This is a preprint of an article published in *JASIST* (Stvilia, B., Mon, L., Yi, Y. (2009). A model for online consumer health information quality. *JASIST, 60*(9), 1781-1791. http://dx.doi.org/10.1002/asi.21115).

A model for online consumer health information quality

Besiki Stvilia[1], Lorri Mon, Yong Jeong Yi

College of Information, Florida State University

101 Louis Shores Building, Tallahassee, FL 32306, USA

{bstvilia, lmon, yjy4617}@fsu.edu

Abstract

This article defines a quality model for online consumer health information consisting of five quality criteria constructs. These constructs are grounded in empirical data from the perspectives of the three main sources in the communication process: health information providers, consumers, and intermediaries, such as Web directory creators and librarians, who assist consumers in finding healthcare information. The article also defines five constructs of Web page structural markers that could be used in information quality evaluation and maps these markers to the quality criteria. Findings from correlation analysis and multinomial logistic tests indicate that use of the structural markers depended significantly on the type of Web page and type of information provider. The findings suggest the need to define genre-specific templates for quality evaluation and the need to develop models for an automatic genre-based classification of health information Web pages. In addition, the study showed that consumers may lack the motivation or literacy skills to evaluate the information quality of health Web pages, which suggests the need to develop accessible automatic information quality evaluation tools and ontologies.

Introduction

A widely used general definition of information quality (IQ) is the information's "fitness for use" (Wang & Strong, 1996). Specific definitions of IQ, however, are context dependent and dynamic. That is, the same information may be evaluated differently in different contexts and at different times (Strong, Lee, & Wang, 1997; Stvilia, Gasser, Twidale, & Smith, 2007). Furthermore, IQ is the key determinant of the quality of decisions and actions. Consequently, the value of IQ is evaluated by the value or cost of decisions, or by the lack of them (Marschak, 1971; Stvilia & Gasser, 2008). In healthcare, where decisions and actions can affect human life and health, the value of IQ can be particularly high (Gustafson & Wyatt, 2004; Institute of Medicine, 1999).

The Web is an important source for people who are seeking healthcare information (Hesse et al., 2005). The Pew Internet Project estimated that at least 75% of U.S. Internet users have searched for health information online, with 8 million Americans seeking health information online in a typical day (Fox, 2006, 2008). Not surprisingly, health information seeking and participation in online support groups are among the most common activities performed by patients. It is estimated that 63 million Americans visited health-related blogs and online community support groups, as well as ratings sites for prescription drugs and other health

[1] Corresponding author

materials (Lewis, 2008, p. 16). With the aging of the country's population, leading to an increased need for healthcare information, and with healthcare providers such as the government, hospitals, and foundations increasingly moving information services to the Web, consumer use of online health information has also increased (Fox, 2006).

Online health information supports decision-making for patients, caregivers, and healthcare providers. Online health searching helps in clarifying unfamiliar medical terms used in diagnoses, finding support groups, and locating possible alternative treatments (Sillence, Briggs, Harris, & Fishwick, 2007). Users of online health care information can include patients with chronic conditions, who are often expected to manage their diseases on their own between planned visits to medical providers such as outpatient clinics; low-income patients, who may not have health insurance; and rural patients, who live far away from healthcare facilities and may have to resort to self-care. The need for streamlining and easing the task of self-care for patients with chronic conditions has been recognized both by the government and in academia, where technologies have been proposed that include adaptive online questionnaires and wireless sensor devices to help in monitoring patients' health parameters or reminding patients of routine procedures (e.g., taking drugs) to avoid preventable complications that may result in emergency care (Harris, Wathen, & Fear, 2006; Sanders Berlin, & Schatz, 2008). Quality Web-based health information assists both patients and their caregivers in keeping up-to-date with research and practices related to the patient's illness, as well as in enabling access to moral and emotional support from the online community of fellow patients and caregivers.

The Web, however, is open to numerous kinds of publishers and information providers, and the quality of health information published on the Web is highly variant (Berland et al., 2001) and highly dynamic. Once published, information requires continuous quality evaluation and maintenance, especially in the dynamic field of healthcare, where the state of knowledge changes at a relatively fast pace (Phelps, 1992). In many research studies, clinicians reviewing the quality of information on health Web sites have found inaccuracies that raise concerns about the quality of information health consumers are encountering on the Internet (e.g., Bernstam et al., 2008; Eysenbach, Powell, Kuss, & Sa, 2002; Hardey, 2001). A survey of doctors on patient use of Internet health information found that the doctors estimated 44% of their patients had health problems because of using Internet information, whereas 85% of their patients were estimated to have benefited from online health information (Potts and Wyatt, 2002). In addition, Potts and Wyatt pointed out that most research on Internet health IQ has been from the perspective of experts and medical professionals, rather than that of health consumers. Bliemel and Hassanein (2007) called for more research on consumer perspectives regarding health IQ.

A typical healthcare information consumer may lack sufficient knowledge and training to evaluate the validity and quality of healthcare Web page content, and instead may have to rely on Web page surface markers and trust indicators to make IQ judgments. To enable more effective and efficient decision making by consumers, it is important to improve end-user healthcare IQ assessment on the Web. In particular, there is a need to develop an empirically grounded and potentially automatable model of healthcare IQ that could allow consumers, caregivers, and others (such as librarians) who assist with online health information searching to identify quality health information more easily and effectively. In developing a model for online health IQ judgments, this research addresses the problem of how to assist healthcare consumers in evaluating online healthcare IQ systematically, thereby supporting patients and caregivers in selecting higher quality Web-based healthcare resources to better meet their information needs.

Related Work

To evaluate IQ effectively and efficiently, an IQ assessment infrastructure is needed that would include IQ criteria, metrics, and reference resources; IQ ontology(s); and evaluation and monitoring services. A number of conceptual IQ criteria sets have been proposed in the general IQ literature. For example, using an empirical approach (a user survey), Wang and Strong (1996) developed a taxonomy of IQ dimensions grouped into 4 categories: (1) *Intrinsic:* Accuracy, Objectivity, Believability, and Reputation; (2) *Accessibility:* Access, Security; (3) *Contextual:* Relevancy, Value-Added, Timeliness, Completeness, Appropriate Amount of Data; (4) *Representational:* Interpretability, Ease of Understanding, Representational Consistency, Concise Representation. In healthcare informatics, Charnock, Shepperd, Needham, and Gann (1999) used a similar approach to develop an IQ assessment instrument or questionnaire by having healthcare domain experts develop a set of questions for a questionnaire divided into 3 sections: Reliability, Coverage, and Overall Quality. After reviewing 79 empirical studies of consumer health information on the Web, Eysenbach et al. (2002) found Accuracy, Completeness, Readability, Design, Disclosures, and References as the most frequently cited quality criteria.

To operationalize conceptual IQ models and criteria effectively through questionnaires or metrics access is needed not only to the information itself, but also to the metadata of the processes of its creation, maintenance, and use (Stvilia, Gasser, Twidale, Smith, 2007). Often, access to the behind-the-scenes metadata and policy information is not available, and members of the government and healthcare community have recognized this problem. The goal of the Healthy People 2010 Information Access Project at the U.S. Department of Health and Human Services (DHHS, 2007) has been to increase the proportion of health-related Web sites that disclose information that can be used for assessing their quality. The project identified six properties or types of metadata essential for carrying out an IQ evaluation of a health Web site: (1) the identity of owners, developers, and sponsors; (2) the purpose of the site; (3) the sources of the content; (4) the privacy and confidentiality of personal information; (5) evaluation or feedback mechanisms, and (6) content update procedures. Interestingly, a survey of 102 Web sites conducted by the same project found that none of the healthcare Web sites provided all this information, and less than 4% of the Web sites disclosed the sources of their content and how it was updated.

Likewise, prior research has revealed inconsistencies in how healthcare consumers evaluate the quality of online information. For example, Fox (2006) found that 75% of health information seekers did not consistently check online healthcare information for basic IQ indicators, such as the publication date or the source of the information. A laboratory experiment by Eysenbach and Köhler (2002) found that although users described a Web site’s source, professional design, formal or official appearance, language, and ease of use as the criteria they used to evaluate the quality of healthcare Web pages, observations in an actual information retrieval experiment indicated that none of the users actually examined Web pages for these quality cues. These findings point to a possible trade-off between quality and cost in terms of the time spent by users in evaluating quality, and indicate the contextual nature of quality evaluation. Earlier studies showed that the same information could be evaluated differently in different circumstances and by members of different age and social groups (Fox & Rainie, 2002; Stvilia & Gasser, 2008). Consumers might not think it worthwhile to check the quality of the pages in an experiment, but they might behave differently if they were patients seeking information that could have a direct impact on their lives. This also points to the need for additional research identifying the consumer value structure for health IQ and consumers’ motivations for IQ assessment, or the lack of it.

Indeed, consumers might not be able to assess the quality of health information directly. Instead, they might rely on indirect quality dimensions, such as reputation and trust markers, to estimate or predict the quality of a Web site. Zucker (1986) identified three sources of trust production: (1) the similarity of personal or social characteristics (Characteristics Based); (2) the record of past performance or encounters (Process Based); and (3) social institutions and intermediaries (Institution Based). Bailey, Gurak, and Konstan (2001) built on the model proposed in the literature by Zucker and others by synthesizing a taxonomy of trust dimensions and sources. The sources of trust they established included Presumptions, Surface Inspections, Experience, and Third-Party Institutions. The dimensions of trust, on the other hand, consisted of Attraction, Dynamism, Expertness, Faith, Intentions, and Localness.

Effects of online trust-building mechanisms on the overall trustworthiness of a Web site might differ in different circumstances. Chang and Cheung (2005) showed that a third-party certification was the most effective way of increasing the trust of consumers in an online vendor when the vendor's reputation was unknown ahead of time. At the same time, another trust-building mechanism, the return policy, had a significant effect on the vendor's trustworthiness only when the consumer was aware of the vendor's reputation. By inference, this might suggest that the effects of trust-building mechanisms may differ with the type of online information provider; consequently, different providers (e.g., the government, commercial sites) might use different cues to convey trust to consumers.

The utility of Web page trust markers and third-party endorsements was investigated in the context of using Web sites to answer library "ready reference" questions. Frické and Fallis (2004) examined the relationships between some quality indicators (e.g., a copyright sign, citations, a lack of advertising) and Web page accuracy. They found that although the presence of a copyright sign and the currency of a Web page correlated positively with accuracy, the rest of the quality indicators did not. Furthermore, their analysis showed that most of the sampled Web pages contained accurate answers to ready reference questions. In an earlier study, the same authors examined the use of quality indicators in health Web pages and found that displaying the Health on the Net (HON) Code logo, having an organization domain, and displaying a copyright were positively correlated with accuracy (Fallis & Frické, 2002). Another related study found positive correlation between Google's PageRank scores and accuracy of health Webpages (Frické, M., Fallis, D., Jones, M., & Luszko, G., 2005). A growing number of companies provide third-party rating services to Web sites (e.g., HON, Trustee, the Internet Content Rating Association (ICRA)), with the majority of them following the Platform for Internet Content Selection framework (Resnick & Miller, 1996), which proposed a general architecture and syntax for labeling Internet content, including the use of controlled vocabulary. However, some concerns about the proliferation of healthcare rating services on the Web have been raised by researchers such as Gagliardi and Jadad (2002), who noted that many healthcare rating services that existed in 1997 had disappeared by 2002, yet dozens of others had appeared in the interim. Further, few of these rating services revealed the criteria behind their ratings of the quality of health Web sites.

There have been attempts to develop applications for evaluating the quality of health Web sites automatically. Wang and Liu (2007) built a tool to detect some of the components (author name, references, etc.) of a Web page that could be used for evaluating its reputation or authority. The same corpus of data was used both for training and evaluating the parser, and components with a high degree of detection errors were removed from the set. Although Wang and Liu reported high precision and recall in the ability of the tool to identify the components, it is not known how well the tool would perform on a different data set. In an earlier study, Griffiths, Tang, Hawking, and Christensen (2005) used a domain-specific set of weighted key words to assess the evidence-based quality of depression Web sites based on content similarity. They reported that the scores produced by the automatic tool were correlated strongly with the manual assessment of the quality of the Web sites. These prior efforts suggest that an automated tool for assisting users with the evaluation of online healthcare Web sites is feasible, and provide some guidance toward a generalized healthcare IQ model.

The quality of information can be further affected not only by changes in the underlying knowledge and reality, but also by changes in the representation of this reality (e.g., malicious edits of a document or Web page), and by changes in the context of the interpretation and use this information (Stvilia et al., 2007). An information provider might not be motivated to document information that is thought to be common knowledge in the provider's local context (Stvilia, Gasser, Twidale, Shreeves, & Cole, 2004). Furthermore, in some instances, the provider might purposefully degrade IQ to trigger desirable actions or responses from the consumer (Stvilia, Twidale, Smith, & Gasser, 2008). Meredith, Emberton, Wood, and Smith (1995) found that, in some instances, physicians might prefer not to provide the patient with the complete information about a particular disease or the possible outcome of a treatment to avoid or reduce anxiety on the part of the patient. Hence, a healthcare information resource (e.g., fact sheet, leaflet, or Web page) evaluated by a professional as being of high quality may not be viewed as such by a patient. The research further showed that healthcare materials often were designed to the context of the provider, rather than that of the patient, in terms of having goals for selling medications, promoting particular treatments, or using vocabulary more conducive to communicating with health professionals rather than health consumers.

Research Questions

The purpose of this study was to define an empirically grounded model of consumer healthcare IQ assessment. The research sought to identify the quality of healthcare information from three perspectives: Web providers of health information in terms of how IQ is signaled on health Web sites; consumers of health information, their health questions, and their perceptions of quality indicators; and intermediaries such as librarians and Web directory creators, whose criteria for IQ are applied in evaluating and selecting health Web sources. In particular, the study aimed to address the following research questions:

- What are the “virtues” or criteria considered to be important when evaluating the quality of healthcare information?
- What are the quality markers that providers may use to signal IQ, and are these markers related?
- What are some of the types of healthcare information Web pages and providers?
- Does the use of quality markers vary with the type of Web page and the type of provider?

Procedures

The study used a mixed methodology with multiple data sources. In particular, the researchers analyzed the healthcare informatics literature to identify the types of activities that use healthcare information, the types of IQ problems, and the sets of quality criteria, markers, and metrics. The findings of the literature analysis were combined with the IQ criteria set from the general framework of IQ measurement developed earlier (Stvilia et al., 2007). This aggregate set comprised an initial conceptual model for healthcare IQ. The model was then used in the analysis of empirical data and was iteratively revised.

The researchers examined a random sample of 150 Web pages selected from a set of links harvested from the Yahoo! Directory’s “Health: Diseases and Conditions” subdirectory. The total number of links harvested at the time of data collection (May 11, 2008) was 10,961. Sampled Web pages were content analyzed and coded for Web page types, provider types, and quality markers. The researchers began with a schema aggregating the IQ indicators referenced in the literature (e.g., Frické & Fallis, 2004; Wang & Liu, 2007) and then modified it iteratively in coding the sample. After resolving coding differences, the final version was used to recode the sample. To categorize Web pages by type, the researchers used open coding, and the pages were analyzed and clustered based on intended uses, functionality, and form (Haas & Grams, 1998; Yoshioka, Herman, Yates, & Orlikowski, 2001).

Additionally, the researchers analyzed a sample of 150 e-mail transcripts of questions asked by health information consumers in the Internet Public Library (IPL; http://ipl.org), an online digital library question-answering (Q&A) service. The sample of health consumer questions was randomly selected from health and medicine questions (320 e-mail communication instances) in the IPL’s Q&A service archives from 2005 to 2007 (17,205 e-mail communication instances). Half the sample was drawn from questions answered by the IPL librarians, and the other half was sampled from transcripts of rejected questions not answered by the librarians because they were “out of scope” for the service (such out-of-scope rejected questions might include asking librarians to diagnose a health problem based on symptoms—questions that should not be handled by librarians, but rather by medical professionals). Two researchers used content analysis to code the question transcripts for the types of questions asked, information sources used, intended uses of the information, and quality criteria as articulated in the questions. Each researcher open coded the complete sample independently. After coding was completed, the resultant schemas were aggregated and differences were resolved, and the researchers used the aggregated final schema to recode the entire sample.

Finally, to gain better insight into the health IQ evaluation behavior of consumers, the researchers surveyed a convenience sample of 108 healthcare information consumers. Survey participants were given the aggregate set of IQ criteria assembled through the previous phases of the study and were asked to rank them in order of importance to their healthcare information-seeking and selection tasks. The survey was then followed by semi-structured in-depth interviews. A sample of 20 survey participants, stratified by age groups from ages 18 to 65, were selected for the follow-up interviews by using a semistructured interview procedure. The sample consisted of 5 participants between the ages 18 and 30, 5 participants between the ages of 30 and 40, 5 participants between the ages of 40 and 50, and 5 participants above age 50. In the interviews, the critical incident technique (Flanagan, 1954) was used, in which participants were asked to recall a specific incident in which they had sought healthcare information, and to describe their judgments of the information found.

Findings

Only 35% of the Yahoo! Directory sample had some form of quality evaluation criteria set posted or linked explicitly to the Web page. The analysis of those Web pages and related Web sites identified the following major approaches the providers might use to define their IQ criteria set: *centrally defined*, *community constructed*, and *outsourced to third-party raters*. The majority of the government Web sites and Web pages referenced the general quality criteria included in the guidelines of the Office of Management and Budget (OMB, 2002). The OMB guidelines implement Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (P.L. 106-554), also known as the "Data Quality Act." The OMB requires government agencies to develop an IQ standard and adhere to that standard in their information dissemination practices. The model of the OMB consists of three general conceptual criteria: objectivity, utility, and integrity (see Figure 1). The OMB guidelines do not include an implementation model for these criteria, however. Individual federal agencies are left to develop specific implementation mechanisms for each of the criteria based on the nature of the information each agency disseminates, its tasks, and its responsibilities. Most of the government healthcare information pages did not specify their IQ criteria, but instead pointed to the quality guidelines posted at the DHHS (2006) Web site. The only exception was MedlinePlus (2007), which specified its own model for quality, consisting of 10 criteria (see Figure 1).

Examples of the community-defined quality model were Wikipedia pages. It is important to note that at the time this article was written, Wikipedia did not have a separate IQ model for its healthcare-related articles, but rather used a general model (see Figure 1). A detailed description of Wikipedia's IQ assurance model can be found elsewhere (Stvilia et al., 2008).

Finally, some of the Web pages carried seals of approval from third-party rating agencies as a sign of adherence to their quality principles. The most frequently occurring seal was from the Health On the Net Foundation (HON). The HON principles (the HON Code) consists of seven general principles, which then are further detailed into specific operationalization guidelines. In addition, the HON Code quality guidelines contain sections for both "closed" (centrally controlled) and "open" (collaborative or community-based) Web sites.

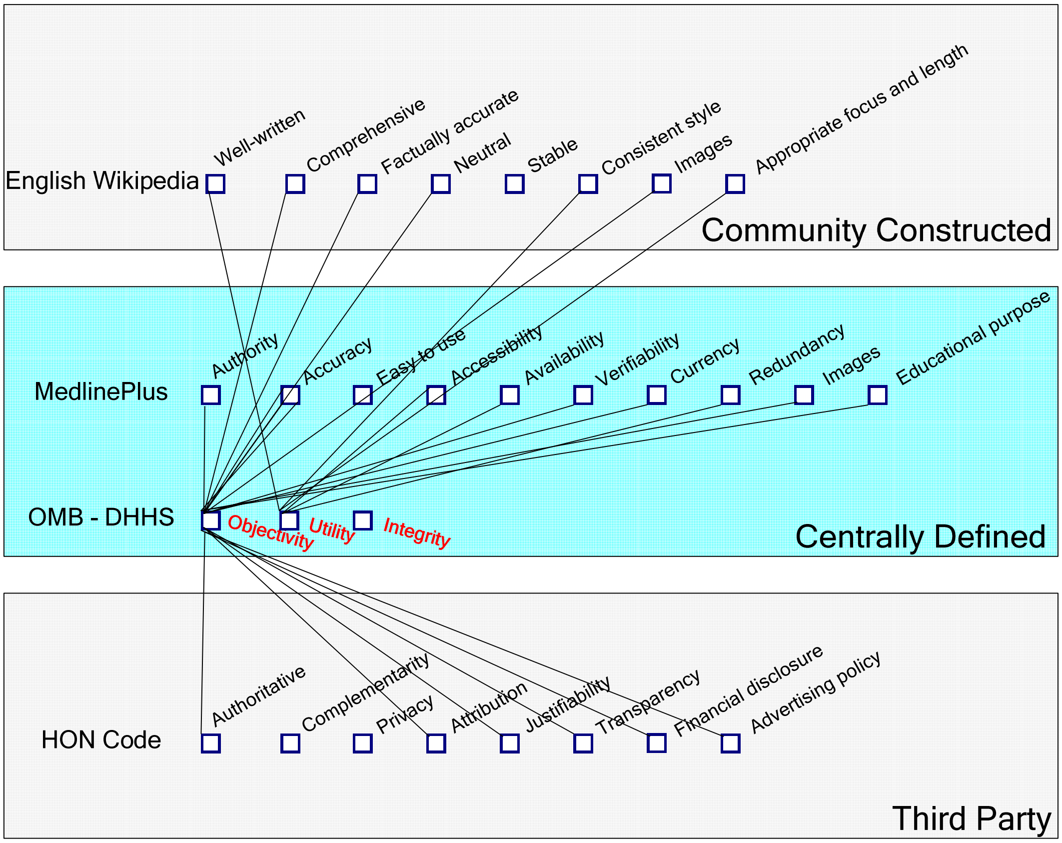

Figure 1. Information quality criteria and dimension mapping. OMB = Office of Management and Budget; DHHS = U.S. Department of Health and Human Services; HON = Health On the Net.

To identify the healthcare IQ criteria used by consumers and information intermediaries, the researchers content analyzed a sample of the IPL's Q&A communication archives. A detailed report of the analysis of Q&A transcripts from the IPL is provided elsewhere (Mon et al., in process). The analysis identified seven IQ criteria referenced by IPL users and volunteers that were relevant to healthcare IQ judgments: *Accuracy*, *Authority*, *Completeness*, *Currency*, *Objectivity*, *Relevancy*, and *Understandability*.

The set of IQ criteria extracted from the Yahoo! and IPL samples was aggregated with the set of IQ criteria proposed by Stvilia et al. (2007) and used to develop the survey questionnaire. The questionnaire was distributed to a convenience sample of 108 participants, aged 18 to 65. In the questionnaire, the participants were asked to rank the criteria in order of importance to their healthcare information-seeking and selection tasks. The questionnaire used a 5-point Likert scale, with a survey completion rate of 74%. Descriptive statistics showing the ranking of the criteria are given in Table 1. On average, the participants ranked the Accuracy dimension as the highest and the Volatility dimension as the lowest.

In addition, when the survey participants were asked what quality markers they checked for on healthcare Web pages, authorship information ranked highest, at 80%. Only 25%—the lowest share of the survey participants—reported checking whether the Web site had an internal or external quality review process. Furthermore, an overwhelming majority (74%) of the survey respondents named general search engines as their way of finding health information, whereas only 9% of the respondents indicated they used Web portals to locate health information (Table 2).

Table 1

The Ranking of Information Quality Dimensions by Importance

Dimension	***N***	**Mean**	**Median**	**Mode**	**Std. deviation**
Accuracy	76	4.66	5	5	0.60
Reliability	76	4.57	5	5	0.60
Credibility	76	4.53	5	5	0.74
Trustworthiness	76	4.51	5	5	0.64
Clarity	76	4.20	4	5	0.78
Objectivity	76	4.08	4	5	0.93
Utility	76	4.03	4	4	0.91
Verifiability	76	3.95	4	3	0.88
Usefulness	76	3.92	4	4	0.83
Integrity	76	3.91	4	4	0.88
Ease of Understanding	76	3.91	4	3	0.93
Consistency	76	3.89	4	4	0.90
Relevance	76	3.83	4	4	0.99
Completeness	76	3.80	4	5	1.05
Currency	76	3.80	4	5	1.12
Authority	76	3.76	4	3	1.03
Lack of Bias	76	3.63	4	3	0.95
Accessibility	76	3.54	3	3	0.97
Ease of Use (Web site)	76	3.47	3	3	1.10
Cohesiveness	76	3.29	3	3	0.95
Volatility	76	3.28	3	3	1.00

To identify consumers' value structure for quality, factor analysis was applied to the survey respondents' rankings of quality criteria. Both the Bartlett and Measure of Sampling Adequacy (MSA) tests for the sample pointed to a significant level of correlation (Bartlett test: $\chi^2 = 757.2$, $p < 0.001$; MSA = 0.812) among the criteria. A scree plot suggested selecting the first 5 components. In addition, because of the sample size (80 participants), the cutoff size for the criteria loadings on the factors was set to 0.65 (Hair, Black, Babin, Anderson, & Tatham, 2005; see Table 3).

Table 2

How Consumers Find Health Information (80 Respondents)

Source	**Percentage**
General search engine	74
Family and friends	12
Web directory or portal	9
Television	9
Physicians	4
Printed periodicals, books	3

Table 3

Factor Loadings for the Information Quality Criteria

	Component				
Dimension	**1**	**2**	**3**	**4**	**5**
Accessibility	**0.69**	0.36	0.10	0.19	0.10
Accuracy	–0.02	0.04	**0.74**	–0.05	0.34
Authority	0.16	0.13	–0.08	**0.80**	0.18
Clarity	0.14	0.26	0.34	0.07	**0.75**
Cohesiveness	**0.76**	0.00	0.11	0.29	0.13
Completeness	0.24	0.03	0.10	0.25	**0.69**
Consistency	**0.78**	0.17	0.15	0.03	0.03
Credibility	0.16	0.25	**0.70**	0.26	–0.17
Currency	0.56	0.48	–0.10	0.06	0.26
Ease of Use (Web site)	0.54	**0.63**	0.14	–0.09	–0.02
Ease of Understanding	0.29	0.49	0.22	0.09	0.18
Integrity	0.57	0.11	0.40	0.07	0.23
Lack of Bias	0.47	0.27	0.02	0.53	0.27
Objectivity	0.14	**0.69**	0.15	0.30	0.04
Relevance	0.57	0.33	–0.13	–0.05	0.48
Reliability	0.34	0.13	**0.68**	–0.03	0.14
Trustworthiness	0.01	0.42	0.31	0.57	–0.17
Usefulness	–0.02	0.54	0.31	0.17	0.17
Utility	0.25	**0.75**	–0.03	0.10	0.11
Verifiability	0.50	0.08	0.19	0.55	0.23
Volatility	**0.69**	0.16	0.12	0.20	0.10

Note. Extraction method: principal component analysis; rotation method: varimax with Kaiser normalization.

The criteria loaded on the first factor were mostly access related (see Table 3). The second factor construct included usefulness criteria. The third factor construct had both accuracy and trust-related criteria, with Accuracy criteria having the highest loading. The fourth factor had a single criterion—Authority. The criteria loaded on the firth factor construct could be categorized as related to Completeness. The criteria constructs were then ranked by the averages of their loading rankings (see Table 4). The Accuracy construct was ranked the highest, followed by the Completeness construct.

Table 4

IQ Criteria Constructs

IQ criteria constructs	**Ranking**	**IQ criteria**
Accuracy	4.41	Accuracy, Credibility, Reliability
Completeness	4.17	Completeness, Clarity
Authority	3.8	Authority
Usefulness	3.75	Ease of Use, Objectivity, Utility
Accessibility	3.57	Accessibility, Cohesiveness, Consistency, Volatility

The analysis of the literature and the content analysis of the Yahoo! Directory sample suggested 23 document components or markers that could be used in IQ evaluation (see Table 5). In addition, the analysis identified seven types or genres of Web pages: *Article*, *Blog*, *Directory*, *Factsheet*, *Instrument*, *Mainpage*, and *Q&A*. The blog type consisted not only of blogs and patient stories or testimonials, but also of accounts and histories of disease or drug discovery and development. Likewise, the Factsheet type combined both disease factsheets and "what to do" guides (e.g., Centers for Disease Control travelers' guides). The

Instrument type included online questionnaires for disease risk assessment and diagnosis. The analysis also suggested five types of consumer health information providers: *commercial*, *nonprofit or community*, *government*, *patient and family*, and *research*. The nonprofit type included charities, associations, professional organizations, and societies. The research type covered both research institutions and individual project-based Web sites.

The different types of Web pages and information providers seemed to use different quality markers. A nonparametric Kruskal-Wallis test showed that the presence or absence of the majority of quality markers in the sample was significantly dependent both on the type of document and the type of provider (see Table 5). The quality markers found to be significantly related to the Web page types in the previous test were then regressed on the Web page types using multinomial logistic regression (model fit likelihood ratio: χ^2 = 217.62; $p < 0.0001$). The Article type was used as a baseline. The regression analysis confirmed that the markers were statistically significant in distinguishing some of Web page types from each other. For instance, the presence of the Copyright marker was a negative predictor of the Directory and Factsheet types over the Article type. In addition, holding all the other variables constant, having the Disclaimer component increased the odds of the Web page being of the Factsheet type rather than the Main Page type. A similar regression analysis of the quality markers to the provider types (model fit likelihood ratio: χ^2 = 199.58; $p <$ 0.0001) with the Nonprofit or Community provider type as a baseline also found the markers to be significant in distinguishing some of the provider types from each other. For example, the Third-Party Review and Disclaimer markers were positive indicators of the Commercial type, whereas the Date of Last Update was a negative indicator. The Copyright and Site Map markers were negatively related to the odds of the provider being of the Government provider type.

Table 5
Kruskal-Wallis Correlation Test of the Quality Markers on the Document and Provider Types (150 Cases)

	Document type			Provider type		
Document marker	χ^2	**df**	**α**	χ^2	**df**	**α**
About	*19.70*	*6*	*0.003*	6.24	4	0.182
Accessibility	10.22	6	0.116	*66.51*	*4*	*0.000*
Advertising Policy	11.89	6	0.064	*11.70*	*4*	*0.020*
Author Affiliation	*14.35*	*6*	*0.026*	6.51	4	0.164
Author Credentials	11.08	6	0.086	*21.21*	*4*	*0.000*
Author Name	*12.48*	*6*	*0.052*	*20.74*	*4*	*0.000*
Contact Us	*18.72*	*6*	*0.005*	6.85	4	0.144
Copyright	12.16	6	0.058	*17.06*	*4*	*0.002*
Date of Creation	11.37	6	0.078	3.60	4	0.463
Date of Last Update	*23.87*	*6*	*0.001*	*12.76*	*4*	*0.013*
Disclaimer	*20.11*	*6*	*0.003*	*9.43*	*4*	*0.051*
Editorial Review Process	*15.61*	*6*	*0.016*	2.50	4	0.644
Formal IQ Criteria	*49.37*	*6*	*0.000*	*60.90*	*4*	*0.000*
Payment	4.66	6	0.588	3.06	4	0.548
Privacy Policy	*18.66*	*6*	*0.005*	*16.21*	*4*	*0.003*
Provider Name	*16.65*	*6*	*0.011*	*64.75*	*4*	*0.000*
Quality Guidelines	*24.93*	*6*	*0.000*	*47.05*	*4*	*0.000*
Reference(s)	*52.02*	*6*	*0.000*	5.85	4	0.211
Search Box	*39.88*	*6*	*0.000*	*17.95*	*4*	*0.001*
Site Map	*12.73*	*6*	*0.048*	*15.49*	*4*	*0.004*
Sponsored Content	4.90	6	0.557	3.72	4	0.445
Statement of Purpose or Mission Statement	7.18	6	0.304	6.38	4	0.173
Third-Party Quality Seal	*17.78*	*6*	*0.007*	*23.13*	*4*	*0.000*
Terms of Use, Policies, and Regulations	*19.17*	*6*	*0.004*	*10.38*	*4*	*0.035*

Although the purpose of the study was not to investigate the feasibility of automatic classification of healthcare Web pages, the researchers nevertheless evaluated the strength of the IQ markers as a set in distinguishing the Web page types (e.g., Article, Blog, Directory) and provider types (e.g., Government, Commercial, Community). The WEKA implementation of the C4.5 Decision Tree classifier was used to classify the Yahoo! sample based on the quality markers. The sample cases were labeled by the Web page and information provider types, and a 10-fold cross-validation was used. In both cases, the classification accuracies were only slightly higher than random selection—60% for the Web page types and 63% for the information provider types.

To identify a possible latent structure underlying the IQ markers and to combine the markers into more independent or orthogonal groups, the researchers applied exploratory factor analysis to the sample. Both the Bartlett and MSA tests pointed to a statistically significant level of correlation (Bartlett test: $\chi^2 = 1{,}239.1$, $p < 0.0001$; MSA = 0.650) among the markers. The cutoff value for factor loadings was set above 0.4, the recommended value for the size of the sample (150 Web pages; Hair et al., 2005). The scree plot suggested selecting the first 5 components, which had the following factor loadings (see Table 6): (1) Provider Name, About, Third-Party Quality Seal, Copyright, Disclaimer, Privacy Policy; Search Box; (2) Author Name, Author Credentials, Author Affiliation; (3) Editorial Review Process, Quality Guidelines, Accessibility, Formal IQ Criteria; (4) Reference(s), Date of Last Update, Contact Us; (5) Terms of Use, Policies, and Regulations; Sponsored Content. The factor constructs did not include the Site Map, Advertising Policy, Payment, Statement of Purpose or Mission Statement, and Date of Creation markers, either because the loadings of these markers on the factors were below the cutoff value or because the markers contributed significantly to more than one factor. It is important to note that the five factors selected captured only slightly more than 50% of the total variance. Hence, a significant amount of variance was not covered by the structure.

Table 6
Factor Loadings for the Quality Markers

	Component				
Document marker	**1**	**2**	**3**	**4**	**5**
About	**0.44**	–0.11	0.32	0.06	0.40
Accessibility	0.09	–0.10	**0.75**	–0.22	–0.04
Advertising Policy	0.36	0.36	–0.22	–0.03	0.39
Author Name	–0.07	**0.91**	–0.06	0.03	–0.06
Author Affiliation	0.09	**0.54**	–0.04	0.32	0.23
Author Credentials	0.13	**0.88**	–0.05	0.00	–0.07
Contact Us	0.23	–0.10	0.16	**–0.47**	0.35
Copyright	**0.57**	0.23	0.01	0.01	0.07
Date of Creation	–0.09	0.09	–0.09	0.27	0.16
Date of Last Update	0.18	0.07	0.41	**0.56**	0.07
Disclaimer	**0.51**	0.05	–0.19	0.35	–0.42
Editorial Review Process	–0.03	0.39	**0.42**	0.23	0.38
Formal IQ Criteria	0.27	–0.15	**0.73**	0.20	–0.20
Payment	0.20	–0.04	–0.24	–0.13	–0.04
Privacy Policy	**0.74**	0.08	0.17	–0.08	0.12
Provider Name	**0.55**	–0.23	0.05	0.08	0.15
Quality Guidelines	0.13	0.02	**0.79**	–0.21	–0.04
Reference(s)	0.30	0.01	0.02	**0.79**	–0.02
Search Box	**0.77**	0.07	0.17	0.06	0.18
Site Map	0.56	0.07	0.16	–0.59	0.14
Sponsored Content	0.09	–0.16	–0.05	–0.01	**0.47**
Statement of Purpose or Mission Statement	–0.25	–0.24	–0.09	–0.01	0.12
Terms of Use, Policies, and Regulations	0.20	0.18	–0.09	0.06	**0.78**
Third-Party Quality Seal	**0.45**	0.35	–0.07	–0.03	0.39

Note. Extraction method: principal component analysis; rotation method: varimax with Kaiser normalization.

The first factor includes seemingly less related, but commonly occurring, markers (see Table 6). Hence, the first set of markers could be labeled as the baseline set. The markers loaded on the second factor indicated the authorship of the Web page content. The third factor included the indicators of a formal IQ assurance process. The fourth factor loadings were associated with content provenance or verifiability. Finally, the fifth factor was associated with content ownership, which could be different from authorship (see Table 7).

Table 7
IQ Marker Constructs

IQ marker constructs	IQ markers
Baseline	Provider Name, About, Third-Party Quality Seal, Copyright, Disclaimer, Privacy Policy, Search Box
Authorship	Author Name, Author Credentials, Author Affiliation
IQ Assurance Process	Editorial Review Process, Quality Guidelines, Accessibility, Formal IQ Criteria
Verifiability	Reference(s), Date of Last Update, Contact Us
Content Ownership	Terms of Use, Sponsored Content

Discussion

The first research question sought to identify the concepts or criteria people consider important when they evaluate the quality of healthcare information on the Web. The study identified three different approaches that providers might use to define their IQ evaluation sets: *community constructed*, *centrally mandated*, or *outsourced to third-party raters.* The mapping of these sets onto each other showed that the sets differed in criteria as well. The set used by an open system did not include the Integrity dimension, whereas the centrally mandated set used by the DHHS and its institutes did not include the Stability criterion. This difference could be attributed to the different organizational models these providers might have for information collections. The English Wikipedia and other open systems view an article as a continuous work in progress, and their relative stability is more an indicator of the community's consensus about its content than a result of security assurance actions. This approach has been considered both a strength and a weakness of the English Wikipedia model (Stvilia et al., 2008). Also, in certain instances (e.g., ongoing "edit wars"), the English Wikipedia does protect articles from unauthorized edits. Government-affiliated health information providers, on the other hand, are closed systems. They do not publish information unless it is mature, and any further modification to it, other than routine maintenance, is unusual and unwelcomed because of the potential impact on the public. Hence, it was expected that the security and integrity of government information had to be protected and that changes could be made only by authorized entities. Indeed, one of the main goals of the Data Quality Act and the Sarbanes-Oxley legislation (U.S. Congress, 2002) was to ensure the validity and integrity of information provided by the government and publicly traded companies throughout the life cycle of that information.

It is interesting to note that the HON Code contains the criterion Complementarity, which says that information should only support, not replace, the guidance of a doctor. The Code also includes Privacy, which was not presented in the other two sets. This could be attributed to the HON Code assessing the provider's overall trustworthiness and compliance with established norms and practices of health information dissemination, whereas the Wikipedia and DHHS sets were focused on IQ evaluation.

In this study, factor analysis was applied to the consumer rankings of quality criteria to identify the underlying model (see Table 4). Interestingly the Accuracy construct of the model combined both the Accuracy criterion itself and the trust-related criteria of Credibility and Reliability. This result is in line with the literature, which indicates that consumers may use source reliability and credibility to assess information accuracy indirectly or to assess quality in general (Bailey, Gurak, & Konstan, 2001). That consumers ranked the accuracy and trust-related quality dimensions high was not surprising and is in an agreement with previous research (e.g., Eysenbach et al., 2002). In addition, it is important to note that the Trustworthiness, Currency, Integrity, and Usefulness criteria fell short of the cutoff value for the factor loadings and were not included in the constructed criteria. Although most of these criteria had conceptually synonymous criteria present in the derived constructs, the Currency criterion did not. That the Currency criterion was missing in the constructed or derived set was unexpected, considering the importance given in the literature to health information being up-to-date (e.g., Eysenbach et al., 2002). This suggests a need to validate the results of this statistical model on another larger sample, which would allow lower cutoff values to be set for factor loadings.

The second research question focused on identifying Web page components or indicators that could be used to signal IQ or for making IQ judgments. The exploratory factor analysis of the structural IQ markers found in the Yahoo! sample indicated the presence of strong underlying relations connecting these markers (see Table 6). These statistical relations can give us a better understanding of the structure of consumer health Web pages and the ways providers may signal quality and trust through structural markers.

Conceptually, the quality markers can be mapped onto and can support metrics for more than one quality criterion (see Figure 2). However, the set of quality markers identified and analyzed in this study reflected the findings of the literature analysis and the practices of the provider. This study did not examine whether and how consumers associated the markers with specific quality criteria. The survey used in the study did not ask consumers to rank markers by their importance in making quality judgments. The researchers thought that a survey might not be an appropriate method for eliciting consumer preferences for the markers. These markers are not standardized and can take different forms and be referred to by different terms. The respondents might have known the marker by a different name, or might not have been familiar with the marker at all, and their response could have been skewed by a definition of the marker the researchers had to supply. Future research could include the study of consumers' understanding and use of quality markers by applying less obtrusive or guiding methods, such as observing consumers "thinking aloud" (Ericsson & Simon, 1993) when making IQ judgments in a laboratory environment.

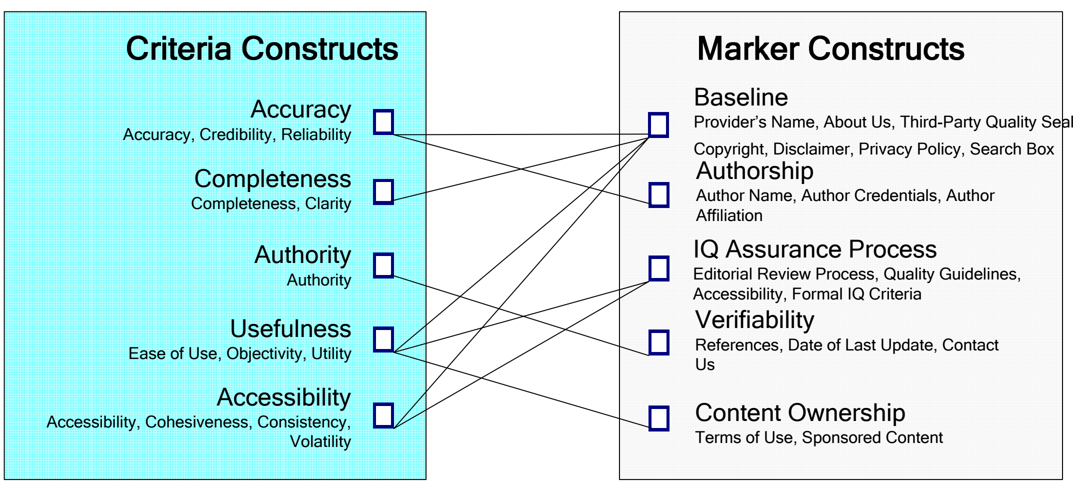

Figure 2. Information quality marker–criteria mapping.

The third question was aimed at identifying healthcare Web page genres and information provider types. The analysis identified seven Web page genres: *Article*, *Blog*, *Directory*, *Factsheet*, *Instrument*, *Mainpage*, and *Q&A*. The analysis also found five types of consumer health information providers: *commercial*, *nonprofit or community*, *government, patient and family*, and *research*. In an earlier study, Khechine, Pascot, and Prémont (2008, p. 20) divided health Web site providers into four main types: "scientific" or "professional" Web sites (medical databases, government health sites); general and nonscientific sites not specialized in health (e.g., electronic reviews); commercial sites with information that also sells health products or services (such as pharmaceutical Web sites); and newsgroups, online forums, and mailing lists, such as discussion groups. Comparison of the two typologies showed that the provider typology defined in this study could fully accommodate the provider types identified by Khechine et al. (2008). The differences in grouping between the two could be attributed to different attributes used for defining the types. Although this study used solely organizational types of providers in defining categories, the other study based the groupings on a confluence of provider types and service types.

The fourth research question was concerned with the degree of relationship between the quality markers and the types or genres of Web pages and the types of providers. The study found that different Web page and provider types did use different sets of quality markers and that the relationships between the types and markers were statistically significant. That is, consumers may not have the same set of quality markers available when assessing the quality of different kinds of health Web pages. The findings also imply the possibility of creating Web page type- or genre-specific templates of quality indicators, which could be used in automatic IQ assessment. As culturally justified and socially typified communication forms, information genres could serve as valuable heuristics for specifying common or shared features, including quality markers. For different classes of information objects, genres and their characteristics (i.e. typical context of use, functionalities, form, components, attributes) could be used to specify a baseline quality model for a particular type of information. This might include applicable quality metrics (e.g., readability level, currency), sources of the metrics (e.g., language, html markup), critical values, and relationships. Indeed, the social world is based on typified activities. Typified activities use typified communication actions, roles, and tools. Information objects can serve as tools in typified actions, and they may require typified IQ. Information typification can be explicit, as with organizational rules or legal acts, or it can be implicit, as with cultural and social norms of communication (Bakhtin, 1986; Orlikowski & Yates, 1994). As a result, identifying the genres of healthcare information, and enumerating their properties, could reduce uncertainty about IQ evaluation and could make IQ evaluation less expensive and more systematic. Furthermore, creating genre-specific templates for healthcare Web pages could be used for educating consumers, intermediaries, and providers about quality or trust indicators, what quality indicators and functionalities each Web page type is expected to have, and how to use those indicators in evaluating or alternatively signaling the quality of the Web page.

In addition, the study found that the set of quality markers identified in the literature might not be sufficient for classifying pages automatically. The accuracy of the classification of the sampled Web pages using the quality markers was only slightly higher than random selection. Further research is needed to define a more complete set of document surface markers and linguistic cues, which may not necessarily be IQ markers and which would enable more accurate automatic classification of consumer health Web pages by genre.

Finally, only 25% of the survey respondents indicated that they checked whether the Web site had a formal IQ policy, whereas only 35% of the sampled Web sites had such a policy. These findings are in line with the existing literature (Eysenbach & Köhler, 2002) and point to a potential IQ literacy problem. Seventy-four percent of the survey respondents indicated that they found online health information resources by using a search engine. The second most frequent source of learning about online health information resources was via family and friends, at 12%. Only 4% of the respondents said they learned about health information resources from a physician, and only 9% used health information portals and directories to find information (see Table 2). Thus, in spite of efforts to establish electronic portals of quality consumer health information (Breckons, Jones, Morris, & Richardson, 2008), general search engines remain by far the most dominant gateways on the path of finding and selecting online health information. By inference, this also suggests that the need persists for consumer health IQ literacy, IQ evaluation knowledge bases (ontologies), and services, and for further research efforts.

Conclusion

This study developed a quality model of consumer health Web pages consisting of constructs of IQ criteria derived by exploratory factor analysis of empirical data. The research also made initial explorations into connecting structural quality markers with Web page and provider types. Developing document type-specific templates of IQ evaluation, which may include genre-specific quality markers and relationships among those markers, metrics, and criteria, could be a significant step toward developing effective, reusable, and automatic IQ evaluation tools.

Analysis of the consumer surveys and interviews revealed that consumers might evaluate health information pages based on virtues (or quality indicators) that have not been included in the IQ literature. One of the respondents identified empathy as a desired virtue in health information. Further research is needed to define potential sources of and metrics for consumer-identified quality indicators, such as the assessment of empathy in healthcare information. Algorithms and methods used in text classification and affect analysis (e.g., Mishne, 2005) might be helpful in developing an automatic classifier and metrics for empathy identification and measurement.

The study has certain limitations. The researchers used a convenience sample of consumers selected from a local ethnic community. It is unknown how much the cultural values of the community may have influenced the respondents' answers and the derived model of quality. Future research may involve replicating this research with a larger and more culturally diverse sample of consumers. Furthermore, the derived model of IQ is based on statistical analysis of a single sample. Future research may involve testing the stability of the model by using confirmatory factor analysis on a different sample.

The study examined general relationships among quality criteria, structural markers, and Web page types or genres. Future work would include defining genre-specific templates and metrics of IQ measurement for consumer health information. This future research may also involve developing models for the automatic classification of health Web pages by genre.

References

Bailey, B., Gurak, L., & Konstan, J. (2001). An examination of trust production in computer-mediated exchange. In *Proceedings of the 7th Conference on Human Factors and the Web.* Retrieved June 4, 2008, from http://agora.cs.illinois.edu/download/attachments/9644168/hfweb-2001.pdf

Bakhtin, M. (1986). The problem of speech genres. In V. McGee (Ed.), *Speech genres and other late essays* (pp. 60–101). Austin: University of Texas Press.

Beghtol, C. (2001). The concept of genre and its characteristics. *Bulletin of the ASIS, 27*(2), 17-19.

Berland, G., Elliott, M. N., Morales, L. S., Algazy, J. I., Kravitz, R. L., Broder, M. S., Kanouse, D. E., Muñoz, J. A., Puyol, J. A., Lara, M., Watkins, K. E., Yang, H., & McGlynn, E. A. (2001). Health information on the Internet: Accessibility, quality, and readability in English and Spanish. *JAMA, 285*(20), 2612–2621.

Bernstam, E. V., Walji, M. F., Sagaram, S., Sagaram, D., Johnson, C. W., & Meric-Bernstam, F. (2008). Commonly cited website quality criteria are not effective at identifying inaccurate online information about breast cancer. *Cancer, 112*(6), 1206–1213.

Bliemel, M., & Hassanein, K. (2007) Consumer satisfaction with online health information retrieval: A model and empirical study. *E-Service Journal, 5*(2), 53–84.

Breckons, M., Jones, R., Morris, J., & Richardson, J. (2008). What do evaluation instruments tell us about the quality of complementary medicine information on the internet. *Journal of Medical Internet Research, 10*(1), e3.

Carletta, J. (1996). Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics, 22*(2), 249–254.

Chang, M., & Cheung, W. (2005). Online trust production: Interactions among trust building mechanisms. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences - HICSS 2005.* (p. 181).

Charnock, D., Shepperd, S., Needham, G., & Gann, R. (1999). DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology and Community Health, 53*, 105–111.

Chen, G., & Choi, B. (2008). Web page genre classification. In *Proceedings of the 2008 ACM Symposium on Applied Computing* (pp. 2353–2357). New York: ACM.

Ericsson, K., & Simon, H. (1993). *Protocol Analysis: Verbal Reports as Data* (2nd ed. ed.). Boston: MIT Press.

Eysenbach, G., & Köhler, C. (2002). How do consumers search for and appraise health information on the World Wide Web? Qualitative study using focus groups, usability tests, and in-depth interviews. *British Medical Journal*, *324*(7337), 573–577.

Eysenbach, G., Powell, J., Kuss, O., & Sa, E. (2002). Empirical studies assessing the quality of health information for consumers on the World Wide Web. *JAMA, 287*(20), 2691–2700.

Fallis, D., & Frické, M. (2002). Indicators of accuracy of consumer health information on the Internet: A study of indicators relating to information for managing fever in children in the home. *Journal of the American Medical Informatics Association, 9*(1), 73–79.

Family Online Safety Institute. (2009). *ICRA tools: Technology, tools and best practice*. Retrieved June 4, 2008, from http://www.fosi.org/icra/

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*(4), 327–359.

Fox S, Rainie L. (2002). *Vital decisions: How Internet users decide what information to trust when they or their loved ones are sick.* Pew Internet and American Life Project. Washington, DC: Pew Research Center. Retrieved from http://www.pewinternet.org/pdfs/PIP_Vital_Decisions_May2002.pdf

Fox, S. (2006). *Online health search 2006.* Washington, DC: Pew Internet & American Life Project. Washington, DC: Pew Research Center. Retrieved June 4, 2008, from http://www.pewinternet.org/PPF/r/190/report_display.asp

Fox, S. (2008). *The engaged e-patient population.* Pew Internet & American Life Project. Washington, DC: Pew Research Center. Retrieved June 4, 2008, from http://www.pewinternet.org/PPF/r/259/report_display.asp

Frické, M., & Fallis, D. (2004). Indicators of accuracy for answers to ready reference questions on the internet. *Journal of the American Society for Information Science and Technology, 55*(3), 238–245.

Frické, M., Fallis, D., Jones, M., & Luszko, G. (2005). Consumer health information on the internet about carpal tunnel syndrome: indicators of accuracy. *The American Journal of Medicine, 118*(2), 168-174.

Gagliardi, A., & Jadad, A. R. (2002) Examination of instruments used to rate quality of health information on the Internet: Chronicle of a voyage with an unclear destination, *British Medical Journal, 324*(7337), 569–573.

Google Co-op. (2009). *Google Co-op page for the Health topic*. Retrieved December 9, 2008, from http://www.google.com/coop/topics/Health

Griffiths, K. M., Tang, T. T., Hawking, D., & Christensen, H. (2005). Automated assessment of the quality of depression websites. *Journal of Medical Internet Research, 7*(5), e59.

Gustafson, D., & Wyatt, J. (2004). Evaluation of ehealth systems and services: We need to move beyond hits and testimonials. *British Medical Journal, 328*(744), 1150.

Haas, S., & Grams, E. (1998). Page and link classifications: Connecting diverse resources. In I. Witten, R. Akscyn, & F. Shipman (Eds.), *Proceedings of the Third ACM Conference on Digital Libraries* (pp. 99–107). New York: ACM.

Hair, J., Black, B., Babin, B., Anderson, R., & Tatham, R. (2005). *Multivariate data analysis.* Upper Saddle River, NJ: Prentice-Hall.

Hardey, M. (2001). ‘E-health’: The Internet and the transformation of patients into consumers and producers of health knowledge. *Information Communication and Society, 4*(3), 388–405.

Harris, R., Wathen, C., & Fear, J. (2006). Searching for health information in rural Canada. Where do residents look for health information and what do they do when they find it? *Information Research, 12*(1). Retrieved December 9, 2008, from http://informationr.net/ir/12-1/paper274.html

Health on the Net Foundation (HON). (2009). *Health On the Net Foundation home page*. Retrieved June 8, 2008, from http://www.hon.ch/

Hesse, B., Nelson, D., Kreps, G., Croyle, R., Arora, N., Rimer, B., & Viswanath, K. (2005). Trust and sources of health information. *Archives of Internal Medicine, 165*(22), 2618–2624.

Institute of Medicine. (1999). *To err is human: Building a safety health system.* Washington, DC: National Academy Press.

Khechine, H., Pascot, D., & Prémont, P. (2008). Use of health-related information from the Internet by English-speaking patients. *Health Informatics Journal, 14*(1), 17–28.

Lewis, M., Jr. (2008). Health 2.0: The numbers. *Medical Economics, 85*(23), 16.

Marschak, J. (1971). Economics of information systems. *Journal of the American Statistical Association, 66*(333), 192–219.

MedlinePlus. (2007). *MedlinePlus quality guidelines*. Retrieved June 8, 2008, from http://www.nlm.nih.gov/medlineplus/criteria.html

Meredith, P., Emberton, M., Wood, C., & Smith, J. (1995). Comparison of patients’ needs for information on prostate surgery with printed materials provided by surgeons. *Quality in Health Care, 4*(1), 18–23.

Mishne, G. (2005). Experiments with mood classification. In *Proceedings of Stylistic Analysis of Text for Information Access Workshop (SIGIR, 2005).* Retrieved June 8, 2008, from http://staff.science.uva.nl/~gilad/pubs/style2005-blogmoods.pdf

Mon, L., Stvilia, B., Yi, Y. (in process). *Health and medicine with the digital library: patterns in questioning and answering from online consumers.*

Office of Management and Budget (OMB). (2002). Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by federal agencies. *Federal Register*, *67*(36), 8451–8460.

Orlikowski, W., & Yates, J. (1994). Genre repertoire: The structuring of communicative practices in organizations. *Administrative Science Quarterly, 39*, 541–574.

Phelps, C. (1992). Diffusion of information in medical care. *The Journal of Economic Perspectives, 6*(3), 23–42.

Potts, H. W. W., & Wyatt, J. C. (2002). Survey of doctors’ experience of patients using the Internet. *Journal of Medical Internet Research, 4*(1). Retrieved December 9, 2008, from http://www.jmir.org/2002/1/e5/

Resnick, P., Miller, J. (1996). PICS: Internet access controls without censorship. *Communications of the ACM, 39*(10), 87-93.

Sanders, B., Berlin, R., & Schatz, B. (2008). Adaptive question asking for population health monitoring. In *Proceedings of the AMIA Annual Meeting.* Retrieved Nov. 9, 2008, from http://www.canis.uiuc.edu/~schatz/healthmonitor-AMIA08.doc

Sillence, E., Briggs, P., Harris, P., & Fishwick, L. (2007). Going online for health advice: Changes in usage and trust practices over the last five years. *Interacting with Computers*, *19*(3), 397–406.

Strong, D., Lee, Y., & Wang, R. (1997). Data quality in context. *Communications of the ACM, 40*(5), 103–110.

Stvilia, B. (2006). *Measuring information quality.* Unpublished doctoral thesis, University of Illinois at Urbana-Champaign.

Stvilia, B., & Gasser, L. (2008). Value based metadata quality assessment. *Library & Information Science Research, 30*(1), 67–74.

Stvilia, B., Gasser, L., Twidale M., B., & Smith L. C. (2007). A framework for information quality assessment. *Journal of the American Society for Information Science and Technology, 58*(12), 1720–1733.

Stvilia, B., Gasser, L., Twidale, M. B., Shreeves, S., & Cole, T. (2004). Metadata quality for Federated Collections. In S. Chengulur-Smith, L. Raschid, J. Long, & C. Seko (Eds.), *Proceedings of the International Conference on Information Quality–ICIQ 2004* (pp. 111–125). Cambridge, MA: MITIQ.

Stvilia, B., Twidale, M., Smith, L. C., & Gasser, L. (2008). Information quality work organization in Wikipedia. *Journal of the American Society for Information Science and Technology, 59*(6), 983–1001.

TRUSTe. (1997-2009). Retrieved June 8, 2008, from http://www.truste.org/

U.S. Congress. (2002). An Act to Protect Investors by Improving the Accuracy and Reliability of Corporate Disclosures Made Pursuant to the Securities Laws, and for Other Purposes. The Sarbanes-Oxley Act, 107th Cong., H.R. 3763.

U.S. Department of Health and Human Services (DHHS). (2006). *HHS information quality Web site: HHS guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated to the public.* Retrieved June 8, 2008, from http://aspe.DHHS.gov/infoquality/Guidelines/index.shtml

U.S. Department of Health and Human Services. (2007). *Report on Objective 11-4: Estimating the proportion of health related websites disclosing information that can be used to assess their quality*. Retrieved June 8, 2008, from http://www.health.gov/communication/healthypeople/obj1104/default.htm

Wang, R., & Strong, D. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5–35.

Wang, Y., & Liu, Z. (2007). Automatic detecting indicators for quality of health information on the Web. *International Journal of Medical Informatics, 76,* 575–582.

Yahoo! (2008). *Yahoo! Directory—Disease and conditions*. Retrieved June 8, 2008, from http://dir.yahoo.com/Health/Diseases_and_Conditions/

Yoshioka, T., Herman, G., Yates, J., & Orlikowski, W. (2001). Genre taxonomy: A knowledge repository of communicative actions. *ACM Transactions on Information Systems (TOIS), 19*(4), 431–456.

Zucker, L. (1986). Production of trust: Institutional sources of economic structure, 1840–1920. *Research in Organizational Behavior, 8,* 53–111.