Quality Evaluation of Consumer Health Information Websites
Found on Google Using DISCERN, CRAAP, and HONcode

Ivan A. Portillo
Catherine V. Johnson
Scott Y. Johnson

Follow this and additional works at: https://digitalcommons.chapman.edu/librarian_articles

Part of the Health Communication Commons, Health Sciences and Medical Librarianship Commons, Information Literacy Commons, and the Other Medicine and Health Sciences Commons
Quality Evaluation of Consumer Health Information Websites Found on Google Using DISCERN, CRAAP, and HONcode

Comments
This is an Accepted Manuscript of an article published in Medical Reference Services Quarterly, volume 40, issue 4, in 2021, available online at https://doi.org/10.1080/02763869.2021.1987799. It may differ slightly from the final version of record.

The Creative Commons license below applies only to this version of the article.

This scholarship is part of the Chapman University COVID-19 Archives.

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial 4.0 License

Copyright
Taylor & Francis
Quality Evaluation of Consumer Health Information Websites found on Google using DISCERN, CRAAP, and HONcode

Ivan A. Portillo\textsuperscript{a}* Catherine V. Johnson\textsuperscript{b} and Scott Y. Johnson\textsuperscript{c}

\textsuperscript{a}Leatherby Libraries, Chapman University, Orange, California, USA; \textsuperscript{b}Wilson Library, University of La Verne, La Verne, California, USA; \textsuperscript{c}M.B. Ketchum Memorial Library, Marshall B. Ketchum University, Fullerton, California, USA

*Corresponding author: iportillo@chapman.edu

Biographical notes on contributors

Ivan A. Portillo, MLIS, AHIP (iportillo@chapman.edu) is a Health Sciences Librarian at Leatherby Libraries, Chapman University, 9401 Jeronimo Road, Irvine, CA 92618 USA.

Catherine V. Johnson, MLIS (cjohnson2@laverne.edu) is a Research and Instruction Librarian at Wilson Library, University of La Verne, 2040 Third Street, La Verne, CA 91750 USA.

Scott Y. Johnson, MLIS, AHIP (scottjohnson@ketchum.edu) is Director of Library Services at M.B. Ketchum Memorial Library, Marshall B. Ketchum University, 2575 Yorba Linda Boulevard, Fullerton, CA 91831 USA.

Please provide ORCiD for each author, if available

ORCID

Ivan A. Portillo - https://orcid.org/0000-0002-8031-1006

Catherine V. Johnson

Scott Y. Johnson
Quality Evaluation of Top-Ranking Consumer Health Information Websites on Google using DISCERN, CRAAP, and HONcode

ABSTRACT

Online health misinformation is a growing problem, and health information professionals and consumers would benefit from an evaluation of health websites for reliability and trustworthiness. Terms from the Google COVID-19 Search Trends dataset were searched on Google to determine the most frequently appearing consumer health information websites. The quality of the resulting top five websites was evaluated. The top five websites that appeared most frequently were WebMD, Mayo Clinic, Healthline, MedlinePlus, and Medical News Today, respectively. All websites, except Medline Plus, received HONcode certification. Based on DISCERN and CRAAP scores, MedlinePlus was found to be the most reliable health website.

Keywords: website evaluation; consumer health information; DISCERN; CRAAP; HONcode, COVID-19

INTRODUCTION

Searching for health information online is one of the most common uses of the internet. The 2013 Pew Study for Internet and Health found that 59% of people use a search engine, such as Google, to begin their health-related research.1 Almost a decade later, a 2021 study by Wolters Kluwer found that two-thirds of Americans trust the health information they find online.2 Taken together, one could infer that Google search results can have an outsized impact on the health information that consumers find and trust. Health misinformation on the internet is a growing problem, especially on social media.3 Vaccines, pandemics, and medical interventions are among
the topics with the largest amount of misinformation online.\textsuperscript{3} Considering the impact of COVID-19 on the daily lives of the world’s population, finding reliable and authoritative health information online is of vital importance.

Currently, there is a dearth of research regarding the overall evaluation of top-ranking consumer health information websites. Many studies on consumer health information-seeking behaviors typically focus on a specific disease or condition, such as pregnancy, and evaluate websites that concentrate on these topics.\textsuperscript{4} Moore and Ayers evaluated postnatal mental health websites based on their own rating criteria and generated a list of five recommended websites.\textsuperscript{5} Studies on orthodontic pain and concussions used evaluation criteria similar to Moore and Ayers’s study.\textsuperscript{6-7} Another study attempted to create a methodological approach for reviewing and analyzing health information websites, focusing on the health practitioner side and not the consumer.\textsuperscript{8} Furthermore, many studies on consumer health information overload exist, suggesting they are uncertain about where to find reliable information.\textsuperscript{9}

This study evaluated the quality of top-ranking health websites in Google’s search results on COVID-19 using the DISCERN instrument and CRAAP (Currency, Relevance, Authority, Accuracy, and Purpose) test. The DISCERN instrument and CRAAP test have both been used in health information research and are validated measures of website evaluation. The results of this research will help health information professionals and consumers choose the most reliable health information websites when they are searching for health information using Google.

The DISCERN instrument “has been designed to help users of consumer health information judge the quality of written information about treatment choices”.\textsuperscript{10} The instrument consists of 16 questions used to evaluate the reliability, relevance, authorship, potential bias, and treatment options described by the consumer health website.\textsuperscript{10} The DISCERN instrument is
accessible to consumers and helpful to differentiate between high and low-quality health information online.\textsuperscript{10} The reliability of the DISCERN instrument has been tested on multiple occasions, and a substantial agreement between independent raters has been found.\textsuperscript{11-13} The DISCERN instrument has been frequently used in research on specific written health topics, such as chronic pain, vascular tumors, and breast or prostate cancer.\textsuperscript{14-17} It has also been used to evaluate misinformation online related to the COVID-19 virus.\textsuperscript{18}

The CRAAP test is a librarian-developed tool used by librarians to teach users about the validity of the sources they find online.\textsuperscript{19-20} The test is a 5-point evaluation tool based on the consistency, relevance, authority, accuracy, and purpose of the source.\textsuperscript{19} It has also been used to evaluate patient health information in pediatric neurosurgery and training materials for psychiatry residents.\textsuperscript{21-22} Recently, Liu re-evaluated the CRAAP test and concluded that it is still an effective evaluation method but could be improved with a critical thinking component.\textsuperscript{23}

Health on the Net Foundation Code of Conduct (HONcode) certification of websites should theoretically allow consumers to avoid the use of evaluation tools, as the seal displayed on approved websites indicates that the website has met specific standards.\textsuperscript{24} The main principles of HONcode certification are clearly listing author qualifications, citing sources and dates, objectivity, transparency, financial disclosure, clearly distinguishing advertising, and confidentiality for users.\textsuperscript{24} Numerous studies have used HONcode certification to assess the reliability of health information websites. Three studies focused on websites pertaining to inflammatory bowel disease, urology, and prostate cancer, and all used HONcode in conjunction with the DISCERN instrument.\textsuperscript{25-27} Despite commonly being used as a measure of reliability, HONcode may not help consumers when it comes to readability, relevancy, or comprehensiveness of their topic of interest.\textsuperscript{28}
MATERIALS AND METHODS

Search term selection
To replicate common health information searches, a list of 65 search terms (see Appendix) was gathered from the COVID-19 Search Trends symptoms dataset. These search terms, taken from a dataset of several hundred symptoms, were chosen to generate a representative sample of health-related search terms that a health consumer may search for on Google. Each term, along with the term “COVID-19”, was searched on Google in March 2021, and a list of the first 10 search results for each term was recorded. The top 10 search results were chosen because that is the typical number of results on the first page in Google.

Website selection
To ensure as much impartiality as possible and avoid targeted results, the web browser was set to incognito/private mode, and the search history and cookies were cleared before conducting searches on Google. According to Google, search results can vary by country and region. Each term was searched in three different regions of the United States [West Coast (Los Angeles), Midwest (Chicago) and East Coast (New York)] using a Virtual Private Network (VPN). After compiling the top 10 search results for all 65 terms in each of the three regions, the researchers determined the top five websites that appeared most frequently in Google.

Website evaluation
The quality of the resulting top five websites was evaluated independently by two authors using the DISCERN instrument and CRAAP test. For consistency, the researchers reviewed the same general topic on each website – diabetes. For the DISCERN instrument, each website was evaluated against the sixteen questions from the DISCERN Instrument. Each question was given
a score of 1-5 (1 for a definite no, 5 for a definite yes, and 2-4 for partially), and a total score out of 80 was calculated. According to the creators of DISCERN, a score from 63-80 is “Excellent”, 51-62 is “Good”, 39-50 is “Fair”, 27-38 is “Poor”, and 15-26 is “Very Poor”. For the CRAAP test, the two authors independently evaluated each website based on five CRAAP criteria, giving each a score from 1-5 (1 for “Unreliable”, 5 for “Excellent”, and 2-4 “Partially”). A total score out of 25 was calculated. A score of 20-25 is “Excellent”, 15-19 is “Good”, 10-14 is “Average”, 5-9 is “Borderline Acceptable”, and 0-4 is “Unacceptable”. The websites were ranked based on the averaged total scores for each website.

The top five resulting websites were also reviewed for HONcode certification. Websites that have received HONcode certification are considered reliable and credible sources of health information. While the websites were not given a score for HONcode certification, it served as another data point to determine a website’s reliability.

**RESULTS**

A total of 65 search terms from the Google COVID-19 Search Trends symptoms dataset were selected and searched on Google. The first 10 search results for each term was retrieved, resulting in 123 unique websites and a total of 1,950 URLs appearing across the three regions searched via VPN. The website that appeared most frequently within the 1,950 search results was WebMD with 216 results. Mayo Clinic was the only other website to appear over 200 times.

The top five websites that appeared most frequently were WebMD, Mayo Clinic, Healthline, MedlinePlus, and Medical News Today, respectively. These five websites were chosen to be evaluated as they appeared at least 40 times in the search results of each region. Wikipedia was the sixth most frequently appearing website, averaging 35 appearances in each
region. Each region had the same top five websites, but with varying number of appearances in the search results (Table 1). Local search results, such as nearby clinics or treatment centers, did appear in the results but not for every search term.

[Table 1 near here]

The top five websites were evaluated independently by two of the authors using the DISCERN instrument and CRAAP tool. The top-rated website using both evaluation tools was MedlinePlus. MedlinePlus was also the only website evaluated that was produced by a government agency. Using the DISCERN instrument, MedlinePlus was the only website to have a quality rating of “Excellent”, with Medical News Today having the lowest quality rating of “Fair”. Using the CRAAP tool, MedlinePlus was also rated as “Excellent”, and the remaining websites received a quality rating of “Good”. Medical News Today received a rating of 14.5 using CRAAP, which fell between “Average” (10-14) and “Good” (15-19).

[Table 2 near here]

All five websites were also visited to verify if they had received HONcode certification. MedlinePlus was the only website that did not have a HONcode seal. The HONcode website confirmed that MedlinePlus has not been certified since 2017.

The results show that the first page of Google search results was populated most frequently by the same five websites. Although WebMD, Healthline, and Medical News Today are presented as individual websites, both Healthline and Medical News Today are owned by Healthline Media Group. MedicineNet, which was the seventh ranked website, is owned and operated by WebMD. Thus, four of the top seven websites are operated by Healthline Media Group and WebMD. Overall, 33% of the total search results are generated by websites produced by two commercial organizations.
Based on domain type, commercial websites with .com suffixes were the most common to appear in the search results. Organization websites (.org) were the next most common, with government agency websites (.gov) being the least commonly found.

[Table 3 near here]

**DISCUSSION**

With Americans seeking information on COVID-19 during the pandemic, many turned to Google as their primary resource for health information. Using DISCERN and CRAAP to evaluate the quality of top-ranking consumer health websites on Google, the authors found that MedlinePlus is the most reliable and trustworthy website. Although this study focused solely on Google search results, the top five websites also appeared multiple times as advertisements, which Google indicated with a bold badge labeled “Ad” next to the website title. Featured news articles, image search options, and Google Knowledge Graphs also appeared as additional sources of health information.

Google Knowledge Graphs are featured boxes located on the right of the results page, offering illustrations and brief summaries on the topic being searched. The medical information in Google Knowledge Graphs is sourced from government agencies, regional hospitals, and other medical professionals. Sixty-one of the 65 terms searched displayed a Google Knowledge Graph. Any websites with the “Ad” badge next to the title, news articles, or links provided by Google Knowledge Graphs were not included in the final counts. However, consumers may still click on these links and would be advised to be critical of the information presented from these links.

Both researchers who reviewed each website are health sciences librarians. All five websites were evaluated based on the terminology and criteria that are explicitly written in both
DISCERN and CRAAP instruments. Interrater reliability of the two researchers was assessed by calculating the intraclass correlation between their final scores of each website, with a two-way mixed model and testing absolute agreement, using SPSS. Inter-rater reliability was excellent, with an intraclass correlation coefficient (ICC) of 0.96 (95% CIs: 0.86-0.99).

When evaluating information found on Google, consumers may also be encouraged to use the evaluation tools used in this study. Individual scores and quality ratings may vary compared to those of the reviewers of this study, but the evaluation criteria and questions being asked in both DISCERN and CRAAP instruments are the most important aspects to consider when evaluating health information. Both evaluation tools are meant to guide users to assess who is producing the website, when it was last updated, the purpose of the website, and who is the targeted audience. Health and medical librarians have offered similar guidance on how to evaluate online health information and have offered guidelines and recommendations supported by their professional associations and special interest groups.34-36

Four of the five top-ranking websites displayed the HONcode seal. If users were to solely rely on a HONcode seal to consider a website trustworthy, they might pass on reliable websites such as MedlinePlus, which is currently not certified. Like MedlinePlus, most websites found on Google are likely not to be certified by the HON foundation.24

CONCLUSION
This study aimed to evaluate the quality of top-ranking health websites in Google’s search results using the DISCERN instrument and CRAAP (Currency, Relevance, Authority, Accuracy, and Purpose) test. The results of the study revealed that MedlinePlus was the most reliable and trustworthy website based on its DISCERN and CRAAP scores. Although MedlinePlus was the
highest scoring website, three other websites appeared more frequently in Google’s search results. Health information professionals and those in the healthcare field may recommend MedlinePlus as the most reliable consumer health information website that frequently appears on Google. However, Google search results are likely to change in the future, and MedlinePlus may not appear in the first page of results.

One of the limitations of this study is that consumers may retrieve different search results depending on their location, browsing history, and privacy settings on their browsers. Previous research has also shown that the most frequently appearing health information websites on Google have shifted and will likely continue to evolve as websites improve their search optimization strategies, companies merge, or websites shutter over time.\(^{37-38}\) The shift in search results over time may also have led to the two commercial organizations identified generating 1/3 of the top search results found on Google. Organizations such as the HON foundation and the Medical Library Association are attempting to keep consumers informed of the most reliable health information websites.\(^{39}\) Health science librarians and health information professionals should continue to pursue efforts to guide consumers to evaluate their sources and consider the relevance and reliability of websites.

**Funding:** The authors reported there is no funding associated with the work reported in this article.

**REFERENCES**


Table 1. How often each website appeared in the Google search results

<table>
<thead>
<tr>
<th>Rank</th>
<th>Website Name</th>
<th>West Coast</th>
<th>Midwest</th>
<th>East Coast</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>WebMD</td>
<td>72</td>
<td>73</td>
<td>71</td>
<td>216</td>
</tr>
<tr>
<td>2</td>
<td>Mayo Clinic</td>
<td>65</td>
<td>70</td>
<td>67</td>
<td>202</td>
</tr>
<tr>
<td>3</td>
<td>Healthline</td>
<td>68</td>
<td>64</td>
<td>63</td>
<td>195</td>
</tr>
<tr>
<td>4</td>
<td>MedlinePlus</td>
<td>58</td>
<td>61</td>
<td>64</td>
<td>183</td>
</tr>
<tr>
<td>5</td>
<td>Medical News Today</td>
<td>48</td>
<td>49</td>
<td>47</td>
<td>144</td>
</tr>
<tr>
<td>6</td>
<td>Wikipedia</td>
<td>35</td>
<td>35</td>
<td>34</td>
<td>104</td>
</tr>
<tr>
<td>7</td>
<td>MedicineNet</td>
<td>38</td>
<td>30</td>
<td>33</td>
<td>101</td>
</tr>
<tr>
<td>8</td>
<td>Cleveland Clinic</td>
<td>27</td>
<td>27</td>
<td>25</td>
<td>79</td>
</tr>
<tr>
<td>9</td>
<td>CDC</td>
<td>25</td>
<td>20</td>
<td>20</td>
<td>65</td>
</tr>
<tr>
<td>10</td>
<td>Harvard Health</td>
<td>15</td>
<td>14</td>
<td>14</td>
<td>43</td>
</tr>
</tbody>
</table>

Table 2. Evaluation results of top-ranking Google websites

<table>
<thead>
<tr>
<th>Rank</th>
<th>Website Name</th>
<th>DISCERN</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DISCERN</td>
<td>Quality</td>
<td>DISCERN</td>
<td>Quality</td>
</tr>
<tr>
<td>1</td>
<td>MedlinePlus</td>
<td>67.5</td>
<td>Excellent</td>
<td>21</td>
<td>Excellent</td>
</tr>
<tr>
<td>2</td>
<td>Mayo Clinic</td>
<td>57.5</td>
<td>Good</td>
<td>18.5</td>
<td>Good</td>
</tr>
<tr>
<td>3</td>
<td>WebMD</td>
<td>53</td>
<td>Good</td>
<td>18</td>
<td>Good</td>
</tr>
<tr>
<td>4</td>
<td>Healthline</td>
<td>51.5</td>
<td>Good</td>
<td>19</td>
<td>Good</td>
</tr>
<tr>
<td>5</td>
<td>Medical News Today</td>
<td>44.5</td>
<td>Fair</td>
<td>14.5</td>
<td>Good</td>
</tr>
</tbody>
</table>
Table 3. Most common websites by domain that appear on Google

<table>
<thead>
<tr>
<th>Domain Type</th>
<th>Most Common Websites by Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial (.com)</td>
<td>791</td>
</tr>
<tr>
<td>Organization (.org)</td>
<td>688</td>
</tr>
<tr>
<td>Government (.gov)</td>
<td>370</td>
</tr>
<tr>
<td>Other</td>
<td>101</td>
</tr>
</tbody>
</table>
Appendix: List of search terms selected from Google COVID-19 Search Trends symptoms dataset

Abdominal pain
Allergy
Anemia
Angioedema
Anxiety
Attention Deficit Disorder
Autoimmune disease
Bradycardia
Bronchitis
Cyst
Depression
Diabetes
Diarrhea
Dizziness
Dysphagia
Dysuria
Eczema
Edema
Fatigue
Fever
Fibromyalgia
Gout
Headache
Heart arrhythmia
Heartburn
Hemorrhoid
Hepatitis A
Hives
Hyperglycemia
Hypertension
Hyperthyroidism
Indigestion
Inflammation
Insomnia
Jaundice
Kidney failure
Lactose intolerance
Laryngitis
Lightheadedness
Lymph
Melena
Meningitis
Menorrhagia
Migraine
Muscle weakness
Myocardial infarction
Nasal congestion
Nausea
Obesity
Panic attack
Pneumonia
Prediabetes
Pulmonary edema
Rhinitis
Shallow breathing
Sleep apnea
Sleep disorder
Sore throat
Stroke
Tinnitus
Tonsillitis
Tumor
Upper respiratory tract infection
Vertigo
Vomiting