



OPEN

Evaluating the readability and quality of AI-generated scoliosis education materials: a comparative analysis of five language models

Mengchu Zhao^{1,6}, Mi Zhou^{2,6}✉, Yexi Han^{3,6}, Xiaomei Song^{4,6}, Youbin Zhou^{5,6} & Haoning He^{1,6}✉

The complexity of scoliosis-related terminology and treatment options often hinders patients and caregivers from understanding their choices, making it difficult to make informed decisions. As a result, many patients seek guidance from artificial intelligence (AI) tools. However, AI-generated health content may suffer from low readability, inconsistency, and questionable quality, posing risks of misinformation. This study evaluates the readability and informational quality of scoliosis-related content produced by AI. We evaluated five AI models—ChatGPT-4o, ChatGPT-o1, ChatGPT-o3 mini-high, DeepSeek-V3, and DeepSeek-R1—by querying each on three types of scoliosis: congenital, adolescent idiopathic, and neuromuscular. Readability was assessed using the Flesch-Kincaid Grade Level (FKGL) and FleschKincaid Reading Ease (FKRE), while content quality was evaluated using the DISCERN score. Statistical analyses were performed in R-Studio. Inter-rater reliability was calculated using the Intraclass Correlation Coefficient (ICC). DeepSeek-R1 achieved the lowest FKGL (6.2) and the highest FKRE (64.5), indicating superior readability. In contrast, ChatGPT-o1 and ChatGPT-o3 mini-high scored above FKGL 12.0, requiring college-level reading skills. Despite readability differences, DISCERN scores remained stable across models (~ 50.5/80) with high inter-rater agreement (ICC = 0.85–0.87), suggesting a fair level quality. However, all responses lacked citations, limiting reliability. AI-generated scoliosis education materials vary significantly in readability, with DeepSeek-R1 being the most accessible. Future AI models should enhance readability without compromising information accuracy and integrate real-time citation mechanisms for improved trustworthiness.

Keywords AI-generated health information, Readability assessment, DISCERN score, Scoliosis, Patient health literacy

Abbreviations

AI	Artificial intelligence
FKGL	Flesch–Kincaid grade level
FKRE	Flesch–Kincaid reading ease
DISCERN	A standardized tool for assessing the quality of written consumer health information

Access to accurate and comprehensible health information is fundamental to decision-making in medical care¹. The growing dependence on digital health resources highlights this necessity: a 2022 global survey revealed that nearly 75% of patients seek online medical information before consulting healthcare professionals². However, the quality and readability of such materials often vary significantly, frequently neglecting to account for disparities in health literacy levels³.

¹Department of Rehabilitation Medicine, Shanghai Sixth People's Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China. ²Allied Health & Human Performance, University of South Australia, Adelaide, Australia.

³Department of Rehabilitation Medicine, Liaobu Hospital of Dongguan City, Dongguan, China. ⁴Department of Nursing, The Second Affiliated Hospital of Soochow University, Suzhou, Jiangsu, China. ⁵College of Intelligent Science and Control Engineering, Jinling Institute of Technology, Nanjing, China. ⁶Mengchu Zhao, Haoning He, Mi Zhou, Yexi Han, Youbin Zhou and Xiaomei Song are listed as co-first authors, having contributed equally to the work.

✉email: mi.zhou@mymail.unisa.edu.au; haoning.he@connect.polyu.hk

Patients with scoliosis and their caregivers necessitate educational resources that are lucid and accessible to comprehend treatment alternatives, including bracing, physiotherapy, and surgical interventions. Although clinicians strive to simplify medical jargon, numerous families are progressively relying on AI-generated content as supplementary information sources⁴. However, AI-generated content often remains difficult to read, which may result in patient misunderstandings regarding treatment options and potentially lead to poor decision-making or decreased adherence. Moreover, the accuracy, and reliability of AI-generated health information is also very important, as prior studies have shown that AI-generated content often lacks appropriate citations and personalization, potentially leading to treatment-related risks^{5–7}. Therefore, the quality of AI-generated educational materials demands rigorous evaluation.

While previous studies have explored the clarity and empathy of AI-generated scoliosis education materials, existing evaluations predominantly depend on subjective assessments, such as user satisfaction ratings⁸. Although useful, these methods may introduce bias and fail to provide an objective measurement of content quality and readability. Building upon prior research, the study adopts a more rigorous, data-driven approach, systematically evaluating AI-generated scoliosis education materials through standardized readability metrics and content quality assessments. The study seeks to elucidate the strengths and limitations of AI-generated health information to enhance the creation of more effective AI-driven educational resources for scoliosis patients and their caregivers.

Methods

Study design

This study systematically evaluated AI-generated patient education materials on scoliosis using validated metrics to overcome this gap. Specifically, we assessed readability and informational integrity in outputs from five prominent natural language processing models: ChatGPT-4o (OpenAI), a multimodal AI system optimized for clinical reasoning; ChatGPT-o1, an earlier iteration of OpenAI's multimodal AI, known for its strong language generation capabilities, ChatGPT-o3 mini—high (OpenAI), a lightweight model released in February 2025, emphasizing resource-efficient text generation DeepSeek-V3 (DeepSeek Inc.), and DeepSeek-R1, which incorporates cognitive architecture to enhance explanatory coherence.

On February 9, 2025, the five models were utilized to respond to common patient questions regarding the three major categories of scoliosis: idiopathic, neuromuscular, and congenital. These inquiries aimed to elicit accessible, user-friendly medical information from AI systems, following the approach outlined by Akkan and Seyyar⁹.

Quality analysis

The DISCERN score assesses the quality of AI-generated responses using two validated tools. The DISCERN score was developed as an instrument for assessing the quality of written patient information related to treatment options¹⁰. It consists of 16 questions in three structured sections: (1) eight questions evaluating the reliability of the information presented, (2) seven questions focusing on the completeness and accuracy of treatment-related details, and (3) a final overall quality rating. Each question is scored on a 1–5 scale, where 1 represents 'very poor quality' and 5 represents 'very high quality.' The total possible score is 80, with classifications as follows: a score exceeding 70 is considered "excellent," while a score above 50 is classified as "fair." Several studies have already used it to assess AI-generated health education materials^{11,12}. In this study, it was used to evaluate the quality of AI-generated responses.

Readability analysis

The readability of the responses was evaluated using two recognized metrics: the Flesch-Kincaid Reading Ease (FKRE) and the Flesch-Kincaid Grade Level (FKGL) (Flesch, 1948). The FKRE assigns a numerical readability score ranging from 0 to 100, with higher scores indicating simpler and more accessible text. A score approaching 100 suggests that the content is very easy to understand, whereas lower scores denote increased complexity. The FKGL is an adaptation of the FKRE that estimates the minimum education level required for comprehension. A higher FKGL score signifies increased complexity, indicating that individuals with lower formal education may struggle to comprehend the topic. The readability was calculated using the WebFX online readability tool¹³.

Statistical analysis

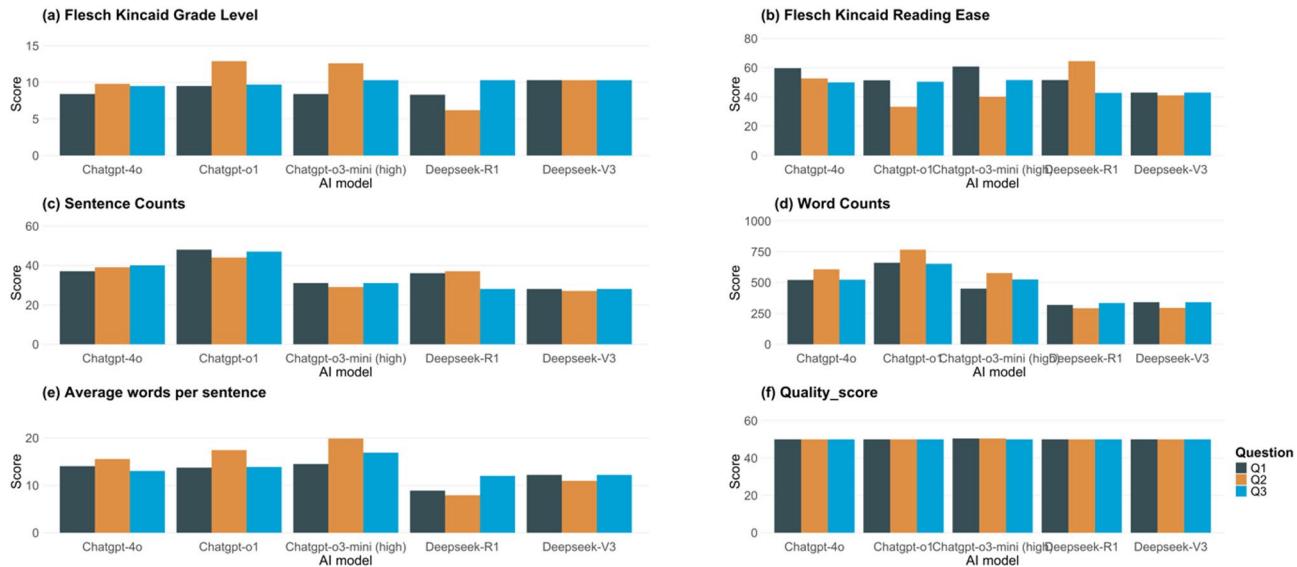
Continuous variables were presented as their raw values. The average DISCERN score by two reviewers was also calculated by two independent authors (DISCERN scores by MZ and XS). The raters have backgrounds in nursing and physical therapy and possess extensive experience in the therapeutic management of scoliosis. Inter-rater reliability was assessed using intraclass correlation coefficients (ICCs). Graphing and statistical tests were performed in R-Studio version 4.2.2¹⁴.

Results

Descriptive analysis

The AI-generated responses can be found in Appendix 1. The AI-generated responses demonstrated a generally high level of reliability and factual accuracy across all five evaluated LLMs. Each model effectively translated medical terminology related to scoliosis into accessible language suitable for readers, covering core components such as definitions, causes, symptoms, diagnostic methods, treatment options, and prognosis for idiopathic, neuromuscular, and congenital scoliosis. ChatGPT-4o and DeepSeek-R1 provided particularly structured and comprehensive explanations, employing clear subheadings, intuitive breakdowns of terms, and user-friendly phrasing. The content generated by ChatGPT-o1 and ChatGPT-03-mini, while also informative, tended to vary

No.	Prompt
Q1	"I need information about 'Idiopathic Adolescent Scoliosis.' I'm not familiar with medical terms. Can you help clarify this for me?"
Q2	"I need information about 'Neuromuscular Scoliosis.' I'm not familiar with medical terms. Can you help clarify this for me?"
Q3	"I need information about 'Congenital Scoliosis'. I'm not familiar with medical terms. Can you help clarify this for me?"

Table 1. Questions for scoliosis terminology clarification.**Fig. 1.** Readability and DISCERN score of the response from ChatGPT and Deepseek for the three questions.

in depth and stylistic consistency, with occasional redundancies and less nuanced explanations. DeepSeek-V3 offered concise but medically sound overviews, particularly excelling in explaining the pathophysiological basis of scoliosis subtypes. All models included appropriate cautionary advice encouraging consultation with healthcare professionals, which enhanced perceived reliability.

Readability

Table 1 of Appendix 2 presents the readability outcomes for responses generated by five models—ChatGPT-4o, ChatGPT-o1, ChatGPT-o3 mini-high, DeepSeek-V3, and DeepSeek-R1—across three prompts (Q1, Q2, Q3). Figure 1 (subplots a–e) visualized the readability results. According to the FKGL, DeepseekR1 generally produced the most accessible texts, evidenced by its lowest score of 6.2 (Q2). Conversely, ChatGPTo3mini (high) and ChatGPTo1 occasionally yielded more challenging content, both above FKGL 12.0 on Q2 (12.6 and 12.9, respectively), suggesting material that may require college-level reading skills. ChatGPT-4o exhibited a moderate level of difficulty, with FKGL values ranging from 8.4 to 9.8, while DeepseekV3 consistently achieved FKGL 10.3 across all prompts.

In reference to the FKRE, ChatGPTo3mini (high) achieved the highest individual score of 60.8 on Q1, indicating relatively easy-to-read content. At the other end, ChatGPTo1 reached the lowest FKRE of 33.3 (Q2), reflecting more intricate text. DeepseekR1 ranged from 42.8 to 64.5, notably featuring the highest reading ease score overall (64.5, Q2). ChatGPT-4o produced moderately accessible responses (FKRE 50.0–59.7), whereas DeepseekV3 maintained a narrower band of 41.1–43.0, indicating a comparatively denser style.

Analyses of sentence and word counts further underscore differences in verbosity and structure. ChatGPTo1 generated the lengthiest responses, comprising up to 767 words (Q2) and as many as 48 sentences (Q1). Despite ChatGPTo3mini (high) reaching a considerable word count (577 in Q2), it tended to use fewer overall sentences (29–31). Meanwhile, DeepseekR1 maintained concise outputs (293–336 words), coupled with shortest average words per sentence (7.92–12). ChatGPT-4o fell between 520 and 607 words and comprised 37 to 40 sentences for each response.

Quality assessment (DISCERN-like score)

As illustrated in Fig. 1 (subplot f), the reviewer-based scores (Reviewer 1, Reviewer 2) and their average remained relatively consistent across all models, with ICC scores ranging from 0.85 to 0.87. The scores hovered around 50.0–50.5 out of a maximum of 80 and were consistently rated as “fair.” No substantive differences were observed in these quality assessments, suggesting comparable performance across all systems on this metric.

Discussion

The study conducts a comparative evaluation of five large language models (LLMs) in generating medical information, emphasizing readability and content quality. The findings reveal significant disparities in readability across the models. Notably, Deepseek-R1 produces the most comprehensible content, whereas ChatGPT-o3-mini (high) and ChatGPT-o1 generate more complex texts that require higher reading proficiency. Notwithstanding these variations in readability, all models attain similar DISCERN scores, suggesting that while sentence structures and complexity differ, the overall quality of medical information remains stable, albeit consistently rated at only a 'fair' level.

Consistent with previous research on LLMs in medical applications, the architecture of AI models significantly influences text readability¹⁵. In our study, ChatGPT-o3-mini (high) and ChatGPT-o1 received elevated FKGL scores, signifying increased syntactic complexity and denser vocabulary. This trend aligns with previous findings that smaller or less optimized LLMs often mitigate their limited contextual understanding by increasing information density, leading to longer sentences and a higher prevalence of medical terminology¹⁶. The two DeepSeek models demonstrate outstanding performance in terms of readability. However, research on these models remains limited, and there is currently no systematic analysis in the literature exploring the potential reasons behind their superior performance. We hypothesize that several factors may contribute to this advantage. First, DeepSeek may employ a more advanced neural network architecture, incorporating more efficient attention mechanisms, deeper model structures, or optimized parameter tuning, allowing it to capture linguistic patterns with greater precision during text comprehension and generation. Second, the model might have been trained on a larger and higher-quality dataset, which could enhance its language understanding capabilities and improve text fluency. Additionally, DeepSeek may leverage fine-tuning strategies specifically designed to optimize readability, further refining the coherence and clarity of the generated text. Finally, optimizations at the inference stage, such as improved sampling methods or decoding strategies, may also contribute to producing text that aligns more closely with natural language conventions. Nevertheless, these remain speculative explanations, and future research should further investigate the specific optimization strategies that contribute to DeepSeek's readability performance.

The observed differences in model behavior highlight a critical design consideration in LLMs: certain models prioritize simplified content to enhance accessibility, whilst others emphasize precision and depth, rendering them more suitable for audiences with elevated health literacy. Nevertheless, excessive simplification can sometimes lead to the omission of essential medical information. For example, it is reported that when AI-generated content is oversimplified, a substantial amount of critical information is omitted. Oversimplified explanations in teaching materials on scoliosis inadequately represent the complexity of the disease. This poses significant challenges for patients, as it may impede their understanding of disease progression, potential consequences, and the necessity for timely intervention. Ambiguous explanations may encourage patients or their parents to underestimate the importance of early treatment, leading to delayed medical intervention.

Additionally, patients' ability to process medical information differs considerably depending on their developmental stage and age. Individuals with limited health literacy, especially younger ones, often struggle to grasp abstract medical concepts. They rely increasingly on intuitive analogies, visual aids (such as diagrams or animations), and narrative storytelling to build foundational knowledge of diseases. However, current LLMs focus primarily on enhancing text readability rather than diversifying information presentation. This constraint suggests that, even if text readability improves, key medical concepts may be ineffectively communicated, especially for complex diseases.

Moreover, the salience and comprehensibility of medical details shift markedly with patients' cognitive maturity and psychosocial development. For example, adolescents prioritize different aspects of medical information compared to adults. They concentrate on the implications of a disease on their daily existence, including whether scoliosis would limit physical activities, alter body posture, or impact social interactions and self-esteem. However, current medical information dissemination strategies are predominantly designed for parents or healthcare professionals, rather than directly targeting adolescents. This "bystander" communication approach often overlooks adolescents' autonomy and emotional needs, leading to a fragmented understanding of their condition or potential resistance to medical advice¹⁷. If AI-generated information overemphasizes the need for medical intervention while neglecting lifestyle or exercise recommendations, adolescents may perceive limited personal choice, hence reducing their acceptance of treatment plans. Therefore, merely simplifying language or enhancing readability is insufficient; the key lies in developing a personalized information framework that caters to different age groups and psychological needs.

Future optimization efforts should transcend traditional text modifications by integrating insights from cognitive science and health communication research. For example, interactive Q&A models could augment adolescents' engagement, while adapting medical content to social media formats may better correspond with their reading preferences. Simultaneously, it is essential to avoid excessive simplification that undermines the scientific integrity of medical information. The ultimate goal is to ensure that audiences understand medical content and may make informed health decisions based on complete and accurate information.

Despite the evident differences in readability among models, the study reveals that their DISCERN scores remain relatively consistent, indicating that content quality is not significantly affected by textual complexity. However, a major limitation common to all models is the absence of explicit source citations—a well-documented issue in AI-generated medical content. One study highlighted that AI medical chatbots frequently produce "hallucinated" citations, referencing inaccurate or non-existent sources. This problem is particularly prevalent in ChatGPT and Bing-generated medical content¹⁸. Additionally, another study substantiated that mainstream AI models perform poorly in the accuracy of scientific citation, especially in medical literature, frequently leading to misleading conclusions¹⁹.

The lack of reliable references diminishes the credibility and verifiability of AI-generated health information and poses potential risks to medical decision-making. Previous study found that general users often fail to distinguish AI-generated medical advice from professional recommendations, displaying high levels of trust in AI-generated content, even when it contains inaccuracies²⁰. This over-reliance on AI-generated health information, along with citation deficiencies, increases the risk of widespread misinformation among the public.

Beyond citation concerns, another critical limitation is the lack of adaptive personalization in AI-generated health content. Current AI models rely on generalized information-generation strategies, failing to dynamically tailor content according to an individual's health literacy, prior knowledge, or cognitive capacity. One study suggest that personalized health education materials markedly enhance patient understanding and adherence to treatment plans, particularly for those with complex medical conditions²¹. Nonetheless, existing AI systems continue to offer "one-size-fits-all" medical information, which not only inadequately addresses diverse patient requirements but may also result in misinterpretations or cognitive overload. This limitation is especially evident in scoliosis-related health communication. Adolescents and their parents often possess divergent apprehensions about scoliosis, such as its effects on daily life, physical appearance, and mental well-being. However, AI-generated texts frequently neglect to adequately address these aspects, reducing the practical relevance of the information.

Limitation

One major limitation of this study is the rapid evolution of AI, meaning that future versions may enhance transparency and clinical applicability, hence rendering the current conclusions time-sensitive. Another is that the quality of AI responses was judged based on expert opinion rather than clinical trials. Future studies should integrate feedback from both physicians and patients to comprehensively evaluate the reliability of AI-generated medical services. Ultimately, the study exclusively compared models from ChatGPT and Deepseek; subsequent work should encompass a wider array of AI models for comparison.

Conclusion

This study evaluated the readability and quality of AI-generated health education content specifically aimed at adolescents with scoliosis and their caregivers. The study highlights notable differences in the readability of AI-generated medical texts. Among the evaluated models, Deepseek-R1 produces the most understandable content, whereas ChatGPT-o3-mini (high) and ChatGPT-o1 often provide more complex responses. Despite these readability disparities, the overall level of content remains consistent across models, with all five LLMs generated content rated as only "fair" in quality, highlighting a need for improvement in trustworthiness and informational depth. Future advancements should focus on enhancing readability without compromising informational integrity, integrating real-time citation mechanisms, and developing AI systems to customize responses based on users' health literacy and individual medical conditions.

Data availability

All data supporting the findings of this study—including prompts, model metadata, generated texts, computed readability/quality metrics, and the scoring rubric—are included in this article and its Supplementary Information files.

Received: 18 February 2025; Accepted: 8 September 2025

Published online: 10 October 2025

References

1. Adegoke, B. O., Odugbosa, T. & Adeyemi, C. Assessing the effectiveness of health informatics tools in improving patient-centered care: A critical review. *Int. J. Chem. Pharm. Res. Updates [online]*. **2** (2), 1–11 (2024).
2. Calixte, R. et al. Social and demographic patterns of health-Related internet use among adults in the united states: A secondary data analysis of the health information National trends survey. *Int. J. Environ. Res. Public Health.* **17** (18), 6856 (2020).
3. Evangelista, L. S. et al. Health literacy and the patient with heart failure—Implications for patient care and research: A consensus statement of the heart failure society of America. *J. Card. Fail.* **16** (1), 9–16 (2010).
4. Zhang, P., Kamel, M. N. & Boulos Generative AI in medicine and healthcare: Promises, opportunities and challenges. *Future Internet.* **15**(9), 286 (2023).
5. Shekar, S. et al. *People Over Trust AI-Generated Medical Responses and View Them to be as Valid as Doctors, Despite Low Accuracy* (2024). [arXiv:abs/2408.15266](https://arxiv.org/abs/2408.15266).
6. Shin, D. et al. Debiasing misinformation: How do people diagnose health recommendations from AI? *Online Inf. Rev.* **48**, 1025–1044 (2024).
7. Moore, I. et al. Doctor AI? A pilot study examining responses of artificial intelligence to common questions asked by geriatric patients. *Front. Artif. Intell.*, **7**, 1438012 (2024).
8. Lang, S. et al. *Is the Information Provided by Large Language Models Valid in Educating Patients about Adolescent Idiopathic Scoliosis? An Evaluation of Content, Clarity, and Empathy* (Spine Deformity, 2024).
9. Akkan, H. & Seyyar, G. K. Improving readability in AI-generated medical information on fragility fractures: The role of prompt wording on chatgpt's responses. *Osteoporos. Int.*, **36**(3), 403–410 (2025).
10. Charnock, D. et al. DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *J. Epidemiol. Commun. Health.* **53**(2), 105–111 (1999).
11. Zhou, M. et al. Evaluating AI-generated patient education materials for spinal surgeries: Comparative analysis of readability and DISCERN quality across ChatGPT and deepseek models. *Int. J. Med. Inform.* **198**, 105871 (2025).
12. Ghanem, Y. K. et al. Dr. Google to dr. ChatGPT: Assessing the content and quality of artificial intelligence-generated medical information on appendicitis. *Surg. Endosc.* **38**(5), 2887–2893 (2024).
13. Readability Test. ; (2025). Available from: <https://www.webfx.com/tools/read-able/>
14. Wickham, H. *ggplot2*. Wiley interdisciplinary reviews: computational statistics, **3**(2): pp. 180–185. (2011).

15. Behers, B. et al. Assessing the readability of patient education materials on cardiac catheterization from artificial intelligence chatbots: an observational Cross-Sectional study. *Cureus*, **16**(7) (2024).
16. Mannhardt, N. et al. *Impact of Large Language Model Assistance on Patients Reading Clinical Notes: A Mixed-Methods Study* (2024). [arXiv:abs/2401.09637](https://arxiv.org/abs/2401.09637).
17. Chen, C. Y., Lo, F. & Wang, R. H. *Roles of Emotional Autonomy, Problem-Solving Ability and Parent-Adolescent Relationships on Self-Management of Adolescents with Type 1 Diabetes in Taiwan* (Journal of pediatric nursing, 2020).
18. Aljamaan, F. et al. Reference hallucination score for medical artificial intelligence chatbots: development and usability study. *JMIR Med. Inform.* **12**, e54345 (2024).
19. Emily, M., Graf, J. A., McKinney, A. B., Dye, L., Lin, L. & Sanchez-Ramos, L. Exploring the limits of artificial intelligence for referencing scientific articles. *Am. J. Perinatol.* **41**(15), 2072–2081 (2024).
20. Shekar, S., Pataranaporn, P., Sarabu, C., Cecchi, G. A. & Maes, P. People over trust AI-generated medical responses and view them to be as valid as doctors, despite low accuracy. arXiv preprint, [arXiv:2403.09755](https://arxiv.org/abs/2403.09755) (2024).
21. Golan, T., Casolino, R., Biankin, A. V., Hammel, P., Whitaker, K. D., Hall, M. J. & Riegert-Johnson, D. L. Germline BRCA testing in pancreatic cancer: improving awareness, timing, turnaround, and uptake. *Ther. Adv. Med. Oncol.* **15**, (2023).

Acknowledgements

We would like to extend our sincere gratitude to the developers of ChatGPT and DeepSeek for their valuable contributions.

Author contributions

McZ contributed to conceptualization, methodology, software, investigation, data curation, writing – original draft, and visualization. HH contributed to methodology, validation, formal analysis, investigation, and data curation. MZ contributed to conceptualization, validation, writing–review & editing. XS contributed to conceptualization, validation, writing–review & editing, supervision, and project administration. YH contributed to the formal analysis and visualization. YZ contributed to the formal analysis and visualization. All the authors have read and approved the final version of the manuscript and agree with the order of presentation of the authors.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declarations

Competing interests

The authors declare no competing interests.

Clinical trial number

Not applicable.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-19370-3>.

Correspondence and requests for materials should be addressed to M.Z. or H.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025