Zortuk *et. al.,* Babcock Univ. Med. J.2025 8(2):428-434
https://doi.org/10.38029/babcockuniv.med.j..v8i2.1085
eISSN: 2756-4657

RESEARCH ARTICLE OPEN ACCESS

Comparative analysis of ChatGPT and Gemini responses on epistaxis: Accuracy and readability

Zortuk O[1ID], Bedel C[2ID], Selvi F[2ID]

[1]Department of Emergency Medicine, Bandırma State Hospital, Balıkesir, Turkey
[2]Department of Emergency Medicine, Health Science University Antalya Training and Research Hospital, Antalya, Turkey

Submitted:31st July 2025
Accepted: 15th August 2025
Published: 31st December 2025
ID: Orcid ID

Abstract
Objective: There is a dearth of literature addressing the utilisation of AI models in patients with epistaxis, and ambiguities exist in the responses these models provide to patient-generated inquiries. This study aimed to evaluate and compare the accuracy and readability of responses regarding frequently asked questions (FAQs) about epistaxis provided by two advanced artificial intelligence (AI) models: ChatGPT-4 Pro and Gemini 2.5 Pro.
Methods: A total of 30 commonly asked questions about epistaxis were retrieved from the publicly accessible Quora platform and submitted separately to ChatGPT and Gemini. Two independent medical experts evaluated the AI-generated responses on a 5-point scale, focusing on accuracy and comprehensibility. Readability was assessed using multiple indices, including the Flesch Reading Ease, Gunning Fog, and SMOG indices, among others. Statistical analyses, including interobserver agreement and t-tests, were conducted using SPSS v27.
Results: The mean evaluation scores from the two observers were 4.18 ± 0.85 and 4.01 ± 0.83, respectively, with excellent interobserver agreement (ICC = 0.877, $p < 0.001$). ChatGPT scored slightly higher (4.18 ± 0.66) than Gemini (4.01 ± 0.91), though the difference was not statistically significant ($p = 0.179$). When readability metrics for artificial intelligence were compared, there was no difference in all parameters except Linear Write Grade Level Formula: 14.17±4.56 vs. 10.35±3.77, $p<<0.001$.
Conclusion: Both ChatGPT and Gemini provided highly accurate and readable responses to questions about epistaxis. These results indicate that AI-based tools can effectively support patient education and clinical communication. However, attention to content readability and regular evaluation is still necessary.

Keywords: Chat-GPT, GEMINI, Readability, Epistaxis, AI in healthcare, Patient education

Plain English Summary

The consistency and readability features of contemporary artificial intelligence applications are undergoing enhancement. While no significant difference was observed in the success rates between ChatGPT and Gemini, ChatGPT was found to be significantly more effective in the Linear Write Grade Level comparison.

Introduction

Epistaxis, also known as a nosebleed, is a common condition frequently encountered in emergency and primary care settings. The severity of these symptoms can range from minor, self-limiting episodes to severe cases that necessitate medical intervention (1). The management of epistaxis necessitates a comprehensive understanding of its aetiology, the implementation of initial interventions, and the consideration of preventive measures to mitigate recurrence. The condition is classified as either

Correspondence:
Cihan Bedel
Department of Emergency Medicine
Health Science University Antalya Training and Research Hospital, Muratpaşa
Antalya, Turkey
+905075641254, cihanbedel@hotmail.com

© BUMJ. 2025 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

anterior or posterior, with anterior being more prevalent and typically less severe (2).

The utilisation of artificial intelligence (AI) holds considerable promise in enhancing the management of epistaxis (3). While extant literature has predominantly centred on conventional management strategies, the integration of AI has the potential to unveil novel avenues for enhancing diagnostic accuracy, risk assessment, and treatment outcomes (4). The utilisation of artificial intelligence (AI) in healthcare settings holds considerable potential for enhancing the efficacy of risk assessment and decision-making processes. By leveraging advanced analytical capabilities, AI can facilitate the interpretation of voluminous electronic health record datasets, thereby contributing to more informed and expeditious decision-making. This approach has been successfully applied to the prediction of gastrointestinal bleeding, where machine learning algorithms have been shown to outperform traditional risk scores (5). The development of analogous applications for epistaxis could facilitate the prediction of bleeding risks and inform treatment decisions. The utilisation of artificial intelligence has the potential to standardise care by providing evidence-based recommendations and treatment algorithms. This could reduce variability in management practices and improve patient outcomes (6). Artificial intelligence-driven educational tools, such as computer-assisted instruction modules, are as effective as traditional seminars in teaching medical students about epistaxis management (7). These tools have the potential to be further developed to include AI-driven simulations and decision-making exercises. The integration of artificial intelligence (AI) into the development of patient education materials has shown promise in enhancing the readability of these materials, a crucial factor in improving patient comprehension and health literacy. AI tools such as ChatGPT have been employed to streamline complex medical information, thereby rendering it more accessible to a broader audience. However, the effectiveness of AI in achieving optimal readability is mixed. Some studies have highlighted significant improvements, while others have pointed out challenges in maintaining accuracy and comprehensiveness (8). The utilisation of AI models has been demonstrated to be effective in the simplification of patient education materials (PEMs), thereby enhancing their readability. For instance, ChatGPT attained 100% post-simplification readability scores in certain instances, signifying a substantial enhancement in the comprehension of medical information (7). In the context of emergency medical conditions, a comparative analysis was conducted to assess the readability of AI-generated brochures by ChatGPT and Google Gemini. The findings revealed that the readability of Gemini's brochures was marginally superior (8). The primary outcomes of interest were the comprehensiveness, accuracy and readability of the suggestions generated by ChatGPT to inform patients about epistaxis. The secondary objective of this exploratory study was to conduct a qualitative analysis of the AI-generated suggestions to identify potential patterns, consistencies, and gaps. As this technology is still in its infancy, the study is exploratory and does not test any pre-existing hypotheses. The existing literature contains a paucity of studies that address the readability and comparison of answers provided by artificial intelligence program queries. The impetus for conducting this study stems from the dearth of literature addressing the utilisation of AI models in patients with epistaxis and the ambiguity of the responses these models provide to patient-generated inquiries concerning the condition.

Materials and Methods

The present study employed artificial intelligence programs. The authors have no affiliation, communication, or financial interest with the companies associated with the program. The study population was determined by compiling answers to frequently asked questions about epistaxis from patients and their families on the publicly accessible Quora website, which was used for educational purposes only (What is meant by epistaxis?, What is the first aid for epistaxis?, Which condition has symptoms of frequent epistaxis bruising and prolonged bleeding?.....etc.) Following the exclusion of duplicate and non-medical inquiries, the study population was ascertained.

Questions that were repetitive, irrelevant, subjective, or unclear were excluded from the analysis. The question-focused approach strengthened the meaningfulness and validity of the collected data by ensuring that the dataset consisted only of questions directly related to the topic. Repetitive, irrelevant, contrary to the general approach, and personalised questions were excluded from the study. Consequently, 30 questions were identified and posed separately to two AI models: The present study will examine the ChatGPT-4 Pro and Gemini 2.5 Pro.

Observer Scoring

The responses furnished by the two AI models were subjected to a rigorous academic evaluation by two independent medical doctors. Each response was evaluated on a scale of 1 to 5, with the following criteria:

a) 5 points awarded for accuracy and comprehensibility for non-specialists.

b) 4 points granted for academic precision, but impracticality for the general public.
c) 3 points allotted for accuracy, but with incomplete information.
d) 2 points designated for a combination of accurate and inaccurate/misleading data; and
e) 1 point assigned for total inaccuracy.

Readability Score

These responses were then assessed using several criteria. The Flesch-Kincaid Calculator was utilised to measure word count, sentence count, readability, and ease of understanding [8].To evaluate the comprehensibility of the responses obtained in our study, the Average Reading Level Consensus (ARLC) method was employed to analyse the mean reading value. This value was calculated and documented using the means of other scales. The initial method is the Flesch reading-ease score (FRES) readability test, which is scored on a scale of 0-100, with higher scores indicating superior readability. Formula: The formula for calculating the number of words per sentence is as follows: 206.835 × (total words / total sentences) × (84.6 × (total syllables / total words)) (10). Another evaluation tool is the Automated Readability Index (ARI), which is calculated using the following formula: ARI = 4.71 × (characters/words) + 0.5 × (words/sentences) – 21.43. Notably, the ARI is considered a particularly efficient formula due to its ability to facilitate rapid and straightforward calculation (11). The Gunning Fog Index (GFI) is a readability formula designed to estimate the number of years of formal education required to understand a text on first reading. This method is particularly useful for assessing the complexity of English texts by considering sentence length and the number of complex words. The index has found widespread application in various fields, including education, healthcare, and translation studies, to assess the accessibility of written materials. The GFI is calculated using the following formula: The GFI is calculated as follows: 0.4 × (words/sentences + 100 × complex words/words). A higher GFI score indicates a less readable text, suggesting that more years of education are required to understand it (12). The Coleman-Liau Index is a formula employed to evaluate the readability of English texts. It is one of several traditional formulas of this kind developed in the 20th century that are still widely used in various applications, including natural language processing (NLP) and text simplification studies (13). The index is held in high esteem for its simplicity and ease of use in evaluating the complexity of written material. The program utilises an algorithm to calculate readability based on the number of characters, words, and sentences in a text. This calculation provides an estimate of the U.S. school grade level required to understand the text. The formula is as follows: The following equation is used to calculate the correlation coefficient (CLI): 0.0588 × L – 0.296 × S – 15.8, where L is the average number of letters per 100 words and S is the average number of sentences per 100 words (13). The Simple Measure of Gobbledygook (SMOG) is a readability formula designed to estimate the years of education needed to understand a piece of writing. This approach is particularly useful in assessing the complexity of texts to ensure they are accessible to the intended audience. The SMOG formula is frequently employed in various disciplines, including political communication and health literacy, to assess the readability of written materials. The SMOG formula is a linguistic analysis tool that calculates the reading level of a text sample by enumerating the number of polysyllabic words (words with three or more syllables) contained therein. The formula is applied to a sample of 30 sentences, and the number of polysyllabic words is used to estimate the reading grade level. The formula is as follows: The SMOG grade is determined by the sum of the squares of the polysyllabic words in 30 sentences, subsequently multiplied by 3 and then raised to the power of 1.4. The FORCAST readability formula is a tool designed to assess the readability of technical job materials by predicting the reading level required for comprehension. The method was developed by John S. Caylor and Thomas G. Sticht. It is distinguished by its simplicity and effectiveness. It demonstrates a high correlation with other well-established readability indices, such as the Flesch and Dale-Chall formulas (14). The FORCAST formula's emphasis on technical materials positions it as a valuable tool for the evaluation of job-related texts. The values correspond to the number of years of education required for a reader to comprehend a text. For instance, the age of 11–12 years corresponds to the level of reading proficiency exhibited by individuals in the 17–18 age range (15).

Statistical Analysis

The collected data were entered into a database and subsequently analysed using SPSS version 27. The data were subsequently classified. Categorical data were defined as percentages and frequencies, while the distribution analysis of numerical data was performed. Data exhibiting a normal distribution were characterised by their mean and standard deviation. An interobserver test was conducted among observers to perform a consistency analysis. To assess the relationship between numerical values, a t-test

was employed, and data with a p-value below 0.05 were deemed statistically significant.

Results

The results obtained from 30 questions evaluated by two different experts showed that the average score of the first expert was 4.18±0.85, while the average response score of the second expert was 4.01±0.83, and good agreement was observed between them (Interclass correlation=0.877, p<0.001). Table 1 shows the agreement analysis.

Table 1: Inter-expert agreement analysis

	Score	Interclass Correlation
Observer-1	4.18±0.85	0.877
Observer-2	4.01±0.83	

ICC > 0.75 indicates excellent agreement

The mean score for Chat GPT was 4.18 ± 0.66, while the mean score for GEMINI was 4.01 ± 0.91 (p = 0.179). As illustrated in Figure 1, a comparison of the scores reveals a significant disparity. The results of the readability comparison are presented in Table 2.

Table 2: Readability comparison

	ChatGPT	Gemini	p-Value
Average Reading Level Consensus	11.26±2.37	11.19±2.24	0.411
Automated Readability Index	11.54±3.15	11.80±3.69	0.853
Flesch Reading Ease	45.83±17.65	48.80±15.89	0.499
Gunning Fog Index	11.78±3.21	11.83±3.38	0.787
Flesch-Kincaid Grade Level	10.09±3.13	10.69±2.84	0.286
Coleman-Liau Readability Index	12.60±2.41	12.14±2.80	0.81
The SMOG Index	9.67±2.73	9.88±2.37	0.216
Original Linear Write Formula	71.60±10.27	68.70±9.15	0.929
Linear Write Grade Level Formula	14.17±4.56	10.35±3.77	<0.001*
FORCAST Readability Formula	11.79±1.38	11.29±1.04	0.104

*2-tailed t-test; For Flesch Reading Ease, higher scores indicate better readability; for other indices, lower scores indicate better readability

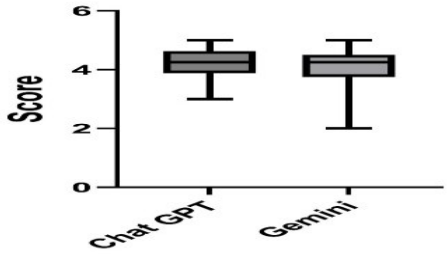


Figure 1: Comparison of median scores

Discussion

ChatGPT GEMINI signifies the amalgamation of ChatGPT, a conversational AI model, with the GEMINI project, which prioritises the development of multilingual, multimodal dialogue interfaces. OpenAI developed ChatGPT, a sophisticated language model that can produce text responses that are indistinguishable from those of human beings (16). This renders it suitable for a variety of applications, including customer service, education, and scientific writing. The objective of the GEMINI project is to develop a flexible platform for interactive dialogue interfaces that can be enhanced by incorporating ChatGPT's capabilities (17). Artificial intelligence-generated content frequently exhibits substandard performance in terms of traditional readability indices, such as the Flesch–Kincaid grade level and the Gunning Fog index. This finding suggests that the content is more challenging than the average adult literacy level (9). Nevertheless, a comparison of AI-generated and human-authored educational content reveals that AI can produce materials that are read faster and with better comprehension, despite comparable quality and interest levels (18).

The availability of accessible online information regarding epistaxis, also known as nosebleeds, is imperative for the effective education and management of patients. Epistaxis is a prevalent otolaryngological emergency, underscoring the imperative for both patients and healthcare providers to comprehend its aetiology, management strategies, and the potential

complications it may entail (19, 20, 21). The quality and accessibility of information can have a substantial impact on patient outcomes, particularly in emergencies. This article explores the readability of online epistaxis resources and the importance of clear communication in managing this condition (22). Plain language summaries, such as those provided by Tunkel et al., aim to make medical information more accessible to the general public, ensuring that patients of all ages can understand the basics of epistaxis and its management (21).

A comprehensive understanding of nasal anatomy is paramount for effective epistaxis management, which necessitates a meticulous awareness of the potential risks and complications associated with treatment. This knowledge is crucial for healthcare providers and patients alike, as it ensures effective management and prevents adverse outcomes (23). The management of epistaxis can range from simple interventions, such as nasal compression and cauterisation, to more complex procedures like arterial embolisation or surgical intervention. It is imperative that patients have access to clear communication and a comprehensive understanding of the procedures they are to undergo. This is essential for ensuring patient compliance and achieving optimal outcomes (24). Despite the abundance of online resources available, the intricacy of medical terminology and the inconsistency in the quality of information can present significant challenges for patients attempting to comprehend their condition. This underscores the importance of providing reliable, clear, and appropriate information to patients and their families (25). The variability in the presentation and severity of epistaxis, ranging from benign anterior bleeds to potentially life-threatening posterior bleeds, necessitates tailored communication strategies to address the specific needs and concerns of patients (26).

ChatGPT is a transformer-based model that has been demonstrated to generate human-like text with a high degree of efficacy. This renders it suitable for a variety of applications, including chatbots, virtual assistants, and language translation systems (27). Scientific Writing and Research: ChatGPT can facilitate scientific writing processes by producing drafts, summarising articles, and translating content. However, its utilisation necessitates regulatory oversight, considering ethical considerations (28). ChatGPT has the potential to enhance communicative competence in English as a second language by facilitating constructive communication in a chat-based environment (29). The GEMINI project is centred on the development of user-friendly dialogue interfaces that support multiple languages and modes of interaction (30, 31). Integration of ChatGPT with GEMINI enables the project to capitalise on ChatGPT's natural language processing capabilities, thereby ensuring the provision of more lifelike and contextually relevant responses in multilingual and multimodal settings (32). While the potential of ChatGPT to enhance user interaction is evident, challenges such as ensuring accuracy, avoiding misinformation, and maintaining ethical standards must be addressed (33). When readability metrics for artificial intelligence were compared, there was no difference in all parameters except Linear Write Grade Level Formula: 14.17±4.56 vs. 10.35±3.77, $p<<0.001$. We believe that these two programs can be used by the patient for informational purposes by keeping them easily stored, as there are significant differences in readability.

The study's most significant limitations are that, although these artificial models can produce reasonable summaries using leading internet data and are constantly evolving, we acknowledge that the answers provided by ChatGPT and GEMINI regarding issues of accuracy and reliability in academic medical publishing may evolve. The article also discusses the potential for bias in AI-powered systems like ChatGPT. It is submitted that the addition of further diseases would have served to enhance the clarity of the text. It is important to note that chatbots are frequently updated, which can result in the deployment of outdated AI versions. In certain instances, the capacity of AI to deliver updated content in the domain of medical science may be hindered by the challenges associated with accessing such information. It explains that human biases from the source data may be carried over into the word embeddings used in training. This can lead to greater biases in the outputs produced.

Conclusion

A comparative analysis of Gemini and ChatGPT revealed no substantial disparities in their ability to provide accurate responses and the comprehensibility of their answers concerning epistaxis.

List of Abbreviations

AI: Artificial Intelligence
ARI: Automated Readability Index
ARLC: Average Reading Level Consensus
CLI: Correlation coefficient
FRES: Flesch reading-ease score
GFI: Gunning Fox Index
NLP: Natural language processing
PEM: Patient education material
SMOG: Simple Measure of Gobbledygook

Declarations

Ethics approval and consent to participate

This study does not involve any human subjects. Therefore, it does not require ethical review.

Consent for Publication

All the authors gave consent for the publication of the work under the Creative Commons Attribution Non-Commercial 4.0 license.

Availability of Data

Access to the data can be obtained by contacting the author.

Conflicts of Interest

The authors declare that they have no conflicts of interest with each other or with other institutions.

Funding

No financial support was provided for this study.

Authors' contributions

ÖZ: Data analysis, writing
CB: İnvestigation, proof reading
FS: Conceptualization, Literature analysis

Acknowledgments

Nil.

References

1. Janapala RN, Tran QK, Patel J, Mehta E, Pourmand A. Efficacy of topical tranexamic acid in epistaxis: a systematic review and meta-analysis. The American journal of emergency medicine. 2022 Jan 1;51:169-75. https://doi.org/10.1016/j.ajem.2021.10.043
2. Ahmed J, Leopard D, Teasdale A, Cheah W, Marnane C. Management of epistaxis: a guide for junior doctors. British Journal of Hospital Medicine. 2021 Jul 2;82(7):1-8. https://doi.org/10.12968/hmed.2021.0071
3. Mukherjee J, Sharma R, Dutta P, Bhunia B. Artificial intelligence in healthcare: a mastery. Biotechnology and Genetic Engineering Reviews. 2024 Nov 1;40(3):1659-708. https://doi.org/10.1080/02648725.2023.2196476
4. Su TH, Wu CH, Kao JH. Artificial intelligence in precision medicine in hepatology. Journal of Gastroenterology and Hepatology. 2021 Mar;36(3):569-80. https://doi.org/10.1111/jgh.15415
5. Dao HV, Nguyen BP, Nguyen TT, Lam HN, Nguyen TT, Dang TT, Hoang LB, Le HQ, Dao LV. Application of artificial intelligence in gastrointestinal endoscopy in Vietnam: a narrative review. Therapeutic Advances in Gastrointestinal Endoscopy. 2024 Dec;17:26317745241306562. https://doi.org/10.1177/26317745241306562
6. Johnson R, Li MM, Noori A, Queen O, Zitnik M. Graph artificial intelligence in medicine. Annual Review of Biomedical Data Science. 2024 Aug 23;7(2024):345-68. https://doi.org/10.1146/annurev-biodatasci-110723-024625
7. Nasra M, Jaffri R, Pavlin-Premrl D, Kok HK, Khabaza A, Barras C, Slater LA, Yazdabadi A, Moore J, Russell J, Smith P. Can artificial intelligence improve patient educational material readability? A systematic review and narrative synthesis. Internal Medicine Journal. 2025 Jan;55(1):20-34. https://doi.org/10.1111/imj.16607
8. Adithya S, Aggarwal S, Sridhar J, Kavya VS, John VP, Singh C. An observational study to evaluate readability and reliability of AI-generated brochures for emergency medical conditions. Cureus. 2024 Aug 31;16(8). https://doi.org/10.7759/cureus.68307
9. Momenaei B, Wakabayashi T, Shahlaee A, Durrani AF, Pandit SA, Wang K, Mansour HA, Abishek RM, Xu D, Sridhar J, Yonekawa Y. Appropriateness and readability of ChatGPT-4-generated responses for surgical treatment of retinal diseases. Ophthalmology Retina. 2023 Oct 1;7(10):862-8. https://doi.org/10.1016/j.oret.2023.05.022
10. Wong K, Levi JR. Partial tonsillectomy: Content and readability of online health information. Annals of Otology, Rhinology & Laryngology. 2017 Mar;126(3):192-8. https://doi.org/10.1177/0003489416681583
11. Stelzer JW, Wellington IJ, Trudeau MT, Mancini MR, LeVasseur MR, Messina JC, Mazzocca AD. Readability assessment of patient educational materials for shoulder arthroplasty from top academic orthopedic institutions. JSES international. 2022 Jan 1;6(1):44-8. https://doi.org/10.1016/j.jseint.2021.08.004
12. Gibson D, Jackson S, Shanmugasundaram R, Seth I, Siu A, Ahmadi N, Kam J, Mehan N, Thanigasalam R, Jeffery N, Patel MI. Evaluating the efficacy of ChatGPT as a patient education tool in prostate cancer: multimetric assessment. Journal of Medical Internet Research. 2024 Aug 14;26:e55939. https://doi.org/10.2196/55939
13. Soliman L, Soliman P, Gallo Marin B, Sobti N, Woo AS. Craniosynostosis: Are online resources readable?. The Cleft Palate Craniofacial Journal. 2024 Jul;61(7):1228-32. https://doi.org/10.1177/10556656231154843
14. Wu G, Lee DA, Zhao W, Wong A, Sidhu S. ChatGPT: is it good for our glaucoma patients?. Frontiers in Ophthalmology. 2023

Nov 16;3:1260415. https://doi.org/10.3389/fopht.2023.1260415
15. Stefu J, Slavych BK, Zraick RI. Patient-reported outcome measures in voice: an updated readability analysis. Journal of Voice. 2023 May 1;37(3):465-e27. https://doi.org/10.1016/j.jvoice.2021.01.028
16. Daungsupawong H, Wiwanitkit V. ChatGPT and Gemini large language models for pharmacometrics with NONMEM: comment. Journal of Pharmacokinetics and Pharmacodynamics. 2024 Aug;51(4):303-4. https://doi.org/10.1007/s10928-024-09926-7
17. Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. Healthc Manage Forum. 2020 Jan;33(1):10-18. https://doi.org/10.1177/0840470419873123
18. Ortiz-Barrios M, Arias-Fonseca S, Ishizaka A, Barbati M, Avendaño-Collante B, Navarro-Jiménez E. Artificial intelligence and discrete-event simulation for capacity management of intensive care units during the Covid-19 pandemic: A case study. Journal of Business Research. 2023 May 1;160:113806. https://doi.org/10.1016/j.jbusres.2023.113806
19. Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Molecular diversity. 2021 Aug;25(3):1315-60. https://doi.org/10.1007/s11030-021-10217-3
20. Tabassom A, Dahlstrom JJ. Epistaxis. 2022 Sep 12. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan–. PMID: 28613768.
21. Tunkel DE, Anne S, Payne SC, Ishman SL, Rosenfeld RM, Abramson PJ, Alikhaani JD, Benoit MM, Bercovitz RS, Brown MD, Chernobilsky B. Clinical practice guideline: nosebleed (epistaxis). Otolaryngology–Head and Neck Surgery. 2020 Jan;162:S1-38. https://doi.org/10.1177/0194599819890327
22. Villwock JA, Jones K. Recent trends in epistaxis management in the United States: 2008-2010. JAMA Otolaryngology–Head & Neck Surgery. 2013 Dec 1;139(12):1279-84. https://doi.org/10.1001/jamaoto.2013.5220
23. Kucik CJ, Clenney T. Management of epistaxis. Am Fam Physician. 2005 Jan 15;71(2):305-11.
24. Bonnici M, Orabi NA, Gannon M, Williams N, Stokes CM, Ramadan HH, Turner MT, Makary CA. Complications and Outcomes of Endovascular Embolization for Intractable Epistaxis: A Systematic Review and Meta-analysis. Annals of Otology, Rhinology & Laryngology. 2023 Oct;132(10):1233-48. https://doi.org/10.1177/00034894221143187
25. El Naamani K, Morse C, Ghanem M, Barbera J, Amllay A, Severance G, Ruiz R, Sweid A, Gooch MR, Herial NA, Jabbour P. Endovascular Embolization for Epistaxis: A Single Center Experience and Meta-Analysis. Journal of Clinical Medicine. 2023 Nov 7;12(22):6958. https://doi.org/10.3390/jcm12226958
26. Moreau S, De Rugy MG, Babin E, Courtheoux P, Valdazo A. Supraselective embolization in intractable epistaxis: review of 45 cases. The Laryngoscope. 1998 Jun;108(6):887-8. https://doi.org/10.1097/00005537-199806000-00018
27. Bedel HA, Bedel C, Selvi F, Zortuk Ö, Karancı Y. Emergency Medicine Assistants in the Field of Toxicology, Comparison of ChatGPT-3.5 and GEMINI Artificial Intelligence Systems. Acta Medica Lituanica. 2024 Dec 4;31(2):294. https://doi.org/10.15388/Amed.2024.31.2.18
28. Lee SW, Choi WJ. Utilizing ChatGPT in clinical research related to anesthesiology: a comprehensive review of opportunities and limitations. Anesthesia and Pain Medicine. 2023 Jul 26;18(3):244-51. https://doi.org/10.17085/apm.23056
29. Salvagno M, Taccone FS, Gerli AG. Can artificial intelligence help for scientific writing? Critical care. 2023 Feb 25;27(1):75. https://doi.org/10.1186/s13054-023-04380-2
30. Künzle P, Paris S. Performance of large Language artificial intelligence models on solving restorative dentistry and endodontics student assessments. Clinical Oral Investigations. 2024 Oct 7;28(11):575. https://doi.org/10.1007/s00784-024-05968-w
31. Jin HK, Lee HE, Kim E. Performance of ChatGPT-3.5 and GPT-4 in national licensing examinations for medicine, pharmacy, dentistry, and nursing: a systematic review and meta-analysis. BMC Medical Education. 2024 Sep 16;24(1):1013. https://doi.org/10.1186/s12909-024-05944-8
32. Günay S, Öztürk A, Yiğit Y. The accuracy of Gemini, GPT-4, and GPT-4o in ECG analysis: a comparison with cardiologists and emergency medicine specialists. The American Journal of Emergency Medicine. 2024 Oct 1;84:68-73. https://doi.org/10.1016/j.ajem.2024.07.043
33. Meral G, Ateş S, Günay S, Öztürk A, Kuşdoğan M. Comparative analysis of ChatGPT, Gemini and emergency medicine specialists in ESI triage assessment. The American Journal of Emergency Medicine. 2024 Jul 1;81:146-50. https://doi.org/10.1016/j.ajem.2024.05.001