

Assessing the Utility of ChatGPT in Simplifying Text Complexity of Patient Educational Materials

Rasika Sudharshan¹, Alena Shen¹, Shreya Gupta¹, Sandy Zhang-Nunes¹

1. Ophthalmology, University of Southern California (USC) Roski Eye Institute, Los Angeles, USA

Corresponding author: Rasika Sudharshan, sudharsh@usc.edu

Review began 02/09/2024

Review ended 02/26/2024

Published 03/01/2024

© Copyright 2024

Sudharshan et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Introduction: AI chatbots are being increasingly used in healthcare settings. There is growing interest in using AI to assist in patient education. Currently, extensive healthcare information is found online but is often too complex to understand. Our objective is to determine if physicians can recommend the free version of ChatGPT version 3.5 (OpenAI, San Francisco, CA, USA) for patients to simplify text from the American Academy of Ophthalmology (AAO) in English and Spanish. This version of ChatGPT was assessed in this study due to its increased accessibility across various patient populations.

Methods: Fifteen articles were chosen from AAO in both languages and simplified with ChatGPT 10 times each. The readability of original and simplified articles was assessed with the Flesch Reading Ease and Gunning Fog Index for English and Fernández Huerta, Gutiérrez, Szigriszt-Pazo, INFLESZ, and Legibilidad- μ for Spanish. Grade levels to assess readability were calculated with Flesch Kincaid Grade Level and Crawford Nivel-de-Grado. Mean, standard deviation, and two-tailed t-tests were performed to assess differences before and after simplification.

Results: Average grade levels before and after simplification were as follows: English 8.43 \pm 1.17 to 8.9 \pm 2.1 (p=0.41) and Spanish 5.3 \pm 0.34 to 4.1 \pm 1.1 (p=0.0001). Spanish articles were significantly simplified per Legibilidad- μ (p=0.003). No significant difference was noted for other scales.

Conclusions: The readability of AAO articles in English worsened without significance but significantly improved in Spanish. This may result from simpler syllable structures and a lesser overall vocabulary in Spanish. With increased testing, physicians can recommend ChatGPT for Spanish-speaking patients to improve health literacy.

Categories: Ophthalmology

Keywords: spanish-speaking patients, ophthalmology education, health literacy, readability, text simplification, ophthalmology, patient education, chatbots, chatgpt, artificial intelligence

Introduction

The Internet is the first source of medical information for almost 75% of individuals in the US [1,2]. While healthcare information is abundantly found online, it is often too complex for patients to understand for many reasons, including lower reading levels, difficulty understanding medical jargon, and limited English proficiency in our population [3-5]. Additionally, with a large population of non-native English speakers in the US, there are increased barriers to health literacy [6-8].

With the advent of chatbots like ChatGPT (OpenAI, San Francisco, CA, USA), users can access information online with greater ease. ChatGPT reportedly gets about 100 million users each week, as per its CEO, Sam Altman [9]. This AI chatbot is advertised as being able to assist with education and clarifying concepts, so it was hypothesized that this capacity could be expanded to medical concepts [10].

Current studies speculate how AI and ChatGPT can be helpful for patient education and have emphasized accessibility for use in populations with limited healthcare access [11,12]. While these studies importantly address improving access to information, our study aims to focus on better comprehension of the abundant materials already available online. As articles on medical advancements continue to get published, there is no tangible change in health literacy, as materials are often far above the average reading level of an adult in the US [3,13,14]. If patients have questions about materials they are referred to, they also do not have an immediate source for reliable answers unless they message their physician or wait until their next visit.

ChatGPT's free version, version 3.5, has been assessed here as a potential solution to this problem as it is more accessible to a larger patient population. Many patients in the US are Spanish-speaking, and it has been consistently reported that patients who have limited English proficiency are at a disadvantage in not only understanding their conditions but also in communicating with their physicians and ultimately in receiving treatment [7,8,13,15]. Any advances we can make in improving health literacy in this population with this freely available resource can potentially improve health outcomes for these patients.

How to cite this article

Sudharshan R, Shen A, Gupta S, et al. (March 01, 2024) Assessing the Utility of ChatGPT in Simplifying Text Complexity of Patient Educational Materials. Cureus 16(3): e55304. DOI 10.7759/cureus.55304

The American Academy of Ophthalmology (AAO) has already begun to address the gap in limited English proficiency by providing patient education materials readily available in Spanish on their website. Though this is a great start, these materials must be available at reading levels catered to this patient population. In 2021, it was found that 62% of Spanish-speaking patients had less than a high school diploma, and for 44% of these patients, elementary school was their highest level of education [7].

This paper discusses the potential for employing chatbot-based simplification techniques in the realm of healthcare communication. By exploring the advantages, challenges, and potential applications of this technology, we aim to shed light on its transformative role in improving health literacy and overall healthcare outcomes.

Materials And Methods

We assessed the ability of the freely available version of ChatGPT, version 3.5, to improve the readability of educational materials provided by the AAO in both English and Spanish. A preliminary prompt, "What do you do when I ask, "Can you make this text easier to read?" or "¿Puedes hacer esto más fácil de leer?" was used to confirm that ChatGPT interpreted our subsequent prompt correctly to simplify the text and make the text more comprehensible. This specific phrasing was chosen to model the natural phrasing a patient might use when interacting with ChatGPT while making the objective of improving readability clear to ChatGPT as well.

Fifteen unique educational articles on the same topics in both languages were entered into ChatGPT following the prompt, "Can you make this text easier to read?" or "¿Puedes hacer esto más fácil de leer?" for each of the English and Spanish articles, respectively. Article topics included LASIK, glaucoma, ptosis, age-related macular degeneration, black eye, astigmatism, blue light, color blindness, conjunctivitis, strabismus, dry eye, eye allergies, eyelashes, eyeglasses, and combined cataract surgery and minimally invasive glaucoma surgery. All topics selected for analysis contained a complete article in both languages on the AAO website. Each prompt was repeated 10 times per article, and uniquely generated responses were recorded.

The readability of the original and simplified articles and all ChatGPT outputs were assessed using widely recognized readability scoring formulas. Readability scales are numerical formulas that analyze the characteristics of text, including sentence length, number of words per sentence, syllable count per word, percentage of multisyllabic words, etc. These formulas vary between languages to accommodate differences in language structures. For English text, the Flesch Reading Ease (FRE) Calculator and Gunning Fog Index (GFI) were used. The FRE formula uses mean sentence length (L) and mean syllable count per word (S). FRE scores range from 0 to 100, with larger values indicating higher reading ease [16].

$$\text{FRE Score} = 206.835 - (1.015 * L - (84.6 * S))$$

The GFI formula uses sentence length (L), number of words (N), and percentage of multisyllabic words (M) to calculate readability. GFI scores range from 0 to 20, with larger numbers indicating more difficult readability. GFI values roughly correlate to the number of years of formal education needed to comprehend a given text.

$$\text{GFI Score} = (0.4 * L) + (100 * M / N)$$

Spanish readability was assessed using the Fernández Huerta (FH), Gutiérrez, and Szigriszt-Pazo (SP) Calculators and the INFLESZ and Legibilidad- μ ($L\mu$) scales [17-20]. FH, SP, INFLESZ, and $L\mu$ scores range from 0 to 100, with values closer to 0 being more difficult to read. The FH score is an adaptation of the English FRE and is calculated from the number of syllables (P) and the number of sentences (C).

$$\text{FH Score} = 206.84 - (0.60 * P) - (1.02 * F)$$

The Gutiérrez score was designed for Spanish texts using the number of characters (L), number of words (P), and number of sentences (F).

$$\text{Gutiérrez Score} = 92.5 - (9.7 * L)/P - 0.35P/F$$

The SP score uses total syllables (S), total number of words (P), and total number of sentences (C) to calculate the intelligibility of text. The INFLESZ scale is another term for the SP score.

$$\text{SP Score} = 206.835 - (62.3 * S)/P - P/F$$

The $L\mu$ score uses the total number of words (n), the average number of letters per word (\bar{x}), and the variance (σ^2) [17].

$$\mu = [n/(n - 1)](\bar{x}/\sigma^2) * 100$$

The grade levels of texts were calculated with Flesch Kincaid Grade Level (FKGL) for English articles and Crawford Nivel-de-Grado (CNG) for Spanish articles. Higher values indicate increased reading difficulty. The FKGL formula uses a mean syllable count per word (S) and a mean word count per sentence (W) [20].

$$FKGL = (0.39 * W) + (11.8 * S) - 15.39$$

The CNG index calculates reading grade level using the number of sentences per 100 words (OP) and the number of syllables per 100 words.

$$CNG \text{ Score} = -0.205OP + 0.049SP - 3.407$$

Mean, standard deviation, and two-tailed t-tests were performed to assess the difference between the mean readability score of ChatGPT output and that of the original article. All calculations were performed in Excel (Microsoft Corporation, Washington, USA).

Results

The average grade level of the 15 original English articles was 8.43±1.17 when analyzed with the FKGL formula. Following simplification on ChatGPT, the grade level increased to 8.80±1.88 (p=0.53). The average FRI score for original articles was 61.8±6.96 and decreased to 55.97±12.38 (p=0.12) following simplification. The average GFI score for original articles was 10.81±1.41, while the score for simplified articles was 10.29±1.56 (p=0.34). None of the readability formulas showed statistically significant differences between original and simplified articles in English, and readability scale results per article are shown in Figure 1 and Table 1.



FIGURE 1: Mean score of individual AAO English articles before and after ChatGPT modification for each readability scoring test

Article numbers correspond as follows: (1) LASIK (laser eye surgery), (2) What is glaucoma?, (3) What is ptosis?, (4) What is macular degeneration?, (5) What is a black eye?, (6) What is astigmatism?, (7) Should you be worried about blue light?, (8) What is color blindness?, (9) Conjunctivitis: What is pink eye?, (10) strabismus in children, (11) What is dry eye?, (12) What are eye allergies?, (13) Why are my eyelashes falling out?, (14) Eyeglasses: How to choose glasses for vision correction, and (15) combined cataracts-glaucoma surgery and MIGS

AAO: American Academy of Ophthalmology, MIGS: micro-invasive glaucoma surgery

Readability Test	Original Articles		Simplified Articles		P-value
	Mean	SD	Mean	SD	
Flesch Kincaid Grade Level	8.4	1.2	8.8	1.9	0.528
Gunning Fog Index	10.8	1.4	10.3	1.6	0.343
Flesch Reading Ease	61.8	7	56	12.4	0.122

TABLE 1: Mean scores and standard deviation of all tested AAO English articles before and after ChatGPT modification for each readability scoring test. P-values represent the difference between original and simplified means

AAO: American Academy of Ophthalmology

The average grade level of the 15 original Spanish articles was 5.3±0.34 using the CNG formula. This was reduced to 4.15±1.10 (p=0.0001) after simplification. The average scores for FH before and after simplification were 67.23±3.65 and 65.65±10.99, respectively (p=0.57). Using the Gutiérrez scale, the original articles averaged 42.54±1.68, and they averaged 41.18±4.81 after simplification (p=0.275). Notably, all texts

were rated as having normal reading difficulty, with scores around 40. The SP/INFLESZ scales yielded an average score of 62.68 ± 3.76 for the original articles and an average score of 60.61 ± 11.47 for the simplified articles ($p=0.476$). The average score on the μ scale was 55.57 ± 5.02 , which increased to 62.34 ± 8.23 after simplification ($p=0.003$). Readability scale results for articles in Spanish are shown in Figure 2 and Table 2.

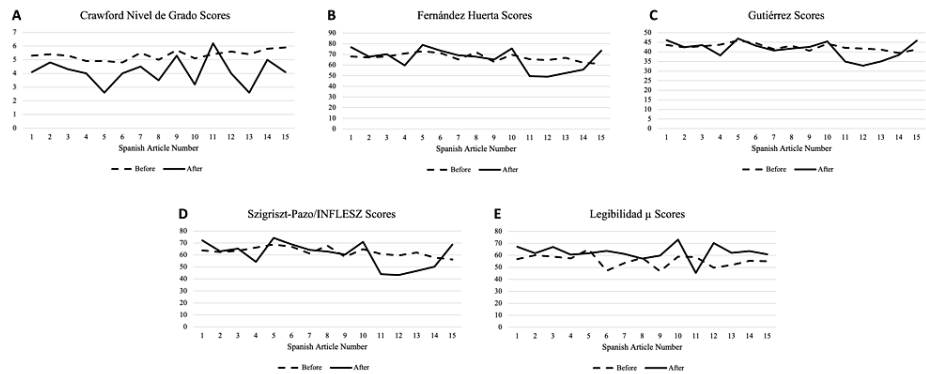


FIGURE 2: Mean score of individual AAO Spanish articles before and after ChatGPT modification for each readability scoring test

Article numbers correspond as follows: (1) LASIK (cirugía ocular con láser), (2) ¿Qué es el glaucoma?, (3) ¿Qué es la Ptoxis?, (4) ¿Qué es la degeneración macular relacionada con la edad?, (5) ¿Qué es un ojo morado?, (6) ¿Qué es el astigmatismo?, (7) ¿Debe preocuparnos el uso de la luz azul?, (8) ¿Qué es el daltonismo?, (9) Conjuntivitis: ¿Qué es el ojo rojo?, (10) Estrabismo infantil, (11) ¿Qué es el ojo seco?, (12) ¿Qué son las alergias de los ojos?, (13) ¿Por qué se caen las pestañas?, (14) Gafas: Cómo elegir gafas para la corrección de la visión, and (15) Cirugía combinada de cataratas y glaucoma y MIGS

AAO: American Academy of Ophthalmology, MIGS: micro-invasive glaucoma surgery

Readability Test	Original Articles		Simplified Articles		P-value
	Mean	SD	Mean	SD	
Crawford Nivel-de-Grado*	5.33	0.34	4.15	1.1	0.0001
Fernández Huerta	67.23	3.65	65.65	10.99	0.57
Gutiérrez	42.54	1.68	41.18	4.81	0.275
Szigriszt-Pazos/INFLESZ	62.68	3.76	60.61	11.47	0.476
Legibilidad- μ *	55.57	5.02	62.34	8.23	0.003

TABLE 2: Mean scores and standard deviation of all tested AAO Spanish articles before and after ChatGPT modification for each readability scoring test. P-values represent the difference between original and simplified means. * $p < 0.005$

AAO: American Academy of Ophthalmology

Discussion

Our results showed that readability was improved in Spanish among this subset of articles but was not improved in English. This is an interesting finding, considering ChatGPT was primarily trained in English. Per ChatGPT itself and OpenAI, this chatbot is advertised as being most proficient in English when compared to the 102 other languages it knows, including Spanish [21]. It is possible that ChatGPT's inability to significantly improve readability for English articles is due to factors like the English language's grammar and syntax, the length of words or phrases that cannot be substituted, or ChatGPT's inability to handle medical terminology. Spanish grammar, for example, does not necessitate the use of a subject at the beginning of a sentence, so removing this extra word from a sentence could alter readability in ways different from English.

Given the relatively recent development of ChatGPT in 2022, there are a limited number of studies

investigating ChatGPT's ability to simplify patient education material. Various articles test across the existing versions of ChatGPT to produce accurate responses to commonly asked patient questions. This was tested in several fields, including glaucoma in ophthalmology, general gastrointestinal disease, and interventional radiology procedures [22-24]. All studies found that ChatGPT can consistently produce accurate patient education information. These studies, paired with our study on the simplification of patient education materials, can prove useful in improving the health literacy of our patients.

Another study published by the American Urological Association compared readability between patient education materials created by urologists and responses from ChatGPT version 3.5. ChatGPT had significantly poorer readability than provider-created articles across all topics that were tested, despite being prompted to provide responses at a sixth-grade reading level (all $p < 0.001$) [25]. It is important to identify opportunities and limitations for chatbot use in patient care settings so we can maximize its impact.

It will be interesting to note if updates to ChatGPT's free platform will include more training in medical knowledge to further assist in clarifying medical concepts.

To ensure that the goal of improving readability was targeted with the final prompt, ChatGPT was asked how it modified text when given varying phrases like "summarize this text" or "rewrite this text." The wording of the phrase was also chosen carefully based on the perceived likelihood of patients using this phrase. To standardize the results, the final prompt was entered into ChatGPT 10 times, since ChatGPT produced responses with varying reading levels each time. The goal of this study was specifically to assess if ChatGPT can improve readability with one prompt, rather than comparing it across multiple prompts. In the future, more prompts can be used to test ChatGPT's ability to improve readability and perform other functions, such as summarizing text.

While our study has provided valuable perspectives on the ability to use ChatGPT in patient education, it is important to address the following limitations: We used ChatGPT version 3.5 since it is available to patients at no cost and can be recommended for patients with financial barriers. However, with many changes in ChatGPT version 4, it is important to explore the capabilities of the paid version of this tool as well. We also only analyzed a subset of articles from the AAO. These articles from the AAO are specifically designed for patient education, so they may have already started at a lower reading level than other sources of information. In the future, our study should be expanded to test a larger set of articles in different specialties and with other types of medical articles found online to validate consistency. These results should also be compared with those of other freely available chatbots to determine better chatbot recommendations for patients. In general, the average reading levels of patient education materials in English range from 7.2 to 13.4 on the FKGL scale in a 20-year analysis of English patient education materials [3]. The average grade level for the English articles used in this study was 8.4 and 5.33 for the Spanish articles. This should be further investigated, as it is possible that ChatGPT is better at improving readability in articles that start at a lower reading level. As per ChatGPT, when prompted to "make this text easier to read," the chatbot focuses on improving readability by simplifying sentence structure, removing jargon, and clarifying vocabulary. However, the quantitative readability scales used in this study may not adequately assess all the dimensions by which ChatGPT is modifying the text to improve human perception of readability. In the future, patient feedback should be incorporated to support this quantitative analysis and determine if readability has improved. Analyses of ChatGPT's responses should also be performed by ophthalmologists to ensure the accuracy of this material.

Conclusions

Using existing readability scales, our findings showed that ChatGPT version 3.5 worsened the readability of AAO articles in English without significance and significantly simplified Spanish articles based on two of five scales. While the study showed no improvements in the readability of English-language materials, it highlights a promising avenue to continue exploring the potential of using chatbots to improve healthcare communication with our patients. These findings shed light on the potential for chatbots to be used to improve health literacy among Spanish-speaking patients. With the current results from this study, ChatGPT version 3.5 can be recommended for Spanish-speaking patients to use with AAO ophthalmology articles. However, further exploration should be conducted to analyze whether advancements in chatbots will change our current recommendation or help to expand this recommendation to more patient populations and medical topics.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Rasika Sudharshan, Alena Shen, Shreya Gupta, Sandy Zhang-Nunes

Acquisition, analysis, or interpretation of data: Rasika Sudharshan, Alena Shen

Drafting of the manuscript: Rasika Sudharshan, Alena Shen, Shreya Gupta

Critical review of the manuscript for important intellectual content: Rasika Sudharshan, Alena Shen, Shreya Gupta, Sandy Zhang-Nunes

Supervision: Sandy Zhang-Nunes

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Finney Rutten LJ, Blake KD, Greenberg-Worisek AJ, Allen SV, Moser RP, Hesse BW: Online health information seeking among US adults: measuring progress toward a healthy People 2020 objective. *Public Health Rep.* 2019, 134:617-25. [10.1177/0033554919874074](https://doi.org/10.1177/0033554919874074)
2. Amante DJ, Hogan TP, Pagoto SL, English TM, Lapane KL: Access to care and use of the Internet to search for health information: results from the US National Health Interview Survey. *J Med Internet Res.* 2015, 17:e106. [10.2196/jmir.4126](https://doi.org/10.2196/jmir.4126)
3. Rooney MK, Santiago G, Perni S, et al.: Readability of patient education materials from high-impact medical journals: a 20-year analysis. *J Patient Exp.* 2021, 8:2374373521998847. [10.1177/2374373521998847](https://doi.org/10.1177/2374373521998847)
4. Talking Points About Health Literacy. (2021). <https://www.cdc.gov/healthliteracy/shareinteract/TellOthers.html>.
5. Eltorai AE, Ghanian S, Adams CA Jr, Born CT, Daniels AH: Readability of patient education materials on the american association for surgery of trauma website. *Arch Trauma Res.* 2014, 3:18161. [10.5812/at.18161](https://doi.org/10.5812/at.18161)
6. Nearly 68 Million People Spoke a Language Other Than English at Home in 2019. (2022). Accessed: October 30, 2023: <https://hispanicad.com/news/nearly-68-million-people-spoke-a-language-other-than-english-at-home-in-2019/>.
7. Christy SM, Cousin LA, Sutton SK, et al.: Characterizing health literacy among Spanish language-preferring Latinos ages 50-75. *Nurs Res.* 2021, 70:344-53. [10.1097/NNR.0000000000000519](https://doi.org/10.1097/NNR.0000000000000519)
8. Hudson S, Rikard RV, Staiculescu I, Edison K: Improving health and the bottom line: the case for health literacy. Building the case for health literacy: proceedings of a workshop. National Academies Press (US), Washington (DC); 2018.
9. ChatGPT Hits 100M Weekly Users. (2023). Accessed: January 09, 2024: <https://www.linkedin.com/news/story/chatgpt-hits-100m-weekly-users-5808204/>.
10. Ray PP: ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys Syst.* 2023, 3:121-54. [10.1016/j.iotcps.2023.04.003](https://doi.org/10.1016/j.iotcps.2023.04.003)
11. Khan RA, Jawaid M, Khan AR, Sajjad M: ChatGPT - reshaping medical education and clinical management. *Pak J Med Sci.* 2023, 39:605-7. [10.12669/pjms.39.2.7653](https://doi.org/10.12669/pjms.39.2.7653)
12. Javaid M, Haleem A, Singh RP: ChatGPT for healthcare services: an emerging stage for an innovative perspective. *BenchCouncil Trans Benchmarks Stand Eval.* 2023, 3:100-5. [10.1016/j.tbench.2023.100105](https://doi.org/10.1016/j.tbench.2023.100105)
13. Hickey KT, Masterson Creber RM, Reading M, Sciacca RR, Riga TC, Frulla AP, Casida JM: Low health literacy: implications for managing cardiac patients in practice. *Nurse Pract.* 2018, 43:49-55. [10.1097/01.NPR.0000541468.54290.49](https://doi.org/10.1097/01.NPR.0000541468.54290.49)
14. Hutchinson N, Baird GL, Garg M: Examining the reading level of internet medical information for common internal medicine diagnoses. *Am J Med.* 2016, 129:637-9. [10.1016/j.amjmed.2016.01.008](https://doi.org/10.1016/j.amjmed.2016.01.008)
15. Wilson E, Chen AH, Grumbach K, Wang F, Fernandez A: Effects of limited English proficiency and physician language on health care comprehension. *J Gen Intern Med.* 2005, 20:800-6. [10.1111/j.1525-1497.2005.0174.x](https://doi.org/10.1111/j.1525-1497.2005.0174.x)
16. Boztas N, Omur D, Ozbilgin S, Altuntas G, Piskin E, Ozkardesler S, Hanci V: Readability of internet-sourced patient education material related to "labour analgesia". *Medicine (Baltimore).* 2017, 96:8526. [10.1097/MD.00000000000008526](https://doi.org/10.1097/MD.00000000000008526)
17. Cantalejo B, Inés M: Readability and health: readability measurement methods and their application to the design of health educational brochures (Article in Spanish). Universidad Autónoma de Madrid, Madrid, Spain; 2007.
18. Readability of Fernández Huerta (Article in Spanish). (2018). Accessed: October 31, 2023: <https://legible.es/blog/lecturabilidad-fernandez-huerta/>.
19. Pazos FS: Predictive systems for legibility of the written message: perspicuity formula (Article in Spanish). Universidad Complutense de Madrid, Madrid, Spain; 2001.
20. Vásquez-Rodríguez L, Cuenca-Jiménez PM, Morales-Esquivel S, and Alva-Manchego F: A benchmark for neural readability assessment of texts in Spanish. Proceedings of the workshop on text simplification, accessibility, and readability. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates; 2022. 188-98. [10.18653/v1/2022.tsar-1.18](https://doi.org/10.18653/v1/2022.tsar-1.18)
21. ChatGPT 3.5. (2024). Accessed: January 09, 2024: <https://chat.openai.com>.
22. Scheschenja M, Viniol S, Bastian MB, Wessendorf J, König AM, Mahnken AH: Feasibility of GPT-3 and GPT-4 for in-depth patient education prior to interventional radiological procedures: a comparative analysis. *Cardiovasc Intervent Radiol.* 2024, 47:245-50. [10.1007/s00270-023-03563-2](https://doi.org/10.1007/s00270-023-03563-2)

23. Kerbage A, Kassab J, El Dahdah J, Burke CA, Achkar JP, Roupael C: Accuracy of ChatGPT in common gastrointestinal diseases: impact for patients and providers. *Clin Gastroenterol Hepatol.* 2023, [10.1016/j.cgh.2023.11.008](https://doi.org/10.1016/j.cgh.2023.11.008)
24. Lee DA, Wu G, Tien K, et al.: Can ChatGPT™, an "intelligent" chatbot, be used to educate our glaucoma patients?. *Invest Ophthalmol Vis Sci.* 2023, 64:379.
25. Shah YB, Ghosh A, Hochberg AR, Rapoport E, Lallas CD, Shah MS, Cohen SD: Comparison of ChatGPT and traditional patient education materials for men's health. *Urol Pract.* 2024, 11:87-94. [10.1097/UPJ.0000000000000490](https://doi.org/10.1097/UPJ.0000000000000490)