There were 2 studies investigating the efficacy of 2 different Internet-based relapse prevention programs: 1 uncontrolled study and 1 RCT. A total of 297 patients were included, of whom 130 were treatment as usual (TAU) patients. Overall, small short-term effects were found for the global of the Eating Disorder Examination Questionnaire (EDE-Q) [90] and EDI-2 total [63,65]. No effects were shown for binging, vomiting, and body mass index (BMI) [63,65]. At follow-up, there was a small effect for BMI for completers. Compared to TAU, small effects were found for BMI (completers) and EDI-2 total at postintervention.

Efficacy of Internet-Based Treatments for Prevention and Early Intervention

Efficacy of Internet-Based Treatments for Body Image Problems and Prevention of Eating Disorders

Adults

Seven studies investigated the efficacy of CBIs in the prevention of eating disorders. A total of 775 participants were included: 224 were controls without intervention, 207 active controls (face-to-face), and 49 video or brochure controls. Primary outcomes studied were mainly body dissatisfaction and thin-ideal internalization. From pre- to postintervention, there were small to medium effects for the Ideal-Body Stereotype Scale-Revised (IBSS-R [91]) [70], as well as medium to large effects for the Body Shape Questionnaire (BSQ [92]) [66,67] and for EDE-Q global [71,72]. From baseline to follow-up, there was a small effect for IBSS-R [69], medium effects for EDE-Q global [71,72], and a large effect for BSQ [66]. At posttreatment, no (IBSS-R) [68,70] to small effects (BSQ) [66,67] of prevention programs were seen in comparison to face-to-face conditions, small (EDE-Q global, IBSS-R) [70,71] to medium effects (BSQ) [67] in comparison to controls without intervention, and medium to large effects (IBSS-R) [68] in comparison to a video and a brochure control condition, respectively [68]. At follow-up, no (BSQ) [66] to small effects (IBSS-R) [69] were observed in comparison to face-to-face conditions, a small effect in comparison to a control condition (EDE-Q global) [71], and no effects in comparison to a video as well as a brochure control condition (IBSS-R) [69].

Adolescents

Heinicke et al [73] reported a medium effect for an adolescent sample from pre- to posttreatment in BSQ that was maintained at follow-up. They also found a medium between-group effect in comparison to a delayed treatment group.

Efficacy of Internet-Based Treatments for Subthreshold Eating Disorders

We identified 2 studies investigating Internet-based treatments for subthreshold eating disorders. A total of 231 participants were randomized, of whom 35 were active controls and 97 were on a waiting list. From pre- to postintervention, a medium effect for binging [75], a small effect for vomiting [75], and medium to large effects for EDE-Q global [74,75] were observed. From baseline to follow-up, large effects for binging [75] and EDE-Q global [74,75], but no effect for vomiting [75] was identified. At postintervention, the Internet intervention differed only slightly from bibliotherapy (EDE-Q global) [74]. In comparison to a waiting list, there were small effects for binging and vomiting [75], and no effect in 1 study, but a medium effect for EDE-Q global in another study [74,75]. At follow-up, no differences were found in comparison to bibliotherapy (EDE-Q global) [74]. Compared to the waiting list, there were small effects for binging and vomiting and a medium effect for EDE-Q global [75].

Efficacy of Internet-Based Treatments to Enhance Motivation to Change in Eating Disorders

Internet-based treatments that aimed at enhancing motivation to change in eating disorders were the subject of 2 studies. A total of 397 participants were included, of whom 109 were waiting list controls. There were small time effects for binging [77] and motivation to gain weight in problematic areas [76], but no time effects for vomiting [76,77] or for motivation to give up binging and vomiting [77] from baseline to postintervention. At follow-up, there was a small effect for the motivation to give up binging, but no effects for binging, vomiting, or motivation for giving up vomiting [77]. A small between-group effect in vomiting was reported, but no between-group effect for motivation to gain weight in problematic areas in comparison to a control group at posttreatment.

Efficacy of Internet-Based Treatments for Carers of People with Eating Disorders

There were 5 intervention studies including 210 carers of people with eating disorders. A control intervention called Beating Eating Disorders (BEAT; patient and carer organization) was offered to 30 carers, whereas 47 received no intervention. From baseline to postintervention and to follow-up, no effect was found for the Level of Expressed Emotion Scale (LEE [93]) (unguided overcoming anorexia online [OAO]), a small effect was found for LEE (guided OAO) [81], and a medium effect was found for the total score of the Hospital Anxiety and Depression Scale (HADS [94]) (guided OAO) [79]. Furthermore, medium effects for the scales “critical to others” and “healthy outlook” of the Parental Attitudes and Criticism Scale (PACS [82]) were seen at postintervention, but only small effects at follow-up [82]. Finally, a large effect for the total score of the Eating Disorder Symptom Impact Scale (EDSIS [95]) was evident at postintervention [78]. In comparison to a control group, there were small effects for “critical to others” and “healthy outlook” at postintervention [82]. Between unguided and guided OAO, a small effect was shown at postintervention and a medium effect at follow-up [81].

Efficacy of Mobile Therapy in the Treatment of Eating Disorders

Four studies evaluated mobile interventions for eating disorders. Three studies (total N=230; n=83 control patients) employed short message service (SMS) text messaging and 1 study used vodcasts (N=31). For SMS text messaging interventions, the baseline versus follow-up effects were small [84] to large [83] for binging, zero [84] to medium [83] for vomiting, and large for EDI-2 total [83]. Until follow-up, a large effect for binging, a medium effect for vomiting, and a large effect for EDI-2 total [83] were reported. For the vodcast intervention, no effect was found for binging and small effects for vomiting and EDE-Q total [86].

Abstinence Rates for Technology-Based Interventions

Abstinence rates ranged between 12% [43] and 46.2% [42] for unguided CBIs (waiting list: 10% to 20% [43]), between 16.6% [45] and 46.6% [45] for guided CBIs (waiting list: 0% [45] to 20.7% [48]), and between 21% [62] and 66.7% [61] for therapist-delivered treatments (waiting list: 0% [60]; face-to-face: 25.8% to 53.3% [62]). In studies evaluating CBIs for subthreshold eating disorders, abstinence rates were between 37% [74] and 45.1% [75] (waiting list: 8% [74] to 26.9% [75]; bibliotherapy: 8% to 34% [74]). Mobile interventions showed abstinence rates between 29.4% [84] and 37.8% [85] (control: 18.1% [85]). Abstinence rates for individual studies are shown in Multimedia Appendix 3.

In the following, results of further relevant aspects of TBIs are presented. Note that not all studies mentioned so far provided data regarding these points.

Intervention Uptake

Twenty studies provided information about non-take-up rates. These ranged from 2.9% [80] (carers) to 50% [82] (carers) with an average of 20.1%. However, reasons for non-take-up were rarely reported. Paxton et al [67] and Hoyle et al [81] indicated that persons did not initiate treatment due to technical difficulties or difficulties accessing a computer or the website, due to seeking alternative treatment, or due to not having enough time to take part in the intervention.

Murray et al [47] found no differences between those who started the program and those who did not, except for the expected personal usefulness of this kind of intervention. In the study by Leung et al [56], patients who took up the intervention had a lower BMI and were generally more educated. There were no differences in age or severity of eating disorder. Leung et al [77] reported that participants working through the motivational enhancement exercises had a slightly higher stage of change than those who did not take part in these.

Acceptance

Twenty-six studies provided data about compliance rates (we defined compliance as the full completion of the intervention regardless of whether participants completed posttreatment and/or follow-up assessment, and we only extracted data from studies that explicitly reported information on this). Between 18.4% [52] and 95.5% [51] completed the program (mean compliance rate in the TBI conditions: 57.6% of patients that were included for uncontrolled studies or RCTs).

Pretorius et al [57] found that higher baseline scores on eating concerns were associated with higher compliance in terms of number of completed sessions. Fichter et al [63] reported that patients with less pronounced compensatory behaviors and a lifetime mood disorder showed better adherence. The results by Leung et al [56] suggested that the motivational stage of change in patients and their treatment expectations predict compliance particularly well. These authors also named reasons for discontinuing: motivational difficulties, not enough energy and time for the program, loss of interest, or nonresponse.

Most of the studies found no differences between completers and noncompleters in sociodemographic or clinical variables or in baseline scores [55,66,67,73,75,77,84]. Serdar et al [70] reported that noncompleters had higher baseline scores in thin-ideal internalization. Ljotsson et al [49] found that completers showed less bulimic episodes at baseline. Likewise, Carrard et al [50] showed that completers and noncompleters differed in frequency of binges and vomiting.

Satisfaction was either measured by a satisfaction scale or assessed by qualitative comments. The majority of participants were satisfied with the programs and rated modules as pleasant, easy, and useful [46,71,72,75,78,96]. Patients also said that they would recommend the programs to others [54,72,74,78,97]. Some patients stated that they preferred participating in a program offered via the Internet over face-to-face treatment [71,73]. However, there were also patients who would recommend it only as an adjunct to face-to-face psychotherapy [52].

Follow-Up Dropouts

Follow-up dropout rates in the studies included in this review ranged from 4.7% [69] to 84.8% [46]. In terms of predictors of dropout, results were inconsistent. There were studies where no predictors could be identified [43], whereas others found that dropout could be predicted by higher anxiety scores, a lower hyperactivity, a lower minimum BMI, and lower reward dependence scores on the Temperament and Character Inventory-Revised (TCI-R [98]) [53]. Graham and Walton [46] reported that dropouts were characterized by higher scores on the “drive for thinness” EDI-3 scale. Fichter et al [63] observed that lack of time or not wanting to deal with the eating disorder prevented patients from taking part in the follow-up measurement. Bauer et al [85] found that follow-up dropouts did not differ from those who completed follow-up assessment in any of the baseline characteristics.

Predictors of Outcome

Several studies reported predictors of outcome of TBIs in eating disorders. A better state of general psychological health was found to predict a better outcome [54]. In relation to primary symptomatology, Marrone et al [99] identified reduction of binge eating as the best indicator of a positive treatment outcome (abstinence) at the 1-year follow-up, whereas for short-term follow-up (3 months), it was the reduction of purging behavior. Moreover, higher scores on the EDI “perfectionism” scale and on the Eating Attitudes Test (EAT) [100], as well as a higher minimum BMI significantly correlated with a better outcome [53]. Johnston et al [44] reported that patients who benefited the most from their intervention were those with high body shame at baseline. Fichter et al [63] found lower scores on the Structured Interview for Anorexic and Bulimic Syndromes (SIAB-EX [101]) “compensatory behavior” subscale and higher spontaneity to be predictive of successful relapse prevention. Mezei et al [65] identified a higher ratio of words related to family of origin, a higher BMI, and lower binging and emotional distress at baseline to be significantly associated with change in EDE-Q total. Furthermore, better compliance was associated with a better outcome also. Adherence and more spontaneity were also identified as long-term predictors for a favorable course, as were higher scores on the “ineffectiveness” scale of the EDI-2. In the “SALUT” program, improvement increased with the number of diary entries and the number of steps patients completed [50,54]. Even in email therapy, patients who displayed more commitment in terms of number of words written had better outcomes [59,60]. Paxton et al [67] and Serdar et al [70] also found a correlation between the number of sessions and outcome, whereas Zabinski et al [71] found no such correlation. Finally, Cardi et al [86] reported that the mean use of the vodcasts was correlated with the BMI change.

Details About Guidance

Six eating disorder studies presented information about therapists’ time and efforts in guided CBIs. It ranged between 45 minutes and 135 minutes per patient [47,79,80,86,102]. Therapists in the study by Leung et al [56] invested approximately 5-10 minutes per email. The intervention study of Ruwaard et al [74] included 25 scheduled therapist feedback moments, which took approximately 13 hours to complete. Regarding content, emails from therapists were predominantly supportive (95.4%) with only 14.7% of the emails containing at least 1 CBT comment and 13.6% at least 1 technical comment [102].

Therapeutic Alliance

Only 1 study on videoconferencing for eating disorder patients investigated therapeutic alliance. Therapists experienced differences between the delivery methods in terms of adherence to therapeutic tasks, adherence to therapeutic goals, and therapeutic bond, whereas patients did not [103].

Cost-Effectiveness

Only 1 study compared cost-effectiveness of face-to-face CBT and CBT delivered via telemedicine [104]. Efficacy was found to be comparable, whereas costs of telemedicine were lower, although still substantial.

Risks and Side Effects of TBIs

Three studies provided information about adverse events occurring in the reviewed TBI studies. Sánchez-Ortiz et al [48] as well as Robinson and Serfaty [60] reported that no major adverse events occurred in their studies. Fichter et al [63] reported that 11 of 258 (4.3%) AN patients showed adverse events (BMI<13.5 during the course of the trial). Fortunately, no serious adverse event (acute suicidality, suicide attempts, or death) was identified.

Discussion

Principal Findings

Interest in research on TBIs for AN and BN has increased during the last decade. Forty studies whose outcome results were published in 45 papers fulfilled the inclusion criteria of this systematic review. Most studies dealt with guided CBIs (“Overcoming Bulimia,” “SALUT”) or programs for the prevention and early intervention of eating disorders. One has to consider that many of these treatments are derived from (guided) self-help interventions based on other media (eg, self-help manuals such as “Overcoming Bulimia”) that have been evaluated by a considerable number of researchers in the past and whose results have been summarized in a current meta-analysis [105]. Other approaches were rarely investigated: only 2 studies looked into videoconferencing and email therapy. Merely 3 studies evaluated unguided CBIs and 2 researched the potential of CBIs in relapse prevention.

Summary of Main Results

With regard to BN, guided CBIs led to improvements in the core symptoms of binging and purging and global eating disorder psychopathology. Patients receiving guided CBIs improved more than controls. Guided CBI was shown to be as effective as guided bibliotherapy. Initial findings suggest that treatment results can also be maintained at follow-up. Furthermore, videoconferencing showed promise in treating patients with BN. Unfortunately, this approach has only been evaluated in 1 RCT so far.

With regard to AN, CBIs might be used for relapse prevention. However, only 1 study has empirically evaluated this kind of intervention in AN patients so far. Several case reports by Yager [106-108] suggest that CBIs (eg, email) may be effective when used as a therapeutic adjunct in the outpatient treatment of AN. In his opinion, CBIs increase adherence, patient satisfaction, and therapeutic alliance, and he did not observe any negative effects. Therefore, he suggests including email contact in the regular therapy process to enhance weekly sessions, to monitor daily food intake, or for crisis intervention.

With regard to adolescents with eating disorders, research findings suggest that CBIs may be a treatment option for bulimic patients. However, results should be confirmed in RCTs and replicated by other research groups before widely recommending it.

Furthermore, CBIs may also be considered in the prevention and early intervention of eating disorders as well as for supporting carers of eating disorder patients. Finally, preliminary evidence suggests that mobile interventions are useful for patients with eating disorders in relapse prevention or as an adjunct to therapy (eg, symptom monitoring).

Efficacy results of TBIs are in-line with studies evaluating CBIs in other mental health disorders. For example, medium effects at posttreatment were also shown in a meta-analysis of CBIs for depressive disorders [16]. Videoconferencing—the intervention that is closest to traditional face-to-face therapy—showed very promising results. Videoconferencing was also shown to yield similar clinical outcomes as face-to-face therapy with regard to a variety of other disorders [109]. Although this intervention provides no benefit for the therapist in terms of time or cost-effectiveness, it might be an attractive alternative treatment for patients living far away from a specialist eating disorder therapist. CBIs for carers of eating disorder patients are a relatively new field of research. In-line with Hu et al [110] who showed that Internet-based interventions were able to reduce stress and to improve well-being of caregivers for others with medical conditions, we found that CBIs may also be helpful for carers of patients with eating disorders.

Although efficacy results in the reviewed studies are promising, high rates of non-take-up, noncompliance, and dropout in the reviewed studies severely hamper the validity of study results. Therefore, findings must be interpreted with some caution. The result that only slightly more than half of the participants (57%) completed the offered TBI parallels the report by Waller and Gilbody [111] that only half of patients completed a full course of CBI for anxiety and depression. However, as adherence and satisfaction are important determinants of therapy outcome, it will take additional efforts to improve acceptance of TBIs for participants. This may be accomplished by educational advertising and by developing specific, individually tailored programs that appeal to participants searching for help. To increase adherence, future interventions should feature the following characteristics: a strong theoretical foundation, perceived personal relevance, tailoring, persuasive technologies, credibility, social networking, and common “push factors,” including human support and/or periodic prompts (by email or telephone), as suggested by Murray [112].

Is Human Contact and Guidance Necessary?

Guidance may be essential for both compliance and outcome of CBIs [113,114]. In the field of research on CBIs for eating disorders, only 3 studies directly compare unguided and guided CBI. Huon [45] found more improvement in the 2 guided groups. In contrast, Murray et al [115] did not find any significant differences between brief therapist guidance and minimal researcher guidance. However, this study was not an RCT, but rather compared the results of 2 consecutive cohorts. Hoyle et al [81] got mixed results on whether further guidance is useful. Eating disorder patients see guidance as a helpful and important element of the intervention [54,97].

Studies of CBIs for various mental health disorders showed that guidance augments efficacy [27]. Palmqvist et al [114] calculated a meta-analytic correlation of ρ=.75 between time invested by therapists and outcome. Outcomes equivalent to face-to-face therapy were also reported [116,117].

Overall, details about guidance are unsatisfactorily reported in most studies on CBIs for eating disorders. Information about the type and qualifications of coaches, about the type of support, as well as the timing, frequency, and overall amount of contact is often missing. However, all this important information is required to be able to compare the different programs and their efficacy and to get a clearer picture of how much and what kind of contact is needed to optimize the interventions.

Therapeutic Alliance

The issue of human contact in guided CBIs also raises concerns about whether online therapy can establish any meaningful therapeutic alliance [118]. The only study in the field of CBIs for eating disorders suggests that videoconferencing establishes an adequate therapeutic alliance [113]. However, there are no findings regarding its quality when nonverbal cues are completely lost, such as in email guidance.

A systematic review investigating therapeutic relationships in e-therapy for mental health suggests that a therapeutic alliance equivalent to that in face-to-face therapy can be established in Internet-based therapy [119]. However, this topic is underresearched because just 11 (1.3%) of 840 reviewed studies reported data on this matter. Furthermore, Andersson et al [120] speculated that despite high alliance ratings, these relationships may not be as important as in face-to-face therapy. In conclusion, much more research is necessary to really understand similarities and differences as well as special features of therapeutic relationships in both delivery methods.

Strengths and Limitations of This Review

One strength of the present review is that we considered the whole spectrum of care from prevention, early intervention, treatment, relapse prevention of eating disorders, to interventions for carers of eating disorders. Furthermore, by including a wide range of study designs beyond just RCTs, a broad range of new types of TBIs were included. This made it possible to give a comprehensive up-to-date picture of the dynamic field of TBIs in eating disorders. Moreover, a wide range of aspects relevant to TBIs, such as acceptance, efficacy, predictors of outcome, need of guidance, and therapeutic alliance, were reviewed. This review is limited by the fact that a meta-analysis could not be performed due to the enormous heterogeneity of studies. Furthermore, diagnoses in studies were made in sometimes more but also less rigorous ways (informal clinical interview, semistructured clinical interview, questionnaire using *DSM* criteria), thereby reducing the chances that all studies had the same diagnostic threshold. Unfortunately, no pure samples of BN patients were available in most studies. Instead, patients with EDNOS or binge eating disorder were also included, so that results must be interpreted with some caution.

There are also several methodological limitations in the studies discussed in this review. We mainly followed the coding for weaknesses as suggested by Newman et al [121] (see table in Multimedia Appendix 1). Sixteen studies were uncontrolled and overall, 55% RCTs with either a control group or 13 with an active treatment as a comparator were still too few. For 11 studies, no follow-up data were available and in 28 studies outcomes were determined only through self-report and not by blinded interviewers. Sample sizes were predominantly moderate and rarely included more than 100 participants; 14 studies had sample sizes of less than 50. In most cases, no power calculation was reported. Fifteen studies did not report on additional psychological treatment, as did 22 studies on pharmacological treatment. Twenty-two studies failed to present information on adherence or quality checks in the treatment as did 18 studies on therapist training regarding TBIs. Most studies did not explicitly investigate outcome predictors or the active ingredients of mostly complex interventions. Outcome measures varied considerably making it hard to compare different study results.

Consequentially, a number of challenges for future research arise that are detailed subsequently.

Recommendations for Future Research

There is need for more high-quality RCTs that adhere to the Consolidated Standards of Reporting Trials (CONSORT) statement [122], with sufficient sample sizes and long-term follow-ups to investigate stability of effects. Future studies should compare TBIs for AN and BN patients not only concerning waiting list or no treatment comparisons, but also with respect to active conditions such as face-to-face psychotherapy and traditional (manual-based) self-help interventions. Furthermore, different modes of TBIs as well as TBIs with different amounts of guidance should be compared with each other to derive information about comparative effectiveness. Future research should also evaluate the helpfulness of TBIs in preparing patients for face-to-face therapy, and explore the potential of combining face-to-face therapy and TBIs. Outcomes of evaluations should be preferably assessed by independent, blinded assessors and not by self-reports alone. To facilitate comparability of treatments, standardized outcome measures should be used. In addition, abstinence and compliance should be defined consistently across studies. Furthermore, reporting of studies should be improved. Studies should include information about additional psychological and/or pharmacological treatments to control for outside effects. Additionally, information about adherence or quality checks on treatment needs to be consistently reported. In addition, detailed descriptions of guidance should be given including information on qualification and on training of the coach/therapist. Modality, frequency, and intensity of guidance, as well as information about therapist's time and efforts should also be specified. To improve the methodological quality and the reporting of studies, researchers should follow the guidelines for executing and reporting Internet intervention research [123].

Many questions regarding the optimal delivery method of TBIs, as well as the optimal dose in terms of frequency, intensity, and duration of interventions, still await qualified answers. Research shows that adding more therapy components does not lead to better results [124]. In this context, Andersson and Titov [125] pointed out that it is important to consider patients' capacity and to avoid overloading. Studies should also try to find out how much time patients have to spend on a treatment program to achieve maximal benefits [24]. Similarly, questions regarding optimal guidance remain unresolved: research needs to identify the usefulness of clinical contact [24] and the ideal frequency and kind of provided support [125]. The issue of what background the therapist should ideally have and how much face-to-face contact a therapist needs to have in order to be empathetic, interested, present, and thus therapeutically effective, also needs clarification [126]. In case of engaging TBIs of adequate quality, therapist expertise may not be as important as in face-to-face therapy [125]. Furthermore, future studies should also include checking adherence to the intervention protocol of therapists who provide guidance and the quality of their guidance.

Usually TBIs are complex interventions (ie, they consist of a number of interacting components) [127]. Therefore, it is essential to explore mechanisms of action of TBIs and how therapy can be optimized by identifying the active ingredients. In addition, mediators and moderators of therapeutic change in TBIs should be explored. Future research will also have to clarify who benefits, who drops out, and what factors increase compliance. Moreover, the impact of sociodemographic factors, such as sex, age, and/or education, needs to be elucidated.

Integrating more motivational components into programs and tailoring the treatment to the AN and BN patients' needs in terms of severity of illness and comorbidity will be imperative to further improve interventions. Research on depression has shown that more severely affected patients respond better to tailored, rather than to nontailored treatment [128]. Patients' preferences should also be considered when designing future interventions (eg, by providing a choice of different treatment modules). Therefore, it is necessary to conduct qualitative studies to explore patients' opinions and wishes concerning TBIs and to explore acceptance.

More research on how to enhance the uptake and utilization of TBIs and how to promote its broader dissemination and implementation in routine care is necessary. TBIs should be integrated into existing health care systems and efforts should be made to enhance their acceptability and adoption by patients as well as by health care delivery teams [129]. Many patients and clinicians still adopt a skeptical attitude toward TBIs, a problem that can be addressed through education [125]. More pragmatic RCTs in naturalistic settings should be performed to evaluate effectiveness and cost-effectiveness of TBIs within a stepped-care program [130]. This kind of research is able to identify where TBIs fit best into existing treatment options and to economically integrate them into our health care system [129]. TBIs may be presented as a first step in a stepped-care process probably followed by or combined with other treatment options such as face-to-face psychotherapy [125]. Cavanagh et al [131] reported a naturalistic nonrandomized trial of CBI for anxiety and/or depression that found CBI to be effective even under routine conditions, suggesting good generalizability.

TBIs also have some disadvantages. Some patients may not be able to access technological interventions [132] or may feel uncomfortable using them [129] due to security issues [125] and privacy concerns [133]. Moreover, the lack of nonverbal cues (eg, in email therapy) can lead to a loss of important therapeutic information and to misinterpretations. Additionally, the lack of immediate exchange may cause a patient who expects immediate response to feel insecure and helpless [133]. Furthermore, patients' safety cannot be guaranteed and, in case of a crisis, there are few options for action. Kiluk et al [134] advocate that future studies should focus more on potential adverse events in TBIs before a final recommendation for TBIs is made. So far, adverse events have rarely been considered in reports of TBIs for eating disorder patients. Therefore, this issue should be more systematically addressed. In addition, future studies should focus on the identification of nonresponders or patients that even deteriorate (eg, by symptom monitoring on a regular basis) and on the development of treatment strategies for these patients.

Last but not least, research needs to determine the limits of TBIs [126]. In-line with Proudfoot et al [135], TBIs for eating disorders may not be indicated in case of a BMI<17.5, severe vomiting, comorbid posttraumatic stress disorder, severe depression, or a psychotic illness.

Conclusions

Unfortunately, the evidence for TBIs in the treatment of AN and BN remains insufficient because many approaches were investigated only once, by 1 research group, or only in uncontrolled studies. Nevertheless, the initial results are encouraging. At this stage, unguided CBIs cannot be recommended for the treatment of AN and BN, whereas guided CBIs may be a promising treatment approach, especially for BN. Videoconferencing may also be an approach worth pursuing further in research as well as in practice. Furthermore, Internet-based relapse prevention for AN inpatients may be an effective way to stabilize treatment success and to bridge the gap between inpatient and outpatient therapy. Guided CBIs seem to be a promising approach even in the treatment of adolescents with BN. Efficacy of email therapy still remains to be seen. CBIs might also be considered for the prevention of eating disorders as well as to support carers of eating disorder patients. Furthermore, evaluation of mobile interventions should be further pursued.

Until now, many TBIs have only been used within the context of research studies and have not become part of routine health care. Before a widespread implementation of TBIs, it is imperative to ascertain that they are also feasible under naturalistic conditions and across settings.

In conclusion, one should stay open-minded about the integration of novel technologies that may enhance psychological prevention and treatment of AN and BN and their carers. TBIs can especially serve as a first step in a stepped-care model. However, patient compliance, which is essential for TBIs to work, is still a major challenge. Future research is needed before widely recommending TBIs for AN and BN.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Studies evaluating efficacy of technology-based psychological treatments for anorexia and bulimia nervosa.

[PDF File (Adobe PDF File), 100KB - jmir_v17i3e85_app1.pdf]

Multimedia Appendix 2

Within- and between-group effect sizes and confidence intervals for included studies (posttreatment and follow-up).

[PDF File (Adobe PDF File), 57KB - jmir_v17i3e85_app2.pdf]

Multimedia Appendix 3

Definition of abstinence in studies included in the review as well as abstinence rates and results of significance test wherever data was available.

[PDF File (Adobe PDF File), 42KB - jmir_v17i3e85_app3.pdf]

References

1. Hoek HW. Incidence, prevalence and mortality of anorexia nervosa and other eating disorders. Curr Opin Psychiatry 2006 Jul;19(4):389-394. [doi: 10.1097/01.yco.0000228759.95237.78] [Medline: 16721169]
2. Machado PP, Machado BC, Gonçalves S, Hoek HW. The prevalence of eating disorders not otherwise specified. Int J Eat Disord 2007 Apr;40(3):212-217. [doi: 10.1002/eat.20358] [Medline: 17173324]
3. American Psychiatric Association. Practice Guideline for the Treatment of Patients with Eating Disorders. 3rd edition. Washington, DC: American Psychiatric Association; 2006.
4. National Collaborating Centre for Mental Health. Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders. London: The British Psychological Society and the Royal College of Psychiatrists; 2004.
5. Watson HJ, Bulik CM. Update on the treatment of anorexia nervosa: review of clinical trials, practice guidelines and emerging interventions. Psychol Med 2013 Dec;43(12):2477-2500. [doi: 10.1017/S0033291712002620] [Medline: 23217606]
6. Halmi KA, Agras WS, Crow S, Mitchell J, Wilson GT, Bryson SW, et al. Predictors of treatment acceptance and completion in anorexia nervosa: implications for future study designs. Arch Gen Psychiatry 2005 Jul;62(7):776-781. [doi: 10.1001/archpsyc.62.7.776] [Medline: 15997019]
7. Lock J, Brandt H, Woodside B, Agras S, Halmi WK, Johnson C, et al. Challenges in conducting a multi-site randomized clinical trial comparing treatments for adolescent anorexia nervosa. Int J Eat Disord 2012 Mar;45(2):202-213 [FREE Full text] [doi: 10.1002/eat.20923] [Medline: 21495052]
8. Mond JM, Hay PJ, Rodgers B, Owen C. Health service utilization for eating disorders: findings from a community-based study. Int J Eat Disord 2007 Jul;40(5):399-408. [doi: 10.1002/eat.20382] [Medline: 17497708]
9. Cachelin FM, Striegel-Moore RH. Help seeking and barriers to treatment in a community sample of Mexican American and European American women with eating disorders. Int J Eat Disord 2006 Mar;39(2):154-161. [doi: 10.1002/eat.20213] [Medline: 16252278]
10. Hart LM, Granillo MT, Jorm AF, Paxton SJ. Unmet need for treatment in the eating disorders: a systematic review of eating disorder specific treatment seeking among community cases. Clin Psychol Rev 2011 Jul;31(5):727-735. [doi: 10.1016/j.cpr.2011.03.004] [Medline: 21501580]
11. Currin L, Schmidt U, Treasure J, Jick H. Time trends in eating disorder incidence. Br J Psychiatry 2005 Feb;186(2):132-135 [FREE Full text] [doi: 10.1192/bjp.186.2.132] [Medline: 15684236]
12. Escobar-Koch T, Banker JD, Crow S, Cullis J, Ringwood S, Smith G, et al. Service users' views of eating disorder services: an international comparison. Int J Eat Disord 2010 Sep;43(6):549-559. [doi: 10.1002/eat.20741] [Medline: 19718664]
13. de la Rie S, Noordenbos G, Donker M, van Furth E. Evaluating the treatment of eating disorders from the patient's perspective. Int J Eat Disord 2006 Dec;39(8):667-676. [doi: 10.1002/eat.20317] [Medline: 16868992]
14. Marks IM, Cavanagh K, Gega L. Computer-aided psychotherapy: revolution or bubble. Br J Psychiatry 2007 Dec;191(6):471-473 [FREE Full text] [doi: 10.1192/bjp.bp.107.041152] [Medline: 18055948]
15. Foroushani PS, Schneider J, Assareh N. Meta-review of the effectiveness of computerised CBT in treating depression. BMC Psychiatry 2011 Aug;11:131 [FREE Full text] [doi: 10.1186/1471-244X-11-131] [Medline: 21838902]
16. 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]
17. Kaltenthaler E, Sutcliffe P, Parry G, Beverley C, Rees A, Ferriter M. The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. Psychol Med 2008 Nov;38(11):1521-1530. [doi: 10.1017/S0033291707002607] [Medline: 18205964]

18. 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 Oct;5(10):e1396 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
19. Van't Hof E, Cuijpers P, Stein DJ. Self-help and Internet-guided interventions in depression and anxiety disorders: a systematic review of meta-analyses. CNS Spectr 2009 Feb;14(2 Suppl 3):34-40. [Medline: 19238128]
20. Cuijpers P, Marks IM, van Straten A, Cavanagh K, Gega L, Andersson G. Computer-aided psychotherapy for anxiety disorders: a meta-analytic review. Cogn Behav Ther 2009 Jun;38(2):66-82. [doi: 10.1080/16506070802694776] [Medline: 20183688]
21. Reger MA, Gahm GA. A meta-analysis of the effects of internet- and computer-based cognitive-behavioral treatments for anxiety. J Clin Psychol 2009 Jan;65(1):53-75. [doi: 10.1002/jclp.20536] [Medline: 19051274]
22. Herbst N, Voderholzer U, Stelzer N, Knaevelsrud C, Hertenstein E, Schlegl S, et al. The potential of telemental health applications for obsessive-compulsive disorder. Clin Psychol Rev 2012 Aug;32(6):454-466. [doi: 10.1016/j.cpr.2012.04.005] [Medline: 22705583]
23. Tumur I, Kaltenthaler E, Ferriter M, Beverley C, Parry G. Computerised cognitive behaviour therapy for obsessive-compulsive disorder: a systematic review. Psychother Psychosom 2007 Jun;76(4):196-202. [doi: 10.1159/000101497] [Medline: 17570957]
24. Amstadter AB, Broman-Fulks J, Zinzow H, Ruggiero KJ, Cercone J. Internet-based interventions for traumatic stress-related mental health problems: a review and suggestion for future research. Clin Psychol Rev 2009 Jul;29(5):410-420 [FREE Full text] [doi: 10.1016/j.cpr.2009.04.001] [Medline: 19403215]
25. Benight CC, Ruzek JI, Waldrep E. Internet interventions for traumatic stress: a review and theoretically based example. J Trauma Stress 2008 Dec;21(6):513-520. [doi: 10.1002/jts.20371] [Medline: 19107724]
26. Barak A, Hen L, Boniel-Nissim M, Shapira N. A comprehensive review and a meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. J Tech Hum Serv 2008 Jul;26(2/4):109-160. [doi: 10.1080/15228830802094429]
27. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
28. Richardson T, Stallard P, Velleman S. Computerised cognitive behavioural therapy for the prevention and treatment of depression and anxiety in children and adolescents: a systematic review. Clin Child Fam Psychol Rev 2010 Sep;13(3):275-290. [doi: 10.1007/s10567-010-0069-9] [Medline: 20532980]
29. Siemer CP, Fogel J, Van Voorhees BW. Telemental health and web-based applications in children and adolescents. Child Adolesc Psychiatr Clin N Am 2011 Jan;20(1):135-153 [FREE Full text] [doi: 10.1016/j.chc.2010.08.012] [Medline: 21092918]
30. Schmidt U, Grover M. Computer-based intervention for bulimia nervosa and binge eating. In: Latner JD, Wilson GT, editors. Self-Help Approaches for Obesity and Eating Disorders: Research and Practice. New York: The Guilford Press; 2007:166-176.
31. Paxton SJ, Franko DL. Body image and eating disorders. In: Cucciare MA, Weingardt KR, editors. Using Technology to Support Evidence-Based Behavioral Health Practices: A Clinician's Guide. New York: Taylor & Francis Group; 2010:151-168.
32. Engel SG, Wonderlich SA. New technologies in treatments for eating disorders. In: Grilo CM, Mitchell JE, editors. The Treatment of Eating Disorders: A Clinical Handbook. New York: Guilford Publications; 2009:500-509.
33. Kaiser NC, Owen JE, Winzelberg AJ. Technological advances in modifying adolescent health risk behaviors. In: DiClemente RJ, Santelli JS, Crosby RA, editors. Adolescent Health: Understanding and Preventing Risk Behaviors. San Francisco, CA: Jossey-Bass; 2009:447-472.
34. Myers TC, Swan-Kremeier L, Wonderlich S, Lancaster K, Mitchell JE. The use of alternative delivery systems and new technologies in the treatment of patients with eating disorders. Int J Eat Disord 2004 Sep;36(2):123-143. [doi: 10.1002/eat.20032] [Medline: 15282683]
35. Marks I, Cavanagh K, Gega L. CP for eating problems. In: Marks I, Cavanagh K, Gega L, editors. Hands-On Help: Computer-Aided Psychotherapy. New York: Psychology Press; 2007:109-136.
36. Dölemeyer R, Tietjen A, Kersting A, Wagner B. Internet-based interventions for eating disorders in adults: a systematic review. BMC Psychiatry 2013 Aug;13:207 [FREE Full text] [doi: 10.1186/1471-244X-13-207] [Medline: 23919625]
37. Bauer S, Moessner M. Harnessing the power of technology for the treatment and prevention of eating disorders. Int J Eat Disord 2013 Jul;46(5):508-515. [doi: 10.1002/eat.22109] [Medline: 23658102]
38. Ferrer-García M, Gutiérrez-Maldonado J. The use of virtual reality in the study, assessment, and treatment of body image in eating disorders and nonclinical samples: a review of the literature. Body Image 2012 Jan;9(1):1-11. [doi: 10.1016/j.bodyim.2011.10.001] [Medline: 22119329]
39. Beintner I, Jacobi C, Taylor CB. Effects of an Internet-based prevention programme for eating disorders in the USA and Germany-a meta-analytic review. Eur Eat Disord Rev 2012 Jan;20(1):1-8. [doi: 10.1002/erv.1130] [Medline: 21796737]

40. Beintner I, Jacobi C, Schmidt U. Participant adherence to the Internet-based prevention program student bodies(TM) for eating disorders - A review. Internet Interventions 2014 Mar;1(1):26-32 [FREE Full text]
41. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. Chichester: John Wiley & Sons; 2009.
42. Bara-Carril N, Williams CJ, Pombo-Carril MG, Reid Y, Murray K, Aubin S, et al. A preliminary investigation into the feasibility and efficacy of a CD-ROM-based cognitive-behavioral self-help intervention for bulimia nervosa. Int J Eat Disord 2004 May;35(4):538-548. [doi: 10.1002/eat.10267] [Medline: 15101069]
43. Schmidt U, Andiappan M, Grover M, Robinson S, Perkins S, Dugmore O, et al. Randomised controlled trial of CD-ROM-based cognitive-behavioural self-care for bulimia nervosa. Br J Psychiatry 2008 Dec;193(6):493-500 [FREE Full text] [doi: 10.1192/bjp.bp.107.046607] [Medline: 19043154]
44. Johnston O, Startup H, Lavender A, Godfrey E, Schmidt U. Therapeutic writing as an intervention for symptoms of bulimia nervosa: effects and mechanism of change. Int J Eat Disord 2010 Jul;43(5):405-419. [doi: 10.1002/eat.20714] [Medline: 19544556]
45. Huon GF. An initial validation of a self-help program for bulimia. Int J Eat Disord 1985 Nov;4(4):573-588. [doi: 10.1002/1098-108X(198511)4:4<573::AID-EAT2260040416>3.0.CO;2-0]
46. Graham L, Walton M. Investigating the use of CD-Rom CBT for bulimia nervosa and binge eating disorder in an NHS adult outpatient eating disorders service. Behav Cogn Psychother 2011 Jul;39(4):443-456. [doi: 10.1017/S1352465810000688] [Medline: 21208485]
47. Murray K, Schmidt U, Pombo-Carril MG, Grover M, Alenya J, Treasure J, et al. Does therapist guidance improve uptake, adherence and outcome from a CD-ROM based cognitive-behavioral intervention for the treatment of bulimia nervosa. Comput Human Behav 2007 Jan;23(1):850-859. [doi: 10.1016/j.chb.2004.11.014]
48. Sánchez-Ortiz VC, Munro C, Stahl D, House J, Startup H, Treasure J, et al. A randomized controlled trial of internet-based cognitive-behavioural therapy for bulimia nervosa or related disorders in a student population. Psychol Med 2011 Feb;41(2):407-417. [doi: 10.1017/S0033291710000711] [Medline: 20406523]
49. Ljotsson B, Lundin C, Mitsell K, Carlbring P, Ramklint M, Ghaderi A. Remote treatment of bulimia nervosa and binge eating disorder: a randomized trial of Internet-assisted cognitive behavioural therapy. Behav Res Ther 2007 Apr;45(4):649-661. [doi: 10.1016/j.brat.2006.06.010] [Medline: 16899213]
50. Carrard I, Rouget P, Fernández-Aranda F, Volkart AC, Damoiseau M, Lam T. Evaluation and deployment of evidence based patient self-management support program for Bulimia Nervosa. Int J Med Inform 2006 Jan;75(1):101-109. [doi: 10.1016/j.ijmedinf.2005.07.031] [Medline: 16115793]
51. Liwowsky I, Cebulla M, Fichter M. New ways to combat eating disorders-evaluation of an internet-based self-help program in bulimia nervosa. MMW Fortschr Med 2006 Aug;148(31-32):31-33. [Medline: 16937874]
52. Nevonen L, Mark M, Levin B, Lindström M, Paulson-Karlsson G. Evaluation of a new Internet-based self-help guide for patients with bulimic symptoms in Sweden. Nord J Psychiatry 2006 Jan;60(6):463-468. [doi: 10.1080/08039480601021993] [Medline: 17162454]
53. Fernández-Aranda F, Núñez A, Martínez C, Krug I, Cappozzo M, Carrard I, et al. Internet-based cognitive-behavioral therapy for bulimia nervosa: a controlled study. Cyberpsychol Behav 2009 Feb;12(1):37-41. [doi: 10.1089/cpb.2008.0123] [Medline: 19006463]
54. Carrard I, Fernandez-Aranda F, Lam T, Nevonen L, Liwowsky I, Volkart AC, et al. Evaluation of a guided internet self-treatment programme for bulimia nervosa in several European countries. Eur Eat Disord Rev 2011 Mar;19(2):138-149. [doi: 10.1002/erv.1043] [Medline: 20859989]
55. Wagner G, Penelo E, Wanner C, Gwinner P, Trofaier ML, Imgart H, Karwautz Andreas F K. Internet-delivered cognitive-behavioural therapy v. conventional guided self-help for bulimia nervosa: long-term evaluation of a randomised controlled trial. Br J Psychiatry 2013 Feb;202(2):135-141 [FREE Full text] [doi: 10.1192/bjp.bp.111.098582] [Medline: 23222037]
56. Leung SF, Joyce LC, Russell J. An open trial of self-help behaviours of clients with eating disorders in an online programme. J Adv Nurs 2013 Jan;69(1):66-76. [doi: 10.1111/j.1365-2648.2012.05988.x] [Medline: 22448710]
57. Pretorius N, Arcelus J, Beecham J, Dawson H, Doherty F, Eisler I, et al. Cognitive-behavioural therapy for adolescents with bulimic symptomatology: the acceptability and effectiveness of internet-based delivery. Behav Res Ther 2009 Sep;47(9):729-736. [doi: 10.1016/j.brat.2009.05.006] [Medline: 19515360]
58. Wagner G, Penelo E, Nobis G, Mayerhofer A, Schau J, Spitzer M, et al. Is technology assisted guided self-help successful in treating female adolescents with bulimia nervosa. Neuropsychiatr 2013 Jun;27(2):66-73. [doi: 10.1007/s40211-013-0062-x] [Medline: 23609487]
59. Robinson PH, Serfaty MA. The use of e-mail in the identification of bulimia nervosa and its treatment. Eur Eat Disord Rev 2001 May;9(3):182-193. [doi: 10.1002/erv.411]
60. Robinson P, Serfaty M. Getting better byte by byte: a pilot randomised controlled trial of email therapy for bulimia nervosa and binge eating disorder. Eur Eat Disord Rev 2008 Mar;16(2):84-93. [doi: 10.1002/erv.818] [Medline: 17879223]

61. Simpson S, Knox J, Mitchell D, Ferguson J, Brebner J, Brebner E. A multidisciplinary approach to the treatment of eating disorders via videoconferencing in north-east Scotland. J Telemed Telecare 2003 Jun;9 Suppl 1:37-38. [doi: 10.1258/135763303322196286] [Medline: 12952717]
62. Mitchell JE, Crosby RD, Wonderlich SA, Crow S, Lancaster K, Simonich H, et al. A randomized trial comparing the efficacy of cognitive-behavioral therapy for bulimia nervosa delivered via telemedicine versus face-to-face. Behav Res Ther 2008 May;46(5):581-592 [FREE Full text] [doi: 10.1016/j.brat.2008.02.004] [Medline: 18374304]
63. Fichter MM, Quadflieg N, Nisslmüller K, Lindner S, Osen B, Huber T, et al. Does internet-based prevention reduce the risk of relapse for anorexia nervosa. Behav Res Ther 2012 Mar;50(3):180-190. [doi: 10.1016/j.brat.2011.12.003] [Medline: 22317754]
64. Fichter MM, Quadflieg N, Lindner S. Internet-based relapse prevention for anorexia nervosa: nine-month follow-up. J Eat Disord 2013 Jul;1:23 [FREE Full text] [doi: 10.1186/2050-2974-1-23] [Medline: 24999404]
65. Mezei Á, Gulec H, Czeglédi E, Fritz A, Túry F. Linguistic characteristics of patients with bulimic symptomatology in an online post-treatment program: an exploratory study. Eat Weight Disord 2015 Mar;20(1):63-70. [doi: 10.1007/s40519-014-0136-1] [Medline: 24962794]
66. Gollings EK, Paxton SJ. Comparison of internet and face-to-face delivery of a group body image and disordered eating intervention for women: a pilot study. Eat Disord 2006 Feb;14(1):1-15. [doi: 10.1080/10640260500403790] [Medline: 16757445]
67. Paxton SJ, McLean SA, Gollings EK, Faulkner C, Wertheim EH. Comparison of face-to-face and internet interventions for body image and eating problems in adult women: an RCT. Int J Eat Disord 2007 Dec;40(8):692-704. [doi: 10.1002/eat.20446] [Medline: 17702020]
68. Stice E, Rohde P, Durant S, Shaw H. A preliminary trial of a prototype Internet dissonance-based eating disorder prevention program for young women with body image concerns. J Consult Clin Psychol 2012 Oct;80(5):907-916 [FREE Full text] [doi: 10.1037/a0028016] [Medline: 22506791]
69. Stice E, Durant S, Rohde P, Shaw H. Effects of a prototype Internet dissonance-based eating disorder prevention program at 1- and 2-year follow-up. Health Psychol 2014 Dec;33(12):1558-1567. [doi: 10.1037/hea0000090] [Medline: 25020152]
70. Serdar K, Kelly NR, Palmberg AA, Lydecker JA, Thornton L, Tully CE, et al. Comparing online and face-to-face dissonance-based eating disorder prevention. Eat Disord 2014 Jan;22(3):244-260. [doi: 10.1080/10640266.2013.874824] [Medline: 24456277]
71. Zabinski MF, Wilfley DE, Calfas KJ, Winzelberg AJ, Taylor CB. An interactive psychoeducational intervention for women at risk of developing an eating disorder. J Consult Clin Psychol 2004 Oct;72(5):914-919. [doi: 10.1037/0022-006X.72.5.914] [Medline: 15482051]
72. Ohlmer R, Jacobi C, Taylor CB. Preventing symptom progression in women at risk for AN: results of a pilot study. Eur Eat Disord Rev 2013 Jul;21(4):323-329. [doi: 10.1002/erv.2225] [Medline: 23520152]
73. Heinicke BE, Paxton SJ, McLean SA, Wertheim EH. Internet-delivered targeted group intervention for body dissatisfaction and disordered eating in adolescent girls: a randomized controlled trial. J Abnorm Child Psychol 2007 Jun;35(3):379-391. [doi: 10.1007/s10802-006-9097-9] [Medline: 17243014]
74. Ruwaard J, Lange A, Broeksteeg J, Renteria-Agirre A, Schrieken B, Dolan CV, et al. Online cognitive-behavioural treatment of bulimic symptoms: a randomized controlled trial. Clin Psychol Psychother 2013 Aug;20(4):308-318. [doi: 10.1002/cpp.1767] [Medline: 22298417]
75. Jacobi C, Völker U, Trockel MT, Taylor CB. Effects of an Internet-based intervention for subthreshold eating disorders: a randomized controlled trial. Behav Res Ther 2012 Feb;50(2):93-99. [doi: 10.1016/j.brat.2011.09.013] [Medline: 22137366]
76. Hötzel K, von Brachel R, Schmidt U, Rieger E, Kosfelder J, Hechler T, et al. An Internet-based program to enhance motivation to change in females with symptoms of an eating disorder: a randomized controlled trial. Psychol Med 2014 Jul;44(9):1947-1963. [doi: 10.1017/S0033291713002481] [Medline: 24128818]
77. Leung SF, Ma J, Russell J. Enhancing motivation to change in eating disorders with an online self-help program. Int J Ment Health Nurs 2013 Aug;22(4):329-339. [doi: 10.1111/j.1447-0349.2012.00870.x] [Medline: 22882349]
78. Binford Hopf RB, 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]
79. 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]
80. 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]
81. Hoyle D, Slater J, Williams C, Schmidt U, Wade TD. Evaluation of a web-based skills intervention for carers of people with anorexia nervosa: a randomized controlled trial. Int J Eat Disord 2013 Sep;46(6):634-638. [doi: 10.1002/eat.22144] [Medline: 23712500]

82. Bruning Brown J, Winzelberg AJ, Abascal LB, Taylor CB. An evaluation of an Internet-delivered eating disorder prevention program for adolescents and their parents. J Adolesc Health 2004 Oct;35(4):290-296. [doi: 10.1016/j.jadohealth.2003.10.010] [Medline: 15450542]
83. Shapiro JR, Bauer S, Andrews E, Pisetsky E, Bulik-Sullivan B, Hamer RM, et al. Mobile therapy: Use of text-messaging in the treatment of bulimia nervosa. Int J Eat Disord 2010 Sep;43(6):513-519. [doi: 10.1002/eat.20744] [Medline: 19718672]
84. Robinson S, Perkins S, Bauer S, Hammond N, Treasure J, Schmidt U. Aftercare intervention through text messaging in the treatment of bulimia nervosa-feasibility pilot. Int J Eat Disord 2006 Dec;39(8):633-638. [doi: 10.1002/eat.20272] [Medline: 16937381]
85. Bauer S, Okon E, Meermann R, Kordy H. Technology-enhanced maintenance of treatment gains in eating disorders: efficacy of an intervention delivered via text messaging. J Consult Clin Psychol 2012 Aug;80(4):700-706. [doi: 10.1037/a0028030] [Medline: 22545736]
86. Cardi V, Clarke A, Treasure J. The use of guided self-help incorporating a mobile component in people with eating disorders: a pilot study. Eur Eat Disord Rev 2013 Jul;21(4):315-322. [doi: 10.1002/erv.2235] [Medline: 23677740]
87. Henderson M, Freeman CP. A self-rating scale for bulimia. The 'BITE'. Br J Psychiatry 1987 Jan;150(1):18-24. [Medline: 3651670]
88. Cooper Z, Cooper PJ, Fairburn CG. The validity of the eating disorder examination and its subscales. Br J Psychiatry 1989 Jun;154(6):807-812. [Medline: 2597887]
89. Garner DM. Eating Disorder Inventory-2: Professional Manual. Odessa: Psychological Assessment Resources; 1991.
90. Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire. Int J Eat Disord 1994 Dec;16(4):363-370. [Medline: 7866415]
91. Stice E, Ziemba C, Margolis J, Flick P. The dual pathway model differentiates bulimics, subclinical bulimics, and controls: Testing the continuity hypothesis. Behav Ther 1996 Sep;27(4):531-549. [doi: 10.1016/S0005-7894(96)80042-6]
92. Cooper PJ, Taylor MJ, Cooper Z, Fairbum CG. The development and validation of the body shape questionnaire. Int J Eat Disord 1987 Jul;6(4):485-494. [doi: 10.1002/1098-108X(198707)6:4<485::AID-EAT2260060405>3.0.CO;2-O]
93. Cole JD, Kazarian SS. The Level of Expressed Emotion Scale: a new measure of expressed emotion. J Clin Psychol 1988 May;44(3):392-397. [Medline: 3384966]
94. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [Medline: 6880820]
95. Sepulveda AR, Whitney J, Hankins M, Treasure J. Development and validation of an Eating Disorders Symptom Impact Scale (EDSIS) for carers of people with eating disorders. Health Qual Life Outcomes 2008 Apr;6:28 [FREE Full text] [doi: 10.1186/1477-7525-6-28] [Medline: 18426597]
96. Pretorius N, Rowlands L, Ringwood S, Schmidt U. Young people's perceptions of and reasons for accessing a web-based cognitive behavioural intervention for bulimia nervosa. Eur Eat Disord Rev 2010 May;18(3):197-206. [doi: 10.1002/erv.985] [Medline: 20443203]
97. Fichter MM, Quadflieg N, Nisslmüller K, Lindner S, Voderholzer U, Wünsch-Leiteritz W, et al. Internetbasierte Ansätze in der Therapie von Essstörungen. Nervenarzt 2011 Sep;82(9):1107-1117. [doi: 10.1007/s00115-010-3228-9]
98. Cloninger CR. The Temperament and Character Inventory - Revised. St Louis, MO: Center for Psychobiology of Personality; 1999.
99. Marrone S, Mitchell JE, Crosby R, Wonderlich S, Jollie-Trottier T. Predictors of response to cognitive behavioral treatment for bulimia nervosa delivered via telemedicine versus face-to-face. Int J Eat Disord 2009 Apr;42(3):222-227. [doi: 10.1002/eat.20603] [Medline: 18951452]
100. Garner DM, Garfinkel PE. The Eating Attitudes Test: an index of the symptoms of anorexia nervosa. Psychol Med 1979 May;9(2):273-279. [Medline: 472072]
101. Fichter MM, Quadflieg N. Strukturiertes Inventar für Anorektische und Bulimische Essstörungen (SIAB). Göttingen: Hogrefe; 1999.
102. Sánchez-Ortiz VC, Munro C, Startup H, Treasure J, Schmidt U. The role of email guidance in internet-based cognitive-behavioural self-care treatment for bulimia nervosa. Eur Eat Disord Rev 2011 Jul;19(4):342-348. [doi: 10.1002/erv.1074] [Medline: 21394832]
103. Ertelt TW, Crosby RD, Marino JM, Mitchell JE, Lancaster K, Crow SJ. Therapeutic factors affecting the cognitive behavioral treatment of bulimia nervosa via telemedicine versus face-to-face delivery. Int J Eat Disord 2011 Dec;44(8):687-691. [doi: 10.1002/eat.20874] [Medline: 22072405]
104. Crow SJ, Mitchell JE, Crosby RD, Swanson SA, Wonderlich S, Lancanster K. The cost effectiveness of cognitive behavioral therapy for bulimia nervosa delivered via telemedicine versus face-to-face. Behav Res Ther 2009 Jun;47(6):451-453 [FREE Full text] [doi: 10.1016/j.brat.2009.02.006] [Medline: 19356743]
105. Beintner I, Jacobi C, Schmidt U. Participation and outcome in manualized self-help for bulimia nervosa and binge eating disorder - a systematic review and metaregression analysis. Clin Psychol Rev 2014 Mar;34(2):158-176. [doi: 10.1016/j.cpr.2014.01.003] [Medline: 24508686]
106. Yager J. E-mail therapy for anorexia nervosa: prospects and limitations. Eur Eat Disord Rev 2003 May;11(3):198-209. [doi: 10.1002/erv.526]

107. Yager J. Using e-mail to support the outpatient treatment of anorexia nervosa. In: Hsiung RC, editor. E-Therapy: Case Studies, Guiding Principles, and the Clinical Potential of the Internet. New York: WW Norton & Company; 2002:39-68.
108. Yager J. E-mail as a therapeutic adjunct in the outpatient treatment of anorexia nervosa: Illustrative case material and discussion of the issues. Int J Eat Disord 2001 Mar;29(2):125-138. [Medline: 11429975]
109. Backhaus A, Agha Z, Maglione ML, Repp A, Ross B, Zuest D, et al. Videoconferencing psychotherapy: a systematic review. Psychol Serv 2012 May;9(2):111-131. [doi: 10.1037/a0027924] [Medline: 22662727]
110. Hu C, Kung S, Rummans TA, Clark MM, Lapid MI. Reducing caregiver stress with internet-based interventions: a systematic review of open-label and randomized controlled trials. J Am Med Inform Assoc 2014 Aug:1-11. [doi: 10.1136/amiajnl-2014-002817] [Medline: 25125686]
111. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
112. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. Med 2.0 2012 Aug;1(2):1-13 [FREE Full text] [doi: 10.2196/med20.1741] [Medline: 25075231]
113. Gellatly J, Bower P, Hennessy S, Richards D, Gilbody S, Lovell K. What makes self-help interventions effective in the management of depressive symptoms? Meta-analysis and meta-regression. Psychol Med 2007 Sep;37(9):1217-1228. [doi: 10.1017/S0033291707000062] [Medline: 17306044]
114. Palmqvist B, Carlbring P, Andersson G. Internet-delivered treatments with or without therapist input: does the therapist factor have implications for efficacy and cost. Expert Rev Pharmacoecon Outcomes Res 2007 Jun;7(3):291-297. [doi: 10.1586/14737167.7.3.291] [Medline: 20528315]
115. Murray K, Pombo-Carril MG, Bara-Carril N, Grover M, Reid Y, Langham C, et al. Factors determining uptake of a CD-ROM-based CBT self-help treatment for bulimia: patient characteristics and subjective appraisals of self-help treatment. Eur Eat Disord Rev 2003 May;11(3):243-260. [doi: 10.1002/erv.519]
116. Hedman E, Andersson G, Ljótsson B, Andersson E, Rück C, Mörtberg E, et al. Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: a randomized controlled non-inferiority trial. PLoS One 2011 Mar;6(3):1-10 [FREE Full text] [doi: 10.1371/journal.pone.0018001] [Medline: 21483704]
117. Bergström J, Andersson G, Ljótsson B, Rück C, Andréewitch S, Karlsson A, et al. Internet-versus group-administered cognitive behaviour therapy for panic disorder in a psychiatric setting: a randomised trial. BMC Psychiatry 2010 Jul;10:54 [FREE Full text] [doi: 10.1186/1471-244X-10-54] [Medline: 20598127]
118. Perle JG, Langsam LC, Nierenberg B. Controversy clarified: an updated review of clinical psychology and tele-health. Clin Psychol Rev 2011 Dec;31(8):1247-1258. [doi: 10.1016/j.cpr.2011.08.003] [Medline: 21963670]
119. Sucala M, Schnur JB, Constantino MJ, Miller SJ, Brackman EH, Montgomery GH. The therapeutic relationship in e-therapy for mental health: a systematic review. J Med Internet Res 2012 Aug;14(4):1-13 [FREE Full text] [doi: 10.2196/jmir.2084] [Medline: 22858538]
120. Andersson G, Paxling B, Wiwe M, Vernmark K, Felix CB, Lundborg L, et al. Therapeutic alliance in guided internet-delivered cognitive behavioural treatment of depression, generalized anxiety disorder and social anxiety disorder. Behav Res Ther 2012 Sep;50(9):544-550. [doi: 10.1016/j.brat.2012.05.003] [Medline: 22728647]
121. Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: is human contact necessary for therapeutic efficacy. Clin Psychol Rev 2011 Feb;31(1):89-103. [doi: 10.1016/j.cpr.2010.09.008] [Medline: 21130939]
122. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010 Mar;340:c332 [FREE Full text] [Medline: 20332509]
123. Proudfoot J, Klein B, Barak A, Carlbring P, Cuijpers P, Lange A, et al. Establishing guidelines for executing and reporting Internet intervention research. Cogn Behav Ther 2011 Jun;40(2):82-97. [doi: 10.1080/16506073.2011.573807] [Medline: 25155812]
124. Danaher BG, Seeley JR. Methodological issues in research on web-based behavioral interventions. Ann Behav Med 2009 Aug;38(1):28-39 [FREE Full text] [doi: 10.1007/s12160-009-9129-0] [Medline: 19806416]
125. 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]
126. Stein B, Schauenburg H, Eichenberg C. Psychotherapie und Internet - eine Herausforderung für uns alle. PiD 2011 Jun;12(2):171-173. [doi: 10.1055/s-0031-1276822]
127. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep;337:a1655 [FREE Full text] [Medline: 18824488]
128. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-869. [doi: 10.1586/ern.12.63] [Medline: 22853793]
129. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013 Aug;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]

130. Kaltenthaler E, Brazier J, De Nigris E, Tumur I, Ferriter M, Beverley C, et al. Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. Health Technol Assess 2006 Sep;10(33):1-168 [FREE Full text] [Medline: 16959169]
131. Cavanagh K, Shapiro DA, Van Den Berg S, Swain S, Barkham M, Proudfoot J. The effectiveness of computerized cognitive behavioural therapy in routine care. Br J Clin Psychol 2006 Nov;45(4):499-514. [doi: 10.1348/014466505X84782] [Medline: 17076960]
132. Schulz R, Lustig A, Handler S, Martire LM. Technology-based caregiver intervention research: Current status and future directions. Gerontechnology 2002;2(1):15-47 [FREE Full text]
133. Luxton DD, June JD, Kinn JT. Technology-based suicide prevention: current applications and future directions. Telemed J E Health 2011 Sep;17(1):50-54. [doi: 10.1089/tmj.2010.0091] [Medline: 21214382]
134. Kiluk BD, Sugarman DE, Nich C, Gibbons CJ, Martino S, Rounsaville BJ, et al. A methodological analysis of randomized clinical trials of computer-assisted therapies for psychiatric disorders: toward improved standards for an emerging field. Am J Psychiatry 2011 Aug;168(8):790-799 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10101443] [Medline: 21536689]
135. Proudfood J, Klein B, Andersson G, Carlbring P, Kyrios M, Munro C, et al. Guided CBT internet interventions: specific issues in supporting clients with depression, anxiety and comorbid conditions. In: Bennett-Levy J, Richards DA, Farrand P, editors. Oxford Guide to Low Intensity CBT Interventions. Oxford: Oxford University Press; 2010:253-264.

Abbreviations

AN: anorexia nervosa
BEAT: Beating Eating Disorders
BED: binge eating disorder
BITE: Bulimic Investigatory Test-Edinburgh
BMI: body mass index
BN: bulimia nervosa
BSQ: Body Shape Questionnaire
CBI: computer- and Internet-based intervention
CBT: cognitive behavioral therapy
CONSORT: Consolidated Standards of Reporting Trials
DSM: Diagnostic and Statistical Manual of Mental Disorders
EAT: Eating Attitudes Test
ED: eating disorder
EDE: Eating Disorder Examination
EDE-Q: Eating Disorder Examination Questionnaire
EDI-2/EDI-3: Eating Disorder Inventory-2/Eating Disorder Inventory-3
EDNOS: eating disorder not otherwise specified
EDSIS: Eating Disorder Symptom Impact Scale
HADS: Hospital Anxiety and Depression Scale
IBSS-R: Ideal-Body Stereotype Scale-Revised
IPT: interpersonal psychotherapy
LEE: Level of Expressed Emotion Scale
OAO: Overcoming Anorexia Online
PACS: Parental Attitudes and Criticism Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SIAB-EX: Structured Interview for Anorexic and Bulimic Syndromes
SMS: short message service
TAU: treatment as usual
TBI: technology-based interventions
TCI-R: Temperament and Character Inventory-Revised

Edited by G Eysenbach; submitted 24.05.14; peer-reviewed by F Fernandez-Aranda, R von Brachel, C Dahlgren, I Beintner; comments to author 17.07.14; revised version received 28.09.14; accepted 20.10.14; published 31.03.15.

Please cite as:
Schlegl S, Bürger C, Schmidt L, Herbst N, Voderholzer U
The Potential of Technology-Based Psychological Interventions for Anorexia and Bulimia Nervosa: A Systematic Review and Recommendations for Future Research
J Med Internet Res 2015;17(3):e85
URL: http://www.jmir.org/2015/3/e85/
doi:10.2196/jmir.3554
PMID:25840591

Original Paper

Know Your Audience: Predictors of Success for a Patient-Centered Texting App to Augment Linkage to HIV Care in Rural Uganda

Mark J Siedner[1], MD, MPH; Data Santorino[2], MD; Jessica E Haberer[1], MD, MS; David R Bangsberg[1], MD, MPH

[1]Center for Global Health, Massachusetts Genneral Hospital, Harvard Medical School, Boston, MA, United States
[2]Department of Pediatrics, Mbarara University of Science and Technology, Mbarara, Uganda

Corresponding Author:
Mark J Siedner, MD, MPH
Center for Global Health
Massachusetts Genneral Hospital
Harvard Medical School
15th Floor
100 Cambridge Street
Boston, MA, 02114
United States
Phone: 1 617 726 4686
Fax: 1 617 724 1637
Email: msiedner@partners.org

Abstract

Background: Despite investments in infrastructure and evidence for high acceptability, few mHealth interventions have been implemented in sub-Saharan Africa.

Objective: We sought to (1) identify predictors of uptake of an mHealth application for a low-literacy population of people living with HIV (PLWH) in rural Uganda and (2) evaluate the efficacy of various short message service (SMS) text message formats to optimize the balance between confidentiality and accessibility.

Methods: The trial evaluated the efficacy of a SMS text messaging app to notify PLWH of their laboratory results and request return to care for those with abnormal test results. Participants with a normal laboratory result received a single SMS text message indicating results were normal. Participants with an abnormal test result were randomized to 1 of 3 message formats designed to evaluate trade-offs between clarity and privacy: (1) an SMS text message that stated results were abnormal and requested return to clinic (“direct”), (2) the same message protected by a 4-digit PIN code (“PIN”), and (3) the message “ABCDEFG” explained at enrollment to indicate abnormal results (“coded”). Outcomes of interest were (1) self-reported receipt of the SMS text message, (2) accurate identification of the message, and (3) return to care within 7 days (for abnormal results) or on the date of the scheduled appointment (for normal results). We fit regression models for each outcome with the following explanatory variables: sociodemographic characteristics, CD4 count result, ability to read a complete sentence, ability to access a test message on enrollment, and format of SMS text message.

Results: Seventy-two percent (234/385) of participants successfully receiving a message, 87.6% (219/250) correctly identified the message format, and 60.8% (234/385) returned to clinic at the requested time. Among participants with abnormal tests results (138/385, 35.8%), the strongest predictors of reported message receipt were the ability to read a complete sentence and a demonstrated ability to access a test message on enrollment. Participants with an abnormal result who could read a complete sentence were also more likely to accurately identify the message format (AOR 4.54, 95% CI 1.42-14.47, *P*=.01) and return to clinic appropriately (AOR 3.81, 95% CI 1.61-9.03, *P*=.002). Those who were sent a PIN-protected message were less likely to identify the message (AOR 0.11, 95% CI 0.03-0.44, *P*=.002) or return within 7 days (AOR 0.26, 95% CI 0.10-0.66, *P*=.005). Gender, age, and socioeconomic characteristics did not predict any outcomes and there were no differences in outcomes between those receiving direct or coded messages.

Conclusions: Confirmed literacy at the time of enrollment was a robust predictor of SMS text message receipt, identification, and appropriate response for PLWH in rural Uganda. PIN-protected messages reduced odds of clinic return, but coded messages were as effective as direct messages and might augment privacy.

Trial Registration: Clinicaltrials.gov NCT 01579214; https://clinicaltrials.gov/ct2/show/NCT01579214 (Archived by WebCite at http://www.webcitation.org/6Ww8R4sKq).

(J Med Internet Res 2015;17(3):e78) doi:10.2196/jmir.3859

KEYWORDS

telemedicine; text messaging; randomized controlled trial; Uganda; HIV

Introduction

The promise of mobile phone-based interventions to improve health care delivery in resource-limited settings has been well described [1,2]. The widespread availability of cellular networks coupled with the exponential growth in mobile phone ownership [3] creates an opportunity to leverage limited human resource capacity in resource-limited settings through improved patient-provider communication, information management, clinical decision making, disease surveillance, as well as monitoring and evaluation [2]. Despite a rapid increase in investments in mHealth programs, there are few reports of successfully implemented mHealth interventions in sub-Saharan Africa [4,5].

End-user characteristics that influence acceptance and use patterns of mHealth interventions are critical to successful implementation [6], particularly in settings with variable literacy and technology experience [7,8]. Although there is much literature on the behavioral science of novel technology acceptance and uptake in resource-rich settings [9,10], similar data are largely lacking from resource-limited settings. Data are even sparser on technology acceptability for low-literacy end users. A handful of studies have evaluated the general acceptability of mobile phone-based interventions in these scenarios [11-15], but there is an important need to better understand the barriers and facilitators of technology acceptance for low-literacy populations for whom many mHealth interventions are intended. Of particular importance to patient end users is attention to privacy and confidentiality [11,12,14,16], which can have health and safety implications for stigmatized health conditions, such as human immunodeficiency virus (HIV) infection [17].

We previously reported on the perceived acceptability of a mobile phone app to improve communication of laboratory results to patients at an HIV clinic in southwestern Uganda [12], and on the results of a clinical trial at the same clinic to improve linkage to care through a combination short message service (SMS) text messaging app and transportation reimbursement intervention [18]. In the clinical trial, participants with abnormal results who received SMS text messages about CD4 count results and a transportation reimbursement had significantly improved time to clinic return and time to antiretroviral therapy (ART) initiation than those in a preintervention control period. Here we report results of a prespecified secondary analysis restricted to participants in the intervention period who were sent SMS text messages with the following specific objectives: (1) to identify predictors of self-reported message receipt, accurate identification, and appropriate return to clinic in response to the SMS text messaging app and (2) to evaluate the comparative efficacy of randomly allocated SMS text formats to optimize the balance between confidentiality and accessibility of health-related SMS text communications for low-literacy end users.

Methods

Study Setting and Participants

Data for this analysis were collected as part of a randomized clinical trial of an SMS text app to notify people living with HIV (PLWH) of their laboratory results and request return to care for those with abnormal results. Full details and preliminary results of the trial have been reported previously (NCT01579214) [18]. We present the results of a secondary analysis restricted to those who were sent at least one SMS text message as part of the intervention arm of the clinical trial. The goal of the analysis was to identify predictors of receipt, comprehension, and appropriate response to an SMS text-based clinical communication intervention. Patients were eligible for the study if they were actively in care at the adult HIV clinic of the Mbarara Regional Referral Hospital in Western Uganda, had access to a mobile phone, lived in the clinic catchment area, and were undergoing CD4 count testing. Clinicians specified an abnormal test threshold for each participant, defined as a result that would prompt early return to care for treatment initiation, regimen change, or clinical evaluation.

Intervention Development

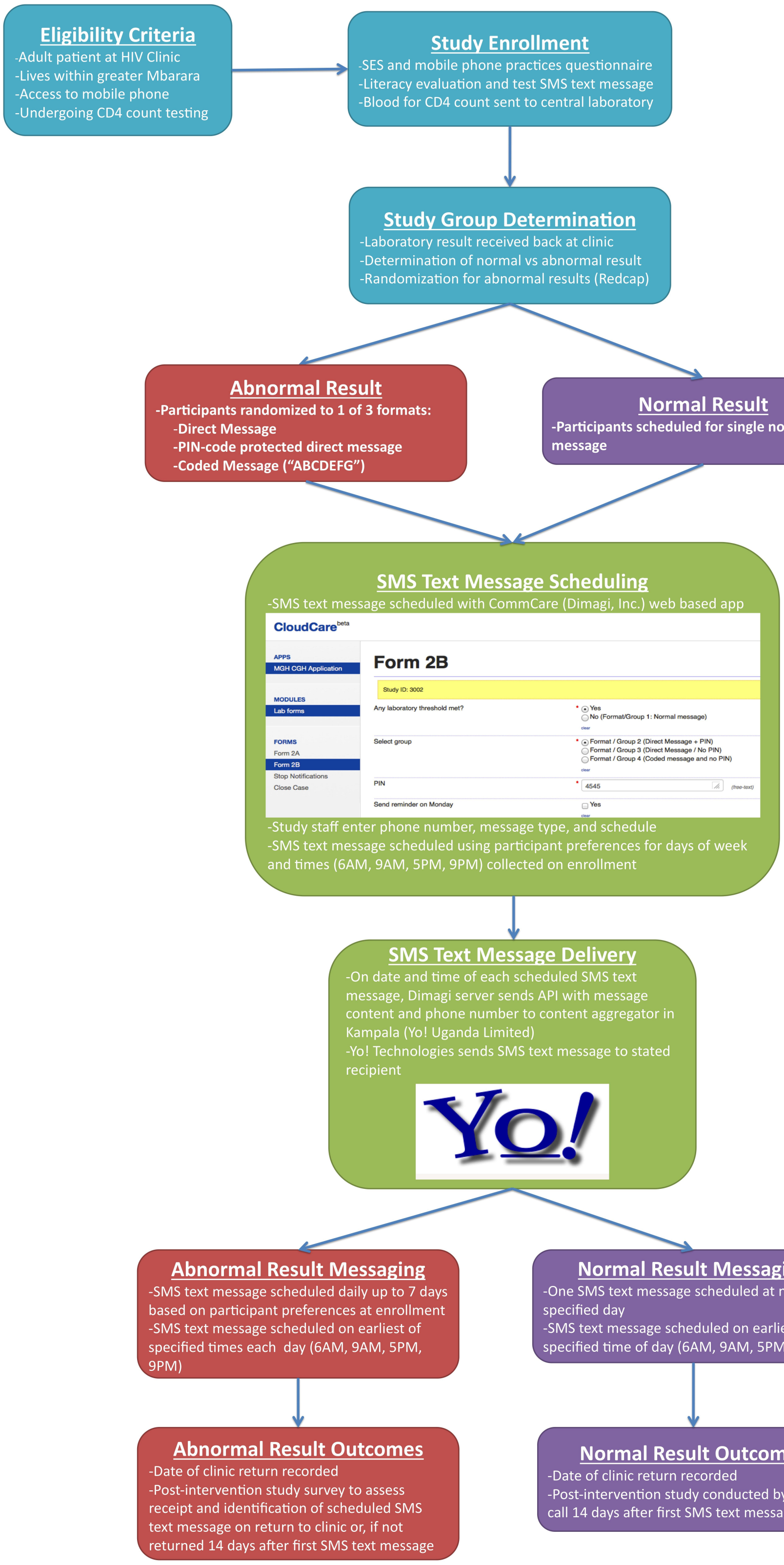
We designed the SMS text-based intervention based on a conceptual framework derived from a preliminary survey of clinic patients, conducted to understand barriers to linkage and acceptability of SMS text message-based health communications [12]. We learned that the two most cited patient-reported barriers to clinic return after abnormal laboratory results were lack of efficient communication with clinical staff and difficulty affording the costs of transportation to clinic [19-21]. Based on this input, we designed a combination intervention to address both of these factors through (1) an SMS text-based communication system to inform patients of their laboratory results and (2) a transportation reimbursement for those with abnormal results if they returned within 7 days of the first message. We involved a multidisciplinary team, including research staff, programmers, clinicians, and patients in development of the SMS text intervention. Key considerations included messaging format that balanced privacy and clarity, and optimization of message timing and frequency. We pilot-tested the intervention with study staff prior to study implementation.

Study Procedures

Participants were approached for enrollment after completion of their clinic visit and blood collection for CD4 count testing. Study staff administered a questionnaire on the day of enrollment to collect data on socioeconomic status and mobile phone use characteristics. As part of the survey, participants were asked to read a complete sentence in the local language (Runyankole). For those who had a mobile phone available on the day of enrollment, a test message was sent and participants were asked to open and read the test message (“ABCDEFG”).

Finally, preferences for receiving SMS text messages were recorded, with options for day of the week and time of day (options included 6 am, 9 am, 5 pm, and 9 pm). Participants were instructed to return to clinic within 7 days of the first abnormal SMS text message. Those who did return within 7 days received a transportation reimbursement (Figure 1).

Figure 1. Study schema.



Form 2B

Study ID: 3002

Any laboratory threshold met? * ● Yes ○ No (Format/Group 1: Normal message)

Select group * ● Format / Group 2 (Direct Message + PIN) ○ Format / Group 3 (Direct Message / No PIN) ○ Format / Group 4 (Coded message and no PIN)

PIN * 4545 (free-text)

Send reminder on Monday ☐ Yes

Study Groups

Per standard clinical protocols, blood samples were sent to an offsite central laboratory where they were processed and a laboratory result form was returned to the clinic. On receipt of test results at the clinic, study staff determined whether test results were below the prestated abnormal threshold to classify each participant as having a normal or abnormal test result. Participants with a normal laboratory result received a single message indicating a normal result and requesting return on the date of their next visit ("Your laboratory result was within the normal range. Please return to clinic on your scheduled date").

Participants with an abnormal result were randomized with the use of the randomization module in Research Electronic Data Capture (REDCap [22]) to 1 of 3 SMS text message formats to evaluate trade-offs between clarity and privacy:

1. Notification that results were abnormal and requesting return to clinic (direct message): "This is an important message from your doctor. You had an abnormal test result. You should return to clinic as soon as possible"
2. An identical message as the direct message, but prompted by an initial requirement to enter a 4-digit personal identification number (PIN message), which was selected by participants on the day of enrollment and given to the patient on a form to take home on the day of enrollment: "Please enter your PIN code to see your message"
3. A coded message which was explained on enrollment to indicate an abnormal test result and signify early return to clinic (coded message): "ABCDEFG"

Participants with abnormal test results were eligible to receive messages daily for up to a maximum of 7 days. The number and timing of messages was determined by their scheduling preferences on enrollment.

Text Message Scheduling and Transmission

Research staff scheduled messages through a Web-based messaging app (CommCareHQ, Dimgi, Inc, Cambridge, MA, USA) on the date of laboratory result receipt. The CommCare app sent an automated application program interface (API) with the mobile phone number of the participants along with the message content to a content aggregator in Kampala (Yo! Uganda Limited, Kampala, Uganda), which relayed the automated SMS text messages to the indicated phone number. All SMS text messages, including both incoming and outgoing, were paid for by the study through use of a short code.

Outcomes Assessment

Study staff recorded the date of clinic return for all participants. Participants who returned within 14 days of the first scheduled message completed an in-person follow-up questionnaire. Research assistants called those who did not return 14 days after the first scheduled message to complete the interview. Questions included whether or not participants received the message, the number of SMS text messages received, and identification of the type of message received.

Statistical Analyses

We used standard data summarization techniques to describe characteristics for the total cohort and by laboratory result subgroups (normal vs abnormal laboratory result). We assessed for predictors of 3 outcomes of interest: (1) reported receipt of at least one SMS text message, (2) accurate identification of the message format delivered, and (3) whether participants returned to care at the appropriate time, defined as within 7 days of the first SMS text message for those with abnormal results or on the date of the scheduled appointment for those with normal results. For predictors of each outcome, we performed stratified analyses by presence or absence of abnormal test results and assessed for statistically significant relationships with the following baseline characteristics: age, gender, CD4 count result, net household income, educational attainment, ability to read a complete sentence, duration of time required to reach the HIV clinic, whether or not they shared their mobile phone with others in the household, ability to access a sample test message on enrollment (for those with a phone available that day), number of messages sent, cellular network used, and—for those with abnormal results—the type of message sent (direct, PIN, or coded). Crude associations between explanatory variables and outcomes were performed with chi-square testing. We assessed for independent predictors of each outcome by fitting multivariable logistic regression model, including age, gender, CD4 count result, literacy, ability to access a test SMS text message on enrollment, and message type. We assessed for an interaction effect between literacy on enrollment and SMS text message format in 2 separate analyses restricted to either those who were sent (1) direct and PIN messages, or those who were sent (2) direct and coded messages. Although we did not include household income, educational attainment, or number of messages sent because they were not significant in any crude analyses, we performed sensitivity analyses with them added to the multivariable models to assess for negative confounding. All statistical analyses were performed with Stata version 13.1 (StataCorp LP, College Station, TX, USA).

Ethical Considerations

The study was reviewed and approved by the ethical review committees of the Mbarara University of Science and Technology, Partners Healthcare, and the Ugandan National Council of Science and Technology. The trial was registered at clinicaltrials.gov (NCT01579214).

Results

We enrolled 385 participants into the trial during the intervention period. The median age of participants was 32 years (IQR 26-39), 65.2% (251/385) were female, more than 60% (240/385, 62.3%) had a primary education or less, median monthly household income was US $80/month (IQR 36-180), and 67.5% (260/385) successfully read a complete sentence on enrollment (Table 1). Nearly half of participants (164/385, 42.6%) shared their mobile phone with others. Of those who had a mobile phone available on the day of enrollment (315/385, 81.8%), approximately three-quarters (247/315, 78.4%) successfully accessed and read the test message. In all, 138 of 385 participants (35.8%) had an abnormal test result. All participants who had a normal result were sent only a single SMS text notification, whereas most with abnormal test results were sent the maximum allowed 7 daily messages (100/138, 72.5%).

Table 1. Cohort characteristics.

Cohort characteristics	Total cohort (N=385)	Abnormal result (n=138)	Normal result (n=247)
Age (years), median (IQR)	32 (26-39)	30 (25-38)	34 (26-4)
Gender (female),n (%)	251 (65.2)	75 (54.3)	176 (71.3)
CD4 count, median (IQR)	409 (272-549)	224.5 (110-293)	504 (419-654)
Estimated household income (US $/month), n (%)			
<40	77 (20.0)	20 (14.5)	57 (23.2)
40-80	81 (21.0)	26 (18.8)	55 (22.3)
80-180	58 (15.1)	30 (21.7)	28 (11.3)
>180	69 (17.9)	26 (18.9)	43 (17.4)
Unknown/unable to estimate	100 (26.9)	36 (26.1)	64 (25.9)
Educational attainment, n (%)			
Less than primary	43 (11.2)	12 (8.7)	31 (12.6)
Primary	197 (51.2)	71 (51.4)	126 (51.2)
Secondary	104 (27.0)	37 (26.8)	67 (27.1)
Beyond secondary	41 (10.6)	18 (13.0)	23 (9.3)
Successfully read a complete sentence in Runyankole at enrollment visit, n (%)	260 (67.5)	97 (70.3)	163 (66.0)
Estimate duration of journey to clinic (minutes), median (IQR)	60 (30-120)	42.5 (20-90)	60 (30-120)
Shared mobile phone, n (%)	164 (42.6)	49 (35.5)	115 (46.6)
Available mobile phone at enrollment visit, n (%)	315 (81.8)	111 (80.4)	204 (82.6)
Successfully accessed test message	247 (78.4)	93 (83.8)	154 (75.5)
Messages sent, n (%)			
1	249 (64.6)	2 (1.4)	247 (100)
2-6	36 (9.4)	36 (26.1)	0
7	100 (26.0)	100 (72.5)	0
Cellular network, n (%)			
MTN	199 (51.6)	68 (49.3)	131 (54.1)
Airtel	111 (28.8)	43 (31.6)	68 (28.1)
Warid	61 (15.8)	23 (16.7)	38 (15.7)
Other	8 (2.1)	3 (2.2)	5 (2.1)
Message type, n (%)			
Normal	247 (64.2)		247 (100)
Direct	46 (11.9)	46 (33.3)	0
PIN	48 (12.5)	49 (35.5)	0
Coded	43 (11.2)	43 (31.2)	0
Study outcomes, n/N (%)			
Received at least 1 SMS text message	250/346 (72.3)	111/138 (80.4)	139/208 (66.8)
Accurately identified transmitted SMS text format	219/250 (87.6)	87/111 (78.4)	132/139 (95.0)
Returned to clinic based on SMS text instructions	234/385 (60.8)	78/138 (56.5)	156/247 (62.8)

For the entire cohort, 72.3% (250/346) reported successful receipt of a message, 87.8% (219/250) of whom correctly identified the message format and 60.0% (231/385) returned to clinic at the requested time. For participants with abnormal tests results, these proportions were 80.4% (111/138), 78.4% (87/111), and 56.5% (78/138), respectively. Although there were no independent predictors in multivariable models for reported receipt of at least one SMS text message, the ability to read a sentence (85%, 82/97 vs 71%, 27/38; *P*=.07) and ability to access and read a test message on enrollment (83%, 77/93 vs 61%, 11/18; *P*=.04) were most closely associated in crude analyses (Table 2). The ability to read a complete sentence on enrollment was independently associated with accurate identification of the message sent (AOR 4.54, 95% CI 1.42-14.47, *P*=.01) and return to clinic within 7 days of the first transmitted SMS text message (AOR 3.81, 95% CI 1.61-9.03, *P*=.002). An ability to access an SMS text message on enrollment was also independently associated with returning to clinic within 7 days of abnormal SMS text notification (AOR 4.90, 95% CI 1.06-22.61, *P*=.04).

In addition to literacy and mobile phone familiarity, SMS text message format was an important predictor of outcomes for participants with abnormal laboratory results. Compared to receipt of a direct message, those with a PIN-protected message were significantly less likely to identify the message sent (AOR 0.11, 95% CI 0.03-0.44, *P*=.002) or return to clinic within 7 days (AOR 0.26, 95% CI 0.10-0.66, *P*=.005) (Table 3). The odds of SMS text message identification and return to clinic were nominally decreased with receipt of a coded versus direct message, but these associations were not statistically significant (Table 3). In restricted analyses comparing either direct versus PIN or direct versus coded messages, we found no statistically significant interaction terms between literacy and message type, suggesting that both literacy and message type were independent predictors of outcomes (Figure 2). Age, gender, household income, educational attainment, and number of messages sent were not associated with any outcomes for participants with abnormal laboratory results in crude analyses. In sensitivity analyses with each of these variables added to the models, we found no substantial differences in our estimates of association with literacy or message format. Lastly, outcomes for those with abnormal test results did not vary meaningfully by telecommunication network used.

We found similar results for predictors of outcomes for participants with normal results (Table 4). Report of receiving an SMS text message was lower for those with normal test results (66.8%, 139/208 vs 80%, 111/138; *P*=.006), whereas those who did receive a message were more likely to appropriately identify the message received (94.9%, 132/139 vs 78.4%, 87/111; *P*<.001). Both an ability to read a sentence on enrollment (74%, 104/140 vs 51%, 35/68; *P*=.001) and the ability to access a test message on enrollment (74%, 96/130 vs 58%, 23/40; *P*=.049) were associated with receipt of a message. Aside from network type, which was associated with accurate identification of a message, we found no other significant predictors of outcomes for those with normal laboratory results.

Table 2. Outcomes for participants with an abnormal laboratory result.

Characteristic	Reported text receipt, n/N (%)	*P*	Accurate message identification, n/N (%)	*P*	Appropriate return to clinic, n/N (%)	*P*
All abnormal result participants	111/138 (80.4)	n/a	87/111 (78.4)	n/a	78/138 (56.5)	n/a
Message type		.81		.004		.045
Direct	38/46 (83)		34/38 (89)		31/46 (67)	
PIN	38/49 (78)		23/38 (60)		21/49 (43)	
Coded	35/43 (81)		30/35 (86)		26/43 (60)	
Age (years)		.64		.50		.98
≤25	26/35 (74)		22/26 (85)		20/35 (57)	
26-32	30/38 (79)		23/30 (77)		22/28 (58)	
33-39	25/30 (83)		21/25 (84)		16/30 (53)	
≥40	30/35 (86)		21/30 (70)		20/35 (57)	
Gender		.57		.73		.89
Female	59/75 (79)		40/52 (77)		42/75 (56)	
Male	52/63 (83)		47/59 (80)		36/63 (57)	
CD4 result		.78		.65		.04
≤100	27/34 (79)		22/27 (81)		25/34 (73)	
101-350	84/103 (82)		65/84 (77)		43/103 (51)	
Estimated household income (US $/month)		.13		.62		.66
<40	14/20 (70)		12/14 (86)		13/20 (65)	
40-80	19/26 (73)		13 /19 (68)		13/26 (50)	
80-180	22/30 (73)		16/22 (73)		15/30 (50)	
>180	24/26 (92)		19/24 (79)		17/26 (65)	
Unknown/unable to estimate	32/36 (89)		27/32 (84)		20/36 (56)	
Educational attainment		.22		.11		.30
Less than primary	7/12 (58)		4/7 (57)		4/12 (33)	
Any primary	59/71 (83)		43/59 (73)		40/71 (56)	
Any secondary	31/37 (84)		28/31 (90)		24/37 (65)	
More than secondary	14/18 (78)		12/14 (86)		10/18 (56)	
Literacy on enrollment		.07		.03		.003
Cannot read a complete sentence	27/38 (71)		17/27 (63)		14/38 (37)	
Reads all of a sentence	82/97 (85)		68/82 (83)		63/97 (65)	
Transportation time to clinic		.21		.32		.32
<30 minutes	30/38 (79)		25/30 (83)		20/38 (53)	
30-59 minutes	28/37 (76)		19/28 (68)		20/37 (54)	
60-119 minutes	29/31 (94)		25/29 (86)		22/31 (71)	
≥129 minutes	24/32 (75)		18/24 (75)		16/32 (50)	
Share mobile phone with household		.79		.87		.41
No	71/89 (80)		56/71 (79)		48/89 (54)	
Yes	40/49 (82)		31/40 (78)		30/49 (61)	
Accessed sample SMS text message on enrollment		.04		.62		.03
No	11/18 (61)		8/11 (73)		6/18 (33)	

Characteristic	Reported text receipt, n/N (%)	*P*	Accurate message identification, n/N (%)	*P*	Appropriate return to clinic, n/N (%)	*P*
Yes	77/93 (83)		61/77 (79)		57/93 (61)	
Messages sent		.45		.70		.84
1-6	29/38 (76)		22/29 (76)		22/38 (58)	
7	82/100 (82)		65/82 (79)		56/100 (56)	
Cellular network		.67		.07		.93
MTN	56/68 (83)		44/56 (79)		38/68 (56)	
Airtel	35/43 (81)		24/35 (69)		23/43 (53)	
Warid	17/23 (74)		17/17 (100)		14/23 (61)	
Other	3/3 (100)		2/3 (67)		2/3 (67)	

Table 3. Multivariable logistic regression models among participants with abnormal test results for predictors of successful receipt of at least one SMS text message, accurate identification of the transmitted SMS text message, and return to clinic within 7 days of message transmission.

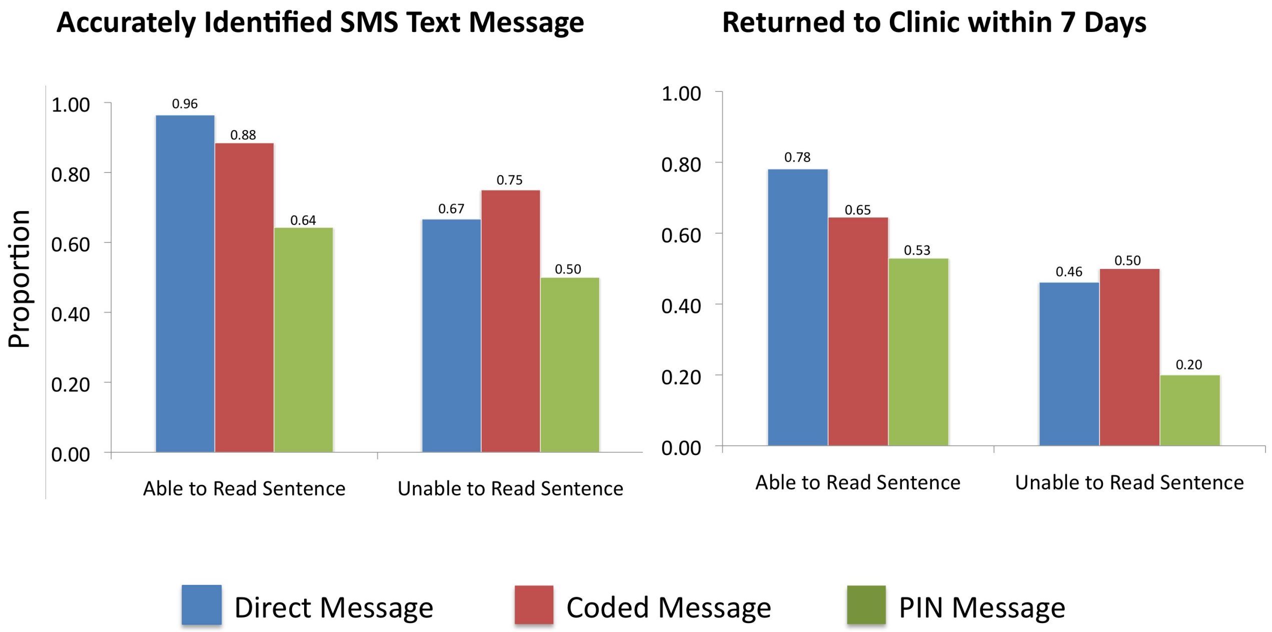
Characteristic	Reported SMS receipt		Accurate SMS identification		Return to clinic <7 days	
	AOR (95% CI)	*P*	AOR (95% CI)	*P*	AOR (95% CI)	*P*
Age (years)						
<26	reference		reference		reference	
26-32	0.97 (0.30-3.13)	.97	0.33 (0.07-1.60)	.17	0.71 (0.24-2.07)	.53
33-39	1.43 (0.40-5.16)	.59	0.98 (0.18-5.27)	.99	0.72 (0.23-2.19)	.56
≥40	1.49 (0.41-5.45)	.54	0.24 (0.05-1.19)	.08	0.66 (0.22-1.95)	.45
Gender (female)	0.95 (0.38-2.37)	.92	1.30 (0.44-3.83)	.63	1.15 (0.52-2.52)	.73
CD4 result						
≤100	reference		reference		reference	
101-350	1.08 (0.39-2.96)	.89	0.51 (0.13-1.96)	.33	0.28 (0.11-0.75)	.01
Read a complete sentence on enrollment	2.14 (0.85-5.39)	.11	4.54 (1.42-14.47)	.01	3.81 (1.61-9.03)	.002
Accessed sample SMS text on enrollment[a]	3.05 (0.76-12.21)	.12	0.63 (0.08-4.68)	.65	4.90 (1.06-22.61)	.04
Randomized SMS text format						
Direct	reference		reference		reference	
PIN	0.76 (0.27-2.17)	.61	0.11 (0.03-0.44)	.002	0.26 (0.10-0.66)	.005
Coded	1.00 (0.31-3.20)	.99	0.38 (0.08-1.80)	.22	0.58 (0.22-1.55)	.28

[a] Restricted to participants with an available mobile phone on enrollment.

Table 4. Outcomes for participants with a normal laboratory result.

Characteristic	Reported SMS receipt, n/N (%)	*P*	Accurate message identification, n/N (%)	*P*	Appropriate return to clinic, n/N (%)	*P*
All normal test result participants	139/208 (66.8)	n/a	132/139 (95.0)	n/a	156/247 (63.2)	n/a
Age (years)		.75		.26		.61
≤25	26/40 (65)		23/26 (88)		34/48 (71)	
26-32	31/42 (74)		31/31 (100)		32/50 (64)	
33-39	37/56 (66)		35/37 (95)		39/66 (59)	
≥40	45/70 (64)		43/45 (96)		51/83 (61)	
Gender		.14		.28		.26
Female	93/146 (64)		87/93 (94)		115/176 (65)	
Male	46/62 (74)		45/46 (98)		41/71 (58)	
Estimated household income (US $/month)		.14		.46		.85
<40	29/48 (60)		26/29 (90)		37/57 (65)	
40-80	34/45 (76)		33/34 (97)		34/55 (62)	
80-180	16/27 (59)		15/16 (94)		17/28 (61)	
>180	28/35 (80)		28/28 (100)		30/43 (70)	
Unknown/unable to estimate	32/53 (60)		30/32 (94)		38/64 (59)	
Educational attainment		.14		.29		.11
Less than primary	15/27 (56)		13/15 (87)		22/31 (71)	
Any primary	64/101 (63)		62/64 (97)		77/126 (61)	
Any secondary	44/61 (72)		41/44 (93)		38/67 (57)	
More than secondary	16/19 (84)		16/16 (100)		19/23 (83)	
Literacy on enrollment		.001		.27		.27
Cannot read a complete sentence	35/68 (51)		32/35 (91)		57/84 (68)	
Reads all of a sentence	104/140 (74)		100/104 (96)		99/163 (61)	
Transportation time to clinic		.36		.95		.37
<30 minutes	32/44 (73)		30/32 (94)		34/48 (71)	
30-59 minutes	32/47 (68)		31/32 (97)		37/54 (69)	
60-119 minutes	38/65 (58)		36/38 (95)		45/77 (58)	
≥129 minutes	37/52 (71)		35/57 (95)		40/68 (59)	
Share mobile phone with household				.34		.67
No	75/116 (75)	.46	70/75 (93)		85/132 (64)	
Yes	64/92 (70)		62/64 (97)		71/115 (62)	
Accessed sample SMS text on enrollment		.049		.52		.17
No	23/40 (58)		21/23 (91)		27/50 (54)	
Yes	96/130 (74)		91/96 (95)		100/154 (65)	
Cellular network		.14		<.001		.99
MTN	84/113 (74)		82/84 (98)		85/131 (65)	
Airtel	35/58 (60)		35/35 (100)		43/68 (63)	
Warid	18/30 (60)		13/18 (72)		24/38 (63)	
Other	2/2 (100)		2/2 (100)		3/5 (60)	

Figure 2. Outcomes by SMS text message format and literacy.



Discussion

Principal Findings

Among PLWH in rural Uganda, confirmed literacy at the time of enrollment, a demonstrated ability to access a sample SMS text message, and absence of a PIN-code protector were robust predictors of receipt, identification, and appropriate response to an SMS text–based laboratory result messaging app. Specifically, PIN-protected messages were poorly accessed and reduced odds of message identification and appropriate clinic return. However, coded messages, which obviate the need for literacy, were as effective as direct messages and might augment privacy. We found no associations between age, gender, educational attainment, household income, or number of messages sent on any outcomes. Many studies have explored the acceptability of SMS text message interventions among PWLH in resource-limited settings, but our study is the first to our knowledge to directly assess the impact of literacy and technology experience on process and clinical outcomes, and demonstrates the importance of such features for mHealth interventions in these settings.

Prior work, largely among high-literacy users in resource-rich settings, has demonstrated that 2 major constructs—perceived ease of use and perceived usefulness—are important predictors of intention to use and use of health technology [9]. Two conceptual models of technology acceptance, the technology acceptance model and the unified theory of technology acceptance and use of technology, have validated the importance of additional mediators, including subjective norms, output quality, and technology experience [23-25]. Studies among patient end users in resource-limited settings are relatively scarce, largely restricted to preintervention assessments, and have demonstrated high potential for acceptance [11,12,14,26,27]. Our study demonstrates that among a patient population with a near 100% acceptability for SMS text-based mHealth interventions [12], rates of SMS text message receipt (73%), accurate SMS text message identification (88%), and appropriate response to SMS text request (63%) were modest. Although we do not report results of ease of use or perceived ease of use directly, our findings that literacy, mobile phone experience, and non-PIN-protected messages all strongly support predictive roles for ease of use as a dominant role for successful technology uptake. Other postintervention assessments in resource-limited settings have demonstrated similar effects. A study assessing technology acceptance of self-service health kiosks in South Africa found that ease of use was the strongest predictor and most correlated with technology anxiety and self-efficacy [28]. A second study from South Africa evaluating acceptability of a mobile phone-assisted personal interview to augment face-to-face maternal health data collection among health care workers found that both ease of use and perceived usefulness improved after a training session [15]. A study of SMS text-based antiretroviral adherence reminders for caregivers of pediatric patients demonstrated low uptake of PIN-protected messages [8]. We attempted to mitigate the complexity of PIN-based messages by allowing participant-selected codes and giving participants a form with the code, but we found similarly poor results for PIN-based messages. Finally, an analysis to determine predictors of response to an SMS text-based ART adherence reminder intervention demonstrated that higher educational attainment predicted improved adherence [29].

Because an important minority of patients in a prestudy interview reported confidentiality concerns about receiving clinic-related information by SMS text message [12], we attempted to assess various message formats to evaluate trade-offs between privacy and clarity. Whereas PIN code-protected messages were negatively associated with successful outcomes in our study, we were able to demonstrate the feasibility of protecting patient confidentiality through use of non-text messages without need for PIN-code protection. Participants who received a message stating “ABCDEFG” had similar rates of message receipt, identification, and early clinic return times as those who received a full instructional text message. Although this method shows promise, it is limited to transmission of qualitative messages (eg, “yes” or “no” information). Prior studies of mHealth interventions have attempted alternative strategies including not mentioning the words “HIV” or “ART” [30] or using nonspecific greetings [31]. One prior study, which did specifically use the term “medication” in an SMS text-based adherence reminder, found no effect of the intervention [32].

Our results have important implications for future mHealth interventions targeted to low-literacy end users. First, text messaging was broadly successful in a rural, resource-limited population with limited education, suggesting that age, gender, educational attainment, and income should not be used as screening criteria for SMS text message interventions. Second, thorough assessments of end-user written literacy and technology experience should be made before and during implementation design. Third, we found that in-person confirmation of mobile phone competency was highly predictive and should be considered for future similar interventions where possible. Fourth, we demonstrate that coded messages can have similar efficacy as text messages, while maintaining confidentiality. Importantly, our study involved transmitting qualitative information (ie, normal or abnormal laboratory results). More complex instructional messages or quantified information will present additional challenges that should be explored in future studies through use of pictorial or other symbolic message formats. Finally, we observed increased rates of reported message receipt among those in the abnormal results group who mostly received 7 daily messages (median 7, IQR 5-7) compared to those in the normal results group who only received a single message, suggesting that repeated messages might increase successful transmission. In contrast, prior work has showed that weekly messages might be modestly preferably to daily messages [30]. However, our study involved only a single notification, as opposed to prior adherence studies that transmitted SMS text message reminders for up to a year.

Limitations

Our study was conducted at a single clinical site with a highly impoverished and low-literacy population. Although this limits the generalizability of our study, it also adds important data about a study population in a low resource setting, which is often the target of mHealth interventions. Our study would have also benefited from evaluation of additional message formats. For example, interactive voice response or direct voice call groups would have added important comparative information; however, they might have challenged the scalability of the intervention.

Future Work

We are pursuing further activities to build on these results. Specifically, we are conducting postintervention qualitative interviews to collect in-depth accounts about ease of use and usefulness of the information, as well as barriers and promoters of uptake of the intervention. Lastly, we have partnered with the clinic data managers and faculty members in computer science at the Mbarara University of Science and Technology to implement an SMS text–based reminder and results messaging system clinic-wide. We are planning a second evaluation following implementation to learn about large-scale effectiveness and scalability.

Conclusions

In summary, we demonstrate that end-user characteristics, particularly literacy and technology experience, are important predictors of an mHealth intervention for PLWH in rural Uganda. We also demonstrate that, although PIN code–protected messages decrease the efficacy of SMS text message information delivery in this population, privacy can be maintained through coded messaging. Future SMS text–based interventions for low-literacy users in similar settings should consider these factors in design and implementation of mHealth interventions. Further evaluation of technology acceptance in this population and similar ones is needed if the potential of mHealth in sub-Saharan Africa is to be realized.

Acknowledgments

This study was funded by the National Institutes of Health (R24 TW007988, T32 AI007433, K23 MH099916), and the Harvard Global Health Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

References

1. Qiang C, Yamamichi M, Hausman V, Altman D. Mobile applications for the health sector. Washington, DC: World Bank; 2011 Dec. URL: http://siteresources.worldbank.org/INFORMATIONANDCOMMUNICATIONANDTECHNOLOGIES/Resources/mHealth_report.pdf [accessed 2015-03-11] [WebCite Cache ID 6Ww8gQfiL]
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. The World in 2013: ICT Facts and Figures. Geneva: International Telecommunication Union; 2013. URL: http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf [accessed 2014-08-10] [WebCite Cache ID 6WkxjqyfK]
4. 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]
5. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. J Health Commun 2012;17 Suppl 1:82-104. [doi: 10.1080/10810730.2011.649160] [Medline: 22548603]
6. 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]
7. Andreatta P, Debpuur D, Danquah A, Perosky J. Using cell phones to collect postpartum hemorrhage outcome data in rural Ghana. Int J Gynaecol Obstet 2011 May;113(2):148-151. [doi: 10.1016/j.ijgo.2010.11.020] [Medline: 21420086]
8. Haberer JE, Kiwanuka J, Nansera D, Wilson IB, Bangsberg DR. Challenges in using mobile phones for collection of antiretroviral therapy adherence data in a resource-limited setting. AIDS Behav 2010 Dec;14(6):1294-1301 [FREE Full text] [doi: 10.1007/s10461-010-9720-1] [Medline: 20532605]
9. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
10. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS quarterly 2003;27(3):425-477.
11. Jennings L, Ong'ech J, Simiyu R, Sirengo M, Kassaye S. Exploring the use of mobile phone technology for the enhancement of the prevention of mother-to-child transmission of HIV program in Nyanza, Kenya: a qualitative study. BMC Public Health 2013;13:1131 [FREE Full text] [doi: 10.1186/1471-2458-13-1131] [Medline: 24308409]
12. Siedner MJ, Haberer JE, Bwana MB, Ware NC, Bangsberg DR. High acceptability for cell phone text messages to improve communication of laboratory results with HIV-infected patients in rural Uganda: a cross-sectional survey study. BMC Med Inform Decis Mak 2012;12:56 [FREE Full text] [doi: 10.1186/1472-6947-12-56] [Medline: 22720901]
13. Tomlinson M, Solomon W, Singh Y, Doherty T, Chopra M, Ijumba P, et al. The use of mobile phones as a data collection tool: a report from a household survey in South Africa. BMC Med Inform Decis Mak 2009;9:51 [FREE Full text] [doi: 10.1186/1472-6947-9-51] [Medline: 20030813]
14. Chang LW, Njie-Carr V, Kalenge S, Kelly JF, Bollinger RC, Alamo-Talisuna S. Perceptions and acceptability of mHealth interventions for improving patient care at a community-based HIV/AIDS clinic in Uganda: a mixed methods study. AIDS Care 2013;25(7):874-880 [FREE Full text] [doi: 10.1080/09540121.2013.774315] [Medline: 23452084]
15. van HA, Norris S, Tollman S, Richter L, Rotheram-Borus MJ. Collecting maternal health information from HIV-positive pregnant women using mobile phone-assisted face-to-face interviews in Southern Africa. J Med Internet Res 2013;15(6):e116 [FREE Full text] [doi: 10.2196/jmir.2207] [Medline: 23748182]
16. Leon N, Schneider H, Daviaud E. Applying a framework for assessing the health system challenges to scaling up mHealth in South Africa. BMC Med Inform Decis Mak 2012;12:123 [FREE Full text] [doi: 10.1186/1472-6947-12-123] [Medline: 23126370]
17. Tsai AC, Bangsberg DR, Kegeles SM, Katz IT, Haberer JE, Muzoora C, et al. Internalized stigma, social distance, and disclosure of HIV seropositivity in rural Uganda. Ann Behav Med 2013 Dec;46(3):285-294 [FREE Full text] [doi: 10.1007/s12160-013-9514-6] [Medline: 23690283]
18. Siedner M, Santorino D, Lankowski A, Kanyesigye M, Bwana M, Haberer J, et al. An SMS intervention to improve HIV linkage to care: a randomized, comparative effectiveness trial. 2014 Presented at: 21st Conference on Retrovirusesand Opportunistic Infections; March 2014; Boston, MA.
19. Ware NC, Idoko J, Kaaya S, Biraro IA, Wyatt MA, Agbaji O, et al. Explaining adherence success in sub-Saharan Africa: an ethnographic study. PLoS Med 2009 Jan 27;6(1):e11 [FREE Full text] [doi: 10.1371/journal.pmed.1000011] [Medline: 19175285]
20. Siedner MJ, Lankowski A, Tsai AC, Muzoora C, Martin JN, Hunt PW, et al. GPS-measured distance to clinic, but not self-reported transportation factors, are associated with missed HIV clinic visits in rural Uganda. AIDS 2013 Jun 1;27(9):1503-1508 [FREE Full text] [doi: 10.1097/QAD.0b013e32835fd873] [Medline: 23435294]
21. Lankowski AJ, Siedner MJ, Bangsberg DR, Tsai AC. Impact of geographic and transportation-related barriers on HIV outcomes in sub-Saharan Africa: a systematic review. AIDS Behav 2014 Jul;18(7):1199-1223. [doi: 10.1007/s10461-014-0729-8] [Medline: 24563115]

22. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
23. Venkatesh V, Thong J, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS quarterly 2012;36(1):157-178.
24. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
25. Nakamura A, Osonoi T, Terauchi Y. Relationship between urinary sodium excretion and pioglitazone-induced edema. J Diabetes Investig 2010 Oct 19;1(5):208-211 [FREE Full text] [doi: 10.1111/j.2040-1124.2010.00046.x] [Medline: 24843434]
26. Mitchell KJ, Bull S, Kiwanuka J, Ybarra ML. Cell phone usage among adolescents in Uganda: acceptability for relaying health information. Health Educ Res 2011 Oct;26(5):770-781 [FREE Full text] [doi: 10.1093/her/cyr022] [Medline: 21536715]
27. Crankshaw T, Corless IB, Giddy J, Nicholas PK, Eichbaum Q, Butler LM. Exploring the patterns of use and the feasibility of using cellular phones for clinic appointment reminders and adherence messages in an antiretroviral treatment clinic, Durban, South Africa. AIDS Patient Care STDS 2010 Nov;24(11):729-734. [doi: 10.1089/apc.2010.0146] [Medline: 21039181]
28. Cohen JF, Sergay SD. An empirical study of health consumer beliefs, attitude and intentions toward the use of self-service kiosks. 2011 Presented at: American Conference on Information Systems (AMCIS); August 4-8, 2011; Detroit, MI.
29. van der Kop ML, Karanja S, Thabane L, Marra C, Chung MH, Gelmon L, et al. In-depth analysis of patient-clinician cell phone communication during the WelTel Kenya1 antiretroviral adherence trial. PLoS One 2012;7(9):e46033 [FREE Full text] [doi: 10.1371/journal.pone.0046033] [Medline: 23049928]
30. Pop-Eleches C, Thirumurthy H, Habyarimana JP, Zivin JG, Goldstein MP, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. AIDS 2011 Mar 27;25(6):825-834 [FREE Full text] [doi: 10.1097/QAD.0b013e32834380c1] [Medline: 21252632]
31. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
32. Mbuagbaw L, Thabane L, Ongolo-Zogo P, Lester RT, Mills EJ, Smieja M, et al. The Cameroon Mobile Phone SMS (CAMPS) trial: a randomized trial of text messaging versus usual care for adherence to antiretroviral therapy. PLoS One 2012;7(12):e46909 [FREE Full text] [doi: 10.1371/journal.pone.0046909] [Medline: 23236345]

Abbreviations

API: application program interface
ART: antiretroviral therapy
HIV: human immunodeficiency virus
PIN: personal identification number
PLWH: people living with HIV
SMS: short message service

Edited by G Eysenbach; submitted 12.09.14; peer-reviewed by W Nilsen, W Brown; comments to author 23.10.14; revised version received 24.10.14; accepted 29.10.14; published 24.03.15.

Please cite as:
Siedner MJ, Santorino D, Haberer JE, Bangsberg DR
Know Your Audience: Predictors of Success for a Patient-Centered Texting App to Augment Linkage to HIV Care in Rural Uganda
J Med Internet Res 2015;17(3):e78
URL: http://www.jmir.org/2015/3/e78/
doi:10.2196/jmir.3859
PMID:25831269

Original Paper

Wikipedia and Medicine: Quantifying Readership, Editors, and the Significance of Natural Language

James M Heilman[1], BSc, MD, CCFP(EM); Andrew G West[2], PhD, MS Eng

[1]Faculty of Medicine, Department of Emergency Medicine, University of British Columbia, Vancouver, BC, Canada
[2]Verisign Labs (Verisign, Inc.), Reston, VA, United States

Corresponding Author:
James M Heilman, BSc, MD, CCFP(EM)
Faculty of Medicine
Department of Emergency Medicine
University of British Columbia
2194 Health Sciences Mall, Unit 317
Vancouver, BC, V6T1Z3
Canada
Phone: 1 4158306381
Fax: 1 6048226061
Email: jmh649@gmail.com

Abstract

Background: Wikipedia is a collaboratively edited encyclopedia. One of the most popular websites on the Internet, it is known to be a frequently used source of health care information by both professionals and the lay public.

Objective: This paper quantifies the production and consumption of Wikipedia's medical content along 4 dimensions. First, we measured the amount of medical content in both articles and bytes and, second, the citations that supported that content. Third, we analyzed the medical readership against that of other health care websites between Wikipedia's natural language editions and its relationship with disease prevalence. Fourth, we surveyed the quantity/characteristics of Wikipedia's medical contributors, including year-over-year participation trends and editor demographics.

Methods: Using a well-defined categorization infrastructure, we identified medically pertinent English-language Wikipedia articles and links to their foreign language equivalents. With these, Wikipedia can be queried to produce metadata and full texts for entire article histories. Wikipedia also makes available hourly reports that aggregate reader traffic at per-article granularity. An online survey was used to determine the background of contributors. Standard mining and visualization techniques (eg, aggregation queries, cumulative distribution functions, and/or correlation metrics) were applied to each of these datasets. Analysis focused on year-end 2013, but historical data permitted some longitudinal analysis.

Results: Wikipedia's medical content (at the end of 2013) was made up of more than 155,000 articles and 1 billion bytes of text across more than 255 languages. This content was supported by more than 950,000 references. Content was viewed more than 4.88 billion times in 2013. This makes it one of if not the most viewed medical resource(s) globally. The core editor community numbered less than 300 and declined over the past 5 years. The members of this community were half health care providers and 85.5% (100/117) had a university education.

Conclusions: Although Wikipedia has a considerable volume of multilingual medical content that is extensively read and well-referenced, the core group of editors that contribute and maintain that content is small and shrinking in size.

(J Med Internet Res 2015;17(3):e62) doi:10.2196/jmir.4069

KEYWORDS

health information systems; consumer health information; information sharing; information networks; information science; Internet; Web 2.0; cooperative behavior

Introduction

Wikipedia is a multilingual, online, open-source encyclopedia that anyone with Internet access can edit. It is available in more than 275 languages and contains more than 32 million articles across a tremendously broad topic space [1]. Although a considerable amount is known about the volume of content, readership, and editor population of Wikipedia as a whole, less is known about these aspects as they pertain to Wikipedia articles in the medical domain. Moreover, non-English language editions are dramatically understudied in comparison to the larger and more popular English version.

In January of 2014, Wikipedia was referred to as “the single leading source of medical information for patients and health care professionals” by the Institute of Medical Science (IMS) Institute for Healthcare Informatics [2]. It is used as a source of health care information by 50% to 70% of physicians [3,4] and has been reported as being the single most used resource by medical students (94%) [5]. A 2013 US survey found people spend more than 52 hours a year searching for health information online, with 22% reporting using Wikipedia [6]. Wikipedia’s readership is also affected by current events, including popular culture [7] or disease outbreaks [8,9]. Because Wikipedia’s health content is extensively read by the general public and in communities of practice, its authorship and reliability are important qualities. Additionally, quantifying topic popularity can help focus improvements toward greater impact.

With respect to measures of quality, the small amount of available research came to differing conclusions [10]. In 2 small samples, Wikipedia’s accuracy was found to be similar to that of UpToDate, eMedicine, and the National Cancer Institute’s Physician Data Query (PDQ) comprehensive cancer database [10]. A narrow look at pharmacological articles assessed Wikipedia’s accuracy to be high based on significant overlap with textbook sources [11]. Other research found a selection of 50 English medical articles to be relatively well cited [12]. Since 2010, the number of health science academic articles using Wikipedia as a citation has increased substantially [13]. Differing research has found Wikipedia’s coverage to be incomplete or less than that of professional sources [10]. A paper examining gastroenterology articles from 2013 found insufficient discussion of the mechanisms of disease [14]. A comparison of pediatric otolaryngology articles between Wikipedia, MedlinePlus, and eMedicine found Wikipedia had a similar accuracy to MedlinePlus, but less than that of eMedicine [15].

In our subsequent analysis, we will report on the amount of medical content on Wikipedia. This includes determining the number of references supporting this content and how this quantity has changed over the past 5 years. Readership for both English and non-English versions in 2013 will be analyzed, along with an attempt to determine how the popularity of Wikipedia’s medical content compares to that of other well-known Internet health care sites. We will determine if the most commonly viewed articles are those that cover major global health problems or more obscure ones. Finally, the size and makeup of the core editor community will be examined, including how this has changed since 2009.

Methods

Amount of Wikipedia Medical Content

To quantify the number of medical articles and the amount of content within them, one must first determine the subset of Wikipedia which is medically relevant. Wikipedia has a category hierarchy that is built collaboratively, similar to how its core content is amassed and refined. These categories are the basis for identifying medical articles, drawn from the tagging work of WikiProject Medicine [16], which identified those English articles that fall within its project’s scope.

Examples of medical articles include medical diseases and syndromes, medical procedures and diagnostic tests, medications and drugs, and articles related to the history of medicine. Some fitness, pathogenic, and microbiology topics are also categorized as medical; notable health care workers also often meet the threshold. However, articles for anatomy, individuals with specific conditions, pharmaceutical companies, and hospitals tend not to be categorized as “medical” because they are usually well covered by other projects [17].

To identify non–English language equivalents for English articles we relied on the interlanguage link infrastructure. Also collaboratively built, these links build a graph of all articles—across all language editions—corresponding to a shared topic. Before 2013, these links were annotated in the articles themselves in a distributed fashion. Throughout 2013, these links were migrated to a centralized location (WikiData) for ease of maintenance. When we measured the amount of content (in bytes) we accounted for this migration otherwise it would appear articles were losing content when, in fact, duplicate content was just being more efficiently stored.

Determining the size of a language’s medical article membership was straightforward aggregation. Our analysis reports only on article content, not the discussion or policy-based pages that surround it. Programmatic access to category and interlanguage data are available via the Wikimedia application program interface (API) [18]. That same API permitted us to obtain an article’s full content at any historical timestamp. We used snapshots from start-2013 and end-2013 to plot the byte growth of medical content, measuring only textual content in this manner.

Citations Supporting Wikipedia’s Medical Content

One marker to estimate the quality of Wikipedia’s content is the number of references present in articles and the reputation of those referenced sources. Leveraging the ability to obtain an article’s full content at any timestamp, we parsed that content for standardized citation templates (ie, the “<ref>” and “{{cite}}” notations). Counting template usage is straightforward and article snapshots at end-of-year 2009-2013 were used to analyze longitudinal trends. The citation templates also contained a “source” field. We used this to analyze the relative citation counts of leading medical journals, bearing in mind that nonstandardized naming and abbreviation conventions (eg, New England Journal of Medicine, NE Journal of Medicine, NE J Med) inhibit precise aggregation. In particular, we highlighted citations to Cochrane reviews because they are a highly regarded source. Parsers based on regular expressions were used in reference counting and source extraction.

Readership of Wikipedia's Medical Content

Readership of specific articles and medical content in total were derived from the hourly page view aggregates [19] made available by the Wikimedia Foundation (WMF). These are large plaintext files in which each line contains a language, article title, and view count—with a single day's volume (24 files) on the order of 10 GB in size. We authored scripts to obtain and process these files nightly, writing daily aggregates to a persistent database table indexed by language and article.

These files report only "desktop" views. However, mobile views were reported at project-scale [1,20] (eg, for all of English/French/Spanish Wikipedia), permitting some rough estimates if one assumed mobile traffic was uniformly proportional across all articles. An examination of the phenomena underlying this collection and broader readership trends was done by West [7].

Our database of daily views can be queried to produce aggregates by language, specific article, or the topics that span multiple language equivalents. To compare Wikipedia's medical readership to that of other common health care websites, we used SimilarWeb [21], a traffic measurement service. We multiplied the "estimated visitors" and "page views per visit" metrics that the service provides to produce a page views statistic comparable to the one reported by the WMF.

To measure topic readership variance between languages we identified a core set of equivalent articles that existed in all Wikipedia's 10 largest language editions. We first analyzed these by topic, finding anomalous popularity patterns and outliers. For an aggregate comparison, we also calculated the Pearson correlation coefficient between all language pairs.

We also wanted to determine if diseases of greater global severity were more frequently viewed Wikipedia topics. To do so, we took the top 20 diseases by disability adjusted life years (DALYs) and the top 20 diseases by years lived with disability (YLDs) for 2012 as reported by the World Health Organization [22], yielding 33 conditions in combination. We then found the 42 corresponding English Wikipedia articles for each disease (some, such as "child behavioral disorders" referred to both "ADHD" and "conduct disorder"). Traffic on these articles was compared against that on a broader set of Wikipedia articles corresponding to diseases, as identified by the presence of a standardized template ("infobox") that concisely summarized disease metadata (eg, a condition's index in various disease databases).

Quantity/Characteristics of Wikipedia's Medical Contributors

Already leveraged for categories and language links, the Wikimedia API also permits one to crawl version histories to gather metadata about an article's editors. Aggregating this across all medical articles (or just those of a particular language), we were able to plot participation at various thresholds. In particular, we identified 274 contributors who made more than 250 edits to medical articles in 2013. In May 2014, we utilized a Wikipedia messaging system to award 271 of these users a "barnstar," a digital form of peer-to-peer recognition. Posted to users' talk pages, the awards contained a request to complete a survey containing 6 questions:

1. What is your highest level of education?
2. Do you currently work in the health care field? Or have you previously?
3. Are you currently studying health care (a student)?
4. What language of Wikipedia do you mostly work on?
5. Did you receive a barnstar?
6. How do you identify your gender?

Question #5 was used to sanity check respondents (because barnstar awards are public, uninvited participants could traverse the survey link). We also posed an open question: "Why do you edit Wikipedia's medical content?"

Results

Amount of Wikipedia Medical Content

Number of Articles

Wikipedia had 155,805 medical articles across 255 natural languages at the end of 2013. A further 31 languages did not contain any medical articles per our methodology. Of the more than 155,000 articles, 29,072 (18.66%) were in English. Although a significant portion of Wikipedia's content (both medical and otherwise) is in English, this imbalance is less than that observed across the broader Internet (Figure 1). In Figure 1, the "world by language" subgraph was based on 2007-2010 data per the aggregation of the Wikipedia community [23], "Internet by language" was derived per W^{3}Techs Web Technology Surveys [24], and "Wikipedia by language (medical portions)" was based on independent calculations of medical articles by language edition. Note that for all independent calculations/figures/graphs presented in this paper, Multimedia Appendix 1 presents raw data and/or extends those presentations.

Table 1 presents the top languages by quantity of medical articles. Going beyond this list, the top 10 languages made up 51.37% (80,043/155,805) of the total articles, whereas the top 25 languages accounted for 74.97% (116,808/155,805). Figure 2 plots the article quantity distribution, showing it to have a power-law distribution (ie, few languages have many articles and vice versa).

Table 1. Wikipedia language editions ranked by number of medical articles and the amount of textual content in each language (in bytes).

Rank	Language	Medical articles, n	Size (MB)
1	English	29,072	241
2	German	7761	66
3	French	6372	57
4	Spanish	6367	51
5	Polish	5999	28
6	Italian	5677	44
7	Portuguese	5269	25
8	Russian	4832	48
9	Dutch	4391	18
10	Japanese	4303	36
11	Arabic	4055	26
12	Swedish	3661	12

Figure 1. Relative amount of population/content by natural language group. The 10 European languages are German, French, Spanish, Polish, Italian, Portuguese, Russian, Dutch, Swedish, and Catalan.

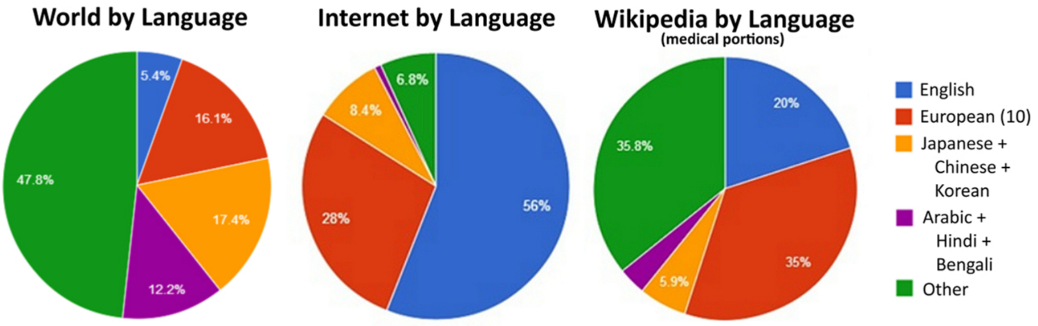
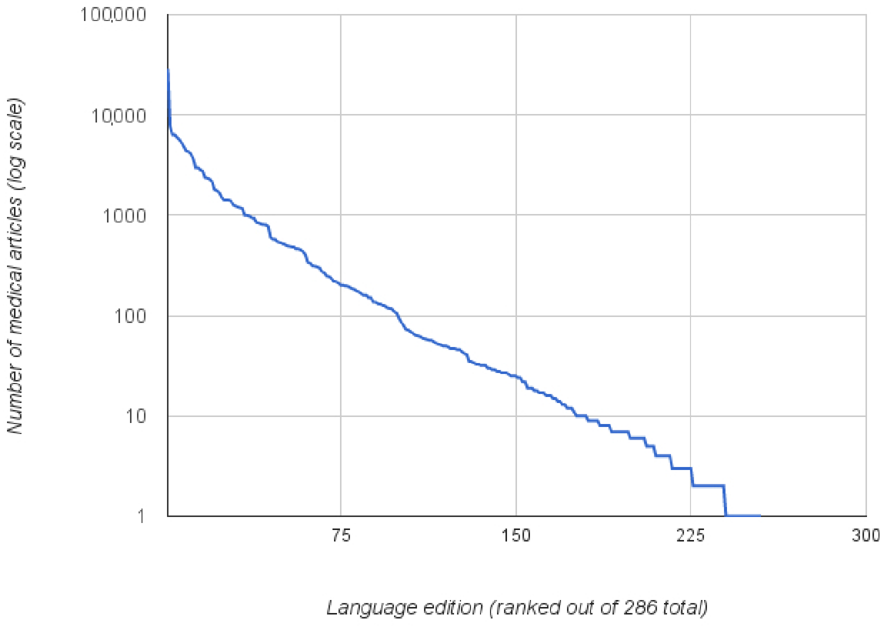


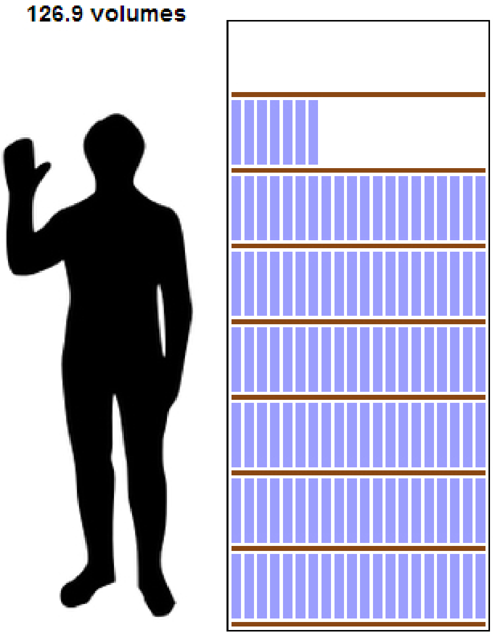
Figure 2. Distribution for the quantity of medical articles in a Wikipedia language edition presented in rank order (note log scale on y-axis).



Bytes of Content

At the end of 2013, Wikipedia had 1016 MB of textual medical content, up 10.19% from 1 year earlier when the total was 922 MB. English medical articles saw the most growth during this period, gaining some 19.7 MB. Assuming the average word has 6 characters, this equates to 3.28 million English words added in 2013. If the total (combined language) 1016 MB of content were printed in textbooks roughly the size of the *Encyclopedia Britannica* at 8 million characters per volume, it would consume 126.9 volumes (Figure 3). English-language medical articles were responsible for 23.72% (241/1016 MB) of all medical content (by bytes). The next largest languages per this metric were German, French, Spanish, Russian, Italian, Japanese, Polish, Arabic, and Portuguese (similar but not identical to Table 1). Together the top 10 languages accounted for 61.22% (622/1016 MB) of all byte content.

Figure 3. Estimated volume of Wikipedia's medical content if printed (attribution of human outline: Linda Salzman Sagan).



Citations Supporting Wikipedia's Medical Content

As a marker for Wikipedia's reliability, we counted the number of references in year-end article versions between 2009 and 2013. This was done for medical portions of both English Wikipedia and all languages (Figure 4). We found that English references more than doubled from 187,107 to 376,123, whereas the increase was more than 2.5 times from 373,558 to 952,053 across all languages. Note that this citation growth ratio significantly outpaced that observed for byte growth.

By parsing a standardized citation format, we were able to determine the journals that were most commonly used as references on Wikipedia were also some of the most respected, including *The Lancet*, *The New England Journal of Medicine*, *Nature*, *British Medical Journal*, *JAMA*, *Science*, and the Cochrane Database of Systematic Reviews. Although a lack of standardized naming/abbreviation conventions prevented precise aggregation, we were able to measure references to a high-quality source. Plain text and citation references to "Cochrane (reviews)" (Figure 5) across all languages increased nearly 3-fold from 2717 in 2009 to 7290 in 2013.

Figure 4. Citations/references appearing in the medical content of English Wikipedia and all Wikipedia languages based on year-end snapshots.

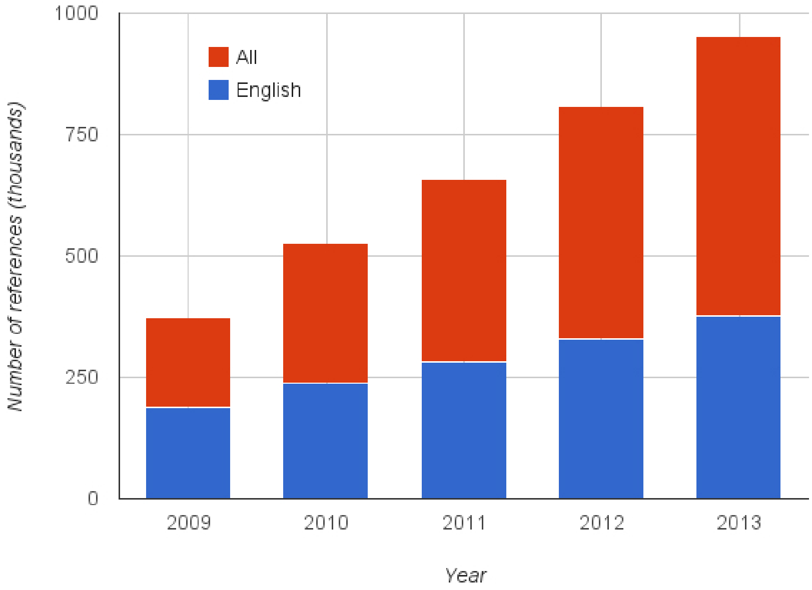
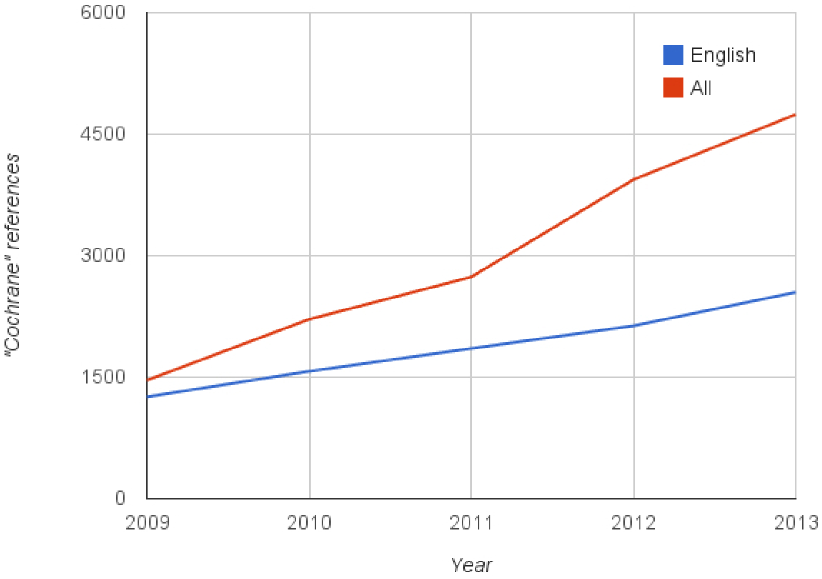


Figure 5. References to “Cochrane (reviews)” in medical content of English Wikipedia and all Wikipedia languages in both plain text and citation formats.



Readership of Wikipedia’s Medical Content

Comparison Between Wikipedia and Other Health Care Websites

Before embarking on traffic comparisons between Wikipedia and other health care sites, we first established Wikipedia’s medical readership in isolation. In 2013, across all languages, Wikipedia’s medical content received 4.88 billion nonmobile views (estimates put the mobile-inclusive total close to 6.5 billion). Approximately 4.56 billion of these were in the top 12 languages (Table 2), with English accounting for 46.72% (2.28/4.88 billion views).

Table 2. Languages sorted by millions of page views to medical content in 2013 and percentage of medical views out of all language views.

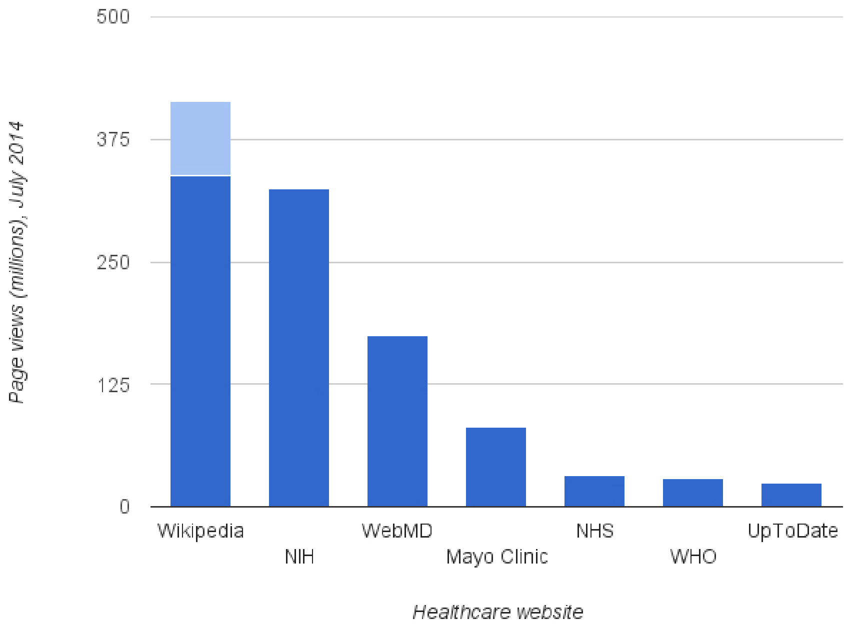
Language	Medical page views (million), n	All language views (million), n (%)
English	2277	91,252 (2.49)
Spanish	659	14,806 (4.45)
German	348	11,067 (3.15)
Japanese	254	12,535 (2.02)
French	219	8305 (2.63)
Portuguese	212	5266 (4.03)
Russian	150	12,072 (1.24)
Polish	147	5691 (2.59)
Italian	147	5738 (2.56)
Dutch	76	2198 (3.44)
Chinese	36	3775 (0.96)
Turkish	34	1646 (2.06)

Medical content accounted for 0.64% (0.029/4.5 million) of all articles on English Wikipedia, yet these received 2.49% (2277/91,252 million) of all English Wikipedia page views. Similar patterns were observed across many language editions, with medical articles receiving far more than the mean expected traffic. As a portion of all content, among prominent languages, medical readership varied from 0.96% (36/3775 million) in Chinese to 4.45% (658/14,806 million) in Spanish; the global percentage across all languages was 2.50% (4.88/195 billion), roughly the same as for English.

Recall that we used the Web monitoring service SimilarWeb [21] to estimate the traffic received at other health care websites. Despite having precise page view data for Wikipedia's medical portions, in the interest of fairness, we also derived Wikipedia's totals from SimilarWeb. That service's sampling methodology likely introduces bias we would prefer to be uniform across all sites under evaluation. The health care sites we examined (National Institutes of Health, WebMD, Mayo Clinic, National Health Service, World Health Organization, UpToDate) host exclusively medical content. In contrast, the traffic statistics SimilarWeb reports for the Wikipedia domain must be scaled down to its medical portion (2.49%).

Figure 6 presents the comparison after such adjustments for July 2014, with the light blue portion capturing that SimilarWeb slightly underreports traffic compared to the WMF data (recalling that neither reports mobile views). Regardless, Wikipedia appears to be the most utilized online health care information resource.

Figure 6. Health care site traffic comparison. Light blue portions represent official Wikimedia Foundation data.



Comparison Among Wikipedia’s Natural Language Editions

The popularity of individual topics across languages varied dramatically. Among the 100 most popular English articles, none were unanimously in the top 100 across 9 other prominent languages (compromising the 10 most popular languages by overall page views in 2013) in which a corresponding article existed. For example, “Down syndrome” was third most popular in German, seventh most accessed in Italian/Polish, and 17th in English. However, it was outside the top 1000 in Russian, Japanese, French, Portuguese, and Chinese. “Asperger syndrome” was 1 of few articles close to being in the top 100 most viewed in all languages, but was nearly 1500th in Russian. Similarly, “tuberculosis” fared well in all languages except French and Polish. “Sexual intercourse”—a typical stronghold of Internet attention—was only in the top 10 most popular articles for English, where it secured third place. Table 3 presents the most popular topics overall and Table 4 further highlights popularity variance.

Table 3. Medical topics with the most traffic summed across languages. View count is for 2013 and the number of languages with a corresponding article is presented.

Topic/Article	Views (million), n	Languages, n
Leonardo da Vinci	17.36	180
Asperger syndrome	15.48	56
Schizophrenia	13.84	74
Bipolar disorder	13.12	55
Sexual intercourse	13.07	96
Tuberculosis	11.86	123
Diabetes mellitus	11.28	113
Autism	10.43	80
HIV/AIDS	9.65	128
Angelina Jolie[a]	9.51	98
Human	9.47	165
Hemorrhoid	9.42	73
Pneumonia	9.26	90
Blood type	9.20	67
Human papillomavirus	8.97	39
Down syndrome	8.83	66

[a] The full article title was “Angelina Jolie Cancer Treatment,” which was a newsworthy topic of 2013. However, that title redirected to a subsection of the “Angelina Jolie” article, effectively making her (broadly popular) article and its (broadly popular) foreign language equivalents part of the medical category, despite the fact many visitors likely arrived there for nonmedical reasons.

Table 4. Topics having most and least variable popularity rank across the top 10 languages.

Topic/Article	Relative variance[a]	Most popular (rank)[b]	Least popular (rank)[b]
Down syndrome	1.000	German (3)	Russian (1506)
Pneumonia	0.973	Russian (1)	German (1516)
Diabetes mellitus	0.964	Spanish (6)	French (1522)
Hypertension	0.953	Chinese (4)	Italian (1522)
Myocardial infarction	0.937	Russian (17)	Polish (1484)
Sixth disease	0.022	Spanish (1012)	English (1402)
Chitosan	0.021	Chinese (609)	Spanish (903)
Candidiasis	0.019	Portuguese (81)	Polish (394)
Mortality rate	0.019	Italian (408)	Japanese (729)
Asbestos	0.005	German (35)	Portuguese (187)

[a] Relative variance is the percentage of the maximum observed variance (ie, it is not an absolute measure, but based on the variance calculated for the “Down syndrome” article).

[b] Popularity rank goes from 1 (most popular) to 1536 (least) because there were 1536 articles in all of the 10 languages used herein; a constraint that helped to normalize these comparisons.

Although sometimes regional or cultural trends were observed (eg, disease effected regions having high popularity for the corresponding article in the local language), a broader explanation of these patterns is a topic for future investigation.

Rather than looking at articles or topics in isolation, we calculated rank similarity between language pairs (Table 5). Working from the set of topics with articles in all the top 10 language editions, we found Portuguese and Spanish visitors (r=.668) had the most similar browsing habits, whereas Russian and English visitors (r=.207) were most dissimilar.

Table 5. Pearson correlation coefficient (eg, “rank similarity”) metric for medical topic popularity in 10 prominent languages.[a]

Language	English	Spanish	Russian	Japanese	German	French	Portuguese	Italian	Polish	Chinese
English	1	.313	.207	.263	.334	.294	.305	.287	.315	.286
Spanish		1	.423	.412	.555	.584	.668	.598	.564	.471
Russian			1	.332	.425	.408	.410	.426	.439	.372
Japanese				1	.529	.470	.412	.479	.486	.532
German					1	.613	.551	.639	.642	.504
French						1	.563	.633	.588	.464
Portuguese							1	.585	.556	.439
Italian								1	.608	.484
Polish									1	.488
Chinese										1

[a] More informally: “how similar is the popularity ordering for topics between 2 languages?” The measure is symmetric.

Correlation of Wikipedia Article Traffic and Disease Prevalence

A 2014 IMS report made the claim that “rarer diseases, which often have fewer available information sources and are less understood by patients and clinicians, show a higher frequency of [Wikipedia] visits than many more common diseases” [2]. Given that English is frequently the language used to search for information on Wikipedia regardless of a person’s country of origin, we used the English traffic data to gain perspective on this claim. We found that the articles associated with the 20 conditions having the greatest YLD and the 20 conditions with the greatest DALYs had an average view count of 1.68 million in 2013. This compares to an average of 189,351 views for the 4791 articles tagged with the disease “infobox” and 78,000 views for the average English medical article. Clearly, globally prevalent and well-known medical conditions tend to receive considerable traffic.

Such macroscale correlation is intuitive, but recent research [9] has also demonstrated the more nuanced capability to utilize traffic data for individual articles in near real time. That work found that the popularity of influenza articles not just correlated with the spread of the disease, but could also be temporally analyzed to create reasonably accurate infection forecasts. The extent to which this applies across the entire article base and the ways the health care community can utilize such rapid signaling are topics for future work.

Quantity/Characteristics of Wikipedia’s Medical Contributors

Year-Over-Year Analysis of Editor Numbers

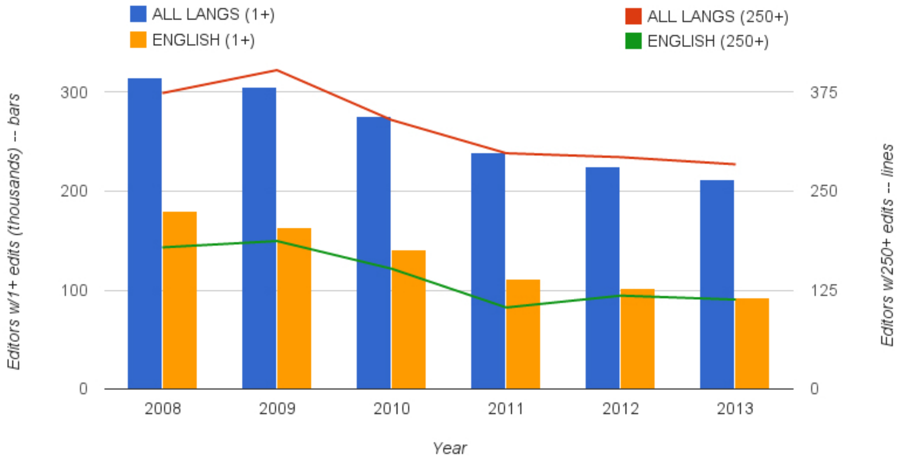
Given Wikipedia’s collaborative nature, it is logical to investigate the editor community that has authored the content of such a frequently accessed resource. Most often, “editors” in this context are users with a persistent account name and log-in credentials. Although one may edit without an account, rarely do such users exhibit the consistent participation on which we focused. Of the 274 top contributors, just 4 edited without an account name.

We measured participation by looking at an editor’s quantity of contributions on medical articles in a given calendar year. The following are some participation thresholds measured across all languages in 2013: ≥5 edits=21,563 editors; ≥25 edits=5573 editors; ≥100+ edits=1237 editors; ≥250 edits=274 editors; ≥1000 edits=39 editors; ≥10,000 edits=1 editor (this paper’s lead author). There were 32 language versions that had at least 1 editor with ≥250 edits in 2013.

We plotted some of these same thresholds on a yearly basis from 2008 to 2013 with breakdowns (Figure 7). We found that at all participation thresholds the number of editors decreased. Over this 5-year span, the decrease in editor numbers was approximately 40% for English Wikipedia, with 10%-20% attrition typical for non-English languages.

Not included in the preceding totals is the work of nonhuman, automated “bot” editors: computer programs that perform much repetitive maintenance. Bots and humans combined made 1,106,575 medical edits in 2013 with 406,003 (36.69%) of those in English. Bots accounted for 24.72% (274k/1107k) of the global total and 10.54% (43k/406k) of the English total, numbers slightly inflated due to the bot-driven migration of interlanguage links as described in the Methods section.

Figure 7. Quantity of editors making ≥1 and ≥250 medical contributions by year across all languages for English and all languages.



Contributor Demographics/Background via Survey

In May of 2014, we sent out a survey to 271 of the 274 top medical editors in 2013. Three users were omitted because they had been blocked from contributing to Wikipedia due to various issues. Of these, 117 (43.2%) responded and their answers are summarized in Table 6.

We found more than half of editors (50.4%, 59/117) were either health care professionals or studying health care. Of the 58 outside of health care, 17 used the open text area to describe their activities as primarily grammatical, formatting, language simplifications, and the removal of vandalism. Fifteen others reported more substantive editing despite lacking formal medical training. In some cases (2 self-reported), contributors were arguably experts despite not being health care providers: 1 was a PhD biochemist and another was a SCUBA diver editing in related medical spaces.

Table 6. Survey responses from 117 top medical editors across all language editions.

Question	n (%)
What is your highest level of education?	
High school or less	17 (14.5)
Bachelor's	39 (33.3)
MS/PhD/MD	61 (52.2)
Do you currently work in the health care field? Or have you previously?	
Yes	59 (50.4)
No	58 (49.6)
Are you currently studying health care (a student)?	
Yes	17 (14.5)
No	100 (85.5)
What language of Wikipedia do you mostly work on?	
English	58 (49.6)
Non-English	59 (50.4)
Did you receive a barnstar? (yes)[a]	117 (100)
How do you identify your gender?	
Male	96 (82.1)
Female	10 (8.5)
Other/omitted	11 (9.4)

[a] Sanity check; surveys with "no" responses were discarded.

Discussion

Principal Results

Wikipedia's medical content is made up of more than 155,000 articles and 1 billion bytes of text across 255 languages. This content is supported by more than 950,000 references and was viewed more than 4.88 billion times in 2013 (with mobile-inclusive estimates at 6.5 billion). Third-party analytics suggests Wikipedia is the most viewed medical resource globally. As of 2013, the core editor community numbered less than 300 and had decreased over the previous 5 years. The members of this community are half health care providers and 85% have a university education.

Limitations

Amount of Wikipedia Medical Content

Our analysis depended heavily on the Wikipedia editor community to establish (1) what constitutes a medically related article and (2) the interlanguage links between corresponding articles. Whether or not something is related to medicine or related "enough" to justify a tagging is a subjective distinction. Interlanguage links are often less ambiguous, but still require a bilingual speaker who is familiar with Wikipedia syntax.

Although subjectivity might shift these bounds slightly, more articles have likely never been considered in these contexts, either because they are undiscovered entirely or they are too emergent, tangential, or unpopular to draw the attention of the editors who typically make category and interlanguage annotations. Although usually quickly restored [25], "vandals" also sometimes destroy tags or links with malicious intent.

Following the very nature of collaborative work, it is our subjective experience that "major" topics are more likely to be correctly tagged and linked than more obscure ones. Thus, tagging and linking inaccuracies likely have a greater impact on article quantity measurements than readership totals. In particular, categorization omissions could be estimated by searching English Wikipedia using a database of terms such as the International Classification of Disease (ICD-10) and verifying that corresponding articles have been appropriately tagged. We leave this as a topic for future research.

Lastly, our analysis used tagged English articles as the starting point for interlanguage link discovery. A medical topic that did not have a corresponding English article version would not be included in our analysis.

Citations Supporting Wikipedia's Medical Content

Wikipedia strives for verifiable content rather than the less agreeable notion of absolute "truth." As such, information drawn from reputable sources upholds the notability and verifiability requirements that Wikipedia promotes.

In this work, we quantify the number of references (and highlight some particularly well-reputed sources) as a proxy for reliability. We recognize that the number of references is just 1 mark of quality. Content may be inaccurate despite having a citation and vice versa. Our data do not look at whether or not the text of Wikipedia accurately reflects the sources in question or if the sources are outdated. Both would be interesting questions to investigate further.

Readership of Wikipedia's Medical Content

Language-scale aggregates regarding Wikipedia readership are influenced by the number of member articles. Thus, previously discussed limitations surrounding category tagging and interlanguage links also cascade into this analysis.

It is important to emphasize that none of our traffic data (Wikipedia or third party) includes readership from mobile devices. These shortcomings in the WMF's collection infrastructure were remedied during our writing in October 2014; mobile readership will be analyzed in future work. Although allowing for fair comparison, this also means we underreport the scale at which other online health care resource operate. Across all English Wikipedia (not just medical portions), mobile views are more than 30% of the total traffic and growing [20]. Thus, readership as we present it may underrepresent the browsing habits of certain economies, languages, and regions (eg, where mobile networks are the only means of connectivity and/or cellular devices are the only affordable means of access) or certain demographics (eg, youth demonstrating a preference for mobile browsing).

Moreover, when comparing Wikipedia's medical readership to other health care websites, one must be mindful of the varying coverage and scope. Although it would be interesting to compare per-topic page views, alternative sites (some proprietary) have not made such granular traffic data publicly available.

In our broad comparison of readership on health care websites, we relied on the third-party service SimilarWeb [21,26-32]. That service's measurement methodology and accuracy is not known. However, it is reassuring that SimilarWeb's page view estimates for the entirety of English Wikipedia differed only by about 3% from the more authoritative data published by the WMF.

To some extent, all information sources find themselves mirrored across the Internet and combined into other sources. However, this occurs more frequently with Wikipedia and government sources because they are freely licensed or in the public domain which encourages reuse. Such transitive/downstream consumption (both online and offline) is difficult to quantify. For example, low-cost "alternative textbook" provider Boundless amasses such open-source content when compiling its texts [33], with some becoming popular in practice [34]. Further, the National Institutes of Health (NIH) and Wikipedia often see their content integrated directly into Google search results and these sources often have high search-engine ranking [35].

Quantity/Characteristics of Wikipedia's Medical Contributors

Our survey to medical editors had a response rate of approximately 43%. This raises the concern that those with the time and willingness to complete the questionnaire are somehow nonrepresentative. Although approximately half of recipients primarily edit a non-English Wikipedia, our survey was available only in English, potentially limiting and biasing the response pool. Our validation question ("Did you receive a barnstar?") also takes respondents at their word in addition to trusting the feedback received for all other questions.

We identified 4 Internet Protocol (IP) accounts that made more than 250 edits assuming that those IP addresses are statically assigned to a single contributor. Dynamic IP assignment (ie, the Dynamic Host Configuration Protocol, DHCP) is common in residential and wireless networks and could have effects such that multiple human users inhabit a single IP over time (causing an overestimation on our part) or that a single user's contributors are unknowingly spread across IP space (an underestimation).

Comparison With Prior Work

The Introduction enumerates some of the prior research that qualitatively relates to this work. A purely quantitative point of reference comes from the parallel work of Farič and Potts [36], who also surveyed English Wikipedia's most active medical editors. That research found 50% of those surveyed had a medical background, 70% were older than 30 years, most were male, and 75% had a college degree. All data points were quite similar to our findings, which additionally considered non-English editors.

Conclusions

Amount of Wikipedia Medical Content

Although Wikipedia has a tremendous amount of medical content, it is primarily concentrated in English and a few major European languages. As a user-generated website, this reflects the populations that are willing and able to contribute. Wikipedia's distribution of content by language, however, better matches global language popularity than the Internet does as a whole. Additionally there are ongoing efforts to improve Wikipedia's medical coverage in non-English languages via a partnership with the not-for-profit Translators Without Borders.

Citations Supporting Wikipedia's Medical Content

Wikipedia is relatively well referenced and by this marker is becoming increasingly reliable over time. Encouragingly, references to high-quality sources, such as The Cochrane Collaboration, are rising at a greater rate than references on the whole.

Readership of Wikipedia's Medical Content

A previous IMS report [2] claimed that Wikipedia is the single most used medical resource on the Internet. Our statistical work herein appears to confirm this assertion, with conservative analysis putting Wikipedia's readership on par with NIH and surpassing that of WebMD (2 sites traditionally atop the health category). With the Internet likely to be the most consulted information medium, Wikipedia may well be the most used medical resource overall.

Our study unexpectedly found strong variance (up to a factor of 4) in the proportional popularity of health content across different languages. The catalyst for this variation is unclear. Is it the case that Spanish speakers care more about their health than Chinese speakers? Or do Chinese populations prefer a different information resource?

We also found that popular topics/articles differed wildly among languages. This has interesting ramifications as emergent language editions try to expand their medical content (either organically or through translation). Simply assuming content that is well read in 1 language will draw audiences in another is insufficient and more careful cultural consideration may be prudent.

Quantity/Characteristics of Wikipedia's Medical Contributors

Although Wikipedia's medical content has tremendous readership, the number of significantly active contributors is few. It is concerning that these editor numbers, at all thresholds, have decreased over the past 5 years. This trend is one exhibited not just by medical contributors, but the overall Wikipedia community. A number of explanations have been proposed for this poor retention and recruitment: (1) deterrents such as stricter reference requirements and more policy, (2) growing competition for participant attention in the open-source and user-generated content communities, (3) xenophobia and a community unwelcoming of new users [37], and (4) the perception that in some languages there remains little "low hanging fruit" to be authored. Understanding and reversing this trend is an area of active research for Wikipedia and its subcommunities.

The community of medical editors, like Wikipedia overall, is male dominated [38]. The reasons are not entirely clear, but some possibilities include technical barriers, lack of self-confidence, minimal social activity, and the adversarial nature of some discussions [39]. Efforts to make Wikipedia more female friendly are also ongoing.

Our survey of Wikipedia's medical contributors found many are health care professionals and most are university educated. Although just 29% of the US population has a Bachelor's degree [40], 85% of Wikipedia's core medical editors have attained one (with more than 50% going beyond that level). Educational levels attained were similar between editors for English and non-English versions. These educational and professional benchmarks put into doubt the claims by some that Wikipedia is "antiexpert" [41].

Acknowledgments

Nikita Belavate, Soumyadeep Bhaumik, Abhijith Jayanthi, Manu Mathew, and Jake Orlowitz are thanked for their help on this project. Data analysis was done on machines hosted at the University of Pennsylvania.

Conflicts of Interest

The authors are active contributors to Wikipedia. They have no financial conflicts of interest related to the subject matter of the article.

Multimedia Appendix 1

Additionally data.

[PDF File (Adobe PDF File), 28KB - jmir_v17i3e62_app1.pdf]

References

1. Zachte E. Wikimedia report card-August 2014. 2014 Aug. URL: http://reportcard.wmflabs.org/ [accessed 2014-10-24] [WebCite Cache ID 6TZK6977p]
2. IMS Health. 2014. Pharma should make better use of social media to engage patients and improve the use of medicines URL: http://www.imshealth.com/portal/site/imshealth/menuitem.c76283e8bf81e98f53c753c71ad8c22a/?vgnextoid=ebc072cc270b3410VgnVCM10000076192ca2RCRD&vgnextchannel=ba11e590cb4dc310VgnVCM100000a48d2ca2RCRD&vgnextfmt=default [WebCite Cache ID 6RI7WRu4P]
3. Heilman JM, Kemmann E, Bonert M, Chatterjee A, Ragar B, Beards GM, et al. Wikipedia: a key tool for global public health promotion. J Med Internet Res 2011;13(1):e14 [FREE Full text] [doi: 10.2196/jmir.1589] [Medline: 21282098]
4. Hughes B, Joshi I, Lemonde H, Wareham J. Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. Int J Med Inform 2009 Oct;78(10):645-655. [doi: 10.1016/j.ijmedinf.2009.04.008] [Medline: 19501017]
5. Allahwala UK, Nadkarni A, Sebaratnam DF. Wikipedia use amongst medical students - new insights into the digital revolution. Med Teach 2013 Apr;35(4):337 [FREE Full text] [doi: 10.3109/0142159X.2012.737064] [Medline: 23137251]
6. Makovsky Integrated Communications. 2014. State of healthcare searches online URL: http://www.makovsky.com/images/Health-infographic_90613_v4.png [accessed 2014-08-31] [WebCite Cache ID 6SFYY9lSW]
7. West AG, Milowent. Wikipedia Signpost. 2013 Feb 04. Examining the popularity of Wikipedia articles: Catalysts, trends, and applications URL: https://en.wikipedia.org/wiki/Wikipedia:Wikipedia_Signpost/2013-02-04/Special_report [accessed 2014-07-23] [WebCite Cache ID 6RID1plVj]
8. McIver DJ, Brownstein JS. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. PLoS Comput Biol 2014 Apr;10(4):e1003581 [FREE Full text] [doi: 10.1371/journal.pcbi.1003581] [Medline: 24743682]

9. Generous N, Fairchild G, Deshpande A, Del Valle SY, Priedhorsky R. Global disease monitoring and forecasting with Wikipedia. PLoS Comput Biol 2014 Nov;10(11):e1003892 [FREE Full text] [doi: 10.1371/journal.pcbi.1003892] [Medline: 25392913]
10. Mesgari M, Okoli C, Mehdi M, Lanamäki A. “The sum of all human knowledge”: A systematic review of scholarly research on the content of Wikipedia. Journal of the Association for Information Science and Technology 2014 Dec 02;66(2):219-245 [FREE Full text] [doi: 10.1002/asi.23172]
11. Kräenbring J, Monzon PT, Gutmann J, Muehlich S, Zolk O, Wojnowski L, et al. Accuracy and completeness of drug information in Wikipedia: a comparison with standard textbooks of pharmacology. PLoS One 2014 Sep;9(9):e106930 [FREE Full text] [doi: 10.1371/journal.pone.0106930] [Medline: 25250889]
12. Haigh CA. Wikipedia as an evidence source for nursing and healthcare students. Nurse Educ Today 2011 Feb;31(2):135-139 [FREE Full text] [doi: 10.1016/j.nedt.2010.05.004] [Medline: 20646799]
13. Bould MD, Hladkowicz ES, Pigford AE, Ufholz LA, Postonogova T, Shin E, et al. References that anyone can edit: review of Wikipedia citations in peer reviewed health science literature. BMJ 2014;348:g1585 [FREE Full text] [Medline: 24603564]
14. Azer SA. Evaluation of gastroenterology and hepatology articles on Wikipedia: are they suitable as learning resources for medical students? Eur J Gastroenterol Hepatol 2014 Feb;26(2):155-163. [doi: 10.1097/MEG.0000000000000003] [Medline: 24276492]
15. Volsky PG, Baldassari CM, Mushti S, Derkay CS. Quality of Internet information in pediatric otolaryngology: a comparison of three most referenced websites. Int J Pediatr Otorhinolaryngol 2012 Sep;76(9):1312-1316. [doi: 10.1016/j.ijporl.2012.05.026] [Medline: 22770592]
16. Wikipedia. WikiProject Medicine URL: http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine [accessed 2014-10-23] [WebCite Cache ID 6TXxqlqYL]
17. Wikipedia. 2014. WikiProject Medicine/Assessment URL: http://en.wikipedia.org/w/index.php?title=Wikipedia:WikiProject_Medicine/Assessment [WebCite Cache ID 6V3jXhmaY]
18. Wikipedia. MediaWiki API help URL: http://en.wikipedia.org/w/api.php [accessed 2014-10-23] [WebCite Cache ID 6TXxXXrgn]
19. Page view statistics for Wikimedia projects. URL: http://dumps.wikimedia.org/other/pagecounts-raw/ [accessed 2014-10-23] [WebCite Cache ID 6TXocLq6m]
20. Zachte E. Page views for Wikimedia, all projects, all platforms, normalized. 2014. URL: http://stats.wikimedia.org/EN/TablesPageViewsMonthlyAllProjects.htm [accessed 2014-08-31] [WebCite Cache ID 6SFYGxqW9]
21. SimilarWeb. 2014 Sep 1. Desktop traffic statics URL: http://www.similarweb.com/ [accessed 2015-02-24] [WebCite Cache ID 6WaaX0NdA]
22. World Health Organization. Estimates for 2000-2012: Disease burden URL: http://www.who.int/healthinfo/global_burden_disease/estimates/en/index2.html [accessed 2014-09-01] [WebCite Cache ID 6SFt8WoqM]
23. Wikipedia. 2014 Sep 1. List of languages by number of native speakers URL: https://en.wikipedia.org/w/index.php?title=List_of_languages_by_number_of_native_speakers [WebCite Cache ID 6SFY2J2bi]
24. W3Techs Web Technology Surveys. 2014. Usage of content languages for websites URL: http://w3techs.com/technologies/overview/content_language/all [accessed 2014-09-01] [WebCite Cache ID 6SFYiWhnZ]
25. West AG. Damage Detection and Mitigation in Open Collaboration Applications [PhD thesis]. Philadelphia, PA: University of Pennsylvania; 2013 May. URL: http://works.bepress.com/cgi/viewcontent.cgi?article=1034&context=andrew_g_west
26. URL: http://www.similarweb.com/website/en.wikipedia.org [accessed 2014-09-01] [WebCite Cache ID 6SG2jo8hD]
27. URL: http://www.similarweb.com/website/nih.gov [accessed 2014-09-01] [WebCite Cache ID 6SG2tpSFG]
28. URL: http://www.similarweb.com/website/webmd.com [accessed 2014-09-01] [WebCite Cache ID 6SG2wYwzE]
29. URL: http://www.similarweb.com/website/mayoclinic.org [accessed 2014-09-01] [WebCite Cache ID 6SG306pg9]
30. URL: http://www.similarweb.com/website/nhs.uk [accessed 2014-09-01] [WebCite Cache ID 6SG32tqhC]
31. URL: http://www.similarweb.com/website/who.int [accessed 2014-09-01] [WebCite Cache ID 6SG35JVc3]
32. URL: http://www.similarweb.com/website/uptodate.com [accessed 2014-09-01] [WebCite Cache ID 6SG37Cq2t]
33. Owen J. Boundless. 2013. Boundless catapults textbooks to a new, personalized level with learning technology URL: http://blog.boundless.com/2013/08/boundless-catapults-textbooks-to-a-new-personalized-level-with-learning-technology/ [accessed 2014-07-24] [WebCite Cache ID 6RIM1XAMT]
34. Weddell C. Boundless. 2014. Feel the love: Educators tell us why they love Boundless URL: http://blog.boundless.com/2014/02/educators-on-boundless-teaching-platform/ [accessed 2014-07-24] [WebCite Cache ID 6RIP7bNAQ]
35. Laurent MR, Vickers TJ. Seeking health information online: does Wikipedia matter? J Am Med Inform Assoc 2009 Jul;16(4):471-479 [FREE Full text] [doi: 10.1197/jamia.M3059] [Medline: 19390105]
36. Farič N, Potts HWW. Motivations for contributing to health-related articles on Wikipedia: an interview study. J Med Internet Res 2014 Dec;16(12):e260 [FREE Full text] [doi: 10.2196/jmir.3569] [Medline: 25498308]
37. Halfaker A, Kittur AN, Riedl J. Don't bite the newbies: How reverts affect the quantity and quality of Wikipedia work. 2011 Presented at: Proceedings of the 7th International Symposium on Wikis and Open Collaboration; Oct 3, 2011; Mountain View, CA. [doi: 10.1145/2038558.2038585]

38. Hill BM, Shaw A. The Wikipedia gender gap revisited: characterizing survey response bias with propensity score estimation. PLoS One 2013 Jun;8(6):e65782 [FREE Full text] [doi: 10.1371/journal.pone.0065782] [Medline: 23840366]
39. Gardner S. Sue Gardner's Blog. 2011 Feb 19. Nine reasons women don't edit Wikipedia (in their own words) URL: http://suegardner.org/2011/02/19/nine-reasons-why-women-dont-edit-wikipedia-in-their-own-words/ [WebCite Cache ID 6V31765KCYl]
40. United States Census Bureau. Educational attainment in the United States: 2013 URL: http://www.census.gov/hhes/socdemo/education/data/cps/2013/tables.html [accessed 2014-09-01] [WebCite Cache ID 6SG4EhBSl]
41. Goodwin J. The authority of Wikipedia. 2009 Jun Presented at: 8th Conference of the Ontario Society for the Study of Argumentation ‘Argument Cultures; June 3-6, 2009; Windsor, ON URL: http://goodwin.public.iastate.edu/pubs/goodwinwikipedia.pdf

Abbreviations

API: application program interface
DALY: disability adjusted life years
DHCP: Dynamic Host Configuration Protocol
IMS: Institute of Medical Science
IP: Internet Protocol
NIH: National Institutes of Health
PDQ: Physician Data Query
WMF: Wikimedia Foundation
YLD: years lived with disability

Edited by G Eysenbach; submitted 23.11.14; peer-reviewed by N Farič, H Potts; comments to author 17.12.14; revised version received 10.01.15; accepted 14.01.15; published 04.03.15.

Please cite as:
Heilman JM, West AG
Wikipedia and Medicine: Quantifying Readership, Editors, and the Significance of Natural Language
J Med Internet Res 2015;17(3):e62
URL: http://www.jmir.org/2015/3/e62/
doi:10.2196/jmir.4069
PMID:25739399

Original Paper

Blogging as a Viable Research Methodology for Young People With Arthritis: A Qualitative Study

Julie Prescott[1], PhD; Nicola J Gray[2], PhD; Felicity J Smith[3], PhD; Janet E McDonagh[4], MD

[1]University of Bolton, Department of Education and Psychology, Bolton, United Kingdom

[2]Green Line Consulting Ltd, Manchester, United Kingdom

[3]UCL, School of Pharmacy, London, United Kingdom

[4]University of Manchester, Centre for Musculoskeletal Research, Manchester, United Kingdom

Corresponding Author:
Julie Prescott, PhD
University of Bolton
Department of Education and Psychology
Deane Road
Bolton, BL3 5AB
United Kingdom
Phone: 44 01204903676
Fax: 44 01204903676
Email: j.prescott@bolton.ac.uk

Abstract

Background: The development of services that are responsive to the needs of users is a health policy priority. Finding ways of engaging young people in research to gain insights into their particular experiences, perspectives, and needs is vital but challenging. These data are critical to improving services in ways that meet the needs of young people.

Objective: Our aim was to evaluate Web-based blogging as a viable method for understanding the daily experiences and condition management strategies of young people with juvenile arthritis.

Methods: To meet the objectives of the study, a qualitative approach was required to gather information on the experiences and perspectives of young people regarding the management of their condition and its daily impact. In collaboration with a group of young people with arthritis, a custom website was developed. This website provided the opportunity for young people (aged 11-19) with arthritis from a United Kingdom pediatric hospital to contribute blogs. It was designed so that young people were free to write about whatever was important to them, but the site also included some structure and prompts to facilitate the writing of blogs. Qualitative analytical procedures were employed, supported by NVivo software.

Results: Engagement in the study by young people was variable in terms of their participation rates, frequency of website visits, and the length of their blogs. Young people used the site in different ways, some responding to the website categories and prompts that the team created, while others used it as a diary to record their experiences and thoughts. In line with principles of qualitative inquiry, the data collection was participant-led. Young people were in control of what, how much, and how often they wrote. However, some young people expressed difficulty regarding knowing what they should blog about. For a number of reasons, discussed here, the blogs may also not be fully reflective of experiences and perspectives of the participants. However, the data obtained provided insights into young people's experiences of living with arthritis and their use of medicines in the context of their daily lives.

Conclusions: Web-based research with young people presents opportunities and challenges for researchers. Web-based blogging methodology has the potential to give young people and parents the space and empowerment to express their own ideas and concerns. However, this project suggests that it might not be the best way to engage a large diverse group of young people and might most effectively be combined with other approaches. Despite these limitations, the study provided valuable data about the experience and impact of living with a long-term condition from the perspectives of young people with arthritis.

(J Med Internet Res 2015;17(3):e61) doi:10.2196/jmir.3608

KEYWORDS

young people; blogging; qualitative research; arthritis

Introduction

In 2003, Internet-based research was viewed as still being in its infancy [1]. More than 10 years on, the Internet is widely employed in many aspects of daily life and is increasingly being used within academic research. In particular, Web-based surveys have become very popular [2,3]. They have been noted for their potential to reach very large audiences, inexpensively, with rapid replies [4,5]. Although for some research, they may be limited by anticipated response rates; in other studies with specified populations, and with the use of more sophisticated software, they can be valuable [6,7].

The potential value of Web-based research for qualitative studies is less clear. The Internet has been used for a diverse range of studies using traditional qualitative data-gathering techniques such as interviews [8,9] and focus group research [10]. Experience is considerably more limited regarding the potential use of blogs [11,12]. A blog (short for weblog) has been described as a Web-based diary in which the blogger can freely record opinions and recount experiences. Blogs can be produced by single or multiple authors, and they provide information to tell a story, and/or represent particular experiences or perspectives on any issue [13]. Bloggers are often part of a blogging community that centers on a common interest (eg, Olive's 2013 study on surfers [12]). By their very nature, it seems that blogs may provide an excellent research opportunity; however, they are currently underutilized in social science and health services research. Hookway [11] compares blogs to offline diary research, suggesting blogs allow for more extensive research opportunities through the potential anonymity of blogs (both the anonymity of the blogger and those that comment on other blogs) and the ability to enable researchers to reach a wider audience. Diary research has been used to investigate, among other issues, health behavior [14-16]. Benefits include the ability to collect sensitive information [14] and the ability to capture an "everchanging event" [14], where the time between an event and its recording is reduced, resulting in less memory impairment and reduced reconstruction of the event compared to other qualitative methods such as interviews and focus groups [17]. Olive [12] found blogging a good means of data collection since there were no location constraints to blogging—as long as you had access to a computer and the Internet—and the process of blogging could also allow the blogger to be reflective.

Web-based resources for gaining information and sharing experiences on medical/health issues are increasingly popular [18], and these too may provide data for research, especially for populations with specific medical/health issues [18]. Malik and Coulson [19] used Web-based methodologies by examining a password-protected Web-based support group to understand the lives and experiences of young people with inflammatory bowel disease (IBD). A particular benefit of Web-based support groups, as featured in Malik and Coulson's [19] study, is that the groups are available 24/7, which enables people to receive support and advice when more traditional resources are unavailable. Web-based support groups also offer people the convenience of support with similar people from within their own homes [20], which may be an important feature for people with long-term conditions who may have difficulty attending face-to-face meetings. Web-based groups also offer anonymity, which may increase self-disclosure [20,21]. Another study [16] of 6 young people aged between 11 and 16 years with IBD used an audio (offline) diary methodology. Researchers claimed that the diary methodology was a viable method for communicating and engaging with young people. Participants were able to record what they wanted, when they wanted, with researchers able to gain an understanding of the lives of a diverse group of young people.

With all Web-based methodologies, researchers must be aware of their sample: their sample's ability to blog, having an interest in blogging, and access to the resources. However, research suggests that many young people aged between 13-19 blog [22,23]. Indeed, Hookway [11] posits that bloggers need low technical competence in order to blog, and therefore it is a methodology that is potentially inclusive. In the United Kingdom, however, inequalities in relation to access and use of the Internet have been identified [24,25] with boys, older children, and middle class children more likely to have access. This suggests there is still a digital divide for young people that should be borne in mind when considering Web-based research with specific groups. Despite inequalities in access, previous studies in the United States [23] and United Kingdom [26] on gender differences in blogging have suggested that females aged 15-17 blog more than their male counterparts. Gender differences have been found to exist in how and why people blog [26,27] and with female-authored blogs being commented on, and noted less [26].

The development of services that are responsive to the needs of users is a health policy priority in the United Kingdom, and the value of qualitative methods in enabling researchers and policy makers to gain insights into the experiences and priorities from the perspective of service users is recognized. In particular, research into the perspectives of young people that can inform service development is very limited [28]. It is recognized, however, that young people with long-term conditions have very distinct needs and concerns in their daily management and preferences regarding service provision and delivery. Finding ways of engaging young people in research—to gain insights into their experiences, needs, and priorities—is vital but challenging [28].

This paper describes Web-based blogging as a method for understanding the daily experiences and condition management of young people with juvenile arthritis. A systematic review of qualitative studies involving the experience of children and young people living with juvenile idiopathic arthritis (JIA) has been conducted [29]. None of the resulting 27 studies, completed up to 2011, involved Web-based blog data, although there were a small number involving video diaries [29]. The overall aim of the research was to investigate the relationships between identity and medication use among adolescents with arthritis and to explore the role of health professionals in delivering services to this group. This paper discusses how using a qualitative Web-based bespoke blogging website, shaped by young people, enabled the research team to meet the overall research objectives and engage with young people with a long-term health condition.

Methods

Study Setting and Participants

Young people (aged 11-19 years) with arthritis who attended a clinic at a major UK pediatric hospital were invited to take part in the study. The clinic served a diverse group of young people from the West Midlands area.

Ethical and Research Approvals

Approval for the study was given by Coventry & Warwickshire Research Ethics Committee in July 2012 (ref 12/WM/0184) and the Hospital Directorate for Research and Development.

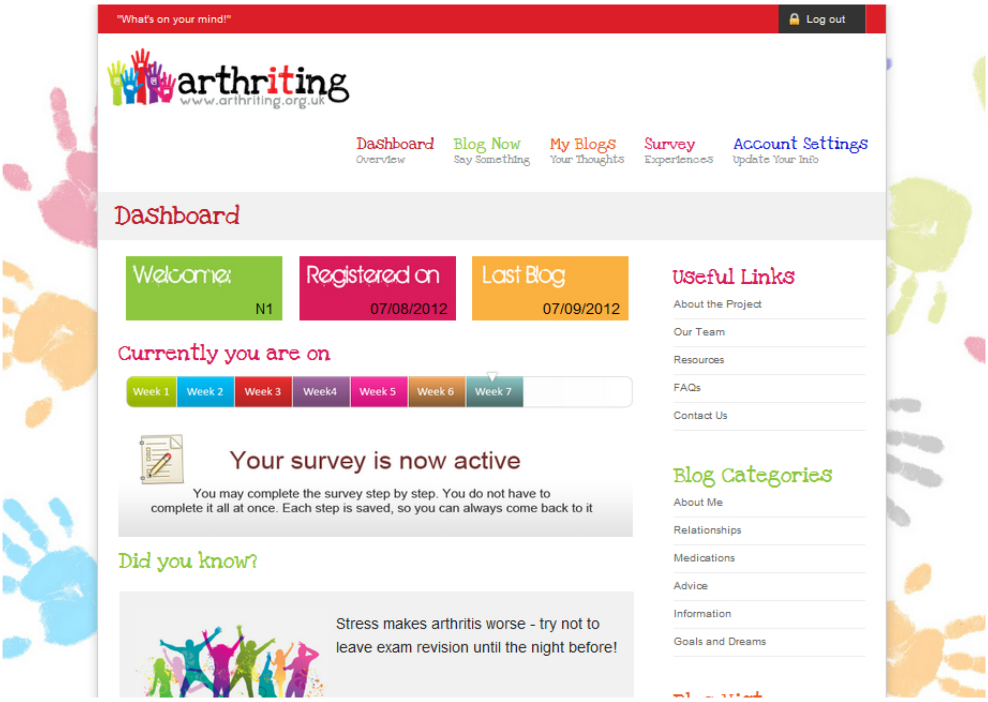
Procedure

Young people and their parents were identified from the patient list of a rheumatology clinic in a large UK pediatric hospital. Eligibility criteria for young people required that they had confirmed child/adolescent onset arthritis, were aged 11-19 years, and had no cognitive/learning disorder that would prevent their ability to provide data. All eligible young people (and their parents) were invited to participate by letter (n=107) and followed up during routine clinic appointments (n=70). Young people under the age of 16 gave their own assent to take part, coupled with a parent's consent. Young people aged 16 and over provided their own consent. Consenting participants were given an information pack about the "Arthriting" website and a personal login code. To maximize relevance and uptake, development of the website was undertaken in partnership with a young people's user group and included stringent security processes to ensure data protection and participant safety (eg, password-protected logins, secure firewalls, regular moderation). The website was also functional on smartphones, recognizing that many young people (and parents) access the Internet this way. Once registered, participants were asked to choose a nickname and password for future logins. On completion of the 2-month period, participants were sent a letter and online shopping voucher to thank them for their time.

Website Design

The young people contributed to the design of the website by attending a discussion with the website builders and members of the research team. Many of its features were decided in partnership with the young people, who had strong ideas regarding the look and operation of the website. One young person suggested the name "Arthriting" and the logo for the project and website. The website had a secure login function for young people to contribute blogs. Once logged in, participants had a project "Dashboard", which showed them where they were in the project and gave them a launching-off point to the blog space (see Figure 1). To ensure confidentiality of data, blogging could not take place as a community, but privately as individuals. Some of the young people expressed a wish to communicate with each other, but due to the potentially sensitive nature of the material and desire of others that blogs were private, a decision was made to have a closed individual blog rather than an open group blog method. Thus, blogs were not open for comment by others.

Figure 1. Arthriting website logged-in dashboard (dummy demo testing user).



"What's on your mind!"

Dashboard

Welcome: N1

Registered on 07/08/2012

Last Blog 07/09/2012

Currently you are on

Your survey is now active

You may complete the survey step by step. You do not have to complete it all at once. Each step is saved, so you can always come back to it

Did you know?

Stress makes arthritis worse - try not to leave exam revision until the night before!

Blog Structure

To achieve the objectives of the project, the research team generated an initial list of blog categories to give young people ideas for the topics and types of messages that they might post (eg, thoughts about identity, the arthritis condition, medication, and the use of health services). These "blog categories" were developed and refined in discussion with the young people.

When using the blogs, participants were able to choose their own font style and color and choose an emoticon to express how they were feeling that day. They were invited to choose a blog category and to write their own title for the blog. They could also indicate their feelings in a “smiley status”. Young people could edit or delete blogs at any time during their 2-month engagement with the project. There were no minimum or maximum blog lengths or number of blog entries that any person could post. We hoped that they would visit and revisit during their 2-month stay.

Data Collection

The young people’s information sheet gave the following guidance about blogging:

> *You will then be invited to write down online what you think of your medicines (both injectable ones as well as those you take by mouth), and how medicines fit—or indeed do not fit—into your everyday lives. It will be like writing a diary but online and like a diary, entries do not have to be every day nor do they have to be lengthy—only just enough to give others a picture of what life is like for you with your arthritis and your medicines. You can take up to 2 months to cover all the issues that you want to tell us about.*

Basic demographic and relevant clinical data (which we asked permission to extract during the consent process) were recorded. These data were age, sex, ethnic group, age at diagnosis, and number of medications.

Data Processing and Analysis

For analysis of the blog data to meet the objectives of this paper, a subgroup of the wider multidisciplinary project team (the authors of this paper) first examined each individual blog to provide a case-by-case overview of the themes of each individual’s entry. From this, a coding frame was developed. As analysis proceeded, this was modified and refined using constant comparison techniques, in which all items of data assigned a particular code were appraised for similarities and divergences from those already coded. To ensure the reliability of analytical procedures, all stages of the data processing, coding, and analysis involved at least 2 members of the research team. Computer software (NVivo) was used to assist in the data management and handling.

Results

Overview

In total, 107 young people were eligible and invited to take part, of which 36 young people completed consent procedures and 25 registered with the site. At least one blog was received from each of 21 young people. In addition, 6 parents also registered and completed at least one blog. The sample of the 21 young bloggers broadly reflected the wider clinic population in terms of age and type of arthritis diagnosis, but white young people and females were overrepresented in our blog sample. This may be viewed as consistent with literature showing that female adolescents are more likely to use online health information sources [25]. Table 1 shows characteristics of the 21 young people who contributed at least one blog.

We asked each young person and parent to choose a nickname for their blog and to avoid linking it closely to their real name. The moderator did not know the identity of the bloggers, but the clinical staff did. Where the clinical staff determined that the nickname was too close to the blogger’s real name, we retained the first and last letter and put two full stops “..” in between, regardless of how many other letters there actually were. Blogs have not been edited by the team, but some examples are extracts from a longer blog.

Table 1. Characteristics of participants.

Characteristics	n
Age at recruitment	
11-15	18
16-19	3
Gender	
Female	17
Male	4
Ethnic group	
White	17
Non-white	4
Age at diagnosis	
<11 years	13
11+ years	8
Time since diagnosis	
1-5 years	8
6-10 years	4
11+ years	8
Not known	1
Type of juvenile arthritis	
JIA	19
Other	2

Young People's Approach to the Site

The following quote illustrates a young person's understanding of the study and their role in it: "Taking medication every day is hard, and living with the pain of arthritis is harder, and over the next few weeks, I hope to blog and give you a better understanding of what it's like 'living with arthritis' and 'being me'." [BeanyBabe96, female, age 16].

There was great variability in the frequency of visits to the site by young people and in the length of blogs, with some young people writing frequently and at length and others writing one or two blogs over the duration of the project. Among the young people who blogged, the average number of blogs posted was 8, and the range 1-36. Among the 6 parents, the average number of blogs posted was 4, and the range 1-12. In total, 187 blog entries were contributed.

The young people approached the site and blogging in different ways. Some were guided by the "blog categories" listed on the site. They worked through these, contributing thoughts and experiences that they felt were relevant. This could provide a young person's perspectives on a broad range of issues. For example, the following were contributed by a female blogger, aged 13, nickname 123456:

> *Hi today I'm going to talk about medication and how I feel about it. Honestly it doesn't really affect me most of the time); it doesn't make me feel different from anyone. I'm glad I take medication because without it I wouldn't be well. However sometimes when I am angry I don't feel like taking it because I can't be bothered and think why should I have to take it. WHY ME????*
>
> *Hey, so my dreams and goals. Well, when I'm older I want to become a doctor. I've always wanted to become a doctor even from a young age, but after I got diagnosed I saw how much the doctors helped and cared for me. I would like to do that for some to. My illness won't affect my future a lot; hopefully it stays in control and then everything will be alright.*
>
> *So relationships...it doesn't really affect me at all I'm still the same person I was before I got diagnosed. My relationships with other are the same, I'm a bit more moody than I used to be which sometimes makes relationships with others hard because I just want to be by myself, but other than that it doesn't affect me at all!!!'*

Other bloggers adopted a more diary-like approach recording their experiences, views, and perspectives with no particular reference to the blog categories. These blogs could provide insights in the context of current priorities and everyday concerns of young people. Thus the blogs may not necessarily explicitly focus on the experience of arthritis and medicines but still do demonstrate how these issues relate to, or impact on, wider aspects of a young person's life. The blog of "flower123", a 14-year-old female blogger, is a fascinating example of this approach. A typical entry recounted experiences at school and social activities with family and friends:

> *Yesterday after school I went to my dance class I had loads of fun !! Today I Finally got my guitar tuned YAY !! I can now hopefully try to teach myself how to play also me and my friend [girl's name] were on Skype playing our guitars (really badly :/) !!!!!*

Although flower123's blogs do not always focus on arthritis and medicines, those that do provide genuine insights into the place, priority, and impact of arthritis and medicines for the young person. There is a series of references to "my injection", identified by the drug name in her very first entry as Enbrel (the brand name for the biological medicine etanercept) but never referred to by name again. These events seemed to have a profound effect on her. We present them here with the blog dates for reference:

> *28thSeptember 2012—Feel rubbish today :(got a cold. Had my injection last night didn't hurt at all !! YAY !!*
>
> *2ndOctober 2012—yesterday I had my first injection of the week and it didn't hurt !! YAY :)*
>
> *10thOctober 2012—on Monday I had my injection stung a bit not as bad as it has done but it's OK now*
>
> *17thOctober 2012—on Monday had my injection it didn't hurt but it was uncomfortable through my leg I could feel it spreading throughout :/*
>
> *23rdOctober 2012—yesterday I had my injection and it didn't hurt a bit YAY !!!!*
>
> *20thNovember 2012—yesterday when I had my injection it hurt a bit :/ it hurt when it was going into my leg :(*
>
> *23rdNovember 2012—but I feel better today still not 100% and I didn't go to school today :/ x had my injection yesterday when I was ill and it hurt :* [flower123, female]

After writing weekly about the first few injections, there was a gap of almost a month before the next blog. Perhaps the intervening weeks had been uneventful and straightforward, as the post-gap entries describe more painful experiences. The blog ended on November 29, so it is unclear whether the situation improved. The blogs of flower123 show that her main priorities were activities with her friends and family; her arthritis seemed not to restrict her. The injections were sometimes troubling, but not limiting.

Depth and Breadth

The quotes above demonstrate how the blogs enabled the collection of data that are contextualized within the lives, priorities, and experiences of the bloggers. These data would be difficult to obtain in either focus groups or interviews. Many of the issues brought up by individuals were repeated in subsequent blogs, sometimes enabling a longitudinal perspective in which researchers could gain insight into the dynamics of the situation from the perspective of the young person, and a temporal view of the impact of both having arthritis and the use of medicines.

The following quote illustrates how, in a single blog, many and diverse issues relating to a young person's experience of living with arthritis may be presented and illustrates the potential for examining the wider context and importance of these issues for the young person. It enables an understanding of health and medicines in the context of wider positive and negative aspects of life:

> *Arthritis is annoying because you try to put it in the back of your head but it keeps coming, every day you have to think about it because when you want to do something like eg: run but then you realise if you fall over it could be a trip to the hospital which I really don't like, not because of the staff or anything like that because you are all lovely but it's just because I have been so many times there and most of the time it has been bad news which makes it a horrible place if you know that when you go there's a chance you will come home upset and worried about the next step like an operation. I get angry when other people moan about little things like your legs ache, but I'm not saying that I feel sorry for myself all the time it's just I would swap with anyone in a heartbeat not to have this, it's horrible all the not doing stuff other people can do, but in a way I'm lucky that I can walk at all it's just that I love sports and getting involved but sometimes I can't do that which is upsetting because I want to do the things I love. It makes me sound ungrateful but I'm not I think my mom and dad are amazing with what they do with me and support me all the time so I am grateful for that but I just get annoyed when people moan because it's better than not running or even being in this earth today. If I could change myself eg: not having arthritis I don't know if I would because it has made me a stronger person and realise that not everything's good in life there are bead things on the way and I have faced some of my fears like operations and needles and I'm proud of myself for that because if I didn't have this condition then I wouldn't have become brave and I wouldn't have faced some of the things that I have faced before.* [B..a, female, age 13]

Overall, the blogs covered a huge variety of issues. The study objectives and the website provided some framework for the topics that bloggers may wish to write about, and this would inevitably impact on the data obtained. Notwithstanding, bloggers did contribute wide-ranging perspectives and experiences. Table 2 provides examples of the range of topics in the blogs.

Table 2. Themes and topics of young people's blogs.

Themes	Examples of topics	Blogs by young people, n
Identity	Self-image	11
	Feeling normal	12
	Limitation in life	12
	Relationships with peers	11
Arthritis	Physical effects	15
	Pain	10
	Psychosocial effects	22
	Beliefs about arthritis	8
Medication	Does it work?	13
	Side-effects and risks	20
	Use of specific (named) medications	19
	Obtaining information	7
Health services	About health providers/professionals	10

Young Person-Parent Dyads—Comparing Perspectives

This study also provided an opportunity to analyze data across a small number of parent-young person dyads. Data from young people and parents may relate to similar or different experiences and issues. Data may also highlight similar and potentially conflicting perspectives. Qualitative analysis may enable these viewpoints to be examined in the contexts of the differing concerns and priorities of young people and their parents. For instance, this parent-young person dyad felt differently about the school's involvement:

> *At school I know they help her either by rearranging the seating plan in the classroom so her neck is ok or by letting her sit out of certain activities if she cannot manage them and doing something else, college however is another big step – feeling a bit scary for her.* [chickflick-parent]
>
> *sometimes school does not understand, like on Tuesday we had assembly and the chairs in the rows are really close together, it hurts my legs to keep them in that position for all that time, mom keeps on telling me to tell them but I feel awkward about it, hopefully college will be different.* [pefkosfan, female, age 15]

Sharing the Experience of Living With Arthritis

For the purposes of the research, the blog site was closed. However, contributions that were effectively giving advice to others may reflect the desire of young people for a forum to interact with others with whom they share experiences:

> *The advice I would give to any young person that has arthritis would be don't give up the things you love just because you've been diagnosed. if you give up on the active things you love then you will start to give up on other things and start using it as an excuse why not to do things.* [Spacecadet, female, age 16]

Although not intended to have any "therapeutic role", several participants commented (both via their blogs and to research team members in clinic) that the experience of blogging had been helpful:

> *well today is the last day of my blog and I think that this has been a great opportunity to talk about my feelings about my arthritis and just general things really :) thank you to everyone who set this up as a really good website. :)* [flower123, female, age 14]

Another young person and their parent commented that the experience of blogging had been therapeutic during a difficult period of adjustment following their diagnosis.

Challenges of Using Blogging as a Research Tool

Overview

Despite the many benefits for using blogging to engage with young people with arthritis as discussed, there were a number of challenges with conducting research through blogging.

Non-Participation

A number of participants who declined to participate in the study gave reasons to the researchers when they were approached. These included not having enough time, not using the Internet very much, being "too lazy", busy with studying for their exams, not liking to write, and the perception of a parent that a young person would not wish to participate.

Some other reasons that were not explicitly stated, but hypothesized by the research team, include not knowing what to write on a blog, preference for outdoor rather than indoor activities, not wanting to think about their illness when they were well, and problems in accessing the site at home, for example, due to permissions that are required to run the website.

The specific reasons given by the young people may indicate how blogging in research of this kind may not be suitable for all young people.

Not Knowing What to Write

Some young people indicated to the research team that they did not know what to write about and were unsure as to whether what they were writing about would be deemed important or significant. The prompts posted by the research team on possible issues to blog about were intended to reduce this anxiety. Perhaps more reference should have been made to these prompts throughout the study, but the research team wanted to avoid "telling" the young people what to write about, as the goal of the project was to find out about was important to them.

Technical Problems

There were also some technical issues. Access to the Arthriting site was restricted by schools and workplaces with firewalls, although participants would have been able to access the site from home or through a smartphone. Access was actually restricted in the pediatric hospital itself, which reduced opportunities to demonstrate the site during the informed consent procedures.

Discussion

Principal Findings

This project offered a novel way of engaging young people in research, by creating a secure individual blog-space. Young people blogged to different extents; not all took up the blogging opportunity, but others enjoyed the experience and their blogs provided valuable insights relevant to our study objectives.

Strengths and Limitations

This methodology did present a number of limitations. First, the sample was self-selecting. Despite all eligible young people being invited to take part, many declined. The sample may reflect only those who feel competent with Web-based activities and who have reasonable literacy skills. We likely missed out on the insights of young people who do not consider Web-based resources relevant to their own condition. A recent study emphasized that not all young people choose to seek information about pain online [30]. Some of the reasons given for non-participation in our study indicate that the methodology may not suit all; this approach might be most effective if combined with other inclusive qualitative methods, such as telephone interviewing. This might promote involvement of a more representative sample. We saw substantial attrition at each recruitment stage of the study, and the expectation of a 2-month engagement might have been too ambitious.

The blogs (ie, the dataset) may also not be fully reflective of experiences and perspectives of the participants. This methodology would not necessarily be expected to provide a systematic and comprehensive dataset. For example, it could be that the young people were more likely to blog when they had problems or issues with their arthritis that made them focus on their condition and perhaps blog more. The blogs, often rich in context, need to be analyzed with this in mind, with each data item providing a perspective important to a young person at a particular point in time and in the context of other aspects of their life and priorities. The data enabled "how" and "why" questions to be explored in the analysis. When discussing the viability and relative strengths of blogging as a research methodology, it is necessary also to consider issues of reliability and validity. If we were to compare this with interview and focus group methodologies, then this method had no interviewer or transcriber bias. In the use of "audio-diaries", Sargeant and Gross [16] similarly described how they achieved "a view of how the disease fitted into individual lives".

In line with principles of qualitative inquiry, the data collection was "participant-led". Young people were in control of their involvement. They determined what they wrote about, how much they wrote, and how often they visited the site. When young people expressed difficulty regarding what they should blog about, the research team—while providing appropriate assistance (possible categories on the website)—wished to avoid influencing their blogs.

In our quest for security, the blogs were for private use by each individual and were not associated with a community where experiences and issues could be shared. This might also have been off-putting to those who wished to communicate with others. Conversely, others may have been willing to disclose feelings and experiences that they may have been reluctant to share in a Web-based community setting. The desire of some young people with a less common condition, like juvenile arthritis, to "meet" and compare experiences could be addressed by a dedicated Web-based forum. Although a young person's discussion forum does exist for young people with arthritis, this can be dominated by older adolescents and adults and is thus not appealing to the age group of "younger young people" that we engaged.

Conclusions

Web-based research with young people presents opportunities and challenges for researchers. Web-based blogging methodology has the potential to give young people and parents the space and empowerment to express their own ideas and concerns. However, this project suggests that it might not be the best way to engage a large diverse group of young people and might most effectively be combined with other approaches. Despite this, the study provided valuable data on the experience and impact of living with a long-term condition from the perspectives of young people with arthritis.

Acknowledgments

This research project was commissioned by the Pharmacy Practice Research Trust (PPRT) and completed in association with Birmingham Children's Hospital and Arthritis Care. The PPRT merged with the Pharmaceutical Trust for Educational and Charitable Objects (PTECO) in December 2012 to form Pharmacy Research UK.

We would like to thank all the young people and parents who took part in the Arthriting project and especially the young advisers who helped to develop the blog website and who coined the title “Arthriting”.

We would like to acknowledge the work of the Research Management Group of the Arthriting Project and specifically of our colleagues Kevin Harvey (University of Nottingham, Karen Shaw (University of Birmingham), Rachel Stephenson (Birmingham Children’s Hospital NHS Foundation Trust), David Terry (Birmingham Children’s Hospital NHS Foundation Trust), Kate Fleck (Arthritis Care), and Rachel Roberts (Pharmacy Research UK).

We would also like to thank the Web designers, Ambinet, who worked with the young advisers and the project team, and Marie McGee (adolescent rheumatology transition co-ordinator at Birmingham Children’s Hospital) who supported the young advisers’ Web development meetings.

Pharmacy Research UK funded the evaluation from which this paper and the data included are derived. The views and opinions expressed herein are those of the authors and do not necessarily reflect those of Pharmacy Research UK.

Authors' Contributions

All authors have contributed to each section within this paper and made additional edits to the paper.

Conflicts of Interest

None declared.

References

1. Hewson C. Conducting research on the Internet. The Psychologist 2003;16(6):290-293.
2. Wright KB. Researching internet-based populations: advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. Journal of Computer-Mediated Communication 2005;10(3):11. [doi: 10.1111/j.1083-6101.2005.tb00259.x]
3. Corley M, Scheepers C. Syntactic priming in English sentence production: Categorical and latency evidence from an Internet-based study. Psychonomic Bulletin & Review 2002 Mar;9(1):126-131. [doi: 10.3758/BF03196267]
4. Kehoe C, Pitkow J. The World Wide Web Journal (Online). 1996. Surveying the territory: GVU&#39;s five WWW user surveys URL: http://www.cc.gatech.edu/gvu/user_surveys/papers/w3j.html [accessed 2014-09-04] [WebCite Cache ID 6SL5Hcrhi]
5. Schmidt WC. World-Wide Web survey research: Benefits, potential problems, and solutions. Behavior Research Methods, Instruments, and Computers 1997 Jun;29(2):274-279. [doi: 10.3758/BF03204826]
6. Cook C, Heath F, Thompson RL. A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys. Educational and Psychological Measurement 2000 Dec 01;60(6):821-836. [doi: 10.1177/00131640021970934]
7. Raziano DB, Jayadevappa R, Valenzula D, Weiner M, Lavizzo-Mourey R. E-mail Versus Conventional Postal Mail Survey of Geriatric Chiefs. The Gerontologist 2001 Dec 01;41(6):799-804. [doi: 10.1093/geront/41.6.799]
8. Beck CT. Benefits of participating in internet interviews: women helping women. Qual Health Res 2005 Mar;15(3):411-422. [doi: 10.1177/1049732304270837] [Medline: 15761108]
9. Davis M, Bolding G, Hart G, Sherr L, Elford J. Reflecting on the experiences of interviewing online: perspectives from the Internet and HIV study in London. AIDS Care 2004;16(8):944-952. [doi: 10.1080/09540120412331292499]
10. Franklin KK, Lowry C. Computer-mediated focus group sessions: naturalistic inquiry in a networked environment. Qualitative Research 2001 Aug 01;1(2):169-184. [doi: 10.1177/146879410100100204]
11. Hookway N. 'Entering the blogosphere': some strategies for using blogs in social research. Qualitative Research 2008 Feb 01;8(1):91-113. [doi: 10.1177/1468794107085298]
12. Olive R. 'Making friends with the neighbours': Blogging as a research method. International Journal of Cultural Studies 2012 Apr 26;16(1):71-84. [doi: 10.1177/1367877912441438]
13. Gauntlett D. Web studies. London: Arnold; 2004.
14. Elliot H. Sociological Research Online. 1997. The use of diaries in sociological research on health experience URL: http://www.socresonline.org.uk/2/2/7.html [accessed 2014-09-04] [WebCite Cache ID 6SL6Dd4vW]
15. Johnson J, Bytheway B. An evaluation of the use of diaries in a study of medication in later life. International Journal of Social Research Methodology 2001 Jan;4(3):183-204. [doi: 10.1080/13645570010029467]
16. Sargeant S, Gross H. Young people learning to live with inflammatory bowel disease: working with an “unclosed” diary. Qualitative Health Research 2001;21(10):1360-1370. [doi: 10.1177/1049732311407211]
17. Verbrugge LM. Health diaries. Med Care 1980 Jan;18(1):73-95. [Medline: 6986517]
18. Pew Research Center. Report by Pew Internet and American Life Project. Washington, DC; 2011. Health topics URL: http://www.pewinternet.org/2011/02/01/health-topics-2/ [WebCite Cache ID 6SL6p2Mdd]

19. Malik S, Coulson NS. The therapeutic potential of the internet: exploring self-help processes in an internet forum for young people with inflammatory bowel disease. Gastroenterol Nurs 2011;34(6):439-448. [doi: 10.1097/SGA.0b013e318237a9ba] [Medline: 22129797]
20. Coulson NS, Knibb RC. Coping with food allergy: exploring the role of the online support group. Cyberpsychol Behav 2007 Feb;10(1):145-148. [doi: 10.1089/cpb.2006.9978] [Medline: 17305463]
21. Tanis M. Health-related on-line forums: what's the big attraction? J Health Commun 2008;13(7):698-714. [doi: 10.1080/10810730802415316] [Medline: 18958781]
22. Bortree DS. Presentation of self on the Web: an ethnographic study of teenage girls' weblogs. Education, Communication & Information 2005 Mar;5(1):25-39. [doi: 10.1080/14636310500061102]
23. Lenhart A, Madden M. Report by Pew Internet and American Life Project. Washington, DC; 2005. Teen content creators and consumers URL: http://www.pewinternet.org/ [accessed 2014-09-18] [WebCite Cache ID 6SgGEuU6Y]
24. Livingstone S, Helsper E. Gradations in digital inclusion: children, young people and the digital divide. New Media & Society 2007 Aug 01;9(4):671-696. [doi: 10.1177/1461444807080335]
25. Gray NJ, Klein JD, Cantrill JA, Noyce PR. Adolescent girls' use of the Internet for health information: issues beyond access. J Med Syst 2002 Dec;26(6):545-553. [Medline: 12385536]
26. Pedersen S, Macafee C. Gender differences in British blogging. Journal of Computer-Mediated Communication 2007;12(4):1472-1492 [FREE Full text] [doi: 10.1111/j.1083-6101.2007.00382.x]
27. Liu EZF, Chang YF. Gender differences in usage, satisfaction, self-efficacy and performance of blogging. British Journal of Educational Technology 2010;41(3):39-43. [doi: 10.1111/j.1467-8535.2009.00939.x]
28. Department of Health. London: Department of Health. 2013. Chief Medical Officer&#8217;s annual report 2012: Our Children Deserve Better URL: http://www.cc.gatech.edu/gvu/user_surveys/papers/w3j.html [accessed 2014-09-04] [WebCite Cache ID 6SL5Hcrhi]
29. Tong A, Jones J, Craig JC, Singh-Grewal D. Children's experiences of living with juvenile idiopathic arthritis: a thematic synthesis of qualitative studies. Arthritis Care Res (Hoboken) 2012 Sep;64(9):1392-1404. [doi: 10.1002/acr.21695] [Medline: 22504867]
30. Henderson EM, Keogh E, Rosser BA, Eccleston C. Searching the internet for help with pain: adolescent search, coping, and medication behaviour. Br J Health Psychol 2013 Feb;18(1):218-232. [doi: 10.1111/bjhp.12005] [Medline: 23126577]

Abbreviations

IBD: inflammatory bowel disease
JIA: juvenile idiopathic arthritis
PPRT: Pharmacy Practice Research Trust
PTECO: Pharmaceutical Trust for Educational and Charitable Objects

---

Edited by G Eysenbach; submitted 16.06.14; peer-reviewed by P Ressler, R Booth; comments to author 25.08.14; revised version received 24.09.14; accepted 08.11.14; published 05.03.15.

Please cite as:
Prescott J, Gray NJ, Smith FJ, McDonagh JE
Blogging as a Viable Research Methodology for Young People With Arthritis: A Qualitative Study
J Med Internet Res 2015;17(3):e61
URL: http://www.jmir.org/2015/3/e61/
doi:10.2196/jmir.3608
PMID:25749691

Original Paper

Improving Self-Help E-Therapy for Depression and Anxiety Among Sexual Minorities: An Analysis of Focus Groups With Lesbians and Gay Men

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

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

Corresponding Author:
Tomas Rozbroj, BA(Hons)
Australian Research Centre in Sex, Health and Society
La Trobe University
215 Franklin Street
Melbourne, 3000
Australia
Phone: 61 9479 8700
Fax: 61 9479 8711
Email: t.rozbroj@latrobe.edu.au

Abstract

Background: E-therapies for depression and anxiety rarely account for lesbian and gay users. This is despite lesbians and gay men being at heightened risk of mood disorders and likely to benefit from having access to tailored self-help resources.

Objective: We sought to determine how e-therapies for depression and anxiety could be improved to address the therapeutic needs of lesbians and gay men.

Methods: We conducted eight focus groups with lesbians and gay men aged 18 years and older. Focus groups were presented with key modules from the popular e-therapy “MoodGYM”. They were asked to evaluate the inclusiveness and relevance of these modules for lesbians and gay men and to think about ways that e-therapies in general could be modified. The focus groups were analyzed qualitatively using a thematic analysis approach to identify major themes.

Results: The focus groups indicated that some but not all aspects of MoodGYM were suitable, and suggested ways of improving e-therapies for lesbian and gay users. Suggestions included avoiding language or examples that assumed or implied users were heterosexual, improving inclusiveness by representing non-heterosexual relationships, providing referrals to specialized support services and addressing stigma-related stress, such as “coming out” and experiences of discrimination and harassment. Focus group participants suggested that dedicated e-therapies for lesbians and gay men should be developed or general e-therapies be made more inclusive by using adaptive logic to deliver content appropriate for a user’s sexual identity.

Conclusions: Findings from this study offer in-depth guidance for developing e-therapies that more effectively address mental health problems among lesbians and gay men.

(J Med Internet Res 2015;17(3):e66) doi:10.2196/jmir.4013

KEYWORDS

Internet therapy; e-therapy; cCBT; mental health; gay men; lesbian; minority stress; depression; anxiety; focus groups

Introduction

E-therapies are programs that utilize the Internet or mobile phones to deliver interactive interventions for preventing and treating depression, anxiety, and other mental health problems. E-therapies most commonly utilize cognitive behavioral therapy (CBT) [1-3], are typically undertaken over several weeks or months, and involve users completing modules or exercises while receiving feedback on their progress. E-therapies show considerable clinical benefits, especially in the treatment of depression and anxiety [2,4,5], and have become an increasingly important part of strategies to address mood disorders in a number of countries, such as Australia [1,6], United Kingdom [7], and New Zealand [8].

An important advantage of e-therapies is their accessibility, especially for marginalized populations [9,10], such as lesbians and gay men [11,12]. Same-sex attracted people often report stigma and other negative experiences as barriers to accessing traditional health care [13-15]. When they do access treatment, they are also less likely than the general population to report feeling satisfied [13,16]. E-therapies have potential to resolve these issues, as they are accessible from anywhere, anonymous, and typically available for free or at low-cost, thus making them an attractive alternative to face-to-face therapy. Furthermore, e-therapies have potential for delivering tailored content to address specific issues faced by lesbians and gay men, such as "coming out", or disclosing their sexual orientation to family, friends, and co-workers, as well as managing experiences of discrimination and the challenges of engaging in stigmatized same-sex relationships.

Improving access to mental health care is further warranted given that same-sex attracted populations have disproportionately high rates of depression and anxiety [13,17-20], which are largely the result of chronic stress that arises from ongoing experiences of stigma, discrimination, and marginalization [21,22]. In Australia, the 2007 National Survey of Mental Health and Well-Being found that same-sex attracted people were two times more likely to experience anxiety and three times more likely to experience depression [23]. In the United Kingdom, the 2012 Stonewall Report found 22% of gay and bisexual men experienced moderate to severe depression, compared to 9% of the population in general [14,18]. In the United States, there is considerable variance between states, but a 2014 American Psychological Association (APA) report concluded that lesbian, gay, and bisexual people are twice as likely on average to report a mental health disorder [17]. A range of studies from a host of countries also shows that lesbians and gay men are at increased risk of substance abuse problems [13,24] and suicide [18,25].

Despite potential for addressing mental health issues among lesbians and gay men, e-therapies currently fail to cater to these populations. Organizations such as the National Health Service (United Kingdom) [15], Beyondblue (Australia) [26], and the World Health Organization [27,28] are advocating for mental health responses that specifically address the needs of non-heterosexual people. However, a recent review of 24 Web- and mobile phone-based e-therapies found only a few e-therapies included any content to address the specific experiences of being same-sex attracted, such as coming out, same-sex relationships, or stigma-related challenges [29]. Less than a fifth (17%) had gay or lesbian-specific mental health referrals such as helplines, and more than half (58%) used language and examples that assumed or suggested users were heterosexual. Thus, currently available e-therapies may not be as effective as they could be for lesbians and gay men, who may not relate to or may even feel alienated by content that appears designed for heterosexual audiences [29].

To inform the development of tailored e-therapies, we explored possible ways in which e-therapies could be tailored to better cater to the needs of lesbians and gay men. We conducted eight focus groups involving lesbians and gay men with the aim of identifying key areas in which e-therapies could be improved to make them more inclusive, relevant, and appealing to lesbians and gay men. It is important to note that many of the findings reported in this article may also apply to other non-heterosexual populations, such as bisexual men and women. However, these and other sexual identity groups face challenges that are often different from those of lesbians and gay men [30,31], and therefore merit separate studies. It was unfortunately beyond the scope of this study to include all sexual identity groups, so we have focused on lesbians and gay men as a starting point.

Methods

Sample

The sample comprised 32 participants, consisting of 14 women and 18 men, and an average of four participants per focus group. All participants indicated that they were lesbian or gay and were aged 18 years or older. A total of 15 participants were aged 35 years or older and 17 were aged between 18 and 34 years. The mean participant age was 34 years (SD 15). The sample was highly educated on average, with 20 reporting having a university degree. At the start of each focus group, participants were asked to indicate their level of knowledge of e-therapies on a five-point scale (1: I have no knowledge at all, 2: I have very little knowledge, 3: I have moderate knowledge, 4: I have good knowledge, 5: I have excellent knowledge). The mean level of knowledge for the sample was 2 and the median was 3.

Recruitment

To take part in the focus groups, participants were required to meet three selection criteria. They had to be aged 18 years or over, fluent in English, and identifying as lesbian or gay. The sample was recruited purposively through a mix of advertising and snowballing. Advertisements were placed on the Web (eg, Facebook, gay and lesbian-related websites, and the La Trobe University website), in print media (eg, magazines that targeted lesbians and gay men), and on a free-to-air radio station that targets lesbians and gay men in Melbourne, Australia. Advertisements were also included in the newsletters of gay and lesbian support organizations. All advertisements directed prospective participants to a registration website where they provided their contact details and confirmation that they met the selection criteria. In total, 111 eligible prospective participants registered their interest. The first author divided registrants into the four demographic categories (see Procedure section), from which he randomly invited registrants to focus groups, with the aim of having approximately four participants per group. Reimbursement of A$30 was offered for participation.

Procedure

This study received ethics approval from La Trobe University Human Research Ethics. The focus groups were held at the Australian Research Centre in Sex, Health and Society at La Trobe University in Melbourne. Eight focus groups were conducted, comprising two groups for each of the following four demographic categories: (1) younger lesbians aged 18 to 34 years (YF), (2) younger gay men aged 18 to 34 years (YM), (3) older lesbians aged 35+ years (OF), and (4) older gay men aged 35+ years (OM).

These four categories were chosen in light of research that shows key differences in the mental health and experiences of sexuality-related stigma among same-sex attracted people according to age [32,33] and sexual identity [34]. An age limit of 35 years was set for the younger group to account for many early-life milestones that may have recently been experienced by this group and can be potential sources of stigma-related challenges, such as coming out for the first time, gaining a higher education, entering the workforce, and establishing long-term relationships. Limiting the number of focus groups to eight was deemed appropriate, as similar ideas were expressed between all groups, thus suggesting that additional focus groups were unlikely to result in further substantive data. Focus groups were conducted between October and November 2013. Each group took up to 90 minutes and was digitally audio-recorded.

The aim of the focus groups was to evaluate the needs and experiences of lesbians and gay men in the provision of depression and anxiety e-therapies. To facilitate discussion, participants were asked to examine a series of case studies drawn from the Australian e-therapy “MoodGYM”, which follows the lives of six characters as it takes users through a structured therapeutic program comprising reading, exercises, and evaluative tools [35,36]. Excerpts from MoodGYM were chosen as stimulus material because it is one of the most prominent e-therapies [37], has been repeatedly proven to be clinically effective, and is included in the Substance Abuse and Mental Health Services Administration (SAMHSA) Registry of Evidence-Based Programs and Practices [38], and its material typifies features found in many prominent e-therapies, such as the use of characters as examples, a CBT modality, and interaction with users’ everyday experiences as part of therapy.

All focus groups were facilitated by the first author, who is male, and has prior training and experience in conducting focus groups with men and women. At the beginning of each focus group, participants were briefed by the facilitator about his research background, key information about e-therapies, and mood disorders. The facilitator did not disclose his sexual orientation to the focus groups. Each focus group was then presented with the following printed MoodGYM case studies, in order: a set of character profiles and stories/examples that featured the characters; a questionnaire designed to identify dysfunctional beliefs and attitudes related to relationships and other life experiences, known as the Warpy Thoughts Scale [39]; and excerpts from a module that addresses mood problems related to romantic relationships, known as the Relationships Module. For each case study, participants were asked to discuss the material with regard to its applicability to lesbians and gay men. Additional topics were also discussed that did not require excerpts from MoodGYM, such as whether referrals were needed to mental health services that explicitly targeted same-sex attracted people. After discussing MoodGYM, participants were then given opportunities to discuss content they thought should be included in e-therapies more generally.

Analysis

The focus group discussions were first transcribed using an independent transcription service. All transcripts were subsequently checked for accuracy by the first author. We then conducted a thematic analysis of the data [40]. The first two authors of this article were the researchers who were primarily responsible for the analysis. Both researchers first read the transcripts independently and then met to discuss the main themes they had identified. Both researchers had arrived at similar broad themes. Following discussion, the first-named researcher developed a more detailed coding system, which informed the final analysis. The first-named researcher then coded all transcripts in detail, and codes were recorded in a spreadsheet. The first-two named researchers met regularly throughout the coding process to review the coding system and to discuss emerging themes. Through this process, four main themes were identified. These were: making e-therapies more inclusive; making e-therapies more relevant to stigma-related challenges faced by lesbians and gay men, including vulnerable subpopulations; and preferences for delivering tailored content. In this article, we report on the main themes and have included quotes to further illustrate key aspects of each theme. The first two named researchers met regularly to discuss the choice of quotes and to ensure that those selected were consistent with the data they represented, were typical of comments made by participants in the focus groups, or provided further relevant detail. Because data was collected in focus groups, it was not always possible to identify the voices of individual participants in audio recordings. The quotes presented are therefore identified according to the group, such as “YF” to indicate a quote from a younger lesbian focus group, rather than according to individual participants. The remaining authors of this article scrutinized the final analysis, including the choice of quotes, and suggested further refinements to guide the interpretation of the findings. For this analysis, participants did not provide feedback on the findings.

Results

Overview

Although we divided focus groups into four demographic groups (as outlined earlier), the groups generally gave similar feedback and suggestions. Unless otherwise specified, all of the findings reported below should be read as applying across the focus groups and not limited to particular age or sexual identity categories.

Making E-Therapies More Inclusive

All focus groups indicated that it was important for e-therapies to acknowledge and represent lesbians and gay men. One of the ways in which participants felt excluded from the MoodGYM examples was through the language used. Focus groups suggested that it would be better if neutral language was used, for example avoiding words like “spouse” and using words like “they” instead of “he” or “she”. Focus groups also suggested that examples, such as images, stories, or scenarios, needed to include same sex-attracted people to be relevant:

> *It's about being included and it's about identifying with the character, and the specific experience of being lesbian is very different from that of a straight woman or a gay man or a heterosexual man. It's important to be represented.* [OM]
>
> *It's heterosexual and I don't relate to it.* [OF]

Groups largely did not support the cast of MoodGYM characters, which they felt did not represent non-heterosexual issues. They had varied views as to how best to achieve a representative cast of characters. Suggestions ranged from using a cast of mixed-sexuality characters to using androgynous characters:

> *I would probably either go with… a mix of gay and straight [characters], or just random men and women that go on dates with people, with androgynous partners.* [YF]
>
> *[Androgynous] would work for me in certain situations. But I think a lot of people would also like to identify with the person… some people would rather have that kind of personalized connection.* [YM]
>
> *I think I would want at least fifty percent to be not heterosexual, like I know that's not in line to statistics in society or whatever but to feel like it wasn't just the stereotypical one or token one or whatever… I'd feel like you'd need quite a breadth of representation.* [YM]

Focus groups also suggested that inclusiveness could be improved by making lesbians and gay men feel welcome at the beginning of the e-therapy program by explicitly mentioning them upon sign-up:

> *I'm thinking that in terms of this service, the e-therapy, from the initial, if they use more inclusive terminology and things… that might help to improve the uptake.* [OF]

Another participant in that group added:

> *It's an enormous first step up. It's like you respect me and you're going to treat me and help me... So if you've included me on the form [referring to early in the e-therapy], therefore you care.* [OF]

Including a rainbow flag, which is a universal indicator of a gay and lesbian inclusive service, was one further suggestion. However, focus groups were reticent about specifically singling out sexuality at the beginning of an e-therapy that is aimed toward general users, instead preferring a broader statement of inclusiveness that included sexuality alongside gender, ethnicity, and other diverse populations.

Though suggestions about how to achieve inclusiveness varied, a consistent underpinning theme for all focus groups was that inclusion is needed. Avoiding language and content that suggests or assumes users are heterosexual was a major theme in all focus groups, and having some direct references to lesbians and gay men, such as including a gay character in examples and stories, was a common preference.

Making E-Therapies More Relevant to Lesbians and Gay Men

Addressing stigma and stigma-related issues, such as coming out, were the most prominent type of content requested by focus groups for making e-therapies more relevant to lesbians and gay men. This included content to assist users to manage and address public stigma, such as discrimination and prejudice, as well as self-stigma, such as experiences of shame or having negative beliefs and attitudes toward their own sexuality [41,42]. Groups recurrently flagged both kinds of stigma as a cause of mood disorder issues, tended to react favorably to MoodGYM examples that they perceived as relevant to addressing stigma, such as some items in the Warpy Thoughts Scale, and made numerous suggestions for content to be added to e-therapies that targeted stigma-related issues. With regard to public stigma, focus groups flagged several forms of stigma that they suggested should be addressed in e-therapies:

> *Homophobia manifests itself in so many ways, and there's so many kinds of homophobia, and you know there's overt homophobia, there's covert homophobia, there's implied homophobia, there's homophobia that happens when no one intends to be homophobic at all… I think a lot of the things that cause depression and anxiety in the gay community are linked to one of those many kinds of homophobia that's out there.* [OM]
>
> *Just being part of a society that is overall really heteronormative and you know like constantly not being part of things and not being necessarily discriminated against as such but like finding lack of representation or just like kind of coming up against things that don't fit you or your family or your friends or the way that you live, that sort of thing.* [YF]

A number of focus groups suggested that e-therapies needed to provide a safe space from stigma in order to be effective:

> *I think if you want something like this to work, you have to provide a forum where people feel like they're in a comfortable environment, and you know I mean a lot of people want some respite from the straight world, and they're particularly going to want that when they're dealing with issues to do with mental health.* [OM]

Coming out was a key issue. Groups indicated that coming out should be addressed as an ongoing process of deciding whether to disclose or conceal, or "coming out" and "coming in", as the Working Therapeutically with LGBTI Clients: A Practice Wisdom Guide has termed it [43]:

> *You're coming out every day; always deciding if it's safe to come out.* [YM]
>
> *You often need to hide your relationship, which would be good to cover [in e-therapy].* [YM]

With regard to internalized stigma, groups also expressed a need to address issues of shame and negative self-perceptions:

> *I've definitely felt that [internalized homophobia] and I think that, you know, if there was more focus*

> *on you know accepting who you are, and you know ways to develop yourself that are outside of, you know, just getting a partner or getting a relationship or whatever then that would be a good way to go I think.* [OF]

> *Internalized homophobia is a big issue… you can catch yourself every now and then thinking 'oh well if I wasn't gay I would be doing this, I would be better and I would be'; it's… how we think how other people see us... And it comes down to just having enough self-worth. Say well… this is me and I am getting on with my life. That's all I would ever want.* [OM]

Participants also highlighted a range of issues that arose from within gay and lesbian communities, and portrayed these communities as sources of stigma and exclusion as well as sources of strength and resilience. This reflects previous research that identifies subcultural groups within gay communities that have potential for marginalization within these communities [44]. Older age, especially among men, can also be stigmatized within the gay community [32,33,45]. Thus, e-therapies may need to avoid addressing stigma as an issue that only comes from outside gay and lesbian communities:

> *I think [ageing] that's one of the biggest issues of all… You know, it's sort of like people are just, 'oh you are only an old poof', this sort of business. You know it's a throwaway thing but it's bullying and it hurts.* [OM]

In response to MoodGYM's relationships module, focus groups commented on the lack of relevance to lesbian and gay relationships:

> *It has to be specific, I mean if I was to go online for relationship breakup assistance, and I was to see an icon of a man and an icon of a woman, I would just think this is aimed at straight people. It's going to have no concept of same-sex relationships, or issues, or dialogue at all, so what's the point?* [OM]

They specifically commented on the large attention given to breakups and tensions within relationships that characterized examples presented within the relationships module. While relevant, focus groups pointed out the lack of attention given to the pressures faced in a healthy but stigmatized relationship, and suggested content and support to address these additional pressures:

> *It's tough to be seen holding your girlfriend's hand in public... you kind of feel constantly on edge in public… It's not the relationship itself; it's the context that it occurs in. And the context is hostile a lot of the time.* [OF]

> *Surprised there's just breakup; there's as much anxiety about starting a relationship. There should be [more than just breakup].* [YM]

Groups further stressed that while the issues they face in the breakup of a relationship are fundamentally the same as for everyone else, they are faced with additional challenges related to the acceptance of their relationships, such as from family:

> *Family should be covered. Aren't I getting married? When can we expect grandchildren?* [OF]

> *My family, they don't view it as a real relationship, to them it's nothing.* [OF]

To further improve the relevance of e-therapies, focus groups also suggested a need for e-therapies to include helplines that cater to the specific needs of lesbians and gay men, and provide a safe space for talking about mental health issues:

> *Yes, [helplines] must be tailored. There's extra issues. You need someone who understands what you're going through. It's duty of care to provide such resources.* [OF]

> *[Helplines] should be targeted, and more available. Perhaps online chat would be good.* [YM]

While including tailored telephone helplines was the most common suggestion, focus groups pointed to other resources such as websites or referrals to face-to-face services that specialize in the mental health of same-sex attracted people. In particular, younger groups highlighted the positive role that online communities have had in breaking down a sense of isolation and in creating support networks, thus further highlighting a need for e-therapies to provide tailored referrals.

Including Vulnerable Subpopulations

When discussing the merits of tailoring e-therapy to lesbians and gay men, focus groups suggested a need for therapies to be inclusive and relevant to several subpopulations. Age was a commonly discussed factor. In particular, the younger focus groups suggested a need for e-therapies to give special attention to young people while the older groups mentioned the importance of addressing aging-related issues. Pressures around bullying at school, coming out, and developing an identity were mentioned as important issues for young people, while the death of a partner and finding gay-friendly aged care facilities were mentioned as issues to address for older people. Rural lesbians and gay men are at greater risk of mental health problems [46] and these populations were also flagged by the focus groups as needing specific attention; the pressures of "being the only queer in the village", as a young female put it. Some participants saw particular value in Web-based modes of communication in reaching these populations:

> *Coming from the country where I was the only open person in a 60 kilometer radius. And I know that, and not because of Grindr [a mobile phone app that facilitates dating among same-sex attracted men], but because I knew everyone in a 60 kilometer radius. And so it was kind of like having the online resources and social media, you know, connecting with people.* [YM]

Focus groups also suggested that e-therapies take into account lesbians and gay men from ethnic minority backgrounds. A common point was that some of those with ethnic minority backgrounds face additional challenges in gaining acceptance from others, particularly if they are from traditional or religious cultures. Gaining acceptance within gay and lesbian communities was also flagged as a challenge for some. More generally, those with an ethnic minority background wanted to feel included in e-therapies:

> *My ethnic background is very tightly aligned with my difficulties with my sexuality and so if I were to see a white female who was same-sex attracted [as e-therapy content], I'd feel like there were a lot of exceptions to me with my case and it might not be a very good tool.* [YF]

Preferences for Delivering Tailored Content

When asked about how best to achieve inclusiveness and relevance, all focus groups suggested that some degree of tailoring was required. Many wanted separate e-therapies that were specifically dedicated to lesbians and gay men:

> *The generalist approaches that works for people, with the gay and lesbian add-on bits, based on what comes out of your research, could be just fine, but I think something specifically tailored would be better.* [OM]

> *I'd be more inclined to use the specific one, [to which another participant added]: It would be more relevant and accurate. I'd be more trusting of one.* [OF]

> *I personally would prefer to use like a specific one if it, like specific to gays and lesbians, if it is linked to me being a lesbian. Like it's an issue I face because I am a lesbian. But if I am having a separate issue, more general to you know, I don't know, my everyday life, which is nothing to do with me being a lesbian, then I would opt for the general one.* [YF]

While tailoring was broadly supported, there was some concern that creating a specific e-therapy for lesbians and gay men could reinforce marginalization by suggesting that lesbians and gay men could not be included within general e-therapies:

> *If you're making a totally separate program and separate website then that just kind of enhances the feeling of 'okay you guys are separate, you guys are different, so we need something special for you'.* [OF]

> *I have worked so hard all my life to have equal friends and relationships. I wouldn't want to go back to 'John and Betty get married and Chris and Philesco over here and they live there completely separate'… Everybody together is what we're trying to aim for.* [OF]

Some focus groups suggested that general e-therapies might instead use adaptive logic in which content on sexuality and sexuality-related issues is directed to a user's sexual identity:

> *Why don't you customize? When you log on you customize your own.* [OF]

> *[Tailoring using adaptive logic] That sounds awesome.* [YF]

Further extending adaptive logic, some focus groups also suggested incorporating the use of avatars: personalized representations of a user. Groups suggested that inclusion and relevance could be enhanced by allowing users to customize their MoodGYM characters to experience e-therapy as a personalized story that responds to their particular identity:

> *It should be like The Sims and you should be able to create your own character.* [YF]

> *That will work for everyone if they have enough kind of menu items to choose from.* [OF]

In all, whether e-therapies use adaptive logic or separate dedicated e-therapies are developed for lesbians and gay men, some form of tailoring that delivers targeted content for addressing issues specifically faced by lesbians and gay men was strongly supported in all focus groups.

Discussion

Principal Findings

Focus group participants offered a range of insights into how e-therapies could be made more applicable to lesbians and gay men. They articulated a need to address stigma-related challenges, including those from within gay and lesbian communities. This encompassed a need for e-therapies to address same-sex relationship issues in ways that take account of challenges such as stigma and rejection of the relationship, lack of support during breakups, and family rejection or misunderstanding. Attention also needs to be given to language, imagery, and examples when making e-therapies more inclusive and relevant to same-sex attracted people. Incorporating content that is inclusive and relevant to particular subpopulations, such as those living in rural areas and from minority ethnic backgrounds, may also deserve consideration. In addition, focus groups pointed to a need for making e-therapies more relevant by including helplines and other resources that provide appropriate follow-up options for lesbians and gay men.

All focus groups expressed a need for e-therapies to involve some form of tailoring to lesbians and gay men. On balance, having e-therapies available that are specifically targeted to non-heterosexual users received the greatest support. However, there was some tension around this. Some participants were concerned that providing separate content for lesbians and gay men draws a problematic segregation between heterosexual and non-heterosexual and may therefore run counter to the principle of inclusiveness. Other participants argued that the needs of lesbians and gay men are significantly different, and that content should therefore be tailored explicitly and specifically in order that e-therapies adequately address the needs of these populations. In fact, having an explicit acknowledgement of lesbians and gay men appears to be particularly important because "lack of acknowledgement" and "invisibility" were mentioned in every focus group and were highlighted as major stressors by participants. A large number of comments in focus groups also referred to a need to be specifically spoken to by e-therapies.

While having separate e-therapies for lesbians and gay men is one option, developers could consider fine-tuning e-therapies aimed at the general population, such as MoodGYM, so that they are also inclusive and relevant to lesbians and gay men. Adaptive logic is one mechanism that can be used to achieve this. Adaptive logic presents different content to users based on their responses to questions asked in the e-therapy. Adaptive logic is already a common feature of e-therapies, but is presently not being used to tailor e-therapies for different sexual orientations. That said, the issue of segregating users also applies here. Thinking about subtle ways of delivering targeted content according to a user's sexual orientation may be advisable rather than having clearly demarcated areas of an e-therapy for heterosexual and non-heterosexual users.

The various suggestions that emerged from the focus groups should be evaluated in the context of participants responding favorably to many aspects of the MoodGYM case studies. The request to provide extra content to address additional stressors on top of largely effective general content is congruent with experiences described by Minority Stress Theory [21,47], which is currently the dominant model to account for the mental health implications of living with a stigmatized identity. Minority Stress Theory proposes that stigmatized populations often withstand additional stress derived from their minority status on top of general life stress that people of all backgrounds may face. Some suggestions, like including references to helplines specific to same-sex attracted people, are low-cost and easy to implement improvements. Others, like the delivery of targeted content on stigma-related stress, are more complex. However, on the whole, we believe that the feedback offered in our focus groups, read in the context of broader recommendations for tailoring mental health strategies to non-heterosexual people, may prove useful for e-therapy developers.

A paper by Lucassen et al [11] is, to our knowledge, the only other published work that qualitatively assesses the views of same-sex attracted people on tailoring computerized CBT programs (cCBT), although it relates to a CD-ROM program, "SPARX", rather than an e-therapy. Lucassen et al identified coming out and other stigma-related challenges such as obtaining family support as major problems faced by their sample, and also identified coping mechanisms, including social support. Their sample also responded positively to a "rainbow" version of the SPARX program, tailored to same-sex attracted people. Major themes and suggestions offered by participants in our study correlate closely with those identified by Lucassen et al. Furthermore, suggestions made by our focus groups fit closely with findings from other research that examined the experiences of same-sex attracted people within broader health care services. For example, issues around non-inclusive language and assumptions about relationship structures that do not account for same-sex attracted people are some of the key areas highlighted in previous research on health care services as problems needing to be addressed [43,48-50], and that also emerged in our focus groups.

Suggestions from our focus groups on how to improve e-therapies are also consistent with broader work aimed at adjusting mental health care for non-heterosexual clients. For example, major policy documents about mental health care, such as those from Stonewall and the UK Department of Health [14,15], and "Going Upstream: A framework for promoting the mental health of lesbian, gay, bisexual, transgender and intersex (LGBTI) people" [28] from Australia, all fundamentally advocate for tailored responses for same-sex attracted people. Recommendations in these documents are closely related to themes and suggestions that emerged in our focus groups, such as the need to address stigma-related challenges, to specifically address same-sex relationships, to account for the needs of vulnerable subpopulations, and to provide relevant referral services.

Limitations and Future Directions

Our sample was highly educated compared to the general Australian population. Education is strongly linked to socioeconomic status, and there is potential for the feedback we received from the focus groups to reflect the ideas and concerns of a higher socioeconomic group. Notably, many participants seemed to already be engaged in broader debates related to lesbian and gay mental well-being and entered our discussion with a depth of existing perspectives to draw on. Future work may need to be conducted with lower socioeconomic groups, as access to mental health services and experiences of stigma-related stress may be different.

Another possible limitation is that data about the participants' history of mood disorders was not sought. The topics discussed were about e-therapies' relevance to the needs and experiences of lesbians and gay men, and it was important to understand how this can best be achieved for the lesbian and gay population in general. Therefore, it would not have been optimal to target only those with previous experience of depression or anxiety. However, it is possible that participants with a history of mood disorders may have had different perspectives on the topics discussed, and future studies may wish to take this into account. Researchers who conduct future studies that are similar to our study may also wish to consider the potential role of ethnicity. We did not collect data on the ethnic backgrounds of participants, but the intersection of ethnic and sexual identities was raised as an issue by participants in our study. Having data on the ethnic backgrounds of participants may therefore be useful for further contextualizing the findings that emerge from future research.

Focus groups were also limited to a discussion of MoodGYM rather than other specific e-therapies. Decisions about which parts of MoodGYM were used as case studies were undertaken with the aim of generating findings that could be generalized to other e-therapies. Furthermore, we believe MoodGYM to be a good prototype for our study because it is prominent, well-evidenced, and is often regarded as a benchmark for other e-therapies. It also uses "scenarios", which are commonly found in well-evidenced e-therapies, but can make it challenging to incorporate content that is inclusive of diverse populations [29]. Nevertheless, our findings may have been a little different had another e-therapy been used as a case study, and any generalization of our findings to other e-therapies entails a degree of speculation.

Finally, we acknowledge that the scope of this study was limited by encompassing only lesbians and gay men. The exclusion of other non-heterosexual groups, as well as sex and gender-diverse groups (such as intersex, gender questioning, and transgender people), was decided because each group has different health outcomes and often has different life experiences to those of lesbians and gay men [30,31]. For example, bisexual men and women sometimes report feeling excluded from both heterosexual and gay communities. Thus, it would be inappropriate to lump all of these groups together, and conducting separate focus groups for each group was beyond the scope of this study. It is therefore recommended that future studies build on our work to encompass other sexual and gender identity groups. Some of the specific suggestions from this study may not necessarily be applicable to these other groups, but the general themes around improving inclusiveness and relevance are still likely to apply.

Conclusions

At present, e-therapies for depression and anxiety fail to cater to the mental health needs of lesbians and gay men, and many actively exclude lesbians and gay men by delivering content that assumes or suggests users are heterosexual. Through a series of focus groups, this study aimed to identify key areas in which e-therapies might be improved. Focus groups raised a number of issues that broadly related to improving the inclusiveness and relevance of e-therapy, such as avoiding language and content that excludes same-sex attracted users, addressing stigma-related challenges by expanding content on relationships, coming out, managing discrimination, and coping with lack of support from family and friends, as well as providing additional resources that specifically cater to same-sex attracted people. Our findings give further weight to broader advocacy and policy work aimed at improving mental health resources and health care for lesbians and gay men. Our findings also build on previous research, such as a recent protocol offered for improving e-therapies for lesbians and gay men [29] and the qualitative work by Lucassen et al [11]. Taken together, findings from both this study and previous work provide e-therapy developers with guidance to improve the inclusiveness and relevance of e-therapy for lesbians and gay men, and to help strengthen the role of e-therapy in reducing high rates of depression and anxiety among these populations.

Acknowledgments

We wish to thank beyondblue for supporting this study with a grant awarded as part of its “National Priority Driven Research Program in Gay, Lesbian, Bisexual, Transgender, and Intersex (GLBTI) People”. We also thank Ruth McNair, Amaryll Perlesz, Kieran O’Loughlin, David Belasic, and Simon Rice for the advice they offered on this project.

Conflicts of Interest

None declared.

References

1. Fuller MSD, Mathews R. APS Professional Practice. Australia: Australian Psychological Society; 2013. Internet supported psychological interventions: A guide to navigating the online world of psychological programs URL: http://www.acpor.ro/files/file/resurse_articole/APS_2012_internet_supported_interventions.pdf [accessed 2015-03-02] [WebCite Cache ID 6WicjEkCo]
2. Cowpertwait L, Clarke D. Effectiveness of web-based psychological interventions for depression: A meta-analysis. Int J Ment Health Addiction 2013 Jan 18;11(2):247-268. [doi: 10.1007/s11469-012-9416-z]
3. Sucala M, Schnur JB, Constantino MJ, Miller SJ, Brackman EH, Montgomery GH. The therapeutic relationship in e-therapy for mental health: a systematic review. J Med Internet Res 2012 Aug;14(4):e110 [FREE Full text] [doi: 10.2196/jmir.2084] [Medline: 22858538]
4. 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]
5. Bell CJ, Colhoun HC, Carter FA, Frampton CM. Effectiveness of computerised cognitive behaviour therapy for anxiety disorders in secondary care. Aust N Z J Psychiatry 2012 Jul;46(7):630-640. [doi: 10.1177/0004867412437345] [Medline: 22327097]
6. Department of Health & Ageing. E - mental health strategy for Australia. Australia; 2012. URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/7C7B0BFEB985D0EBCA257BF0001BB0A6/$File/emstrat.pdf [accessed 2015-03-01] [WebCite Cache ID 6WidZ5yrx]
7. National Institute for Health and Care Excellence. NICE technology appraisal guidance 97: Computerised cognitive behaviour therapy for depression and anxiety. 2006. URL: http://www.nice.org.uk/guidance/ta97 [accessed 2015-03-01] [WebCite Cache ID 6Widhxn5V]
8. Sparx e-therapy for young people launched 2014.: New Zealand Ministry of Health; 2014. URL: http://www.health.govt.nz/news-media/news-items/sparx-e-therapy-young-people-launched [accessed 2014-06-10] [WebCite Cache ID 6QDZ76ByT]
9. Spurgeon JA, Wright JH. Computer-assisted cognitive-behavioral therapy. Curr Psychiatry Rep 2010 Dec;12(6):547-552. [doi: 10.1007/s11920-010-0152-4] [Medline: 20872100]
10. Rochlen AB, Zack JS, Speyer C. Online therapy: review of relevant definitions, debates, and current empirical support. J Clin Psychol 2004 Mar;60(3):269-283. [doi: 10.1002/jclp.10263] [Medline: 14981791]
11. Lucassen MFG, Hatcher S, Stasiak K, Fleming T, Shepherd M, Merry SN. The views of lesbian, gay and bisexual youth regarding computerised self-help for depression: An exploratory study. Advances in Mental Health 2014 Dec 17;12(1):22-33. [doi: 10.5172/jamh.2013.12.1.22]

12. Abbott JM, Klein B, Ciechomski L. Best practices in online therapy. Journal of Technology in Human Services 2008 Jul 03;26(2-4):360-375. [doi: 10.1080/15228830802097257]
13. Leonard W, Pitts M, Mitchell A, Lyons A, Smith A, Patel S, et al. Private lives 2: The second national survey of the health and wellbeing of gay, lesbian, bisexual and transgender (GLBT) Australians. Australia: Australian Research Centre in Sex, Health and Society, La Trobe University; 2012. URL: http://www.glhv.org.au/files/PrivateLives2Report.pdf [accessed 2015-03-02] [WebCite Cache ID 6WieaExbx]
14. Ashworth A. Sexual orientation: A guide for the NHS. United Kingdom: Stonewall Institute; 2013. URL: http://www.stonewall.org.uk/documents/stonewall_guide_for_the_nhs_web.pdf [accessed 2015-03-01] [WebCite Cache ID 6Wieit4iO]
15. Department of Health: Equality & Human Rights Group. Sexual orientation: A practical guide for the NHS. United Kingdom: Department of Health; 2009. URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_095634 [accessed 2015-03-01] [WebCite Cache ID 6Wiew5cZV]
16. Guasp A, Taylor J. Stonewall health briefing: Experiences of healthcare. United Kingdom: Stonewall Institute; 2012. URL: http://www.stonewall.org.uk/documents/experiences_of_healthcare.pdf [accessed 2015-03-02] [WebCite Cache ID 6WifdCACD]
17. Mental health disparities: Lesbian / gay / bisexual / transgender. United States: American Psychological Association; 2014. URL: http://www.apa.org/pi/ses/resources/publications/factsheet-lgbt.pdf [accessed 2015-03-01] [WebCite Cache ID 6Wifnthg7]
18. Guasp A, Taylor J. Stonewall health briefing: Mental health. United Kingdom: Stonewall Institute; 2012. URL: http://www.stonewall.org.uk/documents/mental_health.pdf [accessed 2015-03-01] [WebCite Cache ID 6WifuzCGf]
19. Balsam KF, Beadnell B, Molina Y. The Daily Heterosexist Experiences Questionnaire: Measuring minority stress among lesbian, gay, bisexual, and transgender adults. Meas Eval Couns Dev 2013 Jan;46(1):3-25 [FREE Full text] [doi: 10.1177/0748175612449743] [Medline: 24058262]
20. Grabski B, Iniewicz G, Mijas M. Mental health of homosexual and bisexual people: Review of research and presentation of phenomenon. Psychiatr Pol 2012;46(4):637-647. [Medline: 23214165]
21. Meyer IH. Minority stress and mental health in gay men. Journal of Health and Social Behavior 1995 Mar;36(1):38. [doi: 10.2307/2137286]
22. Angermeyer M, Matschinger H. The stigma of mental illness: effects of labelling on public attitudes towards people with mental disorder. Acta Psychiatr Scand 2003 Oct;108(4):304-309. [doi: 10.1034/j.1600-0447.2003.00150.x]
23. National survey of mental health and wellbeing: Summary of results. Australia: Australian Bureau of Statistics; 2007. URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4326.02007?OpenDocument [accessed 2015-03-01] [WebCite Cache ID 6WigHbVns]
24. Green KE, Feinstein BA. Substance use in lesbian, gay, and bisexual populations: an update on empirical research and implications for treatment. Psychol Addict Behav 2012 Jun;26(2):265-278 [FREE Full text] [doi: 10.1037/a0025424] [Medline: 22061339]
25. Haas AP, Eliason M, Mays VM, Mathy RM, Cochran SD, D'Augelli AR, et al. Suicide and suicide risk in lesbian, gay, bisexual, and transgender populations: review and recommendations. J Homosex 2011;58(1):10-51 [FREE Full text] [doi: 10.1080/00918369.2011.534038] [Medline: 21213174]
26. Beyondblue. 2012. Research grant programs 2002-2012 URL: http://www.beyondblue.org.au/resources/research/research-grant-programs-2002-12 [accessed 2014-09-29] [WebCite Cache ID 6SwCvKJxU]
27. Improving the health and well-being of lesbian, gay, bisexual and transgender persons: A report by the secretariat. Switzerland: World Health Organization; 2013. URL: http://apps.who.int/gb/ebwha/pdf_files/EB133/B133_1_annotated-en.pdf [accessed 2015-03-01] [WebCite Cache ID 6WihO07hv]
28. Leonard W, Metcalf A. Going upstream: A framework for promoting the mental health of lesbian, gay, bisexual, transgender and intersex (LGBTI) people. Australia: National LGBTI Health Alliance; 2014. URL: http://www.beyondblue.org.au/docs/default-source/default-document-library/bw0257-going-upstream-online-o-lgbti-mental-health-promotion-framework.pdf?sfvrsn=2 [accessed 2015-03-01] [WebCite Cache ID 6WihXdM0R]
29. Rozbroj T, Lyons A, Pitts M, Mitchell A, Christensen H. Assessing the applicability of e-therapies for depression, anxiety, and other mood disorders among lesbians and gay men: analysis of 24 web- and mobile phone-based self-help interventions. J Med Internet Res 2014 Jul;16(7):e166 [FREE Full text] [doi: 10.2196/jmir.3529] [Medline: 24996000]
30. Friedman MR, Stall R, Silvestre AJ, Mustanski B, Shoptaw S, Surkan PJ, et al. Stuck in the middle: longitudinal HIV-related health disparities among men who have sex with men and women. J Acquir Immune Defic Syndr 2014 Jun 1;66(2):213-220. [doi: 10.1097/QAI.0000000000000143] [Medline: 24662298]
31. Kuyper L, Fokkema T. Minority stress and mental health among Dutch LGBs: examination of differences between sex and sexual orientation. J Couns Psychol 2011 Apr;58(2):222-233. [doi: 10.1037/a0022688] [Medline: 21401219]
32. Lyons A, Croy S, Barrett C, Whyte C. Growing old as a gay man: how life has changed for the gay liberation generation. Ageing and Society 2014 Aug 27;FirstView:1-22. [doi: 10.1017/S0144686X14000889]
33. Lyons A, Pitts M, Grierson J. Growing old as a gay man: psychosocial well-being of a sexual minority. Research on Aging 2012 May 03;35(3):275-295. [doi: 10.1177/0164027512445055]

34. Fredriksen-Goldsen KI, Kim HJ, Barkan SE, Muraco A, Hoy-Ellis CP. Health disparities among lesbian, gay, and bisexual older adults: results from a population-based study. Am J Public Health 2013 Oct;103(10):1802-1809 [FREE Full text] [doi: 10.2105/AJPH.2012.301110] [Medline: 23763391]
35. National Institute for Mental Health Research. 2014. MoodGYM URL: https://moodgym.anu.edu.au/welcome [accessed 2014-09-24] [WebCite Cache ID 6Sok1fJ8a]
36. Christensen H, Griffiths KM, Groves C. MoodGYM training program: clinicians manual. Australia: Centre for Mental Health Research, ANU; 2004.
37. National Institute for Mental Health Research. 2014. Beacon URL: https://beacon.anu.edu.au/ [accessed 2015-03-01] [WebCite Cache ID 6Wii4Ebci]
38. The National Registry of Evidence-based Programs and Practices.: Substance Abuse and Mental Health Services Administration; 2012. MoodGYM URL: http://www.nrepp.samhsa.gov/ViewIntervention.aspx?id=319 [accessed 2014-09-29] [WebCite Cache ID 6SwEx67Zo]
39. Parslow RA, Christensen H, Griffiths KM, Groves C. The Warpy Thoughts Scale: a new 20-item instrument to measure dysfunctional attitudes. Cogn Behav Ther 2006 Jun;35(2):106-116. [doi: 10.1080/16506070500372279] [Medline: 16754265]
40. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
41. Griffiths KM, Christensen H, Jorm AF. Predictors of depression stigma. BMC Psychiatry 2008;8:25 [FREE Full text] [doi: 10.1186/1471-244X-8-25] [Medline: 18423003]
42. Corrigan PW, Watson AC. Understanding the impact of stigma on people with mental illness. World Psychiatry 2002 Feb;1(1):16-20 [FREE Full text] [Medline: 16946807]
43. Bradstreet B, Calver S, Kent R, Klein T, Mitchell MB, Parker D, et al. Working therapeutically with LGBTI clients: A practice wisdom resource. Australia: National LGBTI Health Alliance; 2014. URL: http://www.beyondblue.org.au/docs/default-source/default-document-library/bw0256-practice-wisdom-guide-online.pdf?sfvrsn=2 [accessed 2015-03-02] [WebCite Cache ID 6WiieGjO5]
44. Lyons A, Hosking W. Health disparities among common subcultural identities of young gay men: physical, mental, and sexual health. Arch Sex Behav 2014 Nov;43(8):1621-1635. [doi: 10.1007/s10508-014-0315-1] [Medline: 25060121]
45. Meisner BA, Hynie M. Ageism with heterosexism: Self-perceptions, identity, and psychological health in older gay and lesbian adults. Gay and Lesbian Issues and Psychology Review 2009;5(1):51-58.
46. Lyons A, Hosking W, Rozbroj T. Rural-urban differences in mental health, resilience, stigma, and social support among young Australian gay men. J Rural Health 2015 Jan;31(1):89-97. [doi: 10.1111/jrh.12089] [Medline: 25183054]
47. Meyer IH, Northridge ME. Prejudice and Discrimination as Social Stressors. In: The health of sexual minorities: public health perspectives on lesbian, gay, bisexual, and transgender populations. New York: Springer; 2007.
48. Murphy BC. Educating mental health professionals about gay and lesbian issues. J Homosex 1992;22(3-4):229-246. [doi: 10.1300/J082v22n03_10] [Medline: 1573259]
49. Morrison S, Dinkel S. Heterosexism and health care: a concept analysis. Nurs Forum 2012;47(2):123-130. [doi: 10.1111/j.1744-6198.2011.00243.x] [Medline: 22512770]
50. Saulnier CF. Deciding who to see: lesbians discuss their preferences in health and mental health care providers. Social Work 2002 Oct 01;47(4):355-365. [doi: 10.1093/sw/47.4.355]

Abbreviations

CBT: cognitive behavioral therapy
CCBT: computerized cognitive behavioral therapy
OF: older female focus group
OM: older male focus group
YF: younger female focus group
YM: younger male focus group

Edited by G Eysenbach; submitted 10.11.14; peer-reviewed by M Lucassen, E McCann; comments to author 01.12.14; revised version received 17.12.14; accepted 10.01.15; published 11.03.15.

Please cite as:
Rozbroj T, Lyons A, Pitts M, Mitchell A, Christensen H
Improving Self-Help E-Therapy for Depression and Anxiety Among Sexual Minorities: An Analysis of Focus Groups With Lesbians and Gay Men
J Med Internet Res 2015;17(3):e66
URL: http://www.jmir.org/2015/3/e66/
doi:10.2196/jmir.4013
PMID:25761775

Original Paper

Web-Based Psychotherapy for Posttraumatic Stress Disorder in War-Traumatized Arab Patients: Randomized Controlled Trial

Christine Knaevelsrud[1,2], PhD; Janine Brand[1,2], MA; Alfred Lange[3], PhD; Jeroen Ruwaard[4], PhD; Birgit Wagner[5], PhD

[1]Department of Clinical Psychology, Freie University, Berlin, Germany
[2]Treatment Center for Torture Victims, Berlin, Germany
[3]Department of Clinical Psychology, Universität Amsterdam, Amsterdam, Netherlands
[4]Department of Clinical Psychology, VU University, Amsterdam, Netherlands
[5]Department Clinical Psychology and Psychotherapy, Medical School Berlin, Berlin, Germany

Corresponding Author:
Christine Knaevelsrud, PhD
Department of Clinical Psychology, Freie University
Habelschwerdter Allee 45
Berlin, 14195
Germany
Phone: 49 30 838557
Fax: 49 30 83851233
Email: christine.knaevelsrud@fu-berlin.de

Abstract

Background: In recent years, armed conflicts in the Middle East have resulted in high rates of exposure to traumatic events. Despite the increasing demand of mental health care provision, ongoing violence limits conventional approaches of mental health care provision. Internet-based interventions for posttraumatic stress disorder (PTSD) have proved feasible and effective in Western countries, but their applicability and efficacy in war and conflict regions remains unknown.

Objective: This study investigated the efficacy of a cognitive behavioral Internet-based intervention for war-traumatized Arab patients, with focus on Iraq.

Methods: A total of 159 individuals with PTSD participated in a parallel group randomized trial. Participants were randomly allocated by a computer-generated sequence to a treatment group (n=79) or a waiting list control group (n=80). The treatment group received 2 weekly 45-minute cognitive behavioral interventions via Internet over a 5-week period (10 sessions in total). The primary outcome was recovery from posttraumatic stress symptoms.

Results: Posttraumatic stress symptoms were significantly reduced from baseline to posttreatment (intention-to-treat analysis) in the treatment group relative to the control group ($F_{1,157}$=44.29, P<.001, d=0.92). Treatment effects were sustained at 3-month follow-up. Completer analysis indicated that 29 of 47 patients (62%) in the treatment group had recovered from posttraumatic stress symptoms at posttreatment (reliable change and Posttraumatic Stress Diagnostic Scale score <20) versus 1 patient (2%) in the control group (OR 74.19, 95% CI 9.93-585.8, P<.001) indicating that the chance of recovering was 74.19 times higher in the treatment than in the control group.

Conclusions: The results indicate, even in unstable and insecure settings with ongoing exposure to human rights violations through war and dictatorships, people with posttraumatic stress symptoms benefit from a cognitive behavioral treatment provided entirely through the Internet. This method of delivery could improve patients' access to humanitarian aid in the form of e-mental health services.

Trial Registration: Australian New Zealand Clinical Trial Registry, ACTRN12611001019998; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=347505 (Archived by WebCite at http://www.webcitation.org/6Wto4HCdH).

(J Med Internet Res 2015;17(3):e71) doi:10.2196/jmir.3582

KEYWORDS

posttraumatic stress disorders; Middle East; war; violence; cognitive therapy; mental health service; Internet

Introduction

In recent decades, war and human rights violations in the Middle East have led to high rates of exposure to traumatic events and to a correspondingly high incidence of posttraumatic stress disorder (PTSD) in the region [1-3]. Under the Ba'ath regime, for example, Iraqis were subjected to widespread, severe human rights abuses, including torture, killings, disappearances, beatings, kidnappings, forced amputation, and rape. With the invasion of Iraq by the US-led coalition forces in 2002, exposure to traumatizing events increased dramatically, with suicide bombers killing significantly more civilians than coalition soldiers [4]. Hick and colleagues [5] reported 1003 suicide bomb events between 2003 and 2010, causing 11% of all Iraqi civilian deaths and 26% of civilian injuries. Between 2004 and 2007, Iraqis were at the highest risk worldwide of dying in a violent conflict [6]. The escalation of violence led to an unprecedented demand for medical and psychotherapeutic support. However, half of the nation's physicians are estimated to have fled Iraq [7]. Physicians, mental health professionals, and health care professionals in general have been frequent targets of kidnappings, bodily mutilation, and random shootings; many health professionals have been killed [8,9]. Given the security situation, most international aid organizations have left the country. At the same time, violence has spread throughout the neighboring countries. Dramatic political developments (ie, the Arab Spring) in the Middle East and North Africa have been associated with state violence and civil war has further increased the demand for additional physical and mental health support structures.

Providing medical or mental health care in regions of war and ongoing violent conflict often puts mental health professionals at great risk. Very few studies have reported on mental health care services provided for survivors of war in developing countries. Although their results have been encouraging, these approaches are available to only small numbers of people, are relatively costly, and require health professionals to be located on site [10,11].

Native-speaking health professionals who are geographically independent of their clients may, therefore, be able to provide vital psychiatric support in underserved conflict regions.

Internet-based delivery of psychotherapeutic interventions has become increasingly established in the Western world. In particular, interventions developed for patients with PTSD have been shown to produce significant reductions in PTSD symptoms and in associated psychopathology, such as depression and anxiety [12-17]. Internet-based approaches may provide a unique treatment alternative in conflict areas where there is an urgent need for psychological care that is easily accessible, independent of the location of the therapist, and relatively inexpensive.

The major aim of this study was to evaluate an online cognitive behavioral therapy (CBT) intervention for posttraumatic stress symptoms in a setting that remains highly unstable, namely Iraq. This study randomly assigned participants to either an Internet-based treatment program or a waiting list control condition. Those in the intervention group received immediate access to an Internet-based program for PTSD treatment, whereas those in the control condition had to wait several weeks before they get access to the same Internet-based program thereafter. We hypothesized that the Internet-based treatment would produce a significantly greater improvement on the outcome compared to the control condition. Additionally, participants in the treatment group were asked to take part in a 3-month follow-up measure, whereas the control group did not.

Methods

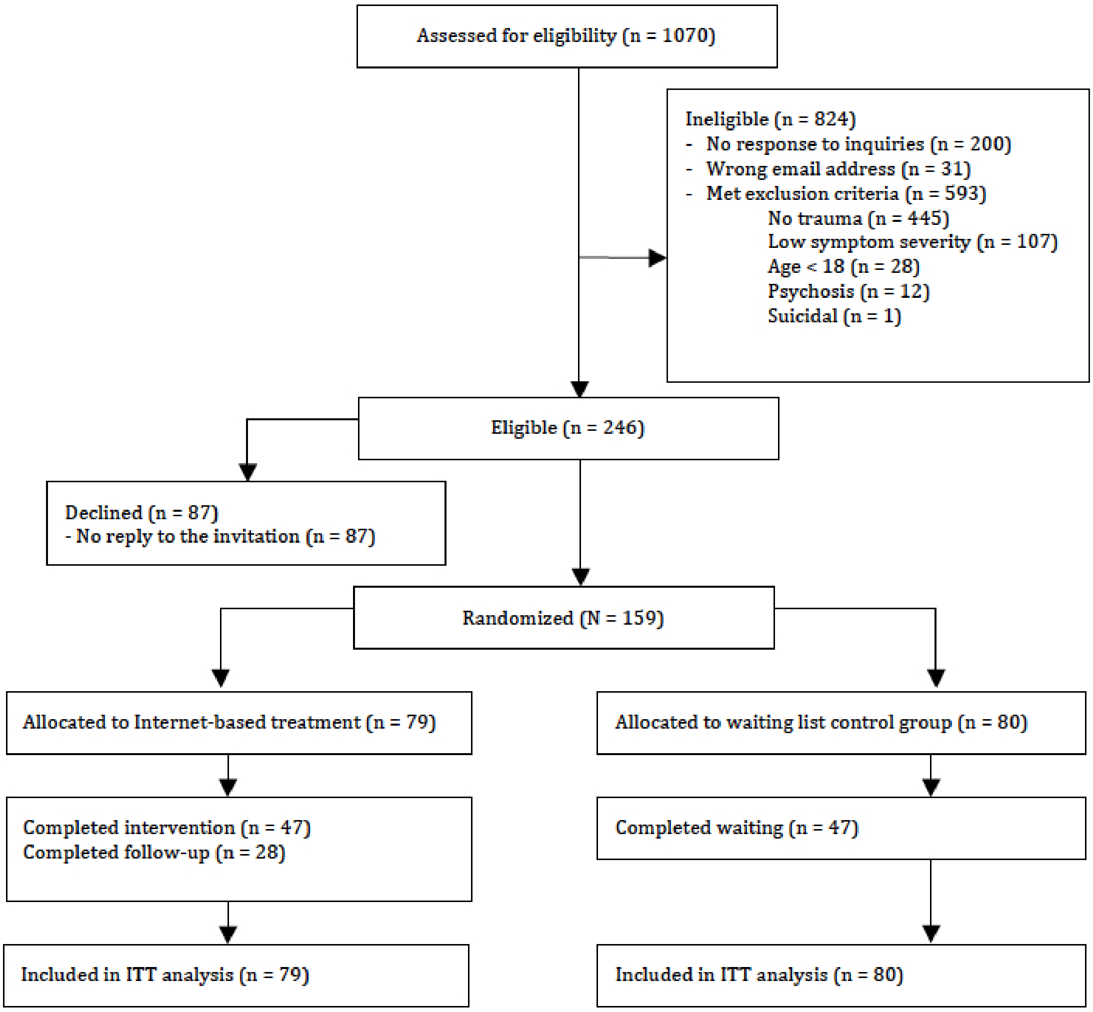
Participants

Participants were 159 Arabic-speaking adults (45 male, 114 female) with clinical levels of posttraumatic stress, aged from 18 to 56 years (mean 28.1, SD 7.43). In all, 48 of 159 patients (30.2%) were married and 41 of 159 participants (25.8%) completed secondary school; 99 of 159 (62%) had a university degree. Concerning the type of trauma, 63 of 159 patients (39.6%) had experienced sexual violence (war-related and sexual abuse) and 24 of 159 (15.1%) had experienced the killing of a family member or close person. Moreover, 30 of 159 participants (18.9%) reported being exposed to violence (eg, kidnapping, witnessing bomb attacks) and war or torture as their index trauma.

To be included in the study, participants had to have a history of trauma according to the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition; *DSM-IV*) criteria accompanied by posttraumatic stress symptoms, knowledge of Arabic, and age between 18 and 65 years. The Posttraumatic Stress Diagnostic Scale (PDS) was used to identify if patients reported the minimum number of symptoms required by *DSM-IV* for each of the symptom clusters (at least 1 intrusion, 3 avoidance, and 2 hyperarousal symptoms). Additionally the minimum score on the PDS to be included in the trial was 11 indicating moderate symptom severity. Applicants were excluded if they met 1 of the following criteria: currently receiving treatment elsewhere, substance abuse or dependence, high risk of suicide, psychotic symptoms, and low symptom severity. Symptom severity of depression was assessed by the Hopkins Symptom Checklist for depression [18] and risk of suicide by the Arabic translation of the Suicide Risk Assessment [19], a 6-item self-report questionnaire designed to capture suicidal tendencies. It consists of questions identifying suicidal plans, previous suicide attempts, and current suicidal intentions. Psychotic symptoms were assessed by the Arabic translation of the Dutch Screening Device for Psychotic Disorder [20]. Because no data are yet available from an Iraqi norm group, the Dutch norm data were used; however, due to intercultural differences, these norm data may have been too conservative.

Of the 1070 people who approached the study, 593 were excluded on the basis of our exclusion criteria (eg, outside age range, non–trauma-related difficulties, lack of Internet access). A total of 159 participants were randomly assigned to the treatment (n=79) or control condition (n=80). Participant flow is illustrated in Figure 1.

Figure 1. Flowchart showing progression of participants through the study.



Procedure

The study was carried out in Berlin (Treatment Center for Torture Victims, Freie University, Berlin). Recruitment for this randomized controlled trial took place from January 2009 to November 2011. Participants were recruited through radio, TV, and newspaper announcements, as well as health-related websites, specifically in Iraq. Information about the study was published regularly on a Facebook page.

The study website [21] provided general information about PTSD, online assessment, and the treatment program (Figure 2). Potential participants were informed about the study and received information about (1) posttraumatic stress reactions, (2) the study and its inclusion and exclusion criteria, (3) the Internet-based treatment, and (4) other treatment alternatives. A detailed description of the 3 treatment modules and the text-based form of the intervention was also given to the participants along with the patient information. Because the pilot study revealed that some patients had doubts about the neutrality of the website and treatment offered [22], participants were explicitly informed that all patient data and texts would be protected by rigorous security measures.

Potential patients logged in and completed the screening questionnaires online (1070 screenings completed). Initial screening was conducted with a fully automated computerized assessment battery including all outcome measures in the trial. These outcomes later served as the pretreatment scores for the included participants. Additional questions regarding exclusion criteria (suicidality, psychotic symptoms), demographics (age, gender, and education), current treatment, and treatment history were included in the online assessment. Whenever any data regarding the exclusion criteria were found to be unclear, participants were contacted by phone and asked to provide additional information about their psychotic symptoms and suicidal thoughts or behaviors (20.3%, 217 of participants were contacted by telephone to gather this information). The excluded individuals received an explanation as to why they had not been included and, if necessary, advice on how to seek help. Participants who met all inclusion criteria following diagnostic assessment and who provided informed consent were randomly assigned to either the Internet-based treatment or a waiting list control condition. Randomization was based on a computer-generated randomization list. Treatment started right after providing informed consent without any latency for those participants who were assigned to the Internet-based therapy.

All data reported in the trial were collected online and participants were given standardized reminders to complete the assessments using the online assessment system. They completed the outcome measures at pretreatment (initial screening), posttreatment (right after the treatment), and 3-month follow-up. For ethical reasons, participants assigned to the control condition received treatment directly after completing the waiting period. Thus, there are no follow-up results available for the control condition. The Ethics Committee of the University of Leipzig approved the study. Researchers and psychotherapists were not masked to the intervention. See Figure 1 for the progression of study participants.

Figure 2. Screenshot Startpage.

معالجة إضطراب ما بعد الصدمة النفسية عن طريق الإنترنت

طريقة العلاج

بدأ مركز علاج الصدمات النفسية العمل مع الأفراد الذين يعانون من إضطراب ما بعد الصدمة النفسية في ألمانيا وهولندا، ولقد ثبت علمياً نجاح هذا العلاج في التعامل مع مشكلات هؤلاء الأفراد عن طريق الإنترنت.

يستخدم المعالجون المختصون العاملون في هذا الموقع برنامجاً مجرباً علمياً. يتكون هذا الأخير من عدة أجزاء ثابتة ، التي يمكن لمختلف الأفراد وفي ظروف مختلفة الإستفادة منها وتوظيفها في حل مشكلاتهم الشخصية، بحيث يمكن للمريض أن يكتب عن تجربته الشخصية من مكان تواجده دون الحاجة إلى زيارة المعالج ودون التعرض للحرج عند رواية الأحداث الشخصية.

الإشتراك المجاني

لأول مرة، يمكن حالياً للشخص المصاب الإنضمام إلى العلاج النفسي عبر الإنترنت باللغة العربية مجاناً. يستغرق العلاج مدة تتراوح بين أربعة وخمسة أسابيع ويمكن لعدد محدود من الأشخاص الإشتراك به مجاناً وبدون مقابل لعرضه في إطار بحث علمي.

يتم تشخيص حالة المتسجلين الراغبين بالعلاج من خلال فحص أسئلة التشخيص النفسي الواردة في إستمارة البيانات. ومقابل إستخدام هذه المعلومات لغرض البحث العلمي (المعلومات المأخوذة ستكون سرية وبدون ذكر أسماء المشاركين وستوضع تحت أرقام مشفرة) سيكون العلاج مجانياً.

يتم تنفيذ هذا البحث في مركز العلاج النفسي في برلين تحت إشراف الدكتورة كرستينة كنيفلسرود الأستاذة في الجامعة

Outcomes

Posttraumatic stress symptoms were assessed by the PDS [23], which asks participants to rate on 17 items the occurrence of symptoms of PTSD (intrusions, avoidance, and hyperarousal) in the past 30 days (from 0=never to 3=nearly always). A diagnosis of PTSD is made only when a patient reports the minimum number of symptoms required by *DSM-IV* for each of the symptom clusters. The PDS yields a sum score measuring the overall severity of posttraumatic stress symptoms. Scores of 20 or higher on the PDS indicate moderate to severe symptom severity [23]. Previous research has confirmed the reliability of the Arabic version of the PDS [24]. The Hopkins Symptom Checklist-25 (HSCL-25) [18] was used to assess depression and anxiety. The HSCL-25 comprises 10 items forming an anxiety subscale and 15 items forming a depression subscale with 2 questions relating to somatic symptoms. Both subscales were used in this trial. Scores range from 1 (little) to 4 (very much). A mean score was calculated by adding all scores and dividing by the number of items. Scores above the cutoff of 1.75 indicate clinically significant distress. The HSCL-25 has proved reliable and valid for measuring depression and anxiety in cross-cultural studies [18].

The somatization subscale of the Symptom Checklist-90-Revised (SCL-90-R) [25] was used to assess somatization. This subscale (SCL) of commonly experienced physical symptoms comprises 12 items each rated on a 5-point Likert scale from 0 (not at all) to 4 (extremely). An overall score is formed by calculating the mean score of all items with higher scores indicating severe somatization symptoms during the last week.

Quality of life was assessed with the EUROHIS-QOL [26], an 8-item scale which measures 4 domains of life quality (psychological, physical, social, and environmental), each represented by 2 items. Responses are given on a 5-point Likert scale ranging from 1 (not at all) to 5 (completely). An overall EUROHIS score is formed by calculating the mean score of all 8 items, with higher scores indicating better quality of life. Schmidt et al [26] have reported the measure to show good reliability and validity across a range of countries. The mean time to complete all questionnaires was 45 minutes with a range from 15 to 60 minutes.

Treatment

Internet-Based Treatment

A Dutch Internet-based CBT manual (*Interapy* [15]) was translated into Arabic and culturally adapted for this treatment program. Translations were conducted by different native-speaking psychotherapists following the guidelines for cross-cultural translations [27]. All texts were translated to Modern Standard Arabic because this is the standard for written language in Arab countries and readable for all participants independent from their dialect. The treatment protocol builds on evidence-based principles of CBT for PTSD [28]. Specifically, confrontation with the traumatic event has proven to be an important element of effective psychotherapy for PTSD and has been found to significantly reduce avoidance behavior. The treatment consisted of 2 weekly structured writing activities assigned each week over a period of 5 weeks. There were 3 treatment phases: (1) self-confrontation with the traumatic event, (2) cognitive restructuring, and (3) social sharing.

The basic structure of 10 writing assignments proved to be acceptable in the pilot study [22]. However, based on the evaluation of the pilot study, a number of substantial changes concerning the content of the modules had to be implemented. Patients' expectations of health care professionals are culturally shaped. Compared to Interapy, this approach uses a more pronounced directive therapeutic stance. In Muslim countries, the health care professional is an authoritative and highly respected figure who gives expert advice. Therefore, straight instructions and responsibility for therapeutic choices are expected. Refusal to give explicit advice and lack of assertion are associated with incompetence and indecisiveness of the therapist and are met with irritation and may even prompt discontinuation of the therapy by a patient.

In the first phase of self-confrontation, the participants were asked to write 4 essays describing the traumatic event and its circumstances in as much detail as possible, in the first person and in the present tense. In contrast to Western trials, participants were explicitly asked not to mentioned specific places or names of persons who were involved due to basic precautionary measures. In the second phase of cognitive restructuring, they had another 4 writing assignments taking the form of a supportive letter to a hypothetical friend who had experienced the same traumatic event. The aim of this phase was to provide new perspectives on the traumatic event. In this module, cultural norms came explicitly into play. Knowledge of the Koran proved extremely helpful. Therapists frequently used quotes and helpful metaphors from the Koran that could inspire patients to take a different perspective and challenge their dysfunctional thoughts.

Generally, the therapists expressed explicit respect toward and appreciation of the concept of the family. However, female participants who had experienced sexual violence were explicitly discouraged from disclosing their traumatic experiences to other family members due to potential serious social consequences of known dishonor (ie, due to experienced sexual violence). The third and final phase of social sharing focused on a symbolic farewell letter (2 assignments) that participants were normally instructed to address to themselves, to a person connected with the traumatic event, or to a friend. In this study, these choices were limited to a letter directed to themselves because of the previously described potentially aversive consequences. All communication with participants was asynchronous. Whenever participants did not conduct their writing assignment, they received a short reminder via email. If no response was received after 2 email reminders, the participant was contacted by phone to encourage them to complete the treatment.

Therapists

The participating therapists were 8 native Arabic-speaking psychotherapists or psychiatrists living in Iraq, Palestine, Syria, the Emirates, or Europe. Therapists were trained in 7-day workshops in Europe that focused on the handling of the treatment manual, special features of Internet-based therapy, and how to solve common problems in an online communication setting. After participating in the workshop, the therapists completed an introduction phase with being monitored continuously by a supervisor who read all texts and observed the treatment process. Only after completing this phase successfully did the therapists start to work independently (participating in weekly supervision sessions, either face-to-face or via Skype). The therapist provided individually written feedback and instructions on the next writing assignment within 1 working day. The therapist time involved in responding to texts ranged from 20 to 50 minutes per text, depending on the therapist's experience with Internet-based therapies.

Control Condition

Participants assigned to the control condition were asked to complete a waiting period of 6 weeks. Afterwards, they received the same Internet-based intervention as the treatment group. Because they received treatment straight after completing the waiting period, no relevant follow-up results are available for the control group.

Statistical Analysis

Statistical analyses were performed with the SPSS version 19.0 for Mac (IBM Corp, Armonk, NY, USA). Data distributions were approximately normal and did not require transformation. As a primary analysis, we performed mixed design ANOVAs with time as the within-subject and condition as the between-subject factors. These analyses were based on an intention-to-treat (ITT) design, including all dropouts to estimate the efficacy of the treatment compared to a waiting list control group. Whenever posttreatment and follow-up scores were not available for a participant, the last observation data were carried forward. Because our repeated-measures variable had only 2 levels, the assumption of sphericity was met and it was not necessary to apply a correction factor to the degrees of freedom [29]. According to Everitt and Howell [29], it is not meaningful to interpret main effects if there are strong interaction effects. Therefore, we abstained from reporting main effects of the ANOVA. In addition to the ITT analysis, we also performed a completer analysis, as proposed by Myers [30]. Chi-square tests and *t* tests were used to determine how similar people who dropped out and people who completed the treatment were, and to assess any differences between the treatment and control groups at baseline. To assess the magnitude of change in mean symptoms between baseline and posttest and between baseline and 3-month follow-up, we calculated effect sizes using Cohen's d for repeated measures. An effect size of d=0.80 for a psychological treatment is typically considered large [31]. Moreover, 2 indicators were used to examine whether there was not only a statistical change, but also a clinically significant effect: the reliable change index (RCI) and the clinically significant change following Jacobson and Truax [32]. The RCI is used to determine whether the change observed goes beyond expected measurement fluctuations. The RCI considers measurement error and its effects on variability of scores and is computed by subtracting the subject's posttest score from his or her pretest score and dividing this value by the standard error of difference between the 2 test scores. Clinically significant change following Jacobson and Truax [32] was determined as scoring below the clinical cutoff (<20 for PDS and <1.75 for the HSCL depression and anxiety subscale).

Role of the Funding Source

The sponsors of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to the study data and had responsibility for the decision to submit for publication.

Results

Baseline Data

Baseline data of the intrusion (mean 8.32, SD 3.90), avoidance (mean 12.65, SD 3.79), and hyperarousal subscales (mean 9.53, SD 3.32) indicated that all patients suffered from posttraumatic stress (mean 30.50, SD 8.10) and 140 of 159 participants (88.0%) scored above a PDS score of 20.0 indicating moderate to severe symptom levels. Participants also reported high levels of anxiety (mean 2.90, SD 0.60) and depression (mean 3.07, SD 0.54). Table 1 summarizes demographic data and sample characteristics of each group. With the exception of disclosure (χ^2_1=4.7, P=.02), there were no systematic group differences in any of the sociodemographic variables or in baseline data at pretreatment. Therefore, we did not include any covariates in the analysis.

To investigate differences in outcome for the people who completed the treatment within the planned time frame and those who took a lot longer, we divided the sample into 2 groups (shorter vs longer treatment duration) using a median-split approach. Results indicate that the 2 groups did not differ in their baseline posttraumatic stress (t_{45}=–1.30, P=.20), anxiety (t_{45}=–1.21, P=.23), and depression level (t_{45}=–1.54, P=.13). Similarly, there were no systematic group differences at posttreatment with respect to these variables (posttraumatic stress symptoms: t_{45}=–0.78, P=.44; anxiety: t_{45}=–1.26, P=.22; and depression: t_{45}=–1.29, P=.20).

Dropout Analysis

In the treatment group, 47 of 79 participants (59%) completed the posttreatment assessment. The treatment group did not differ from the control group in terms of attrition rate (χ^2_1=0.0, P=.92). The reasons for treatment dropout were often difficult to discern because 20 of 32 participants (63%) did not respond to emails or telephone calls. However, we ascertained that 3 participants were lost due to difficulties with electricity and Internet access. Two patients had to terminate the treatment due to hospitalization and 2 were referred to local psychiatrists. Two participants preferred face-to-face therapy, whereas another 3 patients completed the treatment but not the posttreatment assessment. Completers and noncompleters did not differ with respect to any sociodemographic variables (eg, age: t_{77}=–0.20, P=.84; gender: χ^2_1=1.5, P=.22; educational level: χ^2_3=7.2, P=.07; marital status: χ^2_3=4.1, P=.25; or professional status: χ^2_2=3.4, P=.19). At baseline, there were no differences in their posttraumatic stress symptoms (t_{77}=0.68, P=.50), anxiety (t_{77}=0.28, P=.78), or depression levels (t_{77}=–0.27, P=.79). Additionally, completers and noncompleters did not differ with respect to the number of traumatic events experienced (t_{77}=0.64, P=.52) or the type of trauma (χ^2_3=0.8, P=.86).

Table 1. Demographic and descriptive characteristics of the treatment and waiting list control groups at baseline (N=159).

Demographic characteristics	Treatment	Control	t_{157}	χ^2 (df)	*P*
Traumatic events, mean (SD)	3.67 (3.01)	3.03 (2.61)	1.68		.10
Age (years), mean (SD)	29.11 (8.20)	27.15 (6.48)	1.45		.15
Age (years), range	18-56	18-43			
Gender, n (%)				1.0 (1)	.31
Female	76 (60)	69 (55)			
Educational level, n (%)				6.1 (3)	.11
Completed secondary school	22 (17)	35 (28)			
University	71 (56)	48 (38)			
Marital status, n (%)				2.7 (3)	.44
Single	68 (53)	56 (45)			
Partnership/married	26 (20)	37 (30)			
Professional status, n (%)				1.3 (2)	.51
Student	23 (18)	31 (25)			
Employed	44 (35)	33 (26)			
Unemployed	33 (26)	36 (29)			
Type of trauma, n (%)				0.3 (3)	.96
Killing of a family member	14 (11)	16 (13)			
Sexual violence related to war/sexual abuse	39 (31)	40 (32)			
Violence/war/torture	20 (16)	18 (14)			
Others (eg, kidnapping, witnessing bomb attacks)	27 (21)	26 (21)			
Time since trauma (primary traumatic event), n (%)				4.4 (5)	.50
Less than 6 months	13 (10)	10 (8)			
6 months to 3 years	22 (17)	18 (14)			
More than 3 years	65 (51)	70 (56)			
Treatment history, n (%)				0.0 (1)	.96
Yes	8 (6)	8 (6)			
Disclosure, n (%)				4.7 (1)	.02
Yes	78 (62)	63 (50)			

Primary Analysis: Intention-to-Treat

Table 2 presents means and standard deviations of baseline, posttreatment, and 3-month follow-up scores for all outcome measures in both experimental conditions. It also presents the group×time interaction of the ANOVAs, whether the group change from pre- to posttreatment is significant, and between-group effect sizes for all outcome measures. As shown, there were significant interaction effects for all outcome variables. This result confirms that the decrease in the primary outcomes intrusion, avoidance, and hyperarousal (PDS) as well as in the secondary outcomes anxiety, depression (HSCL-25), and somatization (SCL) was significantly higher in the treatment group than in the control group. Furthermore participants receiving the Internet-based treatment showed a significant increase in satisfaction with life relative to the control group ($F_{1,157}$=44.20, *P*<.001). Despite the high attrition rate of 41% (32/79 in the treatment group), the 3-month follow-up indicated that the results remained stable after treatment. Individuals in the control condition showed no improvement in trauma-related symptoms (PDS subscales: intrusion, avoidance, hyperarousal) or in the secondary outcomes anxiety, depression, somatization, or satisfaction with life. In the treatment group, the ITT analysis revealed moderate to large within-group effect sizes from baseline to posttreatment for the primary outcomes of intrusion (*d*=0.78), avoidance (*d*=0.81), and hyperarousal (*d*=0.86). Large effect sizes were found for the depression (*d*=0.92) and anxiety (*d*=0.84) subscales, and moderate effect sizes for the somatization (*d*=0.42) and quality of life (*d*=0.76) subscales. Furthermore, moderate to large effect sizes (*d*=0.40-0.84) were calculated from baseline to follow-up. In the control condition, all effect sizes were close to zero (*d*=–0.08 to 0.13).

To test for interdependence among dependent variables, we performed a MANOVA for repeated measures (intrusion, avoidance, hyperarousal) with time (pre- and posttreatment) as the within-subject factor and condition as the between-subject factor revealing a significant interaction effect ($F_{1,157}$=44.29, *P*<.001). A MANOVA for repeated measures (anxiety and depression) also yielded a significant interaction effect ($F_{1,157}$=38.61, *P*<.001).

Secondary Analysis: Completer

Table 3 presents means and standard deviations of baseline, posttreatment, and follow-up scores for all outcome measures for the subgroup of participants who completed the treatment. It also presents the group×time interaction indicating whether the group change from pre- to posttreatment was significant as well as between-group effect sizes for all outcome measures. As shown, there were significant interaction effects for all outcome variables. As hypothesized, the effect sizes of these factors were larger than in the ITT analysis. With the exception of the secondary outcomes life satisfaction and depression, the completers showed ongoing arithmetical improvement in all outcomes up to the 3-month follow-up. However, these changes did not reach statistical significance. The results of the completer analysis provide further confirmation that both trauma-related symptoms and secondary outcomes, such as anxiety, depression, and somatization, decreased in all patients receiving the treatment. In contrast, there was no significant change from baseline to posttreatment in the control condition. Very large within-group effect sizes were found for posttraumatic stress symptoms (intrusion, avoidance, hyperarousal) from pre- to posttreatment (*d*=1.42-1.53) and 3-month follow-up (*d*=1.95-2.07) in the treatment condition. Large effect sizes were also found for the depression (*d*=1.69) and anxiety (*d*=1.44) subscales, and for quality of life (*d*=1.30) and somatization (*d*=0.68) at posttreatment and follow-up (*d*=1.69, *d*=1.67, *d*=1.09, and *d*=0.83, respectively). Only small effect sizes were found in the control condition (*d*=–0.15 to 0.25).

Table 2. Results of mixed design ANOVAs for the treatment and waiting list control groups at baseline, posttreatment, and 3-month follow-up: intention-to-treat analysis.

Outcome measures	Groups×pre-post, mean (SD)		$F_{1,157}$	*P*	*d*
	Treatment (n=79)	Control (n=80)			
PDS					
Intrusion			30.74	<.001	0.72
Baseline	8.32[a] (3.98)	8.33[a] (3.84)			
Posttreatment	5.09[b] (4.33)	8.06[a] (3.89)			
Follow-up	5.37[b] (4.59)	—			
Avoidance			34.26	<.001	0.92
Baseline	12.49[a] (3.77)	12.80[a] (3.82)			
Posttreatment	8.78[b] (5.38)	13.04[a] (3.78)			
Follow-up	8.97[b] (5.15)	—			
Hyperarousal			28.58	<.001	0.68
Baseline	9.54[a] (3.07)	9.52[a] (3.58)			
Posttreatment	6.42[b] (4.19)	9.07[a] (3.64)			
Follow-up	6.53[b] (4.09)	—			
Total score			44.29	<.001	0.92
Baseline	30.35[a] (8.16)	30.65[a] (8.10)			
Posttreatment	20.29[b] (12.45)	30.17[a] (8.70)			
Follow-up	20.87[b] (12.37)	—			
HSCL-25					
Anxiety			28.30	<.001	0.79
Baseline	2.88[a] (0.60)	2.92[a] (0.61)			
Posttreatment	2.29[b] (0.81)	2.85[a] (0.61)			
Follow-up	2.39[b] (0.76)	—			
Depression			40.66	<.001	1.03
Baseline	3.04[a] (0.58)	3.10[a] (0.50)			
Posttreatment	2.36[b] (0.90)	3.11[a] (0.50)			
Follow-up	2.54[b] (0.85)	—			
SCL: Somatization			11.68	<.001	0.56
Baseline	1.44[a] (0.84)	1.50[a] (0.90)			
Posttreatment	1.09[b] (0.81)	1.56[a] (0.88)			
Follow-up	1.11[b] (0.82)	—			
EUROHIS: Life satisfaction			44.20	<.001	0.84
Baseline	2.33[a] (0.73)	2.33[a] (0.79)			
Posttreatment	2.97[b] (0.95)	2.27[a] (0.71)			
Follow-up	2.72[b] (0.82)	—			

[a] Means within column grouping that share this superscript do not differ at *P*=.05.

[b] Means within column grouping that share this superscript do not differ at *P*=.05.

Table 3. Results of mixed design ANOVAs for the treatment and waiting list control groups at baseline, posttreatment, and 3-month follow-up: completer analysis.

Outcome measures	Groups×pre-post, mean (SD)		$F_{1,92}$	P	η^2_p
	Treatment (n=47)[a]	Control (n=47)			
PDS					
Intrusion			43.89	<.001	1.50
Baseline	8.28[b] (4.09)	8.21[b] (3.51)			
Posttreatment	2.85[c] (3.01)	7.77[b] (3.56)			
Follow-up	1.93[c] (2.42)	—			
Avoidance			45.50	<.001	1.59
Baseline	12.70[b] (3.71)	13.13[b] (3.85)			
Posttreatment	6.47[c] (5.04)	13.53[b] (3.75)			
Follow-up	5.29[c] (3.77)	—			
Hyperarousal			42.41	<.001	1.42
Baseline	9.89[b] (2.98)	10.47[b] (2.94)			
Posttreatment	4.64[c] (3.87)	9.70[b] (3.26)			
Follow-up	3.68[c] (3.03)	—			
Total score			74.85	<.001	1.77
Baseline	30.87[b] (8.13)	31.81[b] (7.13)			
Posttreatment	13.96[c] (10.75)	31.00[b] (8.36)			
Follow-up	10.89[c] (7.91)	—			
HSCL-25					
Anxiety			40.71	<.001	1.56
Baseline	2.90[b] (0.64)	3.04[b] (0.51)			
Posttreatment	1.91[c] (0.74)	2.91[b] (0.54)			
Follow-up	1.85[c] (0.62)	—			
Depression			59.58	<.001	2.04
Baseline	3.03[b] (0.59)	3.16[b] (0.48)			
Posttreatment	1.88[c] (0.77)	3.18[b] (0.47)			
Follow-up	1.94[c] (0.70)	—			
SCL					
Somatization			12.40	<.001	0.81
Baseline	1.45[b] (0.91)	1.40[b] (0.84)			
Posttreatment	0.87[c] (0.79)	1.52[b] (0.81)			
Follow-up	0.75[c] (0.78)	—			
EUROHIS					
Life satisfaction			63.99	<.001	1.52
Baseline	2.33[b] (0.76)	2.30[b] (0.82)			
Posttreatment	3.39[c] (0.87)	2.20[b] (0.68)			
Follow-up	3.16[c] (0.76)	—			

[a] Treatment group reduced to n=28 at 3-month follow-up due to dropout.

[b] Means within column grouping that share this superscript do not differ at *P*=.05.

[c] Means within column grouping that share this superscript do not differ at *P*=.05.

Reliable Change and Clinical Significance

Table 4 presents the percentage of patients showing reliable change at posttreatment (completers' data). According to this criterion, the proportion of participants showing reliable change in the treatment group was significantly higher than in the control group on all outcome measures. Additionally, we assessed whether treatment led to a change in participants' diagnostic category. At baseline, 74 of 79 participants (94%) in the treatment group and 75 of 80 patients (94%) in the control group scored at or above the clinical cutoff (20) for posttraumatic stress. At posttreatment the percentage was reduced to 34% (16/47) in the treatment group and to 89% (42/47) in the control group. At the 3-month follow-up, only 14% (4/28) of the treatment group participants scored above the clinical cutoff. With regard to the secondary outcome anxiety, 94% (74/79) in the treatment group and all patients in the control group scored above the cutoff at the beginning of the study (1.75). At posttreatment, this was reduced to 45% (21/47) in the treatment group, whereas 98% (46/47) of the control group still showed above-threshold symptoms. At 3-month follow-up, 13 of 28 participants (46%) in the treatment group showed anxiety symptoms above the cutoff score. Similar results were found on the depression subscale; 96% (76/79) of the treatment and 98% (78/80) of the control group participants scored above the cutoff (1.75) at baseline. All control group patients still suffered from depression symptoms at posttreatment, whereas the percentage was reduced to 40% (19/47) in the treatment group. At 3-month follow-up, this increased to 54% (15/28) of the treatment group participants. Furthermore, 29 of 47 patients (62%) recovered from posttraumatic stress symptoms (reliable change and PDS score<20) in the intervention group compared to 1 (2%) in the control group at posttreatment (OR 74.19, 95% CI 9.93-585.8, *P*<.001), indicating that the chance of recovering was 74.19 times higher in the treatment group than in the control group.

Table 4. Percentage of patients showing reliable change at posttreatment.

Outcome measures	Group, n (%)		χ^2_1	*P*
	Treatment (n=47)	Control (n=47)		
PDS				
Intrusion	21 (45)	1 (2)	23.7	<.001
Avoidance	26 (55)	4 (9)	23.7	<.001
Hyperarousal	27 (57)	7 (15)	18.4	<.001
Total score	35 (74)	3 (6)	45.2	<.001
HSCL-25				
Anxiety	31 (66)	5 (11)	30.4	<.001
Depression	35 (74)	4 (9)	42.1	<.001

Technology Approval und Feasibility

Using the Distress/Endorsement Validation Scale (DEVS [33]) we asked participants about their experience of the treatment. Despite the short-term nature of the intervention, 78% (37/47) of the participants considered the duration to be sufficient; 74% (35/47) spent up to 2 hours a week on the writing therapy. Furthermore, 87% (41/47) of participants regarded the therapy as clearly understandable and as an effective method for reducing tension and exhaustion, 74% (35/47) experienced a marked decrease in their symptoms, and 76% (36/47) would recommend the treatment to others.

Discussion

The aim of this study was to investigate whether it is possible to produce significant and sustained reduction of posttraumatic stress in participants living in an unstable conflict region using a brief Internet-delivered intervention. We observed significant reductions in posttraumatic stress symptom severity in all symptom clusters, and the effect sizes were of a similar magnitude to those reported for Western samples using the same treatment protocol [15-17,34-39]. In addition, the treatment had significant benefits with respect to symptoms of depression and anxiety and quality of life. Although many of the patients continued to experience difficulties in terms of exposure to life-threatening situations and severe human rights violations during the course of the treatment, they nevertheless benefited psychologically from the intervention.

The attrition rate of 41% (32 of 79 participants in the treatment group and 33 of 80 patients in the control group) was relatively high. A potential explanation for this may be that ongoing violence and economic insecurity keeps patients in a constant state of hyperalertness that calls for a shift of attention to primary day-to-day needs. We tried to determine reasons for dropout by emailing an additional questionnaire, but few people replied. A number of participants questioned the neutrality of the website and the treatment and were concerned that the intervention was supported by a foreign secret service (eg, CIA, Mossad). Others reported technical problems or a lack of privacy at home to write undisturbed by family members. The Dutch

Internet-based PTSD study conducted before 2003 [15] also reported a relatively high dropout rate of 43%. An analysis by Lange and colleagues [15] showed that about 59% of the dropouts had stopped the treatment because of technical problems. This is likely to be a key issue in conflict regions, where slow or instable Internet connections or power cuts are still a common problem, and may add to dropout. Despite the attrition rate, the ITT analysis still revealed large effect sizes for posttraumatic stress and depression. Given the high dropout rate, however, ITT analysis may be an overly conservative approach and may have underestimated the effectiveness of this novel approach as indicated by the very large effect sizes of treatment observed in the completer analysis.

A number of limitations demand further comment. First of all, assessment of psychopathology was exclusively based on self-rating questionnaires. A clinical interview would have facilitated more accurate information and clinical diagnosis, and this should be implemented in future trials.

The use of using a waiting list trial design poses substantial limitations on the validity and generalizability of the results. Because the waiting list control condition received treatment after the waiting period, effects in the follow-up intervals can only be estimated based on within-group effect sizes. Furthermore, an active control condition using an alternative evidence-based treatment protocol would have produced more valid data concerning the specific efficacy of this treatment approach. In addition, we found a gender bias because 74% (35 of 47 participants in the treatment group) of completers were female. These figures are comparable with Western treatment samples, but clearly not representative of the general population in this region. Women frequently experience rape or sexual abuse by armed groups in wars and civil conflicts. In Arab countries, women exposed to sexual abuse are often considered to be dishonored; therefore, many of them do not seek help for their psychological problems. Often, they do not risk confiding in others because this leaves them vulnerable to stigma and ostracism and could have life-threatening consequences. The anonymity of the Internet may encourage these women to seek therapeutic treatment. Finally, our sample was very well educated. This is in-line with other Internet-based samples [40,41]. It seems that at present, Internet-based interventions do not generally manage to engage less well-educated people in an intervention [42], independent of the type of program and country of origin.

Future research should expand treatment delivery techniques, especially focusing on conflict regions. New developments in the context of communication technology report an increasing use of mobile apps (ie, mobile phones) in the treatment of mental health problems. As mobile phone usage in the Middle East is growing, apps could be an additional way to extend treatment accessibility.

Finally, PTSD is only 1 outcome of chronic exposure to traumatic events in postconflict countries. In their extensive meta-analysis of 181 studies on psychological consequences of war and deportation, Steele and colleagues [43] found that rates of depression (30.8%) match those for PTSD (30.6%). A randomized controlled trial testing an Internet-based CBT treatment for depression in Arabic postconflict countries is currently underway.

To our knowledge, this is the first randomized controlled trial of any form of psychological intervention for people in the Middle East suffering from PTSD. The written narratives produced by our participants impressively document the toll of decades of dictatorship and war. Seventy psychiatrists are currently registered in Iraq, which has a population of 30 million. It will take decades before local services are able to provide sufficient health care for the country’s citizens. Our findings show, even when living in difficult conditions, people with posttraumatic stress symptoms benefit from a CBT provided entirely through the Internet. Findings on the program’s uptake and applicability confirm that new technologies can be used to provide humanitarian aid in the form of e-mental health services, even in areas that remain highly unstable. The opportunity to disseminate evidence-based, accessible interventions in regions where human rights violations are common represents an important contribution to humanitarian aid and has substantial public health implications.

Acknowledgments

The study was sponsored by the German Foreign Ministry; Misereor eV, Germany.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [44].

[PDF File (Adobe PDF File), 82KB - jmir_v17i3e71_app1.pdf]

References

1. el Sarraj E, Punamäki RL, Salmi S, Summerfield D. Experiences of torture and ill-treatment and posttraumatic stress disorder symptoms among Palestinian political prisoners. J Trauma Stress 1996 Jul;9(3):595-606. [Medline: 8827659]

2. Laban CJ, Gernaat HB, Komproe IH, Schreuders BA, De Jong Joop TVM. Impact of a long asylum procedure on the prevalence of psychiatric disorders in Iraqi asylum seekers in The Netherlands. J Nerv Ment Dis 2004 Dec;192(12):843-851. [Medline: 15583506]
3. Al-Jawadi AA, Abdul-Rhman S. Prevalence of childhood and early adolescence mental disorders among children attending primary health care centers in Mosul, Iraq: a cross-sectional study. BMC Public Health 2007;7:274 [FREE Full text] [doi: 10.1186/1471-2458-7-274] [Medline: 17910748]
4. Hicks MH, Dardagan H, Guerrero SG, Bagnall PM, Sloboda JA, Spagat M. Violent deaths of Iraqi civilians, 2003-2008: analysis by perpetrator, weapon, time, and location. PLoS Med 2011;8(2):e1000415 [FREE Full text] [doi: 10.1371/journal.pmed.1000415] [Medline: 21358813]
5. Hicks MH, Dardagan H, Bagnall PM, Spagat M, Sloboda JA. Casualties in civilians and coalition soldiers from suicide bombings in Iraq, 2003-10: a descriptive study. Lancet 2011 Sep 3;378(9794):906-914. [doi: 10.1016/S0140-6736(11)61023-4] [Medline: 21890055]
6. Geneva Declaration. Global Burden of Armed Violence. Geneva: Geneva Declaration Secretariat; 2008. URL: http://www.genevadeclaration.org/fileadmin/docs/Global-Burden-of-Armed-Violence-full-report.pdf [accessed 2015-03-04] [WebCite Cache ID 6WmMKBz7Z]
7. Webster PC. Iraq's health system yet to heal from ravages of war. Lancet 2011 Sep 3;378(9794):863-866. [Medline: 21898919]
8. Al Sheibani BI, Hadi NR, Hasoon T. Iraq lacks facilities and expertise in emergency medicine. BMJ 2006 Oct 21;333(7573):847 [FREE Full text] [doi: 10.1136/bmj.38986.476782.68] [Medline: 17053243]
9. Burnham GM, Lafta R, Doocy S. Doctors leaving 12 tertiary hospitals in Iraq, 2004-2007. Soc Sci Med 2009 Jul;69(2):172-177. [doi: 10.1016/j.socscimed.2009.05.021] [Medline: 19501443]
10. Schaal S, Elbert T, Neuner F. Narrative exposure therapy versus interpersonal psychotherapy. A pilot randomized controlled trial with Rwandan genocide orphans. Psychother Psychosom 2009;78(5):298-306. [doi: 10.1159/000229768] [Medline: 19628958]
11. Neuner F, Onyut PL, Ertl V, Odenwald M, Schauer E, Elbert T. Treatment of posttraumatic stress disorder by trained lay counselors in an African refugee settlement: a randomized controlled trial. J Consult Clin Psychol 2008 Aug;76(4):686-694. [doi: 10.1037/0022-006X.76.4.686] [Medline: 18665696]
12. Benight CC, Ruzek JI, Waldrep E. Internet interventions for traumatic stress: a review and theoretically based example. J Trauma Stress 2008 Dec;21(6):513-520. [doi: 10.1002/jts.20371] [Medline: 19107724]
13. Litz BT, Engel CC, Bryant RA, Papa A. A randomized, controlled proof-of-concept trial of an Internet-based, therapist-assisted self-management treatment for posttraumatic stress disorder. Am J Psychiatry 2007 Nov;164(11):1676-1683. [doi: 10.1176/appi.ajp.2007.06122057] [Medline: 17974932]
14. Klein B, Mitchell J, Abbott J, Shandley K, Austin D, Gilson K, et al. A therapist-assisted cognitive behavior therapy internet intervention for posttraumatic stress disorder: pre-, post- and 3-month follow-up results from an open trial. J Anxiety Disord 2010 Aug;24(6):635-644. [doi: 10.1016/j.janxdis.2010.04.005] [Medline: 20447802]
15. Lange A, Rietdijk D, Hudcovicova M, van de Ven JP, Schrieken B, Emmelkamp PMG. Interapy: a controlled randomized trial of the standardized treatment of posttraumatic stress through the internet. J Consult Clin Psychol 2003 Oct;71(5):901-909. [doi: 10.1037/0022-006X.71.5.901] [Medline: 14516238]
16. Knaevelsrud C, Wagner B, Karl A, Mueller J. New treatment approaches: integrating new media in the treatment of war and torture victims. Torture 2007;17(2):67-78 [FREE Full text] [Medline: 17728484]
17. Wagner B, Knaevelsrud C, Maercker A. Internet-based cognitive-behavioral therapy for complicated grief: a randomized controlled trial. Death Stud 2006 Jun;30(5):429-453. [doi: 10.1080/07481180600614385] [Medline: 16610157]
18. Mollica RF, Wyshak G, de Marneffe D, Khuon F, Lavelle J. Indochinese versions of the Hopkins Symptom Checklist-25: a screening instrument for the psychiatric care of refugees. Am J Psychiatry 1987 Apr;144(4):497-500. [Medline: 3565621]
19. Arnoldi J, van de Ven JP, Schrieken B, Lange A. De Risico Taxatielijst-eigenschappen van een kort screeningsintrument Suicide Risk Assessment-a short screening tool [unpublished manuscript University of Amsterdam]. Amsterdam: University of Amsterdam; 2000.
20. Lange A, Schrieken B, Blankers M, Van de Ven JP, Slot M. Constructie en validatie van de Gewaarwordingenlijst (GL) een hulpmiddel bij het signaleren van een verhoogde kans op psychosen. Directieve Therapie 2000;20:162-173.
21. Ilajnafsy. URL: http://ilajnafsy.bzfo.de/www/Ilajnafsy.org/ [accessed 2015-03-04] [WebCite Cache ID 6WmLK0PQG]
22. Wagner B, Schulz W, Knaevelsrud C. Efficacy of an Internet-based intervention for posttraumatic stress disorder in Iraq: a pilot study. Psychiatry Res 2012 Jan 30;195(1-2):85-88. [doi: 10.1016/j.psychres.2011.07.026] [Medline: 21813187]
23. Foa EB. The Posttraumatic Diagnostic Scale (PDS) Manual. Minneapolis: National Computer Systems; 1995.
24. Norris AE, Aroian KJ. Avoidance symptoms and assessment of posttraumatic stress disorder in Arab immigrant women. J Trauma Stress 2008 Oct;21(5):471-478. [doi: 10.1002/jts.20363] [Medline: 18956451]
25. Derogatis LR. SCL-90-R: Self-Report Symptom Inventory. In: Internationale Skalen für Psychiatrie: Collegium Internationale Psychiatriae Scalarum. Weinheim: Beltz; 1986.
26. Schmidt S, Mühlan H, Power M. The EUROHIS-QOL 8-item index: psychometric results of a cross-cultural field study. Eur J Public Health 2006 Aug;16(4):420-428 [FREE Full text] [doi: 10.1093/eurpub/cki155] [Medline: 16141303]

27. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993 Dec;46(12):1417-1432. [Medline: 8263569]
28. Bradley R, Greene J, Russ E, Dutra L, Westen D. A multidimensional meta-analysis of psychotherapy for PTSD. Am J Psychiatry 2005 Feb;162(2):214-227. [doi: 10.1176/appi.ajp.162.2.214] [Medline: 15677582]
29. Everitt B, Howell D. Bayesian methods for categorical data. Encyclopedia of Statistics in Behavioral Science 2005;1:139-146.
30. Myers WR. Handling missing data in clinical trials: an overview. Therapeutic Innovation & Regulatory Science 2000 Apr 01;34(2):525-533. [doi: 10.1177/009286150003400221]
31. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: L Erlbaum Associates; 1988.
32. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [Medline: 2002127]
33. Devilly GJ. An approach to psychotherapy toleration: the Distress/Endorsement Validation Scale (DEVS) for clinical outcome studies. J Behav Ther Exp Psychiatry 2004 Dec;35(4):319-336. [doi: 10.1016/j.jbtep.2004.08.001] [Medline: 15530846]
34. Lange A, van de Ven JP, Schrieken B, Emmelkamp PM. Interapy, treatment of posttraumatic stress through the Internet: a controlled trial. J Behav Ther Exp Psychiatry 2001 Jun;32(2):73-90. [Medline: 11764063]
35. Lange A, van de Ven JP, Schrieken B. Interapy: treatment of post-traumatic stress via the internet. Cogn Behav Ther 2003;32(3):110-124. [doi: 10.1080/16506070302317] [Medline: 16291543]
36. Knaevelsrud C, Maercker A. Does the quality of the working alliance predict treatment outcome in online psychotherapy for traumatized patients? J Med Internet Res 2006;8(4):e31 [FREE Full text] [doi: 10.2196/jmir.8.4.e31] [Medline: 17213049]
37. Knaevelsrud C, Maercker A. Internet-based treatment for PTSD reduces distress and facilitates the development of a strong therapeutic alliance: a randomized controlled clinical trial. BMC Psychiatry 2007;7:13 [FREE Full text] [doi: 10.1186/1471-244X-7-13] [Medline: 17442125]
38. Wagner B, Knaevelsrud C, Maercker A. Internet-based treatment for complicated grief: concepts and case study. Journal of Loss and Trauma 2005;10:409-432.
39. Wagner B, Maercker A. A 1.5-year follow-up of an Internet-based intervention for complicated grief. J Trauma Stress 2007 Aug;20(4):625-629. [doi: 10.1002/jts.20230] [Medline: 17721955]
40. Griffiths KM, Christensen H. Review of randomised controlled trials of Internet interventions for mental disorders and related conditions. Clinical Psychologist 2011;10:16-29.
41. Titov N, Andrews G, Kemp A, Robinson E. Characteristics of adults with anxiety or depression treated at an internet clinic: comparison with a national survey and an outpatient clinic. PLoS One 2010;5(5):e10885 [FREE Full text] [doi: 10.1371/journal.pone.0010885] [Medline: 20526371]
42. Lampe LA. Internet-based therapy: too good to be true? Aust N Z J Psychiatry 2011 Apr;45(4):342-343. [doi: 10.3109/00048674.2011.560138] [Medline: 21306182]
43. Steel Z, Chey T, Silove D, Marnane C, Bryant RA, van OM. Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement: a systematic review and meta-analysis. JAMA 2009 Aug 5;302(5):537-549. [doi: 10.1001/jama.2009.1132] [Medline: 19654388]
44. Eysenbach G, Consort- E. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

CBT: cognitive behavioral therapy
DEVS: Distress/Endorsement Validation Scale
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders (4th Edition)
HSCL-25: Hopkins Symptom Checklist-25
ITT: intention-to-treat
PDS: Posttraumatic Stress Diagnostic Scale
PTSD: posttraumatic stress disorder
RCI: reliable change index
SCL-90-R: Symptom Checklist-90-Revised

Edited by G Eysenbach; submitted 04.06.14; peer-reviewed by T Berger, H Gläsmer, N Morina, S Langrial; comments to author 17.07.14; revised version received 01.09.14; accepted 26.11.14; published 20.03.15.

Please cite as:
Knaevelsrud C, Brand J, Lange A, Ruwaard J, Wagner B
Web-Based Psychotherapy for Posttraumatic Stress Disorder in War-Traumatized Arab Patients: Randomized Controlled Trial
J Med Internet Res 2015;17(3):e71
URL: http://www.jmir.org/2015/3/e71/
doi:10.2196/jmir.3582
PMID:25799024

Original Paper

Efficacy of a Web-Based, Crowdsourced Peer-To-Peer Cognitive Reappraisal Platform for Depression: Randomized Controlled Trial

Robert R Morris[1], PhD; Stephen M Schueller[2], PhD; Rosalind W Picard[1], ScD

[1]MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA, United States

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

Corresponding Author:
Robert R Morris, PhD
MIT Media Lab
Massachusetts Institute of Technology
E14-348A
75 Amherst St
Cambridge, MA, 02139
United States
Phone: 1 6172530611
Fax: 1 6172530369
Email: rmorris@media.mit.edu

Abstract

Background: Self-guided, Web-based interventions for depression show promising results but suffer from high attrition and low user engagement. Online peer support networks can be highly engaging, but they show mixed results and lack evidence-based content.

Objective: Our aim was to introduce and evaluate a novel Web-based, peer-to-peer cognitive reappraisal platform designed to promote evidence-based techniques, with the hypotheses that (1) repeated use of the platform increases reappraisal and reduces depression and (2) that the social, crowdsourced interactions enhance engagement.

Methods: Participants aged 18-35 were recruited online and were randomly assigned to the treatment group, “Panoply” (n=84), or an active control group, online expressive writing (n=82). Both are fully automated Web-based platforms. Participants were asked to use their assigned platform for a minimum of 25 minutes per week for 3 weeks. Both platforms involved posting descriptions of stressful thoughts and situations. Participants on the Panoply platform additionally received crowdsourced reappraisal support immediately after submitting a post (median response time=9 minutes). Panoply participants could also practice reappraising stressful situations submitted by other users. Online questionnaires administered at baseline and 3 weeks assessed depression symptoms, reappraisal, and perseverative thinking. Engagement was assessed through self-report measures, session data, and activity levels.

Results: The Panoply platform produced significant improvements from pre to post for depression (*P*=.001), reappraisal (*P*<.001), and perseverative thinking (*P*<.001). The expressive writing platform yielded significant pre to post improvements for depression (*P*=.02) and perseverative thinking (*P*<.001), but not reappraisal (*P*=.45). The two groups did not diverge significantly at post-test on measures of depression or perseverative thinking, though Panoply users had significantly higher reappraisal scores (*P*=.02) than expressive writing. We also found significant group by treatment interactions. Individuals with elevated depression symptoms showed greater comparative benefit from Panoply for depression (*P*=.02) and perseverative thinking (*P*=.008). Individuals with baseline reappraisal deficits showed greater comparative benefit from Panoply for depression (*P*=.002) and perseverative thinking (*P*=.002). Changes in reappraisal mediated the effects of Panoply, but not the expressive writing platform, for both outcomes of depression (ab=-1.04, SE 0.58, 95% CI -2.67 to -.12) and perseverative thinking (ab=-1.02, SE 0.61, 95% CI -2.88 to -.20). Dropout rates were similar for the two platforms; however, Panoply yielded significantly more usage activity (*P*<.001) and significantly greater user experience scores (*P*<.001).

Conclusions: Panoply engaged its users and was especially helpful for depressed individuals and for those who might ordinarily underutilize reappraisal techniques. Further investigation is needed to examine the long-term effects of such a platform and whether the benefits generalize to a more diverse population of users.

Trial Registration: ClinicalTrials.gov NCT02302248; https://clinicaltrials.gov/ct2/show/NCT02302248 (Archived by WebCite at http://www.webcitation.org/6Wtkj6CXU).

(J Med Internet Res 2015;17(3):e72) doi:10.2196/jmir.4167

KEYWORDS

Web-based intervention; crowdsourcing; randomized controlled trial; depression; cognitive behavioral therapy; mental health; social networks

Introduction

Major depressive disorder is a debilitating and costly illness. In the United States alone, depression affects as many as 6.6%-10.3% of the population each year [1,2] and creates a huge economic burden, costing tens of billions of dollars [3]. To address a problem of this magnitude, innovative solutions are needed. Self-guided treatments, such as those delivered via the Web, show promise [4] and have the potential to reduce many of the practical and emotional barriers that typically prevent depressed individuals from seeking traditional psychotherapy [5]. In practice, however, many self-guided interventions suffer from high attrition rates and low levels of engagement. A recent review of self-guided, Web-based treatments found a median completion rate of 56% [6]. Open trials show even higher rates of attrition [7]. Low levels of engagement can be especially problematic and might be one of the reasons that self-guided treatments produce smaller gains than supported methods [4].

To address problems related to engagement and adherence, self-guided treatments can be augmented with external support from clinicians or coaches. Mohr et al, for example, found greater adherence to a self-guided depression intervention when participants were provided weekly 5-10 minute phone calls from an assigned coach [8]. While this approach holds promise, its ability to scale widely may be limited. Potential barriers to access include cost, availability of coaches, and scheduling logistics. Further, many individuals seek out Web-based treatments as an alternative to interacting with a clinician and may not be comfortable seeking support from other professionals [9], even trained coaches. Ideally, individuals should be intrinsically motivated to engage with intervention technologies on their own, without prompting from outside clinicians, coaches, or researchers.

Online peer support networks are extremely popular and are known to naturally engage users. Indeed, Horrigan reports that over 84% of American Internet users have visited an online community group at least once [10]. Anonymous peer-to-peer support apps also attract a wide audience. *Whisper,* the anonymous, confessional peer-to-peer app, attracted well over a billion page views a month in 2014 [11]. However, there remains a paucity of rigorous, controlled studies on the efficacy of online support groups and peer-to-peer support apps for mental health [12]. Future work is needed to determine whether these platforms are as helpful as they claim to be. For some individuals, unmoderated Internet support platforms may actually be detrimental. For example, Kaplan et al showed that individuals who participated frequently on unstructured, online mental health forums reported greater psychological distress over time [13]. Mixed findings with regard to discussion forums and peer-to-peer support apps are not surprising given the lack of oversight on the content provided in these resources.

Still, there may be a way to adapt these platforms, creating peer-to-peer interactions that are structured and moderated to reinforce evidence-based clinical techniques. It may be possible, for instance, to create an intervention that is as engaging and personalized as typical peer-to-peer platforms, while still providing the therapeutic content found in self-guided, clinical programs.

The aim of this paper is to introduce such a system, outline its design and its putative benefits, and evaluate its potential to reduce depression systems and foster engagement. In this paper, we present *Panoply*—a peer-to-peer platform that provides cognitive reappraisal and socioaffective support, anytime, anywhere. In lieu of clinician oversight, Panoply coordinates supportive reappraisals from online crowd helpers, all of whom are trained on demand, as needed. Panoply incorporates recent advances in crowdsourcing and human computation to ensure that interactions are timely [14] and vetted for quality [15].

While Panoply incorporates many novel design features, the overall user experience was built to resemble existing peer support apps: users can post content, respond to others, and get notifications when new interactions have taken place. These interactions provide natural triggers for engagement and are designed to bring users back to the platform again and again [16]. As with the most successful peer-to-peer apps, Panoply is also aligned with how individuals typically engage with technologies today. Users are increasingly likely to “snack” on apps, visiting them frequently, but in short bursts [17,18]. Therefore, instead of bundling app content into long weekly sessions that require repeated, lengthy periods of sustained attention, Panoply accommodates multiple levels of commitment. It offers tutorials and other didactic exercises, but all the content is self-contained in short, bite-sized chunks. Everything can be absorbed piece-meal, if necessary, without requiring extended time commitments on the part of the user. Taken together, these design choices highlight the importance of adapting interventions to current norms of technological interaction and consumption, rather than defaulting to holdovers from face-to-face therapy sessions or “psychological skeuomorphs” [19]. Panoply is also interactive, personal, social, and supportive and therefore includes design features that have been recommended for building engaging Web-based cognitive-behavioral therapy (CBT) interventions [20].

The primary therapeutic approach behind Panoply draws from CBT. Much of CBT’s efficacy relies on teaching people compensatory skills [21], and research supports that cognitive skills are an important mediator of symptom change [22]. A critical skill taught in CBT for depression is cognitive reappraisal—an adaptive emotion regulatory technique that involves reinterpreting the meaning of a thought or situation to change its emotional trajectory [23]. Cognitive restructuring, a form of reappraisal, is one of the most common components in Web-based treatments for depression [24-27]. On Panoply, users are taught reappraisal skills and are trained to think more flexibly and objectively about the stressful events and thoughts that upset them. They learn these techniques experientially, in relation to their own day-to-day problems and negative self-beliefs. They also learn by acting as respondents in the system and applying these techniques to other people. Users do not simply consume reappraisal assistance passively, they actively provide it to others as a way to rehearse and practice this technique, over and over.

While some elements of the Panoply design have been described and analyzed elsewhere [28,29], this paper is the first to introduce the complete peer-to-peer design and assess its effects within a randomized controlled trial design. In this paper, we examine the hypotheses that repeated use of this platform will reduce depression symptoms and that the social, interactive design will promote engagement. We also examine whether reappraisal mediates changes in depression symptoms and perseverative thinking for the Panoply and expressive writing platforms.

Methods

Study Design and Participants

We conducted a parallel-arm randomized controlled trial (RCT), assigning participants to either the Panoply intervention or an active control intervention (online expressive writing). The study was approved by the MIT Committee on the Use of Humans as Experimental Subjects (ref. no. 1311006002). Recruitment took place between April and June 2014. To be eligible for the trial, participants needed to be native English speakers between the ages of 18 and 35. This age range was selected because users in this age group are more likely to have experience with anonymous, social messaging platforms, and this study sought to find initial support for this platform rather than investigating widespread implementation.

Participants were recruited from various universities, Internet websites (craigslist, research portals), and through social media channels (Facebook, Twitter). Participants signed up on the Web, by submitting their emails on the study recruitment website. The study was advertised as an opportunity to try a new Web-based stress reduction app. It was open to the general public and depression status was not an inclusion criterion. Depression was not mentioned in any of the recruitment materials. Participants were not paid directly for participation in the study. Instead, all participants who completed the baseline and follow-up assessments were offered a chance to win an iPad Mini (valued at US $300). Use of the platform was not a factor for being eligible to win the iPad Mini.

Procedures

Participants who submitted their emails were assigned unique, anonymous study IDs and were randomized to condition on a 1-to-1 ratio. Randomization occurred prior to any screening procedures and before participants received descriptions of their assigned intervention in the consent form. Obtaining consent after randomization can increase the probability of individuals participating in a research study [30] and has better generalizability to real-world settings in which a platform like Panoply would be used. This approach also prevented our control participants from feeling unmotivated, simply because they felt they had been assigned a less exciting, less social app. With this approach, we were less likely to encounter the “resentful moralization problem”, which is a bias that can occur when consent is provided prior to randomization [31,32]. However, this approach might result in more attrition in the stage immediately following randomization as users at this point have expressed only preliminary interest in this study. When barriers to entry are low in studies of Internet interventions, dropout, especially early on, might be very high, possibly biasing results when traditional methods of dealing with missing data are applied [33]. Typically, intent-to-treat analysis examines all those who are randomized, but that might be inappropriate when randomization is conducted prior to consent. As such, our study analyzed individuals who were consented.

Randomization and email correspondence were automatically coordinated through scripts we wrote in the Python programming language. After completing the consent form and baseline assessments, participants were asked to use their study IDs to create an anonymous account on their assigned platform. They were told to use the app for at least 25 minutes per week, for 3 weeks. To best approximate real usage with an unmoderated app, participants were not given any further instructions about how to use their assigned system. Instead, participants were told to use the app in ways that best fit their schedules and interests. Participants in both groups received four automated emails throughout the study reminding them to use their assigned app. After 3 weeks, participants were emailed a link to the follow-up assessments. The online assessments were hosted by SurveyGizmo and required a unique study ID to log in. This prevented multiple submissions. Incomplete survey data were not included in the analyses.

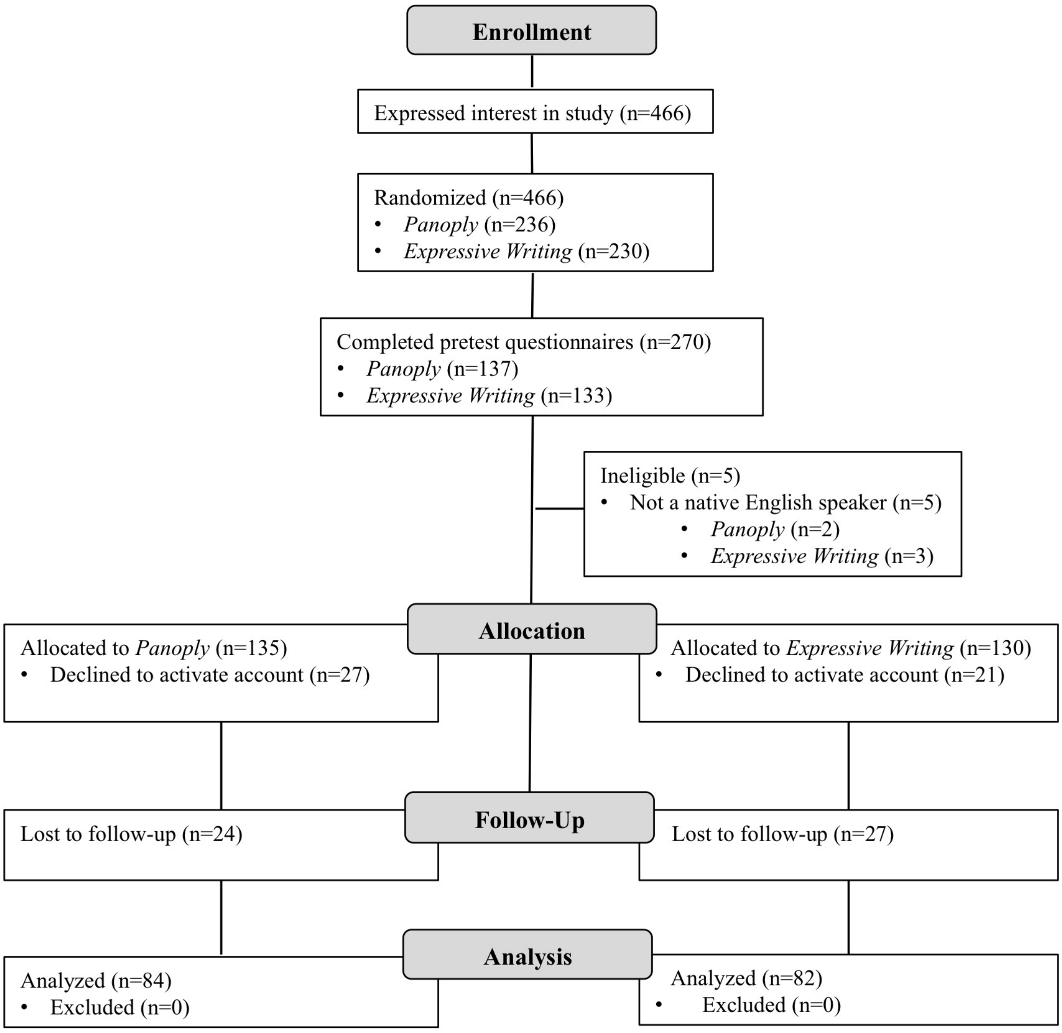
As shown in the CONSORT diagram (Figure 1), 270 individuals completed the online consent form and baseline questionnaires. Five were excluded from the study after consenting because they reported being non-native English speakers. See Multimedia Appendix 1 for the CONSORT checklist [34].

A total of 217 individuals activated an account (Panoply=108, expressive writing=109). Of these, 166 (76.5%) completed the follow-up questionnaires. No significant differences existed in the rates of dropout between the control or treatment interventions at any stage of the study.

Three individuals reported dropping out prematurely because they were not truly interested in a stress reduction app but wanted to explore the new technology. The social media advertisements were broadcast from MIT’s Media Lab, which has a reputation for high-tech innovation, and it is likely that this recruiting channel attracted tech-curious individuals who were not actually in need of an intervention. It is likely that others dropped out for similar reasons. Other reasons for dropout included not having enough time to adhere to the recommended 25/min per week guidelines (n=2), being out of town (n=1), or not having reliable access to a desktop computer (n=1). The remaining 41 individuals who did not activate an account could not be reached for comment and did not respond to our emails. Those who activated an account but did not complete follow-up assessments (n=51) could also not be reached for comment, despite three separate attempts to reach them by email.

Though all study procedure emails were automated, participants could email the experimenters directly during the study if they needed clarifications about the procedures or if they had technical difficulties using their assigned app. To be able to answer specific questions about either the control or treatment app, experimenters were not blind to the random assignment of participants. However, during the course of the study, only four participants emailed the experimenters to request technical support or procedure clarification.

Figure 1. CONSORT flow diagram.



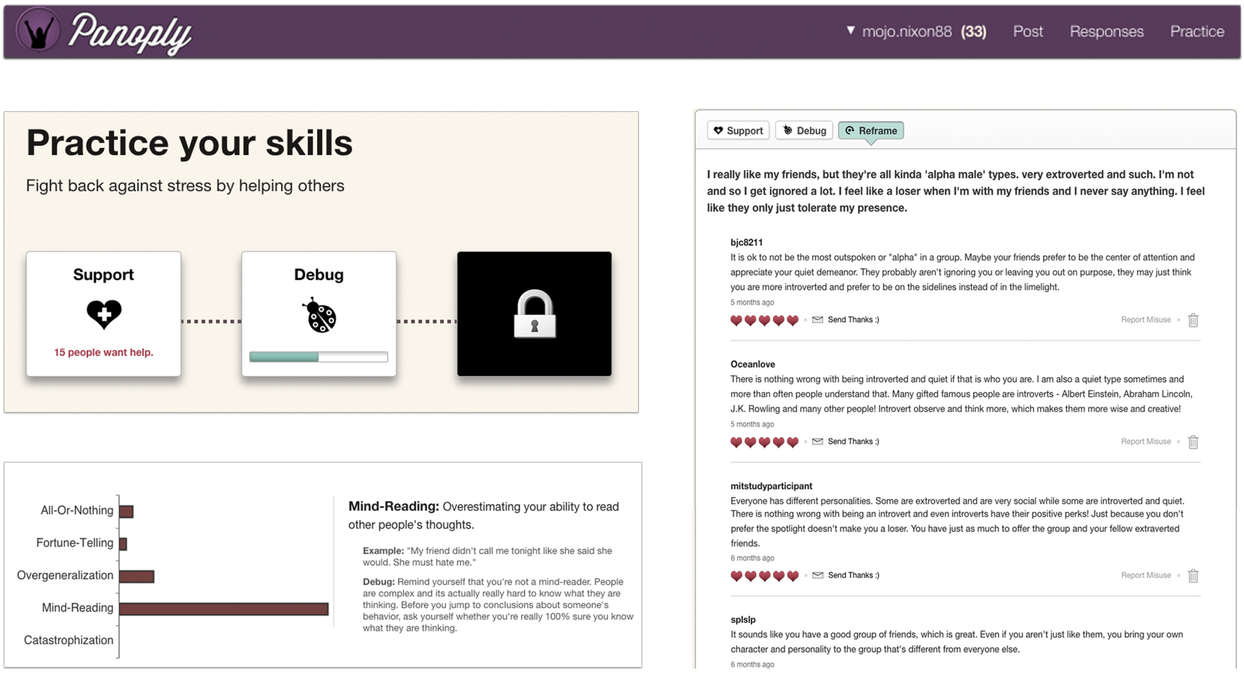
Interventions

Treatment Group (Panoply)

Overview

Panoply is a peer-to-peer, Web-based cognitive reappraisal program developed at the MIT Media Lab. Unlike other Web-based psychoeducation platforms (eg, [35,36]), Panoply does not rely on static, didactic content to teach therapeutic techniques. Rather, the Panoply platform is a dynamic, social, and interactive platform. Panoply users can post content, respond to others, receive responses, and get feedback on their performance (Figure 2). Panoply also offers additional structure, training, and moderation to ensure that all interactions on the site are aligned with evidence-based therapeutic techniques. In the sections that follow, we examine the three core behaviors that occur on Panoply: posting content, responding to others, and receiving responses.

Figure 2. Screenshots from the Panoply platform, illustrating the tutorial panel, the debug dashboard, and the reframe responses.



Practice your skills

Fight back against stress by helping others

Mind-Reading: Overestimating your ability to read other people's thoughts.

Example: "My friend didn't call me tonight like she said she would. She must hate me."

Debug: Remind yourself that you're not a mind-reader. People are complex and its actually really hard to know what they are thinking. Before you jump to conclusions about someone's behavior, ask yourself whether you're really 100% sure you know what they are thinking.

I really like my friends, but they're all kinda 'alpha male' types. very extroverted and such. I'm not and so I get ignored a lot. I feel like a loser when I'm with my friends and I never say anything. I feel like they only just tolerate my presence.

bjc8211
It is ok to not be the most outspoken or "alpha" in a group. Maybe your friends prefer to be the center of attention and appreciate your quiet demeanor. They probably aren't ignoring you or leaving you out on purpose, they may just think you are more introverted and prefer to be on the sidelines instead of in the limelight.
5 months ago

Oceanlove
There is nothing wrong with being introverted and quiet if that is who you are. I am also a quiet type sometimes and more than often people understand that. Many gifted famous people are introverts - Albert Einstein, Abraham Lincoln, J.K. Rowling and many other people! Introvert observe and think more, which makes them more wise and creative!
5 months ago

mitstudyparticipant
Everyone has different personalities. Some are extroverted and are very social while some are introverted and quiet. There is nothing wrong with being an introvert and even introverts have their positive perks! Just because you don't prefer the spotlight doesn't make you a loser. You have just as much to offer the group and your fellow extraverted friends.
6 months ago

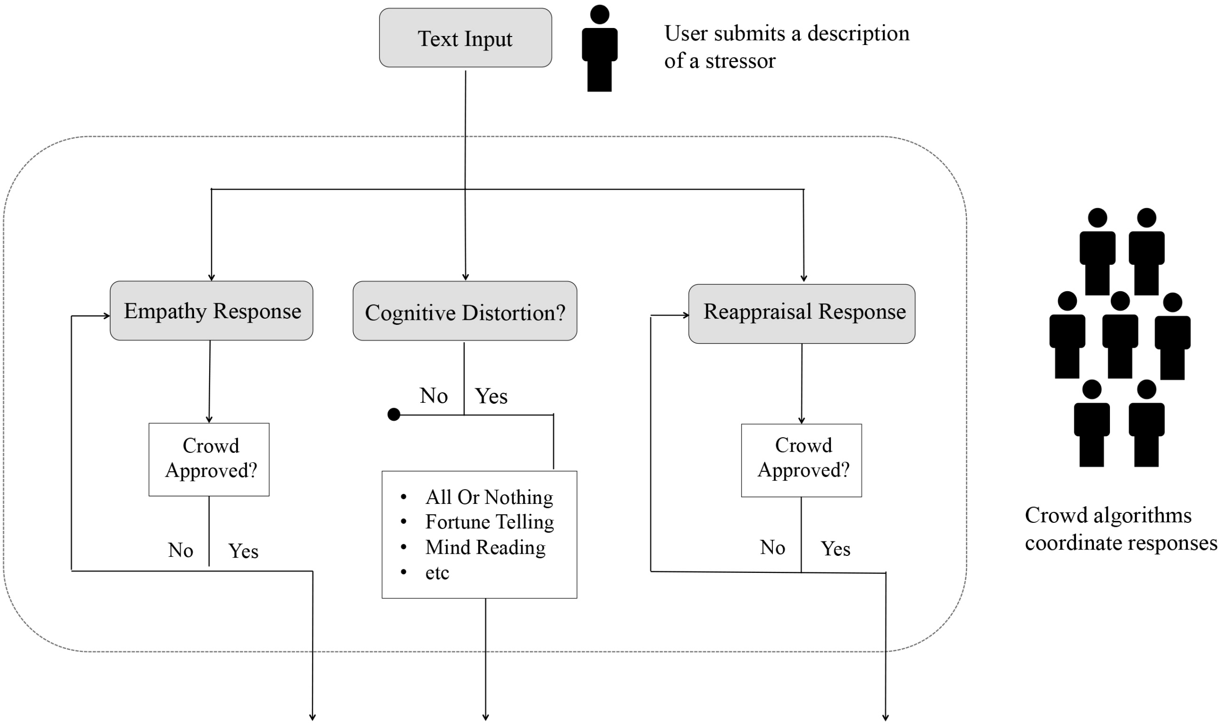
splslp
It sounds like you have a good group of friends, which is great. Even if you aren't just like them, you bring your own character and personality to the group that's different from everyone else.
6 months ago

Posting Content

The initial activity on Panoply involves posting short descriptions of negative thoughts and situations (500 characters maximum). When posting, users are asked to first describe a stressful situation, using one to two sentences. Next, they are asked to record any automatic negative thoughts they might have about the situation. A short tutorial helps first-time users understand the difference between negative situations and the automatic negative thoughts associated with them.

Once a user posts on Panoply, a sequence of crowdwork is automatically set into motion (Figure 3). First, crowd helpers from Amazon’s Mechanical Turk service (MTurk) are hired to review each post. Any post that contains offensive material, off-topic content, or language related to self-harm is excluded from the system. In the case of language related to self-harm, an automated email is immediately sent to the author of the post. The email includes links to mental health resources and reminds the poster that the system is a self-help tool, not to be used for crisis-related situations. Once a post is approved, it is automatically delivered to several sets of trained respondents.

Figure 3. Multiple sets of crowd workers are coordinated to compose and curate responses on the system. The cognitive distortions are identified and labeled, but not subject to crowd review.



Responding to Others

Before being given a chance to respond to the post of a peer, each respondent is trained to use a specific therapeutic technique. Respondents are taught to (1) offer empathy, (2) identify cognitive distortions, or (3) help users reframe negative situations in ways that are more positive and adaptive. Also, responses are vetted by other crowd helpers before being returned to the person who made the original post. If a response is deemed inappropriate or abusive, it is immediately discarded. All of the aforementioned interactions are coordinated entirely through Panoply's automation. The user needs to only submit their post to start this sequence of crowd work.

Respondents in our study were a mixture of other Panoply users and paid workers from MTurk. Workers from MTurk are used for several reasons. First, because Panoply is not yet a large, peer-to-peer system, MTurk provides a temporary stand-in for a large user base, helping ensure that users receive lots of responses extremely quickly (the median response time on Panoply was 9 minutes). Second, MTurk workers can be hired at marginal cost to efficiently moderate content ($0.01) and composes responses ($0.10-0.14) on the system. In our study, the average weekly cost per Panoply user was US $1.04. This could eventually drop to zero. Indeed, if Panoply were to attract a large and active user base, MTurk workers would probably not be needed.

Training for both MTurk workers and Panoply site users occurs on demand, as needed, and involves short, 3-5 minute training modules. Users are introduced to a specific therapeutic technique, are shown positive and negative exemplars of responses, and complete an interactive quiz to assess comprehension. After successfully completing the training, users are given the opportunity to practice the technique by responding to a post from a real Panoply user.

MTurk workers receive a small amount of payment for each response they compose. Panoply users, by contrast, contribute for free. They are told that each time they respond to others they get to practice techniques that are important for managing stress and negative emotions. They are reminded that teaching others can be an exceptionally great way to learn. This concept, a form of peer-based learning, has been studied at length in pedagogical research [37]. To our knowledge, this peer-based learning approach has rarely, if ever, been used in the context of Web-based depression interventions.

Responses on Panoply fall into three categories: support, debug, and reframe. These categories were drawn from evidence-based practices for depression and have been examined on earlier versions of the Panoply system [28]. Support responses offer emotional support and active listening. Debug responses help users identify and dispute cognitive distortions ("bugs"). Reframe responses offer alternative, more positive ways of thinking about the stressful situation. Respondents are not asked to use any one particular reappraisal strategy but instead are given a bulleted list of tactics to consider in case they need inspiration. These prompts were culled from reappraisal taxonomies and strategies cited in the emotion regulation research literature [38,39].

Receiving Responses

The response panel on Panoply features a button-based navigation bar that lets users switch between the three types of responses generated by the crowd. When a user visits the response panel, support messages are displayed first by default and are listed in a newsfeed format (the most recent appearing at the top). Users can rate these responses as they appear in the interface. If particularly moved, users can compose a short note, thanking the respondent for their contribution.

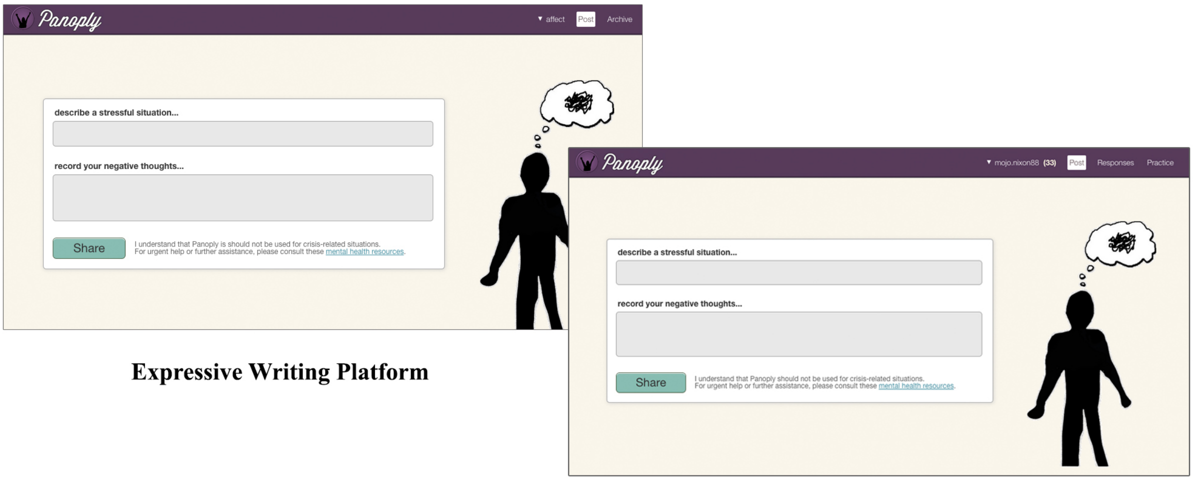
The debug section features a graphical dashboard of cognitive distortions (“bugs”) identified by the crowd. The dashboard includes personalized suggestions that describe how to restructure the specific distortions that were observed by the crowd.

The reframe section features short messages, displayed in the same newsfeed format as the support responses. Unlike the other response categories, however, the reframes are not revealed outright. When users get notified of new responses in this category, they are not able to view them initially. Instead, users must compose a reappraisal for themselves before they can access responses from the crowd. The hope is that users will be naturally moved to complete reappraisals for themselves because they know that doing so unlocks interesting new social content.

Control Group (Online Expressive Writing)

The visual and interface design for the control condition was built to mirror the Panoply intervention. The instructions for describing stressful situations and negative thoughts were exactly the same (Figure 4). However, users in this condition did not receive feedback from the crowd and were not given the opportunity to provide feedback for others. We did not expect this condition to be inert because writing expressively about negative experiences is a well-studied and efficacious intervention it its own right and can help reduce depression symptoms [40]. A meta-analysis has found that expressive writing in various formats can improve physical and psychological health outcomes [41]. Our writing condition was a useful control because although it matched Panoply on nonspecific factors (eg, Web design, user registration, composing negative thoughts), it did not contain reappraisal training or crowdsourced interactions. However, it did allow users to engage in a similar process of entering content, thus allowing for comparisons that control for the effects of being online and putting feelings into words.

Figure 4. The control and treatment platforms were matched on non-specific factors, including visual design, user interface design, and logotype.



I understand that Panoply is should not be used for crisis-related situations.
For urgent help or further assistance, please consult these mental health resources.

I understand that Panoply should not be used for crisis-related situations.
For urgent help or further assistance, please consult these mental health resources.

Expressive Writing Platform

Panoply Platform

Usability Testing

Both platforms underwent several usability studies, both in the lab and over the Web. For lab-based studies, an experimenter was present at all times and the sessions were moderated using techniques such as “concurrent thinking aloud”, “retrospective thinking aloud”, and “retrospective probing” [42]. MTurk workers were also recruited online to help identify potential design flaws. These studies helped identify user experience issues and points of confusion around site navigation and other user interface components (eg, buttons, links), enabling us to refine the usability before the RCT was conducted.

Outcome Measures

Participants completed assessments online, both at baseline and at 3-weeks’ follow-up. The primary outcome measure was the Center for Epidemiologic Studies Depression Scale (CES-D) [43], a 20-item self-report scale that assesses symptoms of depression. Secondary outcome measures included reappraisal frequency, as assessed by the Emotion Regulation Questionnaire - Reappraisal (ERQ-R) [23], and maladaptive rumination, as assessed by the Perseverative Thinking Questionnaire (PTQ) [44].

Center for Epidemiologic Studies Depression Scale

The CES-D [43] is a 20-item self-report scale that assesses symptoms of depression. Respondents are asked to indicate the extent to which they have felt various depression symptoms over the past week. The questions address symptoms such as loss of appetite, depressed mood, and feelings of loneliness. A score of 16 or higher suggests a high level of depression and is often used as a cut-off to determine clinically relevant symptoms.

Emotion Regulation Questionnaire – Reappraisal

The ERQ is a 10-item questionnaire that assesses individual differences in the habitual use of two emotion regulation strategies: cognitive reappraisal and expressive suppression. It produces scores for both reappraisal and suppression. For the purposes of this study, we analyzed only the reappraisal scores. Reappraisal is considered an adaptive regulatory strategy and is associated with positive psychological functioning, including increased positive affect, well-being, and interpersonal functioning [23].

Perseverative Thinking Questionnaire

Depressive rumination is a cognitive style that involves repetitive elaboration of the symptoms and causes of distress. In essence, it is an unproductive form of reappraisal. Instead of recasting a situation in ways that lead to positive recontextualizations and problem solving insights, depressive rumination mires individuals in circular reinterpretations that serve only to magnify distress. Rumination is considered a risk factor for depression and suicide and is thought to play a causal role in the development and maintenance of depressive illness. The PTQ [44] is a 15-item scale that assesses three components of rumination: its repetitiveness, its unproductiveness, and its tendency to capture mental capacity.

User Experience Questionnaire

To assess engagement, we administered an online version of the User Experience Questionnaire (UEQ) [45]. This self-report measure examines a product's ability to promote an engaging user experience. The UEQ includes 26 pairs of contrasting attributes (eg, "pleasant vs unpleasant"; "motivating vs demotivating") that are ordered along a 7-point bipolar Likert scale. For each test item, the Likert points represent gradations between the two labels. Selections indicate which of the two labels applies best to the technology being assessed.

Behavioral Activity Levels

We also examined activity levels on both the treatment and control apps. These data were logged automatically by our server and by Google Analytics.

Adherence rate for module completion, while a common metric for many online mental health interventions, does not apply to Panoply or the expressive writing intervention. Neither platform utilized the kind of psychoeducation modules that are typically found in Web-based CBT interventions. Further, Donkin et al recently examined the relationship between various engagement metrics and outcome in an online intervention for depression and found that the total number of modules completed was less important than the level of activity observed per login [6].

Therefore, for behavioral measures of engagement, we examined usage level and assessed the amount of activity observed per login. Specifically, we compared the average number of words written by individuals in the treatment versus control group. This is a useful metric because both the Panoply and the expressive writing interventions involve a considerable amount of writing. Writing is the only task activity one can perform on the expressive writing task. Similarly, with the exception of the "debug" exercise and the training modules, all activities on Panoply require writing. A Python script was used to compute the number of words submitted by individuals in the treatment and control conditions.

To assess the frequency and duration of logins, "sessions" data from Google Analytics were used. A "session" is defined as the period of time a user interacts with a site. Google sessions expire as soon as a user is inactive for 30 minutes.

Analytic Plan

We used chi-square and *t* tests to evaluate whether randomization yielded equivalent demographic and symptom characteristics at baseline for the treatment and control groups. Chi-square tests were also used to compare rates of dropout between the two interventions. For engagement analyses, *t* tests were used to compare UEQ scores and word counts across the treatment and control interventions.

For psychological outcomes, we conducted a set of planned analyses to explore the overall effects of the platforms, as well as specific moderators and mediators. The stages of these analyses follow from our primary hypotheses. To reduce unnecessary multiple tests, we progressed to the next stage of analysis given significant findings at each stage. First, we explored changes within and between each condition. Our primary analyses compared the difference between the groups at post-test using linear regression models controlling for baseline levels of the dependent measure. Because Panoply was designed to teach reappraisal skills, we hypothesized that Panoply would result in greater improvements for those with deficits in this skill (as measured by reappraisal on the ERQ at baseline), and we posited that reappraisal might be a useful mechanism of action within the Panoply condition. Secondary analyses then added moderator variables as interactions within these models to explore if people with different characteristics at baseline benefited more or less from the intervention. For depression status, participants were separated into two groups based on normative values on the CES-D. Following the standard cutoff for clinically meaningful symptoms, we classified individuals scoring 16 or higher as depressed. Based on this categorization, 47 Panoply participants and 44 expressive writing participants were classified as depressed. For reappraisal, participants were dichotomized into two groups (high and low reappraisers) based on a median split. This resulted in 41 low reappraisers for Panoply and 28 low reappraisers for the writing platform. Lastly, because we believed that change in reappraisal is the key skill taught through Panoply, we explored whether reappraisal was a mediator of changes in depressive symptoms and perseverative thinking. We examined mediation using Preacher and Hayes (2008) bootstrapping procedure and SPSS macro. This procedure produces the bias-corrected and accelerated bootstrapped confidence intervals of the product of the direct pathways between condition and the mediator (a) and the mediator and the outcome (b) to estimate the indirect effect (ab).

All participants who activated an account and completed follow-up assessments were included in the analyses. Some participants, however, were lost to follow-up and were not included in the analysis of outcomes. As only two assessment time points were obtained, methods of data imputation were not used.

Results

Participant Characteristics

No significant differences in baseline characteristics or dropout rates were observed between the control and treatment interventions (Table 1). The sample was 71.7% female, with a mean age of 23.7. Participants were well educated: 88.6% reported having had at least some college education and 45.8% reported having a 4-year college degree or higher.

Table 1. Baseline characteristics of the participants.

Characteristics	Total sample (n=166)	Treatment (n=84)	Control (n=82)	t or χ^2	P
Demographics					
Age, mean (SD)	23.7 (5.3)	23.5 (5.2)	23.9 (5.5)	-0.4[b]	.70
Female, n (%)	119 (71.7)	62 (73.8)	57 (69.5)	0.20[c]	.66
Higher education,[a] n (%)	76 (45.8)	37 (44.1)	39 (47.6)	0.09[c]	.77
Baseline scores, mean (SD)					
CES-D	19.0 (10.4)	19.4 (10.2)	18.6 (10.6)	0.52[b]	.61
ERQ-R	26.4 (6.9)	26.0 (6.9)	26.7 (6.7)	0.72[b]	.47
PTQ	47.5 (10.9)	46.8 (10.7)	48.2 (11.1)	0.87[b]	.39

[a] Equivalent to a 4-year Bachelor's degree or higher.

[b] t_{164}.

[c] χ^2_1.

Psychological Outcomes

We first examined changes in primary (depression symptoms) and secondary (reappraisal, perseverative thinking) outcomes within each condition. Table 2 displays baseline and post-intervention scores on all outcomes for each condition. Participants in the Panoply condition reported significant changes on all dependent measures, whereas those in the expressive writing condition reported significant changes in depression and perseverative thinking.

We then tested whether differences existed between the conditions at post-intervention controlling for baseline levels. No significant differences existed between the control and treatment conditions at post-test on depressive symptoms, controlling for baseline depression: t_{164}=-0.82, P=.41, beta=-1.02, 95% CI -3.47 to 1.43. Similarly, no significant differences between conditions were found for perseverative thinking: t_{164}=-0.56, P=.57, beta=-.82, 95% CI -3.69 to 2.05. Panoply users did, however, report significantly greater levels of reappraisal compared to users of the expressive writing condition controlling for baseline levels: t_{164}=2.29, P=.02, beta=1.98, 95% CI 0.27-3.68.

Second, we wanted to assess whether Panoply was more useful for users with certain characteristics. Indeed, we observed that depression status was a significant moderator of both post-test depressive symptoms, t_{164}=-2.28, P=.02, beta=-5.53, 95% CI -10.3 to -0.76, and perseverative thinking, t_{164}=-2.70, P=.008, beta=-7.79, 95% CI -13.47 to -2.10. The same was true for high versus low reappraisers. Reappraisal scores at baseline were a significant moderator of depressive symptoms, t_{164}=3.12, P=.002, beta=7.64, 95% CI 2.80-12.47, and perseverative thinking, t_{164}=3.19, P=.002, beta=9.14, 95% CI 3.49-14.80. Participants who were depressed or low reappraisers at baseline benefited more from Panoply compared to expressive writing. The importance of depression status and reappraisal in terms of predicting who benefits, in addition to the fact that Panoply had a stronger effect on change in reappraisal than expressive writing, suggests that change in reappraisal might be an important mediator of the benefits of Panoply, so we investigated that further.

Using Preacher and Hayes' (2008) bootstrapping method, we examined whether improvement in reappraisal was a mediator of the effect of Panoply. Figure 5 displays the results of this analysis for depressive symptoms. The effect of treatment on change in reappraisal (a) was statistically significant (B=.39, SE 0.16, P=.02, 95% CI 0.08-0.70). The effect of change in reappraisal on change in depression (b) was also statistically significant (B=-2.55, SE 0.63, P<.001, 95% CI -3.78 to -1.32). The indirect effect of Panoply on changes in depression via changes in reappraisal was statistically significant (ab=-1.04, SE 0.58, 95% CI -2.67 to -0.12). These results suggest that change in reappraisal may be a specific mechanism of Panoply compared to the writing condition in reducing depressive symptoms.

We also assessed whether change in reappraisal was a mediator of changes in perseverative thinking. Figure 6 displays the results of this analysis. The (a) pathway is the same as the previous analysis. The effect of change in reappraisal on change in perseverative thinking (b) was also statistically significant

(*B*=-2.69, SE 0.68, *P*<.001, 95% CI -4.02 to -1.36). The indirect effect of Panoply on changes in depression via changes in reappraisal was statistically significant (ab=-1.02, SE 0.61, 95% CI -2.88 to -0.20). These results suggest that change in reappraisal may be a specific mechanism of Panoply compared to the writing condition in reducing perseverative thinking.

Table 2. Changes from baseline to post-intervention by condition.

	Pre, mean (SD)	Post, mean (SD)	Change, mean (SD)	95% CI	*T* [a]	*P*	*d*
Panoply							
CES-D	19.38 (10.16)	15.79 (9.53)	-3.60 (9.86)	-5.74 to -1.45	3.34	.001	-0.36
ERQ-R	25.99 (6.91)	28.92 (6.22)	2.93 (6.11)	1.60 to 4.25	4.39	<.001	0.48
PTQ	46.76 (10.70)	42.35 (11.04)	-4.42 (10.13)	-6.62 to -2.22	4.00	<.001	-0.44
Expressive Writing							
CES-D	18.55 (10.60)	16.33 (10.38)	-2.20 (8.24)	-6.03 to -0.41	2.44	.02	-0.24
ERQ-R	26.74 (6.65)	27.32 (6.78)	0.57 (6.85)	-0.93 to 2.08	0.76	.45	0.06
PTQ	48.23 (11.11)	44.21 (13.12)	-4.02 (9.54)	-6.12 to -1.93	3.82	<.001	-0.44

[a] *df*=83 for *Panoply* and 81 for Expressive Writing.

Figure 5. Mediation of change in reappraisal on change in depression.

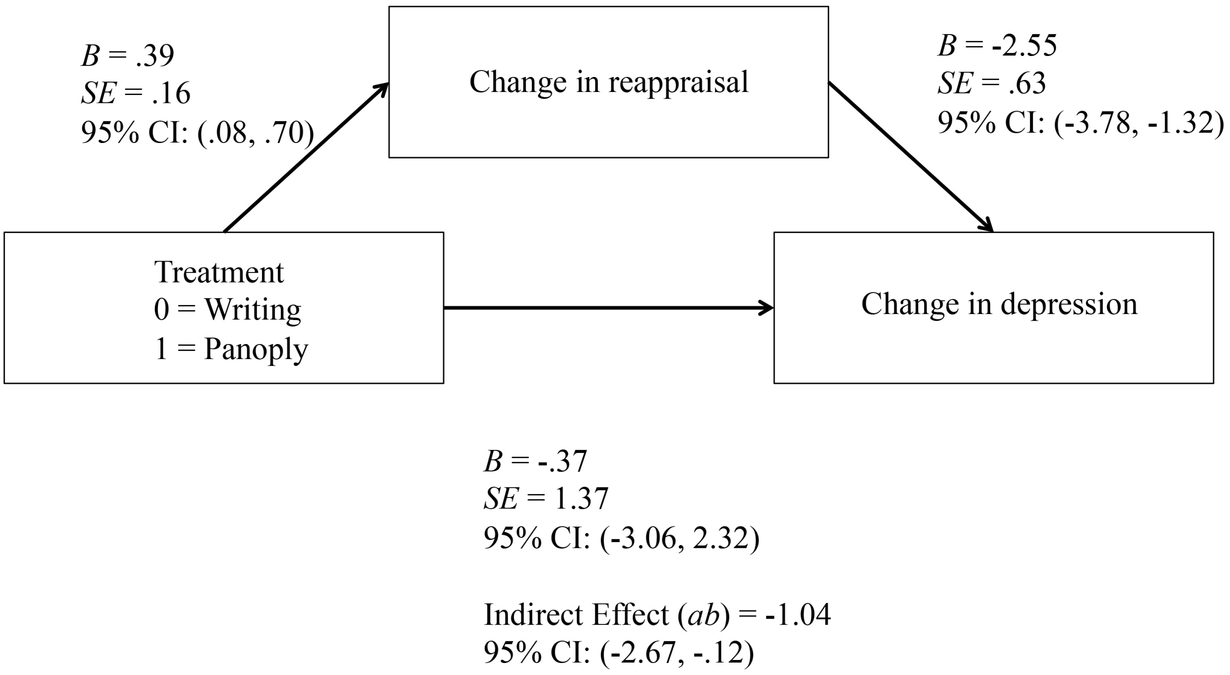
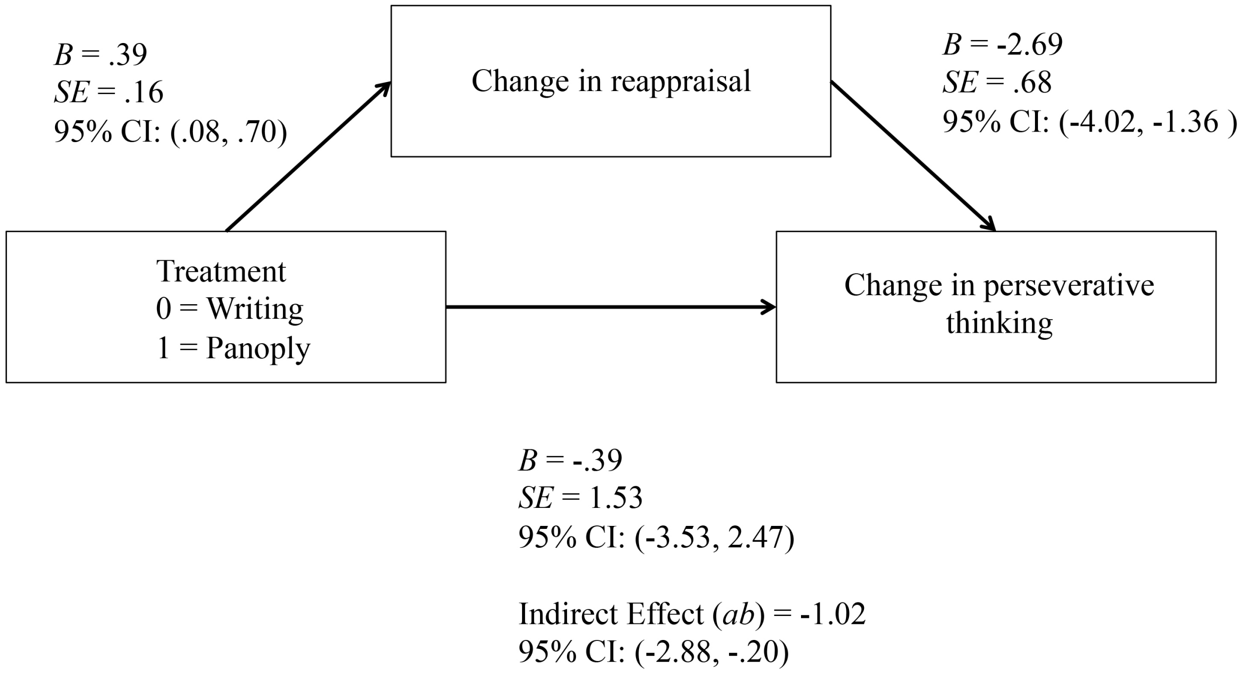


Figure 6. Mediation of change in reappraisal on change in perseverative thinking.



Engagement Outcomes

Engagement measured by the UEQ was significantly higher for the Panoply platform (mean 137.10, SD 20.93) than the expressive writing platform (mean 122.29, SD 20.81), (t_{164}=4.57, *D*=.71, 95% CI 1.02-0.39, *P*<.001).

There was a significant difference in activity level between conditions ($t_{164=}$-4.1, *D*=.62, 95% CI 0.31-0.94, *P*<.001), with individuals in the Panoply condition writing significantly more words (mean 1013.28, SD 1145.14) than those in the expressive writing condition (mean 433.85, SD 609.59). This is striking, given that writing was not the only activity available to Panoply users. They could also identify distortions and review crowd-generated feedback—two additional activities that were used frequently but were not included in the word count metric.

Descriptive statistics from the Google session data revealed that users in the Panoply condition logged 21 sessions over the 3-week deployment on average. Their average time per session was 9 minutes and 18 seconds per session. By comparison, users in the expressive writing condition logged an average of 10 sessions, spending an average of 3 minutes and 10 seconds per session. Thus, the Panoply group averaged a total of over 195 minutes over the course of the study, a considerably longer amount than was suggested (75 minutes). Inferential statistics could not be computed, because the Google session data were not revealed at the level of the user.

Usage Patterns

In addition to directly comparing measures of engagement between the treatment and control groups, we also captured general usage patterns for each platform (see Table 3). For the purposes of this analysis, we included everyone who activated an account and accessed the site (N=214), even if they did not return to complete follow-up assessments.

Of note is the fact that individuals assigned to the writing condition submitted considerably more posts than those in the Panoply condition. There are several possible explanations for this. First, those in the writing condition had only one task to do. Their attentions were never diverted elsewhere; the number of posts they wrote reflects their entire contribution to the site. By comparison, those in the Panoply condition could divide their time between submitting posts, responding to others, and reviewing responses from the crowd. Second, those in the Panoply condition may have been more tentative about submitting posts, simply because they had an audience. To the extent that submitting posts was therapeutic, participants in the Panoply condition, on average, received less than half the dose of those in the writing condition. Future designs of Panoply should offer additional incentives for users to post more frequently if this activity is determined to be helpful. For example, users might be given the option to record negative thoughts privately, should they want to reframe their thoughts on their own, without any input from the crowd.

Table 3. Average usage patterns for the Panoply and expressive writing apps for 3 weeks.

Activity	Group	Mean frequency (SD)
Posts		
	Panoply	2.72 (2.76)
	Writing	6.62 (8.76)
Responses		
Support	Panoply	8.94 (11.95)
Debug	Panoply	10.82 (13.72)
Reframe	Panoply	5.99 (8.71)

Adverse Events

One study participant composed several troubling and off-topic posts on the Panoply platform. MTurk workers detected this behavior and an automated email was sent, reminding the participant that Panoply is a self-help tool, not a formal mental health resource. Links to mental health resources were also emailed automatically to this participant. After consulting with the MIT IRB, we decided to prevent this participant from posting any further content. This individual was not withdrawn from the study, however, and was still allowed to compose responses to other Panoply users. None of the responses this individual made to others were flagged as off-topic, malicious, or otherwise inappropriate.

Discussion

Psychological Outcomes

Overall, participants allocated to the Panoply platform received greater clinical benefits than those assigned to the writing task. While the two platforms did not diverge significantly at post-test with respect to depression or perseverative thinking, Panoply users reported significantly high levels of reappraisal. Further, follow-up analyses suggest that individuals with elevated depression symptoms stand to benefit more from a platform like Panoply than from expressive writing. Panoply produced significantly less depression and perseverative thinking for individuals with high depression scores at baseline. A similar pattern was found for individuals who scored low on reappraisal at baseline.

As opposed to many Web-based interventions for depression that attempt to teach a variety of strategies drawn from CBT [8,24-27], Panoply specifically targets cognitive reappraisal. A benefit of this approach is that it allows testing the specific mechanism of change corresponding to that behavior change principle. Follow-up analyses suggest that the benefits accrued from Panoply appeared to be mediated by changes in reappraisal skills. Specific, targeted interventions such as Panoply, with a well-understood mechanism of action, can offer personalized treatment approaches, providing a powerful resource for those who do not typically use reappraisal skills to regulate emotions.

Engagement Outcomes

Panoply was engineered to be an engaging mental health intervention. The final system incorporated many features that were specifically designed to enhance user experience. Indeed, it was hoped that many users would find the crowdsourced interactions particularly novel, motivating, and exciting. Therefore, it was hypothesized that Panoply would score higher on both self-reported user experience and behavioral measures of activity, relative to the expressive writing condition. These hypotheses were confirmed.

Limitations

There are several methodological shortcomings that limit the generalizability of this study. First, the duration of the study was extremely short (3 weeks). While it is encouraging that Panoply managed to confer benefits in this short time, it is unclear how enduring these improvements might be. Additional long-term follow-ups are needed. Also, the study was limited to individuals aged 18-35. Additional research is needed to examine whether similar effects might be observed for other populations of users. Our sample was also largely female, and future studies will need to seek a more balanced gender distribution. Moreover, while expressive writing was a useful control comparison for many reasons, future studies should compare Panoply to more traditional, Web-based CBT programs.

In addition to these methodological limitations, there were some design shortcomings. For the purposes of this study, Panoply was built to target reappraisal skills first and foremost. While this design enabled us to test specific hypotheses about how reappraisal might mediate therapeutic outcomes, it may have limited the therapeutic potential of the platform. Future versions of this type of platform should address other techniques besides just cognitive reappraisal. For instance, Panoply could be extended to address some of the behavioral components of CBT. Behavioral interventions from positive psychology could also be incorporated in future versions, as described by Morris & Picard [29]. Finally, all interactions with Panoply were made through a Web browser, optimized for use on a laptop or desktop. To increase engagement, the platform could be redesigned for mobile use, to better align with contemporary technology usage patterns.

Conclusions

In this paper, we introduced and evaluated a Web-based, peer-to-peer cognitive reappraisal platform designed to promote reappraisal and thus reduce depression symptoms. We found that repeated use of our system produced significant benefits, particularly for depressed individuals and for those who typically underutilize reappraisal strategies. Furthermore, we believe Panoply conferred benefits because it taught reappraisal skills.

On the platform, users gained exposure to reappraisal by (1) receiving reappraisal assistance from the crowd and (2) by repeatedly reframing the thoughts and situations of others on the network. Our mediation analyses suggest that reappraisal helped reduce depression and perseverative thinking for the Panoply platform, but not the expressive writing platform. This supports our hypothesis that Panoply's unique features are especially helpful for building reappraisal skills.

We also found that the platform engaged its users. Indeed, the Panoply platform inspired well over twice as much activity as the control platform. Further work is needed to assess whether these findings might extend to a wider, more diverse set of individuals. The longevity of these effects should also be examined. Measuring engagement over time is an important area for future research. For example, future studies should examine the rates at which individuals revisit intervention platforms on their own, as needed. Unlike other interventions that offer a limited amount of psychoeducation modules, our platform offers a potentially inexhaustible source of varied social content. As long as individuals continue to submit posts on the platform, there remain interesting new opportunities to practice therapeutic techniques. Interventions like ours, that can theoretically be revisited again and again, without appearing stale, could have unique benefits. Further research is needed to address these possibilities.

Acknowledgments

We thank the study participants and MTurk workers for taking part in the trial. This work was supported by the MIT Media Lab member consortium. SMS was supported by a grant from the National Institute of Mental Health K08 MH102336 (PI: SMS).

Conflicts of Interest

Since the work was completed, RRM has taken steps towards forming a company related to peer-produced mental health interventions.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.1 [46].

[PDF File (Adobe PDF File), 3MB - jmir_v17i3e72_app1.pdf]

References

1. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry 1994 Jan;51(1):8-19. [Medline: 8279933]
2. Kessler RC, Berglund P, Demler O. The epidemiology of major depressive disorder: results from the national comorbidity survey replication (NCS-R). J Am Med Assoc 2003;289(23):3095-3105. [doi: 10.1001/jama.289.23.3095]
3. Wang PS, Simon G, Kessler RC. The economic burden of depression and the cost-effectiveness of treatment. Int J Methods Psychiatr Res 2003;12(1):22-33. [Medline: 12830307]
4. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
5. Cartreine JA, Ahern DK, Locke SE. A roadmap to computer-based psychotherapy in the United States. Harv Rev Psychiatry 2010 Mar;18(2):80-95. [doi: 10.3109/10673221003707702] [Medline: 20235773]
6. Donkin L, Hickie IB, Christensen H, Naismith SL, Neal B, Cockayne NL, et al. Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: randomized controlled trial. J Med Internet Res 2013 Oct;15(10):e231 [FREE Full text] [doi: 10.2196/jmir.2771] [Medline: 24135213]
7. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
8. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One 2013 Aug;8(8):e70086 [FREE Full text] [doi: 10.1371/journal.pone.0070086] [Medline: 23990896]
9. Schneider J, Sarrami FP, Grime P, Thornicroft G. Acceptability of online self-help to people with depression: users' views of MoodGYM versus informational websites. J Med Internet Res 2014 Mar;16(3):e90 [FREE Full text] [doi: 10.2196/jmir.2871] [Medline: 24681717]
10. Horrigan JB. Online communities: networks that nurture long-distance relationships and local ties. URL: http://www.pewinternet.org/2001/10/31/online-communities/ [accessed 2014-12-20] [WebCite Cache ID 6Uxs2DJl4]
11. Shontell, A. Secret-sharing app Whisper is nearing 3 billion monthly pageviews because it does something Facebook can't URL: http://www.businessinsider.com/whisper-2013-12 [accessed 2014-12-19] [WebCite Cache ID 6UwF5M6Nr]
12. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? J Med Internet Res 2009;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]

13. Kaplan K, Salzer MS, Solomon P, Brusilovskiy E, Cousounis P. Internet peer support for individuals with psychiatric disabilities: A randomized controlled trial. Soc Sci Med 2011 Jan;72(1):54-62. [doi: 10.1016/j.socscimed.2010.09.037] [Medline: 21112682]
14. Bernstein MS, Brandt J, Miller R, Karger DR. Crowds in two seconds: enabling realtime crowd-powered interfaces. New York: ACM Press; 2011 Presented at: 24th Symposium on User Interface Software and Technology; 2011; New York, New York. [doi: 10.1145/2047196.2047201]
15. Bernstein MS, Little G, Miller RC, Hartmann B, Ackerman MS, Karger DR, et al. Soylent: a word processor with a crowd inside. New York: ACM Press; 2010 Presented at: 23rd Symposium on User Interface Software and Technology; 2010; New York, New York. [doi: 10.1145/1866029.1866078]
16. Fogg B. A behavior model for persuasive design. 2009 Presented at: 4th International Conference on Persuasive Technology; 2009; New York, New York. [doi: 10.1145/1541948.1541999]
17. Oulasvirta A, Tamminen S, Roto V, Kuorelahti J. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. New York: ACM Press; 2005 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 2005; Portland, OR. [doi: 10.1145/1054972.1055101]
18. Vaish R, Wyngarden K, Chen J, Cheung B, Bernstein MS. Twitch crowdsourcing: crowd contributions in short bursts of time. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM Press; 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 2014; Toronto, ON. [doi: 10.1145/2556288.2556996]
19. Schueller SM, Munoz RF, Mohr DC. Realizing the potential of behavioral intervention technologies. Current Directions in Psychological Science 2013 Dec 03;22(6):478-483. [doi: 10.1177/0963721413495872]
20. Doherty G, Coyle D, Sharry J. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. New York: ACM Press; 2012 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 2012; Austin, TX. [doi: 10.1145/2207676.2208602]
21. Barber JP, DeRubeis RJ. On second thought: Where the action is in cognitive therapy for depression. Cogn Ther Res 1989 Oct;13(5):441-457. [doi: 10.1007/bf01173905]
22. DeRubeis RJ, Evans MD, Hollon SD, Garvey MJ, Grove WM, Tuason VB. How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. J Consult Clin Psychol 1990 Dec;58(6):862-869. [Medline: 2292637]
23. Gross JJ, John OP. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. J Pers Soc Psychol 2003 Aug;85(2):348-362. [Medline: 12916575]
24. Andersson G, Bergström J, Holländare F, Carlbring P, Kaldo V, Ekselius L. Internet-based self-help for depression: randomized controlled trial. Br J Psychiatry 2005 Nov;187:456-461 [FREE Full text] [doi: 10.1192/bjp.187.5.456] [Medline: 16260822]
25. Perini S, Titov N, Andrews G. Clinician-assisted Internet-based treatment is effective for depression: randomized controlled trial. Aust N Z J Psychiatry 2009 Jun;43(6):571-578. [doi: 10.1080/00048670902873722] [Medline: 19440890]
26. Ruwaard J, Schrieken B, Schrijver M, Broeksteeg J, Dekker J, Vermeulen H, et al. Standardized web-based cognitive behavioral therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. Cogn Behav Ther 2009 Dec;38(4):206-221. [doi: 10.1080/16506070802408086] [Medline: 19221919]
27. Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cogn Behav Ther 2011 Dec;40(4):251-266. [doi: 10.1080/16506073.2011.616531] [Medline: 22060248]
28. Morris RR, Picard RW. Crowdsourcing collective emotional intelligence. Presented at: Collective Intelligence; 2012; Cambridge, MA.
29. Morris RR, Picard RW. Crowd-powered positive psychological interventions. The Journal of Positive Psychology 2014 May 06;9(6):509-516. [doi: 10.1080/17439760.2014.913671]
30. Armstrong PW, Watts DG. Clinical trials:randomization before consent. Biomedicine 1981 Sep;34(2):65-66. [Medline: 7284544]
31. Torgerson DJ, Sibbald B. Understanding controlled trials. What is a patient preference trial? BMJ 1998 Jan 31;316(7128):360 [FREE Full text] [Medline: 9487173]
32. Bradley C. Designing medical and educational intervention studies. A review of some alternatives to conventional randomized controlled trials. Diabetes Care 1993 Feb;16(2):509-518. [Medline: 8432226]
33. Leykin Y, Aguilera A, Torres LD, Pérez-Stable EJ, Muñoz RF. Interpreting the outcomes of automated internet-based randomized trials: example of an international smoking cessation study. J Med Internet Res 2012 Feb;14(1):e5 [FREE Full text] [doi: 10.2196/jmir.1829] [Medline: 22314016]
34. Eysenbach G, Consort-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

35. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J Med Internet Res 2004 Dec 22;6(4):e46 [FREE Full text] [doi: 10.2196/jmir.6.4.e46] [Medline: 15631970]
36. Geraedts AS, Kleiboer AM, Wiezer NM, van MW, Cuijpers P. Short-term effects of a web-based guided self-help intervention for employees with depressive symptoms: randomized controlled trial. J Med Internet Res 2014 May;16(5):e121 [FREE Full text] [doi: 10.2196/jmir.3185] [Medline: 24800966]
37. Cohen PA, Kulik JA, Kulik CC. Educational outcomes of tutoring: A meta-analysis of findings. American Educational Research Journal 1982 Jan 01;19(2):237-248. [doi: 10.3102/00028312019002237]
38. McRae K, Ciesielski B, Gross JJ. Unpacking cognitive reappraisal: goals, tactics, and outcomes. Emotion 2012 Apr;12(2):250-255. [doi: 10.1037/a0026351] [Medline: 22148990]
39. Jamieson JP, Nock MK, Mendes WB. Mind over matter: reappraising arousal improves cardiovascular and cognitive responses to stress. J Exp Psychol Gen 2012 Aug;141(3):417-422 [FREE Full text] [doi: 10.1037/a0025719] [Medline: 21942377]
40. Gortner EM, Rude SS, Pennebaker JW. Benefits of expressive writing in lowering rumination and depressive symptoms. Behav Ther 2006 Sep;37(3):292-303. [doi: 10.1016/j.beth.2006.01.004] [Medline: 16942980]
41. Frattaroli J. Experimental disclosure and its moderators: a meta-analysis. Psychol Bull 2006 Nov;132(6):823-865. [doi: 10.1037/0033-2909.132.6.823] [Medline: 17073523]
42. Rubin J, Chisnell D. Handbook of usability testing how to plan, design, and conduct effective tests, second edition. Indianapolis, Ind: Wiley Pub; 2008.
43. Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
44. Ehring T, Zetsche U, Weidacker K, Wahl K, Schönfeld S, Ehlers A. The Perseverative Thinking Questionnaire (PTQ): validation of a content-independent measure of repetitive negative thinking. J Behav Ther Exp Psychiatry 2011 Jun;42(2):225-232 [FREE Full text] [doi: 10.1016/j.jbtep.2010.12.003] [Medline: 21315886]
45. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. : Springer-Verlag; 2008 Presented at: Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society on HCI and Usability for Education and Work; 2008; Austria.
46. Eysenbach G, Consort-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

CBT: cognitive-behavioral therapy
CES-D: Center for Epidemiologic Studies Depression Scale
ERQ-R: Emotion Regulation Questionnaire – Reappraisal
MTurk: Amazon's Mechanical Turk Service
PTQ: Perseverative Thinking Questionnaire
RCT: randomized controlled trial
UEQ: User Experience Questionnaire

Edited by G Eysenbach; submitted 22.12.14; peer-reviewed by M Czerwinski, M Salzer; comments to author 15.01.15; revised version received 04.02.15; accepted 12.02.15; published 30.03.15.

Please cite as:
Morris RR, Schueller SM, Picard RW
Efficacy of a Web-Based, Crowdsourced Peer-To-Peer Cognitive Reappraisal Platform for Depression: Randomized Controlled Trial
J Med Internet Res 2015;17(3):e72
URL: http://www.jmir.org/2015/3/e72/
doi:10.2196/jmir.4167
PMID:25835472

Original Paper

Social Media as a Sensor of Air Quality and Public Response in China

Shiliang Wang[1], BS; Michael J Paul[1], MS Eng; Mark Dredze[2], PhD

[1]Johns Hopkins University, Department of Computer Science, Baltimore, MD, United States
[2]Johns Hopkins University, Human Language Technology Center of Excellence, Baltimore, MD, United States

Corresponding Author:
Michael J Paul, MS Eng
Johns Hopkins University
Department of Computer Science
3400 N Charles St
Baltimore, MD, 21218
United States
Phone: 1 410 516 4237
Fax: 1 410 516 5050
Email: mpaul@cs.jhu.edu

Abstract

Background: Recent studies have demonstrated the utility of social media data sources for a wide range of public health goals, including disease surveillance, mental health trends, and health perceptions and sentiment. Most such research has focused on English-language social media for the task of disease surveillance.

Objective: We investigated the value of Chinese social media for monitoring air quality trends and related public perceptions and response. The goal was to determine if this data is suitable for learning actionable information about pollution levels and public response.

Methods: We mined a collection of 93 million messages from Sina Weibo, China's largest microblogging service. We experimented with different filters to identify messages relevant to air quality, based on keyword matching and topic modeling. We evaluated the reliability of the data filters by comparing message volume per city to air particle pollution rates obtained from the Chinese government for 74 cities. Additionally, we performed a qualitative study of the content of pollution-related messages by coding a sample of 170 messages for relevance to air quality, and whether the message included details such as a reactive behavior or a health concern.

Results: The volume of pollution-related messages is highly correlated with particle pollution levels, with Pearson correlation values up to .718 (n=74, $P<.001$). Our qualitative results found that 67.1% (114/170) of messages were relevant to air quality and of those, 78.9% (90/114) were a firsthand report. Of firsthand reports, 28% (32/90) indicated a reactive behavior and 19% (17/90) expressed a health concern. Additionally, 3 messages of 170 requested that action be taken to improve quality.

Conclusions: We have found quantitatively that message volume in Sina Weibo is indicative of true particle pollution levels, and we have found qualitatively that messages contain rich details including perceptions, behaviors, and self-reported health effects. Social media data can augment existing air pollution surveillance data, especially perception and health-related data that traditionally requires expensive surveys or interviews.

(J Med Internet Res 2015;17(3):e22) doi:10.2196/jmir.3875

KEYWORDS

air pollution; public health surveillance; social media; data mining; text mining; natural language processing

Introduction

Recent studies have demonstrated the utility of social media data sources for a wide range of public health goals. Studies have focused on epidemiological surveillance systems for influenza [1,2] and allergies [3], tracking health behaviors such as smoking [4-6] and exercise [7], identifying mental health trends [8,9], and measuring health perceptions and sentiment [10,11]. These are just some of the many health topics discussed on the social media service Twitter [12], demonstrating the ability of social media to complement traditional public health methods, often providing trends faster than traditional surveillance and insights that are difficult to detect through traditional mechanisms.

However, most work to date has focused on Twitter, emphasizing health topics of major concern in the United States, with little work concerning health issues important in other countries. Only recently has attention been given to studying health in Chinese social media, primarily for the purpose of influenza surveillance [13-16]. Our recent study [17] analyzed the diversity of health content in messages from Sina Weibo (abbreviated as Weibo), a microblogging site popular in China. While many of the health topics were similar to those identified on Twitter (eg, influenza, common cold, exercise, and vision health), some topics were unique to China. Most notably, Chinese social media users often discussed pollution and air quality in China, a major Chinese public health issue [18], which receives less attention on Twitter [19].

Air pollution can have tremendous health consequences, such as increased respiratory and cardiovascular disease [20,21]. Air pollution is a major concern in China, where pollution levels are rising alongside rapid urbanization and industrialization [22,23]. Addressing air pollution requires localized surveillance of pollutant levels. Additionally, it is important to understand public awareness, concern, attitudes, health effects, and behavioral response to air pollution [24]. Researchers have investigated public perceptions of risk regarding pollutants [25], emotional and affective responses to air pollution [26], and behavioral responses to pollution, for example, to understand whether people are taking averting action such as staying indoors [27]. This knowledge is important for guiding public policy efforts to reduce pollution, for informing researchers building accurate models of pollution health effects, and for directing the public on how to best respond and protect themselves. So far, these studies have relied on traditional public health methods, such as surveys, for obtaining necessary data.

In this paper, we investigate whether social media data can be used to identify air quality trends and public response in China. Mining social media offers the potential for these trends to be identified in real time and on a massive scale. We mined Weibo messages for statements about air quality and pollution. We demonstrated two epidemiological uses of these data. First, we compared the volume of air quality messages with fine particle pollution in 74 Chinese cities to evaluate the effectiveness of social media for complementing air quality sensors. Second, we conducted a manual coding analysis of a sample of messages to evaluate the ability for measuring public perception, awareness, and response to pollution, a first step toward quantifying the impact of environmental factors on health.

Methods

Data

We collected 93 million messages from Weibo using Weibo's public API. Starting with a small set of randomly selected seed users, we downloaded the 100 most recent messages from each user, then proceeded recursively to download data for the user's followers. All messages were collected in December 2013 but the messages were written as far back as 2009. Since we obtained the most recent messages for each user, the bulk are from 2013 (68.42%, 63,789,097/93,225,579). To focus on the health aspect of air quality and pollution, we selected a set of 917,708 messages obtained by filtering using a list of 1282 health-related terms from a Chinese medical dictionary [28], as well as terms added manually, such as terms related to air pollution: pollution (污染), lungs (肺部), and smog (烟雾). These data were originally collected as part of a broad study into health topics in Chinese social media [17]. The text was preprocessed by removing punctuation, common "stop words", and infrequent words, and performing Chinese word segmentation (see [17] for details).

Weibo requires that users provide city and province upon registration, which is included in the downloaded data. Additionally, each user account has a verified attribute that designates whether it is an individual user (as well as celebrities), a government account, a company account, the media, or others.

To aid additional work on this topic, we are making publicly available the health keywords used to filter Weibo messages, the statistics computed from Weibo for each city and filter, and a list of the Weibo message IDs used in this study along with the filters they matched (see Multimedia Appendix 1). While we are unable to provide the raw Weibo messages per the terms of service, the Weibo public API can be used to directly download messages given the IDs.

Identifying Air Quality Messages

We experimented with two methods for identifying messages related to air quality or pollution. First, we used a simple keyword-based filter in which we selected messages that contain one of four relevant terms: pollution (污染), air (空气), breathe (呼吸), and cough (咳嗽).

Second, we used Latent Dirichlet Allocation (LDA) [29], a probabilistic topic model, to filter messages that belonged to topics relevant to air quality or pollution. A topic model is a probabilistic model of text data, which has two sets of parameters: each document has a discrete distribution over "topics" and each topic has a discrete distribution over words. When estimating the parameters of this model, the topic-specific word distributions typically give high probability to words that tend to occur together in documents. Each topic can therefore be interpreted as a topically or semantically coherent group of words. These parameters are wholly inferred from a raw text corpus, allowing the model to learn topics specific to data of interest.

The LDA model parameters were estimated after 1000 iterations of Gibbs sampling, using 100 topics on our health Weibo dataset. We found two topics whose high-probability words were potentially relevant to air quality, shown in Figure 1 as word clouds. The words in the figure represent the 25 highest-probability words in each topic. Larger words are more probable. The words have been translated from the original Chinese text. The first topic ("AQ") includes many words related to air quality, while the second topic ("PO") is more generally about pollution. Since these words are derived from a fully automated method, they contain many words readily recognizable as relevant to the topic, whereas a few are not as clear.

We used these two topics to filter Weibo messages by selecting messages where at least one token was assigned to the given topic by the sampler.

We experimented with combining our two filtering mechanisms—keyword-based and topic model-based filters—by taking their intersection, selecting messages that both contain a particular topic and a particular keyword.

Finally, we experimented with filtering out messages that contained URLs, under the assumption that these messages are likely to be sharing news media rather than personal experiences [30].

Figure 1. Two pollution-related topics learned from a probabilistic topic model. The left topic is about air quality, and the right topic is about pollution in general.

Comparison to Air Quality Measurements

We compared the volume of air quality messages with fine particle pollution (PM2.5) measurements for 74 Chinese cities from 2013. We compared to the average daily value across 2013 as well as the maximum of all daily values. The data came from the State Environmental Protection Department, which began air quality monitoring in 2012 for these 74 cities [31,32]. Fine particles are those less than 2.5 micrometers in diameter, detected with automated monitoring systems that run continuously with at least 85% uptime, beginning August 2013. The sensing methods are described in [33] (Chinese only). We focus on fine particle pollution because it poses a greater health risk than coarse particle pollution [34].

For each of the cities, we computed the volume of social media activity as the number of messages from the city after filtering for relevance, divided by the total number of messages from the city across the entire dataset. This normalization technique has been previously used for obtaining rates from Twitter data [2]. We measured the Pearson correlation (n=74) between the Weibo volumes and the city PM2.5 values.

Analysis of Message Content

We coded 170 randomly selected messages. We labeled whether the message discussed air quality or air pollution, and if so whether it described a firsthand experience by the user (rather than a general awareness), and if so whether the user reported a change in behavior (eg, wearing a mask), and whether the user expressed concern for his or her health. If a message discussed air quality, we also labeled whether the user requested that action (eg, by the government or community) be taken to improve air quality.

Of messages expressing a health concern, we noted any specific symptoms or health conditions explicitly identified in the message that were perceived to be a result of poor air quality.

Messages were coded independently by two annotators and disagreements were resolved after discussion with a third annotator. We measured the agreement between the two primary annotators using Cohen’s kappa score.

Message Classification

Finally, we experimented with a supervised machine learning approach for identifying relevant messages, using the 170 coded messages as training data. While the messages were not coded for the purpose of training a model, this is a natural experiment to try because messages were labeled with details about relevance.

We used a cascade approach similar to that of Lamb et al for influenza in social media [30], first classifying messages for relevance to air quality, and then classifying messages indicating a firsthand experience (rather than a more general awareness). The first classifier (relevance) was trained on all 170 messages, while the second classifier (firsthand experience) was trained on the subset of messages labeled as relevant. The two classifiers were constructed as logistic regression models using 1-, 2-, and 3-gram word features.

We applied the classifiers to the full set of messages and, as with the other filters, we measured the correlation between the volume of messages identified by the classifiers with the government data.

Results

Data Statistics

Of the 917,708 messages that were filtered for all health-related keywords, 405,467 messages came from the 74 cities with PM2.5 data, with an average of 5479 messages per city (median 3079).

Almost all user accounts, 99.31% (432,862/435,873), were considered "individual" users (not government, business, or media). Government accounts were 0.14% (613/435,873) of users, 0.49% (2147/435,873) were companies, and 0.06% (251/435,873) were media accounts. Thus our data represents individual users as opposed to organizations or governments.

In total, regardless of location, the four keywords matched 75,912 messages, the AQ topic matched 15,763 messages, and the PO topic matched 45,172. For the air quality comparison, we filtered these messages based on the 74 available cities, while the analysis of message content drew from the total dataset.

Comparison to Air Quality Measurements

Table 1 shows the correlations between the volume of filtered messages in each of the 74 cities and the PM2.5 values. None of the differences between correlations when using the maximum daily value (MDV) versus the average daily value (ADV) are statistically significant, but the highest correlations are with ADV. Figure 2 shows a scatter plot of these values for our best filter.

Of the individual keyword filters, "air" has the highest correlations, while of the topic model filters, the AQ topic correlates best. Additionally, we discovered that the correlations can be improved further by combining the best topic model (AQ) with the best keywords ("air" and "pollution"). Combining the AQ topic with "pollution" yields the highest correlation.

If we exclude messages that contain URLs, the correlations improve in all cases except with the PO topic filter.

The highest correlation achieved is with the AQ+"pollution" filter on messages without URLs, at .703 (*P*<.001).

Table 1. Correlation of messages matching each filter in 74 cities to the average (ADV) and maximum (MDV) daily PM2.5 values in 2013.

Filter	Including URLs			Without URLs		
	Number of messages	Corr. (ADV)[a]	Corr. (MDV)[b]	Number of messages	Corr. (ADV)	Corr. (MDV)
AQ[c] topic	7665	.546	.545	5866	.583	.565
PO[d] topic	21,902	.361	.421	17,696	.286	.387
"air"	6321	.552	.593	4949	.610	.637
"pollution"	15,809	.458	.474	12,044	.606	.633
"breathe"	4807	.351	.257	4454	.361	.290
"cough"	12,437	−.005	−.151	11,921	.027	−.023
AQ+"air"	4133	.564	.557	3103	.623	.579
AQ+"pollution"	4866	.630	.619	3766	.703	.657

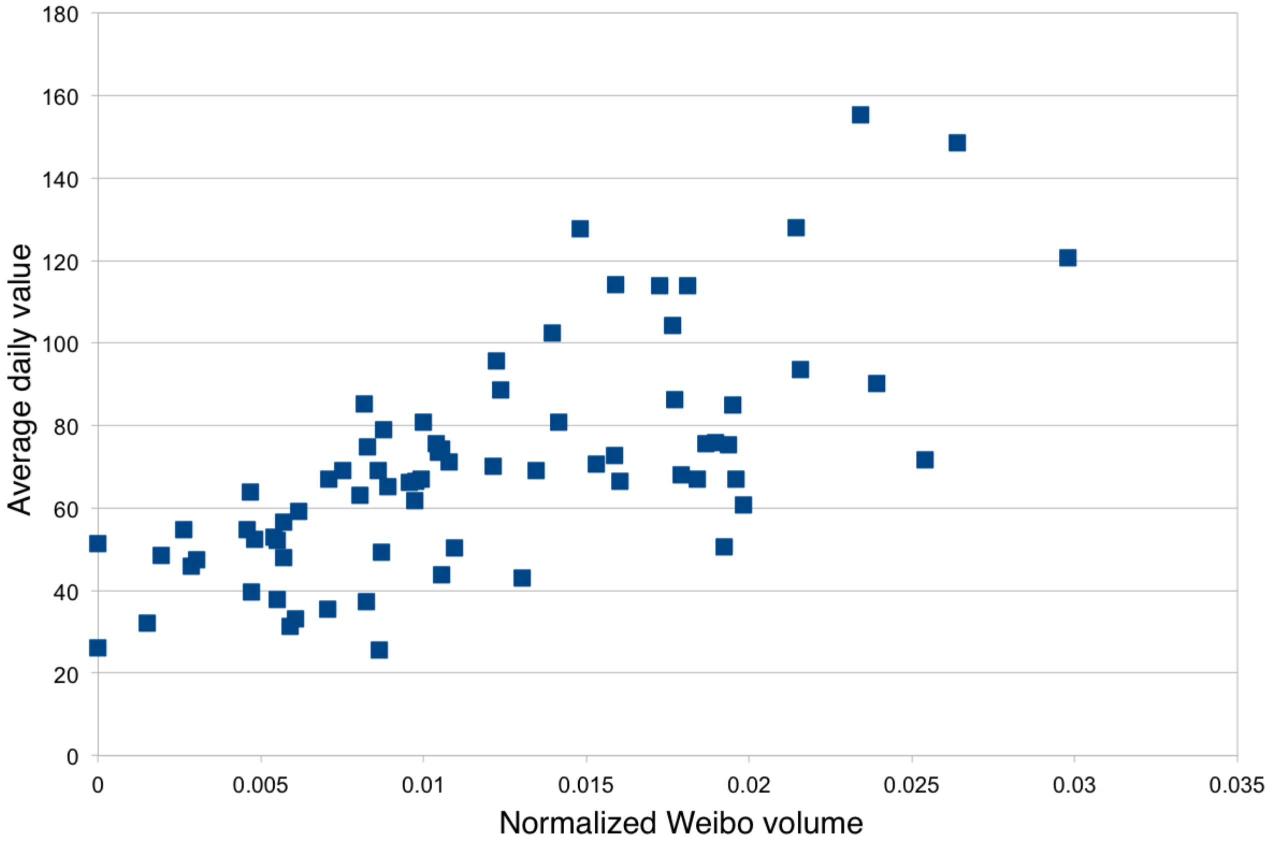
[a] Corr. (ADV): Correlation, average daily value

[b] Corr. (MDV): Correlation, maximum daily value

[c] AQ: air quality

[d] PO: pollution

Figure 2. Scatter plot showing average daily PM2.5 values (y-axis) and the Weibo rate for 74 cities using our most correlated filter, AQ+”pollution” (r=.703).



Analysis of Message Content

We analyzed 170 messages from the AQ topic, which had the highest correlation of the two topics. We did not filter for specific keywords so that we could get a broader set of messages. To target messages that were more strongly relevant to the topic, we selected messages such that the message’s topic distribution assigned more than a 0.1 probability of the document being about this topic, which yields messages with at least two tokens of this topic on average.

These results are summarized in Figure 3. We found that 114 (67.1%) messages sampled through this filter were actually relevant to air quality or air pollution. Of those 114 messages, 90 indicated a firsthand experience (79.0%). Of those 90 messages, 32 (36%) mentioned a reactive behavior, and 17 (19%) expressed a concern for the user’s health.

Three (2.6%) out of 114 relevant messages requested that action by taken to improve the air quality. One message declared a need to reduce carbon emissions, while the other two more generally called for cleaner air (one was directed at the government).

The most common reactive behavior was wearing a face mask, while other behaviors include washing clothes and staying indoors.

Of the 17 messages expressing a health concern, five reported a cough, three reported a sore throat, and two reported dry or peeling skin. Various health conditions were also reported: rhinitis (four messages), allergic rhinitis (one), pharyngitis (one), and asthma (one).

A common pattern that we noticed in firsthand messages that did not belong to the more specific categories (reactive behavior or health concern) is the expression of emotions such as anger or sadness; however, we chose not to quantify this characteristic because it is difficult to define concretely.

Table 2 shows annotator agreement scores from the initial annotations, before disagreements were resolved. Annotator agreement percentages ranged from 78% to 97%. There was very high agreement on whether messages were relevant to air quality, whether the user requested action to improve quality, and whether the user expressed a reactive behavior.

There was less high agreement about whether messages were a firsthand experience, which was sometimes ambiguous and difficult to determine. The lowest agreement was on whether the user expressed a health concern. Annotator divergence primarily stemmed from disagreement over whether a general discomfort should be classified as a health concern. For example, many users expressed discomfort breathing and thus wore a mask. After discussion, we did not count such messages as health concerns, unless health concerns were explicitly stated. Table 3 shows examples of messages that illustrate the various annotations.

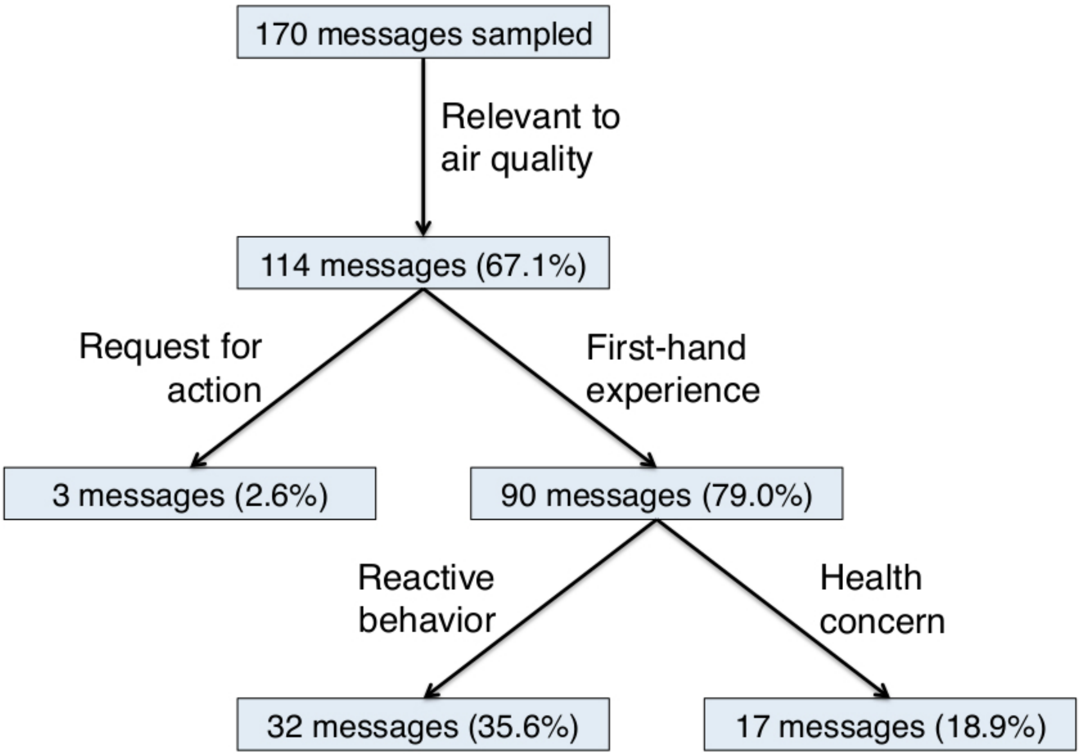
Table 2. Percentage of annotated messages matching the criteria, along with annotator agreement statistics for each question.

Code	Agreement, n (%)	Agreement (kappa)
Relevant to air quality, n=170	160 (94.1)	.869
Request for action, n=107	104 (97.2)	.557
Firsthand experience, n=107	87 (81.3)	.363
Reactive behavior, n=78	73 (93.6)	.864
Health concern, n=78	61 (78.2)	.429

Table 3. Examples of messages with various labels (the original Chinese Weibo is shown, followed by an English translation).

Label	Message
Not about pollution	累昏厥了。牢笼一般的机场巴士,传说中根本不叫花钱的物价,空气里的尿骚味以及灰蒙蒙的天。无论哪顿饭除了咖喱还是咖喱。 I was tired and fainting. The high price, the urine-scented air, and the heavy, gray day made the airport bus feel like a cage. Plus, every meal on the airport bus was curry.
About pollution, not a firsthand experience	老外说: 这幅画表达的是污染程度的北京。PM爆表。 A foreigner said that this picture shows the serious pollution of Beijing. The PM value is too high.
Request for action	不能在空气质量重度污染时才想起低碳行动! Don't wait until the air has already been heavily polluted to start reducing carbon.
Firsthand, reactive behavior	今晚想出去跑步,一查空气指数,还是轻度污染,在家避毒吧。 I want to go running this evening. However, it is lightly polluted based on the air pollution index, so I have to stay at home.
Firsthand, health concern (+ reactive behavior)	三天前开始咳嗽。一定是北京污染的天气有关, 以后出门戴口罩[生病]。 I start coughing three days ago. It must be caused by the pollution in Beijing! I will wear a mask when I go outside [sick].

Figure 3. Summary of annotation results on sample of 170 messages. Tree structure indicates which codes are dependent on their parent codes. Different branches are not mutually exclusive.



Message Classification

We evaluated the classifiers with 10-fold cross validation. The first classifier, trained on 170 messages with 114 positive for relevance, achieved a cross-validation accuracy of .794 (precision .794, recall .947). The second classifier, trained on 114 messages with 90 positive for firsthand experience, achieved a cross-validation accuracy of .718 (precision .689, recall .867).

Because these classifiers were trained on messages that were already filtered by the AQ topic, we then applied the classifiers to the subset of messages containing the AQ topic, similar to the AQ+"pollution" filter. Using this filter, the correlations with pollution data are .718 (ADV) and .664 (MDV). These are both higher than the best correlation reported in Table 1, though not by significant margins.

Discussion

Principal Findings

It is encouraging that even simple content filtering produced moderately high correlations with existing surveillance data. This suggests that lightweight methods can be used for social media-based air quality monitoring. We also showed that better text modeling, through topic modeling and supervised classification, can further improve the correlations. It is perhaps surprising that the supervised classifier did not greatly improve the correlation over simply combining the "pollution" keyword filter with the AQ topic. This may be because the training set of 170 messages was small. In a recent study concurrent with our own, Mei et al found machine learning to help identify air pollution trends in social media [35].

Some of the keyword filters did quite poorly. While "cough" and "breathe" are related to air quality, they are related to much more popular topics as well, yielding filters with low specificity. This demonstrates the benefit of basic natural language processing via topic models. Topic models, which make probabilistic inferences about the topic composition of a message, led to improved correlations when combined with the keyword filters. These models make use of the entire context of a message, which can provide a better relevance model than individual words or phrases. Topic models can also introduce noise, since the models are unsupervised, which we believe is why combining the topic model with a highly relevant keyword like "pollution" correlates better than either filter alone.

Another point to consider is that our filters identify whether a message is *about* air quality, but not *what* the quality is. A promising research direction is to infer a scalar value of air quality based on message content. Natural language processing techniques used for sentiment analysis—the task of quantifying the degree to which text expresses a positive or negative sentiment [36]—could perhaps apply here. For example, certain words like "terrible" or "worst" might indicate worse air quality than simply "bad".

Additionally, we hypothesize that there is a potentially much larger number of messages that could be mined. Extrapolating from a 67% relevance rate of 170 out of 15,763 messages, we estimate there are at least 10,000 messages about air quality in our crawled dataset. Additionally, our entire dataset contains only 93 million messages, a much smaller dataset than those typically used in Twitter research today; for example, Paul and Dredze [3] used a general collection of 2 billion messages to study health topics on Twitter. Since Weibo has more registered users than Twitter, we expect data collection targeted at obtaining air quality messages would obtain a much larger collection.

Our coding results suggest a promising direction in using Weibo messages to understand health concerns, behavioral responses, and health impacts of environmental factors. We found users reporting on all three. While previous work suggests that users will report on well-being during an air pollution crisis [37], we are the first to show that Chinese users make relevant statements on social media services. By building systems that automatically identify these three issues and aggregate them over many users, we could greatly expand traditional surveillance capabilities and inform health interventions.

We believe social media-derived information will be especially advantageous for measuring public perception and response. This is information that cannot be captured with physical sensors, and instead relies traditionally on surveys, panels, and interviews. However, measuring the perceived level of pollution can be just as important as the objective level, as the perceived level is a stronger predictor of willingness to reduce pollution [26]. In this sense, social media reports are more akin to citizen complaints than physical sensors. Although not objective, citizen complaints can complement physical surveillance, and complaints often result in follow-up investigations by regulators in China [38]. Formal complaints likely have different characteristics than the informal complaints found in social media. Social media complaints tend to be general, and the threshold of perceived pollution before writing a complaint is quite likely lower in social media. As social media becomes a more common outlet for citizen complaints, the relationship between these systems of complaint will be important to understand [39].

As with perception, behavioral response to pollution is hard to measure, and often is not measured at all, but rather inferred or assumed [27]. There is therefore a clear knowledge gap that social media data can help fill. A type of behavioral response of particular interest is response to public awareness campaigns or health advisories regarding pollution [40]. Because awareness campaigns and advisories may take place over a short period of time—in some cases, just a single day—it can be difficult to measure their outcome. This has motivated researchers to use Web-derived data to measure the effectiveness of such campaigns, for example using Web search activity to understand World Tobacco Day [41] and Breast Cancer Awareness Month [42]. Having shown that social media users report their perceptions and behaviors regarding air quality, we believe that this data could similarly be used to understand the effectiveness of pollution advisories. This falls into a broader trend of using digital data to support research in behavioral medicine [43].

Finally, we found that many users report perceived health effects of pollution, including specific conditions such as asthma and symptoms such as cough. Previous research has shown that self-reports of health status can be combined with reports of air pollution exposure to understand the associated health effects [44,45]. Social media data, including our Weibo collection, offer passive self-reporting at a much larger scale than what can be collected through traditional, active methods, such as interviews. Such data can augment our understanding of environmental health effects, especially because social media reports include people who experience symptoms but do not seek care, and thus fall outside of what is captured in medical records. Furthermore, research on the health effects of pollution often focuses on more serious outcomes such as disease, while social media reports contain evidence of milder but still important effects, such as discomfort and irritability [24]. Indeed, reports of discomfort were so common in our dataset that we decided to exclude them from consideration as a "health concern" in our coding analysis, yet there is clearly potential for the data to help quantify these effects.

An important limitation to consider is the effect of government censorship on using social media for informatics in China. Studies have shown that collective action and mentions of certain politically sensitive topics are subject to censorship [46-48], but it is not clear whether this would affect pollution-related media. More research is required on this topic, but it is clear from our analysis that many messages describing experiences with air quality can be found in this data source.

Conclusions

To conclude, our findings show that social media messages in China contain a variety of relevant firsthand user reports of air quality, and the volume of these messages correlates with air particle pollution levels in 74 Chinese cities. This was a proof of concept study. Our goals were to understand the content of air quality-related messages, through a qualitative coding of a sample of messages, and to validate the relevance of the messages, by correlating the social media data with existing surveillance data. Our results validate both the quality of these messages and suggest that mining their content can deliver important epidemiological insights into environment health.

Acknowledgments

We thank Brian Schwartz of the Johns Hopkins University Department of Environmental Health Sciences for helpful comments on the draft. We thank Angie Chen for annotation assistance. Michael Paul is supported by a PhD fellowship from Microsoft Research.

Conflicts of Interest

Michael Paul serves on the advisory board for Sickweather, a website that mines health trends from social media data. Mark Dredze reports receipt of compensation for travel for talks at various academic, corporate, and governmental entities and consulting for Directing Medicine, Progeny Systems, and Sickweather.

Multimedia Appendix 1

Health keywords used to filter Weibo messages, statistics computed from Weibo for each city and filter, and list of Weibo message ids used along with the filters they matched.

[ZIP File (Zip Archive), 8MB - jmir_v17i3e22_app1.zip]

References

1. Culotta A. Towards detecting influenza epidemics by analyzing Twitter messages. 2010 Presented at: KDD Workshop on Social Media Analytics; 2010; Washington, DC.
2. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. PLoS ONE 2013;8(12):e83672. [doi: 10.1371/journal.pone.0083672]
3. Paul MJ, Dredze M. You are what you tweet: analyzing Twitter for public health. 2011 Presented at: Int Conf Weblogs Soc Media; 2011; Barcelona.
4. Prier KW, Smith MS, Giraud-Carrier C, Hanson CL. An exploration of tobacco-related tweets as a test topic. 2011 Presented at: 4th Int Conf Soc Comput Behav Model Predict; 2011; College Park, MD p. 18-25.
5. Myslín M, Zhu SH, Chapman W, Conway M. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res 2013;15(8):e174 [FREE Full text] [doi: 10.2196/jmir.2534] [Medline: 23989137]
6. Huang J, Kornfield R, Szczypka G, Emery SL. A cross-sectional examination of marketing of electronic cigarettes on Twitter. Tob Control 2014 Jul;23 Suppl 3:iii26-iii30 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051551] [Medline: 24935894]
7. Zhang N, Campo S, Janz KF, Eckler P, Yang J, Snetselaar LG, et al. Electronic word of mouth on Twitter about physical activity in the United States: exploratory infodemiology study. J Med Internet Res 2013;15(11):e261 [FREE Full text] [doi: 10.2196/jmir.2870] [Medline: 24257325]
8. De Choudhury M. Role of social media in tackling challenges in mental health. 2013 Presented at: 2nd International Workshop on Socially-Aware Multimedia; 2013; Barcelona.
9. Coppersmith G, Harman C, Dredze M. Measuring post traumatic stress disorder in Twitter. 2014 Presented at: International Conference on Weblogs and Social Media; 2014; Ann Arbor, MI.

10. Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. PLoS Comput Biol 2011 Oct;7(10):e1002199 [FREE Full text] [doi: 10.1371/journal.pcbi.1002199] [Medline: 22022249]
11. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. PLoS One 2011;6(5):e19467 [FREE Full text] [doi: 10.1371/journal.pone.0019467] [Medline: 21573238]
12. Paul MJ, Dredze M. Discovering health topics in social media using topic models. PLoS One 2014;9(8):e103408 [FREE Full text] [doi: 10.1371/journal.pone.0103408] [Medline: 25084530]
13. Fung IC, Fu KW, Ying Y, Schaible B, Hao Y, Chan CH, et al. Chinese social media reaction to the MERS-CoV and avian influenza A(H7N9) outbreaks. Infect Dis Poverty 2013;2(1):31 [FREE Full text] [doi: 10.1186/2049-9957-2-31] [Medline: 24359669]
14. Sun X, Ye J, Ren F. Real time early-stage influenza detection with emotion factors from Sina Microblog. 2014 Presented at: 5th Work South Southeast Asian NLP; 2014; Dublin, Ireland.
15. Zhang F, Luo J, Li C, Wang X, Zhao Z. Detecting and analyzing influenza epidemics with social media in China. In: Advances in Knowledge Discovery and Data Mining. Switzerland: Springer International Publishing; 2014:90.
16. Salathé M, Freifeld CC, Mekaru SR, Tomasulo AF, Brownstein JS. Influenza A (H7N9) and the importance of digital epidemiology. N Engl J Med 2013 Aug 1;369(5):401-404. [doi: 10.1056/NEJMp1307752] [Medline: 23822655]
17. Wang S, Paul MJ, Dredze M. Exploring health topics in Chinese social media: An analysis of Sina Weibo. 2014 Presented at: AAAI Work World Wide Web Public Heal Intell; 2014; Quebec City, Quebec.
18. Gong P, Liang S, Carlton EJ, Jiang Q, Wu J, Wang L, et al. Urbanisation and health in China. Lancet 2012 Mar 3;379(9818):843-852 [FREE Full text] [doi: 10.1016/S0140-6736(11)61878-3] [Medline: 22386037]
19. Riga M, Karatzas K. Investigating the relationship between social media content and real-time observations for urban air quality and public health. 2014 Presented at: Proc 4th Int Conf Web Intell Min Semant; 2014; Thessaloniki, Greece.
20. Brunekreef B, Holgate ST. Air pollution and health. Lancet 2002 Oct 19;360(9341):1233-1242. [doi: 10.1016/S0140-6736(02)11274-8] [Medline: 12401268]
21. Dockery DW, Pope CA. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 1994;15:107-132. [doi: 10.1146/annurev.pu.15.050194.000543] [Medline: 8054077]
22. He K, Huo H, Zhang Q. Urban air pollution in China: current status, characteristics, and progress. Annu Rev Energy Environ 2002 Nov;27(1):397-431. [doi: 10.1146/annurev.energy.27.122001.083421]
23. Chan CK, Yao X. Air pollution in mega cities in China. Atmospheric Environment 2008 Jan;42(1):1-42. [doi: 10.1016/j.atmosenv.2007.09.003]
24. Evans GW, Jacobs SV. Air pollution and human behavior. J Soc Issues 1981;37(1):95-125. [doi: 10.1111/j.1540-4560.1981.tb01059]
25. Zhang J. Environmental hazards in the Chinese public's eyes. Risk Anal 1994 Apr;14(2):163-167. [Medline: 8008925]
26. Zeidner M, Shechter M. Psychological responses to air pollution: some personality and demographic correlates. Journal of Environmental Psychology 1988 Sep;8(3):191-208. [doi: 10.1016/S0272-4944(88)80009-4]
27. Bresnahan BW, Dickie M, Gerking S. Averting behavior and urban air pollution. Land Econ 1997;73(3):340-357.
28. Sogou Thesaurus. URL: http://wubi.sogou.com/dict/cell.php?id=272 [accessed 2014-09-15] [WebCite Cache ID 6ScEJE1MG]
29. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003;3:993-1022.
30. Lamb A, Paul MJ, Dredze M. Separating fact from fear: tracking flu infections on Twitter. 2013 Presented at: Proc NAACL-HLT; 2013; Atlanta, Georgia p. 789-795.
31. China National Environmental Monitoring Center. URL: http://www.cnemc.cn/publish/totalWebSite/0666/newList_1.html [accessed 2014-09-16] [WebCite Cache ID 6SdP0C68Y]
32. Greenpeace Press Release. URL: http://www.greenpeace.org/china/zh/news/releases/climate-energy/2014/01/PM25-ranking/ [accessed 2014-09-16] [WebCite Cache ID 6SdOsmmAf]
33. Specifications and test procedures for ambient air quality continuous automated monitoring system for PM10 and PM2.5. URL: http://kjs.mep.gov.cn/hjbhbz/bzwb/dqhjbh/jcgfffbz/201308/t20130802_256852.htm [accessed 2014-11-29] [WebCite Cache ID 6USElviqD]
34. United States Environmental Protection Agency. Fine particle designations. URL: http://www.epa.gov/pmdesignations/faq.htm [accessed 2014-11-29] [WebCite Cache ID 6UheRhMIn]
35. Mei S, Li H, Fan J, Zhu X, Dyer CR. Inferring air pollution by sniffing social media. 2014 Presented at: IEEE / ACM Int Conf Adv Soc Netw Anal Min; 2014; Beijing.
36. Pang B, Lee L. Opinion mining and sentiment analysis. Found Trends Inf Retr 2008;2(1--2):1-135 [FREE Full text]
37. Zhang MW, Ho C, Fang P, Lu Y, Ho RC. Usage of social media and smartphone application in assessment of physical and psychological well-being of individuals in times of a major air pollution crisis. JMIR Mhealth Uhealth 2014;2(1):e16 [FREE Full text] [doi: 10.2196/mhealth.2827] [Medline: 25098255]
38. Dasgupta S, Wheeler D. Citizen complaints as environmental indicators: evidence from China. Washington, DC: The World Bank; 1997.

39. Frias-Martinez V, Sae-Tang A, Frias-Martinez E. To call, or to tweet? Understanding 3-1-1 citizen complaint behaviors. 2014 Presented at: Sixth Int Conf on Social Computing; 2014; Stanford, CA.
40. Semenza JC, Wilson DJ, Parra J, Bontempo BD, Hart M, Sailor DJ, et al. Public perception and behavior change in relationship to hot weather and air pollution. Environ Res 2008 Jul;107(3):401-411. [doi: 10.1016/j.envres.2008.03.005] [Medline: 18466894]
41. Ayers JW, Althouse BM, Allem JP, Ford DE, Ribisl KM, Cohen JE. A novel evaluation of World No Tobacco day in Latin America. J Med Internet Res 2012;14(3):e77 [FREE Full text] [doi: 10.2196/jmir.2148] [Medline: 22634568]
42. Glynn RW, Kelly JC, Coffey N, Sweeney KJ, Kerin MJ. The effect of breast cancer awareness month on internet search activity--a comparison with awareness campaigns for lung and prostate cancer. BMC Cancer 2011;11:442 [FREE Full text] [doi: 10.1186/1471-2407-11-442] [Medline: 21993136]
43. Ayers JW, Althouse BM, Dredze M. Could behavioral medicine lead the web data revolution? JAMA 2014 Apr 9;311(14):1399-1400. [doi: 10.1001/jama.2014.1505] [Medline: 24577162]
44. Duhme H, Weiland SK, Keil U, Kraemer B, Schmid M, Stender M, et al. The association between self-reported symptoms of asthma and allergic rhinitis and self-reported traffic density on street of residence in adolescents. Epidemiology 1996 Nov;7(6):578-582. [Medline: 8899382]
45. Peabody JW, Riddell TJ, Smith KR, Liu Y, Zhao Y, Gong J, et al. Indoor air pollution in rural China: cooking fuels, stoves, and health status. Arch Environ Occup Health 2005;60(2):86-95. [doi: 10.3200/AEOH.60.2.86-95] [Medline: 16983861]
46. King G, Pan J, Roberts M. How censorship in China allows government criticism but silences collective expression. 2012 Presented at: APSA 2012 Annu Meet Pap; 2012; Washington, DC.
47. Bamman D, O'Connor B, Smith N. Censorship and deletion practices in Chinese social media. First Monday 2012 Mar 02;17(3):e0 [FREE Full text] [doi: 10.5210/fm.v17i3.3943]
48. King G, Pan J, Roberts ME. Reverse-engineering censorship in China: randomized experimentation and participant observation. Science 2014;345(6199):981. [doi: 10.1126/science.1251722]

Abbreviations

ADV: average daily value
AQ: air quality
LDA: Latent Dirichlet Allocation
MDV: maximum daily value
PO: pollution

Edited by G Eysenbach; submitted 18.09.14; peer-reviewed by A Yates, JP Allem; comments to author 12.11.14; revised version received 11.12.14; accepted 12.12.14; published 26.03.15.

Please cite as:
Wang S, Paul MJ, Dredze M
Social Media as a Sensor of Air Quality and Public Response in China
J Med Internet Res 2015;17(3):e22
URL: http://www.jmir.org/2015/3/e22/
doi:10.2196/jmir.3875
PMID:25831020

Original Paper

Information Presentation Features and Comprehensibility of Hospital Report Cards: Design Analysis and Online Survey Among Users

Uwe Sander[1*], MD; Martin Emmert[2*], MSc, PhD; Jochen Dickel[3*], Dipl-Des; Nina Meszmer[2*], BA; Benjamin Kolb[1*], MSc

[1]Department of Information and Communication, Faculty for Media, Information and Design, University of Applied Sciences and Arts Hannover, Hannover, Germany

[2]Institute of Management (IFM), School of Business and Economics, Friedrich-Alexander-University Erlangen-Nuremberg, Nuremberg, Germany

[3]Faculty of Media, University of Applied Sciences FHM, Bielefeld, Germany

[*]all authors contributed equally

Corresponding Author:
Martin Emmert, MSc, PhD
Institute of Management (IFM)
School of Business and Economics
Friedrich-Alexander-University Erlangen-Nuremberg
Lange Gasse 20
Nuremberg, 90403
Germany
Phone: 49 911 5302 ext 253
Fax: 49 911 5302 114
Email: Martin.Emmert@fau.de

Abstract

Background: Improving the transparency of information about the quality of health care providers is one way to improve health care quality. It is assumed that Internet information steers patients toward better-performing health care providers and will motivate providers to improve quality. However, the effect of public reporting on hospital quality is still small. One of the reasons is that users find it difficult to understand the formats in which information is presented.

Objective: We analyzed the presentation of risk-adjusted mortality rate (RAMR) for coronary angiography in the 10 most commonly used German public report cards to analyze the impact of information presentation features on their comprehensibility. We wanted to determine which information presentation features were utilized, were preferred by users, led to better comprehension, and had similar effects to those reported in evidence-based recommendations described in the literature.

Methods: The study consisted of 5 steps: (1) identification of best-practice evidence about the presentation of information on hospital report cards; (2) selection of a single risk-adjusted quality indicator; (3) selection of a sample of designs adopted by German public report cards; (4) identification of the information presentation elements used in public reporting initiatives in Germany; and (5) an online panel completed an online questionnaire that was conducted to determine if respondents were able to identify the hospital with the lowest RAMR and if respondents' hospital choices were associated with particular information design elements.

Results: Evidence-based recommendations were made relating to the following information presentation features relevant to report cards: evaluative table with symbols, tables without symbols, bar charts, bar charts without symbols, bar charts with symbols, symbols, evaluative word labels, highlighting, order of providers, high values to indicate good performance, explicit statements of whether high or low values indicate good performance, and incomplete data ("N/A" as a value). When investigating the RAMR in a sample of 10 hospitals' report cards, 7 of these information presentation features were identified. Of these, 5 information presentation features improved comprehensibility in a manner reported previously in literature.

Conclusions: To our knowledge, this is the first study to systematically analyze the most commonly used public reporting card designs used in Germany. Best-practice evidence identified in international literature was in agreement with 5 findings about German report card designs: (1) avoid tables without symbols, (2) include bar charts with symbols, (3) state explicitly whether high or low values indicate good performance or provide a "good quality" range, (4) avoid incomplete data (N/A given as a value), and (5) rank hospitals by performance. However, these findings are preliminary and should be subject of further evaluation. The implementation of 4 of these recommendations should not present insurmountable obstacles. However, ranking hospitals by performance may present substantial difficulties.

(J Med Internet Res 2015;17(3):e68) doi:10.2196/jmir.3414

KEYWORDS

public reporting; report cards; information presentation

Introduction

Background

In recent years, many health care systems have implemented strategies to improve the quality of care [1]; nevertheless, quality deficits and variability still remain [1-6]. In general, patients are unlikely to be aware of the existence of quality differences [7,8]. One reason for this is the limited amount of publicly available information on the quality of health care providers [9]. Therefore, improving the transparency of information about health care provider quality is a major challenge [10]. It is assumed that this will improve overall quality by steering patients toward better-performing health care providers [11,12] and by incentivizing providers to improve quality [8,9,13-15].

Public reporting in Germany is partly regulated by law. Since 2005, German hospitals have had to publish quality reports online to help patients and physicians make informed choices of hospitals. The AQUA Institute has been commissioned to make further improvements in quality assurance [16]. In 2011, 1666 hospitals had to participate in an “external quality assurance” process. Since the beginning of 2012, the quality reports of individual hospitals include 182 of 430 quality measures [16]. A possible quality shortfall at a hospital can trigger an evaluation, including a structured quality dialog, which allows a group of experts to conduct a qualitative investigation of discrepant results at individual hospitals. In 2010, a total of 21,053 discrepant results were identified in 4,064,320 datasets. Of these, 8.0% were evaluated as qualitatively discrepant through the structured quality dialog. The AQUA Institute does not report on individual hospitals, but German public reporting portals draw on the information provided in the quality reports of individual hospitals.

Nevertheless, there are several barriers to effective public reporting in Germany. Friedemann et al [17] analyzed quality reports of individual German hospitals and concluded that they were neither readable nor understandable for most patients. In international studies, 1 of the barriers most frequently discussed is that consumers do not understand the formats in which the information is presented. Hibbard [18] noted that we need to find more effective ways to present data to consumers. Similarly, Sinaiko et al [19] argued that the current report cards need to be substantially expanded and refined. Kullgren and Werner [20] added that the problem with current public reporting is partly caused by the limitations imposed by the design of report cards.

Although the number of public reporting websites is likely to continue to rise [21], many argue that in their current state they might confuse consumers. Rothberg and colleagues [22] even argued that it would be better to report nothing at all rather than misleading information. Similarly, Emmert et al [23] stated that patients or physicians should not yet use such information to choose an individual physician. In the rush to make providers accountable, enthusiasm has often outstripped science [22]. Several researchers have pointed to the tremendous diversity in the presentation of quality data [21]. The websites and related reports vary widely in terms of ease of access, ease of use, usefulness of information, timeliness of updates, and credibility [24].

Research Questions

After reviewing the available international recommendations, we analyzed the 10 most commonly used German public report cards and addressed 3 questions:

1. What information presentation elements were utilized?
2. Which led to better comprehension?
3. Which had similar effects to those reported in the evidence-based recommendations described in literature?

We focused on elements of information design that are used to communicate ideas, illustrate information, or express relationships visually, such as pictures, symbols, and colors [25].

Methods

Overview

The study consisted of 5 steps: (1) identification of best-practice evidence about the presentation of information on hospital report cards, (2) selection of a single risk-adjusted quality indicator for the study, (3) selection of a sample of the public report card designs used by German hospitals, (4) identification of information presentation elements used in public reporting initiatives in Germany, and (5) conduct of an online-based survey.

Identification of Best-Practice Evidence About the Presentation of Information on Hospital Report Cards

A literature search was conducted in April 2013 by searching the Medline (via PubMed) and Cochrane Library databases using the “Abstract/Title/Keywords” search field: Search (“Public Report$ OR Publicly Report$ OR Publicly Release$ OR Public Disclos$ OR Information Disseminat$ OR Report Card$ OR Consumer Report$ OR Quality Report$ OR Comparative Report$ OR Reporting Instrument$”); limitations: English, German, and Spanish; published in the last 10 years; field: “Title/Abstract/Keywords.” Only peer-reviewed journal articles and 3 types of review were included: (1) studies which compared health care report cards in terms of specific criteria,

(2) studies which theoretically or empirically assessed or discussed the distinguishing features of public report cards and provided evidence on how performance data should be published or presented, and (3) studies which summarized or discussed best practice in public reporting for health care.

Selection of a Single Risk-Adjusted Quality Indicator for the Study

Several risk-adjusted outcome quality measures were available from the German Hospital Quality Report 2011 [17]. The selection of the quality indicator for the study was done by assessing case numbers, the number of hospitals using the measure, and its role in quality assurance. We selected an elective procedure because publicly available information on elective treatments might be expected to help patients identify a good health care provider. We selected a risk-adjusted mortality rate (RAMR) measure because these are considered useful indicators of hospital quality [26-28].

Selection of a Sample of the Public Report Card Designs Used by German Hospitals

To identify relevant hospital websites, we used a Google search by searching on several keywords (eg, *Kliniksuche* [clinic search], *Krankenhaussuche* [hospital search], *Gute Klinik* [good clinic], *Klinikvergleich* [compare clinic]). The most frequently visited hospital rating websites were identified using the Alexa analysis tool.

Identification of Information Presentation Elements Used in Public Reporting Initiatives in Germany

The presentations of the selected quality indicators in 10 report cards were captured with screenshots that included interactive features (mouseovers). Information presentation elements were categorized with nVivo 10 by 2 authors using the previously identified literature-derived categories. Additional categories were added when no literature-derived categories were available.

Conduct of an Online-Based Survey

We applied an online-based cross-sectional study by surveying an online panel to address the following questions:

1. Were respondents able to identify the hospital with the lowest RAMR?
2. Were respondents' hospital choices associated with particular information design elements?

Consultation took place with an online panel in Germany in August-September 2013; each participant received €1 per finished survey. The panelists were recruited through several recruitment channels including online recruitment, direct mailing, and offline recruitment. The panel members were invited by email to participate (the invitation contained a link to the online survey). The survey was administered and conducted by Norstat Germany Ltd (formerly ODC Services Ltd), a fieldwork agency for survey research.

The online questionnaire consisted of 3 parts. First, a short introduction described the background and study objectives. Second, a random selection of 3 of the 10 hospital report cards was presented to prevent training effects (eg, overfamiliarization with presentation techniques leading to biased results). Respondents were asked to select the best quality hospital, to justify their choice in response to an open-ended question, and to assess the comprehensibility of the website (range of 1=very comprehensible to 7=very incomprehensible). Third, respondents provided sociodemographic data. The questionnaire was piloted and descriptive analyses were conducted using SPSS version 21.0 (IBM Corp, Armonk, NY, USA). The statistical significance of differences between responses was calculated using chi-square tests and *t* tests. Analysis of the open-ended questions was conducted by 2 authors coding and categorizing answers independently using nVivo 10; discrepancies were discussed to achieve a consensus.

Results

Identification of Best-Practice Evidence About the Presentation of Information on Hospital Report Cards

Our search yielded 2506 hits in the Cochrane Library and 1827 in PubMed. After elimination of duplicates, 4302 articles were screened by title and abstract, resulting in exclusion of 4018 articles. In addition, 7 studies were identified through reference search, expert consultation, and Internet searches, giving 291 articles for full-text review. Of these, 13 articles published between 2001 and 2013 met the inclusion criteria. Ten articles described observations in the United States, 2 in the Netherlands, and 1 in Germany.

Table 1 shows the literature-derived categories for the presentation of information on health care report cards. Because this study focused on outcome quality measures, only the information presented on these measures is described here. The methodological quality of the literature used is described in Multimedia Appendix 1.

Table 1. Features of the presentation of information on health care report cards from previous studies.

Category	Recommendations and results
Evaluative table with symbols	Consider using a table design such as the “evaluative table with stars” rather than a bar chart [29]
	Evaluative tables using words or stars are superior to numerical tables [29]
	Physicians preferred formats that used traffic light symbols to code the value of indicators (numerical table with traffic lights) [30]
Tables without symbols	Graphic displays were more helpful to users than text-only tables [31]
Bar charts	Bar charts were commonly used (43% of public reporting websites) [32]
Bar charts without symbols	Comprehension was lowest when data were presented in bar charts [30]
	Standard bar charts were not well-liked by respondents and led to the lowest levels of comprehension [29]
Bar charts with symbols	Symbols and bar charts should be used [31]
	A combination of bar charts and star ratings facilitated correct interpretation by users [32]
	Adding stars to bar charts increases comprehension significantly [33]
Symbols	Participants liked to use symbols to identify the best surgeon [31]
	Physicians preferred formats that used symbols (eg, traffic lights) [30]
	Star-only formats should be used in preference to numerical values [34]
	Only important information should be made easier to evaluate using symbols [35]
Evaluative word labels	Adding evaluative labels to bar charts did not increase comprehension [33]
Highlighting	Color-coding important information improves comprehension [36]
	Highlighting information about quality resulted in greater understanding [37]
	Presentation formats which highlighted key messages increased comprehension [38]
Order of providers	Physicians prefer presentation formats that combine individual indicator values with evaluative features such as rankings [30]
	Comprehension of respondents who were low in numeracy was significantly improved by the ordered compared to the unordered condition [35]
	Providers should be ranked by performance [12]
	Ranking plans by performance significantly decreased errors in interpreting data [33]
	Ranking by performance increased the frequency with which users chose higher-performing services [15]
	Providers should be ranked in descending order of quality, as this was valued by participants and increased their comprehension [36]
	One of the more powerful display strategies is to rank providers in terms of performance [33]
	When providers were ordered alphabetically participants were more likely to make effective use of the data (ie, choose the best provider) than when providers were ordered by performance [32]
High values indicate good performance	Performance data should be displayed such that high values always represent high performance [35]
	Numeric tables and bar charts often led respondents to conclude that the worst performing nursing homes (those with the higher percentages) were the best, notwithstanding the warning label at the top [29]
State explicitly whether high or low values indicate good performance	It should be stated explicitly whether high or low values indicate good performance, regardless of the direction of the scale [8,36]
Incomplete data (“N/A” as a value)	Incomplete data (missing values) have a negative influence on provider assessment and the potential to influence a decision [29]

Selection of a Single Risk-Adjusted Quality Indicator for the Study

For our investigation, we selected the risk-adjusted quality indicator coronary angiography and percutaneous coronary intervention (PCI). This procedure was performed 715,469 times in 841 German hospitals in 2011. In 2011, 6369 of 276,866 (2.30%) patients died after PCI; 2.24% mortality had been expected, resulting in a RAMR of 1.03 [17].

Selection of a Sample of the Public Report Card Designs Used by German Hospitals

A total of 63 hospital public reporting websites were identified by a Google search. Several report cards were eliminated because they did not present outcome quality measures or used presentation formats identical to those of sites already included in the sample. Of the remaining report cards, the 10 most frequently visited were used as a sample (Multimedia Appendix 2): Portal A [39], B [40], C [41], D [42], E [43], F [44], G [45], H [46], I [47], and K [48].

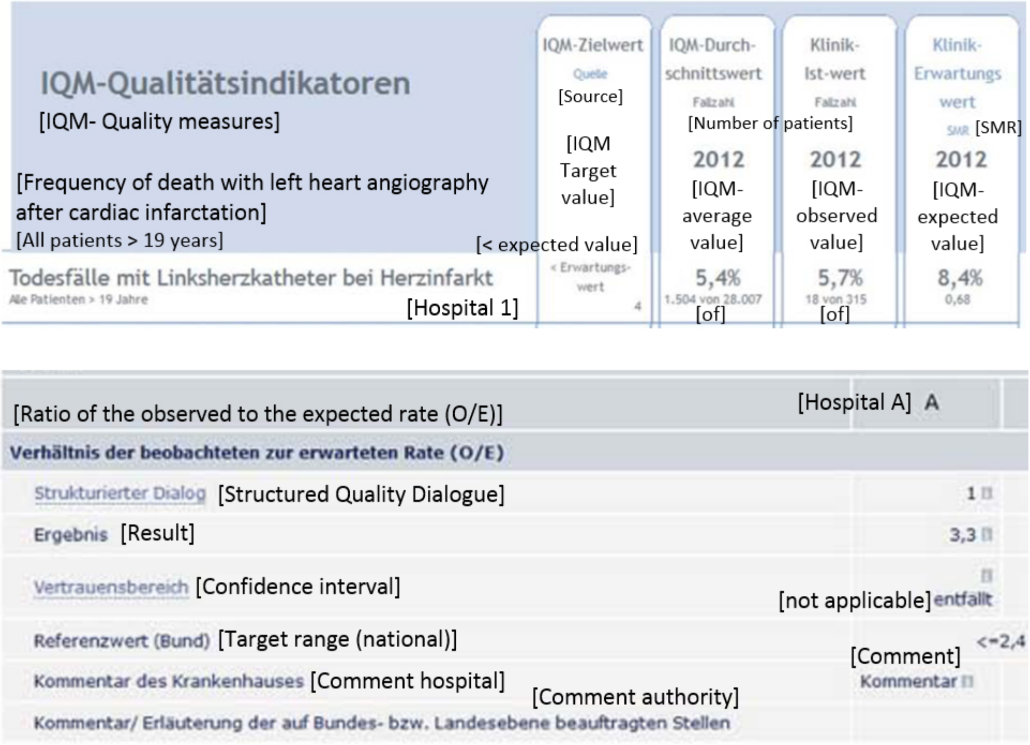
Identification of Information Presentation Elements Used in Public Reporting Initiatives in Germany

The formats used to present information about RAMR in coronary catheterization by the 10 public reporting websites (see Multimedia Appendix 2) that we studied are summarized in Table 2. Tables (5 sites) and bar charts (5 sites) were equally popular; 4 sites presented reports with incomplete or missing data ("N/A" as a value). Symbols such as traffic lights were commonly used (7 sites), sometimes in combination with bar charts (4 sites) or tables (5 sites) (see Figure 1). Five report cards used low values (for the mortality rate) to indicate good performance and we identified 2 report cards which indicated a "good quality" range for the RAMR.

Table 2. Features used in the presentation of risk-adjusted mortality rates for coronary catheterization by 10 German portals.

Elements of information presentation	n	Portals
Table with symbols	3	B, I, K
Table without symbols	2	C, E,
Bar chart without symbols	1	G
Bar chart with symbols	4	A, D, F, H
Bar chart with traffic light symbols	3	A, D, H
Bar chart with thumb symbols	1	F
Symbols only	0	—
Evaluative word labels	0	—
Highlighting	0	—
Providers ranked by performance	2	D, H
High values indicating good performance	0	—
Explicit statement about whether high or low values indicate good performance	5	A, D, G, H, I
No statement about scale direction, but a "good quality" range identified	2	A, H
Incomplete data ("N/A" as a value)	4	B, C, F, K

Figure 1. Tables without symbols. Top: Portal E; bottom: Portal G. Only results for hospital 1 are displayed. English translations in brackets.



IQM-Qualitätsindikatoren [IQM- Quality measures]	IQM-Zielwert Quelle [Source] [IQM Target value]	IQM-Durchschnittswert Fallzahl [Number of patients] 2012 [IQM-average value]	Klinik-Ist-wert Fallzahl 2012 [IQM-observed value]	Klinik-Erwartungswert SMR [SMR] 2012 [IQM-expected value]
Todesfälle mit Linksherzkatheter bei Herzinfarkt [Frequency of death with left heart angiography after cardiac infarctation] Alle Patienten > 19 Jahre [All patients > 19 years] [Hospital 1]	< Erwartungswert [< expected value] 4	5,4% 1.504 von [of] 28.007	5,7% 18 von [of] 315	8,4% 0,68

[Ratio of the observed to the expected rate (O/E)]	[Hospital A] A
Verhältnis der beobachteten zur erwarteten Rate (O/E)	
Strukturierter Dialog [Structured Quality Dialogue]	1
Ergebnis [Result]	3,3
Vertrauensbereich [Confidence interval]	[not applicable] entfällt
Referenzwert (Bund) [Target range (national)]	<=2,4
Kommentar des Krankenhauses [Comment hospital]	[Comment] Kommentar
Kommentar/ Erläuterung der auf Bundes- bzw. Landesebene beauftragten Stellen [Comment authority]	

Conduct of an Online-Based Survey

In total, 3064 respondents started the online survey and 2027 completed it (completion rate=66.16%), taking a mean 14.51 (SD 9.39) minutes. The overall mean age of respondents was 41.57 (SD 15.87) years; 978 of 2027 respondents were female (48.24%) and 50.71% (1028/2027) graduated from high school or obtained a technical university entrance qualification (see Table 3). A total of 96.69% (1960/2027) used the Internet at least once a day. Table 3 also shows the results of the Arbeitsgemeinschaft Online Forschung (AGOF) Internet Facts 2014-07 survey of the German population who used the Internet in the past 3 months [49]. Comparing our survey results to those of the AGOF, the strongest difference appears to be in the demographics of the 2 surveys, specifically in the rate of persons without school qualifications or with secondary general school. In our survey, this rate was 11.69% (237/2027) compared to 35.2% in the AGOF survey. Differences in age, gender, and household size were weaker.

Table 3. Overview of the study sample in comparison with Internet users in Germany.[a]

Demographics	Study sample (N=2027)	Internet users in Germany (N=106,677)
Age (years)		
Mean (SD)	41.57 (15.87)	—
Range, n (%)		
≤20	255 (12.58%)	(13.6%)
21-30	334 (16.47%)	(17.1%)
31-40	355 (17.51%)	(16.3%)
41-50	484 (23.87%)	(20.6%)
51-60	328 (16.18%)	(16.8%)
≥61	270 (13.32%)	(15.5%)
Gender (female), n (%)	978 (48.24%)	(47.5%)
Household size, n (%)		
1 person	456 (22.50%)	(16.8%)
2 persons	725 (35.76)	(33.7%)
3 persons and more	845 (41.69%)	(49.5%)
Education, n (%)		
Still at school	67 (3.31%)	(4.7%)
Without school qualification or secondary general school	237 (11.69%)	(35.2%)
Intermediate secondary school or equivalent qualification	694 (34.24%)	(30.6%)
High school graduation/technical university entrance qualification	1028 (50.72%)	(34.2%)

[a] As measured by the Arbeitsgemeinschaft Online Forschung (AGOF) Internet Facts 2014-07 survey of the German population who used the Internet in the last past 3 months [49].

Were Respondents Able to Identify the Hospital With the Lowest Risk-Adjusted Mortality Rate?

Table 4 shows the results of the survey. Three of the 10 report cards were presented to each respondent (N=2027) for a total of 6081 observations. For each site, a mean 60.68% (1230/2027) of respondents successfully identified the hospital with the lowest RAMR, and 6.81% (138/2027) (range 0.8%-14.0%) selected the hospital with the highest RAMR. In approximately 14.60% (296/2027) of all observations, respondents stated they were unable to provide an answer. Only 32.02% (649/2027) of respondents selected the hospital with the lowest RAMR on all 3 randomly assigned report cards, whereas 14.01% (284/2027) did not identify the hospital with the lowest RAMR on any report card.

Table 4. Respondents were asked to select the best quality hospital: overview of the selected hospitals (N=6081 observations).

Portal used	Which hospitals did the respondents select?, n (%)					
	Hospital 1	Hospital 2	Hospital 3	Hospital 4	Hospital 5	Could not answer
Portal A	123 (2.02)	4695 (77.21)[a]	282 (4.64)	208 (3.42)	195 (3.21)[b]	581 (9.55)
Portal B	224 (3.68)[b]	2238 (36.8)	480 (7.89)	144 (2.37)	1909 (31.39)[a]	1088 (17.89)
Portal C	766 (12.60)[b]	341 (5.61)	1669 (27.45)[a]	274 (4.51)	857 (14.09)	2174 (35.75)
Portal D	4674 (76.86)[a]	207 (3.40)	165 (2.71)	499 (8.21)	145 (2.38)[b]	391 (6.43)
Portal E	238 (3.91)	255 (4.19)	622 (10.23)[b]	166 (2.73)	3956 (65.06)[a]	842 (13.85)
Portal F	65 (1.07)	864 (14.21)	84 (1.38)[b]	36 (0.59)	4441 (73.03)[a]	585 (9.62)
Portal G	3925 (64.55)[a]	541 (8.90)[b]	239 (3.93)	109 (1.79)	157 (2.58)	1113 (18.3)
Portal H	4129 (67.90)[a]	347 (5.71)	681 (11.20)	297 (4.88)	49 (0.81)[b]	572 (9.41)
Portal I	584 (9.60)	663 (10.9)[b]	3537 (58.16)[a]	209 (3.44)	222 (3.65)	870 (14.31)
Portal K	851 (13.99)[b]	91 (1.5)	4020 (66.11)[a]	321 (5.28)	122 (2.01)	666 (10.95)

[a] Hospital with the lowest RAMR.

[b] Hospital with the highest RAMR.

Were Respondents' Hospital Choices Associated With Particular Information Design Elements?

Overview

To explore whether respondents' hospital choices were influenced by the design, we asked the following questions:

1. Did respondents who used websites that included a given feature choose the hospital with the lowest RAMR significantly more or less often than respondents using portals not including these features?
2. How did respondents who used the portals including this feature rate the overall comprehensibility of the site?
3. Based on their answers to the open-ended question about reasons for their decision, did respondents regard this design feature as useful (appreciative comments) or confusing (critical comments)?

Table Without Symbols

Two of 10 portals presented tables without symbols (Table 3). Respondents using these portals chose the hospital with the lowest RAMR significantly less often than respondents using other portals did (46.18%, 575/1245 vs 64.50%, 3119/4836, *P*<.001). They also rated the comprehensibility of the presentation significantly lower (mean 3.07, SD 1.85 vs mean 3.77, SD 1.92; Table 5). This corresponded with negative comments about tables given in response to the open-ended question. Most responses to the open-ended question that mentioned tables without symbols were disapproving:

> *I cannot decide as the table is not understandable even after reading the explanations.*
>
> *I think this table is very confusing; it is difficult to interpret the various measures.*
>
> *This presentation is much too incomprehensible; therefore, I cannot choose 1 of the hospitals.*

Table 5. Choice of the hospital with the lowest risk-adjusted mortality rate (RAMR).

Information presentation feature	Feature included			Feature not included			Choice of lowest RAMR		Comprehensibility	
	Respondents, n	Selected hospital with the lowest RAMR, n (%)	Comprehensibility,[a] mean (SD)	Respondents, n	Selected hospital with the lowest RAMR, n (%)	Comprehensibility, [a] mean (SD)	χ^2 (*df*)	*P*	*t* (*df*)	*P*
Table without symbols	1245	575 (46.18)	3.07 (1.85)	4836	3119 (64.50)	3.77 (1.92)	139.2 (1)	<.001	-11.657 (6079)	<.001
Table with symbols	1787	928 (51.93)	3.58 (1.84)	4294	2766 (64.42)	3.65 (1.96)	82.5 (1)	<.001	-1.300 (6979)	.19
Bar chart without symbols	608	392 (64.47)	2.99 (1.84)	5473	3302 (60.33)	3.70 (1.92)	3.9 (1)	.047	-8.626 (6079)	<.001
Bar chart with symbols	2441	1799 (73.70)	4.11 (1.92)	3640	1895 (52.06)	3.31 (1.86)	286.9 (1)	<.001	16.289 (6079)	<.001
Bar chart with traffic light symbols	1814	1341 (73.93)	4.25 (1.93)	4267	2353 (55.14)	3.36 (1.86)	188.3 (1)	<.001	16.774 (6079)	<.001
Bar chart with thumb symbols	627	458 (73.05)	3.70 (1.83)	5454	3236 (59.33)	3.62 (1.94)	44.4 (1)	<.001	1.030 (6079)	.30
Providers ranked by performance	1221	883 (72.32)	4.29 (1.91)	4860	2811 (57.84)	3.46 (1.89)	85.8 (1)	<.001	13.620 (6079)	<.001
Explicit statement about whether higher or lower values indicate better performance	3017	2079 (68.91)	3.85 (1.97)	3064	1615 (52.71)	3.41 (1.86)	167.3 (1)	<.001	9.112 (6079)	<.001
No statement about scale direction, but range for good quality presented	1220	884 (72.46)	4.04 (1.91)	4861	2810 (57.81)	3.52 (1.92)	87.8 (1)	<.001	8.440 (6079)	<.001
*Incomplete data (N/A labels)	2445	1212 (49.57)	3.32 (1.88)	3636	2483 (68.30)	3.84 (1.93)	214.2 (1)	<.001	-10.436 (6079)	<.001

[a] Based on a 7-point Likert scale with a range of 1=not at all comprehensible to 7=very comprehensible.

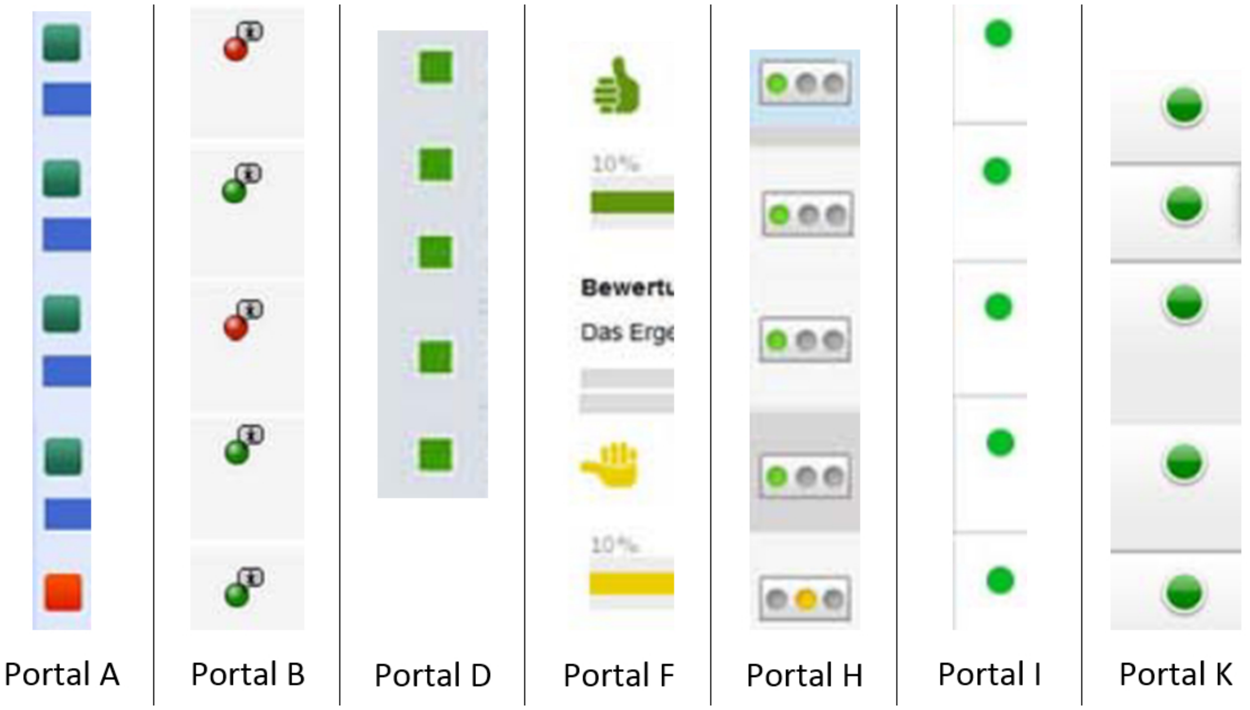
Table With Symbols

Three of 10 portals presented tables with evaluative symbols (Table 2). Respondents using these portals chose the hospital with the lowest RAMR significantly less often than respondents using other portals (51.93%, 928/1787 vs 64.42%, 2766/4294; *P*<.001) and rated the comprehensibility of the presentation lower (without meeting statistical significance) (Table 5). This corresponded to mainly negative comments about tables and positive ones about symbols given in response to the open-ended question (Table 6). In combination with tables or bar charts, 6 of 10 portals (Table 5) used green, yellow, or red symbols; 1 (Portal F) presented thumbs (thumbs up or thumbs down) symbols in these colors (see Figure 2). Red (3/51) and yellow (4/51) symbols were displayed less often than green ones (44/51, 86%). Three portals presented green symbols for all 5 displayed hospitals.

Table 6. Responses to the open-ended question about information presentation.

Information presentation feature	Responses, n
Table without symbols	39 (38 incomprehensible)
Table with symbols	Helpful green symbol: n=59; table: n=36 (35 incomprehensible)
Bar chart without symbols	25 (23 incomprehensible)
Bar chart with symbols	Symbol helpful: n=79; bar chart (incomprehensible): n=39; bar chart (helpful): n=24
Bar chart with traffic light symbols	Bar chart: n=50 (30 incomprehensible, 20 helpful); symbol: n=30 (28 helpful)
Bar chart with thumb symbols	Helpful thumb: n=51; bar chart (incomprehensible): n=9; bar chart (helpful): n=4
Providers ranked by performance	Ranking (helpful): n=3
Explicit statement about whether higher or lower values indicate better performance	Higher values as a reason for hospital choice: n=67
No statement about scale direction, but range for good quality presented	—
Incomplete data (N/A labels)	Complaints about incomplete or missing data: n=59

Figure 2. Symbols used by 7 of the 10 portals.



Bar Chart Without Symbols

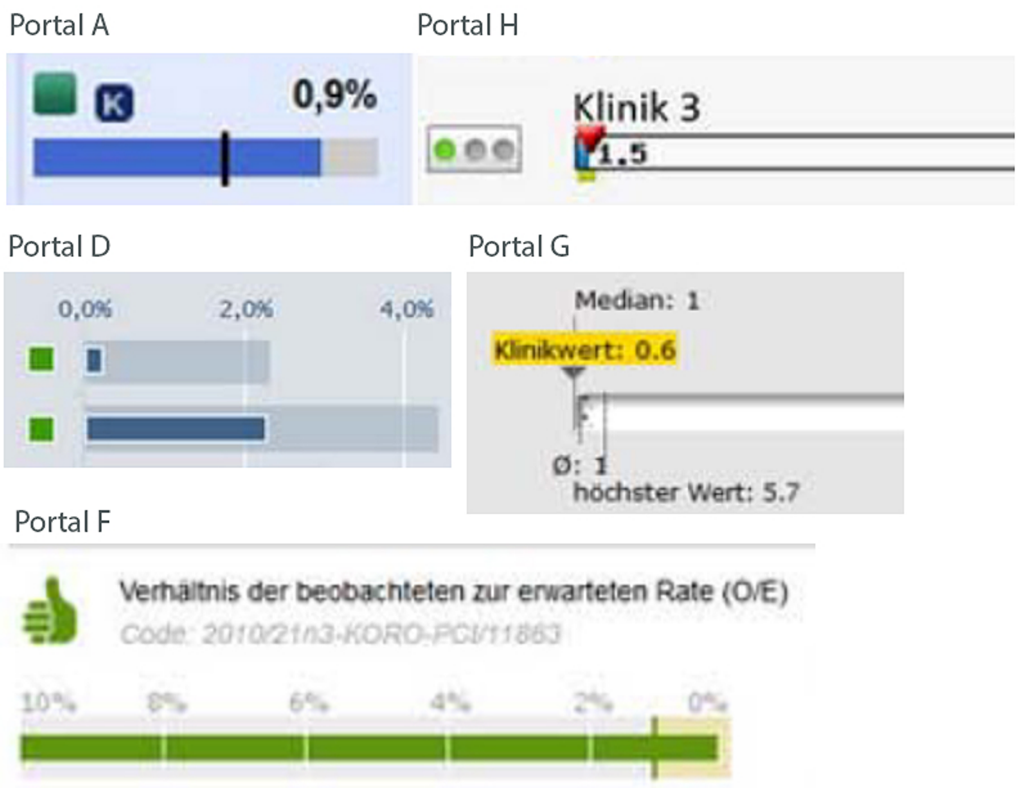
One report card used bar charts without symbols (Table 3 and Figure 3). Respondents using this portal chose the hospital with the lowest RAMR significantly more often than respondents using other portals (64.5%, 392/608 vs 60.3%, 3302/ 5473; *P*=.047), but rated the comprehensibility of the presentation lower (mean 2.99, SD 1.84 vs mean 3.70, SD 1.92; *P*<.001) (Table 5). This corresponded to mostly negative comments about bar charts given in response to the open-ended question (Table 6).

The bars displayed on Portals G and H were too narrow for evaluation. Portal G presented 4 numbers (benchmarking information, median, mean, highest mortality rate for the hospital) in the bar chart, which may have led to information overload. All bar charts in the German portals used longer bars to indicate lower quality, which led to some confusion.

Most responses to the open-ended question that mentioned bar charts without symbols were disapproving:

> *Diagram is unclear and incomprehensible.*
>
> *According to this graphic, I would choose clinic 1, even without really knowing which one is better. The graphics are not very useful.*
>
> *It has the lowest value, but I find the charts very difficult to understand because I cannot work out what they mean.*

Figure 3. Bar chart presentation taken from 5 portals.



Bar Chart With Symbols

Four of 10 portals used bar charts with symbols (Table 2 and Figure 2). Respondents using these portals chose the hospital with the lowest RAMR significantly more often than respondents using other portals (73.70%, 1799/2441 vs 52.06%, 1895/3640; *P*<.001) and rated the comprehensibility of the presentation higher (mean 4.11, SD 1.92 vs mean 3.31, SD 1.86; *P*<.001) (Table 5). Comments about symbols were mostly positive, but comments about bar charts were mixed (Table 6). Respondents testing Portal A (bar chart with symbols) referred to the bar chart with a blue bar for their decision (see Figure 2), which attracted 10 approving comments, such as:

> *I didn't rely on percentages, but the longest blue bar in the diagram.*
>
> *Longest bar and lowest result.*

But, more comments (n=19) were disapproving, such as:

> *What is the meaning of the bar—is a lot of blue good?*
>
> *0.9% is the smallest number, even if the blue bar is the longest.*
>
> *What does this blue bar mean?*

Bar Chart With Traffic Light Symbols

Three of 10 portals used bar charts with traffic light symbols (Table 3 and Figure 2). Respondents using these portals chose the hospital with the lowest RAMR significantly more often than respondents using other portals (73.70%, 1341/1814 vs 52.06%, 2353/4267; *P*<.001); they also rated the comprehensibility of the presentation higher (mean 4.25, SD 1.93 vs mean 3.36, SD 1.86; *P*<.001) (Table 5). Comments about symbols were mostly positive, but comments about bar charts were mixed (Table 6).

Bar Chart With Thumb Symbols

One of 10 portals used bar charts with thumb symbols (Table 3 and Figure 2). Respondents using this portal chose the hospital with the lowest RAMR significantly more often than respondents using other portals (73.05%, 458/627 vs 59.33%, 3236/5454, *P*<.001), but assigned a similar rating to the comprehensibility of the presentation (mean 3.70, SD 1.83 vs mean 3.62, SD 1.94; *P*=.30) (Table 5). Comments about symbols were positive, but the few comments about bar charts were mixed (Table 6).

Ranking Providers by Performance

Two of the 10 portals ranked providers by performance (Table 2 and Figure 2). Respondents using these portals chose the hospital with the lowest RAMR significantly more often than respondents using other portals (72.32%, 883/1221 vs 57.84%, 2811/4860; *P*<.001) and rated the comprehensibility of the presentation higher (mean 4.29, SD 1.91 vs mean 3.46, SD 1.89; *P*<.001) (Table 5).

Explicit Statement That Low Values Indicate Good Performance

Five of 10 portals explicitly stated that lower values indicated better performance (Table 3). Respondents using these portals chose the hospital with the lowest RAMR significantly more often than respondents using other portals (68.91%, 2079/3017 vs 52.71%, 1615/3064; *P*<.001) and rated the comprehensibility of the presentation higher (mean 3.85, SD 1.97 vs mean 3.41, SD 1.86; *P*<.001) (Table 5).

No Explicit Statement About Scale Direction, But "Good Quality" Range Identified

Two of 10 portals explicitly gave a "good quality" range (Table 3 and Figure 2). Respondents using these portals chose the hospital with the lowest RAMR significantly more often than respondents using other portals (72.46%, 884/1220 vs 57.81%, 2810/4861; *P*<.001) and rated the comprehensibility of the presentation higher (mean 4.04, SD 1.91 vs mean 3.52, SD 1.92; *P*<.001) (Table 5).

Incomplete Data (N/A as a Value)

Four of the 10 portals presented incomplete data (N/A as a value) (Table 3) for 1 or more measures (confidence intervals, frequency of cases treated, frequency of mortality, mortality rates, comments about hospital or quality controlling agency). Respondents using these portals chose the hospital with the lowest RAMR significantly less often than respondents using other portals (49.57%, 1212/2445 vs 68.30%, 2483/3636; *P*<.001) and rated the comprehensibility of the presentation lower (mean 3.32, SD 1.87 vs mean 3.84, SD 1.93; *P*<.001) (Table 5). A total of 59 responses to the open-ended question complained about incomplete or missing data (Table 6). Comments of respondents about missing values on Portal B (hospitals 4 and 5 had missing data in a number of cases; hospital 5 had the lowest RAMR) included:

> *As I do not know how many cases hospital 4 or 5 have, I decided against them.*
>
> *Hospitals 4 and 5 seem to be suspect as they do not show the frequency of cases as the basis for the observed to expected rate.*

Because the 2 remaining best-practice elements, evaluative word labels and highlighting best providers, were not included in any of the report cards we studied, we were not able to not investigate their effect on hospital choice.

Discussion

Overview

The 3 research questions answered are as follows:

1. What information presentation elements were utilized?
2. Which led to better comprehension?
3. Which information presentation elements had similar effects to those reported in the evidence-based recommendations described in the literature?

What Information Presentation Elements Were Utilized?

We identified 10 elements that were used in the presentation of RAMR for coronary catheterization by 10 German portals (see Table 2).

Which Led to Better Comprehension?

Report cards using the following 7 presentation elements were more comprehensible, because respondents choose the hospital with the lowest RAMR significantly more often: (1) bar chart without symbols, (2) bar chart with symbols, (3) bar chart with traffic light symbols, (4) bar chart with thumb symbols, (5) providers ranked by performance, (6) an explicit statement about whether higher or lower values indicated better performance, and (7) no statement about scale direction, but a presented range for good quality.

Furthermore, respondents rated presentations as more comprehensible when they contained the following 4 elements: (1) bar chart with symbols, (2) bar chart with traffic light symbols, (3) providers ranked by performance, (4) explicit statement about whether higher or lower values indicated better performance, and (5) no statement about scale direction, but range for good quality presented.

Report cards using the following 2 presentation elements were less comprehensible, because respondents using these elements choose the hospital with the lowest RAMR significantly less often. This was true for tables without symbols or for incomplete data (N/A labels). Moreover, respondents rated presentations as less comprehensible when they contained tables without symbols or incomplete data (N/A labels).

Which Information Presentation Elements Had Similar Effects to Those Reported in the Evidence-Based Recommendations Described in the Literature?

Based on the 13 identified international reports about the presentation of information on report cards, 14 information presentation elements were identified. These studies relied on a variety of methods, including in-depth or cognitive interviews and focus groups, plus online, telephone, or paper-based surveys as well as controlled laboratory experiments (see Multimedia Appendix 1). This variety of methods made it possible to obtain insights from different perspectives about the behavior of users, but also made it more difficult to systematically compare the results. Additionally, samples used in the studies varied widely; sizes ranged from 59 to 2052. Sample selections also had limitations, such as the use of convenience samples, limited geographic locations, low response rates, or a possible selection bias because of questioning by mail and Internet. These limitations have to be taken into account when conclusions are drawn from those studies.

Did the Respondents to Our Survey Make Choices and Give Reasons That Corresponded to the Results of the 13 Identified International Reports?

Our respondents, like those in the study by Gerteis et al [29], made more interpretive errors with the numeric formats (table without symbols) than with the graphical formats (bar charts). They had more difficulty understanding numeric tables than other presentation formats, a finding similar to that of Donelan et al [31].

Similar to Geraedts et al [30], we observed that the hospital with the lowest mortality was chosen more often when tables with symbols were used than when tables without symbols were used. In our survey, tables with symbols were rated better for comprehension (mean 3.58, SD 1.84) than bar charts without symbols (mean 2.99, SD 1.84). However, we cannot fully endorse Gerteis et al's [29] recommendation (in a study limited to 90 respondents) that tables with symbols should be used rather than a standard bar chart without symbols. In our survey, bar charts without symbols more frequently resulted in the choice of the hospital with the lowest mortality (64.47%) than tables with symbols (51.93%), but as discussed previously, missing data rather than the presentational format may have been the major reason.

Similar to Damman et al [32], we found that bar charts were used quite often and, like Gerteis et al [29], we found that graphical formats (standard bar chart without symbols) were not liked by respondents. However, we did not confirm the findings of Geraedts et al, who presented the formats of presentations to physicians, that bar charts without symbols did not assist the comprehension of data on hospital quality [30] or were the format least well understood by participants [29]. In our sample, more respondents who used a bar chart without symbols (64.47%) chose the hospital with the lowest mortality rate than respondents who did not (60.33%). As Figure 3 shows, Portal G displayed bars that were too narrow for evaluation, but numbers were also included. Thus, respondents may still have been able to evaluate the presentation by evaluating the numerical value instead of using the bar chart ("It has the lowest value, but I find the charts very difficult to understand").

The respondents in our sample who used portals that displayed bar charts with symbols were more likely than average to choose the hospital with the lowest mortality rate, a result which is in accordance with other studies [31-33]. Two portals ranked hospitals by performance, as recommended in several studies [12,15,30,33,35,36]. This was also valued by our participants.

Our results support the recommendation that whether high or low values indicate good performance should be explicitly stated [8,36]. Incomplete data (N/A given as a value) had a negative influence on provider assessment, as Gerteis et al [29] suggested.

To summarize, 5 of 14 information presentation elements were found to be associated with better comprehension as reported in literature: (1) bar charts with symbols, (2) explicit statement about whether higher or lower values indicated better performance, and (3) providers ranked by performance. A reduction of user comprehension was associated with (4) incomplete data (N/A labels) and (5) tables without symbols, again in accordance with literature-based evidence. Reasons have been given for the 2 discrepant findings on (6) bar charts without symbols (improvement of comprehension observed in our German survey, but not described in international literature) and (7) tables with symbols (improvement of comprehension described in international literature, but not observed in our German survey).

Conclusions

To our knowledge, this is the first study to systematically analyze the most commonly used public report card designs in Germany. It is also the first study to analyze German report cards using real-world applications [15] instead of controlled laboratory studies. The best-practice evidence found in 13 international studies led to 14 findings about information presentation elements. However, due to limitations of these studies, conclusions have to be drawn carefully. Five of these findings were in agreement with our findings about German report card designs: (1) avoid tables without symbols, (2) include bar charts with symbols, (3) state explicitly whether high or low values indicate good performance or provide a "good quality" range, (4) avoid incomplete data (N/A given as a value), and (5) rank hospitals by performance. However, due to limitations of our study as described subsequently, these recommendations are preliminary and should be subject to further evaluation.

The implementation of 4 of these findings should not present an insurmountable obstacle to public reporting instruments because they can be achieved by redesigning the format. However, ranking hospitals by performance may present substantial difficulties because ordering by performance is often resisted by providers [7]. Ordering makes report sponsors responsible for determining what constitutes a meaningful difference in mortality rates. The benchmarking information available in Germany might provide a basis for ranking providers. However, ranking might also require measures of statistical significance to be provided that have the potential to confuse users [8,31,36,50,51].

Externally validated measures of hospital quality could also be used as the basis for ranking hospitals and are readily available in Germany [17]; indeed, most German portals use these results to assign their traffic light symbols. However, because only approximately 0.1% of quality measures are finally validated as being qualitatively discrepant, the ranking of hospitals in these terms is difficult. As a respondent put it: "All the symbols for hospital quality are green."

Limitations

Several limitations are relevant for this study. This report only investigated the design of report cards. Other important features of report cards (eg, the quality of explanations, use of technical terms, and provision of a quality framework) will be the subject of further studies. Respondents in our study only saw a small part of the public reporting portals because we focused on the presentation of 1 RAMR measure. Five report cards (Portals B, E, F, G, I) did not allow a performance comparison of hospitals in the way we presented it in this study, but were adapted to enable respondents to make the comparison we were investigating. We limited our search for recommendations on best practice in the presentation of information in public report cards to the PubMed and Cochrane medical databases and did not take into account other databases, such as PsychINFO.

Because our study was designed as a Web survey, the results presented might be influenced by self-selection of the study participants. Better-educated persons were overrepresented. Because education levels influence comprehension of public report cards, this might have influenced our results; certain features of presented information might have been misunderstood depending on the education of respondents. Hibbard et al [52] stated that higher education levels were related to improved performance. Emmert et al [53] found that those with a higher level of education tended to comprehend public reporting better. Damman et al [32] stated that consumers' educational level was positively related to the correct interpretation. Donelan et al [33] found that respondents with at least some college education were significantly more likely to identify surgeons with the lowest risk-adjusted mortality, compared with respondents having no college education.

Acknowledgments

We would like to thank Laura Polster, Elena Kreß, Petra Nölp, and Mareike von der Ahe for their assistance in coding open-ended questions, and Markus Reichert for his assistance in the literature research. This study was financed by the Dr Theo and Friedl Schoeller Research Center at the Friedrich-Alexander-University of Erlangen-Nuremberg.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Methodological quality of used literature.

[PDF File (Adobe PDF File), 130KB - jmir_v17i3e68_app1.pdf]

Multimedia Appendix 2

Screenshots of report cards.

[PDF File (Adobe PDF File), 1MB - jmir_v17i3e68_app2.pdf]

References

1. Lebrun LA, Shi L, Zhu J, Sharma R, Sripipatana A, Hayashi AS, et al. Racial/ethnic differences in clinical quality performance among health centers. J Ambul Care Manage 2013;36(1):24-34. [doi: 10.1097/JAC.0b013e3182473523] [Medline: 23222010]
2. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003 Jun 26;348(26):2635-2645. [doi: 10.1056/NEJMsa022615] [Medline: 12826639]
3. Asch SM, Kerr EA, Keesey J, Adams JL, Setodji CM, Malik S, et al. Who is at greatest risk for receiving poor-quality health care? N Engl J Med 2006 Mar 16;354(11):1147-1156. [doi: 10.1056/NEJMsa044464] [Medline: 16540615]
4. Laschet H. ÄrzteZeitung 21. The German GBA targets five challenging indications for treatment URL: http://www.aerztezeitung.de/politik_gesellschaft/versorgungsforschung/article/829958/qualitaet-gba-knoepft-fuenf-problem-indikation.html [accessed 2015-03-07] [WebCite Cache ID 6Wr8Mqhxw]
5. Merchant RM, Yang L, Becker LB, Berg RA, Nadkarni V, Nichol G, American Heart Association Get With the Guideline-Resuscitation Investigators. Variability in case-mix adjusted in-hospital cardiac arrest rates. Med Care 2012 Feb;50(2):124-130 [FREE Full text] [doi: 10.1097/MLR.0b013e31822d5d17] [Medline: 22249921]
6. Tsai J, Rosenheck RA. Racial differences among supported housing clients in outcomes and therapeutic relationships. Psychiatr Q 2012 Mar;83(1):103-112. [doi: 10.1007/s11126-011-9187-x] [Medline: 21811835]
7. Hibbard JH, Sofaer S. Best Practices in Public Reporting No. 1: How To Effectively Present Health Care Performance Data To Consumers. Rockville, MD: Agency for Healthcare Research and Quality; Jun 2010.
8. Hibbard JH, Cunningham PJ. How engaged are consumers in their health and health care, and why does it matter? Res Brief 2008 Oct(8):1-9. [Medline: 18946947]
9. Porter ME, Teisberg EO. Redefining competition in health care. Harv Bus Rev 2004 Jun;82(6):64-76, 136. [Medline: 15202288]
10. Lagu T, Lindenauer PK. Putting the public back in public reporting of health care quality. JAMA 2010 Oct 20;304(15):1711-1712. [doi: 10.1001/jama.2010.1499] [Medline: 20959582]
11. Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. Med Care 2009 Jan;47(1):1-8. [doi: 10.1097/MLR.0b013e3181808bb5] [Medline: 19106724]
12. Emmert M, Gemza R, Schöffski O, Sohn S. [Public reporting in health care: the impact of publicly reported quality data on patient steerage]. Gesundheitswesen 2012 Jun;74(6):e25-e41. [doi: 10.1055/s-0031-1285857] [Medline: 21866496]
13. Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. Ann Intern Med 2008 Jan 15;148(2):111-123. [Medline: 18195336]
14. Hibbard JH, Peters E. Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. Annu Rev Public Health 2003;24:413-433. [doi: 10.1146/annurev.publhealth.24.100901.141005] [Medline: 12428034]
15. Werner RM, Asch DA. The unintended consequences of publicly reporting quality information. JAMA 2005 Mar 9;293(10):1239-1244. [doi: 10.1001/jama.293.10.1239] [Medline: 15755946]
16. AQUA-Institut G. German Hospital Quality Report. German Hospital Quality Report 2011;2012:50-53.

17. Friedemann J, Schubert HJ, Schwappach D. [On the comprehensibility of German hospital quality reports: systematic evaluation and need for action]. Gesundheitswesen 2009 Jan;71(1):3-9. [doi: 10.1055/s-0028-1086010] [Medline: 19173143]
18. Grantmakers In Health, Washington, DC, USA. Considering quality: engaging consumers to make better health care decisions. Issue Brief (Grantmakers Health) 2007 Jan(27):i-v, 1. [Medline: 17621688]
19. Sinaiko AD, Eastman D, Rosenthal MB. How report cards on physicians, physician groups, and hospitals can have greater impact on consumer choices. Health Aff (Millwood) 2012 Mar;31(3):602-611 [FREE Full text] [doi: 10.1377/hlthaff.2011.1197] [Medline: 22392672]
20. Kullgren JT, Werner RM. Counterpoint: will public reporting of health-care quality measures inform and educate patients? No. Chest 2011 Nov;140(5):1117-20; discussion 1120. [doi: 10.1378/chest.11-2094] [Medline: 22045876]
21. Damman OC, van den Hengel YKA, van Loon AJM, Rademakers J. An international comparison of web-based reporting about health care quality: content analysis. J Med Internet Res 2010;12(2):e8 [FREE Full text] [doi: 10.2196/jmir.1191] [Medline: 20439252]
22. Rothberg MB, Benjamin EM, Lindenauer PK. Public reporting of hospital quality: recommendations to benefit patients and hospitals. J Hosp Med 2009 Nov;4(9):541-545. [doi: 10.1002/jhm.481] [Medline: 19514092]
23. 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]
24. Amini A, Birnbaum DW, Black B, Hyman DA. Public reporting of hospital infection rates: ranking the states on credibility and user friendliness. Stud Health Technol Inform 2013;183:87-92. [Medline: 23388261]
25. Emmerson J. Visualizing Information for Advocacy: An Introduction to Information Design. Berlin: Tactical Technology Collective; 2008.
26. Consortium of Chief Quality Officers. Using hospital standardized mortality ratios for public reporting: a comment by the consortium of chief quality officers. Am J Med Qual 2009;24(2):164-165. [doi: 10.1177/1062860608326543] [Medline: 19033214]
27. The Lancet. Public reporting of surgical outcomes. Lancet 2011 Apr 2;377(9772):1126. [doi: 10.1016/S0140-6736(11)60446-7] [Medline: 21459198]
28. Dimick JB, Staiger DO, Birkmeyer JD. Ranking hospitals on surgical mortality: the importance of reliability adjustment. Health Serv Res 2010 Dec;45(6 Pt 1):1614-1629 [FREE Full text] [doi: 10.1111/j.1475-6773.2010.01158.x] [Medline: 20722747]
29. Gerteis M, Gerteis JS, Newman D, Koepke C. Testing consumers' comprehension of quality measures using alternative reporting formats. Health Care Financ Rev 2007;28(3):31-45 [FREE Full text] [Medline: 17645154]
30. Geraedts M, Hermeling P, de CW. Communicating quality of care information to physicians: a study of eight presentation formats. Patient Educ Couns 2012 Jun;87(3):375-382. [doi: 10.1016/j.pec.2011.11.005] [Medline: 22177585]
31. Donelan K, Rogers RS, Eisenhauer A, Mort E, Agnihotri AK. Consumer comprehension of surgeon performance data for coronary bypass procedures. Ann Thorac Surg 2011 May;91(5):1400-5; discussion 1405. [doi: 10.1016/j.athoracsur.2011.01.019] [Medline: 21524448]
32. Damman OC, Hendriks M, Rademakers J, Spreeuwenberg P, Delnoij DMJ, Groenewegen PP. Consumers' interpretation and use of comparative information on the quality of health care: the effect of presentation approaches. Health Expect 2012 Jun;15(2):197-211. [doi: 10.1111/j.1369-7625.2011.00671.x] [Medline: 21615637]
33. Hibbard JH, Peters E, Slovic P, Finucane ML, Tusler M. Making health care quality reports easier to use. Jt Comm J Qual Improv 2001 Nov;27(11):591-604. [Medline: 11708039]
34. Hibbard JH, Slovic P, Peters E, Finucane ML. Strategies for reporting health plan performance information to consumers: evidence from controlled studies. Health Serv Res 2002 Apr;37(2):291-313 [FREE Full text] [Medline: 12035995]
35. Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. Med Care Res Rev 2007 Apr;64(2):169-190. [doi: 10.1177/10775587070640020301] [Medline: 17406019]
36. Mazor KM, Dodd KS. A qualitative study of consumers' views on public reporting of health care-associated infections. Am J Med Qual 2009;24(5):412-418. [doi: 10.1177/1062860609335971] [Medline: 19525369]
37. Uhrig JD, Harris-Kojetin L, Bann C, Kuo TM. Do content and format affect older consumers' use of comparative information in a Medicare health plan choice? Results from a controlled experiment. Med Care Res Rev 2006 Dec;63(6):701-718. [doi: 10.1177/1077558706293636] [Medline: 17099122]
38. Harris-Kojetin LD, McCormack LA, Jaël EF, Sangl JA, Garfinkel SA. Creating more effective health plan quality reports for consumers: lessons from a synthesis of qualitative testing. Health Serv Res 2001 Jul;36(3):447-476 [FREE Full text] [Medline: 11482584]
39. krankenhaussuche.berlin.de. URL: http://krankenhaussuche.berlin.de/ [accessed 2015-03-01] [WebCite Cache ID 6WhcS7jqi]
40. BKK. URL: https://www.bkk-klinikfinder.de/ [accessed 2015-03-07] [WebCite Cache ID 6Wr9BsWtz]
41. Deutsches Krankenhaus Verzeichnis. URL: http://www.deutsches-krankenhaus-verzeichnis.de/de/19/Suche.html [accessed 2015-03-07] [WebCite Cache ID 6Wr9F66yb]
42. Hamburger Krankenhausspiegel. URL: http://www.hamburger-krankenhausspiegel.de/ [accessed 2015-03-07] [WebCite Cache ID 6Wr9JYcsx]
43. IQM. URL: http://www.initiative-qualitaetsmedizin.de/ [accessed 2015-03-07] [WebCite Cache ID 6Wr9MnNjV]

44. Krankenhaus.de. URL: http://www.krankenhaus.de/ [accessed 2015-03-07] [WebCite Cache ID 6Wr9OaHkD]
45. Qualitaetskliniken.de. URL: http://www.qualitaetskliniken.de/ [accessed 2015-03-07] [WebCite Cache ID 6Wr9QgIKt]
46. Klinikfuehrer Rheinland. URL: http://www.klinikfuehrer-rheinland.de/ [accessed 2015-03-07] [WebCite Cache ID 6Wr9TWWbB]
47. TK-Klinikfuehrer. URL: https://www.tk.de/tk/klinikfuehrer/114928 [accessed 2015-03-07] [WebCite Cache ID 6Wr9W2WD8]
48. Weisse Liste. URL: http://www.weisse-liste.de/ [accessed 2015-03-07] [WebCite Cache ID 6Wr9Yi7rf]
49. Internet Facts 2014-07. Arbeitsgemeinschaft Online Forschung (AGOF) 2014 URL: http://www.agof.de/aktuelle-studie-internet/ [accessed 2015-03-07] [WebCite Cache ID 6Wr959LST]
50. Dudley A, Hibbard J, Shaller D. Model Public Report Elements: A Sampler. Rockville, MD: Agency for Healthcare Research and Quality; 2010:10-0088.
51. Shahian DM, Edwards FH, Jacobs JP, Prager RL, Normand SLT, Shewan CM, et al. Public reporting of cardiac surgery performance: Part 1--history, rationale, consequences. Ann Thorac Surg 2011 Sep;92(3 Suppl):S2-11. [doi: 10.1016/j.athoracsur.2011.06.100] [Medline: 21867789]
52. Hibbard JH, Slovic P, Peters E, Finucane ML, Tusler M. Is the informed-choice policy approach appropriate for medicare beneficiaries? Health Affairs 2001 May 01;20(3):199-203. [doi: 10.1377/hlthaff.20.3.199]
53. Emmert M, Hessemer S, Meszmer N, Sander U. Do German hospital report cards have the potential to improve the quality of care? Health Policy 2014 Dec;118(3):386-395. [doi: 10.1016/j.healthpol.2014.07.006] [Medline: 25074783]

Abbreviations

AGOF: Arbeitsgemeinschaft Online Forschung
PCI: percutaneous coronary intervention
RAMR: risk-adjusted mortality rate

Edited by G Eysenbach; submitted 21.03.14; peer-reviewed by M Geraedts, C Schaefer; comments to author 21.08.14; revised version received 03.10.14; accepted 22.01.15; published 16.03.15.

Please cite as:
Sander U, Emmert M, Dickel J, Meszmer N, Kolb B
Information Presentation Features and Comprehensibility of Hospital Report Cards: Design Analysis and Online Survey Among Users
J Med Internet Res 2015;17(3):e68
URL: http://www.jmir.org/2015/3/e68/
doi:10.2196/jmir.3414
PMID:25782186

Review

Effects of Home Telemonitoring Interventions on Patients With Chronic Heart Failure: An Overview of Systematic Reviews

Spyros Kitsiou[1], PhD; Guy Paré[2], MSRC, PhD; Mirou Jaana[3], MPH, PhD

[1]College of Applied Health Sciences, Department of Biomedical and Health Information Sciences, University of Illinois at Chicago, Chicago, IL, United States

[2]Research Chair in Information Technology in Health Care, Department of Information Technology, HEC Montréal, Montréal, QC, Canada

[3]Telfer School of Management, University of Ottawa, Ottawa, ON, Canada

Corresponding Author:
Spyros Kitsiou, PhD
College of Applied Health Sciences
Department of Biomedical and Health Information Sciences
University of Illinois at Chicago
1919 West Taylor Street (AHSB)
Chicago, IL, 60612
United States
Phone: 1 3123553519
Fax: 1 3129968342
Email: skitsiou@uic.edu

Abstract

Background: Growing interest on the effects of home telemonitoring on patients with chronic heart failure (HF) has led to a rise in the number of systematic reviews addressing the same or very similar research questions with a concomitant increase in discordant findings. Differences in the scope, methods of analysis, and methodological quality of systematic reviews can cause great confusion and make it difficult for policy makers and clinicians to access and interpret the available evidence and for researchers to know where knowledge gaps in the extant literature exist.

Objective: This overview aims to collect, appraise, and synthesize existing evidence from multiple systematic reviews on the effectiveness of home telemonitoring interventions for patients with chronic heart failure (HF) to inform policy makers, practitioners, and researchers.

Methods: A comprehensive literature search was performed on MEDLINE, EMBASE, CINAHL, and the Cochrane Library to identify all relevant, peer-reviewed systematic reviews published between January 1996 and December 2013. Reviews were searched and screened using explicit keywords and inclusion criteria. Standardized forms were used to extract data and the methodological quality of included reviews was appraised using the AMSTAR (assessing methodological quality of systematic reviews) instrument. Summary of findings tables were constructed for all primary outcomes of interest, and quality of evidence was graded by outcome using the GRADE (Grades of Recommendation, Assessment, Development, and Evaluation) system. Post-hoc analysis and subgroup meta-analyses were conducted to gain further insights into the various types of home telemonitoring technologies included in the systematic reviews and the impact of these technologies on clinical outcomes.

Results: A total of 15 reviews published between 2003 and 2013 were selected for meta-level synthesis. Evidence from high-quality reviews with meta-analysis indicated that taken collectively, home telemonitoring interventions reduce the relative risk of all-cause mortality (0.60 to 0.85) and heart failure-related hospitalizations (0.64 to 0.86) compared with usual care. Absolute risk reductions ranged from 1.4%-6.5% and 3.7%-8.2%, respectively. Improvements in HF-related hospitalizations appeared to be more pronounced in patients with stable HF: hazard ratio (HR) 0.70 (95% credible interval [Crl] 0.34-1.5]). Risk reductions in mortality and all-cause hospitalizations appeared to be greater in patients who had been recently discharged (≤28 days) from an acute care setting after a recent HF exacerbation: HR 0.62 (95% CrI 0.42-0.89) and HR 0.67 (95% CrI 0.42-0.97), respectively. However, quality of evidence for these outcomes ranged from moderate to low suggesting that further research is very likely to have an important impact on our confidence in the observed estimates of effect and may change these estimates. The post-hoc analysis identified five main types of non-invasive telemonitoring technologies included in the systematic reviews: (1) video-consultation, with or without transmission of vital signs, (2) mobile telemonitoring, (3) automated device-based telemonitoring, (4) interactive voice response, and (5) Web-based telemonitoring. Of these, only automated device-based telemonitoring and mobile telemonitoring were effective in reducing the risk of all-cause mortality and HF-related hospitalizations. More research data are required for interactive voice response systems, video-consultation, and Web-based telemonitoring to provide robust conclusions about their effectiveness.

Conclusions: Future research should focus on understanding the process by which home telemonitoring works in terms of improving outcomes, identify optimal strategies and the duration of follow-up for which it confers benefits, and further investigate whether there is differential effectiveness between chronic HF patient groups and types of home telemonitoring technologies.

(J Med Internet Res 2015;17(3):e63) doi:10.2196/jmir.4174

KEYWORDS

home telemonitoring; telemedicine; telehealth; remote monitoring; remote consultation; heart failure; chronic diseases; continuity of patient care; physiologic monitoring; ambulatory monitoring; home care services; review; umbrella review; systematic review; meta-analysis

Introduction

Heart failure (HF) is a chronic and life-threatening condition that places a substantial burden on health care systems worldwide with high rates of hospitalizations, readmissions, and outpatient visits [1]. Approximately 1-2% of the adult population in the western world has HF, with the prevalence rising to ≥10% among persons 65 years of age or older [2]. Developed countries devote 1%-2% of health care expenditures towards HF, while in the United States alone the estimated direct and indirect annual cost of HF was approximately US $39.2 billion in 2010 [3,4]. Improving the management of the ever-growing population of patients with HF is a high and growing priority for cardiovascular health services.

Home telemonitoring (HT) is a form of non-invasive, remote patient monitoring that has gained attention as a promising strategy to improve the care and management of patients with chronic HF. It can be particularly helpful for those living in remote and rural areas, the elderly and frail who are housebound, as well as those at high risk of deterioration [5]. HT involves the use of electronic devices and telecommunication technologies (eg, monitoring devices, hand-held or wearable technologies, and intelligent sensors) for the digital transmission of physiological and other disease-related data from the patient's home to a health care center providing care and clinical feedback. By allowing clinical data to be remotely collected on a regular basis, HT can enable early detection of clinical decompensation in patients with HF, allowing for timely intervention to prevent mortality events or further deterioration of the patient's condition necessitating hospitalization and use of more resources.

HT has attracted a large amount of research over the years both in the form of primary studies and systematic reviews. The extant literature contains published results from numerous trials investigating the clinical, structural, behavioral, or economic effects of HT interventions on patients with chronic diseases [6]. Similarly, the number of published systematic reviews aimed at summarizing the available evidence from primary studies on HT has increased considerably [7-9]. A recent critical review indicated that among all chronic conditions, HF has attracted the highest number of primary studies and systematic reviews [7]. Growing interest on the effects of home telemonitoring on patients with HF has led to a rise in the number of reviews addressing the same or very similar research questions with a concurrent increase in discordant findings in terms of direction and magnitude of HT effects. Differences in scope, methods of analysis, results, and quality of systematic reviews can cause great confusion and make it difficult for policy makers to access the review-level evidence, and for researchers to know where gaps in the evidence exist. Overviews of systematic reviews are an efficient way to gather and summarize in a single source the best available evidence on the effectiveness of interventions. They serve as a useful starting point for decision makers to unpack the evidence towards finding solutions to improve practice and identify areas where new research is needed [10].

This overview aims to collect, appraise, and summarize evidence from multiple systematic reviews examining the effects of HT interventions on patients with HF with a view to providing policy makers and practitioners with the evidence they need to make informed decisions related to the telemonitoring of HF patients. It also aims to identify research gaps in this area and suggest avenues for future research.

Methods

Overview

As shown below, the Cochrane Collaboration methodology [11] and available methodological guidelines for overviews of reviews [12,13] were rigorously applied throughout this study.

Search Methods

A systematic search of MEDLINE, EMBASE, CINAHL, and the Cochrane Library (Cochrane Database of Systematic Reviews, DARE, and the Health Technology Assessment Database) was conducted on December 13, 2013, and updated on December 3, 2014, using key terms and clinical query filters [14,15] to identify all relevant, peer-reviewed systematic reviews of HT interventional studies for patients with chronic HF, published since January 1, 1996. No language restrictions were applied. Details of the full search strategy are presented in Multimedia Appendix 1. Searches were supplemented by manual searching of two relevant scientific journals (ie, *Journal of Telemedicine and Telecare* and *Telemedicine and e-Health*) and screening of relevant reviews' bibliographies.

Selection of Systematic Reviews

In the preliminary stage, all titles and abstracts were screened by the first author according to the pre-specified inclusion criteria described in Table 1. References that clearly did not meet all of the criteria were excluded. Full-text articles were then retrieved and independently assessed for inclusion by the first 2 authors. Any disagreements were resolved through discussion. The third author was available for arbitration in case of persistent disagreements. Reviews were excluded if they studied the effects of HT on patients with other chronic or long-term conditions but did not report findings for HF separately.

Table 1. Inclusion criteria for the selection of relevant systematic reviews.

Criteria categories	Description of inclusion criteria
Study type	Systematic reviews (with or without meta-analysis) of original, interventional studies. Following the definitions used by the Cochrane Collaboration and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement, a systematic review was defined as a review that attempts to search, identify, appraise, and collate all empirical evidence that fits pre-specified eligibility criteria to answer a clearly stated set of objectives or specific research question(s), using explicit and systematic methods with a view to minimizing risk of bias.
Publication type	Full, peer-reviewed articles published in English.
Population	Patients with definitive diagnosis of HF.
Intervention	HT defined as the use of non-invasive devices in conjunction with information and communication technologies to monitor and electronically transmit physiological, biometric, and/or disease-related data (eg, arterial blood pressure, weight, cardiac rate, medications, symptoms) from the patient at home to the health care provider responsible for monitoring remotely the patient’s health status.
Comparisons	Standard (usual) care or other non–home telemonitoring approaches.
Outcomes	Primary or secondary outcomes pertaining to the clinical, structural, behavioral, or economic effects of HT. More specifically, systematic reviews reporting at least one of the following outcomes and having met the abovementioned criteria were eligible for inclusion: mortality, all-cause hospitalizations, HF-related hospitalizations, emergency department visits, clinic/outpatient visits, quality of life, cost-effectiveness, patient satisfaction, acceptability, and compliance/adherence.

Data Extraction

Data relating to key characteristics of the reviews, including information about the objectives, participants, intervention features, outcomes assessed, and comparisons performed; as well as the quality of included studies, quality of the reviews, pooled effect sizes for outcomes meta-analyzed, and main conclusions were extracted using an electronic form that was developed for the purposes of this review. Data were extracted by the first author and were verified for accuracy by the second. Differences were resolved in group meetings.

Quality Assessment of Included Reviews

The methodological quality of the included systematic reviews was independently assessed by the first two authors using the Assessment of Multiple Systematic Reviews (AMSTAR) tool [16]. AMSTAR is a validated instrument that uses 11 items to assess the degree to which review methods are unbiased. To ensure consistency of assessment between the 2 assessors, we developed decision support rules for scoring each criterion (Multimedia Appendix 2). Based on the results of the critical appraisal, reviews were categorized into three categories: “low” (score 0 to 3); “middle” (score 4 to 7); and “upper” (score 8 to 11). These groups reflect the existence of “major”, “moderate”, and “minor or no methodological limitations” in the included reviews, respectively.

Quality of Evidence in Included Systematic Reviews

Quality of evidence in included systematic reviews was assessed by outcome using the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) system criteria [17-27]. GRADE identifies five key elements that influence quality of evidence and can be used for rating down one’s confidence in the estimates of intervention effects. These are (1) risk of bias (limitations in the design and execution of primary studies included in the reviews) [20], (2) inconsistency (statistical heterogeneity between estimates of effect across studies) [23], (3) indirectness (applicability of participants, interventions, and outcomes to the clinical question under consideration) [24], (4) imprecision (impact of random error as reflected by the relative confidence interval of the summary effect estimate) [22], and (5) publication bias (publication or non-publication of research findings depending on the direction and statistical significance of the results of the primary studies) [21]. Assessing and combining these components determine the quality of evidence for each outcome of interest as “high” (ie, further research is very unlikely to change our confidence in the estimate of effect), “moderate” (ie, further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate), “low” (ie, further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate), or “very low” (ie, any estimate of effect is very uncertain). In cases where reviews had not performed a risk of bias assessment, we performed this task independently using the six criteria of the Cochrane Collaboration tool [28]. This facilitated application of the GRADE system and development of standardized summary of findings tables by outcome [11,29], as explained below, using the GRADEpro software (version 3.6). When a review had performed risk of bias and quality of evidence assessments, we collected this information during the data extraction process.

Analysis and Synthesis

To summarize the evidence on the effectiveness of HT interventions, we designed a 3-step process [30,31]. First, as a means of evaluating the comparability of the included systematic reviews and the extent to which reviews overlapped in terms of included studies, we carried out a bibliographic analysis that cross-linked individual systematic reviews with cited HT studies (Multimedia Appendix 3). Following the methodology used in Martel et al [32], we then calculated the ratio of cited to total pre-existing HT studies for each systematic review, taking into account the lag time between the reported end-date of search for identification of interventional studies and the actual publication date of the review. The mean ratio and 95% confidence intervals (CI) were calculated for both randomized controlled trials (RCTs) and observational studies to reflect the overall degree of overlap between the reviews. Second, we closely examined the objectives, participants, interventions, comparisons, and outcomes (PICO) characteristics of each review and subsequently categorized the reviews into homogeneous groups according to their common elements and relevant comparisons. We then recalculated the ratio of cited to total HT primary studies in order to evaluate the degree of overlap between reviews in the same group. These two steps were designed with a view to identify differences across reviews in terms of scope, range of included studies, and ways that HT interventions were split or lumped together, and to allow us formulate meaningful interpretation of the extracted data and results by disaggregating the evidence as appropriate. Third, within each group of the developed classification scheme, we ranked the reviews according to their methodological quality (AMSTAR score), documented the consistency of findings and conclusions across all reviews, and constructed standardized summary of findings tables following the Cochrane Collaboration guidelines [11,29] to summarize the effects of HT interventions from the most direct evidence (ie, from reviews that achieved the highest methodological quality score in each group of the taxonomy). In presenting the effects of HT, we focused on outcomes that were reported in more than 50% of the systematic reviews.

Results

Description of the Included Systematic Reviews

As shown in Figure 1, our search (up to December 13, 2013) yielded 4683 citations after removal of 771 duplicate references. After screening titles and abstracts, we retrieved 65 reviews in full text for further assessment. The references of these articles were manually screened to identify any relevant reviews that were not originally captured by our search strategy. This process yielded 8 more reviews. After further assessment, we excluded 55 articles that did not meet our eligibility criteria. Details about the primary reasons for exclusion and the full references of the excluded articles are provided in Multimedia Appendix 4. After completion of the screening process, we identified 15 systematic reviews for inclusion (18 references; 3 systematic reviews had duplicate publications) [33-50].

General characteristics about the population, interventions, and comparison groups included in the 15 systematic reviews along with the main conclusions of each review are summarized in Table 2. Reviews were published between 2003 and 2013, with more than half (n=9) published in 2009 or later. All reviews were published in peer-reviewed journals. However, two systematic reviews were initially published as Health Technology Assessment reports and one as a Cochrane review. Five reviews (33%) contained meta-analysis for at least one primary outcome of interest, while the remaining 10 used narrative synthesis. Seven reviews (based on the corresponding author) originated in North America (Canada=5 and United States=2), 6 in Europe (United Kingdom=3, Greece=1, Netherlands=1, Spain=1), and 2 in Australia. Four systematic reviews (27%), in addition to HT investigated the effects of structured telephone support (STS) on patients with HF [33,34,37,41]. However, findings from these distinctively different interventions were analyzed and reported separately. Therefore, data extracted and presented in Table 2 from these reviews pertain only to the HT studies and not the STS ones.

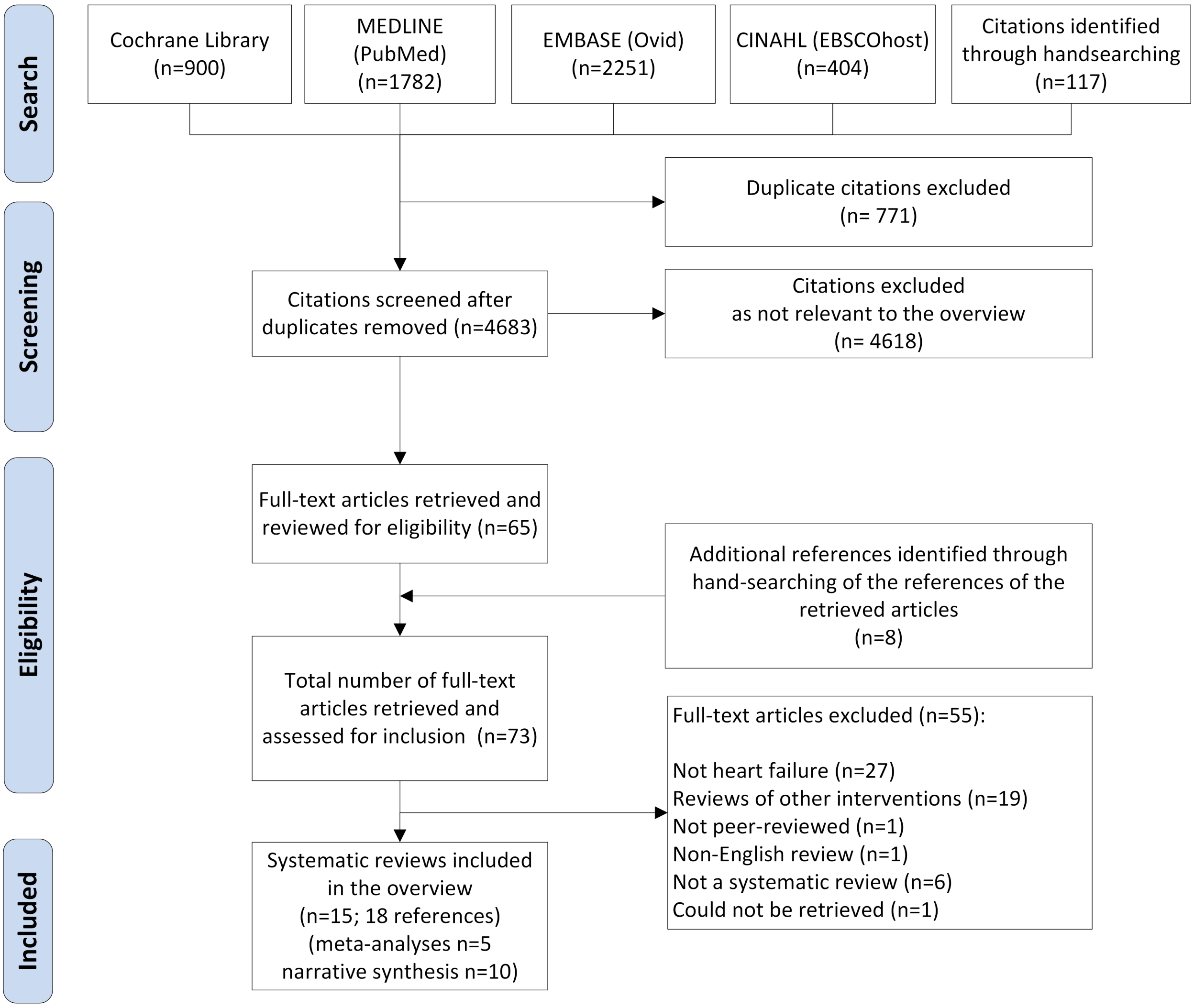
The number of HT studies identified and included in the reviews ranged from 4 to 42: mean 17.53 (SD 11.4). The total number of identified studies was 105; 38 (36%) referring to RCTs and 67 (64%) to observational studies. The citation patterns of these publications are presented in Multimedia Appendix 3. Seven reviews (47%) included only RCTs [33,34,36,37,41,43,47]. The remaining reviews included studies with quasi-experimental and cohort designs. The overall degree of overlap between the included reviews expressed as the mean ratio of cited to total pre-existing HT studies was 0.40 (95% CI 0.29-0.52) for RCTs, ranging from 0.11-1.00, and 0.25 (95% CI 0.1-0.42) for observational studies, ranging from 0.08-0.50.

Table 2. Characteristics of included systematic reviews.

Authors (year)	Years searched	Number and design of HT studies	Population (mean age; disease severity)	Intervention (length of follow-up)	Control group	Main conclusions
Clark et al (2007) [33]	2002 to May 2006	5 RCTs	807 patients (mean age range 57-75; NYHA class I-IV)	HT without home visits (Follow-up: 3-16 months)	Usual care	HT reduced all-cause mortality and HF-related hospitalizations Results were mixed for quality of life and costs
Inglis et al, 2010 [34,48]	2002 to Nov. 2008	14 RCTs	2710 patients (mean age range 57-78 years; NYHA class I-IV; most II-IV)	HT without home visits (Follow-up: 3-15 months)	Usual care	HT reduced the risk of all-cause mortality and HF-related hospitalizations HT improved quality of life and reduced costs No consistent impact on length of stay
Polisena et al, 2010 [35,49]	1998-2008	21 studies (11 RCTs, 10 observational)	3082 patients (mean age range 52-79; NYHA class I-IV; most III-IV)	HT with or without home visits (Follow-up: 1-12 months)	Usual care	HT reduced mortality and hospitalizations Patient quality of life with HT was similar or better than with usual care
Clarke et al, 2011 [36]	1969 to Oct. 2009	13 RCTs	3480 patients (mean age range 55-85 years; NYHA class I-IV)	HT interventions with or without home visits (Follow-up: 3-15 months)	Usual care	HT reduced all-cause mortality and HF hospitalizations HT in conjunction with nurse home visiting and specialist unit support can be effective in the clinical management of patients with HF and help improve their quality of life
Pandor et al, 2013 [37,50]	2002 to Jan. 2012	20 RCTs [10 RCTs of recently discharged patients (≤28 days) + 10 RCTs of patients with stable HF]	6561 patients [1918 recently discharged patients (mean age range 57-78 years; NYHA class: I-IV; most II-IV); 4643 patients with stable HF (mean age not summarized; NYHA class: I-IV)]	HT without home visits using patient-initiated external electronic devices with transfer of physiological data from the patient to the health care provider by landline or mobile phone, cable network or broadband technology (Follow-up: 3-12 months, recently discharged patients; 6-22 months, patients with stable HF)	Usual care	HT with medical support provided during office hours showed beneficial trends in reducing all-cause mortality for recently discharged patients with HF. However, these effects were statistically inconclusive Where usual care is below average or suboptimal, the impact of remote monitoring is likely to be greater
Louis et al, 2003 [38]	1966-2002	24 studies (6 RCTs, 12 observational)	2629 patients (mean age range 53-82 years; NYHA class: I-IV; most II-IV)	HT of patients using special telecare devices in conjunction with a telecommunication system (Follow-up: 2-18 months)	Usual care, home visits, and/or nurse telephone support	HT improved mortality, yet adequately powered multicenter RCTs are required to further evaluate the potential benefits and cost-effectiveness of this intervention
Martínez et al, 2006 [39]	1966 to April 2004	42 studies (13 reports of 10 RCTs, 29 observational)	2303 patients (5 studies did not specify number of participants) (mean age range 48-83; NYHA class I-IV; most II-IV)	HT using peripheral devices for measuring and automatically transmitting physiological data (Follow-up: 1-24 months)	Usual care, home nurse visits, pre/post HT	Reduces hospital readmissions, length of stay, mortality, emergency visits, and costs It is viable, easy to use, and is widely accepted by patients and health professionals
Paré et al, 2007 [40]	1990-2006	16 studies (7 reports of 5 RCTs, 9 observational)	Not summarized	HT as an automated process for the transmission of patient health status data (Follow-up: 1 to 36 months)	Usual care, home visits, pre/post HT	Promising patient management Future studies need to build evidence related to its clinical effects, cost effectiveness, impacts on services utilization, and acceptance by health care providers

Authors (year)	Years searched	Number and design of HT studies	Population (mean age; disease severity)	Intervention (length of follow-up)	Control group	Main conclusions
Chaudhry et al, 2007 [41]	1966 to Aug. 2006	4 RCTs	774 patients with HF (mean age range 59-70 years; NYHA class I-IV)	HT with or without home visits (Follow-up: not summarized)	Usual care, home visits	HT may be an effective strategy for disease management in high-risk heart failure patients , but the evidence base is currently quite limited
Seto 2008 [42]	up to April 2007	10 studies (5 RCTs, 4 observational, 1 survey)	1394 patients with HF (mean age range 58-74 years; NYHA not summarized)	HT with a component of home physiological measurement (Follow-up: 2-36 months)	Usual care, home visits, pre/post HT	All studies found cost reductions (range: 1.6% to 68.3%) mostly related to reduced hospitalization expenditures
Dang et al, 2009 [43]	1966 to Apr. 2009	9 RCTs	2020 adult patients with HF (mean age range 53-79 years; NYHA class II-IV)	Home telehealth remote monitoring (ie automated or physiologic monitoring of signs and symptoms; two-way video monitoring with or without physiologic monitoring; Internet, Internet Protocol, or Web-based technologies or image capture and transfer) (Follow-up: 3-12 months)	Usual care, home visits	Telemonitoring is a promising strategy. More research required to determine the ideal patient population, technology, and parameters, frequency and duration of telemonitoring, and the exact combination of case management and close monitoring that would assure consistent and improved outcomes with cost reductions in HF
Maric et al, 2009 [44]	Up to Aug. 2007	42 studies; 52 references (12 RCTs, 30 observational)	4290 patients (9 studies did not specify number of participants) (mean age and NYHA class not summarized)	HT using modalities that transmit data to health care professionals to assist in self-monitoring (eg, telephone-based touch pad, website based modalities, video consultations, and other technology-assisted devices) (Follow-up: 1-18 months)	Usual care, home visits, nurse telephone support, pre/post HT	Most studies demonstrated improvements in outcome measures, including improved QoL and decreased hospitalizations. However, not all studies reported the same improvements and in several cases the sample sizes were relatively small
Paré et al, 2010 [45]	1966-2008	17 studies (13 reports of 10 RCTs, 4 observational)	Not summarized	HT interventions in which physiological and biological data are transferred from the patients' home to the telemonitoring center to monitor patients, interpret the data, and make clinical decisions (Follow-up: not summarized)	Usual care, home visits, pre/post HT	Many studies failed to show a reduction in either mortality or hospitalization rates, although a few trials have reported a trend towards shorter lengths of stay in hospital. Due to the equivocal nature of current findings of HT involving patients with HF, larger trials are still needed to confirm the clinical effects of this technology for these patients.
Kraai et al, 2011 [46]	Up to November 2010	14 studies (4 RCTs, 10 observational)	2005 patients (mean age range 50-78; NYHA not summarized)	Noninvasive remote monitoring with external equipment to measure physiologic data such as weight and blood pressure (Follow-up: not summarized)	Usual care, home visits, nurse telephone support, pre/post HT	In general, patients seemed to be satisfied or very satisfied with HT
Giamouzis et al, 2012 [47]	2001 to Nov. 2011	12 RCTs	3877 patients (mean age range 57-78; NYHA class I-IV; most II-IV)	HT with at least one device that measured physiological data provided by the researchers for home use (Length of follow-up: 6 to 26 months)	Usual care	Currently available trial results tend to be in favor of HT HT was highly acceptable by HF patients

Figure 1. Selection process of the systematic reviews.



Population

As shown in Table 2, the majority of systematic reviews reported the mean age (12/15, 80%) and New York Heart Association (NYHA) class of the participants (10/15, 67%). The highest reported mean age in a study was 85 years [36] and the lowest was 48 years [39]. The NYHA class of participants in all but one review ranged from I to IV. Six reviews reported that most participants in the included studies were NYHA class II-IV [34-35,37-39,47]. One review [37] focused primarily on HF patients who had been recently discharged (within 28 days) from an acute care setting after a recent HF exacerbation and conducted additional analyses to assess whether or not the results from these studies differed markedly from the results of studies that included patients with stable HF. The remaining reviews did not make this distinction and combined in their analysis studies with both recently discharged and stable HF patients.

Intervention

Reviews consisted of a family of complex HT interventions, rather than a single type, involving various telehealth devices (eg, videoconferencing equipment, automated telemonitoring stations, mobile phones, and interactive voice or symptom response systems), technological approaches for data collection and transmission (eg, modem, broadband, or mobile phone transmission; Web-based or telephone touch-pad data entry), as well as other chronic disease management strategies (eg, education and home visits). However, reporting of the active ingredients of these interventions was often poor in the included reviews and not consistently performed for all primary studies (eg, [36,38-40,46,47]). While the majority of systematic reviews (10/15, 67%) summarized and reported relevant information about the type of physiological measures monitored in the included trials [33,34,37,38,41-44,47,49], and in fewer cases the frequency with which data transmission or communication between the patients and clinicians occurred [33,34,37,41,43,49], most reviews (8/15, 53%) did not provide sufficient details about the technology used in each of the included primary studies [33,34,36-40,45-47] (Table 3). Furthermore, virtually all reviews treated HT as a “black box” making no attempts to investigate whether technological differences between HT interventions are associated with different effects. In five reviews [37,38,45-47], the description of the intervention extracted from the trials was reduced into just a few words presented in a summary table or within the text, while three reviews did not provide any information at all about the included interventions [36,39-40]. Also, out of the 15 systematic reviews, only two classified and analyzed HT interventions according to the technologies used [41,44]. Lack of sufficient details about the included interventions coupled with the way that systematic reviews have analyzed the effects of HT up to now, make it difficult to determine what types of HT technologies have been used in the primary studies included in the systematic reviews, what proportion of studies included in the systematic reviews have used each type of HT technology, and more importantly, which HT technologies are more effective. We think this issue deserves more attention. This is why we have decided to extract additional data and conduct post-hoc analyses. Results are presented and discussed at the end of this section.

Table 3. Characterization of HT technologies in systematic reviews.

Author (year)	Does the review present information from all studies about the types of HT technologies used in the intervention group?	Does the review present information from all studies about the types of physiological parameters monitored in the intervention group?	Does the review classify and analyze studies according to the different types of HT technologies?
Inglis et al (2010) [34]	Yes	Yes	No
Clark et al (2007) [33]	Yes	Yes	No
Giamouzis et al (2012) [47]	No	Yes	No
Pandor et al (2013) [37]	No	Yes	No
Polisena et al (2010) [35]	Yes	Yes	No
Clarke et al (2011) [36]	No	No	No
Chaudhry et al (2007) [41]	Yes	Yes	Yes
Dang et al (2009) [43]	Yes	Yes	No
Louis et al (2003) [38]	No	Yes	No
Martinez et al (2006) [39]	No	No	No
Paré et al (2010) [45]	No	No	No
Kraai et al (2011) [46]	No	No	No
Maric et al (2009) [44]	Yes	Yes	Yes
Seto (2008) [42]	Yes	Yes	No
Paré et al (2007) [40]	No	No	No

Comparison

Six systematic reviews included RCTs in which the control group received usual (routine) care. This involved provision of post-discharge multidisciplinary care without intensified follow-up or home visits. Usual care varied across studies from in-person follow-up visits to a general practitioner or primary care provider to attendance at an outpatient clinic-based disease management program [33,34,36,37,47,49]. The remaining reviews, in addition to usual care, included RCTs that compared HT with other relevant comparators (eg, home visits or telephone support).

Outcomes

Despite differences in the scope and range of included studies, most reviews reported on a number of similar outcomes (Table 4). Most frequently reported outcomes included all-cause mortality (11/15, 73%), hospital re-admissions (13/15, 87%), costs (12/15, 80%), and length of stay (9/15, 60%). Six reviews made a distinction between all-cause and HF-related hospitalizations, while the remaining reviews did not. This methodological deficiency gave rise to unit-of-analysis errors, as authors often mixed study results from these two different outcomes. Other commonly reported outcomes comprised the impact of HT interventions on health care costs, quality of life, and length of stay in hospital. Compliance, acceptability, emergency room visits, and patient satisfaction were reported in less than half of the reviews and thus were excluded from our analysis.

Table 4. Outcomes reported by the systematic reviews[a].

Author (year)	ACM	ACH	HFH	HOSP	Costs	QoL	LoS	CA	ACC	ER	PS	OV
Inglis et al (2010) [34]	X	X	X		X	X	X	X	X			
Clark et al (2007) [33]	X	X	X		X	X		X	X			
Giamouzis et al (2012) [47]	X			X	X							
Pandor et al (2013) [37]	X	X	X			X	X	X	X		X	
Polisena et al (2010) [35]	X	X	X			X	X			X	X	X
Clarke et al (2011) [36]	X	X	X		X	X	X	X	X	X		
Chaudhry et al (2007) [41]	X	X	X		X							
Dang et al (2009) [43]	X	X	X		X		X			X		X
Louis et al (2003) [38]	X			X	X			X	X			
Martinez et al (2006) [39]	X			X	X	X	X		X			
Paré et al (2010) [45]	X			X			X					
Kraai et al (2011) [46]											X	
Maric et al (2009) [44]				X	X	X				X		
Seto (2008) [42]					X							
Paré et al (2007) [40]				X	X	X	X	X		X		

[a] ACM: all-cause mortality; ACH: all-cause hospitalizations; HFH: HF-related hospitalizations; HOSP: Hospitalizations (indicates reviews that did not make a distinction between all-cause and HF-related hospitalizations); Costs: Cost Savings; QoL: Quality of life; LoS: Length of stay; CA: Compliance; ACC: Acceptability; ER: Emergency room visits; PS: Patient satisfaction; OV: Outpatient visits.

Methodological Quality of Included Systematic Reviews

The methodological quality of the reviews varied considerably, and many of them had important limitations (Table 5). For example, only 5 reviews assessed risk of bias in the primary studies (Q7), while quality of evidence (Q8) was used in the interpretation of results in just four reviews. Only three reviews were found to be of high quality, achieving a score of 8 or more on the 11-point AMSTAR scale (3/15, 20%). These reviews indicated minimal bias in their design and execution. Four reviews were of moderate quality, scoring between 4 and 7 points, and eight reviews were of low quality scoring less than 4 points. Of the moderate quality reviews, most received a rating of 5 or less, which suggests that these reviews along with the ones of low quality may be at risk of important bias.

Table 5. Methodological quality of systematic reviews based on AMSTAR criteria and scores.[a,b]

Author (year)	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Total
Inglis et al (2010) [34]	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	10
Pandor et al (2013) [37]	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	9
Polisena et al (2010) [35]	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	8
Clark et al (2007) [33]	N	Y	Y	Y	N	Y	Y	N	Y	Y	N	7
Chaudhry et al (2007) [41]	N	CA	Y	CA	N	Y	Y	Y	Y	N/A	N	5
Dang et al (2009) [43]	N	CA	Y	N	Y	Y	N	N	Y	N/A	N	4
Paré et al (2010) [45]	Y	Y	N	N	Y	N	N	N	Y	N/A	N	4
Louis et al (2003) [38]	N	CA	Y	Y	N	Y	N	N	N	N/A	N	3
Martinez et al (2006) [39]	N	CA	Y	Y	N	Y	N	N	N	N/A	N	3
Giamouzis et al (2012) [47]	N	CA	Y	N	N	Y	N	N	N	N/A	N	2
Clarke et al (2011) [36]	N	CA	Y	N	CA	N	N	N	N	Y	N	2
Kraai et al (2011) [46]	N	CA	N	N	N	Y	N	N	Y	N/A	N	2
Maric et al (2009) [44]	N	N	N	N	N	Y	N	N	Y	N/A	N	2
Seto (2008) [42]	N	CA	N	N	N	Y	N	N	Y	N/A	N	2
Paré et al (2007) [40]	N	CA	Y	N	N	N	N	N	N	N/A	N	1

[a] Q1: A priori design; Q2: Duplicate study selection and data extraction; Q3: Search comprehensiveness; Q4: Inclusion of grey literature; Q5: Included and excluded studies provided; Q6: Characteristics of the included studies provided; Q7: Scientific quality of the primary studies assessed and documented; Q8: Scientific quality of included studies used appropriately in formulating conclusions; Q9: Appropriateness of methods used to combine studies' findings; Q10: Likelihood of publication bias was assessed; Q11: Conflict of interest – potential sources of support were clearly acknowledged in both the systematic review and the included studies.

[b] "Y" (Yes): Criterion met; "N" (No): Criterion not met; CA: Cannot answer; N/A: Not applicable. We awarded one point to each item that scored "yes" and summed these to calculate a total score for each review.

Classification of Reviews

Although ostensibly the 15 systematic reviews may appear to be similar, a closer examination of their PICO characteristics and criteria used for the selection of HT studies revealed several differences. For example, some reviews had a focused scope of inquiry and used narrow inclusion criteria to examine the effects of HT interventions without home visits for clinical assessment or educational purposes versus usual care. Other reviews had a broader scope of inquiry and examined the effects of HT interventions (with or without home visits) versus a variety of comparison interventions, including usual care, nursing telephone support, and/or home care. To overcome such heterogeneity, we carefully examined the population, intervention(s), and comparison group(s) that each review addressed, and subsequently classified reviews into homogeneous groups. Table 6 presents the taxonomy that was developed and used to facilitate the synthesis process.

One of the reviews [37] incorporated a network meta-analysis to determine the effectiveness of different remote monitoring strategies versus usual care on adult HF patients who had been recently discharged home (≤28 days) from an acute care setting and also (in additional analyses) on patients with stable HF. Remote monitoring strategies included STS, HT with clinical support during offices hours, and HT with clinical support provided 24/7. This review provided evidence for multiple comparison pairs and thus, findings from each comparison that was relevant to our overview were extracted and examined separately. For Comparison 2 (Table 6), it should be noted that we extracted the results of the sensitivity analysis that Pandor et al [37] performed, based on which a study that appeared to be an outlier (because patients in the control group received better-than-usual support and optimal medical treatment) was excluded.

Given the above classification of the reviews (Table 6), the ratio of cited to total pre-existing HT studies was recalculated to evaluate the degree of overlap between reviews in the same comparison group. Comparison groups 2, 3, and 4 were excluded from this analysis because they each comprised only one systematic review. Overall, the analysis yielded a mean ratio of 0.87 (95% CI 0.29-1.44) for RCTs in Comparison group 1, a mean ratio of 0.66 (95% CI 0.43-0.90) for RCTs in Comparison group 5, and a mean ratio of 0.48 (95% CI 0.28-0.67) for RCTs in Comparison group 6. The ratio of cited to total published observational studies remained the same, as virtually all reviews that included studies with observational designs were classified in the same group (Comparison 6). The observed increase in the mean ratio of cited to total pre-existing RCTs in each group compared to the overall mean ratio presented earlier, further supports the theoretically appealing classification of systematic reviews into groups and validates the methodological approach of summarizing the effects of HT interventions from the most direct evidence in each group to account for the overlap and methodological quality of the reviews.

Table 6. Taxonomy of HT systematic reviews according to key elements.

Comparisons	Systematic reviews	AMSTAR score	Review characteristics
1	Inglis et al (2010) [34]	10	Population: Stable and recently discharged patients
	Clark et al (2007) [33]	7	Intervention: HT with clinical support provided during office hours or 24/7, without home visits for clinical assessment or educational purposes
	Giamouzis et al (2012) [47]	2	Comparator group: Usual care
2	Pandor et al (2013) [37]	9	Population: Recently discharged patients (≤28 days) Intervention: HT with clinical support provided during office hours only, without home visits for clinical assessment or educational purposes Comparator group: Usual care
3	Pandor et al (2013) [37]	9	Population: Patients with stable heart failure Intervention: HT with clinical support provided during office hours only, without home visits for clinical assessment or educational purposes Comparator group: Usual care
4	Pandor et al (2013) [37]	9	Population: Patients with stable heart failure Intervention: HT with clinical support provided 24/7, without home visits for clinical assessment Comparator group: Usual care
5	Polisena et al (2010) [35]	8	Population: Stable and recently discharged patients
	Clarke et al (2011) [36]	2	Intervention: HT with clinical support provided during office hours or 24/7, with or without home visits for clinical assessment or educational purposes Comparator group: Usual care
6	Chaudhry et al (2007) [41]	5	Population: Stable and recently discharged patients
	Dang et al (2009) [43]	4	Intervention: HT with clinical support provided during office hours or 24/7, with or without home visits for clinical assessment or educational purposes
	Paré et al (2010) [45]	4	Comparator groups: usual care, home visits, nursing telephone support, and/or pre/post HT
	Martinez et al (2006) [39]	3	
	Louis et al (2003) [38]	3	
	Kraai et al (2011) [46]	2	
	Maric et al (2009) [44]	2	
	Seto (2008) [42]	2	
	Paré et al (2007) [40]	1	

Home Telemonitoring Effects

All-Cause Mortality

Eleven reviews examined the effects of HT on all-cause mortality. Of them, five pooled study results into a meta-analysis [33-37]. The remaining six summarized the available evidence narratively using various qualitative or semi-quantitative techniques [38,39,41,43,45]. All reviews across the six taxonomy groups concluded that HT is effective in reducing the risk of all-cause mortality. As shown in Table 7, relative risk reductions in meta-analyses that achieved the highest AMSTAR score in each group of the taxonomy ranged from 15%, with hazard ratio (HR) 0.85 (95% credible interval [CrI] 0.59-1.2) [37] to 40%, with risk ratio (RR) 0.60 (95% CI 0.45-0.81) [35]. The absolute risk reduction (ARR) in mortality with HT ranged from 1.4% to 6.5%, respectively. Similar relative and absolute risk reductions were also observed in the other two meta-analyses that achieved lower AMSTAR scores [33,36]. The strongest evidence (moderate quality) comes from the Cochrane review that compared the effects of HT without home visits or intensified attendance at cardiology clinics with usual care. In the pooling of all-cause mortality data from 11 RCTs involving 2710 patients with HF, HT resulted in statistically significant risk reduction of 34% (RR 0.66 [95% CI 0.54-0.81]); an ARR of 5.2% (95% CI -2.9 to -7.1), equating to a number needed to treat of 19 to postpone one death. Evidence from a more recent review [37] suggests that improvements in survival rates with HT are more pronounced in patients who have been recently discharged from the hospital (HR 0.62 [95% CrI 0.42-0.89]; ARR 5% [95% CI -1.4 to -7.8]) than patients without any acute event or deterioration in the past 28 days before randomization (HR 0.85 [95% CrI 0.59-1.20]; ARR 1.4% [95% CI -3.9 to 1.9]). However, quality of evidence was low for this finding, suggesting that further research is very likely to have an important impact on this estimate.

All-Cause Hospitalizations

With respect to all-cause hospitalizations, most reviews reported beneficial effects with HT interventions. However, the magnitude and uncertainty of the reported estimates varied considerably across reviews due to differences in the inclusion, classification, and analysis of HT studies. Relative effects in meta-analyses that favored HT over usual care ranged from 0.99 (95% CI 0.88-1.11) [36] to 0.67 (95% CrI 0.42-0.97) [37]. The ARR in all-cause hospitalizations with HT varied from 4.7%-13.8%. As shown in Table 8, the largest relative and absolute risk reduction was seen in recently discharged patients with HF, receiving HT with clinical support during office hours. Results from two separate meta-analyses of RCTs that included patients with stable HF (ie, without any acute event or deterioration in the past 28 days prior to randomization) yielded markedly different results. HT with clinical support provided during office hours increased the risk of all-cause hospitalizations in patients with stable HF by 17% (HR 1.17 [95% CrI 0.89-1.59], with an absolute risk increase of 4.7% [95% -3.2 to 14.8]). While HT with 24/4 clinical support yielded beneficial but marginally lower relative effects than those observed in recently discharged patients: HR 0.84 (95% CrI 0.54-1.15); ARR 5.7% (-1.8 to 4.8%). Results were statistically inconclusive owing to the small number of patients and events in the included RCTs. Despite the generally favorable effects of HT observed in the reviews, quality of evidence was consistently low for the outcome of all-cause hospitalizations due to risk of bias in the primary studies and statistical heterogeneity in the meta-analyses.

Heart Failure Hospitalizations

A total of seven reviews examined the impact of HT interventions on HF-related hospitalizations. All of them reported beneficial effects with HT. Relative risk reductions in the four reviews that incorporated meta-analysis ranged from 14% (HR 0.86 [95% CI 0.61-1.21)] [37] to 36% (HR 0.64 [95% CrI 0.34-1.14]), while the ARR in HF-related hospitalizations extended from 3.7%-8.2%. As shown in Table 9, the strongest evidence (moderate quality) comes from a high-quality meta-analysis of 4 large RCTs (N=1570 patients), which found statistically significant risk reductions of 21% in HF-related hospitalizations with HT versus usual care, equating to an ARR of 6% (95% CI-1.7 to -9.7) [34]. Findings from a recent review [37] suggest that improvements in HF-related hospitalizations might be more pronounced in patients with stable HF receiving telemonitoring with clinical support during 24/7 (HR 0.64 [95% CrI 0.34-1.14]). HT interventions with clinical support during office hours yielded a smaller relative effect for recently discharged patients than patients with stable HF (HR 0.86 [95% CrI 0.61-1.21] versus HR 0.70 [95% CrI 0.34-1.50]). However, results were statistically inconclusive and the overall quality of evidence was found to be low, suggesting that these estimates should be interpreted with caution until more evidence accumulates.

Table 7. Summary of findings from the meta-analyses with the most direct evidence in each group for the outcome of all-cause mortality.[a-g]

Outcome: Comparison	Number of participants (studies), Follow-up	Quality of evidence, GRADE[a-d]	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator	Risk difference with HT (95% CI)
All-cause mortality: Comparison 1 [34]					
Population: Stable and recently discharged patients Intervention: Home telemonitoring with clinical support provided during office hours or 24/7, without home visits for clinical assessment or educational purposes Comparator group: usual care	2710 (11 studies) 3-8 months	MODERATE due to risk of bias[e]	RR 0.66 (0.54 to 0.81)	154 per 1000	52 fewer per 1000 (from 29 fewer to 71 fewer)
All-cause mortality: Comparison 2 [37]					
Population: Recently discharged patients (≤28 days) Intervention: Home telemonitoring with clinical support provided during office hours, without home visits for clinical assessment or educational purposes Comparator group: usual care	1234 (8 studies) 3-12 months	LOW due to risk of bias and imprecision[e,f]	HR 0.62 (0.42 to 0.89)[g]	139 per 1000	50 fewer per 1000 (from 14 fewer to 78 fewer)
All-cause mortality: Comparison 3 [37]					
Population: Patients with stable heart failure Intervention: Home telemonitoring with clinical support provided during office hours, without home visits for clinical assessment or educational purposes Comparator group: usual care	1501 (7 studies) 6-12 months	LOW due to risk of bias and imprecision[e,f]	HR 0.85 (0.59 to 1.2)[g]	99 per 1000	14 fewer per 1000 (from 39 fewer to 19 more)
All-cause mortality: Comparison 4 [37]					
Population: Patients with stable heart failure Intervention: Home telemonitoring with clinical support provided 24/7, without home visits for clinical assessment or educational purposes Comparator group: usual care	1258 (3 studies) 12-24 months	LOW due to risk of bias and imprecision[e,f]	HR 0.85 (0.58 to 1.27)[g]	143 per 1000	20 fewer per 1000 (from 57 fewer to 35 more)
All-cause mortality: Comparison 5 [35]					
Population: Stable and recently discharged patients Intervention: Home telemonitoring with clinical support provided during office hours or 24/7, with or without home visits for clinical assessment or educational purposes Comparator group: usual care	1200 (5 studies) 3-12 months	MODERATE due to risk of bias[e]	RR 0.60 (0.45 to 0.81)	164 per 1000	65 fewer per 1000 (from 31 fewer to 90 fewer)

[a] High quality: Further research is very unlikely to change our confidence in the estimate of effect.

[b] Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

[c] Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

[d] Very low quality: We are very uncertain about the estimate.

[e] Most trials did not provide details of random sequence generation, allocation concealment, and blinding of data analysts or assessors (see Multimedia Appendix 5).

[f] The optimal information size criterion was not met by the meta-analysis (power <80%).

[g] 95% credible intervals (Bayesian meta-analysis).

Table 8. Summary of findings from the meta-analyses with the most direct evidence in each group for the outcome of all-cause hospitalizations.[a-h]

Outcome: Comparison	Number of participants (studies) Follow-up	Quality of evidence GRADE[a-d]	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator	Risk difference with HT (95% CI)
All-cause hospitalization: Comparison 1 [34]					
Population: Stable and recently discharged patients Intervention: Home telemonitoring with clinical support provided during office hours or 24/7, without home visits for clinical assessment or educational purposes Comparator group: usual care	2343 (8 studies) 6-12 months	LOW due to risk of bias, inconsistency, and imprecision[e-g]	RR 0.91 (0.84 to 0.99)	521 per 1000	47 fewer per 1000 (from 5 fewer to 83 fewer)
All-cause hospitalization: Comparison 2 [37]					
Population: Recently discharged patients (≤28 days) Intervention: Home telemonitoring with clinical support provided during office hours, without home visits for clinical assessment or educational purposes Comparator group: usual care	831 (5 studies) 6-12 months	LOW due to risk of bias, inconsistency, imprecision[e-g]	HR 0.67 (0.42 to 0.97)[h]	569 per 1000	138 fewer per 1000 (from 11 fewer to 271 fewer)
All-cause hospitalization: Comparison 3 [37]					
Population: Patients with stable heart failure Intervention: Home telemonitoring with clinical support provided during office hours, without home visits for clinical assessment or educational purposes Comparator group: usual care	1267 (5 studies) 6-12 months	LOW due to risk of bias, imprecision[e,f]	HR 1.17 (0.89 to 1.59)[h]	357 per 1000	47 more per 1000 (from 32 fewer to 148 more)
All-cause hospitalization: Comparison 4 [37]					
Population: Patients with stable heart failure Intervention: Home telemonitoring with clinical support provided 24/7, without home visits for clinical assessment or educational purposes Comparator group: usual care	1258 (3 studies) 12-24 months	LOW due to risk of bias, inconsistency, imprecision[e-g]	HR 0.84 (0.54 to 1.15)[h]	474 per 1000	57 fewer per 1000 (from 181 fewer to 48 more)
All-cause hospitalization: Comparison 5 [35]					
Population: Stable and recently discharged patients Intervention: Home telemonitoring with clinical support provided during office hours or 24/7, with or without home visits for clinical assessment or educational purposes Comparator group: usual care	787 (3 studies) 3-12 months	LOW due to risk of bias, imprecision[e,f]	RR 0.79 (0.66 to 0.94)	438 per 1000	92 fewer per 1000 (from 26 fewer to 149 fewer)

[a] High quality: Further research is very unlikely to change our confidence in the estimate of effect.

[b] Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

[c] Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

[d] Very low quality: We are very uncertain about the estimate.

[e] Most trials did not provide details of random sequence generation, allocation concealment, and blinding of data analysts or assessors (see Multimedia Appendix 5).

[f] The optimal information size criterion was not met by the meta-analysis (power <80%).

[g] Serious unexplained inconsistency/heterogeneity (I^2>70%). Point estimates and confidence intervals between RCTs varied considerably in magnitude and direction.

[h] 95% credible intervals (Bayesian meta-analysis).

Table 9. Summary of findings from the meta-analyses with the most direct evidence in each group for the outcome of HF-related hospitalizations.

Outcome: Comparison	Number of participants (studies) Follow up	Quality of evidence (GRADE)[a-d]	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator	Risk difference with HT (95% CI)
HF-related hospitalizations: Comparison 1 [34]					
Population: Stable and recently discharged patients Intervention: Home telemonitoring with clinical support provided during office hours or 24/7, without home visits for clinical assessment or educational purposes Comparator group: usual care	1570 (4 studies) 8-12 months	MODERATE due to risk of bias[e]	RR 0.79 (0.67 to 0.94)	285 per 1000	60 fewer per 1000 (from 17 fewer to 94 fewer)
HF-related hospitalizations: Comparison 2 [37]					
Population: Recently discharged patients (≤28 days) Intervention: Home telemonitoring with clinical support provided during office hours, without home visits for clinical assessment or educational purposes Comparator group: usual care	755 (2 studies) 6-8 months	LOW due to risk of bias and imprecision[e,f]	HR 0.86 (0.61 to 1.21)[g]	315 per 1000	37 fewer per 1000 (from 109 fewer to 52 more)
HF-related hospitalizations: Comparison 3 [37]					
Population: Patients with stable heart failure Intervention: Home telemonitoring with clinical support provided during office hours, without home visits for clinical assessment or educational purposes Comparator group: usual care	432 (2 studies) 12 months	LOW due to risk of bias and imprecision[e,f]	HR 0.70 (0.34 to 1.5)[g]	221 per 1000	61 fewer per 1000 (from 139 fewer to 91 more)
HF-related hospitalizations: Comparison 4 [37]					
Population: Patients with stable heart failure Intervention: Home telemonitoring with clinical support provided 24/7, without home visits for clinical assessment or educational purposes Comparator group: usual care	1170 (3 studies) 12-24 months	LOW due to risk of bias and imprecision[e,f]	HR 0.64 (0.34 to 1.14)[g]	251 per 1000	82 fewer per 1000 (from 157 fewer to 30 more)

[a] High quality: Further research is very unlikely to change our confidence in the estimate of effect.

[b] Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

[c] Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

[d] Very low quality: We are very uncertain about the estimate.

[e] Most trials did not provide details of random sequence generation, allocation concealment, and blinding of data analysts or assessors (see Multimedia Appendix 5).

[f] The optimal information size criterion was not met by the meta-analysis (power <80%).

[g] 95% credible intervals (Bayesian meta-analysis).

Cost Savings

Eleven reviews examined the effects of HT interventions on cost savings, but none of them pooled results into a meta-analysis due to the inconsistency in cost-analysis methods used in the original studies. One review of both RCTs and observational studies that focused explicitly on cost savings as an outcome found direct cost reductions to the health care system from HT compared to usual care, which ranged between 1.6% and 68.3% [42]. Eight reviews concurred that the impact of HT interventions on health care costs appeared to be positive in more cases than not, but in general results were statistically inconclusive and varied depending on the context and specific national health system of the study [34,36,38,39,41,43,44,47]. Identified cost reductions, taken individually and collectively, were mainly associated with savings from reduced expenditures on hospitalizations and, to a lesser extent, from home visits and patient travel costs. Although none of the systematic reviews used standardized instruments or validated methods to formally appraise the quality of economic evaluations in the original studies, most reviews criticized the methodologies adopted in these studies and strongly recommended that future research rigorously conduct cost-effectiveness assessments of HT in adequately powered RCTs.

Quality of Life

Eight reviews included health-related quality of life as an outcome measure. All of them summarized the available evidence qualitatively owing to the different assessment instruments used in the primary studies. Overall, reviews concluded that HT improved quality of life. However, such inferences were not supported by the study-level evidence they presented. For example, in one of the high-quality reviews, the authors concluded that HT improves quality of life, but only three of the seven studies that reported data for this outcome found positive and statistically significant improvements. In most reviews, data extraction and reporting pertaining to this outcome was inadequate, since authors focused on the statistical significance of study results rather than the direction and magnitude of effect [33-35,39,44].

Length of Stay

The impact of HT on hospital length of stay due to exacerbated HF events and/or any cause hospitalization was examined in eight reviews. Results for this outcome were tabulated and summarized narratively. Three reviews, all of which focused on RCTs, concluded that the impact of HT on length of stay was ambivalent [34,36,37]. However, the remaining systematic reviews reached different conclusions [35,39,40,43,45]. Reviews that incorporated observational studies in their analysis and a broader set of interventional studies with various comparator groups (eg, home visits and nurse telephone support) concluded that HT reduces length of stay [39,40,45].

Opening the Black Box of Home Telemonitoring Technologies

Description

As described earlier, the extent to which various technological devices may have an impact on the effectiveness of HT has not been investigated systematically in previous reviews. Despite the different forms and generations of HT that have emerged over the years as a result of the continuous technological advances and efforts to improve remote monitoring of patients with HF, most systematic reviews have treated HT as a black box, paying little or no attention to the technology component.

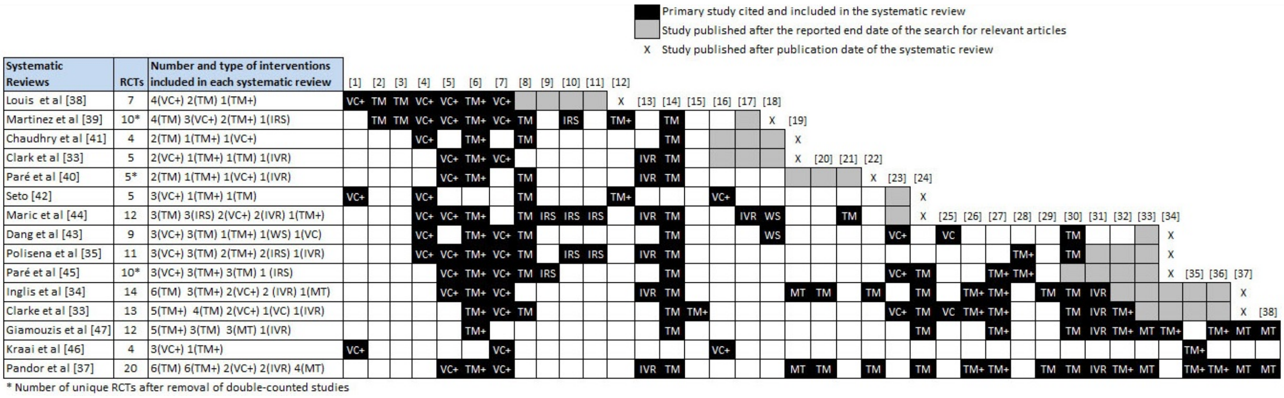
To investigate this issue further, we extracted data from the primary studies included in the 15 reviews and conducted a series of post-hoc analyses. Our main goal was to gain further insights into the various types of HT technologies in use and investigate the link between HT technologies and clinical effectiveness. Put simply, we explored the following question: Does HT technology matter? Our results are presented in the following two subsections.

Toward a Taxonomy of Home Telemonitoring Interventions

Data Extraction and Post-Hoc Analyses

Building on the citation matrices presented earlier (Multimedia Appendix 3), we retrieved all primary studies included across the 15 reviews, extracted relevant information from each study about the different technologies and monitoring approaches in use, and subsequently classified HT interventions into groups according to the technology in use (Figures 2 and 3). We were able to extract data from all 105 primary studies, with the exception of 8 observational studies that could not be retrieved. These are denoted in Figure 3 with the letter "U" (Unknown). By carefully reading through the detailed descriptions of the interventions provided in the primary studies, we identified five main types of HT interventions as follows.

Figure 2. Citation matrix of previously published RCTs included in the 15 systematic reviews (all references are available in Multimedia Appendix 3).

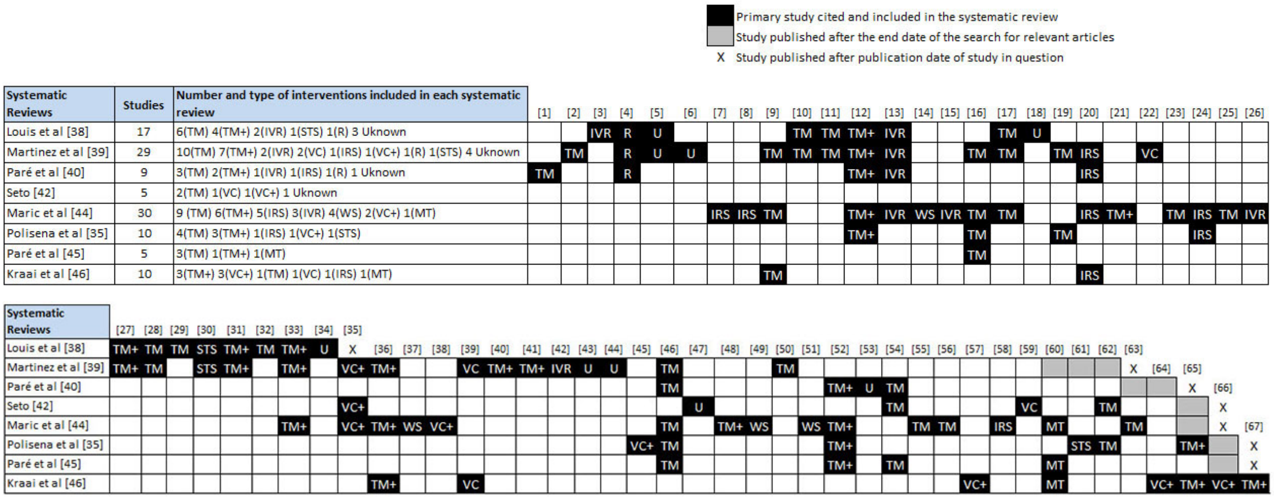


■ Primary study cited and included in the systematic review
▨ Study published after the reported end date of the search for relevant articles
X Study published after publication date of the systematic review

Systematic Reviews	RCTs	Number and type of interventions included in each systematic review	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]	[33]	[34]	[35]	[36]	[37]	[38]
Louis et al [38]	7	4(VC+) 2(TM) 1(TM+)	VC+	TM	TM	VC+	VC+	TM+	VC+	▨	▨	▨	▨	X																										
Martinez et al [39]	10*	4(TM) 3(VC+) 2(TM+) 1(IRS)		TM	TM	VC+	VC+	TM+	VC+	TM		IRS		TM+		TM			▨	X																				
Chaudhry et al [41]	4	2(TM) 1(TM+) 1(VC+)				VC+		TM+		TM						TM		▨	▨	▨	X																			
Clark et al [33]	5	2(VC+) 1(TM+) 1(TM) 1(IVR)					VC+	TM+	VC+						IVR	TM		▨	▨	▨	X																			
Paré et al [40]	5*	2(TM) 1(TM+) 1(VC+) 1(IVR)					VC+	TM+		TM					IVR	TM					▨	▨	▨	X																
Seto [42]	5	3(VC+) 1(TM+) 1(TM)	VC+			VC+				TM				TM+				VC+							▨	X														
Maric et al [44]	12	3(TM) 3(IRS) 2(VC+) 2(IVR) 1(TM+)				VC+	VC+	TM+		TM	IRS	IRS	IRS		IVR	TM			IVR	WS			TM		▨	X														
Dang et al [43]	9	3(VC+) 3(TM) 1(TM+) 1(WS) 1(VC)				VC+		TM+	VC+	TM						TM				WS					VC+		VC					TM			▨	X				
Polisena et al [35]	11	3(VC+) 3(TM) 2(TM+) 2(IRS) 1(IVR)				VC+	VC+	TM+	VC+	TM		IRS	IRS		IVR	TM														TM+		TM	▨	▨	▨	X				
Paré et al [45]	10*	3(VC+) 3(TM+) 3(TM) 1 (IRS)					VC+	TM+	VC+	TM	IRS					TM									VC+	TM			TM+	TM+		▨	▨	▨	▨	X				
Inglis et al [34]	14	6(TM) 3(TM+) 2(VC+) 2 (IVR) 1(MT)					VC+	TM+	VC+						IVR	TM					MT	TM		TM		TM		TM+	TM+		TM	TM	IVR	▨	▨	▨	▨	▨	X	
Clarke et al [33]	13	5(TM+) 4(TM) 2(VC+) 1(VC) 1(IVR)						TM+	VC+	TM						TM	TM+								VC+	TM	VC	TM+	TM+			TM	IVR	TM+	▨	▨	▨	▨	X	
Giamouzis et al [47]	12	5(TM+) 3(TM) 3(MT) 1(IVR)						TM+								TM										TM			TM+			TM	IVR	TM+	MT	TM+		TM+	MT	MT
Kraai et al [46]	4	3(VC+) 1(TM+)	VC+						VC+									VC+																			TM+			
Pandor et al [37]	20	6(TM) 6(TM+) 2(VC+) 2(IVR) 4(MT)					VC+	TM+	VC+						IVR	TM					MT	TM		TM		TM		TM+	TM+		TM	TM	IVR	TM+	MT		TM+	TM+	MT	MT

* Number of unique RCTs after removal of double-counted studies

Figure 3. Citation matrix of previously published observational studies included in the 15 systematic reviews (all references are available in Multimedia Appendix 3).



■ Primary study cited and included in the systematic review
▨ Study published after the end date of the search for relevant articles
X Study published after publication date of study in question

Systematic Reviews	Studies	Number and type of interventions included in each systematic review	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]
Louis et al [38]	17	6(TM) 4(TM+) 2(IVR) 1(STS) 1(R) 3 Uknown			IVR	R	U					TM	TM	TM+	IVR				TM	U								
Martinez et al [39]	29	10(TM) 7(TM+) 2(IVR) 2(VC) 1(IRS) 1(VC+) 1(R) 1(STS) 4 Uknown		TM		R	U	U			TM	TM	TM	TM+	IVR			TM	TM		TM	IRS		VC				
Paré et al [40]	9	3(TM) 2(TM+) 1(IVR) 1(IRS) 1(R) 1 Uknown	TM			R								TM+	IVR							IRS						
Seto [42]	5	2(TM) 1(VC) 1(VC+) 1 Uknown																										
Maric et al [44]	30	9 (TM) 6(TM+) 5(IRS) 3(IVR) 4(WS) 2(VC+) 1(MT)							IRS	IRS	TM			TM+	IVR	WS	IVR	TM	TM			IRS	TM+		TM	IRS	TM	IVR
Polisena et al [35]	10	4(TM) 3(TM+) 1(IRS) 1(VC+) 1(STS)												TM+				TM			TM					IRS		
Paré et al [45]	5	3(TM) 1(TM+) 1(MT)																TM										
Kraai et al [46]	10	3(TM+) 3(VC+) 1(TM) 1(VC) 1(IRS) 1(MT)									TM											IRS						

Systematic Reviews	[27]	[28]	[29]	[30]	[31]	[32]	[33]	[34]	[35]	[36]	[37]	[38]	[39]	[40]	[41]	[42]	[43]	[44]	[45]	[46]	[47]	[48]	[49]	[50]	[51]	[52]	[53]	[54]	[55]	[56]	[57]	[58]	[59]	[60]	[61]	[62]	[63]	[64]	[65]	[66]	[67]
Louis et al [38]	TM+	TM	TM	STS	TM+	TM	TM+	U	X																																
Martinez et al [39]	TM+	TM		STS	TM+		TM+		VC+	TM+			VC	TM+	TM+	IVR	U	U		TM				TM										▨	▨	▨	X				
Paré et al [40]																				TM						TM+	U	TM									▨	▨	X		
Seto [42]									VC+												U							TM					VC			TM			▨	X	
Maric et al [44]							TM+		VC+	TM+	WS	VC+								TM		TM+	WS		WS	TM+			TM	TM		IRS		MT			TM		▨	X	
Polisena et al [35]																			VC+	TM						TM+									STS	TM			TM+	▨	X
Paré et al [45]																				TM						TM+		TM						MT						▨	X
Kraai et al [46]										TM+			VC																		VC+			MT				VC+	TM+	VC+	TM+

Video Consultation

Seven RCTs and 9 observational studies across 14 systematic reviews involved real-time, two-way synchronous communication between patients and caregivers via the use of specialized videoconferencing equipment. In most of these studies (6 RCTs, 6 observational), which are denoted with the acronym "VC+", patients received scheduled video-consultations using specialized videoconferencing equipment and peripheral devices that were either interconnected with the main videoconferencing unit or were capable of independently transmitting basic physiological measures through telephone (trans-telephonic transmission) or broadband connection. In the studies that used videoconferencing equipment with integrated electronic stethoscopes and blood pressure monitors, nurses used a headset at the receiving station to hear cardiac sounds and video-consult with patients about items such as weight, symptoms, and observance. This process allowed in-depth assessment and triage at pre-scheduled times (eg, daily or twice weekly) without patients having to leave home. In the remaining studies (1 RCT and 3 observational), the intervention involved use of stand-alone videoconferencing equipment without transmission of vital signs. These studies are denoted as "VC".

Automated Device-Based Telemonitoring

A total of 23 RCTs and 37 observational studies across 15 systematic reviews utilized various non-invasive electronic devices, to support remote patient monitoring and automated transmission of vital signs and/or symptoms. These interventions were further categorized into three mutually exclusive subgroups, based on the technological characteristics of the intervention, as well as methods of data collection and transmission.

The first subgroup of studies (10 RCTs and 20 observational) employed patient-initiated electronic devices, such as wired or wireless weight scales and blood pressure monitors, capable of measuring and automatically transmitting data from the patient at home to a Web-based server through dedicated telecommunication stations (eg, modems or broadband connection boxes). These studies are denoted as "TM". The second subgroup (10 RCTs and 12 observational studies) also involved the use of patient-initiated electronic devices for monitoring and transmission of vital signs, but contrary to the first subgroup, these devices were interconnected with individualized symptom response devices that prompted patients about a heart-healthy diet, physical activity, and medication compliance, and requested answers (Yes/No) to relevant questions about HF-related symptoms. Patient responses along with monitored vital signs were recorded and automatically uploaded to a central server through a telephone line or broadband connection. Clinicians were able to monitor the patients, provide advice, and update treatment regimens by accessing data via standard browser interfaces. Studies belonging to this subgroup are denoted as "TM+". The third subgroup of studies (3 RCTs and 5 observational) cited across 6 systematic reviews involved the use of stand-alone, interactive symptom response devices without capability of transmission of physiological measures or vital signs. Studies in this subgroup are denoted as "IRS".

Web-Based Telemonitoring

One RCT and 4 observational studies cited across 2 systematic reviews examined the effectiveness of secured websites for the purpose of patient remote monitoring. Interventions involved the use of personal computers or laptops with Internet access, allowing patients gain secure access to a Web-based portal to manually enter their vital signs obtained through stand-alone peripheral devices (eg, blood pressure monitors, weight scales, sphygmomanometer, and pedometer), answer a range of questions about symptoms, and receive feedback, as well as educational material. Web-based telemonitoring studies are denoted with the acronym "WS".

Interactive Voice Response Systems

A total of 3 RCTs and 5 observational studies cited across 9 systematic reviews utilized automated interactive voice response systems (similar to the ones used by airline companies and banks) that required manual input of data by the patient using the telephone keypad of their home or mobile telephone to send information concerning vital signs and symptoms to health care providers. Each parameter was entered by the patient in reply to a computer-generated question asked by a recorded voice. The system generated alerts to the study nurse when pre-specified symptoms or physiologic changes were detected. Studies that implemented an automated interactive voice response system are denoted as “IVR”.

Mobile Telemonitoring

Four RCTs and 1 observational study across 6 systematic reviews (denoted in Figures 2 and 3 as “MT”) involved telemonitoring of patients with HF via the use of mobile devices such as mobile phones and personal digital assistants (PDAs). In 3 RCTs, patients were instructed to use on a daily basis external electronic devices that were interconnected via Bluetooth to their mobile phone or PDA allowing measurement and automatic forwarding of physiological parameters and vital signs (eg, electrocardiogram, blood pressure, weight) to a central server. In the other two studies, the process of measurement and transmission was not automated. After establishing an Internet connection using the micro browser of their mobile device, patients had to manually enter their vital signs and physiological parameters using data-entry templates generated in Wireless Markup Language (WML).

Three observational studies cited across 4 systematic reviews involved the use of other communication platforms that could not be classified under any of the above categories. Two used structured telephone support only (denoted in Figure 3 as “STS”), while a third study used a pager to transmit computer-generated reminders to the patient to take medications, weigh themselves, and measure their blood pressure and heart rate. Patients were then contacted by phone once a week to verbally communicate to the nurse their physiological data. This study is denoted in Figure 3 as “R” (Reminders).

Exploratory Analysis of the Effectiveness of Home Telemonitoring Technologies

Description

Having identified the main types of HT interventions included in the systematic reviews, we now turn our attention to the question of whether technology has an impact on the effectiveness of HT and whether it can explain variations in the direction and magnitude of the observed effect sizes with respect to all-cause mortality, all-cause hospitalizations, and HF-related hospitalizations.

Using data from the three systematic reviews in our sample with the highest quality assessment scores in Table 6 [34,35,37], we conducted a series of exploratory subgroup meta-analyses for each type of HT intervention that had at least 2 studies. More specifically, we identified all RCTs of both recently discharged and stable HF patients included in the systematic reviews (Figure 2), extracted all relevant data by outcome (ie, number of events and patients in each group), cross-checked and validated all data for accuracy, and then grouped these studies according to the specific type or category of HT technology used. All subgroup analyses were restricted to RCTs that compared HT without home visits to usual care. Meta-analyses were performed using risk ratios, intention-to-treat analysis, the Mantel-Haenszel statistical method, and a random-effects analysis model to account for functional differences between interventions.

Four categories of HT interventions were identified among the 20 RCTs that were cross-referenced in the three systematic reviews. These included (1) automated device-based telemonitoring (TM and TM+), (2) mobile telemonitoring (MT), (3) automated interactive voice response (IVR), and (4) video-consultation with trans-telephonic vital signs monitoring (VC+). Table 10 presents the summary effect size for each type of HT intervention. Forest plots for each meta-analysis are presented in Multimedia Appendix 6. Wherever a meta-analysis could not be carried out due to insufficient number of studies, we present the results of the single trials instead.

Automated Device-Based Telemonitoring

Meta-analysis of 12 RCTs comparing device-based telemonitoring with usual care showed a statistically significant relative reduction of 35% in all-cause mortality (RR 0.65 [0.54-0.79], *P*<.001). Similarly, the relative risk of HF-related hospitalizations was reduced by 23% (RR 0.77 [0.64-0.91], *P*=.003). However, the number of studies reporting this outcome was significantly smaller (5 RCTs). The impact of automated device-based telemonitoring on all-cause hospitalization was also positive but statistically inconclusive (RR 0.89 [0.76-1.05], *P*=.17). Studies that involved telemonitoring of vital signs and symptoms through individualized symptom response systems (TM+) yielded slightly smaller effects with respect to all-cause mortality and all-cause hospitalizations than studies involving telemonitoring of vital signs only (TM). However, this difference was not statistically significant.

Mobile Telemonitoring

The effectiveness of mobile telemonitoring versus usual care was examined in 4 RCTs. All studies showed a beneficial trend in reducing all-cause mortality (RR 0.67 [0.35-1.26], *P*=.21). Similarly, the relative risk of HF-related hospitalizations was reduced with mobile telemonitoring by 28% (RR 0.72 [0.42-1.26], *P*=.25). However, results for both of these outcomes were statistically inconclusive owing to the small number and size of RCTs in this group.

Interactive Voice Response

The effectiveness of IVR was examined in only two RCTs. Contrary to other types of HT interventions, use of IVR systems was not associated with reductions in the relative risk of all-cause mortality, all-cause hospitalizations, and HF hospitalizations in either of these trials.

Table 10. Effects of HT according to the type of technology used.

Types of HT technologies	All-cause mortality	All-cause hospitalizations	HF hospitalizations
Automated device-based telemonitoring (TM & TM+)	RR 0.65 [0.54-0.79], *P*<.001, I^2=0% (12 RCT)	RR 0.89 [0.76-1.05], *P*=.17, I^2=66% (10 RCT)	RR 0.77 [0.64-0.91], *P*=.003, I^2=25% (5 RCT)
Telemonitoring of vital signs (TM)	RR 0.64 [0.51-0.80], *P*<.001, I^2=0% (6 RCT)	RR 0.81 [0.64-1.03], *P*=.08, I^2=76% (6 RCT)	RR 0.73 [0.58-0.91], *P*=.005, I^2=32% (4 RCT)
Telemonitoring of vital signs and symptoms (TM+)	RR 0.70 [0.47-1.04], *P*=.08, I^2=12% (6 RCT)	RR 1.04 [0.90-1.21], *P*=.58, I^2=0% (4 RCT)	RR 0.87 [0.66-1.13], *P*=.29[a]
Mobile telemonitoring (MT)	RR 0.67 [0.35-1.26], *P*=.21, I^2=44% (4 RCT)	RR 0.99 [0.76-1.29], *P*=.94, I^2=54% (3 RCT)	RR 0.72 [0.42-1.26], *P*=.25, I^2=48% (2 RCT)
Interactive voice response (IVR)	RR 1.09 [0.57-2.07], *P*=.80 I^2=0% (2 RCT)	RR 1.18 [0.87-1.60], *P*=.29[a]	RR 1.03 [0.65-1.61], *P*=.91[a]
Video-consultation with vital signs monitoring (VC+)	RR 0.95 [0.35-2.53], *P*=.91, I^2=0% (2 RCT)	RR 1.06 [0.97-1.16], *P*=.22[a]	No studies available
All types of HT combined (TM, TM+, MT, IVR, VC+	RR 0.73 [0.62-0.85], *P*<.001, I^2=0% (20 RCT)	RR 0.95 [0.85-1.06], *P*=.38, I^2=66% (15 RCT)	RR 0.79 [0.69-0.91], *P*=.001, I^2=19% (RCT=8)

[a] Meta-analysis could not be performed. Only 1 RCT provided data.

Videoconferencing With Vital Signs Monitoring

Only two RCTs investigated the effectiveness of videoconferencing with trans-telephonic monitoring of vital signs. No significant differences were found for all-cause mortality (RR 0.95 [0.35-2.53], *P*=.91) owing to the small number of patients and events. Neither of these studies reported results for the outcome of HF-related hospitalizations.

All Types of Home Telemonitoring Interventions Combined

As a final step, we pooled data from all the RCTs in an effort to further explore the effectiveness of HT, by combining the findings of the systematic reviews with the highest methodological quality to increase statistical power and precision. As shown at the bottom of Table 10, when considered collectively, HT interventions without home visits are associated with a statistically significant, relative risk reduction of all-cause mortality (RR 0.73 [0.62-0.85], *P*<.001) and HF-related hospitalizations (RR 0.79 [0.69-0.91], *P*=.001) of 27% and 21%, respectively. However, there was no significant reduction in the relative risk of all-cause hospitalizations (RR 0.95 [0.85-1.06], *P*=.38). Results pertaining to all-cause hospitalizations were also associated with high statistical heterogeneity (I^2 66%) due to differences in both direction and magnitude of effects between the included studies.

Discussion

Summary of the Evidence

This overview appraised and summarized evidence from 15 systematic reviews assessing the effects of HT interventions on patients with HF. To our knowledge, it is the first synthesis of systematic reviews to take a broad perspective on evidence-based telemonitoring in HF. It is also the first overview to investigate the nature of the link between different types of HT technologies and outcomes.

The systematic reviews included in our evidence synthesis covered a broad family of complex HT interventions rather than a standardized type of HT, involving various technologies and monitoring approaches that were supplemented with various other components in the context of comprehensive care programs (eg, clinical advice via telephone, patient education, and in some cases home visits). Despite ostensibly being reviews of the same body of literature as their research objectives suggest, we identified several key differences between them with respect to the scope of inquiry, study selection criteria, classification, and analysis of HT studies.

To best organize and synthesize the evidence, we performed a citation analysis and developed a taxonomic structure to categorize the included reviews into homogeneous groups according to their common elements and PICO characteristics. Subsequently, we appraised the methodological quality of the reviews and constructed summary of findings tables to present the effects of HT interventions from the most direct evidence, that is, from reviews that achieved the highest methodological quality score in each classification group. Limitations in the quality of evidence were formally reflected in the summary of findings tables by outcome, using the evidence grading system developed by the GRADE group [17] and in the analysis by interpreting results and formulating statements about the effectiveness of HT in light of the risk of bias in the primary studies. We also conducted a series of post-hoc analyses to develop a preliminary taxonomy of HT technologies and then investigate the link between these technologies and HT effectiveness.

Looking both collectively and individually across the included systematic reviews, this overview demonstrates that there is no high-quality evidence for or against the effectiveness of HT interventions for HF patients. There is moderate quality evidence that HT interventions with clinical support provided during office hours or 24/7 reduce the risk of all-cause mortality and HF-related hospitalizations compared to usual care. Yet the bulk of the literature consists of low-quality and inconsistent evidence about the beneficial effects of HT on all-cause hospitalizations [34,37]. Risk reductions in mortality and all-cause hospitalizations appear to be greater in patients who have been recently discharged from an acute care setting after an HF exacerbation and are at high risk of re-hospitalization or sudden death, while improvements in HF-related hospitalizations appear to be more pronounced with 24/7 HT on patients with stable HF. However, these results should be interpreted with caution and be considered as hypothesis-generating in future trials and systematic reviews, given the large uncertainty (imprecision) in the estimates of effect. Evidence about cost-effectiveness remains limited, and there are no reliable data on the long-term benefits and economic implications of HT interventions [51]. Despite current indications in the literature that HT can generate cost savings for health care providers and national health care systems [42], the economic evidence base is still weak and fails to meet generally accepted standards of economic analysis. With respect to the effects of HT on hospital length of stay and quality of life, there is no consistent evidence from which to draw robust conclusions.

The results of the exploratory post-hoc analyses we conducted show that the majority of interventions included in the 15 systematic reviews (62% of RCTs and 55% of observational studies) involved the use of non-invasive, patient-initiated electronic devices and/or interactive response systems capable of measuring and automatically transmitting vital signs, physiological data, and/or symptoms from the patient at home to the health care professionals providing care and clinical feedback. Other less frequent types of HT interventions included in the systematic reviews involved the use of video-consultation equipment (18% of RCTs and 13% of observational studies), mobile telemonitoring through mobile phones and PDAs (10% of RCTs), and automated interactive voice response systems requiring manual data entry by the patient (8% of RCTs and 7% of observational studies). Therefore, it can be argued that the results of the systematic reviews included in this overview reflect for the most part the effectiveness of “automated device-based HT interventions” and, to a lesser extent, the effects of interventions involving other technologies and monitoring approaches. In fact, the effects of the other types of HT technologies identified in our analysis (eg, videoconferencing, mobile telemonitoring and interactive voice response) have largely been masked in prior systematic reviews due to the fact that virtually all of them have treated HT as a “single-type intervention”. This is further supported by the results of the exploratory meta-analyses we conducted, which show that not all types of HT technologies are equally effective. Yet, when pooled together into one large group of HT interventions, the category with the most trials that has the largest impact on the results (ie, automated device-based telemonitoring) masks valuable insights about the effects of the other interventions. For example, mobile telemonitoring, which has emerged as a relatively new approach due to the ubiquitous nature of mobile devices and cell phones despite the small number of available studies, is associated with beneficial trends showing promise in reducing mortality and HF-related hospitalizations. None of the prior systematic reviews included in our study identified or commented on this. On the other hand, interventions using interactive voice response systems and video-consultations were not associated with beneficial effects on all-cause mortality and hospitalizations. However, when trials from these distinctively different interventions were pooled together with the dominant group of studies (automated device-based telemonitoring), the relative risk reduction of all-cause mortality and HF-related hospitalizations remained statistically significant. This indicates that the “one size fits all” approach that has been used so far in prior systematic reviews and meta-analyses in the field of HT may not be appropriate.

Systematic Reviews Published Since Completion of the Main Search

On December 3, 2014, we re-ran our search strategy to identify new systematic reviews published after our main search was completed. We identified 3 reviews [52-54], two of which contained meta-analysis for at least one outcome of interest. One review examined the effects of HT interventions on patients with HF and conducted several subgroup meta-analyses of RCTs containing 40 or more patients to determine which HT model is more effective and for which patient population [52]. The other two reviews assessed the effectiveness of several other “disease management” and “transitional care interventions” (eg, structured telephone support, home visiting programs, cognitive training, and invasive telemonitoring interventions) in addition to non-invasive HT. Outcomes were analyzed and reported separately for each intervention and hence, both reviews were deemed eligible for inclusion. The methodological quality of the two reviews [52,54] was found to be low (AMSTAR=2), contrary to the third review [53], which was conducted for the Agency for Healthcare Research and Quality [55] and met most of the AMSTAR criteria achieving a score of 9. All-cause and HF-related hospitalizations were reported in 2 of the 3 reviews [53,54], while findings pertaining to all-cause mortality were reported in all three.

The systematic review with the highest AMSTAR score, contrary to the other reviews included in our main analysis, concluded that HT interventions are not effective in reducing the overall risk of all-cause mortality, all-cause hospitalizations, and HF-related hospitalizations over a period of 6 months compared to usual care: *all-cause mortality at 3-6 months*: RR 0.93 (95% CI 0.25-3.48) 3 RCTs, 564 patients, “low quality of evidence”; *all-cause hospitalizations at 30 days*: RR 1.02 (95% CI 0.64-1.63) 1 RCT, 168 patients, “insufficient quality of evidence”; *all-cause hospitalizations at 3-6 months*: RR 1.11 (95% CI 0.87-1.42) 3 RCTs, 434 patients, “moderate quality of evidence”; and *HF-related hospitalizations at 3-6 months*: RR 1.70 (95% CI 0.82-3.51) 1 RCT, 182 patients, “moderate quality of evidence” [53,55]. However, this review differed in scope from all previously published systematic reviews in that it included only RCTs of adult patients recruited during or within only 1 week of an index hospitalization for HF. Also, the required timing of outcome measurement had to occur no more than 6 months from the index hospitalizations in order for RCTs to be eligible for inclusion. The use of such narrow scope and eligibility criteria limits the applicability and external validity (generalizability) of this review, the results of which should be interpreted with caution as they rely on a very small number of RCTs (≤3 per outcome), insufficient for drawing meaningful conclusions about the effectiveness of HT interventions on recently discharged (≤7 days) patients with HF.

In the second systematic review that contained meta-analysis, Nakamura et al [52] sought to investigate which HT model is more effective in reducing all-cause mortality in patients with HF. In this line of thought, they conducted a series of subgroup analyses across 13 RCTs (3337 patients) by age, severity of illness, measurement frequency, medication management, and speed of intervention. According to the findings of this review, studies in which clinical intervention was performed within one day of a change in the patient's vital signs (termed by the authors as "rapid intervention") had statistically significantly lower mortality rates compared to the group of studies in which clinical intervention took place later (RR 0.59 vs 0.88, *P*=.05). Also, the risk for all-cause mortality was found to be lower in HT studies that (1) had high measurement frequency of vital signs (more than twice a week vs ≤ once a week: RR 0.62 vs 0.89, *P*=.07), (2) included patients with a mean age of 65 years or over (RR 0.63 vs 0.71, *P*=.60), (3) had 70% or more of patients classified under NYHA class III or IV (RR 0.63 vs 0.86, *P*=.13), and (4) included a medication management component (RR 0.65 vs 0.85, *P*=.19) [52]. However, it is important to note that these findings are observational in nature and suffer from important limitations [56,57], including possible bias introduced through confounding by other study-level characteristics; misclassification of certain RCTs providing insufficient or no information at all for some categories; and arbitrary selection of cut-off points without any supporting evidence from sources other than the included RCTs, suggesting possible data dredging. Indeed, it is difficult to explain or justify the authors' motivation for the selection of the cut-off points used to classify studies into subgroups. Also, it is not possible to discern which of the investigated characteristics explain, and to what extent, the observed differences in the magnitude of effects (quantitative interaction) between the included RCTs, when several studies involving frequent measurement of vital signs and "rapid intervention" by clinicians, also included older patients with more severe HF (stages III and IV). In light of these and many more limitations associated with the nature of these observational investigations, the findings of this systematic review should be interpreted with extreme caution and at best be considered as hypothesis-generating rather than hypothesis testing.

The third systematic review identified by our recent search included 14 RCTs (5021 patients), within the broader scope of HF disease management programs, evaluating the efficacy of non-invasive HT support [54]. Using vote counting by statistical significance as the main method of analysis, the authors of this review found that only 2 RCTs demonstrated a significantly positive effect on all-cause mortality, and only 3 RCTs significantly reduced all-cause and HF-related hospitalizations. Therefore, it was concluded that current evidence supporting the efficacy of HF disease management programs (including non-invasive HT interventions) demonstrates highly inconsistent results, and therefore one approach applied to a broad spectrum of different patient types may not be effective. However, it should be borne in mind that vote counting by statistical significance is inadequate to answering the question of whether there is any evidence of an effect [56]. Furthermore, vote counting has a notorious record for being misleading (p. 252 [58]), as in the case of this review where many of the included RCTs were not sufficiently powered to reach statistically significant results, leading the authors to the perception that these studies yielded "conflicting results", although the treatment effects in these RCTs were actually similar or even larger than the ones in the studies that were statistically significant.

Our search also identified 2 recent publications that conducted post-hoc subgroup analysis of the results contained in the Cochrane systematic review included in our main results. The main objectives of these studies was to determine whether age is a factor in the success or failure of remote monitoring interventions in HF (including HT) [59] and the extent to which technological differences have an impact on the primary outcomes of interest [60]. Similarly to this overview, the study by Conway et al [60] underscores the important need to characterize HT interventions according to the technology component in use and to investigate the link between technology and HT effectiveness. However, in contrast to Conway et al [60], this overview relied on 15 systematic reviews and 105 studies to derive a preliminary taxonomy of HT technologies. Therefore it provided a richer and more comprehensive classification of telemonitoring technologies and incorporated twice as many RCTs in the meta-analysis. Briefly, the main inferences that were extracted from the two studies identified are as follows. Older people (≥70 years) with heart failure seem to benefit from HT interventions (all-cause mortality: RR 0.56 [95% CI 0.41-0.76] 4 RCTs; all-cause hospitalizations: RR 0.89 [95% CI 0.80-1.00] 3 RCTs), despite a popularly held belief of the opposite among clinicians [59]. Given the observational nature of this analysis, however, the authors stated that "discrimination by age alone may not be appropriate when inviting participation in a remote monitoring service for HF" [59]. Furthermore, evidence of systematic bias identified in the body of literature towards recruitment of individuals younger than the epidemiological average constitutes a significant problem that should be addressed in future RCTs, given the fact that HF becomes more prevalent as age is increased [59]. With respect to the impact of technological differences, the authors found that unlike other (broadly defined) HT technologies, interactive voice response systems requiring manual data entry by the patients may not be effective in reducing mortality and hospitalizations [60]. However, the number of studies included in the subgroup analysis was insufficient to draw definitive conclusions. Therefore, incorporation of new evidence in systematic reviews from recent RCTs is expected to provide further insights. This finding is consistent with the results of our post-hoc meta-analysis. Finally, Conway et al [60] argue that consideration should be given to measuring more than weight in telemonitoring interventions, as change in weight may not be sensitive enough to detect worsening of HF.

Overall Completeness and Applicability of the Evidence

While reductions in mortality and HF-related hospitalizations found in the systematic reviews included in our main analysis are particularly encouraging and HT as a research area has witnessed considerable growth over the years expanding its evidence base, there still remain important uncertainties around the general applicability and long-term efficacy of HT interventions due to several gaps and methodological weaknesses in prior research. First, most outcome data included in the systematic reviews are drawn from interventional studies that are clinically heterogeneous in terms of duration of follow-up, measures transmitted to the care providers (eg, weight, blood pressure, symptoms, and electrocardiogram), types of HT modalities used, frequency of data transmission, as well as diagnostic criteria used for the selection of patients with HF [37]. Second, the definition of usual care and the health services provided to patients in the control group also differed between primary studies in terms of intensity, clinical visits, patient education/training, or telephone support calls, depending on the country, area, and health care organization where the study was conducted and the model of care that was implemented [36]. Third, primary studies included in the systematic reviews were performed at different intervals over an extended period of time (12 years), during which both usual care and HT technologies have markedly evolved, witnessing important improvements. The impact of these temporal changes on the treatment effects, as well as the age and clinical/pragmatic differences between the primary studies may have been an important confounding factor in the observed results. As Gurne et al [61] note, some of the very first studies of HT included in the systematic reviews were conducted in the late 1990s (see Multimedia Appendix 3), when beta-blockers were not used as consistently as they are today in patients with HF. Also, the delivery of usual care for HF has improved over the last 15 years in many developed countries with the progressive introduction of multidisciplinary care, patient education, counseling services, home visits, and self-management programs led by specialist nurses—all of which have been shown to reduce mortality and hospitalizations [62]. The extent to which improvements in the conventional methods for delivering care may have minimized the gap between HT and standard care remains unclear. One of the frequently discussed challenges in the reviews was that in most primary studies the control group was not clearly described, compromising the reviewers' ability to understand the context the study was conducted in and how it might translate to other settings.

When interpreting the effects of HT interventions, besides the different types of technologies, it is also important to consider the technological advances that have occurred over the years (eg, in analytics, user-interfaces, and devices) and the different generations of HT technologies that have been developed. The sophistication of the technology, aside from changes in the models of care, is likely to have played an important role in outcomes. For example, as Anker et al maintain [8], first-generation HT systems, used in some of the early trials included in the reviews, were mainly "non-reactive data collection and analysis systems" that connected to external devices (eg, blood pressure and pulse monitors) utilizing conventional telephone lines to transfer physiological measures from the patient's home to a central server accessed by clinicians. Data transfer was generally asynchronous and the care providers could not respond instantaneously. Furthermore, these systems did not provide any patient advice, education, or automated feedback. Second-generation HT systems were more interactive from a patient perspective. They used approximately the same assessment measures (weight, heart rate, blood pressure, etc) but utilized patient medication reminders, educational components, as well as feedback mechanisms. They also involved additional and more sophisticated sensors for real-time transmission of vital signs and symptoms to the care providers [8]. Although delays in detection of patient deterioration and clinical intervention could potentially occur in cases where the systems were active only during office hours, it is likely that their impact on patient outcomes was more direct than that of first-generation systems [8]. Third-generation and fourth-generation HT technologies, which provide constant analytical and decision-making structures involving mobile phones, new sensors, as well as invasive and non-invasive devices that can measure heart, lung, and/or fluid retention more accurately, might deliver even greater health gains than HT systems of previous generations [63]. Despite recent attempts by researchers (eg, [37,44]), including attempts of this overview, to separate the effectiveness of different HT interventions and modalities and identify the type of patient population that benefits the most, there is still a lack of sufficient and high-quality studies to clearly indicate which types of HT technologies and strategies provide optimum clinical benefit, under what circumstances, and for which patient subgroup. The duration for which HT would continue to confer benefits also remains unclear. A frequently cited challenge, which we encountered too during the post-hoc analysis, is that most primary studies do not provide sufficient contextual information about the intervention and control group(s). Furthermore, results are presented in a manner that does not allow stratification of the benefits across strategies, stages of illness, and patient population [34,37]. Consequently, uncertainties remain around the determinants of successful HT programs. Subgroup differences ideally require individual-level data, and meta-analyses of individual-level data simply do not exist in the field of HT. An additional inhibiting factor that has been cited by several researchers [7,9,64] includes the lack of a commonly accepted taxonomy for classifying HT interventions into meaningful groups according to the technology in use and other key characteristics (eg, intensity and complexity, health care professionals involved in the delivery of clinical feedback, and response time). It is our hope that the taxonomy of HT technologies provided in this overview will serve as a valuable resource and also as a starting point for those that conduct systematic reviews and clinical trials in the area of HT. We strongly encourage researchers who start a systematic review to build on our classification scheme to explore the extent to which differences in the technologies used by HF patients have an impact on HT outcomes.

Quality of Evidence

In this overview, we identified and formally reflected in the summary of findings tables a number of serious limitations that we encountered during the appraisal of the primary studies and meta-analyses, which subsequently led us to rate down our confidence (quality of evidence) in the estimates of effect by outcome, following the methodological guidelines suggested by the GRADE group [17-27].

First, a high proportion of RCTs (>50%) included in the systematic reviews we examined did not provide sufficient details about random sequence generation, allocation concealment, attrition, and blinding of data collectors or outcome assessors, while in several studies there were significant differences in the baseline comparability of important prognostic factors [34,37]. As well, one third of the trials contributing data to the primary outcomes of interest received commercial funding from HT solution providers [37]. Receipt of such funding has been shown in other scientific fields to systematically bias the results in favor of the products made by the companies that fund the research [65]. Overall, as shown in Multimedia Appendix 5, most trials included in the reviews we examined contained important limitations in the design and/or execution. Therefore, quality of evidence was rated down by one level (from high to moderate) in all primary outcomes to reflect that most of the relevant evidence about the effectiveness of HT comes from studies with high or unclear risk of bias.

Second, given that the evidence base consists mainly of small trials that usually are not adequately powered to detect meaningful differences in outcomes, several meta-analyses included in this overview (eg, Comparisons 2, 3, and 4) did not meet the optimal information size criterion [22] required to establish a high level of confidence and therefore, lacked precision. The 95% credible intervals of the pooled effect crossed the line of “no effect” (1.0) and included appreciable benefit (HR<0.75) or harm (>1.25), or even both, suggesting that the effectiveness of HT in a randomly chosen study can vary substantially if the upper versus the lower boundary of the credible intervals represented the truth. Owing to the large uncertainty in the pooled estimates of effect, quality of evidence was rated down for imprecision [22].

Third, a high degree of statistical heterogeneity (eg, I^2>50%) in study results pertaining to all-cause hospitalizations was reported in many meta-analyses. Some trials included in the reviews found HT to be associated with substantially beneficial effects (RR 0.36), while others showed that HT increased the relative risk of all-cause hospitalizations versus usual care (RR 1.18). However, none of the reviews was able to identify potential effect modifiers that might explain the observed heterogeneity.

In short, risk of bias in the primary studies coupled with large and unexplained inconsistencies or imprecision in meta-analyses, inevitably decrease one’s confidence (quality of evidence) in the estimates of HT effects.

Potential Biases in the Overview Process

This overview adopted and applied rigorous methods suggested by the Cochrane Collaboration [11] with a view to minimizing the impact of bias arising from different sources within and across systematic reviews, as well as the overview process itself. Strengths of our approach include the use of sensitive and comprehensive search methods to identify all relevant reviews, the duplicated process applied in study selection, data extraction, and methodological quality appraisal, as well as the use of the GRADE system to rate the quality of evidence for each primary outcome.

Nonetheless, overviews are inevitably constrained by the quality and reporting characteristics of the systematic reviews, the quality of evidence within reviews, and the time lag between the publication of original studies and the reviews. Taking published systematic reviews as the sole evidence source and not searching for original trials that have not been identified by the included reviews increases the potential effect of publication lag and increases the chance that some evidence has not been considered in the review process [10]. However, the inclusion of a recent systematic review [37] that was both comprehensive and of high methodological quality mitigates this issue to a large extent.

Citation analyses performed in this overview indicated that there was overlap between the included reviews. Many interventional studies have contributed to multiple systematic reviews. Therefore, when interpreting the results of this overview, it is important not to treat the included systematic reviews as independent observations, but rather see them as a different way to address similar research questions to determine whether different review teams draw broadly similar conclusions about the effectiveness of HT for patients with chronic HF.

Implications for Policy and Practice

This overview provides a comprehensive analysis and synthesis that can be used as an evidence map to inform practitioners and policy makers about the effectiveness of HT interventions for patients with HF. Clinicians, health care policy makers, and clinical guideline developers, who rely on systematic reviews and RCTs to help them make informed choices among alternative interventions for the management of HF patients, may use the summary of findings tables of this overview as an entry point to HT evidence. Quality appraisal results can also be used to identify reviews and primary studies of high quality with minimal flaws in both their design and execution that can be trusted to support decision making or address specific questions and details not covered in this overview.

The positive findings associated with HT interventions may present useful resources for policy makers as they address timely issues involving the process and outcomes of care. Mortality, which represents an outcomes measure, has always been used as an indicator of performance and quality of care. Hospital readmissions related to HF also represent a popular indicator of the evolution of a patient’s condition, which is directly linked to the process of care. As such, it can be argued that there is evidence of quality improvement in patient care and a potential alleviation of pressure on hospitals in terms of patient hospitalizations and admissions related to HF conditions, which may otherwise free up places for other patients. Nonetheless, in light of the abovementioned limitations in the quality of evidence of prior research, we concur with Stroetmann et al

[66] that making the case for investment in HT applications at a national or international level, requires robust evaluations of the benefits and cost-effectiveness of HT applications under “routine conditions” in different contexts and settings toward the creation of a more convincing evidence base, not only to show that HT works, but also to show in what organizational context it works, for whom, and at what cost. Therefore, from a policy perspective, it is critical to take into consideration the findings of this overview and formulate appropriate policies and funding mechanisms that will support careful evaluations of the socioeconomic impacts of HT in real conditions, greater awareness and exchange of information between key stakeholders about the potential benefits of HT, opportunities for disseminating best-practices, and initiatives that bring policy responsibilities together to support better collaboration and coordination across sectors [66].

Health care decision makers and practitioners who are faced with implementing HT programs in community settings need to consider the complexity of these programs when interpreting the results of the systematic reviews. It is important to recognize that HT technologies are tools that facilitate early detection of deterioration signs. The key to the success of these programs is not the technology itself, but the coordination of care that needs to be in place along the continuum of health services delivered for HF patients within a health care system [62]. The effects of HT will most likely be better when the technology is used as part of a comprehensive and integrated care package that involves various multidisciplinary program components recommended by clinical guidelines [62], for example, patient education, appropriate pharmacological treatment, and psychological support. There is evidence suggesting that tailoring the interventions to those who have been recently discharged from the hospital due to HF exacerbation and are at high risk of sudden death or re-hospitalization may be beneficial to the effectiveness of the treatment strategy. However, health care decision makers should be cautious about implementing these approaches until further evidence accumulates and corroborates these findings.

Implications for Research

As shown in this overview, there exists a considerable body of evidence evaluating the effectiveness of HT interventions for patients with HF. Researchers conducting both primary studies and systematic reviews should consider the breadth of knowledge that has been created over the years and attempt to address existing gaps in order to inform future deployment and configuration of HT services for patients with HF. For example, new trials should select a small set of potentially mediating variables or risk factors highlighted in previous research studies (eg, HF severity, age, psychological support) and empirically test them within multifactorial designs or, alternatively, explore their impact on outcomes and publish results in meaningful ways as to allow stratification of the benefits of HT programs across subgroups of patients with HF [34]. Future research should also focus on carrying out direct comparisons between different HT technologies and delivery methods to elucidate whether there is differential effectiveness between HT strategies. Collection and reporting of rich contextual information pertaining to the features or components of HT interventions that contribute to variation in outcomes will facilitate a better understanding of the process by which HT works, improve the available evidence base, and maximize the meaningfulness of research findings.

The results from the post-hoc analysis conducted in this overview along with the recent findings of Conway et al [60] have significant implications for future research and provide important methodological insights that need to be considered in conducting future systematic reviews and meta-analyses evaluating the effects of HT interventions. Future systematic reviews should compare the effects of different HT technologies and interventions to provide specific insights on which approaches provide more effective management of HF patients. Development and use of a wide-ranging taxonomy that can adequately classify all types and aspects of HT interventions from the most comprehensive to those that are more simple and selective in what they offer can facilitate more robust comparisons and syntheses of results across studies and can enable interpretation of outcomes with reference to specific monitoring applications and components [64]. It is our hope that the preliminary taxonomy of HT technologies provided in this overview will serve as a valuable starting point toward accomplishing this goal.

Overall, there is a great need to shift our research focus from the basic evaluation question of “is HT effective?” to “what features or components of HT are effective, which patients benefit more from these interventions, under what circumstances, for how long, and why?” This shift requires use of multidisciplinary research designs and methodologies capable of untangling the often complex set of factors that may influence the effects of HT [64]. Realist reviews for instance [67], which attempt to provide an explanatory analysis of how and why complex interventions work (or not) in particular contexts [68,69], can help further advance our conceptual understanding about the impact of human behavior and interactions on the outcomes of telemonitoring interventions.

Finally, given our observation that 80% (12/15) of the systematic reviews assessed in this overview had moderate or major methodological limitations, researchers are strongly encouraged to closely adhere to the available methodological and reporting guidelines for systematic reviews [70-72] and consider the AMSTAR evaluation criteria [16] in order to improve the methodological rigor and reporting quality of their work. Similarly, at the primary study level, more carefully designed trials with longer observation periods, adequate power to detect differences in outcomes, and comprehensive economic evaluations are needed to provide conclusive answers on the effectiveness, cost-effectiveness, viability, and long-term impacts of HT interventions.

Conclusions

Overviews of systematic reviews use explicit research methods to collect and synthesize in a single source a comprehensive body of published evidence on the effectiveness of interventions. This overview identified and summarized available evidence from 15 systematic reviews on the effectiveness of HT interventions for patients with HF. It also conducted a post-hoc analysis to offer further insights into the various types of HT technologies included in the systematic reviews and investigate the link between HT technologies and HT effectiveness. The results from the principal analysis of this overview suggest that compared with usual care, HT interventions improve survival rates and reduce the risk of HF-related hospitalizations. Patients who have been recently discharged (≤28 days) from an acute care setting and are at high risk of re-hospitalization or sudden death appear to benefit more from HT programs compared to patients with stable HF, but this finding needs to be confirmed in large and rigorously designed RCTs. Overall, the favorable effects of HT reported in previous systematic reviews are based on moderate or low-quality evidence. The results of the post-hoc analyses suggest that only interventions involving automated device-based telemonitoring and mobile telemonitoring are effective in reducing the risk of all-cause mortality and HF-related hospitalizations. However, these findings should be interpreted with caution and be considered as hypothesis generating rather than hypothesis testing due to the exploratory nature of our investigation. More research data are required for interactive voice response systems, video-consultation, and Web-based telemonitoring to provide robust conclusions about their effectiveness. Future research should investigate further which HT strategies provide optimal outcomes, under what circumstances, and for which patient subgroup by adopting multidisciplinary methodologies capable of untangling the often complex set of factors that influence the effects of HT interventions.

Acknowledgments

We gratefully acknowledge the Canada Research Chairs Program for providing financial support for this research. We would also like to thank Nancy Santesso RD, MLIS, PhD, at McMaster University, for her generous assistance with the review methods, quality of evidence assessment, and application of the GRADE criteria.

Authors' Contributions

All 3 authors actively participated in the design and development of the review. SK and GP contributed to developing the protocol, review methods, search strategy, selection criteria, taxonomy of the systematic reviews and outcomes; performed the quality appraisals of the systematic reviews and post-hoc meta-analyses of the primary studies; and were responsible for writing the manuscript. MJ was involved in drafting the manuscript and revising it critically for important intellectual content. All authors approved the manuscript prior to its submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 100KB - jmir_v17i3e63_app1.pdf]

Multimedia Appendix 2

AMSTAR operationalization.

[PDF File (Adobe PDF File), 105KB - jmir_v17i3e63_app2.pdf]

Multimedia Appendix 3

Citation analysis and references of all primary studies included in the systematic reviews.

[PDF File (Adobe PDF File), 186KB - jmir_v17i3e63_app3.pdf]

Multimedia Appendix 4

Excluded articles and reasons for exclusion.

[PDF File (Adobe PDF File), 137KB - jmir_v17i3e63_app4.pdf]

Multimedia Appendix 5

Risk of bias assessment in original studies by outcome.

[PDF File (Adobe PDF File), 161KB - jmir_v17i3e63_app5.pdf]

Multimedia Appendix 6

Post-hoc meta-analyses.

[PDF File (Adobe PDF File), 225KB - jmir_v17i3e63_app6.pdf]

References

1. Bui AL, Horwich TB, Fonarow GC. Epidemiology and risk profile of heart failure. Nat Rev Cardiol 2011 Jan;8(1):30-41 [FREE Full text] [doi: 10.1038/nrcardio.2010.165] [Medline: 21060326]
2. Mosterd A, Hoes AW. Clinical epidemiology of heart failure. Heart 2007 Sep;93(9):1137-1146 [FREE Full text] [doi: 10.1136/hrt.2003.025270] [Medline: 17699180]
3. Liao L, Allen LA, Whellan DJ. Economic burden of heart failure in the elderly. Pharmacoeconomics 2008;26(6):447-462. [Medline: 18489197]
4. Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De SG, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Executive summary: heart disease and stroke statistics--2010 update: a report from the American Heart Association. Circulation 2010 Feb 23;121(7):948-954 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192666] [Medline: 20177011]
5. Lainscak M, Blue L, Clark AL, Dahlström U, Dickstein K, Ekman I, et al. Self-care management of heart failure: practical recommendations from the Patient Care Committee of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2011 Feb;13(2):115-126. [doi: 10.1093/eurjhf/hfq219] [Medline: 21148593]
6. Wootton R. Twenty years of telemedicine in chronic disease management--an evidence synthesis. J Telemed Telecare 2012 Jun;18(4):211-220 [FREE Full text] [doi: 10.1258/jtt.2012.120219] [Medline: 22674020]
7. Kitsiou S, Paré G, Jaana M. Systematic reviews and meta-analyses of home telemonitoring interventions for patients with chronic diseases: a critical assessment of their methodological quality. J Med Internet Res 2013;15(7):e150 [FREE Full text] [doi: 10.2196/jmir.2770] [Medline: 23880072]
8. Anker SD, Koehler F, Abraham WT. Telemedicine and remote management of patients with heart failure. Lancet 2011 Aug 20;378(9792):731-739. [doi: 10.1016/S0140-6736(11)61229-4] [Medline: 21856487]
9. Conway A, Inglis SC, Chang AM, Horton-Breshears M, Cleland JGF, Clark RA. Not all systematic reviews are systematic: a meta-review of the quality of systematic reviews for non-invasive remote monitoring in heart failure. J Telemed Telecare 2013 Sep;19(6):326-337. [doi: 10.1177/1357633X13503427] [Medline: 24163297]
10. Worswick J, Wayne SC, Bennett R, Fiander M, Mayhew A, Weir MC, et al. Improving quality of care for persons with diabetes: an overview of systematic reviews - what does the evidence tell us? Syst Rev 2013;2:26 [FREE Full text] [doi: 10.1186/2046-4053-2-26] [Medline: 23647654]
11. Becker LA, Oxman AD. Overviews of Reviews. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. Chichester, UK: John Wiley & Sons, Ltd; 2008:607-631.
12. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. BMC Med Res Methodol 2011;11(1):15 [FREE Full text] [doi: 10.1186/1471-2288-11-15] [Medline: 21291558]
13. Whitlock EP, Lin JS, Chou R, Shekelle P, Robinson KA. Using existing systematic reviews in complex systematic reviews. Ann Intern Med 2008 May 20;148(10):776-782. [Medline: 18490690]
14. US National Library of Medicine. Search strategy used to create the systematic reviews subset on PubMed URL URL: http://www.nlm.nih.gov/bsd/pubmed_subsets/sysreviews_strategy.html [accessed 2013-12-12] [WebCite Cache ID 6V1OUJdru]
15. Lee E, Dobbins M, Decorby K, McRae L, Tirilis D, Husson H. An optimal search filter for retrieving systematic reviews and meta-analyses. BMC Med Res Methodol 2012;12:51 [FREE Full text] [doi: 10.1186/1471-2288-12-51] [Medline: 22512835]
16. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10 [FREE Full text] [doi: 10.1186/1471-2288-7-10] [Medline: 17302989]
17. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011 Apr;64(4):383-394. [doi: 10.1016/j.jclinepi.2010.04.026] [Medline: 21195583]
18. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 2011 Apr;64(4):395-400. [doi: 10.1016/j.jclinepi.2010.09.012] [Medline: 21194891]
19. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011 Apr;64(4):401-406. [doi: 10.1016/j.jclinepi.2010.07.015] [Medline: 21208779]
20. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol 2011 Apr;64(4):407-415. [doi: 10.1016/j.jclinepi.2010.07.017] [Medline: 21247734]

21. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. J Clin Epidemiol 2011 Dec;64(12):1277-1282. [doi: 10.1016/j.jclinepi.2011.01.011] [Medline: 21802904]
22. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol 2011 Dec;64(12):1283-1293. [doi: 10.1016/j.jclinepi.2011.01.012] [Medline: 21839614]
23. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. J Clin Epidemiol 2011 Dec;64(12):1294-1302. [doi: 10.1016/j.jclinepi.2011.03.017] [Medline: 21803546]
24. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. J Clin Epidemiol 2011 Dec;64(12):1303-1310. [doi: 10.1016/j.jclinepi.2011.04.014] [Medline: 21802903]
25. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol 2011 Dec;64(12):1311-1316. [doi: 10.1016/j.jclinepi.2011.06.004] [Medline: 21802902]
26. Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, et al. GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. J Clin Epidemiol 2013 Feb;66(2):158-172. [doi: 10.1016/j.jclinepi.2012.01.012] [Medline: 22609141]
27. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ 2004 Jun 19;328(7454):1490 [FREE Full text] [doi: 10.1136/bmj.328.7454.1490] [Medline: 15205295]
28. Higgins JPT, Altman DG. Assessing Risk of Bias in Included Studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. Chichester, UK: John Wiley & Sons, Ltd; 2008:187-241.
29. Schünemann HJ, Oxman AD, Higgins JPT, Vist GE, Glasziou P, Guyatt GH. Presenting Results and ‘Summary of Findings’ Tables. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. Chichester, UK: John Wiley & Sons, Ltd; 2008:335-357.
30. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. CMAJ 1997 May 15;156(10):1411-1416 [FREE Full text] [Medline: 9164400]
31. Moja L, Fernandez del Rio MP, Banzi R, Cusi C, D'Amico R, Liberati A, et al. Multiple systematic reviews: methods for assessing discordances of results. Intern Emerg Med 2012 Dec;7(6):563-568. [doi: 10.1007/s11739-012-0846-1] [Medline: 22941412]
32. Martel G, Duhaime S, Barkun JS, Boushey RP, Ramsay CR, Fergusson DA. The quality of research synthesis in surgery: the case of laparoscopic surgery for colorectal cancer. Syst Rev 2012;1:14 [FREE Full text] [doi: 10.1186/2046-4053-1-14] [Medline: 22588035]
33. Clark RA, Inglis SC, McAlister FA, Cleland JGF, 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]
34. 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]
35. Polisena J, Tran K, Cimon K, Hutton B, McGill S, Palmer K, et al. Home telemonitoring for congestive heart failure: a systematic review and meta-analysis. J Telemed Telecare 2010;16(2):68-76. [doi: 10.1258/jtt.2009.090406] [Medline: 20008054]
36. Clarke M, Shah A, Sharma U. Systematic review of studies on telemonitoring of patients with congestive heart failure: a meta-analysis. J Telemed Telecare 2011;17(1):7-14. [doi: 10.1258/jtt.2010.100113] [Medline: 21097564]
37. Pandor A, Thokala P, Gomersall T, Baalbaki H, Stevens JW, Wang J, et al. Home telemonitoring or structured telephone support programmes after recent discharge in patients with heart failure: systematic review and economic evaluation. Health Technol Assess 2013 Aug;17(32):1-207, v [FREE Full text] [doi: 10.3310/hta17320] [Medline: 23927840]
38. Louis AA, Turner T, Gretton M, Baksh A, Cleland JJG. A systematic review of telemonitoring for the management of heart failure. Eur J Heart Fail 2003 Oct;5(5):583-590. [Medline: 14607195]
39. Martínez A, Everss E, Rojo-Alvarez JL, Figal DP, García-Alberola A. A systematic review of the literature on home monitoring for patients with heart failure. J Telemed Telecare 2006;12(5):234-241. [doi: 10.1258/135763306777889109] [Medline: 16848935]
40. Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. J Am Med Inform Assoc 2007;14(3):269-277 [FREE Full text] [doi: 10.1197/jamia.M2270] [Medline: 17329725]
41. 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]
42. Seto E. Cost comparison between telemonitoring and usual care of heart failure: a systematic review. Telemed J E Health 2008 Sep;14(7):679-686. [doi: 10.1089/tmj.2007.0114] [Medline: 18817497]

43. Dang S, Dimmick S, Kelkar G. Evaluating the evidence base for the use of home telehealth remote monitoring in elderly with heart failure. Telemed J E Health 2009 Oct;15(8):783-796. [doi: 10.1089/tmj.2009.0028] [Medline: 19831704]
44. Maric B, Kaan A, Ignaszewski A, Lear SA. A systematic review of telemonitoring technologies in heart failure. Eur J Heart Fail 2009 May;11(5):506-517. [doi: 10.1093/eurjhf/hfp036] [Medline: 19332417]
45. Paré G, Moqadem K, Pineau G, St-Hilaire C. Clinical effects of home telemonitoring in the context of diabetes, asthma, heart failure and hypertension: a systematic review. J Med Internet Res 2010;12(2):e21 [FREE Full text] [doi: 10.2196/jmir.1357] [Medline: 20554500]
46. Kraai IH, Luttik MLA, de Jong RM, Jaarsma T, Hillege HL. Heart failure patients monitored with telemedicine: patient satisfaction, a review of the literature. J Card Fail 2011 Aug;17(8):684-690. [doi: 10.1016/j.cardfail.2011.03.009] [Medline: 21807331]
47. Giamouzis G, Mastrogiannis D, Koutrakis K, Karayannis G, Parisis C, Rountas C, et al. Telemonitoring in chronic heart failure: a systematic review. Cardiol Res Pract 2012;2012:410820 [FREE Full text] [doi: 10.1155/2012/410820] [Medline: 22720184]
48. Inglis SC, Clark RA, McAlister FA, Stewart S, Cleland JGF. Which components of heart failure programmes are effective? A systematic review and meta-analysis of the outcomes of structured telephone support or telemonitoring as the primary component of chronic heart failure management in 8323 patients: Abridged Cochrane Review. Eur J Heart Fail 2011 Sep;13(9):1028-1040. [doi: 10.1093/eurjhf/hfr039] [Medline: 21733889]
49. Tran K, Polisena J, Coyle D, Coyle K, Kluge EHW, Cimon K, et al. Home telehealth for chronic disease management [Technology report number 113]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2008. URL: http://www.cadth.ca/en/products/health-technology-assessment/publication/865 [accessed 2014-02-20] [WebCite Cache ID 6WaW1n4zs]
50. Pandor A, Gomersall T, Stevens JW, Wang J, Al-Mohammad A, Bakhai A, et al. Remote monitoring after recent hospital discharge in patients with heart failure: a systematic review and network meta-analysis. Heart 2013 Dec;99(23):1717-1726. [doi: 10.1136/heartjnl-2013-303811] [Medline: 23680885]
51. Mistry H. Systematic review of studies of the cost-effectiveness of telemedicine and telecare. Changes in the economic evidence over twenty years. J Telemed Telecare 2012 Jan;18(1):1-6. [doi: 10.1258/jtt.2011.110505] [Medline: 22101609]
52. Nakamura N, Koga T, Iseki H. A meta-analysis of remote patient monitoring for chronic heart failure patients. J Telemed Telecare 2014 Jan;20(1):11-17. [doi: 10.1177/1357633X13517352] [Medline: 24352899]
53. Feltner C, Jones CD, Cené CW, Zheng ZJ, Sueta CA, Coker-Schwimmer EJL, et al. Transitional care interventions to prevent readmissions for persons with heart failure: a systematic review and meta-analysis. Ann Intern Med 2014 Jun 3;160(11):774-784. [doi: 10.7326/M14-0083] [Medline: 24862840]
54. Gorthi J, Hunter CB, Mooss AN, Alla VM, Hilleman DE. Reducing Heart Failure Hospital Readmissions: A Systematic Review of Disease Management Programs. Cardiol Res 2014 Jan;5(5):126-138. [doi: 10.14740/cr362w]
55. Feltner C, Jones CD, Cene CW, Zheng ZJ, Sueta CA, Coker-Schwimmer EJL, et al. Transitional Care Interventions To Prevent Readmissions for People With Heart Failure. Comparative Effectiveness Review No. 133. (Prepared by the Research Triangle Institute–University of North Carolina Evidence-based Practice Center under Contract No. 290-2012-00008-I). AHRQ Publication No. 14-EHC021-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014. URL: http://www.effectivehealthcare.ahrq.gov/ehc/products/510/1910/heart-failure-readmission-report-130527.pdf [accessed 2014-12-04] [WebCite Cache ID 6WaWeCBJ1]
56. Deeks JJ, Higgins J, Altman DG. Analysing Data and Undertaking Meta‐Analyses. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. Chichester, UK: John Wiley & Sons, Ltd; 2008:243-296.
57. Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Ann Intern Med 1992 Jan 1;116(1):78-84. [Medline: 1530753]
58. Borenstein M, Hedges L, Rothstein H. Introduction to meta-analysis. Oxford: Wiley; 2009.
59. Inglis SC, Conway A, Cleland JGF, Clark RA. Is age a factor in the success or failure of remote monitoring in heart failure? Telemonitoring and structured telephone support in elderly heart failure patients. Eur J Cardiovasc Nurs 2014 Mar 29:1-8. [doi: 10.1177/1474515114530611] [Medline: 24681423]
60. Conway A, Inglis SC, Clark RA. Effective technologies for noninvasive remote monitoring in heart failure. Telemed J E Health 2014 Jun;20(6):531-538. [doi: 10.1089/tmj.2013.0267] [Medline: 24731212]
61. Gurné O, Conraads V, Missault L, Mullens W, Vachierys JL, Van MW, Belgian Working Group on Heart Failure. A critical review on telemonitoring in heart failure. Acta Cardiol 2012 Aug;67(4):439-444. [Medline: 22997998]
62. McMurray JJV, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, Task Force for the DiagnosisTreatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology, ESC Committee for Practice Guidelines. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail 2012 Aug;14(8):803-869. [doi: 10.1093/eurjhf/hfs105] [Medline: 22828712]
63. ESC Press Office. Individualised treatment for heart failure is rarely available outside hospital - New telemonitoring systems could ‘radically’ change the situation.

http://www.escardio.org/about/press/press-releases/pr-09/Pages/Individualised-treatment-HF-hospital.aspx: ESC Press release Heart Failure Congress 2009; 2009 May 30. URL: http://www.escardio.org/about/press/press-releases/pr-09/Pages/Individualised-treatment-HF-hospital.aspx [accessed 2014-09-20] [WebCite Cache ID 6WaYcS3L3]

64. McLean S, Sheikh A, Cresswell K, Nurmatov U, Mukherjee M, Hemmi A, et al. The impact of telehealthcare on the quality and safety of care: a systematic overview. PLoS One 2013;8(8):e71238 [FREE Full text] [doi: 10.1371/journal.pone.0071238] [Medline: 23977001]
65. Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 2003 Aug 20;290(7):921-928. [doi: 10.1001/jama.290.7.921] [Medline: 12928469]
66. Stroetmann K, Kubitschke L, Robinson S, Stroetmann V, Cullen KD. How can telehealth help in the provision of integrated care?.: World Health Organization; 2010. URL: http://www.euro.who.int/__data/assets/pdf_file/0011/120998/E94265.pdf [accessed 2014-09-19] [WebCite Cache ID 6WaYzv88T]
67. Paré G, Trudel M, Jaana M, Kitsiou S. Synthesizing information systems knowledge: A typology of literature reviews. Information & Management 2015 Mar;52(2):183-199. [doi: 10.1016/j.im.2014.08.008]
68. Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, et al. Can we systematically review studies that evaluate complex interventions? PLoS Med 2009 Aug;6(8):e1000086 [FREE Full text] [doi: 10.1371/journal.pmed.1000086] [Medline: 19668360]
69. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy 2005 Jul;10 Suppl 1:21-34. [doi: 10.1258/1355819054308530] [Medline: 16053581]
70. Higgins JPG, Green S. Cochrane handbook for systematic reviews of interventions. Chichester, England: Wiley-Blackwell; 2008.
71. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
72. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]

Abbreviations

AMSTAR: Assessment of Multiple Systematic Reviews
ARR: absolute risk reduction
CI: confidence interval
Crl: credible interval
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
HF: heart failure
HR: hazard ratio
HT: home telemonitoring
PICO: population, intervention, comparison, outcomes
PRISMA: preferred reporting items for systematic reviews and meta-analyses
RCT: randomized controlled trial
RR: risk ratio

Edited by G Eysenbach; submitted 23.12.14; peer-reviewed by AL Vuorinen, G Hindricks; comments to author 16.01.15; accepted 07.02.15; published 12.03.15.

Please cite as:
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;17(3):e63
URL: http://www.jmir.org/2015/3/e63/
doi:10.2196/jmir.4174
PMID:25768664

Review

Information and Communication Technology–Enabled Person-Centered Care for the “Big Five” Chronic Conditions: Scoping Review

Sabine E Wildevuur[1,2], MSc; Lianne WL Simonse[3], MSc, PhD

[1]Talma Institute, Social Sciences, VU University Amsterdam, Amsterdam, Netherlands

[2]Institute for Art, Science and Technology, Creative Care Lab, Waag Society, Amsterdam, Netherlands

[3]Faculty of Industrial Design Engineering, Product Innovation Management Department, Delft University of Technology, Delft, Netherlands

Corresponding Author:
Sabine E Wildevuur, MSc
Talma Institute
Social Sciences
VU University Amsterdam
VU Amsterdam/FSW/SOC
De Boelelaan 1081
Amsterdam, 1081 HV
Netherlands
Phone: 31 205986888
Fax: 31 0205986800
Email: s.wildevuur@vu.nl

Abstract

Background: Person-centered information and communication technology (ICT) could encourage patients to take an active part in their health care and decision-making process, and make it possible for patients to interact directly with health care providers and services about their personal health concerns. Yet, little is known about which ICT interventions dedicated to person-centered care (PCC) and connected-care interactions have been studied, especially for shared care management of chronic diseases. The aim of this research is to investigate the extent, range, and nature of these research activities and identify research gaps in the evidence base of health studies regarding the “big 5” chronic diseases: diabetes mellitus, cardiovascular disease, chronic respiratory disease, cancer, and stroke.

Objective: The objective of this paper was to review the literature and to scope the field with respect to 2 questions: (1) which ICT interventions have been used to support patients and health care professionals in PCC management of the big 5 chronic diseases? and (2) what is the impact of these interventions, such as on health-related quality of life and cost efficiency?

Methods: This research adopted a scoping review method. Three electronic medical databases were accessed: PubMed, EMBASE, and Cochrane Library. The research reviewed studies published between January 1989 and December 2013. In 5 stages of systematic scanning and reviewing, relevant studies were identified, selected, and charted. Then we collated, summarized, and reported the results.

Results: From the initial 9380 search results, we identified 350 studies that qualified for inclusion: diabetes mellitus (n=103), cardiovascular disease (n=89), chronic respiratory disease (n=73), cancer (n=67), and stroke (n=18). Persons with one of these chronic conditions used ICT primarily for self-measurement of the body, when interacting with health care providers, with the highest rates of use seen in chronic respiratory (63%, 46/73) and cardiovascular (53%, 47/89) diseases. We found 60 relevant studies (17.1%, 60/350) on person-centered shared management ICT, primarily using telemedicine systems as personalized ICT. The highest impact measured related to the increase in empowerment (15.4%, 54/350). Health-related quality of life accounted for 8%. The highest impact connected to health professionals was an increase in clinical outcome (11.7%, 41/350). The impacts on organization outcomes were decrease in hospitalization (12.3%, 43/350) and increase of cost efficiency (10.9%, 38/350).

Conclusions: This scoping review outlined ICT-enabled PCC in chronic disease management. Persons with a chronic disease could benefit from an ICT-enabled PCC approach, but ICT-PCC also yields organizational paybacks. It could lead to an increase in health care usage, as reported in some studies. Few interventions could be regarded as “fully” addressing PCC. This review will be especially helpful to those deciding on areas where further development of research or implementation of ICT-enabled PCC may be warranted.

(J Med Internet Res 2015;17(3):e77) doi:10.2196/jmir.3687

KEYWORDS

patient-centered care; person-centered care; telemedicine; chronic disease; neoplasms; cardiovascular diseases; chronic respiratory tract diseases; diabetes mellitus; stroke; disease management; self-care; decision making; eHealth

Introduction

Information and communication technology (ICT) offers a means to support the self-management of chronic diseases and “empowerment” of patients, primarily through the Internet. Chronic diseases—also known as noncommunicable diseases—generally progress slowly over a long time. According to the World Health Organization (WHO), the “big 5” chronic diseases are diabetes mellitus, cardiovascular and chronic respiratory diseases, cancer, and stroke [1]. In Western society, chronic diseases make up the largest proportion of diseases and this is expected to grow further as a result of an aging society putting pressure on the sustainability of the health care system. By successfully adapting to a chronic illness and self-managing the disease, people are able to handle their life with some degree of independence despite their medical condition and are capable of participating in social activities including work and feel healthy despite their limitations [2,3]. For “connected” home health care during disease management, patients are expected to increasingly use eHealth services in codecision with their health care providers and thus play an active role in managing their own disease. eHealth offers a promising way to connect chronic patients and their health care providers, thereby ensuring that both chronic patients and health care providers are more involved in the long-term care needed for chronic diseases. In-depth research has been conducted to show that patients use health-related virtual communities and electronic support groups to keep themselves informed on treatment decisions and to manage their health [4,5]. However, these ICT applications focus on situations in which a health care professional is not necessarily engaged. Moreover, the impact of these types of ICT and person-centered care (PCC) interventions on (health-related) quality of life is unknown. In our study, we explored the extent to which ICT applications have been used to support self-management of 1 of the 5 chronic diseases—in situations where both a health care professional and patient are involved—and determine their impact.

Because health systems and services have become overly biometrics-oriented, disease-focused, technology-driven, and doctor-dominated, WHO advocates putting patients at the center of health care addressing PCC as a key dimension of health care quality [6]. The 21st century is envisaged as the century of PCC, especially in the care of the chronically ill [3]. The term PCC was initially used in the field of elderly care, where practitioners sought to provide better services to particularly frail and vulnerable people. Nowadays, more and more health care professionals, policy makers, and managers envision that patients could benefit from a person-centered approach to care in which the patient is no longer the passive target of a medical intervention, but is instead actively involved in his or her care [7].

Ekman [7] distinguished 3 routines of PCC activities:

1. Initiating the partnership: patient narratives;
2. Working the partnership: shared decision making; and
3. Safeguarding the partnership: documenting the narrative.

A narrative is defined as a spoken or written account of connected events. Modern medicine is generally disease-oriented and evidence-based; PCC starts with the person’s subjective experience of his or her illness and its impact on daily life [8]. Ekman et al [7] stated that the narrative is the starting point for building a collaborative, equalitarian partnership between the provider (care and treatment expert) and the patient (person expert) that encourages and empowers patients to actively take part in finding solutions to their problems.

Initial studies on PCC are promising and suggest that an implemented PCC approach shortens hospital stays and improves quality of care [9,10]. Given the growing interest in the topic of PCC [11], the term is slowly entering the scholarly discussions around ICT interventions. eHealth—supported by ICT—could encourage patients to take an active part in their health care and the decision-making process, “empower” them, and support a person-centered approach [7]. Connecting patients and health care professionals would not only improve the (technical) system of communicating, but also trigger social innovations of care models in which new ways of interacting and deciding improve the quality and efficiency of the organization [12]. However, we do not know to what extent ICT-enabled PCC exists, has been studied, and proved to be effective in terms of medical and organizational outcomes, such as cost efficiency. In this study, we focused on identifying the gaps in this field.

With the introduction of the Internet, Web-based technology has been applied to health-related ICT systems, with an initial focus on fields such as telemedicine and telemonitoring, and more recently in Medicine 2.0 approaches applying Web 2.0 technologies [13]. Telemedicine is defined in the Medical Subject Heading (MeSH) from 1993 as follows: “Telemedicine is the delivery of health services via remote telecommunications. This includes interactive consultative and diagnostic services.” Telemonitoring represents a patient management approach combining various information technologies for monitoring patients at a distance [14]. These advances have led to reviews on the specific technology of mHealth and eHealth [15,16]. With respect to PCC, scoping reviews on care management have been conducted within the fields of reproductive medicine and chronic heart failure [9,17]. However, no such studies on the combination of ICT and PCC management have been found.

We studied ICT interventions concerning the whole range of Internet technologies introduced since the inception of the Internet in 1989, from telemedicine to the new semantic and Web-based technologies of Health 2.0 and Medicine 2.0 technologies, including the recent evolution to smartphone communication with app technologies, and how these eHealth technologies are linked to PCC and with what impact. Through shared decision making, clinicians can help patients understand the importance of the information, measurements, and preferences in making the decisions that are best for them [18]. The innovations in ICT could support PCC routines, self-management and empowerment, and enable persons to codecide about their medical treatments [19].

Given the lack of a general overview of the extent and nature of published research involving the subset of ICT interventions in PCC for chronic conditions, the aim of this study was to contribute by exploring existing studies and to draw conclusions regarding the overall state of research activities and discover research gaps. The objective of this paper is to provide a review of the literature and to scope the field with respect to 2 research questions. Firstly, which ICT interventions have been used to support patients and health care professionals in PCC management of the big 5 chronic diseases? Secondly, what is the impact of these interventions, such as on health-related quality of life and cost efficiency?

This paper addresses the methods of the scoping review and its 5 different stages. Results provides overviews of the primary studies on PCC-ICT with participation of persons with a chronic condition of diabetes mellitus, cardiovascular and chronic respiratory diseases, cancer, or stroke. The state of knowledge in the health care domain is reported in terms of volume and nature and in relation to the outcomes reported. In Discussion, the review results are interpreted and compared with prior work. In addition, theoretical and practitioner implications of the study are described.

Methods

Scoping Review Study

Overview

We employed a rigorous literature review procedure by adopting the scoping review method. This is an appropriate method to systematically scan and evaluate which studies are within or out of the scope of the research area that is explored for evidence [20,21]. We considered other types of literature review methods, such as systematic review, meta-analysis, and structured reviews, which share similar activities such as the collection, evaluation, and presentation of available research evidence in a systematic manner. However, we chose to carry out a scoping review study because it best fit our research purpose with the emphasis placed on the scoping technique to "map" relevant literature in the field of interest rather than collecting similar evidence for a highly focused research question. The method is effective in identifying gaps in the evidence base where no research has been conducted and identifying emerging results in new fields of research; the methodological framework of Arksey and O'Malley was followed [22]. Five stages of scoping and review were carried out: (1) identify the research question, (2) identify relevant studies, (3) select studies, (4) chart the data, and (5) collate, summarize, and report the results.

Stage 1: Identifying the Research Question

The research question was conceived from people's high expectations regarding the potential impact of ICT innovations on heavily overburdened health care organizations, specifically in combination with an increase in self-management of diseases of long duration such as chronic diseases. We postulated that ICT could help persons with chronic conditions to interact directly with their health care providers about their personal health concerns and thereby empower them in the self-management of their personal health (information) and care plan. To search for evidence that might support our postulate, we formulated the following questions:

1. Which ICT interventions have been used to support chronic patients with the big 5 chronic conditions and their health care providers in PCC?
2. What is the impact of these interventions on health-related quality of life, and cost efficiency?
3. What other relevant study outcomes have been reported?

In the context of this paper, patients are defined as "individuals who are interacting directly with health care providers and services about personal health concerns" [23]. Starting from the point of view of ICT-enabled self-care and seeing "empowerment" as a possible outcome of applying this to the field of management of chronic diseases, we chose the definition of person-centered care coined by Ekman [7].

We consider eHealth as the use of ICT for health, as stated by the WHO initiative Global Observatory for eHealth. Our research builds on and contributes to the eHealth field, as defined by Eysenbach [22]: "eHealth is an emerging field in the intersection of medical informatics, public health, and business referring to health services and information delivered or enhanced through the Internet and related technologies. In a broader sense, the term characterizes not only a technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking to improve health care locally, regionally, and worldwide by using information and communication technology" [13].

For the purpose of our study, we initially defined PCC-ICT as a category of Internet technology that connects patients to health care professionals and enables them to interact and exchange information, including multimedia data such as audio (voice), video, and images. The PCC-ICT category covered different modes of Web communication including dedicated telemonitoring and/or telemedicine systems, Internet-based systems, telephone, and mobile phones. It excluded electronic patient record systems, public health information, and clinical and decision support systems for health care professionals only.

The outcome terms health-related quality of life and cost efficiency in the research question were only mentioned because they received considerable attention from health care managers and scholars, but in the scoping review were not restricted to these 2 outcomes. On the contrary, we conducted this scoping study to explore, extract, and describe all relevant outcomes used in the studies.

Stage 2: Identifying Relevant Studies

To identify original studies suitable for answering the research questions, we searched EMBASE, PubMed, and the Cochrane Library. To determine the relevant search words (keywords differ between databases), a medical information specialist devised an initial search strategy based on the research questions and definitions of key concepts, and on 10 seed articles [4,5,15,16,24-29]. The strategy was refined in the light of other published scoping review searches and other relevant sources [15,30,31]. A medical information specialist checked the final search syntaxes.

The search syntax was composed of "person-centered care," "ICT," and their synonyms, including the different types of spelling (US and UK), such as "person-centred care," "self-care," "self-management," "e-health," "Web 2.0," "decision support techniques," "videoconferencing," "cellular phone," "remote consultation," "user-computer interface," "Internet," and "telemedicine" combined with "chronic disease," "diabetes mellitus," "cardiovascular," and "chronic respiratory diseases," "cancer," and "stroke" and their synonyms.

Only those studies published between January 1989 and December 2013 were included. The start date of 1989 was chosen because the Internet went public in 1989. The end date was the last date on which we accessed the medical electronic databases. The search was limited to studies in English because of the costs and time involved in translating material in foreign languages such as French, German, Polish, Spanish, Russian, and Chinese. The search excluded letters, editorials, and news items. To manage the digital output from the search, we used EndNote software. The EndNote database comprised 9380 references with links to the digital portable document formats (PDFs) of the studies stored in the source database of the journals.

Stage 3: Study Selection

For the selection of studies, inclusion and exclusion criteria were developed and applied iteratively over 4 rounds of duplicate screening involving all authors as reviewers (Table 1).

First Review Step

In the first round, we screened titles and abstracts and excluded studies published before 1989, studies in which no ICT was involved, nonrelevant studies in which "mobile" was used in the sense of mobility (eg, mobile teams), nonrelevant studies focusing on mobile phones and the risk of brain damage as a result of mobile phone usage, studies on preventive care and public care involving screening and prevention activities, and studies on acute diseases (eg, acute stroke). Furthermore, studies focusing on children as the main target group were excluded because children do not manage their health on their own. Retained for inclusion were all articles addressing topics of direct relevance to the research questions: the big 5 chronic diseases, chronic care, PCC, ICT intervention, and an outcome measurement of some sort, including health-related quality of life and cost efficiency. In this round, the database was subdivided into the 5 chronic diseases.

Second Review Step

In the second round, the articles in each subdivision of chronic diseases in EndNote were reviewed based on title and abstract. The researchers determined whether the studies included connected-care communication of some sort that involved both the patients and their health care professionals, and ICT intervention (including telephone) to facilitate communication and interaction. In this round, we established separate folders of groups in EndNote based on our inclusion and exclusion criteria. In the separate folders, we excluded addressed applications for only health care professionals, community applications for online self-help groups, self-tests (diagnoses) for patients, and intramural health care settings. Even though we do consider that these applications are important, they do not meet our criterion that these applications should be part of an established relationship and collaboration between a patient and his or her health care professional.

Third Review Step

The third screen involved extracting the data by reviewing the abstracts and full text of the articles within each of the 5 databases on chronic diseases. The classification scheme for extracting the data addressed the categories: time (year of publication); origin (country); type of ICT intervention (mode of communication, data type, users); type of connected care (type of disease management, mode of PCC), and outcomes (person outcomes, health care professional outcomes, organization outcomes, and technical outcomes). General information such as gender, age, and background were not included since this scoping review study was intended to "map" relevant literature in the field of interest rather than collect evidence for a highly focused research question.

Additional criteria were developed iteratively to retain a set of articles. For example, telephones were identified as the first connection devices that made remote health care service between patient and health care provider possible. For exclusion, the criterion was developed on personal health records and other medical record applications for uses other than PCC-ICT self-management.

Table 1. Inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Collection of studies for the research data base	Publications after the invention of Internet (1978) onwards from 1989 when the first studies have been reported on e-health, in which Internet technology is applied in the health domain	Publications before 1989
	Publications in English language	Publications in other languages than English
	Published studies in EMBASE, PubMed and Cochrane Library.	Letters, editorials, news items and conference abstracts
First review step	Persons coping with one or more of the “big five” of chronic diseases	Persons coping with an acute diseases, such as acute stroke
	Chronic Care for persons already diagnosed with a chronic disease	Preventive Care and Public Care involving screening and prevention activities.
	Person centered self management and self care involved	Children (since they are taken care of by their parents)
	ICT involved	No ICT involved in the study
		Mobile in sense of mobility (mobile teams)
		The risk on brain damage through the use of cellular phone
	Medical study relating outcomes to ICT-intervention	Managerial study outcomes of for example cost estimation comparisons, or proposed strategies, care models etc
	Theoretical study outcomes such as frameworks	
	Study outcomes measuring Health related quality of life (HRQL) and Quality of Life (QoL)	
	Study outcomes measuring Cost efficiency	
	Study outcomes measuring other impact and performance factors	
	Documenting, monitoring and interaction applications for person-centered care	
Second review step	Connected care communication: multiple target groups as users of the application	One target group of the Health care application
	Related to a person or patient	No patients mentioned or involved
	Minimal two users involved; a patient person with chronic condition and health care professional	Health care professional applications
		Patient community applications
	Home health care setting: care activities at home connected to care activities at other health care settings	Care only in a hospital or other intramural setting
		Self tests at home (for example, self-diagnosis)
Third review step	Telephone as device to connect patients with caregivers, in combination with remote health care services	
	Share personal health concerns	
	Manage own personal health information	Personal health records and other medical records applications for other usage
	Manage personal care plan	
	Virtual Reality for rehabilitation of stroke	
	Twitter	
	Skype	

Fourth Review Step

This step involved analyzing the systematic reviews to determine which studies had been carried out in an evidence-based manner, meeting the inclusion and exclusion criteria of the scoping review. Some of the studies that were described as “a systematic review” did not in fact meet the criteria set for a systematic review. However, these studies were not excluded because a scoping review does not assess the quality of the studies [22].

An additional exclusion criterion was that publications that did not include a full study (ie, ones that consisted merely of protocols or structured abstracts) were left out. Another additional inclusion criterion concerned technologies that were incidentally described within the searched databases. These included the usage of Skype, social media such as Twitter, or robot assistance for rehabilitation. We expect these applications—and others such as wearable devices—to be studied and described more frequently in the years to come and thus included them in the existing list of criteria.

The reviewers met a couple of times at every reviewing round to discuss the selection of studies and to refine the inclusion and exclusion criteria. The criteria were used in an iterative way, meaning that where necessary the reviewing procedure was repeated to ensure that the references were covered in a comprehensive way.

Stage 4: Charting the Data

For the critical fourth stage, we carefully crafted the classification schemes in such a way that ICT interventions used for PCC were classified into mutually exclusive and cumulatively exhaustive categories. This required a number of iterations in refinement and modification of the categories to ensure reliability of the study classification.

To answer the research questions, we created charts for:

1. Study context identification: time and geographical origin of the study;
2. Process intervention studied: types and modes of ICT intervention used for connected-care activities;
3. Per targeted population of patients with a chronic condition: the monitoring, documenting, and interacting devices per connected-care activity; and
4. Study outcome measures.

We coded ICT interventions into 4 categories: (1) telephone-based, (2) mobile phone–based, (3) Internet-based, and (4) dedicated telemonitoring/telemedicine system–based. Distinctions in applications for each of these types of hardware and software were made according to their data source (telephone, smartphone, Internet, or telemonitoring/telemedicine) and their primary function: documenting, interacting, and/or monitoring. In this stage, we compared the studies first within the chronic care management domain and then across the different chronic disease domains. We undertook this process manually using tabulation charts in Excel as visual aids. Each study was charted in a table per chronic care management activity and the ICT intervention for either activity to connect disease management and/or support person–centered activities. We determined patterns of commonalities and differences among the ICT interventions and care management activities.

Stage 5: Collating, Summarizing, and Reporting the Results

Having charted the information, we were able to numerically analyze the included studies. We then answered the research questions based on the analysis overviews. Through the systematic reporting and charting of the data, we were also able to make comparisons across ICT interventions, identify contradictory evidence regarding specific interventions, and identify research gaps in the existing research evidence.

Results

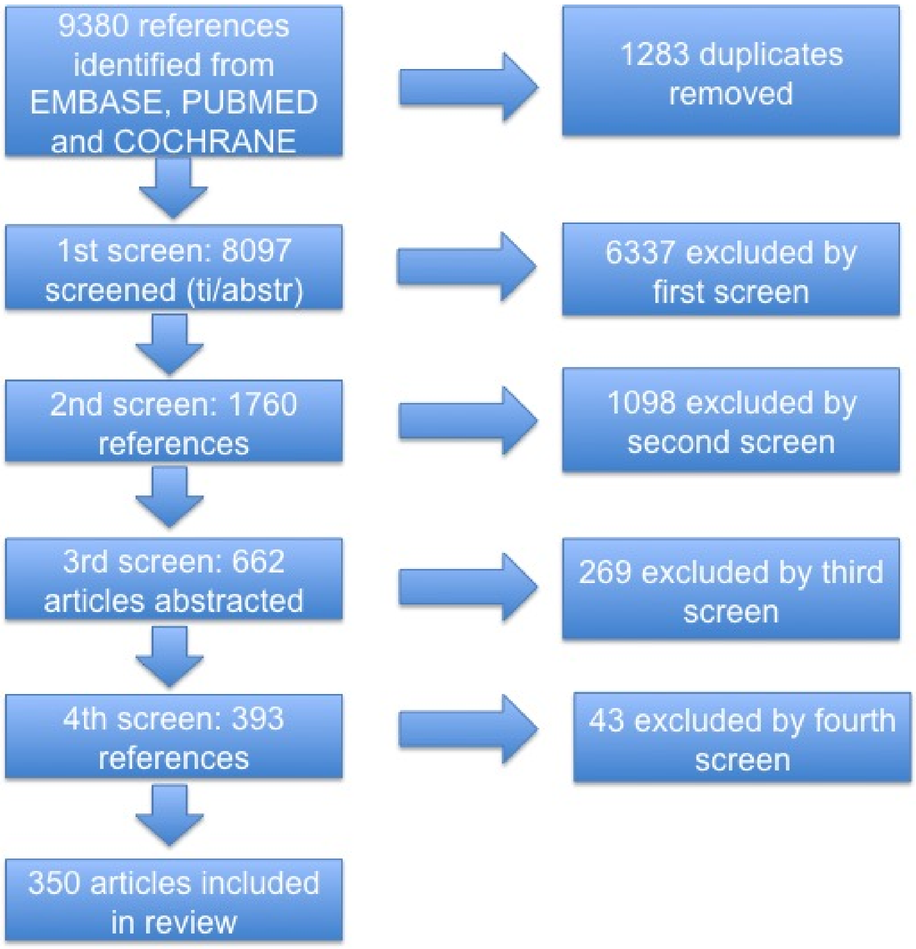
Overview

From the initial 9380 search results from EMBASE (n=6702), PubMed (n=1866), and the Cochrane Library (n=812), we identified 350 studies that qualified for inclusion (Figure 1). Classified according to the participation of persons with a chronic condition, the number of studies by condition was as follows: diabetes mellitus (n=103), cardiovascular disease (n=89), chronic respiratory disease (n=73), cancer (n=67), and stroke (n=18) (acute stroke was excluded) (Multimedia Appendix 1).

First, we excluded the duplicates (n=1283). In the first screening round, we excluded both the ones published before 1989 and nonrelevant studies (n=6337) according to the exclusion criteria we developed in this round (see Figure 1), leaving 1760 studies.

In the second screening round, 1098 articles appeared not to address the research scope in terms of the inclusion criteria. In the third screening round, 269 were excluded, leading to 393 articles. Finally, in the fourth screening round, we charted the systematic reviews from the separate EndNote folder from the 105 systematic reviews; 62 were eventually included and 43 excluded. We were left with 350 studies.

Figure 1. Search and screening results.



Study Characteristics

In characterizing the included studies by origin, 40 countries are represented: Europe (147/350, 42.0%), North America (138/350, 39.4%), Pacific region (39/350, 11.1%), Asia (20/350, 5.7%), Middle East (8/350, 2.3%), and Latin America (3/350, 0.9%) (see Table 2).

Figure 2 shows that almost 10 years after the starting point of our search strategy (1989), the number of studies published annually increased until 2013. Within the domain of cancer, the first studies were identified on connected care in 1997. Around 2005, attention to connected care seemed to rise, with an even more substantial increase in publication volume from 2010 onward. The trend lines differed for the big 5 chronic conditions, with chronic respiratory conditions showing the steadiest increase and the others more fluctuation; cardiovascular conditions showed a slight decrease in recent years.

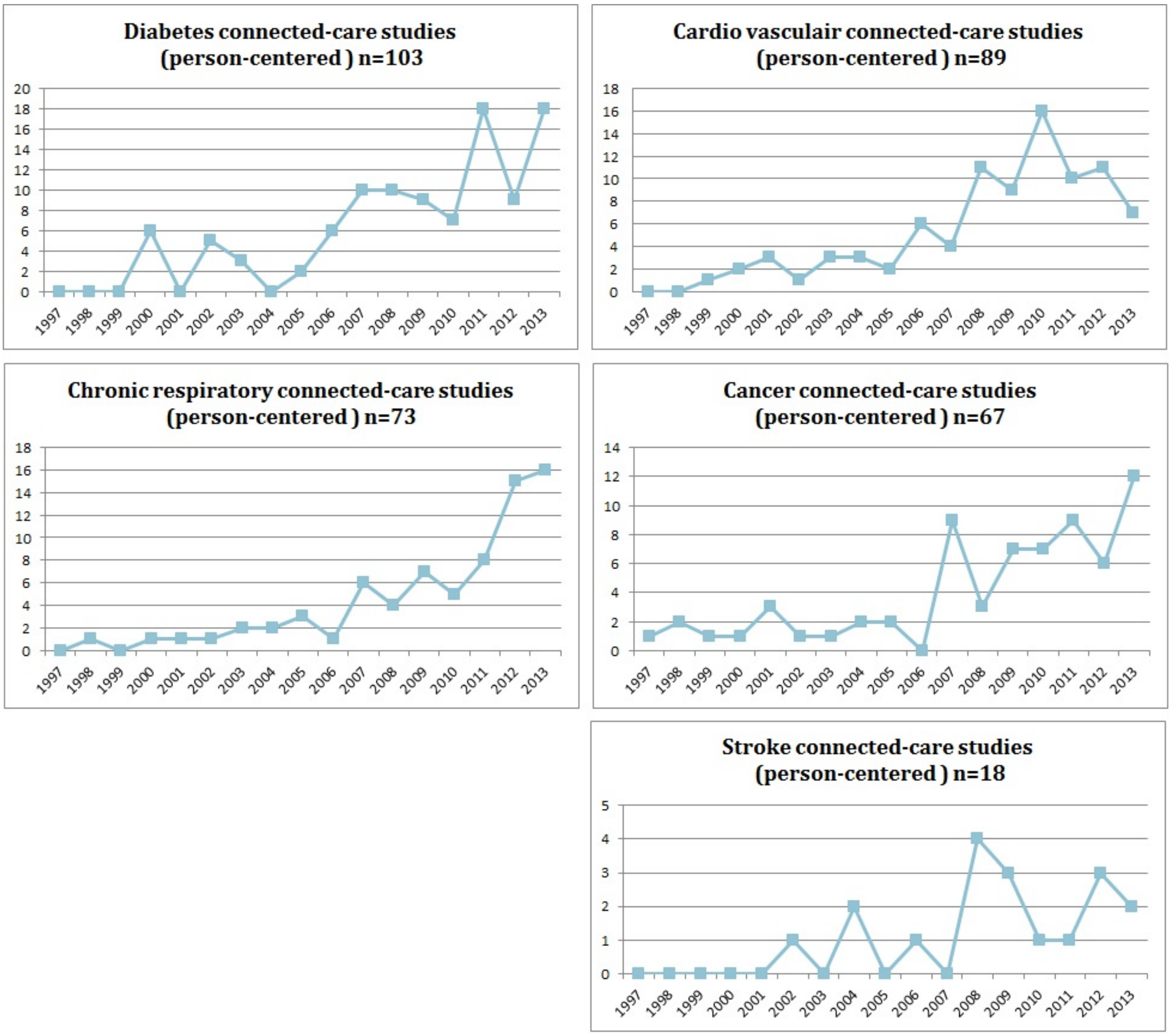
Table 2. Representation of the studies according to geographical location.[a]

Continent and country	Overall, n (%)	Chronic disease, n (%)				
		Diabetes	Cardiovascular	Chronic respiratory	Cancer	Stroke
	N=350	n=103	n=89	n=73	n=67	n=18
Europe						
Netherlands	24 (6.9)	5 (1.4)	5 (1)	10 (3)	4 (1)	
Belgium	1 (0.3)		1 (0)			
United Kingdom	47 (13.4)	10 (2.9)	11 (3)	14 (4)	12 (3)	
Germany	17 (4.9)	5 (1.4)	10 (3)	1 (0)		1 (0)
France	3 (0.9)	1 (0.3)	1 (0)		1 (0)	
Italy	21 (6.0)	5 (1.4)	7 (2)	4 (1)	1 (0)	4 (1)
Poland	6 (1.7)	3 (0.9)	3 (1)			
Austria	1 (0.3)	1 (0.3)				
Switzerland	2 (0.6)		1 (0)			1 (0)
Spain	8 (2.3)	1 (0.3)	3 (1)	3 (1)	1 (0)	
Denmark	5 (1.4)			5 (1)		
Finland	3 (0.9)	1 (0.3)			1 (0)	1 (0)
Sweden	3 (0.9)				3 (1)	
Portugal	1 (0.3)			1 (0)		
Cyprus	1 (0.3)		1 (0)			
Norway	4 (1.1)		2 (1)		2 (1)	
Total	147 (42.0)	32 (9.1)	45 (12)	38 (11)	25 (7)	7 (2)
North America						
United States	116 (33.1)	48 (13.7)	23 (7)	13 (4)	29 (8)	3 (1)
Canada	22 (6.3)	2 (0.6)	9 (3)	7 (2)	3 (1)	1 (0)
Total	138 (39.4)	50 (14.3)	32 (9)	20 (6)	32 (9)	4 (1)
Asia						
China	4 (1.1)	2 (0.6)		2 (1)		
Taiwan	5 (1.4)	2 (0.6)		2 (1)		1 (0)
Korea	4 (1.1)	3 (0.9)			1 (0)	
Thailand	1 (0.3)	1 (0.3)				
Hong Kong	3 (0.9)			1 (0)		2 (1)
Japan	1 (0.3)		1 (0)			
India	1 (0.3)		1 (0)			
Sri Lanka	1 (0.3)	1 (0.3)				
Total	20 (5.7)	9 (2.6)	2 (1)	5 (1)	1 (0)	3 (1)
Middle East						
Saudi Arabia	2 (0.6)	1 (0.3)	1 (0)			
Israel	4 (1.1)	1 (0.3)	3 (1)			
Iran	2 (0.6)	2 (0.6)				
Total	8 (2.3)	4 (1.1)	4 (1)			
Latin America						
Chile	1 (0.3)	1 (0.3)				
Brazil	1 (0.3)	1 (0.3)				

Continent and country	Overall, n (%)	Chronic disease, n (%)				
		Diabetes	Cardiovascular	Chronic respiratory	Cancer	Stroke
	N=350	n=103	n=89	n=73	n=67	n=18
Panama	1 (0.3)	1 (0.3)				
Total	3 (0.9)	3 (0.9)				
Pacific						
Australia	37 (10.6)	5 (1.4)	9 (3)	10 (3)	9 (3)	4 (1)
New Zealand	2 (0.6)		1 (0)	1 (0)		
Total	39 (11.1)	5 (1.4)	10 (3)	11 (3)	9 (3)	4 (1)

[a] More than 1 country possible due to consortia (n=5). Percentages estimated by total number of studies (N=350).

Figure 2. Number of studies conducted over the years for the “big five” chronic conditions.



Information and Communication Technology Enabling Person-Centered Care for Diabetes Mellitus

From the total number of relevant studies concerning PCC in which patients with a diabetes condition were central (n=66), ICT applied for self-body measurement (eg, with a glucose meter device and monitoring system) was the most used (48/66, 47%) (Table 3).

The second in line was personal lifestyle sharing (16/103, 15.5%), which is distinctively related to diabetes care. ICT for shared treatment decisions ranks low, with 2 studies found (2/103, 1.9%). In broader terms, connected care for diabetes, the most studied ICT intervention, addressed education with transfer of diabetes knowledge from the care provider to the patient (26/103, 25.2%).

The most commonly used ICT intervention employed telemonitoring/telemedicine systems (42/103, 40.8%). Internet interventions ranked second (29/103, 28.2%), whereas mobile phone interventions accounted for 24.3% (25/103) and telephone interventions for 16.5% (17/103). Interestingly, text messaging was used in 9.7% (10/103) of the studies. Glucose monitoring devices plus systems ranked third (13/103, 12.6%). These and other modes of telemedicine systems were mostly used (21/103, 20.2%) for the PCC activity of self-body measurement versus 9% used for physical care with measurement by the physician. Within the category of mobile phone interventions, 9% used a monitoring app. A personal e-diary app is worth mentioning as a personal means of sharing health information with the care provider. Within the category of Internet interventions, most of the interventions concerned monitoring, but were integrated with interaction via the Web app (6%). Another type of Web-based app used more than once was a documenting Web app that was mostly employed for sharing lifestyle information with the health care professional. The last category of charted ICT intervention studies indicated a high research interest (17%) in low-tech technology interventions via telephone, the first connecting device enabling remote care management by means of follow-up telephone calls made by nurses.

Information and Communication Technology Enabling Person-Centered Care for Chronic Cardiovascular Diseases

The studies in which cardiovascular patients participated showed a clear preference (71%, 63/89) for telemonitoring/telemedicine system interventions applied for PCC, self-measurement of the body (38%, 34/89), versus interventions applied for connected physical care (17%), and education (9%) as a way to connect the patient and the health care professionals providing advice and service (see Table 4). In almost one-quarter of the studies (23%, 21/89), telephone interventions—in particular nurse telephone calls (21%, 19/89)—were used.

Of the 89 studies in the scope of cardiovascular conditions, the most studied person-centered care activity was self-measurement of the body (38%, 34/89). We discovered 3 studies addressing self-rehabilitation exercises by a virtual clinic app on the Internet and a telemonitoring system. A telemedicine system was used for shared treatment decision making in only 1% (1/89) of the studies.

Another 8% (7/89) of the studies dealt with the use of remote monitoring and a cardiac implant device for self-measurement of the body. These are remote monitoring apps for implanted cardiac pacing systems, which enable persons and health care professionals to (self) monitor the heart implant and are used specifically for cardiovascular patients. Mobile phone interventions were used in 8% (7/89) of the studies; in 7% (6/89) of the studies, monitoring apps were used to self-measure the body, whereas 1% (1/89) were used for educational purposes. We also found combinations of interventions (eg, telephone support by nurses for the educational part and a monitoring device with a Web app for self-monitoring and sharing the data with health professionals).

Table 3. Information and communication technology intervention used for person-centered care for diabetes (n=103).[a]

ICT Intervention	Overall, n (%)	Connected-care activity, n (%)				Person-centered care activity, n (%)			
		Consult	Physical care	Behavior therapy	Education	Self-measurement	Lifestyle sharing	Shared decisions	
	n=103	n=3	n=8	n=10	n=26	n=48	n=16	n=2	
Telephone intervention	17 (16.5)		1 (1)	2 (2)	4 (4)	6 (6)	4 (4)		
Nurse telephone calls	8 (7.8)		1 (1)	2 (2)		2 (2)	3 (3)		
Automated telephone calls	8 (7.8)				4 (4)	4 (4)			
Nurse call center	1 (1.0)						1 (1)		
Mobile phone intervention	25 (24.3)	1 (1)		2 (2)	6 (6)	11 (11)	5 (5)		
Smartphone calls	1 (1.0)					1 (1)			
Text messaging	10 (9.7)			1 (1)	4 (4)	2 (2)	3 (3)		
Video messages	1 (1.0)				1 (1)				
e-Dairy messaging app	2 (1.9)					1 (1)	1 (1)		
Monitoring app	9 (8.7)				1 (1)	7 (7)	1 (1)		
Interaction app	2 (1.9)	1 (1)		1 (1)					
Internet intervention	29 (28.2)	2 (2)	2 (2)	3 (3)	7 (7)	10 (10)	4 (4)	1 (1)	
Secure messaging app	1 (1.0)	1 (1)							
Health knowledge base	3 (2.9)				2 (2)			1 (1)	
Documenting app	4 (3.9)			1 (1)			3 (3)		
Personal health record app	3 (2.9)			1 (1)	1 (1)	1 (1)		1 (1)	1)1(
Monitoring app	3 (2.9)					3 (3)			
Monitoring device + app	2 (1.9)		1 (1)		1 (1)				
Interaction app	3 (2.9)	1 (1)				1 (1)	1 (1)		
Monitoring + interaction app	6 (5.8)		1 (1)		2 (2)	3 (3)			
Monitoring video conferencing	2 (1.9)					2 (2)			
Virtual clinic	2 (1.9)			1 (1)	1 (1)				
Telemedicine intervention	42 (40.8)		5 (5)	3 (3)	9 (9)	21 (20)	3 (3)	1 (1)	
Video phone visits	1 (1.0)				1 (1)				
Monitoring device + system	13 (12.6)		1 (1)			10 (10)	2 (2)		
Telemonitoring system	7 (6.8)		2 (2)			5 (5)			
Telemedicine system	21 (20.4)		2 (2)	3 (3)	8 (8)	6 (6)	1 (1)	1 (1)	

[a] More than 1 (person-centered) connected-care management activity possible. Percentage estimated by total number of studies (n=103).

Table 4. Information and communication technology intervention used for person-centered care for chronic cardiovascular conditions (n=89).[a]

ICT Intervention	Overall, n (%)	Connected-care activity, n (%)				Person-centered care activity, n (%)				
		Consult	Medication	Physical care	Education	Self-measurement	Rehabilitation exercises	Lifestyle sharing	Shared decisions	Self-care plan
	n=89	n=10	n=3	n=19	n=20	n=47	n=3	n=1	n=1	n=1
Telephone intervention	21 (24)	5 (6)	1 (1)	4 (4)	8 (9)	2 (2)		1 (1)		
Nurse telephone calls	19 (21)	4 (4)	1 (1)	3 (3)	8 (9)	2 (2)		1 (1)		
Nurse call center	2 (2)	1 (1)		1 (1)						
Mobile phone intervention										
Monitoring app	7 (8)				1 (1)	6 (7)				
Internet intervention	13 (15)	3 (3)	1 (1)		3 (3)	5 (6)	1 (1)			
Secure messaging app	1 (1)				1 (1)					
Health knowledge base	2 (2)				2 (2)					
Documenting app	1 (1)	1 (1)								
Personal health record app	1 (1)	1 (1)								
Monitoring app	3 (3)		1 (1)			2 (2)				
Monitoring device + app	1 (1)					1 (1)				
Monitoring + interaction application	3 (3)	1 (1)				2 (2)				
Virtual clinic	1 (1)						1 (1)			
Telemedicine intervention	63 (71)	2 (2)	1 (1)	15 (17)	8 (9)	34 (38)	2 (2)		1 (1)	1 (1)
Monitoring system + cardiac implant	7 (8)					7 (8)				
TV channel system	1 (1)				1 (1)					
Telemonitoring system	33 (37)	1 (1)		13 (15)	1 (1)	16 (18)	2 (2)			
Telemedicine system	22 (25)	1 (1)	1 (1)	2 (2)	6 (7)	11 (12)			1 (1)	1 (1)

[a] More than 1 (person-centered) connected-care management activity possible. Percentage estimated by total number of studies (n=89).

Information and Communication Technology Enabling Person-Centered Care for Chronic Respiratory Conditions

Within the group of studies addressing patients with chronic respiratory disease (n=73) (Table 5), more than half of the studies (52%, 38/73) used Internet interventions; this was the highest score among the 5 categories of chronic diseases and much higher than in the category of studies addressing cardiovascular patients (13%), for example. In 28% (21/73) of the studies, telemonitoring/telemedicine system interventions were used. One-fifth (20%, 15/73) used telephone interventions with nurse telephone calls accounting for 16% (12/73) and mobile phones for 14% (11/73). Furthermore, we extracted and added Skype as a new Web intervention (1%, 1/73).

Concerning person-centered care, we found 2 studies (3%, 2/73) in which a self-care plan was created and decided on, 1 with a monitoring videoconference app on the Internet and 1 with a telemedicine system. Also noteworthy, only 5% (4/73) of the studied telemedicine system interventions were used for consultations and few (7%, 5/73) of the telephone interventions focused on physical care or education. This indicates that the transformation to PCC is still in its early phase in disease management for chronic respiratory diseases. Similar to the category of patients with cardiovascular diseases, the category of patients with chronic respiratory diseases also had an additional PCC activity—compared to the other groups of chronic patients—namely exercises for self-rehabilitation, which is of importance for this group; 8% (6/73) of the studies were targeted toward this type of activity.

Table 5. Information and communication technology interventions used for person-centered care for chronic respiratory conditions (n=73).[a]

ICT Intervention	Overall, n (%)	Connected-care activity, n (%)					Person-centered care activity, n (%)		
		Consult	Medication	Physical care	Behavior therapy	Education	Self-measurement	Rehabilitation exercises	Self-care plan
	n=73	n=5	n=3	n=7	n=4	n=12	n=46	n=6	n=2
Telephone intervention	15 (21)			6 (8)	2 (3)	5 (7)		2 (3)	
Nurse telephone calls	12 (16)			4 (5)	2 (3)	5 (7)		1 (1)	
Automated telephone calls	2 (3)			1 (1)				1 (1)	
Nurse call center	1 (1)			1 (1)					
Mobile phone intervention	11 (15)		1 (1)		1 (1)	2 (3)	6 (8)	1 (1)	
Text messaging	3 (4)		1 (1)			1 (1)	1 (1)		
Monitoring app	6 (8)				1 (1)	1 (1)	4 (5)		
Music app	1 (1)							1 (1)	
Interaction app	1 (1)						1 (1)		
Internet intervention	38 (52)	1 (1)	2 (3)	1 (1)	1 (1)	5 (7)	26 (36)	1 (1)	1 (1)
Health knowledge base	3 (4)				1 (1)	2 (3)			
Documenting app	1 (1)		1 (1)						
Personal health record app	1 (1)			1 (1)					
Monitoring app	4 (5)						4 (5)		
Monitoring device + app	14 (19)		1 (1)				12 (16)	1 (1)	
Monitoring + interaction app	13 (18)	1 (1)				3 (4)	9 (12)		
Monitoring video conferencing	1 (1)								1 (1)
Skype video app	1 (1)						1 (1)		
Telemedicine intervention	21 (29)	4 (5)					14 (19)	2 (3)	1 (1)
Monitoring device + system	1 (1)	1 (1)							
Telemonitoring system	8 (11)	1 (1)					6 (8)	1 (1)	
Telemedicine system	12 (16)	2 (3)					8 (11)	1 (1)	1 (1)

[a] More than 1 (person-centered) connected-care management activity possible. Percentage estimated by total number of studies (n=73).

Information and Communication Technology Enabling Person-Centered Care for Cancer

The category of cancer care was the broadest one, with ICT interventions used and studied for a wide variety of self-care and connected-care activities (Table 6). Cancer care management appeared to be one of the most progressive ones in our findings, as involved cancer patients actively used the latest ICT innovations and the highest number of studies (n=23) related to shared management activities in PCC (excluding self-measurement of body activities that are not typically used among those with cancer). Only here did we find studies in which patients and their care providers used the social medium Twitter for PCC for social support.

Table 6. Information and communication technology interventions used for person-centered care for cancer (n=67).[a]

ICT Intervention	Overall, n (%)	Connected-care activity, n (%)						Person-centered care activity, n (%)						
		Consult	Medication	Physical care	Behavior therapy	Palliative care	Education	Self-measurement	Knowledge sharing	Self-reporting symptoms	Lifestyle sharing	Shared decisions	Self-care plan	Social support
n=67	n=12	n=4	n=12	n=4	n=1	n=17	n=1	n=5	n=5	n=4	n=3	n=3	n=2	
Telephone intervention	16 (24)	4 (6)	1 (1)	1 (1)	2 (3)		4 (6)				3 (4)		1 (1)	
Nurse telephone calls	15 (22)	4 (6)		1 (1)	2 (3)		4 (6)				3 (4)		1 (1)	
Automated telephone calls	1 (1)		1 (1)											
Mobile phone intervention	8 (12)		1 (1)	4 (6)						1 (1)		1 (1)		1 (1)
Monitoring app	6 (9)		1 (1)	3 (4)						1 (1)		1 (1)		
Interaction app	1 (1)			1 (1)										
Twitter	1 (1)													1 (1)
Internet intervention	30 (45)		2 (3)	2 (3)	2 (3)	1 (1)	9 (13)	1 (1)	4 (6)	4 (6)	1 (1)	1 (1)	2 (3)	1 (1)
Health knowledge base	9 (13)			1 (1)	1 (1)		4 (6)		2 (3)					1 (1)
Documenting app	1 (1)									1 (1)				
Personal health record app	6 (9)		1 (1)				2 (3)			1 (1)	1 (1)		1 (1)	
Monitoring device + app	1 (1)					1 (1)								
Interaction app	2 (3)				1 (1)		1 (1)							
Monitoring + interaction app	7 (10)		1 (1)	1 (1)				1 (1)		2 (3)		1 (1)	1 (1)	
Skype video app	1 (1)						1 (1)							
Virtual clinic	1 (1)						1 (1)							
Internet support groups	2 (3)								2 (3)					
Telemedicine intervention	20 (30)	8 (12)		5 (7)			4 (6)		1 (1)			1 (1)		
Video phone visits	3 (4)	3 (4)												
Monitoring device + system	1 (1)											1 (1)		
Telemedicine system	16 (24)	5 (7)		5 (7)			4 (6)		1 (1)					1 (1)

[a] More than 1 (person-centered) connected-care management activity possible. Percentage estimated by total number of studies (n=67).

The most studied interventions for persons with a cancer condition (n=67) addressed Internet interventions for educational purposes (12%, 8/67) of connected care, especially Web portals (6%, 4/67) versus 7% (5/67) of sharing cancer knowledge in a person-centered way. This differs from the other conditions, where intervention for self-measurement of the body ranked the highest. An explanation could be that body measurements related to cancer are performed by health professionals (for the most part), illustrated by our findings that 7% (5/67) of the studied telemedicine interventions focused on physical care. A new type of PCC activity emerged from our data: 5% of the Internet interventions were applied for self-reporting symptoms. This self-care activity was distinctive and not found among studies concerning the other types of chronic diseases. A monitoring app on mobile phones, a documenting app, a personal health record app, and 2 monitoring plus interaction apps supported the self-reporting of symptoms. Overall, the total number of PCC activities (n=23) included 3 studies on shared decision making and 3 studies on self-care plan creation. Compared to connected-care activities, nurse telephone calls were used in several cases, both for consultation (6%, 4/67) and education (6%, 4/67) versus PCC personal lifestyle sharing (4%, 3/67) and cancer knowledge sharing (6%, 4/67).

Information and Communication Technology Enabling Person-Centered Care for Stroke

The fewest studies of all the chronic condition conditions were encountered in the stroke category (n=18) (Table 7). Six studies addressed PCC, of which 5 studies supported a new distinctive type of care activity: self-rehabilitation therapy. Interestingly, high-tech innovations involving a robot assistant and virtual reality gaming were studied for self-rehabilitation. Even though stroke is regarded, according to the WHO, as one of the 5 leading chronic diseases, many of the studies on connected care by applying ICT are conducted in the field of acute stroke and mobile teams. We only included the studies when acute stroke turned into a chronic situation. Most studied ICT interventions were telemedicine/telemonitoring systems (12/18, 67%) addressing physical care, consultation, and education. A note on the reliability of this data: not all 12 studies were precise in what type of ICT intervention was used and some referred to it in very general terms, such as “telemedicine system.” We suggest keeping this in mind when interpreting these results.

Table 7. Information and communication technology interventions used for person-centered care for stroke (n=18).[a]

ICT Intervention	Overall, n (%)	Connected-care activity, n (%)					Person-centered care activity, n (%)	
		Consult	Physical care	Behavior therapy	Cognitive therapy	Education	Self- measurement body	Rehabilitation exercises
	n=18	n=2	n=6	n=3	n=1	n=3	n=1	n=5
Telephone intervention								
Nurse telephone calls	1 (6)			1 (6)				
Mobile phone intervention								
Video messages	1 (6)			1 (6)				
Internet intervention	7 (39)							3 (17)
Monitoring device + app	1 (6)						1 (6)	
Monitoring video conferencing	4 (22)		1 (6)	1 (6)		1 (6)		1 (6)
Gaming (virtual reality)	2 (11)							2 (11)
Telemedicine intervention	12 (67)	2 (11)	5 (28)		1 (6)	2 (11)		2 (11)
Robot assistant	1							1 (6)
Telemonitoring system	4 (22)		4 (22)					
Telemedicine system	7 (39)	2 (11)	1 (6)		1 (6)	2 (11)		1 (6)

[a] More than 1 (person-centered) connected-care management activity possible. Percentage estimated by total number of studies (n=18).

Outcomes

Overview

In addition to the (health-related) quality of life and costs efficiency outcomes, we found 33 outcome indicators (see Table 8). Extracted from the studied text, we classified these outcome indicators under 4 category definitions. We defined person outcomes (12 indicators), connected to health professional outcomes (7 indicators), organization outcomes (9 indicators), technical outcomes (5 indicators), and no outcomes. This last category included 7% (25/350) of the studies, in which no outcome measurement was found.

Table 8. Outcomes of the information and communication technology interventions for person-centered care and connected-care management.

Outcomes	Overall N=350		Diabetes n=103		Cardiovascular n=89		Respiratory n=73		Cancer n=67		Stroke n=18	
	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)
No outcomes	25 (7)		11 (3)		6 (2)		3 (1)		5 (1)			
Person outcomes	224 (64)	45 (13)	79 (23)	9 (3)	45 (13)	12 (3)	45 (13)	19 (5)	42 (12)	4 (1)	13 (4)	1 (0)
Quality of life	46 (13)	6 (2)			28 (8)	1 (0)	8 (2)	3 (1)	8 (2)	2 (1)	2 (1)	
Health-related quality of life	28 (8)	18 (5)	8 (2)	5 (1)	3 (1)		14 (4)	12 (3)	3 (1)	1 (0)		
Mental health–related quality of life	1 (0)		1 (0)									
Mortality (less)	1 (0)	11 (3)			1 (0)	11 (3)						
Self-efficacy	19 (5)	4 (1)	8 (2)	1 (0)	5 (1)		1 (0)	2 (1)	3 (1)	1 (0)	2 (1)	
Empowerment (self-care)	54 (15)	2 (1)	22 (6)		7 (2)		11 (3)	2 (1)	11 (3)		3 (1)	
Physical condition	49 (14)		30 (9)				10 (3)		5 (1)		4 (1)	
Metabolic control	9 (3)	1 (0)	9 (3)	1 (0)								
Pain reduction	1 (0)								1 (0)			
Behavior change	3 (1)		1 (0)		1 (0)		1 (0)					
Mental health condition	13 (4)	2 (1)		2 (1)					11 (3)		2 (1)	
Loneliness		1 (0)										1 (0)
Connected to health professional outcomes	81 (23)	12 (3)	11 (3)	3 (1)	32 (9)	1 (0)	23 (7)	7 (2)	6 (2)	0 (0)	9 (3)	1 (0)
Medication adherence	12 (3)		3 (1)		5 (1)		4 (1)					
Treatment adherence	8 (2)				4 (1)		3 (1)				1 (0)	
Clinical outcomes	41 (12)	8 (2)	5 (1)	3 (1)	17 (5)		11 (3)	4 (1)	1 (0)		7 (2)	1 (0)
Effectives of intervention		4 (1)				1 (0)		3 (1)				
Documentation quality	5 (1)				1 (0)		1 (0)		2 (1)		1 (0)	
Communication quality	9 (3)		3 (1)		2 (1)		2 (1)		2 (1)			
Health knowledge	6 (2)				3 (1)		2 (1)		1 (0)			
Organization outcomes (care model)	73 (21)	59 (17)	18 (5)	6 (2)	26 (7)	31 (9)	12 (3)	19 (5)	11 (3)	3 (1)	6 (2)	0 (0)
Cost efficiency	38 (11)	6 (2)	9 (3)	1 (0)	15 (4)	1 (0)	7 (2)	4 (1)	7 (2)			
(Time) efficiency	11 (3)		5 (1)		4 (1)				1 (0)		1 (0)	
Quality effectiveness	4 (1)	1 (0)	1 (0)		2 (1)		1 (0)	1 (0)				
Productivity	1 (0)								1 (0)			
Less hospitalization	6 (2)	43 (12)		2 (1)	2 (1)	26 (7)	4 (1)	13 (4)		2 (1)		
Reduced comanagement	1 (0)				1 (0)							
Implementation enablers / barriers (including ethical)	6 (2)	6 (2)	2 (1)	3 (1)	2 (1)	2 (1)		1 (0)			2 (1)	
Improve office visits / replace face-to-face consult	1 (0)	3 (1)	1 (0)			2 (1)				1 (0)		
Improve access difficulties	5 (1)								2 (1)		3 (1)	
Technical outcomes	91 (26)	2 (1)	19 (5)	0 (0)	22 (6)	0 (0)	22 (6)	1 (0)	23 (7)	1 (0)	5 (1)	0 (0)
Feasibility	35 (10)		7 (2)		6 (2)		9 (3)		10 (3)		3 (1)	
Usability	21 (6)		5 (1)		6 (2)		4 (1)		6 (2)			

Outcomes	Overall N=350		Diabetes n=103		Cardiovascular n=89		Respiratory n=73		Cancer n=67		Stroke n=18	
	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)	Pos n (%)	Neg n (%)
Satisfaction	28 (8)	1 (0)	7 (2)		7 (2)		8 (2)		5 (1)	1 (0)	1 (0)	
Safety	6 (2)				2 (1)		1 (0)		2 (1)		1 (0)	
Commercial feasibility	1 (0)	1 (0)			1 (0)			1 (0)				
Total outcomes: positive impact	469 (134)	118 (34)	127 (36)	18 (5)	125 (36)	44 (13)	102 (29)	46 (13)	82 (23)	8 (2)	33 (9)	2 (1)

Person Outcomes

A total of 15.4% (54/350) of the studies measured a positive impact on empowerment (self-care) closely followed by improvement in physical condition (14.0%, 49/350). The increase in quality of life and health-related quality of life accounted for 13.1% (46/350) and 8.0% (28/350), respectively, and self-efficacy for 5.1% (18/350).

Three person outcome indicators were found to be distinctive for one of the 5 chronic conditions: metabolic control was measured in 10 diabetes studies and lower mortality in 11 cardiovascular studies, whereas improvement in mental health was reported in 11 cancer studies and 2 stroke studies. Overall, 76.9% (269/350) of the studies reported on person outcomes, with 64.0% (224/350) reporting a positive impact versus 12.9% (45/350) reporting a negative or no impact. These findings confirm the importance of measuring the person-centeredness of the ICT intervention, for which these 5 outcome indicators are currently commonly used.

Connected to Health Professional Outcomes

The impact for being connected to the health care professional by ICT was found to be the highest on a familiar clinical outcomes indicator. Of the total studies, 11.7% (41/350) reported an increase in clinical outcome versus a decrease in 2.0% (7/350) of the studies. Interestingly, “medication adherence” and “treatment adherence” emerged as outcome indicators in a few studies. In relation to PCC, a few other studies suggested that “documentation quality” and “communication quality” should be used to measure the concept of acquiring better insight into the patient. Overall, one-quarter of the studies (93/350, 26.6%) reported on professional outcomes connected to health, with 23.1% (81/350) reporting a positive impact versus 3.4% (12/350) reporting a negative or no impact.

Organization Outcomes

Remarkably, the most studied impact on organization outcome was not cost efficiency itself, but the related impact of less hospitalization (43/350, 12.3%), closely followed by cost efficiency (38/350, 10.9%). Time efficiency was a third outcome indicator appearing in a few studies (11/350, 3.1%). Overall, 37.7% (132/350) of the studies on connected-care and person-centered ICT interventions reported on organization outcomes. To a certain extent a positive impact was reported (20.9%, 73/350), which was challenged by a relatively large number of studies that reported a negative outcome; 59 studies (16.9%) reported a negative impact regarding organization outcomes. Most reported were both a decrease in cost efficiency (1.7%, 6/350) and an increase in hospitalization (1.7%, 6/350).

Technical Outcomes

As far as technical outcomes related to the implementation of the ICT innovation were concerned, the most measured outcome was technical feasibility (10.0%, 35/350) followed by satisfaction (8.0%, 28/350) with the ICT intervention. Important for PCC, usability was measured in 6.0% of the studies (21/350).

In sum, a positive outcome indicator was reported 469 times (134%) versus a negative outcome indicator 118 times (34%). As a percentage of the total 350 studies, we found a relatively more positive impact in studies on diabetes (36.3%, 127/350) and cancer conditions (23.4%, 82/350) versus a relatively more negative impact in the studies on cardiovascular (12.6%, 44/350) and chronic respiratory conditions (13.1%, 46/350).

Discussion

Principal Findings

Shared decision making, personal information sharing, and setting up a care plan enabled by ICT seem to be relatively new. This indicates that the state of knowledge in the PCC field of interest is still emerging, meaning there are many research opportunities to contribute. The type of ICT mostly used by persons with a chronic condition for interacting with health care providers is ICT for self-measurement of the body (n=143) (Table 9); the highest rankings were found in studies on diabetes (n=48) and cardiovascular (n=47) and chronic respiratory diseases (n=46). These are in striking contrast with the lowest ranking; only 1 study was found on self-measurement of the body within the group on cancer and only 1 on stroke. Given these types of chronic diseases (cancer and stroke), physical measurements and check-ups likely require the health care provider to use specialized professional equipment. Instead, shared care management activities are enabled by the person-centered ICT for cancer (n=23) and stroke (n=5). Overall, we found 60 studies (17%) on this type of shared decision-making ICT. Cancer ranked first in number of studies followed by diabetes (n=18).

We note that hardly any of these interventions could be regarded as “fully” addressing the 3 routines of PCC for activities related to initiating the partnership (patient narratives), working the partnership (shared decision making), and safeguarding the partnership (documenting the narrative) [7].

Table 9. Person-centered care and information and communication technology interventions used for the big 5 chronic connected-care activities (CCA), person-centered self-measurement (PCM), and person-centered shared management (PCS).[a]

PCC-ICT interventions used	Total	Telephone	Mobile phone app	Internet app	Telemedicine system
Overall	398 (114)	70 (20)	52 (15)	117 (33)	159 (45)
CCA	195 (56)	51 (15)	20 (6)	50 (14)	74 (21)
PCM	143 (41)	8 (2)	23 (7)	43 (12)	69 (20)
PCS	60 (17)	11 (3)	9 (3)	24 (7)	16 (5)
Diabetes					
CCA	47 (13)	7 (2)	9 (3)	14 (4)	17 (5)
PCM	48 (14)	6 (2)	11 (3)	10 (3)	21 (6)
PCS	18 (5)	4 (1)	5 (1)	5 (1)	4 (1)
Cardiovascular					
CCA	52 (15)	18 (5)	1 (0)	7 (2)	26 (7)
PCM	47 (13)	2 (1)	6 (2)	5 (1)	34 (10)
PCS	6 (2)	1 (0)	0 (0)	1 (0)	4 (1)
Chronic respiratory					
CCA	31 (9)	13 (4)	4 (1)	10 (3)	4 (1)
PCM	46 (13)	0 (0)	6 (2)	26 (7)	14 (4)
PCS	8 (2)	2 (1)	1 (0)	2 (1)	3 (1)
Cancer					
CCA	50 (14)	12 (3)	5 (1)	16 (5)	17 (5)
PCM	1 (0)	0 (0)	0 (0)	1 (0)	0 (0)
PCS	23 (7)	4 (1)	3 (1)	13 (4)	3 (1)
Stroke					
CCA	15 (4)	1 (0)	1 (0)	3 (1)	10 (3)
PCM	1 (0)	0 (0)	0 (0)	1 (0)	0 (0)
PCS	5 (1)	0 (0)	0 (0)	3 (1)	2 (1)

[a] More than 1 (person-centered) connected-care management activity possible.

Furthermore, our findings suggested that the most commonly used personalized ICT interventions involved telemonitoring or telemedicine systems (n=159) followed by Web-based applications on the Internet (n=117). In approximately one-fifth of the studies, the telephone (n=70) was used to connect patient and physician, mostly for consultation and education. For example, in the case of cardiovascular conditions, we found 18 studies on telephone intervention for connected care and in 19 studies on persons with diabetes, the telephone was used, often in combination with Internet-based interventions. In addition, the use of mobile phone apps ranked the highest in diabetes care activities (n=25).

The usage of social media, such as Twitter, was only incidentally mentioned in the reviewed studies, even though eHealth app and medical Internet-based interventions are paying increasing attention to social media [32]. A possible explanation could be that Twitter is less used in the relationship between a patient and his or her health care professional, which is the starting point for this scoping review, and more for accessing health information in general.

When comparing ICT-enabled PCC innovations used in different chronic diseases, several results stand out. First, in the case of cardiac patients, high-tech innovations connect remote monitoring software to implant devices (53%), such as pacemakers. Second, persons with a chronic stroke condition are beginning to use (serious) gaming and robot devices, specifically for rehabilitation purposes, which is a necessary treatment immediately after a stroke incident. Because technology is becoming smaller and cheaper, the possibilities of “wearable” smart technologies are increasing, and we expect to see more of these technologies in the future. Third, virtual clinics provide self-rehabilitation exercises. This technology combines a virtual clinic app on the Internet with telemonitoring systems.

The impact of PCC-ICT interventions on quality of life and health-related quality of life are positive (Table 9). Several studies claim an increase in quality of life (46/350, 13.1%) or health-related quality of life (28/350, 8.0%). It seems that enabling a person to manage his or her own disease through ICT leads to an improvement in the perceived quality of an individual’s daily life (quality of life) and an increase in the measurement of an individual's well being affected over time by the disease, disability, or disorder (health-related quality of life).

The impact on cost and efficiency seems to be positive but less conclusive. Some studies reported positive impacts (38/350, 10.9%). Some of the studies, however, indicated negative impacts, either an increase in hospital (re)admission (6/350, 1.7%) or rise in costs (6/350, 1.7%). Our study suggests that not only could a person with a chronic disease benefit from an ICT-enabled PCC approach, but also that ICT-PCC yields organizational paybacks, although not in all cases. It could also lead, as was reported in some studies, to an increase in health care usage.

Other relevant study outcomes suggest that organizational barriers stand in the way of implementation of ICT-PCC, which is also supported by previous studies [26].

Limitations

Although we covered a considerable number of studies, the search was limited to medical databases. Due to this system restriction, there is a chance that we have missed possible related articles in other domains, such as information systems research, social studies, and organizational change management research.

We realize that conducting a scoping study comes with limitations. We acknowledge the fact that the quantitative overview typical for scoping review results, unlike systematic reviews, does not appraise the quality of evidence in the primary research reports with a detailed analysis of a smaller and similar number of studies [22]. Because scoping reviews do not assess the quality of the studies, we included those studies in the review even though they did not reach the quality standard of some peer-reviewed journals. This also sets limitations on the results, which could only be described in general terms such as a "telemonitoring device." Furthermore, we acknowledge the publication bias of a tendency to publish positive results that could yield a distorted overview of the scope of conducted research on ICT-enabled PCC. Lastly, our interpretations are limited to outcomes reported in the English language.

Comparison With Prior Works

This scoping review mapped ICT-PCC interventions that are applied in chronic disease management to support patients to take an active part in their care and the decision-making process, and make it possible for patients to interact directly with health care providers and services about their personal health concerns. Our study distinguished 13 extracted care activities of connected care and PCC, which build on previous literature review studies on PCC and/or ICT, such as the one conducted by Aarts et al [17]. They extracted 2 care activities, namely the provision of support and education to patients and the promotion of mental health for patient-focused Internet interventions within the discipline of reproductive medicine.

Corresponding to our findings of less hospitalizations are the findings of the scoping review on the effects of PCC for patients with chronic heart failure in hospital settings. Ekman et al [9] found that a fully implemented PCC approach shortens hospital stays and maintains functional performance in patients hospitalized for worsening congestive heart failure without increasing the risk of readmission or jeopardizing patients' health-related quality of life. However, Ekman's study did not involve ICT and the focus was fully on the impact of a PCC approach. A comparable conclusion was drawn in a meta-analysis conducted on the outcomes of an Internet intervention and eHealth counseling on risk factors linked to certain chronic diseases [16].

Theoretical Implication and Further Research Suggestions

This scoping research study has contributed to the growing scholarly interest in PCC and ICT interventions for self-management (of chronic conditions) by providing an overview of the extent and nature of the existing literature and evidence base involving the subset of ICT interventions in PCC for chronic conditions. Sixty relevant health studies have been identified regarding the big 5 chronic diseases to support patients and health care providers in the online and personalized management of these diseases.

For future research, we have 3 suggestions: first, given that hardly any of the studies showed a fully PCC-ICT approach, a logical next step is a qualitative study addressing the selection of the studies we found. Such a study can add qualitative insight and lead to placing an emphasis on building a framework. Second, given the 35 outcome indicators we identified, further research on the definition and measurement can help to further develop an evidence base for PCC and ICT for self-management of chronic disease. Third, we pose 2 challenging questions for further research:

1. How can ICT-enabled PCC be implemented in network organizations to support self-management of chronic patients in a person-centered care manner?
2. What does this mean for innovative care models?

Practitioner and Managerial Implication

Concerning the impact of ICT-enabled PCC, this scoping review study found that empowerment (self-care) of the patient was the main outcome (15%) of the ICT interventions, followed by physical condition (14%), quality of life (13%), and health-related quality of life (8%). For the health care professional, the impact was highest when looking at clinical outcomes (14%). We also found a decrease in clinical outcomes in 2% of the studies. Regarding the most studied impact in the organization, we concluded that the outcome is less hospitalization (12%) and cost efficiency (11%). As far as the ICT intervention is concerned, the impact of feasibility (10%) is high. We also did find negative outcomes within the overall chronic disease categories: health-related quality of life decrease (5%), cost efficiency decrease (2%), and increase in hospitalization (2%).

Conclusions

Hardly any of the interventions could be regarded as "fully" PCC meeting the 3 routines of initiating the partnership (patient narratives), working the partnership (shared decision making), and safeguarding the partnership (documenting the narrative). This review will be especially helpful to those deciding on areas where the further development of research or implementation of ICT for PCC may be warranted.

The scoping review investigated the extent, range, and nature of research activities regarding ICT interventions that have been studied to support patients and health care professionals in PCC management of the big 5 chronic diseases. From the initial 9380 search results, we identified 350 studies that qualified for inclusion. The largest share of ICT interventions studied sought to support patients in self-measurement of the body. The highest impact of ICT interventions (15%) of the studies on patients was measured on the increase of empowerment (self-care) closely followed by improvement in physical condition (14%), increase in quality of life (13%), health-related quality of life (8%), and self-efficacy (5%). Only 6% of the studies measured usability. This is disturbing since usability is an important fact for the acceptability of ICT by its users, and the lack of attention paid to usability in the reviewed studies indicates that there would be much to be gained from this.

The scoping review suggests that not only can persons with a chronic disease benefit from an ICT-enabled PCC approach, but also that ICT-PCC yields organizational paybacks, although not in all cases. It could also lead, as was reported in some studies, to an increase in health care usage. Other relevant study outcomes suggest that organizational barriers stand in the way of implementation of ICT-PCC, which is also supported by previous studies.

The impact of being connected to the health care professional by ICT is found to be the highest (12%) on a familiar clinical outcomes indicator versus a decrease in 2% of the studies. Remarkably, the most studied impact on organization outcome is not cost efficiency itself, but the related impact of less hospitalization (12%) closely followed by cost efficiency (11%).

Persons with a chronic disease are beginning to use (serious) gaming, social media, wearable technology, and robot devices for the management of diseases. Because technology overall is becoming smaller and cheaper, the possibilities of these smart technologies are increasing and we expect to see more of these technologies in the future.

Acknowledgments

This research has been supported by COMMIT/, a Dutch public-private research community in information and communication science and relates to the research project entitled “Care Model Design for Person Centered ICT-interventions in Home Healthcare” funded by NWO-KIEM. The authors would like to acknowledge Medical Information Specialist A Tillema of the Medical Library of Radboud University Nijmegen for her valuable support and contribution to the scoping study, and R Otten, Medical Information Specialist of the VU University Medical Library for his support in the last stage of the research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Overview of studies (n=350), qualified for inclusion.

[PDF File (Adobe PDF File), 233KB - jmir_v17i3e77_app1.pdf]

References

1. World Health Organization, Alwan A. Global status report on noncommunicable diseases 2010. Geneva: World Health Organization; 2011. URL: http://whqlibdoc.who.int/publications/2011/9789240686458_eng.pdf?ua=1 [WebCite Cache ID 6WsXeEbu7]
2. Huber M, Knottnerus JA, Green L, Horst HVD, Jadad AR, Kromhout D, et al. How should we define health? BMJ 2011 Jul 26;343(jul26 2):d4163-d4163. [doi: 10.1136/bmj.d4163]
3. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]
4. Frost J, Massagli M. PatientsLikeMe the case for a data-centered patient community and how ALS patients use the community to inform treatment decisions and manage pulmonary health. Chron Respir Dis 2009;6(4):225-229. [doi: 10.1177/1479972309348655] [Medline: 19858352]
5. 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]
6. World Health Organization Western Pacific Region. People-centred health care: A policy framework. Geneva: World Health Organization; 2007. URL: http://www.wpro.who.int/health_services/people_at_the_centre_of_care/documents/ENG-PCIPolicyFramework.pdf [accessed 2015-03-08] [WebCite Cache ID 6WsXyGuIb]
7. Ekman I, Swedberg K, Taft C, Lindseth A, Norberg A, Brink E, et al. Person-centered care--ready for prime time. Eur J Cardiovasc Nurs 2011 Dec;10(4):248-251. [doi: 10.1016/j.ejcnurse.2011.06.008] [Medline: 21764386]

8. Ekman I, Britten N, Bordin J, Codagnone C, Eden S, Forslund D, et al. The person-centred approach to an ageing society. EJPCH 2013 Jun 11;1(1):132-137. [doi: 10.5750/ejpch.v1i1.644]
9. Ekman I, Wolf A, Olsson LE, Taft C, Dudas K, Schaufelberger M, et al. Effects of person-centred care in patients with chronic heart failure: the PCC-HF study. Eur Heart J 2012 May;33(9):1112-1119 [FREE Full text] [doi: 10.1093/eurheartj/ehr306] [Medline: 21926072]
10. Olsson LE, Karlsson J, Ekman I. The integrated care pathway reduced the number of hospital days by half: a prospective comparative study of patients with acute hip fracture. J Orthop Surg Res 2006;1:3 [FREE Full text] [doi: 10.1186/1749-799X-1-3] [Medline: 17150123]
11. Olsson LE, Jakobsson UE, Swedberg K, Ekman I. Efficacy of person-centred care as an intervention in controlled trials - a systematic review. J Clin Nurs 2013 Feb;22(3-4):456-465. [doi: 10.1111/jocn.12039] [Medline: 23231540]
12. van Meeuwen DPD, van Walt Meiijer QJ, Simonse LWL. Care models of eHealth services: A case study on the Design of a Business Model for an Online Precare Service. J Med Internet Res Protoc 2015;4(1):e32. [doi: 10.2196/resprot.3501]
13. 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]
14. Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. Journal of the American Medical Informatics Association 2007;14(3):269-277. [doi: 10.1197/jamia.M2270]
15. 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]
16. Liu S, Dunford SD, Leung YW, Brooks D, Thomas SG, Eysenbach G, et al. Reducing blood pressure with Internet-based interventions: a meta-analysis. Can J Cardiol 2013 May;29(5):613-621. [doi: 10.1016/j.cjca.2013.02.007] [Medline: 23618507]
17. Aarts JWM, van den Haak P, Nelen WLDM, Tuil WS, Faber MJ, Kremer JAM. Patient-focused internet interventions in reproductive medicine: a scoping review. Hum Reprod Update 2012 Apr;18(2):211-227 [FREE Full text] [doi: 10.1093/humupd/dmr045] [Medline: 22108381]
18. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. N Engl J Med 2012 Mar 1;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
19. International College of Person-Centered Medicine. Geneva Declaration on Person-centered Care for Chronic Diseases. The International Journal of Person Centered Medicine 2012;2:153-154. [doi: 10.5750/2Fijpcm.v2i2.206]
20. Anderson S, Allen P, Peckham S, Goodwin N. Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. Health Res Policy Syst 2008;6:7 [FREE Full text] [doi: 10.1186/1478-4505-6-7] [Medline: 18613961]
21. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010;5(1):1-9. [doi: 10.1186/1748-5908-5-69]
22. Arksey HOML, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
23. Gruman J, Rovner MH, French ME, Jeffress D, Sofaer S, Shaller D, et al. From patient education to patient engagement: implications for the field of patient education. Patient Educ Couns 2010 Mar;78(3):350-356. [doi: 10.1016/j.pec.2010.02.002] [Medline: 20202780]
24. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. JMIR 2008;10(3):22. [doi: 10.2196/jmir.1030]
25. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS medicine 2013;10(1):1001362. [doi: 10.1371/journal.pmed.1001363]
26. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
27. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
28. Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011(10):CD001431. [doi: 10.1002/14651858.CD001431.pub3] [Medline: 21975733]
29. Verhoeven F, Tanja-Dijkstra K, Nijland N, Eysenbach G, van Gemert-Pijnen L. Asynchronous and synchronous teleconsultation for diabetes care: a systematic literature review. J Diabetes Sci Technol 2010 May;4(3):666-684 [FREE Full text] [Medline: 20513335]
30. García-Lizana F, Sarría-Santamera A. New technologies for chronic disease management and control: a systematic review. J Telemed Telecare 2007;13(2):62-68. [doi: 10.1258/135763307780096140] [Medline: 17359568]
31. Nijland N, van Gemert-Pijnen J, Boer H, Steehouder MF, Seydel ER. Evaluation of internet-based technology for supporting self-care: problems encountered by patients and caregivers when using self-care applications. J Med Internet Res 2008;10(2):e13 [FREE Full text] [doi: 10.2196/jmir.957] [Medline: 18487137]

32. Nickels A, Dimov V. Innovations in technology: social media and mobile technology in the care of adolescents with asthma. Current Allergy Asthma Report 2012 Dec 01;12(6):607-612. [doi: 10.1007/s11882-012-0299-7] [Medline: 22976493]

Abbreviations

CCA: connected-care activities
HRQL: health-related quality of life
ICT: information and communication technology
MeSH: Medical Subject Heading
PCC: person-centered care
PCM: person-centered self-measurement
PCS: person-centered shared management
PDF: portable document format
WHO: World Health Organization

Edited by G Eysenbach; submitted 08.07.14; peer-reviewed by I Ekman, O Kulyk; comments to author 21.08.14; revised version received 26.09.14; accepted 20.10.14; published 27.03.15.

Please cite as:
Wildevuur SE, Simonse LWL
Information and Communication Technology–Enabled Person-Centered Care for the "Big Five" Chronic Conditions: Scoping Review
J Med Internet Res 2015;17(3):e77
URL: http://www.jmir.org/2015/3/e77/
doi:10.2196/jmir.3687
PMID:25831199

Original Paper

Web-Based Self-Reported Height, Weight, and Body Mass Index Among Swedish Adolescents: A Validation Study

Sandra Ekström[1], MSc; Inger Kull[1,2,3], RN, PhD; Sara Nilsson[4], MSc; Anna Bergström[1], PhD

[1]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[2]Department of Clinical Science and Education, Karolinska Institutet, Stockholm, Sweden
[3]Sachs' Children's Hospital, South General Hospital Stockholm, Stockholm, Sweden
[4]Center for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden

Corresponding Author:
Sandra Ekström, MSc
Institute of Environmental Medicine
Karolinska Institutet
BOX 210
Stockholm, 17177
Sweden
Phone: 46 852487010
Fax: 46 8524304571
Email: Sandra.Ekstrom@ki.se

Abstract

Background: Web-collected height and weight are increasingly used in epidemiological studies; however, the validity has rarely been evaluated.

Objective: The aim of the study was to validate self-reported height, weight, and corresponding body mass index (BMI) among Swedish adolescents aged approximately 16 years. A secondary aim was to investigate possible prediction factors for validity of self-reported BMI.

Methods: The study included 1698 adolescents from the population-based cohort BAMSE. Height and weight were collected through a Web-based questionnaire and subsequently measured using standard procedures. Differences between reported and measured height, weight, and corresponding BMI were compared by t tests and agreement was evaluated by Pearson correlation and Bland-Altman plots. Multivariable linear regression analysis was used to investigate whether lifestyle and demographic factors predicted validity of self-reported BMI.

Results: On average, weight was underestimated by 1.1 kg and height was overestimated by 0.5 cm, leading to an underestimation of BMI by 0.5 kg/m2. Correlation coefficients were .98 for height, .97 for weight, and .94 for BMI, and highly significant. Females underestimated weight to a higher extent than males and overweight and obese participants underestimated weight to a higher extent than normal-weight participants, which resulted in higher underestimation of BMI. Underweight participants, on the contrary, overestimated weight and correspondingly BMI. Overall, a high proportion of participants were classified into the correct BMI category; however, among overweight and obese participants, only 60.2% (139/231) and 46% (20/44) were correctly classified, respectively. In the multivariable prediction model, only gender and BMI status significantly predicted discrepancy between reported and measured BMI.

Conclusions: Web-collected BMI may be used as a valid, quick, and cost-effective alternative to measured BMI among Swedish adolescents. The accuracy of self-reported BMI declines with increasing BMI and self-reported BMI should not be used to estimate the prevalence of overweight or obesity.

(J Med Internet Res 2015;17(3):e73) doi:10.2196/jmir.3947

KEYWORDS

adolescent; body height; body weight; body mass index; validity; Internet

Introduction

Self-reported Web-based weight and height have become widely used in epidemiological research as a quick and cost-efficient way to assess body mass index (BMI, kg/m^2). BMI status is used for investigating disease associations, evaluating interventions, and monitoring obesity trends. The accuracy of Web-based self-reported weight and height is essential; however, few studies have validated these against measured weight and height [1-4].

The validity of self-reported weight and height might vary between populations depending on age, gender, and cultural factors [5-10]. Usually, high correlations between self-reported (using interviews or mailed questionnaires) and measured BMI are observed; however, systematic biases between the two frequently exist. Generally, weight is underestimated whereas height is slightly overestimated, leading to underestimation of BMI and misclassification (eg, of overweight as normal weight) [3,5,6,8]. Overweight/obese people tend to underestimate weight to a higher extent compared to normal-weight people, and women tend to underestimate weight more than men [8,9].

During adolescence, self-reported BMI might be particularly biased. Adolescents experience rapid growth and may have less knowledge about their current weight and height. Also, factors such as body image and social desirability might influence adolescents to report idealized or socially accepted numbers [11].

Using the Web has been suggested to increase the validity of sensitive questions such as weight because it provides anonymity and creates distance between the participant and the researcher [12]. The Web also provides other benefits because it enables direct detection of missing values, elimination of data entry errors, and allows for quick administration to a large number of participants [12]. The Web might be particularly useful for adolescents because they are accustomed to computers and may find Web-based questionnaires easier and more attractive than paper questionnaires. The increasing number of Web-based surveys could, on the other hand, lead to carelessness when filling out these types of questionnaires.

Web-based self-reported height and weight have rarely been evaluated among adolescents [2,13] and, to our knowledge, no study has been done in a European setting. The aim of this study was to validate Web-based self-reported weight, height, and corresponding BMI against measured weight, height, and BMI in Swedish adolescents aged approximately 16 years. A second aim was to investigate determinants for possible discrepancies between self-reported and measured BMI.

Methods

Study Design and Study Population

The study population included participants from the prospective birth cohort BAMSE, originally consisting of 4089 children born between 1994 and 1996 in predefined areas of Stockholm. Children in the BAMSE study have been followed repeatedly from birth up to age 16 years, with the main objective to investigate lifestyle and environmental factors associated with the development of allergic diseases. Detailed description of the BAMSE study has been published elsewhere [14,15].

At approximately 16 years of age (range 15.7-19.0 years), adolescents answered a Web-based questionnaire containing questions on various health outcomes and lifestyle factors, including weight and height. After approximately 2 weeks, participants who did not answer the questionnaire were sent a reminder, including a paper-based questionnaire that could be answered instead of the Web-based version. In total, 3115 of 4089 adolescents (76.18%) answered the 16-year follow-up questionnaire of which 2847 (91.40%) answered the Web-based version of the questionnaire.

After answering the questionnaire, participants were invited to a clinical examination where measurements of weight, height, and blood pressure were taken. Weight was measured without shoes and with light indoor clothes to the nearest 0.1 kg using an electronic scale (Seca 799). Height was measured twice without shoes to the nearest 0.1 cm using a wall-mounted wooden stadiometer with a measurement range of 700 to 2050 mm. All measurements were performed by trained nurses and documented according to standard protocols. The study was approved by the regional ethical review board in Stockholm.

Definition and Classification of Variables

The mean of the 2 height measurements was computed and used in the analyses. BMI was calculated as body weight in kilograms divided by height in meters squared (kg/m^2). Underweight, normal weight, overweight, and obese were defined using the age- and sex-specific cut-off values developed for children younger than 18 years [16,17]. For participants older than 18 years, underweight was defined as BMI <18.5 kg/m^2, normal weight as BMI 18.5-24.9 kg/m^2, overweight as BMI 25-29.9 kg/m^2, and obese as BMI ≥30 kg/m^2.

Parental socioeconomic status (professional or manual labor worker) and parental ethnicity (any parent born outside of Scandinavia, including Sweden, Norway, Denmark, or Finland) were obtained at the 8-year follow-up questionnaire in BAMSE. Lifestyle factors and self-rated health were reported by the adolescent in the 16-year questionnaire, whereas blood pressure was measured at the 16-year clinical investigation. Blood pressure was measured 3 times by using an oscillometric monitor (Omron M6 Professional) according to standard procedures recommended by the Swedish Pediatric Society. The mean value of the last 2 measurements was used in the analyses. Questions about physical activity and sedentary time for both summer and winter in the past 12 months were asked. The mean value of the seasons was used in the analyses. Vigorous physical activity included time (hours/week) spent on activities such as lifting heavy weights, aerobics, or high-speed bicycling. Sedentary time included time (hours/day) (outside of school) spent on watching TV, computer use, playing computer or video games, and/or reading. Unreasonable values (>35 hours/week of vigorous physical activity [n=26] and >15 hours/day sedentary time [n=17]) were excluded. Sleep was categorized into an average of <8 or ≥8 hours/night and fruit and vegetable intake were combined into 1 variable that was categorized into every day (≥1/day for fruit and/or ≥7/week for vegetables) or less than every day. Tobacco use included cigarettes and snuffs (regular or irregular use). Self-rated health was categorized into completely healthy compared to fairly healthy/not very healthy. Pubertal status was defined according to a scoring system developed by Petersen et al [18] and categorized into pre/early, mid, or late/post puberty based on questions on body hair development, linear growth spurt, and pubic hair growth (both males and females); voice change and beard growth (males only); and breast development and menarche (females only).

Adolescents (N=1698) where included in the present analyses if (1) self-reported weight and height was reported through the Web (268 paper-based answers excluded), (2) self-reported and measured height and weight were available (additional 471 excluded), (3) self-reported weight and height were collected prior to measured weight and height (additional 19 excluded), (4) the difference between the 2 height measurements was not greater than 0.5 cm (additional 15 excluded), and (5) the time span between self-reported and measured weight and height did not exceed 8 weeks (additional 644 excluded).

Statistical Methods

Differences in demographic and lifestyle factors were evaluated using the *t* test (continuous variables) and the chi-square test (categorical variables). Mean values of self-reported and measured weight, height, and BMI were compared by using the *t* test and agreement was evaluated by Pearson correlation coefficients. Test for trend of differences across BMI categories were tested using the Wilcoxon rank sum test (Cuzick's trend test) [19]. Absolute agreement between self-reported and measured BMI was investigated by plotting the difference between self-reported and measured BMI against measured BMI (Bland-Altman plot) [20]. Because weight and height were measured to 1 decimal place but reported without decimals, a sensitivity analysis was performed using measured values of weight and height rounded to 1 decimal place.

To identify factors that might explain differences between self-reported and measured BMI, a prediction model was built by using linear regression with a backward selection technique using a *P* value <.2 from the log-likelihood ratio test to select the final model [21]. The following potential demographic and lifestyle factors, coded and categorized as previously mentioned, were included in the full model: gender, age, parental ethnicity, parental socioeconomic status, BMI status (normal weight as referent), pubertal status, vigorous physical activity, sedentary time, sleep duration, fruit and vegetable intake, tobacco use, and self-perceived health. To be able to compare potential models, only participants with complete information on all the preceding variables were included in the prediction analysis (n=1337).

All analyses were performed using the statistical software Stata version 13 (StataCorp LP, College Station, TX, USA).

Results

Description of Study Population

Characteristics of the included study participants are shown in Table 1. Of the 1698 participants, 889 (52.36%) were females and 809 (47.64%) were males. Age when answering the questionnaire ranged from 15.7 to 18.9 years with a mean of 16.5 years and the majority (1517/1698, 89.34%) between 16 and 17 years. The mean time between reported and measured height and weight was 4.6 weeks.

The majority of the participants at 8 years (1458/1619, 90.06%) had at least one parent working as a professional worker and 267 of 1622 (16.46%) had a parent that was born outside of Scandinavia. According to measured weight and height of 1698 participants, 114 (6.71%) were classified as underweight, 1309 (77.09%) as normal weight, 231 (13.60%) as overweight, and 44 (2.59%) as obese.

Compared to females, males were more likely to be overweight and obese and had higher systolic blood pressure. Almost all females (799/813, 98.3%) were in late-/postpuberty, whereas the corresponding proportion for males was 54.8% (382/697). Males reported more vigorous activity, sedentary time, and sleep than females. Less than half of the males (386/799, 48.3%) reported eating fruit or vegetables every day, whereas 64.6% (574/888) of the females did so. Tobacco was used among 238 of 1696 (14.03%) participants and did not differ statistically across gender. The majority of the participants (1382/1693, 81.63%) considered themselves to be completely healthy.

A comparison between the study participants and adolescents that did not fulfill the inclusion criteria (Multimedia Appendix 1) showed that included participants were slightly younger and reported a somewhat lower amount of vigorous physical activity. Also, they considered themselves as completely healthy to a higher extent.

Table 1. Description of the study population (N=1698) derived from a prospective birth cohort born in Stockholm in 1994-1996.

Characteristics	All (N=1698)[a]	Females (n=889)[a]	Males (n=809)[a]	*P* [b]
Age (years),[c] mean (SD)	16.5 (0.3)	16.5 (0.4)	16.5 (0.3)	.88
Time between self-reported and measured BMI (weeks), mean (SD)	4.6 (1.7)	4.6 (1.7)	4.7 (1.7)	.73
Parental socioeconomic status at 8 years, n (%)				
Professional worker	1458 (90.06)	767 (90.2)	691 (89.9)	.80
Any parent born outside of Scandinavia, n (%)	267 (16.46)	148 (17.4)	119 (15.5)	.31
BMI status,[d] n (%)				
Underweight	114 (6.71)	58 (6.5)	56 (6.9)	
Normal weight	1309 (77.09)	712 (80.1)	613 (73.8)	
Overweight	231 (13.60)	104 (11.7)	134 (15.7)	
Obese	44 (2.59)	15 (1.7)	31 (3.6)	.004
Blood pressure (mm Hg), mean (SD)				
Systolic	121.3 (11.9)	115.9 (9.7)	127.4 (11.2)	<.001
Diastolic	67.2 (7.0)	67.0 (6.0)	67.3 (7.2)	<.38
Pubertal status, n (%)				
Pre/early	26 (1.72)	1 (0.1)	25 (3.6)	
Mid	303 (20.07)	13 (1.6)	290 (41.6)	
Late/post	1181 (78.21)	799 (98.3)	382 (54.8)	<.001
Vigorous physical activity(h/week), mean (SD)	4.8 (4.2)	4.1 (3.6)	5.7 (4.7)	<.001
Sedentary time (h/day), mean (SD)	4.0 (2.2)	3.6 (1.9)	4.4 (2.3)	<.001
Sleep 8 h/day, n (%)	948 (56.03)	462 (52.0)	486 (60.5)	<.001
Fruit and vegetable consumption every day, n (%)	960 (56.91)	574 (64.6)	386 (48.3)	<.001
Tobacco use, n (%)	238 (14.03)	116 (13.1)	122 (15.1)	.22
Consider themselves completely healthy, n (%)	1382 (81.63)	715 (80.5)	667 (82.9)	.23

[a] Numbers for each variable might not add up to total due to missing information in some variables.

[b] *P* value for comparing females and males.

[c] When answering the questionnaire.

[d] Based on measured weight and height.

Differences Between Self-Reported and Measured Weight, Height, and Body Mass Index

The mean self-reported and measured height, weight, and corresponding BMI for the total group and separated by gender is shown in Table 2. There were significant differences between self-reported and measured weight, height, and BMI for both girls and boys (P<.001). On average, self-reported height was 0.5 cm higher than measured height and self-reported weight was 1.1 kg lower than measured weight. This corresponded to a mean underestimation of BMI by 0.5 kg/m^2. Females underestimated weight and BMI to a higher extent than males, whereas males overestimated height to a higher extent than females. Using rounded values of measured weight and height resulted in very similar numbers and did not affect the results in any considerable way.

Table 2. Mean self-reported and measured height, weight, and BMI by gender.

Anthropometrics	Self-reported, mean (SD)	Measured, mean (SD)	Difference, mean (SD)	*P*
Total (N=1698)				
Height (cm)	173.6 (9.0)	173.1 (9.0)	0.5 (1.8)	<.001
Weight (kg)	63.9 (11.0)	65.0 (11.5)	–1.1 (2.9)	<.001
BMI (kg/m^2)	21.1 (2.8)	21.6 (3.1)	–0.5 (1.1)	<.001
Females (n=889)				
Height (cm)	167.9 (6.1)	167.4 (6.2)	0.4[a](1.5)	<.001
Weight (kg)	59.0 (8.7)	60.5 (9.2)	–1.5[b](2.5)	<.001
BMI (kg/m^2)	20.9 (2.8)	21.6 (3.0)	–0.6[b](1.0)	<.001
Males (n=809)				
Height (cm)	179.9 (7.2)	179.3 (7.2)	0.6 (2.1)	<.001
Weight (kg)	69.2 (10.7)	69.9 (11.7)	–0.7 (3.2)	<.001
BMI (kg/m^2)	21.3 (2.8)	21.7 (3.2)	–0.4 (1.1)	<.001

[a] Significantly different from males (*P*=.02).

[b] Significantly different from males (*P*<.001).

Figure 1 shows histograms of differences between self-reported and measured height, weight, and corresponding BMI. Differences ranged from –9.4 to 19.6 cm for height, –24.9 to 16.3 kg for weight, and –7.5 to 5.4 kg/m^2 for BMI. The differences were relatively normally distributed, although somewhat negatively skewed for weight and BMI. A higher proportion of females compared to males recalled their height within 1 and 2 cm, respectively (550/889, 61.9% and 788/889, 88.6% among females; 415/809, 51.3% and 620/809, 76.6% among males). No such differences were observed for comparable weight or BMI categories (data not shown).

The Pearson correlation coefficient between self-reported and measured BMI was .94 (*P*<.001) (Figure 2). For height and weight, the correlation coefficients were .98 and .96, respectively. Similar values were obtained for males and females.

Dose-response associations (*P*<.001) between higher BMI categories (defined by measured weight and height) and larger overreporting of height and underreporting of weight and corresponding BMI was observed (Table 3). On average, normal-weight participants underreported weight by 0.8 kg, overweight participants by 2.8 kg, and obese participants by 5.1 kg. In contrast, underweight participants overreported weight by on average 0.9 kg, whereas there was no difference between reported and measured height. Figure 3 shows a graphical illustration of the difference between self-reported and measured BMI in relation to measured BMI (Bland-Altman plot). There was a clear trend of increasing underestimation of BMI with greater measured BMI values.

Table 3. Mean of self-reported and measured height, weight, and BMI by measured BMI status.

BMI status[a]	Self-reported, mean (SD)	Measured, mean (SD)	Difference, mean (SD)	*P*
Underweight (n=114)				
Height (cm)	173.1 (8.4)	173.2 (8.7)	0.0 (2.0)	.87
Weight (kg)	52.1 (6.5)	51.3 (5.4)	0.9 (2. 6)	<.001
BMI (kg/m^2)	17.3 (1.0)	17.1 (0.7)	0.3 (0.8)	<.001
Normal weight (n=1307)				
Height (cm)	173.5 (8.9)	173.0 (8.9)	0.5 (1.6)	<.001
Weight (kg)	61.9 (8.4)	62.8 (8.1)	–0.8 (2.4)	<.001
BMI (kg/m^2)	20.5 (1.7)	20.9 (1.7)	–0.4 (0.9)	<.001
Overweight (n=231)				
Height (cm)	174.5 (9.5)	173.5 (9.3)	1.0 (2.0)	<.001
Weight (kg)	75.5 (9.4)	78.3 (9.0)	–2.8 (3.5)	<.001
BMI (kg/m^2)	24.7 (1.8)	25.9 (1.3)	–1.2 (1.3)	<.001
Obese (n=44)				
Height (cm)	175.3 (9.5)	174.2 (9.2)	1.1[b](2.7)	.009
Weight (kg)	90.9 (8.9)	96.0 (8.6)	–5.1[b](5.2)	<.001
BMI (kg/m^2)	29.7 (3.5)	31.7 (2.9)	–2.0[b](1.8)	<.001

[a] Based on measured weight and height.

[b] *P*<.001 across BMI groups.

Table 4 shows the number of adolescents classified as underweight, normal weight, overweight, and obese according to self-reported and measured BMI, respectively. In total, 1467 of 1698 (86.40%) of the adolescents were classified into the correct BMI category. Among normal-weight participants, 1227 of 1309 (93.74%) were classified correctly, whereas corresponding numbers were 81 of 114 (71.1%) for underweight, 139 of 231 (60.2%) for overweight, and 20 of 44 (46%) for obese. When categorizing participants into nonoverweight or overweight, 1592 of 1698 (93.78%) were classified correctly (1407/1423, 98.88% among nonoverweight and 185/275, 67.3% among overweight).

Table 4. Frequencies of children in categories of BMI status according to measured and self-reported weight and height.

Measured	Self-reported, n				Total
	Underweight	Normal weight	Overweight	Obese	
Underweight	81	33	0	0	114
Normal weight	66	1227	16	0	1309
Overweight	0	90	139	2	231
Obese	0	0	24	20	44
Total	147	1350	179	22	1698

Figure 1. Histogram of difference between self-reported and measured a) height, b) weight, and c) BMI.

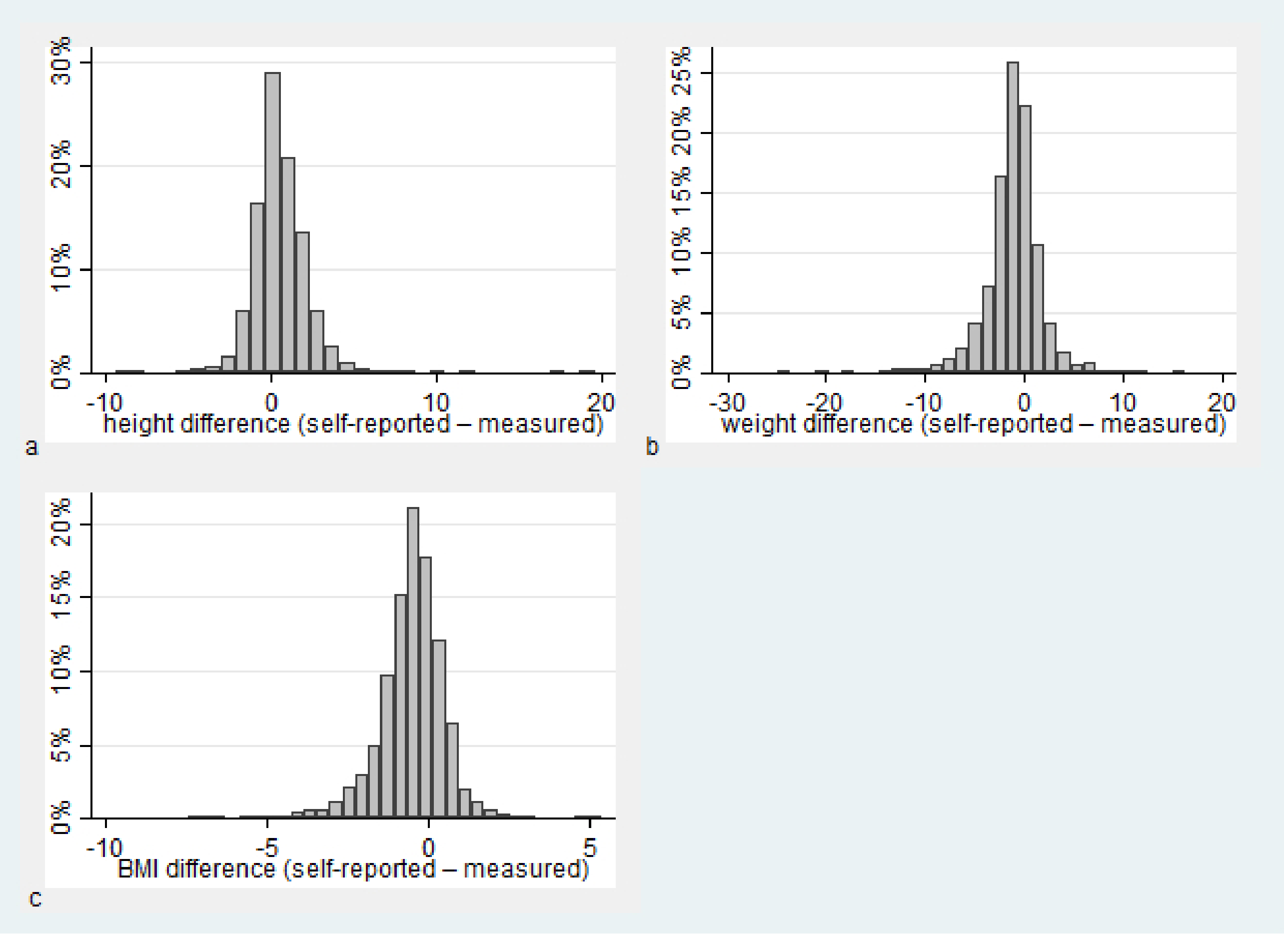


Figure 2. Scatterplot of self-reported (y-axis) and measured (x-axis) BMI (kg/m2), r=.94, P <.001 (N=1698).

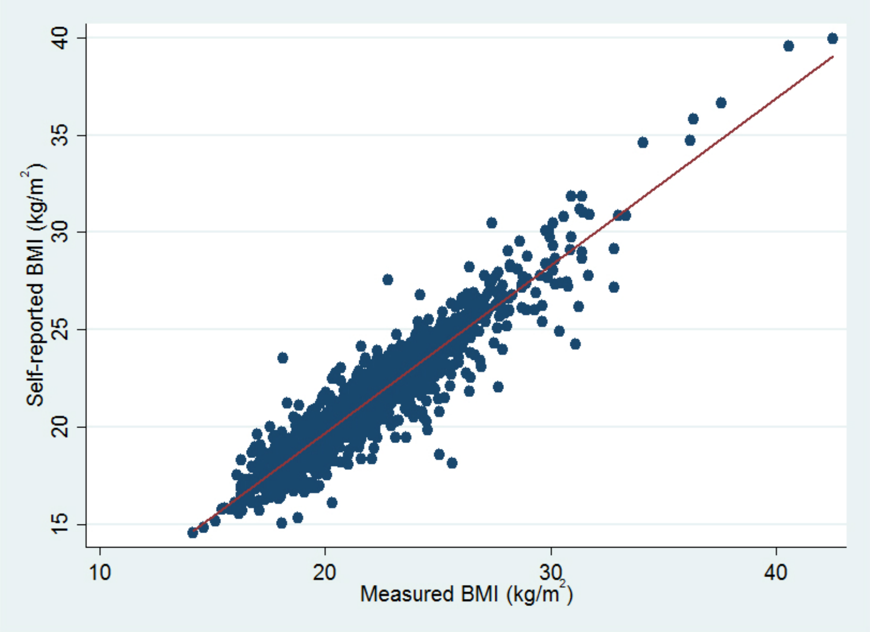
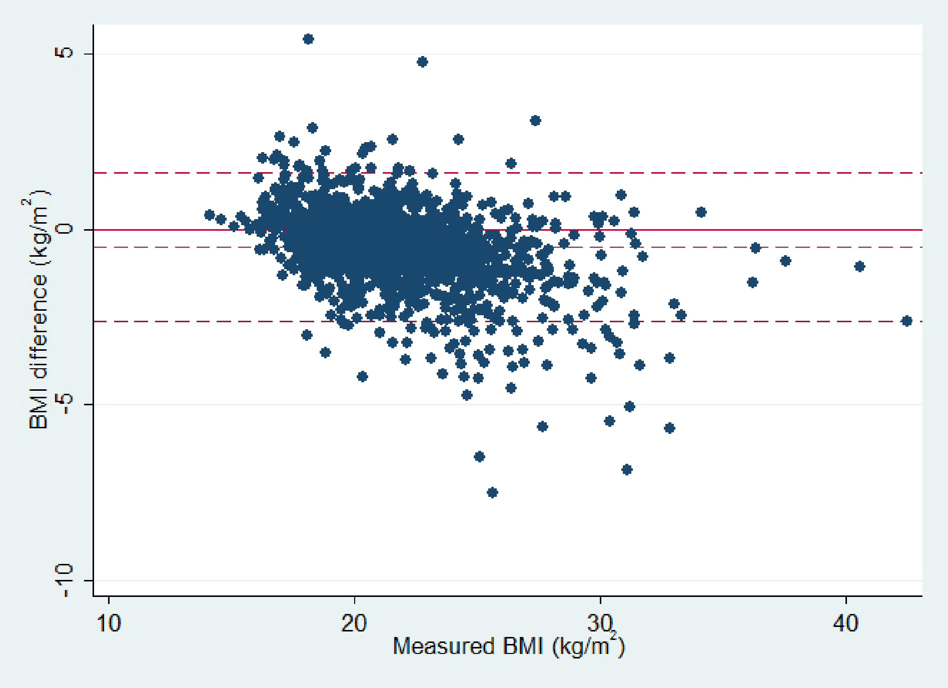


Figure 3. Bland-Altman plot of difference between self-reported and measured BMI (y-axis) in relation to measured BMI (x-axis). The solid line shows difference equal to zero and the dashed lines show mean difference ±2 SD.



Prediction Factors for Discrepancy Between Self-Reported and Measured Body Mass Index

To identify factors that predict validity of self-reported BMI, a multivariable linear regression model was built using a backward selection method. The final model (n=1337) included gender; BMI categories (based on measured BMI): underweight, overweight, and obese (normal weight as reference); vigorous physical activity (hours/week); and sedentary time (hours/day). Of these variables, only gender and BMI categories significantly predicted differences between self-reported and measured BMI. After mutual adjustment, the estimated mean difference showed that males underreported BMI to a lower extent than females (difference 0.4 kg/m^2) and that overweight and obese participants underreported BMI to a higher extent compared to normal-weight participants (difference 0.8 and 1.5 kg/m^2, respectively). Underweight participants, on the contrary, significantly overreported BMI compared to normal-weight participants (difference 0.6 kg/m^2). Vigorous physical activity and sedentary time were borderline significant predictors of difference between reported and measured BMI (more time associated with larger underreporting), although the absolute effect was very small.

Discussion

Principal Findings

We examined the validity of Web-based self-reported weight and height among Swedish adolescents aged approximately 16 years. Correlations between self-reported and measured height and weight were high; however, on average, weight was somewhat underestimated and height was slightly overestimated leading to an underestimation of BMI. Females underreported weight to a higher extent than males and overweight/obese participants underreported weight to a higher extent than normal-weight participants did. Overall, a high proportion of participants were classified into the correct BMI category; however, among overweight and obese participants, only 139 of 231 (60.2%) and 20 of 44 (46%) were correctly classified, respectively. Several lifestyle factors were tested as prediction factors for validity of self-reported BMI; however, after mutual adjustment, only gender and BMI categories significantly predicted discrepancy from measured BMI.

Comparison With Other Studies

Few studies have evaluated Web-based self-reported weight and height in adolescents. Storey et al [2] validated nutrients, physical activity, weight, and height among 459 adolescents aged 11-15 years in Canada using the Web-based questionnaire Web-SPAN. They found a mean difference between self-reported and measured weight (-2.5 kg), which was larger compared to our study. Similar to us, they found height to be slightly, although nonsignificantly, overestimated. Gender differences and the influence of BMI status were not investigated in this study [2]. Another rather small study [13] validated Web-based weight and height in 137 middle school and 242 high school Korean children. They observed mean differences between self-reported and measured weight and height that were similar to our study (weight difference from -1.1 to -1.7 kg and BMI difference from -0.5 to -0.7 kg/m^2, depending on gender and school grade). No differences were found between boys and girls. However, the study was not population-based and participants were told beforehand that their weight and height were going to be measured.

The results of our study are also comparable to validation studies using paper questionnaires [10,11,22-26], although varying results have been reported across countries, age groups, and gender. Sherry et al [27] summarized 11 validation studies on

US adolescents aged 12-19 years and concluded that self-reported data underestimate overweight prevalence and that there is a gender- and weight status-dependent bias. Compared to our findings, most of the studies included in this review observed somewhat larger discrepancies and lower correlations between estimated and measured weight and height. Two large studies in European settings [11,25], on the other hand, found somewhat smaller differences between reported and measured weight (-0.7 and -0.8 kg) compared to our population.

Most previous studies in adolescents observe larger underreporting of weight and BMI in females compared to males [11,25,28-30]. However, even if validity of self-reported BMI seems to be lower for females, the precision is possibly higher. In this study, this was supported by a lower standard deviation for weight differences and a higher proportion of females that could recall their height within 1 or 2 cm. Fonseca et al [22] similarly found that although mean differences between self-reported and measured BMI were larger in females, variability was lower. Jayawardene et al [26] additionally showed that being female was associated with correct reporting of BMI, whereas being male was associated with overreporting of BMI (using predefined cut-offs for under-, over-, and correct reporting).

Gender differences in the validity of self-reported BMI might have several explanations. Male adolescents experience faster growth and may, therefore, have less knowledge about current weight and height. However, anxiety over their own body size may be more common among females and the social desirability bias might be greater. A more narrow range but larger mean difference between reported and measured BMI would support the hypothesis that females in general are more aware of weight and height than males, although more often systematically underreport weight by 1 or 2 kg.

Several studies [11,13,25,26,31] have observed increasing underreporting with increasing BMI, whereas underweight participants in contrast tend to overreport weight. Consequently, high correlations between self-reported and measured BMI are generally reported whereas the sensitivity for overweight and obesity is low. This indicates that, for this age group, self-reported BMI might be better used as a continuous variable (eg, z score for children younger than 18 years) rather than divided into categories of overweight.

In this study, several lifestyle factors, such as physical activity, fruit and vegetable intake, and tobacco use, were tested as potential explanatory factors for the discrepancy between self-reported and measured BMI. The best fitting model included gender, BMI categories, physical activity, and sedentary time; however, only gender and BMI categories significantly predicted discrepancy between reported and measured BMI. Previous studies have shown conflicting results regarding determinant factors for validity of self-reported BMI. Jayawardene et al [26] found no difference in reporting capacity with physical activity, whereas screen time and fast food consumption were associated with overreporting of BMI. Bae et al [13], on the other hand, found that students who engaged in a high amount of moderate physical activity underestimated weight. Previous studies have shown that self-perceived body image [11,29], weight-loss intentions [13,26], or expressed concern over the own body silhouette and paying close attention to own figure [25] are associated with underreporting. These concerns might, or might not, be independent from overweight and may reflect a tendency to report toward social norms. In this study, we lacked information about these factors; however, we found no association between self-perceived health and reporting capacity.

Differences between studies might be explained by several factors. In this study, there were relatively few overweight participants, which possibly increases validity, compared to populations with higher overweight prevalence. Another factor may be the Swedish tradition of monitoring child growth development through school and health care, which could make families and children more aware of current weight and height. The validity of reported weight and height is likely also affected by the age of the population. Adolescents in the present study were older than in some of the compared studies and perhaps more interested or informed about their weight and height.

Strengths and Limitations

The main strength of the present study is the population-based design and the relatively large study sample. Height and weight were measured according to standard protocols by trained nurses and were checked for quality controls. In addition, information on a wide range of lifestyle factors (eg, blood pressure, physical activity, and dietary factors) have been collected which made it possible to test whether these affected the validity.

Using the Web for data collection has many advantages, including quick responses, easier administration, and reduction of manual data errors. Web-based questionnaires are more frequently replacing paper-based questionnaires in epidemiological studies; therefore, validation studies are needed to assess the accuracy.

Some limitations need to be addressed. Firstly, measurements were not obtained directly after self-reports. This could have led to small changes in height and weight during the period between the questionnaire and the measurements. To limit bias, we excluded participants with more than 8 weeks between self-reported and measured weight and height.

Secondly, when answering the questionnaire, participants were informed about the clinical investigation. However, it was not stated specifically that the investigation would include height and weight measurements; therefore, it is unlikely to affect the results.

Furthermore, weight was measured with light clothes on. This might explain some of the difference between self-reported and measured weight if participants usually weigh themselves without clothes. Weight could also fluctuate during the day and during the menstrual cycle for females. However, on a group level, these factors would balance each other out.

Lastly, as in all validation studies, the results may not be generalized to other populations as validity might vary with age and ethnicity. To be able to use and interpret self-reported height and weight, it is important to carry out validation studies in different countries and populations. To evaluate whether Web-based questionnaires are more valid than paper questionnaires or interviews, there is a need for more studies, ideally comparing randomized groups of interviews, Web questionnaires, and paper questionnaires.

Conclusions

Overall agreement between Web-based self-reported and measured weight and height indicate that Web-based BMI can be used as a valid, quick, and cost-effective substitute to measured BMI among Swedish adolescents. The accuracy of self-reported BMI declines with increasing BMI and self-reported BMI should not be used to estimate overweight/obesity prevalences.

Acknowledgments

We would like to thank all participants and staff in the BAMSE study. The study was funded by the Swedish Research Council, the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Heart and Lung Foundation, the Swedish Research Council for Health, Working Life and Welfare, the Swedish Research Council Formas, the Swedish Asthma and Allergy Foundation, the European Commission’s Seventh Framework 29 Program MeDALL under grant agreement No 261357, and the Strategic Research Program (SFO) in Epidemiology at Karolinska Institutet.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Comparison of lifestyle and demographic characteristics among adolescents included in the study population and total number of adolescents answering the 16 year questionnaire.

[PDF File (Adobe PDF File), 100KB - jmir_v17i3e73_app1.pdf]

References

1. Pursey K, Burrows TL, Stanwell P, Collins CE. How accurate is web-based self-reported height, weight, and body mass index in young adults? J Med Internet Res 2014;16(1):e4 [FREE Full text] [doi: 10.2196/jmir.2909] [Medline: 24398335]
2. Storey KE, McCargar LJ. Reliability and validity of Web-SPAN, a web-based method for assessing weight status, diet and physical activity in youth. J Hum Nutr Diet 2012 Feb;25(1):59-68. [doi: 10.1111/j.1365-277X.2011.01181.x] [Medline: 21615806]
3. Lassale C, Péneau S, Touvier M, Julia C, Galan P, Hercberg S, et al. Validity of web-based self-reported weight and height: results of the Nutrinet-Santé study. J Med Internet Res 2013;15(8):e152 [FREE Full text] [doi: 10.2196/jmir.2575] [Medline: 23928492]
4. Bonn SE, Trolle Lagerros Y, Bälter K. How valid are Web-based self-reports of weight? J Med Internet Res 2013;15(4):e52 [FREE Full text] [doi: 10.2196/jmir.2393] [Medline: 23570956]
5. Huber LRB. Validity of self-reported height and weight in women of reproductive age. Matern Child Health J 2007 Mar;11(2):137-144. [doi: 10.1007/s10995-006-0157-0] [Medline: 17066316]
6. Nyholm M, Gullberg B, Merlo J, Lundqvist-Persson C, Råstam L, Lindblad U. The validity of obesity based on self-reported weight and height: Implications for population studies. Obesity (Silver Spring) 2007 Jan;15(1):197-208. [doi: 10.1038/oby.2007.536] [Medline: 17228048]
7. Dijkshoorn H, Ujcic-Voortman JK, Viet L, Verhoeff AP, Uitenbroek DG. Ethnic variation in validity of the estimated obesity prevalence using self-reported weight and height measurements. BMC Public Health 2011;11:408 [FREE Full text] [doi: 10.1186/1471-2458-11-408] [Medline: 21624122]
8. Wen M, Kowaleski-Jones L. Sex and ethnic differences in validity of self-reported adult height, weight and body mass index. Ethn Dis 2012;22(1):72-78 [FREE Full text] [Medline: 22774312]
9. Dahl AK, Hassing LB, Fransson EI, Pedersen NL. Agreement between self-reported and measured height, weight and body mass index in old age--a longitudinal study with 20 years of follow-up. Age Ageing 2010 Jul;39(4):445-451 [FREE Full text] [doi: 10.1093/ageing/afq038] [Medline: 20453247]
10. Zhou X, Dibley MJ, Cheng Y, Ouyang X, Yan H. Validity of self-reported weight, height and resultant body mass index in Chinese adolescents and factors associated with errors in self-reports. BMC Public Health 2010;10:190 [FREE Full text] [doi: 10.1186/1471-2458-10-190] [Medline: 20384994]
11. Kurth BM, Ellert U. Estimated and measured BMI and self-perceived body image of adolescents in Germany: part 1 - general implications for correcting prevalence estimations of overweight and obesity. Obes Facts 2010 Jun;3(3):181-190. [doi: 10.1159/000314638] [Medline: 20616608]
12. 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]

13. Bae J, Joung H, Kim JY, Kwon KN, Kim Y, Park SW. Validity of self-reported height, weight, and body mass index of the Korea Youth Risk Behavior Web-based Survey questionnaire. J Prev Med Public Health 2010 Sep;43(5):396-402 [FREE Full text] [doi: 10.3961/jpmph.2010.43.5.396] [Medline: 20959710]
14. Wickman M, Kull I, Pershagen G, Nordvall SL. The BAMSE project: presentation of a prospective longitudinal birth cohort study. Pediatr Allergy Immunol 2002;13 Suppl 15:11-13. [Medline: 12688617]
15. Ekström S, Magnusson J, Kull I, Lind T, Almqvist C, Melén E, et al. Maternal body mass index in early pregnancy and offspring asthma, rhinitis and eczema up to 16 years of age. Clin Exp Allergy 2015 Jan;45(1):283-291 [FREE Full text] [doi: 10.1111/cea.12340] [Medline: 24807420]
16. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ 2000 May 6;320(7244):1240-1243 [FREE Full text] [Medline: 10797032]
17. Cole TJ, Flegal KM, Nicholls D, Jackson AA. Body mass index cut offs to define thinness in children and adolescents: international survey. BMJ 2007 Jul 28;335(7612):194 [FREE Full text] [doi: 10.1136/bmj.39238.399444.55] [Medline: 17591624]
18. Petersen AC, Crockett L, Richards M, Boxer A. A self-report measure of pubertal status: Reliability, validity, and initial norms. J Youth Adolesc 1988 Apr;17(2):117-133. [doi: 10.1007/BF01537962] [Medline: 24277579]
19. Cuzick J. A Wilcoxon-type test for trend. Stat Med 1985;4(1):87-90. [Medline: 3992076]
20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
21. Kirkwood BRSJ. Essential Medical Statistics. Oxford: Blackwell Science; 2003.
22. Fonseca H, Silva AM, Matos MG, Esteves I, Costa P, Guerra A, et al. Validity of BMI based on self-reported weight and height in adolescents. Acta Paediatr 2010 Jan;99(1):83-88. [doi: 10.1111/j.1651-2227.2009.01518.x] [Medline: 19878130]
23. Rasmussen M, Holstein BE, Melkevik O, Damsgaard MT. Validity of self-reported height and weight among adolescents: the importance of reporting capability. BMC Med Res Methodol 2013;13:85 [FREE Full text] [doi: 10.1186/1471-2288-13-85] [Medline: 23805955]
24. Legleye S, Beck F, Spilka S, Chau N. Correction of body-mass index using body-shape perception and socioeconomic status in adolescent self-report surveys. PLoS One 2014;9(5):e96768 [FREE Full text] [doi: 10.1371/journal.pone.0096768] [Medline: 24844229]
25. Béghin L, Huybrechts I, Ortega FB, Coopman S, Manios Y, Wijnhoven TMA, et al. Nutritional and pubertal status influences accuracy of self-reported weight and height in adolescents: the HELENA Study. Ann Nutr Metab 2013;62(3):189-200. [doi: 10.1159/000343096] [Medline: 23485769]
26. Jayawardene W, Lohrmann D, YoussefAgha A. Discrepant body mass index: behaviors associated with height and weight misreporting among US adolescents from the National Youth Physical Activity and Nutrition Study. Child Obes 2014 Jun;10(3):225-233. [doi: 10.1089/chi.2014.0002] [Medline: 24828965]
27. Sherry B, Jefferds ME, Grummer-Strawn LM. Accuracy of adolescent self-report of height and weight in assessing overweight status: a literature review. Arch Pediatr Adolesc Med 2007 Dec;161(12):1154-1161. [doi: 10.1001/archpedi.161.12.1154] [Medline: 18056560]
28. Brener ND, Mcmanus T, Galuska DA, Lowry R, Wechsler H. Reliability and validity of self-reported height and weight among high school students. J Adolesc Health 2003 Apr;32(4):281-287. [Medline: 12667732]
29. Brettschneider AK, Rosario AS, Ellert U. Validity and predictors of BMI derived from self-reported height and weight among 11- to 17-year-old German adolescents from the KiGGS study. BMC Res Notes 2011;4:414 [FREE Full text] [doi: 10.1186/1756-0500-4-414] [Medline: 22005143]
30. Clarke P, Sastry N, Duffy D, Ailshire J. Accuracy of self-reported versus measured weight over adolescence and young adulthood: findings from the national longitudinal study of adolescent health, 1996-2008. Am J Epidemiol 2014 Jul 15;180(2):153-159. [doi: 10.1093/aje/kwu133] [Medline: 24944288]
31. Lee B, Chung SJ, Lee SK, Yoon J. Validation of self-reported height and weight in fifth-grade Korean children. Nutr Res Pract 2013 Aug;7(4):326-329 [FREE Full text] [doi: 10.4162/nrp.2013.7.4.326] [Medline: 23964321]

Abbreviations

BMI: body mass index

Edited by G Eysenbach; submitted 16.10.14; peer-reviewed by C Hawke; comments to author 23.01.15; revised version received 29.01.15; accepted 04.02.15; published 18.03.15.

Please cite as:
Ekström S, Kull I, Nilsson S, Bergström A
Web-Based Self-Reported Height, Weight, and Body Mass Index Among Swedish Adolescents: A Validation Study
J Med Internet Res 2015;17(3):e73
URL: http://www.jmir.org/2015/3/e73/
doi:10.2196/jmir.3947
PMID:25791395

Original Paper

Use of a Web-Based Physical Activity Record System to Analyze Behavior in a Large Population: Cross-Sectional Study

Hideyuki Namba[1,2], PhD, MEng; Yosuke Yamada[2], PhD; Mika Ishida[3], RD; Hideto Takase[3], PhD; Misaka Kimura[4], PhD

[1]Department of Health and Nutrition, Wayo Women's University, Ichikawa, Chiba, Japan
[2]National Institute of Health and Nutrition, Tokyo, Japan
[3]Kao Corporation, Tokyo, Japan
[4]Faculty of Health and Medical Sciences, Kyoto Gakuen University, Kyoto, Japan

Corresponding Author:
Hideyuki Namba, PhD, MEng
Department of Health and Nutrition
Wayo Women's University
2-3-1 Konodai
Ichikawa, Chiba, 272-8533
Japan
Phone: 81 47 371 2160
Fax: 81 47 371 2556
Email: h-nanba@wayo.ac.jp

Abstract

Background: The use of Web-based physical activity systems has been proposed as an easy method for collecting physical activity data. We have developed a system that has exhibited high accuracy as assessed by the doubly labeled water method.

Objective: The purpose of this study was to collect behavioral data from a large population using our Web-based physical activity record system and assess the physical activity of the population based on these data. In this paper, we address the difference in physical activity for each urban scale.

Methods: In total, 2046 participants (aged 30-59 years; 1105 men and 941 women) participated in the study. They were asked to complete data entry before bedtime using their personal computer on 1 weekday and 1 weekend day. Their residential information was categorized as urban, urban-rural, or rural. Participant responses expressed the intensity of each activity at 15-minute increments and were recorded on a Web server. Residential areas were compared and multiple regression analysis was performed.

Results: Most participants had a metabolic equivalent (MET) ranging from 1.4 to 1.8, and the mean MET was 1.60 (SD 0.28). The median value of moderate-to-vigorous physical activity (MVPA, ≥3 MET) was 7.92 MET-hours/day. A 1-way ANCOVA showed that total physical activity differed depending on the type of residential area ($F_{2,2027}$=5.19, *P*=.006). The urban areas (n=950) had the lowest MET-hours/day (mean 37.8, SD, 6.0), followed by urban-rural areas (n=432; mean 38.6, SD 6.5; *P*=.04), and rural areas (n=664; mean 38.8, SD 7.4; *P*=.002). Two-way ANCOVA showed a significant interaction between sex and area of residence on the urban scale ($F_{2,2036}$=4.53, *P*=.01). Men in urban areas had the lowest MET-hours/day (MVPA, ≥3 MET) at mean 7.9 (SD 8.7); men in rural areas had a MET-hours/day (MVPA, ≥3 MET) of mean 10.8 (SD 12.1, *P*=.002). No significant difference was noted in women among the 3 residential areas. Multiple regression analysis showed that physical activity consisting of standing while working was the highest contributor to MVPA, regardless of sex.

Conclusions: We were able to compile a detailed comparison of physical activity because our Web-based physical activity record system allowed for the simultaneous evaluation of physical activity from 2046 Japanese people. We found that rural residents had greater total physical activity than urban residents and that working and transportation behaviors differed depending on region type. Multiple regression analysis showed that the behaviors affected MVPA. People are less physically active while working, and sports and active transportation might be effective ways of increasing physical activity levels.

(J Med Internet Res 2015;17(3):e74) doi:10.2196/jmir.3923

KEYWORDS

motor activity; behavior; Internet; computer systems; metabolic equivalent; Japan

Introduction

Despite known hazards, physical inactivity continues to be a major risk for the development of chronic diseases [1]. It is associated with an increased incidence of obesity, diabetes, cardiovascular disease, osteoporosis, and cancer [2-4]. Therefore, instruments that allow for accurate measurement of physical activity in the general population are needed. Further, accurately determining what factors affect physical activity is crucial. Measurement of physical activity via the Internet may help to compensate for the weaknesses of previous measurement methods.

In epidemiological studies, questionnaires are the most frequently used instruments with which physical activity and energy expenditure are estimated [5]. However, the validity of traditional self-reported questionnaires is lower than that of the gold standard measurement technique: the doubly labeled water (DLW) method [6]. Accelerometers are readily available and have adequate memory for long-term data collection. Although the accelerometer is an excellent tool for physical activity assessment, it is not accurate for cycling, swimming, or activities involving only the upper limbs [5]. Furthermore, its cost precludes it from being suitable for epidemiological studies that require physical activity or energy expenditure measurements in large populations. The DLW method is one of the most accurate and valid systems for evaluating total energy expenditure under free-living conditions [7]. However, it is too costly for epidemiological studies. Therefore, it is necessary to develop new and inexpensive methods for evaluating the physical activity of large populations.

Three recent studies compared physical activity measurements between Web-based systems and the DLW method and found that the validity of the Web-based systems was equivalent to that of traditional questionnaires [8-10]. The simplified recording of physical activity every 15 minutes showed high validity compared with the DLW method with respect to energy expenditure [11,12]. We also developed a Web-based physical activity measurement system and compared its accuracy with that of the DLW method [10]. In our system, participants chose their behavior from a menu on the left of the screen and then filled in a table on the right side based on recollection of their behaviors for each 15-minute interval throughout a 24-hour period.

The use of Web-based measurement systems has been suggested as an easier method for collecting self-reported physical activity data than traditional questionnaires. These measurement systems are low cost, practical, and data can be collected quickly (input time of less than 10 minutes). Web-based measurement systems can measure daily physical activity in many people simultaneously because they are compatible with multiple Web browsers and computer operating systems that can be used to complete the assessments via the Internet. The successful and systematic collection of demographic and lifestyle data are central to any epidemiological study [13]. Studies using questionnaires have shown that neighborhoods and social environments are associated with physical activity [14-16]. In the United States, reports showed that rural residents had lower levels of physical activity than urban residents, and this was suggested to be an important contributor to their relatively poorer health [17,18]. However, another recent study that used accelerometer-measured physical activity indicated that total physical activity was higher in rural residents compared with urban residents, but that the reverse was true when considering high-intensity activity [19]. Although epidemiological studies that use accelerometers to collect physical activity data from large populations of people could be done by a national organization, they could also be done using the Internet at a much lower cost. Because comparing urban and rural physical activity has yielded conflicting results, examining subcategories of behavior might be informative.

The purpose of this study was to collect behavioral records from a large population using our Web-based physical activity system and compare the physical activity of people living in residential environments that differed in how urban/rural they were.

Methods

Recruitment

A total of 2566 people aged 30-59 years, including 1438 men (56.04%) and 1128 women (43.96%), participated in this study. A request to participate in this survey was made to 9134 of approximately 4,450,000 people who had registered with an Internet research company between July 12 and August 8, 2012. When registering, people were required to provide personal information, such as their residential area, age, and sex. We requested that the Internet company send surveys to 9134 of their registered customers based on our criteria for residence, age, and sex. Of the 9134 people who received a cross-sectional survey, 2566 responded (28.09% cooperation rate). The activity records included 1 weekday and 1 weekend day.

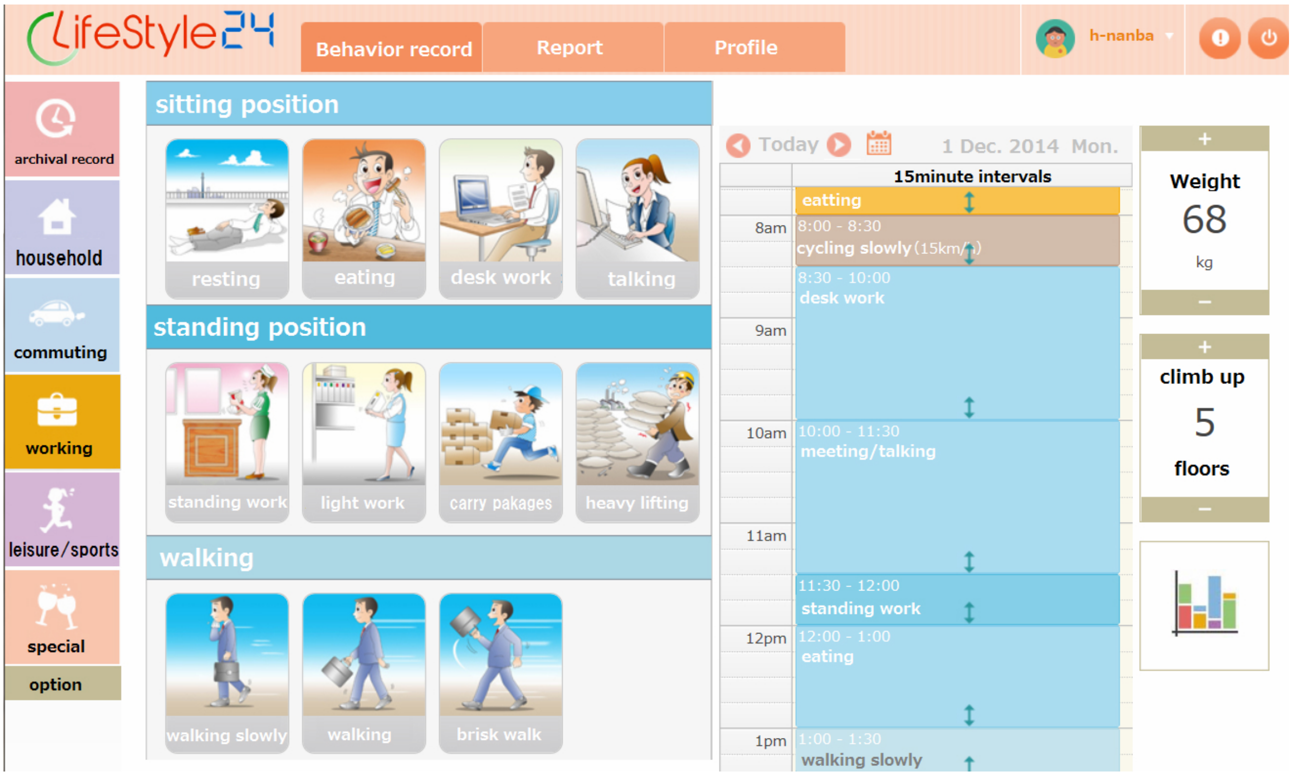
All participants were categorized by age, sex, and residential area. Residential areas were divided into urban (cities around Tokyo), urban-rural (cities with a population >300,000), or rural (municipalities with a population <50,000). The Internet research company provided a URL address in an email and requested that each participant fill out a survey. Each participant received the equivalent of approximately US $3 for completing the survey. The study was conducted with approval by the ethics committees of Fukuoka University in Fukuoka prefecture and Wayo Women’s University in Chiba, Japan. All participants read an explanation of the purpose and content of the study by email and indicated their understanding before providing written consent (via email) to participate.

Web-Based Physical Activity Records

Figure 1 shows the screen of the Web-based physical activity records system [20]. Participants were requested to respond before bedtime on 2 days (1 weekday and 1 weekend day) using their personal computer. The Web-based record used the diary method [11] to collect the recalled activities over the last 24 hours. Using this method, the participants recorded preset activities performed in each of 4 different categories (household, transportation, work-related, and leisure/sports) for each 15-minute interval of a 24-hour day for cross tabulation. Participants could choose from 91 activities among the 4 categories. We used the intensity of each activity as measured in metabolic equivalent (MET) [21] to quantify participant responses. Thus, the data used for analysis were two 24-hour records of MET values, each with 96 values (every 15 minutes).

The mean MET over a 24-hour period was calculated as (weekday × 5 + weekend day × 2) / 7 using equations listed in Figure 2. We also calculated the mean moderate-to-vigorous physical activity (MVPA, MET-hours/day) over a 24-hour period using a similar formula (Figure 2). In this case, we summed MET ≥3, as defined by the EPAR2006 guidelines for MVPA in Japan [22].

Figure 1. Screenshot of the Web-based physical activity record system.



eatting

8:00 - 8:30 cycling slowly (15km/h)

8:30 - 10:00 desk work

10:00 - 11:30 meeting/talking

11:30 - 12:00 standing work

12:00 - 1:00 eating

1:00 - 1:30 walking slowly

Figure 2. Equations for calculating mean metabolic equivalents (METS) and moderate-to-vigorous physical activity (MVPA) over a 24-hour period.

Mean METs over a 24-hour period:

$$MET_{24h\ average} = \frac{(5 \times \sum_{1}^{96} weekday\ MET_i) + (2 \times \sum_{1}^{96} weekend\ day\ MET_i)}{7 \times 96}$$

Mean MVPA (MET-hours/day) over a 24-h period:

$$MET\,3 = \begin{cases} 0 & (MET < 3) \\ MET & (MET \geq 3) \end{cases}$$

$$MVPA = \sum_{1}^{96} MET3_i \;/4$$

$$MVPA_{average} = \frac{(5 \times MVPA_{weekday}) + (2 \times MVPA_{weekend\ day})}{7}$$

Statistical Analysis

All statistical analyses were performed using SPSS for Windows 20.0 (IBM Corp, Armonk, NY, USA). All data are shown as mean and standard deviation (SD). Comparison of residential area type (a general characteristic) was performed by 1-way ANOVA and chi-square test. Comparison of the total physical activity, MVPA, and the 4 categories of physical activity in each region was performed by 1-way ANOVA and 1-way ANCOVA, and multiple comparisons were subsequently carried out by Tukey honestly significant difference test. A significant difference was defined as *P*<.05. We also compared time spent commuting and working depending on residential area using a 1-way ANCOVA. Multiple regression analysis was performed to investigate the impact of the behaviors on total physical activity (MET-hours) and on the 24-hour MVPA (MET-hours) for the total sample, men, and women. The standardized coefficients (βs) were calculated.

Results

Table 1 displays the characteristics of the study population. As intended, we recruited a heterogeneous group (3 residential areas, both sexes, and a wide age range), and this is demonstrated in the table. Of the 2566 participants, 520 (20.26%) were excluded from analysis because they did not provide complete datasets (one 24-hour weekday and one 24-hour weekend day). The regional breakdown for the remaining 2046 participants was 950 in urban areas (46.43%), 432 in urban-rural areas (21.11%), and 664 in rural areas (32.45%).

Table 1. Characteristics of the study population (N=2046).

Characteristics	Urban (n=950)	Urban-Rural (n=432)	Rural (n=664)	*P*
Age (years), mean (SD)	43.9 (7.9)	43.4 (7.2)	43.6 (7.4)	.47
Weight (kg), mean (SD)	61.3 (12.7)	62.0 (12.6)	62.3 (13)	.27
BMI (kg/m^2), mean (SD)	22.3 (3.5)	22.7 (3.5)	22.3 (3.7)	.13
Gender, n (%)				
Men	511 (53.8)	220 (50.9)	374 (56.3)	.21
Women	439 (46.2)	212 (49.1)	290 (43.7)	
Age range (years), n (%)				
30-39	331 (34.8)	156 (36.1)	221 (33.3)	.003
40-49	326 (34.3)	174 (40.3)	280 (42.2)	
50-59	293 (30.8)	102 (23.6)	163 (24.5)	
Custom of periodical exercise, n (%)				
No	621 (65.4)	272 (63.0)	443 (66.7)	.09
>30 minutes twice a week	329 (34.6)	160 (37.0)	221 (33.3)	
Occupation, n (%)				
Employee	522 (54.9)	238 (55.1)	342 (51.5)	.07
Self-employed	59 (6.2)	28 (6.5)	62 (9.3)	
Full-time housewife	142 (14.9)	56 (13.0)	78 (11.7)	
Housewife (includes part-time job)	114 (12.0)	47 (10.9)	82 (12.3)	
Student	4 (0.4)	1 (0.2)	0 (0.0)	
Unemployed	50 (5.3)	28 (6.5)	37 (5.6)	
Other	59 (6.2)	34 (7.9)	63 (9.5)	
Household income (million yen), n (%)				
<3	123 (12.9)	86 (19.9)	121 (18.2)	<.001
3-5	204 (21.5)	129 (29.9)	198 (29.8)	
5-7	245 (25.8)	103 (23.8)	175 (26.4)	
7-10	217 (22.8)	68 (15.7)	103 (15.5)	
>10	161 (16.9)	46 (10.6)	67 (10.1)	

There were no significant differences among the 3 regions with respect to mean age (*P*=.47), weight (*P*=.27), or BMI (*P*=.13). Detailed characteristics of the study population were also investigated. These included gender, age, exercise habits, occupation, and household income in each respective region. We found no differences among the 3 regions with respect to gender, exercise habits, or occupation. However, some significant differences were found with respect to age. Among the age group between 50-59 years, 30.8% (293/950) were from urban areas, 23.6% (102/432) from urban-rural areas, and 24.5%

(163/664) from rural areas (*P*=.003). Additionally, household income was significantly higher in urban areas than in other regions (*P*<.001).

Figure 3 shows the distribution of the 24-hour mean MET values among all participants. Most participants had a mean MET between 1.4 and 1.8, and the mean 24-hour MET was 1.60 (SD 0.28). Figure 4 shows the MVPA distribution among all participants. Most participants had an MVPA between 0 and 10 MET-hours; the median 24-hour MVPA value was 7.92 MET-hours (IQR 3.6-14.7).

Table 2 compares total physical activity, MVPA, the 4 categories of physical activity, and detailed categorical activity across region type. One-way ANOVA revealed significant differences in total physical activity (24-hour mean MET) depending on area of residence ($F_{2,2043}$=7.26, *P*<.001), and a 1-way ANCOVA showed that when adjusted for age, gender, occupation, household income, and marriage, the total physical activity (MET-hours/day) differed across region type ($F_{2,2027}$=5.19, *P*=.006). Multiple comparison analysis showed that urban areas had the lowest 24-hour mean MET (mean 1.57, SD 0.25), followed by urban-rural areas (mean 1.62, SD 0.27; *P*=.02), and rural areas (mean 1.63, SD 0.31; *P*<.001). No significant differences were found in MVPA across regions.

A 1-way ANOVA of the 4 categories of activity showed that transportation ($F_{2,2043}$=32.87, *P*<.001) and working ($F_{2,2043}$=17.75, *P*<.001) significantly differed depending on region type. When we broke the transportation category down into subcategories, a 1-way ANCOVA revealed that the amount of time spent walking to work ($F_{2,2028}$=115.8, *P*<.001), on public transit ($F_{2,2028}$=179.4, *P*<.001), and sitting in a car ($F_{2,2028}$=89.51, *P*<.001) significantly differed across region type. Multiple comparison analysis showed that urban areas had the longest walking time, followed by urban-rural and rural regions (urban: mean 24.6, SD 25.1 min; urban-rural: mean 7.8, SD 24.9 min; *P*<.001; rural: mean 7.4, SD 25.0 min; *P*<.001). Time spent sitting in cars was lowest in urban areas (mean 21.9, SD 52.0 min; *P*<.001), followed by urban-rural areas (mean 51.5, SD 51.9 min) and rural areas (mean 52.8, SD 51.9 min). There were no significant differences in time spent sitting at work among the 3 region categories. Time spent standing at work was longest in rural areas (mean 89.7, SD 152.4 min; *P*<.002) vs urban-rural, followed by urban-rural areas (mean 67.0, SD 125.5 min; *P*<.001) vs urban, and urban areas (mean 44.7, SD 106.2 min; *P*<.001) vs rural.

Two-way ANCOVA adjusted for age, occupation, household income, and marriage showed a significant interaction between sex and area of residence on the urban scale ($F_{2,2036}$=4.53, *P*=.01), and a significant main effect was observed for sex ($F_{1,2036}$=20.98, *P*<.001). Men in urban areas had the lowest MET-hours/day (mean 7.9, SD 8.7); men in rural areas had a mean MET-hours/day of 10.8 (SD 12.1, *P*=.002). No significant difference was noted in women among the 3 residential areas.

Figure 5 shows the results of the multiple regression analysis for MVPA (MET-hours/day, ≥3.0 MET) against the subcategories of activities, and plots the resulting standardized coefficients (βs). We analyzed the total sample (R^2=.95, *P*<.001, standard error of the estimate [SEE]=2.08, n=2046), just the men (R^2=.96, *P*<.001, SEE=2.24, n=1105), and just the women (R^2=.94, *P*<.001, SEE=1.86, n=941).

The independent variable here is physical activity (MET-hours/day) and includes varying intensities. In the total sample, the highest contribution was made by working while standing, followed in order by vigorous work, sport and exercise, cooking, cleaning the house, cycling as transportation, washing the car and gardening, childcare, time spent walking on the way to work, and time spent walking during work. Physical activities that were weighted differently between men and women were vigorous working, cooking, cleaning the house, and childcare.

Table 2. Comparison of physical activity in each region type.

Types of physical activity	Urban, mean (SD) (n=950)	Urban-Rural, mean (SD) (n=432)	Rural, mean (SD) (n=664)	*F* (*df*)	*P*
Total physical activity					
24-Hour mean MET	1.57 (0.25)	1.62 (0.27)	1.63 (0.31)	7.26 (2,2043)	<.001
24-Hour MET-hours/day	37.9 (6.0)	38.8 (6.5)	39.1 (7.5)	7.26 (2,2043)	<.001
24-Hour MET-hours/day[a]	37.8 (6.0)	38.6 (6.5)	38.8 (7.4)	5.19 (2,2027)	.006
Moderate-to-vigorous physical activity					
24-Hour MET-hours/day ≥3.0 MET	9.8 (8.5)	10.7 (9.4)	11.3 (10.6)	5.58 (2,2043)	.04
24-Hour MET-hours/day[a] ≥3.0 MET	10.0 (9.3)	10.4 (9.3)	11.2 (9.3)	3.23 (2,2038)	.04
Categories of physical activity					
Household physical activity					
MET-hours/day	5.6 (6.5)	6.5 (7.3)	5.9 (6.4)	2.73 (2,2043)	.07
Time (min)	118.0 (135.6)	134.0 (150.0)	121.6 (132.0)	2.02 (2,2043)	.13
Transportation physical activity					
MET-hours/day	4.1 (3.2)	3.1 (3.1)	3.0 (2.7)	32.87 (2,2043)	<.001
Time (min)	99.7 (78.9)	71.1 (67.8)	73.8 (69.4)	34.16 (2,2043)	<.001
Working physical activity					
MET-hours/day	8.0 (7.4)	9.5 (9.0)	10.6 (10.4)	17.75 (2,2043)	<.001
Time (min)	253.0 (202.0)	269.1 (202.7)	277.8 (204.7)	3.06 (2,2043)	.047
Leisure/Sports activity					
MET-hours/day	9.8 (5.5)	9.4 (4.6)	9.3 (4.9)	2.67 (2,2043)	.07
Time (min)	394.6 (187.4)	393.6 (190.4)	389.0 (192.8)	0.18 (2,2043)	.84
Detail of transportation					
Walking time while commuting[a] (min)	24.6 (25.1)	7.8 (24.9)	7.4 (25.0)	115.8 (2,2028)	<.001
Public transit while commuting[a] (min)	45.6 (42.4)	7.6 (42.3)	11.9 (42.2)	179.4 (2,2028)	<.001
Sitting in a car while commuting[a] (min)	21.9 (52.0)	51.5 (51.9)	52.8 (51.9)	89.51 (2,2028)	<.001
Detail of working					
Sitting time of working[a] (min)	203.3 (196.9)	213.2 (196.3)	192.7 (185.2)	1.89 (2,2028)	.15
Standing time of working[a] (min)	44.7 (106.2)	67.0 (125.5)	89.7 (152.4)	17.77 (2,2028)	<.001

[a] Adjusted for age, gender, occupation, household income, and marriage.

Figure 3. Distribution of 24-hour mean metabolic equivalents (MET) for 2046 participants.

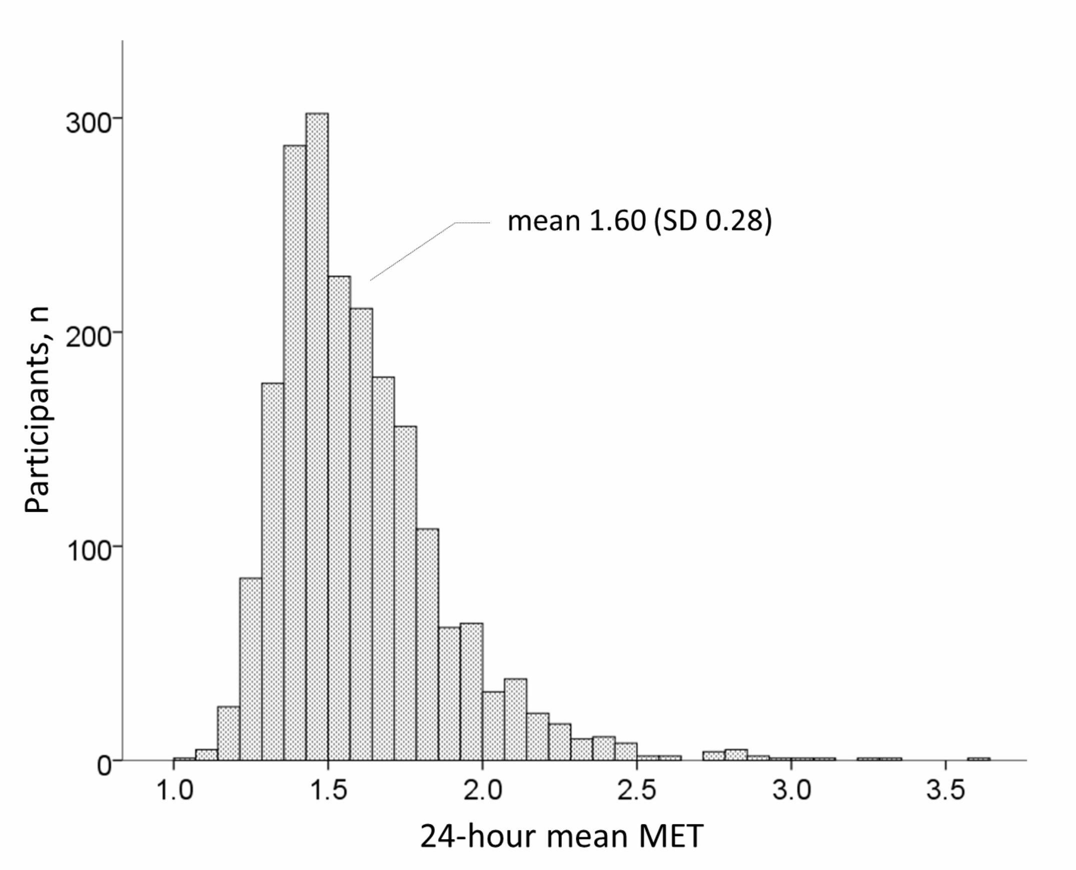


Figure 4. Distribution of 24-hour moderate-to-vigorous physical activity (MVPA) for 2046 participants.

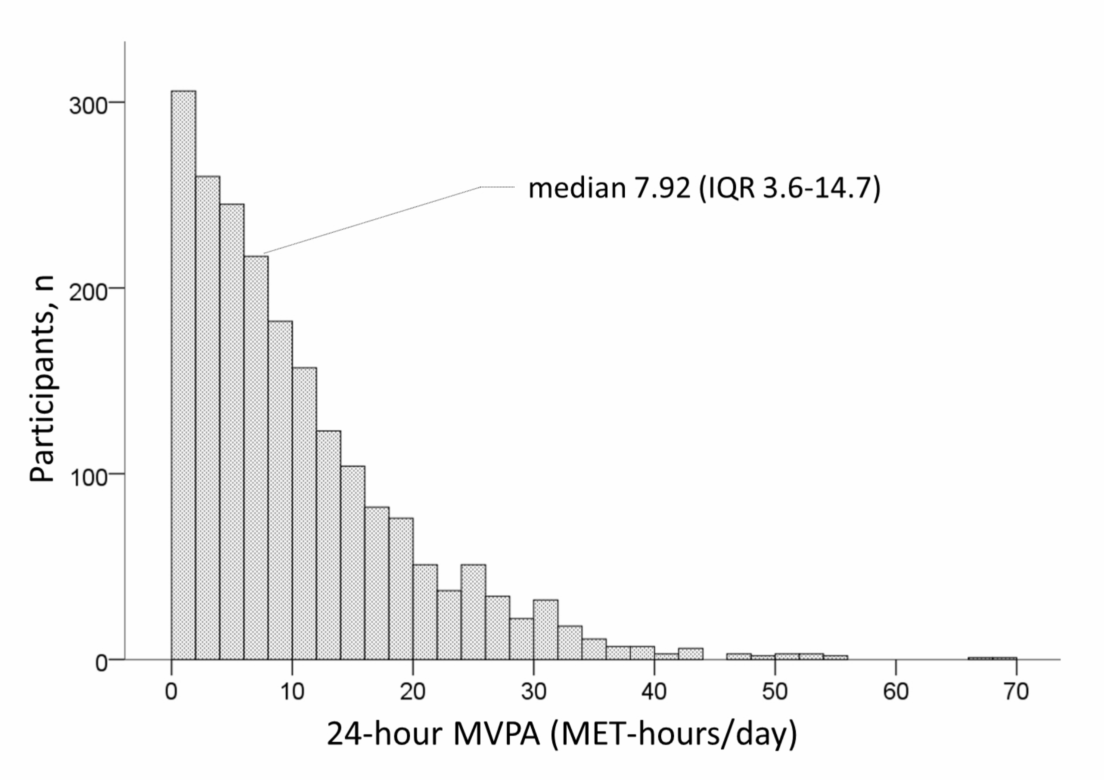
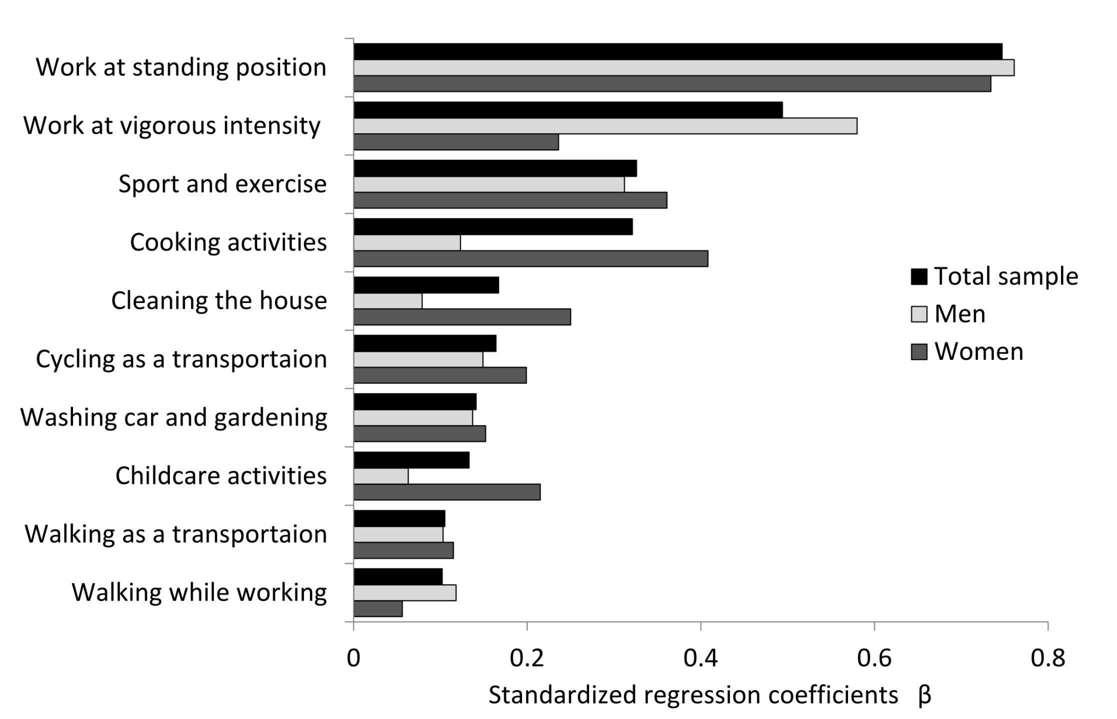


Figure 5. Multiple regression analysis for moderate-to-vigorous physical activity in men and women.



Discussion

In this study, we succeeded in simultaneously collecting data via the Internet from 2046 people using the Web-based physical activity system that we developed. The data were used to perform a behavior analysis and comparison of physical activity by residential area type. The percentage of people who fully cooperated with the request to participate in data collection was only 22.4%. Therefore, although large-scale collection of physical activity data is possible using Internet research, challenges remain.

The accuracy of the simplified physical activity record has been confirmed by the DLW method [11,12] and good agreement was observed using the Web-based physical activity measurement system [10]. Although recording behavior using paper requires many procedures, using a Web-based system allows participants to complete the record in a relatively short time and responses from many participants can be collected and managed by the data server simultaneously. Figure 3 shows that the mean 24-hour MET value is similar to the mean value obtained from Japanese participants using the DLW method [23]. Figure 3 indicates that the median 24-hour mean MVPA (MET-hours/day) was 7.92, which we consider to be higher than the mean 25.0 (SD 14.7) MET-hours/week (ie, approximately 3.5 MET-hours/day) reported using an accelerometer in 1837 Japanese participants [24]. Therefore, the MVPA value using the present system may have been overestimated. For example, 10-minute activities would be counted as 15 minutes because our minimum time interval was 15 minutes. Furthermore, participants may have overestimated their actual behavior through inaccurate recall. However, the system described here allows assessment of swimming, cycling, and climbing activities, which cannot be measured using an accelerometer. The risk of overestimating physical activity using the recall method has been previously described [6], but further studies of diary methods are needed to clarify the efficacy of assessment using accelerometers and global positioning systems.

Comparing the activity data with how urban/rural the residential area was showed that people in urban areas had lower MET-hours values than those in rural and urban-rural areas. Associations of neighborhood walkability attributes with commuting by walking have been confirmed in Australia [14]. With respect to environment, physical activity has been associated with transportation, population density, recreational activities, and connectivity [25]. In this study, although the time spent walking to work among participants in urban areas was a mean 24 (SD 25.3) minutes, it was only a mean 7.4 (SD 25.3) minutes in rural areas. Time spent commuting by car or public transportation has been shown to differ significantly between urban and rural residents. That people in urban areas had higher physical activity related to transportation than did those living in rural areas is consistent with data from the National Health and Nutrition Examination Survey (NHANES) [26].

The duration of time spent standing while working was shorter in urban areas than in urban-rural and rural areas. This can be explained by the fact that urban areas exhibited lower MET values and potentially less time standing. The duration of time spent sitting while working among men was longer in urban than in rural areas. This can be explained by the fact that men in urban areas exhibited lower MVPA values and potentially longer sedentary times. Therefore, there is a possibility that work behavior, when such behavior accounts for most of the day, may affect the physical activity of the day. However, no difference in the MVPA according to city size was observed in women. Occupational activity significantly contributes to overall physical activity in British adults [27]. Researchers in the United Kingdom have shown that deindustrialization is related to a decrease in physical activity in a region [28]. Therefore, when work behavior accounts for most activity within a day, it should have a high impact on the physical activity for that day.

Multiple regression analysis was performed to determine which behaviors affect the MVPA. In the total sample, standing while working was the category that most strongly affected MVPA. Behaviors related to working (standing while working, vigorous working, and walking while working) accounted for 3 of the top 10 behaviors that affected MVPA. Household and childcare activities (childcare, cleaning, and cooking) accounted for another 3, leisure and sport activities (sports/exercise, washing car and gardening) accounted for 2, and commuting (walking and bicycling) accounted for the remaining 2 (Figure 5). Therefore, to increase physical activity, activity while working must be intentionally performed, and doing sports, exercising, commuting, or doing household activities in daily life is important. The idea that work behavior strongly influences physical activity is a perspective consistent with previous research [27-29]. Characteristic physical activities for men included working vigorously, and those for women were cooking, cleaning, and childcare. These activities might be difficult components to change in Japan because they depend on gender cultural roles. Therefore, recommendations regarding sports activities and changes in commuting behavior (ie, switch from driving to cycling or walking) might be useful for preventing lifestyle-related diseases and for improving physical fitness. However, because a recent study demonstrated that health risks are associated with sitting for a long period despite an adequately high MVPA [30], reducing time spent sitting might be an additional point to consider if one wants to increase the 24-hour mean MET.

Some limitations of the present study should be noted. In Japan, more than 95% of people aged 30-49 years and 91.4% of people aged 50-59 years have access to the Internet, whereas only 81.7% of people (all ages) have access to a personal computer [31]. Because the system described here can only be used on a personal computer, the potential use of smartphones should be explored in further studies. The low cooperation rate in this Internet survey might have led to self-selection bias in which those who chose to cooperate were typically either more or less active than those who did not. Combining the Internet version with conventional mailing may be an effective way to improve the cooperation rate.

In conclusion, the present Web-based physical activity records system allowed for the simultaneous evaluation of physical activity in 2046 Japanese individuals. The system also ensured that behavioral data were collected for each 15-minute interval within a 24-hour period. Prior Web-based activity record systems have been developed to select the behavior record of character information [10]; we have developed a method by behavior record illustrations on the Web screen in the new system. New Web and smartphone systems that use illustrations may allow for the spread of such systems to a large number of users. We compared the physical activity of residents in urban vs rural regions, and examined 4 categories of physical activity. We found that working and transportation behavior differed depending on region type. Multiple regression analysis showed that physical activity while working was the factor that most affected MVPA, whereas men and women differed in how much vigorous working, cooking, cleaning, and childcare affected their MVPAs.

Future research should investigate the development of systems for measuring behavioral change based on behavior records. The characteristics of the present Web-based physical activity records system are that it can evaluate details of individual lifestyles over a 24-hour period through the recall of activity in 15-minute intervals. It is important to note that behavior can be changed to improve individual lifestyles.

Acknowledgments

The authors thank all participants for their cooperation in this study. This work was supported by a Grant-in-Aid for Young Scientists (B) to the first author (No. 24700740).

Conflicts of Interest

The authors have applied for a patent for PHYSICAL ACTIVITY MEASURING SYSTEM (Japanese Unexamined Patent Application Publication No. 2013-085611).

References

1. World Health Organization. Global Recommendations on Physical Activity for Health. Geneva, Switzerland: World Health Organization; 2010.
2. Lakka HM, Laaksonen DE, Lakka TA, Niskanen LK, Kumpusalo E, Tuomilehto J, et al. The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. JAMA 2002 Dec 4;288(21):2709-2716. [Medline: 12460094]
3. Lakka TA, Laaksonen DE. Physical activity in prevention and treatment of the metabolic syndrome. Appl Physiol Nutr Metab 2007 Feb;32(1):76-88. [doi: 10.1139/h06-113] [Medline: 17332786]
4. Pate RR, Pratt M, Blair SN, Haskell WL, Macera CA, Bouchard C, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995 Feb 1;273(5):402-407. [Medline: 7823386]

5. Montoye HJ, Kemper HC, Saris WH, Washburn RA. Measuring Physical Activity and Energy Expenditure. Champaign, IL: Human Kinetics; 1996.
6. Neilson HK, Robson PJ, Friedenreich CM, Csizmadi I. Estimating activity energy expenditure: how valid are physical activity questionnaires? Am J Clin Nutr 2008 Feb;87(2):279-291 [FREE Full text] [Medline: 18258615]
7. Schoeller DA, Ravussin E, Schutz Y, Acheson KJ, Baertschi P, Jéquier E. Energy expenditure by doubly labeled water: validation in humans and proposed calculation. Am J Physiol 1986 May;250(5 Pt 2):R823-R830. [Medline: 3085521]
8. Bonn SE, Trolle LY, Christensen SE, Möller E, Wright A, Sjölander A, et al. Active-Q: validation of the web-based physical activity questionnaire using doubly labeled water. J Med Internet Res 2012;14(1):e29 [FREE Full text] [doi: 10.2196/jmir.1974] [Medline: 22356755]
9. Bexelius C, Löf M, Sandin S, Trolle LY, Forsum E, Litton JE. Measures of physical activity using cell phones: validation using criterion methods. J Med Internet Res 2010;12(1):e2 [FREE Full text] [doi: 10.2196/jmir.1298] [Medline: 20118036]
10. Namba H, Yamaguchi Y, Yamada Y, Tokushima S, Hatamoto Y, Sagayama H, et al. Validation of Web-based physical activity measurement systems using doubly labeled water. J Med Internet Res 2012;14(5):e123 [FREE Full text] [doi: 10.2196/jmir.2253] [Medline: 23010345]
11. Koebnick C, Wagner K, Thielecke F, Moeseneder J, Hoehne A, Franke A, et al. Validation of a simplified physical activity record by doubly labeled water technique. Int J Obes (Lond) 2005 Mar;29(3):302-309. [doi: 10.1038/sj.ijo.0802882] [Medline: 15672111]
12. Yamada Y, Noriyasu R, Yokoyama K, Osaki T, Adachi T, Itoi A, et al. Association between lifestyle and physical activity level in the elderly: a study using doubly labeled water and simplified physical activity record. Eur J Appl Physiol 2013 Oct;113(10):2461-2471. [doi: 10.1007/s00421-013-2682-z] [Medline: 23801238]
13. Ekman A, Litton JE. New times, new needs; e-epidemiology. Eur J Epidemiol 2007;22(5):285-292. [doi: 10.1007/s10654-007-9119-0] [Medline: 17505896]
14. Owen N, Cerin E, Leslie E, duToit L, Coffee N, Frank LD, et al. Neighborhood walkability and the walking behavior of Australian adults. Am J Prev Med 2007 Nov;33(5):387-395. [doi: 10.1016/j.amepre.2007.07.025] [Medline: 17950404]
15. Inoue S, Ohya Y, Odagiri Y, Takamiya T, Kamada M, Okada S, et al. Perceived neighborhood environment and walking for specific purposes among elderly Japanese. J Epidemiol 2011;21(6):481-490 [FREE Full text] [Medline: 22001543]
16. Rind E, Jones A, Southall H. How is post-industrial decline associated with the geography of physical activity? Evidence from the Health Survey for England. Soc Sci Med 2014 Mar;104:88-97 [FREE Full text] [doi: 10.1016/j.socscimed.2013.12.004] [Medline: 24581066]
17. Patterson PD, Moore CG, Probst JC, Shinogle JA. Obesity and physical inactivity in rural America. J Rural Health 2004;20(2):151-159. [Medline: 15085629]
18. Martin SL, Kirkner GJ, Mayo K, Matthews CE, Durstine JL, Hebert JR. Urban, rural, and regional variations in physical activity. J Rural Health 2005;21(3):239-244. [Medline: 16092298]
19. Fan JX, Wen M, Kowaleski-Jones L. Rural-urban differences in objective and subjective measures of physical activity: findings from the National Health and Nutrition Examination Survey (NHANES) 2003-2006. Prev Chronic Dis 2014;11:E141 [FREE Full text] [doi: 10.5888/pcd11.140189] [Medline: 25144676]
20. Namba H, Yamada Y, Kimura M. Physical activity measurement system. URL: http://lifestyle24.jp/hp/ [accessed 2014-10-07] [WebCite Cache ID 6T94XRYDv]
21. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 2011 Aug;43(8):1575-1581. [doi: 10.1249/MSS.0b013e31821ece12] [Medline: 21681120]
22. Ishikawa-Takata K, Tabata I. Exercise and Physical Activity Reference for Health Promotion 2006 (EPAR2006). J Epidemiol 2007 Sep;17(5):177. [Medline: 17827865]
23. Ishikawa-Takata K, Tabata I, Sasaki S, Rafamantanantsoa HH, Okazaki H, Okubo H, et al. Physical activity level in healthy free-living Japanese estimated by doubly labelled water method and International Physical Activity Questionnaire. Eur J Clin Nutr 2008 Jul;62(7):885-891. [doi: 10.1038/sj.ejcn.1602805] [Medline: 17522602]
24. Murakami H, Kawakami R, Ohmori Y, Miyatake N, Morita A, Miyachi M. Translating from 23METs-h/wk as physical activity reference value for Japanese to daily step counts. Jpn. J. Phys. Fitness Sports Med 2012;61(2):183-191. [doi: 10.7600/jspfsm.61.183]
25. McCormack GR, Shiell A. In search of causality: a systematic review of the relationship between the built environment and physical activity among adults. Int J Behav Nutr Phys Act 2011;8:125 [FREE Full text] [doi: 10.1186/1479-5868-8-125] [Medline: 22077952]
26. Fan JX, Wen M, Kowaleski-Jones L. Rural-urban differences in objective and subjective measures of physical activity: findings from the National Health and Nutrition Examination Survey (NHANES) 2003-2006. Prev Chronic Dis 2014;11:E141 [FREE Full text] [doi: 10.5888/pcd11.140189] [Medline: 25144676]
27. Allender S, Foster C, Boxer A. Occupational and nonoccupational physical activity and the social determinants of physical activity: results from the Health Survey for England. J Phys Act Health 2008 Jan;5(1):104-116. [Medline: 18209257]

28. Rind E, Jones A, Southall H. How is post-industrial decline associated with the geography of physical activity? Evidence from the Health Survey for England. Soc Sci Med 2014 Mar;104:88-97 [FREE Full text] [doi: 10.1016/j.socscimed.2013.12.004] [Medline: 24581066]
29. Tanaka C, Tanaka S. Contribution of non-locomotive activity to habitual physical activity in Japanese workers. Jpn J Phys Fitness Sports Med 2012;61(4):435-441 [FREE Full text] [doi: 10.7600/jspfsm.61.435]
30. Matthews CE, George SM, Moore SC, Bowles HR, Blair A, Park Y, et al. Amount of time spent in sedentary behaviors and cause-specific mortality in US adults. Am J Clin Nutr 2012 Feb;95(2):437-445 [FREE Full text] [doi: 10.3945/ajcn.111.019620] [Medline: 22218159]
31. Ministry of Internal Affairs and Communications. Results of the Telecom Use Trend Survey. 2014 Jun 27. URL: http://www.soumu.go.jp/johotsusintokei/statistics/data/140627_1.pdf [accessed 2015-03-12] [WebCite Cache ID 6WzJG3KsW]

Abbreviations

DLW: doubly labeled water
MET: metabolic equivalent
MVPA: moderate-to-vigorous physical activity
SEE: standard error of the estimate

Edited by G Eysenbach; submitted 07.10.14; peer-reviewed by S Bonn, K Chen, A Burdorf, A Moller; comments to author 03.11.14; revised version received 25.12.14; accepted 04.02.15; published 19.03.15.

Please cite as:
Namba H, Yamada Y, Ishida M, Takase H, Kimura M
Use of a Web-Based Physical Activity Record System to Analyze Behavior in a Large Population: Cross-Sectional Study
J Med Internet Res 2015;17(3):e74
URL: http://www.jmir.org/2015/3/e74/
doi:10.2196/jmir.3923
PMID:25794109

Original Paper

Do Web-Based and Clinic Samples of Gay Men Living With HIV Differ on Self-Reported Physical and Psychological Symptoms? A Comparative Analysis

Richard Harding[1], BSc, MSc, PhD; Fiona Lampe[2], BSc (Hons), PhD; Tim Molloy[3], BSc; Lorraine Sherr[2], BSc(Psych)(Hons), MS(c), PhD

[1]Cicely Saunders Institute, Department of Palliative Care, Policy & Rehabilitation, King\\\'s College London, London, United Kingdom

[2]University College London, Department of Infection and Population Health, London, United Kingdom

[3]GMFA, London, United Kingdom

Corresponding Author:
Richard Harding, BSc, MSc, PhD
Cicely Saunders Institute
Department of Palliative Care, Policy & Rehabilitation
King\\\'s College London
Bessemer Road
London, SE5 9PJ
United Kingdom
Phone: 44 (0)2078485518
Fax: 44 (0)207 8485517
Email: richard.harding@kcl.ac.uk

Abstract

Background: Although the Internet is commonly used to recruit samples in studies of human immunodeficiency virus (HIV)-related risk behaviors, it has not been used to measure patient-reported well-being. As the burden of long-term chronic HIV infection rises, the Internet may offer enormous potential for recruitment to research and interventions.

Objective: This study aimed to compare two samples of gay men living with HIV, one recruited via the Web and the other recruited in outpatient settings, in terms of self-reported physical and psychological symptom burden.

Methods: The Internet sample was recruited from a UK-wide Web-based survey of gay men with diagnosed HIV. Of these, 154 respondents identified themselves as resident in London and were included in this analysis. The HIV clinic sample was recruited from five HIV outpatient clinics. Of these participants, 400 gay men recruited in London clinics were included in this analysis.

Results: The Web-based sample was younger than the clinic sample (37.3 years, SD 7.0 vs 40.9 years, SD 8.3), more likely to be in paid employment (72.8%, 99/136 vs 60.1%, 227/378), less likely to be on antiretroviral therapy (ART) (58.4%, 90/154 vs 68.0%, 266/391), and had worse mean psychological symptom burden compared to the clinic sample (mean scores: 1.61, SD 1.09 vs 1.36, SD 0.96) but similar physical symptom burden (mean scores: 0.78, SD 0.65 vs 0.70, SD 0.74). In multivariable logistic regression, for the physical symptom burden model, adjusted for age, ethnicity, employment status, and ART use, the recruitment setting (ie, Web-based vs clinic) was not significantly associated with high physical symptom score. The only variable that remained significantly associated with high physical symptom score was employment status, with those in employment being less likely to report being in the upper (worst) physical symptom tertile versus the other two tertiles (adjusted OR 0.41, 95% CI 0.28-0.62, *P*<.001). For the psychological symptom burden model, those recruited via the Web were significantly more likely to report being in the upper (worst) tertile (adjusted OR 2.20, 95% CI 1.41-3.44, *P*=.001). In addition, those in employment were less likely to report being in the upper (worst) psychological symptom tertile compared to those not in employment (adjusted OR 0.32, 95% CI 0.21-0.49, *P*<.001).

Conclusions: Our data have revealed a number of differences. Compared to the clinic sample, the Web-based sample had worse psychological symptom burden, younger average age, higher prevalence of employment, and a lower proportion on ART. For future research, we recommend that Web-based data collection should include the demographic variables that we note differed between samples. In addition, we recognize that each recruitment method may bring inherent sampling bias, with clinic populations differing by geographical location and reflecting those accessing regular medical care, and Web-based sampling recruiting those with greater Internet access and identifying survey materials through specific searches and contact with specific websites.

(J Med Internet Res 2015;17(3):e57) doi:10.2196/jmir.3800

KEYWORDS

HIV; pain; symptoms; mental health; methods; recruitment; sampling; Internet

Introduction

Research protocols that utilize electronic and Web-based methods of participant recruitment to research and intervention participation and associated data collection activities have become increasingly common. The method has become particularly well used in behavioral surveillance research studies among persons living with human immunodeficiency virus (HIV) infection in high-income countries. Web-based methods have been used in various ways in epidemiological HIV studies investigating risk behavior [1], to conduct Web-based interventions [2], and to determine the role of the Internet as a phenomenon itself in sexual behavior [3].

A review of the methodological implications of Web-based HIV behavioral surveillance methods summarized the advantages as convenience, reduced costs in the management of tools and data collection and entry, ease of tool modification, anonymity, and reduced social desirability bias [4]. The identified disadvantages were the impossibility of implementing a sampling frame, the potential for limited generalizability due to convenience and self-selected sampling, the challenges of calculating a response rate, and the biases of access to (and familiarity with) Internet use.

A review of HIV behavioral research among men who have sex with men (MSM) found equivocal evidence in the literature as to whether men who use the Internet are more likely to report risk behavior, although those who use the Internet for sex are more likely to be younger, report sex with women, have had a sexually transmitted infection, and to use a public sex environment [5]. The specific websites used for recruitment may also introduce sampling bias, with different demographic profiles between sites [6].

A comparison of London MSM recruited via the Web and in the community found that the Web-based sample were younger and were less likely to exclusively have sex with men, to be in a relationship with a man, to have received higher education, or to have been tested for HIV [7]. A further study comparing a sample of UK-wide MSM recruited online to a second study that conducted a national random probability population sample found no significant differences in reported ethnicity, education, social class, country of birth, alcohol consumption, injecting drug use, or age of first male sexual partner. However, there was strong evidence that men in the Internet sample were younger, less likely to live in London, less likely to report being in good health, and more likely to be working [8]. The evidence to date has been of MSM and the general population, but it has not addressed the specific characteristics of people with diagnosed HIV.

There have been methodological advances in the design and implementation of Web-based recruitment and data collection to investigate behavioral aspects of HIV infection (principally primary prevention and risk behavior). However, the utility of using these methods to investigate disease-oriented variables and patient self-report burden of disease has not been explored. As much health research and delivery becomes oriented to long-term and chronic conditions (where the patient-reported experience is an important area of inquiry, and those living with long-term chronic conditions have greater potential to use the Internet as compared to rapidly declining conditions), there may be great potential for use of the Internet for recruitment and data collection for both research and care activities. The rise in the use of patient-reported outcome measures (PROMS) to improve equity and quality in health care [9] gives great potential for Web-based self-reporting of health states. Self-reported physical and psychological symptom prevalence and intensity is an important area of clinical HIV research, as recent evidence has shown that symptoms are associated with sexual risk-taking (with higher psychological symptom burden associated with risk taking) [10], with poor adherence to antiretroviral treatment (associated with higher psychological and global symptom burden) [11], treatment switching (associated with higher psychological and physical symptom burden) [12], viral rebound (predicted by higher physical, psychological and global symptom burden) [13], poorer quality of life (associated with higher physical and psychological burden) [14], and suicidal ideation (associated with higher physical and psychological burden) [15].

This study aimed to compare two samples of gay men living with HIV, one recruited via the Web and the other recruited in HIV outpatient settings, in terms of self-report physical and psychological symptom burden. The outcome of interest was the self-report 7-day period prevalence and burden of physical and psychological symptoms.

Methods

Overview

The study is a secondary analysis of two datasets (one recruited via the Web and one in outpatient clinics). Participants were gay men with diagnosed HIV who were resident in London, United Kingdom.

Settings and Recruitment Methods

We summarize the two prior study designs here. The Web-based sample [16] was recruited through banners on gay-interest and health information websites for a UK-wide Web-based survey of men who identified as gay with diagnosed HIV (N=347). Of these, 154 respondents identified themselves as resident in London and were included in this analysis.

The HIV clinic sample [15] was recruited in a cross-sectional study conducted in five HIV outpatient clinics in London and the south east of England. Consecutive HIV-diagnosed attending patients were approached and invited to participate by filling out a self-completed questionnaire, and 778 were completed and returned (77%, 778/1010 of all patients; 86.0%, 778/905 of those eligible to receive a questionnaire). Of the participants, 400 men who identified as gay recruited in London clinics were included in this analysis. All data were anonymized, and no identifiable information was available during data merging.

Common Questionnaire Items Between Samples

In both studies, symptoms were measured using the Memorial Symptom Assessment Scale−Short Form (MSAS-SF), a patient self-report scale that measures the 7-day period prevalence of 26 physical and 6 psychological symptoms. This standardized symptom questionnaire captures the presence of each symptom and associated distress (for physical symptoms) or frequency (for psychological symptoms) and has often been reported in studies of people living with HIV [17-20]. Three summary subscale indices can be derived: Physical Symptom Distress (MSAS-Phys), Psychological Symptom Distress (MSAS-Psych), and Global Distress Index (MSAS-GDI) [14]. Each of these 3 subscales has a possible score range of 0-4.

Respondents in both studies gave demographic data on age (analyzed as a continuous variable), ethnicity (categorized as white/non-white), education (categorized as university/non-university), employment (categorized as currently in paid employment or not), and current antiretroviral therapy (ART) use (yes/no).

Analysis

For each sample (clinic and Web-based), we present descriptive analyses for the demographic characteristics and MSAS-SF variables (total number of symptoms, global distress subscale, physical distress subscale, and psychological distress subscale). The demographic variables were compared between samples (clinic vs Web-based) using *t* tests for continuous variables, and chi-square tests for categorical variables. The symptom subscales were compared between samples using *t* tests; additionally, the subscales were categorized into tertiles and compared between samples using chi-square test for trend. Subsequently, in light of the associations between demographic characteristics and sample source (ie, Web-based vs not Web-based), we constructed two multivariable logistic regression models with the MSAS-SF physical and psychological symptom scores as the dependent variables, categorized as the upper (worst) tertile of each score versus the remaining two tertiles. These models aimed to determine whether sample source is independently associated with poorer physical and psychological outcomes when adjusting for demographic factors. Variables significant at the 10% level ($P<.10$) in univariate analysis were entered into the multivariable logistic models [20]. Results from these models are presented as odds ratios (OR) with 95% confidence intervals (CI).

Results

Sample Comparison

The comparison of characteristics between the two samples is presented in Table 1, demonstrating the statistically significant differences identified between samples. The Web-based sample was younger, more likely to be in paid employment, less likely to be on ART, and had worse mean psychological symptom burden compared to the clinic sample. The test for trend revealed that the Web-based sample had a comparatively lower proportion in the worst tertile for physical symptoms, but a higher proportion in the worst psychological tertile, compared to the clinic group.

The multivariable analysis assessed the effect of setting on symptom score (predicting having a score in the highest tertile compared to the other two) after adjusting for all factors with $P<.1$ in univariable analysis (ie, age, ethnicity, employment, ART use). For the physical symptom burden model, the recruitment setting (ie, clinic vs Web-based) was not significant after adjusting for the other factors. The only variable that remained significantly associated with high physical symptom score was employment status, with those in employment being less likely to report being in the upper (worst) physical symptom tertile (adjusted OR 0.41, 95% CI 0.28-0.62, $P<.001$). For the psychological symptom burden model, those recruited via the Web were significantly more likely to report being in the upper (worst) tertiles (adjusted OR 2.20, 95% CI 1.41-3.44, $P=.001$) compared to the clinic sample. In addition, those in employment were less likely to report being in the upper (worst) psychological symptom tertile compared to those not in employment (adjusted OR 0.32, 95% CI 0.21-0.49, $P<.001$). See Table 2.

Table 1. Univariate comparison of the clinic versus Web-based samples.

	Clinic setting (n=400)	Web-based setting (n=154)	Test comparison	Degrees of freedom
Age in years, mean (SD) median				
Missing: clinic n=6; Web-based n=1				
	40.9 (SD 8.3) 40.0	37.3 (SD 7.0) 37.0	*t*=4.77 *P*<.001	530
Education, n (%)				
Missing: clinic n=5; Web-based n=11				
Below university	196 (49.6)	62 (43.1)	χ^2=1.65	
University	199 (50.4)	81 (56.3)	*P*=.20	1
Ethnicity, n (%)				
Missing: clinic n=6; Web-based n=0				
White	346 (87.8)	143 (92.9)	χ^2=2.93	
Non-white	48 (12.2)	11 (7.1)	*P*=.087	1
Employment, n (%)				
Missing: clinic n=22; Web-based n=18				
Not in employment	151 (39.9)	37 (27.2)	χ^2=7.00	
In employment	227 (60.1)	99 (72.8)	*P*=.008	1
Current ART [a] use, n (%)				
Missing: clinic n=9; Web-based n=0				
Not on ART	125 (32.0)	64 (41.6)	χ^2=4.49	
On ART	266 (68.0)	90 (58.4)	*P*=.034	1
MSAS [b] Global, mean (SD) median				
Missing: clinic n=0; Web-based n=12				
	1.15 (SD 0.79) 1.12	1.25 (SD 0.86) 1.23	*t*=−1.29 *P*=.20	529
MSAS Physical, mean (SD) median				
Missing: clinic n=0; Web-based n=17				
	0.78 (SD 0.65) 0.73	0.70 (SD 0.74) 0.47	*t*=1.19 *P*=.23	517
MSAS Psychological, mean (SD) median				
Missing: clinic n=0; Web-based n=4				
	1.36 (0.96) 1.33	1.61 (SD 1.09) 1.65	*t*=−2.69 *P*=.007	522
MSAS Global (tertile groups), n (%)				
<0.72	137 (34.3)	45 (31.7)		
0.72-1.52	136 (34.0)	45 (31.7)		
>1.52	127 (31.8)	52 (36.6)	χ^2(trend)=0.87 *P*=.35	2
MSAS Physical (tertile groups), n (%)				
<0.27	131 (32.8)	55 (40.1)		
0.27-1.0	126 (31.5)	47 (34.3)		
>1.0	143 (35.8)	35 (25.5)	χ^2(trend)=4.66 *P*=.031	2

	Clinic setting (n=400)	Web-based setting (n=154)	Test comparison	Degrees of freedom
MSAS Psychological (tertile groups), n (%)				
<0.87	139 (34.8)	47 (31.3)		
0.87-1.87	149 (37.3)	42 (28.0)		
>1.87	112 (28.0)	61 (40.7)	χ^2(trend)=4.32 *P*=.038	2

[a] ART: antiretroviral therapy

[b] MSAS: Memorial Symptom Assessment Scale

Table 2. Multivariable logistic analysis with (1) physical, and (2) psychological symptoms as independent variable, predicting upper tertile (ie, worst) symptom burden score.

		Physical symptoms as dependent variable (n=482)		Psychological symptoms as dependent variable (n=495)	
		Adjusted odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Setting					
	Clinic[a]	1		1	
	Web-based	0.81 (0.50-1.30)	*P*=.38	2.20 (1.41-3.44)	*P*=.001
Age (years)		1.00 (0.97-1.02)	*P*=.91	0.98 (0.95-1.00)	*P*=.98
Ethnicity					
	Non-white[a]	1		1	
	White	1.40 (0.73-2.69)	*P*=.30	1.22 (0.64-2.24)	*P*=.54
Employment					
	Not employed[a]	1		1	
	Employed	0.41 (0.28-0.62)	*P*<.001	0.32 (0.21-0.49)	*P*<.001
ART [b] use					
	No[a]	1		1	
	Yes	1.39 (0.90-2.15)	*P*=.13	1.20 (0.78-1.85)	*P*=.40

[a] reference category

[b] ART: antiretroviral therapy

Discussion

Principal Findings

Our data have revealed a number of differences between samples recruited via the Web and in clinic settings, in terms of both demographics and self-report psychological and physical symptom burden. First, as has been found with behavioral studies of gay men, the Web-based sample was younger than the clinic sample. Second, we found that those recruited via the Web were more likely to be in employment, which may reflect the costs associated with Internet connectivity. It may also be that those in employment face greater challenges in attending clinics regularly, or that the clinic sample had a higher proportion of participants with health problems linked to not being employed. Third, the clinic sample was more likely to be on ART. This may be because clinic sampling may under-sample those with poor or erratic attendance who are not on treatment or those with mild immunosuppression not yet on treatment and who feel less need to attend for care. In terms of the self-report symptom burden, participants in the Web-based sample were less likely to have high physical symptom burden; this difference was largely explained by the differences between the samples in demographic characteristics. However, the Web-based sample was more likely to have higher psychological symptom burden; this difference was not attenuated in adjusted analysis. There are several potential explanations for this latter finding. Psychological problems are more prevalent among HIV-infected populations (even among those on ART [21]) than the general population and those with other long-term conditions. It may be that the Web-based sample was actively seeking resources via the Web regarding the problems associated with living with HIV, and therefore accessed the survey materials. If the Web-based sample is, as we hypothesized above, less engaged with clinical services, this may be because their poorer mental health prevents them from attending regular care. Alternatively, it is possible that participants were more likely to report their psychological symptoms in the Web-based study due to a greater perceived anonymity of this study, or that those recruited via clinics are receiving better psychological care. In terms of intervention, while Web-based recruitment to intervention may enable those with worse psychological distress to be introduced to studies, it does not imply acceptability of Web-based interventions nor that those recruited in clinics may be more or less likely to take up Web-based interventions.

Our findings that the Web-based sample was statistically significantly younger and more likely to be employed are in line with previous studies [5,7]. However, our finding that they are also more likely to be in treatment and have lower psychological symptom burden is entirely new.

Limitations

There are a number of limitations to our study. First, although we were able to compare a well-defined population by analyzing data from only those men in the Web-based survey with a London postal code to those accessing care at a London clinic, we note that men may travel to access care. Second, although we had a high response rate for the clinic survey (86%), we do not have a response rate for the Web-based survey.

For future research, we recommend that Web-based recruitment should include collection of the demographic variables that we note differed between samples: age, ART use, and employment. The effect of adjustment for these factors in analyses can then be examined. In addition, we recognize that each recruitment method may bring inherent sampling bias, with clinic populations differing by geographical populations served and reflecting those accessing regular medical care, and Web-based sampling recruiting those with greater Internet access and identifying survey materials through specific searches and contact with specific websites. Furthermore, our Web-based sampling did not allow specification of a sampling frame.

Conclusions

Patient-reported symptom data can feasibly be collected though Web-based recruitment as well as through clinic-based questionnaire studies. There may be some specific advantages of Web-based studies when investigating stigmatized problems such as psychological burden of HIV disease, where social desirability may bias traditional face-to-face recruitment methods. We conclude that the Web-based sample had a higher psychological symptom burden (but not physical symptom burden) and that therefore they differ clinically from the sample recruited in clinics. This offers great opportunity to recruit people with HIV to interventions to improve mental health, which is a highly prevent and burdensome problem in this population.

Acknowledgments

We would like to thank the volunteers of GMFA who conducted the Web-based survey and GlaxoSmithKline for an unrestricted educational grant for the clinic survey.

Conflicts of Interest

None declared.

References

1. Prestage G, Jin F, Bavinton B, Scott SA, Hurley M. Do differences in age between sexual partners affect sexual risk behaviour among Australian gay and bisexual men? Sex Transm Infect 2013 Dec;89(8):653-658. [doi: 10.1136/sextrans-2012-050947] [Medline: 23843453]
2. Miranda J, Côté J, Godin G, Blais M, Otis J, Guéhéneuc YG, et al. An Internet-based intervention (Condom-Him) to increase condom use among HIV-positive men who have sex with men: protocol for a randomized controlled trial. JMIR Res Protoc 2013;2(2):e39 [FREE Full text] [doi: 10.2196/resprot.2723] [Medline: 24132072]
3. Wong ML, Koh TT, Tjahjadi S, Govender M. Men seeking sex online practise riskier sexual behaviours than men frequenting brothels: survey findings from Singapore. Sex Transm Infect 2014 Aug;90(5):401-407. [doi: 10.1136/sextrans-2013-051366] [Medline: 24633651]
4. Zhang D, Bi P, Hiller JE, Lv F. Web-based HIV/AIDS behavioral surveillance among men who have sex with men: potential and challenges. Int J Infect Dis 2008 Mar;12(2):126-131 [FREE Full text] [doi: 10.1016/j.ijid.2007.06.007] [Medline: 17884663]
5. Chiasson MA, Parsons JT, Tesoriero JM, Carballo-Dieguez A, Hirshfield S, Remien RH. HIV behavioral research online. J Urban Health 2006 Jan;83(1):73-85 [FREE Full text] [doi: 10.1007/s11524-005-9008-3] [Medline: 16736356]
6. Bull SS, McFarlane M, Lloyd L, Rietmeijer C. The process of seeking sex partners online and implications for STD/HIV prevention. AIDS Care 2004 Nov;16(8):1012-1020. [doi: 10.1080/09540120412331292426] [Medline: 15511733]
7. Elford J, Bolding G, Davis M, Sherr L, Hart G. Web-based behavioral surveillance among men who have sex with men: a comparison of online and offline samples in London, UK. J Acquir Immune Defic Syndr 2004 Apr 1;35(4):421-426. [Medline: 15097159]
8. Evans AR, Wiggins RD, Mercer CH, Bolding GJ, Elford J. Men who have sex with men in Great Britain: comparison of a self-selected internet sample with a national probability sample. Sex Transm Infect 2007 Jun;83(3):200-5; discussion 205 [FREE Full text] [doi: 10.1136/sti.2006.023283] [Medline: 17135330]
9. Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ. The routine use of patient reported outcome measures in healthcare settings. BMJ 2010;340:c186 [FREE Full text] [Medline: 20083546]

10. Harding R, Clucas C, Lampe FC, Norwood S, Leake DH, Fisher M, et al. Behavioral surveillance study: sexual risk taking behaviour in UK HIV outpatient attendees. AIDS Behav 2012 Aug;16(6):1708-1715. [doi: 10.1007/s10461-011-0023-y] [Medline: 21850443]
11. Sherr L, Lampe F, Norwood S, Leake DH, Harding R, Johnson M, et al. Adherence to antiretroviral treatment in patients with HIV in the UK: a study of complexity. AIDS Care 2008 Apr;20(4):442-448. [doi: 10.1080/09540120701867032] [Medline: 18449821]
12. Clucas C, Harding R, Lampe FC, Anderson J, Date HL, Johnson M, et al. Doctor-patient concordance during HIV treatment switching decision-making. HIV Med 2011 Feb;12(2):87-96. [doi: 10.1111/j.1468-1293.2010.00851.x] [Medline: 20561081]
13. Lampe FC, Harding R, Smith CJ, Phillips AN, Johnson M, Sherr L. Physical and psychological symptoms and risk of virologic rebound among patients with virologic suppression on antiretroviral therapy. J Acquir Immune Defic Syndr 2010 Aug;54(5):500-505. [doi: 10.1097/QAI.0b013e3181ce6afe] [Medline: 20150819]
14. Harding R, Clucas C, Lampe FC, Date HL, Fisher M, Johnson M, et al. What factors are associated with patient self-reported health status among HIV outpatients? A multi-centre UK study of biomedical and psychosocial factors. AIDS Care 2012;24(8):963-971. [doi: 10.1080/09540121.2012.668175] [Medline: 22519889]
15. Sherr L, Lampe F, Fisher M, Arthur G, Anderson J, Zetler S, et al. Suicidal ideation in UK HIV clinic attenders. AIDS 2008 Aug 20;22(13):1651-1658. [doi: 10.1097/QAD.0b013e32830c4804] [Medline: 18670226]
16. Harding R, Molloy T, Easterbrook P, Frame K, Higginson IJ. Is antiretroviral therapy associated with symptom prevalence and burden? Int J STD AIDS 2006 Jun;17(6):400-405. [doi: 10.1258/095646206777323409] [Medline: 16734963]
17. Wakeham K, Harding R, Bamukama-Namakoola D, Levin J, Kissa J, Parkes-Ratanshi R, et al. Symptom burden in HIV-infected adults at time of HIV diagnosis in rural Uganda. J Palliat Med 2010 Apr;13(4):375-380. [doi: 10.1089/jpm.2009.0259] [Medline: 20384499]
18. Harding R, Lampe FC, Norwood S, Date HL, Clucas C, Fisher M, et al. Symptoms are highly prevalent among HIV outpatients and associated with poor adherence and unprotected sexual intercourse. Sex Transm Infect 2010 Dec;86(7):520-524. [doi: 10.1136/sti.2009.038505] [Medline: 20551235]
19. Harding R, Selman L, Agupio G, Dinat N, Downing J, Gwyther L, et al. Prevalence, burden, and correlates of physical and psychological symptoms among HIV palliative care patients in sub-Saharan Africa: an international multicenter study. J Pain Symptom Manage 2012 Jul;44(1):1-9. [doi: 10.1016/j.jpainsymman.2011.08.008] [Medline: 22658471]
20. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
21. Lowther K, Selman L, Harding R, Higginson IJ. Experience of persistent psychological symptoms and perceived stigma among people with HIV on antiretroviral therapy (ART): a systematic review. Int J Nurs Stud 2014 Aug;51(8):1171-1189. [doi: 10.1016/j.ijnurstu.2014.01.015] [Medline: 24602830]

Abbreviations

ART: antiretroviral therapy
HIV: human immunodeficiency virus
MSM: men who have sex with men
MSAS-SF: Memorial Symptom Assessment Scale−Short Form
MSAS-PHYS: Memorial Symptom Assessment Scale−Physical Distress Subscale
MSAS-PSY: Memorial Symptom Assessment Scale−Psychological Distress Subscale
MSAS-GDI: Memorial Symptom Assessment Scale−Global Distress Subscale

Edited by G Eysenbach; submitted 21.08.14; peer-reviewed by J Downing, C Aicken; comments to author 09.10.14; revised version received 04.12.14; accepted 21.01.15; published 19.03.15.

Please cite as:
Harding R, Lampe F, Molloy T, Sherr L
Do Web-Based and Clinic Samples of Gay Men Living With HIV Differ on Self-Reported Physical and Psychological Symptoms? A Comparative Analysis
J Med Internet Res 2015;17(3):e57
URL: http://www.jmir.org/2015/3/e57/
doi:10.2196/jmir.3800
PMID:

Original Paper

Emergency Physicians' Views of Direct Notification of Laboratory and Radiology Results to Patients Using the Internet: A Multisite Survey

Joanne Callen[1*], BA, MPH, PhD; Traber Davis Giardina[2*], MA, MSW, PhD; Hardeep Singh[2*], MD, MPH; Ling Li[1*], MBiostats, PhD; Richard Paoloni[3*], MBBS, MMed; Andrew Georgiou[1*], MSc, PhD; William B Runciman[4*], MBBCh, PhD; Johanna I Westbrook[1*], MHA, PhD

[1]Centre for Health Systems and Safety Research, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

[2]Houston VA HSR&D Center for Innovations in Quality, Effectiveness and Safety, Michael E DeBakey Veterans Affairs Medical Center, Department of Medicine, Baylor College of Medicine, Houston, TX, United States

[3]Emergency Department, Concord Repatriation and General Hospital, Sydney, Australia

[4]The Joanna Briggs Institute, Department of Anaesthesia and Intensive Care, University of Adelaide, Adelaide, Australia

[*]all authors contributed equally

Corresponding Author:
Joanne Callen, BA, MPH, PhD
Centre for Health Systems and Safety Research
Australian Institute of Health Innovation
Macquarie University
Level 6, 75 Talavera Road
Macquarie University
Sydney, 2109
Australia
Phone: 61 9850 2425
Fax: 61 8088 6234
Email: joanne.callen@mq.edu.au

Abstract

Background: Patients are increasingly using the Internet to communicate with health care providers and access general and personal health information. Missed test results have been identified as a critical safety issue with studies showing up to 75% of tests for emergency department (ED) patients not being followed-up. One strategy that could reduce the likelihood of important results being missed is for ED patients to have direct access to their test results. This could be achieved electronically using a patient portal tied to the hospital's electronic medical record or accessed from the relevant laboratory information system. Patients have expressed interest in accessing test results directly, but there have been no reported studies on emergency physicians' opinions.

Objective: The aim was to explore emergency physicians' current practices of test result notification and attitudes to direct patient notification of clinically significant abnormal and normal test results.

Methods: A cross-sectional survey was self-administered by senior emergency physicians (site A: n=50; site B: n=39) at 2 large public metropolitan teaching hospitals in Australia. Outcome measures included current practices for notification of results (timing, methods, and responsibilities) and concerns with direct notification.

Results: The response rate was 69% (61/89). More than half of the emergency physicians (54%, 33/61) were uncomfortable with patients receiving direct notification of abnormal test results. A similar proportion (57%, 35/61) was comfortable with direct notification of normal test results. Physicians were more likely to agree with direct notification of normal test results if they believed it would reduce their workload (OR 5.72, 95% CI 1.14-39.76). Main concerns were that patients could be anxious (85%, 52/61), confused (92%, 56/61), and lacking in the necessary expertise to interpret their results (90%, 55/61).

Conclusions: Although patients' direct access to test results could serve as a safety net reducing the likelihood of abnormal results being missed, emergency physicians' concerns need further exploration: which results are suitable and the timing and method of direct release to patients. Methods of access, including secure Web-based patient portals with drill-down facilities providing test descriptions and result interpretations, or laboratories sending results directly to patients, need evaluation to ensure patient safety is not compromised and the processes fit with ED clinician and laboratory work practices and patient needs.

(J Med Internet Res 2015;17(3):e60) doi:10.2196/jmir.3721

KEYWORDS

Internet; patient safety; electronic health records; patient empowerment; diagnostic tests; emergency care; radiology

Introduction

Information and communication technology (ICT) is an essential component for facilitating communication between health professionals within and across health care settings [1,2]. Patients are also increasingly using the Internet to communicate with health care providers and access general and personal health information using email and patient portals tied to hospital-based electronic medical records [3-10]. Given the increasing use of ICT to assist communication, any investigation of suboptimal communication between patients and providers should explore ways in which the Internet could improve information exchange. Breakdown in communication has been identified as a common problem with test result follow-up [11-15]. In the emergency department (ED), up to 75% of test results are missed and the potential impact on patient outcomes includes missed cancer diagnoses [16].

More systematic processes for test result follow-up are needed, including the use of electronic information systems [13,17-20]. It has also been proposed that patient involvement could assist in mitigating missed test results [21-23]. Patients could act as a safety net and be directly notified of their results by the treating clinician transmitting them by email, text messaging (short message service, SMS), letter, or telephone. Alternatively, patients could access test results electronically directly from the laboratory or via a patient portal tied to their hospital electronic medical record (EMR). Physicians would still play a major role in test follow-up and action regarding further treatments, etc. There are moves toward legislating direct notification of tests results in the United States. Since 2004, the Mammography Quality Standards Act has required that a summary of the mammogram report, written in layperson terms, be sent directly to the patient within 30 days [24]. In February 2014, the US Department of Health and Human Services issued a regulation specifying that laboratories must release test reports directly to patients if requested [25].

Studies have shown that patients support direct notification of normal and abnormal test results [5,21,26-29]. However, studies that have explored physicians' views have only included primary care practitioners [30-36]. Because the ED is a critical area where results are missed, it is important to gauge physicians' views on direct notification of test results to patients. There are no published studies of emergency physicians' opinions of patients' direct access to test results.

The aim of this study was to explore emergency physicians' current practices of abnormal test result notification and their attitudes to direct patient notification of clinically significant abnormal and normal test results.

Our research questions were 2-fold:

1. What are emergency physicians' current practices for patient notification of abnormal test results in terms of methods, timing of, and responsibilities for notification?
2. What are emergency physicians' views about direct patient notification of normal and abnormal test results?

Methods

Study Design, Setting, and Population

A cross-sectional paper-based survey was administered to all senior emergency physicians (site A: n=50; site B: n=39) in the EDs of 2 large public metropolitan teaching hospitals in Australia (Table 1). The sites were selected using convenience sampling. Senior emergency physicians were defined as department directors, specialists in emergency medicine, registrars, and senior medical officers and were included based on their involvement and responsibility in relation to test ordering and test result follow-up. The study was approved by the ethics committees from both hospital study sites and the University of New South Wales, Sydney, Australia.

Table 1. Characteristics of 2 hospital emergency department (ED) sites.

Characteristics	Hospital site	
	A	B
Hospital beds, n	758	543
Annual inpatient discharges, n	83,898	45,055
Annual ED attendances, n	61,939	35,687
Annual ED discharges home, n	40,713	23,019
Hospitalizations for which the ED is responsible, %	25%	28%
ED admission rates, %	34%	36%

Survey Instrument

A survey consisting of 22 closed-ended questions was developed based on literature related to patients' and physicians' access to clinical information [34,37]. We pilot-tested the survey face-to-face with 2 senior hospital specialists (1 ED physician and 1 rheumatology specialist) from Australia and 10 US primary care practitioners. Based on their feedback, modifications were made to improve clarity of the survey. Fifteen of the survey questions had 7-point Likert response scales with options ranging from strongly agree to strongly disagree.

A detailed description of the design and development of the survey instrument is included in the paper on primary care physicians' attitudes to directly releasing results to patients [38]. An identical survey was used in both studies. Questions included: current practices and institutional policies for patient notification of abnormal test results (timing, methods, and responsibilities), attitudes toward direct notification of normal and abnormal results, possible physician concerns including lack of patient expertise in interpreting results, previous experiences with missed abnormal test results by themselves or others, patient care decision-making attitudes, and participant demographic details.

For the purposes of the study, direct notification was defined as the automatic release of test results directly to the patient regardless of whether or not the ordering physician had reviewed the results. Direct notification could be achieved by mail, telephone, fax, SMS text message, or by patients accessing their EMR using a patient portal. *Abnormal test results* were defined for this study as clinically significant abnormal results. This includes those that are not immediately life-threatening but required short-term follow-up; for example, newly elevated glucose blood levels or a chest x-ray with a new shadow.

Data Collection

The survey was self-administered by emergency physicians at the 2 study hospitals between July 1 and September 30, 2012. The survey took each physician approximately 15 minutes to complete. After obtaining staff rosters, researchers approached each ED physician from the study population and asked them to participate in the study. Each physician was informed of the confidential nature of the results and each survey contained a participant information sheet outlining the nature of the study and that completion of the survey implied consent. The survey was left with the ED physician and they were asked to place the completed survey in a dedicated secure box in the ED secretary's office. The researchers returned at the end of each week to collect the completed surveys and follow-up with nonrespondents. Those physicians who had failed to complete the survey were reminded and given another survey if the previous one had been mislaid.

Statistical Analysis

Data were analyzed using SPSS version 21.0 (IBM Corp, Armonk, NY, USA) and descriptive statistics calculated. For ease of interpretation, we recoded the dependent variable responses into dichotomous categories. For descriptive analyses and logistic regression, the responses of agree, moderately agree, or strongly agree and disagree, moderately disagree, and strongly disagree were collapsed into agree and disagree, respectively. The category of neither agree nor disagree was retained for descriptive analyses.

We used logistic regression to identify predictors related to emergency physicians' comfort with direct notification of abnormal test results to patients and their agreement with direct notification of normal test results. Exact methods for logistic regression models were adopted because the sample size of the study was too small for asymptotic methods [39]. The category of neither agree nor disagree was excluded for logistic regression because the number of physicians who responded to this category was small and exact analysis for an ordinal logistic regression model was not available computationally. Two exact logistic regression models were developed using a stepwise forward selection of covariate method [39]. To avoid overfitting the models, we included no more than 2 explanatory variables in each model [40]. We included only variables that were clinically relevant as possible predictors. Predictors in the final models were considered significant at the $P<.05$ level. Odds ratios and their 95% confidence intervals were presented for the predictors included in the final models. The regression analysis was generated using SAS/STAT version 9.3 (SAS Institute, Inc, Cary, NC, USA).

Results

Characteristics of Respondents

A total of 89 emergency physicians were invited to complete the survey and 61 complied giving an overall response rate of 69% (61/89) (Table 2). The response rate for hospital A was 74% (37/50) and for hospital B was 62% (24/39).

Table 2. Characteristics of emergency physicians from the 2 study hospitals (N=61).

Participant characteristics	Total, n (%)
Gender	
Female	30 (49)
Male	31 (51)
Age (years)	
20-29	20 (33)
30-39	17 (28)
40-49	16 (26)
50-59	5 (8)
60-69	3 (5)
Position [a]	
Senior emergency physician	21 (34)
Registrar	25 (41)
SRMO	15 (25)
Years in practice	
<5	25 (41)
5-10	12 (20)
11-15	10 (16)
16-20	4 (7)
>21	10 (16)

[a] Senior emergency physicians are board certified specialists in emergency medicine; registrars and senior resident medical officers (SRMOs) are physicians in senior fellowship or residency positions.

Emergency Physicians' Attitudes Toward and Concerns About Direct Notification of Test Results

Approximately half of the emergency physicians (54%, 33/61) were not comfortable with patients receiving direct notification of abnormal results, although a similar proportion (57%, 35/61) agreed with direct notification of normal results (Table 3). Most (57%, 35/61) agreed that overall a direct notification system would reduce the number of patients lost to follow-up. Physicians' major concerns relating to direct notification of abnormal results were confusion (92%, 56/61), lack of expertise necessary to interpret test results (90%, 55/61), and patient anxiety (85%, 52/61). Most also expressed concerns that the patient may seek information that could be unreliable (57%, 35/61). Most physicians were not concerned that direct notification would interfere with the practice of medicine (90%, 55/61) or increase their workload (92%, 56/61). The majority of physicians disagreed with releasing abnormal results directly to patients for human immunodeficiency virus (HIV) (84%, 51/61), cancer screening (84%, 51/61), and chest x-rays (59%, 36/61). Tests which most thought suitable for direct notification were lipid profile (64%), blood glucose (59%), and urinalysis (57%). Approximately half the respondents felt comfortable with releasing abnormal complete blood counts (49%, 30/61), thyroid function tests (48%, 29/61), and electrolytes (44%, 27/61) directly to patients (Table 3).

Emergency Physicians' Current Practices and Responsibilities for Direct Notification of Abnormal Results

Most respondents agreed (75%, 46/61) that there were standard policies and procedures for notification of abnormal results in their EDs (Table 4). However, there were mixed responses regarding who should be responsible for notifying patients of results with some physicians (43%, 26/60) stating that it was not always clear who should notify patients of abnormal results. The majority agreed that the ordering physician (or their assigned delegate) should be solely responsible for notifying patients (65%, 39/60) and the majority also agreed that the primary care provider should be responsible for following up abnormal results regardless of who ordered the test (65%, 39/60). Emergency physicians' practices in relation to methods of patient notification of abnormal test results varied with respondents stating that they always (47%, 25/53) or sometimes (43%, 23/53) phoned the patient. The majority of emergency physicians indicated that they never emailed (100%, 44/44), never faxed (91%, 40/44), and never waited for the patient to contact them (91%, 40/44) regarding abnormal results. In relation to decision making, most respondents (58%, 35/60) thought that they shared responsibility with the patient for deciding on treatments.

Table 3. Emergency physicians' attitudes toward direct notification.

Opinions and concerns about direct notification of test results to patients	Scale, n (%)			n (%)
	Agree/yes	Neither agree or disagree	Disagree/no	
Attitudes to direct notification of test results to patients (n=61)				
I am comfortable with patients receiving direct notification of abnormal test results	24 (39)	4 (7)	33 (54)	
Do you agree that there should be direct patient notification of normal results?	35 (57)	8 (13)	18 (30)	
Overall, a direct notification system would reduce the number of patients lost to follow-up	35 (57)	13 (21)	13 (21)	
Overall, a direct notification system would reduce physician workload	19 (32)	15 (25)	26 (43)	
Concerns regarding direct notification of abnormal test results [a] to patients (n=61)				
Patient anxiety about test results	52 (85)		9 (15)	
Patient confusion about test results	56 (92)		5 (8)	
Patients lack expertise necessary to interpret the results	55 (90)		6 (10)	
Patient may seek unreliable information	35 (57)		26 (43)	
Patient may seek care without consulting their primary care provider	29 (48)		32 (53)	
Interferes with the practice of medicine	6 (10)		55 (90)	
Physician workload increase	5 (8)		56 (92)	
I have no concerns	2 (3)		59 (97)	
If direct notification became the norm, which abnormal test results [a] would you be comfortable with releasing directly to patients (n=61)				
Complete blood count	30 (49)		31 (51)	
Electrolyte panel	27 (44)		34 (56)	
Blood glucose	36 (59)		25 (41)	
Chest x-ray	25 (41)		36 (59)	
Lipid profile (TC, HDL, LDL, TG)	39 (64)		22 (36)	
Thyroid blood tests (TSH, T4, TPO)	29 (48)		32 (53)	
HIV	10 (16)		51 (84)	
Urinalysis	35 (57)		26 (43)	
Cancer screening tests (eg, mammography, PAP smear)	10 (16)		51 (84)	
Please specify at what time interval, after the result became available, would you be comfortable with direct notification of abnormal test results [a] to patients (n=60) [b]				
24 hours				31 (52)
48 hours				16 (27)
7 days				5 (8)
14 days				0 (0)
30 days				1 (2)
Other				7 (12)

[a] Abnormal test results are clinically significant abnormal results such as newly elevated blood glucose or chest x-ray with new shadow.

[b] One missing response.

Table 4. Emergency physicians' current practices and responsibilities for direct notification of abnormal results.[a]

Emergency physicians' current practices and responsibilities for direct notification of abnormal results	n (%)	Scale, n (%)		
		Always	Sometimes	Never
As part of your usual practice when do you (or staff delegated by you) typically notify patients of abnormal test results? (n=61)				
<24 hours	50 (82)			
24 hours-1 week	11 (18)			
>1 week	0 (0)			
In my practice, there are standardized policies and procedures for notification of abnormal test results (n=61)				
Agree	46 (75)			
Neither agree nor disagree	7 (12)			
Disagree	8 (13)			
The physician who ordered the test or their assigned delegate should be solely responsible for notifying patients of abnormal test results (n=60)				
Agree	39 (65)			
Neither agree nor disagree	0 (0)			
Disagree	21 (35)			
The assigned primary care provider for the care of the patient should always be responsible for following up abnormal test results regardless of who ordered the test (n=60)				
Agree	39 (65)			
Neither agree nor disagree	2 (3)			
Disagree	19 (32)			
It is not always clear who should notify patients of abnormal test results (n=60)				
Agree	26 (43)			
Neither agree nor disagree	9 (15)			
Disagree	25 (42)			
When it comes to make decision about my patients' care, I am most comfortable when (n=60)				
I make the decision about treatment	0 (0)			
I make the final decision but consider the patient's opinion	22 (37)			
The patient and I share responsibility for deciding treatment	35 (5)			
The patient makes the final decision but considers my opinion	3 (5)			
The patient makes the final selection with little input from me	0 (0)			
Once you have seen an abnormal test result how do you (or staff delegated by you) notify patients?[b]				
Phone		25 (47)	23 (4)	5 (9)
Email		0 (0)	0 (0)	44 (100)
Fax		1 (2)	3 (7)	40 (91)
Letter		3 (7)	23 (50)	20 (44)
Wait until next appointment		0 (0)	5 (12)	38 (88)
Schedule a follow-up appointment		3 (7)	14 (32)	27 (61)
Wait for the patient to contact you		1 (2)	3 (7)	40 (91)

[a] Abnormal test results are clinically significant abnormal results such as newly elevated blood glucose or chest x-ray with new shadow.

[b] Not all respondents answered each question.

Emergency Physicians' Experiences With Missed Test Results

In the previous 12 months, 21% (13/61) of emergency physicians said they had missed an abnormal result which led to delayed patient care, although half said they did not know whether they had missed a result or not (51%, 31/61). When asked if their colleagues had missed an abnormal result, a higher proportion responded yes (44%, 27/61). Participants reported radiology results to be the most frequently missed (38%, 23/61) followed by microbiology (13%, 8/61) and chemistry (13%, 8/61).

Results From Exact Logistic Regression Models

The final exact logistic regression model showed that emergency physicians who were not concerned about whether patients might seek unreliable information were 4 times more likely to be comfortable with direct notification of abnormal test results than those who were not concerned about this issue (OR 4.56, 95% CI 1.04-24.30). If direct notification became the norm, physicians would feel more comfortable with releasing abnormal test results on blood glucose than on other test results (OR 15.74, 95% CI 2.84-171.16).

Emergency physicians' agreement on whether a direct notification system would reduce physician workload was the main factor associated with agreement on direct patient notification of normal test results. Those who agreed that a direct notification system would reduce physician workload were more likely to agree with a direct patient notification of normal results than those who disagreed that it would reduce workloads (OR 5.72, 95% CI 1.14-39.76). Those who were neutral about this issue were more likely to agree on direction notification than those who disagreed, but this difference was not statistically significant (OR 3.84, 95% CI 0.72-27.81).

Table 5. Logistic regression exploring predictors of emergency physician comfort with direct patient notification of abnormal test results and agreement with patient notification of normal test results.

Parameter	OR (95% CI)	*P*
Feel comfortable with patients receiving direct notification of abnormal test results[a]		
Concerned that patients may seek unreliable information		.04
No (reference)	4.56 (1.04-24.30)	
Yes		
Feel comfortable with releasing abnormal test results on blood glucose, if direct notification became the norm		<.001
Yes	15.74 (2.84-171.16)	
No (reference)		
Agree that there should be direct patient notification of normal results		
Overall, a direct notification system would reduce physician workload		.03
Agree	5.72 (1.14-39.76)	.003
Neutral	3.84 (0.72-27.81)	.14
Disagree (reference)		

[a] Abnormal test results are clinically significant abnormal results such as newly elevated blood glucose or chest x-ray with new shadow.

Discussion

Principal Results

Although most emergency physicians in our study thought that a direct test notification system would reduce the number of patients lost to follow-up, just over half did not support direct notification of abnormal test results. The main concerns expressed were that it would result in patient anxiety and confusion, and that patients lacked the necessary expertise to interpret test results. Physicians were more likely to be comfortable with direct notification of abnormal test results if they were not concerned about whether patients might seek potentially unreliable information. However, the majority of emergency physicians in our study were comfortable with direct notification of normal test results and they were more likely to be supportive if they thought this would reduce their workload. Approximately 1 in 5 respondents reported missing results in the previous 12 months, whereas half did not know whether they had or not. Radiology results were cited as the most frequently missed. There were diverse responses regarding who should be responsible for notifying patients of abnormal results with a number of respondents stating that this was not always clear.

Two key results from our study, which have not been previously published, focus attention on the unique ED environment, where physicians normally have no continuing relationship with their patients. Most respondents in our study notified patients of abnormal test results via telephone, sometimes they used a letter, and email was never used. ED physicians also showed mixed views regarding who was responsible for notifying the patient of an abnormal result. Lines of responsibility for test notification seem unclear because the majority of respondents in our study thought the physician who ordered the test should be responsible, but the majority also agreed that the primary care provider was responsible as well. These new findings underscore the need to design test notification systems to suit the ED context.

Limitations

The data were collected from senior emergency physicians from 2 EDs. Emergency physicians from regional and smaller hospitals and junior emergency physicians may have different attitudes. We did not include a question in the survey about the possible responsibility of laboratories (pathology or imaging) with regard to notifying patients of test results, which meant we did not garner emergency physicians' opinions on this. Qualitative interviews with emergency physicians, laboratory personnel, and information technology staff could have elicited richer information on barriers and facilitators to patients having direct access to their test results.

Comparison With Prior Work

Overview

Studies have shown that patients overwhelmingly support direct notification of their test results [5,9,21,26-29] and when given access to their hospital medical records via a patient portal, test results are often cited as the most frequently accessed information [41-43]. One study surveyed 304 ED patients regarding their use of the Internet and found that the majority were willing to view laboratory results via the Internet [44]. However, when physicians are asked their views of direct notification, they often perceive barriers [5,32,34,36,45]. In contrast to the generally negative opinions from surveys and qualitative studies asking physicians for their views, pilot studies trialing patients' electronic access to their EMR have reported positive attitudes from physicians [30,31,33,35,46,47]. An example is the positive findings reported from the Open Notes project [31,37,48].

Physicians' Concerns About Patients' Anxiety if Results Are Directly Notified

Physicians' perceived concerns about direct access to test results raising anxiety levels of patients is an often reported issue [32,34,45,49] and was also a finding from our study. However, studies which measured whether patients' access to medical records and test results caused undue anxiety found no increase in their anxiety levels [23,35,50-52] and have reported positive outcomes for patients. A systematic review to determine the effect of providing patients with access to their medical records, although not specifically reporting on test results, concluded overall that access reduced or had no effect on patient anxiety [53].

Physicians' Concerns About Patients' Lack of Clinical Knowledge if Results Are Directly Notified

A concern expressed by emergency physicians in our study was that patients lacked the necessary expertise or knowledge to interpret test results; emergency physicians were more likely to be comfortable with release of abnormal results if they were not concerned about patients seeking potentially unreliable information. A pilot study exploring patient access to electronic laboratory results via a patient portal attempted to address this issue by providing easily accessible test result reference information for patients [33]. Each test result had an information button with a hyperlink to general reference information about the result, and they reported that patients in 1 primary practice who viewed results also viewed reference information in 50% of cases [33]. Other studies have reported patients' lack of understanding and confusion when receiving written histopathology reports following endoscopy [54] and mammography [55], which underscores the need to provide further background/reference information to patients to enhance comprehension of the test result report.

Physicians' Concerns About the Timing of Reporting Results Directly to Patients

Timing of release of information was identified as important in our study with the majority of emergency physicians indicating that they would prefer a 48-hour delay for release of abnormal results to allow them time to contact the patient first. Patients have indicated that timeliness of notification of test results is important [28,29,54,55] and this was taken into account in the Wald et al pilot study [33], as an expert panel of clinicians identified the set of results for direct release and their timing rules (immediate release or a 2-day embargo).

Physicians' Concerns About Direct Notification of Abnormal Results

Our study found whether a result is normal or abnormal is an issue for physicians in relation to reporting test results directly to patients. Most emergency physicians were not comfortable with releasing results that might have a significant emotional impact on patients, such as abnormal HIV, mammography, and PAP smears. This finding is supported by results from a survey of primary care physicians by Sung et al [32], which indicated physicians were significantly more interested in reporting normal rather than abnormal results ($P<.001$). That study found that the level of interest in direct reporting of results declined progressively depending on the perceived emotional impact of the result on the patient, from low (dual energy x-ray absorptiometry scan) to intermediate (genital herpes testing) to high (breast biopsy) ($P<.001$) [32].

Physicians' Concerns Regarding Direct Notification of Radiology Results

Radiology result follow-up remains a vulnerable area [15,56,57]. Although allowing patients direct access to certain radiology results is gaining some support from radiologists [56], some have expressed concerns about imaging results raising patients' anxiety because the terms used would be unfamiliar and there might be too many requests from patients to meet with radiologists to seek further information and explain the test results [45]. If direct online reporting was instituted, mixed views from physicians and radiologists about how much information patients should have (ie, the full report versus just the conclusions) have been reported [45]. Our study found similar equivocal results in relation to patients' direct access to radiology reports: the majority were not comfortable with patients directly accessing abnormal chest x-rays (although 41% agreed), and the majority (83%) also did not agree with patients being sent abnormal cancer screening tests such as mammography and PAP smear results. In our study, physicians reported that radiology was the most frequently missed test result; if there is no process of systematic feedback to physicians regarding missed results it is difficult to understand how they learn about test results which they, or a colleague, have missed [58-60].

Physicians’ Views Regarding Workload if Results are Directly Notified to Patients

Any intervention introduced into the emergency care context must weigh up the potential impact on the workload of busy emergency clinicians. Our study found that the majority of emergency physicians were not concerned that direct notification of results to patients would increase their workload. In relation to direct notification of normal test results, they were more likely to agree with the process if they believed it would reduce their workload. Other studies have supported this finding [30,32,61].

Physicians’ Views Regarding Responsibility for Test Result Follow-Up

Our study showed that there were mixed views regarding responsibility for notifying patients of abnormal results with a lack of clarity regarding the responsibilities of the ordering physician and the primary care provider. Results pending are particularly prevalent for discharged patients from the ED who often have a short stay. Responsibility for follow-up of results, which may or may not be listed on the ED discharge summary, is impacted by unclear lines of responsibility for follow-up between local medical officers and hospital doctors, combined with emergency physicians’ lack of a continuing relationship with the patient [14,19,62,63]. Electronic discharge summaries can play a role in improving information transfer between hospital and community settings. However, critical information can still be missing [64,65]. Others studies have made recommendations for test follow-up responsibility and these need to be assessed to ensure they apply in the ED context [14,66]. The issue of whether patients’ direct access to test results challenges the physicians’ role as an information gatekeeper has been raised by some [30,49]. However, the majority of emergency physicians in our study did not think that direct notification would interfere with the practice of medicine, so concerns regarding physician role adjustments may be overstated.

How Emergency Physicians Currently Communicate Abnormal Results to Patients

Our study found that most ED physicians notified patients about abnormal test results by telephone. Other studies have shown that patients prefer a direct phone call from the physician for abnormal results [22,26,29,67]; however, studies of email communications between patients and physicians for result notification have also reported positive experiences from patients [28,68]. Physicians still express some concerns, such as managing clinical issues by email and integrating email into office work processes [4].

Conclusions

Future work needs to determine if direct notification of test results to patients leads to improved follow-up of abnormal results. Methods of ensuring patients can access test results directly, including secure Web-based patient portals with drill-down capabilities providing test descriptions and result interpretations need to be evaluated in terms of patient outcomes, cost, and patient usability across socioeconomic groups. Implications of patients “pulling” results from a laboratory information system using a patient portal tied to a hospital-based EMR versus test results being “pushed” automatically to patients needs evaluation.

The fast-paced ED environment presents a number of unique challenges for test result follow-up. Although notification of test results to ED patients has promise, it is important to ensure that methods for direct notification suit the environment and work practices of ED clinicians and laboratories and meet patient needs. Efforts should be directed toward establishing a clear set of recommendations regarding which test results should be directly notified to which patients, methods of notification, and the timing of notification. Emergency physicians’ concerns regarding anxiety, confusion, and lack of patients’ expertise to interpret results should be addressed in order to promote wider test results access to patients.

Acknowledgments

We would like to acknowledge the contributions of Louise Robertson, Michael Stewart, Julie Li, George Touli, and Elia Vecellio. This study is part of an Australian Research Council Discovery Project Grant (DP120100297) to investigate whether technology can make communication in complex systems safer and more efficient.

Dr Giardina is supported by the VA Health Services Research postdoctoral fellowship through the VA Office of Academic Affiliations. Dr Singh is supported by the VA Health Services Research and Development Service (CRE 12-033; Presidential Early Career Award for Scientists and Engineers USA 14-274), the VA National Center for Patient Safety and the Agency for Health Care Research and Quality (R01HS022087). This work is also supported in part by the Houston VA HSR&D Center for Innovations in Quality, Effectiveness and Safety (CIN 13–413).

Authors' Contributions

JC, HS, TG, AG, RP, WR, LL, and JW designed the study. JC, AG, and WR obtained research funding. JC, HS, TG, LL, AG, RP, WR, and JW supervised conduct of the study and data collection. JC, AG, RP, LL, and JW undertook recruitment of participating hospitals and clinicians and managed the data, including quality control. JC, AG, JW, and LL provided statistical advice and analyzed the data. JC drafted the manuscript. All authors contributed substantially to its revision and all authors approve the final version of the manuscript submitted. JC takes responsibility for the paper as a whole.

Conflicts of Interest

None declared.

References

1. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
2. Bates DW, Gawande AA. Improving safety with information technology. N Engl J Med 2003 Jun 19;348(25):2526-2534. [doi: 10.1056/NEJMsa020847] [Medline: 12815139]
3. Hobbs J, Wald J, Jagannath YS, Kittler A, Pizziferri L, Volk LA, et al. Opportunities to enhance patient and physician e-mail contact. Int J Med Inform 2003 Apr;70(1):1-9. [Medline: 12706177]
4. Patt MR, Houston TK, Jenckes MW, Sands DZ, Ford DE. Doctors who are using e-mail with their patients: a qualitative exploration. J Med Internet Res 2003;5(2):e9 [FREE Full text] [doi: 10.2196/jmir.5.2.e9] [Medline: 12857665]
5. Hassol A, Walker JM, Kidder D, Rokita K, Young D, Pierdon S, et al. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. J Am Med Inform Assoc 2004;11(6):505-513 [FREE Full text] [doi: 10.1197/jamia.M1593] [Medline: 15299001]
6. Zhou YY, Kanter MH, Wang JJ, Garrido T. Improved quality at Kaiser Permanente through e-mail between physicians and patients. Health Aff (Millwood) 2010 Jul;29(7):1370-1375 [FREE Full text] [doi: 10.1377/hlthaff.2010.0048] [Medline: 20606190]
7. Ralston JD, Rutter CM, Carrell D, Hecht J, Rubanowice D, Simon GE. Patient use of secure electronic messaging within a shared medical record: a cross-sectional study. J Gen Intern Med 2009 Mar;24(3):349-355 [FREE Full text] [doi: 10.1007/s11606-008-0899-z] [Medline: 19137379]
8. Lin CT, Wittevrongel L, Moore L, Beaty BL, Ross SE. An Internet-based patient-provider communication system: randomized controlled trial. J Med Internet Res 2005 Aug;7(4):e47 [FREE Full text] [doi: 10.2196/jmir.7.4.e47] [Medline: 16236699]
9. Haun JN, Lind JD, Shimada SL, Martin TL, Gosline RM, Antinori N, et al. Evaluating user experiences of the secure messaging tool on the Veterans Affairs' patient portal system. J Med Internet Res 2014;16(3):e75 [FREE Full text] [doi: 10.2196/jmir.2976] [Medline: 24610454]
10. 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]
11. Gandhi TK, Kachalia A, Thomas EJ, Puopolo AL, Yoon C, Brennan TA, et al. Missed and delayed diagnoses in the ambulatory setting: a study of closed malpractice claims. Ann Intern Med 2006 Oct 3;145(7):488-496. [Medline: 17015866]
12. Matheny ME, Gandhi TK, Orav EJ, Ladak-Merchant Z, Bates DW, Kuperman GJ, et al. Impact of an automated test results management system on patients' satisfaction about test result communication. Arch Intern Med 2007 Nov 12;167(20):2233-2239. [doi: 10.1001/archinte.167.20.2233] [Medline: 17998497]
13. Poon EG, Gandhi TK, Sequist TD, Murff HJ, Karson AS, Bates DW. "I wish I had seen this test result earlier!": Dissatisfaction with test result management systems in primary care. Arch Intern Med 2004 Nov 8;164(20):2223-2228. [doi: 10.1001/archinte.164.20.2223] [Medline: 15534158]
14. Singh H, Vij MS. Eight recommendations for policies for communicating abnormal test results. Jt Comm J Qual Patient Saf 2010 May;36(5):226-232. [Medline: 20480756]
15. Callen JL, Westbrook JI, Georgiou A, Li J. Failure to follow-up test results for ambulatory patients: a systematic review. J Gen Intern Med 2012 Oct;27(10):1334-1348 [FREE Full text] [doi: 10.1007/s11606-011-1949-5] [Medline: 22183961]
16. Callen J, Georgiou A, Li J, Westbrook JI. The safety implications of missed test results for hospitalised patients: a systematic review. BMJ Qual Saf 2011 Feb;20(2):194-199 [FREE Full text] [doi: 10.1136/bmjqs.2010.044339] [Medline: 21300992]
17. Singh H, Arora HS, Vij MS, Rao R, Khan MM, Petersen LA. Communication outcomes of critical imaging results in a computerized notification system. J Am Med Inform Assoc 2007;14(4):459-466 [FREE Full text] [doi: 10.1197/jamia.M2280] [Medline: 17460135]
18. Singh H, Thomas EJ, Sittig DF, Wilson L, Espadas D, Khan MM, et al. Notification of abnormal lab test results in an electronic medical record: do any safety concerns remain? Am J Med 2010 Mar;123(3):238-244 [FREE Full text] [doi: 10.1016/j.amjmed.2009.07.027] [Medline: 20193832]
19. Roy CL, Rothschild JM, Dighe AS, Schiff GD, Graydon-Baker E, Lenoci-Edwards J, et al. An initiative to improve the management of clinically significant test results in a large health care network. Jt Comm J Qual Patient Saf 2013 Nov;39(11):517-527. [Medline: 24294680]
20. Georgiou A, Hordern A, Dimigen M, Zogovic B, Callen J, Schlaphoff G, et al. Effective notification of important non-urgent radiology results: a qualitative study of challenges and potential solutions. J Med Imaging Radiat Oncol 2014 Jun;58(3):291-297. [doi: 10.1111/1754-9485.12156] [Medline: 24460883]
21. Meza JP, Webster DS. Patient preferences for laboratory test results notification. Am J Manag Care 2000 Dec;6(12):1297-1300 [FREE Full text] [Medline: 11151807]

22. Grimes GC, Reis MD, Budati G, Gupta M, Forjuoh SN. Patient preferences and physician practices for laboratory test results notification. J Am Board Fam Med 2009;22(6):670-676 [FREE Full text] [doi: 10.3122/jabfm.2009.06.090078] [Medline: 19897696]
23. Spodik M, Goldman J, Merli K, Walker C, Alpini B, Kastenberg D. Providing an endoscopy report to patients after a procedure: a low-cost intervention with high returns. Gastrointest Endosc 2008 Jan;67(1):103-111. [doi: 10.1016/j.gie.2007.08.035] [Medline: 18155431]
24. US Department of Health & Human Services. 2004. Mammography Quality Standards Act (MQSA) URL: http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/Regulations/ucm110823.htm [accessed 2015-02-03] [WebCite Cache ID 6W4zn1Lwj]
25. Federal Register: The Daily Journal of the United States Government. 2014. CLIA Program and HIPAA Privacy Rule; Patients' Access to Test Reports URL: https://www.federalregister.gov/articles/2014/02/06/2014-02280/clia-program-and-hipaa-privacy-rule-patients-access-to-test-reports [accessed 2015-02-03] [WebCite Cache ID 6W4zrQAbM]
26. Baldwin DM, Quintela J, Duclos C, Staton EW, Pace WD. Patient preferences for notification of normal laboratory test results: a report from the ASIPS Collaborative. BMC Fam Pract 2005 Mar 8;6(1):11 [FREE Full text] [doi: 10.1186/1471-2296-6-11] [Medline: 15755328]
27. Keren R, Muret-Wagstaff S, Goldmann DA, Mandl KD. Notifying emergency department patients of negative test results: pitfalls of passive communication. Pediatr Emerg Care 2003 Aug;19(4):226-230. [Medline: 12972818]
28. Couchman GR, Forjuoh SN, Rascoe TG, Reis MD, Koehler B, van Walsum KL. E-mail communications in primary care: what are patients' expectations for specific test results? Int J Med Inform 2005 Jan;74(1):21-30. [doi: 10.1016/j.ijmedinf.2004.08.005] [Medline: 15626633]
29. Leekha S, Thomas KG, Chaudhry R, Thomas MR. Patient preferences for and satisfaction with methods of communicating test results in a primary care practice. Jt Comm J Qual Patient Saf 2009 Oct;35(10):497-501. [Medline: 19886088]
30. Earnest MA, Ross SE, Wittevrongel L, Moore LA, Lin CT. Use of a patient-accessible electronic medical record in a practice for congestive heart failure: patient and physician experiences. J Am Med Inform Assoc 2004;11(5):410-417 [FREE Full text] [doi: 10.1197/jamia.M1479] [Medline: 15187074]
31. 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]
32. Sung S, Forman-Hoffman V, Wilson MC, Cram P. Direct reporting of laboratory test results to patients by mail to enhance patient safety. J Gen Intern Med 2006 Oct;21(10):1075-1078 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00553.x] [Medline: 16836627]
33. Wald JS, Burk K, Gardner K, Feygin R, Nelson E, Epstein M, et al. Sharing electronic laboratory results in a patient portal--a feasibility pilot. Stud Health Technol Inform 2007;129(Pt 1):18-22. [Medline: 17911670]
34. Ross SE, Todd J, Moore LA, Beaty BL, Wittevrongel L, Lin CT. Expectations of patients and physicians regarding patient-accessible medical records. J Med Internet Res 2005;7(2):e13 [FREE Full text] [doi: 10.2196/jmir.7.2.e13] [Medline: 15914460]
35. Bartlett C, Simpson K, Turner AN. Patient access to complex chronic disease records on the Internet. BMC Med Inform Decis Mak 2012;12:87 [FREE Full text] [doi: 10.1186/1472-6947-12-87] [Medline: 22867441]
36. Kittler AF, Carlson GL, Harris C, Lippincott M, Pizziferri L, Volk LA, et al. Primary care physician attitudes towards using a secure web-based portal designed to facilitate electronic communication with patients. Inform Prim Care 2004;12(3):129-138. [Medline: 15606985]
37. Walker J, Leveille SG, Ngo L, Vodicka E, Darer JD, Dhanireddy S, et al. Inviting patients to read their doctors' notes: patients and doctors look ahead: patient and physician surveys. Ann Intern Med 2011 Dec 20;155(12):811-819 [FREE Full text] [doi: 10.7326/0003-4819-155-12-201112200-00003] [Medline: 22184688]
38. Giardina TD, Callen J, Georgiou A, Westbrook JI, Greisinger A, Esquivel A, et al. Releasing Test Results Directly to Patients: A Multisite Survey of Physician Perspectives. Patient Education and Counseling 2015 Feb:1. [doi: 10.1016/j.pec.2015.02.011]
39. Hosmer DWS, Lemeshow S, Sturdivant S. Applied Logistic Regression. 3rd edition. Hoboken, NJ: Wiley; 2013.
40. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996 Dec;49(12):1373-1379. [Medline: 8970487]
41. Ralston JD, Carrell D, Reid R, Anderson M, Moran M, Hereford J. Patient web services integrated with a shared medical record: patient use and satisfaction. J Am Med Inform Assoc 2007;14(6):798-806 [FREE Full text] [doi: 10.1197/jamia.M2302] [Medline: 17712090]
42. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95 [FREE Full text] [doi: 10.1197/jamia.M1833] [Medline: 16221943]
43. Cimino JJ, Patel VL, Kushniruk AW. What do patients do with access to their medical records? Stud Health Technol Inform 2001;84(Pt 2):1440-1444. [Medline: 11604964]
44. Bond MC, Klemt R, Merlis J, Kopinski JE, Hirshon JM. Computer access and Internet use by urban and suburban emergency department customers. J Emerg Med 2012 Jul;43(1):159-165. [doi: 10.1016/j.jemermed.2011.03.034] [Medline: 22142670]

45. Johnson AJ, Frankel RM, Williams LS, Glover S, Easterling D. Patient access to radiology reports: what do physicians think? J Am Coll Radiol 2010 Apr;7(4):281-289. [doi: 10.1016/j.jacr.2009.10.011] [Medline: 20362944]
46. Nordfeldt S, Ängarne-Lindberg T, Berterö C. To use or not to use--practitioners' perceptions of an open web portal for young patients with diabetes. J Med Internet Res 2012;14(6):e154 [FREE Full text] [doi: 10.2196/jmir.1987] [Medline: 23137767]
47. Nordqvist C, Hanberger L, Timpka T, Nordfeldt S. Health professionals' attitudes towards using a Web 2.0 portal for child and adolescent diabetes care: qualitative study. J Med Internet Res 2009;11(2):e12 [FREE Full text] [doi: 10.2196/jmir.1152] [Medline: 19403464]
48. 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]
49. Davis GT, Singh H. Should patients get direct access to their laboratory test results? An answer with many questions. JAMA 2011 Dec 14;306(22):2502-2503. [doi: 10.1001/jama.2011.1797] [Medline: 22122864]
50. Gravis G, Protière C, Eisinger F, Boher JM, Tarpin C, Coso D, et al. Full access to medical records does not modify anxiety in cancer patients: results of a randomized study. Cancer 2011 Oct 15;117(20):4796-4804 [FREE Full text] [doi: 10.1002/cncr.26083] [Medline: 21607939]
51. Wiljer D, Leonard KJ, Urowitz S, Apatu E, Massey C, Quartey NK, et al. The anxious wait: assessing the impact of patient accessible EHRs for breast cancer patients. BMC Med Inform Decis Mak 2010;10:46 [FREE Full text] [doi: 10.1186/1472-6947-10-46] [Medline: 20809950]
52. Wiljer D, Urowitz S, Apatu E, Leonard K, Quartey NK, Catton P. Understanding the support needs of patients accessing test results online. PHRs offer great promise, but support issues must be addressed to ensure appropriate access. J Healthc Inf Manag 2010;24(1):57-63. [Medline: 20077927]
53. Davis GT, Menon S, Parrish DE, Sittig DF, Singh H. Patient access to medical records and healthcare outcomes: a systematic review. J Am Med Inform Assoc 2014;21(4):737-741. [doi: 10.1136/amiajnl-2013-002239] [Medline: 24154835]
54. Karnieli-Miller O, Adler A, Merdler L, Rosenfeld L, Eidelman S. Written notification of test results: meanings, comprehension and implication on patients' health behavior. Patient Educ Couns 2009 Sep;76(3):341-347. [doi: 10.1016/j.pec.2009.07.021] [Medline: 19660891]
55. Dolan NC, Feinglass J, Priyanath A, Haviley C, Sorensen AV, Venta LA. Measuring satisfaction with mammography results reporting. J Gen Intern Med 2001 Mar;16(3):157-162 [FREE Full text] [Medline: 11318910]
56. Berlin L. Communicating results of all outpatient radiologic examinations directly to patients: the time has come. AJR Am J Roentgenol 2009 Mar;192(3):571-573. [doi: 10.2214/AJR.08.1954] [Medline: 19234250]
57. Singh H, Thomas EJ, Mani S, Sittig D, Arora H, Espadas D, et al. Timely follow-up of abnormal diagnostic imaging test results in an outpatient setting: are electronic medical records achieving their potential? Arch Intern Med 2009 Sep 28;169(17):1578-1586 [FREE Full text] [doi: 10.1001/archinternmed.2009.263] [Medline: 19786677]
58. Schiff GD. Minimizing diagnostic error: the importance of follow-up and feedback. Am J Med 2008 May;121(5 Suppl):S38-S42. [doi: 10.1016/j.amjmed.2008.02.004] [Medline: 18440354]
59. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. Am J Med 2008 May;121(5 Suppl):S2-23. [doi: 10.1016/j.amjmed.2008.01.001] [Medline: 18440350]
60. Croskerry P. The feedback sanction. Acad Emerg Med 2000 Nov;7(11):1232-1238. [Medline: 11073471]
61. Liederman EM, Morefield CS. Web messaging: a new tool for patient-physician communication. J Am Med Inform Assoc 2003;10(3):260-270 [FREE Full text] [doi: 10.1197/jamia.M1259] [Medline: 12626378]
62. Greenes DS, Fleisher GR, Kohane I. Potential impact of a computerized system to report late-arriving laboratory results in the emergency department. Pediatr Emerg Care 2000 Oct;16(5):313-315. [Medline: 11063357]
63. Kachalia A, Gandhi TK, Puopolo AL, Yoon C, Thomas EJ, Griffey R, et al. Missed and delayed diagnoses in the emergency department: a study of closed malpractice claims from 4 liability insurers. Ann Emerg Med 2007 Feb;49(2):196-205. [doi: 10.1016/j.annemergmed.2006.06.035] [Medline: 16997424]
64. O'Leary KJ, Liebovitz DM, Feinglass J, Liss DT, Evans DB, Kulkarni N, et al. Creating a better discharge summary: improvement in quality and timeliness using an electronic discharge summary. J Hosp Med 2009 Apr;4(4):219-225. [doi: 10.1002/jhm.425] [Medline: 19267397]
65. Callen JL, Alderton M, McIntosh J. Evaluation of electronic discharge summaries: a comparison of documentation in electronic and handwritten discharge summaries. Int J Med Inform 2008 Sep;77(9):613-620. [doi: 10.1016/j.ijmedinf.2007.12.002] [Medline: 18294904]
66. Schiff GD. Eight questions for getting beyond "getting results". Jt Comm J Qual Patient Saf 2010 May;36(5):224-225. [Medline: 20480755]
67. Fazili J, Ilagan M, Phipps E, Braitman LE, Levine GM. How gastroenterologists inform patients of results after lower endoscopy. Am J Gastroenterol 2001 Jul;96(7):2086-2092. [doi: 10.1111/j.1572-0241.2001.03967.x] [Medline: 11467636]
68. Houston TK, Sands DZ, Jenckes MW, Ford DE. Experiences of patients who were early adopters of electronic communication with their physician: satisfaction, benefits, and concerns. Am J Manag Care 2004 Sep;10(9):601-608 [FREE Full text] [Medline: 15515992]

Abbreviations

ED: emergency department
EMR: electronic medical record
ICT: information and communication technology

Edited by G Eysenbach; submitted 22.07.14; peer-reviewed by M Wong, F Kammerer; comments to author 16.10.14; revised version received 15.12.14; accepted 21.01.15; published 04.03.15.

Please cite as:
Callen J, Giardina TD, Singh H, Li L, Paoloni R, Georgiou A, Runciman WB, Westbrook JI
Emergency Physicians' Views of Direct Notification of Laboratory and Radiology Results to Patients Using the Internet: A Multisite Survey
J Med Internet Res 2015;17(3):e60
URL: http://www.jmir.org/2015/3/e60/
doi:10.2196/jmir.3721
PMID:25739322

Original Paper

“It’s Got to Be on This Page”: Age and Cognitive Style in a Study of Online Health Information Seeking

Emily M Agree[1], MA, PhD; Abby C King[2], PhD; Cynthia M Castro[2], PhD; Adrienne Wiley[3], MSW, MID; Dina LG Borzekowski[4], EdD

[1]Johns Hopkins University, Departments of Sociology and Population, Family, and Reproductive Health, Baltimore, MD, United States

[2]Stanford Prevention Research Center, School of Medicine, Stanford University, Palo Alto, CA, United States

[3]Kauffman Associates, Washington, DC, United States

[4]Dept of Behavioral and Community Health, School of Public Health, University of Maryland, College Park, MD, United States

Corresponding Author:
Emily M Agree, MA, PhD
Johns Hopkins University
Departments of Sociology and Population, Family, and Reproductive Health
530 Mergenthaler hall
3400 North Charles Street
Baltimore, MD, 21218
United States
Phone: 1 410 516 5832
Email: emily.agree@jhu.edu

Abstract

Background: The extensive availability of online health information offers the public opportunities to become independently informed about their care, but what affects the successful retrieval and understanding of accurate and detailed information? We have limited knowledge about the ways individuals use the Internet and the personal characteristics that affect online health literacy.

Objective: This study examined the extent to which age and cognitive style predicted success in searching for online health information, controlling for differences in education, daily Internet use, and general health literacy.

Methods: The Online Health Study (OHS) was conducted at Johns Hopkins School of Public Health and Stanford University School of Medicine from April 2009 to June 2010. The OHS was designed to explore the factors associated with success in obtaining health information across different age groups. A total of 346 men and women aged 35 years and older of diverse racial and ethnic backgrounds participated in the study. Participants were evaluated for success in searching online for answers to health-related tasks/questions on nutrition, cancer, alternative medicine, vaccinations, medical equipment, and genetic testing.

Results: Cognitive style, in terms of context sensitivity, was associated with less success in obtaining online health information, with tasks involving visual judgment most affected. In addition, better health literacy was positively associated with overall success in online health seeking, specifically for tasks requiring prior health knowledge. The oldest searchers were disadvantaged even after controlling for education, Internet use, general health literacy, and cognitive style, especially when spatial tasks such as mapping were involved.

Conclusions: The increasing availability of online health information provides opportunities to improve patient education and knowledge, but effective use of these resources depends on online health literacy. Greater support for those who are in the oldest cohorts and for design of interfaces that support users with different cognitive styles may be required in an age of shared medical decision making.

(J Med Internet Res 2015;17(3):e79) doi:10.2196/jmir.3352

KEYWORDS

eHealth; Internet; health literacy; age groups; field dependence-independence

Introduction

Background

Making informed decisions about health and health care is a key part of enhanced patient care in the twenty-first century. Shared decision making between patients and providers is increasingly the preferred model for health care delivery [1,2]. Improved patient involvement in medical decision making allows more informed decisions overall, as well as better patient perception of medical risks, greater patient engagement, and increased cooperation and compliance among patients [3,4]. For more involvement, patients not only need access to scholarly and popular online health information, but they must also be able to understand the retrieved information [5]. Critical to shared decision making is patients' successful online health seeking where they find "transparent and credible information about the relative value and risk of various medical diagnostic and therapeutic interventions" [6].

Since the Internet's introduction, a common activity has been seeking health information online. At the time of the first Health Information National Trends Survey (HINTS) in 2005, 58% of Internet users had used the Internet to search for health information for themselves [7]. By the fourth wave of HINTS in 2012, this had risen to 77.6% [8]. Websites and online forums offer opportunities to obtain and evaluate information about almost every imaginable health condition, treatment, and/or test. Yet, our knowledge of how individuals retrieve and understand online health information is still limited; we also know little about how individuals' use of online tools affects treatment choices and health outcomes [9]. Research shows that more intensive Internet users prefer greater involvement in medical decisions, but even among these intensive users individual ability to obtain and process online information varies widely [10]. For example, older individuals experiencing cognitive changes and sensory limitations tend to have difficulties interacting with computers and accessing online resources.

Online Health Literacy and Aging

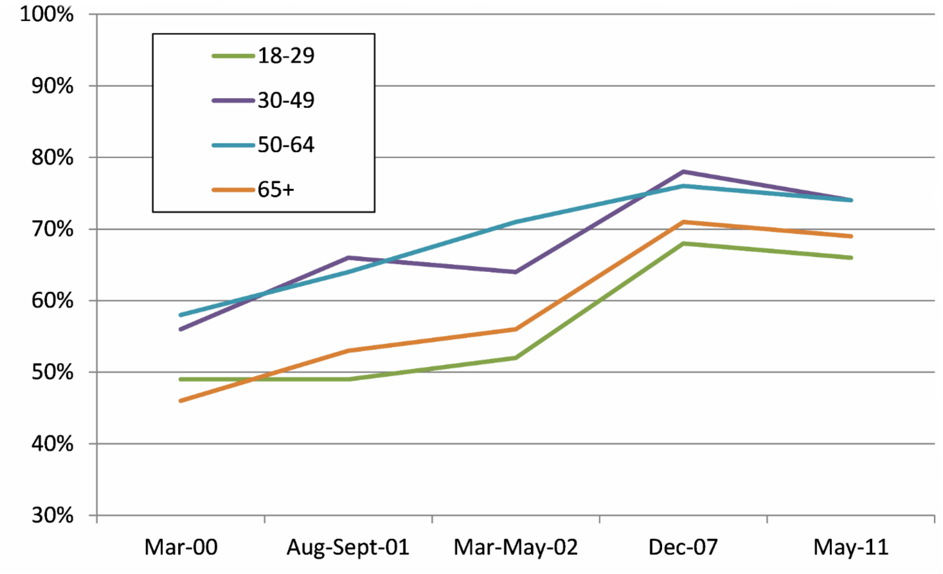
The ability of individuals to participate in informed decisions about their health care depends on the degree to which they have the capacity to obtain, process, and understand health information (health literacy) [11,12]. Online health literacy (sometimes called *eHealth literacy*) is simply the ability of individuals to obtain, process, and understand health information from the Internet.

Older adults are a particularly vulnerable group, characterized by low health literacy and poor health outcomes [13]. Gazmararian and colleagues [14] reported one-third of Medicare managed-care enrollees had difficulty reading written information from health care providers, with the lowest health literacy among those aged 85 years and older. Low health literacy among older adults is compounded by the fact that older adults experience more chronic diseases, take more medications, and visit health care providers more often than younger adults. Often, health information presented to older adults is complex, reflecting multiple health concerns and conditions. As we age, our search for health information can be more demanding; as a result, the Internet's role as a primary repository becomes critically important [15].

Age differences in computer use, skills, and Internet use are well established. Older cohorts are far less likely than younger ones to use computers regularly, more rarely rely on the Internet for information, and report less ease in locating information on the "net" [16]. Despite these differences, the closing of the age-based digital divide is greatly anticipated as new cohorts, more familiar with computers, enter old age. The proportion of adults using the Internet has steadily risen among adults of all ages. For those aged 65 and older, the rate has risen from 13% to almost 60% in just 10 years [17]. The proportion of Internet users who look for health information online also increased over the same period (Figure 1). Those at the oldest and youngest ages are the least likely to seek health information online—the oldest perhaps because they are less trusting of online resources and the youngest because they are less in need of health information, in general.

Basic computer skills—as well as the ability to discriminate among online resources and understand and use information—are paramount for user success obtaining online health information. Age-related declines in sensory abilities and cognition affect visual acuity, especially the ability to discriminate important information in a graphically challenging visual field. Such declines also serve as key factors influencing the usefulness of online resources for older persons [18].

Figure 1. Percentage of Internet users who looked for health information online. Tabulations are drawn from the Pew Internet and American Life Project spreadsheet “Usage Over Time” [18].



Cognitive Style

Cognitive style reflects the different ways individuals solve problems; people vary in how they acquire and process information [19]. A central distinction in the measure of cognitive style is between individuals who are context independent or context sensitive. Those who are context independent are better able to abstract a visual element from its background than those who are context sensitive. Context-independent people often have a more analytic approach to learning; they tend to outperform context-sensitive people in structured learning environments [20]. Persons who may be described as context sensitive observe less differentiation between an object and its context. They tend to be more global in their approach to information and more responsive to visual cues, such as color coding of information [21,22]. Studies show that both context-independent and context-sensitive individuals can perform equally well, but approach learning and use different tools [20,23].

How cognitive style affects success in navigating the Internet deserves increased examination, given the increasing dependency of decision making on online information. Not all websites are the same. Many use increasingly complex interfaces and rely on multimedia content to convey information [24]. Search engine interfaces and the presentation of search results also have received little attention in terms of design and organization [25]. Although accessibility guidelines address sensory and physical deficits, website and search engines are not designed to accommodate differences in learning styles. Some evidence suggests that context sensitivity may increase with age, as cognitive changes affect visual-spatial learning [26]. A better understanding of age-related changes in the perception of screen content, which can affect success in locating and understanding online health information, is needed.

In this study, we examined the extent to which cognitive style (as dichotomized by context independence vs sensitivity) matters for success with online health seeking, controlling for differences in age, education, health literacy, and Internet experience.

Methods

Study Design

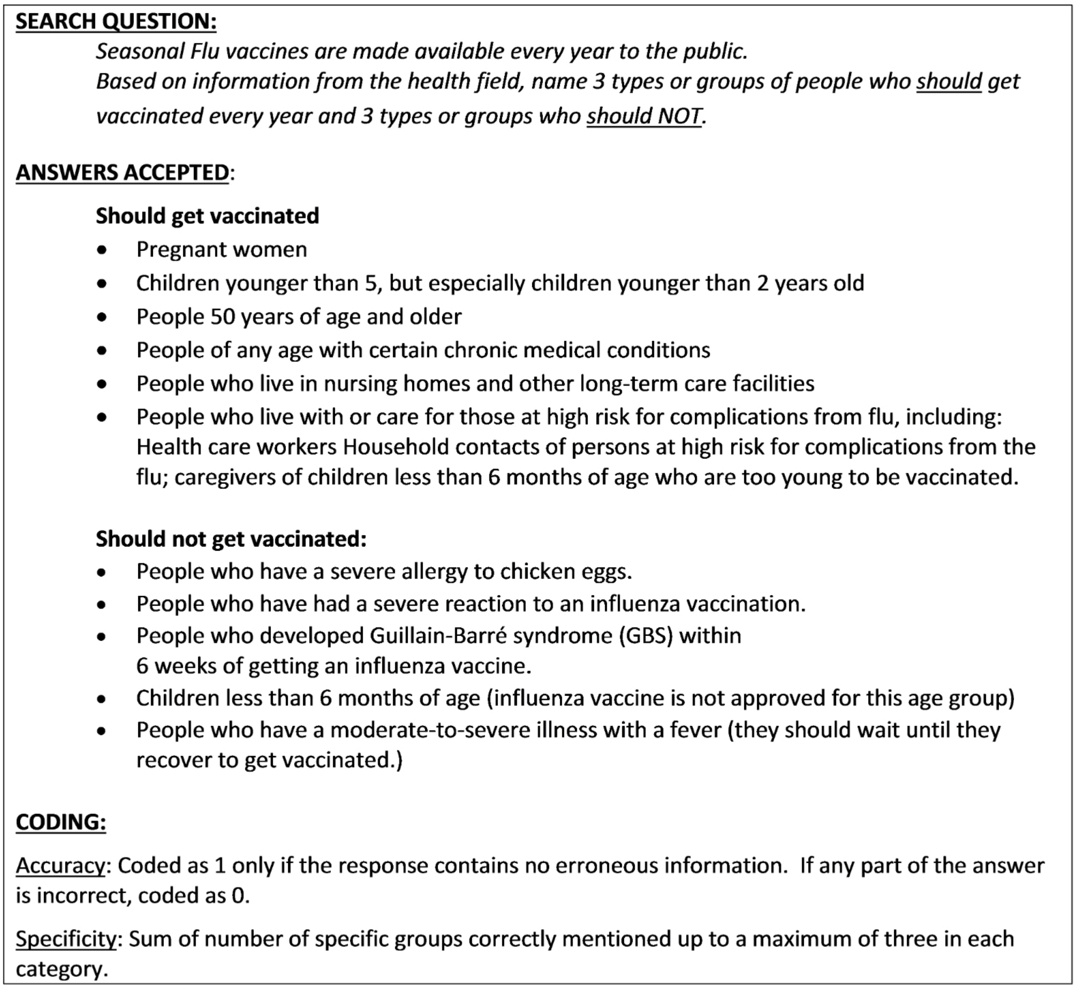
The Online Health Study (OHS) was developed to explore age differences in the strategies used by adults to successfully navigate the Internet for health information and to understand how older adults’ online health literacy compares with that of younger adults. The OHS was designed to examine demographic, cognitive, and environmental factors associated with success obtaining online health information.

The study was conducted at the Johns Hopkins School of Public Health and Stanford University School of Medicine. From April 2009 to June 2010, 346 men and women aged 35 years and older of diverse racial-ethnic backgrounds participated. Participants were recruited from the community and screened for eligibility using a Web-based interface constructed by the research team. Such screening allowed appropriate representation in the sample from different demographic groups and ensured that potential participants had suitable levels of Internet skill to work through the study protocols. Participants who completed all study procedures received a US $35 gift card.

After completing consent procedures, participants provided information on their socioeconomic and demographic backgrounds, health status, and experience with computers and other media. Participants also completed a Rapid Estimate of Adult Literacy in Medicine (REALM) to provide a general measure of health literacy and the Witkin Group Embedded Figures Test (GEFT) to assess context sensitivity or independence.

Participants did a practice search task to familiarize themselves with the protocol and warm-up to using the research computers. Then, participants answered 6 health-related questions by performing online searches on the Hopkins or Stanford project computers. Search time was limited to 15 minutes per task and sessions averaged 60-90 minutes overall. After each online search task, participants reported their answer, which was transcribed and later coded by 2 assistants for response accuracy and specificity. Online search tasks reflected typical and realistic tasks. Health topics covered in the search tasks included diet/nutrition guidelines, skin cancer, alternative medicine, vaccine recommendations, assistive health technology, and over-the-counter genetic testing. In addition to varying on subject matter, each of the tasks required different levels of health literacy and computing skills, including reading texts, reviewing charts, locating health resources on maps, performing simple computations, and evaluating diverse health opinions. For example, the nutrition question asked participants to name 2 heart-healthy foods. This required reading recommendations and lists. The question for assistive technology involved online mapping skills, as participants were asked to locate a store near a specific address where grab bars could be purchased. Figure 2 offers an example of one of the OHS search tasks and the coding guidelines used to score accuracy and success.

Figure 2. Example of an online health task and coding scheme for accuracy and specificity.



SEARCH QUESTION:

Seasonal Flu vaccines are made available every year to the public.
Based on information from the health field, name 3 types or groups of people who should get vaccinated every year and 3 types or groups who should NOT.

ANSWERS ACCEPTED:

Should get vaccinated

- Pregnant women
- Children younger than 5, but especially children younger than 2 years old
- People 50 years of age and older
- People of any age with certain chronic medical conditions
- People who live in nursing homes and other long-term care facilities
- People who live with or care for those at high risk for complications from flu, including: Health care workers Household contacts of persons at high risk for complications from the flu; caregivers of children less than 6 months of age who are too young to be vaccinated.

Should not get vaccinated:

- People who have a severe allergy to chicken eggs.
- People who have had a severe reaction to an influenza vaccination.
- People who developed Guillain-Barré syndrome (GBS) within 6 weeks of getting an influenza vaccine.
- Children less than 6 months of age (influenza vaccine is not approved for this age group)
- People who have a moderate-to-severe illness with a fever (they should wait until they recover to get vaccinated.)

CODING:

Accuracy: Coded as 1 only if the response contains no erroneous information. If any part of the answer is incorrect, coded as 0.

Specificity: Sum of number of specific groups correctly mentioned up to a maximum of three in each category.

Analytic Approach

To examine the relative importance of age, education, health literacy, Internet use, and cognitive style for accuracy and success in online health searches, we analyzed a sample of 323 participants with complete data (93.1% of the 347 participants enrolled in the study) from both sites (128 from Johns Hopkins and 195 from Stanford).

Outcome measures included both overall accuracy (the number of accurate answers on each of the 6 search tasks) with a potential range of 0-6, and success—a scale that combined both accuracy (coded 0-1 for each item) and specificity (coded 0-2 for each item). The potential range for the success scale was 0-18. Models also were estimated for both outcomes with individual search tasks (range of 0-1 for accuracy and 0-2 for success). STATA 11 (StataCorp LP, College Station, TX, USA) was used to estimate generalized linear models (GLM), logistic regression, and ordinal logistic regression models as appropriate to the specific outcome.

Key predictors included an assessment of health literacy using the REALM [27]. The REALM is a vocabulary-based health literacy measure in which each participant reads aloud a list of 66 medical terms; scores are based on correctness of pronunciation. To accommodate the diversity of our study sample, each answer was coded by 2 different assistants as either “0” (incorrect), “1” (incorrect because of a non-English accent), or “2” (correct). The potential range of our coding of the REALM was 0-132.

Cognitive style was measured using the GEFT [19,28] The GEFT involves identification of a visual figure, such as a triangle, within a more complex line drawing. Correct answers are summed in a total score. Those who scored 12 or higher were classified as context independent (treated as the reference group in models), and those scoring 11 or lower were considered to be context sensitive. The GEFT has been validated across settings and age groups; it is also a nonverbal test, which reduces confounding effects of educational attainment. The GEFT complements the use of the REALM for health literacy, which relies on verbal abilities and vocabulary knowledge for assessment.

Results

Participants' ages ranged from 35-90 years, with a stratified design that ensured representation across broad age and sex groups (see Table 1). The sample had a mean age of 55.5 (SD 1.6) years, with approximately 25% (80/323, 24.8%) of the participants aged 60-69 years and 16.4% (53/323) aged 70 or older. More than half (193/323, 59.8%) were female. The sample was not designed to be representative of the entire population, but more closely resembles the general population of Internet users. Most participants were highly educated (65%, 208/323 with college or higher education), reported daily Internet use (88.3%, 285/323), and had good health literacy (REALM mean 128.5, SD 3.6). The mean score on the GEFT was 10.9 (SD 3.0), with 58.8% (190/323) of the sample classified as context sensitive.

Table 1. Participants in the Online Health Study (N=323).

Variable	Participants
Age, n (%)	
35-59 years	190 (58.8)
60-69 years	80 (24.8)
≥70 years	53 (16.4)
Sex, n (%)	
Male	130 (40.2)
Female	193 (59.8)
Education, n (%)	
High school or less	113 (35.2)
College or more	208 (64.8)
Daily Internet use, n (%)	285 (88.3)
Health literacy, mean (SD)	
REALM score	128.5 (3.6)
Cognitive style, n (%)	
Context sensitive	190 (58.8)
Context independent	133 (41.2)

On average, participants answered 4.1 of 6 search tasks with no errors and 29 of 323 participants (9.0%) gave accurate responses to all 6 questions. Performance on individual tasks varied substantially, with the highest percentage giving error-free answers for the question about heart-healthy foods (95.4%, 308/323) and the lowest on the flu question (32.5%, 105/323) (see Table 2). For success in online health seeking, which combined accuracy with the specificity of the responses, participants averaged a score of 8.2 out of 12. Our success measure showed similar variation across question types, with the most success on naming 2 heart-healthy foods and the worst for the question on whether herbs can help with memory loss. Although no question was specifically more difficult than another, they did test different types of health literacy and skills so that variation can be attributed in part to different types of Internet search skills.

Table 2. Scores for accuracy and success for the online health task (N=323).

Topic	Search task/question	Accuracy, %	Success, mean (SD)
Nutrition	Name 2 heart-healthy foods	95.4	1.8 (0.5)
Cancer	How do you identify skin cancer?	81.7	1.6 (0.6)
Alternative medicine	Can herbs help memory?	59.1	0.7(0.6)
Vaccinations	Who should get flu shots?	32.5	1.2 (0.6)
Assistive technology	Where can you find a store selling grab bars?	75.9	1.6 (0.8)
Genetic testing	Should genetic tests be sold over the counter?	70.0	1.3 (0.8)
Overall (all 6 items)		9.0	8.2 (1.8)

Table 3 shows regression models for success in online health seeking. In the first model, those in the oldest age group had significantly lower success scores relative to younger participants. There were no significant sex differences, although having a college degree and daily Internet use both associated positively with more successful health information searches. In Model 2, we added the REALM measure of health literacy, also significantly associated with better success in searching. In the final model, we added the indicator for cognitive style, which demonstrated a strong negative association with success. When controlling for education, health literacy, current Internet use, and context sensitivity throughout the models, the oldest individuals exhibited a significant disadvantage in the successful attainment of online health information.

Table 3. Coefficients for generalized linear models predicting overall search success.

Variables	Success, coefficient (95% CI)		
	Model 1	Model 2	Model 3
Age			
35-59 years	Reference	Reference	Reference
60-69 years	–0.36 (–0.84, 0.11)	–0.37 (–0.84, 0.10)	–0.32 (–0.79, 0.15)
≥70 years	–0.71 (–1.28,–0.14)	–0.86 (–1.43, –0.29)	–0.78 (–1.34, –0.21)
Sex			
Male	Reference	Reference	Reference
Female	0.22 (–0.19, 0.62)	0.17 (–0.23, 0.57)	0.23 (–0.17,0.63)
Education			
High school or less	Reference	Reference	Reference
Some college or more	0.51 (0.07, 0.94)	0.44 (0.00,0.87)	0.36 (–0.07, 0.80)
Daily Internet use	0.94 (0.35, 1.54)	0.80 (0.21, 1.40)	0.80 (0.02, 0.13)
Health literacy			
REALM score		0.09 (0.04, 0.15)	0.08 (0.22,1.39)
Cognitive style			
Context independent			Reference
Context sensitive			–0.58 (–0.99, –0.18)
Constant	7.12 (6.47, 7.77)	–4.29 (–11.31, 2.72)	–2.36 (–9.45, 4.72)

Table 4 displays ordered logistic regression models for success for each specific search task. Outcomes varied with respect to sex, daily Internet use, health literacy, and cognitive style when successfully seeking online health information. Those in the oldest age group showed a negative association with overall success for all the tasks, although the effects for age were significant only for a specific task involving mapping skills (wherein participants needed to locate a store selling grab bars). Being female was only significantly associated with the ability to report visual indicators of cancerous skin growths: women were more than twice as likely to be more successful than men, controlling for other variables. Having a higher education was positively associated with success across several—although not all—tasks. It was not associated with success for questions about heart-healthy food or public guides for identifying cancerous skin growths. Surprisingly, during the search task that asked participants to investigate the ability of herbal supplements to reduce memory loss, those with more education were less likely to be successful than those with less education.

Table 4. Odds ratios for ordered logistic regression models predicting success with specific online health-seeking tasks.

Variables	Online health-seeking task, OR (90% CI)					
	Nutrition	Cancer	Alternative medicine	Vaccinations	Mapping	Genetic testing
Age						
35-59 years	Reference	Reference	Reference	Reference	Reference	Reference
60-69 years	0.80 (0.47, 1.36)	0.92 (0.57, 1.50)	0.71 (0.46, 1.11)	1.21 (0.76, 1.92)	0.59 (0.34, 1.00)	0.85 (0.55, 1.32)
≥70	0.59 (0.31, 1.10)	0.76 (0.42, 1.38)	0.60 (0.35, 1.03)	0.78 (0.44, 1.40)	0.44 (0.24, 0.83)	0.71 (0.42, 1.22)
Female	0.80 (0.50, 1.29)	2.33 (1.54, 3.54)	0.88 (0.60, 1.29)	1.08 (0.73, 1.61)	0.92 (0.58, 1.45)	1.35 (0.93, 1.96)
Education						
High school or less	Reference	Reference	Reference	Reference	Reference	Reference
College or more	0.98 (0.59, 1.62)	1.09 (0.69, 1.73)	0.34 (0.22, 0.53)	1.67 (1.08, 2.59)	2.00 (1.23, 3.27)	2.08 (1.40, 3.10)
Daily Internet use	1.77 (0.96, 3.25)	2.27 (1.28, 4.04)	3.73 (2.06, 6.75)	1.32 (0.72, 2.43)	0.84 (0.43, 1.62)	1.10 (0.64, 1.90)
REALM	1.02 (0.96, 1.09)	1.04 (0.99, 1.11)	0.95 (0.90, 1.01)	1.07 (1.01, 1.13)	1.04 (0.98, 1.11)	1.11 (1.06, 1.17)
Cognitive style						
Context independent	Reference	Reference	Reference	Reference	Reference	Reference
Context sensitive	0.54 (0.33, 0.88)	0.57 (0.37, 0.88)	0.73 (0.50, 1.08)	0.87 (0.58, 1.29)	0.59 (0.37, 0.96)	1.00 (0.68, 1.46)

Health literacy and cognitive style were significant for different outcomes. A higher REALM score was positively associated with success on the search task about the flu vaccine and over-the-counter genetic tests. Context sensitivity was negatively associated with naming heart-healthy foods, identifying cancerous skin growths, and locating a store that sells grab bars. Context-sensitive participants were roughly half as likely to be successful compared to context-independent individuals

Discussion

The increasing availability of online health resources would suggest improved patient education and knowledge; however, this outcome requires successful online health literacy. Patients need to retrieve and comprehend online health information in order for it to positively impact decision making and health care. Results from this study suggest that the oldest health seekers may be at a disadvantage compared to younger cohorts, even after controlling for technology use, education, health literacy, and cognitive style, especially when spatial tasks such as mapping are involved.

Cognitive style was hypothesized to be particularly important for success in online health seeking. Our results reveal that context sensitivity was associated with less success in obtaining online health information, with specific tasks involving visual judgment and mapping most affected. These results are consistent with the idea that individuals who are context sensitive will indeed be most greatly affected by tasks that involve spatial abilities (such as using mapping tools to find a store or understanding the visual illustrations of skin cancer). They also seem to perform worse in the relatively straightforward identification of heart-healthy foods, in part because they tend to be less specific and thus more vague about the types of food items that are recommended.

In addition, better health literacy seems important both for overall success in searching and specifically for questions requiring prior health knowledge, such as flu vaccine recommendations or the sale of over-the-counter genetic tests. Interestingly, those with greater education were less likely to answer the question about alternative medicine correctly. This is consistent with other research that has shown that complementary and alternative medicine has been taken up at greater rates by those with higher education levels who are more skeptical of the allopathic medical profession [29].

Study limitations should be acknowledged. Although the GEFT has been used extensively in studies of hypermedia learning especially to examine differences in website design and other graphical interfaces [30], it also can be affected by participants' visual acuity and other abilities. As with previous studies, we employ a dichotomous indicator of context sensitivity, but it may be more appropriate to examine cognitive style as a continuum [31]. The extent to which Internet searchers have prior knowledge about a topic can also affect their approaches to navigating the Internet [32]. Those with more expertise tend to seek out specific information, whereas novices benefit from a broader overview that orients them to the subject. Finally, this research evaluated participants' online health seeking during a single monitored session. There is evidence that patients develop avenues for online health information over time, employing iterative processes that involve multiple sessions over longer time periods [33]. Some OHS study group participants may have been more successful in their online health seeking if given more time to search online or assisted by a family member or health provider.

The present study raises issues concerning online health communication, suggesting that vulnerable populations may need targeted assistance if this is to be a primary source of information in decision making. Older cohorts as well as those with different cognitive styles had difficulties finding accurate, detailed information. When posed very specific questions, some people can easily find information; others are lost, even when the information seems readily apparent on a particular webpage.

It is difficult to control the online health information to which people will be exposed. Websites change continuously and search results can yield different outcomes over a course of a single day. Health educators and providers can potentially serve as a resource, providing guidance to reliable and credible websites that maintain appropriate standards of content and usability [34]. They could also help interpret search findings, especially when patients obtain inaccurate, vague answers to the health questions. However, this recommendation requires that health educators and providers have sufficient time and skills to serve these functions, which may not be the case. Other community resources (such as librarians or adult educators) could possibly help deliver this support to those with low online health literacy.

Adaptations to health information websites could reduce distracting content and highlight key sections, as this has been shown to help persons with context sensitivity [35]. Websites about skin cancer could supplement visual information with verbal descriptions in order not to lose patients whose cognitive style is more context sensitive. In addition, educational research has shown that context-sensitive individuals benefit from additional levels of structure and guidance in information materials [36]. Thus, indexes (or more recent tools such as drop-down menus) might be useful additions to websites that would help these users [30].

However, the rising dominance of Google search engines as the primary interface for searching means that indexes and other website design tools may be receding in importance [37]. Thus, the promotion of standards for health information website design, similar to the accessibility standards aimed to help Internet users with disabilities, should be only one aspect of design taken into account. Rather than focusing on individual websites, information technologists could design applications such as browser add-ons or mobile apps that improve online interfaces and thus reduce some of the cognitive demands involved in searching. In light of the growing importance of and reliance on online health information-seeking across all age groups, such improvements may prove necessary and critical for improved public health.

Acknowledgments

This research has been supported by NIH grant R01AG026430, with additional contributions from the Hopkins Population Center (R24HD042854) and the Hopkins Center for Population Aging and Health (P30AG34460.) The authors would like to thank Michele Trieb and several research assistants for their dedication to preparing data for this analysis.

Conflicts of Interest

None declared.

References

1. Kon AA. The shared decision-making continuum. JAMA 2010 Aug 25;304(8):903-904. [doi: 10.1001/jama.2010.1208] [Medline: 20736477]
2. Deber RB, Kraetschmer N, Irvine J. What role do patients wish to play in treatment decision making? Arch Intern Med 1996 Jul 8;156(13):1414-1420. [Medline: 8678709]
3. LeBlanc A, Kenny DA, O'Connor AM, Légaré F. Decisional conflict in patients and their physicians: a dyadic approach to shared decision making. Med Decis Making 2009;29(1):61-68. [doi: 10.1177/0272989X08327067] [Medline: 19196706]
4. Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011(10):CD001431. [doi: 10.1002/14651858.CD001431.pub3] [Medline: 21975733]
5. Krist AH, Woolf SH. A vision for patient-centered health information systems. JAMA 2011 Jan 19;305(3):300-301. [doi: 10.1001/jama.2010.2011] [Medline: 21245186]
6. Cassel CK, Guest JA. Choosing wisely: helping physicians and patients make smart decisions about their care. JAMA 2012 May 2;307(17):1801-1802. [doi: 10.1001/jama.2012.476] [Medline: 22492759]
7. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]
8. National Cancer Institute. Health Information National Trends Survey (B4). Washington DC: National Institutes of Health In the past 12 months, have you done the following things while using the Internet? URL: http://hints.cancer.gov/question-details.aspx?dataset=41&method=mail&qid=757 [accessed 2014-02-20] [WebCite Cache ID 6NXAga5K7]
9. Gustafson DH, Wyatt JC. Evaluation of ehealth systems and services. BMJ 2004 May 15;328(7449):1150 [FREE Full text] [doi: 10.1136/bmj.328.7449.1150] [Medline: 15142895]
10. Xie B, Wang M, Feldman R, Zhou L. Internet use frequency and patient-centered care: measuring patient preferences for participation using the health information wants questionnaire. J Med Internet Res 2013;15(7):e132 [FREE Full text] [doi: 10.2196/jmir.2615] [Medline: 23816979]

11. Nielsen-Bohlman L, Panzer AM, Kindig DA. Health Literacy: A Prescription to End Confusion. Washington, DC: National Academies Press; 2004.
12. US Department of Health and Human Services. Healthy People 2010: Understanding and Improving Health. 2nd ed. Washington, DC: US Government Printing Office; 2000.
13. Berkman ND, Dewalt DA, Pignone MP, Sheridan SL. Literacy and health outcomes: summary. Evidence Report/Technology Assessment No 87. Rockville, MD: Agency for Health Care Research and Quality; 2004.
14. Gazmararian JA, Baker DW, Williams MV, Parker RM, Scott TL, Green DC, et al. Health literacy among Medicare enrollees in a managed care organization. JAMA 1999 Feb 10;281(6):545-551. [Medline: 10022111]
15. Voelker R. Seniors seeking health information need help crossing "digital divide". JAMA 2005 Mar 16;293(11):1310-1312. [doi: 10.1001/jama.293.11.1310] [Medline: 15769951]
16. Stronge AJ, Rogers WA, Fisk AD. Web-based information search and retrieval: effects of strategy use and age on search success. Hum Factors 2006;48(3):434-446. [Medline: 17063960]
17. Zickuhr K, Madden M. Older Adults and Internet Use. Washington, DC: Pew Internet & American Life Project; 2012 Jun 6. URL: http://www.pewinternet.org/Reports/2012/Older-adults-and-internet-use.aspx [accessed 2014-02-11] [WebCite Cache ID 6NJOh4D8W]
18. Laberge JC, Scialfa CT. Predictors of web navigation performance in a life span sample of adults. Hum Factors 2005;47(2):289-302. [Medline: 16170939]
19. Witkin HA, Goodenough DR. Cognitive styles: essence and origins. Field dependence and field independence. Psychol Issues 1981(51):1-141. [Medline: 7012884]
20. Liu M, Reed W. The relationship between the learning strategies and learning styles in a hypermedia environment. Computers in Human Behavior 1994 Dec;10(4):419-434. [doi: 10.1016/0747-5632(94)90038-8]
21. Witkin HA, Moore CA, Goodenough D, Cox PW. Field-dependent and Field-independent cognitive styles and their educational implications. Review of Educational Research 1977 Jan 01;47(1):1-64. [doi: 10.3102/00346543047001001]
22. Thompson B, Melancon JG. Measurement characteristics of the group embedded figures test. Educational and Psychological Measurement 1987 Sep 01;47(3):765-772. [doi: 10.1177/0013164487047003330]
23. DeTure M. Cognitive style and self-efficacy: predicting student success in online distance education. American Journal of Distance Education 2004 Mar;18(1):21-38. [doi: 10.1207/s15389286ajde1801_3]
24. Huang AH, Yen DC, Hung S. Research issues of Internet-integrated cognitive style. Computers in Human Behavior 2012 Sep;28(5):1547-1554 [FREE Full text] [doi: 10.1016/j.chb.2012.03.014] [Medline: 25319526]
25. Kerkmann F, Lewandowski D. Accessibility of web search engines. Library Review 2012 Aug 31;61(8/9):608-621. [doi: 10.1108/00242531211292105]
26. Schwartz DW, Karp SA. Field dependence in a geriatric population. Percept Mot Skills 1967 Apr;24(2):495-504. [doi: 10.2466/pms.1967.24.2.495] [Medline: 6040222]
27. Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med 1993 Jun;25(6):391-395. [Medline: 8349060]
28. Witkin HA, Oltman PK. Cognitive syle. Int J Neurol 1967;6(2):119-137. [Medline: 4885959]
29. Neiberg RH, Aickin M, Grzywacz JG, Lang W, Quandt SA, Bell RA, et al. Occurrence and co-occurrence of types of complementary and alternative medicine use by age, gender, ethnicity, and education among adults in the United States: the 2002 National Health Interview Survey (NHIS). J Altern Complement Med 2011 Apr;17(4):363-370 [FREE Full text] [doi: 10.1089/acm.2009.0157] [Medline: 21495904]
30. Chen SY, Macredie RD. Cognitive styles and hypermedia navigation: Development of a learning model. J Am Soc Inf Sci 2001;53(1):3-15. [doi: 10.1002/asi.10023]
31. Cook DA. Revisiting cognitive and learning styles in computer-assisted instruction: not so useful after all. Acad Med 2012 Jun;87(6):778-784. [doi: 10.1097/ACM.0b013e3182541286] [Medline: 22534603]
32. Chen SY, Fan JP, Macredie RD. Navigation in hypermedia learning systems: experts vs. novices. Computers in Human Behavior 2006 Mar;22(2):251-266. [doi: 10.1016/j.chb.2004.06.004]
33. 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]
34. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. Cyberpsychol Behav 2004 Oct;7(5):497-510. [Medline: 15667044]
35. Silvana MA, Oliveira R, Fortes D. Context sensitive accessibility aid to middle-aged adults and elderly users in web systems. In: Proceedings of the 29th ACM international conference on Design of communication. 2011 Oct 03 Presented at: SIGDOC; 2011; New York. [doi: 10.1145/2038476.2038482]
36. Cook DA. Learning and cognitive styles in web-based learning: theory, evidence, and application. Acad Med 2005 Mar;80(3):266-278. [Medline: 15734809]
37. Sharit J, Hernández MA, Czaja SJ, Pirolli P. Investigating the roles of knowledge and cognitive abilities in older adult information seeking on the web. ACM Trans Comput Hum Interact 2008;15(1):3 [FREE Full text] [doi: 10.1145/1352782.1352785] [Medline: 20011130]

Abbreviations

GEFT: Group Embedded Figures Test
GLM: generalized linear model
HINTS: Health Information National Trends Survey
OHS: Online Health Study
REALM: Rapid Estimate of Adult Literacy in Medicine

Edited by G Eysenbach; submitted 25.02.14; peer-reviewed by M Masoni, C Bergeron; comments to author 17.07.14; revised version received 29.12.14; accepted 22.01.15; published 24.03.15.

Please cite as:
Agree EM, King AC, Castro CM, Wiley A, Borzekowski DLG
"It's Got to Be on This Page": Age and Cognitive Style in a Study of Online Health Information Seeking
J Med Internet Res 2015;17(3):e79
URL: http://www.jmir.org/2015/3/e79/
doi:10.2196/jmir.3352
PMID:25831483

Original Paper

eHealth Literacy and Web 2.0 Health Information Seeking Behaviors Among Baby Boomers and Older Adults

Bethany Tennant[1], PhD; Michael Stellefson[2], PhD; Virginia Dodd[3], MPH, PhD; Beth Chaney[4], MCHES, PhD; Don Chaney[4], MCHES, PhD; Samantha Paige[2], MPH; Julia Alber[2], MPH

[1]ICF International, 9300 Lee Highway, Fairfax, VA, United States
[2]Center for Digital Health and Wellness, Department of Health Education and Behavior, University of Florida, Gainesville, FL, United States
[3]Department of Community Dentistry and Behavioral Science, University of Florida, Gainesville, FL, United States
[4]Department of Health Education & Promotion, East Carolina University, Greenville, NC, United States

Corresponding Author:
Michael Stellefson, PhD
Center for Digital Health and Wellness
Department of Health Education and Behavior
University of Florida
PO Box 118210
Gainesville, FL,
United States
Phone: 1 352 294 1805
Fax: 1 352 392 1909
Email: mstellefson@ufl.edu

Abstract

Background: Baby boomers and older adults, a subset of the population at high risk for chronic disease, social isolation, and poor health outcomes, are increasingly utilizing the Internet and social media (Web 2.0) to locate and evaluate health information. However, among these older populations, little is known about what factors influence their eHealth literacy and use of Web 2.0 for health information.

Objective: The intent of the study was to explore the extent to which sociodemographic, social determinants, and electronic device use influences eHealth literacy and use of Web 2.0 for health information among baby boomers and older adults.

Methods: A random sample of baby boomers and older adults (n=283, mean 67.46 years, SD 9.98) participated in a cross-sectional, telephone survey that included the eHealth literacy scale (eHEALS) and items from the Health Information National Trends Survey (HINTS) assessing electronic device use and use of Web 2.0 for health information. An independent samples *t* test compared eHealth literacy among users and non-users of Web 2.0 for health information. Multiple linear and logistic regression analyses were conducted to determine associations between sociodemographic, social determinants, and electronic device use on self-reported eHealth literacy and use of Web 2.0 for seeking and sharing health information.

Results: Almost 90% of older Web 2.0 users (90/101, 89.1%) reported using popular Web 2.0 websites, such as Facebook and Twitter, to find and share health information. Respondents reporting use of Web 2.0 reported greater eHealth literacy (mean 30.38, SD 5.45, n=101) than those who did not use Web 2.0 (mean 28.31, SD 5.79, n=182), $t_{217.60}=-2.98$, $P=.003$. Younger age ($b=-0.10$), more education ($b=0.48$), and use of more electronic devices ($b=1.26$) were significantly associated with greater eHealth literacy ($R^2=.17$, R^2adj =.14, $F_{9,229}=5.277$, $P<.001$). Women were nearly three times more likely than men to use Web 2.0 for health information (OR 2.63, Wald= 8.09, df=1, $P=.004$). Finally, more education predicted greater use of Web 2.0 for health information, with college graduates (OR 2.57, Wald= 3.86, df =1, $P=.049$) and post graduates (OR 7.105, Wald= 4.278, df=1, $P=.04$) nearly 2 to 7 times more likely than non-high school graduates to use Web 2.0 for health information.

Conclusions: Being younger and possessing more education was associated with greater eHealth literacy among baby boomers and older adults. Females and those highly educated, particularly at the post graduate level, reported greater use of Web 2.0 for health information. More in-depth surveys and interviews among more diverse groups of baby boomers and older adult populations will likely yield a better understanding regarding how current Web-based health information seeking and sharing behaviors influence health-related decision making.

(J Med Internet Res 2015;17(3):e70) doi:10.2196/jmir.3992

KEYWORDS

social media; aging; health literacy; Web 2.0

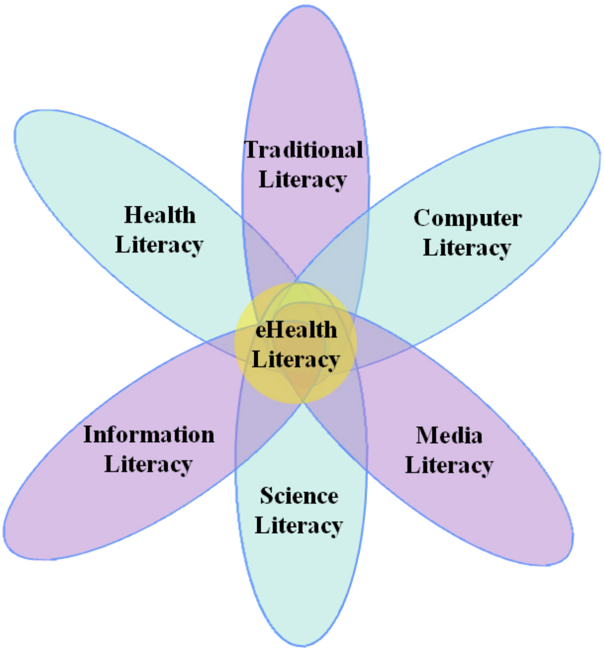
Introduction

Over the past several decades, inequities in Internet availability and accessibility have diminished due to technological advances and lower-cost access to broadband Internet. Currently, over 2.8 billion people use the Internet worldwide [1], with estimates indicating that nearly 90% of adults regularly access the Internet for information [2]. Greater access to the Internet has increased the availability of health information [3-5], yet many Internet users continue to face challenges accessing relevant and literacy-sensitive health and medical content that is of high quality [4,6-13]. Individuals without adequate skills to navigate the Internet may also unknowingly access health information that is inaccurate and potentially dangerous to their overall health [11,12,14]. This phenomenon is especially problematic for the aging population who is at particularly high risk for disability and chronic disease [15]. Compared to their younger counterparts, older adults are more likely to have lower health literacy that negatively impacts health care access, chronic disease management, and health status [16,17].

Although older adults are traditionally “late adopters” of technology, research conducted by the Pew Research Center’s Internet and American Life Project indicates that more than half (59%) of adults 65 years and over [18], and 88% of baby boomers between 50 and 64 years access the Internet [19]. Approximately 74% of older adults and 88% of baby boomers use a cellular device, and an increasing number are now beginning to use advanced digital devices with mobile Internet access [20]. One common reason that baby boomers and older adults use these electronic devices is to seek out relevant Web-based health information [21]. For example, in a recent study by Medlock et al [22], researchers found that the Internet was a trusted source of health information among older adults, especially for learning more about the prognosis, symptoms, and treatment options for personal health issues.

While the older adult population is becoming more and more reliant on the Internet to locate and obtain health-related information and services [2,6,23], baby boomers and older may struggle to possess adequate eHealth literacy [12,24]. eHealth literacy is defined as the ability to seek, find, understand, and appraise health information from electronic resources and apply that knowledge to solving a health problem or making a health-related decision [25]. The construct of eHealth literacy represents a foundational skill set that combines six forms of literacy that extend beyond traditional definitions of health literacy and numeracy to include: (1) traditional, (2) information, (3) media, (4) health, (5) scientific, and (6) computer (Figure 1) [25].

Figure 1. eHealth literacy Lily Model [25].



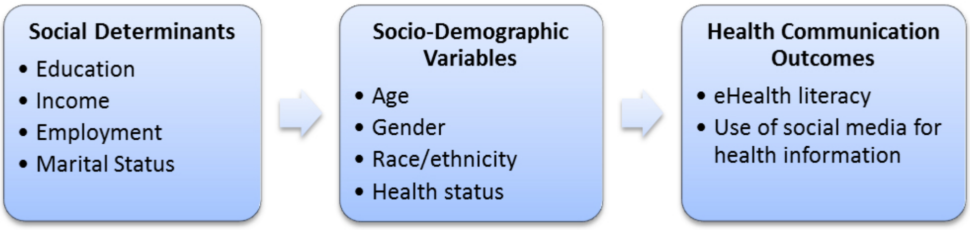
Two context-specific domains of eHealth literacy that are particularly noteworthy to measure among Internet users beyond the age of 50 years include *health* and *computer* (or *digital*) literacy [24,26,27]. Health literacy is defined as the degree to which individuals can obtain, process, and understand the basic health information and services they need to make appropriate health decisions [28]. Results from a systematic literature review that examined the role that health literacy plays in the treatment of older adults indicated that lower health literacy is associated with increased health care costs, more medication errors, ineffective and undesirable patient-provider communication, as well as inefficient use of health care services [29]. Computer (or digital) literacy involves an individual’s ability to adapt to new technologies productively and efficiently to solve problems or answer questions through the operation of an electronic device [25,26]. Computer literacy is especially important to consider among the aging population because adroit use of technology may help reduce cognitive decline among older adults 50 to 89 years [30]. Research suggests that determinants of computer literacy include knowledge about technology [21], exposure to electronic devices [31], and the type and number of electronic devices that are used [25]. Low computer literacy in older adults often precludes these populations from successfully accessing and deciphering high-quality sources of Web-based health information [6,26,32]. Both health and computer literacy are not static; rather, they are influenced by an individual's health status, motivation, education level, and changes in technology [25]. Without adequate attention to health and computer (or digital) literacy among the aging population, there is a risk of reopening the digital divide, solidifying current health disparities, and perpetuating inequities that result in behavioral risk factors that compromise patient safety and reduce health outcomes among these vulnerable populations.

While the Internet has traditionally been used as a one-way health communication channel (ie, Web 1.0) [33], the concept of "participative Internet" (ie, Web 2.0) has risen in popularity due to the advent of social networking, which facilitates multi-way conversations about health [24,34-36]. Web 2.0 has transformed health communication patterns, allowing users to add information or content on the Web [37] and collaborate with others on issues related to health care [34-36]. Although baby boomers and older adults have traditionally been identified as "passive consumers" of health information on using Web 1.0 [38], Web 2.0 provides new opportunities for promoting health and preventing behavioral risk factors associated with chronic disease. A recent study suggests that adults between the ages of 50 and 60 years living with a compromised health status utilize the Internet for health care purposes because they want to be active in their health care decision-making [39]. For example, some older adults use email and interactive communication tools on the Internet to promote cancer screening to their peers [40]. Virtual discussion-based forums for patient engagement now also target individuals living with long-term health problems [31,35,41,42], which has caused the number of customized Internet applications for chronic disease-related behavioral risk management to grow [42,43]. Individuals with a primary health care provider, chronic disease, and those who are younger are more likely to use social networking sites for health-related activities [44]. However, baby boomers and older adults report not accessing or utilizing Web 1.0 and Web 2.0 for a number of reasons, including the high cost of devices and Internet access, insufficient knowledge about device function, and poor perceived self-efficacy [31,45].

Currently, there is a dearth of information regarding which sociodemographic and social determinant variables, other than age, education, and income [38], are associated with eHealth literacy and use of Web 2.0 for health care purposes among baby boomers and older adult populations. Preliminary research suggests that eHealth literacy is negatively associated with age among low income homebound adults above 60 years of age [31], but the literature is not definitive regarding relationships between social determinants, sociodemographic variables, eHealth literacy, and use of social networking sites for health promotion [36,37]. Some research indicates that baby boomers are significantly more likely than older adults to use health information websites, email, automated call centers, medical video conferencing, texting, and podcasts for health care purposes [46], but it is unclear whether aging populations have confidence in their ability to utilize these technologies to find and evaluate Web-based health information. The Structural Influence Model of Health Communication (SIMHC) postulates that different forms of media and different genres within a medium (ie, using the Internet as a one-way communication channel, Web 1.0 vs using social media as a two-way communication channel, Web 2.0) may differentially influence health information seeking and sharing behaviors among different populations [47]. Further, SIMHC posits that media communications influences health by raising awareness, focusing attention on health, highlighting relevant health issues, providing health information, and reinforcing health-related knowledge, attitudes, and behaviors [48].

The demographic composition of the United States will undoubtedly be reshaped by the baby boomer generation in the coming decades. Baby boomers already make up a large proportion of the population [49], and by 2029, 20% of the United States population will be over the age of 65 years [49]. With increased age comes concomitant demands for health care resources; therefore, it is important to examine whether people in the baby boomer and older adult age group are confident in their ability to access and effectively navigate Web-based health resources to obtain quality health information that will allow for informed decision making. At the current time, it is unknown whether or not health status or electronic device use is associated with eHealth literacy and/or use of Web 2.0 for health promotion among adults 50 years and older [50]. Therefore, the purpose of this study was to determine the extent to which social determinants such as electronic device use and sociodemographic variables included in the SIMHC (Figure 2) were associated with distinct health communication outcomes (ie, eHealth literacy and use of Web 2.0 to find and evaluate health information), in baby boomers and older adults.

Figure 2. Adapted Structural Influence Model of Health Communication.



Methods

Recruitment

In February 2013, a cross-sectional telephone survey was conducted as part of the state of Florida Consumer Confidence Index (F-CCI) Survey, administered by the University of Florida Bureau of Economic and Business Research (BEBR). The BEBR conducts and disseminates demographic and economic research on residents of the State of Florida to inform public policy [51]. At least 500 households in the state are surveyed on a monthly basis, using the random digit dialing (RDD) method. A minimum of 10 call attempts are placed per household. Dillman supports the use of telephone surveys for collecting data among aging populations who often feel reassured that they are speaking with an actual person on the telephone, as opposed to answering questions via other modes (eg, paper-based, Internet) [52]. Throughout February 2013, telephone surveys were administered Monday through Friday between 9 AM and 9 PM, Saturdays between 12 PM and 6 PM, and Sundays between 3 PM and 9 PM. A total of 6695 telephone calls were placed during this time period, and a total of 493 individuals agreed to complete the telephone survey.

Participants

Respondents were included in main analyses if they (1) reported being 50 years of age or older, and (2) had ever accessed the Internet or sent/received email messages. The youngest baby boomers just recently turned 50 years of age , having been born between 1946 and 1964 [49]. A total of 393 respondents in the sample reported being 50 years of age or over, yet 110 respondents responded “no” to the following question adopted from the Health Information National Trends Survey (HINTS) [53]: “Do you ever go online to access the Internet or World Wide Web, or to send and receive email?” Therefore, data from a total of 283 respondents was analyzed in this study. Human subjects approval was secured from the university’s Institutional Review Board (IRB) prior to administering the telephone survey or analyzing any participant data.

Measurement

Electronic Device Use

Electronic device use was measured using one item adapted from the HINTS survey [53]. Participants were asked, “In the past 12 months, have you used the Internet on any of the following devices to look for health or medical information for yourself?” Respondents could select any devices from the following list: (1) desktop computer, (2) laptop computer, (3) cell phone, or (4) mobile handheld device like an e-reader or tablet.

eHealth Literacy

eHealth literacy was measured using the eHealth Literacy Scale (eHEALS) [54]. The eHEALS determines consumers’ combined knowledge, confidence, and perceived skills finding, evaluating, and applying electronic health information to health problems [54]. The measure consists of 8-items scored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores on the eHEALS indicates higher eHealth literacy (total score range=5-40). The internal consistency of the data collected using the eHEALS in this study was high (Cronbach alpha=.90), and comparable to reliability estimates reported in previous studies [54,55].

Use of Web 2.0 for Health Information

Use of social media (Web 2.0) for health information was measured using one item adapted from the HINTS survey [53]. Participants were asked, “In last 12 months, have you used the Internet for any of the following reasons to locate or share health information?” Respondents could select all reasons for using the Internet: (1) participated in a Web-based-support group, (2) used a social networking site like Facebook/Twitter/ LinkedIn, or (3) wrote in a Web-based diary or blog.

Sociodemographic and Social Determinant Variables

Sex (male or female), age (in years), race (Caucasian/white, non-Caucasian/white), ethnicity (Hispanic/non-Hispanic), education (less than high school, high school/GED, some college, college graduate, post-graduate), income (US$) (less than $20,000, $20,000-$49,999, $50,000-$99,999, $100,000 or more), and marital status (married, separated, divorced, widowed, never been married) were all assessed. Perceived health status was also measured using the following scale: (1) poor, (2) fair, (3) good, (4) very good, and (5) excellent.

Statistical Analysis

SPSS version 21.0 was used to compute frequency and descriptive statistics to summarize sociodemographic and social determinant characteristics, frequency statistics for each eHEALS item, and the number of respondents reporting use of Web 2.0 for health information. An independent samples t test was performed to compare eHealth literacy among users and non-users of Web 2.0 for health information. Given that specific technologies and tools must be considered when attempting to examine the use of Web 2.0 in health promotion [56], we also examined whether use of discrete Web 2.0 tools (ie, social networking websites, Web-based support groups, blogs) was associated with eHealth literacy. A multiple linear regression was also conducted to determine whether use of multiple electronic devices (number of digital devices used), sociodemographic variables (sex, age, income, race, ethnicity, education, marital status), and perceived health status as a social determinant predicted overall eHEALS scores. Finally, a multiple logistic regression was conducted to determine whether these predictor variables were associated with the use/non-use of Web 2.0 for health information. Use of Web 2.0 for health information was dummy coded as “0” for participants who had never used social media for seeking or sharing health information and “1” for participants who had reported use of social media for seeking or sharing health information. Analyses were considered statistically significant at the $P<.05$ alpha level (two-tailed).

Results

Participant Characteristics

Table 1 describes the characteristics of study participants reporting use of the Internet (n=283). Respondents ranged in age from 50 to 91 years (mean 67.46 years, SD 9.98). Slightly over half of the respondents were male (155/283, 54.8%), and the vast majority identified their race as Caucasian/white (252/283, 89.0%). A small proportion of respondents (16/283, 5.7%) identified as Hispanic. The majority of respondents (186/283, 65.7%) reported being married, yet 15.9% (45/283) were widowed, 12.7% (36/283) were divorced or separated, and 3.9% (11/283) reported never being married. Over 90% of participants (263/283, 92.9%) reported completing high school and over three-quarters attended college (215/283, 75.9%). The largest number of respondents fell into the $20,000 to $49,999 annual income bracket (82/283, 29.0%), followed by $50,000 to $99,999 (80/283, 28.3%), and ≥$100,000 (58/283, 20.5%). Over 50% of participants (165/283, 58.3%) reported "very good" or "excellent" health status.

Electronic Device Use for Accessing Web-Based Health Information

A little over half of the respondents accessed the Internet through a desktop computer to search for Web-based health information (143/283, 50.5%), and over 40% reported use of a laptop computer (120/283, 42.4%). More than 20% of respondents (58/283, 20.5%) reported using a mobile phone, and 14.5% (41/283) reported use of a tablet computer. Less than half of the respondents (124/283, 43.8%) reported using one electronic device to search for health information, and 30.4% (86/283) reported use of two or more devices.

Use of Social Media (Web 2.0) for Health Information

Table 2 describes use of Web 2.0 for health information among respondents. Over one-third of respondents (35.7%, 101/283) reported using Web 2.0 to locate or share health information over the past 12 months. However, almost 90% of Web 2.0 users (90/101, 89.1%) reported using only one type of social media for this purpose. Most Web 2.0 users (96/101, 95.0%) reported using popular social media sites such as Facebook and Twitter. Far fewer reported belonging to Web-based support groups (11/101, 10.9%) or contributing to Web-based health diaries/blogs (6/101, 5.9%).

eHealth Literacy

Total scores on the eHEALS ranged from 11 to 40 (mean 29.05, SD 5.75). Figure 3 illustrates the response frequencies for each eHEALS item. Over 70% of respondents agreed with the following five statements on the eHEALS: "I have the skills I need to evaluate the health resources I find on the Internet" (204/283, 72.1%); "I know how to use the health information I find on the Internet to help me" (215/283, 76.0%); "I know how to use the Internet to answer my health questions" (218/283, 77.0%); "I know how to find helpful resources on the Internet" (215/283, 76.0%); and "I know where to find helpful health resources on the Internet" (201/283, 71.0%). Two statements with the greatest level of disagreement were related to confidence using Web-based health information to make health decisions (81/283, 28.6%) and the ability to distinguish between high- and low-quality health resources on the Internet (61/283, 21.6%).

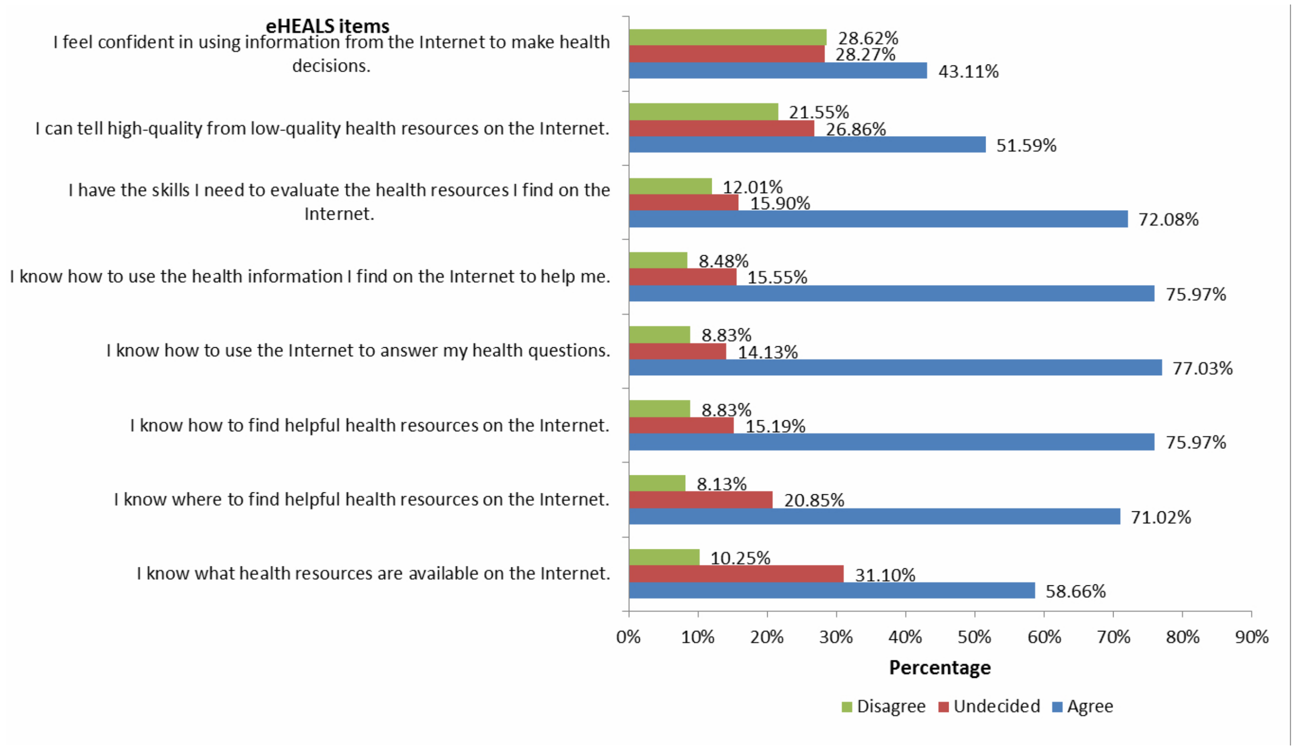
Table 1. Sociodemographic and health status characteristics of study participants (n=283).

Demographics		n (%)
Sex		
	Female	128 (45.2)
	Male	155 (54.8)
Marital status		
	Married	186 (65.7)
	Widowed	45 (15.9)
	Never married	11 (3.9)
	Divorced or separated	36 (12.7)
	No response	5 (1.8)
Ethnicity		
	Yes, Spanish or Hispanic	16 (5.7)
	No, Spanish or Hispanic	264 (93.3)
	No response	3 (1.1)
Race		
	White	252 (89.0)
	Black	10 (3.5)
	Asian or Pacific Islander	1 (0.4)
	American Indian or Alaska native	3 (1.1)
	Other	6 (2.1)
	Multi-racial or mixed race	7 (2.5)
	No response	4 (1.4)
Education		
	Less than high school graduate	19 (6.7)
	High school graduate/GED	48 (17.0)
	Some college/associates degree	82 (29.0)
	College graduate	70 (24.7)
	Postgraduate	63 (22.2)
	No response	1 (0.4)
Income (US$)		
	Less than $19,999	30 (10.6)
	$20,000 to $49,999	82 (29.0)
	$50,000 to $99,999	80 (28.3)
	More than $100,000	58 (20.5)
Health status		
	Excellent	62 (21.9)
	Very good	103 (36.4)
	Good	71 (25.1)
	Fair	30 (10.6)
	Poor	14 (4.9)
	No response	3 (1.1)

Table 2. Frequency and percentage of baby boomers and older adults who used Web 2.0 to locate or share health information (n=283).

In last 12 months, have you used the Internet for any of the following reasons to locate or share health information?		n (%)
Popular social media		
	No	187 (66.1)
	Yes	96 (33.9)
Web-based support group		
	No	272 (96.1)
	Yes	11 (3.9)
Blogs		
	No	277 (97.9)
	Yes	6 (2.1)
Report using at least one of these types of social media		
	No	182 (64.3)
	Yes	101 (35.7)

Figure 3. Frequency of responses to 8-item eHEALS (n=283).



Relationship Between Use of Social Media (Web 2.0) for Health Information and eHealth Literacy

There was a statistically significant difference in total eHEALS scores among users (mean 30.38, SD 5.45, n=101) and non-users (mean 28.31, SD 5.79, n=182) of Web 2.0 for health information, $t_{217.60}=-2.98$, $P=.003$. Respondents reporting use of Web 2.0 reported greater eHealth literacy than those who did not use Web 2.0. Users of popular social networking sites such as Facebook, Twitter, and LinkedIn for health information, had greater eHealth literacy (mean 30.22, SD 5.49, n=96) than non-users (mean 28.45, SD 5.80, n=187) for health information, $t_{201.28}=-2.20$, $P=.01$. Similar to users of popular social networking sites, respondents who reported prior use of Web-based support groups for health-related purposes reported greater eHealth literacy (mean 31.82, SD 3.06, n=11) than those reporting no such involvement (mean 28.94, SD 5.81, n=272), $t_{13.12}=-2.91$, $P=.01$. However, there was no statistically significant difference between users and non-users of Web-based diaries/blogs for health-related purposes, $t_{5.40}=-1.80$, $P=.13$.

Predictors of eHealth Literacy

Prior to conducting the multiple linear regression analysis to determine whether sociodemographic, health status, and electronic device use were associated with eHealth literacy, data were examined for multicollinearity. Both the variance inflation factors (VIF) (≤1.48) and tolerance statistics (≤0.89) met recommended cut-off points of less than 10 and greater than

0.10 respectively [57]. These results indicated that the regression model was not adversely compromised. Overall, the model accounted for 18.2% of the variance in eHEALS scores, which was statistically significant, R^2=.18, R^2_{adj} =.14, $F_{9,229}$=5.28, P<.001. Table 3 presents a summary of the regression coefficients generated by the analysis. Statistically significant predictors of eHealth literacy included age, education, and total number of electronic devices used to seek out health information. As age (b=−0.10) increased by 1 year, total eHEALS score decreased by .10 points. This indicated that, on average, the youngest baby boomers of age 50 years were likely to score approximately 1.56 points higher on the eHEALS scale than older adults who were 65 years of age. In addition, as education level (b=0.48) increased, total eHEALS scores increased by .48 points. Finally, holding all other factors in the regression model constant, the use of more electronic devices to access Web-based health information (b=1.26) was significantly associated with greater eHealth literacy. Sex, marital status, race, ethnicity, income, and health status were not significantly associated with eHealth literacy.

Table 3. Multiple linear regression predicting eHealth literacy (eHEALS).

Model[a]	*B*	*SE B*	β
Constant	26.74	3.90[b]	
Sex	1.07	0.73	.10
Age	−0.10	0.04	−.19[b]
Marital status	−0.26	0.37	−.05
Ethnicity	0.32	1.40	.01
Race	0.04	0.35	.01
Education level	0.48	0.18	.18[b]
Income	0.23	0.50	.03
Health status	0.02	0.32	.01
Total number of electronic devices used to seek health information[c]	1.26	0.31	.25[b]

[a] Model R^2=.18, R^2_{adj} =.14.

[b] P<.01, two-tailed.

[c] Participants were asked to report whether or not they used the following electronic devices to seek out health information: desktop, laptop, cell phone, or mobile tablet.

Predictors of Web 2.0 Use for Health Information

Prior to conducting the multiple logistic regression analysis to determine whether sociodemographic, health status, and electronic device use were associated with Web 2.0 use for health information, data were examined for multicollinearity. Both the variance inflation factors (VIF) (≤1.23) and tolerance statistics (≤0.79) met their respective cut-off points of less than 10 and greater than 0.10 [57] indicating that the independent variables could reasonably be entered into multivariable analyses. In the multiple logistic regression analysis, the predictor variables were able to distinguish between use and non-use of social media for health information ($\chi^2_{19,283}$=51.47, P=.001) by explaining a significant amount of variance in the model (Nagelkerke R^2 =.26). Table 4 lists the logistic regression coefficients for each predictor variable with associated 95% confidence intervals. Five of the predictor variables were significantly associated with use of Web 2.0 for health information: sex (b=0.97), possessing a baccalaureate (b=0.94) or post-graduate (b=1.96) degree, and self-reported use of one (b=1.30) or more than one (b=1.80) electronic device to find health information. Women were nearly three times more likely than men to use Web 2.0 for health information (OR 2.63, Wald= 8.09, df=1, P=.004), even after controlling for all other factors in the model. More education also predicted use of Web 2.0 for health information, with older college graduates over two times more likely than non-high school graduates to use Web 2.0 (OR 2.57, Wald= 3.86, df=1, P=.049). Respondents reporting a post graduate-level education were seven times more likely than non-high school graduates to use Web 2.0 for health information (OR 7.11, Wald=4.23, df=1, P=.04). In addition, when all other factors were held constant, respondents reporting use of one electronic device to search for health information were more than three times more likely to use Web 2.0 for health information than non-users of an electronic device for health information (OR 3.68, Wald=8.86, df=1, P=.003). Respondents reporting use of two or more electronic devices were more than six times more likely to report using Web 2.0 as compared to non-users (OR 6.06, Wald= 15.93, df=1, P=.001). Age, race, ethnicity, marital status, high school graduation, some college education, income, and health status did not significantly predict use of Web 2.0 for health information among respondents.

Table 4. Logistic regression predicting use of Web 2.0 for health information.

Sociodemographic variable		*B*	*SE B*	Exp (β)	95% CI
Constant		0.94	1.68	2.57	
Age		−0.03	0.02	0.98	0.95-1.02
Sex		0.97	0.34	2.63[b]	1.35-5.13
Ethnicity		0.03	0.66	1.03	0.29-3.75
Race		−0.28	0.17	0.77	0.56-1.08
Marital status [c]					
	Widowed	−0.25	0.49	0.78	0.30-2.05
	Never married	−0.52	0.60	0.59	0.18-1.94
	Divorced or separated	−0.22	0.84	0.80	0.16-4.13
Education [d]					
	High school graduate	0.33	0.59	1.39	0.44-4.39
	Some college	0.72	0.48	2.05	0.81-5.21
	4 years of college	0.94	0.48	2.57[a]	1.00-6.59
	Post graduate	1.96	0.95	7.11[a]	1.11-45.56
Income (US$) [e]					
	$20,000 to $49,999	0.85	0.57	2.35	0.77-7.16
	$50,000 to $99,999	−0.40	0.47	0.67	0.27-1.69
	Over $100,000	−0.05	0.42	0.96	0.42-2.17
Health status		−0.20	0.15	0.82	0.61-1.09
Use of one electronic device for health information[f]		1.30	0.44	3.68[b]	1.56-8.68
Use of multiple electronic devices for health information[e]		1.80	0.45	6.06[b]	2.50-14.69

[a] *P*<.05 two-tailed.

[b] *P*<.01

[c] Reference category: Now married

[d] Reference category: Did not graduate high school

[e] Reference category: Less than $10,000

[f] Single electronic device use defined as self-reported use of 1 electronic device (ie, desktop, laptop, cell phone, tablet) to find Web-based health information.

[e] Multiple electronic device use defined as self-reported use of ≥2 electronic devices (ie, desktop, laptop, cell phone, tablet) to find Web-based health information.

Discussion

Principal Findings

Most sociodemographic variables (eg, gender, race/ethnicity, health status) and social determinants (eg, income, employment, marital status) examined in this study were not significant predictors of eHealth literacy or use of Web 2.0 for health information among baby boomers and older adults. However, education level, advanced age, and the extent to which electronic devices were used did appear to affect eHealth literacy. Level of education, electronic device use, and being female significantly influenced the use of Web 2.0 for health-related information.

eHealth Literacy

The present study found that the majority of baby boomers and older adults used the Internet to find health information, and believed the Internet was useful for helping to make health decisions. While eHealth literacy scores decreased with age, they were comparable to scores reported in similar populations [12,31]. Overall, respondents in this study felt quite confident in their ability to use the Internet to find resources and answer questions about their health, yet they were less confident in their ability to evaluate Web-based health information. This finding is supported by Manafó and Wong [58], who reported that older adults lack confidence in their ability to discriminate between low- and high-quality health information. Research suggests that effective and user-friendly health promotion applications should be developed according to the intended audience's eHealth literacy level [25]. Unfortunately, there are few eHealth literacy interventions that exist to increase user confidence among aging populations [58]. However, it is likely that baby boomers and older adults will display more confidence using eHealth tools over time, as individuals in these populations continue to adopt more technologically advanced digital devices [18]. Future research in the aging populations should focus on investigating how improved search functionality and e-communication skills may increase self-efficacy for finding age-appropriate, trustworthy health information on the Internet.

eHealth literacy was found to be influenced by age, education, and number of electronic devices used to search for health information in this study. Previous research notes that demographics, educational background, and technology use uniquely influences health literacy [28] and eHealth literacy [25] in the general population. Although having a higher level of education has been associated with more frequent use of the Internet for health information [12,59] and greater overall eHealth literacy in some instances [60], previous research indicates that more education is not always predictive of better eHealth literacy [12,55]. However, findings from this study suggest that baby boomers and older adults with more education have higher self-reported eHealth literacy. Because of the inconsistent findings regarding the association between education level and eHealth literacy in the aging population, further research is needed to further explore these relationships.

Use of Web 2.0 for Health Information in Baby Boomers and Older Adults

Over one-third (35.7%) of respondents in this study indicated that they used some form of Web 2.0 to locate or share health information. This proportion is similar, yet slightly less, than the number reported in a 2010 Pew Research Center's Internet and American Life Project survey, which found that 42% of Internet users over the age of 50 years had used Web-based social networking tools for general purposes in the past year [42]. Facebook and Twitter are Web 2.0 platforms most commonly used among individuals younger than 50 years old [61], yet the vast majority of older Web 2.0 users in this study reported use of these popular social networking technologies.

Internet and Web 2.0 users with a history of feeling socially isolated are less likely to perceive themselves as socially isolated when compared to individuals who do not utilize Internet or Web 2.0 [61,62]. The aging population may experience greater social support as a result of using Web 2.0 to connect with more powerful support networks [42]. The versatile elements of social media (eg, games, chat, shopping, health information) may provide these populations with more socially supportive venues to learn about their own health conditions and communicate with others who may be going through similar experiences. Future research should continue to explore the specific purposes that baby boomers and older adults have for using Web 2.0 to answer their health-related questions and/or communicate with others about their chronic health conditions.

Sex, education, and use of electronic devices to seek out health information significantly predicted use of Web 2.0 for health-related purposes among this random sample of baby boomers and older adults. Women were almost three times more likely to use Web 2.0 for health information than men. However, sex-related differences in the use of Web 2.0 for health information have been inconsistent in the general adult population. Chou and colleagues [37] found sex was not associated with social media use among US adults aged 18 years and older, and Elkin [63] noted that men were actually more likely than women to use social media to research health and wellness issues in a sample of adults between the ages of 18 and 80 years. Results from another national survey of US adults 18 years and older, indicated that women use social media to find health information 22% more often than men [64]. Some researchers have suggested that women are the primary health information seekers not only for themselves but also for loved ones, which may motivate their drive to find health information on the Internet [65]. Additional research is needed to clarify the precise role that sex plays in use of Web 2.0 for seeking and sharing health information among baby boomers and older adults, including the design and evaluation of Web 2.0 applications that target gender-specific health and informal caregiving needs.

Interestingly, age was not a significant predictor of utilizing Web 2.0 for health information, although it was a significant predictor of eHealth literacy. Kontos and colleagues found younger age to be the "primary driving factor" of social networking use among US adults, with use of social networking decreasing with age [66]. Although use of Web 2.0 for health information may decline with age, findings from this study suggest that the use of Web 2.0 for health information may bridge some generational gaps that extend beyond the baby boomer generation. Norman and Skinner suggest that the "more an individual uses technology, the more likely they are to develop skills in using that technology as a tool" [54]. Some researchers speculate that the phenomenon known as "the graying of social networking sites" may provide enumerable opportunities for providing health information to baby boomers and older adults in need of resources for health promotion and disease prevention [67].

In this study, race and ethnicity were not statistically significant predictors of Web 2.0 use for health information. Large, cross-sectional surveys have noted that Caucasians/whites, African Americans, and Latinos who use the Internet are all equally likely to use social networking applications for health-related purposes [68]. Kontos et al reported greater social networking use among racial/ethnic minorities and those with lower education and income levels [66]. Also, Chou and colleagues reported that African American Internet users in the United States are actually more likely than Caucasian Internet users to use social media for health communication [37]. While race and ethnicity were not significant predictors of Web 2.0 use for health information in the current study, these racial and ethnic minority groups were grossly underrepresented in the sample. Future research should investigate the use of Web 2.0 for health information among aging populations with diverse racial and ethnic backgrounds.

Use of Web 2.0 for Health Information and eHealth Literacy

Respondents who used popular Web 2.0 websites (eg, Facebook, Twitter) and Web-based support groups for health-related purposes reported higher eHealth literacy than those who did not. While access to the Internet does not guarantee that individuals will be able to find, understand, evaluate, and act on Web-based health information [4,69], data from this study suggests that baby boomers and older adults who reported use of Web 2.0 for health-related purposes perceived themselves to have higher eHealth literacy [31]. LeRouge and colleagues report that barriers associated with utilizing technology among baby boomers and older adults are specific to the type of technology or device being used [46]. For example, baby boomers and older adults believe the mobile phone is an appropriate health-information technology, but they need more training to use it effectively for health-related purposes. While eHealth literacy has been described as a “learning system” of six discrete types of literacy that is not amenable to division [69], an updated definition of “eHealth literacy 2.0” is needed to account for the evolution of technology and the participative, social context of Web-based health information [24]. Computer (digital) and media literacies may actually be larger “petals” of eHealth literacy for baby boomer and older adult populations who need training and support to benefit from eHealth innovations. It is possible that baby boomers and older adults who learn to utilize Web 2.0 to locate and evaluate health information may gain Web-based social experiences that translate into better computer and media literacy skills. Therefore, to improve the ability of baby boomers and older adults to effectively access and utilize Web 2.0 for health care purposes, theory-based eHealth literacy interventions that apply high-quality research designs (eg, randomized controlled trials) should be evaluated in the aging population, particularly to measure effects on media and computer (digital) literacy related to health [70].

Limitations

The current study possessed several limitations. The cross-sectional research design limits the researchers from establishing causation when considering the interrelationships between sociodemographic variables, social determinants, and health communication outcomes. In addition, the use of self-reported telephone surveys may have led participants to provide socially desirable responses [52]. For example, the interviewer was unable to provide respondents with visual cues or written definitions of potentially unfamiliar technical terms such as social media. Furthermore, the types of questions asked were somewhat restricted in scope, which resulted in data that was less rich than if in-person interviews were conducted with more exploratory, open-ended questions. The absence of visual and social cues may have also resulted in the loss of contextual and nonverbal data (eg, body language), which could have compromised responses and response interpretation [71]. Use of a follow-up Web survey would allow users to view the definitions of the social media tools discussed (eg, blogs) and likely decreased the demand on individuals’ cognitive burden [72]. Unlike during in-person interviews, the interviewer cannot see the individual to gauge their understanding of an item, and therefore may not provide clarification when it is needed [73]. It would be valuable to conduct an in-depth qualitative study of older adults who access Web-based health information using Web 2.0 tools.

Another limitation of this cross-sectional study was a lack of survey items that measured frequency of Internet and Web 2.0 use for health information. Previous research noted that more frequent access to computers and the Internet was associated with higher eHealth literacy [12,31] and more positive health behavior change [74]. Among the aging population, it is possible that frequency of Internet use and type of Internet access could affect both eHealth literacy and the use of Web 2.0 for health information. In future research, it will be important to explore the perceptions of older adults who access different types of Web 2.0 with variable frequency. Baby boomers and older adults who use popular social media may consider themselves to be simply involved in informal Web-based support groups (ie, they might not consider their affiliation with social media group to be official in nature). Among aging populations, it is unclear whether frequency of interaction and engagement (like, dislikes, comments, etc) on Web 2.0 is truly an active ingredient causing greater eHealth literacy. It may be that simple membership on social media sites/pages devoted to health might improve perceived knowledge and skills related to eHealth literacy.

Additionally, the eHEALS instrument is based on an individual’s perception of personal knowledge and skills related to eHealth literacy [75] rather than demonstrated eHealth literacy competencies. While the eHEALS is a valuable instrument for assessing Web 1.0 skills, it is unclear how accurately it measures use of Web 2.0 technologies to find and evaluate health information [24]. When eHEALS was developed, social media was still in its infancy; therefore, in this study, items assessing the use of social media for health information were adapted from HINTS [53]. It should be noted that eHealth literacy is an evolving concept that requires greater inquiry [76,77], and there may be a need for a more comprehensive survey instrument that assesses health information seeking and sharing using all types of Internet applications. This type of instrument should be culturally sensitive enough to administer in diverse populations, and may focus on types of Web-based health information sought, perceived goals of Web-based health information searching, and the use of different social media tools to communicate with others about health [78]. To date, an instrument of this kind has yet to be developed and validated.

One final limitation of this study was related to the landline sampling method that was employed, which excluded over one-third of the state population that owns only a mobile phone [79]. Individuals with mobile phones are more likely to be eHealth literate with greater social media savvy [80]. The landline sampling method could have also led to selection bias, as evidenced by the lower than expected minority representation in this study. Alternative sampling methods than the ones employed in this study may be needed to reach underrepresented populations such as Hispanic adults who are more likely than non-Hispanic white adults to be living in households with only wireless telephones [81].

Conclusions

Web 2.0 has become a leading health communication platform and will continue to attract adult users of all ages; thus, it is important to continue to understand the impact of Web 2.0 on health information seeking and sharing among baby boomers and older adults. By 2015, use of the mobile Internet is projected to overtake conventional broadband Internet accessed through desktop computers [82]. As mobile apps continue to rapidly transform health care for seniors [67,83], future research should examine how mobile apps are being accessed among aging populations. The cross-sectional data from this study provides important new insights into select sociodemographics and social determinants that are associated with eHealth literacy levels and Web 2.0 use for health information in these populations. Specially, lower age and more education predicted higher eHealth literacy, and more education and being of female gender was associated with greater use of Web 2.0 for health information. Future interventions should consider providing access to tailored training opportunities based on age, education level, and gender, to improve use of advanced electronic devices to access Web-based health information. More in-depth qualitative studies with older populations are needed to better understand how and why aging populations use the Internet and Web 2.0 applications to locate and evaluate health information to make health-related decisions and solve health-related problems.

Acknowledgments

This work was supported, in part, by the National Institutes of Health (NIH) National Center for Advancing Translational Sciences (NCATS), Clinical and Translational Science Awards (CTSA) awards to the University of Florida UL1TR000064 and KL2TR000065.

Conflicts of Interest

None declared.

References

1. Internet World Stats. Internet users in the world distribution by world regions – Q4 2014. URL: http://www.internetworldstats.com/stats.htm [accessed 2014-09-19] [WebCite Cache ID 6ShuzygAZ]
2. Pew Internet & American Life Project. Health Fact Sheet. URL: http://www.pewinternet.org/fact-sheets/health-fact-sheet/ [accessed 2014-02-13] [WebCite Cache ID 6NMVzSY1A]
3. 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]
4. Knapp C, Madden V, Marcu M, Wang H, Curtis C, Sloyer P, et al. Information seeking behaviors of parents whose children have life-threatening illnesses. Pediatr Blood Cancer 2011 May;56(5):805-811. [doi: 10.1002/pbc.22674] [Medline: 21370415]
5. Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. J Med Internet Res 2010;12(5):e60 [FREE Full text] [doi: 10.2196/jmir.1463] [Medline: 21169162]
6. Bodie GD, Dutta MJ. Understanding health literacy for strategic health marketing: eHealth literacy, health disparities, and the digital divide. Health Mark Q 2008 Jan;25(1-2):175-203. [doi: 10.1080/07359680802126301] [Medline: 18935884]
7. Bonnar-Kidd KK, Black DR, Mattson M, Coster D. Online physical activity information: will typical users find quality information? Health Commun 2009 Mar;24(2):165-175. [doi: 10.1080/10410230802676763] [Medline: 19280460]
8. Connolly KK, Crosby ME. Examining e-Health literacy and the digital divide in an underserved population in Hawai'i. Hawaii J Med Public Health 2014 Feb;73(2):44-48 [FREE Full text] [Medline: 24567867]
9. Hargittai E. Digital Na(t)ives? Variation in internet skills and uses among members of the “Net Generation”. Sociological Inquiry 2010 Jan;80(1):92-113. [doi: 10.1111/j.1475-682X.2009.00317.x]
10. Lorence D, Park H. Group disparities and health information: a study of online access for the underserved. Health Informatics J 2008 Mar;14(1):29-38. [doi: 10.1177/1460458207086332] [Medline: 18258673]
11. Comerci GD. Eating disorders in adolescents. Pediatr Rev 1988 Aug;10(2):37-47. [Medline: 3222184]
12. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
13. Zajac IT, Flight IHK, Wilson C, Turnball D, Cole S, Young G. Internet usage and openness to internet-delivered health information among Australian adults aged over 50 years. Australas Med J 2012 May 31;5(5):262-267. [doi: 10.4066/AMJ.2012.1065]
14. US Food and Drug Administration. How to evaluate health information on the internet. 2013. URL: http://www.fda.gov/drugs/resourcesforyou/consumers/buyingusingmedicinesafely/buyingmedicinesovertheinternet/ucm202863.htm [accessed 2014-09-18] [WebCite Cache ID 6Sgs5QZQ8]
15. Centers for Disease Control and Prevention. Healthy aging. 2011. URL: http://www.cdc.gov/chronicdisease/resources/publications/aag/aging.htm [accessed 2011-07-07] [WebCite Cache ID 5zzGL8QpS]
16. World Health Organization. Health literacy: the solid facts. 2013. URL: http://www.euro.who.int/__data/assets/pdf_file/0008/190655/e96854.pdf [accessed 2014-10-27] [WebCite Cache ID 6Tdo2QPb8]

17. Sudore RL, Mehta KM, Simonsick EM, Harris TB, Newman AB, Satterfield S, et al. Limited literacy in older people and disparities in health and healthcare access. J Am Geriatr Soc 2006 May;54(5):770-776. [doi: 10.1111/j.1532-5415.2006.00691.x] [Medline: 16696742]
18. Smith A. Older adults and technology use.: Pew Internet & American Life Project; 2014. URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2014-04-23] [WebCite Cache ID 6P34DTel4]
19. Internet user demographics.: Pew Internet & American Life Project; 2014. URL: http://www.pewinternet.org/data-trend/internet-use/latest-stats/ [accessed 2015-02-02] [WebCite Cache ID 6W2qz0j1B]
20. Mobile technology fact sheet.: Pew Internet & American Life Project URL: http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/ [accessed 2015-02-02] [WebCite Cache ID 6W2r7LrF1]
21. Wagner N, Hassanein K, Head M. Computer use by older adults: A multi-disciplinary review. Computers in Human Behavior 2010 Sep;26(5):870-882. [doi: 10.1016/j.chb.2010.03.029]
22. Medlock S, Eslami S, Askari M, Arts DL, Sent D, de Rooij SE, et al. Health information-seeking behavior of seniors who use the Internet: a survey. J Med Internet Res 2015;17(1):e10 [FREE Full text] [doi: 10.2196/jmir.3749] [Medline: 25574815]
23. White RE. Health information technology will shift the medical care paradigm. J Gen Intern Med 2008 Apr;23(4):495-499 [FREE Full text] [doi: 10.1007/s11606-007-0394-y] [Medline: 18373152]
24. Norman C. eHealth literacy 2.0: problems and opportunities with an evolving concept. J Med Internet Res 2011 Dec;13(4):e125 [FREE Full text] [doi: 10.2196/jmir.2035] [Medline: 22193243]
25. Norman CD, Skinner HA. eHealth Literacy: Essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
26. Xie B. Effects of an eHealth literacy intervention for older adults. J Med Internet Res 2011 Nov;13(4):e90 [FREE Full text] [doi: 10.2196/jmir.1880] [Medline: 22052161]
27. Weinstein RS, Lopez AM. Health literacy and connected health. Health Aff (Millwood) 2014 Jun;33(6):1103-1104. [doi: 10.1377/hlthaff.2014.0396] [Medline: 24889963]
28. Institute of Medicine. Health literacy: A prescription to end confusion. 2004. URL: http://www.iom.edu/~/media/Files/Report%20Files/2004/Health-Literacy-A-Prescription-to-End-Confusion/healthliteracyfinal.pdf [accessed 2014-09-18] [WebCite Cache ID 6Sgrwus9y]
29. Zamora H, Clingerman EM. Health literacy among older adults: a systematic literature review. J Gerontol Nurs 2011 Oct;37(10):41-51. [doi: 10.3928/00989134-20110503-02] [Medline: 21634314]
30. Xavier AJ, d'Orsi E, de Oliveira CM, Orrell M, Demakakos P, Biddulph JP, et al. English Longitudinal Study of Aging: can Internet/E-mail use reduce cognitive decline? J Gerontol A Biol Sci Med Sci 2014 Sep;69(9):1117-1121. [doi: 10.1093/gerona/glu105] [Medline: 25116923]
31. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
32. Xie B. Lifelong Interactions: Older adults, health information, and the Internet. ACM Interactions 2008;15(4):44-46.
33. Aghaei S. Evolution of the world wide web: from Web 1.0 to Web 4.0. IJWesT 2012 Jan 31;3(1):1-10. [doi: 10.5121/ijwest.2012.3101]
34. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
35. Gibbons MC, Fleisher L, Slamon RE, Bass S, Kandadai V, Beck JR. Exploring the potential of Web 2.0 to address health disparities. J Health Commun 2011;16 Suppl 1:77-89. [doi: 10.1080/10810730.2011.596916] [Medline: 21843097]
36. Fox S, Jones S. The social life of health information.: Pew Internet & American Life Project; 2009. URL: http://www.pewinternet.org/2009/06/11/the-social-life-of-health-information/ [accessed 2014-06-06] [WebCite Cache ID 6Q8fer6YP]
37. Chou WYS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009 Nov;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
38. Hardt JH, Hollis-Sawyer L. Older adults seeking healthcare information on the Internet. Educational Gerontology 2007 Jun 13;33(7):561-572. [doi: 10.1080/03601270701364628]
39. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014;16(12):e262 [FREE Full text] [doi: 10.2196/jmir.3706] [Medline: 25470306]
40. Cutrona SL, Roblin DW, Wagner JL, Gaglio B, Williams AE, Torres SR, et al. Adult willingness to use email and social media for peer-to-peer cancer screening communication: Quantitative interview study. JMIR Res Protoc 2013;2(2):e52 [FREE Full text] [doi: 10.2196/resprot.2886] [Medline: 24287495]
41. Capel S, Childs S, Banwell L, Heaford S. Access to information and support for health: some potential issues and solutions for an ageing population. Health Informatics J 2007 Dec;13(4):243-253. [doi: 10.1177/1460458207079824] [Medline: 18029402]
42. Madden M. Older adults and social media.: Pew Internet & American Life Project; 2010. URL: http://www.pewinternet.org/2010/08/27/older-adults-and-social-media/ [accessed 2014-09-18] [WebCite Cache ID 6Sgt3vD50]

43. Stellefson M, Chaney B, Barry AE, Chavarria E, Tennant B, Walsh-Childers K, et al. Web 2.0 chronic disease self-management for older adults: a systematic review. J Med Internet Res 2013 Feb;15(2):e35 [FREE Full text] [doi: 10.2196/jmir.2439] [Medline: 23410671]
44. Thackeray R, Crookston BT, West H. Correlates of health-related social media use among adults. J Med Internet Res 2013;15(1):e21 [FREE Full text] [doi: 10.2196/jmir.2297] [Medline: 23367505]
45. Gatto SL, Tak SH. Computer, internet, and e-mail use among older adults: benefits and barriers. Educational Gerontology 2008 Aug 22;34(9):800-811. [doi: 10.1080/03601270802243697]
46. LeRouge C, Van SC, Seale D, Wright K. Baby boomers' adoption of consumer health technologies: survey on readiness and barriers. J Med Internet Res 2014;16(9):e200 [FREE Full text] [doi: 10.2196/jmir.3049] [Medline: 25199475]
47. Ackerson LK, Viswanath K. The social context of interpersonal communication and health. J Health Commun 2009;14 Suppl 1:5-17. [doi: 10.1080/10810730902806836] [Medline: 19449264]
48. Viswanath K, Ramanadhan SR, Kontos EZ. Mass media. In: Galea S, editor. Macrosocial Determinants of Population Health. New York: Springer; 2007:275-294.
49. Colby SL, Ortman JM. The baby boom cohort in the United States: 2012-2060. 2014. URL: http://www.census.gov/prod/2014pubs/p25-1141.pdf [accessed 2015-02-04] [WebCite Cache ID 6W694xxDJ]
50. Coughlin J, D'Ambrosio LA, Reimer B, Pratt MR. Older adult perceptions of smart home technologies: implications for research, policy & market innovations in healthcare. In: Conference Proceedings IEEE Engineering in Medicine and Biology Society. 2007 Presented at: 29th Annual International Conference of the IEEE EMBS; 2007; Lyon, France p. 1810-1815.
51. Bureau of Economic and Business Research. URL: http://www.bebr.ufl.edu/ [accessed 2014-09-18] [WebCite Cache ID 6SgrzZqxu]
52. Dillman DA. Mail and internet surveys: the tailored design method. Hoboken, NJ: Wiley; 2007.
53. Health Information National Trends Survey. HINTS Questions: Internet use. 2012. URL: http://hints.cancer.gov/topic.aspx?section=Internet+Use [accessed 2012-12-02] [WebCite Cache ID 6Cc2vqEg8]
54. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
55. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
56. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
57. Pallant J. SPSS Survival Manual: A step by step guide to data analysis using SPSS. New York: Open University Press; 2010.
58. Manafò E, Wong S. Assessing the eHealth literacy skills of older adults: A preliminary study. Journal of Consumer Health On the Internet 2012 Oct;16(4):369-381. [doi: 10.1080/15398285.2012.701163]
59. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
60. van der Vaart R, Drossaert CHC, de Heus M, Taal E, van de Laar MAFJ. Measuring actual eHealth literacy among patients with rheumatic diseases: a qualitative analysis of problems encountered using Health 1.0 and Health 2.0 applications. J Med Internet Res 2013 Feb;15(2):e27 [FREE Full text] [doi: 10.2196/jmir.2428] [Medline: 23399720]
61. O'Brien C. Participation in online communities and psychosocial well-being among older adults. 2012. URL: http://www.matherlifewaysinstituteonaging.com/wp-content/uploads/2012/03/MLWOrangePaper_InternetUse_2.pdf [accessed 2014-09-18] [WebCite Cache ID 6SgsDRMOf]
62. Social networking fact sheet.: Pew Internet & American Life Project; 2014. URL: http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/ [accessed 2015-03-02] [WebCite Cache ID 6W6JApNDP]
63. Elkin N. How America searches: health and wellness. 2008. URL: http://www.icrossing.com/sites/default/files/how-america-searches-health-and-wellness.pdf [accessed 2012-02-05] [WebCite Cache ID 65DJnxuVZ]
64. Duggan M, Brenner J. The demographics of social media users – 2013.: Pew Internet & American Life Project; 2012. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2013/PIP_SocialMediaUsers.pdf [accessed 2014-09-18] [WebCite Cache ID 6SgtPtlLF]
65. Baur C. An analysis of factors underlying e-health disparities. Camb Q Healthc Ethics 2008;17(4):417-428. [doi: 10.1017/S0963180108080547] [Medline: 18724881]
66. Kontos EZ, Emmons KM, Puleo E, Viswanath K. Communication inequalities and public health implications of adult social networking site use in the United States. J Health Commun 2010;15 Suppl 3:216-235 [FREE Full text] [doi: 10.1080/10810730.2010.522689] [Medline: 21154095]
67. Boulos MNK. Using social media for improving health literacy. 2012. URL: http://ecaalyx.org/tmp/KamelBoulosChapter_social-media-for-health-literacy_WHO_THE-SOLID-FACTS_2012.pdf [accessed 2012-08-08] [WebCite Cache ID 69l4M2uVC]

68. Jones S, Fox S. Generations online in 2009.: Pew Internet & American Life Project; 2009. URL: http://www.pewinternet.org/2009/01/28/generations-online-in-2009/ [accessed 2014-09-16] [WebCite Cache ID 6ScRAUY2y]
69. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
70. Watkins I, Xie B. eHealth literacy interventions for older adults: a systematic review of the literature. J Med Internet Res 2014;16(11):e225 [FREE Full text] [doi: 10.2196/jmir.3318] [Medline: 25386719]
71. Novick G. Is there a bias against telephone interviews in qualitative research? Res Nurs Health 2008 Aug;31(4):391-398 [FREE Full text] [doi: 10.1002/nur.20259] [Medline: 18203128]
72. Dillman DA, Smyth JD, Christian LM. Internet, phone, mail, and mixed-mode surveys: the tailored design method, 4th edition. Hoboken, NJ: Wiley; 2014.
73. Check J, Schutt R. Research Methods in Education. Thousand Oaks, CA: Sage Publications, Inc; Oct 27, 2011.
74. Ayers SL, Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. Health (London) 2007 Jul;11(3):327-347. [doi: 10.1177/1363459307077547] [Medline: 17606698]
75. van Deursen AJAM, van Dijk JAGM. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011 Apr;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
76. Chan CV, Kaufman DR. A framework for characterizing eHealth literacy demands and barriers. J Med Internet Res 2011 Nov;13(4):e94 [FREE Full text] [doi: 10.2196/jmir.1750] [Medline: 22094891]
77. van DA, van DJ, Peters O. Rethinking internet skills: The contribution of gender, age, education, internet experience, and hours online to medium- and content-related internet skills. Poetics 2011 Apr;39(2):125-144. [doi: 10.1016/j.poetic]
78. Miller LM, Bell RA. Online health information seeking: the influence of age, information trustworthiness, and search challenges. J Aging Health 2012 Apr;24(3):525-541. [doi: 10.1177/0898264311428167] [Medline: 22187092]
79. Blumberg SJ, Luke JV, Ganesh N, Davern ME, Bouderaux MH. Natl Health Stat Report. 2012 Oct. Wireless substitution: state-level estimates from the national health interview survey, 2010-2011 URL: http://www.cdc.gov/nchs/data/nhsr/nhsr061.pdf [accessed 2015-03-09] [WebCite Cache ID 6Wu2WQBN5]
80. Drury G. Opinion piece: Social media: Should marketers engage and how can it be done effectively? J Direct Data Digit Mark Pract 2008 Mar;9(3):274-277. [doi: 10.1057/palgrave.dddmp.4350096]
81. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
82. Charlton G. Will mobile internet replace desktop?. 2012. URL: https://econsultancy.com/blog/11186-will-mobile-internet-replace-desktop-infographic [accessed 2014-09-18] [WebCite Cache ID 6SgtkEKtd]
83. Madden M, Zickuhr K. 65% of online adults use social networking sites.: Pew Internet & American Life Project; 2011. URL: http://www.pewinternet.org/2011/08/26/65-of-online-adults-use-social-networking-sites/ [accessed 2014-09-18] [WebCite Cache ID 6SgskO7Vb]

Abbreviations

BEBR: University of Florida Bureau of Economic and Business Research
eHEALS: eHealth Literacy Scale
F-CCI: Florida Consumer Confidence Index
HINTS: Health Information National Trends Survey
OR: odds ratio
RDD: random digit dialing
SIMHC: Structural Influence Model of Health Communication
SPSS: Statistical Package for the Social Sciences
VIF: variance inflation factors

Edited by G Eysenbach; submitted 29.10.14; peer-reviewed by M Silver, S Childs; comments to author 18.11.14; revised version received 06.02.15; accepted 23.02.15; published 17.03.15.

Please cite as:
Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, Alber J
eHealth Literacy and Web 2.0 Health Information Seeking Behaviors Among Baby Boomers and Older Adults
J Med Internet Res 2015;17(3):e70
URL: http://www.jmir.org/2015/3/e70/
doi:10.2196/jmir.3992
PMID:25783036

Original Paper

Cost Analysis of In-Home Telerehabilitation for Post-Knee Arthroplasty

Michel Tousignant[1,2], PT, PhD; Hélène Moffet[3,4], PhD; Sylvie Nadeau[5,6], PhD; Chantal Mérette[3,7], PhD; Patrick Boissy[1,2], PhD; Hélène Corriveau[1,2], PT, PhD; François Marquis[3,8], MD; François Cabana[2,9], MD; Pierre Ranger[5,6,10], MD; Étienne L Belzile[3,8], MD; Ronald Dimentberg[11,12], MD

[1]Research Centre on Aging, University Institute of Geriatrics of Sherbrooke, Sherbrooke, QC, Canada
[2]Université de Sherbrooke, Sherbrooke, QC, Canada
[3]Université Laval, Quebec, QC, Canada
[4]Centre for Interdisciplinary Research in Rehabilitation and Social Integration, Quebec, QC, Canada
[5]Université de Montréal, Montreal, QC, Canada
[6]Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal, Montreal, QC, Canada
[7]Centre de recherche de l'Institut universitaire en santé mentale de Québec, Quebec, QC, Canada
[8]Centre hospitalier universitaire de Québec, Quebec, QC, Canada
[9]Centre hospitalier universitaire de Sherbrooke, Sherbrooke, QC, Canada
[10]Hôpital Jean-Talon, Montreal, QC, Canada
[11]McGill University, Montreal, QC, Canada
[12]St Mary's Hospital, Montreal, QC, Canada

Corresponding Author:
Michel Tousignant, PT, PhD
Research Centre on Aging, University Institute of Geriatrics of Sherbrooke
1036 Belvédère Sud
Sherbrooke, QC, J1H4C4
Canada
Phone: 1 819 780 2220 ext 45351
Fax: 1 819 829 7141
Email: michel.tousignant@usherbrooke.ca

Abstract

Background: Rehabilitation provided through home visits is part of the continuum of care after discharge from hospital following total knee arthroplasty (TKA). As demands for rehabilitation at home are growing and becoming more difficult to meet, in-home telerehabilitation has been proposed as an alternate service delivery method. However, there is a need for robust data concerning both the effectiveness and the cost of dispensing in-home telerehabilitation.

Objective: The objective of this study was to document, analyze, and compare real costs of two service delivery methods: in-home telerehabilitation and conventional home visits.

Methods: The economic analysis was conducted as part of a multicenter randomized controlled trial (RCT) on telerehabilitation for TKA, and involved data from 197 patients, post-TKA. Twice a week for 8 weeks, participants received supervised physiotherapy via two delivery methods, depending on their study group allocation: in-home telerehabilitation (TELE) and home-visit rehabilitation (VISIT). Patients were recruited from eight hospitals in the province of Quebec, Canada. The TELE group intervention was delivered by videoconferencing over high-speed Internet. The VISIT group received the same intervention at home. Costs related to the delivery of the two services (TELE and VISIT) were calculated. Student's *t* tests were used to compare costs per treatment between the two groups. To take distance into account, the two treatment groups were compared within distance strata using two-way analyses of variance (ANOVAs).

Results: The mean cost of a single session was Can $93.08 for the VISIT group (SD $35.70) and $80.99 for the TELE group (SD $26.60). When comparing both groups, real total cost analysis showed a cost differential in favor of the TELE group (TELE minus VISIT: -$263, 95% CI -$382 to -$143). However, when the patient's home was located less than 30 km round-trip from the health care center, the difference in costs between TELE and VISIT treatments was not significant (*P*=.25, .26, and .11 for the <10, 10-19, and 20-29 km strata, respectively). The cost of TELE treatments was lower than VISIT treatments when the distance was 30 km or more (30-49 km: $81<$103, *P*=.002; ≥50 km: $90<$152, *P*<.001).

Conclusions: To our knowledge, this is the first study of the actual costs of in-home telerehabilitation covering all subcosts of telerehabilitation and distance between the health care center and the patient’s home. The cost for a single session of in-home telerehabilitation compared to conventional home-visit rehabilitation was lower or about the same, depending on the distance between the patient’s home and health care center. Under the controlled conditions of an RCT, a favorable cost differential was observed when the patient was more than 30 km from the provider. Stakeholders and program planners can use these data to guide decisions regarding introducing telerehabilitation as a new service in their clinic.

Trial Registration: International Standard Registered Clinical Study Number (ISRCTN): 66285945; http://www.isrctn.com/ISRCTN66285945 (Archived by WebCite at http://www.webcitation.org/6WlT2nuX4).

(J Med Internet Res 2015;17(3):e83) doi:10.2196/jmir.3844

KEYWORDS

telemedicine; economics; cost analysis; knee arthroplasty

Introduction

The rapid growth in the number of total knee arthroplasty (TKA) surgeries per year and shorter hospital stays have greatly increased demand for outpatient rehabilitation services for patients returning home more quickly after hospital discharge. However, this increasing need has not been followed by an increase in accessibility to rehabilitation services after TKA, in-home care, or outpatient clinics. As part of a worldwide situation [1], in the province of Quebec, Canada, it has been shown that actual rehabilitation services cannot meet needs, both because the services are not accessible everywhere and because of a lack of health care professionals [2-4]. To increase accessibility to rehabilitation, alternative approaches to delivering these services must be considered [5].

Teletreatment, which is defined as care and treatment delivered remotely between patients and health care professionals with the support of communication and information technology [6], is a recent approach that can be relevant to rehabilitation patients. As opposed to teleconsultation, teletreatment has an essential characteristic, which is the repeated delivery of synchronous services through videoconferencing to provide real-time interactions between patients and clinicians. Another characteristic is that teletreatment may be delivered between two health care centers or between a health care center and the patient’s home under direct supervision (ie, in-home teletreatment). In this paper, we focus on in-home teletreatment.

Our main interest is to increase access to, and efficiency of, rehabilitation services after TKA through the use of teletreatment between the health care center and patient’s home. In this context, the feasibility of providing in-home telerehabilitation following joint replacement surgery has been studied previously [7-8]. Those pilot projects showed the feasibility of telerehabilitation for TKA. However, the small sample size or lack of a control group limited the generalizability of those results. No real cost analysis was included in those studies. Only a theoretical estimate of the costs associated with teletreatment compared to home visits suggested that teletreatment had the potential to be less costly than home visits [7]. However, that theoretical estimate cannot be generalized and a rigorous cost analysis is needed to provide robust data on the cost of teletreatment. In a recent literature review [9], 35 articles were found, where the majority were about cost-effectiveness of telemedicine in general. Some, but not all, studies demonstrate that telemedicine can reduce the costs. Up to now, no cost analysis on teletreatment between the health care center and the patient’s home using an audiovideo synchronous system has been done.

This cost analysis was embedded in the telerehabilitation for knee arthroplasty (TelAge) randomized controlled trial (RCT) on the effectiveness of in-home telerehabilitation. The TelAge project is a multicenter randomized trial, involving three research centers located in Quebec City, Sherbrooke, and Montreal, and well as eight hospitals [10]. Thus, the purpose of this paper was to compare costs related to in-home teletreatment versus conventional home-visit rehabilitation following TKA.

Methods

TelAge Randomized Controlled Trial Design

The project was approved by ethics committees of each hospital and the participants’ written consent was obtained prior to their inclusion. The RCT was registered with the International Standard Registered Clinical Study Number (ISRCTN) registry (ISRCTN66285945).

Participants were recruited from the surgical waiting lists of orthopedic surgeons in each hospital. To be included in the study, they had to meet the following inclusion criteria: (1) were waiting for a primary TKA after a diagnosis of osteoarthritis, (2) returning back home after hospital discharge, (3) living in an area served by high-speed Internet services (at least 512 kbps in upload), and (4) living within a 1-hour driving distance from the treating hospital. Patients were excluded if they met the following criteria: (1) had health conditions that could interfere with tests or the rehabilitation program, including lower limb surgery in the last 9 months, (2) were planning a second lower limb surgery within 4 months, (3) had cognitive or collaboration problems, (4) had major postoperative complications, or (5) had weight-bearing restrictions for a period longer than 2 weeks after surgery.

After their surgery, patients were block randomized by random generator with sealed envelopes into either the in-home telerehabilitation (TELE) experimental group or the home-visit (VISIT) control group. For both groups, the goal of the physiotherapy program was to recover function—only the delivery method differed. The duration and frequency of the supervised physiotherapy sessions were standardized for both groups and consisted of two 45-minute sessions per week for 8 weeks. Patients were assessed four times by a blind assessor—twice before and twice after the intervention—to evaluate rehabilitation efficacy, along the following timeline: (1) at baseline (ie, 1 month or less before TKA) (E1), (2) 1 to 2 days before discharge from hospital (E2), (3) 2 months postdischarge (E3), and (4) 4 months postdischarge (E4). Figure 1 shows the study design, including the cost evaluation period. Results obtained for efficacy demonstrated the noninferiority of in-home telerehabilitation and support its use as an effective alternative to face-to-face service delivery for TKA patients after hospital discharge [11].

The telerehabilitation platform included various components to provide a user-friendly videoconferencing experience for both the clinician and the patient at home. The core was the videoconferencing system (Tandberg 550 MXP), which used an H.264 video codec and incorporated a pan-tilt-zoom (PTZ) camera with wide-angle lens and omnidirectional microphone. On the patient side, the system was mounted over a 20-inch LCD screen, which displayed the video received from the other end. Audio was played using external speakers placed on both sides of the screen, as the internal LCD screen speakers are rarely sufficient to provide a satisfactory experience. Video and audio data were encrypted and transmitted over a high-speed Internet connection—at the time of the TelAge RCT, minimum bandwidth for both upload and download was 512 kbps. On the clinician side, a computer with dedicated software (TeRA) was added to the videoconference link to enable user-friendly control and monitoring of videoconferencing sessions, near- and far-end camera controls, built-in clinical tests, photo and video recordings, and acquisition and display of external sensors and peripherals (Figure 2).

Figure 1. TelAge clinical randomized controlled trial study design including cost evaluation period.

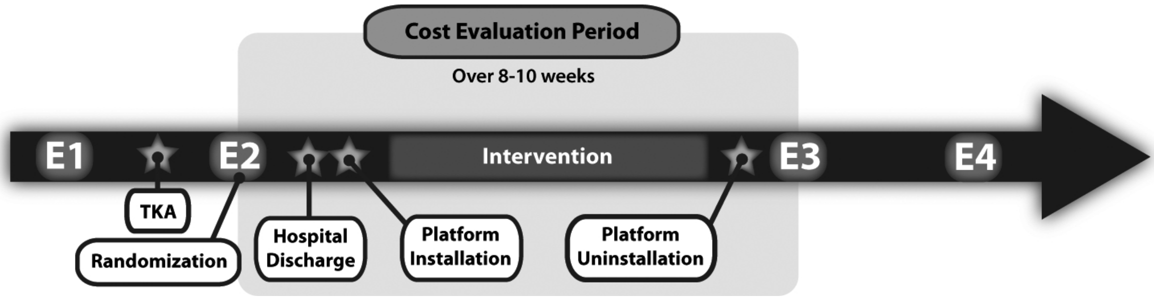
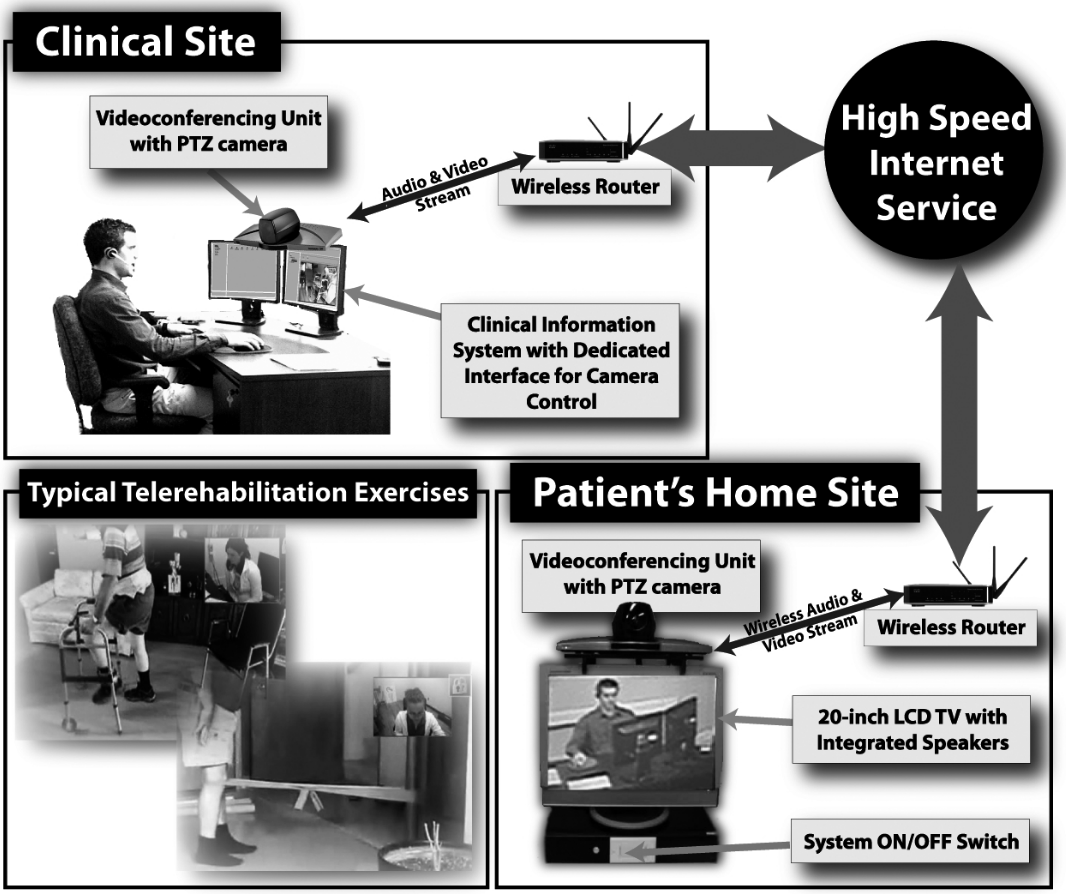


Figure 2. Telerehabilitation platform.



Cost Measurement

Costs were measured during the intervention period—between E2 and E3—for both groups. All cost-related data were directly collected by the following: (1) physiotherapists after each session (ie, intervention, travel, technical problems), and (2) technicians after each installation/uninstallation (ie, travel, technical problems). Cost-related data were collected using a standardized cost sheet during the period from hospital discharge to 2 months postdischarge (ie, the end of rehabilitation intervention). Other costs, such as Internet service and technical/clinical equipment amortization, were calculated using the real cost of the service and purchase price of the equipment, respectively, at the time of the TelAge RCT.

This economic evaluation followed international guidelines for conducting a cost analysis alongside a clinical RCT [12]. The economic analysis was based on a health sector perspective, meaning that only costs relevant to the health center were considered, and not costs relevant to the patient. In addition, only costs related to the delivery of the two services—TELE and VISIT—were taken into account [12]. Costs were divided into two categories: (1) costs related to the clinical aspects, and (2) costs related to the technology. For each of these cost categories, direct costs were defined as essential for the delivery of the clinical intervention, and indirect costs as being related to the intervention without being part of it.

More precisely, direct clinical time included duration of the interaction between physiotherapist and participant to deliver the intervention in both groups. Indirect clinical time was defined as the following: (1) time without direct contact with the participant, but which was required to optimize the intervention (ie, scheduling, file preparation, planning the treatment, writing the report and follow-up notes, phone follow-up with orthopedic surgeon), (2) travel time and kilometer (km) allocation for use of the physiotherapist's car—only for the VISIT group, and (3) clinical equipment amortization, which is the theoretical time duration of use of the technology. Round-trip distance from the health care center (ie, hospital where the research center involved in the project was located) to the patient's home was calculated for each patient in both groups.

Direct technology costs—only for the TELE group—are the costs in Canadian dollars for acquiring clinician and patient telerehabilitation platform kits as illustrated in Figure 2. The cost of the equipment was amortized assuming 3 years of use at the rate of 17 hours per week (ie, about 50% of a 35-hour week). Indirect technology costs included the following: (1) time spent on transportation and installation and uninstallation of the platform at home (ie, technician's salary and kilometer allocation), (2) the cost to provide a new connection or upgrade high-speed Internet from the Internet service providers for the duration of the intervention (ie, connection fee plus 2 months of Internet service), and (3) technician's time and travel to deal with technical problems before/during/after the physiotherapists' sessions.

The standard hourly salary in effect in the public health system in Quebec at the time of the TelAge RCT was applied to the time spent by the physiotherapists and technicians—Can $45.22/hour and Can $22.09/hour, respectively. Travel expenses were calculated as distance in kilometers multiplied by Can $0.40. In addition, to determine the cost of the whole treatment (ie, 16 sessions), decision makers may be interested in the net costs of giving one TELE or one VISIT treatment, for instance, the cost per treatment. For that analysis, only costs of treatments actually received were taken into account. Moreover, the distribution of costs in each delivery-service method was analyzed. Table 1 shows the breakdown of clinical costs for both study groups.

Table 1. Breakdown of costs per treatment for telerehabilitation and home visits.

Cost category	Description	Duration	Professional cost[a] calculation	
			TELE	VISIT
Clinical				
Direct time				
	Contact with the participant	Hours	Hours x $45.22[b]	Hours x $45.22
Indirect time				
	Planning the session	Hours	Hours x $45.22	Hours x $45.22
	Follow-up with orthopedic surgeon	Hours	Hours x $45.22	Hours x $45.22
	Writing the report and follow-up notes	Hours	Hours x $45.22	Hours x $45.22
	Travel	Hours, km	N/A[c]	Hours x $45.22, km x $0.40
	Clinical equipment amortization	16 treatments	$0.63 per treatment ($10/16 treatments)	$0.63 per treatment ($10/16 treatments)
Technology				
Installation/uninstallation of technology			Travel time (hours) x $22.16[d], km x $0.40	N/A
Internet costs (high speed)[e]			$7.63 per intervention	N/A
Technical equipment amortization[f]			$4.88 per intervention	N/A
Technical problems				
Technician			Hours x $22.16	N/A
Physiotherapist			Hours x $45.22	N/A

[a] All costs are in Canadian dollars.

[b] Physiotherapist mean hourly salary in the public system in Quebec.

[c] Not applicable (N/A).

[d] The hourly rate at the higher echelon of the salary scale of a technician.

[e] Cost for 2 months (16 sessions): $122.08/16 sessions = $7.63 per session.

[f] Cost for 1 clinician kit ($5760) and 1 patient kit ($7200) amortized over 3 years based on 50% of usage per week (17 hours): $5760 + $7200 = $12,960. Hours telerehabilitation: 3 years x 52 weeks/year x 17 hours (50% usage)/week = 2652 hours. Cost per treatment: $12,960/2652 hours = $4.88.

Sensitivity Analysis

As time spent on travel is an important cost factor for the VISIT group, the distance between the health care center and patient's home is an important variable that must be taken into account when comparing costs between VISIT and TELE treatments. To consider the impact of this variable, a sensitivity analysis was performed on the basis of the distance between the patient's home and health care center for the physiotherapist delivering the treatment (VISIT) or the technician installing/uninstalling the technology (TELE). Another way to look at this sensitivity analysis is to consider that time spent on travel may differ in metropolitan, urban, and rural areas, for instance, travelling 5 km in a metropolitan area can take more time than in urban and rural areas. Knowing that, the analyses were done separately for each site.

Statistical Analyses

All analyses were performed using SAS/STAT software, version 9.3 of the SAS System for Windows (SAS Institute Inc, Cary, NC, USA). First, the two groups of patients were compared using sociodemographic characteristics. For the continuous variables, Student's *t* tests were conducted using the TTEST procedure. For the categorical variables, chi-square tests were computed using the FREQ procedure. Then the treatment costs were compared between the TELE and VISIT groups by Student's *t* tests using the TTEST procedure. Finally, considering the fact that travel time is an important cost factor for the VISIT group, the round-trip distance from the health care center to the patient's home was calculated for each patient in both groups. Patients were then grouped into five distance strata: <10, 10-19, 20-29, 30-49, ≥50 km. In addition, analyses were conducted per site (ie, Montreal, Quebec City, and Sherbrooke) to take into account metropolitan, urban, and rural areas. Two-way analyses of variance (ANOVAs) with interaction were conducted to compare costs per treatment between the two groups and between distance strata in the total sample and for each site. When the interaction term was significant, post hoc comparisons were examined to compare the two groups within each distance stratum. When appropriate, distance strata were grouped to simplify interpretation of the results. SAS's general linear model (GLM) procedure was used with SLICE option in the least squares means (LSMEANS) statement to obtain the post hoc comparisons.

Results

Participants

A total of 258 patients needing TKA were recruited for the TelAge RCT [10]. Postsurgery, 53 (20.5%) participants were excluded and the 205 (79.5%) who remained were block randomized to the TELE group (104/205, 50.7%) or VISIT group (101/205, 49.3%). Of the 104 TELE group participants, 6 (5.8%) were excluded before the first follow-up assessment (E3) because they were unsatisfied with randomization (3), had poor Internet connection (2), or had self-perception of complete recovery (1). In the TELE group, only 1 of the participants was excluded following the randomization because no physical therapist was available. Moreover, 1 participant in the TELE group was excluded from the analysis because he did not attend any of the rehabilitation sessions. The cost analysis was thus performed with 97 and 100 participants in the TELE and VISIT groups, respectively. The sociodemographic characteristics of the final sample used in the cost analysis and the participants' locations expressed as distance to the treatment center are presented in Table 2.

The characteristics of the participants were comparable in each group for every sociodemographic variable measured, except for participants living alone (*P*=.04), where about twice as many participants in the TELE group lived alone as in the VISIT group. Also, participants in the TELE group lived farther away from the health care center than those in the VISIT group—mean round-trip distance for the TELE group was 59 km (SD 67) and for the VISIT group was 34 km (SD 35) (*P*=.002).

Table 2. Descriptive statistics at baseline of the sample included in the cost analysis (n=197).

Sociodemographic characteristic	TELE (n=97), mean (SD) or n (%)	VISIT (n=100), mean (SD) or n (%)	*P* value[a]
Age (years), mean (SD)	65 (8)	67 (8)	.11
Body mass index (kg/m^2), mean (SD)	34.6 (7.6)	33.1 (5.6)	.12
Comorbidity Index (%), mean (SD)	23 (12)	21 (9)	.16
Functional ability before TKA (WOMAC[b] in %), mean (SD)	53 (19)	54 (17)	.74
Round-trip distance from health care center to patient's home (km), mean (SD)	59 (67)	34 (35)	.002
Sex (males), n (%)	40 (41)	55 (55.0)	.05
Operated knee (right), n (%)	45 (46)	51 (51.0)	.52
Previous lower limb surgery, n (%)	44 (45)	49 (49.0)	.66
Living alone, n (%)	20 (21)	10 (10.0)	.04

[a] *P* values are from Student's *t* tests for continuous variables and from chi-square tests for categorical variables.

[b] Western Ontario and McMaster Universities Arthritis Index (WOMAC).

Costs Related to Treatment Delivery (Treatments Received and Cancelled)

In the randomized trial, some participants from the TELE group and VISIT group did not receive all 16 scheduled treatments due to medical, technical, or personal reasons. Average treatment frequency was 16.0 (SD 0.2) for the VISIT group and 15.3 (SD 2.1) for the TELE group (*P*=.004). Because there are expenses associated with cancelled treatments (ie, travel to the home without giving the treatment, technician time to solve technological problems), the cost analysis took into account real costs, for instance, the cost of treatment given plus cost of cancelled treatment (Table 3). Real cost analysis shows a cost differential in favor of the TELE group (TELE minus VISIT: Can -$263, 95% CI -$382 to -$143), meaning that for each participant's total intervention, telerehabilitation saves the health care system 18% of the costs incurred for conventional rehabilitation following TKA.

Table 3. Costs of telerehabilitation and home visits.

Cost	TELE ($) (n=97), mean (SD)	VISIT ($) (n=100), mean (SD)	TELE-VISIT ($), difference (95% CI)	(TELE-VISIT)/VISIT (%), difference/VISIT (95% CI)	*P* value
Total cost[a]	1224 (241)	1487 (553)	-263 (-382 to -143)	-18 (-26 to -10)	<.001
Cost per treatment[b]	80.99 (26.60)	93.08 (35.70)	-12.09 (-20.91 to -3.27)	-13 (-23 to -4)	.008

[a] Total cost is for the total intervention, which includes all received and cancelled treatments.

[b] Cost per treatment is for one treatment and is based only on treatments received.

Cost per Treatment

Cost per treatment shows that the mean cost of treatment for the VISIT group was Can $93.08 (SD $35.70) versus Can $80.99 (SD $26.60) for TELE group. The differential cost between the two groups was 13% in favor of the TELE group (TELE minus VISIT: Can -$12.09, 95% CI -$20.91 to -$3.27) (Table 3).

Total Cost Distribution for Each Group

Table 4 outlines the distribution related to direct and indirect categories of costs. As expected, total direct clinical costs—TELE $36.59, VISIT $41.65—were quite similar. However, even though rehabilitation session length was standardized in the protocol, treatment duration was longer in the VISIT group—TELE 47.6 (SD 10.2) minutes, VISIT 54.4 (SD 12.1) minutes (*P*<.001). Indirect costs for the VISIT group were higher than for the TELE group because of the costs associated with travel to the patient's home—TELE $11.51, VISIT $51.43. However, technology costs for the TELE group totaled $32.88, which to some extent counterbalanced the travel costs for the VISIT group.

As expected, the travel cost for the VISIT group is an important cost-driven variable—49% of the total cost. Interestingly, costs related to technology support during teletreatment, outside of the expected time for equipment installation/uninstallation, did not significantly affect the cost—1.5% of total cost.

At a micro level, it is somewhat surprising to see discrepancies between some subcosts. For example, therapists spent the same proportion of time with the patient in both groups—TELE 44%, VISIT 44%—but a larger proportion of time was spent in the TELE group preparing for the treatment—TELE 5%, VISIT 2%—and on follow-up—TELE 9%, VISIT 4%.

Sensitivity Analysis

There was a difference in costs per treatment between the two groups, and this difference was not the same in the five distance strata (*P*<.001 for interaction) (Table 5). When the patient's home was less than 30 km round-trip from the health care center, the difference in costs between TELE and VISIT treatments was not significant (*P*=.25, .26, and .11 for the <10, 10-19, and 20-29 km strata, respectively). However, TELE treatments were less costly than VISIT treatments when the distance was 30 km or more (30-49 km, Can $81<$103, *P*=.002; ≥50 km, $90<$152, *P*<.001).

For urban and rural areas, TELE treatments were less expensive than VISIT treatments when the round-trip distance was 50 km or more (Can $97<$144, *P*<.001) (Table 6). For metropolitan areas, TELE treatments were cheaper than VISIT treatments when the round-trip distance was 30 km or more (Montreal, Can $80<$152, *P*<.001; Quebec City, $82<$108, *P*=.001).

Table 4. Breakdown of costs per treatment (telerehabilitation and home visits).

Cost category			TELE		VISIT	
			Can $	%	Can $	%
Clinical						
	Direct					
		Treatment	35.85	44	41.02	44
		Clinical equipment amortization[a]	0.74	1	0.63	1
		Total	36.59	45	41.65	45
	Indirect					
		Time related to preparation for the treatment	4.48	5	2.26	2
		Follow-up with patients	7.03	9	3.55	4
		Travel to participant's home (time + km allocation)	N/A[b]	N/A	45.62	49
		Total	11.51	14	51.43	55
	Total clinical costs		48.10	59	93.08	100
Technological						
	Direct					
		Technical equipment amortization[a]	5.81	7.2	N/A	N/A
		Installation/uninstallation of technology	18.12	22.4	N/A	N/A
		Internet costs	7.63	9.4	N/A	N/A
		Technical problem resolution time and travel time (technician)	1.25	1.5	N/A	N/A
		Technical problems after sessions by the physiotherapist	0.08	0.1	N/A	N/A
	Total technological costs		32.88	41	N/A	N/A
Total costs			80.99	100	93.08	100

[a] The clinical and technical equipment amortization costs are slightly different from those in Table 1 for the TELE group because some patients did not receive all 16 scheduled treatments.

[b] Not applicable (N/A).

Table 5. Cost sensitivity analysis based on round-trip distance from health care center to patient's home in the total sample (n=197).

Distance (km)	Costs per treatment (Can $)				*P* value[a]
	TELE		VISIT		
	n (%)	Mean (SD)	n (%)	Mean (SD)	
<10	12 (6.1)	71.0 (9.8)	21 (10.7)	63.9 (7.4)	.248
10-19	24 (12.2)	69.5 (8.5)	27 (13.7)	74.8 (10.5)	.260
20-29	14 (7.1)	73.6 (10.1)	19 (9.6)	83.1 (8.3)	.112
30-49	12 (6.1)	81.3 (13.1)	13 (6.6)	102.7 (19.5)	.002
≥50	35 (17.8)	90.2 (24.4)	20 (10.2)	151.6 (30.6)	<.001

[a] Two-way ANOVA with interaction was conducted to compare costs per treatment between the two groups and between the five distance strata (*P*<.001 for interaction). *P* values shown in the table come from post hoc comparisons.

Table 6. Cost sensitivity analysis based on round-trip distance from health care center to patient's home, calculated separately for each site.

Site		Costs per treatment (Can $)				Interaction *P* value[a]	Post hoc *P* value
		TELE		VISIT			
	Distance (km)	n (%)	Mean (SD)	n (%)	Mean (SD)		
Urban and rural area							
Sherbrooke (n=109)							
	< 50	34 (31.2)	69.8 (8.3)	46 (42.2)	71.4 (10.4)	<.001	.684
	≥ 50	20 (18.3)	97.3 (28.8)	9 (8.3)	143.5 (26.1)		<.001
Metropolitan areas							
Montreal (n=37)							
	< 30	4 (11)	74.6 (10.4)	8 (22)	92.4 (12.3)	.002	.192
	≥ 30	13 (35)	80.2 (9.0)	12 (32)	152.0 (34.9)		<.001
Quebec City (n=50)							
	< 30	14 (28)	73.7 (10.6)	16 (32)	75.0 (9.0)	.001	.815
	≥ 30	11 (22)	81.7 (15.7)	9 (18)	107.7 (28.6)		.001

[a] For each site, a two-way ANOVA with interaction was conducted to compare costs per treatment between the two groups and between the distance strata. Results shown in the table are *P* values for interaction between groups and distance strata, and *P* values from post hoc comparisons.

Discussion

Principal Findings

The purpose of this study was to document costs for in-home teletreatment versus conventional home-visit rehabilitation following TKA based on an RCT. To our knowledge, this is the first cost analysis of in-home telerehabilitation involving an assessment of the cost structure for this type of service delivery versus traditional in-home, face-to-face rehabilitation. One unique facet of our study is the inclusion of all subcosts related to telerehabilitation (ie, equipment amortization, installation/uninstallation, technical problems related to teletreatment). Our results show that delivering physiotherapy by telerehabilitation post-TKA is less expensive than providing the same service in home visits if the distance between the health care center and patient's home is more than 30 km. Overall, when we take into account all direct and indirect costs of the two approaches, teletreatment between the health care center and patient's home is less costly than home visits by at least 18%.

In the context of cost constraints in the health care system, is the cost savings of 18% sufficient to introduce in-home telerehabilitation? This result did not specifically take into account one major cost-driven factor related to home visits, which is proximity of the participant's home to the health care center (ie, sensitivity analyses). The inherent hypothesis that comes to mind is that proximity of health care centers to patients' homes can significantly influence the cost of the VISIT group—the farther away a participant lives, the higher the costs related to home-visit rehabilitation. For example, in the VISIT group, 16 round-trips were needed to provide treatment in the home by the physiotherapist, but only two round-trips were needed for the TELE group (ie, installation/uninstallation) by the technician. To take into consideration the proximity of the participant's home to the health care center, a sensitivity analysis clearly showed that if travel distance is less than 30 km, round-trip, differential costs between the two alternatives decline. On the other hand, if round-trip travel distance is longer than 30 km, the cost differential increases dramatically, reaching 123% for 50 km or more. This is an important result for those who want to implement telerehabilitation—to benefit from cost savings, telerehabilitation must involve patients who live at least a 30 km drive, round-trip, from the clinic if the resources are available. In addition, it would be more efficient to use specialized health care providers to administer treatments than to spend their time travelling, as was done in the VISIT group.

Moreover, if we consider traffic in metropolitan, urban, and rural areas as a confounding variable to proximity of the participant's home to the health care center, we observed the same tendency, with the difference being that in the urban and rural area (Sherbrooke), a difference of 50 km was needed to keep the cost savings of the TELE versus the VISIT group. This study shows that costs were not higher for the home-visit group, which could be of importance to stakeholders and managers.

As a novel piece of the puzzle to understand costs of telerehabilitation compared to home visits, examination of the subcosts of the two approaches produced some interesting results. For that analysis, only costs of the treatment actually delivered were considered. Direct treatment time was not the same in both groups—it was less in the TELE group. Even though there was a standardization of the treatment in both groups in the research protocol, this difference can be attributed to taking less time to discuss with the patient or less time between exercises because of the virtual mode of communication. However, we saw that the proportion of time spent on preparing for the treatment and following up after therapy was more than double in the TELE group (Table 4). We attribute this result to a virtual session inducing more caution on the part of the therapist. Being in front of the screen waiting for the time scheduled with the patient may lead to taking more time to prepare for the session as opposed to driving a car to go to the patient's home. Further research should assess these aspects in a qualitative study embedded in another trial.

As expected, travel drove the costs of home visits (ie, 49% of total costs). The technology of the TELE group was also costly (ie, 41% of total), but not enough to counterbalance travel costs. Looking at the subcategories of the 41% technology costs, 7.2% of the costs were attributable to equipment amortization and 9.4% to telecommunication (ie, high-speed Internet). At the time of the TelAge RCT (2008-2013), the videoconferencing equipment purchase costs and telecommunication costs were much higher than for comparable equipment and services today. Comparable videoconferencing equipment is available for less than half the purchase price used in our analyses. Telecommunication costs for activation fees and bandwidth are significantly lower as accessibility and use of high-speed Internet have now reached unprecedented levels, even with older adults, and bandwidth costs have declined. As the cost per treatment for the technological factors of in-home telerehabilitation was only 16% of the total costs, this would not appear to be a determining variable in our results. However, the exponential evolution of these technological factors in today's technological context (ie, lower cost of videoconferencing, ubiquity, and performance of telecommunication networks) could impact the cost differential between the TELE group and the face-to-face VISIT group services delivery by lowering the barrier to entry of the TELE option when scaling up in-home visit rehabilitation.

Regarding internal validity, information bias was minimized by using specific procedures. First, data collection was standardized with cost forms and information was collected immediately after each intervention (ie, physiotherapy session). Second, the standardized cost forms were pretested. This assured us of the specific details required to calculate all subcosts. For each subcost, the person involved in service delivery (ie, technician, engineer, physiotherapist, study coordinator, etc) completed the appropriate form the same day that the intervention occurred. Following this rigorous process, memory bias should have been minimized. Finally, to decrease missing data, the site coordinator checked the data for completeness. With all these precautions, we are confident in the validity of the cost data.

Selection bias was controlled by the randomization procedure in the TelAge RCT leading to two equivalent groups regarding sociodemographic variables [10]. The only exception was the mean distance per round-trip between the health care center and the patient's home, which was greater in the TELE group. If this discrepancy had an effect, it should have tended to increase the 18% cost difference observed in the TELE group.

Clinical Implications of This Study

We already know that teletreatment for TKA is efficient [10]. The results of the TelAge clinical trial demonstrated that the TELE group treatment was at least as clinically effective as home visits. That was a key issue in order to justify implementation of telerehabilitation. The cost analysis conducted in this trial supports the use of teletreatment in cost-saving conditions. If teletreatments are given the same way as in this study, cost savings can range from 18% to 123%, depending on travel time from the health care center to the patient's home. The cost savings of at least 18% is a new way to increase budget for rehabilitation services. For example, 18% of a $1 million budget means cost-savings of $180,000. That means that at least two more health professionals can be hired over the next year. Moreover, with sensitivity analysis, this increase may be more important for some facilities. With well-known budget constraints being experienced in health care, worldwide, we must investigate innovative ways of managing and stretching budgets. We believe that health care centers could use teletreatment for a subpopulation of their patients. We think that the hypothesis that 50% of the scheduled treatment time can be dedicated to telerehabilitation—because not all patients can/want/must receive teletreatment in a general hospital—is conservative and reliable. We need to understand that telerehabilitation is a compromise for the therapist—he/she may not want to treat all his/her patients in a virtual way, perhaps only those who cannot reach outpatient clinics or those for whom home care is not feasible.

Conclusions

In the context of increasing access to rehabilitation, our study demonstrated the cost savings of teletreatment compared to home visits for patients post-TKA. Managers can now decide whether to implement telerehabilitation based on robust clinical and economic data. Given the recent attention that telerehabilitation is drawing on an international level, it would be interesting in future research to carry out a large-scale study about one of the major practical issues related to the provision of telemedicine services to a specific patient population.

Acknowledgments

The authors would like to thank all participants, study personnel, and orthopedic surgeons for their contributions to this project. This research was supported by funding from the Canadian Institutes of Health Research (CIHR).

Conflicts of Interest

None declared.

References

1. World Report on Disability. Geneva, Switzerland: WHO Press, World Health Organization; 2011. URL: http://www.unicef.org/protection/World_report_on_disability_eng.pdf [accessed 2015-03-06] [WebCite Cache ID 6WpYMdlNb]
2. Chez Soi: Le Premier Choix. La Politique de Soutien à Domicile. chemin Sainte-Foy, Québec: Ministère de la Santé et des Services sociaux du Québec; 2003. URL: http://publications.msss.gouv.qc.ca/acrobat/f/documentation/2002/02-704-01.pdf [accessed 2014-05-19] [WebCite Cache ID 6PgyacXT0]
3. Lingard EA, Berven S, Katz JN, Kinemax Outcomes Group. Management and care of patients undergoing total knee arthroplasty: variations across different health care settings. Arthritis Care Res 2000 Jun;13(3):129-136. [Medline: 14635286]
4. Naylor J, Harmer A, Fransen M, Crosbie J, Innes L. Status of physiotherapy rehabilitation after total knee replacement in Australia. Physiother Res Int 2006 Mar;11(1):35-47. [Medline: 16594314]
5. Landry MD, Jaglal SB, Wodchis WP, Cooper NS, Cott CA. Rehabilitation services after total joint replacement in Ontario, Canada: can 'prehabilitation' programs mediate an increasing demand? Int J Rehabil Res 2007 Dec;30(4):297-303. [doi: 10.1097/MRR.0b013e3282f14422] [Medline: 17975449]
6. Bashshur R, Shannon G, Krupinski E, Grigsby J. The taxonomy of telemedicine. Telemed J E Health 2011 Jul;17(6):484-494. [doi: 10.1089/tmj.2011.0103] [Medline: 21718114]
7. Tousignant M, Boissy P, Corriveau H, Moffet H, Cabana F. In-home telerehabilitation for post-knee arthroplasty: A pilot study. Int J Telerehab 2009 Sep 04;1(1):9-16. [doi: 10.5195/ijt.2009.5997]
8. Russell TG, Buttrum P, Wootton R, Jull GA. Low-bandwidth telerehabilitation for patients who have undergone total knee replacement: preliminary results. J Telemed Telecare 2003;9 Suppl 2:S44-S47. [doi: 10.1258/135763303322596246] [Medline: 14728759]
9. de la Torre-Díez I, López-Coronado M, Vaca C, Aguado JS, de Castro C. Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: a systematic review. Telemed J E Health 2015 Feb;21(2):81-85. [doi: 10.1089/tmj.2014.0053] [Medline: 25474190]
10. Drummond MF, Sculpher MJ, Torrance G, O'Brien B, Stoddard G. Methods for the Economic Evaluation of Health Care Programmes. 3rd edition. Oxford, UK: Oxford University Press; Jun 20, 2005.
11. Moffet H, Tousignant M, Nadeau S, Mérrette C, Boissy P, Corriveau H, et al. In-Home Telerehabilitation compared to Face-to-Face Rehabilitation after Total Knee Arthroplasty: A Non-Inferiority Randomized Controlled Trial. (forthcoming).
12. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic Evaluation in Clinical Trials. 2nd edition. Oxford, UK: Oxford University Press; 2014.

Abbreviations

ANOVA: analysis of variance
CIHR: Canadian Institutes of Health Research
E1: assessment time, at baseline
E2: assessment time, 1 to 2 days before discharge from hospital
E3: assessment time, 2 months postdischarge
E4: assessment time, 4 months postdischarge
GLM: general linear model
ISRCTN: International Standard Registered Clinical Study Number
km: kilometer
LSMEANS: least squares means
N/A: not applicable
PTZ: pan-tilt-zoom
RCT: randomized controlled trial
TelAge: telerehabilitation for knee arthroplasty
TELE: in-home telerehabilitation (experimental) treatment
TKA: total knee arthroplasty
VISIT: home-visit (control) treatment
WOMAC: Western Ontario and McMaster Universities Arthritis Index

Edited by G Eysenbach; submitted 04.09.14; peer-reviewed by D Putrino, R Lloréns, K Laver; comments to author 06.02.15; revised version received 20.02.15; accepted 04.03.15; published 31.03.15.

Please cite as:
Tousignant M, Moffet H, Nadeau S, Mérette C, Boissy P, Corriveau H, Marquis F, Cabana F, Ranger P, Belzile ÉL, Dimentberg R
Cost Analysis of In-Home Telerehabilitation for Post-Knee Arthroplasty
J Med Internet Res 2015;17(3):e83
URL: http://www.jmir.org/2015/3/e83/
doi:10.2196/jmir.3844
PMID:25840501

Original Paper

Ranking Adverse Drug Reactions With Crowdsourcing

Assaf Gottlieb[1], PhD; Robert Hoehndorf[2], PhD; Michel Dumontier[3], PhD; Russ B Altman[4], MD, PhD

[1]Department of Genetics, Stanford University, Stanford, CA, United States
[2]Computer, Electrical and Mathematical Sciences & Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[3]Stanford Center for Biomedical Informatics Research, Stanford University, Stanford, CA, United States
[4]Departments of Genetics and Bioengineering, Stanford University, Stanford, CA, United States

Corresponding Author:
Russ B Altman, MD, PhD
Departments of Genetics and Bioengineering
Stanford University
Shriram Center Room 209 443 Via Ortega MC 4245
Stanford, CA, 94305
United States
Phone: 1 (650) 725 3394
Fax: 1 (650) 725 3863
Email: russ.altman@stanford.edu

Abstract

Background: There is no publicly available resource that provides the relative severity of adverse drug reactions (ADRs). Such a resource would be useful for several applications, including assessment of the risks and benefits of drugs and improvement of patient-centered care. It could also be used to triage predictions of drug adverse events.

Objective: The intent of the study was to rank ADRs according to severity.

Methods: We used Internet-based crowdsourcing to rank ADRs according to severity. We assigned 126,512 pairwise comparisons of ADRs to 2589 Amazon Mechanical Turk workers and used these comparisons to rank order 2929 ADRs.

Results: There is good correlation (rho=.53) between the mortality rates associated with ADRs and their rank. Our ranking highlights severe drug-ADR predictions, such as cardiovascular ADRs for raloxifene and celecoxib. It also triages genes associated with severe ADRs such as epidermal growth-factor receptor (EGFR), associated with glioblastoma multiforme, and SCN1A, associated with epilepsy.

Conclusions: ADR ranking lays a first stepping stone in personalized drug risk assessment. Ranking of ADRs using crowdsourcing may have useful clinical and financial implications, and should be further investigated in the context of health care decision making.

(J Med Internet Res 2015;17(3):e80) doi:10.2196/jmir.3962

KEYWORDS

pharmacovigilance; adverse drug reactions; drug side effects; crowdsourcing; patient-centered care; alert fatigue

Introduction

Pharmacovigilance plays a crucial role in the continuing evaluation of drug safety. Adverse drug reactions (ADRs) contribute to excess length of hospitalization time, extra medical costs, and attributable mortality [1,2]. Thus, assessment of the impact of ADRs on drug risk-benefit assessment has gained significant interest in recent years as several risk-benefit methodologies have been suggested for assessing drug safety and efficacy [3,4]. Two factors are essential for risk assessment: the prevalence of the ADR in the population (ie, frequency) and the severity of the ADR in terms of medical (morbidity and mortality) or financial consequences. Risk estimates focus mainly on ADR frequency, as there is no publicly available resource that provides estimates of relative severity of ADRs. Thus, these methods either handle a single ADR at a time [3] or assign equal weights for all the drug ADRs [5]. However, not all ADRs are of equal interest: life-threatening ADRs require more attention, while minor ADRs may not. Although a few severe life-threatening ADRs are well recognized, including liver failure, cardiac arrest, and others, there is presumably a gradation of severity from these down to the most benign. Of course, patients' subjective perception of the severity of an ADR varies widely, and so a ranking of ADRs is fundamentally a personal activity when it comes to individual patient decisions. Nonetheless, a ranking of ADRs based on perceived severity is a useful starting point for risk-benefit assessment and for patient-centered care, and is the focus of this paper.

Ranking large sets of ADRs is challenging; theoretical analyses have provided a framework for such evaluations [6,7]. Tallarida et al asked 53 physicians to assign weights to seven severity classes, but their study contained ADRs specific to only two drug classes (treating hypertension and rheumatoid arthritis), and thus has limited general utility. In a subsequent work, the authors showed consistent ranking between the 53 physicians and 56 non-professional rankers. Encouraged by this result, we sought to crowdsource rankers to obtain a more comprehensive ADR ranking. In order to accomplish this, we divided this complex task into simpler microtasks (pairwise comparisons), well-suited to a crowdsourcing platform, such as the Amazon Mechanical Turk (MTurk). MTurk is a crowdsourcing microtask platform (microtasking refers to tasks that are divided into multiple smaller subtasks) that allows human workers to perform tasks in return for payment (see Methods for extended description). Previous evaluations have shown that MTurk can be as reliable as traditional survey methods, and that the use of control validation questions can markedly improve reliability and reduce variability [8]. To the best of our knowledge, crowdsourcing has not been used for pharmacovigilance applications yet.

Our goal was to rank the ADRs by severity from a population (non-expert, non-clinician) perspective. We ranked a list of 2929 ADRs by assigning 126,512 ADR pairwise comparisons to 2589 individuals and processing the comparisons with an optimization algorithm to rank the ADR severities.

ADRs are reported in drug labels following clinical trials. Additional drug-ADR associations can be inferred, both empirically, through reporting systems such as the US Food and Drug Administration (FDA) Adverse Events Reporting System (AERS), or based on computational predictions (using drug similarity [9], genetic overlap [10], and pathway analysis [11]), followed by pharmacoepidemiological studies to confirm these predictions. These associations are numerous: the average FDA label lists 100 ADRs and some prediction work suggests that the number of ADRs may be doubled. A severity ranking would be useful to triage ADR predictions for further investigations. In the discussion, we list three additional uses for our ranked list, including reduction of “alert fatigue”, whereby alerts are ignored [12], identification of ADRs that suggest a larger component of patient preference, and association of individual genes with ADR severity. We make the raw data upon which our analysis is based and the resulting ADR ranking publicly available.

Methods

Data

ADRs were retrieved from the “SIDER2” side effect database (October 2012 version, listing total of 4192 ADRs) [13]. Predicted ADRs-drug associations mined from the FDA Legacy AERS were retrieved from [14] (“OFFSIDES” off-label side effect database). Gene-ADR associations assembled from literature were retrieved from [15] and predicted gene-ADR associations based on inferred pharmacodynamics pathways were retrieved from [11] (DrugRouter). ADRs from SIDER2, AERS, and OFFSIDES are coded using the medical dictionary for regulatory activities (MedDRA) terminology [16].

The AERS data files were downloaded from the FDA website [17]. The data covered the interval first quarter, 2004, through third quarter, 2012. All files were imported into an SQLite database and fields were checked against a list of allowed values drawn from documentation supplied by the FDA in the download. Three formatting errors were corrected manually. Individual safety reports were aggregated into cases, removing duplicate reports per case (follow-ups). The top 100 prescribed drugs in 2013 were retrieved from [18].

Semantic similarity between ADRs was computed using the Human Phenotype Ontology (HPO) [19], downloaded on May 30, 2014, using the Semantic Measures Library v0.8 [20].

What is Amazon Mechanical Turk (MTurk)?

MTurk is a platform for task creation, labor recruitment, and compensation. “Requesters” create and publish “human intelligence tasks” (HITs) and “workers” complete these tasks. The tasks are ones that can be completed using a computer and typically require a short time to complete, with a corresponding small compensation. Prior to posting a task, the requester sets the compensation amount (Amazon charges an additional 10% commission). Workers can browse and choose from available tasks and are paid upon successful completion of each task. Requesters can also reject subpar work. In this case, the rejected workers do not receive payment and it also negatively affects the worker record as requesters may limit their tasks to workers with low rejection rates.

Ranking 2929 Adverse Drug Reactions

We retrieved a set of 2929 common ADRs (expressed in the MedDRA terminology) from drug labels, as represented in the SIDER2 database [13]. All 2929 ADRs were also reported in the FDA AERS, which collects ADR reports from clinicians, patients, and pharmaceutical companies. We used the Amazon MTurk crowdsourcing platform to assign 126,512 ADR comparison tasks to 2589 individual rankers (“workers” in MTurk terminology) (Figure 1 A). Each worker was presented with up to 15 sets of 10 pairwise ADR comparisons and was requested to select, for each pair, which ADR is more severe. The user interface provided clickable links to Google queries with the ADR name in order to help workers learn about ADRs expressed in medical terminology with which they were not familiar (Multimedia Appendix 1 displays an example set of comparisons presented to workers).

The workers were required to possess satisfactory task completion records, rejected in less than 5% of past tasks (95% approval rate), and be located in the United States, as a proxy to English proficiency. In order to identify reliable workers, each worker task of 10 pairs included three pre-defined quality control pairs with expected answers and seven randomly chosen pairs. These quality control pairs were constructed by pairing all combinations from a manually selected set of severe ADRs and a set of mild ADRs.

Using the pre-defined set of quality control comparisons, we removed inconsistent workers who did not answer these appropriately, resulting in 124,513 usable pairwise comparisons (57,901 unique comparisons, multiple comparisons were made for consistency evaluation, see Multimedia Appendix 2 for raw comparisons).

In construction of the pairwise comparisons, we took the following measures in order to maximize the tested pairs and reduce as much as possible potential biases: (1) the tasks were distributed on different weekdays over a period of 1 month, and (2) using an initial crude ranking computed from the first batch of comparisons, we randomly selected the ADR pairs that were not too easy (comparing a severe and a mild ADR) or equivalent (ADRs with very close ranks), as equivalent ADRs are harder to compare and have the potential to frustrate the MTurk workers in being forced to choose.

A quality control batch of pairwise comparisons (14,645 pairs) was repeated three times to assess reproducibility. It was also constructed to maximize the number of pairs that can be tested for triangular inequality (ie, for ADRs A, B, and C, test A vs B, B vs C, and A vs C).

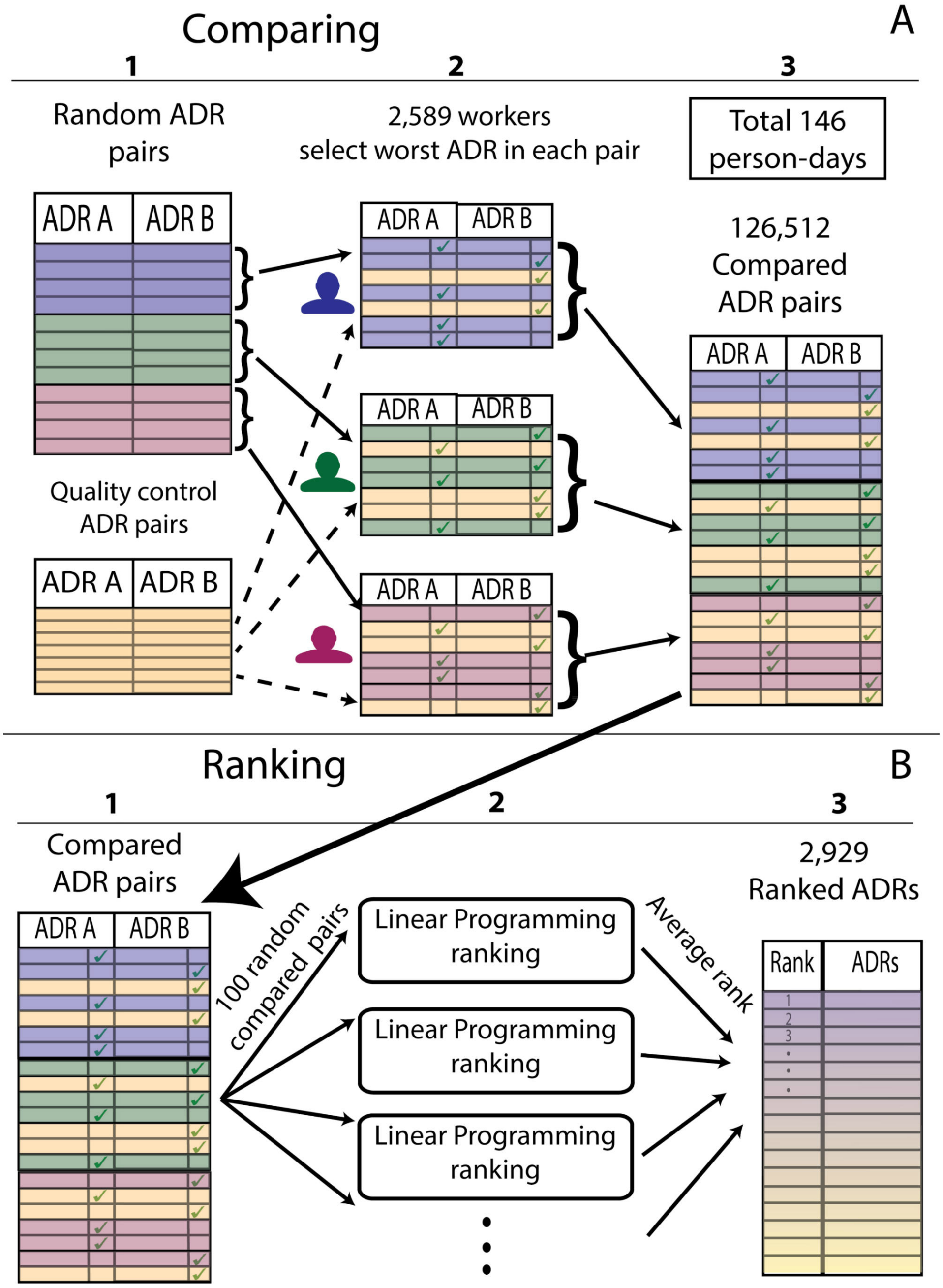
Each task, consisting of 10 pairwise comparisons, took 5 minutes to complete on average, yielding US $0.45 per worker (half a dollar including Amazon's fee). The entire ranking totaled in 146 person days at a cost of US $6,300. A more detailed description and worker statistics are found in Multimedia Appendix 3.

Ranking Adverse Drug Reactions

We formulated a linear programming scheme to compute a ranked list of the ADRs from the pairwise comparisons (illustrated in Figure 1 B). Due to time and budget constraints, we were able to sample only a small fraction (1.5%) of the nearly 4.3 million possible ADR pairs. In order to rank the ADRs based on the pairwise comparisons, we used linear programming that attempted to retain as much of the original rankings of the workers (in the minimization of the utility function) while ensuring that the ADRs obeyed the triangular inequality, that is, for each ADR triplet A, B, and C, we denote more severe as "greater than", so if A>B and B>C then it follows that A>C. The linear programming optimization function and constraints are found in Multimedia Appendix 3, Ranking ADRs section.

The linear programming was implemented in MATLAB using IBM CPLEX package version 12.6 [21].

Figure 1. MTurk task construction (A) and ranking process (B). (A) Random list of pairwise comparisons and list of predefined quality control pairs are constructed (1). Each worker receives unique set of 7 random ADR pairs to compare and 3 quality control pairs for performance evaluation (2). Results are collected and merged (3). (B) Ranked pairs are sampled (1), sent to a linear programming task (2), and ranking of each sample merged to a global ranking (3).



Consistency of Adverse Drug Reaction Rankings

Overview

We estimated the consistency of pairwise comparisons using a batch of comparisons that was constructed for quality control purposes. It was repeated three times and included multiple ADR triplets that were tested for triangular relationships. Specifically, for each ADR (A), we included 10 comparisons that formed 10 testable triangular relations (ie, for ADRs B and C, we included the three comparisons A vs B, A vs C, and B vs C).

We tested the reproducibility of the ranking across the three repeated batches. Only 16% of the workers participated in more than one of the repeated batches (13% in two batches, and 3% in all three batches).

Adverse Drug Reaction Ranks Are Associated With Relative Deaths From Adverse Events Reporting System (AERS) Reports

We counted the number of reports associated with an ADR in the AERS and the number of reports specifying one of the six outcomes (death, disability, life-threatening, required intervention to prevent permanent impairment/damage, hospitalization, and congenital anomaly). The rate of each outcome per ADR is the number of reports with that outcome divided by the total number of reports for that ADR, including reports with non-specific outcome tagged as “other serious” (25%) and reports with no outcome specified (20%).

In order to extract the major outcomes associated with the severity ranking, we used the lasso regression method [22] with a 10-fold cross validation. Relative death rate was the leading factor (producing 4% increase over the best mean standard error).

Results

Ranking 2929 Adverse Drug Reactions

We ranked a set of 2929 common ADRs from the SIDER2 database [13] using workers from the Amazon MTurk crowdsourcing platform (see Figure 1 A, and Methods). We used a pre-defined set of quality control comparisons with expected answers to remove inconsistent workers. A total of 90% of workers answered correctly on all the quality control pairs. We formulated a linear programming scheme to compute a ranked list of the ADRs from these pairwise comparisons (see Figure 1 B and Methods). The most severe ADRs were cardiac arrest and metastatic bone cancer and the least severe were euphoric or elevated mood. We list the 20 most and least severe ADRs in Table 1 and the full ranked list in Multimedia Appendix 4.

Table 1. Top- and bottom-ranked ADRs.

Rank	Top-ranked severe ADRs	Rank	Bottom-ranked mild ADRs
1	Cardiac arrest	2910	Growth of eyelashes
2	Bone cancer metastatic	2911	Eye rolling
3	Left ventricular failure	2912	Night sweats
4	HIV infection[a]	2913	Chapped lips
5	Anal cancer	2914	Nasal congestion
6	Lung cancer metastatic	2915	Agitation
7	Hemorrhage intracranial	2916	Excitability
8	Chronic myeloid leukemia	2917	Breath odor
9	Coma	2918	Hair growth abnormal
10	Breast cancer	2919	Hot flush
11	Multi-organ failure	2920	Sleep talking
12	Cardiopulmonary failure	2921	Blister
13	Cardiac death	2922	Tongue dry
14	Chronic leukemia	2923	Moaning
15	Cardio-respiratory arrest	2924	Discomfort
16	Pulmonary embolism	2925	Decreased appetite
17	Completed suicide	2926	Dry mouth
18	Metastatic renal cell carcinoma	2927	Early morning awakening
19	Hepatic angiosarcoma	2928	Euphoric mood
20	Anaplastic thyroid cancer	2929	Elevated mood

[a] HIV: Human immunodeficiency virus infection, while not caused by a drug, is associated in with several drugs in SIDER.

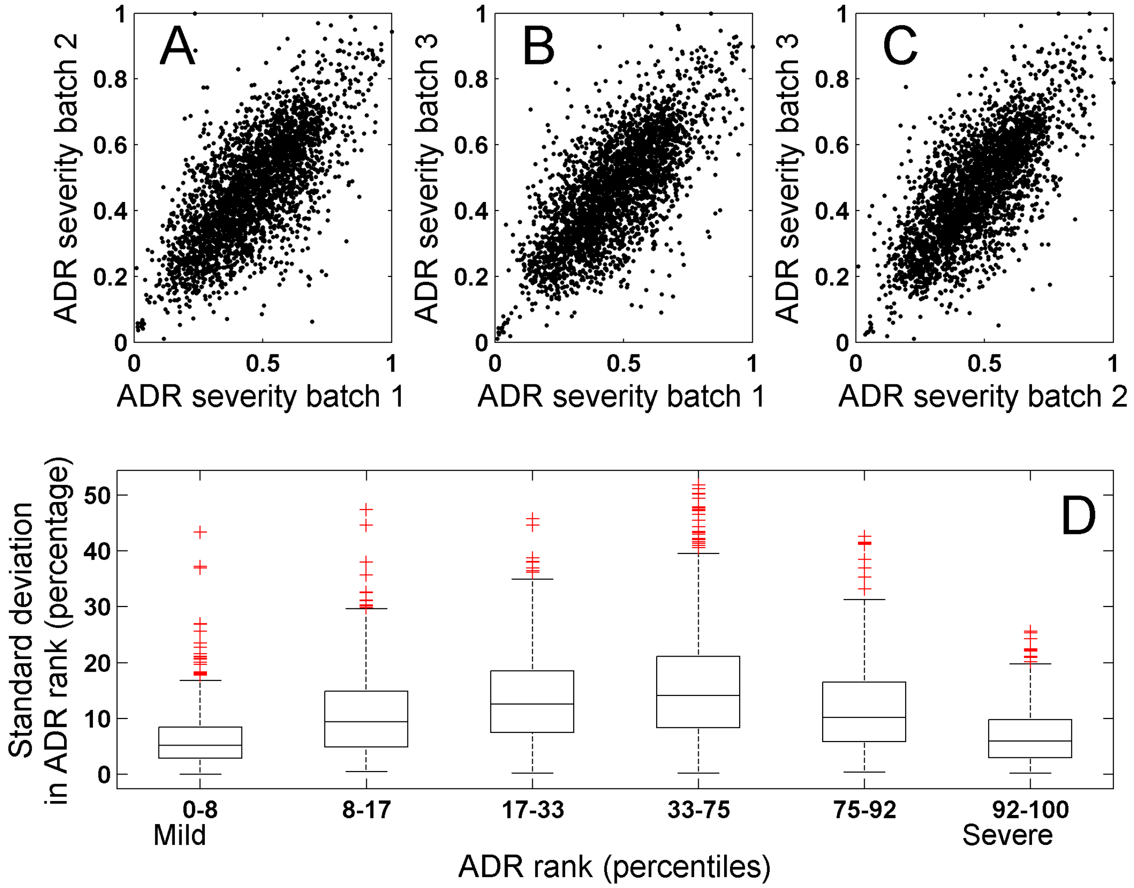
Consistency of Adverse Drug Reaction Rankings

We estimated the consistency of pairwise comparisons by repeating a quality control batch of pairwise comparisons three times. The batch included multiple ADR triplets that were tested for triangular relationships. Only 10% (SD 0.3%) of these ADR triplets violated the triangular inequalities (total of 23,071-26,245 triplets in each batch repeat, variation is due to exclusion of workers judged inconsistent on pre-defined quality control pairwise comparisons).

We next tested the reproducibility of the ranking across the three repeated batches (see Methods). Among pairs compared by at least three different workers from these three duplicate batches, 58% had full agreement. Despite this agreement, the Spearman correlation coefficient between the ranking independently computed from the three duplicate batches was .71 (SD .009, $P<.001$) (Figure 2 A-C). Based on the standard deviation of ranks across the three repeated batches, a one-sided analysis of variance identified six robust classes of ADRs ($P<.001$) (see Figure 2 D and Multimedia Appendix 3 for details). Among the ADRs with highly variable rank, we find hemolysis, tracheoesophageal fistula, actinic keratosis, suicidal ideation, and chronic otitis media. ADRs with the least variable rank included furuncle, moaning, chapped lips, and discomfort.

Finally, ADRs sharing high semantic similarity exhibited smaller difference in their severity ranks (Pearson correlation ρ=−.94, *P*<.001) (see Multimedia Appendix 5).

Figure 2. Correspondence between duplicate quality control batches. Ranking correlation between duplicate batches 1-3 (A-C) and a box-plot of the standard deviation in rank scores across the 3 batches as a function of the score (D).



Severe Adverse Drug Reactions Are Associated With More Deaths in the FDA Adverse Events Reporting System (AERS)

AERS contains reports on adverse event submitted to the FDA. Some of the reports include a specific outcome of the ADR (55% of the reports including ADRs in our set). These specific outcomes are death, disability, life-threatening, required intervention to prevent permanent impairment/damage, hospitalization, and congenital anomaly. We found a significant correlation between the relative death rate in AERS reports (ie, the relative number of deaths out of all ADR reports) and our severity rank for the ADR (ρ=.53, *P*<.001) (Figure 3 A). Additionally, life-threatening and hospitalization outcomes were moderately correlated with our ranking (ρ=.35 and ρ=.34, respectively, *P*<.001 for both) (Figure 3 B and C). The other possible outcomes were not strongly correlated with our ADR severity (see Figure 3 D-F and also the Discussion). Death rate is also the highest contributing factor in a lasso regression [22] of the AERS outcomes percentages against the ADR score (Methods). There are exceptions that illustrate the limits of this evaluation: idiopathic pneumonia syndrome (an ADR associated with high dose carmustine chemotherapy [23]) was ranked at the 48th percentile in severity but displays a high relative death rate (84%). Conversely, breast and thyroid cancers are ranked among the 99th percentile of severe ADRs, but have lower than 10% reported mortality rate in the AERS. We found no significant correlation between the ADR severity ranking and patient demographic traits (age and gender) in the AERS reports. Figure 4 displays a cloud of the most and least severe ADRs, sized by their relative mentioning in the AERS system, displaying wide variation across different ADRs of similar ranking.

Figure 3. Correlation between ADR rank and outcomes. Severe ADRs tend to have significantly higher death rate (A), moderate correlation with life-threatening (B), and hospitalization (C), and negligible correlation with congenital anomaly (D), required intervention to prevent permanent impairment/damage (E), and disability (F).

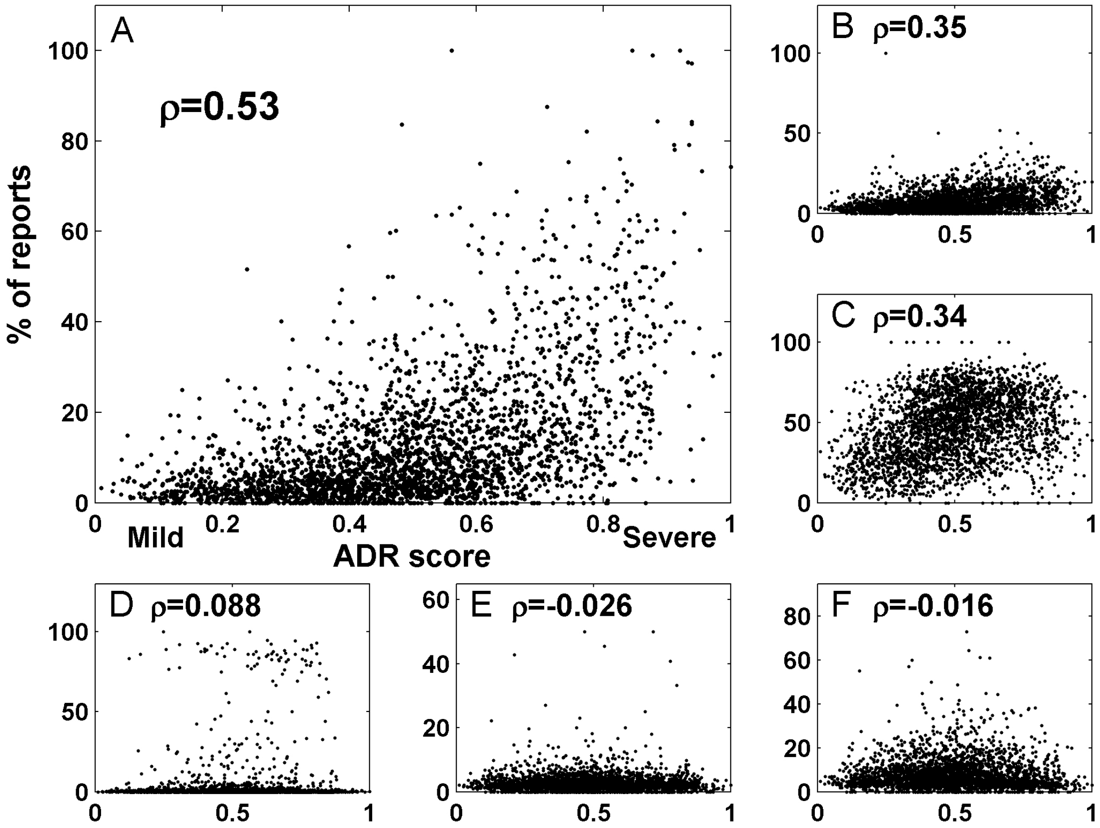
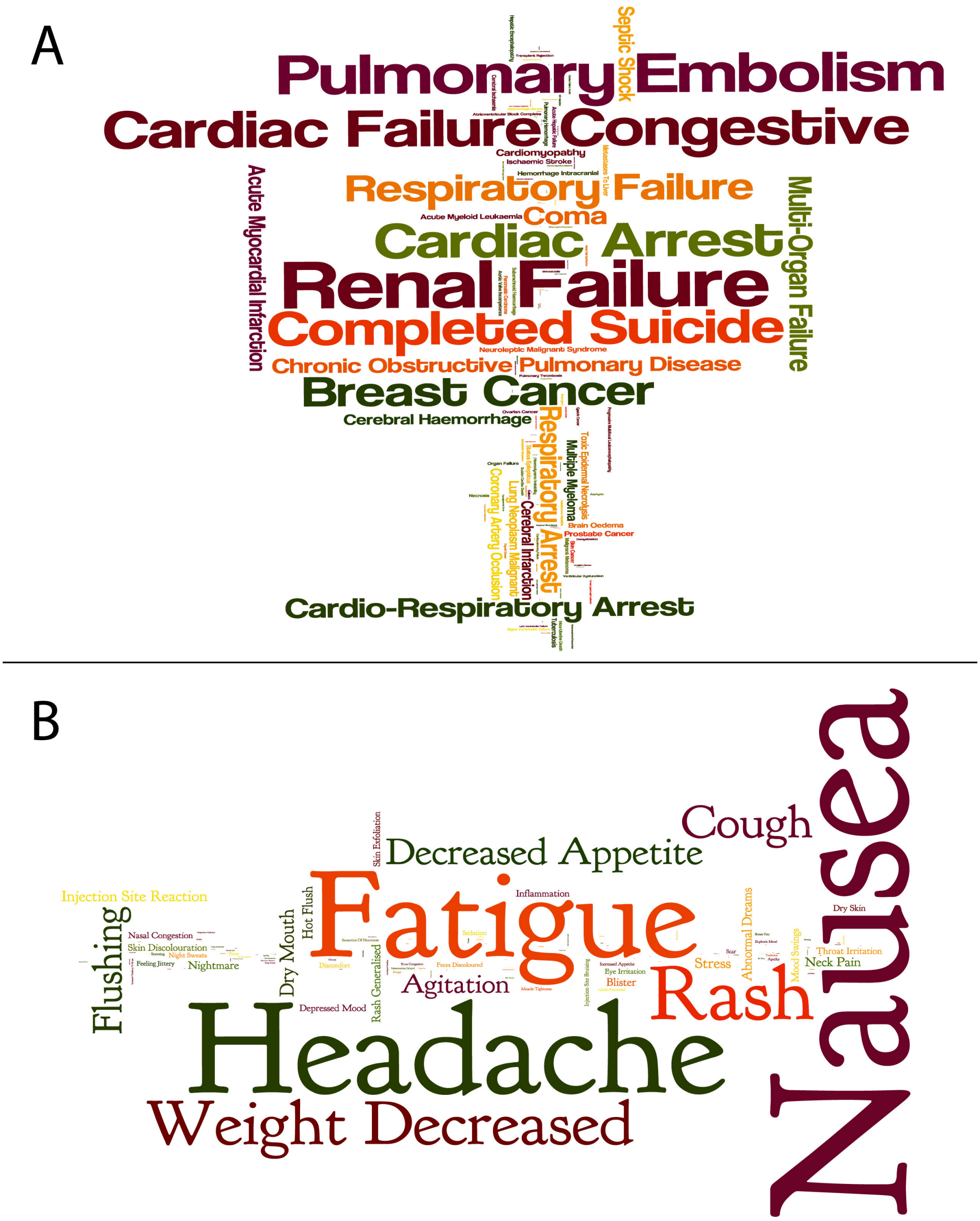


Figure 4. Term clouds for top 95 percentile ADRs (A) and bottom 5 percentile (B). Term size is proportionate to the relative number of reports in the FDA AERS.

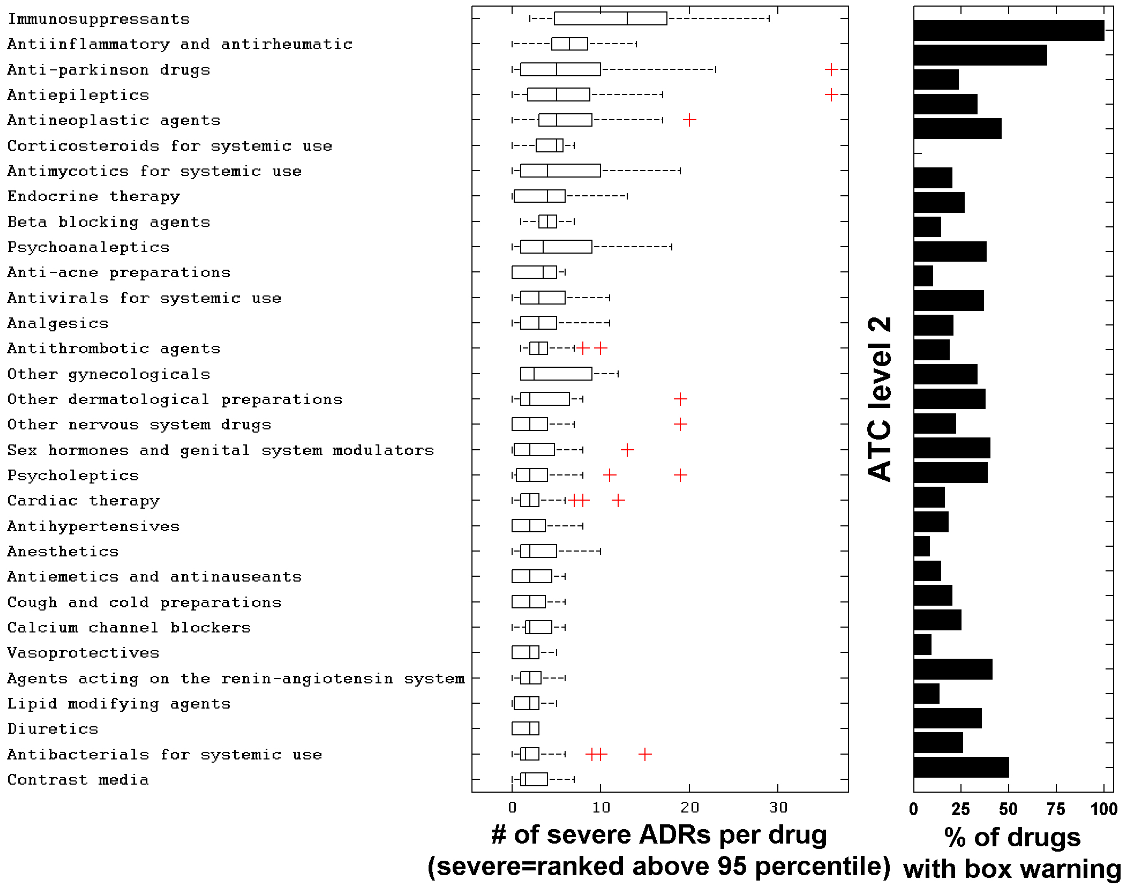


Some Therapeutic Classes Have More Severe Types of Adverse Drug Reactions Than Others

Drug risk assessment is affected by the severity of its associated ADRs and by their frequency in the population. In order to evaluate the reliability of ADR frequencies, we surveyed drug labels for 65 severe and frequent drug-ADR associations, where we define severe ADRs as those ranked above the 95th percentile and frequent drug-ADR association as those reported with larger than 1% frequency in the SIDER database. The frequency information in those labels was largely insufficient to estimate the marginal frequency above a control (ie, a placebo). Only two associations (3%) were compared to a control group that underwent a procedure (orchiectomy) instead of receiving a different drug. The reported frequency was significantly higher than that control (5% occurrence for congestive cardiac failure and chronic obstructive pulmonary disease after administration of zoladex, vs 1% for the control, *P*<.001). We thus disregarded frequency information and focused on the most severe ADRs, assuming they are essential to highlight until their frequencies are determined.

We associated ADRs with a set of therapeutic drug classes by aggregating the drug-ADR associations according to therapeutic class, as defined by the second level of the drug Anatomical Therapeutic Chemical (ATC) Classification System. We counted the number of different severe ADRs per drug as mapped in SIDER. Aggregated across the ATC classes, we identified classes with high variability among drugs in terms of the number of associated severe ADRs (Figure 5). The median number of severe ADRs per therapeutic class is positively correlated to the fraction of drugs having FDA box warnings for that category (ρ=.64, *P*<.001). We highlight two classes that display high numbers of associated severe ADRs (median ≥5 severe ADRs) and large variability (SD>9). Immunosuppressants include drugs associated with only two (azathioprine) or three (leflunomide) severe ADRs as well as drugs associated with a high number of severe ADRs (lenalidomide associated with 29 and tacrolimus with 19). The severe ADRs that are associated with the highest number of immunosuppressants are necrosis, renal failure, and congestive cardiac failure. Anti-Parkinson drugs include drugs associated with no severe ADR (orphenadrine, biperiden, and benztropine) as well as drugs associated with multiple severe ADRs (pramipexole, associated with 36 and ropinirole, associated with 23). The most common severe ADRs among anti-Parkinson drugs are cardiac arrest, coma, renal failure, skin cancer, and cerebral ischemia.

Figure 5. Severity of ATC classes. Box plot of ATC class severity measured by number of severe ADRs in each class (severe defined by top 95 percentile of the ranks) and percentage of drugs with black box warning in that class. Only classes that include more than 2 drugs with ADR information and have at least more than 3 severe ADRs are displayed.



Ranking Can Prioritize Predicted Adverse Drug Reactions

A recent study predicted drug-ADR associations using a statistical analysis of AERS (438,801 drug-ADR pairs [14]). An experimental validation of such a large set of predictions is impractical, and yet regulators need to prioritize such associations for further investigation. Since ADR frequencies cannot be determined from the spontaneous reporting in the AERS, we highlight severe ADRs occurring in highly prescribed drugs, for which even low frequencies would have significant impact. We focused on the top 100 prescribed drugs in 2013 [18], and identified 53 drugs that have novel severe ADRs in the OFFSIDES databases (Multimedia Appendix 6). We highlight two drugs with the largest number of novel severe ADRs. First, pulmonary embolism is listed on the drug label raloxifene, an elective estrogen receptor modulator. OFFSIDES database predicts an associated ADR, pulmonary thrombosis, and additionally carotid artery and cerebral thrombosis. Second, celecoxib, a non-steroidal anti-inflammatory drug, is associated with myocardial infarction, acute coronary syndrome, and transient ischemic attack, which are a result of the predicted coronary artery occlusion and cardiomyopathy. Furthermore, the predicted cardiomyopathy often leads to the reported cardiac failure.

Triaging Genes Associated With Adverse Drug Reactions

During drug development it is useful to identify genes and pathways that are associated with ADRs; it may be even more useful to quantitatively compare these using our severity ranking. Accordingly, we used gene-ADR associations assembled from literature [15] and used the integrated ADR severity for each gene to score its “adverse reaction risk”.

Multimedia Appendix 7 lists the genes and their most severe associated ADR. Our analysis highlights previously associated genes and severe ADRs [15], as shown in Table 2. We also highlight three genes predicted by [11] to be associated with neuroleptic malignant syndrome (NMS): HTR2A, NGPR, and COMT.

Table 2. Genes reported to be associated with severe adverse drug reactions (ADRs) (top 10 percentiles).

Gene	ADR (Percentile)	Reference
EGFR	Glioblastoma multiforme (95)	[24]
SCN1A	Epilepsy (93)	[25]
VDR	Chronic renal failure (91)	[26]
TNF	Multiple sclerosis (91)	[27]
RYR1	Malignant hyperthermia (90)	[28]

Discussion

Principal Results

We ranked the severity of 2929 ADRs using a crowdsourcing platform. This ranking helps highlight drug classes based on the severity of their associated ADRs, triage predicted drug-associated ADRs for further investigation, and associate genes with a severity score based on their association with ADRs, with some implications for drug design. Although our ranking is consistent and reproducible, we cannot claim that it is optimal. A broader sampling of the potential ADR space (perhaps including professionals and patients who have experienced these effects) or a more sophisticated ranking method might improve the quality of the ranking. We include the raw pairwise comparison data (Multimedia Appendix 2) and our rankings in order to enable alternative analyses.

Limitations

Our ranking is based on a non-expert and inexperienced understanding and interpretation of ADR severity. Our analysis includes both point events and interval events, and these were compared without (1) reference to their different time courses, or (2) variations in severity between different instances of the same ADR—the MTurk workers were simply asked to decide if one ADR was better or worse than another, integrating all considerations. The high performance on the quality control ADR pairs (marked in Multimedia Appendix 2) and the consistency of the ranking shows that they generally comprehended the medical terms (possibly through using the provided Google query links). The average completion time of a survey (comprised of 10 pairwise comparisons) was 5.33 minutes, which is higher than the average time (4 minutes) required for a biomedical scientist to complete such surveys. Last, we included only workers from the United States, but our method may be biased by other demographic traits of the workers. While we have no access to such information in our study, we estimate from other sources that the average age is 33-35 years old and 63%-72% females (19).

As mentioned above, we identified some ADRs with discordance between our estimated severity and their mortality rate in the AERS reports. There are two reasons for such discrepancies: (1) a misunderstanding by laymen of the true severity of an ADR (eg, the word “cancer” may get a high ranking, regardless of its survival statistics), and/or (2) a bias in the associated death rates in the AERS system. We are unable to distinguish these, and it is likely that both contribute, highlighting areas for potential improvement.

There is no correlation between the outcome rates of disability, required intervention to prevent permanent impairment/damage, or congenital anomaly to our ADR ranking. After manual examination of the ADRs with high rates for these three types of outcomes, we identified that for the first two, disability and “required intervention” outcomes, a lack of context caused ADRs with high rates to be classified as mild. For example, grimacing or rectal cramps are associated with more than 55% disability rate, and may be frequent disability co-occurring ADRs. Similarly for “required intervention”, light anesthesia (>42% rate) and hyposmia (>25% rate) are moderate without context. In the case of congenital anomaly, many of the anomalies are not life threatening and thus were ranked low (eg, supernumerary nipple, low set ears, or ear malformation).

Finally, we used the list of ADRs appearing in SIDER and the FDA AERS systems “as-is”. Some of the ADRs in our list may not be directly caused by drugs but are associated with drugs (eg, infections may be more frequent as a side effect of the drug, or may simply co-occur with diseases that the drug treats). We retained these ADRs, as they provide important insight regarding how individuals perceive their relative severity.

Implications

We highlight drug therapeutic classes that display large variability between their drug members in terms of occurrences of severe ADRs, suggesting staying vigilant in regard to the effect of drug choice on ADR occurrence in patients. We also highlight genes associated with severe ADRs, which should be subject for further investigations.

Among the potential applications for a ranked list of ADRs, we suggest that mapping these ADRs to drug-drug interactions could aid in reducing “alert fatigue” stemming from too frequent alerts, which often emerge on relatively mild events. This phenomenon may cause physicians to dismiss these alerts and could possibly be attenuated if the alerts focused mostly on major adverse event [29]. Certain ADRs deviate more than others in rank, suggesting that their perceived severity is more of a personal preference (Multimedia Appendix 4). This information could identify cases where patient preferences should be weighted more strongly when making a prescribing decision.

Finally, we focused on the severity of ADRs, but ADR frequency is also crucial for assessment of drug risk. These ADR frequencies require proper control to correct for background frequencies. Carefully constructed clinical trials that allow extracting statistically significant frequencies in a rigorous way should be given high priority.

Conclusions

We believe that our ranking of ADRs may have useful clinical and financial implications, and should be further investigated in the context of health care decision making.

Acknowledgments

We would like to thank Nir Ailon for helpful suggestions for the linear programming, Steve Bagley for supplying the LAERS files, and the thousands of Mechanical Turk workers. Funding for RBA and AG was provided by NIH LM05652, GM102365, and GM61374. MD is supported by NIH U54 HG004028. The study was approved by the Institutional Review Board of Stanford.

Conflicts of Interest

None declared.

Multimedia Appendix 1

An example of a comparison presented to an MTurk worker.

[PNG File, 248KB - jmir_v17i3e80_app1.png]

Multimedia Appendix 2

Table S1. The MTurk workers pairwise comparisons used to compute the ranking.

[XLSX File (Microsoft Excel File), 1MB - jmir_v17i3e80_app2.xlsx]

Multimedia Appendix 3

Supplementary methods, figures, and Multimedia Appendix legends.

[PDF File (Adobe PDF File), 1MB - jmir_v17i3e80_app3.pdf]

Multimedia Appendix 4

Table S2. Ranked list of ADRs with their reported frequency.

[XLSX File (Microsoft Excel File), 148KB - jmir_v17i3e80_app4.xlsx]

Multimedia Appendix 5

Correlation between ADR semantic similarity and mean difference in severity scores, computed for 793 ADRs.

[PNG File, 29KB - jmir_v17i3e80_app5.png]

Multimedia Appendix 6

Table S3. Top prescribed drug in 2013 that have novel severe ADRs in OFFSIDES database.

[XLSX File (Microsoft Excel File), 309KB - jmir_v17i3e80_app6.xlsx]

Multimedia Appendix 7

Table S4. Genes and their most severe associated ADRs.

[XLSX File (Microsoft Excel File), 22KB - jmir_v17i3e80_app7.xlsx]

References

1. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. JAMA 1995 Jul 5;274(1):29-34. [Medline: 7791255]
2. Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. JAMA 1997;277(4):301-306. [Medline: 9002492]

3. Shaffer ML, Watterberg KL. Joint distribution approaches to simultaneously quantifying benefit and risk. BMC Med Res Methodol 2006;6:48 [FREE Full text] [doi: 10.1186/1471-2288-6-48] [Medline: 17038184]
4. Guo JJ, Pandey S, Doyle J, Bian B, Lis Y, Raisch DW. A review of quantitative risk-benefit methodologies for assessing drug safety and efficacy-report of the ISPOR risk-benefit management working group. Value Health 2010 Aug;13(5):657-666. [doi: 10.1111/j.1524-4733.2010.00725.x] [Medline: 20412543]
5. Holden WL, Juhaeri J, Dai W. Benefit-risk analysis: examples using quantitative methods. Pharmacoepidemiol Drug Saf 2003 Dec;12(8):693-697. [doi: 10.1002/pds.794] [Medline: 14762986]
6. Tallarida R, Murray RB, Eiben C. A scale for assessing the severity of diseases and adverse drug reactions. Application to drug benefit and risk. Clin Pharmacol Ther 1979 Apr;25(4):381-390. [Medline: 428184]
7. Tallarida RJ, Smith TM, Johnson AF, Blodgett RC. Non-physicians and physicians assess severity of disease states and adverse drug reactions : Application to drug benefit-risk measurement. Pharmaceutical Medicine 1984;1(1):41-46.
8. Buhrmester M, Kwang T, Gosling SD. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? Perspectives on Psychological Science 2011 Feb 03;6(1):3-5. [doi: 10.1177/1745691610393980]
9. Atias N, Sharan R. An algorithmic framework for predicting side effects of drugs. J Comput Biol 2011 Mar;18(3):207-218. [doi: 10.1089/cmb.2010.0255] [Medline: 21385029]
10. Huang L, Wu X, Chen JY. Predicting adverse side effects of drugs. BMC Genomics 2011 Dec 23;12 Suppl 5:S11 [FREE Full text] [doi: 10.1186/1471-2164-12-S5-S11] [Medline: 22369493]
11. Gottlieb A, Altman RB. Integrating systems biology sources illuminates drug action. Clin Pharmacol Ther 2014 Jun;95(6):663-669. [doi: 10.1038/clpt.2014.51] [Medline: 24577151]
12. Kesselheim AS, Cresswell K, Phansalkar S, Bates DW, Sheikh A. Clinical decision support systems could be modified to reduce 'alert fatigue' while still minimizing the risk of litigation. Health Aff (Millwood) 2011 Dec;30(12):2310-2317 [FREE Full text] [doi: 10.1377/hlthaff.2010.1111] [Medline: 22147858]
13. Kuhn M, Campillos M, Letunic I, Jensen LJ, Bork P. A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol 2010;6:343 [FREE Full text] [doi: 10.1038/msb.2009.98] [Medline: 20087340]
14. Tatonetti N, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. Sci Transl Med 2012 Mar 14;4(125):125ra31 [FREE Full text] [doi: 10.1126/scitranslmed.3003377] [Medline: 22422992]
15. Kuhn M, Al BM, Campillos M, Jensen LJ, Gross C, Gavin AC, et al. Systematic identification of proteins that elicit drug side effects. Mol Syst Biol 2013;9:663 [FREE Full text] [doi: 10.1038/msb.2013.10] [Medline: 23632385]
16. Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). Drug Saf 1999 Feb;20(2):109-117. [Medline: 10082069]
17. The Adverse Event Reporting System (AERS): Older Quarterly Data Files. 2014. URL: http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/ucm083765.htm [accessed 2014-10-21] [WebCite Cache ID 6TUzsQGjw]
18. Brooks M. Medscape. 2014 Jan 30. Top 100 selling drugs of 2013 URL: http://www.medscape.com/viewarticle/820011 [accessed 2014-10-21] [WebCite Cache ID 6TV0IQ419]
19. Köhler S, Doelken SC, Mungall CJ, Bauer S, Firth HV, Bailleul-Forestier I, et al. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. Nucleic Acids Res 2014 Jan;42(Database issue):D966-D974 [FREE Full text] [doi: 10.1093/nar/gkt1026] [Medline: 24217912]
20. Harispe S, Ranwez S, Janaqi S, Montmain J. The semantic measures library and toolkit: fast computation of semantic similarity and relatedness using biomedical ontologies. Bioinformatics 2014 Mar 1;30(5):740-742 [FREE Full text] [doi: 10.1093/bioinformatics/btt581] [Medline: 24108186]
21. IBM Corp. 2010. IBM ILOG AMPL Version 12.2 User's Guide URL: http://ampl.com/BOOKLETS/amplcplex122userguide.pdf [accessed 2015-03-11] [WebCite Cache ID 6Wwyajqyy]
22. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) 1996;58(1):267-288 [FREE Full text]
23. Rubio C, Hill ME, Milan S, O'Brien ME, Cunningham D. Idiopathic pneumonia syndrome after high-dose chemotherapy for relapsed Hodgkin's disease. Br J Cancer 1997;75(7):1044-1048 [FREE Full text] [Medline: 9083341]
24. Penar PL, Khoshyomn S, Bhushan A, Tritton TR. Inhibition of epidermal growth factor receptor-associated tyrosine kinase blocks glioblastoma invasion of the brain. Neurosurgery 1997 Jan;40(1):141-151. [Medline: 8971836]
25. Ogiwara I, Miyamoto H, Morita N, Atapour N, Mazaki E, Inoue I, et al. Nav1.1 localizes to axons of parvalbumin-positive inhibitory interneurons: a circuit basis for epileptic seizures in mice carrying an Scn1a gene mutation. J Neurosci 2007 May 30;27(22):5903-5914 [FREE Full text] [doi: 10.1523/JNEUROSCI.5270-06.2007] [Medline: 17537961]
26. Ross EA, Tian J, Abboud H, Hippensteel R, Melnick JZ, Pradhan RS, et al. Oral paricalcitol for the treatment of secondary hyperparathyroidism in patients on hemodialysis or peritoneal dialysis. Am J Nephrol 2008;28(1):97-106. [doi: 10.1159/000109398] [Medline: 17914251]
27. Jurewicz A, Matysiak M, Tybor K, Kilianek L, Raine CS, Selmaj K. Tumour necrosis factor-induced death of adult human oligodendrocytes is mediated by apoptosis inducing factor. Brain 2005 Nov;128(Pt 11):2675-2688 [FREE Full text] [doi: 10.1093/brain/awh627] [Medline: 16219674]

28. Zhao F, Li P, Chen SR, Louis CF, Fruen BR. Dantrolene inhibition of ryanodine receptor Ca2+ release channels. Molecular mechanism and isoform selectivity. J Biol Chem 2001 Apr 27;276(17):13810-13816 [FREE Full text] [doi: 10.1074/jbc.M006104200] [Medline: 11278295]
29. Phansalkar S, van der Sijs H, Tucker AD, Desai AA, Bell DS, Teich JM, et al. Drug-drug interactions that should be non-interruptive in order to reduce alert fatigue in electronic health records. J Am Med Inform Assoc 2013 May 1;20(3):489-493 [FREE Full text] [doi: 10.1136/amiajnl-2012-001089] [Medline: 23011124]

Abbreviations

ADR: adverse drug reaction
AERS: Adverse Events Reporting System
ATC: Anatomical Therapeutic Chemical Classification System
FDA: Food and Drug Administration
HIT: human intelligence tasks
HIV: human immunodeficiency virus
HPO: Human Phenotype Ontology
MedDRA: medical dictionary for regulatory activities
MTurk: Amazon Mechanical Turk

Edited by G Eysenbach; submitted 21.10.14; peer-reviewed by K Johnson; comments to author 07.01.15; revised version received 17.01.15; accepted 04.02.15; published 23.03.15.

Please cite as:
Gottlieb A, Hoehndorf R, Dumontier M, Altman RB
Ranking Adverse Drug Reactions With Crowdsourcing
J Med Internet Res 2015;17(3):e80
URL: http://www.jmir.org/2015/3/e80/
doi:10.2196/jmir.3962
PMID:25800813

Original Paper

Patient Use of Email for Health Care Communication Purposes Across 14 European Countries: An Analysis of Users According to Demographic and Health-Related Factors

Nikki Newhouse[1,2], BA(Hons); Francisco Lupiáñez-Villanueva[3,4], PhD; Cristiano Codagnone[4,5], PhD; Helen Atherton[1], PhD

[1]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom
[2]UCL Interaction Centre, University College London, London, United Kingdom
[3]Department of Information and Communication Science, Universitat Oberta de Catalunya, Barcelona, Spain
[4]Applied Social Science and Behavioural Economics Research Group, Universitat Oberta de Catalunya, Barcelona, Spain
[5]Dipartimento di Scienze Sociali e Politiche, Universitá degli Studi di Milano, Milan, Italy

Corresponding Author:
Nikki Newhouse, BA(Hons)
Nuffield Department of Primary Care Health Sciences
University of Oxford
Gibson Building
Radcliffe Observatory Quarter, Woodstock Road
Oxford, OX2 6GG
United Kingdom
Phone: 44 (0)1865617770
Fax: 44 (0)1865289412
Email: nicola.newhouse@phc.ox.ac.uk

Abstract

Background: The use of the Internet for health purposes is growing steadily, yet the use of asynchronous communication tools for health care purposes remains undeveloped. The introduction of email as a method of communication in health care has the potential to impact on both patients and health care professionals.

Objective: This study aims to describe the characteristics of people who have sent or received an email to or from their doctor, nurse, or health care organization, by country and in relation to demographics, health care resource use, and health status factors.

Methods: We conducted a secondary analysis of data (N=14,000) collected from the online Citizens and Information Communication Technology for Health survey, a project undertaken in 2011 by the Institute for Prospective Technology Studies of the European Commission’s Joint Research Centre. The survey was developed to understand and characterize European citizens’ use of information communication technologies for health. Descriptive and statistical analyses of association were used to interpret the data.

Results: Denmark reported the highest level of emails sent/received (507/1000, 50.70%). The lowest level reported was by participants in France (187/1000, 18.70%). Men used email communication for health care more than women, as did respondents in the 16-24 age group and those educated to tertiary level or still within the education system. As self-reported health state worsens, the proportion of people reporting having sent or received an email within the context of health care increases. Email use, poor health, multimorbidity, and number of visits to a physician are positively correlated.

Conclusions: The use of email communication within the context of European health care is extremely varied. The relationship between high email use, poor health, doctor visits, and multimorbidity is especially pertinent: provision of asynchronous communication for such groups is favored by policymakers. Low reported email use by country may not necessarily reflect low interest in using email for health care: local health policies and technical infrastructures may be significant factors in the delay in implementation of alternative forms of routine health communication.

(J Med Internet Res 2015;17(3):e58) doi:10.2196/jmir.3700

KEYWORDS

eHealth; patient-doctor communication; electronic mail; Internet; Europe; chronic illness; patient activation; health services; survey

Introduction

Background

The use of the Internet for health purposes is growing steadily as people increasingly go online to access factual and experiential information and to share their own health and illness experiences [1-3]. While health-related Internet use has increased steadily, the use of asynchronous communication tools for health care purposes among European populations remains relatively undeveloped and unexplored. Asynchronous communication refers to interaction that is non-concurrent, such as email, as opposed to the synchronous, real-time communication offered by phone consultation, for example. Email is a commonly used method of communication globally. It has become a major element of day-to-day life for many people, both at work and in their personal lives. However, the uptake in the health care sector has been much lower than for other sectors, and email is not routinely used as a way for patients to contact their health care organization or professional [4,5] despite a reportedly high interest in associated eHealth services among European citizens [6]. Various national policies encourage email use by patients in the health care setting [7-11], for example in Denmark, where patients are able to engage in electronic communication with their family doctor via the official Danish health website [12]. In the United States, health maintenance organizations such as Kaiser Permanente have embraced email for communicating with their patients, offering patient portals where patients can log in and use Web messaging to send an email to their health care professional, make appointments, and receive test results [13,14]. At present, relatively little is known about how patients use email to interact with health care, but this information is likely to be valuable in determining the success of proposed policies and impact of email use. Key variables that may influence the uptake of email communication between patient and health professional include patient health status, the "digital divide", and the potential impact on the use of health care services.

Health Status

A recent study of the effect of asynchronous communication between health care providers and chronically ill patients found positive effects on health behavior, health outcomes, and patient satisfaction [15]. Indeed, it is well established that health status is directly connected to certain health-related online behaviors. Although individuals with chronic or multimorbid health conditions are less likely to have access to the Internet and to be less able to use it, evidence suggests that they are more likely to be engaged with the self-management of their own health [16,17]. Individuals with chronic and multimorbid health conditions are more likely to gather and share information about their health and to follow it up by seeking advice from a health professional, friend, family member, or peer [2]. The number of people living with chronic illnesses is increasing exponentially, and they are a primary target group for policy makers. They are also the group who have perhaps the most to gain from asynchronous communication opportunities due to their need for frequent contact with health care professionals.

The perception by health care professionals is that the open channel of communication offered by email may become congested by the "worried well", those who value convenient access but have a substantially lesser need for this access than those suffering from multimorbidities [18]. This has the potential to impact on both utilization of health resources and overall clinician workload [19] and may be a key factor in clinicians' reluctance to engage with asynchronous communication.

The Digital Divide

The digital divide in the context of using email for health care communication refers to the opportunity gap that emerges when a service is available to those with access to the Internet and online services, in this case email, and unavailable to those who do not have this access. However, the resulting divide is nuanced beyond *having* and *not having* access: even among regular Internet users there is a recognized division in relation to *how* people use the Internet and for what purposes, for instance in how frequently they can or want to access the Internet and their level of computer literacy [20]. Although evidence is mixed, certain groups are deemed to be at an immediate disadvantage: older people, those with lower levels of education, and those not in regular employment [21,22].

As stated above, those with multimorbidities may benefit from having an additional communication route. But as multimorbidity and chronic illness may be more likely to affect older age groups, it is unclear whether this group could adequately benefit since older people are perceived to be less likely to want to engage with the Internet for health care purposes [23]. At present, little is known about the impact on this divide of introducing email as a method of communication in health care, despite this information having a potential impact on how health care is distributed.

Health Care Utilization

There is the possibility that introducing an additional channel for communication in health care may lead to an increased number of contacts with health care services. Where email provides a communication directly with the health care professional, there is mixed evidence on how introducing it impacts on numbers of contacts by patients with their health care providers. It is widely stated by policy makers that introduction of these alternative methods of communication will reduce other types of contacts, and there is some evidence to support this [24]. However, there is also evidence showing that introduction of email for communicating with a health care professional increases the overall number of contacts by patients across other methods of communication [25]. At present, little is known about the behaviors of patients using email for their health care, in relation to visits made. This information is important in relation to planning of workload and funding, as well as determining whether the service brings what it intends to.

Aim of the Study

The introduction of email as a method of communication in health care, be it between a patient and health care professional, or between patient and health care organization, has the potential to impact greatly on both patients and health care professionals by influencing the use of health systems, the relationship between patient and doctor, and the ways in which people manage their own health. An evidence-based understanding of current practice and trends in this area is crucial as we transition towards a future in which patients' remote access to health care and encouragement of large-scale self-management will be policy priorities for many countries. This study describes the characteristics of an Internet-using population who have sent or received an email to or from their doctor, nurse, or health care organization, by country and in relation to demographic and health care resource use and health status factors.

Methods

Survey Instruments and Ethics

We conducted a secondary analysis of data collected from the *Citizens and Information Communication Technology for Health* survey, a project undertaken in 2011 by the Institute for Prospective Technology Studies of the European Commission's Joint Research Centre. This online survey was developed from a theoretical framework of the social determinants of information and communication technology for health, translated into native languages in 14 European Union member countries. The survey was developed to understand and characterize European citizens' use of information and communication technology for health. Technical, methodological, and legal considerations were carefully addressed in the context of designing and implementing the survey. These considerations ensured anonymity and confidentiality of individual responses [26-28]. The survey was conducted in accordance with European Society for Opinion and Marketing Research ethical guidelines [29]. At the time the survey was carried out, 2 of the authors (FLV, CC) were employed by the European Commission.

The questionnaire was structured in 5 blocks—Block A: Health status and health care and social care services use; Block B: Health attitude and Health information sources; Block C: Internet and Information and Communication Technologies uses; Block D: Health-related use of Information and Communication Technologies and the Internet; and Block E: Sociodemographic profile of participants.

Sample and Data Collection

The target population was citizens aged 16-74 years old who had used the Internet in the previous 3 months. The survey was conducted online in Austria, Belgium, Germany, Denmark, Estonia, Finland, France, Italy, Netherlands, Sweden, Slovenia, Slovakia, Spain, and the United Kingdom with a proportional allocation of 1000 interviews per country. A random sample was used, with quotas for gender and age (16-24, 25-54, 55-74) to ensure a representative sample of participants. Table 1 summarizes the sampling information.

Table 1. Sampling information.

Population	Citizens aged from 16-74 years old who had used the Internet in the previous 3 months.
Geographical coverage	Austria, Belgium, Germany, Denmark, Estonia, Finland, France, Italy, Netherlands, Sweden, Slovenia, Slovakia, Spain, United Kingdom
Sample size	1000 interviews per country; 14,000 interviews in total
Quotas	Country; Gender (Female/Male); Age Group (16-24, 25-54, 55-74)
Sampling errors	+0.85% for overall data and +3.16% for country-specific data. In all cases, a maximum indeterminate probability (P=q=50), for a confidence level of 95.5% is applicable for each one of the reference populations.
Weighting	Proportional allocation for each country, to be able interpret the data at a country level; Weighting by population in each country to be able to interpret the overall data
Sampling	Individuals have been sampled in a completely random manner.

Table 2 summarizes the main sociodemographic characteristics. These results are broadly comparable to the characteristics of the Internet population in each country [30]. In order to interpret the overall data, country-specific differences have to be accounted for. The weighting factor was calculated by dividing the proportion of each country's population to its total population by the proportion of individuals in each country's sample to the total sample.

The main survey variable we were interested in was "Regarding health and Information and Communication Technologies, specifically the Internet, how often have you sent or received an email from your doctor, nurse or health care organization?" Participants were asked how often they had done this (every day/almost every day, at least once a week, at least every month, less than once a month, never, or I was not aware of it), and we split responses to this question into yes and no categories, with all the "never" and "I was not aware of it" responses comprising no, and the other groups merged to comprise yes. This allowed us to compare use and non-use of email with other key variables of interest. No definition of the term "email" was included in the survey. It is assumed that respondents understood the term in accordance with its use in common parlance.

Table 2. Sample sociodemographic characteristics (N=14,000).

		n (%)
Gender		
	Female[a]	7210 (51.50)
	Male[a]	6789 (48.50)
Age		
	16-24	2777 (19.84)
	25-54	8708 (62.20)
	55-74	2515 (18.96)
Completed education		
	Primary or lower secondary education (ISCED[b] 0, 1, or 2)	2128 (15.20)
	Upper secondary education (ISCED[b] 3 or 4)	6439 (45.99)
	Tertiary education (ISCED[b] 5 or 6)	5433 (38.81)
Employment status		
	Employed or self-employed	8189 (58.49)
	Unemployed	1335 (9.54)
	Student (not in labor force)	2007 (14.34)
	Other (not in labor force)	2469 (17.64)

[a] n=13999; 1 do not know/did not answer.

[b] UNESCO International Standard.

Statistical Analysis

Data analyses were completed using SPSS version 20.0. Chi-square tests were used to determine if there were any differences between use and non-use of email. An analysis of residuals was performed to determine the sources of significant findings. Under the null hypothesis that the 2 variables are independent, the adjusted residuals will have a standard normal distribution, that is, have a mean of 0 and standard deviation of 1. An adjusted residual that is more than 1.96 (2.0 is used by convention) indicates that the number of cases in that cell is significantly larger than would be expected if the null hypothesis were true, with a significant level of .05. An adjusted residual that is less than -2.0 indicates that the number of cases in that cell is significantly smaller than would be expected if the null hypothesis were true.

Results

Characteristics of Participants

Just over a quarter of participants (25.38%, 3553/14000) reported sending or receiving an email from their doctor, nurse or health care organization (Table 3). Participants largely reported good health: health status "good" or "very good" (74.29%, 10400/14000). Only 25.57% (3580/14000) reported not having a health problem; 58.11% (7849/13506) reported not experiencing any long-standing illness or health problem, and 65.87% (7849/13863) reported that they were not undergoing long-term medical treatment.

Table 3. Characteristics of participants.

Characteristics	n (%)
Assessment of own health status	
Very bad	131 (0.94)
Bad	891 (6.36)
Neither good or bad	2578 (18.41)
Good	7521 (53.72)
Very good	2879 (20.56)
Number of health problems reported	
None	3580 (25.57)
1	3992 (28.51)
2	3011 (21.51)
More than 2	3417 (24.41)
Current long-standing illness or health problem [a]	
Yes	5657 (41.89)
No	7849 (58.11)
Undergoing long term medical treatment [b]	
Yes	4732 (34.13)
No	9131 (65.87)
Number of visits to the doctor in the last 12 months [c]	
None	1544 (11.03)
1-2	4278 (30.56)
3-4	2945 (21.04)
5-6	2202 (15.73)
More than 6	3030 (21.64)
Sent or received an email from your doctor, nurse, or health care organization	
Yes	3553 (25.38)
No	10447 (74.62)

[a] n=13,506; 494 do not know/did not answer.

[b] n=13,863; 137 do not know/did not answer.

[c] n=13,999; 1 do not know/did not answer.

Country and Email Use

Of those reporting sending or receiving an email from their doctor, nurse, or health care organization, there was a statistically significant difference among countries: Denmark reported the highest level of emails sent/received at 50.70% (507/1000). The lowest level reported was by participants in France at 18.70% (187/1000). Respondents from Denmark, Estonia, Italy, and Sweden are more likely to use email within the context of health care than those in France, Belgium, Spain, Slovakia, Slovenia, and United Kingdom (Table 4).

Table 4. Sent/received an email from your doctor, nurse, or health care organization, by country.

Country[a,b]	Sent/received an email, n (%)	Adjusted residual
Austria	283 (28.30)	0.7
Belgium	222 (22.20)	-3.8
Germany	284 (28.40)	0.7
Denmark	507 (50.70)	17.1
Estonia	316 (31.60)	3.1
Finland	297 (29.70)	1.7
France	187 (18.70)	-6.4
Italy	363 (36.30)	6.5
Netherlands	253 (25.30)	-1.6
Spain	247 (24.70)	-2.0
Sweden	308 (30.80)	2.5
Slovakia	187 (18.70)	-6.4
Slovenia	198 (19.80)	-5.6
United Kingdom	186 (18.60)	-6.5

[a] Proportional allocation for each country (N=1000).

[b] χ^2_{13}=494.359; *P*=.000.

Demographic Characteristics and Email Use

More men than women had used email (29.11%, 2099/7210) versus 21.42%, 1454/6789). Highest use was reported in the 16-24 age group (30.00%, 833/2777) and in those educated to tertiary level (27.00%, 1467/5433). Lowest use was reported in the 55-74 age group (20.16%, 507/2515) and in those educated to primary or lower secondary level (22.98%, 489/2128). Students reported the highest level of use (28.95%, 581/2077), with the lowest levels reported by those in the “Other” group, that is, people outside of the labor force (20.66%, 510/2468) (Table 5).

Table 5. Sent/received an email from your doctor, nurse, or health care organization, by demographic characteristic

Characteristics		Sent/received an email	
		Yes, n (%) adjusted residual	No, n (%) adjusted residual
Gender [a]			
	Female	1454 (21.42) -10.5	5335 (78.58) 10.5
	Male	2099 (29.11) 10.5	5111 (70.89) -10.5
Age [b]			
	16-24	833 (30.00) 6.2	1944 (70.00) -6.2
	25-54	2213 (25.41) 0.1	6495 (74.59) -0.1
	55-74	507 (20.16) -6.6	2008 (79.84) 6.6
Completed education [c]			
	Primary or lower secondary education (ISCED[d] 0, 1, or 2)	489 (22.98) -2.8	1639 (77.02) 2.8
	Upper secondary education (ISCED[d] 3 or 4)	1598 (24.81) -1.4	4842 (75.19) 1.4
	Tertiary education (ISCED[d] 5 or 6)	1467 (27.00) 3.5	3966 (73.00) -3.5
Employment status [e]			
	Employed or self employed	2184 (26.67) 4.2	6005 (73.33) -4.2
	Unemployed	278 (20.82) -4.0	1057 (79.18) 4.0
	Student (not in labor force)	581 (28.95) 4.0	1426 (71.05) -4.0
	Other (not in labor force)	510 (20.66) -5.9	1958 (79.34) 5.9

[a] χ^2_1=109.332; *P*=.000.

[b] χ^2_2=67.455; *P*=.000.

[c] χ^2_2=15.109; *P*=.001.

[d] UNESCO International Standard.

[e] χ^2_3=64.299; *P*=.000.

Health Status, Health Resource Utilization, and Email Use

The highest level of email use is reported in those who state that their general health is very bad (40.46%, 53/131). Respondents with more than 2 health problems also report the highest level of email use (33.63%, 1149/3417), indicating that the poorer a person's health, the more likely they are to have used email in this way. As self-reported health state worsens, the proportion of people reporting having sent or received an email increases (Table 6).

As seen in Table 6, the relationship between email use for health care and number of visits to the doctor was varied. Over 60% of respondents who reported visiting the doctor 5-6 times and more than 6 times in the last 12 months had also communicated with their health care provider by email. Those who reported having more than 6 visits in the last 12 months reported the highest use of email (30.33%, 919/3030). Those who reported not visiting the doctor at all in the preceding 12 months reported the lowest level of email use (15.54%, 240/1544). As visit number increases, so does the proportion of people reporting having sent or received an email and vice versa.

Table 6. Sent/received an email from your doctor, nurse, or health care organization, by health status and health care utilization.

		Sent/received an email	
		Yes, n (%) adjusted residual	No, n (%) adjusted residual
Assessment of own health status [a]			
	Very bad	53 (40.46) 4.0	78 (59.54) -4.0
	Bad	292 (32.77) 5.2	599 (67.23) -5.2
	Neither good or bad	680 (26.38) 1.3	1898 (73.62) -1.3
	Good	1863 (24.77) -1.8	5658 (75.23) 1.8
	Very good	665 (23.10) -3.2	2214 (76.90) 3.2
Number of health problems reported [b]			
	None	761 (21.25) -6.6	2820 (78.75) 6.6
	1	890 (22.29) -5.3	3103 (77.71) 5.3
	2	754 (25.04) -0.5	2257 (74.96) 0.5
	More than 2	1149 (33.63) 12.7	2268 (66.37) -12.7
Current long-standing illness or health problem [c]			
	Yes	1665 (29.82) 9.0	3991 (70.56) -9.0
	No	1771 (22.56) -9.0	6078 (77.44) 9.0
Undergoing long term medical treatment [d]			
	Yes	1411 (29.82) 8.9	3321 (70.18) -8.9
	No	2090 (22.89) -8.9	7040 (77.11) 8.9
Number of visits to the doctor in the last 12 months [e]			
	None	240 (15.54) -9.4	1304 (84.46) 9.4
	1-2	1003 (23.44) -3.5	3276 (76.56) 3.5
	3-4	736 (24.99) -0.6	2209 (75.01) 0.6
	5-6	656 (29.78) 5.2	1547 (70.22) -5.2
	More than 6	919 (30.33) 7.1	2111 (69.67) -7.1

[a] χ^2_4=52.178; *P*=.000.

[b] χ^2_3=175.235; *P*=.000.

[c] χ^2_1=81.893; *P*=.000.

[d] χ^2_1=79.214; *P*=.000.

[e] χ^2_4=149.294; *P*=.000.

Discussion

Summary

This study provides the first Europe-wide exploration of email use by patients for health care communication purposes. The most prolific users of email were men, those aged 16-24, and students. Higher numbers of people in poor health used email relative to those reporting good health and no ongoing conditions.

Country and Email Use

The high level of email communication reported in Denmark is consistent with their health policy. It is compulsory for all doctors in Danish primary care services to offer their patients email contact and online services, and structures are in place to provide reimbursement for this use. The Danish public national health portal, Sundhed.dk, has been established for more than a decade and successfully integrates a variety of features with the explicit aim of both facilitating the smooth delivery of national health care priorities and maintaining Denmark's reputation as a world leader in the digital health arena [12]. The high levels of use in Denmark are not typical of European implementation. The level of use reported in this study for the United Kingdom (18.6) is similar to prevalence estimates in other UK-based surveys that currently estimate use in general practice settings at between 20-23% [31]. The use of email by country is hugely varied and indicative of hugely disparate eHealth communication priorities and strategies among the European countries surveyed.

Demographic Characteristics and Email Use

Highest use of email for health care was reported among the youngest and most educated groups, and lowest use was reported among the oldest, least educated, and unemployed respondents. These findings corroborate the general perception that certain groups are deemed to be at an immediate digital disadvantage and may be disenfranchized as a result. The higher use of email by men in the sample is a case in point and is a particularly interesting finding as it contests the comprehensive perception that women are more likely to use the Internet for health care purposes. As eHealth research becomes increasingly nuanced, evidence of gender differences in the context of particular health-related online behaviors becomes apparent. The perceived importance of online health information as a credible resource is indeed particularly strong among young women [1] and yet, in this study, significantly more men used the opportunity to actively communicate with a health care professional by email.

Health Status, Heath Resource Utilization, and Email

Concerns that opportunities for email communication encourage inappropriate use by the "worried well" were unfounded among this sample. Health status was consistently negatively associated with email use, with over 40% of those who rated their health as being very bad having used email to communicate with a health professional, compared to just over 23% of those who rated their health as being very good. This may be due to increased need for contact with health care services among those people with poor health and possibly reflects email being used as an alternative to other forms of contact. There is also the possibility that it reflects a desire by those with multimorbidity to have repeated and frequent contact with their health care professional. People living with a chronic condition who have access to the Internet are significantly more likely than other online adults to gather and act upon health-related behavior [2].

Strengths and Limitations

This research provides key insights into the use of email for health care communication in Europe, particularly in the context of establishing impact on health care resource use. The specific value of email communication to specific groups of people is clear, despite overall use of email for health care communication being relatively low. It should be noted that low use of email does not necessarily indicate low interest. Research exploring expectations of the future use of digital health services consistently highlights substantial interest among European citizens in using email for health care but also stresses that this interest comes with reservations, largely around the ability of local technical infrastructure to cope, patient confidentiality, and a negative perception of clinicians' willingness to use alternative forms of communication [32]. Indeed, in countries where the implementation of a coherent digital health strategy has been particularly slow (eg, Poland), interest remains high but expectation has fallen sharply [6].

These data feature only responses from digitally literate individuals. So although we are able to describe email use only among this population of Internet users, we know that across the European countries included, some 73% of people are online [30] and so the survey covers a large proportion of the population. Coupled with the quota sampling approach, the sample is as representative as can be expected for an online survey. The age range of participants was 16-74 years, thus people in the very oldest age groups were excluded. As these are also likely to be lowest users of the Internet by age group, it is possible that the impact of the digital divide is underestimated in this study.

There is likely to be some confounding of the results because we were looking at so many different factors, and it is important to view these results within the scope of what is possible to ascertain through secondary analysis of questionnaire data. Potential confounds may not be related to demographics or may be related to variables not addressed by the questionnaire at all. Without further research, it is unwise to extrapolate. Sample size means that it is not surprising that there was a significant difference between groups; they may reflect differences between groups, but not necessarily meaningful differences. The variable of interest was not designed as a yes/no question in the survey, and so the results must be viewed in light of this. The "no" includes both people actively deciding not to use it, and those who were not aware of it. The "yes" group includes people having used it at varying frequencies. The two groups in each category may be very different. The variable itself includes communication with a doctor, nurse, or health care organization. Respondents' understanding of the term "email" may have differed. In addition, this contact may be disparate, with emails to a doctor or nurse being different in nature than those to an organization. However, it is presumed that intention to use the technology for communication relating to health care would be the same, as would associated processes.

Conclusions

We can conclude that the use of email for communication in the context of health care is of particular value to specific groups of patients despite relatively low use. Additionally, low use does not necessarily indicate a lack of interest or willingness to engage with health care in this way. Increased use of email is not associated with increased visits to a physician among the "worried well" but is associated with increased overall virtual and physical engagement among those with chronic and multimorbid conditions. Qualitative studies need to be conducted in order to develop our understanding of this phenomenon. Recognizing and understanding the nuances of email communication is crucial in ensuring that any use of email in health care is equitable. It is important to develop guidance around best practice in the use of email, and there are clear opportunities for communities and countries to learn from each other's success. Implemented carefully, email communication could become an important tool for health care professionals, which may allow aspects of consultation to move beyond traditional settings.

Acknowledgments

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

We would like to thank those who participated in the survey, as well as Ioannis Maghiros and Fabienne Abadie for their work on the original survey. Helen Atherton is funded by a National Institute for Health Research (NIHR) School for Primary Care Research (SPCR) fellowship.

Conflicts of Interest

None declared.

References

1. Fox S, Duggan M. http://www.pewinternet.org/. Health Online 2013 URL: http://www.pewinternet.org/2013/01/15/health-online-2013/Archived [WebCite Cache ID 6QwHHlKiu]
2. Fox S, Duggan M. Pew Internet Research. The Diagnosis Difference URL: http://www.pewinternet.org/files/old-media/Files/Reports/2013/PewResearch_DiagnosisDifference.pdf [accessed 2014-07-09] [WebCite Cache ID 6QwHu07mH]
3. O'Neill B, Ziebland S, Valderas J, Lupiáñez-Villanueva F. User-generated online health content: a survey of Internet users in the United Kingdom. J Med Internet Res 2014;16(4):e118 [FREE Full text] [doi: 10.2196/jmir.3187] [Medline: 24784798]
4. Liddell A, Adshead S, Burgess E. Technology in the NHS: Transforming the patient's experience of care. 2008. URL: http://www.kingsfund.org.uk/sites/files/kf/Technology-in-the-NHS-Transforming-patients-experience-of-care-Liddell-Adshead-and-Burgess-Kings-Fund-October-2008_0.pdf [accessed 2014-07-09] [WebCite Cache ID 6QwI4nKkJ]
5. Dixon A. Engaging patients in their health: how the NHS needs to change. 2008. URL: http://www.kingsfund.org.uk/sites/files/kf/Engaging-Patients-in-their-Health-How-the-NHS-needs-to-change-Anna-Dixon-Kings-Fund-October-2008.pdf [accessed 2014-07-09] [WebCite Cache ID 6QwIBWWDQ]
6. Santana S, Lausen B, Bujnowska-Fedak M, Chronaki C, Kummervold PE, Rasmussen J, et al. Online communication between doctors and patients in Europe: status and perspectives. J Med Internet Res 2010;12(2):e20 [FREE Full text] [doi: 10.2196/jmir.1281] [Medline: 20551011]
7. Finnish Ministry of Social Affairs and Health. eHealth Roadmap. 2007. URL: http://ehealth-strategies.eu/database/documents/Finland_eHealth-ERA_country_report.pdf [WebCite Cache ID 6QwIPEXWV]
8. The Netherlands Ministry of Public Health Welfare and Support. National policy document on health: Health close to people. URL: http://www.government.nl/files/documents-and-publications/leaflets/2012/05/11/health-close-to-people/health-close-to-people.pdf [WebCite Cache ID 6QwIb1Prw]
9. Australian Government. Improving Primary Health Care for All Australians. 2011. URL: http://www.yourhealth.gov.au/internet/yourHealth/publishing.nsf/Content/improving-primary-health-care-for-all-australians-toc/$FILE/Improving%20Primary%20Health%20Care%20for%20all%20Australians.pdf [accessed 2014-07-11] [WebCite Cache ID 6QzJpem7z]
10. The Department of Health. The power of information: Putting all of us in control of the health and care information we need. 2012. URL: https://www.gov.uk/government/publications/giving-people-control-of-the-health-and-care-information-they-need [accessed 2014-07-11] [WebCite Cache ID 6QzJxQpAd]
11. Health Canada. Healthcare System: eHealth. 2010. URL: http://www.hc-sc.gc.ca/hcs-sss/ehealth-esante/index-eng.php [accessed 2014-07-11] [WebCite Cache ID 6QzKLnkz7]
12. Danish Ministry of Health. eHealth in Denmark: eHealth as a part of a coherent Danish health care system. 2012. URL: http://www.sum.dk/~/media/Filer%20-%20Publikationer_i_pdf/2012/Sundheds-IT/Sundheds_IT_juni_web.ashx [accessed 2014-07-11] [WebCite Cache ID 6QzKX4N9T]
13. Silvestre AL, Sue VM, Allen JY. If you build it, will they come? The Kaiser Permanente model of online health care. Health Aff (Millwood) 2009;28(2):334-344 [FREE Full text] [doi: 10.1377/hlthaff.28.2.334] [Medline: 19275988]
14. Hassol A, Walker JM, Kidder D, Rokita K, Young D, Pierdon S, et al. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. J Am Med Inform Assoc 2004;11(6):505-513 [FREE Full text] [doi: 10.1197/jamia.M1593] [Medline: 15299001]
15. de Jong CC, Ros WJ, Schrijvers G. The effects on health behavior and health outcomes of Internet-based asynchronous communication between health providers and patients with a chronic condition: a systematic review. J Med Internet Res 2014;16(1):e19 [FREE Full text] [doi: 10.2196/jmir.3000] [Medline: 24434570]
16. Ayers SL, Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. Health (London) 2007 Jul;11(3):327-347. [doi: 10.1177/1363459307077547] [Medline: 17606698]
17. Bansil P, Keenan NL, Zlot AI, Gilliland JC. Preventing Chronic Disease. Health-related information on the Web: results from the HealthStyles Survey, 2002-2003 URL: http://www.cdc.gov/pcd/issues/2006/apr/05_0155.htm [accessed 2014-07-11] [WebCite Cache ID 6QzLWzzz5]

18. Goodyear-Smith F, Wearn A, Everts H, Huggard P, Halliwell J. Pandora's electronic box: GPs reflect upon email communication with their patients. Inform Prim Care 2005;13(3):195-202. [Medline: 16259859]
19. Hanna L, May C, Fairhurst K. The place of information and communication technology-mediated consultations in primary care: GPs' perspectives. URL: http://fampra.oxfordjournals.org/content/29/3/361.full.pdf [accessed 2014-07-11] [WebCite Cache ID 6QzLnZtsR]
20. Eisenberg MB. Journal of Library & Information Technology. Information Literacy: Essential Skills for the Information Age URL: http://ifl.sagepub.com/content/39/3/262.full.pdf+html [accessed 2014-07-11] [WebCite Cache ID 6QzME78oj]
21. Sequist TD. J Clin Outcomes Manage. Increasing Consumer Use of the Internet for Health Information URL: http://www.turner-white.com/memberfile.php?PubCode=jcom_feb06_internet.pdf [accessed 2014-07-11] [WebCite Cache ID 6QzMIpF0J]
22. Car J, Sheikh A. Email consultations in health care: 1--scope and effectiveness. BMJ 2004 Aug 21;329(7463):435-438 [FREE Full text] [doi: 10.1136/bmj.329.7463.435] [Medline: 15321902]
23. Roland M, Paddison C. Better management of patients with multimorbidity. BMJ 2013;346:f2510. [Medline: 23641032]
24. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
25. 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]
26. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002;4(2):E13 [FREE Full text] [doi: 10.2196/jmir.4.2.e13] [Medline: 12554560]
27. Dillman D, Smyth JD, Christian LM. Internet, mail, and mixed-mode surveys: the tailored design method. Hoboken, NJ: Wiley & Sons; 2009.
28. Braithwaite D, Emery J, De LS, Sutton S. Using the Internet to conduct surveys of health professionals: a valid alternative? Fam Pract 2003 Oct;20(5):545-551 [FREE Full text] [Medline: 14507796]
29. ESOMAR Guideline for Online Research. 2011. URL: http://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR_Guideline-for-online-research.pdf [accessed 2014-07-11] [WebCite Cache ID 6QzNZztxG]
30. Internet use in households and by individuals in 2011. URL: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-11-066/EN/KS-SF-11-066-EN.PDF [accessed 2014-07-11] [WebCite Cache ID 6QzNdCdwA]
31. Whitefield L. Netscape navigators. 2013. URL: http://www.ehi.co.uk/insight/analysis/1038 [accessed 2014-07-11] [WebCite Cache ID 6QzOF8W43]
32. Atherton H, Pappas Y, Heneghan C, Murray E. Experiences of using email for general practice consultations: a qualitative study. Br J Gen Pract 2013 Nov;63(616):e760-e767 [FREE Full text] [doi: 10.3399/bjgp13X674440] [Medline: 24267859]

Edited by G Eysenbach; submitted 11.07.14; peer-reviewed by P Kummervold, J Kronenfeld; comments to author 25.11.14; revised version received 17.12.14; accepted 09.01.15; published 06.03.15.

Please cite as:
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
URL: http://www.jmir.org/2015/3/e58/
doi:10.2196/jmir.3700
PMID:25798912

Original Paper

Mapping Publication Trends and Identifying Hot Spots of Research on Internet Health Information Seeking Behavior: A Quantitative and Co-Word Biclustering Analysis

Fan Li[1], MS; Min Li[2], B Med; Peng Guan[3], PhD; Shuang Ma[4], MMed; Lei Cui[1], MS

[1]Department of Medical Informatics, China Medical University, Shenyang, China
[2]Heping Hospital, Qiqihar, China
[3]Department of Epidemiology, School of Public Health, China Medical University, Shenyang, China
[4]Department of Neurology, The Affiliated Shengjing Hospital of China Medical University, Shenyang, China

Corresponding Author:
Lei Cui, MS
Department of Medical Informatics
China Medical University
No.77 Puhe Road, Shenyang North New Area
Shenyang, 110122
China
Phone: 86 24 31939518
Fax: 86 24 31939517
Email: cmuinfo@163.com

Abstract

Background: The Internet has become an established source of health information for people seeking health information. In recent years, research on the health information seeking behavior of Internet users has become an increasingly important scholarly focus. However, there have been no long-term bibliometric studies to date on Internet health information seeking behavior.

Objective: The purpose of this study was to map publication trends and explore research hot spots of Internet health information seeking behavior.

Methods: A bibliometric analysis based on PubMed was conducted to investigate the publication trends of research on Internet health information seeking behavior. For the included publications, the annual publication number, the distribution of countries, authors, languages, journals, and annual distribution of highly frequent major MeSH (Medical Subject Headings) terms were determined. Furthermore, co-word biclustering analysis of highly frequent major MeSH terms was utilized to detect the hot spots in this field.

Results: A total of 533 publications were included. The research output was gradually increasing. There were five authors who published four or more articles individually. A total of 271 included publications (50.8%) were written by authors from the United States, and 516 of the 533 articles (96.8%) were published in English. The eight most active journals published 34.1% (182/533) of the publications on this topic. Ten research hot spots were found: (1) behavior of Internet health information seeking about HIV infection or sexually transmitted diseases, (2) Internet health information seeking behavior of students, (3) behavior of Internet health information seeking via mobile phone and its apps, (4) physicians' utilization of Internet medical resources, (5) utilization of social media by parents, (6) Internet health information seeking behavior of patients with cancer (mainly breast cancer), (7) trust in or satisfaction with Web-based health information by consumers, (8) interaction between Internet utilization and physician-patient communication or relationship, (9) preference and computer literacy of people using search engines or other Web-based systems, and (10) attitude of people (especially adolescents) when seeking health information via the Internet.

Conclusions: The 10 major research hot spots could provide some hints for researchers when launching new projects. The output of research on Internet health information seeking behavior is gradually increasing. Compared to the United States, the relatively small number of publications indexed by PubMed from other developed and developing countries indicates to some extent that the field might be still underdeveloped in many countries. More studies on Internet health information seeking behavior could give some references for health information providers.

(J Med Internet Res 2015;17(3):e81) doi:10.2196/jmir.3326

KEYWORDS

information seeking behavior; Internet; health information; bibliometric analysis; co-word analysis; biclustering; hot spots; publication status

Introduction

In recent decades, Internet technology has developed rapidly and has become an important part of the daily lives of many people around the world. Meanwhile, interest in the medium as a communication tool for health-related information is growing rapidly [1-4].

The Internet has huge potential to meet the health information needs and enhance the health literacy of people because of its abundant resources, convenient access, low cost, interactivity, continuing evolution, and so on. The proliferation of the medium has, it can be argued, changed the way that people use information to protect their health [5]. To launch global health promotion campaigns and meet the increasingly urgent health information needs of everyone, more and more institutions including governments, academic organizations, medical and educational departments, and business corporations have established health information portals, with the result that Web-based health information gets richer and richer. Accordingly, people increasingly seek health information via the Internet, the fastest growing carrier of health information and the largest medical library in the world. As early as 2004, it was determined that about 4.5% of Internet searches in the world were related to health [6]. As the Internet develops, however, there is far more data than before. The latest (2013) national survey by the Pew Internet & American Life Project shows that clinicians are still the top source of health information in the United States, but Web-based information, curated by peers, is a significant supplement. Among US adults, 81% use the Internet, 59% say they have looked online for health information in the past year, and 35% say they have gone online specifically to try to figure out what medical condition they or someone else might have [7]. The statistical results indicated that more and more Internet users, not satisfied with health information provided by medical professionals, search for health information on the Internet and also enjoy the efficiency and convenience of the Internet.

In academia, health information seeking behavior is defined by scholars as “the behavior that the public or consumers exhibit when acquiring health information” [8]. As people increasingly use the Internet to seek health information in recent years, theoretical and empirical studies on Internet health information seeking behavior have been conducted by researchers. Findings of such studies are valuable for practice, such as (1) construction of all types of health or medical information websites, (2) optimization of search language of Web-based health information, (3) development of tools enhancing the users’ ability to read professional health content, (4) formulation of health information dissemination strategy aimed at specific populations, and (5) improvement of the efficiency and quality of Web-based health information services [9].

Despite the fact that scholars have published papers on Internet health information seeking behavior worldwide, there have been only several reviews about the research advancements in this field. In a short review of literature by Younger, for example, some of the existing evidence was collected from 1995 to 2009 to establish whether there were any significant differences in the ways and reasons why doctors and nurses sought out Web-based information and to establish how nurses and doctors located information online [10]. A literature review by the European Centre for Disease Prevention and Control concentrated on research articles published in English on Internet health information seeking behavior by adults from 2006 to 2010. It mainly documented Internet accessibility and usage patterns, outlined Internet health information consumer profiles, identified Internet sources of health information, outlined health professionals’ Internet use, and ascertained challenges for health professionals posed by Internet use [11].

Although the above reviews can reflect the research status on Internet health information seeking behavior to some extent, the contents and the points of view are quite different among different scholars and the bibliometric research across a long-term span is lacking. The bibliometric method, a type of quantitative analysis, has been widely used for the determination of scientific research evolution in recent years. The statistical indicators that measure the contribution of scientific publications within a given topic or research field can represent the research trends and hot spots [12]. A research hot spot refers to a focus of research for which researchers have carried out many studies and published many related papers. By computing the frequencies and relationship of words reflecting the content of articles that appear in a field, the hot spots of the field can usually be identified [13,14]. Co-word analysis is a type of bibliometric method to identify hot spots and find knowledge in academic literature, proposed as early as the late 1970s by French bibliometric scientists [15]. Its principle is as follows: if two professional terms expressing a particular research subject appear in the same articles simultaneously, these two terms may have a certain intrinsic relationship. The more frequently these two terms occur in the same articles, the closer their relationship is. According to this “distance”, the important keywords of a subject are classified further to sum up the research focus and structure of a discipline by modern statistical techniques, such as cluster analysis, factor analysis, multidimensional scaling analysis, or multivariate analysis. It is well known that cluster analysis has been widely used to extract research themes of a field. For example, with clustering algorithms, Raghupathi et al identified the sub-fields of research of health information systems [16], and Schuemie et al characterized the domain of medical informatics [17]. Biclustering, unlike traditional clustering, allows simultaneous clustering of the rows and columns of a matrix, not only to cluster the global information, but also to find local information efficiently in high dimensional data [18]. In 1972, Hartigan first proposed the idea that simultaneous clustering could be performed for the rows and columns of a matrix [19]. Until 2000, Cheng and Church formally presented the algorithm and model of biclustering [18].

Since then, more and more excellent algorithms and models of biclustering have been developed. In recent years, biclustering analysis was introduced to the bibliometric field. Cui et al applied biclustering to analyze Chinese education status on medical informatics from the two aspects of institutes and themes [20]. Yu et al revealed the research subject areas and hot spots in biomedical informatics by biclustering analysis [21]. Fang et al performed biclustering to explore high producing authors and research features of library science and informatics in China [22]. Their research results showed that major research hot spots and representative publications or studies in one subject area could be captured with the biclustering method.

To the authors' knowledge, at present, there have been few bibliometric articles on Internet health information seeking behavior. In the present study, a comprehensive analysis on outer characteristics and content patterns of relevant publications was carried out to reveal the research history and status in this field. Specifically, co-word biclustering analysis was utilized to identify the research hot spots of Internet health information seeking behavior. We hope this paper can provide some reference for future research on Internet health information seeking behavior.

Methods

Data Collection

Relevant articles were identified by searching PubMed without the restriction of language or publication year. PubMed was selected as the data source for two reasons: (1) PubMed is a free authoritative medical literature database of the National Library of Medicine on the Web, with articles on health and medical information seeking behaviors, and (2) articles from Medline (a subset of PubMed) are indexed with MeSH (Medical Subject Headings) terms, a set of normalized words that can reflect the contents of articles; based on those words, the co-word clustering analysis can be performed. The search was carried out on September 30, 2014 to ensure the search results were as current as possible. The search strategy is detailed as follows (for details on how PubMed translates this search strategy, see Multimedia Appendix 1): #1 Internet information seeking behavior; #2 Internet information seeking behaviour; #3 ((search*[ti] OR seek*[ti]) OR (behaviour[ti] OR behavior[ti])) AND (Internet[ti] OR net*[ti] OR online[ti] OR web*[ti]); #4 #1 OR #2 OR #3.

Based on the above search strategy, 2741 publications were found in PubMed. The titles and abstracts of the publications were screened according to relevance and selection criteria. The inclusion criteria were (1) the contents of papers primarily focus on Internet health information seeking behavior, and (2) all study designs. The exclusion criteria were (1) studies focus on accuracy or validity of information resources rather than Internet health information seeking behavior (eg, the article entitled "Infant teething information on the World Wide Web: taking a byte out of the search" has been excluded, because the purpose of this study was to describe and evaluate the quality of infant teething information on selected popular parenting websites [23]), (2) studies on other behaviors, such as suicidal behavior [24], Internet-based problem shopping behavior [25], Internet addictive behavior [26], regular gaming behavior [27], and so on, (3) studies on seeking other information unrelated to health, and (4) studies on other networks such as Bayesian networks [28] and metabolic networks [29] rather than the Internet. Two researchers independently reviewed and evaluated the studies and reached consensus on the inclusion for analysis. The concordance rate between them was 0.90, indicating a strong agreement [30]. Any discrepancies were discussed with reference to the research objective until consensus was reached. Initially, 494 related articles were identified. Then, we counted the number of related articles in each journal. It was found that the top three journals were, in order, the Journal of Medical Internet Research, the Journal of Health Communication, and the Journal of the Medical Library Association: JMLA, for which manual searches were also conducted and another 39 related papers were identified. Finally, a total of 533 papers were included in this study. Each publication downloaded from PubMed contained the following key eligibility items: title, author, institution, country, source, publication year, and MeSH terms. These data were saved as two files in XML and MEDLINE formats, respectively.

Data Extraction and Analysis

Bibliographic Item Co-Occurrence Matrix Builder (BICOMB) (developed by Professor Cui from China Medical University and available freely online) [31], GoPubMed [32], and Microsoft Excel were employed to determine the distribution of the publication year, countries, authors, languages, journals, and the frequency ranking of major MeSH terms of the included publications. In this study, the most active journals were identified according to Bradford's Law (Bradford's Law of Scattering was first described by Bradford in 1934). The law indicates that journals in a given subject area can be sorted by the number of articles into three parts, each with about one-third of all articles: (1) a core of a few journals, (2) a second zone, with more journals, and (3) a third zone, with the bulk of journals. The number of journals in each group will be proportional to 1: n: n^2. Although Bradford's Law is not statistically accurate, librarians commonly use it as a guideline in the research of core journals [33].

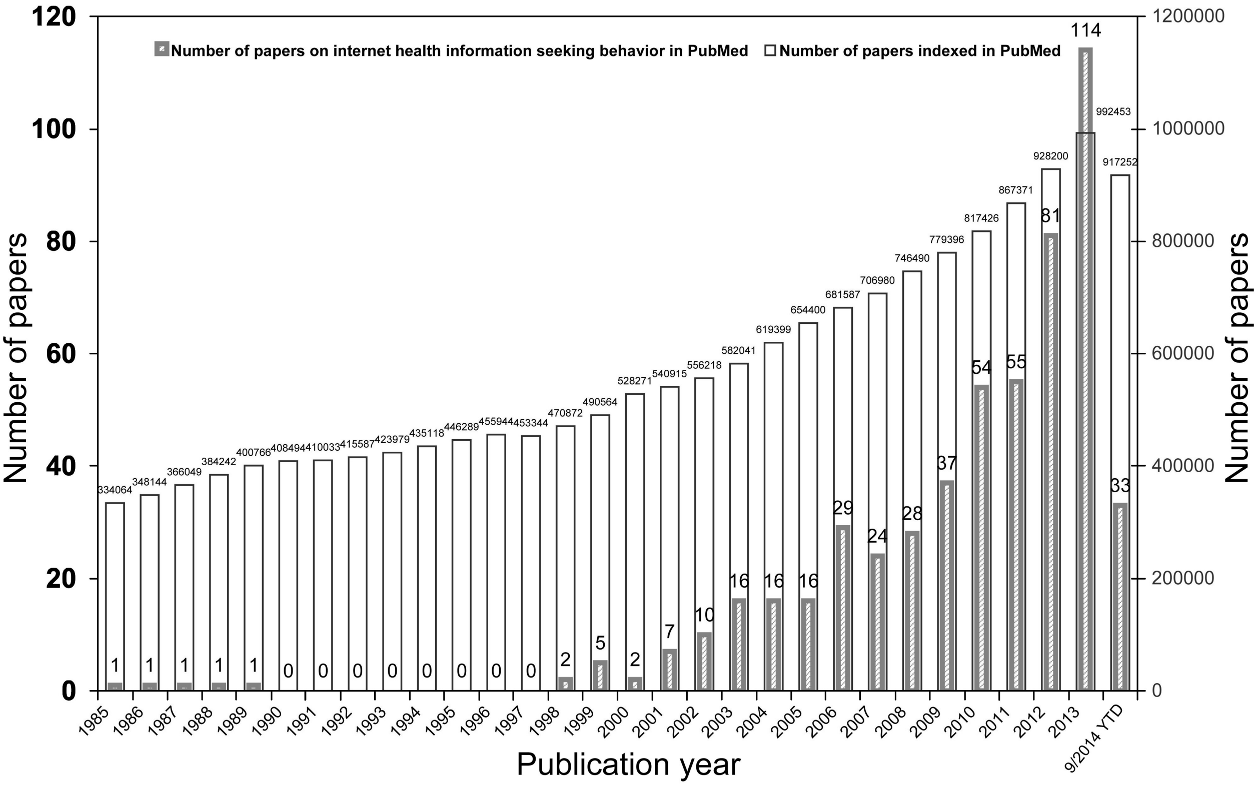
Considering the papers included in this study had to be related to Internet health information seeking behavior, the two highly frequent MeSH terms, "Internet" and "Information Seeking Behavior", were deleted for they would be meaningless in the content analysis. The trends of the remaining highly frequent major MeSH terms over time were illustrated visually. In order to explore the hot spots of Internet health information seeking behavior, biclustering for highly frequent major MeSH terms and included papers was performed. By biclustering, in this study, the relationship among highly frequent words and the relationship between highly frequent words and source articles could be gained. Subsequently, a binary matrix with highly frequent major MeSH terms as the rows and source articles as the columns was built from BICOMB for further biclustering by using the software "gCLUTO", version 1.0 (Graphical CLUstering TOolkit, a graphical front-end for the CLUTO data clustering library, developed by Rasmussen, Newman, and Karypis from University of Minnesota) [34]. Parameters of biclustering in gCLUTO were set according to those appropriate for biclustering analysis based on literature [35]. Repeated bisection was chosen for clustering method, cosine for similarity function, and *I*2 for criterion function of clustering. In the hopes of being able to distinguish the optimal number of clusters, we reran the biclustering with different numbers of clusters. The biclustering result of the matrix of highly frequent major MeSH terms-source articles was shown through mountain visualization and matrix visualization. With the help of semantic relationships among MeSH terms and the content of the representative papers in each cluster, the basic framework of research hot spots of Internet health information seeking behavior was drawn and analyzed.

Results

Growth of the Literature

Based on the search strategy and inclusion criteria, a total of 533 publications (see Multimedia Appendix 2) were included in this study. The average number of these publications per year was about 18. The first article was published in 1985 [36]. In Figure 1, the distribution of the publication year of articles on Internet health information seeking behavior is shown, and, in parallel, that of all the publications indexed in PubMed is also illustrated for the aim of comparison.

Figure 1. Temporal distribution of papers on Internet health information seeking behavior in PubMed and all papers indexed in PubMed.



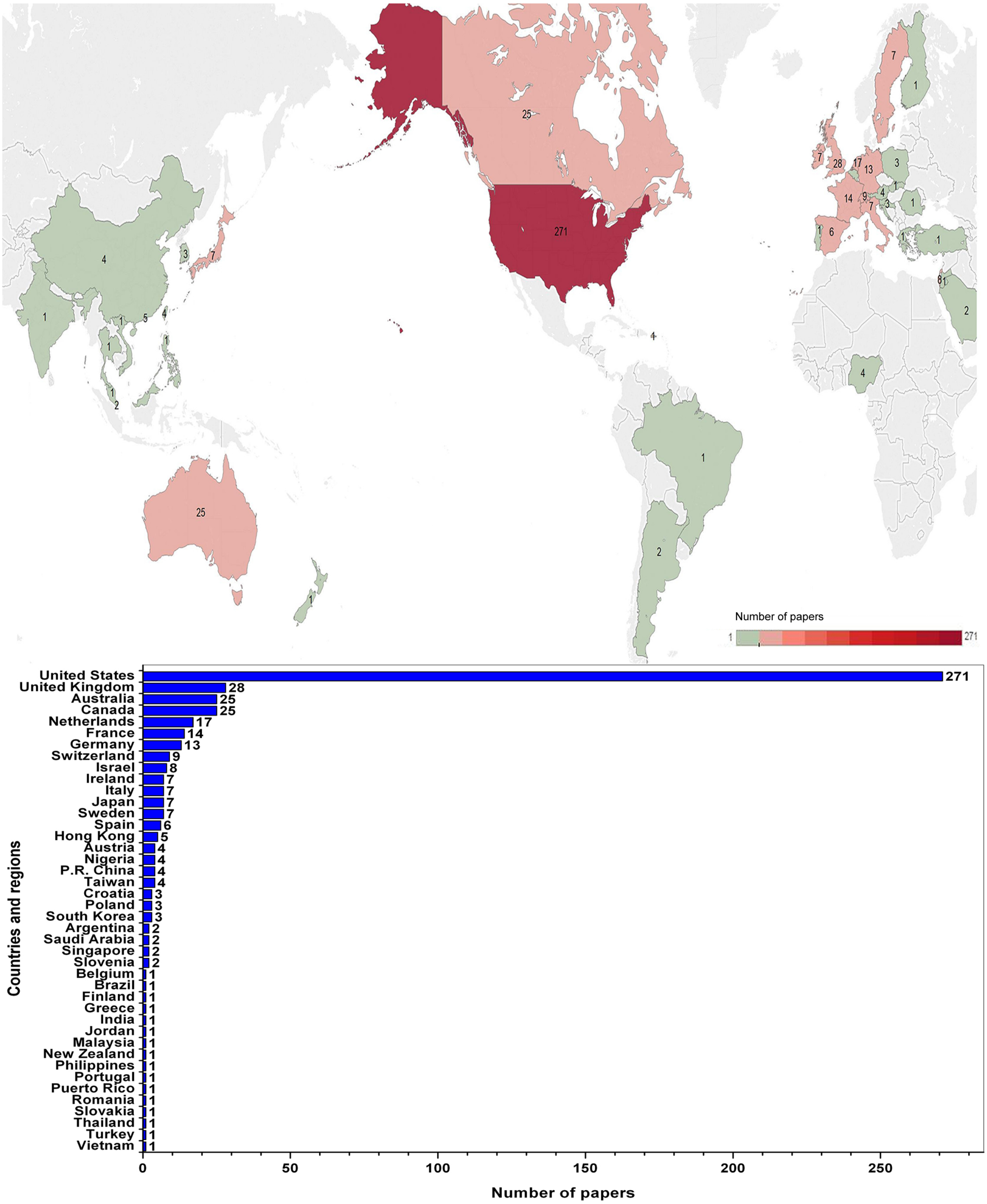
Distribution of Countries, Authors, and Languages

Of all articles on Internet health information seeking behavior from PubMed, 93.4% (498/533) provided the addresses of authors. So, according to the rough statistics, articles on Internet health information seeking behavior indexed in PubMed originated from at least 42 different countries or regions. Figure 2 illustrates the number of research outputs on Internet health information seeking behavior in different countries. The number in the map is the amount of related publications for every country or territory. The United States ranked first with 271 publications (50.8%).

Among all 1758 authors involved in this topic, five authors published four or more articles individually. Park ranked first for publishing five articles, of which he published one article as the first author. Dickerson published four articles for second place, yet he had three articles as the first author. The authors of the first publication on Internet health information seeking behavior were Tolle and Hah from the United States.

Most of these articles were published in English (96.8%, 516/533), the remaining articles were published in Spanish (1.1%, 6/533), German (0.8%, 4/533), French (0.6%, 3/533), Polish (0.4%, 2/533), Hebrew (0.2%, 1/533), Swedish (0.2%, 1/533), and Japanese (0.2%, 1/533), respectively. (Note, one article was published in both English and French [37].)

Figure 2. Geographic distribution of research output on Internet health information seeking behavior.



Most Active Journals

Altogether, 253 journals have been involved in the field. From 1985 to 2014, the eight most active journals published 182 publications on Internet health information seeking behavior, accounting for 34.1% of all 533 publications. Table 1 displays the eight most productive journals, which are considered as the core journals in the research area of Internet health information seeking behavior under Bradford's Law.

Table 1. Most active journals on the topic of Internet health information seeking behavior (PubMed sourced until September 2014) (n=533).

No.	Top journals	Publications n (%)
1	Journal of Medical Internet Research	67 (12.6)
2	Journal of Health Communication	41 (7.7)
3	Journal of the Medical Library Association: JMLA	16 (3.0)
4	American Medical Informatics Association (AMIA): Annual Symposium proceedings / AMIA Symposium. AMIA Symposium	14 (2.6)
5	CyberPsychology & Behavior: the impact of the Internet, multimedia and virtual reality on behavior and society	13 (2.4)
6	International Journal of Medical Informatics	13 (2.4)
7	PLOS ONE	9 (1.7)
8	Health Communication	9 (1.7)
Total		182 (34.1)

Research Hot Spots of Internet Health Information Seeking Behavior

For publications from 1985 to September 2014, there were 495 major MeSH terms with a cumulative frequency of 1577 times excluding the two terms, “Internet” and “Information Seeking Behavior”. After discussion, a major MeSH term with a frequency of 10 or more times’ occurrence was defined as a highly frequent one. Then, 30 highly frequent major MeSH terms were extracted from the included publications with a cumulative percentage of 45.34% (715/1577) (Table 2). The annual distribution of these MeSH terms is displayed in Figure 3, where the circle indicates there is at least one publication in the corresponding year. The bigger the circle is, the more publications there are in the corresponding year. According to the situation of co-occurrence of these highly frequent MeSH terms in the same articles, a matrix was established with highly frequent major MeSH terms as the row names and source articles as the column names. The matrix (localized view in Table 3) presents the availability of the major MeSH terms in the source articles. A “1” in the cells indicates that the major MeSH term is present in the article, while “0” means absent.

Biclustering was performed with the different numbers of clusters; the biclustering result of the matrix of highly frequent major MeSH terms-source articles is shown as mountain visualization and matrix visualization. Figure 4 illustrates the mountain visualization and the highly frequent MeSH terms in each cluster when these MeSH terms were divided into six clusters. The purpose of the mountain visualization is to visually aid the user in understanding the contents of a high-dimensional dataset and the efficacy effect of clustering. In Figure 4, each cluster is represented as a peak in the 3D terrain labeled by the cluster number (0 through 5, total six clusters in this study). A peak’s location on the plane, volume, height, and color are used to portray information about the associated cluster. The most informative attribute of a peak is its location on the plane with respect to other peaks. The distance between a pair of peaks on the plane represents the relative similarity of their clusters. The height of each peak is proportional to the cluster’s internal similarity. The volume of a peak is proportional to the number of MeSH terms contained within the cluster. Last, the color of a peak represents the internal standard deviation of a cluster’s objects. Red represents low deviation, whereas blue represents high deviation. According to the authors’ knowledge, each independent cluster should cover at least 30 publications and there should be no triplet peaks in the mountain visualization, thus these highly frequent MeSH terms were divided into six clusters. Figure 5 shows the matrix visualization, in which the row labels stand for highly frequent major MeSH terms, and the column labels are PubMed Unique Identifiers (PMIDs) of source articles, respectively on the right and the bottom of the matrix. Colors are used to graphically represent the values present in the matrix. The color of each grid represents the relative occurrence frequency of a MeSH term in an article. The increasingly darker shade of red represents larger values and the white color represents the values near zero. The rows of the primary matrix (Table 3) are reordered by biclustering, so that rows of the same cluster are aggregated; black horizontal lines separate these clusters. The matrix visualization reveals that 30 highly frequent major MeSH terms are clustered to six clusters. The left hierarchical tree depicts the relationships between highly frequent major MeSH terms, and the hierarchical tree on the top displays the relationships between articles. Also, it shows the corresponding articles in which each highly frequent MeSH term of each cluster occurs. A closer reading of the representative articles of each cluster contributed to the work of identifying and summarizing the themes of each cluster.

Moreover, some clusters could be divided into smaller topics based on the following criteria discussed by the research group: (1) the semantic relationship among the MeSH terms within one bigger cluster, (2) the year of MeSH terms being introduced in MeSH vocabulary, and (3) the categories in which the MeSH terms are located. The MeSH terms such as “Students”, “Parents”, and “Physicians” are located under the category of “Persons” or “Anthropology”. “Cellular Phone” and “Social Media” were the two newly introduced MeSH terms. Each of these smaller topics was thus summarized to a separate hot spot. Therefore, 10 hot topics in total were found in the field of Internet health information seeking behavior, as follows: (1) behavior of Internet health information seeking about HIV infection or sexually transmitted diseases (Cluster 0), (2) Internet health information seeking behavior of students (Cluster 0), (3) behavior of Internet health information seeking via mobile phone and its apps (Cluster 0), (4) physicians' utilization of Internet medical resources (Cluster 1), (5) utilization of social media by parents (Cluster 1), (6) Internet health information seeking behavior of patients with cancer (mainly breast cancer) (Cluster 2), (7) trust in or satisfaction with Web-based health information by consumers (Cluster 3), (8) interaction between Internet utilization and physician-patient communication or relationship (Cluster 3), (9) preference and computer literacy of people using search engines or other Web-based systems (Cluster 4), and (10) attitude of people (especially adolescents) when seeking health information via the Internet (Cluster 5).

Table 2. Highly frequent major MeSH[a] terms from the included publications on Internet health information seeking behavior (n=1577).

No.	MeSH terms	Frequency n (%[b])	Cumulative percentage, %
1	Consumer Health Information	77 (4.88)	4.88
2	Information Services	67 (4.25)	9.13
3	Patient Education as Topic	55 (3.49)	12.62
4	Information Storage and Retrieval	50 (3.17)	15.79
5	Health Knowledge, Attitudes, Practice	50 (3.17)	18.96
6	Health Education	44 (2.79)	21.75
7	Patient Acceptance of Health Care	35 (2.22)	23.97
8	Neoplasms	33 (2.09)	26.06
9	Information Dissemination	23 (1.46)	27.52
10	Parents	21 (1.33)	28.85
11	Attitude to Computers	18 (1.14)	29.99
12	Medical Informatics	17 (1.08)	31.07
13	Physician-Patient Relations	16 (1.01)	32.09
14	Health Behavior	16 (1.01)	33.10
15	Social Support	16 (1.01)	34.12
16	Consumer Satisfaction	15 (0.95)	35.07
17	Search Engine	15 (0.95)	36.02
18	Breast Neoplasms	14 (0.89)	36.91
19	Attitude to Health	14 (0.89)	37.79
20	Online Systems	13 (0.82)	38.62
21	Computer Literacy	12 (0.76)	39.38
22	Consumer Participation	11 (0.70)	40.08
23	Social Media	11 (0.70)	40.77
24	Adolescent Behavior	11 (0.70)	41.47
25	Trust	11 (0.70)	42.17
26	Physicians	10 (0.63)	42.80
27	Patient Satisfaction	10 (0.63)	43.44
28	HIV Infections	10 (0.63)	44.07
29	Students	10 (0.63)	44.71
30	Cellular Phone	10 (0.63)	45.34

[a] MeSH: Medical Subject Headings

[b] Proportion of the frequency among 1577 times' appearance.

Table 3. Highly frequent major MeSH[a] terms-source articles matrix (localized).

No.	Major MeSH terms	PubMed Unique Identifiers of source articles				
		10052399	10402805	10590585	…	9934530
1	Consumer Health Information	0	0	0	…	0
2	Information Services	0	0	0	…	0
3	Patient Education as Topic	1	1	1	…	0
4	Information Storage and Retrieval	0	0	0	…	1
…	…	…	…	…	…	…
29	Students	0	0	0		0
30	Cellular Phone	0	0	0	…	0

[a] MeSH: Medical Subject Headings

Figure 3. Annual distribution of highly frequent major MeSH terms from the included publications on Internet health information seeking behavior.

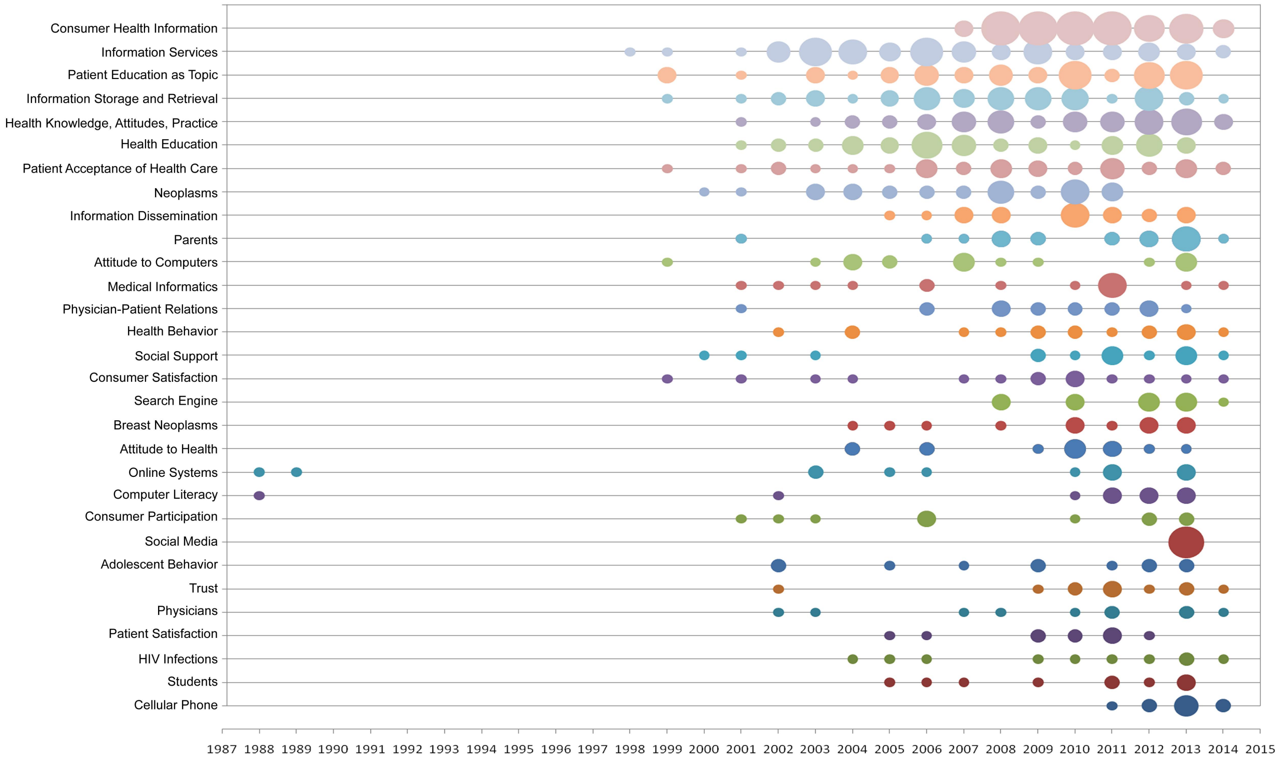
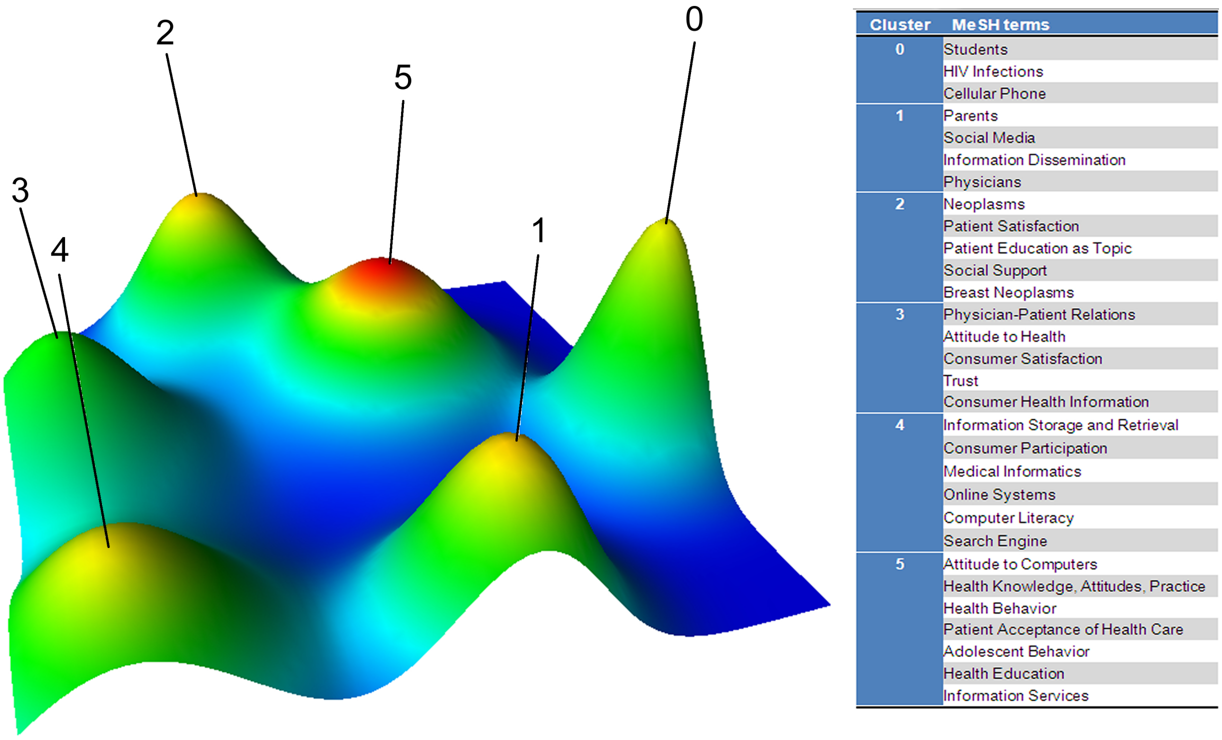
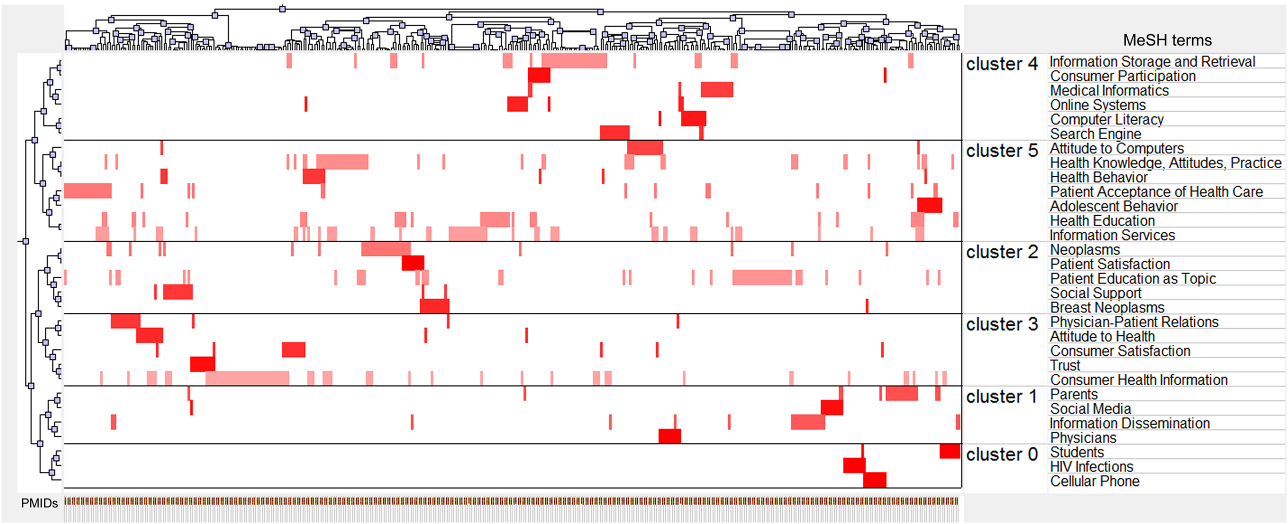


Figure 4. Mountain visualization of biclustering of highly frequent major MeSH terms and articles on Internet health information seeking behavior.



Cluster	MeSH terms
0	Students HIV Infections Cellular Phone
1	Parents Social Media Information Dissemination Physicians
2	Neoplasms Patient Satisfaction Patient Education as Topic Social Support Breast Neoplasms
3	Physician-Patient Relations Attitude to Health Consumer Satisfaction Trust Consumer Health Information
4	Information Storage and Retrieval Consumer Participation Medical Informatics Online Systems Computer Literacy Search Engine
5	Attitude to Computers Health Knowledge, Attitudes, Practice Health Behavior Patient Acceptance of Health Care Adolescent Behavior Health Education Information Services

Figure 5. Visualized matrix of biclustering of highly frequent major MeSH terms and PubMed Unique Identifiers (PMIDs) of articles on Internet health information seeking behavior.



Discussion

Principal Findings

In the context of the development of the Internet and its application in the field of health, academic communities have paid increasing attention to the Internet health information seeking behavior of people. With that in mind, the present study revealed an increase in publications and 10 hot spots of research on Internet health information seeking behavior from a quantitative and content point of view. Compared to the trend of all publications in PubMed, a more significant increase was found, suggesting that the research output increase of the Internet health information seeking behavior subfield was not proportional to the research output increase corresponding to the whole medical field. As expected, the number of publications was less before 2000 and started to increase during the period from 2001 to 2010s. The maximal number of publications (21.4%, 114/533) was found in 2013. In the publications included, the first one on Internet health information seeking behavior can be traced back to 1985; the authors were Americans, suggesting that investigation in this field may have originated in the United States. The authors of this article examined user search patterns on the CATLINE database [36]. Currently, the United States is still the top producer of papers in the field. In terms of all producer countries, the major contributions were from a small number of developed countries in North America, Europe, and Australia. This may be due to the fact that there is more awareness and a higher level of Internet health information seeking behavior research in those countries, where researchers have relative superiority in English as well. For the authors' native language of Chinese, relevant articles on the target topic were also searched in the China Academic Journals Full-Text Database, China Doctoral Dissertations Full-Text Database, China Masters' Theses Full-Text Database, and China Proceedings of Conference Full-Text Database in China National Knowledge Infrastructure (CNKI), which provides comprehensive and current Chinese information on a worldwide scale, including dissertations, proceedings, academic papers, and books from a variety of publishers, research institutions, and information sources in China. In total, 21 articles (17 journal articles, one conference article, and three graduate theses) were found up to September 2014; the first batch was published in 2008. Moreover, compared to the United States, authors from other developed or developing countries/regions (such as United Kingdom, Australia, New Zealand, and India), where English is the official language or native language, did not publish an equivalent number of related articles, so it is reasonable to infer to some extent that the field is still underdeveloped in many countries. Given the free advantage of PubMed, the results of the present study could indicate some relative preferences in the related scientific community who highly rely on PubMed for accessing professional information.

In this study, it was found that the eight most active journals had published 182 related publications (accounting for 34.1% of all related publications), and could be considered as the core journals in the research area of Internet health information seeking behavior. Among the list of the most active journals, the Journal of Medical Internet Research is the top most productive journal. Furthermore, it can be found that publications on Internet health information seeking behavior were mostly published in the journals on medical Internet research, health communication, behavior, psychology, and medical informatics, suggesting that the field crosses multiple disciplines and needs to include knowledge such as communication science, medicine, health promotion, social marketing, psychology, information technology, etc.

Methodologically, the biclustering method and visualization using gCLUTO software was employed to find hot topics on Internet health information seeking behavior in the present study. Altogether, 10 research hot spots of Internet health information seeking behavior were found, as follows:

1. Behavior of Internet health information seeking about HIV infection or sexually transmitted diseases. For instance, in Southern California, a probability sample of 195 YMSM (young men who have sex with men) using Grindr (a mobile phone app) were administered a Web-based survey to assess patterns of and motivations for Grindr use in order to inform development and tailoring of mobile phone-based HIV prevention for YMSM. The results showed that 70% of YMSM expressed a willingness to participate in a mobile phone app-based HIV prevention program. Development and testing of mobile phone apps for HIV prevention delivery has the potential to engage YMSM in HIV prevention programming, which can be tailored based on use patterns and motivations for use [38].

2. Internet health information seeking behavior of students. A cross-sectional study, for example, examined how personality traits such as sensation-seeking and impulsive decision making affected Taiwanese college students' intentions to seek Web-based information [39].

3. Behavior of Internet health information seeking via mobile phone and its apps. A case in this point is the study that people used mobile phone apps for bariatric surgery [40].

4. Physicians' utilization of Internet medical resources. For example, Shabi et al determined the extent, purpose, determinants, and the impact of the utilization of Internet medical databases among physicians [41].

5. Utilization of social media by parents. For example, in Gabbert et al's study, they studied the experiences of parents of preterm infants using social networking sites and the potential of such sites for gathering information and facilitating personal exchange [42].

6. Internet health information seeking behavior of patients with cancer (mainly breast cancer). For example, a randomized controlled experiment examined which search facility for Web-based stories resulted in the satisfaction and search success of patients with breast cancer. It was found that having access to the story topics' search facility clearly had the most positive effect on patient satisfaction and search success [43]. To better understand cancer patients' Web-based information and support seeking behaviors, another study explored how various social and psychological characteristics predicted different levels of engagement with an online breast cancer support group: posters, lurkers, and nonusers. Results showed that the patterns of engagement with the cancer support group differed according to the patients' characteristics [44].

7. Trust in or satisfaction with Web-based health information by consumers. For example, a study built on theoretical perspectives of trust such as personal capital-based, social capital-based, and transfer-based, examined various correlates of consumer trust in Web-based health information [45]. Another study examined the relationship among consumers' motivation, perceived quality, satisfaction, and intention to repeat-search eHealth information [46].

8. Interaction between Internet utilization and physician-patient communication or relationship. Provider-patient communication is an important factor influencing patients' satisfaction and health outcomes. Traditional physician-patient relationships are being challenged by the Internet health information seeking behavior of people. Hou et al drew upon the uses and gratification theory to examine how individuals' perception of communication with health care providers was associated with their Internet use for health-related activities. The researchers found that as individuals perceived their communication with providers to be less patient-centered, they were more likely to engage in various types of Web-based health activities, such as using websites for healthy lifestyles, searching for health care providers, and seeking health information [47]. Another study has examined how Internet use for health information affected the frequency of contact with health professionals [48].

9. Preference and computer literacy of people using search engines or other Web-based systems. A study investigated the impact of preference for information on the search behavior of general consumers seeking health information, their perceptions of search tasks, and user experience with search systems. Those with a high preference were found more likely to use more general queries when searching for specific factual information and to develop more complex mental representations of health concerns of an exploratory nature and try different combinations of concepts to explore these concerns. High-preference users were also more demanding on the system. Health information search systems should be tailored to fit individuals' information preferences [49].

10. Attitude of people (especially adolescents) when seeking health information via the Internet. For example, a paper explored UK and US adolescents' perceptions and experiences using the Internet to find information about health and medicines [50].

From the annual distribution of highly frequent MeSH terms, it can be found that some topics such as Internet health information seeking behavior of patients with cancer are always the research hot spots. We also found the topic on behavior of Internet health information seeking via mobile phone and its apps was quite popular since 2011. That may be the result of the rapid development of mobile Internet in recent years.

Limitations

Also, we recognize that there are several potential limitations that may encourage further research efforts. First, Figure 1 shows that there were no related articles from 1990 to 1997. While it may be objective, it may also be attributable to the fact that the included research output of our target field was only represented by the publications in a single database, PubMed, and external publications about this topic could not be analyzed. But these issues will be solved in future studies depending on the progress of text data mining techniques and more open access literature databases such as PsycINFO. We hope for the accessibility of more literature databases in the future, and then we can utilize the literature to do more studies with better quality. Second, 6.6% (35/533) of articles did not present complete affiliation data, so these articles were not assigned to a country. This might bring the bias of underestimation for some countries. Still, for identifying the country of origin, the method used was according to the corresponding author address, although often used, also not allowing recognition of transnational research. Last, although co-word biclustering is a quite useful method for identifying hot spots of one field, in this study, the results may be also affected by factors such as the accuracy of indexing documents with MeSH terms and the time when a MeSH term is introduced to the MeSH vocabulary.

Conclusions

In this study, content analysis by co-word biclustering was employed to visualize research hot spots in the field of Internet health information seeking behavior. Ten research hot spots were found, which could provide some hints for researchers when launching new projects.

Other bibliometric measures also show the research status of this field; it was found that the research output on Internet health information seeking behavior is gradually increasing. The top most productive journal is the Journal of Medical Internet Research, which is the chief journal publishing articles on Internet health information seeking behavior. In addition, compared to the United States, the relatively small number of publications indexed by PubMed from other developed and developing countries indicates to some extent that the field might be still underdeveloped in many countries. More studies on Internet health information seeking behavior could give some references for health information providers.

Acknowledgments

The authors are grateful that the study was supported in part by the National Natural Science Foundation of China (Grant No.81301930). The authors would also like to thank all reviewers for their valuable comments.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study search strategy and corresponding search details of PubMed.

[PDF File (Adobe PDF File), 41KB - jmir_v17i3e81_app1.pdf]

Multimedia Appendix 2

List of included publications on Internet health information seeking behavior in this study.

[PDF File (Adobe PDF File), 166KB - jmir_v17i3e81_app2.pdf]

References

1. Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. BMC Public Health 2007 Apr 10;7:53 [FREE Full text] [doi: 10.1186/1471-2458-7-53] [Medline: 17425798]
2. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008 Aug 25;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
3. Korp P. Health on the Internet: implications for health promotion. Health Educ Res 2006 Feb;21(1):78-86 [FREE Full text] [doi: 10.1093/her/cyh043] [Medline: 15994845]
4. Lorence D, Park H. Group disparities and health information: a study of online access for the underserved. Health Informatics J 2008 Mar;14(1):29-38. [doi: 10.1177/1460458207086332] [Medline: 18258673]
5. Costa-Font J, Mossialos E, Rudisill C. When is the Internet a valued communication device for health information in Europe? Economics of Innovation and New Technology 2009 Jul;18(5):429-445. [doi: 10.1080/10438590802547159]
6. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. Cyberpsychol Behav 2004 Oct;7(5):497-510. [Medline: 15667044]
7. Fox S, Duggan M. Pew Research Center.: Pew Internet & American Life Project; 2013. Health online 2013 URL: http://www.pewinternet.org/~/media//Files/Reports/PIP_HealthOnline.pdf [accessed 2014-02-13] [WebCite Cache ID 6NKiv0fLu]
8. Liu XL. Doc88.com. 2012 May. The research on patients' health information seeking behavior in the internet environment URL: http://www.doc88.com/p-7512993858536.html [accessed 2015-03-14] [WebCite Cache ID 6X10GiOwu]
9. Liu YJ, Lan XY. Advances in research of foreign network health information users and their inspirations. Chin J Med Libr and Inform Sci 2011;20(7):38-41. [doi: 10.3969/j.issn.1671-3982.2011.07.013]
10. Younger P. Internet-based information-seeking behaviour amongst doctors and nurses: a short review of the literature. Health Info Libr J 2010 Mar;27(1):2-10. [doi: 10.1111/j.1471-1842.2010.00883.x] [Medline: 20402799]
11. Higgins O, Sixsmith J, Barry M, Domegan C. European Centre for Disease Prevention and Control. A literature review on health information-seeking behaviour on the web: a health consumer and health professional perspective URL: http://www.ecdc.europa.eu/en/publications/Publications/Literature%20review%20on%20health%20information-seeking%20behaviour%20on%20the%20web.pdf [accessed 2014-02-13] [WebCite Cache ID 6NKjTKyWi]
12. Su H, Lee P. Mapping knowledge structure by keyword co-occurrence: a first look at journal papers in Technology Foresight. Scientometrics 2010 Jun 22;85(1):65-79. [doi: 10.1007/s11192-010-0259-8]
13. Chen HL, Zhao GQ, Xu NY. The analysis of research hotspots fronts of knowledge visualization based on CiteSpace II. In: Fong J, editor. Hybrid Learning: 5th International Conference, ICHL 2012, Guangzhou, China, August 13-15, 2012, Proceedings (Lecture Notes in Computer Science / Theoretical Computer Science and General Issues). Berlin: Springer; 2012:57-68.
14. Zhang Q. Review on the word frequency analysis is employed to discover the development of science research in China. Knowledge, Learning and Management 2011(2):95-98,128.
15. Law J, Bauin S, Courtial J, Whittaker J. Policy and the mapping of scientific change: A co-word analysis of research into environmental acidification. Scientometrics 1988 Sep;14(3-4):251-264. [doi: 10.1007/BF02020078]
16. Raghupathi W, Nerur S. Research themes and trends in health information systems. Methods Inf Med 2008;47(5):435-442. [doi: 10.3414/ME0516] [Medline: 18852917]
17. Schuemie MJ, Talmon JL, Moorman PW, Kors JA. Mapping the domain of medical informatics. Methods Inf Med 2009;48(1):76-83. [Medline: 19151887]
18. Cheng Y, Church GM. Biclustering of expression data. Proc Int Conf Intell Syst Mol Biol 2000;8:93-103. [Medline: 10977070]
19. Hartigan JA. Direct clustering of a data matrix. Journal of the American Statistical Association 1972 Mar;67(337):123-129. [doi: 10.1080/01621459.1972.10481214]
20. Cui L, Li X. A new approach to identity research focus of medical informatics departments in China. Geomatics and Information Science of Wuhan University 2010;35(Spec):49-54.
21. Yu Y, Xu ZJ, Wang K, Wang W. Text data mining in biomedical informatics based on biclustering method. Library and Information Service 2012;56(18):133-136.
22. Fang L, Cui L. Exploration on part highly-producing authors and the research features in Chinese library and information field by biclustering. In: Proceedings of the 17th Academic Conference of Chinese Medical Informatics. 2011 Presented at: 17th Academic Conference of Chinese Medical Informatics of Chinese Medical Association; 2011; Changsha, China p. 1-6.
23. Kozuch M, Peacock E, D'Auria JP. Infant teething information on the world wide web: taking a byte out of the search. J Pediatr Health Care 2015;29(1):38-45. [doi: 10.1016/j.pedhc.2014.06.006] [Medline: 25088276]
24. Hart EE, Williams CL, Davidson JA. Suicidal behaviour, social networks and psychiatric diagnosis. Soc Psychiatry Psychiatr Epidemiol 1988 Oct;23(4):222-228. [Medline: 3149025]
25. Rose S, Dhandayudham A. Towards an understanding of Internet-based problem shopping behaviour: The concept of online shopping addiction and its proposed predictors. J Behav Addict 2014 Jun;3(2):83-89 [FREE Full text] [doi: 10.1556/JBA.3.2014.003] [Medline: 25215218]

26. Critselis E, Janikian M, Paleomilitou N, Oikonomou D, Kassinopoulos M, Kormas G, et al. Internet gambling is a predictive factor of Internet addictive behavior. J Behav Addict 2013 Dec;2(4):224-230 [FREE Full text] [doi: 10.1556/JBA.2.2013.4.5] [Medline: 25215204]
27. Müller KW, Janikian M, Dreier M, Wölfling K, Beutel ME, Tzavara C, et al. Regular gaming behavior and internet gaming disorder in European adolescents: results from a cross-national representative survey of prevalence, predictors, and psychopathological correlates. Eur Child Adolesc Psychiatry 2014 Sep 5:e. [doi: 10.1007/s00787-014-0611-2] [Medline: 25189795]
28. Adabor ES, Acquaah-Mensah GK, Oduro FT. SAGA: A hybrid search algorithm for Bayesian Network structure learning of transcriptional regulatory networks. J Biomed Inform 2015 Feb;53:27-35. [doi: 10.1016/j.jbi.2014.08.010] [Medline: 25181467]
29. Kweon O, Kim SJ, Kim DW, Kim JM, Kim HL, Ahn Y, et al. Pleiotropic and epistatic behavior of a ring-hydroxylating oxygenase system in the polycyclic aromatic hydrocarbon metabolic network from Mycobacterium vanbaalenii PYR-1. J Bacteriol 2014 Oct;196(19):3503-3515. [doi: 10.1128/JB.01945-14] [Medline: 25070740]
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
31. Cui L, Liu W, Yan L, Zhang H, Hou YF, Huang YN, et al. Development of a text mining system based on the co-occurrence of bibliographic items in literature. New Technology of Library and Information Service 2008(8):70-75.
32. GoPubMed. URL: http://www.gopubmed.com/web/gopubmed/ [accessed 2014-02-13] [WebCite Cache ID 6NKf3PgrI]
33. Nash-Stewart CE, Kruesi LM, Del Mar CB. Does Bradford's Law of Scattering predict the size of the literature in Cochrane Reviews? J Med Libr Assoc 2012 Apr;100(2):135-138 [FREE Full text] [doi: 10.3163/1536-5050.100.2.013] [Medline: 22514511]
34. Karypis Lab. gCLUTO-Graphical Clustering Toolkit URL: http://glaros.dtc.umn.edu/gkhome/cluto/gcluto/download [accessed 2014-02-13] [WebCite Cache ID 6NKhEtfCT]
35. Wang L. Doc88.com. Shenyang: China Medical University; 2012. Application of biclustering on classification of scientific researchers based on literature profiling URL: http://www.doc88.com/p-9713775461927.html [accessed 2015-03-14] [WebCite Cache ID 6X0zsQx0t]
36. Tolle JE, Hah S. Online search patterns: NLM CATLINE database. J Am Soc Inf Sci 1985 Mar;36(2):82-93. [Medline: 10300015]
37. Richardson CG, Hamadani LG, Gotay C. Quantifying Canadians' use of the Internet as a source of information on behavioural risk factor modifications related to cancer prevention. Chronic Dis Inj Can 2013 Jun;33(3):123-128 [FREE Full text] [Medline: 23735451]
38. Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. AIDS Behav 2014 Feb;18(2):285-296. [doi: 10.1007/s10461-013-0671-1] [Medline: 24292281]
39. Lu HY, Palmgreen PC, Zimmerman RS, Lane DR, Alexander LJ. Personality traits as predictors of intentions to seek online information about STDs and HIV/AIDS among junior and senior college students in Taiwan. Cyberpsychol Behav 2006 Oct;9(5):577-583. [doi: 10.1089/cpb.2006.9.577] [Medline: 17034325]
40. Connor K, Brady RRW, Tulloh B, de Beaux A. Smartphone applications (apps) for bariatric surgery. Obes Surg 2013 Oct;23(10):1669-1672. [doi: 10.1007/s11695-013-0950-y] [Medline: 23609330]
41. Shabi IN, Shabi OM, Akewukereke MA, Udofia EP. Physicians utilisation of internet medical databases at the tertiary health institutions in Osun State, south west, Nigeria. Health Info Libr J 2011 Dec;28(4):313-320. [doi: 10.1111/j.1471-1842.2011.00962.x] [Medline: 22051130]
42. Gabbert TI, Metze B, Bührer C, Garten L. Use of social networking sites by parents of very low birth weight infants: experiences and the potential of a dedicated site. Eur J Pediatr 2013 Dec;172(12):1671-1677. [doi: 10.1007/s00431-013-2067-7] [Medline: 23913310]
43. Overberg R, Otten W, de Man A, Toussaint P, Westenbrink J, Zwetsloot-Schonk B. How breast cancer patients want to search for and retrieve information from stories of other patients on the internet: an online randomized controlled experiment. J Med Internet Res 2010;12(1):e7 [FREE Full text] [doi: 10.2196/jmir.1215] [Medline: 20215101]
44. Han JY, Kim JH, Yoon HJ, Shim M, McTavish FM, Gustafson DH. Social and psychological determinants of levels of engagement with an online breast cancer support group: posters, lurkers, and nonusers. J Health Commun 2012;17(3):356-371 [FREE Full text] [doi: 10.1080/10810730.2011.585696] [Medline: 22085215]
45. Ye Y. Correlates of consumer trust in online health information: findings from the health information national trends survey. J Health Commun 2011 Jan;16(1):34-49. [doi: 10.1080/10810730.2010.529491] [Medline: 21086209]
46. Lee YJ, Park J, Widdows R. Exploring antecedents of consumer satisfaction and repeated search behavior on e-health information. J Health Commun 2009 Mar;14(2):160-173. [doi: 10.1080/10810730802659830] [Medline: 19283539]
47. Hou J, Shim M. The role of provider-patient communication and trust in online sources in Internet use for health-related activities. J Health Commun 2010;15 Suppl 3:186-199. [doi: 10.1080/10810730.2010.522691] [Medline: 21154093]
48. Lee CJ. Does the internet displace health professionals? J Health Commun 2008;13(5):450-464. [doi: 10.1080/10810730802198839] [Medline: 18661387]

49. Zhang Y. The effects of preference for information on consumers' online health information search behavior. J Med Internet Res 2013;15(11):e234 [FREE Full text] [doi: 10.2196/jmir.2783] [Medline: 24284061]
50. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. Health information-seeking behaviour in adolescence: the place of the internet. Soc Sci Med 2005 Apr;60(7):1467-1478. [doi: 10.1016/j.socscimed.2004.08.010] [Medline: 15652680]

Abbreviations

BICOMB: Bibliographic Item Co-Occurrence Matrix Builder
MeSH: Medical Subject Headings
YMSM: young men who have sex with men

Edited by G Eysenbach; submitted 15.02.14; peer-reviewed by M Schuemie, B St. Jean, S Kim, J Moreland, Q Li; comments to author 29.08.14; revised version received 06.12.14; accepted 04.02.15; published 25.03.15.

Please cite as:
Li F, Li M, Guan P, Ma S, Cui L
Mapping Publication Trends and Identifying Hot Spots of Research on Internet Health Information Seeking Behavior: A Quantitative and Co-Word Biclustering Analysis
J Med Internet Res 2015;17(3):e81
URL: http://www.jmir.org/2015/3/e81/
doi:10.2196/jmir.3326
PMID:25830358

Corrigenda and Addenda

Metadata Correction: Internet Search and Krokodil in the Russian Federation: An Infoveillance Study

Andrey Zheluk[1], BAppSc, GDip(Hist), GDip(IT), MBA; Casey Quinn[2], BCom, MPhil, PhD; Peter Meylakhs[3], BA(psyc), BA(soc), PhD

[1]Menzies Centre for Health Policy, University of Sydney, Sydney, Australia
[2]PRMA Consulting Ltd, Fleet, Hampshire, United Kingdom
[3]Laboratory for Internet Studies, National Research University, Higher School of Economics, St Petersburg, Russian Federation

Corresponding Author:
Andrey Zheluk, BAppSc, GDip(Hist), GDip(IT), MBA
Menzies Centre for Health Policy
University of Sydney
D02 Victor Coppleson Building
University of Sydney, 2006
Sydney,
Australia
Phone: 61 29036522
Fax: 61 293512818
Email: andreyzheluk@gmail.com

Related Article:

Correction of: http://www.jmir.org/2014/9/e212/

(J Med Internet Res 2015;17(3):e76) doi:10.2196/jmir.4431

The authors of “Internet Search and Krokodil in the Russian Federation: An Infoveillance Study” (http://www.jmir.org/2014/9/e212/) have overlooked an error in the metadata section during the submission and proofreading process. The author Peter Meylakhs carried the affiliation "Internet Studies Lab, Higher School of Economics, St Petersburg, Russian Federation", which should have been "Laboratory for Internet Studies, National Research University Higher School of Economics, St Petersburg, Russian Federation". This error has been corrected in the online version of the paper on the JMIR website on March 20, 2015, together with publishing this correction notice. A correction notice has been sent to PubMed and the correct full-text has been resubmitted to Pubmed Central and other full-text repositories.

Edited by G Eysenbach; submitted 16.03.15; this is a non–peer-reviewed article;accepted 16.03.15; published 20.03.15.

Please cite as:
Zheluk A, Quinn C, Meylakhs P
Metadata Correction: Internet Search and Krokodil in the Russian Federation: An Infoveillance Study
J Med Internet Res 2015;17(3):e76
URL: http://www.jmir.org/2015/3/e76/
doi:10.2196/jmir.4431
PMID:30578230

Corrigenda and Addenda

Correction: Designing and Evaluating an Interactive Multimedia Web-Based Simulation for Developing Nurses' Competencies in Acute Nursing Care: Randomized Controlled Trial

Sok Ying Liaw[1], PhD; Lai Fun Wong[1], BHSc; Sally Wai-Chi Chan[2], PhD; Jasmine Tze Yin Ho[3], MHS; Siti Zubaidah Mordiffi[3], PhD; Sophia Bee Leng Ang[3], MBBS, MMed (Anaesth); Poh Sun Goh[1], MBBS, FRCR, MHPE; Emily Neo Kim Ang[1], DNurs

[1]National University of Singapore, Singapore, Singapore
[2]The University of Newcastle, Newcastle, Australia
[3]National University Hospital, Singapore, Singapore

Corresponding Author:
Sok Ying Liaw, PhD
National University of Singapore
Level 2, Clinical Research Centre, Block MD 11
Singapore, 117597
Singapore
Phone: 65 65167451
Fax: 65 67767135
Email: nurliaw@nus.edu.sg

Related Article:

Correction of: http://www.jmir.org/2015/1/e5/

(J Med Internet Res 2015;17(3):e75) doi:10.2196/jmir.4419

The authors of "Designing and Evaluating an Interactive Multimedia Web-Based Simulation for Developing Nurses Competencies in Acute Nursing Care: Randomized Controlled Trial" (http://www.jmir.org/2015/1/e5/) have overlooked an error in the Methods section during the submission and proofreading process. The phrase "a minimum of 5 years of work experience" should have been "less than 5 years of work experience". This error has been corrected in the online version of the paper on the JMIR website on March 23, 2015, together with publishing this correction notice. A correction notice has been sent to PubMed and the correct full-text has been resubmitted to Pubmed Central and other full-text repositories.

Edited by G Eysenbach; submitted 15.03.15; this is a non–peer-reviewed article;accepted 16.03.15; published 23.03.15.

Please cite as:
Liaw SY, Wong LF, Chan SWC, Ho JTY, Mordiffi SZ, Ang SBL, Goh PS, Ang ENK
Correction: Designing and Evaluating an Interactive Multimedia Web-Based Simulation for Developing Nurses' Competencies in Acute Nursing Care: Randomized Controlled Trial
J Med Internet Res 2015;17(3):e75
URL: http://www.jmir.org/2015/3/e75/
doi:10.2196/jmir.4419
PMID:30578199

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org