

Validating Artificial Intelligence before the Bedside: ChatGPT-4 vs Mini in Guideline-Based Hypertension Content

Cláudio Tinoco Mesquita,^{1,2,3,4} Fernanda Azevedo Silva,^{1,2} Ana Beatriz Costa do Couto^{2,3}

Universidade Federal Fluminense,¹ Niterói, RJ – Brazil

Hospital Universitário Antônio Pedro, EBSERH, Health, Science & Education Lab,² Niterói, RJ – Brazil

Programa de Pós-Graduação em Ciências Cardiovasculares – Universidade Federal Fluminense,³ Niterói, RJ – Brazil

Hospital Pró-Cardíaco,⁴ Rio de Janeiro, RJ – Brazil

Short Editorial related to the article: Assessment of ChatGPT-4.0 versus ChatGPT-Mini in Generating Guideline-Based Hypertension Content

The rapid integration of artificial intelligence (AI) into medical practice is a reality. From clinical decision support and study design to chatbot tools and the production of educational materials, large language models (LLMs) are increasingly used in daily practice. In this context, studies like the study of Ataídes et al.,¹ evaluating the performance of LLM-based materials, are crucial.

Hypertension is one of the leading causes of premature death; therefore, the World Health Organization (WHO) invests in prevention campaigns. It is a disease that affects multiple organs, causing a major impact on the entire body. The condition is defined by a systolic blood pressure (BP) of 140 mmHg or higher, or a diastolic BP of 90 mmHg, according to Brazilian Society of Cardiology guidelines.² The number of individuals diagnosed with hypertension increases each year, and studies indicate that the population's eating habits are characterized by high consumption of ultra-processed foods.³

The study comparing ChatGPT-mini and ChatGPT-4.0 in the generation of hypertension educational content addresses a very important aspect of AI in medicine: the quality and safety of patient-directed information. By systematically assessing accuracy, completeness, structural quality (EQIP), consistency, and alignment with clinical guidelines, the authors provided a comprehensive framework for evaluating AI-generated health content.¹

At first glance, comparing two specific model versions may appear time-sensitive in a field characterized by rapid technological turnover. Indeed, as of February 2026, GPT-4-based models have been fully discontinued within the ChatGPT ecosystem, with newer generations now in use.

However, this does not diminish the relevance of the study. On the contrary, it reinforces its central message.

The authors must be congratulated because the true value of their work lies not in determining which model performs better, but in demonstrating that no AI-generated content should be assumed reliable without rigorous validation.⁴ Even advanced models capable of producing fluent, coherent, and seemingly authoritative text may generate inaccuracies, omit critical information, or present recommendations that are misaligned with established guidelines.⁵ Importantly, this study underscores a very important principle that must be followed with any specific LLM model or version: the responsibility for the content delivered to patients remains with healthcare professionals and institutions—not with the algorithm^{4,6} (Figure 1).

Finally, ethical use of AI is critical, especially when LLMs are used in patient management. Patients should be informed when content is AI-assisted, and institutions must establish clear processes for review, approval, and periodic updating of educational materials.⁷

In conclusion, while technologies will continue to evolve, the fundamental requirement remains unchanged: before AI reaches the patient, it must be carefully tested, validated, and contextualized.

Acknowledgments

Dr. Cláudio Tinoco receives the following funding: CNPq Productivity Grant and Rio de Janeiro State Scientist Grant from FAPERJ.

Keywords

Artificial Intelligence; ChatGPT; Hypertension; Clinical Guidelines; Validation

Mailing Address: Cláudio Tinoco Mesquita •

Universidade Federal Fluminense, Hospital Universitário Antônio Pedro – Health, Science & Education Lab – Av. Marquês do Paraná, 303. Postal Code 24033-900. Niterói, RJ – Brazil

E-mail: claudiotinocomesquita@id.uff.br

Manuscript received March 25, 2026, revised manuscript March 25, 2026, accepted April 01, 2026

DOI: <https://doi.org/10.36660/abc.20260254i>

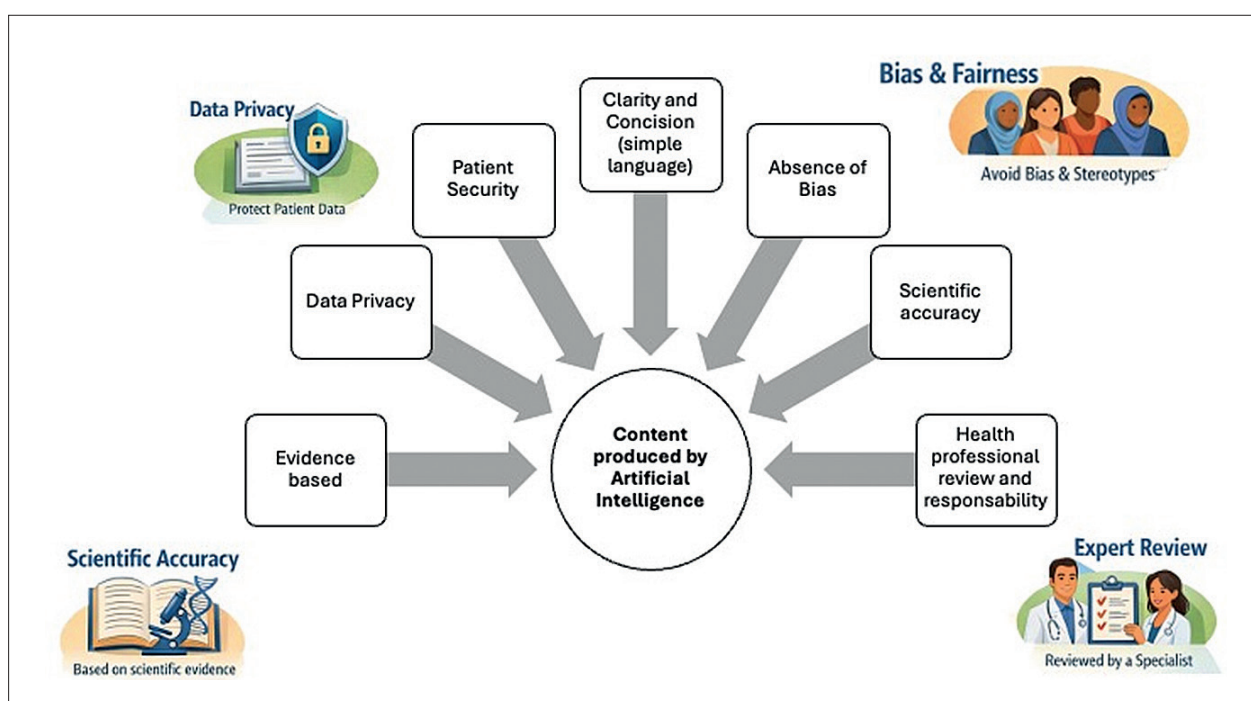


Figure 1 – This figure summarizes the main domains to assess AI-generated content for patients: scientific accuracy and evidence base, expert review and responsibility, clarity and readability, absence of bias, patient safety, and data privacy. Together, these elements reinforce that AI outputs require systematic validation and human oversight before clinical use.

References

1. Ataides RJC, Campos MAG, Souza JVP, Rocha RC, Lacalle AA, Vieira CB, et al. Assessment of ChatGPT-4.0 versus ChatGPT-Mini in Generating Guideline-Based Hypertension Content. *Arq Bras Cardiol.* 2026;123(2):e20250498. doi: 10.36660/abc.20250498.
2. Brandão AA, Rodrigues CIS, Bortolotto LA, Armstrong ADC, Mulinari RA, Feitosa ADM, et al. Brazilian Guidelines of Hypertension - 2025. *Arq Bras Cardiol.* 2025;122(9):e20250624. doi: 10.36660/abc.20250624.
3. Alhuneafat L, Al Ta'ani O, Arriola-Montenegro J, Al-Ajloun YA, Naser A, Chaponan-Lavalle A, et al. The burden of Cardiovascular Disease in Latin America and the Caribbean, 1990-2019: An Analysis of the Global Burden of Disease Study. *Int J Cardiol.* 2025;428:133143. doi: 10.1016/j.ijcard.2025.133143.
4. Griffin AC, Javier SJ, Golding M, Runnels TW, Matthias MS, Shimada SL, et al. Designing Patient-Friendly Messages: Tutorial on Applying Human-Centered, Self-Determination Theory with AI Considerations. *J Med Internet Res.* 2025;27:e78173. doi: 10.2196/78173.
5. Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical Considerations of Using ChatGPT in Health Care. *J Med Internet Res.* 2023;25:e48009. doi: 10.2196/48009.
6. Mello MM, Char D, Xu SH. Ethical Obligations to Inform Patients about Use of AI Tools. *JAMA.* 2025;334(9):767-70. doi: 10.1001/jama.2025.11417.



This is an open-access article distributed under the terms of the Creative Commons Attribution License