Readability Assessment of Patient Education Materials on Major Otolaryngology Association Websites

Jean Anderson Eloy, MD¹,²,³, Shawn Li, MD¹, Khushabu Kasabwala¹, Nitin Agarwal², David R. Hansberry, PhD², Soly Baredes, MD¹, and Michael Setzen, MD, FAAP⁴,⁵

Abstract

Objective. Various otolaryngology associations provide Internet-based patient education material (IPEM) to the general public. However, this information may be written above the fourth- to sixth-grade reading level recommended by the American Medical Association (AMA) and National Institutes of Health (NIH). The purpose of this study was to assess the readability of otolaryngology-related IPEMs on various otolaryngology association websites and to determine whether they are above the recommended reading level for patient education materials.

Study Design and Setting. Analysis of patient education materials from 9 major otolaryngology association websites.

Methods. The readability of 262 otolaryngology-related IPEMs was assessed with 8 numerical and 2 graphical readability tools. Averages were evaluated against national recommendations and between each source using analysis of variance (ANOVA) with post hoc Tukey’s honestly significant difference (HSD) analysis. Mean readability scores for each otolaryngology association website were compared.

Results. Mean website readability scores using Flesch Reading Ease test, Flesch-Kincaid Grade Level, Coleman-Liau Index, SMOG grading, Gunning Fog Index, New Dale-Chall Readability Formula, FORCAST Formula, New Fog Count Test, Raygor Readability Estimate, and the Fry Readability Graph ranged from 20.0 to 57.8, 9.7 to 17.1, 10.7 to 15.9, 11.6 to 18.2, 10.9 to 15.0, 8.6 to 16.0, 10.4 to 12.1, 8.5 to 11.8, 10.5 to 17.0, and 10.0 to 17.0, respectively. ANOVA results indicate a significant difference (P < .05) between the websites for each individual assessment.

Conclusion. The IPEMs found on all otolaryngology association websites exceed the recommended fourth- to sixth-grade reading level.

Keywords

readability, Internet, otolaryngology, patient education materials, Internet-based patient education material, comprehension, Flesch-Kincaid Grade Level, Flesch Reading Ease Score, otolaryngology association websites, readability scores

Received April 13, 2012; revised June 7, 2012; accepted July 6, 2012.

The Internet is a booming database that has revolutionized the distribution of health care information. Eighty percent of US adults search online for health-related information, making it the third most popular Internet activity behind checking e-mail and using a search engine.¹ In addition, as smartphone and mobile network technology continues to improve, more patients are able to find answers to their questions instantaneously.² Although this accessibility and abundance of resources has the potential to improve health care overall, the level of benefit is limited by the reader’s understanding of the information presented.

Many studies have supported a correlation between understanding and health outcome.³,⁴ In 1999, the American Medical Association (AMA) found that health literacy was a stronger predictor of a person’s health than age, income, education level, and race.⁶ Schillinger et al⁷ showed that inadequate health literacy is independently associated with worse glycemic control and higher rates of retinopathy. More recently, Bostock and

¹Department of Otolaryngology–Head & Neck Surgery, University of Medicine and Dentistry of New Jersey, New Jersey Medical School, Newark, New Jersey, USA
²Department of Neurological Surgery, University of Medicine and Dentistry of New Jersey, New Jersey Medical School, Newark, New Jersey, USA
³Center for Skull Base and Pituitary Surgery, University of Medicine and Dentistry of New Jersey–New Jersey Medical School, Newark, New Jersey, USA
⁴Rhinology Section, North Shore University Hospital, Manhasset, New York, USA
⁵Department of Otolaryngology, New York University School of Medicine, New York, New York, USA

Corresponding Author:
Jean Anderson Eloy, MD, Vice Chairman, Director of Rhinology and Sinus Surgery, Department of Otolaryngology–Head and Neck Surgery, UMDNJ–New Jersey Medical School, 90 Bergen St, Suite 8100, Newark, NJ 07103, USA
Email: Jean.Anderson.Eloy@gmail.com
Steepe found poorer understanding to be directly associated with increased mortality in the elderly population.

One approach to the literacy problem is to ensure proper readability of patient education materials. Readability is a measure of how easily a text is read and can be used as a tool to predict comprehension. The average reading ability in the United States is at the eighth-grade level, and the AMA and National Institutes of Health (NIH) recommend health care materials be written between a fourth- and sixth-grade level. Therefore, as medical organizations and societies embrace the online health care phenomenon by contributing verified education material for patients, the resources must be carefully composed to meet these qualifications. With the help of various established readability assessment tools, we previously analyzed the readability of patient education material found on the American Academy of Otolaryngology–Head and Neck Surgery (AAO-HNS) professional webpage and found its contents to be above the recommended reading level. In this study, we examine the readability of patient education materials on all major otolaryngology association websites.

**Materials and Methods**

In March 2012, we browsed through the websites of major otolaryngology associations and downloaded all available Internet-based patient education material (IPEM). These organizations are the American Academy of Facial Plastic and Reconstructive Surgery (AAFPRS), the American Academy of Otolaryngic Allergy (AAOA), the AAO-HNS, the American Board of Facial Plastic and Reconstructive Surgery (ABFPRS), the American Head and Neck Society (AHNS), the American Osteopathic Colleges of Ophthalmology and Otolaryngology–Head and Neck Surgery (AOCOO-HNS), the American Rhinologic Society (ARS), the Deafness Research Foundation (DRF), and the National Institute of Deafness and Chronic Disease (NIDCD). Table 1 provides a list of organizations as well as the total number of pages retrieved from each website.

A total of 262 educational links were reviewed. Only materials directed toward patients and that were found under the patient section of the websites were included. Materials that were in other parts of these associations’ websites (ie, directed toward health care providers) were excluded. Each document was reformatted to plain text in Microsoft Word (Microsoft Corp, Redmond, Washington). Some items were deleted from the original material as part of the exclusion criteria. These include figures, figure legends, captions, multimedia, copyright notice disclaimers, acknowledgments, author information, any information related to webpage navigation, citations, and references. The final edited articles were then analyzed individually using Readability Studio Professional Edition Version 2012.1 (Oleander Software, Ltd, Vandalia, Ohio). The level of readability was determined using 8 numeric and 2 graphical scales.

The Flesch Reading Ease (FRE) test uses sentence length and syllable count to determine a score from 0 to 100. Higher scores indicate greater ease of reading. Scores of 0 to 30 correspond to a reading level of very difficult; 30 to 50, difficult; 50 to 60, fairly difficult; 60 to 70, standard; 70 to 80, fairly easy; 80 to 90, easy; and 90 to 100, very easy. This formula is widely used in the evaluation of medical literature.

The Flesch-Kincaid Grade Level (FKGL), SMOG test, Coleman-Liau Index (CLI), Gunning Fog Index (GFI), New Fog Count (NFC), New Dale-Chall (NDC), and FORCAST use different formulas to generate a score. This score is interpreted as the academic grade level needed to comprehend the analyzed text. The FKGL looks at sentence length and syllable count. The SMOG test uses sentence length and number of complex words (words with 3 or more syllables). The CLI uses sentence length and character count. The GFI is calculated based on number of sentences and complex words. The NFC uses number of complex words, number of easy words (words with under 3 syllables), and number of sentences. The NDC uses sentence length and number of unfamiliar words. Words are deemed unfamiliar if they do not appear in a list of 3000 common words known to most fourth-grade students. Finally, the FORCAST formula analyzes a 150-word sample of text from the document and gives a score based on the number of monosyllabic words found.

The Raygor Readability Estimate (RRE) and Fry Readability Graph also present an estimated academic grade

### Table 1. Otolaryngology Associations Used

<table>
<thead>
<tr>
<th>Organization</th>
<th>Abbreviation</th>
<th>Number of Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Academy of Facial Plastic and Reconstructive Surgery</td>
<td>AAFPRS</td>
<td>17</td>
</tr>
<tr>
<td>American Association of Otolaryngic Allergy</td>
<td>AAOA</td>
<td>2</td>
</tr>
<tr>
<td>American Academy of Otolaryngology–Head and Neck Surgery</td>
<td>AAO-HNS</td>
<td>152</td>
</tr>
<tr>
<td>American Board of Facial Plastic and Reconstructive Surgery</td>
<td>ABFPRS</td>
<td>1</td>
</tr>
<tr>
<td>American Head and Neck Society</td>
<td>AHNS</td>
<td>4</td>
</tr>
<tr>
<td>American Rhinologic Society</td>
<td>ARS</td>
<td>34</td>
</tr>
<tr>
<td>Deafness Research Foundation</td>
<td>DRF</td>
<td>4</td>
</tr>
<tr>
<td>National Institute of Deafness and Chronic Disease</td>
<td>NIDCD</td>
<td>43</td>
</tr>
<tr>
<td>American Osteopathic Colleges of Ophthalmology and Otolaryngology–Head and Neck Surgery</td>
<td>AOCOO-HNS</td>
<td>5</td>
</tr>
</tbody>
</table>
level of documents but in graphical format. The RRE score is based on average number of sentences and long (6 or more characters) words per hundred words. The Fry Readability Graph uses average number of sentences and syllables per hundred words. Both tools plot the 2 independent variables assessed and determine the reading level based on a point of intersection. Further description of the readability tools and their formulas can be found in Table 2.

This study qualifies as exempt status as per the “nonhuman subject research” protocol set by the Institutional Review Board of University of Medicine and Dentistry of New Jersey–New Jersey Medical School UMDNJ.

Table 2. Readability Tools and Calculations Used in the Analysis of Otolaryngology Association Websites

<table>
<thead>
<tr>
<th>Assessment Scale</th>
<th>Variables</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flesch Reading Ease (FRE)</td>
<td>Average number of syllables per word (SY) and average number of words per sentence (W)</td>
<td>Flesch Reading Ease = 206.835 – ((84.6 × SY) - (1.015 × W))</td>
</tr>
<tr>
<td>Flesch Kincaid Grade Level (FKGL)</td>
<td>Average number of syllables per word (SY) and average number of words per sentence (W)</td>
<td>Flesch Kincaid Grade = (11.8 × SY) + (.39 × W) – 15.59</td>
</tr>
<tr>
<td>SMOG Readability Formula</td>
<td>Average number of words with 3 or more syllables (C) and average number of sentences (S)</td>
<td>SMOG = 1.0430 × √C + 3.1291</td>
</tr>
<tr>
<td>Coleman-Liau Index (CLI)</td>
<td>Average number of letters per 100 words (L) and average number of sentences per 100 words (S)</td>
<td>Coleman-Liau Index = (0.0588 × L) – (0.296 × S) – 15.8</td>
</tr>
<tr>
<td>Gunning Fog Index (GFI)</td>
<td>Number of sentences (S), number of words per sentence (W), number of words with 3 or more syllables (C)</td>
<td>Gunning Fog Index = 0.4 × (W/S + ((C/W) + 100))</td>
</tr>
<tr>
<td>New Fog Count (NFC)</td>
<td>Number of easy words (E), number of complex words (C), number of sentences (S)</td>
<td>New Fog Count = (((E) + (3 × C))/S) – 3)/2</td>
</tr>
<tr>
<td>New Dale-Chall (NDC)</td>
<td>Number of words (W), number of sentences (S), and unfamiliar words (U)</td>
<td>New Dale-Chall = (0.0496 × (W/S)) + (0.1579 × (U/W)) + 3.6365</td>
</tr>
<tr>
<td>FORCAST</td>
<td>Number of single-syllable words in a 150-word sample (SS)</td>
<td>FORCAST = 20 – (SS/10)</td>
</tr>
<tr>
<td>Raygor Readability Estimate (RRE)</td>
<td>Average number of sentences and long (six or more characters) words per 100 words</td>
<td>1. Select a 100-word passage from the selection 2. Count the number of sentences, estimated to nearest 0.1 3. Count the number of words that are six or more letters 4. Find the point on the chart (3 samples recommended for best results)</td>
</tr>
<tr>
<td>Fry Readability Graph</td>
<td>Average number of sentences and syllables per 100 words</td>
<td>1. Extract a 100-word passage from the selection 2. Count the number of sentences in each passage (Count a half sentence as 0.5) 3. Count the number of syllables in each passage 4. Find the point on the chart (3 samples recommended for best results)</td>
</tr>
</tbody>
</table>

Statistical Methods

One-way analysis of variance (ANOVA) using OriginPro (Northampton, Massachusetts) was performed to determine differences in assessment scales metrics between the various professional websites. A post hoc Tukey’s honestly significant difference (HSD) analysis was performed for ANOVA results with a \( P < .05 \).

Results

Two hundred sixty-two IPEMs from the otolaryngology association websites were downloaded and analyzed for their level of readability using 10 different readability...
scales. The mean website FRE scores ranged from 20 to 57.8 (Figure 1). Based on 9 readability assessments, all of the websites included in the analysis had mean readability levels greater than the AMA- and NIH-recommended sixth-grade reading level and estimated eighth-grade reading level of the average American. The mean FKGL scores ranged between 9.7 and 17.1. The mean CLI scores ranged from 10.7 to 15.9. The SMOG grading scales demonstrated a mean reading level from 11.6 to 18.2. The mean GFI scores ranged between 10.9 and 15.0. The mean NDC grade levels were between 8.6 and 16.0. The mean FORCAST grade levels for each website ranged from 10.4 to 12.1. The mean NFC readability scores ranged from 8.5 to 11.8 (Figure 2). Readability assessments using the RRE plotting system demonstrated that the mean readability scores fell between a range of 10.5 and 17.0 (Figure 3). Analysis with the Fry Readability Graph plotting system showed a range of 10.0 to 17.0 (Figure 4).

Analysis of variance results indicated a significant difference \( (P < .05) \) between the websites (AAFPRS, AAOA, AAO-HNS, AHNS, AOCOO-HNS, ARS, DRF, NIDCD) for each individual assessment (CLI, FKGL, FRE, FORCAST, GFI, NDC, NFC, SMOG, Fry, and RRE). Further investigation using Tukey’s HSD post hoc analysis found several differences \( (P < .05) \) between the websites for the different assessment scales. In particular, the ARS website was found to be more difficult to read than the DRF and NIDCD as measured by the CLI, FKGL, FRE, FORCAST, Fry, GFI, NDC, and SMOG. The AAFPRS was found to be even more difficult to read than the ARS when evaluated with the FRE scale. Furthermore, the AAFPRS was written at a significantly higher level than the DRF and NIDCD when assessed with the NDC scale. In addition, the NFC found a statistically significant difference between the ARS and NIDCD websites. The AAOA was more difficult to read than the DRF when analyzed with the FKGL, FRE, and SMOG scales. The ARS is written at a level higher than the AOCOO-HNS as evaluated by the NDC scale. We found no difference between the websites when evaluating with the RRE scales \( (P > .05) \).

**Discussion**

The Internet is an easily accessible resource where patients can find information about symptoms, diagnoses, and the management of diseases. Many otolaryngology organizations have embraced the opportunity for patient education by providing comprehensive, peer-reviewed educational material on their respective websites. Although this is certainly a step forward in the improvement of otolaryngology care, authors must ensure that the information is clear and easy to understand. Since the average American adult reads at an eighth-grade level, website content must be written at an even simpler level so that most of the general public can understand the information provided. Prior studies in otolaryngology, urology, orthopedics, and neurosurgery have found the readability of patient education materials to be significantly higher than the recommended reading level for such materials. Likewise, studies on heart disease, cancer, stroke, chronic obstructive pulmonary disease, and diabetes demonstrated similar results.
These findings point to the fact that this issue is significant, widespread, and in fact not just limited to otolaryngology.

In using readability as an estimate for comprehension, we found that all available documentation was more advanced than the recommended fourth- to sixth-grade level. This suggests a limitation on the usefulness of the Internet as an educational database. Articles that exceed the average reading level may be too difficult to read for much of the intended audience. As a result, patients who look online for answers may misconstrue the information presented. In addition, they may feel unsatisfied with these answers and continue to search in poorly supervised forums, blogs, or unofficial websites, leading to the spread of false information. In this regard, the Internet may be counterproductive in increasing patient understanding and improving health care.

One possible explanation for the high readability scores in these IPEMs is the inherent complexity of medicine and, specifically, otolaryngology. Complicated, multisyllable words are frequently encountered, which can increase an article’s readability score. Often, the use of jargon is unavoidable. However, most of the readability tools used in this study do account for other factors such as sentence length, number of commonly used words, and number of monosyllabic words.
Some formulas also exclude proper nouns in their definition of “complex words,” which lessens the impact of complicated eponyms and drug names. It may be prudent to carefully evaluate and address these issues when writing educational materials to make them as easy to understand as possible.

In addition, ANOVA analysis of the websites’ readability scores shows a significant difference between the otolaryngology associations. This may be due to the varying complexity of topics addressed by each association or to a difference in authorship. It would be interesting to explore the credentials of the authors and their respective training in writing patient education materials. However, regardless of the reasons behind such a discrepancy, this difference suggests a lack of standardization in the composition of IPEMs across websites.

It is important to keep in mind that the readability scores were generated from formulas and may not be absolute in their representation of comprehension. For instance, there are many nontextual ways to present information that can drastically improve the understanding of presented topics. This study does not address the use of graphs, pictures, tables, figures, or multimedia, nor does it address the effect of formatting and presentation. These additional tools are frequently used by the websites we explored and can be very effective at conveying challenging materials. As a future direction, it would be interesting to explore the perceived difficulty of reading materials through direct patient feedback.

Some recommendations to improve readability of health education materials while maintaining accuracy of the information presented include using shorter sentences (limited to 8-10 words), using lay term substitutes, changing complex medical terminologies with simpler words, and including simple and clear supporting multimedia. Following these changes, adequate methods will be needed to assess the theoretical impact of improving readability in these societies’ websites. An example could be acquiring feedback in the form of surveys from individuals using online patient education materials from these websites.

**Conclusion**

Otolaryngology associations have embraced the online health care revolution by providing valuable patient education materials on their websites. However, we found that the readability scores for all materials exceeded the recommended reading level as defined by the AMA and NIH. In addition, the websites differed among themselves in average readability.

**Author Contributions**

Jean Anderson Eloy, conception, design, data acquisition, analysis, interpretation, drafting and revision of article, final approval; Shawn Li, data acquisition, analysis, drafting of article, final approval; Khushabu Kasabwala, data acquisition, revision, final approval; Nitin Agarwal, analysis, revision, final approval; David R. Hansberry, analysis, revision, final approval; Soly Baredes, analysis, revision, final approval; Michael Setzen, analysis, revision, final approval.

**Disclosures**

Competing interests: None.
Sponsorships: None.
Funding source: None.
References