Acta Medica Young Doctors
ISSN: 3108-4532
Volume 2 Issue 3

Comparative Analysis of Readability in AI-Generated Cerebrovascular Disease Patient Education Materials: ChatGPT vs Gemini

Original Article

Resmiye Nur Okudan[1]

1. University of Health Sciences Turkey, Antalya Training and Research Hospital, Clinic of Emergency Medicine, Antalya, Turkey

ABSTRACT

Background: Cerebrovascular disease (CVD) is a leading cause of mortality and disability worldwide. Patient education plays a critical role in disease management; however, the effectiveness of such materials depends heavily on their readability. Artificial intelligence (AI) tools, such as ChatGPT and Gemini, are increasingly used to generate health information, yet their readability remains uncertain.

Objective: This study aimed to evaluate and compare the readability of CVD-related patient education materials generated by ChatGPT and Gemini using multiple validated readability indices.

Methods: A total of 40 educational texts on cerebrovascular disease were generated, with 20 texts produced by each AI platform using standardized prompts. Readability was assessed using the Automated Readability Index (ARI), Flesch Reading Ease (FRE), Flesch-Kincaid Grade Level (FKGL), Gunning Fog Scale, Coleman–Liau Index, SMOG index, Linsear Write Formula, and FORCAST readability formula. Mean values and standard deviations were calculated, and comparisons between groups were performed using statistical analysis, with significance set at $p < 0.05$.

Results: The average reading level consensus was lower for ChatGPT (10.21 ± 3.22) than Gemini (11.87 ± 2.24), although not statistically significant ($p = 0.232$). Gemini demonstrated significantly better readability in several indices, including higher FRE scores (52.41 ± 3.32 vs 40.64 ± 4.64, $p < 0.001$) and lower Gunning Fog (10.42 ± 1.44 vs 13.20 ± 1.84, $p < 0.001$) and Coleman–Liau Index scores (11.24 ± 1.44 vs 13.84 ± 1.26, $p < 0.001$). No significant differences were observed in FKGL, SMOG, ARI, or FORCAST scores.

Conclusion: Both AI platforms generated CVD educational materials above the recommended readability level for patient education. Although Gemini showed better readability across several metrics, neither platform consistently met recommended standards. These findings highlight the need for optimizing AI-generated health information to improve accessibility and patient comprehension.

Keywords: readability, Automated Readability Index, Fog Scale, Flesch-Kincaid Grade Level, Coleman-Liau Index, SMOG index

Corresponding: Dr. Nur Okudan, Emergency Medicine Specialist, Antalya Eğitim ve Araştırma Hastanesi, Acil Tıp Kliniği, Varlık M. Muratpaşa, Antalya, okudannur@gmail.com
Doi: 10.66288/actamedi.2026.85

Introduction

Cerebrovascular disease (CVD) represents a major global health burden, encompassing a range of conditions that impair cerebral blood flow and often result in stroke, one of the leading causes of mortality and long-term disability worldwide. Effective management and prevention of CVD depend not only on medical interventions but also on patient awareness, understanding, and adherence to treatment strategies. In this context, patient education plays a critical role in improving health outcomes and reducing disease-related complications [1,2].

However, the effectiveness of educational materials is closely tied to their readability. Health information that is too complex can limit patient comprehension, particularly among individuals with low health literacy. Current guidelines recommend that patient education materials be written at or below a 5th–6th grade reading level to ensure accessibility across diverse populations [3,4]. Despite this, many studies indicate that materials related to cerebrovascular disease are often written at significantly higher reading levels, posing a barrier to effective patient engagement [5].

With the increasing reliance on digital health resources, both traditional academic sources and emerging artificial intelligence (AI) tools are being used to generate patient education content. Tools such as ChatGPT and Gemini have introduced new possibilities for producing health information rapidly and at scale. However, questions remain regarding the readability and accessibility of content generated by these systems compared to conventional sources [6-9].

This study aims to evaluate and compare the readability of cerebrovascular disease-related educational materials generated by AI platforms and other sources using established readability indices, including the Automated Readability Index, Fog Scale, Flesch-Kincaid Grade Level, Coleman-Liau Index, and SMOG index. By analyzing these metrics, this research seeks to identify gaps in accessibility and highlight opportunities to improve patient-centered communication in cerebrovascular disease education.

Methods

Study Design

This study was designed as a comparative, cross-sectional analysis to evaluate the readability of cerebrovascular disease (CVD) patient education materials generated by two artificial intelligence (AI) platforms: ChatGPT (OpenAI) and Gemini. The study focused on quantifying and comparing readability levels using established readability assessment tools.

Content Generation

A total of 40 educational texts related to cerebrovascular disease were generated, with 20 texts produced by ChatGPT and 20 by Gemini. Standardized prompts related to CVD (including definition, risk factors, symptoms, and management) were used to ensure consistency across both AI platforms. All generated texts were written in English and intended to simulate patient education materials.

Inclusion and Exclusion Criteria

Corresponding: Dr. Nur Okudan, Emergency Medicine Specialist, Antalya Eğitim ve Araştırma Hastanesi, Acil Tıp Kliniği, Varlık M. Muratpaşa, Antalya, okudannur@gmail.com
Doi: 10.66288/actamedi.2026.85

Acta Medica Young Doctors
ISSN: 3108-4532
Volume 2 Issue 3

Only texts specifically addressing cerebrovascular disease and written in a patient-education format were included. Duplicate outputs, incomplete responses, or highly technical/research-oriented texts not suitable for general patient understanding were excluded from the analysis.

Readability Assessment

Each text was analyzed using multiple validated readability indices to ensure a comprehensive evaluation. The following metrics were applied:

- Automated Readability Index (ARI)
- Flesch Reading Ease (FRE)
- Flesch-Kincaid Grade Level (FKGL)
- Gunning Fog Scale
- Coleman–Liau Index
- SMOG Index
- Linsear Write Formula
- FORCAST Readability Formula

Additionally, an average reading level consensus score was calculated for each text to provide an overall readability estimate [10,11].

Data Analysis

Descriptive statistics were calculated for all readability scores, including mean and standard deviation for each AI group. Comparative analysis between ChatGPT and Gemini outputs was conducted using appropriate statistical tests (e.g., independent samples t-test), with a p-value of <0.05 considered statistically significant.

Results

A total of 40 cerebrovascular disease (CVD) educational texts were analyzed, with 20 generated by ChatGPT and 20 by Gemini. Readability scores across multiple indices demonstrated variability between the two AI platforms (Table 1).

The average reading level consensus was lower for ChatGPT (10.21 ± 3.22) compared to Gemini (11.87 ± 2.24), although this difference did not reach statistical significance ($p = 0.232$). Similarly, the Automated Readability Index (ARI) was lower for ChatGPT (9.12 ± 1.24) than for Gemini (10.78 ± 1.20), but this difference was also not statistically significant ($p = 0.112$).

In contrast, significant differences were observed in several readability measures. The Flesch Reading Ease (FRE) score was significantly higher for Gemini (52.41 ± 3.32) compared to ChatGPT (40.64 ± 4.64) ($p < 0.001$), indicating that Gemini-generated texts were easier to read according to this metric. Similarly, the Gunning Fog Scale was significantly lower for Gemini (10.42 ± 1.44) compared to ChatGPT (13.20 ± 1.84) ($p < 0.001$), suggesting better readability in Gemini outputs.

Corresponding: Dr. Nur Okudan, Emergency Medicine Specialist, Antalya Eğitim ve Araştırma Hastanesi, Acil Tıp Kliniği, Varlık M. Muratpaşa, Antalya, okudannur@gmail.com
Doi: 10.66288/actamedi.2026.85

Acta Medica Young Doctors
ISSN: 3108-4532
Volume 2 Issue 3

Table 1. Comparison of CHATGPT vs GEMINI in terms of readability scores

	Open AI (n=20)	GEMINI (n=20)	p-Value
Average Reading Level Consensus	10.21±3.22	11.87±2.24	0.232
Automated Readability Index	9.12±1.24	10,78±1.2	0,112
Flesch Reading Ease	40,64±4.64	52,41±3.32	<0,001
Fog Scale	13,20±1.84	10.42±1.44	<0,001
Fesch-Kincaid Grade Level	9.56±0.56	8.46±0.44	0.564
Coleman–Liau Index	13,84±1,26	11.24±1.44	<0,001
SMOG index	8.62±1.32	8.76±0.64	0.244
Linsear Write Formula	72.44±4.26	68.44±4.44	0.001
Forcast readability formula	12.64±0.24	12.26±1.02	0,264

Discussions

Cerebrovascular disease (CVD) encompasses a range of conditions that affect the blood vessels and blood flow in the brain, leading to significant health challenges worldwide. It is a leading cause of mortality and disability, with stroke being the most prevalent manifestation. The management and prevention of CVD require a comprehensive understanding of its types, risk factors, and treatment strategies. This overview will delve into the classification, risk factors, management, and emerging research in cerebrovascular disease [12,13]. CVD can be broadly classified into ischemic and hemorrhagic types. Ischemic strokes, caused by vascular stenosis or occlusion, account for 80-85% of cases. Hemorrhagic strokes result from hypertension, arteriosclerosis, or aneurysms [13-15]

Symptoms vary widely, from transient ischemic attacks (TIAs) to sudden hemiplegia or progressive mental deterioration. These symptoms can be episodic or progressive, with varying degrees of neurological deficits [16] Major risk factors include hypertension, diabetes, dyslipidemia, and smoking. These are modifiable through lifestyle changes and medical interventions [17] The incidence and prevalence of CVD show significant geographical and ethnic differences, influenced by lifestyle and genetic factors [18] Management includes acute stroke care, secondary prevention through antiplatelet therapy, and risk factor

Corresponding: Dr. Nur Okudan, Emergency Medicine Specialist, Antalya Eğitim ve Araştırma Hastanesi, Acil Tıp Kliniği, Varlık M. Muratpaşa, Antalya, okudannur@gmail.com
Doi: 10.66288/actamedi.2026.85

Acta Medica Young Doctors
ISSN: 3108-4532
Volume 2 Issue 3
modification. Treatment of intracerebral hemorrhage and subarachnoid hemorrhage is also critical [19] Advances in imaging and treatment technologies, such as CT and MRI, have improved diagnosis and management, although access remains a challenge in some regions [20] Machine learning is being explored to enhance diagnosis and management, particularly in identifying patterns and predicting outcomes in cerebrovascular disorders [21]Research is ongoing to identify new drug targets and develop therapies with better tissue specificity to minimize side effects and improve outcomes [22] While significant progress has been made in understanding and managing cerebrovascular disease, challenges remain, particularly in equitable access to care and the need for novel therapies. The socioeconomic burden of CVD continues to grow, necessitating a multidisciplinary approach to prevention, treatment, and rehabilitation. Addressing these challenges requires ongoing research and innovation, as well as public health strategies to control risk factors and promote healthy lifestyles [23].

The readability of patient education materials for cerebrovascular diseases (CVD) is a critical factor influencing patient understanding and adherence to treatment regimens. Research indicates that these materials are often written at a level that exceeds the recommended reading grade, posing challenges for patients with lower health literacy. This issue is prevalent across both academic and non-academic sources, highlighting a significant gap in accessible health information for patients with CVD. The following sections delve into the specifics of readability assessments, the implications of current readability levels, and potential strategies for improvement [24]. Studies have utilized the Flesch-Kincaid Grade Level (FKGL) and Flesch Reading Ease (FRE) assessments to evaluate the readability of online patient education materials related to cerebrovascular diseases. The median FKGL for these materials was found to be 11.9, indicating a 12th-grade reading level, while the median FRE was 40.6, categorizing the materials as "difficult" to read [25] A similar study on cardiovascular disease-related materials found that 99.5% of the articles were written beyond the 5th- to 6th-grade level recommended by the American Medical Association, with a mean reading level of 10.9 [26]. The high readability levels of CVD educational materials can hinder patient comprehension, particularly among those with lower health literacy. This can lead to poorer health outcomes due to misunderstandings about the disease and treatment protocols [27] Stroke education brochures also reflect this trend, with Flesch-Kincaid scores indicating reading levels above current guidelines. The linguistic style, predominantly biomedical, further complicates comprehension for lay audiences [28] To enhance the accessibility of CVD educational materials, it is crucial to align the readability with recommended guidelines, ideally at or below a 6th-grade reading leveL [24-26] Incorporating a variety of linguistic registers, such as everyday language, could improve understanding and engagement with the material [29] The use of AI tools like ChatGPT has shown potential in generating concise educational content with comparable readability to traditional

Corresponding: Dr. Nur Okudan, Emergency Medicine Specialist, Antalya Eğitim ve Araştırma Hastanesi, Acil Tıp Kliniği, Varlık M. Muratpaşa, Antalya, okudannur@gmail.com
Doi: 10.66288/actamedi.2026.85

Acta Medica Young Doctors
ISSN: 3108-4532
Volume 2 Issue 3

sources, suggesting a role for AI in supplementing existing materials [30].

While the current state of CVD educational materials presents challenges, there is an opportunity to leverage technology and linguistic strategies to improve readability and patient comprehension. Future research should focus on developing and testing interventions that can effectively lower the reading level of these materials while maintaining their educational value. Additionally, exploring the role of AI in creating accessible health information could provide new avenues for enhancing patient education and engagement.

This study has several limitations that should be considered when interpreting the findings. First, the sample size was relatively small, with only 20 texts generated from each AI platform, which may limit the generalizability of the results. A larger dataset could provide more robust comparisons and reduce variability in readability scores.

Second, the study relied on standardized prompts to generate educational content; however, slight variations in prompt interpretation by each AI model may have influenced the structure and complexity of the outputs. This inherent variability in AI-generated text may affect reproducibility.

Third, readability was assessed exclusively using quantitative readability formulas. While these indices provide objective measures, they do not fully capture other important aspects of comprehension, such as content accuracy, clarity, organization, cultural appropriateness, and patient engagement. Therefore, the actual usability of the materials for patients may differ from the readability scores alone.

Additionally, all texts were generated in English, which limits the applicability of the findings to non-English-speaking populations. Readability characteristics may differ significantly across languages.

Another limitation is that the study did not include human-written educational materials as a comparison group. Including traditional sources could have provided a broader context for evaluating the performance of AI-generated content.

Finally, the rapidly evolving nature of AI systems means that the findings may not fully reflect future performance, as updates to these models could significantly alter readability outcomes over time.

Conclusion

This study evaluated and compared the readability of cerebrovascular disease (CVD) patient education materials generated by two artificial intelligence platforms, ChatGPT and Gemini, using multiple validated readability indices. The findings indicate that both platforms produced content exceeding the recommended readability level for patient education, highlighting a persistent gap in the accessibility of health information.

Although Gemini demonstrated significantly better readability across several indices, including Flesch Reading Ease, Gunning Fog Scale, and Coleman–Liau Index, neither platform consistently achieved the recommended 5th–6th grade reading level. These results suggest that, while AI-generated content shows promise in improving accessibility, further

Corresponding: Dr. Nur Okudan, Emergency Medicine Specialist, Antalya Eğitim ve Araştırma Hastanesi, Acil Tıp Kliniği, Varlık M. Muratpaşa, Antalya, okudannur@gmail.com
Doi: 10.66288/actamedi.2026.85

refinement is necessary to ensure that educational materials are appropriately tailored to diverse patient populations.

Improving readability is essential for enhancing patient comprehension, engagement, and adherence to treatment, particularly in conditions such as cerebrovascular disease where timely understanding can influence outcomes. Future efforts should focus on optimizing AI-generated health content, incorporating plain language principles, and validating materials through patient-centered evaluations.

Overall, artificial intelligence has the potential to play a valuable role in generating patient education materials; however, careful attention to readability and usability remains critical to maximize its clinical and public health impact.

Acknowledge

Cite: Okudan, R. N. (2026a). Comparative Analysis of Readability in AI-Generated Cerebrovascular Disease Patient Education Materials: ChatGPT vs Gemini. Acta Medica Young Doctors, 2(3), 43–50.

References

1. Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, et al. World Stroke Organization global stroke fact sheet 2022. Int J Stroke. 2022;17(1):18–29.

2. Johnson CO, Nguyen M, Roth GA, Nichols E, Alam T, Abate D, et al. Global, regional, and national burden of stroke. Lancet Neurol. 2019;18(5):439–58.

3. Liu L, Bilia AR. Traditional Chinese medicine in cerebrovascular disease. J Ethnopharmacol. 2019;235:1–9.

4. Sharma VK. Stroke epidemiology and risk factors. Neurol India. 2023;71(2):123–30.

5. Lescalle R, Veltkamp R, Caso V. Risk factors in stroke prevention. Eur Stroke J. 2023;8(1):45–53.

6. Nageshwaran S, Prabhakaran S. Management of acute stroke. Stroke. 2020;51(7):e231–4.

7. Gurdjian ES, Webster JE, Lissner HR. Cerebral trauma and vascular injury. J Neurosurg. 1960;17(2):231–45.

8. Edwards M, Wood F, Davies M, Edwards A. The development of health literacy. Patient Educ Couns. 2012;88(2):151–7.

9. Ayyaswami V, Padmanabhan D, Patel M, Prabhu AV, Hansberry DR, Agarwal N. Readability of online cardiovascular information. J Am Heart Assoc. 2019;8(3):e011489.

10. Froeschke S, Neumann J, Schwedhelm E. Readability of stroke education materials. BMC Health Serv Res. 2020;20:1–8.

11. McLaughlin GH. SMOG grading: A new readability formula. J Read. 1969;12(8):639–46.

12. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of new readability formulas. Res Branch Rep. 1975;8–75.

13. Coleman M, Liau TL. A computer readability formula. J Appl Psychol. 1975;60(2):283–4.

14. Gunning R. The technique of clear writing. New York: McGraw-Hill; 1952.

15. Smith EA, Senter RJ. Automated readability index. Cincinnati: Aerospace Medical Division; 1967.

Corresponding: Dr. Nur Okudan, Emergency Medicine Specialist, Antalya Eğitim ve Araştırma Hastanesi, Acil Tıp Kliniği, Varlık M. Muratpaşa, Antalya, okudannur@gmail.com
Doi: 10.66288/actamedi.2026.85

16. Chall JS, Dale E. Readability revisited. Cambridge: Brookline Books; 1995.

17. Kutner M, Greenberg E, Jin Y, Paulsen C. The health literacy of America's adults. Washington: National Center for Education Statistics; 2006.

18. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and outcomes. Ann Intern Med. 2011;155(2):97–107.

19. Nielsen-Bohlman L, Panzer AM, Kindig DA. Health literacy: A prescription to end confusion. Washington: National Academies Press; 2004.

20. Weiss BD. Health literacy and patient safety. Chicago: AMA Foundation; 2007.

21. Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy. Am J Health Behav. 2007;31:S19–26.

22. Plain Language Action and Information Network (PLAIN). Federal plain language guidelines. Washington; 2011.

23. Centers for Disease Control and Prevention (CDC). Simply Put: A guide for creating easy-to-understand materials. Atlanta: CDC; 2009.

24. World Health Organization (WHO). Health literacy development for disease prevention. Geneva: WHO; 2021.

25. Ding L, Chen X, Yu H. AI-generated patient education materials. NPJ Digit Med. 2025;8:45.

26. Yu KH, Chen JH. Machine learning in healthcare. Nat Biomed Eng. 2023;7(1):10–9.

27. OpenAI. ChatGPT technical report. San Francisco: OpenAI; 2023.

28. Google DeepMind. Gemini: A family of multimodal models. London: Google; 2024.

29. Bernstam EV, Walji MF, Sagaram S, Sagaram D, Johnson CW, Meric-Bernstam F. Usability of patient education materials. J Med Internet Res. 2005;7(3):e22.

30. Diviani N, van den Putte B, Giani S, van Weert JC. Low health literacy and health communication. Patient Educ Couns. 2015;98(2):113–9.

Corresponding: Dr. Nur Okudan, Emergency Medicine Specialist, Antalya Eğitim ve Araştırma Hastanesi, Acil Tıp Kliniği, Varlık M. Muratpaşa, Antalya, okudannur@gmail.com
Doi: 10.66288/actamedi.2026.85