

Patient Information in Graves' Disease and Thyroid-Associated Ophthalmopathy: Readability Assessment of Online Resources

Matthew R. Edmunds,^{1,2} Alastair K. Denniston,^{1,3} Kristien Boelaert,^{3,4}
Jayne A. Franklyn,^{3,4} and Omar M. Durrani²

Background: The Internet is a vital source of information for patients hoping to learn more about their disease. Health literacy of the general population is known to be poor, with the U.S. Department of Health and Human Services (USDHHS) recommending that patient-oriented literature be written at a fourth- to sixth-grade reading level to optimize comprehensibility. In this study we assessed the readability of online literature specifically for Graves' disease (GD) and thyroid-associated ophthalmopathy (TAO).

Methods: Readability of the content of the top 20 English-language GD patient-oriented online resources and top 30 of the equivalent TAO resources returned by Google search was analyzed. Web pages were identified using the Google search terms "Graves' disease" and "Thyroid-Associated Ophthalmopathy," respectively. Extraneous text (e.g., hyperlinks, affiliations, disclaimers) was removed. Relevant text proceeded to readability analysis using four validated measures: Flesch Reading Ease Score, Flesch-Kincaid Grade Level, Simple Measure of Gobbledygook, and Gunning-Fog Index. Readability was compared with USDHHS standards.

Results: Overall, median word count (with interquartile range [IQR] and range) was 990 (IQR 846, 195–3867), with a median of 18 words per sentence (IQR 4.0, 7.5–28). Median Flesch Reading Ease Score was 46 (IQR 13, 24–64), Flesch-Kincaid Grade Level 11 (IQR 3.0, 7.2–17), Simple Measure of Gobbledygook 13 (IQR 2.0, 9.6–17), and Gunning-Fog Index 13 (IQR 3.0, 9.2–19), each equivalent to a reading level of >11th grade and "difficult" on the USDHHS classification. None of the web pages evaluated had readability scores in accordance with published guidelines. There was no significant difference with the country of origin of each web page, website commercial status, or with pages predominantly focused on GD or TAO.

Conclusions: Readability scores for online GD and TAO patient-focused materials are inferior to those recommended. Screening of this online material, as well as subsequent revision, is crucial to increase future patient knowledge, satisfaction, and compliance.

Introduction

THYROID-ASSOCIATED OPHTHALMOPATHY (TAO) is a common inflammatory condition of the orbit associated with autoimmune thyroid disease, particularly Graves' disease (GD) (1). TAO is a complex condition with a poorly understood etiology, often requiring medical and surgical management of the underlying thyroid abnormality (1). In addition, discussion may be required regarding the risks and benefits of immunosuppression or orbital radiotherapy, smoking cessation advice, and perhaps explanation of options

for surgical rehabilitation (1). With this extensive information to absorb and understand, together with clinical time constraints and understandable anxiety, many patients may forget details or initially misunderstand, going on to seek further information, commonly through the Internet (2). Indeed, it is known that educational material gained outside the initial face-to-face consultation can be helpful in aiding recall of information and improving patient knowledge (3).

The Internet has revolutionized patient education (4). Studies from earlier in this decade reported that 40–50% of the population uses the Internet to gain health information (5,6),

¹Academic Unit of Ophthalmology and ⁴School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom.

²Birmingham and Midland Eye Centre, Birmingham, United Kingdom.

³University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom.

although this proportion fell to around a third in those aged over 60 years (7). It is conceivable that Internet use for health education purposes has expanded further over the past several years, although it is known that, even in 2006, 98 million U.S. citizens searched for health information on the Internet (2). However, although the quantity of health information on the Internet is extensive, there is variable quality and reliability, and it can be difficult for patients to know which web pages provide accurate and up-to-date information among the many available (8). In addition, the information is useful only if patients have sufficient health literacy to comprehend what is written.

Health literacy is “the ability to perform basic reading and numerical tasks required to function in the healthcare environment” (2). Reading level is a crucial component of health literacy, and the readability of an online resource is therefore a critically important factor determining whether the information may be comprehended (4,9). It is estimated that 34% of adults in the United States have only basic or below basic reading literacy (10) with an inverse relationship between age and reading skill level, even when adjusted for cognitive impairment or years of schooling (11). Furthermore, it has been determined that the average reading age in the United States is between a seventh- and eighth-grade level (12,13).

In response to these apparent high levels of inadequate adult health literacy, the U.S. Department of Health and Human Services (USDHHS) recommended that health information for patients be written at or below the sixth-grade level. Information is categorized as being of “easy” readability if written at this level, “average” if assessed between seventh- and ninth-grade readability, and “difficult” if above this (Table 1) (2). Previous studies have evaluated the readability of online patient information for a variety of medical conditions, including heart disease, cancer, stroke, chronic obstructive pulmonary disease, diabetes mellitus, thyroid surgery, craniofacial conditions, age-related macular degeneration, orthopedic conditions, dermatological problems, abdominal aortic aneurysm, epilepsy, and Parkinson’s disease (Table 2) (2,4,8,9,12–20). Each of these studies determined that the readability of most online patient education materials,

using a variety of readability scores, is less than that recommended by USDHHS. This is important because such information may not be understood or, perhaps worse, may be misinterpreted by patients (14). It has previously been shown that suboptimal health literacy may impact on patient healthcare, with poorer treatment compliance and possible increased risk of hospitalization (9). Furthermore, better health education is associated with improved outcomes, lesser patient anxiety, and higher patient satisfaction (13).

There are no current data on the readability of web pages specifically for patients with GD and TAO. Given that these are common, complex conditions, requiring multidisciplinary management by endocrinologists, ophthalmologists, otolaryngologists, and radiation oncologists, an assessment of the online patient information available is important. The evaluation of this information will result in recommendations or amendments to improve readability. The aim of this study was therefore to assess the readability of online patient information literature specifically related to GD and TAO.

Materials and Methods

The readability of the content of the first 20 English-language GD patient-oriented online resources and first 30 equivalent TAO resources returned by a Google search was analyzed. Web pages were identified in November 2012 using Google search terms “Graves’ disease” and “Thyroid-Associated Ophthalmopathy,” respectively. The country of origin of each web page was recorded, as was its commercial or not-for-profit status. No web pages intended for healthcare professionals were included for assessment. A Google search was undertaken for the TAO web pages first, before proceeding to GD searches. Web pages that were present in the TAO search and then found again in the GD search were not re-analyzed, in order to avoid duplication. In keeping with previous published studies, all text not contributing to patient education (e.g., disclaimers, acknowledgements, hyperlinks, citations, affiliations, copyright notices, author information) was removed, with only relevant patient information text proceeding to analysis with an online readability calculator (www.harrymclaughlin.com) providing three distinct readability scores: Flesch-Kincaid Grade Level (FKGL), Gunning-Fog Index (GFOG), and Simple Measure of Gobbledygook (SMOG). The same text was evaluated in Microsoft Word 2010 for word count, number of words per sentence, and a further readability measure, Flesch Reading Ease Score (FRES). All web searches, relevant text extraction, readability scoring, and subsequent analysis were undertaken by one researcher (M.R.E.).

The four readability measures utilized are based on formulas assessing sentence length and word complexity. They are well validated and have previously been employed in a wide range of studies (Table 2). FRES is based on a 100-point score, with higher numbers signifying text that is easier to read. FKGL, SMOG, and GFOG represent the U.S. school-grade level (number of years of education) necessary to understand the material (Table 1). For example, a score of 9.0 indicates that a pupil in the 9th grade (14–15 years of age) can understand such a document. FKGL, SMOG, and GFOG correlate inversely with FRES, with higher FRES indicating lower FKGL, SMOG, or GFOG (19).

D’Agostino-Pearson test determined that all readability data were nonparametric in nature. All data are therefore

TABLE 1. FLESCH READING EASE SCORE WITH EQUIVALENT GRADE LEVEL; FLESCH-KINCAID GRADE LEVEL, SIMPLE MEASURE OF GOBBLEDYGOOK, OR GUNNING-FOG INDEX SCORE; U.K. SCHOOL AGE; AND U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES READABILITY RATING

FRES	FKGL/SMOG/ GFOG score	U.S. grade level	U.K. school age (years)	USDHHS readability
0–29		Above college		Difficult
30–49	>12	College	17+	
50–59	10–12	10th to 12th	15–17	Average
60–69	8–9	8th to 9th	12–15	
70–79	7	7th		
80–89	6	6th	9–12	Easy
90–100	5	5th		

Readability scores for patient-oriented health information are all recommended to be in the USDHHS “easy” classification. Adapted from Patel *et al.* (13).

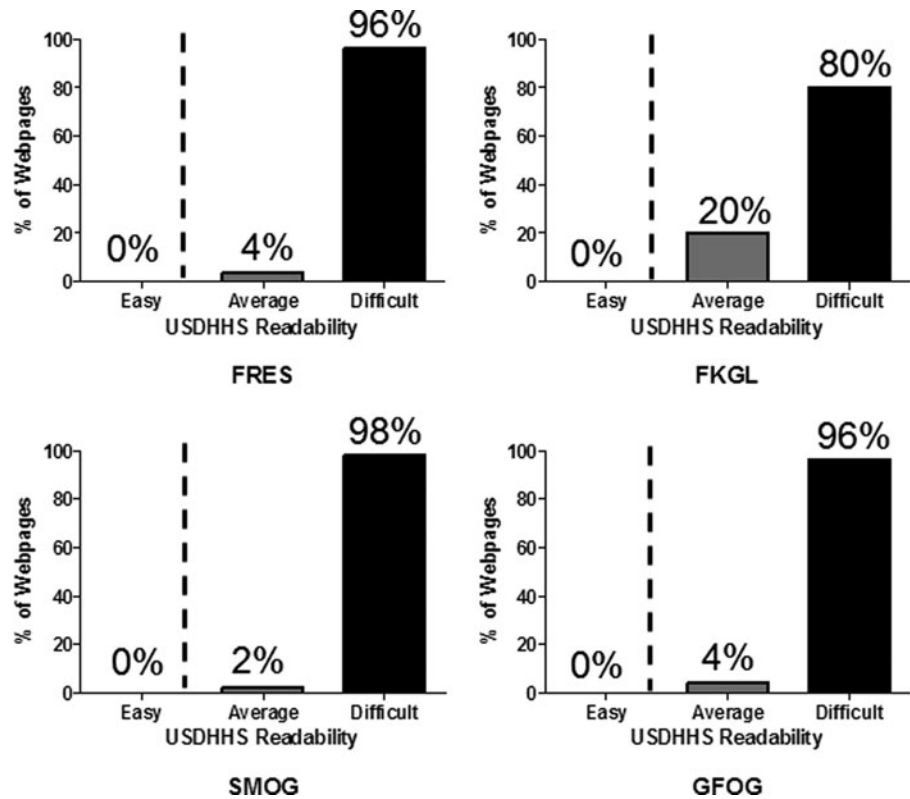
FKGL, Flesch-Kincaid Grade Level; FRES, Flesch Reading Ease Score; GFOG, Gunning-Fog Index; SMOG, Simple Measure of Gobbledygook; USDHHS, U.S. Department of Health and Human Services.

TABLE 2. SUMMARY OF THE PREVIOUS LITERATURE EVALUATING READABILITY OF ONLINE PATIENT-ORIENTED RESOURCES FOR A RANGE OF MEDICAL AND SURGICAL CONDITIONS

<i>Authors</i>	<i>Diagnosis</i>	<i>No. of web pages</i>	<i>Word count</i>	<i>FRES</i>	<i>FKGL</i>	<i>SMOG</i>	<i>GFOG</i>
Current study	Graves' disease and TAO	50	990 (846)	46 (13)	11 (3.0)	13 (2.0)	13 (3.0)
Rennie <i>et al.</i> (8) ^a	Age-related macular degeneration	26				16.2	
Walsh and Volsko (14) ^b	Heart disease	20			10.05±2.24	12.50±1.88	13.10±3.42
	Cancer	20			10.45±1.43	12.65±2.08	13.50±3.05
	Stroke	20			11.20±1.11	13.95±1.61	15.75±2.95
	COPD	20			9.85±2.25	11.80±2.44	13.95±4.30
	Diabetes	20			11.55±0.76	14.40±1.47	16.05±2.31
Sabharwal <i>et al.</i> (17) ^c	Orthopedic (AAOS website)	426			10.4 [10.2–10.6]		
Elliott and Shneker (18) ^b	Epilepsy	595		47.5 (15.7)	11.0 (3.0)	12.8 (2.4)	
Fitzsimmons <i>et al.</i> (2) ^c	Parkinson's disease	100	1450 [1081–1818]		12.1 [11.5–12.7]	14.6 [14.2–15.1]	
Tulbert <i>et al.</i> (4) ^d	Dermatology	60	2185 (388–21,097)	50 (26.8–67.5)	9.7 (6.9–12)		
Bailey <i>et al.</i> (16) ^d	Abdominal aortic aneurysm	50		39 (29–47)			
Kasabwala <i>et al.</i> (19) ^d	Rhinologic conditions	34		35.3 (9.0–53)	13.4 (10.0–18.3)	15.0 (12.6–18.7)	14.9 (11.7–19 +)
Patel <i>et al.</i> (13) ^d	Thyroid surgery	31		50.5 (29.3–67.8)	10.4 (6.9–14.9)	13.0 (11.8–14.5)	13.5 (10.6–18.0)

Comparison with median readability data (and interquartile ranges) in the current study. Data values are reported as: ^amean; ^bmean±SD; ^cmean [95% confidence interval]; ^dmean (range). AAOS, American Academy of Orthopaedic Surgeons; COPD, chronic obstructive pulmonary disease; TAO, thyroid-associated ophthalmopathy.

FIG. 1. Distribution of each of the USDHHS readability ratings for each of the readability scores of all web pages examined. USDHHS recommends readability of patient information resources to be in the “easy” category. FKGL, Flesch-Kincaid Grade Level; FRES, Flesch Reading Ease Score; GFOG, Gunning-Fog Index; SMOG, Simple Measure of Gobbledygook; USDHHS, U.S. Department of Health and Human Services.



presented with median values (and interquartile ranges [IQRs]), and between-group analysis was undertaken with the Mann–Whitney test. Spearman’s rank correlation was used to determine the relationship between different facets of readability. All analysis was undertaken using GraphPad Prism 5.0 (GraphPad Software, San Diego, CA) and readability categorized according to USDHHS standards.

Results

Of the 50 web pages assessed, 58% were of U.K. origin, 36% from the United States, and 6% from other countries. Most (76%) were from not-for-profit websites and 24% from commercial websites. Overall, median word count (with IQR and range) was 990 (IQR 846, 195–3867), with 18 mean words per sentence (IQR 4.0, 7.5–28). Median FRES was 46 (IQR 13, 24–64), FKGL 11 (IQR 3.0, 7.2–17), SMOG 13 (IQR 2.0, 9.6–17), and GFOG 13 (IQR 3.0, 9.2–19) (Table 2). All web pages examined were assessed as containing accurate information, of relevance specifically to GD and TAO patients. The range of different readability scores and their corresponding USDHHS readability category for all web pages are presented in Figure 1. There was no significant difference in any of the readability measures by the country of web page origin, whether the web page was from a commercial or not-for-profit source, or between web pages specifically for GD or TAO patients. There was no correlation between word count and any of the readability measures, but, as expected, words per sentence correlated significantly with all readability scores investigated.

Discussion

This is the first study to assess the readability of online patient-oriented material for both GD and TAO. The quantity

of online literature related to GD and TAO is vast, with 4,610,000 and 2,910,000 online resources being returned, respectively, by Google search. However, it is uncertain whether the information available is accurate, reliable, or readable as web pages are not subject to peer or lay review. These abundant online resources are only useful if patients are able to understand and act upon the contents. Although 60% of web pages were based on a search for TAO and 40% for GD, in reality (and understandably) the content of web pages overlapped considerably, often comprising information related to both diagnoses. In addition, the majority of the web pages were of U.K. origin. This is likely because Google.co.uk was used for all searches, rather than Google.com. However, the results are widely applicable to both U.K. and U.S. patients and physicians as there was no significant difference in readability statistics between U.K. and non-U.K. web pages.

Our study is in keeping with previous articles in finding that the reading level of the majority of online patient-focused information for GD and TAO is too high for the majority of the population, each web page evaluated falling short of published guidelines. None of the web pages examined, with any of the readability scores, was within recommended USDHHS readability guidelines of “easy” and all median scores were in the “difficult” USDHHS category. Only 4% of web pages were in the “average” USDHHS readability bracket as defined by FRES of 60 or more. In terms of an “average” rating by grade level of less than 9th, FKGL had 20%, SMOG 2%, and GFOG 4%. It is important to note that there were no significant differences in median readability scores with web page type (commercial or not-for-profit), country of origin, or whether predominantly related to GD or TAO. This is contrary to a previous study (2), which found commercial web pages to be easier to read, although other studies have found no such

significant difference (8). It is important to note that none of the sites were assessed as containing inaccurate or dangerous information.

The results for each condition were limited to the first 30 TAO sites and first 20 GD sites identified by Google, as it was believed that these were the most likely to be viewed by a patient, and also to avoid selection bias. We essentially wanted to recreate the experience of a typical patient searching for information relevant to them. The alternative was to have randomly chosen web pages from those returned by the Google search terms. Previous studies have found a correlation between reading difficulty and Google ranking, with higher ranked web pages possibly being easier to read (2). Google ranking is based on a number of factors, including the relevance of the web page to the search term used and number of highly ranked links to other web pages, blogs, or discussion forums. However, in our study, there was no difference in readability scores between the top 10 web pages for GD and TAO overall as compared with those of the bottom 10, out of the 50 resources that were assessed.

It is acknowledged that each of the readability scores has limitations. All of the measures estimate only the ease with which particular words can be read, not if they can be understood. Indeed, it has previously been highlighted that while a shorter word may not be more understandable, a longer word may not be more difficult to understand (19). It was also noted that some web pages that subjectively appeared to be of excellent readability and provided accurate, relevant information had poor readability scores. For example, one of the passages within a particular web page was, "If the muscles become swollen, then this can affect how well the muscles are able to control the ability of the eyes to move together. This can cause double vision. This is a symptom which occurs when the eyes do not quite point in the same direction and therefore each sends a slightly different message to the brain. When this happens our brains 'see' two of everything." This returns readability scores of FRES 69.9, FKGL 7.7, SMOG 9.7, and GFOG 10.3—that is, "average" readability, but still above USDHHS-recommended readability guidelines. Even when the passage is simplified further, mainly by shortening sentences, to "If the eye muscles swell, the eyes may not move together well. This can cause double vision. This occurs when the eyes do not point in the same direction and each sends a different message to the brain. When this happens, our brains 'see' two of everything," the readability scores improve only to an extent where FRES is 68.3, FKGL 5.3, SMOG 8.5, and GFOG 8.1. That is, even with these amendments, only FKGL is within USDHHS readability guidelines. To further put this into context, a 989-word excerpt from one of the popular *Twilight* series of books scored FRES 72.3, FKGL 6.1, SMOG 8.2, and GFOG 6.9.

By the same token, the readability scores may be confounded by medical terms used in some web pages, with polysyllabic medical terminology such as "methimazole," "thyrocyte," and "hyaluronic acid" being returned as being of poor readability even if they had been well defined and explained within the web page before further use. For example, when a passage about antithyroid treatment is compared before and after the word "methimazole" is removed, the passage without the drug name, but otherwise unchanged, has an FRES that is 20 points higher and FKGL, SMOG, and GFOG two to three U.S. school-grade levels lower.

It may be that some of the perceived readability of particular web pages is due to aspects of presentation and layout, with the inclusion of meaningful illustrations and figures, which are also known to contribute to comprehension and information recall (4). The readability assessment software does not take account of these features, and these extra dimensions, likely to contribute to comprehension of information, were not assessed in this study. In addition, it may be argued that a certain degree of complexity of information is required. If the written material is too "simplified," then the web page may be of insufficient depth for some patients, particularly if they have become familiar with relevant concepts and terminology of care having already attended perhaps multiple different specialists over a prolonged period.

For those wishing to evaluate their own patient-oriented information before making the content publicly available, we recommend the use of online readability assessment software, such as that used in our study. As well as providing readability scores, the output of such software often also provides specific sentence-by-sentence guidance on how to improve readability of literature produced. Following this, a number of agencies have identified common pitfalls in writing information specifically for patients and provide guidance on how to go about addressing these. In particular, the National Institutes of Health (www.nlm.nih.gov/medlineplus/etr.html) recommend simple strategies such as finding alternatives for complex words; avoiding jargon, abbreviations, and acronyms; keeping sentences to only 10–15 words; and providing simple pronunciation guides. These resources and each of these suggestions should be borne in mind whenever patient-oriented information is prepared (21).

In summary, the readability of GD and TAO web pages specifically for patient information purposes is not in keeping with estimates of the reading level and health literacy of the adult population. Clinicians treating GD and TAO patients should be aware that the online literature may be limited in its readability. We recommend that all current GD and TAO web pages are independently audited for their readability by lay reviewers and readability experts, with amendment as necessary. It may also be beneficial, in future, to have minimum readability stipulations. Individual clinicians specializing in the management of GD and TAO may wish to make their own assessment of the appropriateness and readability of particular web pages and provide the relevant links for recommended web pages to their patients to guide their reading.

Acknowledgments

The Academic Unit of Ophthalmology is supported by the Birmingham Eye Foundation (U.K. Registered Charity 257549). M.R.E. is funded by the Wellcome Trust.

Author Disclosure Statement

No competing financial interests exist.

References

1. Bahn RS 2010 Graves' ophthalmopathy. *N Engl J Med* 362:726–738.
2. Fitzsimmons PR, Michael BD, Hulley JL, Scott GO 2010 A readability assessment of online Parkinson's disease information. *J R Coll Physicians Edinb* 40:292–296.

3. O'Connor AM, Rostom A, Fiset V, Tetroe J, Entwistle V, Llewellyn-Thomas H, Holmes-Rovner M, Barry M, Jones J 1999 Decision aids for patients facing health treatment or screening decisions: systematic review. *BMJ* **319**:731–734.
4. Tulbert BH, Snyder CW, Brodell RT 2011 Readability of patient-oriented online dermatology resources. *J Clin Aesthet Dermatol* **4**:27–33.
5. Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW 2002 Patients' use of the Internet for medical information. *J Gen Intern Med* **17**:180–185.
6. Baker L, Wagner TH, Singer S, Bundorf MK 2003 Use of the Internet and e-mail for health care information: results from a national survey. *JAMA* **289**:2400–2406.
7. Flynn KE, Smith MA, Freese J 2006 When do older adults turn to the Internet for health information? Findings from the Wisconsin Longitudinal Study. *J Gen Intern Med* **21**:1295–1301.
8. Rennie CA, Hannan S, Maycock N, Kang C 2007 Age-related macular degeneration: what do patients find on the Internet? *J R Soc Med* **100**:473–477.
9. Baker GC, Newton DE, Bergstresser PR 1988 Increased readability improves the comprehension of written information for patients with skin disease. *J Am Acad Dermatol* **19**:1135–1141.
10. National Center for Education Statistics, U.S. Department of Education 2007 The Condition of Education. U.S. Government Printing Office, Washington, DC.
11. Gazmararian JA, Baker DW, Williams MV, Parker RM, Scott TL, Green DC, Fehrenbach SN, Ren J, Koplan JP 1999 Health literacy among Medicare enrollees in a managed care organization. *JAMA* **281**:545–551.
12. Patel PP, Hoppe IC, Ahuja NK, Ciminello FS 2011 Analysis of comprehensibility of patient information regarding complex craniofacial conditions. *J Craniofac Surg* **22**:1179–1182.
13. Patel CR, Cherla DV, Sanghvi S, Baredes S, Eloy JA 2013 Readability assessment of online thyroid surgery patient education materials. *Head Neck* **35**:1421–1425.
14. Walsh TM, Volsko TA 2008 Readability assessment of Internet-based consumer health information. *Respir Care* **53**:1310–1315.
15. Misra P, Kasabwala K, Agarwal N, Eloy JA, Liu JK 2012 Readability analysis of Internet-based patient information regarding skull base tumors. *J Neurooncol* **109**:573–580.
16. Bailey MA, Coughlin PA, Sohrabi S, Griffin KJ, Rashid ST, Troxler MA, Scott DJ 2012 Quality and readability of online patient information for abdominal aortic aneurysms. *J Vasc Surg* **56**:21–26.
17. Sabharwal S, Badarudeen S, Unes Kunju S 2008 Readability of online patient education materials from the AAOS web site. *Clin Orthop Relat Res* **466**:1245–1250.
18. Elliott JO, Shneker BF 2009 A health literacy assessment of the epilepsy.com website. *Seizure* **18**:434–439.
19. Kasabwala K, Misra P, Hansberry DR, Agarwal N, Baredes S, Setzen M, Eloy JA 2013 Readability assessment of the American Rhinologic Society patient education materials. *Int Forum Allergy Rhinol* **3**:325–333.
20. Boullos MN 2005 British Internet-derived patient information on diabetes mellitus: is it readable? *Diabetes Technol Ther* **7**:528–535.
21. U.S. National Library of Medicine and National Institutes of Health. How to write easy-to-read health materials. www.nlm.nih.gov/medlineplus/etr.html (accessed June 10, 2013).

Address correspondence to:

*Matthew R. Edmunds, MD, MRCP, MRCOphth
Birmingham and Midland Eye Centre
Dudley Road
Birmingham B18 7QH
United Kingdom*

E-mail: m.r.edmunds@bham.ac.uk