ISSN: 2581 – 3633

International Journal of Medical Science and Advanced Clinical Research (IJMACR)

Available Online at: https://www.ijmacr.com

Volume – 9, Issue – 5, September – 2026, Page No.: 273 – 277

AI Literacy Should Begin at Entry to Medical School: A Case for Entry-Level AI Safety Education

[1]Zeeya Modi, S S Ravishankar Vidya Mandir, Godhavi, Ahmedabad, India

[2]Yushi Modi, Medical Student, Monash University Malaysia, Malaysia

[3]Het Modi, Nursing Student, Anglia Ruskin University, Cambridge, United Kingdom

[4]Girishkumar Modi, Department of Anesthesiology, Tawam Hospital, Al Ain, United Arab Emirates

Corresponding Author: Zeeya Modi, S S Ravishankar Vidya Mandir, Godhavi, Ahmedabad, India.

How to citation this article: Zeeya Modi, Yushi Modi, Het Modi, Girishkumar Modi, "AI Literacy Should Begin at Entry to Medical School: A Case for Entry-Level AI Safety Education", IJMACR– September – 2026, Volume – 9, Issue – 5, P. No. 273 – 277.

Type of Publication: Short Communication

Conflicts of Interest: Nil

Abstract

Generative artificial intelligence (AI) is already part of how many students learn before medical schools have established consistent expectations for its use. Medical students use large language models to explain concepts, summarise information, generate practice questions and assist with academic work, but formal training in responsible use remains inconsistent. Rules governing assignments are necessary but cannot by themselves address inaccurate outputs, fabricated references, bias, privacy risks or over-reliance on AI-generated explanations. This Short Communication argues that a narrow, mandatory programme in foundational AI safety literacy should be delivered at entry to medical school, ideally before or during orientation, before unstructured use becomes routine. The proposed entry-level curriculum is built around five safety competencies: verification of medical claims, protection of confidential information, transparent disclosure of AI assistance, recognition of bias and missing context, and preservation of independent reasoning. A 60–90-minute introductory module is proposed as a pragmatic starting model using practical exercises, without promoting greater AI use or requiring technical instruction in machine learning. More advanced competencies can then be revisited longitudinally during research and clinical training. Foundational AI safety literacy should therefore be treated as an entry point to broader professional AI literacy in medicine.

Keywords: Artificial Intelligence, Generative AI, Medical Education, Medical Students, AI Literacy, Large Language Models.

Introduction

A student beginning medical school today enters a learning environment very different from that of only a few years ago. Generative artificial intelligence (AI) can

explain a biochemical pathway, create practice questions, summarise an article, translate terminology and produce polished prose within seconds. The educational question is therefore no longer simply whether students should use these tools. It is whether medical schools will teach students how to use them responsibly before unstructured use becomes a habit.

Evidence suggests that AI use among medical students is already substantial. In a 2024 survey of 415 students from at least 28 US medical schools, 52% reported using ChatGPT for medical-school coursework[1]. A 2025 multicentre Canadian survey found that 78.9% of respondents used generative AI and 92.8% were willing to learn how to integrate it into future clinical practice[2]. Among first-year medical students in Japan, 84.7% reported using generative AI, while only 49.2% had previously learned about it, mainly through informal sources[3]. A 2026 systematic review and meta-analysis of 26 studies involving 20,963 medical students found that 78.0% supported AI curriculum integration and 83.1% considered AI training necessary, yet only 36.4% felt confident applying AI clinically[4]. This optimism–competence gap supports introducing basic AI literacy at the entry to medical training rather than waiting until students begin clinical work.

Academic-integrity rules are necessary, but prohibition is not education. A student can comply with assessment rules and still use AI poorly: accepting a plausible but incorrect explanation, relying on a fabricated citation, summarising a paper without reading the original, or entering sensitive information into a public platform. An updated scoping review of AI in undergraduate medical education found rapid growth in educational applications but no standardised approach to AI competencies or ethical frameworks[5]. Medical schools therefore need to move beyond asking only, “Is AI allowed?” and begin teaching, “When is AI appropriate, how should its output be checked, and when should it not be used?”

Entry-level AI safety literacy does not require programming, machine-learning mathematics or a semester-long course. The aim should be deliberately narrow: to establish safe habits before routine use develops. Existing initiatives demonstrate that AI training can be implemented in early medical education. Levingston and colleagues described a first-year medical-student AI literacy mini-course covering AI fundamentals, prompt engineering, educational applications and ethical issues[6]. The distinctive contribution of the present proposal is a brief, universal and mandatory safety induction delivered to every incoming student, ideally before or during orientation, rather than a broader technical or prompt-engineering course. It is intended as the foundational safety component of broader AI literacy, regardless of whether students plan to use AI frequently.

Five competencies could form that baseline.

Table 1: Five Foundational AI Safety Competencies for Incoming Medical Students

Competency	Incoming students should learn to…	Example assessment
Verification	Check important AI-generated medical claims and citations against authoritative textbooks, clinical guidelines, or primary literature.	Identify and verify an inaccurate medical claim or fabricated citation.

Privacy	Avoid entering identifiable patient information or other confidential material into unapproved public AI systems.	Identify information that should not be entered into a public AI platform.
Transparency	Disclose AI assistance when required and distinguish legitimate support from misrepresentation of independent work.	Decide whether AI assistance requires disclosure in an academic scenario.
Bias awareness	Question missing context, representation, and systematic bias in AI-generated outputs.	Identify bias, under-representation, or missing clinical context in an AI response.
Reasoning preservation	Attempt problems independently, identify uncertainty, and then use AI to explain, critique, or extend reasoning rather than substitute for it.	Demonstrate an independent-first approach before consulting AI.

Abbreviation: AI, artificial intelligence.

These competencies lend themselves to a brief practical programme. As a pragmatic starting model rather than an evidence-established intervention, a 60–90-minute session could present an AI-generated medical explanation containing subtle errors and ask students to verify it against an authoritative source. Students could test an apparently credible citation that does not exist, compare an AI summary with the original guideline or paper, and work through examples of information that should not be entered into public systems. Short scenarios could assess when AI use is appropriate, when disclosure is required, where bias or missing context is present, and whether an independent-first reasoning approach has been followed. Such training would teach behaviours rather than allegiance to any particular platform, making it more durable as technologies change. Implementation should be aligned with local university policies, assessment requirements and healthcare data-governance standards, including distinctions between public generative-AI platforms and institutionally approved systems. The optimal timing, duration and educational format remain to be established prospectively; future studies should evaluate effects on knowledge, applied behaviour and retention over time.

These principles are consistent with the direction of medical-education guidance. Triola and Rodman have argued that generative AI requires coordinated curricula, policy and governance rather than isolated experimentation[7]. The Association of American Medical Colleges has emphasised human judgement, transparency, equitable access, education, interdisciplinary curriculum development, data privacy and ongoing evaluation in its principles for responsible AI use[8]. World Health Organization guidance warns that large multimodal models may produce false, inaccurate, biased or incomplete information and stresses governance for health-related applications[9]. UNESCO similarly advocates a human-centred approach to generative AI in education, including privacy protection and development of human capacity[10]. Across these frameworks, responsible use is something to be taught rather than assumed.

The most important reason for introducing this training early is not that AI is necessarily harmful to reasoning. The evidence is more nuanced. Generative AI can support personalised learning and efficiency, while reviews

continue to identify risks related to hallucination, bias, privacy, plagiarism and over-reliance[11]. A 2026 longitudinal study of 372 medical students in supervised AI-assisted diagnostic training found that greater AI literacy was prospectively associated with higher subsequent critical-thinking scores, although its observational design cannot establish that AI literacy caused the improvement[12]. The relevant educational question is therefore not whether AI strengthens or weakens reasoning in every circumstance, but how students are taught to use it.

Consider two students facing the same difficult concept. One immediately requests an AI-generated explanation and transfers the answer into notes. The other first attempts the problem, identifies the point of uncertainty, asks AI for an alternative explanation and then checks that explanation against an authoritative source. Both have used the same technology, but educationally they have done different things. Without deliberate learning strategies, AI may reduce opportunities for retrieval, comparison and management of uncertainty that contribute to the development of clinical reasoning. The educational goal should therefore not be maximal AI use. It should be calibrated AI use.

A reasonable counterargument is that incoming students are already overloaded with orientation material and may not yet understand enough medicine to evaluate AI outputs. That concern argues for limiting the scope of entry-level training, not delaying it. At this stage, the objective is a safety induction focused on verification, privacy, transparency, bias awareness and preservation of independent reasoning. More advanced competencies can then be revisited longitudinally as students encounter research methods, clinical reasoning, documentation, patient communication and decision support.

Early teaching may also support fairness. Without institutional guidance, differences in technical familiarity, access to advanced systems or informal mentoring may contribute to unequal learning opportunities. A common introductory programme, consistent with calls for equitable access in responsible AI education[8], can establish the same baseline for everyone: which uses are acceptable, which require disclosure, what information must not be entered, which outputs require verification and which assessments must remain entirely independent. Entry-level education should also identify situations in which students should not use AI: when testing unaided recall, when developing an initial line of reasoning, when an assessment requires independent work, when an appropriate authoritative source such as a guideline, textbook or primary source is readily available, or when information cannot safely be shared with the system. Knowing when not to use AI is itself a foundational safety skill.

Medical education has always taught students to work with tools while remaining responsible for their decisions. Generative AI deserves particular caution because it can simulate explanation, dialogue and confidence, making critical appraisal more important, not less. A brief entry-level safety induction, followed by longitudinal teaching throughout the curriculum, offers a practical starting point. The goal is to educate doctors who can use AI without surrendering judgement to it. That education should begin at entry to medical school.

References

1. Ganjavi C, Eppler M, O'Brien D, Ramacciotti LS, Ghauri MS, Anderson I, et al. ChatGPT and large language models (LLMs) awareness and use: a prospective cross-sectional survey of U.S. medical

students. PLOS Digit Health. 2024;3(9):e0000596. doi:10.1371/ journal.pdig.0000596.

2. Tran C, Hryciw BN, Moore SW, Chaput A, Seely AJE. Perceptions and use of generative artificial intelligence in medical students: a multicenter survey. J Med Educ Curric Dev. 2025; 12:23821205251391969. doi: 10.1177/ 23821205251391969.
3. Tajima H, Kasai H, Shikino K, Shimizu I, Ito S. Perceptions and intentions to use generative AI among first-year medical students in Japan: cross-sectional survey study. JMIR Med Educ. 2025;11:e77552. doi:10.2196/77552.
4. Mastour H, Mirzaei S, Sohrabi S, et al. Is it time to introduce an artificial intelligence curriculum in undergraduate medical education? Medical students' perspectives: a systematic review and meta-analysis. BMC Med Educ. 2026; 26:1380. doi:10.1186/s12909-026-09630-9.
5. Simoni J, Urtubia-Fernandez J, Mengual E, et al. Artificial intelligence in undergraduate medical education: an updated scoping review. BMC Med Educ. 2025; 25:1609. doi:10.1186/s12909-025-08188-2.
6. Levingston H, Anderson MC, Roni MA. From theory to practice: artificial intelligence (AI) literacy course for first-year medical students. Cureus. 2024;16(10): e70706. doi:10.7759/cureus.70706.
7. Triola MM, Rodman A. Integrating generative artificial intelligence into medical education: curriculum, policy, and governance strategies. Acad Med. 2025;100 (4):413–418. doi:10.1097/ ACM.0000000000005963.
8. Association of American Medical Colleges. Principles for the Responsible Use of Artificial Intelligence in and for Medical Education. Version 2.0. Washington, DC: Association of American Medical Colleges; 2025.
9. World Health Organization. Ethics and Governance of Artificial Intelligence for Health: Guidance on Large Multi-Modal Models. Geneva: World Health Organization; 2024.
10. Miao F, Holmes W. Guidance for Generative AI in Education and Research. Paris: UNESCO; 2023.
11. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel). 2023;11(6):887. doi:10.3390/healthcare11060887.
12. Xin Y, Yan D, Shuren L, et al. AI literacy mediates AI assisted diagnosis participation and critical thinking among medical students under supervision. NPJ Digit Med. 2026; 9:344. doi:10.1038/s41746-026-02521-9.