Development of a patient-oriented tool for evaluating the quality of breast cancer information on the Internet

HSU Wen-Chin¹, BATH Peter A.

Health Informatics Research Group, Department of Information Studies, University of Sheffield, Sheffield, UK.

Abstract. The aim of this study was to develop a tool for evaluating the quality of breast cancer information on the Internet from the perspective of patients and their families. A specific tool, Breast Cancer tool (BC tool), was developed based on the information needs of women with breast cancer and their families reported in the literature. The BC tool and other 3 generic tools (HON, IQ tool, Discern) were used to assess 40 breast cancer websites. The reliability and validity of each tool was examined and the time spent reviewing the websites was measured. The four tools were shown to have acceptable reliability (Cronbach’s α >0.7), convergent validity, especially the BC tool which was capable of distinguishing whether a website offers sufficient information for women and their families. However, the BC tool took more time than the other tools to use, suggesting relatively low feasibility. The results of this study reinforce the importance of developing specific tools from perspectives of patients and their family members.

Keywords. Evaluation tool, Reliability, Validity, Feasibility, Information quality

1. Introduction

The importance of developing a patient-orientated tool for evaluating medical information on the Internet has been highlighted [1][2]. Although numerous tools have been developed to guide the layperson away from potentially harmful websites [3], these tools are usually designed to assess generic types of health information using pre-defined quality criteria. Several criticisms have been directed at this kind of tool: first, even though websites score well using the quality criteria, the content may be quite poor [4]; second, most of these evaluation tools are designed from the perspective of health professionals, rather than the patients and their family members; third, there is a lack of consensus about quality of health information on the Internet; fourth, although many organizations have established quality criteria to guide and assess health-related

¹ Corresponding Author: Wen-Chin Hsu, Health Informatics Research Group, Department of Information Studies, University of Sheffield, Sheffield, UK. Email: lip06wh@sheffield.ac.uk.
website content (e.g. HON Code, American Medical Association), the reliability and validity of these evaluation tools are unknown.

To overcome these problems, several studies (e.g. Bouchier and Bath, 2003 [5]) have suggested that there is a need to develop specific evaluation tools that assess whether the information meets the needs of particular groups, e.g., for patients/carers affected by multiple sclerosis [2], Alzheimer’s disease [5], and diabetes [6]. No tool was found that had been developed for breast cancer, the most common cause of cancer deaths in women worldwide [7]. Relatively few studies have tested the reliability and validity of their evaluation instruments [8]. The aim of this study was to develop and test a tool for evaluating the quality of breast cancer information on the Internet based on the reported information needs of patients and their families.

2. Methods

The purpose of the Breast Cancer (BC) tool was to determine whether the information provided on a website meets the needs of patients and families. To establish the evaluation criteria of the BC tool, a comprehensive search was conducted using five medical databases (Medline, EMBASE, PsycInfo, CINAHL, Web of Knowledge) using the search terms ‘(information needs) AND (breast cancer) AND (family OR husband OR partner OR women OR spouse)’ for articles published in English between January 1996 and May 2006. This period (1996-2006) was chosen because a previous review [9] had identified the information needs of women with breast cancer and their family members published between 1988 and 1996. The following studies were included: review-type articles focusing on the information needs of, or the information sources and providers for, women with breast cancer and/or their family members. The search resulted in a total of 369 articles, and each article was examined. Twenty-one articles related to the focus of the study were identified and further classified into several categories to help identify whether the target websites provides sufficient information for people at different stages of breast cancer. These categories included information about the disease, symptoms and diagnosis, treatment and breast re-construction, recovery, etc. The items in these categories were then formulated into 47 questions for assessing whether a web-site contained this information. An example of one question is ‘Does the site provide information about the stage of breast cancer?’ For each question a three-point scale (Yes, Partly/Not sure, No) was used to score the extent to which information on that specific issue was present.

Three generic tools (HON code [10], Information Quality (IQ) Tool [11] and Discern [12]) were selected to compare their reliability, validity and feasibility with those of the BC tool because they were developed by authoritative organizations and have satisfactory internal consistency and inter-rater reliability (IQ Tool and Discern) or been successfully applied by previous researchers (HON code) [2][13][14]. The characteristics of these tools are shown in Table 1.

A search of breast cancer websites was conducted on using three search engines with the term ‘breast cancer’. The top 100 web sites identified from each search engine were reviewed, resulting in an initial sample of 300 websites, of which 260 sites (86%) were excluded because they were either broken links, duplicated websites, provided
little or no breast cancer information, or a fee or registration was required for accessing
the information. A final sample of 40 breast cancer websites was included for analysis.

Table 1: Characteristics of the tools used in the study

<table>
<thead>
<tr>
<th>Title of the tool</th>
<th>Developer</th>
<th>No. of questions</th>
<th>Quality Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>HON Code</td>
<td>Health on the Net Foundation (2006)</td>
<td>17*</td>
<td>Authority, Purpose of the website, Privacy, Reference and dated,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Justification of claims, website contact details, Disclosure of funding sources, Advertising policy</td>
</tr>
<tr>
<td>Information Quality</td>
<td>Health Summit Working Group (1999)</td>
<td>21</td>
<td>Authorship, sponsorship, currency, accuracy, confidentiality, navigability</td>
</tr>
<tr>
<td>(IQ) Tool</td>
<td></td>
<td></td>
<td>Reliability, objectivity, information quality of treatment choice</td>
</tr>
<tr>
<td>DISCERN</td>
<td>University of Oxford and the British Library (1997)</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Breast Cancer tool (BC tool)</td>
<td>Authors in this study</td>
<td>47</td>
<td>Information needs of women with breast cancer and their families</td>
</tr>
</tbody>
</table>

* The HON Code consisted of 14 questions. However, some of these questions are open-ended question (e.g. What/How), which are not easily applied to assessing websites. Therefore, considering the convenience of evaluation, the original version of the HON Code was rewritten as 17 closed-ended questions (e.g. Do/Are/Is)

Each website was evaluated in sequence using the three general and one specific tools tool (BC tool) by one of the authors (Hsu). The time spent on each evaluation tool was recorded in order to test the feasibility of using the tools, i.e., whether they were too long for individuals to use them feasibly. Data were inputted into Microsoft Office™ Excel 2003 and imported into Statistical Package for the Social Sciences™ (SPSS) version 13.0 for statistical analyses. Cronbach’s Alpha Coefficient was used to test the internal reliability of each tool. Kendall’s Coefficient of Concordance and Kendall’s Tau B were chosen to test the convergent validity of the four evaluation tools. The time spent to review each web-site across the four tools was compared.

3. Results

The time taken to assess the websites was: IQ tool (mean = 6.43 minutes; SD=2.62); HON Code (mean = 8.24; range = 1.35-28.12; SD= 4.94); DISCERN (mean = 7.36; range = 3.55-16.51; SD = 2.60); BC tool (mean = 14.98; range = 8.10-34.00; SD = 5.41). There was a significant difference among the four tools in the time taken to review the web-sites (p<0.05). The internal consistency of the 4 tools was as follows: IQ tool (Cronbach’s alpha = 0.766); HON code (0.817); Discern (0.816); BC tool (0.876). The agreement among the tools for ranking the web-sites was significant (p<0.001), but the degree of concordance was fair (Kendall’s w=0.426). The concordance among the tools reduced to w =0.301 when the BC tool was excluded. Kendall’s Tau B was used to examine further the agreement between each tool. The correlations between all of the tools was significant (p<0.05), although there were no strong correlations (i.e., > 0.7). The strongest correlation was between the HON and IQ tool at 0.504. There was moderate correlation between the following tools: Discern and HON (0.458), Discern and IQ tool (0.448), Discern and BC tool (0.430). The
weakest correlations were between the BC tool and the HON (0.255), and the BC tool and the IQ tool (0.238).

4. Discussion

Feasibility refers to how long it takes to examine the websites, and whether the examiner had difficulty in using the tools [6]. In this study, the time spent reviewing the websites using each tool was measured to test for differences in the mean time spent reviewing the websites. The observed differences, particularly for the longer time taken using the BC tool, were probably because the BC tool contained considerably more questions than the other three tools, and because the four tools respectively examine different aspects of quality of the websites. The BC tool aims to examine whether the content of the websites meet the needs of women with breast cancer and their families while the others are designed to evaluate generic quality criteria, such as accuracy and objectivity. Examining some criteria can be time-consuming, e.g., the question in the BC tool: “Does the site provide information about the prevention of the treatment’s side effects?”. In some cases, a web-site only provided information about the treatment’s side effects and nothing about their prevention, and consequently, nothing was found to relate to the question after a lot of time had been spent reviewing the website. Compared with the 3 generic tools, the criteria in the BC tool were also harder to assess, being more subjective in nature.

Although the internal consistency of each tool was satisfactory (Cronbach’s α > 0.7) in this study, meeting the requirements for acceptable reliability suggested by Bland and Altman (1997) [15], there were differences in the values of Cronbach’s α for the tools used in this study and the results reported in previous studies, shown in Table 2. For example, the reliability of the HON code tool was higher in this study (0.817) than in Harland and Bath’s study of web-sites concerning multiple sclerosis (0.537). The reliability of the IQ tool however, was lower in this study (0.766) than in Harland and Bath’s study (0.842), but higher than in Ademiluyi et al.’s study of smoking cessation information (0.634). The reliability of the Discern tool was higher in this study (0.816) than in Ademiluyi et al.’s study (0.777). The reliability of the BC tool (0.876) was lower than that of the MS tool developed by Harland and Bath (0.930) [2].

<table>
<thead>
<tr>
<th>Evaluation tools used</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>HON code</td>
<td>0.817</td>
</tr>
<tr>
<td>IQ tool</td>
<td>0.766</td>
</tr>
<tr>
<td>Discern</td>
<td>0.816</td>
</tr>
<tr>
<td>BC tool</td>
<td>0.876</td>
</tr>
<tr>
<td>MS tool</td>
<td>-</td>
</tr>
</tbody>
</table>
These comparisons are relevant because this confirmed that the result of this study is reliable. That is, IQ tool and Discern were separately used in previous studies [2,14], and had shown to have acceptable reliability (Cronbach’s α >0.7).

There are several possible explanations for these differences. First, the selected websites were different. This study focuses on breast cancer websites while Ademiluyi et al. [14] and Harland and Bath [2] separately concentrate on smoking cessation information and Multiple Sclerosis websites respectively. Second, the numbers of sampled websites was different. Ademiluyi et al. [14] sampled 89 websites while Harland and Bath [2] and this study sampled 40 websites. All of these samples constitute less than 100 cases, the minimum number which Kline (2000) [16] suggested to be essential to minimize the standard error of reliability statistics. Moreover, Ademiluyi et al. [14] also concluded that the sample size used in their study may be too small to ascertain the relationships between the scales’ total scores, leading to a Type II error. The numbers of questions varies. The HON code employed by Harland and Bath [2] consisted of 9 questions, whereas the HON code used in this study was formulated into 14 questions. A greater number of questions are likely to increase the reliability if the tool is internally consistent. Third, the evaluation procedures were dissimilar. Two examiners were involved in the evaluation produce conducted by Ademiluyi et al. [14] while there was only one examiner participating in the evaluation produce conducted by Harland and Bath [2], and in this study. Furthermore, the sequence of using the tools to assess the websites may also affect the results. Fourth, the methods of recording the answers are different. In this study, the answers recorded in the HON code are divided into three categories (Yes/Not sure/Partly/No). However, a five-point-Likert-scale is used in Discern.

Overall, all of the four tools measure different quality concepts. The HON and IQ tools measure more similar concepts of quality than the other tools. The quality concepts that the BC tools measure appear to be very different from those of the other tools. This is not surprising because, basically, the BC tool measures different concepts of quality, regarding the information needs of women and their families, rather than the generic quality criteria.

Several limitations must be taken into account when explaining the results in this study. First, because only one reviewer evaluated all of the breast cancer websites, it is not possible to draw definite conclusions about its suitability as a quality assurance tool. Second, because of the dynamic nature of the Internet the results produced in this study may not be generalized to current population of breast cancer web-sites. Third, while the BC tool was designed to determine how well the web-sites met the cognitive needs of women with breast cancer and their families, e.g., information for decision-making, they do not give a clear indication of how well they might meet their affective needs, e.g., for emotional support. Finally, the BC tool was designed to assess the content of information on web-sites, and while it may be useful as a guide for health care professionals to evaluate web-sites before recommending them to individuals, it may be less useful for individuals themselves to use.

Due to time and financial restrictions, the BC tool was only used to assess the breast cancer websites once, and was not refined in the light of the results. Future researchers could modify this tool by using factor analysis, to identify the most important factors contributing to information content, and thus increase its feasibility for regular use.
5. Conclusion

The BC tool had good reliability and demonstrated some degree of convergent validity with other generic tools to assess the content of web-sites. The tool provides a means to evaluate how well web-sites meet the information needs of women with breast cancer and their families based on previous studies of information need in this group. Evaluating web-sites using tools customized for specific diseases complements the use of more generic tools and organizations such as MedCIRCLE, which evaluate and certify health information on the web. More generally, this study shows how a tool for evaluating the quality of information provided on the websites can be developed from the research literature. In practice, the breast cancer tool can be used by breast cancer organizations and support groups who wish to evaluate web-sites before providing access to them for their members. The tool could also be used by organisations who wish to improve the quality of websites providing information specifically relevant to breast cancer. This approach could be applied to develop tools for other diseases.

6. References