Assessment in medical education: rethinking the reasons for introducing AI

L'évaluation en éducation médicale : repenser les raisons de l'introduction de l'IA

Shigeki Matsubara[1]

[1]Department of Obstetrics and Gynecology, Jichi Medical University, Tochigi, Japan
Correspondence to: Shigeki Matsubara, M.D., Ph.D. Emeritus and Visiting Professor, Department of Obstetrics and Gynecology, Jichi Medical University, 3311-1 Shimotsuke, Tochigi 329-0498, Japan; phone: +81-285-58-7376, Fax: +81-285-44-8505; email: matsushi@jichi.ac.jp

Dear Editors,

Feldman et al.'s discussion, "Artificial Intelligence can transform formative assessment in medical education"[1] although thought-provoking, feels slightly foreign.

First, I cannot imagine the situation where the answers after an assessment are not opened. As they state, it deprives the chance of learning correct answers, one of the main purposes of assessment. Students' claim after seeing the answers prompts teachers to recognize inappropriate questions. A not-open policy also deprives teachers of the opportunity to learn.

Second, is making new questions a simply burdensome job? I made approximately 40–100 new questions annually for 20 years. I always searched for "good materials" for assessment. What should be assessed is fundamentally the same; however, questioning it from a different perspective is like a brain game, akin to everyday surgery for a "new" patient with various contrivances.

Third regards AI use. Proposing AI to generate a test means an "AI-first" strategy, not a human-first strategy. In medical writing, AI-first is criticized,[2] and caution is warranted here as well. Regardless, AI should be introduced for the benefit of students, not to lessen the burden of teachers.

Their context sounds: making new questions is a burden, some questions will be reused, thus answers cannot be open, so AI can make questions. The discussion should be reversed: let's explore how AI may generate "better" questions for "students." The context of St-Onge's editorial on AI and assessment[3] was in line with this.

This retired Japanese obstetrics–gynecology professor, not an education specialist, has long engaged in medical education. Sometimes, a non-specialist voice might help.

Conflicts of Interest: The author declares that there is no conflict of interest regarding this article.
Funding: None
Edited by: Marcel D'Eon (editor-in-chief)

References

1. Feldman J, Gilchrist C, Leung F-H. Artificial Intelligence can transform formative assessment in medical education. *Can Med Ed J.* 2024. https://doi.org/10.36834/cmej.79759
2. Matsubara S. Response: comparing letters written by humans and ChatGPT: a preliminary study. *Int J Gynaecol Obstet.* 2025;169(3):1260-1261. https://doi.org/10.1002/ijgo.70085.
3. St-Onge C. Embracing the next frontier in assessment. *Can Med Ed J.* 2025. https://doi.org/10.36834/cmej.81477