



Comments on "Use of ChatGPT in Urology and its Relevance in Clinical Practice": Toward a Responsible **AI Framework for Middle-Income Settings**

Juan	Martín	Monto	/a Osori	o ¹
------	--------	-------	----------	----------------

¹ Pontificia Universidad Javeriana, Bogotá, Colombia

To the editor,

We read with interest the article by Braga et al. evaluating ChatGPT's accuracy and reproducibility in pediatric urology (1). The study highlights important limitations: although ChatGPT can provide partially correct answers, it also produces incomplete, ambiguous, or potentially misleading responses—particularly regarding diagnostic definitions and treatment recommendations. We also noted a minor typographical error in the abstract ("those rwere evaluated qualitatively..."), which does not affect the results but may merit correction in the online version (1).

Another consideration is the rapid evolution of large language models (LLMs). The version evaluated by Braga et al. was most likely ChatGPT-3.5. Since then, GPT-5 has been released, offering improvements in contextual comprehension, medical reasoning, and response consistency. As a result, performance findings may change over short periods, underscoring the importance of continuous evaluation when clinical accuracy is critical.

These observations reinforce a broader point: deploying LLMs in clinical practice—particularly in middleincome countries-requires careful, evidence-based integration. In Colombia, the Interoperabilidad de la Historia Clínica Electrónica (IHCE), an Interoperable Electronic Health Record, provides an opportunity to structure and validate AI tools responsibly. Building on this context, we propose a five-point framework to guide the safe and effective adoption of AI in urology:

- 1. Prioritize validated, clinically relevant use cases-for example, prostate mpMRI interpretation and Gleason grading—where evidence is strongest (2, 3).
- 2. Enforce rigorous evaluation standards by adhering to TRIPOD+AI for predictive models and established AI trial reporting guidance (4).
- 3. Leverage national interoperability infrastructures (e.g., IHCE) to ensure data traceability, standardization, and privacy.
- 4. Adopt ethical safeguards recommended by the WHO for AI in health, including human oversight, transparency, and bias mitigation (5).
- 5. Implement short-term multicenter pilots measuring diagnostic accuracy, time to clinical decision, and Al-associated adverse events.

By learning from Braga et al's findings, addressing minor editorial issues, and accounting for rapid model evolution, urology in middle-income settings can advance from experimental enthusiasm to real-world, responsible integration. We believe these suggestions will resonate with IBJU's readership and provide a practical roadmap for countries facing similar challenges.

The Author

REFERENCES

- Braga AVNM, Nunes NC, Santos EN, Veiga ML, Braga AANM, de Abreu GE, et al. Use of ChatGPT in Urology and its Relevance in Clinical Practice: Is it useful? Int Braz J Urol. 2024;50(2):192-198. doi:10.1590/S1677-5538.IBJU.2023.0570.
- Saha A, Bosma JS, Twilt JJ, van Ginneken B, Bjartell A, Padhani AR, et al.; PI-CAI consortium. Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study. Lancet Oncol. 2024;25(7):879-887. doi:10.1016/S1470-2045(24)00220-1.
- Bulten W, Pinckaers H, van Boven H, Vink R, de Bel T, van Ginneken B, et al. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. Lancet Oncol. 2020;21(2):233-241. doi:10.1016/S1470-2045(19)30739-9.

Correspondence address: Juan Martín Montoya Osorio, MD

Pontificia Universidad Javeriana Carrera 19a #85-78 Bogotá, 110221, Colombia

E-mail: montoyajmartin@gmail.com

CONFLICT OF INTEREST

None declared.

- Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. doi:10.1136/bmj-2023-078378.
- World Health Organization. Ethics and governance of artificial intelligence for health: Guidance on large multimodal models [Internet]. Geneva: WHO; 2024 [cited 2025 Aug 18]. Available at: https://iris.who.int/bitstream/hand le/10665/375579/9789240084759-eng.pdf>.

ARTICLE INFO

🔟 Juan M. Montoya-Osorio

https://orcid.org/0000-0001-6925-5915

Submitted for publication: August 15, 2025

Accepted after revision: August 18, 2025

Published as Ahead of Print: August 30, 2025