Abstract: With the Internet becoming a growing source of health information, evaluating the quality of the web-based health information is essential. A number of consumer-oriented instruments were designed to assess the quality of health information. However, existent tools are incomprehensive, unvalidated, not easy to use, and difficult to understand. The main objective of this study is to provide a framework for systematically developing validated and usable instruments to evaluate the quality of web-based health information. The framework incorporates item-to-total correlation and factor analysis. Next, weighting scores for the quality criteria are assigned based on the application of the Kano model to classify quality criteria as basic, performance or attractive. The development process ends with testing the instrument in terms of its usability, readability, reliability and validity. Among the implications and contributions of this research is improving the quality of the evaluation instruments and making the process of developing assessment tools more standardized and transparent.

Key Words: E-Health, Information quality, Evaluation, Kano, Satisfaction with information quality, Validity

INTRODUCTION

Information quality is an abstract concept that can be evaluated by “its fitness for use” [46] based on the context of using the information [43]. The concept of information quality was found as a key determinant of quality of decisions in different settings such as e-commerce, [32, 44], e-government [1], e-learning [39], and e-health [12]. With regards to healthcare, information quality is of particular high value as the decisions taken based on information quality may affect human life and health. The e-Health Code of Ethics defines health information as:

“Health information includes information for staying well, preventing and managing disease, and making other decisions related to health and health care. It includes information for making decisions about health products and health services. It may be in the form of data, text, audio, and/or video. It may involve enhancements through programming and interactivity.” [45, para 1].

With the fast development and increasing use of the World Wide Web, health information seeking online has grown substantially to form one of the most popular activities. There are about 64000 health websites and lists [41]. In 2010, about 76% of U.S. adults looked online for health information online from 67% in 2009 [23]. Consumers search online for a variety of health information such as medical diseases, health problems, medical treatments and nutrition [36]. However, such growth in web-based health resources was accompanied with inadequate information quality [13, 20]. Patients may access websites that include inaccurate and no credible information which may lead to confusion and poor decision making. A case of hydrazine sulfate poisoning was caused by inaccurate and exaggerated claims of treatment effectiveness.
and lack of information on side effects [20]. A study by the U.S. Department of Health and Human Services found that only about 4% of frequently visited health-related websites offered the source of their information [36].

Several strategies have been designed to help health information seekers access high-quality information, including codes of conduct, health portals and evaluation instruments [48]. The development of evaluation instruments has received the greatest attention since it has a number of benefits. First, evaluation instruments can educate health information consumers on the criteria of information quality. Second, as health portals\(^1\) (e.g., Healthfinder) form one important resource for health information, evaluation instruments can be used as a tool for choosing high-quality health websites to link to the health portals. Third, information intermediations such as librarians and health-care providers, can apply the evaluation instruments in guiding the health information seekers to high-quality websites. Evaluation instruments work on the premise that they can identify quality sites on the assumption that sites that conform to indicators of quality are likely to contain accurate information [7]. Information accuracy is defined as the information that is based on the gold standard information in the field. As information accuracy assessment by healthcare consumers who lack the domain knowledge is not possible, using the instruments may help in predicting information accuracy. Increasingly, healthcare providers and healthcare information seekers are faced with the challenge of assessing web-based health information. Physicians are unwilling to recommend specific websites before spending a great deal of effort on evaluating online content using a rating tool. A considerable number of evaluation instruments were developed to be used by healthcare professionals and consumers. However, these instruments suffer from incompleteness [12], difficult in application [7], questionable validity and reliability [12, 13] and inappropriate readability level [3]. In this paper we propose a multi-step process to develop better evaluation instruments by incorporating different techniques for assessing the evaluation instrument itself.

**REVIEW OF LITERATURE**

According to Wilson and his colleagues [48], there are five categories of tools to rate health information quality, of which three were suitable for application by healthcare consumers: codes of conduct, quality labels and the evaluation instruments\(^2\), where the latter is the focus of this paper. An evaluation instrument is defined as a rating tool that is composed of one or more criteria, each may consists of one or more attributes. An attribute is an item of information that must be assessed in order to evaluate compliance with the criterion [3]. For example, the widely known evaluation instrument, DISCRN [9] is a 16-item tool that provides health information consumers with guidelines for appraising information using two general meta-criteria, the reliability and the quality of information on treatment choices. Therefore, in order to assess the reliability dimension, the user has to answer eight questions of: (1) are the aims clear, (2) are the aims achieved, (3) is the content relevant, (4) is it clear what sources were used, (5) is it clear when the information used was published, (6) is the content balanced and unbiased, (7) does it provide details of additional support and information, (8) does it refer to areas of uncertainty. Similarly, the Information Quality Tool (IQT) [34] that is developed to assess the content of web-based smoking cessation websites, evaluates information quality on dimensions of authorship (items 1-7), sponsorship (items 8-10), currency (items 11-13) accuracy (items 14-15), confidentiality (item 18) and navigability (items 19-21). Several attempts have been made to develop consumer-oriented evaluation instruments. Some of these instruments serve as general guidelines for any type of health information (e.g., DISCERN), while other tools were developed to assess whether the information meets the needs of a

---

1 Health portals or gateways are collections of sites that have been prescreened, evaluated and approved by the governing organizations
2 Some studies term evaluation instruments as user guide systems
particular groups of specific health topic (i.e., patients affected by specific disease). Among the domain-specific tools, instruments to rate the information quality of multiple-sclerosis [22], Alzheimer’s disease [8], diabetes [40], breast cancer [27], online medicines information [37], and genetics [41]. In contrast to the large number of health evaluation tools, they suffer from a number of issues that limit their use and therefore their longevity.

**Metrics to assess the evaluation instruments**

Despite the large number of evaluation instruments, several criticisms have been directed toward the existent evaluation instruments in terms of following issues.

**Comprehensiveness**

Each evaluation instrument should contain a number of meta-criteria identified by major health organizations. For example, the Health on the Net Foundation (HONcode) [5] eight quality criteria to be included in evaluation instruments: authoritative, complementarity, privacy, attribution, justifiability, transparency, financial disclosure, and advertising policy. Similarly, the Health Improvement Institute and Consumer Reports WebWatch (HIICRW) [25] defined nine key criteria that any assessment tool should include: content relevance, content accessibility, content selection, content validity, content interchange, site transparency, links, quality assurance and safeguards. Another example is the code of conduct of the American medical Association (AMA) [for more details see 38]. One of the recognized problems in current evaluation tools is including relatively few of the items described by such initiatives [7, 27].

**Readability**

The extent to which a written material provides useful information depends highly on the readability of the information communicated. Readability level of a written document can be assessed using a number of formulas such as the Flesch-Kincaid Grade Level [33] and Flesch Reading ease [15] scores. Readability formulas assess the difficulty of the reading level by examining factors of word length, sentence length and complexity of vocabulary [41]. Past research found that six evaluation instruments out of seven are of readability level that exceeds the recommended readability level of the 8th grade [3]. Thus, average users find it difficult to understanding how to use the evaluation tools.

**Usability**

Terms of ease of use or feasibility were used in literature to test the usability of the evaluation instruments [3, 7, 27]. Some attempts tend to measure this factor objectively by examining the time that is taken in using the evaluation tools [e.g., 27], while others test the usability of the instrument subjectively by asking the users to rate a number of questions related to the domain of ease of use [e.g., 7]. The usability of an instrument depends on a number of factors such as number of questions, quality criteria included and type of questions. Instruments to be usable need to have (i) a limited number of quality criteria, (ii) assessable criteria and measures, (iii) and an instrument need to be available online [3]. In order to make evaluation instruments useful to healthcare consumers, it is important to ensure its usability. Among the 273 evaluation instruments, only 80 (29%) instruments publicly disclosed their criteria and only 54 of them had a reasonable number of quality attributes [3]. Other studies, found that some of the evaluation instruments are no more available online [17].

**Reliability and Validity**

A limited number of the evaluation instruments include information describing the development process and to what extent its reliability and validity were tested. The study of [7] found that only the DISCERN and the Minervation tools are the ones that provide information about their reliability and validity. Also,
[17] found that 98 tools have been used to rate the quality of online health information were apparently unvalidated.

The abovementioned recognized problematic issues in evaluation instruments call for a framework that can be applied in constructing instruments with standardization in process and better transparency in development. Additionally, the instrument process should include validation techniques that link the quality criteria to an output measure (e.g., usefulness, satisfaction, behavior change intentions). Satisfaction with health information quality is a key factor that influences the use of health websites for self-education and health behavior change [11]. The next section presents the Kano model [30] on consumer satisfaction. The Kano model is used to classify quality attributes and assigning different weighting scores based on their relation to the consumer satisfaction.

**The Kano Model**

Based on the Herzberg’s motivation-hygiene theory [26], Kano, a Japanese management researcher was motivated to develop his model that is linearly linked to satisfaction (see Figure 1). According to the Kano model [30], there are three levels of quality expectations for products and services that must be met. The three levels are: (i) Basic; (ii) Performance and (iii) Attractive features. Basic quality is the minimum acceptable to the consumers and encompasses those attributes that consumers take for granted. Their presence is not noticed, but their absence will result in complaints and thereby, dissatisfaction. Performance quality is the consciously stated needs. The presence of performance criteria is noted, while their absence results in disappointment or disadvantage. Lastly, the Attractive quality criteria are the ones that delight consumers. For example, when online users shop online (i.e., online shopping), they expect the website to have a visual design, downloads fast and not to have any dead links [48]. These quality criteria are taken as granted and their absence result in complaining and switching behavior [44]. Features such as ease of navigation and comprehensiveness of information are performance measures [48] that are consciously stated needs and influence significantly their satisfaction [44]. On the other hand, features that consumers do not have a conscious need for are considered attractive features. An example of an attractive feature in the online shopping setting would be multimedia usage in the website, which its presence will be noted and would result in delighting the online shoppers [48].

As the case of the website quality features, it is expected that some of the information quality criteria are of greater importance than others. Thus, when measuring the quality of information, different criteria should be assigned different weighting scores corresponding to its importance. Thus, it is expected that integrating the Kano model in the development process would result in more accurate evaluation of the information quality.
A Framework for Assessing Health Information Quality: The Application of the Kano Model

The main purpose of developing evaluation instruments is to provide information consumers with instruments that can help them in selecting high-quality health information. Figure 2 is a diagram that describes a procedure for constructing health information quality evaluation instruments. The proposed framework is composed of a number of steps, and an iterative process.

Health domain specification

At this phase, instruments’ developers should specify the domain of using the instrument (e.g., nutrition, disease prevention, healthy life habits... etc.). This step is important as it will identify the literature to be used for generating list of quality attributes.
**Information quality criteria generation**

Generating the list of the quality criteria should be conducted using the existent literature and knowledge [10]. Developers of the instrument are encouraged to consider conducting focus group research with healthcare providers, health information consumers and patients to include different perspectives. Each quality criteria should be operationalized by a number of items. This issue is important when using the internal consistency as a measure for the instrument’s reliability [10, 21].

**Verification of comprehensiveness**

The HONcode [5] provides guidelines to ensure the comprehensiveness of the evaluation instruments. Thus, experts need to examine the generated list if quality criteria to ensure the inclusion of the eight dimensions of: authority, complementary, confidentiality, attribution, justifiability, transparency of authorship, honesty in advertising, and editorial policy. Another approach is following the recommendations of the Health Improvement Institute and Consumer Reports WebWatch (HIICRW) [25]. The HIICRW defines nine key criteria that any assessment tool should include: content relevance, content accessibility, content selection, content validity, content interchange, site transparency, links, quality assurance and safeguards. Based on the intended use of the evaluation instrument, a panel of experts should review the pool of items for clarity, conciseness, and relevance to construct [10].

**Application of the Kano model**

The Kano model [30] should be applied by collecting the appropriate subject’s perspectives on the type of each of the quality criteria. This will enable the instrument developers to distinguish among the basic, performance and attractive quality criteria. Each of the quality criteria should be assigned a weighting score based on its type as basic, performance or attractive. At this phase, subjects would be presented with the criteria and its definition. For each quality criteria a pair of questions is formulated to which the consumer can answer in one of five different ways [30]. The first question concerns the reaction of the consumer if the health information has that quality criterion (functional form of the question), the second concerns his reaction if the health information does not have that quality criterion (dysfunctional form of the question). Based on combing the two answers, the quality criterion can be classified as either basic, performance or attractive. For criteria types of basic, a value of three is assigned to its weight score, for criteria of type performance a weight of the value two is assigned. The attractive criteria are assigned a weight of the value of one. This step would help the instrument developers to better classify the quality criteria and assigning weighting score that corresponds to its type.

**Initial administration to a development sample**

After generating the list of quality criteria and identifying its type as basic, performance or attractive, the initial instrument should be tested by administrating it to some subjects. The sample of subjects should be large and representative to the targeted population [35]. Testing the instrument would result in enhanced version by eliminating some items for redundancy and ambiguity.
Domain Specification

Quality criteria generation

Verification of comprehensiveness

Application of the Kano model

Data collection

Items purification

If Items per criteria < 3

Generating quality items

Data collection

Reliability and validity

Usability assessment

Readability assessment

Documentation

Recommended techniques

- Literature review
- Focus groups
- Delphi method

HONcode

Reliability analysis
- Item-to-Total Correlation
- Factor analysis

Coefficient alpha
- Criterion validity

Time taken in using instrument
- Subjective measures

Flesch Reading Ease
- Flesch-Kincaid Grade Level

Figure 2. Proposed framework for developing better information quality instruments
**Data collection**

Next to identifying the meta-criteria of quality to include in the instrument, the items should be arranged in a questionnaire format. Data collection should result in acceptable response rate and number of complete responses that would enable analytical work [10].

**Items purification**

At this phase, analytical techniques such as reliability analysis and factor analysis are used to drop and purify measures. It is essential to maintain at least three items per quality criteria [21]. In case, quality criterion is represented by less than three items, instrument developers need to elicit other items that are related to the quality criteria. Then, item purification should be conducted again to ensure the factor structure of the criteria [21]. This iterative process continues until factors are represented with a sufficient number of items and stable structure [21]. Next step requires collecting another set of data. The second phase of data collection is conducted to test the instrument in terms to its reliability, validity, usability and readability.

**Reliability assessment**

The internal reliability of the evaluation tool can be assessed using the Cronbach’s alpha test [10]. Cronbach’s alpha value should not be less than the cut-off value of 0.70 [35]. Another measure for internal reliability is the composite reliability index [24]. Values of the composite reliability index should not be less than the recommended 0.70 cut-off value [24].

**Validity assessment**

Validation of the constructs is a critical step to ensure the generation of a scientifically valid knowledge. Valid measures “represent the essence of content upon which the entity or construct is focused. They are unitary. They are not easily confused with other constructs. They predict well. If they are supposed to manipulate the experience of subjects, they do so.” [42, p. 5]. The demonstration of the instrument’s validity is even more important when researchers attempt to understand the links between measures of quality criteria to other variables. In testing validity, the three types of content validity, convergent validity and divergent validity should be assessed [10]. Face validity is ascertained as the criteria items are generated using existing studies, focus groups and the Delphi method. In testing the convergent validity, a number of statistical analyses can be used such as reliability analysis, correlation analysis and the Multitrait-Multimethod Matrix [10]. Discriminant validity is the “degree to which two conceptually similar constructs are distinct” [21, p. 118]. One way to assess the discriminant validity of the constructs is by using the confidence interval test [24]. This test involves calculating the confidence interval of plus or minus two (±2) of the standard error around the between the factors. The discriminant validity is determined if the confidence interval does not include the value of one [24].

**Usability assessment**

As the developed instrument to be used by healthcare information consumers, healthcare providers and patients, it is important to ensure its ease of use. Usability of the instrument can be assessed objectively by measuring the time taken in using the instrument [7]. Also, based on methods adopted from the
human-computer interaction body, usability testing should include the subjective evaluation of the instrument [4]. Evaluation instruments not only need to be usable, but also accessible online to its intended users.

**Readability assessment**

Readability level is tested by using at least the readability formulas of Flesch-Kincaid Grade level [33] and the Flesch Reading Ease Index [15]. In calculating readability level of web-based text, it is important to edit the text and remove websites addresses, headers, dashes, parentheses, slashes and other punctuation. Also, lists should be re-written as sentences. The recommended score for the Flesch-Kincaid Reading level for consumer-oriented health materials is the 8th grade [18]. The Flesch Reading Ease score is one of the most widely used scoring systems to measure readability of the documents. Documents scoring 70 or above are described as “easy” and are written at the grade school level. A score of 60 to 70 is described as “standard” and written at approximately the high school level. Scores below 60 are described as “fairly difficult”, “difficult,” or “very difficult” as the score decreases. For consumer-oriented information, it is recommended to maintain a score of 60 or greater [3]. Other formulas can also be used including the SMOG Readability Formula, the Fry Readability Graph, the Gunning-Fog formula, and the Lexile Framework [for more details see 13]. However, other factors may affect the readability level that which not considered in readability formulas (e.g., use of active voice, language, motivational, and culturally specific information) that are subjective in nature [14, 47]. Therefore, readability testing should be conducted objectively using readability formulas and subjectively testing the comprehension with actual users.

**Documentation**

This phase of the instrument’s construction aims onto adding transparency to the development process and providing guidelines on using the tool. Instrument developers need to document in details the quality domain, purpose of using the tool, quality criteria included, expert and customer-oriented inputs, and validation and reliability assessment. The documentation would be a set of guidelines for healthcare consumers on how to use the instruments and measure the quality of health information. In case the instrument is modified, the documentation should be updated to reflect such modification and the reasons behind it. Another phase of readability testing should be applied to the documentation to ensure that the guidelines are readable to its intended audience.

**PLANNED EMPIRICAL VALIDATION OF THE FRAMEWORK**

The next step in our research is to empirically test the proposed framework and the use of the Kano model in classifying the information quality criteria. The major research hypotheses are:

1. Criteria of information quality can be classified onto three types of basic features, performance features and attractive features.
2. Healthcare information seekers can be classified onto dissatisfied, satisfied and delighted based on the following:
   a. Subjects are dissatisfied due to the absence of the basic criteria which are the minimal information quality required to meet their needs when using health websites.
b. Subjects are satisfied due to the presence of the basic and performance criteria of information quality when using health websites.

c. Subjects are delighted due to the presence of basic, performance, and attractive criteria of information

3. The proposed framework will result in a better instrument that would meet the required metrics in terms of instrument’s comprehensives, usability, readability and validity.

IMPLICATION FOR RESEARCH

The multi-step proposed methodology for constructing evaluation instruments is expected to add transparency and standardization to the development process. This research is a significant step contributing to efforts to improve the quality and ethical standards of health information online. Using the proposed methodology of building instruments would ensure the comprehensiveness and validity of the instrument. This framework would also support research on understanding the influence on information quality attributes on patient health outcomes. One avenue of such application is to test the actual impact of the identified quality criteria on patients’ participation in web-based health behavior change programs and the effectiveness of these programs. Additionally, if our framework proves to be useful in identifying basic, performance, and attractive quality criteria, health professionals and practitioners would know what health information quality criteria increase the satisfaction of health information consumers and which ones delight them. As such, researchers and professionals in health communication would understand what criteria of health information quality are essential to avoid the information consumers’ dissatisfaction on the net.

REFERENCES


