Quality and Readability of English-Language Internet Information for Voice Disorders

Abigail J. Dueppen, Monica L. Bellon-Harn, Nandhakumar Radhakrishnan, and Vinaya Manchaiah

Summary: Purpose. The purpose of this study is to evaluate the readability and quality of English-language Internet information related to vocal hygiene, vocal health, and prevention of voice disorders. This study extends recent work because it evaluates readability, content quality, and website origin across broader search criteria than previous studies evaluating online voice material.

Method. Eighty-five websites were aggregated using five different country-specific search engines. Websites were then analyzed using quality and readability assessments. The entire web page was evaluated; however, no information or links beyond the first page was reviewed. Statistical calculations were employed to examine website ratings, differences between website origin and quality and readability scores, and correlations between readability instruments.

Result. Websites exhibited acceptable quality as measured by the DISCERN. However, only one website obtained the Health On the Net certification. Significant differences in quality were found among website origin, with government websites receiving higher quality ratings. Approximate educational levels required to comprehend information on the websites ranged from 8 to 9 years of education. Significant differences were found between website origin and readability measures with higher levels of education required to understand information on websites of nonprofit organizations.

Conclusion. Current vocal hygiene, vocal health, and prevention of voice disorders websites were found to exhibit acceptable levels of quality and readability. However, highly rated Internet information related to voice care should be made more accessible to voice clients through Health On the Net certification.

Key Words: Internet-based health information–Vocal hygiene–Vocal health–Voice disorders–Accessibility.

Internet-based health information stands as the second form of sought after resources following direct clinician advice. The percentage of Americans using online resources for health-related topics has increased from 62% to almost 80% in the last 15 years. Apart from consumers and patients, physicians and medical students access Internet-based health-care information. This trend has changed patients’ approach to health care and has increased their involvement to make informed decisions regarding intervention.

However, with the abundance of health information available online, patients have expressed concern over conflicting details reported about their condition. Further, health-care information reportedly carries material that is not easy to read and understand. The quality of content provided across many health-care areas is also under question. In a recent report, Armstrong-Heimsoth et al reported a discrepancy in the number of Americans accessing information related to health versus their ability to understand it.

Internet-based health information is frequently evaluated on readability, content quality, and website origin. According to the National Adult Literacy Survey, the average reading grade level for American adults has been reported to be around seventh grade. Recommendations from experts in the field of health literacy mandate that information available to the public be written at about a fifth or sixth reading grade level. Many readability measures exist; however, there is no standard for choosing readability formulas.

Quality of content is a critical component of website evaluation. Although there are various criteria for quality assessment, two methods increasing in popularity for health-care information include the Health On the Net (HON) certification and results from the DISCERN instrument. HON is a nongovernmental organization that promotes and guides useful and reliable online health information. HON certification is an indication that the web developers adhere to recommended guidelines for quality information and formatting on a website. Unfortunately, few web developers are aware of it. The DISCERN instrument rates websites on various quality indicators and provides a quantitative method to assess Internet-based health-care information. DISCERN evaluates the quality of written information about treatment choices, and facilitates the production of high quality evidence-based patient information. The instrument provides a method to judge publication reliability, quality of treatment information, and an overall quality rating.

Finally, website origin is evaluated because it may have some bearing toward the quality, accuracy, and readability of information provided. For example, websites from commercial, university, and government origins may have varying levels of quality or readability.
Accessibility of information related to communication disorders

Although Internet-based health information has been evaluated in the past two decades, only recently has work examining the accessibility of information in communication disorders been conducted. In a systematic review of the literature, Laplante-Lévesque and Thorén reported poor readability of Internet information on hearing and hearing loss.22 Manchaiah et al evaluated the quality and readability of 134 websites with information for tinnitus.21 Only 13.5% of websites had HON certification. DISCERN scores indicated low quality and high readability levels. An association between website origin and HON certification was noted. Azios et al recently evaluated quality and readability of 43 aphasia treatment websites.24 Websites exhibited low quality, with few websites obtaining HON certification or clear, thorough information as measured by DISCERN. Reported readability scores were also poor. Significant differences were found between website origin and readability measures, with higher levels of education required to understand information on websites of nonprofit organizations.

Recent work has examined accessibility of Internet-based health information related to voice disorders. Eloy et al reviewed Interned-based patient education materials from 262 websites linked to major associations related to otolaryngology.25 Many patients of otolaryngologists receive information on the care of the voice in office with pamphlets. Topics typically cover how to take care of the voice pre- or post surgery, management of reflux, or general vocal hygiene principles. Based on 10 different scales on readability, information reported exceeded the recommended reading level. Similarly, Ting and Hu assessed the quality and readability of Internet-based information related to thyroplasty.26 They analyzed the top 50 websites using the DISCERN instrument to assess quality, Flesch Reading Ease Score, and the Flesch–Kincaid Grade Level score to determine readability. Further analysis was performed by grouping the websites as major or minor, based on the sources used, and professional-targeted or patient-oriented. The results indicated that the information presented was suboptimal in quality and at a higher level of reading. Although the major website group had better quality scores, they were more difficult to read. The readability scores of patient-oriented websites indicated lower levels of reading ability were required to read the material than professional-targeted websites. However, DISCERN scores were lower for patient-oriented sites, albeit not statistically different from professional sites. This discrepancy could indicate that information directed to patients may be easier to read but not of good quality. Likewise, Balakrishnan et al analyzed the top 50 websites carrying patient education materials related to vocal fold paralysis for readability and understandability.27 The researchers attempted to show the difference in readability and understandability because of the fact that medical terminologies may have varied syllable lengths, which could skew the results of readability scores. The researchers employed six readability tests that use formulas involving sentence length, number of words, and number of syllables (ie, Flesch–Kincaid Reading Ease, Flesch–Kincaid Grade Level, Gunning-Fog Score, Coleman–Liau Index, Simple Measure of Gobbledygook [SMOG] Index, and Automated Readability Index). Patient Education Material Assessment Tool (PEMAT), a tool developed by the Agency for Healthcare Research and Quality, was used to assess readability and actionability.28 PEMAT is used to determine understandability of online articles by evaluating the text based on layout, content, length of the material, and inclusion of any visual aids. Balakrishnan et al reported a strong negative correlation between readability and understandability in that lower readability grade levels were related to higher understandability scores.27 Moreover, materials bearing the same readability level had a wide range of PEMAT scores, which suggests that readability alone does not qualify to rank patient education materials.

Study aim

Recent research indicates variability in the quality and readability of health information on the Internet. Clearly, there is the need for continued evaluation so that stakeholders (eg, web developers, clinicians) are aware and develop accessible Internet-based information for all people. The purpose of this study is to evaluate the readability and quality of English-language Internet information related to vocal hygiene, vocal health, and prevention of voice disorders. This study extends recent work because it evaluates readability, content quality, and website origin across broader search criteria than previous studies evaluating online voice material.

METHOD

Search strategy

There is no ethical approval required for this study. Using the search strategy provided by Laplante-Lévesque and colleagues, a panel of 18 speech-language pathology graduate students and six faculty in speech-language pathology were asked to provide key words that might be used when searching for information related to vocal health (see Appendix).29 The panel had completed coursework and clinical training related to voice and voice disorders. They provided 31 unique key words they considered to be most likely used when searching for information on the Internet. Identified words were entered in Google Trends (www.google.com/trends), which compiles relative frequency of key words in the search engine over time. Per Laplante-Lévesque and colleagues, Google customizes its search engine for countries across the world to provide more meaningful search results.29 Google automatically identifies Internet protocol addresses and redirects the user to its country-specific search engine. This means that the results of Google searches vary from one country to the other. The three most frequent key words (vocal health, vocal hygiene, and how to take care of voice) were identified and entered in five country-specific Google search engines (Australia, Canada, India, United Kingdom, and United States) along with the terms. Fifteen searches were completed (ie, three key words for each of five country’s search engines). For each search, the first 20 websites that met inclusion criteria were included in the analysis because they were deemed to be the most widely used. Websites repeated across each country’s search engine were analyzed only once.
Inclusion and exclusion criteria
Websites were included if they provided information directed to voice and voice disorders. Websites that Google identified as advertisements, news, images, and videos were excluded. Websites that linked to scholarly peer-reviewed journal articles or books only were excluded. Websites with only navigation links and no substantive content, description, or explanation were not included. Website origin was classified as commercial, non-profit organization, or government.

Quality assessment
Website quality is related to the degree a site can be effectively used by a targeted group of users for a specific purpose. Two measures used in the quality assessment of the information in the websites included HON certification and DISCERN. HON certification was conducted by using the website http://www.hon.ch/ to inquire if a target voice treatment website was listed.

The first and third author rated each website using the 16-item DISCERN instrument for quality rating. Raters were clinical experts in voice and voice disorders who were trained by individuals proficient with the DISCERN instrument. Following guided practice and consensus on the procedure, the raters independently rated the websites. Rating of each item was performed on a scale from one to five using the following criteria. A rating of one was achieved if the answer to the item was a definite no and the quality criterion had not been met. A rating of two to four was achieved if some part of the item was answered and the quality criterion had been partially met. A rating of five was achieved if the answer to the item was a definite yes and the quality criterion had been completely met. See Table 1 for a list the items rated in the DISCERN instrument.

Readability assessment
Readability of the website text was assessed using Oleander Software (http://www.oleandersolutions.com/). Many readability measures exist and there is no standard for choosing readability formulas. In this study, Flesch Reading Ease (FRE), Flesch-Kincaid Grade Level Formula (F-KGL), and the SMOG were used because these represent three of the most widely used readability measures. The FRE measures readability of text written between grade 5 and college-graduate level. The FRE has withstood the test of time and has been highly correlated with other readability measures. Scores between 60 and 70 are considered acceptable. Scores lower than 60 represent higher reading levels and higher scores represent lower reading levels. The F-KGL ranks text readability based on US grade school level. The availability of F-KGL as a Microsoft Word tool makes it a convenient option for health-care professionals. The SMOG estimates years of education a person needs to understand writings and represents one of the most exacting readability measures available.

Data analysis
Data were analyzed using IBM SPSS Statistics version 22 (Armonk, NY). In the first instance, descriptive statistics were explored. t Test was used to explore the difference in DISCERN and readability measure scores based on website origin. The inter-rater agreement for DISCERN was tested by interclass correlation coefficient. Pearson correlation was performed to study the association between different readability measures. Assumptions of normality were tested where relevant, and data were transformed to achieve normality if necessary. An alpha level of 0.05 was used to determine significance for all statistical analyses, and Bonferroni correction was applied for multiple comparisons.

<table>
<thead>
<tr>
<th>TABLE 1. DISCERN Quality Criteria for Health Information on Treatment Choices (Items, Mean, and Standard Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items</td>
</tr>
<tr>
<td>1. Are the aims clear?</td>
</tr>
<tr>
<td>2. Does it achieve its aims?</td>
</tr>
<tr>
<td>3. Is it relevant?</td>
</tr>
<tr>
<td>4. Is it clear what sources of information were used to compile the publication (other than author or procedure)?</td>
</tr>
<tr>
<td>5. Is it clear when the information used or reported in the publication was reported?</td>
</tr>
<tr>
<td>6. Does it balanced and unbiased?</td>
</tr>
<tr>
<td>7. Does it provide details of additional sources of support and information?</td>
</tr>
<tr>
<td>8. Does it refer to areas of uncertainty?</td>
</tr>
<tr>
<td>9. Does it describe how each treatment works?</td>
</tr>
<tr>
<td>10. Does it describe the benefits of each treatment?</td>
</tr>
<tr>
<td>11. Does it describe the risks of each treatment?</td>
</tr>
<tr>
<td>12. Does it describe what would happen if no treatment is used?</td>
</tr>
<tr>
<td>13. Does it describe how the treatment choices affect overall quality of life?</td>
</tr>
<tr>
<td>14. Is it clear that there may be more than one possible treatment choice?</td>
</tr>
<tr>
<td>15. Does it provide support for shared decision-making?</td>
</tr>
<tr>
<td>16. Based on the answers to all of the above questions, rate the overall quality of the publication as a source of information about treatment choices.</td>
</tr>
</tbody>
</table>
RESULTS

Origin
The search resulted in 85 websites after applying the inclusion and exclusion criteria. These websites were subjected to quality and readability assessments. Of the 85 websites, 31 (36.5%) were from nonprofit organizations, 47 (55.3%) were of commercial origin, and seven (8.2%) were of government.

Quality
Health On Net (HON) certification
Only 1 of 85 websites (1.2%) had obtained HON certification, which was of commercial origin.

DISCERN scores
The scores for each DISCERN item ranged from 1 to 5, with higher scores indicating better quality. Items 1–15 provide information about a specific aspect of quality, whereas item 16 is an overall quality measure based on scores from the initial 15 items. For all items, the low rating (ie, 1) indicates serious or extensive shortcomings, moderate scores (ie, 2–4) indicate potentially important but no serious shortcomings, and high ratings (ie, 5) indicate minimal shortcomings. The DISCERN scores ranged from mean scores of 2.45–4.96. Table 1 provides details of DISCERN scores for each item. Note that the total DISCERN scores were not normally distributed and presented a negative skew. Hence, logarithmic transformation of data was performed before the statistical analysis. However, the original DISCERN scores have been reported for the ease of interpretation.

Quality criterion was met at partial levels on all items, which indicated that websites had potentially important but no serious shortcomings. The mean score of item 16 revealed that the overall quality of websites as a source of information about treatment options was acceptable. Items 1 (ie, Are the aims clear?), 9 (ie, Does it describe how each treatment works?), 10 (ie, Does it describe the benefits of each treatment?), 14 (ie, Is it clear that there may be more than one possible treatment choice?), 15 (ie, Does it provide support for shared decision-making?) had the highest DISCERN scores, whereas items 5 (ie, Is it clear when the information used or reported in the publication was reported?) and 12 (ie, Does it describe the risks of each treatment?) had the lowest DISCERN scores.

Figure 1 provides mean and standard deviation of DISCERN scores for voice websites based on its origin. The t test results showed there is significant difference in DISCERN scores between websites of nonprofit and commercial organizations \([t(76) = 2.75, P = 0.007]\) and commercial and government organizations \([t(52) = −3.65, P = 0.001]\). However, no significant differences were found between websites of nonprofit and government organizations \([t(36) = −1.46, P = 0.15]\).

The inter-rater agreement for DISCERN was high as indicated by the interclass correlation coefficient of 0.95. This was comparable with earlier reports, which have reported interclass correlation coefficient of 0.82 and 0.88.29,35

Readability
Three readability measures were assessed for each of the websites (ie, FRE, F-KGL, SMOG). Table 2 provides the mean and standard deviation of readability scores for each of these domains in terms of website origin. The FRE average score was 61.53 (range between 17 and 81). The results suggest that average people required at least 8–9 years of education to read and understand the information on those websites, although some websites required higher reading levels (ie, 10th–12th grade). t Test results suggested that the FRE score varied significantly between websites of nonprofit organizations and commercial origin \([t(76) = −2.76, P = 0.01]\), but did not vary between websites of nonprofit organizations and government origin \([t(36) = −0.95, P = 0.35]\) and also between commercial and government \([t(52) = 0.65, P = 0.52]\). However, interpreting the t test results with Bonferroni corrected significance for multiple comparisons (ie, 9 comparisons and 0.05 alpha indicating a corrected alpha level of 0.005) indicated no difference in reading level required for websites of different origin.

Mean F-KGL indicated that websites were written at approximately an eighth grade level and did not vary significantly between websites of nonprofit organizations and commercial

![FIGURE 1. Mean DISCERN scores according to website origin. Error bars indicate standard deviation.](image-url)
In that line of thought, although online mate-

Mean SMOG score indicated that websites were written at a level for individuals with between 10 and 11 years of education. Scores did not vary significantly between websites of nonprofit organizations and commercial origin \( t(36) = 0.79, P = 0.43 \), and also between commercial and government \( t(52) = -0.08, P = 0.93 \) origin websites.

Moreover, the three readability measures were highly correlated. The F-KGL and SMOG were positively correlated (Pearson correlation = 0.96, \( P < 0.01 \)), whereas negative but statistically significant correlations were observed between F-KGL and FRE score (Pearson correlation = -0.91, \( P < 0.01 \), and also between FRE score and SMOG (Pearson correlation = -0.91, \( P < 0.01 \)).

### DISCUSSION

This study assessed readability, content quality, and website origin of English-language Internet information available for individuals seeking treatment information related to vocal hygiene, vocal health, and prevention of voice disorders. Access to Internet-based health information facilitates independent application of vocal hygiene and vocal health recommendations. Of the 85 websites, 31 were from nonprofit organizations, 47 were of commercial origin, and 7 were of government origin. This study is inconsistent with other research that has demonstrated that a large portion of Internet health-care information is at reading levels that are too high for the majority of the population. As noted, the recommended average reading level at which websites should be written is fifth or sixth grade reading level.\(^{14,15} \)

In this study, FRE results indicate generally acceptable reading levels. A difference was identified between nonprofit and commercial websites, suggesting that some nonprofit websites may be more difficult to read. Results of F-KGL and SMOG analysis suggest that an eighth grade reading level and a minimum of 10–11 years of education is required to read and understand the information on websites dedicated to voice treatment. This did not differ based on website origin. Reading levels on some websites are slightly higher than the recommended levels as well as the average US adult reading level of seventh to eighth grade.\(^{36} \)

In sum, although websites are written at readability levels that are generally accessible, some may include material that is more difficult to read than others.

<table>
<thead>
<tr>
<th>Website Origin Mean (SD)</th>
<th>All Websites</th>
<th>Nonprofit</th>
<th>Commercial</th>
<th>Government</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flesch Reading Ease Score</strong></td>
<td>61.53 (10.25)</td>
<td>57.81 (9.06)</td>
<td>64.02 (10.76)</td>
<td>61.29 (7.25)</td>
</tr>
<tr>
<td><strong>Flesch-Kincaid Grade Level Formula</strong></td>
<td>8.56 (1.88)</td>
<td>8.94 (2.69)</td>
<td>8.33 (2.07)</td>
<td>8.40 (1.32)</td>
</tr>
<tr>
<td><strong>Simple Measure of Gobbledygook (SMOG)</strong></td>
<td>10.92 (1.46)</td>
<td>11.27 (1.34)</td>
<td>10.72 (1.56)</td>
<td>10.69 (0.84)</td>
</tr>
</tbody>
</table>

Regardless of origin, maintaining adequate readability levels across treatment websites is a priority. The National Institutes of Health (www.nlm.nih.gov/medlineplus/etr.html) recommends strategies for writing easy-to-read health materials. In this study, the websites that included bulleted points with general, easily applicable recommendations that would pertain to many professional voice users and to those seeking vocal advice and treatment may be the most accessible to the largest group of people. Health professionals and voice patients alike may want access to this information on the internet. As discussed earlier, Ting and Hu suggested that there is a discrepancy of readability between patient-oriented and professional-oriented websites that could indicate that information directed to patients may be easier to read, but not of good quality.\(^{36} \)

Additionally, Balakrishnan et al suggested that readability alone does not qualify to rank patient education materials.\(^{27} \)

With regard to quality, only 1.2% of the websites had HON certification. As noted, website developers should be aware of this type of website certification to ensure a standard level of quality. Results of this study are consistent with previous research in that a small percentage of websites have this certification. However, DISCERN scores indicate that most of the websites have acceptable levels of quality, which is inconsistent with other reports in communication disorders.\(^{24,26} \)

Although online material evaluated in this study included aims that were met, the depth and scope of information may not be sufficient. Most of the information shared was common knowledge (eg, increase water intake, avoid drying agents, rest the voice when it feels tired). Further, many of the websites in this study did not describe any negative effects or risks of treatment nor did the websites site their sources and dates of publication. In sum, although content quality is acceptable, improvements can be made.

Laplante-Lévesque and colleagues listed websites dedicated to hearing information that could be used as models for improving existing or developing websites.\(^{29} \)

In that line of thought, voice treatment websites that had the highest levels of quality for content and readability could be used as models. See Table 3 for the highest ranked websites for content quality and readability. Each website received either a 4 or a 5 in overall quality from both raters. The F-KGL for each website is also included for readability.

### Clinical implications

Highly rated information related to voice care on the internet currently available should be made more accessible to voice clients.
through HON certification. Those who work with voice clients (otolaryngologists, speech-language pathologists, etc) can use the material in this study to understand their new client’s informational basis on their first visit. This should create an opportunity to highlight the best principles of vocal hygiene and overall care of the voice. Also, these professionals can heighten their client’s awareness of certain websites and what the client should look for in the overall quality of vocal health information to navigate all that is currently available on the internet.

Limitations and future directions
Presentation, layout, illustrations, and photographs within web pages contribute to website quality and readability. The analysis used in this study did not factor variables other than the text. Hasan and Abuelrub suggest organization elements and ease of use should be considered in website quality evaluation in addition to content and design. Further, the DISCERN tool may not have fully captured the quality of the websites. Other tools should be used and evaluated to determine whether DISCERN is correlated with other measures.

Another limitation was that individuals with voice disorders were not included in website evaluation. Future studies should include their perceptions of website materials so that materials are designed with the end-user in mind. This is critical because evidence-based practice ensures that we are considering the desires and needs of the individual experiencing challenges. Because the way we access information is constantly changing (eg, via mobile devices, tablets, watches), there is a need to examine different types of device access and the potential impact it has on information accessibility. This study examined websites only. Future studies should include other digital technologies.

Acknowledgment
This work was supported in whole or in part by a grant from the Lamar University Visionary Grant Initiative.

APPENDIX

Search Strategy
A panel of six faculty in speech-language pathology and 18 speech-language pathology graduate students were asked to provide key words that might be used when searching for information related to vocal health.

<table>
<thead>
<tr>
<th>Key words survey question:</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are three (3) search terms you would use in a search engine, such as Google, if you were looking for more information on vocal health?</td>
</tr>
</tbody>
</table>

The key words the faculty and students identified are listed below. The key words most identified by both students and faculty are in italics:

- How to keep VF healthy
- Tips on vocal exercises/Vocal workouts
- How to prevent vocal nodules
- Vocal hygiene
- Vocal health
- How to improve vocal health
- Normal voice quality
- Vocal norms
- What should I do for my voice
- How to keep voice clear
- Voice remedy
- How to maintain vocal health
- Normal causes of vocal problems
- Problems with vocal health
- What’s wrong with my voice
- How to take care of voice
- What do healthy VF look like
- Breathing exercises for vocal tension
- Texas Voice Center
- ASHA Vocal Health
- How to not have a raspy voice
- Bad ways to talk that can hurt my voice
- Are my VF normal
- VF and health problems
- Vocal wellness
- Good vocal habits
- Voice problems
- How to maintain a strong/healthy voice
- Prevention voice
- Healthy voice
- Vocal care

The three italicized key words were entered into five country-specific Google search engines. These countries are:

- United States of America
- United Kingdom
- India
Australia
Canada
Fifteen separate searches were completed on April 22, 2017.
Key words vocal health in Google United States of America
Key words vocal hygiene in Google United States of America
Key words how to take care of voice in Google United States of America
Key words vocal health in Google United Kingdom
Key words vocal hygiene in Google United Kingdom
Key words how to take care of voice in Google United Kingdom
Key words vocal health in Google India
Key words vocal hygiene in Google India
Key words how to take care of voice in Google India
Key words vocal health in Google Australia
Key words vocal hygiene in Google Australia
Key words how to take care of voice in Google Australia
Key words vocal health in Google Canada
Key words vocal hygiene in Google Canada
Key words how to take care of voice in Google Canada

REFERENCES