Analysis of Trustworthiness and Readability of English and Spanish Hypo- and Hyperthyroid-Related Online Patient Education Information

Chitra Parikh, BA[1] and Adam M Ostrovsky, BS[1] (iD)

Abstract

This study aims to determine the readability and trustworthiness of English and Spanish hypo- and hyperthyroid-related online information. Google searches were conducted for four search terms: hypothyroidism, Hashimoto's Disease, hyperthyroidism, and Graves' Disease. For each search term, the first 10 websites were analyzed with a total of 40 websites analyzed. Readability formulas were used to determine English and Spanish readability. Trustworthiness was determined using HONcode status, JAMA Benchmark Criteria, and NLM Trustworthy Score. Overall readability largely exceeded recommended grade levels. Only 1 website (2.5%) presented information below the eighth-grade reading level based on overall Readability Consensus score, while 31 websites (77.5%) exceeded this threshold for all measures. The mean (SD) English readability grade level was 9.6 (3.44); the mean (SD) Spanish grade was 8.5 (4.58). No significant relationships were found between the JAMA Benchmark Criteria, NLM Trustworthy Score, HONcode status, and readability. 67.5% of websites analyzed (n = 27) were certified with the Health on the Net Foundation's code of conduct. Websites about common thyroid-related conditions have overall poor readability. The availability of resources for Spanish-speaking patients is also poor. Steps should be taken to ensure that online health-related materials are comprehensible. Physicians should recognize that patients may have few trustworthy and easy-to-understand sources to access information. The readability and trustworthiness of sources should be considered when providing patients suggested sources for further reading. It may be particularly helpful for physicians to utilize websites with favorable readability scores such as the American Thyroid Association website.

Keywords

readability, trustworthiness, health literacy, hypothyroidism, hyperthyroidism, Hashimoto's, graves'

Introduction

In the US, the percent of internet users seeking healthcare-related information online has risen over the past two decades from between 56% and 79% to 93% in 2022,[1,2,3] currently making it the second most common way patients access such information apart from talking to healthcare providers. Apps—eg, Instagram and TikTok—and social media—such as Facebook—are also powerful purveyors of healthcare information, and currently round out the list of the top 10 ways that such information is transmitted.[4] Over 90 million Americans, however, are estimated to have limited health literacy, interfering in their ability to successfully understand and make actionable judgments based on the information provided to them.[5] Health literacy is the single best predictor of a patient's health.[6] Limited health literacy has been linked with worse chronic illness management,[7]

decreased ability to participate in shared medical decision-making,[8] and lower adherence to therapies.[9]

Thyroid disease is estimated to impact 20 million Americans,[10] yet evidence is scarce for the readability and trustworthiness of English-language websites relaying information related to thyroid-related illnesses. Current literature is largely limited to searches regarding thyroid nodules,[11] thyroid surgery,[12] and thyroid cancer.[13] The readability of online resources concerned with the largest causes of

[1] Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA

Corresponding Author:
Adam M Ostrovsky, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA.
Email: adam.ostrovsky@students.jefferson.edu

hypothyroidism and hyperthyroidism—Hashimoto's thyroiditis and Graves' disease, respectively—does not yet appear to be well characterized. A lack of mainstream sources delivering high-quality health information at accessible reading levels can leave behind millions of patients and manifest itself via the inefficient and inappropriate utilization of health care resources.

Current American Medical Association (AMA), National Institutes of Health (NIH), and Centers for Disease Control and Prevention (CDC) guidelines state that patient education materials should be written at no higher than a sixth to eighth-grade level.[6] However, studies across other specialties have shown that online educational materials for patients about common diagnoses are written at higher grade levels, hindering readability.[14,15] To compare to these standards, this study attributes grade-levels to calculated readability scores in order to offer a direct means of comparison to the AMA, NIH, and CDC guidelines.

This study aimed to determine the trustworthiness and readability of available online resources associated with hypothyroidism, hyperthyroidism, Hashimoto's Disease, and Graves' Disease in both English and Spanish. Readability scores were then compared to the recommended grade levels presented by the AMA, NIH, and CDC to analyze the landscape of available online information. In doing so, this paper aims to identify gaps in trustworthiness and readability based on grade level that can be targeted to increase access to resources for common thyroid disorders.

Methods

Website Selection

An "unbiased" Google search independent of previous location and search history[16] was performed on February 26, 2022, with the first 10 English-language websites for each search term ("hypothyroidism," "Hashimoto's Disease," "hyperthyroidism," and "Graves' Disease") being examined. The four equivalent Spanish-language terms ("hipotiroidismo," "enfermedad de Hashimoto," "hipertiroidismo," and "enfermedad de Graves") were also searched, with results often being fewer in number and overlapping with their English-language website counterparts. In total, 40 unique websites were analyzed. Websites were sorted by domain (commercial, government, organization) except for those affiliated with healthcare providers, which were organized separately.

Readability Assessment

Readability was assessed by seven validated readability formulas: the Flesch Reading Ease score (FRE), Gunning FOG Index (GFOG), Flesch–Kincaid Grade Level (FKGL), The Coleman–Liau Index (CLI), The Simple Measure of Gobbledygook Index (SMOG), the Automated Readability Index (ARI), and the Linsear Write Formula (LWF). An overall Readability Consensus was provided by a website

based on all seven of the above-mentioned formulas.[17] Details of each measuring instrument, including how each works and what exactly each is measuring, are also available on this website. The Rix Readability Formula was used as a validated tool for obtaining the readability of resources written in Spanish when they were found to be available.[18]

Trustworthiness Assessment

Website trustworthiness was evaluated by certification status in the Health on the Net Foundation code of conduct—HONcode Certification[19]—the National Library of Medicine (NLM) Trustworthy Score,[20] and the JAMA Benchmark Criteria.[21] The NLM Trustworthy score incorporates how updated a website is, the authority of the author, and the accuracy of the sources cited. The NLM Trustworthy score is assigned based on recency and currency of webpages, the publisher's author, and an evaluation of cited sources' accuracy. Scores range from 0 to 6 with a maximum of two points being attributed to each of the above three categories.

The JAMA Benchmark Criteria similarly examines aspects of websites including the author's biographical information, any conflicts of interest they may have, and how current and accurate displayed content is. It is graded on a 4-point scale, with one point each towards meeting the criteria for authorship, attribution, disclosure, and currency; a higher score indicates more reliability. Each survey can be found above at its respective source.

Statistical Analysis

GraphPad Prism 9.4.0 was used for data analysis. Descriptive statistics were collected for each web page's readability scores, and the Kruskal–Wallis ANOVA followed by Dunn's Multiple Comparison Test was used to compare the readability between different (Commercial, Organization, Government, and Healthcare Provider) website types. Differences in trustworthiness between the four search terms were similarly assessed. Welch's t-test was used to compare readability between HONcode Certified and non-Certified websites. Results were considered statistically significant at $P < .05$. Excel (MS Excel 2016, v.16.62) was used to plot point estimates of mean web page readability across the different readability tests with 95% CIs.

Results

40 unique websites were identified and included for analysis. Organization websites composed the largest portion of the sample (13, 32.5%), followed by government websites (10, 25.0%), commercial websites (8, 20.0%), and those affiliated with healthcare providers (8, 20.0%). 27, or 67.5% of all websites, were HONcode Certified, with government websites making up 9 (33.3%) of certified sites, followed by commercial websites (7, 25.9%), organization websites (7, 25.9%), and healthcare provider sites (3, 11.1%).

Table 1. Website Readability Categorization by Website Ownership and HON Certification for Hypo- and Hyperthyroid Related Online Health Information.

Website Type	No. (%)	Flesch Reading Ease Score	Gunning FOG	Flesch-Kincaid Grade level	Coleman-Liau Index	SMOG	Automated Readability Index	Linsear Write Formula	Consensus
		Readability Score, Median (Range)							
Commercial	8 (20.0)	52.9 (18.7-56.0)	10.4 (9.6-14.3)	8.7 (7.9-12.0)	13.0 (12.0-20.0)	7.9 (7.0-8.6)	8.6 (7.7-11.4)	5.7 (3.5-8.9)	9.0 (8.0-11.0)
Healthcare Provider	9 (22.5)	50.4 (38.4-66.6)	10.8 (8.0-13.0)	8.9 (6.1-10.8)	13.0 (11.0-15.0)	8.0 (6.3-9.3)	8.7 (6.1-10.2)	6.7 (4.5-8.7)	9.0 (7.0-10.0)
Government	10 (25)	50.8 (39.3-59.2)	10.4 (9.5-14.8)	9.0 (7.9-12.4)	13.0 (11.0-14.0)	7.8 (7.2-10.8)	8.5 (7.5-13.1)	6.3 (5.6-12.9)	9.0 (8.0-13.0)
Organization	13 (32.5)	47.9 (7.5-61.5)	11.7 (8.7-20.3)	10.0 (7.5-17.1)	12.0 (10.0-18.0)	8.9 (6.5-14.5)	9.9 (7.2-17.3)	9.0 (4.7-16.3)	10.0 (8.0-17.0)
p-value		0.92	0.35	0.36	0.54	0.17	0.36	0.02*	0.29
HON Certification	No. (%)								
		Readability Score, Mean (SD)							
Certified	27 (67.5)	50.24 (5.71)	11.01 (1.41)	9.22 (1.29)	12.78 (0.96)	8.20 (1.05)	9.05 (1.50)	7.08 (2.30)	9.33 (1.28)
Not Certified	13 (32.5)	44.65 (16.20)	12.05 (2.98)	10.22 (2.77)	13.38 (2.84)	8.75 (2.06)	10.17 (2.78)	8.21 (3.55)	10.15 (2.60)
p-value		0.27	0.27	0.26	0.48	0.39	0.21	0.33	0.32

In general, readability was above the recommended grade levels for most websites in this study. Only one website (2.5%) published information below an eighth-grade reading level as per the overall Readability Consensus score, in comparison to 31 websites (77.5%) that were written with a complexity surpassing the eighth-grade threshold across all readability measures studied. The remaining eight websites surpassed the eighth-grade threshold across some but not all readability scales. The mean (SD) English readability grade-level score was 9.6 (3.44), while the mean (SD) Spanish score was 8.5 (4.58). No significant relationships were found between the JAMA Benchmark Criteria (mean = 3.23/4, SD = 0.92), NLM Trustworthy Score (mean = 4.95/6, SD = 1.13), HONcode status, and readability. Kruskal–Wallis ANOVA and subsequent Dunn's Multiple Comparison Tests showed that website category did not significantly impact readability scores (Table 1). Further, most readability tests tended to show a mean readability above the eighth-grade threshold, and only the LWF approached a mean readability score near the sixth grade (Figure 1).

Discussion

While ample online resources are available outlining common thyroid conditions, the readability and trustworthiness of these materials can significantly impact a patient's understanding of their condition. This study's results indicate that a majority of online sources regarding thyroid disorders are written and published at a level above that recommended by the AMA, NIH, and CDC. Previous studies in other specialties have similarly found that the readability of pertinent online materials was largely too complex, including websites related to common diagnoses in internal medicine[14] and online information about breast cancer.[15]

Patients' health literacy plays an important role in their ability to successfully access, understand, and comply with medical information. Patients with low health literacy have been reported to have higher emergency department use and return rates and are more likely to be hospitalized.[22] Poor health literacy is estimated to contribute more than 73 billion dollars of additional burden to the US healthcare system.[23] The healthcare cost of Medicaid patients with limited literacy is about four times that of those with adequate health literacy.[24] In general, lower-income, uninsured, and sicker consumers are less satisfied with health care information and consult fewer sources of health information compared to middle-income, insured, and healthier consumers.[4] Clinically, the implications of this research suggest that it is important for physicians to recognize that their patients may have few trustworthy, easy-to-understand sources for which to turn for information. Clinicians may choose to provide patients with a list of suggested sources for further reading that considers their calculated readability and trustworthiness scores.

It is implausible and impractical to suggest that patients do not seek health information online. As such, it is paramount to direct patients to reliable and accessible sources.
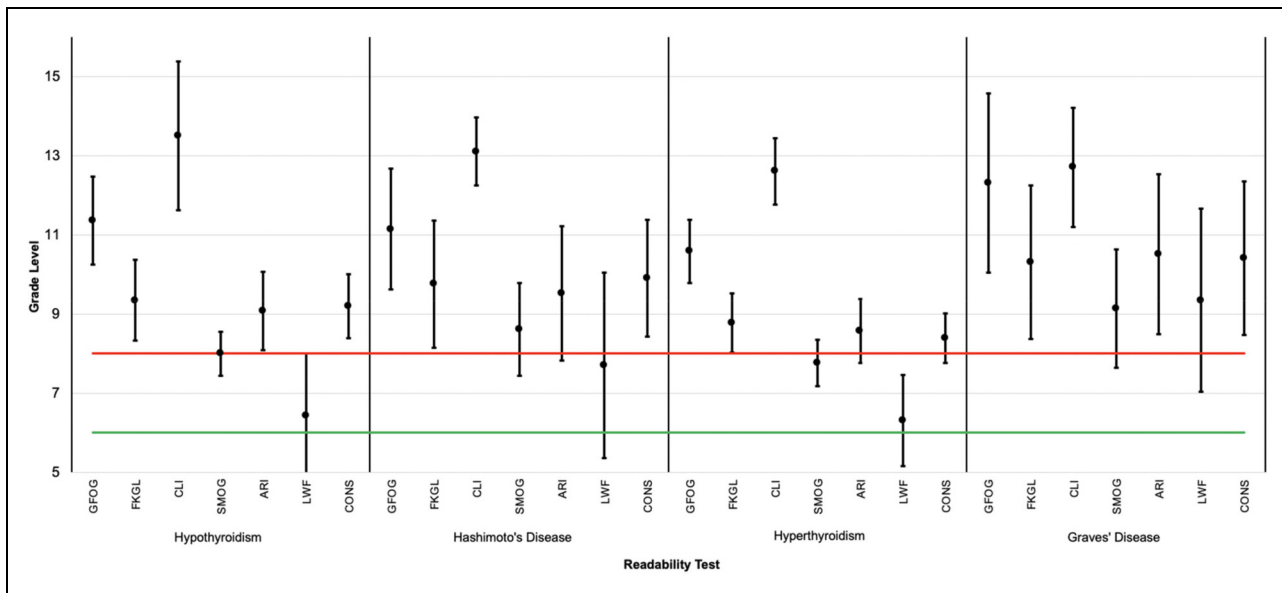
Figure 1. Readability scores for each search term are presented in terms of readability tests and their equivalent grade levels. The source for this figure is the author's analysis of data for the search terms and readability tests included in the analysis. *GFOG*, Gunning FOG Index; *FGKL*, Flesch–Kincaid Grade Level; *CLI*, The Coleman–Liau Index; *SMOG*, Simple Measure of Gobbledygook; *ARI*, Automated Readability Index; *LWF*, Linsear Write Formula; *CONS*, Readability Consensus.

Resources provided by the American Thyroid Association most closely adhered to the recommended reading thresholds. Through targeted communication and counseling, clinicians could recommend that patients review the resources with better calculated readability scores.

Limitations

Several limitations were identified. First, only a single search engine (Google) was used, though it was chosen given that it handles approximately 88% of American and 92% of international search queries, granting insight into what a majority of users are searching.[25] Second, only a single search browser (Google Chrome) was used but was similarly chosen for its dominance in worldwide market share of 65%.[25] Third, analyses were restricted to only the top ten websites per search, given that the intention was only to assess the readability of information most likely to be selected and ultimately read by the average user.[26] Finally, readability formulas have received criticism ranging from questioning their inter-formula reliability, to the very means by which they work. In the context of health content, previous work has demonstrated that even when applied to the same text, certain formulas disagreed by up to five grade levels.[27] Another commentary by Schriver[28] points out that using sentence length and word frequency to discriminate between readability levels may not truly accomplish the stated goals. Medical terminology can have a lot of long words that may not necessarily increase the difficulty of the text, but simultaneously come with no replacement. Nevertheless, the readability level by grade level of widespread online materials regarding the most common thyroid diagnoses is largely above AMA, NIH, and CDC recommendations.

Though the specific grade level of different readability scores may vary, all formulas are consistent in characterizing the text as too difficult to read. This perhaps points to a role of readability formulas in predicting the "difficulty ceiling" of a piece of writing, giving authors an initial impression of where their work stands. While the trustworthiness of the information overall is satisfactory, effort should be spent on increasing readability to make information more accessible to a wider audience of patients.

CRediT Authorship Contribution Statement

Chitra Parikh: Conceptualization, Writing – Original Draft, Writing – Review and Editing, Visualization, Validation, Supervision. **Adam Ostrovsky**: Conceptualization, Methodology, Formal Analysis, Writing – Original Draft, Writing – Review and Editing.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iD

Adam M Ostrovsky 🔟 https://orcid.org/0000-0001-9503-4841

References

1. Hesse BW, Nelson DE, Kreps GL, et al. Trust and sources of health information: the impact of the internet and its implications for health care providers: findings from the first Health Information National Trends Survey. *Arch Intern Med.* 2005;165(22):2618-24. doi:10.1001/archinte.165.22.2618

2. Ybarra ML, Suman M. Help seeking behavior and the internet: a national survey. *Int J Med Inform.* 2006;75(1):29-41. doi:10.1016/j.ijmedinf.2005.07.029

3. Kilo Health. Digital health trends in 2023: how we seek to heal online. The Changing Outlook to Digital Health. Published 2023. Accessed April 11, 2023. https://kilo.health/digital-health/.

4. Duke C. C., Stanik C., Beaudin-Seiler B., et al. (2017). (rep.). Right place, right time improving access to health care information for vulnerable patients. Robert Wood Johnson Foundation. Retrieved May 14, 2022. https://altarum.org/sites/default/files/uploaded-publication-files/USE_RPRT_Consumer_Perspectives_Final.pdf.

5. Improving Electronic Health Records Patient Education Materials. Improving electronic health records patient education materials|Digital Healthcare Research. 2022. Retrieved June 14, 2022. https://digital.ahrq.gov/ahrq-funded-projects/improving-electronic-health-records-patient-education-materials

6. Badarudeen S, Sabharwal S. Assessing readability of patient education materials: current role in orthopaedics. *Clin Orthop Relat Res.* 2010;468(10):2572-80. https://doi.org/10.1007/s11999-010-1380-y

7. Schillinger D, Grumbach K, Piette J, et al. Association of health literacy with diabetes outcomes. *JAMA.* 2002;288(4):475-82. https://doi.org/10.1001/jama.288.4.475

8. Kim SP, Knight SJ, Tomori C, et al. Health literacy and shared decision making for prostate cancer patients with low socioeconomic status. *Cancer Investig.* 2001;19(7):684-91. https://doi.org/10.1081/cnv-100106143

9. Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. *Patient Educ Couns.* 2016;99(7):1079-86. https://doi.org/10.1016/j.pec.2016.01.020

10. American Thyroid Association. General information/press room. Accessed April 11, 2023. https://www.thyroid.org/media-main/press-room/.

11. Barnes JA, Davies L. Reading grade level and completeness of freely available materials on thyroid nodules: there is work to be done. *Thyroid.* 2015;25(2):147-56. doi:10.1089/thy.2014.0352

12. Patel CR, Cherla DV, Sanghvi S, Baredes S, Eloy JA. Readability assessment of online thyroid surgery patient education materials. *Head Neck.* 2013;35(10):1421-5. doi:10.1002/hed.23157

13. Chang KL, Grubbs EG, Ingledew PA. AN analysis of the quality of thyroid cancer websites. *Endocr Pract.* 2019;25(10):1003-11. doi:10.4158/EP-2018-0595

14. Hutchinson N, Baird GL, Garg M. Examining the Reading level of internet medical information for common internal medicine diagnoses. *Am J Med.* 2016;129(6):637-9. https://doi.org/10.1016/j.amjmed.2016.01.008

15. Li Y, Zhou X, Zhou Y, et al. Evaluation of the quality and readability of online information about breast cancer in China. *Patient Educ Couns.* 2021;104(4):858-64. https://doi.org/10.1016/j.pec.2020.09.012

16. Ottaway Digital. Unbiased/neutral google search results. 2021. Retrieved August 5, 2022. https://www.ottawaydigital.com/unbiased-search.php

17. Free readability formulas: Free readability tools: Readability calculators. Free readability formulas: Free readability tools: readability calculators. 2022. Retrieved August 5, 2022. https://readabilityformulas.com/

18. Anderson J. Lix and Rix: variations on a little-known readability Index. *Journal of Reading.* 1983;26(6):490-6. http://www.jstor.org/stable/40031755

19. Health on the Net. HONcode Search. 2022. Retrieved August 5, 2022. https://www.hon.ch/en/search.html

20. NNLM. Evaluating health websites. 2022. Retrieved August 5, 2022. https://old.nnlm.gov/initiatives/topics/health-websites

21. Silberg WM. *Assessing, controlling, and assuring the quality of medical information on the internet.* JAMA. 1997. https://jamanetwork.com/journals/jama/fullarticle/415407. Retrieved August 5, 2022, from.

22. Griffey RT, Kennedy SK, McGowan D, Goodman L, Kaphingst M, A K. Is low health literacy associated with increased emergency department utilization and recidivism? *Acad Emerg Med.* 2014;21(10):1109-15. https://doi.org/10.1111/acem.12476

23. Kirsch I, Jungeblut A, Jenkins L, Kolstad A (1993). Adult literacy in America: A first look at the findings of the National Adult Literacy Survey. National Center for Education Statistics (NCES) Home Page, a part of the U.S. Department of Education. Retrieved August 5, 2022. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=93275

24. Rosenbaum AJ, Pauze D, Pauze D, et al. Health literacy in patients seeking orthopaedic care: results of the literacy in musculoskeletal problems (LIMP) project. *Iowa Orthop J.* 2015;35:187-92.

25. StatCounter Global Stats. Browser Market Share Worldwide. 2022. Retrieved August 5, 2022. https://gs.statcounter.com/

26. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. *Cyberpsychol Behav.* 2004;7(5):497-510. https://doi.org/10.1089/cpb.2004.7.497

27. Wang LW, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: application, results, and recommendations. *Res Social Adm Pharm.* 2013;9(5):503-16. doi:10.1016/j.sapharm.2012.05.009

28. Schriver KA. Readability formulas in the new millennium - ACM digital library. Accessed April 12, 2023. https://dl.acm.org/doi/pdf/10.1145/344599.344638.