2004

Health information seals of approval: What do they signify?

Jacquelyn Burkell
The University of Western Ontario, jburkell@uwo.ca

Follow this and additional works at: https://ir.lib.uwo.ca/fimspub
Part of the Library and Information Science Commons

Citation of this paper:
Health Information Seals of Approval: What Do They Signify?

Jacquelyn Burkell, PhD
Assistant Professor
Faculty of Information and Media Studies
North Campus Building, Room 240
The University of Western Ontario
London, Ontario, Canada N6A 5B7

Email: jburkell@uwo.ca
Tel: 519-661-2111 ext 88506
(address for correspondence)

Running Head: Burkell
Abstract

Much of the health information available to consumers on the Internet is incomplete, out of date, and even inaccurate. Seals of approval or trustmarks have been suggested as a strategy to assist consumers to identify high quality information. Little is known, however, about how consumers interpret such seals. This study addresses this issue by examining assumptions about the quality criteria that are reflected by a seal of approval. This question is of particular importance because a wide variety of quality criteria have been suggested for online health information, including core aspects of quality such as accuracy, currency, and completeness, proxy indicators of quality such as the disclosure of commercial interests, and indicators that reflect the quality of the site or the interaction it affords, such as the availability of a search mechanism. The results of this study suggest that seals of approval are assumed to certify information quality primarily with respect to core quality indicators, aspects that subjects both consider to be important and feel relatively less able to evaluate for themselves (compared to their ability to rate proxy indicators of information or indicators of site or interaction quality). This assumption is largely inconsistent with practice: most seals of approval involve assessment of proxy indicators of information quality, rather than direct assessment of content. These results identify a problem that certification or accreditation bodies must address, since unless and until consumer expectations are congruent with evaluation practice, seals of approval seem to promise more than they deliver.
Introduction

A vast array of health-related information for consumers on virtually any topic is available on the Internet, and surveys of Internet users in the U.S. indicate that over half have gone online to seek health information (Rainie & Packel, 2001; PEW Research Center, 2002). One significant concern of both providers and seekers of health information on the Internet is the relatively poor quality of that information, or at least the growing belief that where quality information does exist it will be impossible to find amidst lower quality, unfiltered information (Eysenbach, 2000a; Shon & Musen, 1999). With regard to health information, this concern is well founded: it is widely documented that health information on the Internet is questionable in terms of coverage, accuracy, and currency (Bichakjian, et al. 2002; Eysenbach et al., 2002; Griffiths & Christensen, 2000; Impicciatore et al., 1997; Latthe et al., 2000; Lissman & Boehnlein, 2001; Pandolfini et al., 2000; Sandvik, 1999). Furthermore, although there is some indication that the quality of health information on the web has improved since the problem was first identified, quality control for online health information remains an issue of concern (Karp and Monroe, 2002; Pandolfini and Bonati, 2002).

Although some suggest there is “nothing radically different about information on the web” (Shepperd & Charnock, 2002: p556; Risk & Dzenowagis, 2001), most researchers agree that Internet information presents significant issues of quality control (Eysenbach & Diegpen, 1998), particularly since research indicates that consumers use Internet health information to make important health care decisions. Internet health information leads consumers to ask questions of health care professionals and to seek second opinions (PEW Research Center, 2000; Potts & Wyatt, 2002). Consumers use Internet health information for self-diagnosis and self-treatment (PEW Research Center, 2000; PEW Research Center, 2002); in fact, the use of this information often replaces or leads to the deferral of a visit to a health care professional (Nicholas et al., 2001; PEW Research Center, 2000; PEW Research Center, 2002). As a result, health care providers are concerned that low quality Internet health information could have negative consequences for consumer health and well-being. Although there are few documented cases of harm resulting from Internet health information (Smith, 2001b), such instances do exist (Kiley, 2002; Crocco et al., 2002), and a Database of Adverse Events Related to the Internet (DAERI) has recently been initiated to
document cases of possible harm caused by health information obtained on the Internet (Eysenbach & Köhler, 2002a).

Despite the general concern about the quality of health information found on the Internet, consumers themselves do surprisingly little to ensure the information they find is of high quality. Internet users only ‘rarely’ or ‘occasionally’ verify information obtained online, although reference information (and health information would clearly fall within this category) is more likely to be verified than commercial or entertainment information (Flanagin & Metzger, 2000). Although consumers indicate that they examine source as a primary criterion when assessing credibility, in practice they do not verify sources or read disclaimers on the web sites they use: in fact, few even recall the sites from which they gather information (Eysenbach & Köhler, 2002b). One possibility is that consumers access information from trusted sites, and therefore do not need to verify sources or information in every case. The data, however, suggest otherwise: most users find health information through an Internet search rather than being directed by a trusted source, and over half access information from unfamiliar web sites (Eysenbach & Köhler, 2002b; Peterson et al., 2003; PEW Research Center, 2000). In practice, consumers explore the first few links provided by general search engines in response to suboptimal search queries (e.g., use of one search term when multiple terms would have been more effective: Eysenbach & Köhler, 2002b). Consumers could use triangulation to identify high quality information, looking for information that is repeated on multiple sites. It appears, however, that this is another potential verification strategy that is not widely used, since fewer than one third consult four or more web sites for a given issue (Eysenbach & Köhler, 2002b; PEW Research Center, 2000). Surprisingly, those consumers who do report verifying Internet information indicate that they are most likely to check if the information is current and complete (though it is unclear how they would make these judgments if they are unfamiliar with the content area), and less likely to check the author’s qualifications or credentials, or check for external recommendations for the site, despite the fact that these latter factors would be easier for consumers to assess (Flanagin & Metzger, 2000).

Obviously, consumers require assistance to identify high quality health information. Although it is important that consumers take a critical approach to all health information, independent evaluation can be of some assistance to help identify high quality information. One strategy that has been suggested is seals of approval, trustmarks, or certificates that can be displayed on health information sites indicating that the
sites has met the standards of the body that grants the certification (Boyer et al., 1998; Eysenbach et al., 2001; Wathen & Burkell, 2002). These quality indicators (which will be collectively termed ‘seals of approval’ throughout the remainder of this paper) have the advantage that they appear directly on websites to which they are awarded, and thus their use as quality indicators requires no additional effort (as does the application of checklists of quality indicators such as the DISCERN checklist, available at http://www.discern.org.uk/) or specialized access (as do health information portals such as MedlinePlus, available at http://medlineplus.gov/) on the part of consumers.

Seals of approval have been used in the online context to identify sites that meet general quality criteria (e.g., the Good Housekeeping website certification: http://magazines.ivillage.com/goodhousekeeping/consumer/institute/articles/0,,284511_290570-3.00.html), and seals have also been used online for specific purposes such as tempering consumer privacy concerns (Miyazaki & Krishnamurthy, 2002; Palmer et al., 2000). A number of seals of approval or trustmarks have been developed specifically for Internet health information. Some reflect self-reported voluntary compliance with a code of conduct (e.g., HONcode, administered by the Health on the Net Foundation, www.hon.ch), others involve third-party verification for which participants pay a fee (e.g., URAC Health Web Site Accreditation, administered by URAC, http://www.urac.org/), and still others involve a collaborative combination of self-report, consumer report, and expert evaluation and do not require a fee (e.g., the MEDcertain project, described at http://www.medcertain.org/). In every case, the goal is to identify Internet health information that meets quality criteria by means of a seal or label that appears on the website itself.

As consumers, we are not unfamiliar with this type of quality indicator: the ‘Good Housekeeping Seal of Approval’, for example, has been used to identify household items that meet the Good Housekeeping standards since 1910. Even in the relatively familiar domain of product advertising, however, the meaning of seals of approval is unclear and their interpretation by consumers is often inconsistent with the actual certification that is being offered. Parkinson (1975) found that advertising seals of approval were felt by consumers to represent “expertise”, “impartiality” and “trustworthiness”, but his results also indicated that consumers tend to attribute more meaning to such seals than is warranted. More recent research indicates that consumers misunderstand the degree to which even familiar seals of approval
support factual, evaluative, and warranty claims (Beltramini & Stafford, 1993; Laric & Sarel, 1981). Nonetheless, in many studies seals of approval are shown to influence the perception of the product or service that is approved (Miyazaki & Krishnamurthy, 2002; Laric & Sarel, 1981; Cheskin Research, 2000; Fogg et al., 2000: but see Beltramini & Stafford, 1993 and Shon & Musen, 2002 for results to the contrary). Thus, consumers appear to interpret seals of approval without fully understanding (or checking to find out) what the seal signifies or on what basis it is awarded.

One particular consequence is that consumers are likely to make assumptions about the criteria that determine the award of seals of approval. When the possible criteria are few and unequivocal, this presents no difficulty; but in the case of Internet health information a wide variety of indicators have been suggested as quality markers (Fallis & Frické, 2002; Jadad & Gagliardi, 1998; Kim et al., 1998; Winker et al., 2000). The ‘JAMA benchmarks’ (Silberg et al., 1997) are representative, and include display of authorship for medical content, source identification (attribution or references), date of last update, and disclosure of ownership, advertising policies, or conflicts of interest. The JAMA benchmarks demonstrate another important aspect of quality indicators for Internet information: the majority of these criteria represent proxy indicators of information quality (such as disclosure of authorship) rather than direct assessment of content. Although there are few formal evaluations of proxy indicators, those studies that have been completed suggest that these indicators bear at best a tentative relationship to information quality (Abbott, 2000; Fallis and Frické, 2002; Griffiths and Christensen, 2002; Martin-Facklam et al., 2002; Meric et al., 2002).

Thus, if consumers are going to guess about the criteria that are signified by a seal of approval (and previous research suggests they do guess), they have lots of options. Moreover, since the different criteria are not interchangeable (in that they reflect different aspects of quality), the choice is important. While proxy indicators such as author identification may well be important in their own right (Eysenbach, 2000b; Eysenbach, 2002), information that satisfies these proxy indicators is not necessarily accurate, complete, current, or unbiased. If consumers assume that seals reflect these core aspects of quality but evaluation instead assesses proxy indicators, there is potential for misinterpretation, and indeed over-interpretation, of the significance of seals of approval for Internet health information.
As Delamothe (2000) points out, we know little about how consumers understand and use seals of approval in this context. In particular, we do not know what seals of approval are taken to signify: when a seal is displayed, what does it mean to a consumer? The research reviewed above suggests that, in general, seal interpretation is driven not by accurate knowledge of the certification process but rather by assumptions regarding the evaluation that underlies the award. This study explores these assumptions with respect to seals of approval for Internet health information. In particular, the study addresses the following questions: which quality criteria do consumers consider to be important, how able do they feel to evaluate the criteria for themselves, and what are their assumptions about the criteria that would be evaluated in order to assign a seal of approval to an Internet health information site?

Method

Identification of Quality Criteria

An initial list of 24 quality indicators for Internet information was developed based on a review by Falshaw (2002), who assembled indicators reported in the published literature in library and information science and medicine as well as online Internet information evaluation guidelines (Alexander & Tate, 1996; Barry & Schamber, 1998; Beck, 2002; Block, 2001; Collins, 1996; Connell & Tipple, 1999; Everhart, 1996; Fallis & Frické, 2002; Grassian, 2000; Hawkins, 1999; Jadad & Gagliardi, 1998; Joseph, 1999; Kapoun, 1998; Lopez, 1998; Minkel, 2000; Schrock, 1998; Schrock, 1999; Smith, 1997; Smith, 2001); Tillman, 2001; Waller 2001). In order to limit the criteria to a reasonable number, similar criteria were grouped together under a single identifier. Thus, for example, all criteria relating to the disclosure of commercial interests were grouped under “Website is free of commercial interests OR commercial interests are declared”. In addition, the list was limited to criteria relating to content, and criteria related to other aspects of service delivery (e.g., privacy protection) were eliminated from the list. The resulting criteria tended to fall into two groups: proxy indicators of information quality (such as lack of spelling, grammar, and typographical errors), and indicators that reflect the quality of the site or the quality of interaction (such as the availability of a search function). None of the criteria gathered from Falshaw (2002) involved direct
assessment of the quality of information, and (as reported above) proxy indicators are not strongly predictive of information quality. As a result, the 24 criteria identified by Falshaw (2002) were augmented by four core aspects of information quality: accuracy, currency, completeness, and lack of bias. The final list of 28 indicators, separated into the three types (core quality criteria, proxy quality indicators, and site or interaction quality criteria), is presented below (see Table 1).

---

Insert Table 1 about here

---

Procedure

Participants were a convenience sample of first year undergraduate students enrolled in an introductory media studies course. The researcher distributed the questionnaires at the beginning of a class session, and students returned the completed questionnaires. A total of 64 questionnaires were returned, of which 3 were unusable because subjects failed to complete all ratings for one or more questions (e.g., all ‘ability to rank’ ratings were left blank), leaving 61 usable responses.

The data were collected by means of a paper and pencil questionnaire. The questionnaire included of three types of ratings for each of the 28 quality criteria described above. Participants were asked rate each quality criterion in response to the following instructions: 1) identify how important you think each criterion is for choosing health information on the web; 2) indicate how confident you are that you could evaluate each criterion for choosing health information on the web; 3) identify how important you think each criterion would be in determining whether a site merits a ‘seal of approval’. Before rating the contribution of the criteria to a ‘seal of approval’, subjects were first shown examples of (hypothetical) seals that might be used to identify health information that met quality criteria (see Figure 2). All ratings were on a scale of 10, with 1 representing the lowest level of the rating (not at all important, certain you could not evaluate yourself, or very unlikely to contribute to a ‘seal of approval’), and 10 representing the
highest level of the rating (very important, certain you could evaluate yourself, or very likely to contribute to a ‘seal of approval’).

Subjects completed each type of rating with respect to every criterion before moving on to the next type of rating. Each subject received the rating tasks in the same order: 1) importance; 2) ability to evaluate; 3) expected contribution to ‘seal of approval’. One random order of the criteria was identified and the criteria were presented in this same order for all rating tasks. The entire package of ratings took approximately 20 minutes to complete.

Results

Table 2 shows, for each quality indicator, the average score (across subjects) for each rating tasks (importance (abbreviated as IMP), ability to rate (abbreviated as ABIL), and expected contribution to a seal of approval certification (abbreviated as SEAL)). For each rating task, the items that ranked in the top five with respect to average rating are identified by table entries in bold. The average scores for IMP ranged from a low of 5.4 (for ‘Web site is aesthetically pleasing’) to a high of 9.6 (for ‘Information is accurate, reliable, and error-free’). ABIL scores ranged from a low of 4.3 (for ‘Information has been reviewed for accuracy’) to a high of 8.2 (for ‘Web site has no spelling, grammatical, or typographical errors’, ‘Information is free of cost’, and ‘Web site is easy to navigate’). The scores for SEAL ranged from a low of 5.5 (for ‘Web site is aesthetically pleasing’) to a high of 9.5 (for ‘Information is accurate, reliable, and error-free’).
A quick perusal of the table reveals a strong similarity between IMP and SEAL ratings, and an almost reverse ordering for ABIL. The same five items receive the highest average ratings for both IMP and SEAL, while four of these items were among the five lowest ratings on the ABIL task. The pattern of correlations between the types of ratings supports these initial impressions. Pairwise correlations revealed a strong positive relationship between IMP and SEAL ($r(28)=.834$, $p<.001$), a strong negative relationship between ABIL and SEAL ($r(28)=-.794$, $p<.001$), and a moderate negative relationship between IMP and ABIL ($r(28)=-.597$, $p<.01$). According to these results, the same criteria tend to be identified as important and expected to contribute to awarding a seal of approval; furthermore, the criteria that receive high ratings for each of these qualities tend to be the core quality criteria of accuracy, completely, and currency (lack of bias, the fourth core quality criterion, also receives high ratings on both scales, but does not rank within the top five). In contrast, respondents are quite uncertain of their own ability to rate information according to these core quality criteria, and instead indicate more confidence in rating proxy quality indicators or criteria relating to site or interaction quality.

The results of an ANOVA comparing average scores for the three types of quality indicators (information quality (abbreviated INF), proxy indicators of information quality (abbreviated PROX), and indicators of site or interaction quality (abbreviated SITE)) across the three rating tasks (IMP, ABIL, SEAL) are largely consistent with this interpretation. The ANOVA results reveal a significant main effect of rating task ($F(1,59)=29.17$, $p<.001$), a significant main effect of type of quality indicator ($F(2,118)=16.99$, $p<.001$), and a significant interaction of rating task and indicator type ($F(1,59)=99.32$, $p<.001$). Further analysis of the interaction was carried out using simple main effects and subsequent tests of means with Bonferroni adjustment for multiple comparisons. For the IMP task, there were significant differences among the three types of indicators ($F(2,120)=97.78$, $p<.001$). For this task, INF indicators received higher ratings than either PROX or SITE indicators, which did not differ significantly from each other (means of 8.9 for INF indicators, 7.2 for PROX indicators, and 6.9 for SITE indicators, $t(61)=10.88$, ...
p<.001 for the comparison of INF and PROX, t(61)=11.21, p<.001 for the comparison of INF and SITE, and t(61)=2.08, n.s. for the comparison of PROX and SITE). For the ABIL task there were also significant differences across the three types of criteria (F(1,59)=49.22, p<.001). All three pairwise comparisons were significant, but the direction of the differences was reversed from that observed for the IMP task, with SITE criteria rated higher than PROX criteria, which were in turn rated higher than INF criteria (means of 5.1 for INF, 6.6 for PROX, and 7.5 for SITE, t(60)=5.47, p<.001 for the comparison of INF and PROX, t(61)=7.76, p<.001 for the comparison of INF and SITE, and t(60)=6.63, p<.001 for the comparison of PROX and SITE). For the SEAL task, there were significant differences among the three types of ratings (F(1,60)=78.01, p<.001), and all three pairwise comparisons were significant, with INF criteria rated higher than PROX criteria, which were in turn rated higher than SITE criteria (means of 8.9, 7.5, and 6.2 respectively, t(61)=7.72, p<001 for the comparison of INF and PROX, t(61)=10.93, p<.001 for the comparison of INF and SITE, and t(61)=6.05, p<.001 for the comparison of PROX and SITE). The results of this analysis are consistent with expectations. Indicators that directly reflect information quality (INF indicators) are rated as most important and most likely to contribute to a seal of approval, while subjects indicate that they feel least able to assess these same indicators.

Finally, a regression analysis was used to examine the degree to which SEAL ratings were determined by IMP and ABIL. The results for the full model (both variables entered) reveals that both IMP and ABIL contribute significantly to the prediction (t=5.47, p<.001 for IMP, t=-4.52, p<.001 for ABIL), and together these two predictors account for 82% of the variance in SEAL ratings (R=.912, F(2,25)=61.97, p<.001, adjusted R²=.819). Furthermore, after accounting for IMP, the partial correlation of ABIL with SEAL is negative (partial correlation=-.67, p<.001). In other words, subjects think that seals of approval reflect first those qualities they consider to be important, and second those criteria they feel less able to evaluate for themselves.

Conclusion

When a consumer sees a seal of approval on a health information website, what does it signify? The results of this study suggest that seals of approval are assumed to identify information that meets core quality criteria such as accuracy, currency, completeness, and lack of bias. In the eyes of study participants,
a seal of approval indicates that the information meets quality criteria they consider to be important, and those they feel less able to rate for themselves.

Seals of approval are intended as information interfaces – tools that assist consumers with the difficult task of sorting information ‘wheat’ from ‘chaff’. A good interface is transparent: that is, it does precisely what the user expects it to do (Mountford, 1990). Interfaces that are transparent allow us to do our work (in this case, finding high quality health information) easily and effectively. The data reported in this paper provide insight into expectations regarding the meaning of seals of approval leads. These results lead to the natural question: do seals of approval meet expectations? In other words, are seals of approval transparent interfaces?

Typically, the answer is ‘no’. HONcode, URAC Health Web Site Accreditation, and MedCertain are designed to certify web health web sites with respect to both the information they provide and the practices (such as privacy protection) associated with the provision of that information. With respect to information quality, these seals of approval tend to rely on proxy indicators of quality rather than direct assessment of content. Thus, the HONcode code of conduct indicates that web sites displaying the HONcode label should comply with standards that include

- Indication of authorship
- Indication of author credentials
- Disclaimer with respect to information use
- Date of last modification and creation date indicated
- Citation of sources for factual information
- Disclosure of commercial interests
- Contact information for author
- Indication of site sponsorship

Although URAC explicitly notes that their accreditation program “is not designed to guarantee a minimum level of quality for health content” (URAC Health Web Site Standards, p 7), it seems likely that consumers would interpret their “Health Site Accreditation” as implying that the site provides trustworthy information, and their criteria include a number of standards that can be interpreted as proxy measures of information quality, including
• The web site does not make claims of therapeutic benefit without reasonable support
• Name of author and information source is provided
• Date content was developed and/or date of last update displayed
• References to information sources provided

Of the three seals discussed here, only the MedCertain uses criteria involve any direct quality assessment, and that is only for the third level of their proposed three-level trustmark system. In the MedCertain approach, first level trustmarks are awarded on the basis of provider claims with respect to proxy measures of information quality, and second level trustmarks are awarded on the basis of consumer evaluation of these same criteria. It is only at the third level of certification that expert assessment of information content is provided. Given this difference between consumer expectation and actual practice, we can only conclude that seals of approval are not transparent interfaces.

This result must be considered with care, since consumer expectations should not necessarily be taken as a practice prescription for seals of approval. There are very good reasons why an accreditation body would choose to certify with respect to proxy indicators rather than through objective assessment of content (Carrol et al., 2002; Eysenbach, 2000b; Eysenbach, 2002; Purcell et al., 2002). Not least among these is the problem that, for most if not all health information, there exists no gold standard that can be used to determine whether the information is in some absolute sense correct. It is possible, therefore, that bringing consumer expectations and the actual significance of seals of approval to congruence is a goal better realized by correcting consumer assumptions than by changing the evaluation practices associated with seals of approval. The data reported here do not point to the solution, but only identify the problem to be solved: as it stands, seals of approval do not deliver what they seem to promise.

As a final note, it is important to address some possible criticisms of this study. First, respondents were a convenience sample of undergraduate students, and thus it cannot be claimed that the sample is formally representative of the population of those seeking health information online. Although this limits the generalizability of the results, the focus of this study is differences between criteria and relationships between the various types of rating tasks, and this type of result is less likely to be influenced by sampling bias (Burkell, 2003). Furthermore, the students who participated in the study are, if anything, more sophisticated than the general public with respect to Internet use and the evaluation of online material, since
a course on this particular issue forms part of their required curriculum. Second, the study examined expectations regarding hypothetical, rather than actual, seals of approval. The responses, therefore, cannot be based on any actual knowledge of the process or criteria for assigning the seals, and it is arguable that understanding of the criteria used to grant a specific real seal of approval (e.g., the HONcode label) would differ from expectations of these generic and hypothetical seals. There are at least two reasons, however, to believe that the results would not be very different if real seals of approval were used. First, even those seals of approval that are most widely used appear on relatively few web sites (e.g., in a recent study, Fallis and Frické (2002) found that only 16% of the health information sites they reviewed displayed the HONcode label), and thus consumers are unlikely to have extensive experience with any seal of approval that would lead to more accurate perception of the conditions under which the seal is granted. Second, earlier research suggests that consumers over-interpret even the most familiar of quality seals (the Good Housekeeping Seal of Approval), suggesting that their interpretation of seals of approval is based more on assumption than on actual knowledge. These results, therefore, are definitely reflective of consumer assumptions regarding new or unknown seals, and likely reflect their assumptions regarding even well-established seals of approval for Internet health information.

Acknowledgements:

The author would like to acknowledge the valuable contribution of an anonymous reviewer, whose thoughtful comments significantly improved the quality of this paper.
References


http://pediatrics.aappublications.org/cgi/reprint/105/1/e1.pdf?ijkey=5e1c638a9d1380bbaacea199a b421eac292e8202 (May 13, 2004).


Burkell - 20


Date: November 17, 2003

Words: 4342
Figure 1: Seals of Approval
Table 1: Quality Criteria

<table>
<thead>
<tr>
<th>Information Quality Criteria</th>
<th>Proxy Quality indicators</th>
<th>Site (or interaction) quality criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information is accurate, reliable and error-free</td>
<td>The web page has no spelling, grammatical, or typographical errors</td>
<td>The web site is accessible with basic software OR links to plug-ins are provided</td>
</tr>
<tr>
<td>Information is unbiased</td>
<td>Information provider is identified</td>
<td>The site is easy to navigate</td>
</tr>
<tr>
<td>Information is complete</td>
<td>Information provider can be contacted</td>
<td>Feedback mechanism is available</td>
</tr>
<tr>
<td>Information is up to date</td>
<td>Web site is sponsored by a reputable organization</td>
<td>Information is appropriate for people who are not medical professionals</td>
</tr>
<tr>
<td></td>
<td>Link to sponsoring organization is provided</td>
<td>A search function is available for the web site</td>
</tr>
<tr>
<td></td>
<td>Information provider qualifications can be verified</td>
<td>Topics covered by the web site are identified (e.g., by site map)</td>
</tr>
<tr>
<td></td>
<td>Web site identifies date of posting to the web and date of last update</td>
<td>The web site is stable and always accessible</td>
</tr>
<tr>
<td></td>
<td>Links are current</td>
<td>The web site is aesthetically pleasing</td>
</tr>
<tr>
<td></td>
<td>Appropriate references are provided so facts can be checked</td>
<td>The information is provided free of cost</td>
</tr>
<tr>
<td></td>
<td>Web site is free of commercial interests OR commercial interests are declared</td>
<td>User support is available if required</td>
</tr>
<tr>
<td></td>
<td>Information has been reviewed for accuracy</td>
<td>Language used is basic and non-technical</td>
</tr>
<tr>
<td></td>
<td>Information provider has definitive knowledge in the area</td>
<td>Site includes a disclaimer describing the limitations of information and use</td>
</tr>
</tbody>
</table>
Table 2: Average Ratings for Quality Criteria (the five highest rated criteria for each task are presented in bold)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Importance</th>
<th>Rating</th>
<th>Seal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information is accurate, reliable, and error-free</td>
<td>9.6</td>
<td>5.0</td>
<td>9.5</td>
</tr>
<tr>
<td>Information is complete</td>
<td>8.7</td>
<td>4.6</td>
<td>9.1</td>
</tr>
<tr>
<td>Information provider has definitive knowledge in the area</td>
<td>9.1</td>
<td>4.6</td>
<td>9.0</td>
</tr>
<tr>
<td>Information has been reviewed for accuracy</td>
<td>8.7</td>
<td>4.3</td>
<td>9.0</td>
</tr>
<tr>
<td>Information is up-to-date</td>
<td>9.0</td>
<td>5.6</td>
<td>8.9</td>
</tr>
<tr>
<td>Information is unbiased</td>
<td>8.4</td>
<td>5.3</td>
<td>7.9</td>
</tr>
<tr>
<td>Appropriate references are provided so facts can be checked</td>
<td>7.8</td>
<td>6.7</td>
<td>7.9</td>
</tr>
<tr>
<td>Information provider qualifications can be verified</td>
<td>7.1</td>
<td>5.1</td>
<td>7.9</td>
</tr>
<tr>
<td>Information provider is identified</td>
<td>7.5</td>
<td>7.2</td>
<td>7.7</td>
</tr>
<tr>
<td>Web site links are current</td>
<td>7.5</td>
<td>6.8</td>
<td>7.6</td>
</tr>
<tr>
<td>Web site has no spelling, grammatical, or typographical errors</td>
<td>7.0</td>
<td>8.2</td>
<td>7.6</td>
</tr>
<tr>
<td>Web site is sponsored by a reputable organization</td>
<td>6.6</td>
<td>5.7</td>
<td>7.4</td>
</tr>
<tr>
<td>Web site is stable and always accessible</td>
<td>7.7</td>
<td>7.3</td>
<td>7.1</td>
</tr>
<tr>
<td>Web site identifies date of posting to the web and date of last update</td>
<td>7.0</td>
<td>7.6</td>
<td>7.1</td>
</tr>
<tr>
<td>Information is appropriate for people who are not medical professionals</td>
<td>7.8</td>
<td>7.5</td>
<td>6.9</td>
</tr>
<tr>
<td>Web site includes a disclaimer describing the limitations of information and use</td>
<td>6.5</td>
<td>6.3</td>
<td>6.8</td>
</tr>
<tr>
<td>Web site is easy to navigate</td>
<td>7.4</td>
<td>8.2</td>
<td>6.5</td>
</tr>
<tr>
<td>Language used is basic and non-technical</td>
<td>7.3</td>
<td>8.1</td>
<td>6.4</td>
</tr>
<tr>
<td>Information is provided free of cost</td>
<td>8.4</td>
<td>7.9</td>
<td>6.3</td>
</tr>
<tr>
<td>Information provider can be contacted</td>
<td>6.0</td>
<td>7.7</td>
<td>6.3</td>
</tr>
<tr>
<td>Web site has a link to sponsoring organization</td>
<td>5.7</td>
<td>7.3</td>
<td>6.3</td>
</tr>
<tr>
<td>Web site has feedback mechanism available</td>
<td>5.8</td>
<td>7.1</td>
<td>6.1</td>
</tr>
<tr>
<td>Web site is free of commercial interests OR commercial interests are declared</td>
<td>6.0</td>
<td>7.3</td>
<td>6.0</td>
</tr>
<tr>
<td>Web site has search function</td>
<td>7.1</td>
<td>7.9</td>
<td>5.8</td>
</tr>
<tr>
<td>Web site includes map or other identification of the topics covered</td>
<td>6.4</td>
<td>7.5</td>
<td>5.8</td>
</tr>
<tr>
<td>Web site provides user support if required</td>
<td>6.2</td>
<td>6.6</td>
<td>5.7</td>
</tr>
<tr>
<td>Web site is accessible with basic software OR links to required plug-ins are provided</td>
<td>6.3</td>
<td>7.3</td>
<td>5.6</td>
</tr>
<tr>
<td>Web site is aesthetically pleasing</td>
<td>5.4</td>
<td>8.2</td>
<td>5.5</td>
</tr>
</tbody>
</table>
Subjects also rated a fourth aspect: the likelihood that they would use the criterion in evaluating health information. This task, however, proved somewhat difficult for subjects and as a result the ratings were difficult to interpret. In light of these issues, this rating task was eliminated from further analyses.

2 The ‘likelihood of use’ task was the third in the package, appearing between ratings of ability to evaluate and contribution to seals of approval.

3 The analysis of variance results reveal that the sphericity assumption is violated for the main effect of rating task and the interaction of task and indicator type. Therefore, these two effects are evaluated with reduced degrees of freedom, using the Lower Bound correction.