Assessment of Online Patient Education Materials From Major Ophthalmologic Associations

Grace Huang, BS; Christina H. Fang, BS; Nitin Agarwal, MD; Neelakshi Bhagat, MD; Jean Anderson Eloy, MD; Paul D. Langer, MD

**IMPORTANCE** Patients are increasingly using the Internet to supplement finding medical information, which can be complex and requires a high level of reading comprehension. Online ophthalmologic materials from major ophthalmologic associations should be written at an appropriate reading level.

**OBJECTIVES** To assess ophthalmologic online patient education materials (PEMs) on ophthalmologic association websites and to determine whether they are above the reading level recommended by the American Medical Association and National Institutes of Health.

**DESIGN, SETTING, AND PARTICIPANTS** Descriptive and correlational design. Patient education materials from major ophthalmology websites were downloaded from June 1, 2014, through June 30, 2014, and assessed for level of readability using 10 scales. The Flesch Reading Ease test, Flesch-Kincaid Grade Level, Simple Measure of Gobbledygook test, Coleman-Liau Index, Gunning Fog Index, New Fog Count, New Dale-Chall Readability Formula, FORCAST scale, Raygor Readability Estimate Graph, and Fry Readability Graph were used. Text from each article was pasted into Microsoft Word and analyzed using the software Readability Studio professional edition version 2012.1 for Windows.

**MAIN OUTCOMES AND MEASURES** Flesch Reading Ease score, Flesch-Kincaid Grade Level, Simple Measure of Gobbledygook grade, Coleman-Liau Index score, Gunning Fog Index score, New Fog Count, New Dale-Chall Readability Formula score, FORCAST score, Raygor Readability Estimate Graph score, and Fry Readability Graph score.

**RESULTS** Three hundred thirty-nine online PEMs were assessed. The mean Flesch Reading Ease score was 40.7 (range, 17.0-51.0), which correlates with a difficult level of reading. The mean readability grade levels ranged as follows: 10.4 to 12.6 for the Flesch-Kincaid Grade Level; 12.9 to 17.7 for the Simple Measure of Gobbledygook test; 11.4 to 15.8 for the Coleman-Liau Index; 12.4 to 18.7 for the Gunning Fog Index; 8.2 to 16.0 for the New Fog Count; 11.2 to 16.0 for the New Dale-Chall Readability Formula; 10.9 to 12.5 for the FORCAST scale; 11.0 to 17.0 for the Raygor Readability Estimate Graph; and 12.0 to 17.0 for the Fry Readability Graph. Analysis of variance demonstrated a significant difference ($P < .001$) between the websites for each reading scale.

**CONCLUSIONS AND RELEVANCE** Online PEMs on major ophthalmologic association websites are written well above the recommended reading level. Consideration should be given to revision of these materials to allow greater comprehension among a wider audience.
reservation of health requires a combination of quality medical care and informed, preventive lifestyle choices. To patients, medical conditions and associated procedures are complex and the terminology is often difficult to understand. Patients are increasingly using Internet resources to supplement health care knowledge. A survey from 2006 estimated that 98 million American adults have used the Internet to find health care information.1

Unfortunately, medical information on the Internet is itself often complex and requires a high level of reading comprehension. Health literacy is defined in an Institute of Medicine report as “the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions.”2 Studies have shown that the reading level of American adults is reportedly between the seventh- and eighth-grade levels.3-5 In addition, the average American adult reads at 3 to 5 grades below the highest grade of schooling completed. The American Medical Association and National Institutes of Health recommend presenting patient education materials (PEMs) at the fourth- to sixth-grade level.3-5 A key aspect of literacy is readability, or the ease with which written materials are read. Material is considered easy to read if written below the sixth-grade level; of average difficulty if written between the seventh- and ninth-grade levels; and difficult if written above the ninth-grade level.3

Seventy percent of Americans who use the Internet to obtain health information have said that it influenced their decision about how to treat an illness or condition.6 These data are significant because patients with low health literacy relying on Internet content to make health decisions could be compromised by lack of comprehension of the information. Patients with lower health literacy have been shown to have less knowledge of their disease, greater risk of hospitalization, inferior treatment compliance, poorer health, and higher mortality than people with adequate health literacy.7 These patients incur medical expenses up to 4 times greater than patients with adequate literacy skills, costing the health care system billions of dollars annually in unnecessary physician visits and hospital stays.8

A recent study from the United Kingdom investigated the readability of a Google web search of 16 ophthalmologic diagnoses.9 None of the webpages examined had a readability score within any of the recommended guidelines. There were several limitations to that study, however. The results for each ophthalmic diagnosis were limited to the first 10 websites identified. Data are more likely to be skewed when a small sample of webpages is used for each condition and if the articles vary in length or readability. In addition, the Google search terms used may not have been those used by patients.10 For example, whereas ophthalmologists may use the term ambylopia, patients may have searched for lazy eye.

In this study, we evaluated PEMs from the websites of 10 major ophthalmologic associations, including the American Academy of Ophthalmology (AAO), American Association of Ophthalmic Oncologists and Pathologists, American Association for Pediatric Ophthalmology and Strabismus (AAPOS), American Glaucoma Society (AGS), American Society of Cataract and Refractive Surgery, American Society of Ophthalmic Plastic and Reconstructive Surgery, American Society of Retina Specialists, American Uveitis Society, Cornea Society, and North American Neuro-Ophthalmology Society (NANOS). The purpose of this study was to analyze the reading level of publicly accessible Internet-based ophthalmologic information articles using 10 assessment tools.

### Methods

From June 1, 2014, through June 30, 2014, we browsed through the websites of major ophthalmologic associations and downloaded all available Internet-based PEMs. These organizations are the AAO, American Association of Ophthalmic Oncologists and Pathologists, AAPOS, AGS, American Society of Cataract and Refractive Surgery, American Society of Ophthalmic Plastic and Reconstructive Surgery, American Society of Retina Specialists, American Uveitis Society, Cornea Society, and NANOS. The Table lists the organizations along with the number of PEMs downloaded from each website. This study qualifies for exempt status as per the nonhuman subject research protocol set by the Institutional Review Board of Rutgers New Jersey Medical School.

Only material directed toward patients that was found under the patient section of the website was downloaded. Media-directed articles such as press releases or articles directed toward clinicians were excluded. The text from each article was pasted as plain text into a new Microsoft Word document (Microsoft Corp.). Text sections of nonmedical information such as copyright notices, author information, citations, references, disclaimers, acknowledgments, and webpage navigation were excluded from assessment. Figures, figure legends, and captions were also excluded.

A readability assessment was then performed using the software package Readability Studio professional edition version 2012.1 for Windows (Oleander Software, Ltd). The readability level was assessed using 8 numerical scales and 2 graphical scales. The Flesch Reading Ease (FRE), Flesch-Kincaid Grade Level (FKGL), Simple Measure of Gobbledygook (SMOG) test, Coleman-Liau Index (CLI), Gunning Fog Index (GFI), New Fog Count (NFC), New Dale-Chall Readability Formula (NDC), and

### Table. Ophthalmologic Associations Used

<table>
<thead>
<tr>
<th>Organization</th>
<th>Documents, No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Academy of Ophthalmology</td>
<td>134</td>
</tr>
<tr>
<td>American Association of Ophthalmic Oncologists and Pathologists</td>
<td>0</td>
</tr>
<tr>
<td>American Association for Pediatric Ophthalmology and Strabismus</td>
<td>117</td>
</tr>
<tr>
<td>American Glaucoma Society</td>
<td>19</td>
</tr>
<tr>
<td>American Society of Cataract and Refractive Surgery</td>
<td>0</td>
</tr>
<tr>
<td>American Society of Ophthalmic Plastic and Reconstructive Surgery</td>
<td>10</td>
</tr>
<tr>
<td>American Society of Retina Specialists</td>
<td>9</td>
</tr>
<tr>
<td>American Uveitis Society</td>
<td>35</td>
</tr>
<tr>
<td>Cornea Society</td>
<td>0</td>
</tr>
<tr>
<td>North American Neuro-Ophthalmology Society</td>
<td>15</td>
</tr>
</tbody>
</table>
difference analysis was performed for analysis of variance on the websites. A post hoc Tukey honestly significant difference scale metrics between the various ophthalmologic association websites. A post hoc Tukey honestly significant difference analysis was performed for analysis of variance results with a significance level of \( P < .05 \).

### Results

Three hundred thirty-nine PEMs from the ophthalmologic association websites were downloaded and analyzed for their level of readability using the 10 assessment techniques (Table). No PEMs were downloaded from the American Association of Ophthalmic Oncologists and Pathologists, American Society of Cataract and Refractive Surgery, or Cornea Society. The mean FRE score of the PEMs was 40.7, ranging from 17.0 to 51.0 (Figure 1). This mean value corresponds to a difficult level of reading. The mean FKGL grade scores ranged from 10.4 to 12.6. The mean CLI scores ranged from 11.4 to 15.8. The mean SMOG scores ranged from 12.9 to 17.7. The mean GFI grade scores demonstrated a mean reading level of 12.4 to 18.7. The mean NDC grade levels were between 11.2 and 16.0. The mean FORCAST grade levels for each website ranged from 10.9 to 12.5. The mean NFC readability scores ranged from 8.2 to 16.0 (Figure 2). Readability of the articles was visually demonstrated using the RRE Graph (Figure 3). The plot showed mean readability scores ranging from 11.0 to 17.0. Readability assessments using the Fry Readability Graph demonstrated that the mean scores ranged between 12.0 and 17.0 (Figure 4).

Analysis of variance results indicated a significant difference \( (P < .001) \) between the websites (AAO, AAPOS, AGS, American Society of Ophthalmic Plastic and Reconstructive Surgery, American Society of Retina Specialists, American Uveitis Society, and NANOS) for each individual readability assessment (CLI, FKGL, FRE, FORCAST scale, GFI, NFC, SMOG test, Fry Readability Graph, and RRE Graph). Further analysis with post hoc Tukey honestly significant difference analysis showed several differences \( (P < .001) \) between the websites for the different assessment scales. The AGS website was found to be more difficult to read than all of the other websites as measured by the NFC and the FRE \( (P < .001) \). In addition, it was found to be more difficult to read than the AAPOS and NANOS websites using the FKGL \( (P < .001) \) and more difficult to read than the AAO and NANOS websites using the GFI \( (P < .001) \). The NANOS website was written at a lower reading level than all of the other websites as evaluated by the FRE scale \( (P < .001) \). We found no difference between the websites when evaluating with the CLI, FORCAST scale, NFC, SMOG test, Fry Readability Graph, and RRE Graph \( (P > .05) \).

### Discussion

Patients are increasingly using the Internet to find health information because of the convenience of obtaining information at any hour at any location and performing this research anonymously. Furthermore, clinicians are more often referring their patients to online sources of PEMs. Patient education and understanding of their medical conditions are critical to optimizing the patient-physician relationship and the patient’s overall health.

---

FORCAST scale use different formulas to generate a readability grade. The FRE uses sentence length and syllable count to calculate a score between 0 and 100. Higher scores indicate greater ease of reading. A score of 90 to 100 indicates a readability level of very easy, whereas a score of 0 to 30 indicates a level of very difficult. The FKGL assessment uses the same independent variables as the FRE to determine grade level. For example, an FKGL score of 8 indicates that the reader requires an eighth-grade level of education to read and understand the article. The SMOG test uses sentence length and number of complex words (words with >3 syllables) and the CLI uses sentence and word counts to determine the grade level of a written document. The GFI assessment uses the total number of sentences and complex words (>3 syllables). The NFC uses number of complex words, number of easy words (<3 syllables), and number of sentences. The NDC considers sentence length and frequency of unfamiliar words. Words are considered unfamiliar if they do not appear on a preset list of 3000 common words recognized and known by the average fourth grader. The FORCAST scale counts the number of single-syllable words in a 150-word sample of text from a document to estimate grade level. The Raygor Readability Estimate (RRE) Graph and the Fry Readability Graph were used to visually display the grade level of reading. The RRE Graph calculates a grade level based on the average number of sentences and long words (>6 characters) per 100 words present in the document. The Fry Readability Graph uses the average number of sentences and syllables per 100 words. Both of these graphs plot the intersection of these 2 independent variables to determine a document’s grade level.

One-way analysis of variance was performed using Microsoft Excel (Microsoft Corp) to determine differences in assessment scale metrics between the various ophthalmologic association websites. A post hoc Tukey honestly significant difference analysis was performed for analysis of variance results with a significance level of \( P < .05 \).
Readability above the average American literacy level is a widespread issue that has been identified broadly throughout PEMs in other specialties and is not just limited to ophthalmology. Prior studies performed in other surgical subspecialties such as otolaryngology,\textsuperscript{20-22} urology,\textsuperscript{23} orthopedic surgery,\textsuperscript{24-26} and neurosurgery\textsuperscript{27} have shown similar results: the reading level of PEMs found on their respective websites is significantly higher than the recommended reading level. Similarly, studies on heart disease, cancer, stroke, chronic obstructive pulmonary disease, and diabetes mellitus found comparable results.\textsuperscript{3,28}

Our analysis of PEMs from major ophthalmologic association websites is consistent with previous studies of ophthalmology-related online PEMs that have shown that readability exceeds the reading level of the average American.\textsuperscript{9-29,30} Analyses using the FKGL, SMOG, CLI, GFI, NDC, NFC, and FORCAST numerical scales showed that PEMs were written at mean grade ranges all greater than the American Medical Association– and National Institutes of Health–recommended fourth- to sixth-grade reading level. The FRE score corresponded to a difficult level to read. In addition, analysis of variance showed that the ophthalmologic association websites differed...
among themselves in average readability. This may be due to the varying complexity of topics that each association addresses and the difference in authorship. Overall, this analysis demonstrated that the PEMs from major ophthalmologic association websites are too advanced for the average American patient to comprehend.

This study has certain limitations despite using 10 different assessment techniques. Readability level is a key, but not sole, component of literacy. Overall readability may also be influenced by factors such as images, content organization, layout, and design. The Suitability Assessment of Materials32 and PMOSE/IKIRSCH33 are instruments that may be used in the future to measure the readability of charts, graphs, video, and audiotaped instructions; however, they have yet to be fully validated in the medical literature.34 New readability measures being developed will take medical vocabulary, cohesion, and style into consideration when evaluating health texts in an attempt to create a gold standard of readability.35 Until the impact of multimedia on patient comprehension is better understood, effort should be focused on the word text of PEMs on patient-oriented websites. Another limitation of this study is that 4 of the numerical scales (FRE, FKGL, SMOG test, and CLI) are influenced by the number of syllables in a word. A shorter word such as drusen or ptosis may not be more understandable, while a longer word may not necessarily be more difficult to understand. This algorithm might underestimate the reading level required to comprehend the text. On the other hand, many terms in ophthalmology such as cataract are inherently technical and may be unavoidably used. Despite the author defining the word, once it is used again in the text it may be assessed again as a difficult word, which would overestimate the reading level. However, in this study, we used other readability assessment tools that examined other parameters in addition to difficult words. Another limitation is that the literacy level of potential ophthalmology patients may be higher than the average American. For example, a study of patient demographic characteristics found that the average patient undergoing cosmetic surgery at a private practice had a college education or beyond.35 Although there is no definite study that shows the same in ophthalmology, this specific subgroup of Americans (ophthalmology patients who use the Internet to obtain health information) may indeed possess higher reading comprehension than the average American. Consequently, one can make the argument that PEMs written at a higher level may be more beneficial to this subgroup. Because it is difficult to determine who is accessing these websites for PEMs, it would seem prudent to target the general population and write PEMs at the recommended grade level. Despite the limitations discussed, however, it is reassuring that all of the readability results were well correlated with each other and were consistent with results from other studies.

Conclusions

Recommendations on methods to improve the presentation of health education information include changing word choice...
and structure. Medical terms may be difficult to avoid in ophthalmology owing to their complexity; however, descriptive or instructive phrases can be substituted to enhance readability. Other difficult word types to avoid are concept words such as normal range, category words such as β-blockers, and value judgment words such as excessive. These types of words may be well defined in the minds of clinicians but not in the minds of patients.36 Simple and clear supporting audio and video multimedia should also be incorporated when appropriate. A potential method to follow up on the implications of this study may include surveys targeting readers of the online PEMs to obtain feedback about the population using these websites. Additional surveys can be used in the future to assess the impact of making these changes in PEMs. Until the impact of these changes in PEMs can be assessed, efforts should be directed at developing more appropriate and targeted PEMs that cater to the needs of the general population. Addressing the problem of limited health literacy on our websites can hopefully improve the understanding of eye disease and positively affect the eye health of individuals and populations.

ARTICLE INFORMATION
Submitted for Publication: July 10, 2014; final revision received December 22, 2014; accepted December 23, 2014.

Author Contributions: Ms Fang and Dr Agarwal had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Huang, Fang, Agarwal, Eloy, Langer. Acquisition, analysis, or interpretation of data: Huang, Fang, Agarwal, Bhagat. Drafting of the manuscript: Huang, Fang, Agarwal. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Huang, Fang, Agarwal. Administrative, technical, or material support: Agarwal. Study supervision: Eloy, Langer.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

REFERENCES