

Original Paper

Psychometric Properties of Patient-Facing eHealth Evaluation Measures: Systematic Review and Analysis

Bonnie J Wakefield¹, PhD; Carolyn L Turvey¹, PhD; Kim M Nazi², PhD; John E Holman¹, MA, MS; Timothy P Hogan³, PhD; Stephanie L Shimada³, PhD; Diana R Kennedy⁴, MHA, MS

¹The Center for Comprehensive Access and Delivery Research and Evaluation, Iowa City Veterans Affairs Healthcare System, Iowa City, IA, United States

²Veterans and Consumers Health Informatics Office, Veterans Health Administration, Washington, DC, United States

³Center for Healthcare Organization and Implementation Research, Edith Nourse Rogers Memorial Veterans Affairs Medical Center, Boston, MA, United States

⁴Department of Health Management and Informatics, University of Missouri, Columbia, MO, United States

Corresponding Author:

Bonnie J Wakefield, PhD

The Center for Comprehensive Access and Delivery Research and Evaluation

Iowa City Veterans Affairs Healthcare System

601 Hwy 6 West

Iowa City, IA, 52246

United States

Phone: 1 319 338 0581

Fax: 1 319 887 4932

Email: wakefieldb@missouri.edu

Abstract

Background: Significant resources are being invested into eHealth technology to improve health care. Few resources have focused on evaluating the impact of use on patient outcomes. A standardized set of metrics used across health systems and research will enable aggregation of data to inform improved implementation, clinical practice, and ultimately health outcomes associated with use of patient-facing eHealth technologies.

Objective: The objective of this project was to conduct a systematic review to (1) identify existing instruments for eHealth research and implementation evaluation from the patient's point of view, (2) characterize measurement components, and (3) assess psychometrics.

Methods: Concepts from existing models and published studies of technology use and adoption were identified and used to inform a search strategy. Search terms were broadly categorized as platforms (eg, email), measurement (eg, survey), function/information use (eg, self-management), health care occupations (eg, nurse), and eHealth/telemedicine (eg, mHealth). A computerized database search was conducted through June 2014. Included articles (1) described development of an instrument, or (2) used an instrument that could be traced back to its original publication, or (3) modified an instrument, and (4) with full text in English language, and (5) focused on the patient perspective on technology, including patient preferences and satisfaction, engagement with technology, usability, competency and fluency with technology, computer literacy, and trust in and acceptance of technology. The review was limited to instruments that reported at least one psychometric property. Excluded were investigator-developed measures, disease-specific assessments delivered via technology or telephone (eg, a cancer-coping measure delivered via computer survey), and measures focused primarily on clinician use (eg, the electronic health record).

Results: The search strategy yielded 47,320 articles. Following elimination of duplicates and non-English language publications (n=14,550) and books (n=27), another 31,647 articles were excluded through review of titles. Following a review of the abstracts of the remaining 1096 articles, 68 were retained for full-text review. Of these, 16 described an instrument and six used an instrument; one instrument was drawn from the GEM database, resulting in 23 articles for inclusion. None included a complete psychometric evaluation. The most frequently assessed property was internal consistency (21/23, 91%). Testing for aspects of validity ranged from 48% (11/23) to 78% (18/23). Approximately half (13/23, 57%) reported how to score the instrument. Only six (26%) assessed the readability of the instrument for end users, although all the measures rely on self-report.

Conclusions: Although most measures identified in this review were published after the year 2000, rapidly changing technology makes instrument development challenging. Platform-agnostic measures need to be developed that focus on concepts important for use of any type of eHealth innovation. At present, there are important gaps in the availability of psychometrically sound measures to evaluate eHealth technologies.

(*J Med Internet Res* 2017;19(10):e346) doi: [10.2196/jmir.7638](https://doi.org/10.2196/jmir.7638)

KEYWORDS

telemedicine; computers; evaluation; use-effectiveness; technology; psychometrics

Introduction

Patient-facing eHealth is a multidisciplinary field focused on the delivery or enhancement of health information and health services through information and communication technologies [1]. eHealth helps consumers engage and collaborate more fully in their health care [2,3], independent of geographic location and also enhances access to health care services by offering novel channels for communication and information flow that complement existing systems [4]. There are many terms related to eHealth, including consumer health informatics, digital health, virtual care, connected care, and telehealth, to list only a few. For purposes of consistency, we use the term “eHealth.”

This paper focuses on patient use of eHealth, which includes personal health records and patient portals accessed via computers or mobile devices, and other telehealth devices designed for use primarily by patients and caregivers, even though some patient-facing technologies (eg, secure patient-provider messaging, mobile apps) are also used by clinicians [5]. Several constructs are important to measure to evaluate patient-facing eHealth technologies. Patient-facing eHealth technologies are used to deliver interventions intended to promote healthy behaviors or effective self-management among consumers. When assessing the efficacy of a behavior-change eHealth intervention, evaluations must address both the intervention and the technology platforms and functions used to deliver the intervention in terms of usability, functionality, and availability of the technology to target users [3]. eHealth may improve the efficiency of and accessibility to clinical and health promotion services for patients. For example, it is anticipated that eHealth may reduce the distance between services and the target user, improving accessibility, or reducing physician or patient workload for a specific task, enhancing efficiency [6-9]. Finally, almost all behavior-change eHealth interventions aim to improve communication in one form or another [10,11].

Although studies using eHealth technologies may include measures that attempt to quantify the characteristics or effect of eHealth interventions, to date, there are no uniform, widely agreed-on measures. More rigorous measurement is needed to determine the full benefit(s) of an eHealth-delivered intervention to both patients and the health care system [12]. Scientific inquiry in other domains has benefited from the development of such standardized measures. At present, various measure compendiums are available that categorize measures of patient-reported outcomes. The Grid-Enabled Measures (GEM) database, for example, was developed starting in 2010 with the purpose of moving social and behavioral science forward by

promoting the use of standardized measures tied to theoretically based constructs and facilitating sharing of data from use of standardized measures [13]. Sponsored by the National Cancer Institute, GEM is an open-source measure compendium that solicits scientific community participation in contributing and selecting measures. Users can add information about constructs, find measures related to constructs, upload new measures, provide feedback on existing measures, and search for and share harmonized data for meta-analyses. In addition to providing useful information such as associated references and information on validity and reliability, the GEM allows researchers to see how often other researchers have used a measure and the feedback and ratings they have provided.

Similarly, the Patient-Reported Outcomes Measurement System (PROMIS) was developed by the National Institutes of Health in an effort to develop, validate, and standardize items that may be used to measure patient-reported outcomes common across medical conditions [14]. PROMIS is collecting and testing items focused on patient-reported outcomes of interest, as opposed to validated instruments. For example, the item banks for physical function, fatigue, and sleep disturbance contain 124, 95, and 27 items, respectively [15]. These item banks are being tested in large populations [16-18].

Both PROMIS and GEM promote use of standardized measures and data analysis across multiple studies and conditions. Although these measures can be an important component of studies focused on use of eHealth technologies, the items and instruments contained in these compendiums do not specifically focus on issues surrounding use of eHealth technology with and by patients. For example, although GEM or PROMIS may include instruments or items that measure patient satisfaction with communication with a physician, they do not include items specific to physician-patient communication when using telehealth or secure messaging, nor do they specifically address technology usability issues. Recent efforts to summarize measures related specifically to technology use include a compendium of health information technology-related survey tools developed by the Agency for Healthcare Research and Quality (AHRQ). The AHRQ compendium includes a wide variety of measures, but the website does not provide detailed information on psychometric properties. Thus, although work is in progress to develop and identify measures that may address eHealth evaluation needs, more work is needed.

Implementation research focuses on structural and organizational characteristics of the environment where an innovation is being or will be used. Within this environment are individuals (patients, providers, administrators) with various characteristics that may hinder or facilitate adoption of the innovation within

the particular environment. In this review, we focus on the innovation (ie, the eHealth intervention) and how features of this innovation will impact implementation. Consistent and well-validated measures will contribute to determining the true benefit of eHealth interventions across studies and over time. Consistently used measures will enable the health care system to collect uniform data on (1) the likelihood of adoption of an eHealth technology; (2) patient, organizational, or health care system barriers and facilitators to adoption; (3) user attitudes toward and/or satisfaction with a technology; (4) the degree to which meaningful user characteristics (eg, health literacy) mediate the relationship between technology use and improved health outcomes (ie, improved self-management of chronic illness, reduced health care utilization), and (5) the return on investment of eHealth technology to assess value.

The objective of this project was to conduct a systematic review to (1) identify existing instruments for eHealth research and implementation evaluation, (2) characterize measurement components, and (3) assess psychometrics. Additionally, this study seeks to highlight current limitations of this body of research.

Methods

Identification of Search Terms

Through a series of investigator meetings, we identified key concepts from existing models, published studies of technology use and adoption, and sociotechnical perspectives on health information technology implementation and evaluation [19-23]. Using these models and studies, our knowledge of the field, and detailed input from an experienced health sciences librarian, we developed a working list of key concepts to focus our search. These were then categorized into five areas: platforms (eg, email), measurement (eg, survey), function/information use (eg, self-management), health care occupations (eg, nurse), and eHealth/telemedicine (eg, mHealth) (Multimedia Appendix 1). Our focus was to identify instruments that could be used for any of these concepts as well as those that may be relevant to only one or two concepts.

Search Strategy

We conducted a systematic search of the literature using the selected search terms. Based on guidance from our health sciences librarian, databases used included MEDLINE, Scopus, PsychInfo, CINAHL, Health and Psychosocial Instruments (HAPI) for articles published through June 2014. Each database was searched using terms included in Multimedia Appendix 1. The search logic followed this format: (A and D and B and C) OR (E and B and C). All terms listed in sets A, B, D, and C were entered and combined using the Boolean operator “and.” Likewise, terms in sets B, C, and E were entered and combined using “and.” The results from these two searches were then combined using the operator “OR.” This logic was used to ensure all possible terms were included and ensured studies included some sort of measurement or evaluations.

Our search strategy also included review of currently funded research projects within the health services research arm of the Veterans Health Administration (VA) system focused on eHealth

(n=56), and existing instrument/measure compendiums (GEM, PROMIS, AHRQ). All search results were transferred to a reference management software database (EndNote); duplicates, articles where the text was not in English, and books were eliminated.

Inclusion Criteria

Our article inclusion criteria were broad to identify the full extent of instruments designed for eHealth research and implementation evaluation. We focused explicitly on instruments that assessed an eHealth-specific construct from the patient’s point of view. Articles were selected if they (1) described development of an instrument, or (2) used an instrument in an evaluation of an eHealth technology that could be traced back to an original publication describing its development, or (3) modified an instrument, and (4) with full text in English language. The review was limited to instruments that reported at least one psychometric property. Excluded were investigator-developed measures or sets of questions without psychometric evaluation, disease-specific assessments delivered via technology or telephone (eg, a cancer-coping measure delivered via computer survey), and measures focused primarily on clinician use (eg, the electronic health record). We limited our review to articles that reported at least one established psychometric property (see Table 1 for psychometric evaluation components).

Data Extraction

Two investigators and a research assistant (BW, JH, AM) independently reviewed 100 article titles followed by an in-depth discussion to establish agreement on inclusion of articles. Next, the review was repeated two times using an additional 100 article titles each time, until agreement was reached on articles to include for further review. All article titles were then reviewed to exclude ineligible articles. The abstracts of the remaining articles were reviewed by a pair of investigators (BW, CT) following an independent review of 20 articles to establish interrater consistency. The remaining abstracts were then independently reviewed and discrepancies between reviewers were resolved by discussion and consensus. Articles that did not meet criteria were excluded (no instrument, use of an instrument, or instrument modification), and remaining articles were retained for full-text review. Articles were then classified as describing the development and testing of an instrument or as using an instrument. For articles using an instrument, reference lists were reviewed to identify citations for the original instrument development.

A data extraction form with definitions for each item was developed by the study team (Table 1) [24]. To establish interrater reliability in data extraction, coauthors were divided into pairs, and were assigned to independently review two articles using the data extraction tool. These reviews were discussed in depth by the whole study team to reach consensus on the definitions used in Table 1. Following minor revisions of the data extraction form, articles from the search were then distributed among the six study investigators for final review and data extraction. The first author then reviewed each article and data extraction information to ensure accuracy.

Table 1. Data extraction elements.

Element	Definition
Construct	Constructs are not directly observable, but may be applied and defined based on observable behavior; many health measures are designed to capture some aspect of an underlying construct. In the authors' own words, what the authors of the scale say they are measuring.
Theoretical foundation	Conception of how attributes exist and relate to one another; theoretical framework; can indicate that a conceptual framework (concepts identified in the framework) was used.
Modification of another instrument by others (alternate forms) abbreviated, short forms, different forms targeting the same construct, translations	State if this article is a modification of the format or administration of an instrument already evaluated for psychometric properties.
# items	Number of items included in the measure.
Item types	Structure of the items: such as Likert-type, categorical (multiple options), open ended, yes/no, visual analog scale, other.
Administration time	Estimated amount of time for completion of the measure.
Administration mode	Assessment completed by self-report vs interviewer/researcher administered.
Active vs passive assessment/obtrusiveness	Data collection which does not involve direct solicitation from the research subject or other participant; indirect ways to obtain the necessary data often relying on technology captured information such as response time, number of navigation errors, etc.
Item development	Briefly overview how items were developed for the original form of the measure (ie, expert generation of items, compilation of items from prior measures).
Scoring	Describe how the measure is scored, include a range of possible scores and other descriptive statistics such as significant threshold scores if available.
Readability	Did the developers test the readability of the measure? Were any readability formulas used (eg, Flesch-Kincaid).
Sensitivity to change	Ability to detect change over time, particularly in response to some intervention; known as responsiveness; floor and ceiling effects.
Reliability: test-retest	Consistency in scores between 2 administrations of the measure separated by time (ie, same subject completes the measure twice).
Reliability: interrater	Consistency between 2 independent observers using the measure (for measures that involve observing subjects)% agreement, kappa.
Reliability: internal consistency	Degree to which all items in the scale correlate with each other taking length of measure into account, indicating the items measure the same underlying construct. Based on a single administration of the measure; Cronbach alpha, Kuder-Richardson, split-half reliability.
Validity: content	Typically, from a review of the literature or review by experts.
Validity: criterion, convergent, concurrent, discriminant	Correlation of the scale with other measures to determine independence from other constructs yet some positive correlation to similar constructs and negative correlation to dissimilar constructs.
Validity: construct	Linking the measure to another known attribute. Factor analysis to identify proposed underlying constructs consistent with proposed theoretic content of the measure.
Sample	Patient population used to develop, validate, or test the measure.
Sample studies using the metric/strength of evidence	Studies using the measure including those that did not present psychometric properties of the measure.
Measure website address	If the measure has an associated website, list the website address here and note the date of last update, if available.
Copyright or fees associated with use of the measure	Requires purchase of the measure or the scoring algorithm?

Results

The search strategy yielded 47,320 articles (PubMed: n=16,968; Scopus: n=24,106; PsychInfo: n=3590; CINAHL: n=2187; HAPI: n=468; GEM: n=1). Following elimination of duplicates and full text not in English language publications (n=14,550) and books (n=27), most articles were excluded through review

of titles (n=31,647). Following a review of the abstracts of the remaining 1096 articles, 68 were retained for full-text review. Of these, 16 described an instrument and six used an instrument; one instrument was drawn from the GEM database, resulting in 23 articles for inclusion in the review (Figure 1). Of these 23 articles, seven were modifications of existing instruments. No additional measures were identified through our VA, PROMIS,

or AHRQ search. Each article was then reviewed by team members, using the data extraction form (Table 1).

We identified common conceptual threads across the 23 instruments. We reviewed the literature to identify salient concepts and constructs from existing technology use models [19-22,25]. Multiple constructs were identified and terminology varied across models. For example, the Technology Acceptance Model includes 16 constructs in four categories (behavioral intention, perceived usefulness, perceived ease of use, and use behavior). Although terminology varied by author and model, categorizations were inferred and grouped. Twelve concepts emerged from this categorization: clinical content, communication, effectiveness, efficiency, frequency/consistency of use, hardware and software, perceived ease of use, policies and procedures, risk and benefits, user preferences, social influence, and usability. Author definitions guided this

categorization. The definition of several of these terms are intuitive (eg, effectiveness), but some are not and are briefly defined here. Efficiency includes the concepts of accuracy, costs, learnability, performance expectations, productivity, quality of use, and workflow. Learnability is an aspect of usability and refers to the ease of learning how to use software. Closely related to learnability is performance expectation, where the end user knows what is expected from them to use the software. Hardware and software aspects include availability, human-computer interface (ie, efficient and desirable interaction between a person and the computer), information display, system maintenance and monitoring, and technical quality. Perceived ease of use incorporates anxiety about and attitude toward using a computer, behavioral intention (the likelihood that an individual will use the computer), computer self-efficacy, engagement, enjoyment, and usefulness.

Figure 1. Flow diagram of search.

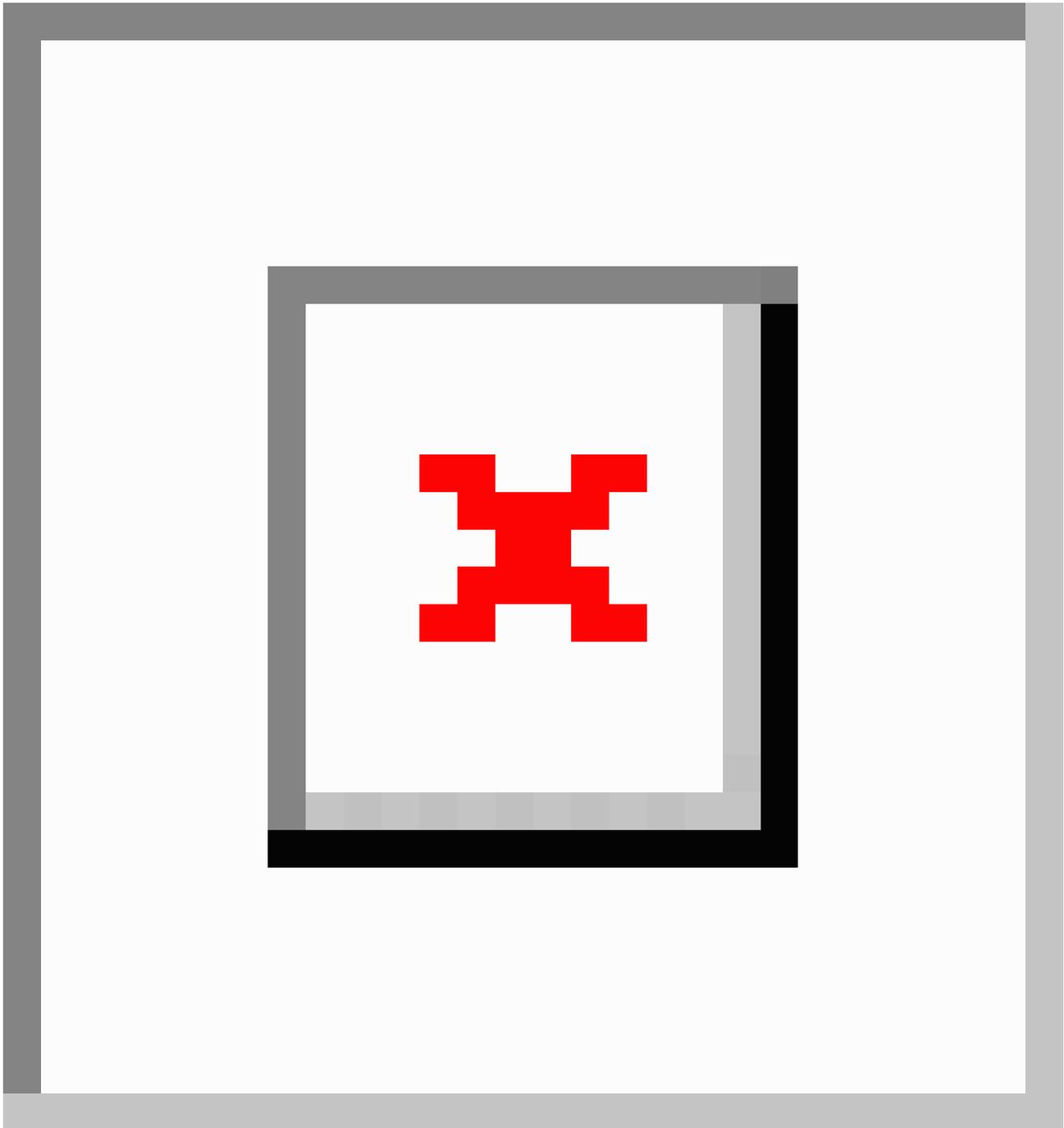


Table 2. Concepts 1 to 6 identified in reviewed instruments (N=23).

Article	Concept and model authors					
	Clinical content [20]	Communication [20,21]	Effectiveness [22]	Efficiency [20-22]	Frequency/consistency of use [21,23]	Hardware and software [19-23]
Atkison, 2007 [29]			X	X		
Bakken, 2006 [30]			X	X		X
Brockmeyer, 2013 [31]						X
Brooke, 1996 [32]			X	X		
Bunz, 2004 [33]			X	X	X	X
Demiris, 2000 [34]			X	X		X
Finkelstein, 2012 [35]	X		X	X	X	
Henkemans, 2013 [36]	X					X
Hudiberg, 1991-1996 [37-40]			X	X		X
Jay & Willis, 1992 [41]			X			
Lewis, 1993 [42]			X	X		X
Lin, 2011 [43]				X		X
Martinez-Caro, 2013 [44]	X				X	X
Montague, 2012 [45]	X		X	X		X
Norman, 2006 [46]	X		X	X	X	X
Pluye, 2014 [47]	X	X		X	X	X
Schnall, 2011 [48]	X	X	X	X		
Tariman, 2011 [49]	X		X	X		X
Wang, 2008 [50]	X		X	X		X
Wehmeyer, 2008 [27]						X
Wolfradt, 2001 [51]		X				X
Xie, 2013 [28]	X		X	X		X
Yip, 2003 [52]	X					X

The 23 articles included in this review were mapped to the 12 identified concepts based on whether the instrument encompassed the concept. The most common constructs addressed by this set of measures were effectiveness, efficiency, hardware and software, perceived ease of use, satisfaction, and usability [19-23] (Tables 2 and 3). Interestingly, although eHealth is a communication technology, only three studies specifically address this aspect. Additionally, to identify potential gaps for future consideration, concepts included in the measures, but not identified in the 12 model concepts, were documented in the crosswalk (last column in Table 3). For example, stress, eHealth literacy, perceived necessity, and others emerged as concepts not identified in the review of existing technology use models. eHealth literacy is defined by Norman and Skinner [26] as “the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem.” Wehmeyer [27] introduced three concepts: symbolism, esthetics, and perceived necessity. Symbolism reflects the meaning or status associated with the device (eg, having a mobile device may signify group membership or a certain social status). Esthetics refers to the appearance of the

device (eg, the perceived beauty of the device may affect the attachment to the device). Finally, the perceived necessity of the device may affect attachment to the device, creating anxiety when the device is not accessible. Xie et al [28] addressed decision-making autonomy, defined as the level of decision making desired when information about health conditions is electronically available.

No instrument included a complete psychometric evaluation (Multimedia Appendix 2). The most frequently assessed property was internal consistency (21/23, 91%). None of the measures were assessed for sensitivity to change, but several authors indicated the instrument was not designed to assess change. Few measures were assessed for test-retest reliability (4/23, 17%) and only one instrument had been tested for interrater reliability. Testing for aspects of validity ranged from 48% (11/23) of measures tested for criterion, convergent, concurrent, or discriminant validity to 78% (18/23) reporting establishing content validity. Approximately half (13/23, 57%) reported how to score the instrument. Only six (26%) assessed the readability of the instrument for end users, although all measures rely on patient self-report.

Table 3. Concepts 7 to 12 identified in reviewed instruments (N=23).

Article	Concepts and model authors						Concepts not included in models
	Perceived ease of use [19,21-23]	Policies & procedures [20]	Risk & benefits [23]	Satisfaction/acceptability/preferences [23]	Social influence [21]	Usability [23]	
Atkison, 2007 [29]	X					X	
Bakken, 2006 [30]	X			X		X	
Brockmeyer, 2013 [31]	X		X	X			
Brooke, 1996 [32]				X		X	
Bunz, 2004 [33]	X			X		X	
Demiris, 2000 [34]	X		X	X		X	
Finkelstein, 2012 [35]	X			X			
Henkemans, 2013 [36]	X		X	X		X	
Hudiberg, 1991-1996 [37-40]	X		X	X	X		Stress
Jay & Willis, 1992 [41]	X				X	X	
Lewis, 1993 [42]	X		X	X		X	
Lin, 2011 [43]	X			X		X	
Martinez-Caro, 2013 [44]	X		X	X			
Montague, 2012 [45]	X		X	X		X	
Norman, 2006 [46]	X			X			eHealth literacy
Pluye, 2014 [47]	X			X			
Schnall, 2011 [48]	X		X				
Tariman, 2011 [49]	X			X			
Wang, 2008 [50]	X			X		X	
Wehmeyer, 2008 [27]	X			X			Symbolism; esthetics; perceived necessity
Wolfradt, 2001 [51]				X			
Xie, 2013 [28]	X			X		X	Decision-making autonomy
Yip, 2003 [52]				X			

Early instruments (prior to the year 2000) [32,37-42] focused on using a computer, reflecting early consumer adoption of personal computers. These measures are not specifically focused on “health” use. During the decade from 2000 to 2009, measures that focused on use of information technology related to health began to emerge, focusing primarily on telehealth [30,34,52]; other measures focused on eHealth literacy [46] and use of eHealth education [29]. Other concepts for which measures were developed included using the Internet [51], use of computers [33], use of mobile devices [27,50], and the effect of video games on engagement [31], although these measures did not specifically focus on “health.” Since 2010, the frequency of “health” themes increased including communication between patients and providers [47,49], patient trust [45], preferences [28], satisfaction [35], and use of technology for care provision [48] or patient self-management [36,48]. One instrument also focused more generally on use of computers [43], and one focused on patient loyalty to online services [44].

Discussion

Principal Findings

Of the 23 articles reviewed, no instrument included a complete psychometric evaluation. The most frequently assessed property was internal consistency. Testing for aspects of validity ranged from 48% (11/23) to 78% (18/23). Approximately half (13/23, 57%) reported how to score the instrument. Only six (26%) assessed the readability of the instrument for end users, although all the measures rely on self-report.

Common theoretical concepts addressed in the instruments were effectiveness, efficiency, hardware/software, perceived ease of use, and satisfaction. A notable exception is that only three instruments focused on communication. Conversely, we identified some concepts addressed in the instruments that have not been included in current theoretical models, including stress, esthetics, eHealth literacy, comfort, and decision-making

autonomy. Current instruments require fuller evaluation of psychometric properties.

Measures that can be applied consistently across technologies and platforms are needed so that distinct platforms that serve the same purpose can be compared. For example, evaluation of an intervention to treat depression could utilize a standard measure of usability, regardless of whether it was a mobile app or Web-based (eg, “It took many tries before I knew how to use the key features of this technology” and “I found the layout of the features very intuitive”), regardless of the platform used to deliver the intervention (eg, mobile app or online program). Using these types of measures, investigators and others implementing eHealth technologies can compare technologies and use this information when selecting a technology.

Our review expands on the AHRQ compendium, which lists available measures but provides less detail about their other attributes. We also investigated whether the psychometric properties of the measures had been established, which is a critical information need when selecting a measure for research or evaluation. However, although most would agree that instruments with psychometric properties are very helpful, there may also be a role for using self-developed questions that may more clearly and directly get at the target construct or a specific patient behavior. The AHRQ compendium is populated with many such instruments and future researchers should carefully consider the trade-offs of using investigator-developed question sets that may specifically address their question of interest versus a more validated instrument that may also need to be modified to fit an eHealth evaluation. Furthermore, investigators may want to consider instruments listed in the AHRQ compendium for further development and psychometric evaluation.

Implementation of eHealth technologies can involve substantial investment in terms of costs and effort. Research on eHealth has also increased dramatically over the past several years, yet studies rarely utilize common methods and/or instruments. The results of this project provide critical insights regarding existing eHealth instruments and identify gaps for which new instruments are needed. Use of common and psychometrically sound instruments can inform future studies so that the results from multiple studies can be compared and synthesized.

Although most the instruments identified in this review were published after the year 2000, rapidly changing technology makes instrument development challenging. Platform-agnostic measures need to be developed that focus on concepts important for use of any type of eHealth innovation. Instrument development as a research enterprise is typically undervalued, relative to more direct practice-relevant research. Instrument development can also be a complex and lengthy process. Thus, funding agencies should consider addressing this gap, given the persistent and expected growth in the deployment of technology to improve care processes and patient outcomes.

Limitations

We did not conduct a comprehensive search for all published uses of the identified instruments as it was beyond the scope of this study. The grey literature (eg, conference abstracts, dissertations, and unpublished studies) were not included in our review. Furthermore, the review potentially missed some published as well as unpublished measures based on keyword choice and/or elimination of articles through review of title or abstract. Finally, our choice of theoretical models used to analyze the selected articles may impose limitations on our findings.

Conclusions

Based on our review, we highlight some of the more useful measures that we believe could be useful in most technology studies. These include the eHealth literacy scale (eHEALS) [46], the Computer-Email-Web Fluency Scale [33], and the System Usability Scale [32]. Additional research is needed to build and further refine measures of literacy such as the eHEALS or Computer-Email-Web Fluency Scale so that researchers have access to a validated measure of user’s comfort with a target technology.

Development of a standard measure of the intuitiveness of the user interface would allow platform-agnostic comparisons between user interfaces (eg, two mobile apps for depression, or comparison of differences between a Web-based and mobile app). Finally, given the explosion of new technologies in the market focused on health behaviors, a standard measure of the relative advantage of a new technology feature when compared to prior methods and/or a standard measure of the degree to which new technology facilitates a target behavior (eg, weight loss, exercise, self-management techniques, or receipt of care) could provide important insights to inform technology adoption strategies.

Advances in eHealth offer tremendous potential to improve access to care, efficiency of care delivery processes, and overall quality. Significant resources are being invested in eHealth technologies, driven in part by meaningful use requirements. Consumer behavioral health interventions are increasingly being made available via multiple platforms (eg, computer vs mobile versions of interventions proven effective for in-person delivery). Identification of useful and valid measures to evaluate these interventions has important potential to contribute to improved implementation, clinical practice, and ultimately population health since insights gleaned from standardized measurement can directly inform system improvements and optimal implementation strategies. In addition, having better measures to evaluate implementation of eHealth technologies will help improve consumers’ experiences with technologies and assess whether use of these technologies is making a measurable difference in quality of care or the patient experience. More longitudinal research will be needed to develop measures that more comprehensively address the wider frame of concepts important for the meaningful implementation of eHealth technologies.

Acknowledgments

The work reported here was supported by the Department of Veterans Affairs Health Services Research & Development Quality Enhancement Research Initiative grant #RRP12-496 and a Career Development Award (CDA 10-210) (Shimada). Assistance was provided by Amy Blevins, University of Iowa Hardin Health Sciences Library, who assisted with the article search; Thomas Houston, MD, for review and input on project design; and Ashley McBurney for project assistance. Study sponsors provided funding, but had no role in the design or conduct of the study; sponsors did not review or approve the manuscript prior to submission. The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search Terms.

[\[PDF File \(Adobe PDF File\), 41KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Detailed Psychometric Properties of Reviewed Instruments.

[\[PDF File \(Adobe PDF File\), 113KB-Multimedia Appendix 2\]](#)

References

1. Eysenbach G. What is e-health? *J Med Internet Res* 2001;3(2):e20 [FREE Full text] [doi: [10.2196/jmir.3.2.e20](https://doi.org/10.2196/jmir.3.2.e20)] [Medline: [11720962](https://pubmed.ncbi.nlm.nih.gov/11720962/)]
2. Klein-Fedyshin MS. Consumer Health Informatics--integrating patients, providers, and professionals online. *Med Ref Serv Q* 2002;21(3):35-50. [doi: [10.1300/J115v21n03_03](https://doi.org/10.1300/J115v21n03_03)] [Medline: [12238015](https://pubmed.ncbi.nlm.nih.gov/12238015/)]
3. Calabretta N. Consumer-driven, patient-centered health care in the age of electronic information. *J Med Libr Assoc* 2002 Jan;90(1):32-37 [FREE Full text] [Medline: [11838457](https://pubmed.ncbi.nlm.nih.gov/11838457/)]
4. Hogan T, Wakefield B, Nazi K, Houston T, Weaver F. Promoting access through complementary eHealth technologies: recommendations for VA's Home Telehealth and personal health record programs. *J Gen Intern Med* 2011 Nov;26 Suppl 2:628-635 [FREE Full text] [doi: [10.1007/s11606-011-1765-y](https://doi.org/10.1007/s11606-011-1765-y)] [Medline: [21989614](https://pubmed.ncbi.nlm.nih.gov/21989614/)]
5. Ahern D, Kreslake J, Phalen J. What is eHealth (6): perspectives on the evolution of eHealth research. *J Med Internet Res* 2006 Mar 31;8(1):e4 [FREE Full text] [doi: [10.2196/jmir.8.1.e4](https://doi.org/10.2196/jmir.8.1.e4)] [Medline: [16585029](https://pubmed.ncbi.nlm.nih.gov/16585029/)]
6. Elbert N, van Os-Medendorp H, van Renselaar W, Ekeland A, Hakkaart-van RL, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. *J Med Internet Res* 2014 Apr 16;16(4):e110 [FREE Full text] [doi: [10.2196/jmir.2790](https://doi.org/10.2196/jmir.2790)] [Medline: [24739471](https://pubmed.ncbi.nlm.nih.gov/24739471/)]
7. Sarkar U, Lyles C, Parker M, Allen J, Nguyen R, Moffet H, et al. Use of the refill function through an online patient portal is associated with improved adherence to statins in an integrated health system. *Med Care* 2014 Mar;52(3):194-201 [FREE Full text] [doi: [10.1097/MLR.000000000000069](https://doi.org/10.1097/MLR.000000000000069)] [Medline: [24374412](https://pubmed.ncbi.nlm.nih.gov/24374412/)]
8. Tang P, Overhage J, Chan A, Brown N, Aghighi B, Entwistle M, et al. Online disease management of diabetes: engaging and motivating patients online with enhanced resources--diabetes (EMPOWER-D), a randomized controlled trial. *J Am Med Inform Assoc* 2013 May 01;20(3):526-534 [FREE Full text] [doi: [10.1136/amiainjnl-2012-001263](https://doi.org/10.1136/amiainjnl-2012-001263)] [Medline: [23171659](https://pubmed.ncbi.nlm.nih.gov/23171659/)]
9. Chen C, Garrido T, Chock D, Okawa G, Liang L. The Kaiser Permanente electronic health record: transforming and streamlining modalities of care. *Health Affairs* 2009;28(2):323-333 [FREE Full text] [doi: [10.1377/hlthaff.28.2.323](https://doi.org/10.1377/hlthaff.28.2.323)] [Medline: [19275987](https://pubmed.ncbi.nlm.nih.gov/19275987/)]
10. Nazi K, Hogan T, Woods S, Simon S, Ralston J. Consumer health informatics: engaging and empowering patients and families. In: Finnell JT, Dixon BE, editors. *Clinical Informatics Study Guide: Text and Review*. New York: Springer; 2016:459-500.
11. de Jong CC, Ros W, 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 Jan 16;16(1):e19 [FREE Full text] [doi: [10.2196/jmir.3000](https://doi.org/10.2196/jmir.3000)] [Medline: [24434570](https://pubmed.ncbi.nlm.nih.gov/24434570/)]
12. Black A, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. *PLoS Med* 2011;8(1):e1000387. [doi: [10.1371/journal.pmed.1000387](https://doi.org/10.1371/journal.pmed.1000387)]
13. Moser R, Hesse B, Shaikh A, Courtney P, Morgan G, Augustson E, et al. Grid-enabled measures: using Science 2.0 to standardize measures and share data. *Am J Prev Med* 2011;40(5 Suppl 2):S134-S143. [doi: [10.1016/j.amepre.2011.01.004](https://doi.org/10.1016/j.amepre.2011.01.004)]

14. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care* 2007;45(Suppl 1):S3-S11. [doi: [10.1097/01.mlr.0000258615.42478.55](https://doi.org/10.1097/01.mlr.0000258615.42478.55)]
15. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks. *J Clin Epidemiol* 2010;63(11):1179-1194. [doi: [10.1016/j.jclinepi.2010.04.011](https://doi.org/10.1016/j.jclinepi.2010.04.011)]
16. Lai J, Cella D, Choi S, Junghaenel D, Christodoulou C, Gershon R, et al. How item banks and their application can influence measurement practice in rehabilitation medicine: a PROMIS fatigue item bank example. *Arch Phys Med Rehabil* 2011;92(10 Suppl):S20-S27. [doi: [10.1016/j.apmr.2010.08.033](https://doi.org/10.1016/j.apmr.2010.08.033)]
17. Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *J Clin Epidemiol* 2011 May;64(5):507-516 [FREE Full text] [doi: [10.1016/j.jclinepi.2010.11.018](https://doi.org/10.1016/j.jclinepi.2010.11.018)] [Medline: [21447427](https://pubmed.ncbi.nlm.nih.gov/21447427/)]
18. Rothrock NE, Hays RD, Spritzer K, Yount SE, Riley W, Cella D. Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol* 2010 Nov;63(11):1195-1204 [FREE Full text] [doi: [10.1016/j.jclinepi.2010.04.012](https://doi.org/10.1016/j.jclinepi.2010.04.012)] [Medline: [20688471](https://pubmed.ncbi.nlm.nih.gov/20688471/)]
19. Holden RJ, Karsh B. 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](https://doi.org/10.1016/j.jbi.2009.07.002)] [Medline: [19615467](https://pubmed.ncbi.nlm.nih.gov/19615467/)]
20. Sittig D, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Health Care* 2010 Oct;19 Suppl 3:i68-i74 [FREE Full text] [doi: [10.1136/qshc.2010.042085](https://doi.org/10.1136/qshc.2010.042085)] [Medline: [20959322](https://pubmed.ncbi.nlm.nih.gov/20959322/)]
21. Venkatesh V, Bala H. Technology Acceptance Model 3 and a research agenda on interventions. *Decision Sci* 2008;39(2):273-315. [doi: [10.1111/j.1540-5915.2008.00192.x](https://doi.org/10.1111/j.1540-5915.2008.00192.x)]
22. Alexander G, Staggers N. A systematic review of the designs of clinical technology: findings and recommendations for future research. *Adv Nurs Sci* 2009;32(3):252-279 [FREE Full text] [doi: [10.1097/ANS.0b013e3181b0d737](https://doi.org/10.1097/ANS.0b013e3181b0d737)] [Medline: [19707093](https://pubmed.ncbi.nlm.nih.gov/19707093/)]
23. Kim J, Park HA. Development of a health information technology acceptance model using consumers' health behavior intention. *J Med Internet Res* 2012 Oct 01;14(5):e133 [FREE Full text] [doi: [10.2196/jmir.2143](https://doi.org/10.2196/jmir.2143)] [Medline: [23026508](https://pubmed.ncbi.nlm.nih.gov/23026508/)]
24. Streiner D, Norman G. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Oxford: Oxford University Press; 1995.
25. Rondan-Cataluña FJ, Arenas-Gaitán J, Ramírez-Correa PE. A comparison of the different versions of popular technology acceptance models. *Kybernetes* 2015;44(5):788-805. [doi: [10.1108/k-09-2014-0184](https://doi.org/10.1108/k-09-2014-0184)]
26. Norman C, Skinner H. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res* 2006 Jun 16;8(2):e9 [FREE Full text] [doi: [10.2196/jmir.8.2.e9](https://doi.org/10.2196/jmir.8.2.e9)] [Medline: [16867972](https://pubmed.ncbi.nlm.nih.gov/16867972/)]
27. Wehmeyer K. User-device attachment scale: development and initial test. *Int J Mob Commun* 2008;6(3):280-295. [doi: [10.1504/ijmc.2008.017511](https://doi.org/10.1504/ijmc.2008.017511)]
28. 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](https://doi.org/10.2196/jmir.2615)] [Medline: [23816979](https://pubmed.ncbi.nlm.nih.gov/23816979/)]
29. Atkinson NL. Developing a questionnaire to measure perceived attributes of eHealth innovations. *Am J Health Behav* 2007;31(6):612-621. [doi: [10.5555/ajhb.2007.31.6.612](https://doi.org/10.5555/ajhb.2007.31.6.612)] [Medline: [17691874](https://pubmed.ncbi.nlm.nih.gov/17691874/)]
30. Bakken S, Grullon-Figueroa L, Izquierdo R, Lee N, Morin P, Palmas W, IDEATel Consortium. Development, validation, and use of English and Spanish versions of the telemedicine satisfaction and usefulness questionnaire. *J Am Med Inform Assoc* 2006;13(6):660-667 [FREE Full text] [doi: [10.1197/jamia.M2146](https://doi.org/10.1197/jamia.M2146)] [Medline: [16929036](https://pubmed.ncbi.nlm.nih.gov/16929036/)]
31. Brockmyer J, Fox C, Curtiss K, McBroom E, Burkhart K, Pidruzny J. The development of the Game Engagement Questionnaire: a measure of engagement in video game-playing. *J Exp Soc Psychol* 2009;45(4):624-634.
32. Brooke J. SUS: a quick and dirty usability scale. In: Jordan P, Thomas B, Weerdmeester B, McClelland I, Bristol P, editors. *Usability Evaluation in Industry*. London: Taylor & Francis; 1996:189-194.
33. Bunz U. The Computer-Email-Web (CEW) Fluency Scale-development and validation. *Int J Hum-Comput Interact* 2004;17(4):479-506.
34. Demiris G, Speedie S, Finkelstein S. A questionnaire for the assessment of patients' impressions of the risks and benefits of home telecare. *J Telemed Telecare* 2000;6(5):278-284. [doi: [10.1258/1357633001935914](https://doi.org/10.1258/1357633001935914)] [Medline: [11070589](https://pubmed.ncbi.nlm.nih.gov/11070589/)]
35. Finkelstein SM, MacMahon K, Lindgren BR, Robiner WN, Lindquist R, VanWormer A, et al. Development of a remote monitoring satisfaction survey and its use in a clinical trial with lung transplant recipients. *J Telemed Telecare* 2012;18(1):42-46. [doi: [10.1258/jtt.2011.110413](https://doi.org/10.1258/jtt.2011.110413)] [Medline: [22067286](https://pubmed.ncbi.nlm.nih.gov/22067286/)]
36. Blanson Henkemans OA, Dusseldorp E, Keijsers J, Kessens J, Neerincx M, Otten W. Validity and reliability of the eHealth analysis and steering instrument. *Med* 2013;2(2):e8 [FREE Full text] [doi: [10.2196/med20.2571](https://doi.org/10.2196/med20.2571)] [Medline: [25075243](https://pubmed.ncbi.nlm.nih.gov/25075243/)]
37. Hudiburg RA. Psychology of computer use: XXXIV. The Computer Hassles Scale: subscales, norms, and reliability. *Psychol Rep* 1995 Dec;77(3 Pt 1):779-782. [doi: [10.2466/pr0.1995.77.3.779](https://doi.org/10.2466/pr0.1995.77.3.779)] [Medline: [8559916](https://pubmed.ncbi.nlm.nih.gov/8559916/)]

38. Hudiburg RA, Ahrens PK, Jones TM. Psychology of computer use: XXXI. Relating computer users' stress, daily hassles, somatic complaints, and anxiety. *Psychol Rep* 1994 Dec;75(3 Pt 1):1183-1186. [doi: [10.2466/pr0.1994.75.3.1183](https://doi.org/10.2466/pr0.1994.75.3.1183)] [Medline: [7892380](https://pubmed.ncbi.nlm.nih.gov/7892380/)]
39. Hudiburg RA, Jones TM. Psychology of computer use: XXIII. Validating a measure of computer-related stress. *Psychol Rep* 1991;69(1):179-182. [doi: [10.2466/pr0.1991.69.1.179](https://doi.org/10.2466/pr0.1991.69.1.179)] [Medline: [1961786](https://pubmed.ncbi.nlm.nih.gov/1961786/)]
40. Hudiburg RA, Necessary JR. Psychology of computer use: XXXV. Differences in computer users' stress and self-concept in college personnel and students. *Psychol Rep* 1996 Jun;78(3 Pt 1):931-937. [doi: [10.2466/pr0.1996.78.3.931](https://doi.org/10.2466/pr0.1996.78.3.931)] [Medline: [8711049](https://pubmed.ncbi.nlm.nih.gov/8711049/)]
41. Jay G, Willis S. Influence of direct computer experience on older adults' attitudes toward computers. *J Gerontol* 1992 Jul;47(4):P250-P257. [Medline: [1624702](https://pubmed.ncbi.nlm.nih.gov/1624702/)]
42. Lewis J. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int J Hum-Comput Int* 1995;7(1):57-78.
43. Lin T. A computer literacy scale for newly enrolled nursing college students: development and validation. *J Nurs Res* 2011;19(4):305-317. [doi: [10.1097/JNR.0b013e318236d03f](https://doi.org/10.1097/JNR.0b013e318236d03f)] [Medline: [22089656](https://pubmed.ncbi.nlm.nih.gov/22089656/)]
44. Martínez-Caro E, Cegarra-Navarro JG, Solano-Lorente M. Understanding patient e-loyalty toward online health care services. *Health Care Manage Rev* 2013;38(1):61-70. [doi: [10.1097/HMR.0b013e31824b1c6b](https://doi.org/10.1097/HMR.0b013e31824b1c6b)] [Medline: [22387970](https://pubmed.ncbi.nlm.nih.gov/22387970/)]
45. Montague E. Validation of a trust in medical technology instrument. *Appl Ergon* 2010;41(6):812-821 [FREE Full text] [doi: [10.1016/j.apergo.2010.01.009](https://doi.org/10.1016/j.apergo.2010.01.009)] [Medline: [20189163](https://pubmed.ncbi.nlm.nih.gov/20189163/)]
46. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. *J Med Internet Res* 2006;8(4):e27 [FREE Full text] [doi: [10.2196/jmir.8.4.e27](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
47. Pluye P, Granikov V, Bartlett G, Grad R, Tang D, Johnson-Lafleur J, et al. Development and content validation of the information assessment method for patients and consumers. *JMIR Res Protoc* 2014 Feb 18;3(1):e7 [FREE Full text] [doi: [10.2196/resprot.2908](https://doi.org/10.2196/resprot.2908)] [Medline: [24550180](https://pubmed.ncbi.nlm.nih.gov/24550180/)]
48. Schnall R, Bakken S. Testing the Technology Acceptance Model: HIV case managers' intention to use a continuity of care record with context-specific links. *Inform Health Soc Care* 2011 Sep;36(3):161-172 [FREE Full text] [doi: [10.3109/17538157.2011.584998](https://doi.org/10.3109/17538157.2011.584998)] [Medline: [21848452](https://pubmed.ncbi.nlm.nih.gov/21848452/)]
49. Tariman J, Berry D, Halpenny B, Wolpin S, Schepp K. Validation and testing of the Acceptability E-scale for web-based patient-reported outcomes in cancer care. *Appl Nurs Res* 2011 Feb;24(1):53-58 [FREE Full text] [doi: [10.1016/j.apnr.2009.04.003](https://doi.org/10.1016/j.apnr.2009.04.003)] [Medline: [20974066](https://pubmed.ncbi.nlm.nih.gov/20974066/)]
50. Wang Y, Wang H. Developing and validating an instrument for measuring mobile computing self-efficacy. *Cyberpsychol Behav* 2008;11(4):405-413. [doi: [10.1089/cpb.2007.0061](https://doi.org/10.1089/cpb.2007.0061)] [Medline: [18721088](https://pubmed.ncbi.nlm.nih.gov/18721088/)]
51. Wolfradt U, Doll J. Motives of adolescents to use the internet as a function of personality traits, personal and social factors. *J Educ Comput Res* 2001;24(1):13-27.
52. Yip M, Chang A, Chan J, MacKenzie A. Development of the Telemedicine Satisfaction Questionnaire to evaluate patient satisfaction with telemedicine: a preliminary study. *J Telemed Telecare* 2003;9(1):46-50. [doi: [10.1258/135763303321159693](https://doi.org/10.1258/135763303321159693)] [Medline: [12641893](https://pubmed.ncbi.nlm.nih.gov/12641893/)]

Abbreviations

- AHRQ:** Agency for Healthcare Research and Quality
CINAHL: Cumulative Index to Nursing and Allied Health Literature
GEM: Grid-Enabled Measures
HAPI: Health and Psychosocial Instruments
PROMIS: Patient-Reported Outcomes Measurement System
VA: Veterans Health Administration

Edited by G Eysenbach; submitted 03.03.17; peer-reviewed by K Claborn, B Xie, J Zheng; comments to author 24.04.17; revised version received 14.06.17; accepted 25.08.17; published 11.10.17

Please cite as:

Wakefield BJ, Turvey CL, Nazi KM, Holman JE, Hogan TP, Shimada SL, Kennedy DR
Psychometric Properties of Patient-Facing eHealth Evaluation Measures: Systematic Review and Analysis
J Med Internet Res 2017;19(10):e346
URL: <http://www.jmir.org/2017/10/e346/>
doi: [10.2196/jmir.7638](https://doi.org/10.2196/jmir.7638)
PMID: [29021128](https://pubmed.ncbi.nlm.nih.gov/29021128/)

©Bonnie J Wakefield, Carolyn L Turvey, Kim M Nazi, John E Holman, Timothy P Hogan, Stephanie L Shimada, Diana R Kennedy. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 11.10.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.