Dr Google: The readability and accuracy of patient education websites for Graves' disease treatment

Amanda C. Purdy, BS,[a] Almoatazbellah Idriss, MD,[b] Susan Ahern, DO,[b] Elizabeth Lin, MD,[b] *and* **Dawn M. Elfenbein, MD, MPH,[a]** *Orange, CA*

Background. *National guidelines emphasize the importance of incorporating patient preferences into the recommendations for the treatment of Graves' disease. Many patients use the Internet to obtain health information, and search results can affect their treatment decisions. This study compares the readability and accuracy of patient-oriented online resources for the treatment of Graves' disease by website affiliation and treatment modality.*

Methods. *A systematic Internet search was used to identify the top websites discussing the treatment of Graves' disease. Readability was measured using 5 standardized tests. Accuracy was assessed by a blinded, expert panel, which scored the accuracy of sites on a scale of 1 to 5. Mean readability and accuracy scores were compared among website affiliations and treatment modalities.*

Results. *We identified 13 unique websites, including 2 academic, 2 government, 5 nonprofit, and 4 private sites. There was a difference in both readability (mean 13.2, range 9.1–15.7,* $P = .003$*) and accuracy (mean 4.04, range 2.75–4.50,* $P = .019$*) based on website affiliation. Government sites (mean readability 11.1) were easier to read than academic (14.3,* $P < .01$*), nonprofit (13.9,* $P < .01$*), and private sites (13.5,* $P < .05$*). Academic sites (mean accuracy 4.50) were more accurate than private sites (3.56,* $P < .05$*).*

Conclusion. *Online patient resources for the treatment of Graves' disease are written at an inappropriately high reading level. Academic sites contain both the most accurate and the most difficult to read information. Private sites represented the majority of our top results but contained the least accurate information. (Surgery 2017;■:■-■.)*

From the Departments of Surgery[a] and Medicine,[b] University of California Irvine Medical Center, Orange, CA

The 3 initial treatment options for Graves' disease include antithyroid medications (ATM), radioactive iodine (RAI) administration, and total thyroidectomy. There are many reasons to choose one treatment modality over another, and there is often more than one acceptable option for each patient. The definitive treatment used most commonly in the United States is RAI; however, thyroidectomy also is considered to be a safe and effective initial treatment option for Graves' disease.[1,2] The most common reasons patients undergo thyroidectomy for Graves' disease include treatment failure of ATM or RAI, moderate to severe Graves' ophthalmopathy, a suspicious nodule, a large goiter, severe reaction to ATM, pregnancy, and patient preference in the absence of any absolute or relative indication.[2-5] The 2016 guidelines of the American Thyroid Association specifically advise physicians to have a conversation with their patients about each treatment modality and to emphasizes the importance of incorporating patient preferences into treatment recommendations.[5]

In a time when increasing numbers of patients are using the Internet to obtain health information, it is unlikely that the physician is the patient's sole source of information.[6] According to a 2010 survey, 74% of Internet users in the United States had obtained health information online.[7] Information

Presented at the 12th Annual Academic Surgical Congress in Las Vegas, NV, February 7–9, 2017.

Accepted for publication July 5, 2017.

Reprint requests: Dawn M. Elfenbein, MD, MPH, Department of Surgery, University of California Irvine Medical Center, 333 City Blvd. West, Suite 1600, Orange, CA 92868. E-mail: delfenbe@uci.edu.

0039-6060/$ - see front matter

http://dx.doi.org/10.1016/j.surg.2017.07.011

obtained online has been shown to influence how patients choose to treat their illnesses, which is why it is essential for patient-directed information on the Internet to be both accurate and written at an appropriate reading level.[6] The American Medical Association recommends a maximum sixth grade reading level for all patient-education materials.[8]

Many prior studies have demonstrated that patient-directed educational material available on the Internet is of variable quality and is written often at an inappropriately high reading level.[9-13] This inappropriate material contributes to literacy-related barriers to health, which has been associated with poor health outcomes.[14,15] Accurate sites written at an appropriate reading level are important particularly for conditions like Graves' disease, where such a large emphasis is placed on shared decision-making. Although prior studies have assessed online, patient-education materials about thyroid nodules,[9] thyroid operation,[12] and Graves ophthalmopathy,[16] no prior studies have evaluated the accuracy and readability of patient education websites discussing specifically treatment options for Graves' disease. The aim of our study was to compare the readability and accuracy of patient education websites for the treatment of Graves' disease by website affiliation and treatment modality.

MATERIALS AND METHODS

Systematic web search and categorization based on website affiliation. The following terms were searched in the popular commercial search sites of Google, Bing, and Yahoo: (1) surgery for Graves' disease, (2) RAI for Graves' disease, and (3) medication for Graves' disease. The top 40 hits of each search were recorded, corresponding to the first 4 pages of the results of the search engine. Individual sites from each search were included only if they were designed for patients, were based in the United States, were written in English, and included at least 200 words discussing the 3 treatment modalities. Websites were excluded if they were designed for health care providers, had access restricted by subscriptions or fees, or presented only videos. Sites meeting inclusion criteria from each search strategy were put into groups according to parent website. The final list of sites evaluated were those 10 sites that came up at least once in the top 40 hits when searching for all 3 treatment modalities.

Websites were then categorized by affiliation. Those associated with an academic institution or ending in ".edu" were considered academic sites, and sites associated with a nonprofit organization or ending in ".org" were labeled as nonprofit. Websites with ".gov" domains were considered government sites. Sites with a ".com" domain that were not associated with an academic institution or nonprofit organization were categorized as private.

Assessment of readability. Readability was measured by 5 standardized tests: the Flesch-Kincaid grade level, the Gunning-Fog score, the Coleman-Liau index, the Simple Measure of Gobbledygook, and the Automated Readability index. These tools use syllables per word and words per sentence as a measure of text complexity.[17] Readability scores correspond to the years of education needed to comprehend written materials. Each document was given 4 mean readability scores: one mean readability score for information on each of the 3 treatment modalities and then an overall website readability score. Readability levels were compared across website affiliations and treatment modalities.

Assessment of accuracy. To deidentify the documents, content from each website was copied and pasted into individual text documents, and all identifying information was removed. Each document was then assigned a random identification number. Deidentified documents were distributed to an expert panel consisting of one endocrine surgeon and 3 endocrinologists. Experts were asked to assess the accuracy of each site using a 5-point accuracy scoring system developed by Storino et al.[13] A score of 1 means <25% of the information is accurate, a score of 2 means 25% to 50% of the information is accurate, a score of 3 means 50% to 75% of the information is accurate, a score of 4 means 76% to 99% of the information is accurate, and a score of 5 means 100% of the information is accurate. Experts rated each website on accuracy of information on ATM, accuracy of information on RIA treatment, and accuracy of information on surgery; in addition, each reviewer gave each website an overall accuracy score.

Statistical analysis. All data were normally distributed, and Stata version 14 (StataCorp 2015, College Station, TX) was used for analysis. Analysis of variance was used to determine if there was a significant difference in readability or accuracy based on website affiliation. Tukey's honestly significant difference test was then used for pairwise comparisons of means to determine, which were significantly different than one another. Assessment of agreement of raters was done using a 2-way mixed-effect, intraclass coefficient model.