Quality of health information on the Internet in pediatric neuro-oncology

Darren R. Hargrave, Ursula A. Hargrave, and Eric Bouffet

Royal Marsden Hospital, Sutton, Surrey, SM2 5PT, UK (D.R.H., U.A.H); and Paediatric Brain Tumour Program, Hospital for Sick Children, Toronto, ON M5G 1X8, Canada (E.B.)

The Internet is now the single largest source of health information and is used by many patients and their families who are affected by childhood brain tumors. To assess the quality of pediatric neuro-oncology information on the Internet, we used search engines to look for information on five common tumor types (brain stem glioma, craniopharyngioma, ependymoma, low-grade glioma, and medulloblastoma). The Web sites were evaluated for content quality by using the validated DISCERN rating instrument. Breadth of content and its accuracy were also scored by a checklist tool. Readability statistics were computed on the highest-rated sites. Of 114 evaluated Web sites, the sources were as follows: institutional, 46%; commercial, 35%; charitable, 15%; support group, 2%; and alternative medicine, 2%. Good interobserver correlation was found for both ratings instruments. The DISCERN tool rated Web sites as excellent (4%), good (7%), fair (29%), poor (39%), or very poor (21%). Only 5% of the Web sites provided one or more inaccurate pieces of information. Web sites were found deficient in topics covering etiology, late effects, prognosis, and treatment choices. Few sites offered information in languages other than English, and readability statistics showed an average required reading level of U.S. grade 12+ (the suggested level being grades 6–8 for an adult audience). The Internet is increasingly being used as a source of oncology information for patients and their families. Health care professionals should be actively involved in developing high-quality information for use in the next generation of Web sites. Neuro-Oncology 8, 175–182, 2006 (Posted to Neuro-Oncology [serial online], Doc. 05-077, February 27, 2006. URL: www.dukeupress.edu/neuro-oncology; DOI: 10.1215/15228517-2005-008)

Keywords: health information, Internet, neuro-oncology, pediatrics, readability

Central nervous system tumors are the leading cause of death from cancer in children, and they result in more morbidity than any other tumor type. High-quality health information is vital for patients and their families and can come from a variety of sources. Studies have confirmed that patients and family alike prefer the primary source of this information to come from their health care team (Levenson et al., 1981; Leydon et al., 2000; Meredith et al., 1996). However, cancer patients and their families report that the information provided by their health care team does not always meet their needs (Houts et al., 1991; Jenkins et al., 2001; Levenson et al., 1983). Several studies of adult oncology patients have established that up to 50% have searched for information about their cancer on the Internet (Carlsson, 2000; Chen and Siu, 2001; Hellawell et al., 2000; Pautler et al., 2001; Vordermark et al., 2000). With this degree of activity, it has become increasingly common for health care professionals to be asked questions based upon the content of Web sites, and there has been concern in the literature about the quality of this information (Crocco et al., 2002; Helft et al., 2003; Roberts 1999; Roberts and Copeland, 2001).

Although a few studies have attempted to objectively rate the quality of health information in oncology (Bichakjian et al., 2002; Biermann et al., 1999), no stud-
ies have looked at Web-based health information for children or the field of neuro-oncology. Therefore, we critically evaluated health information on the Internet, from a patient or family perspective, by examining the content and readability of pediatric neuro-oncology information.

**Methods**

Using the search strategy of Biermann et al. (1999), we searched for health information on five types of childhood brain tumor using the phrases brain stem glioma, craniopharyngioma, ependymoma, low grade glioma, and medulloblastoma. In addition, the general phrase childhood brain tumour was used because this may be a search strategy used by patients and caregivers. The six search terms were entered separately into the six most commonly used English language search engines (at the time of the study): AOL, Google, Lycos, MSN, Netscape, and Yahoo.

Only the first 30 results from each search engine were used for initial evaluation, as previous searches had found that results lower down the relevancy lists were often duplications of earlier results. It was also presumed that lay searchers would be unlikely to scroll through pages and pages of results. The first 30 listed Web sites on each search engine were then visited and initially assessed for suitability to be rated. Web sites did not necessarily provide high-quality criteria for consumer health information in various media (written and online). This tool was used to evaluate the health information relevant to the five tumor types, where available, on every site visited.

Each site was rated independently by the authors, two of whom (D.R.H. and E.B.) are pediatric neuro-oncologists, the third (U.A.H.) being a lay searcher who had no prior knowledge of the subject area and was a competent but nonexpert Internet user. An average of the combined scores was used in the final assessment.

The DISCERN tool does not provide a mechanism for assessing in detail the accuracy of the information provided. Therefore, we adapted the Checklist Rating System Instrument, as described by Bichakjian et al. (2002), to record key neuro-oncology information and record its accuracy for the content topics shown in Fig. 2. For this element of the Web site evaluation, a prior knowledge of the subject was required, so only the pediatric neuro-oncologists (D.R.H. and E.B.) performed this assessment. The knowledge base for the assessment of accuracy came from standard texts such as *Principles and Practice of Pediatric Oncology* (Pizzo and Poplack,

![Combined Rating Scores](https://academic.oup.com/neuro-oncology/article-abstract/8/2/175/1582249/176)

**Fig. 1.** Combined results of the DISCERN rating tool questions (1 = low, 5 = high).

**Quality Rating**

As our aim was to evaluate sites from the perspective of the patient and the patient’s family, we rated all evaluable Web sites with the DISCERN instrument. Developed jointly by workers at Oxford University and the British Library, DISCERN was designed to be used by health consumers and does not require previous knowledge of the subject. The DISCERN instrument is a validated rating tool (Charnock et al., 1999) that is freely available online (www.discern.org.uk) and can be used by health consumers or professionals alike, and a downloadable instruction handbook is available from the DISCERN Web site. A series of 15 questions (see Fig. 1) are asked about the content, and the user rates each one on a scale of 1 to 5 (higher being better). A total score gives an overall rating of the information source and aims to provide high-quality criteria for consumer health information in various media (written and online). This tool was used to evaluate the health information relevant to the five tumor types, where available, on every site visited.
A median of 5147 results were generated for each search term per search engine. By using the first 30 results from each of the six search engines for all six search terms, a total of 1080 Web sites were included for the initial assessment. Of the 1080, 62% were duplicates (i.e., the same Web site listed more than once), which left 478 sites to be visited. Only 3% failed to download, but 75% were found to be unsuitable for full assessment because they did not contain sufficient or suitable information to enable a rating to be undertaken (i.e., 10% were links to other sites, 7% were personal stories, 6% were support groups or message boards, and 50% were unsuitable academic sites, e.g., pathology or radiology slide shows).

A total of 114 Web sites (17 brain stem glioma, 20 craniopharyngioma, 29 ependymoma, 25 low-grade glioma, and 23 medulloblastoma) were fully assessed by two authors using both rating tools. Of the 114 sites assessed, the sources were 46% institutional, 35% commercial, 15% charitable, 2% support group, and 2% alternative medicine. Only two languages (Spanish and German) were offered in addition to the search language (English), and of the 11 multilingual Web sites, nine were based on the NCI Web pages, which provide English and Spanish translation.

**Statistical Methods**

Simple descriptive statistics were used to analyze the data. Simple linear regression and Pearson’s correlation coefficient were used to assess the relationship between the DISCERN and the Checklist Rating System Instrument scores and Web site variables. The interobserver variability for both rating methods was measured by using either weighted or unweighted kappa statistics. The SPSS software statistics package (SPSS, Chicago, Ill.) was used in the analysis.

**Readability of Content**

We used the inbuilt readability statistics in the word processor software program Microsoft Word 2003 (Microsoft Corporation, Redmond, Wash.) as a simple and reproducible measure of the readability of Web site content. The top-five-ranking Web sites for each tumor type, as scored by the DISCERN tool, were cut and pasted into the word processor, and all free text was scored by using the Flesch Reading Ease score and the Flesch-Kincaid Grade Level score. The Flesch Reading Ease score rates text on a 100-point scale. The higher the score, the easier it is to understand.

It is suggested that to achieve “plain English” a score should be around 65 (D’Alessandro et al., 2001; Klingbeil et al., 1995). The Flesch-Kincaid Grade Level score rates text on a U.S. school-grade level (i.e., a score of 8.0 means that an eighth-grader can understand the text). It has been suggested that health information should aim for a grade level of 6.0 to 8.0 (D’Alessandro et al., 2001; Davis et al., 1994).

**Results**

A median of 5147 results were generated for each search term per search engine. By using the first 30 results from each of the six search engines for all six search terms, a total of 1080 Web sites were included for the initial assessment. Of the 1080, 62% were duplicates (i.e., the same Web site listed more than once), which left 478 sites to be visited. Only 3% failed to download, but 75% were found to be unsuitable for full assessment because they did not contain sufficient or suitable information to enable a rating to be undertaken (i.e., 10% were links to other sites, 7% were personal stories, 6% were support groups or message boards, and 50% were unsuitable academic sites, e.g., pathology or radiology slide shows).

A total of 114 Web sites (17 brain stem glioma, 20 craniopharyngioma, 29 ependymoma, 25 low-grade glioma, and 23 medulloblastoma) were fully assessed by two authors using both rating tools. Of the 114 sites assessed, the sources were 46% institutional, 35% commercial, 15% charitable, 2% support group, and 2% alternative medicine. Only two languages (Spanish and German) were offered in addition to the search language (English), and of the 11 multilingual Web sites, nine were based on the NCI Web pages, which provide English and Spanish translation.

**Interobserver Rating Statistics**

A weighted kappa statistic was applied to all the scores for each of the 15 DISCERN questions for all 114 Web sites rated by two authors. The weighted kappa statistics for the DISCERN instrument were 0.54 (95% confidence interval [CI], 0.46–0.62) for raters E.B. versus U.A.H., 0.58 (95% CI, 0.5–0.66) for D.R.H. versus U.A.H., and 0.65 (95% CI, 0.56–0.74) for E.B. versus D.R.H. This represents an acceptable (although not substantial) agreement for a scaled instrument. A nonweighted kappa of
0.88 for D.R.H. versus E.B. was generated for the Checklist Rating System Instrument, representing an acceptable agreement for interobserver ratings.

**Rating Instrument Scores**

For each type of tumor, the Web sites were ranked both according to their total DISCERN score and according to the number of content topics that were present and accurate according to the Checklist Rating System Instrument. The Pearson correlation between the rankings/scores of the two instruments that were used was statistically significant, $r = 0.57$ (95% CI, 0.42–0.69), with a two-tailed $P$ of less than 0.0001. This result suggests good correlation between the two methods, and the content-rating results are therefore displayed according to the DISCERN instrument only because this was our primary assessment tool.

By using the total DISCERN score, from the 15 questions, the Web sites have been grouped into categories of excellent (63–75), good (51–62), fair (39–50), poor (27–38), and very poor (15–26) in content, and this is displayed per tumor type in Fig. 3 and by Web site category in Fig. 4. Web sites were categorized by DISCERN score as excellent (4%), good (7%), fair (29%), poor (39%), and very poor (21%). The mean overall DISCERN score was 34.5 (poor). We attempted subgroup analysis to ascertain any relationships between ratings scores and Web site types. This analysis involved small sample sizes, which may affect the results.

There was no statistically significant difference in rating for Web sites as classified by tumor type. No statistical correlation could be found for a rating score and the order in which a Web site was found on an individual search engines’ listing (see Table 1).

The mean rating scores from all 114 Web sites for each of the 15 DISCERN questions are listed in Fig. 1. The median score for all questions combined was only 2 out of 5. No question had a mean score of 4 or above. The lowest scoring questions related to source of information, alternative sources of information, areas of uncertainty, no treatment, quality of life, choices of treatment, and shared decision making.

The Checklist Rating System Instrument showed that overall only 54% of the content topics selected by the authors as important were included in the Web sites (see Fig. 2). Again, there were no significant differences between Web site scores on the basis of tumor type. The content topic areas best covered were definition of tumor type, signs and symptoms, and types of treatment (all >70% present). However, the topics of etiology, tumor staging, prognosis, late effects, management of recurrence, and tumor-specific research were found in less than 40% of Web sites. There were one or more inaccuracies in 3% of the 114 Web sites evaluated, and the number of inaccuracies varied from 0 to 14 (median, 0; mean, 5) per Web site or 0 to 8% (median, 0%; mean, 3%) per topic. Most inaccuracies were found in the content topics of treatment (surgery, radiotherapy, and chemotherapy) and prognosis.

**Readability Scores**

The median Flesch Reading Score was 20.4 (mean, 27.7; range, 2.3–72.2) for the 25 top-ranking Web sites (five sites regarding the five tumor types) as rated by DISCERN. The recommended score for plain English is 60–70; only two sites achieved this, both from the same hospital and designed for adolescents and children. The median for the Flesch-Kincaid Grade Level was 12 (mean, 11.8; range, 6.8–12), and 64% of Web sites were grade 12. The version of this scoring system as used in the word-processing software described provides only U.S. education grades up to grade 12 (and not college/postgraduate level) and therefore will tend to underrepresent the difficulty of most sites. The recommended level for health literature for adults is suggested as U.S. grades 6–8; again, only the two sites already mentioned achieved this. There was excellent correlation between the two scoring systems ($P < 0.0001$).
Discussion

The staggering growth of the Internet has provided its huge global audience with potentially the world’s most powerful information source. However, this growth has been largely unregulated, and there is concern about the quality of the information that the Internet provides. The American Medical Association has produced guidelines for offering this information online in Medical and Health Information Sites on the Internet (AMA, 2006) and states that the Internet has the potential to speed the transformation of the patient–physician relationship from that of physician authority ministering advice and treatment . . . to that of shared decision making between patient and physician but also states that several substantial barriers exist in realizing this relationship, including equitable access to information, imbalance between patient health literacy and the information provided, extreme variability in the quality of the content (AMA, 2006) and the potential for commercial interests to influence content. This study confirms that within the area of pediatric neuro-oncology these barriers are a reality but that there are examples of good-quality health information on the Internet. Equitable access to information requires a simple means of finding quality information and multilingual content, but we determined that this may be problematic. Internet search engines are the commonest method for navigating the World Wide Web, and we encountered a number of problems in using this strategy. Like us, a patient or family member is faced with thousands of listed links that often lead to duplicate or irrelevant Web pages. There is little guidance from the ordering of search results with respect to quality, as our study found no correlation between order and ratings score (although Google does have some relevancy to its listing order). Upon finding a useful site, a viewer whose first language is not English is unlikely to have access to translation. This is a concern because the sites most often originate within multicultural countries and appear not to provide equitable access for their whole population.

Previous studies have found that an imbalance between patient health literacy and the information provided certainly exists on the Internet, with the majority of such sites written at a reading level that is too complex for most adults in developed countries (Berland et al., 2001; D’Alessandro et al., 2001; Graber et al., 1999). This study confirms these previous results, with

Table 1. The five highest-ranked Web sites according to DISCERN score*

<table>
<thead>
<tr>
<th>DISCERN Ranking</th>
<th>Web Site</th>
<th>URL</th>
<th>Average Google Listing Order</th>
<th>Estimated Hits over 30 Days†</th>
<th>Estimated Traffic Ranking§</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>National Cancer Institute</td>
<td><a href="http://www.cancer.gov">www.cancer.gov</a></td>
<td>11</td>
<td>667,600</td>
<td>8401</td>
</tr>
<tr>
<td>2</td>
<td>eMedicine</td>
<td><a href="http://www.emedicine.com">www.emedicine.com</a></td>
<td>2</td>
<td>1,242,300</td>
<td>3644</td>
</tr>
<tr>
<td>3</td>
<td>Childhood Brain Tumour Foundation</td>
<td><a href="http://www.childhoodbraintumor.org">www.childhoodbraintumor.org</a></td>
<td>3</td>
<td>Not available</td>
<td>Not available</td>
</tr>
<tr>
<td>4</td>
<td>St. Jude Children’s Research Hospital</td>
<td><a href="http://www.stjude.org">www.stjude.org</a></td>
<td>10</td>
<td>127,400</td>
<td>77,193</td>
</tr>
<tr>
<td>5</td>
<td>Children’s Hospital Boston</td>
<td><a href="http://www.childrenshospital.org">www.childrenshospital.org</a></td>
<td>22</td>
<td>140,300</td>
<td>67,945</td>
</tr>
</tbody>
</table>

*The total number of Web sites ranked was 114.
†Data came from www.trafficestimate.com.
§Data came from www.alexa.com (with Yahoo being the Web site ranked number 1).
the Web sites demonstrating an average Flesch Reading Ease score of less than 30 (very difficult) and a U.S. Kincaid Grade Level of 12+ (high school graduate level or higher). This is considerably more difficult than the recommended level for health literature aimed at adults. Unfortunately, the sites with the best content often had the highest reading levels, perhaps negating their usefulness. When one considers that our patients range from young children to adolescents, who may also wish to look at information about their own disease, these readability statistics are of great concern because they effectively exclude the very population that suffers from the tumors they discuss.

There is an extreme variability in the quality of the content of health information in pediatric neuro-oncology on the Internet. The results of our ratings assessment indicate that very few sites achieve high standards by the criteria of the DISCERN tool. Only 11% of sites rated as excellent or good, whereas 60% were rated as poor or very poor. These poor ratings were mainly the result of not providing sufficient details to enable a health information consumer to judge the validity of the information, for example, no details about contributors or lack of referencing. This is a vital part of good practice in providing health information to the public, and this data should always be available. The DISCERN tool also places major significance on the availability of information with respect to shared decision making in health care, such as alternative treatment options and effects on quality of life. Although it is understandable and appropriate that some Web sites may specifically avoid providing information that may be controversial, debatable, or potentially upsetting (e.g., prognosis, treatment of relapse, or specific late effects of therapy), it is useful to direct the viewer to appropriate sources for this information, that is, a more comprehensive Web site, such as the National Cancer Institute or, of course, the patient’s own health care team.

A major worry about medical information on the Internet is inaccuracy. However, the results of our Checklist Rating System Instrument were reassuring in documenting only a few inaccuracies (mean, 5%). These, though, appeared to be clustered in a few sites with a high number of inaccuracies, which meant that any viewer of these particular Web sites could gain contradictory information from the patient’s own health care team that could lead to significant distress. Some of the inaccuracies were simply due to poor presentation, for example, “30% of brain stem gliomas are cured” (no explanation to distinguish the rarer low-grade focal tumors with a good prognosis from the more common diffuse intrinsic pontine glioma with a dismal outcome). More worrying were two examples of Web sites with obvious agendas, which were antiorthodox medicine. These included statements such as “Oncologists must be stopped from using children for experimentation” and “Be certain to undertake biopsy only if it is absolutely essential. Don’t rely on the word of orthodox doctors alone; consult also with progressive or holistic physicians.” Despite these examples, other Web sites described the role of complimentary or alternative therapies in relation to orthodoxy treatment in a balanced and well-informed manner.

Only two sites consistently rated as good or excellent across all five tumor types. The U.S. government site for the National Cancer Institute, which contained comprehensive, well-referenced, and updated information, which was provided in English and Spanish (an adaptation of this site by the University of Bonn also translated this into German). Second was the sponsored site eMedicine, which also provided comprehensive, well-referenced, and updated information by invited experts in the field and also usually provided more than one opinion for each tumor. Both sites provide two levels of information: one version for health professionals and one for patients.

There was no correlation between the type of Web site and either quality of information as rated by the DISCERN tool or the number of inaccuracies found. This may provide some reassurance as pertains to the potential for commercial interests to influence content, although this remains a possible concern.

A possible criticism of this study is that we evaluated all Web sites providing tumor health information, whether or not health consumers were their stated audience. This obviously means that some sites may be judged unfairly when compared to those where providing patient information is the primary aim. However, the study aimed to replicate the situation faced by a patient or family member using a search engine to access more knowledge about a brain tumor, and as stated above, it is unlikely that they will encounter only such dedicated sites. For this reason, we believe there needs to be a shared responsibility of health information consumers and Web site developers.

The general public must be made aware that the Internet is largely unregulated, that information may not be reliable, and that they must be cautious in viewing content not specified for the purpose of patient information. Equally, all those who provide information on the Web must be aware that it may be viewed by those other than their intended audience and that it is wise to state clearly the aim and limitations of any health information-containing Web site.

There has been an extensive debate about whether health information on the Internet should be controlled, rated, or validated in some way and about the possible means of how this could be achieved (Gagliardi and Jadad, 2002; Jadad and Gagliardi, 1998; Kim et al., 1999; Purcell et al., 2002; Shepperd and Charnock, 2002). We chose to use a validated ratings instrument as our primary tool, which was designed mainly to enable health consumers to judge health information. The DISCERN tool provided acceptable interobserver agreement for a graded instrument and also correlated with our second tool, the Checklist Rating System Instrument. However, the time it takes to learn how to use and apply the tool makes the authors doubt that this would be widely accepted by the general public. This means that other methods would need to be employed (e.g., simple consumer guidelines, voluntary codes of conduct, filtering tools, or third-party rating schemes) if some degree of external control is deemed necessary (Durkin, 2000; Edgar et al., 2002; Eysenbach et al.,
2000; Wilson, 2002). Alternatively, one can choose to believe that members of the general public are quite capable of distinguishing between a well-designed, clear, evidence-based Web site and a standard one. However, this means that high-quality, well-maintained Web sites must be available, and this is a major task and an expensive exercise. Ideally, this would be undertaken by national and international cancer groups and funded directly by government as part of public health care.

We decided not to assess the other less tangible elements of the Internet structure such as the architecture of sites, the use of links, and multimedia use (e.g., audio, animation, and video). Although a potentially major advantage over other media types, these elements are difficult to assess objectively (and are not accessible for improved generation of Web sites).

References


Levenson, P.M., Pfefferbaum, B., Silberberg, Y., and Copeland, D.R. (1981) Sources of information about cancer as perceived by adoles-


