Quality of Web-based information on obsessive compulsive disorder

Background: The Internet is increasingly used as a source of information for mental health issues. The burden of obsessive compulsive disorder (OCD) may lead persons with diagnosed or undiagnosed OCD, and their relatives, to search for good quality information on the Web. This study aimed to evaluate the quality of Web-based information on English-language sites dealing with OCD and to compare the quality of websites found through a general and a medically specialized search engine.

Methods: Keywords related to OCD were entered into Google and OmniMedicalSearch. Websites were assessed on the basis of accountability, interactivity, readability, and content quality. The “Health on the Net” (HON) quality label and the Brief DISCERN scale score were used as possible content quality indicators. Of the 235 links identified, 53 websites were analyzed.

Results: The content quality of the OCD websites examined was relatively good. The use of a specialized search engine did not offer an advantage in finding websites with better content quality. A score \( \geq 16 \) on the Brief DISCERN scale is associated with better content quality.

Conclusion: This study shows the acceptability of the content quality of OCD websites. There is no advantage in searching for information with a specialized search engine rather than a general one. Practical implications: The Internet offers a number of high quality OCD websites. It remains critical, however, to have a provider–patient talk about the information found on the Web.

Keywords: Internet, quality indicators, anxiety disorders, OCD, search engine

Introduction

Obsessive compulsive disorder (OCD) is an anxiety disorder associated with considerable impairment in quality of life and functioning.\(^1\) The disorder, classified in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Revision (DSM-V), Obsessive-Compulsive and Related Disorders, is known for a high percentage of severe cases.\(^2\) Available treatments include pharmacological management and cognitive and behavior therapy,\(^3,4\) with treatment sometimes requiring a combination of multiple strategies.\(^5\) The worst prognosis for the disorder is associated with an earlier age of onset and a longer duration of illness. It has been suggested that patients with OCD may benefit from prolonged continuous treatment.\(^6\)

The disorder, due to its frequency, its burden, and possible shame related to seeking help may lead persons with diagnosed or undiagnosed OCD, and their relatives, to search for good quality information on the Web. It has previously been shown that people with psychiatric disorders,\(^7\) and particularly people with OCD, frequently use the Internet as an information source.\(^8\) Approximately 80% of Internet users in developed countries search for information about their health, typically...
about diseases, symptoms, and treatments.\textsuperscript{9} Therefore, it is important for websites to present high quality information on OCD.

Recognition of the central role of the Internet as an information source on health has been put into perspective by the increasing concern about the variable quality of this information.\textsuperscript{10} Although a few studies found rather good content quality,\textsuperscript{11,12} most found that the overall quality of content on health-related websites seems poor.\textsuperscript{13–19} Moreover, one study evaluating the quality of Dutch websites related to OCD concluded that the quality is generally poor.\textsuperscript{20} When facing an abundance of information on health-related websites and the questionable and variable quality of the content, the Internet user is likely to be a little lost.

In response to these concerns, a number of initiatives have been developed to establish quality criteria for health-related websites and to help users to find those that are of good quality. These initiatives include quality labeling such as performed by the Health on the Net Foundation (HON), with a focus on ethical standards related to online publishing;\textsuperscript{21} DISCERN,\textsuperscript{22,23} a questionnaire to help laypersons distinguish good from bad treatment-related information (using questions such as “Is the information balanced and unbiased?”); a six-item Brief version of DISCERN;\textsuperscript{24,25} and specialized search engines.\textsuperscript{26} A Brief DISCERN cutoff score of ≥16 was previously associated with better quality content of health-related websites.\textsuperscript{27}

The present study aimed to assess the content quality of English-language OCD-related websites, to determine the content quality indicators, and to compare the quality of websites found from a search using a general search engine versus one using a specialized search engine. As suggested by previous studies on other topics, we expected the quality of the information to be rather poor. Furthermore, we expected that a specialized search engine (OmniMedicalSearch) to be better than a general search engine (Google) in finding better content quality websites.

**Materials and methods**

A typical search was performed to produce a list of websites similar to one that would be generated by a common user with limited medical or Internet knowledge.

**Selection of websites**

Keyword searches and website evaluations were performed in December 2011 by the first author. Two similar searches were made using the following search engines: Google, the most commonly used general search engine, and OmniMedicalSearch,\textsuperscript{26} a medically specialized meta search engine that gives results from different sources and questions up to 12 different medical search engines. The following queries were entered into the two search engines: “obsessive compulsive disorder,” “OCD,” “obsessive compulsive disorder AND help,” “obsessive compulsive disorder AND treatment.”

Given that most people rarely look beyond the first 20 links returned from a search,\textsuperscript{27} we decided to assess at least the first 20 links from each request, or more if more were available, in order to obtain at least 20 websites for each query. In the present study, due to the insufficient number of OCD pertinent links among the first 20 links, we continued the search until the inclusion of 20 websites per query if available.

The links were excluded from analysis for the following reasons: they were inaccessible (invalid address), they had already been assessed in the current study (ie, to avoid repetition), access to the website required payment or a password, they were newsgroup or open forum sites, they did not correspond to a real website (external links, books, articles), or there was no information in English.

**Evaluation of websites**

The websites were divided into five categories according to their statement of affiliation: commercial, nonprofit organization, university, governmental, or personal pages. The presence of the HON logo was also recorded.

Websites were assessed with a standardized assessment tool based on previous studies.\textsuperscript{11,28} As previously shown, the components of the assessment tool have good inter-rater reliability Silberg ($r=0.841; P<0.05$), readability index – Flesch-Kincaid ($r=0.881; P<0.05$), grade level score – Flesch-Kincaid ($r=0.835; P<0.01$), Abbott’s esthetic criteria ($r=0.751; P<0.05$), DISCERN ($r=0.942; P<0.01$), content quality ($r=0.851; P<0.01$), and interactivity ($r=0.865; P<0.01$).\textsuperscript{11} The assessment tool included the following outcomes.

**Accountability**

Accountability was estimated with a scale of nine items (Silberg scale),\textsuperscript{29} including authorship (names of authors, affiliation, and references), attribution (sources and references), disclosure (property of site, sponsorship, and advertising), and currency (date of creation, modification of site, and updating in the last 6 months). A total score (Silberg) ranging from 0 to 9 (1 point for each item if present) was calculated for each site.
Interactivity
Interactivity was measured with the adapted version of the Abbott scale, which estimates the presence of an internal search engine, the presence of audio or video support, questionnaires of satisfaction or testing users’ knowledge, spaces of support such as forums, and the possibility of sending complaints and requests to the webmasters or to the authors (1 point for each item if present).

Presentation and esthetics
Presentation and esthetics were estimated with “Abbott esthetic criteria,” as adapted by Kisely et al. This score estimates the presence of titles/subtitles, diagrams, and hyperlinks, as well as the absence of advertising (1 point for each item).

Readability
Readability was evaluated with the Flesch-Kincaid grade level score and the Flesch-Kincaid readability index, using the following link: http://www.online-utility.org/english/readability_test_and_improve.jsp. The first score provides an idea of the level of studies by estimating the difficulty of the text in comparison with a US grade level, with higher scores reflecting higher levels of difficulty. The Flesch-Kincaid readability index varies from 0 to 100, with higher scores translating into better readability.

Content quality
This was estimated according to availability of information in connection with the following seven questions.
1. How do I know whether I have OCD? (symptoms)
2. Can I estimate the severity of my disease? (severity)
3. What are the effective treatments for OCD? (availability of treatments)
4. Who can I contact to treat my OCD? (caregivers)
5. What are the various sorts of useful psychotropic drugs for OCD, and what are their side effects? (medicines: types and side effects)
6. How long should I undergo medical treatment? (duration of treatment)
7. What are effective psychotherapies in the treatment of OCD? (psychotherapies)

The information found on websites in connection with these questions was compared with the consensus developed by the Task Force of the World Federation of Societies of Biological Psychiatry. For every request, the coverage (to what extent the question was addressed) and correctness (to what extent the answer was right) were scored on a 3-point scale (0= absent, 1= minimal, 2= sufficient). A total content quality score, ranging from 0 to 28, was calculated by combining the scores of the coverage and correctness scales.

The Brief version of the DISCERN instrument
The Brief DISCERN instrument was used as a potential indicator to estimate the quality of the information about the choice of treatment. The Brief DISCERN includes six items on a five-point scale (1= not at all, 5= completely). The first two items identify the transparency of information sources; the other four estimate the quality of information about treatment. A cutoff score equal to or higher than 16 was proposed to help laypersons detect “good content quality websites.”

Global score
This was computed and defined as the sum of Silberg, interactivity, Abbott’s esthetic criteria, and content quality.

Statistical analyses
Statistical analyses were performed by SPSS (version 18.0; IBM Corporation, Armonk, NY, USA). An initial exploratory analysis involved the calculation of proportions, as well as means and standard deviations, of the abovementioned outcome measures.

Student’s t-tests were used to test the equality of the means of these outcomes in websites with and without the HON label, in websites with Brief DISCERN scores that were ≥16 compared with scores of <16, and in websites exclusively found either by Google or by OmniMedicalSearch.

One-way analyses of variance and/or Kruskal–Wallis tests where appropriate were conducted to test whether any differences existed among the means (average outcome values) for the groups of affiliation (nonprofit organization, commercial, university, government, and personal pages) on the one hand, and for the search engines that found the website (Google, OmniMedicalSearch, or both) on the other. For all analyses, a significance level of P<0.05 was used.

Results
At the end of the queries, 235 links were assessed, 107 related to Google, and 128 related to OmniMedicalSearch. The assessed links corresponded to 125 websites (80 with Google and 45 with OmniMedicalSearch). After exclusion of repeated websites between search engines, we retained 53 sites to be analyzed (Figure 1). Some of the assessed websites (37.7%) were found on both Google and OmniMedicalSearch, 35.8%...
42. http://www.mayoclinic.org/psychiatry/services/obsessive_compulsive_disorder
47. http://www.ocdtribe.com/

Figure 1 List of the websites.
were found only on OmniMedicalSearch, and 26.4% only on Google.

The origins of the sites were mostly nonprofit (73.6%) and commercial organizations (18.9%). University, government, and personal pages were grouped under the category “other” because of the smallness of the sample size, corresponding to 7.5% of the websites assessed. Among the 53 included sites, 12 (22.6%) had the HON label. Additional content in foreign languages was found as follows: 18.9% in Spanish, 9.4% in French, 1.9% in German, and 3.8% in other languages.

The means and standard deviation scores of the websites are described in Table 1. Content quality scores were highly variable, ranging from 5 to 28. Its mean score was 16.6 ± 4.8, which was higher than an average of 14 out 28 (50% of the maximum score). The Brief Discern scores were also highly variable and ranged from 6 to 30. Its mean score was 19.1 ± 5.5 out of 30, higher than the proposed cutoff of 16.

Most of the websites (80.8%) recommended seeking clarification from a health professional. Local health resources were mentioned in 37.7% of the assessed websites. The mean level of interactivity was low at 1.7 ± 1.4 out of 6. Only a small proportion (13.2%) of the websites offered photo or video illustration. An intra-site search engine appeared in 56.6% of the websites, a support group in 24.5%, and the possibility of sending queries to a webmaster in 36.5%.

The mean Silberg score was 5.6 ± 2.0 (minimum 2, maximum 9). Most of the sites clearly specified when the site had been created or modified (78.8%) and mentioned sources or references (64.2%). The average Flesch-Kincaid reading grade was 11.0 ± 1.2, higher than 8, the recommended level for standard documents. Comparison between sites having the HON level with those that do not have this label only revealed one significant difference (Table 2). Brief DISCERN scores were significantly higher on sites with the HON label than on sites without it (t=-2.1 and P=0.04).

Regarding the quality of websites with type of affiliation as factor (non-profit organization versus commercial versus other), there were no statistically significant differences on the main outcomes except for Aestheticism-Abbott which was just significant (Table 3).

However, when assessing the quality of the websites in relation to the search engine used (Google only versus OmniMedicalSearch only versus Google and OmniMedicalSearch), we observed significant between-group differences for Silberg scores, content quality and global scores. The use of Bonferroni post-hoc tests showed that Silberg scores were higher in sites found with both Google and OmniMedicalSearch than in sites found through Google or OmniMedicalSearch alone. Content quality and global scores were higher for sites found with both Google and OmniMedicalSearch than in sites found through OmniMedicalSearch alone (Table 4). Additional analyses comparing sites found only by Google to those found only by OmniMedicalSearch did not reveal significant difference for any outcomes studied. In other words, websites found only on OmniMedicalSearch did not show better scores than those found via a Google query. Moreover, websites with a Brief DISCERN score of ≥16 had higher global scores, higher

### Table 1 Means and SDs of scores of website quality indicators

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silberg (0–9)</td>
<td>5.6 (2)</td>
</tr>
<tr>
<td>Interactivity (0–6)</td>
<td>1.7 (1.4)</td>
</tr>
<tr>
<td>Estheticism – Abbott (0–4)</td>
<td>2.1 (0.8)</td>
</tr>
<tr>
<td>Readability index – Flesch-Kincaid (0–100)</td>
<td>39.6 (14.4)</td>
</tr>
<tr>
<td>Grade level score – Flesch-Kincaid (1–12)</td>
<td>11 (1.2)</td>
</tr>
<tr>
<td>Content quality (0–28)</td>
<td>16.6 (4.8)</td>
</tr>
<tr>
<td>Global score (0–47)</td>
<td>26.1 (6.6)</td>
</tr>
<tr>
<td>Brief DISCERN (6–30)</td>
<td>19.1 (5.5)</td>
</tr>
</tbody>
</table>

**Abbreviation:** SD, standard deviation.

### Table 2 Comparison of sites having the HON label to those without it by t-tests

<table>
<thead>
<tr>
<th>Measure</th>
<th>With HON label (n=12)</th>
<th>Without HON label (n=41)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silberg</td>
<td>6.4 (2.3)</td>
<td>5.4 (1.9)</td>
<td>0.1</td>
</tr>
<tr>
<td>Interactivity</td>
<td>1.7 (1.2)</td>
<td>1.8 (1.5)</td>
<td>0.8</td>
</tr>
<tr>
<td>Estheticism – Abbott</td>
<td>1.9 (0.8)</td>
<td>2.2 (0.8)</td>
<td>0.3</td>
</tr>
<tr>
<td>Readability index – Flesch-Kincaid</td>
<td>38.9 (13.1)</td>
<td>39.9 (14.9)</td>
<td>0.8</td>
</tr>
<tr>
<td>Grade level score – Flesch-Kincaid</td>
<td>11 (1.3)</td>
<td>11.2 (0.9)</td>
<td>0.5</td>
</tr>
<tr>
<td>Content quality</td>
<td>17.8 (5.0)</td>
<td>16.2 (4.7)</td>
<td>0.3</td>
</tr>
<tr>
<td>Global score</td>
<td>27.8 (7.1)</td>
<td>25.5 (6.4)</td>
<td>0.3</td>
</tr>
<tr>
<td>Brief DISCERN</td>
<td>21.9 (4.5)</td>
<td>18.2 (5.6)</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**Abbreviation:** HON, Health on the Net.

### Table 3 Comparison of quality measures by affiliation by one-way analysis of variance

<table>
<thead>
<tr>
<th>Measure</th>
<th>Nonprofit organization</th>
<th>Commercial</th>
<th>Other</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silberg</td>
<td>5.8 (1.9)</td>
<td>5.4 (2.3)</td>
<td>4 (2.2)</td>
<td>0.2</td>
</tr>
<tr>
<td>Interactivity</td>
<td>1.7 (1.4)</td>
<td>2.0 (1.7)</td>
<td>1.0 (0.0)</td>
<td>0.2</td>
</tr>
<tr>
<td>Estheticism – Abbott</td>
<td>2.2 (0.7)</td>
<td>1.6 (0.7)</td>
<td>2.5 (1.0)</td>
<td>0.05</td>
</tr>
<tr>
<td>Readability index – Flesch-Kincaid</td>
<td>41.8 (14.5)</td>
<td>37.2 (11.8)</td>
<td>25.0 (12)</td>
<td>0.07</td>
</tr>
<tr>
<td>Grade level score – Flesch-Kincaid</td>
<td>10.8 (1.3)</td>
<td>11.5 (0.5)</td>
<td>12.0 (0.0)</td>
<td>0.1</td>
</tr>
<tr>
<td>Content quality</td>
<td>17 (4.6)</td>
<td>17.0 (4.0)</td>
<td>11.7 (6.6)</td>
<td>0.1</td>
</tr>
<tr>
<td>Global score</td>
<td>26.8 (6.3)</td>
<td>25.9 (5.8)</td>
<td>19.2 (8.9)</td>
<td>0.09</td>
</tr>
<tr>
<td>Brief DISCERN</td>
<td>19.3 (5.7)</td>
<td>20.6 (4.3)</td>
<td>13.5 (4.4)</td>
<td>0.09</td>
</tr>
</tbody>
</table>
content quality, and higher Silberg scores than those with a Brief DISCERN score of <16 (Table 5).

### Discussion

The present study aimed to assess English-language websites on OCD to determine content quality indicators and to compare the results found via a general search engine with those found via a specialized search engine. In contrast to results of previous studies on health- and mental health-related websites,

<table>
<thead>
<tr>
<th>Measure</th>
<th>Google</th>
<th>OmniMedicalSearch</th>
<th>Google + OmniMedicalSearch</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silberg</td>
<td>5.1 (2.2)</td>
<td>4.5 (1.7)</td>
<td>7 (1.3)</td>
<td>&lt;0.0005</td>
</tr>
<tr>
<td>Interactivity</td>
<td>2.3 (1.8)</td>
<td>1.5 (1.2)</td>
<td>1.6 (1.2)</td>
<td>0.2</td>
</tr>
<tr>
<td>Estheticism – Abbott</td>
<td>2.0 (0.08)</td>
<td>1.9 (0.07)</td>
<td>2.4 (0.8)</td>
<td>0.1</td>
</tr>
<tr>
<td>Readability index – Flesch-Kincaid</td>
<td>41.4 (11.5)</td>
<td>37.1 (13.8)</td>
<td>40.8 (16.8)</td>
<td>0.6</td>
</tr>
<tr>
<td>Grade level score – Flesch-Kincaid</td>
<td>10.9 (1.3)</td>
<td>11.4 (0.09)</td>
<td>10.7 (1.4)</td>
<td>0.2</td>
</tr>
<tr>
<td>Content quality</td>
<td>17.1 (4.4)</td>
<td>14.3 (4.7)</td>
<td>18.3 (4.5)</td>
<td>0.03</td>
</tr>
<tr>
<td>Global score including content quality</td>
<td>26.5 (7.2)</td>
<td>22.3 (5.9)</td>
<td>29.3 (4.9)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

The websites identified by a query on a medically specialized search engine were found to be no better than those found through a Google search, whereas those found by both the Google and the medically specialized search engine had higher content quality scores. As previously shown, the Brief DISCERN is likely to be an interesting content quality indicator.

Our study contains several limitations. The search methods used in this article to identify websites do not cover all methods patients may use. It is possible that some Internet users have different search methods from ours or use different keywords or different search engines. On the other hand, the results provide a snapshot of the situation in a limited period of time (December, 2011). Furthermore, the study did not take into account possible links between website use (frequency and type of use) and quality measures. Further collaborations with website owners and webmasters may lead to a better understanding of links between website content and consumer use. Further studies may also assess websites related to other topics or populations such as adolescent, elderly, migrants, women, or men’s specific needs.

This study brings to light the acceptability of the content quality of OCD-related websites. A possible method to improve OCD websites would be to associate content quality with better interactivity or, as suggested elsewhere, to promote the development of informative versus interactive websites.

The Internet offers a number of good content quality OCD websites. The use of a specialized search engine does not offer an advantage in finding websites with better content quality. The Brief DISCERN could facilitate the identification of good information on the Web by patients and general consumers. It remains critical, however, to have a provider–patient talk about the information found on the Web, as suggested elsewhere,

### Table 5 Comparison of sites by Brief DISCERN (score <16 versus score ≥16) using t-tests

<table>
<thead>
<tr>
<th>Measure</th>
<th>&lt;16 (n=17)</th>
<th>≥16 (n=36)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silberg</td>
<td>4.4 (1.7)</td>
<td>6.2 (1.9)</td>
<td>0.001</td>
</tr>
<tr>
<td>Interactivity</td>
<td>1.4 (0.9)</td>
<td>1.9 (1.6)</td>
<td>0.2</td>
</tr>
<tr>
<td>Estheticism – Abbott</td>
<td>1.9 (0.6)</td>
<td>2.3 (0.8)</td>
<td>0.08</td>
</tr>
<tr>
<td>Readability index – Flesch-Kincaid</td>
<td>42.3 (16.2)</td>
<td>38.4 (13.5)</td>
<td>0.4</td>
</tr>
<tr>
<td>Grade level score – Flesch-Kincaid</td>
<td>10.7 (1.6)</td>
<td>11.2 (1.0)</td>
<td>0.3</td>
</tr>
<tr>
<td>Content quality</td>
<td>12.5 (4.0)</td>
<td>18.5 (3.8)</td>
<td>&lt;0.0005</td>
</tr>
<tr>
<td>Global score</td>
<td>20.1 (4.8)</td>
<td>28.9 (5.4)</td>
<td>&lt;0.0005</td>
</tr>
</tbody>
</table>
Disclosure

The authors report no conflicts of interest in this work.

References


