International Journal of Technology Assessment in Health Care

www.cambridge.org/thc

From task validation to workflow assurance in health technology assessment

Gaurav Agarwal[1] and Sharmila Chaudhuri[2]

[1]Department of Economics, Centre for Distance and Online Education, Galgotias University, India and [2]Department of Physiotherapy, School of Allied Health Sciences, Galgotias University, India

Letter

Cite this article: Agarwal G, Chaudhuri S (2026). From task validation to workflow assurance in health technology assessment. *International Journal of Technology Assessment in Health Care*, **42**(1), e73, 1
https://doi.org/10.1017/S0266462326104012

Received: 30 June 2026 UTC
Revised: 30 June 2026 UTC
Accepted: 12 July 2026 UTC

Corresponding author:
Gaurav Agarwal;
Email: gaurav7777agarwal@gmail.com

We read with great interest the ongoing *Dialogue on Artificial Intelligence in Health Technology Assessment (HTA)*, particularly the article by Fleurence and Chhatwal on implementing generative artificial intelligence (GenAI) into HTA practice (1) and the subsequent letter by Yu et al., which extends these principles to bedside AI and clinical data governance (2). Together, these contributions converge on an important principle: GenAI should be validated according to well-defined, bounded tasks rather than viewed as a technology that is uniformly "ready" or "not ready" for HTA.

We fully agree with this perspective and wish to extend it one step further. Current validation approaches primarily evaluate GenAI within individual tasks, such as literature screening, data extraction, or report drafting. HTA recommendations, however, are generated through interconnected workflows in which the output of one stage becomes the input to the next. As Yu et al. note, translating AI capability into clinical practice requires human oversight and appropriate governance (2). We suggest that it also requires attention to how AI-assisted outputs interact across successive stages of the HTA process.

This distinction has practical implications. For example, omission of a clinically important study during literature screening may influence subsequent evidence synthesis, alter inputs to an economic model, and ultimately affect reimbursement recommendations, even when downstream tasks are themselves performed accurately. Strong performance for individual components therefore does not necessarily guarantee the reliability of the complete HTA workflow.

Future evaluations could usefully complement task-specific benchmarking with workflow-level validation, examining how outputs propagate across chained tasks and whether the overall assessment remains robust. Such an approach would provide a practical extension of the framework proposed by Fleurence and Chhatwal by identifying where human oversight and accountability should be exercised throughout the HTA pathway, rather than only within individual tasks (1).

The Dialogue has convincingly established that GenAI should be validated task by task. We suggest that the next methodological step is to evaluate how these validated tasks perform collectively as an integrated workflow, thereby strengthening trust, accountability, and the reliability of AI-enabled HTA.

Data availability. Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Acknowledgments. None.

Author contribution. Conceptualization: G.A., S.C.; Formal analysis: G.A.; Investigation: G.A.; Writing – original draft: G.A.; Writing – review and editing: S.C.
All authors read and approved the final manuscript.

Funding statement. The authors declare that no external funding was received for this work.

Competing interest. The authors declare no competing interests.

Compliance with ethical standard. Research involving human participants and/or animals.
This research did not require the involvement of human or animal subjects. Therefore, ethical approval for this study was not required according to local regulations and institutional policies.

Informed consent. Not applicable.

Use of generative AI. During the preparation of this manuscript, the authors used ChatGPT and Claude Sonnet 4.6 to assist with language refinement, organization, and improving readability. The authors critically reviewed, revised, and approved all content and take full responsibility for the accuracy and integrity of the manuscript.

References

1. **Fleurence R**, **Chhatwal J**. Trust, validation, and accountability: Implementing uses of generative artificial intelligence into health technology assessment practice. *Int J Technol Assess Health Care*. 2026;**42**(1):e41. https://doi.org/10.1017/S0266462326103638
2. **Yu Z**, **Cheng W**, **Qin F**, **Xiong W**, **Hu Y**. From HTA to the bedside: Aligning GenAI validation with clinical reality. *Int J Technol Assess Health Care*. 2026;**42**(1):1–4. https://doi.org/10.1017/S026646232610378X

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.