Readability and Understandability of Online Vocal Cord Paralysis Materials

Vini Balakrishnan, MS¹, Zachariah Chandy¹, Amy Hseih¹, Thanh-Lan Bui, MA¹, and Sunil P. Verma, MD¹

No sponsorships or competing interests have been disclosed for this article.

Abstract

Objective. Patients use several online resources to learn about vocal cord paralysis (VCP). The objective of this study was to assess the readability and understandability of online VCP patient education materials (PEMs), with readability assessments and the Patient Education Materials Evaluation Tool (PEMAT), respectively. The relationship between readability and understandability was then analyzed.

Study Design and Setting. Descriptive and correlational design.

Method. Online PEMs were identified by performing a Google search with the term “vocal cord paralysis.” After scientific webpages, news articles, and information for medical professionals were excluded, 29 articles from the first 50 search results were considered. Readability analysis was performed with 6 formulas. Four individuals with different educational backgrounds conducted understandability analysis with the PEMAT. Fleiss’s Kappa interrater reliability analysis determined consistency among raters. Correlation between readability and understandability was determined with Pearson’s correlation test.

Results. The reading level of the reviewed articles ranged from grades 9 to 17. Understandability ranged from 29% to 82%. Correlation analysis demonstrated a strong negative correlation between materials’ readability and understandability ($r = -0.462, P < .05$).

Conclusion. Online PEMs pertaining to VCP are written above the recommended reading levels. Overall, materials written at lower grade levels are more understandable. However, articles of identical grade levels had varying levels of understandability. The PEMAT may provide a more critical evaluation of the quality of a PEM when compared with readability formulas. Both readability and understandability should be used to evaluate PEMs.

Keywords
readability, PEMAT, understandability, vocal cord paralysis, patient education materials, readability formula, Internet, patient education
that contribute to the ease of reading and comprehen-
sion.10,15-19 In view of these limitations, other tools have 
been proposed to evaluate PEMs.20-26 One such tool is the 
Patient Education Materials Assessment Tool (PEMAT), 
which was recently developed by the Agency for Healthcare 
Research and Quality.27 The PEMAT evaluates and 
compares the understandability and actionability of written 
materials as a whole and is designed to be used by layper-
sons and health professionals alike.27

The objective of this study was to evaluate the quality of 
online VCP PEMs using both readability assessments and 
the PEMAT. The primary hypothesis was that the readabil-
ity of online materials related to VCP would be significantly 
higher than the recommended guidelines established by the 
AMA and the NIH. Another hypothesis was that materials 
of lower reading grade levels would be more understandable 
than those of higher readability grade levels. To evaluate 
these 2 hypotheses, the readability of each PEM was com-
pared and correlated to its understandability.

Materials and Methods

The Google search engine was used to perform an Internet 
search for PEMs in January 2015. The term “vocal cord 
paralysis” was used to perform the search. The first 50 
search results obtained were analyzed. News articles, per-
sonal experiences, celebrity experiences, scientific web-
pages (ScienceDirect, PubMed), and webpages targeted 
toward medical professionals were excluded from analysis. 
Finally, 29 pages had material directed toward patients and 
were included in the study.

Text from each article was copied into a separate plain-
text document. Text deemed not relevant to patient educa-
tion was deleted, including copyright notices, disclaimers, 
acknowledgements, and references. The readability of each 
article was determined with an online readability calculator 
(http://readability-score.com), which performed the follow-
ning readability tests: Flesch-Kincaid Reading Ease, Flesch-
Kincaid Grade level, Gunning-Fog Score, Coleman-Liau 
Index, SMOG Index, and Automated Readability Index. These 
tests use formulas with variables such as sentence length, 
number of words, and number of syllables to estimate 
the average reading grade level required to comprehend the 
text.10-12 The Flesch-Kincaid Reading Ease test used a formula 
to score text between 0 and 100, with a higher score indicating 
that the material was easier to read. The score obtained 
was then converted to a reading level to allow comparison with 
other readability test scores. The various readability formulas 
differ considerably, which can lead to significant differences in 
the reported reading grade level of materials.28 To reduce the 
variations in scores reported by readability formulas, the mean 
of the reading level scores was used as the readability score.

To determine the understandability of each article, the 
PEMAT was used. Four individuals with differing levels of 
education and medical training scored each article following 
the PEMAT guidelines. The PEMAT evaluated text based on 
layout, content, length of material, and presence of visual 
 aids. Individual understandability items on the PEMAT were 
scored 0 (disagree), 1 (agree), or N/A (not applicable).27 For 
each article, the final score for understandability was calcu-
lated by adding the points for each item, dividing it by the 
total possible points, and multiplying it by 100. This gave a 
percentage value for understandability. Articles with higher 
scores were interpreted to be easier to understand than those 
with lower scores. Fleiss’ Kappa interrater reliability analysis 
was performed to determine the consistency among raters. 
Finally, the Pearson correlation coefficient between the read-
ability and understandability scores was calculated with SPSS 
22.0 for Macintosh (IBM Corp, Armonk, New York).

Results

Twenty-nine VCP articles were identified as PEMs from the 
first 50 Google search results. All PEMs were written well 
above the fourth- to sixth-grade reading level recommended 
by the NIH and the AMA. Table 1 demonstrates that the 
reading level for all PEMs, ranging from grades 9 to 17.

The mean understandability score for all PEMs combined 
was 53% (SD = 13.89; 95% confidence interval = 47.86-
57.97), with a range from 29% to 82%. As represented in 
Figure 1, 95% of articles displayed information in a clear, 
logical sequence, and 84% did not include information that 
distracted from the main content. However, only 39% used 
common everyday terms to convey information, and only 
40% of articles defined medical terms when they were used. 
Although 66% of articles grouped information into short 
sections, only 54% of articles had informative headers for 
the sections. None of the articles provided a summary at the 
end of the text. Only 21% of articles used illustrations to 
make written material more informative. Captions and titles 
that clearly described the visual aids were applied in only 2 
articles. Fleiss’ Kappa interrater reliability analysis deter-
mined that the reliability for the 4 raters was 0.61 (95% 
confidence interval = 0.4605-0.7611) suggesting substantial 
agreement among the 4 raters based on the standards pro-
posed by Landis and Koch.29

Finally, a strong negative correlation between readability 
and understandability scores was found (r = -0.462, P = 
.011; Figure 2).

Discussion

VCP is a common condition encountered in an otolaryngol-
yogy practice. Unfortunately, the multiple etiologies and 
treatment options of VCP may be difficult for patients to 
understand. As such, this study investigated the readability 
and understandability of online resources, which patients are 
likely to turn to for supplementary information on VCP.

The reading grade levels of the PEMs analyzed ranged 
from grades 9 to 17 which was well above the fourth- to 
sixth-grade reading levels recommended by the AMA and 
the NIH. This was not surprising, as readability assess-
ments performed by other authors showed similar results 
on such topics as thyroid diseases, parathyroid diseases, 
and acoustic neuromas.5,6,10-19

While readability is an important factor for analyzing 
PEMs, readability formulas have inherent flaws. For
### Table 1. Reading Grade Levels and PEMAT Scores of Patient Educational Materials Related to Vocal Cord Paralysis.

<table>
<thead>
<tr>
<th>Material</th>
<th>Grade Level</th>
<th>PEMAT Score, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical News Today</td>
<td>9.20</td>
<td>62.42</td>
</tr>
<tr>
<td>Cincinnati Hospital</td>
<td>9.36</td>
<td>74.58</td>
</tr>
<tr>
<td>MSKCC</td>
<td>9.12</td>
<td>67.33</td>
</tr>
<tr>
<td>Baylor College</td>
<td>8.78</td>
<td>81.74</td>
</tr>
<tr>
<td>Chicago Tribune</td>
<td>10.32</td>
<td>45.74</td>
</tr>
<tr>
<td>Free Dictionary</td>
<td>10.14</td>
<td>43.74</td>
</tr>
<tr>
<td>Osborne Head and Neck Institute</td>
<td>11.24</td>
<td>63.32</td>
</tr>
<tr>
<td>Mayo Clinic</td>
<td>11.38</td>
<td>43.83</td>
</tr>
<tr>
<td>Scientific American</td>
<td>12.46</td>
<td>44.66</td>
</tr>
<tr>
<td>Uroplasty</td>
<td>11.92</td>
<td>56.10</td>
</tr>
<tr>
<td>Cleveland Clinic</td>
<td>11.56</td>
<td>35.33</td>
</tr>
<tr>
<td>Virginia Mason</td>
<td>11.68</td>
<td>45.91</td>
</tr>
<tr>
<td>Mt Sinai Hospital</td>
<td>11.60</td>
<td>70.83</td>
</tr>
<tr>
<td>Fauquier ENT</td>
<td>11.70</td>
<td>70.33</td>
</tr>
<tr>
<td>UC Davis Voice</td>
<td>13.12</td>
<td>32.83</td>
</tr>
<tr>
<td>NIDCD</td>
<td>12.54</td>
<td>68.33</td>
</tr>
<tr>
<td>American Academy</td>
<td>13.44</td>
<td>58.25</td>
</tr>
<tr>
<td>American Speech and Language Association</td>
<td>12.76</td>
<td>56.33</td>
</tr>
<tr>
<td>NorthShore University</td>
<td>13.60</td>
<td>33.25</td>
</tr>
<tr>
<td>abouthealth.com</td>
<td>14.14</td>
<td>59.24</td>
</tr>
<tr>
<td>Emory Health Care</td>
<td>13.74</td>
<td>52.17</td>
</tr>
<tr>
<td>Wikipedia</td>
<td>15.42</td>
<td>29.25</td>
</tr>
<tr>
<td>Children's Hospital of Philadelphia</td>
<td>15.48</td>
<td>56.83</td>
</tr>
<tr>
<td>NYU Voice Center</td>
<td>15.38</td>
<td>35.33</td>
</tr>
<tr>
<td>UCI Medical Center</td>
<td>14.52</td>
<td>60.50</td>
</tr>
<tr>
<td>Medscape</td>
<td>16.04</td>
<td>41.58</td>
</tr>
<tr>
<td>Merck Manuals</td>
<td>16.04</td>
<td>48.83</td>
</tr>
<tr>
<td>Massachusetts General Hospital</td>
<td>15.52</td>
<td>53.83</td>
</tr>
<tr>
<td>Pacific ENT</td>
<td>17.12</td>
<td>42.13</td>
</tr>
</tbody>
</table>

Abbreviation: PEMAT, Patient Education Materials Assessment Tool.

### Figure 1. The number of articles that satisfy items on Patient Education Materials Assessment Tool.
example, difficult-to-read texts are defined as having longer sentences and words with >3 syllables, which may be misleading when one is evaluating medical literature. Medical terms are often multisyllabic and repeated throughout the course of text and thus may unnecessarily and significantly increase the reading level of a document. With these flaws, relying on a readability score alone may not be the most accurate method to assess the value of PEMs.

The PEMAT is an alternative method to evaluate the quality of PEMs. The PEMAT is a new instrument that systematically assesses the understandability of PEMs based on a variety of parameters, including clarity of purpose, simplicity of wording, organization, and use of visual aids. In comparison with readability formulas that assess PEMs based solely on sentence structure and word length, the PEMAT may provide a more critical evaluation of the quality of PEMs.

In this study, the mean understandability of the articles analyzed was 53%, with a range from 29% to 82%. The creators of the PEMAT defined understandable materials as a score ≥70% with the PEMAT questionnaire. In this study, the mean understandability score of VCP PEMs was well below the threshold. Only 4 of the 29 articles were considered understandable.

A benefit of the PEMAT is that it provides a framework for authors to improve the understandability of PEMs. For example, it suggests that articles should have an introduction that describes the purpose of the article and a summary of the key points at the end. Other strategies include using easily understandable language, active voice, and consistently defining medical terms when they are used. Finally, the tool emphasizes the use of illustrations with clear titles and captions.

The PEMAT has limitations as well. For example, 1 item on the PEMAT questionnaire evaluates the use of common everyday language in the article of interest. This is a subjective value and may be biased on the basis of the reader’s education level, native language, and medical background. This study attempted to overcome this limitation by having 4 individuals with different levels of education and medical background rate each article. The interrater reliability analysis demonstrated substantial agreement among the 4 raters.

Another limitation was that the PEMAT does not assess accuracy and completeness of the material, which was not accounted for in this study.

Correlation analysis demonstrated a strong negative correlation between readability and understandability. This indicated that lower readability grade levels were associated with higher understandability scores. However, further analysis revealed that articles of the same reading level had a wide range of PEMAT understandability scores. For example, Figure 2 demonstrates that 6 articles with a grade 12 reading level had PEMAT understandability scores ranging from 35% to 70%. The PEMAT may provide a more critical assessment of PEMs as compared with readability formulas. Findings from this study demonstrated that there is a need for more readable and understandable VCP PEMs.

This study has limitations, which are important to acknowledge. For example, the use of computer-based analysis of readability may be controversial because it has been shown to overestimate the difficulty level of materials. However, computer scoring—being faster and incorporating less personal bias than hand-scored the materials—is currently the standard of evaluation for readability in medical literature, as evidenced by its use in numerous articles.

**Conclusion**

Numerous resources on VCP are available online for patients. The reading grade levels of these materials are much higher than that recommended by the AMA and NIH. This study found a strong negative correlation between an article’s readability and understandability, demonstrating that materials written at lower reading levels were more understandable. However, articles of the same reading grade level had a wide range of understandability. Considering the limitations of readability formulas, authors should not rely on readability scores alone to determine if materials are appropriate for patient education. The PEMAT provides a more critical assessment of articles and should be used in conjunction with readability assessments to evaluate PEMs in the future. This may be especially important for diseases like VCP, with complex anatomy, etiologies, and treatment options.
Author Contributions

Vini Balakrishnan, conception and design, data collection, data analysis, drafting the manuscript, editing and revising the manuscript, final submission; Zachariah Chandy, data collection, data analysis, editing and revising the manuscript, final approval; Amy Hseih, data collection, editing the manuscript, final approval; Thanh-Lan Bui, data collection, editing the manuscript, final approval; Sunil P. Verma, oversight and management, conception and design, data analysis, revising and editing the manuscript, final approval.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

References