AI and Ethics (2026) 6:219
https://doi.org/10.1007/s43681-026-01041-2

REVIEW

Uncertain hands: ethical implications of AI dependency in medical education

Enoch Chi Ngai Lim[1] · **Chi Eung Danforn Lim**[1,2,3]

Received: 26 June 2025 / Accepted: 6 February 2026

Abstract
As artificial intelligence (AI) systems are increasingly adopted into educational frameworks, the medical field is undergoing a radical revolution. Incorporating a poem, this review analyses the ethical concerns of technology dependency among medical students who grew up with smartphones and digital devices, focusing on both sides of the AI integration debate. From literature review and case analyses, we address some emerging ethical issues like loss of clinical intuition, empathic deficits, and diagnostic overconfidence bias. There are undeniable advantages to AI technologies, such as improved efficiency in learning and increased access to medical knowledge globally; however, reliance on algorithms raises critical issues regarding the preservation of basic skills in patient-centred care. We suggest an ethical boundary framework that focuses on integration, emphasising assistance rather than secession, to avoid technological disintegration that impedes the development of high-level thinking among clinicians. The results indicate a need for a balance between leveraging AI's capabilities and developing humanistic clinical skills in preparing graduates for complex patient interactions.

Keywords Artificial intelligence · Medical education · Ethics · Clinical reasoning · Empathy

1 Introduction

In quiet rooms where learners grow,

A new voice tells them what to know.

It's sharp, it's fast, it rarely slips -

A guide built strong with code and scripts.

One student paused - unsure, afraid -

The screen had spoken, plans were laid.

But something felt not quite complete -

The face in front, too still, too sweet.

These verses capture a defining moment in contemporary medical education: the emergence of a generation of learners who have never known a world without algorithmic assistance. Today's medical students represent the first group being trained clinically in a world where technology, specifically artificial intelligence (AI), is deeply embedded in the education framework and training infrastructure [1, 2].

Over the last decade, AI has introduced numerous advanced features to educational technologies, including virtual patient simulators, intelligent tutoring systems, automated evaluation systems, and diagnostic decision support tools [3, 4]. This transformation came with Generation Z. They are digital natives who grew up with smartphones and have unlimited information resources available at their fingertips 24/7 [5]. In comparison to earlier generations, these learners approach medical training with preconceived

✉ Chi Eung Danforn Lim
chi.lim@westernsydney.edu.au

Enoch Chi Ngai Lim
enoch@smsg.au

[1] Translational Research Department, Specialist Medical Services Group, Earlwood, Australia

[2] NICM Health Research Institute, Western Sydney University, Westmead, Australia

[3] School of Life Sciences, University of Technology Sydney, Ultimo, Australia

Published online: 25 March 2026

notions that are quite contrasting regarding how information is accessed, utilised, and applied.

Nonetheless, unprecedented advancements in technology designed for AI and learners who are digital natives raise serious ethical concerns regarding the practice of medicine. As one of the most advanced AI systems achieves pattern recognition, suggests diagnoses, and recommends treatments with increasing accuracy, the question arises as to whether future clinicians will possess the critical reasoning skills, clinical intuition, and empathetic engagement essential to quality medical care [6, 7].

The aim of this article is to explore the ethical considerations stemming from over-reliance on AI in medical teaching tools, with a specific focus on how digitally native students engage with algorithmic instructions. We analyse both positive effects and risks of this integration while developing an appropriate model for ethical application that respects fundamental humanitarian principles of medical practices.

1.1 Methodological approach

This article adopts a narrative review and analytical perspective approach rather than a systematic or scoping review. Relevant literature was identified through targeted searches of PubMed, Scopus, Education Resources Information Center (ERIC), and Google Scholar, focusing on publications from 2019 to 2025. Search terms included combinations of artificial intelligence, medical education, ethics, clinical reasoning, and empathy. Sources were selected based on relevance to ethical dimensions of AI use in medical education, with priority given to peer-reviewed reviews, empirical studies, and international guidance documents. The aim was not exhaustive coverage but conceptual synthesis to inform ethical reflection.

2 The promise and peril of AI integration

2.1 Benefits of AI in medical education

It draws the map, it gives the plan,

But doesn't see the trembling hand.

The use of AI in medical education has a unique potential to optimise educational outcomes and widen access to quality training. Adaptive approaches powered by AI technology facilitate the creation of tailored content and strategies for each student, enabling the acquisition of learning objectives and skills through [8, 9]. With the help of virtual patient simulators, students can practise clinical reasoning and decision-making in safe environments through repeated simulations of unusual but intricate cases that are infrequently encountered in traditional rotations [3]. In addition, AI technologies can automate evaluations, significantly improving accuracy, objectivity, and consistency while also reducing the time required to provide results to learners. Automated marking of practical assessments for clinical skills, interactions with standardised patients or even written papers will lessen assessors' workload while enhancing efficiency [4]. A more responsive assessment system enhances feedback loops, which accelerate skill acquisition while targeting areas of weaker performance relative to other evaluative strategies.

AI has the potential to mitigate cognitive biases relevant to medical education and practice. It can recommend differential diagnoses and evidence-based suggestions that help both learners and educators identify and overcome common cognitive oversights such as anchoring bias, confirmation bias or premature closure [10, 11]. This is especially useful for fostering critical thinking among novice learners who are more vulnerable to these cognitive biases. Perhaps most importantly, AI has the capacity to democratise access to expertise in low-resource areas. Learners located in rural or underserved regions have equal access to AI-assisted learning tools as their peers studying at well-resourced urban medical schools, which helps reduce educational inequities and advance global health [12, 13].

2.2 Ethical concerns and risks

It cannot read the subtle cues.

Of pain wrapped up in patient news,

Nor weigh a silence, held with care,

When stories live beyond what's there.

Along with the positive impacts, the increased use of AI in medical education raises several ethical issues that require thorough reflection. One of the most fundamental concerns is regarding the merging or loss of clinical intuition and the growing reliance on algorithmic frameworks for decision-making [14, 15]. The ethical analysis in this section is informed by established principles of medical ethics, beneficence, non-maleficence, autonomy, and justice, while also drawing on relational and virtue-based ethical perspectives relevant to education and professional formation.

Recognising and sensing when something is amiss with a patient, despite normal test results, or identifying patterns

warranting certain diagnoses, has always been regarded as 'the sign of mastery' of clinically expert practice [16]. This intuitive talent stems from encountering many patients and engaging in deep contemplation about the medical services provided. There is concern that students trained under the paradigm of automated recommendation will be unable to acquire this essential skill because they are accustomed to receiving prompt, algorithm-based instructions [14].

The empathy gap is yet another precise concern for ethics. The practice of medicine as a clinical profession has the distinguishing characteristic of requiring a personal connection with patients, as it involves alleviating their fears and providing comfort during periods of vulnerability [17, 18]. No matter how advanced AI systems might be, they cannot match the human ability to understand culture and deal with emotional intelligence empathetically. These interpersonal skills are vital in all areas of life, and students who become overly reliant on AI-mediated interactions may struggle to acquire these essential skills [19].

These issues add to an already overstretched collection of concerns viewed through the lens of ethics. Students tend to receive extreme confidence boosts when working with AI systems because algorithms give recommendations based on preset rules. This tends to suspend critical examination altogether, leading students to trust suggestions put forth by an algorithm excessively. In medicine, a field often marked by uncertainty, this false confidence is particularly challenging to manage, as coping with uncertainty is a crucial clinical competency that one should possess [5]. Students who become accustomed to interacting under drastically different circumstances due to the certainty provided by AIs will have a very low tolerance for the appropriate level of uncertainty. The application of AI in medical education widens equity and justice issues. It is possible that the trained datasets for AI systems do not sufficiently capture a wide range of patients, which results in reinforcing or perpetuating existing disparities in healthcare [20, 21]. Students taught using biased AIs may inadvertently provide substandard care to underrepresented groups. Moreover, the availability of sophisticated AI tools may deepen the already existing gaps in education and widen the disparity between well-resourced and under-resourced institutions [22].

3 Current evidence and illustrative observations

The use of AI in education has noteworthy implications, both positive and negative, as seen in its recent applications in the medical field. Some of the premier medical institutions have started using virtual patient simulations powered by AI to augment clinical teaching. Preliminary assessments indicate that students engaging with these platforms exhibit increased diagnostic accuracy and enhanced clinical reasoning skills compared to students participating in traditional case-based learning [1, 3].

Qualitative analysis of free-text responses from trainee doctors revealed concerns about the impact of AI technologies on their ability to develop clinical judgment and practical skills, highlighting potential challenges to autonomous practice [14]. One junior doctor reflected: "*Doctors would de-skill, when AI fails them the clinical skills would be the only thing they rely on….need to find the balance between both*".

Instructor observations confirm the concerns raised by students. Survey studies indicate that students who trained with AI tools inappropriately relied on them to guide their reasoning independently and tended to exhibit reduced self-assurance when not given the option of AI support [15]. Integrating AI ethics into medical education, while often guided by principle-based approaches, must also attend to virtue-based and relational dimensions of ethics—such as professional character, moral judgment, and responsibility—to prepare students for the ethical complexities of AI in clinical practice [7].

Cross-cultural studies reveal significant differences in how regions utilise AI technology. Some regions have adopted a more cautious approach to integrating AI into medical education, emphasising ethical literacy over technical instruction [6]. A cross-sectional study conducted in Beijing, China, investigated the understanding, trust, and attitudes that medical students hold towards AI in relation to medical education. The study revealed that, although students perceived AI as a valuable resource for educational efficiency, they expressed concerns related to privacy, data security, and ethics—particularly issues of data governance, accountability, and fairness in algorithmic decision-making. These factors shaped both their perceptions and acceptance of AI technologies within their academic settings [23].

4 Toward ethical AI integration

4.1 Guiding principles

Let it assist - but not define.

The fragile path of care and time.

The integration of an ethical approach regarding AI use in medical education needs to strike a balance between the humanistic values of medicine and the technologies that accompany it. We suggest four basic principles to foster

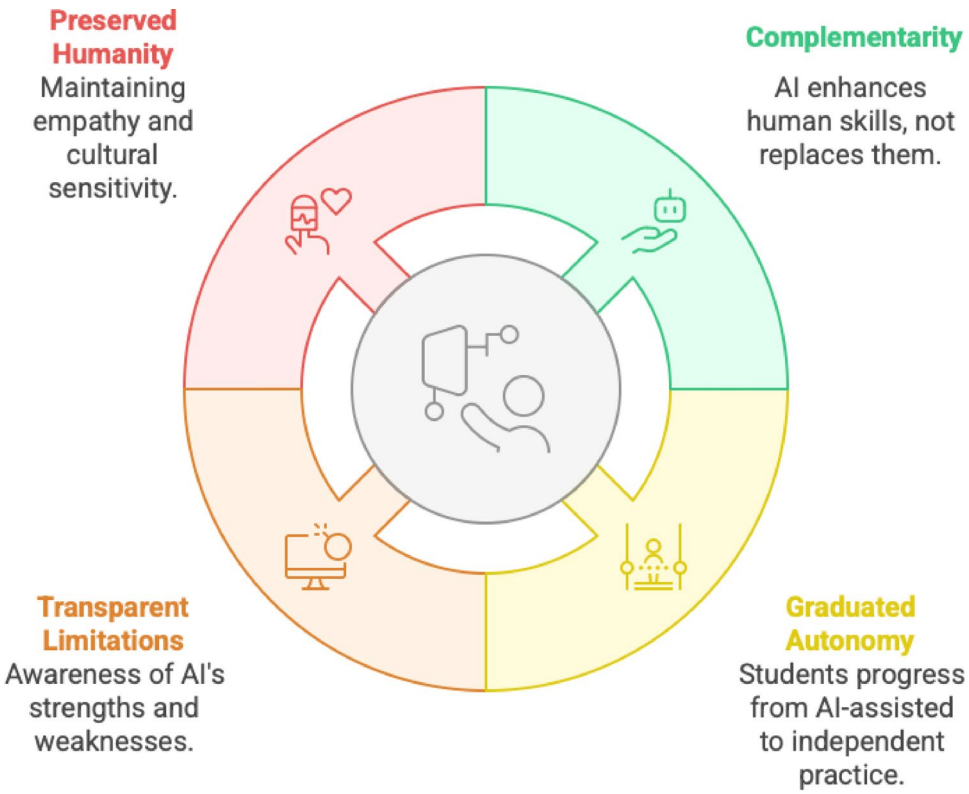

Fig. 1 Four guiding principles to integrate AI use in medical education

this integration, as illustrated in Fig. 1. These four guiding principles are not proposed as replacements for established ethical frameworks, but as applied educational translations of them. Complementarity aligns with beneficence and non-maleficence; graduated autonomy reflects respect for learner autonomy and professional development; transparent limitations relate to justice and accountability; and preserved humanity resonates with virtue ethics and relational ethics in healthcare education.

The principle of complementarity emphasises that AI should enhance, not replace, human clinical skills. This perspective acknowledges that there are distinct capabilities integral to effective medical practice, intuitive as well as analytical, and that AI's strengths lie in data fetching and processing while humans excel in empathy, vital creativity, and multi-layered ethical thought [1]. Programs aimed at training individuals should foster the development of both sets of capabilities in a complementary manner. This principle may be evaluated through assessments comparing student performance with and without AI assistance in clinical reasoning tasks.

As embodied in the principle of graduated autonomy, students should progress from more AI-assisted, structured learning contexts to gain increasing independence step by step. The early-stage use of AI tools accelerates the learning process; however, students must develop adequate confidence and competence in independent clinical reasoning [2]. This shift aligns with the evolution within traditional medical education, transitioning from supervised practice towards independent practice. Progression can be evaluated using staged assessments that progressively remove AI support across training phases. In practice, a staged curriculum model may involve early-phase students using AI for foundational knowledge acquisition and guided case exploration; mid-phase learners engaging in parallel AI-assisted and independent reasoning; and senior students expected to demonstrate diagnostic reasoning and patient communication without AI reliance during high-stakes assessments.

The principle of transparent limitations allows the students to be aware of both the strengths and weaknesses of any AI system. In particular, understanding the gaps in recommended responses, considering recommenders' biases, and judging if an AI-provided response is actually useful are crucial elements which make up the ability to evaluate algorithms [6]. System capabilities alongside AI literacy should be treated equally to traditional clinical skills. Evaluation may include learners' ability to articulate AI limitations during case discussions or examinations.

The principle of preserved humanity is the argument that technology should not erode essential human qualities of caring for patients. Educational curricula must still focus on developing empathy and its related components - cultural sensitivity, active listening, and applied ethics, understood here as the practical enactment of ethical reasoning in real clinical interactions rather than abstract moral theory [17]. These abilities will only grow in significance as more

analytically focused functions shift to AIs. Measures such as observed structured clinical examinations (OSCEs) and validated empathy scales may be used to assess this domain.

4.2 Practical recommendations

For Educators: Medical educators need to incorporate teaching about AI tools in ways that foster human-centred skills, ensuring they address both their capabilities and limitations. Evaluations should measure students' competencies in the use of AI alongside their clinical judgment as a stand-alone system. Regular, instruction-free clinical examinations should be conducted to reinforce the ability to diagnose algorithms independently [1]. Faculty members should also model the use of appropriate AI applications while also overseeing instructional overload and AI dependency. To effectively support learners in this space, faculty must prioritise their learning to focus on AI literacies first [2].

For Students: Questioning suggestions and understanding the rationale behind their algorithms requires active mental engagement at critical levels; therefore, it is essential to teach AI literacy so that medical students can learn and critically evaluate algorithmic recommendations. Encouraging self-explanation around clinical reasoning, as it diverges from an AI-dependent approach, enriches reflective practice [5]. Students must cultivate "analog" skills by performing patient interviews, taking histories, and communicating without relying on machines. The ability to tolerate ambiguity and make decisions with partial information frames is an essential remaining clinical skill [10].

For Technology Developers: Developers of AI systems must ensure transparency in the algorithmic processes involved, offering explanations that bolster rather than undermine rational evaluation. The user interface should be designed in a manner that does not invite unquestioning acceptance of recommendations [7]. Regarding the monitoring of discriminatory algorithm outputs, biases must be addressed from the earliest stages of conceiving and creating representative training datasets through to continuous assessment [20]. There are also information gaps regarding how displaying confidence intervals alongside limitations on system predictions can help users avoid overconfidence bias.

For Institutions: Medical schools need to formulate policies concerning AI applications that protect foundational principles of medical education while accommodating technological advancement. Issues related to academic integrity, privacy of information, and equitable distribution of AI resources should be covered [6]. Support from the institution should encompass monitoring the results of AI integration evaluations focused on students' clinical skills and lifelong learning competencies. Structures for expressing concern about the impact of AI on educational processes should be accessible to instructors and learners [24].

5 Conclusion

For even now, it's hard to see.

If care survives in circuitry.

Adopting artificial intelligence in medical education presents both a remarkable opportunity and a significant challenge within the healthcare field. As algorithm-driven navigation becomes increasingly common among students, the medical teaching community finds itself walking a tightrope between adopting innovation and preserving the humanistic aspects of quality medicine. The ethical aspects surrounding this shift extend far beyond concerns of pedagogical effectiveness; they delve deeply into philosophy regarding clinical skills, professional self-image, and the dynamics of the relationship between physicians and their patients. From our analysis, achieving successful integration will require an active initiative towards achieving complementarity, rather than dependence, where AI is employed to enhance, rather than degrade, human clinical function. Moving forward requires unrelenting vigilance, reassessment, and enduring resolve, while shaping aspiring practitioners who blend compassion with critical thought in healing, backed by technology without overshadowing the essential human ties that are fundamental to medicine.

Appendix: Complete poem

Uncertain hands: when the machine speaks

In quiet rooms where learners grow,

A new voice tells them what to know.

It's sharp, it's fast, it rarely slips -

A guide built strong with code and scripts.

It draws the map, it gives the plan,

But doesn't see the trembling hand.

No look, no pause, no shifting eye -

Just patterns clear and neat and dry.

It cannot read the subtle cues

Of pain wrapped up in patient news,

Nor weigh a silence, held with care,

When stories live beyond what's there.

One student paused - unsure, afraid -

The screen had spoken, plans were laid.

But something felt not quite complete -

The face in front, too still, too sweet.

What if it misses culture's thread,

A quiet fear, a word not said?

Not all is written, not all is shared -

And some are harmed when none have cared.

We trust the code, we trust the speed,

But what of things machines don't read?

The weight of doubt, the space to feel -

Are these the cost of something real?

Let it assist - but not define

The fragile path of care and time.

For even now, it's hard to see

If care survives in circuitry.

Original poem by author, exploring themes central to this analysis

Author contributions CRediT author statementConceptualization: ENCL, CEDLData curation: ENCLFormal analysis: ENCLFunding acquisition: CEDLInvestigation: ENCLMethodology: ENCL, CEDLProject administration: ENCLSupervision: CEDLVisualization: ENCL, CEDLWriting original draft: ENCLWriting-review & editing: ENCL, CEDL.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Nil.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest None of the authors declared Financial and Non-Financial Relationships and Activities, and Conflicts of Interest regarding this manuscript.

Ethical approval Waived for narrative review—Non-applicable.

References

1. Gordon, M., et al.: A scoping review of artificial intelligence in medical education: BEME Guide 84. Med. Teach. **46**, 446–470 (2024). https://doi.org/10.1080/0142159X.2024.2314198
2. Tolsgaard, M.G., et al.: The fundamentals of Artificial Intelligence in medical education research: AMEE Guide 156. Med. Teach. **45**, 565–573 (2023). https://doi.org/10.1080/0142159X.2023.2180340
3. Knopp, M.I., et al.: AI-enabled medical education: threads of change, promising futures, and risky realities across four potential future worlds. JMIR Med. Educ. **9**, e50373 (2023). https://doi.org/10.2196/50373
4. Sapci, A.H., Sapci, H.A.: Artificial intelligence education and tools for medical and health informatics students: systematic review. JMIR Med. Educ. **6**, e19285 (2020). https://doi.org/10.2196/19285
5. Lerchenfeldt, S., et al.: Twelve tips for interfacing with the new generation of medical students: iGen. Med. Teach. **43**, 1249–1254 (2021). https://doi.org/10.1080/0142159X.2020.1845305
6. Masters, K.: Ethical use of artificial intelligence in health professions education: AMEE Guide 158. Med. Teach. **45**, 574–584 (2023). https://doi.org/10.1080/0142159X.2023.2186203
7. Weidener, L., Fischer, M.: Proposing a principle-based approach for teaching AI ethics in medical education. JMIR Med. Educ. **10**, e55368 (2024). https://doi.org/10.2196/55368
8. Abd-Alrazaq, A., et al.: Large language models in medical education: opportunities, challenges, and future directions. JMIR Med. Educ. **9**, e48291 (2023). https://doi.org/10.2196/48291
9. Lee, J., Wu, A.S., Li, D., Kulasegaram, K.M.: Artificial intelligence in undergraduate medical education: a scoping review. Acad. Med. **96**, 62–S70 (2021). https://doi.org/10.1097/ACM.0000000000004054
10. Wang, J., Redelmeier, D.A.: Cognitive biases and artificial intelligence. NEJM AI. **1**, AIcs2400639 (2024). https://doi.org/10.1056/AIcs2400639
11. Siala, H., Wang, Y.: SHIFTing artificial intelligence to be responsible in healthcare: a systematic review. Soc. Sci. Med. **296**, 114782 (2022). https://doi.org/10.1016/j.socscimed.2022.114782
12. Masters, K.: Artificial intelligence in medical education. Med. Teach. **41**, 976–980 (2019). https://doi.org/10.1080/0142159X.2019.1595557
13. Yakhkind, A., Cooney, R.: Generation Z in medical education: not your parent's doctors anymore. In *Transformations of Medical Education and Practice Impacting Critical Care in the New Millennium* 1–15Springer, (2024). https://doi.org/10.1007/978-3-031-69686-2_1
14. Banerjee, M., et al.: Impact of artificial intelligence on clinical training: perceptions of postgraduate trainee doctors in London, UK hospitals. BMC Med. Educ. **21**, 429 (2021). https://doi.org/10.1186/s12909-021-02870-x
15. Civaner, M.M., et al.: Artificial intelligence in medical education: a cross-sectional needs assessment. BMC Med. Educ. **22**, 772 (2022). https://doi.org/10.1186/s12909-022-03852-3
16. Jackson, P., et al.: Artificial intelligence in medical education—perception among medical students. BMC Med. Educ. **24**, 804 (2024). https://doi.org/10.1186/s12909-024-05760-0

17. Yang, F., et al.: Qualitative insights into empathy in medical education: perspectives from students, doctors, and educators. BMC Med. Educ. **24**, 1337 (2024). https://doi.org/10.1186/s12909-024-05337-x
18. Andersen, F.A., et al.: Revisiting the trajectory of medical students' empathy, and impact of gender, specialty preferences and nationality: a systematic review. BMC Med. Educ. **20**, 52 (2020). https://doi.org/10.1186/s12909-020-1964-5
19. Fritsch, S.J., et al.: Attitudes and perception of artificial intelligence in healthcare: a cross-sectional survey among patients. Digit. Health. **8**, 20552076221116772 (2022). https://doi.org/10.1177/20552076221116772
20. Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S.: Dissecting racial bias in an algorithm used to manage the health of populations. Science. **366**, 447–453 (2019). https://doi.org/10.1126/science.aax2342
21. Chen, H., Xuan, H., Cai, J., Liu, M., Shi, L.: The impact of empathy on medical students: an integrative review. BMC Med. Educ. **24**, 455 (2024). https://doi.org/10.1186/s12909-024-05448-5
22. Plochocki, J.H.: Several ways generation Z may shape the medical school landscape. J. Med. Educ. Curric. Dev. **6**, 2382120519884325 (2019). https://doi.org/10.1177/2382120519884325
23. Duan, S., Liu, C., Rong, T., et al.: Integrating AI in medical education: a comprehensive study of medical students' attitudes, concerns, and behavioral intentions. BMC Med. Educ. **25**, 599 (2025). https://doi.org/10.1186/s12909-025-07177-9
24. Grunhut, J., Marques, O., Wyatt, A.T.M.: Needs, challenges, and applications of artificial intelligence in medical education curriculum. JMIR Med. Educ. **8**, e35587 (2022). https://doi.org/10.2196/35587