

Generative AI in medical education: a milestone or tombstone?

Fanny Sunesson 

Department of Internal Medicine, Landskrona Lasarett, Landskrona, Sweden

Abstract

With the bloom of AI tools happening in front of us, did we, the medical community, even had a moment to assess the potential impact of AI on the professional skills and knowledge of our future doctors? As medical students continue to “de-skill” themselves by outsourcing critical reading and thinking to AI tools, how will they differentiate a summary of medical facts from a summary of medical-sounding AI-generated lies? How will they handle advising patients who are confused by the suggestions generated by their gen-AI “medical consultants”? Rather than using gen AI as a study tool or incorporating gen AI technologies in medical education, we seem to desperately need to incorporate into our lives some education about AI usage.

Keywords: medical education • medical school • AI • artificial intelligence

Citation

Sunesson F. Generative AI in medical education: a milestone or tombstone? Eur J Transl Clin Med. 2026;9(2):00-00.

DOI: [10.31373/ejtcmm/241682](https://doi.org/10.31373/ejtcmm/241682)

Thus the first ultraintelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control. It is curious that this point is made so seldom outside of science fiction. It is sometimes worthwhile to take science fiction seriously.

– Irving John Good (1966) [1]

Have you ever been fooled by artificial intelligence (AI)? For example, when reading news briefs or watching short video recordings on your smartphone? You are not alone. A terri-

ble example of this was when 3 researchers from the Karolinska Institutet (Sweden) appeared in fake AI-generated videos that were spread to patients across social media platforms. Among them was prof. Kerstin Bismar, a well-known diabetes researcher, who appeared to be convincing patients to stop taking their prescribed drugs (because they were “dangerous”) and instead to buy “the diabetes cure” offered online [2]. Unfortunately, researchers confirmed that deception of humans is one of the skills acquired by AI [3]. Despite this (and the well-publicized lawsuits regarding the involvement of gen

Corresponding author:

Fanny Sunesson, Department of Internal Medicine, Landskrona Lasarett, Landskrona, Sweden

e-mail: fanny.sunesson@skane.se

Available online: ejtcmm.gumed.edu.pl

Copyright © Medical University of Gdańsk



AI in its users' suicides), increasingly more people stop searching for information and instead ask their favorite generative (gen) AI tool to find and summarize it for them [4-5]. Among those people are patients and their doctors, along with medical students and their teachers [6].

There are many questions: do we know when and how to use AI in medical education? Certainly AI and gen AI have changed the landscape in which we teach and learn medicine, but do we, the medical community, know how to work in this new landscape? Do medical students have the skills necessary to use gen AI in an ethical and purposeful way? Do medical schools have tools and procedures for when AI is used to cheat on exams? With the bloom of AI tools happening in front of us, did the medical community even had a moment to assess the potential impact of AI on the professional skills and knowledge of our future doctors? According to some critics, the problem lies elsewhere, because we didn't know how to teach medicine effectively in the pre-ChatGPT era either [7].

One of the earliest uses of AI in medicine was processing radiological images and facilitating their interpretation [8]. It has since rapidly evolved to interpret laboratory test results, calculate drug dosages, accelerate the discovery of new drug molecules, prepare complex treatment plans and even assist in robotic surgery. Medical students interact with AI in their classroom as well, e.g. in the form of the increasingly more human-like mannequins that simulate patients at the bedside or the automated "diagnoses" that often are printed on the upper margin of 12-lead electrocardiograms.

As I write these words, not even 3 years have passed since November 30th 2022, the day the OpenAI company publicly released its Chat Generative Pre-trained Transformer (ChatGPT). Now there are exponentially more gen AI tools available online. Each of the popular gen AI tools (e.g. ChatGPT, Gemini (Google LLC) and Copilot (Microsoft)) is able to generate full-length essays (or scientific articles) with references that can reasonably pass the scrutiny of existing plagiarism detection tools. Apps such as *ScreenApp*, *Jamworks* and *Mindgrasp* record both online and in-person lectures, which they can automatically transcribe word-for-word and then can use the transcript to generate notes, flashcards, practice questions and individual study plans. If studying without having to take any notes sounds strange to you, then unfortunately there are more bad news. Thanks to gen AI, solving test questions has never been easier: simply photograph the multiple-choice (MCQ) or open-ended question and the gen AI "study" tool (e.g. *Mindko* (formerly *College Tools*), *Socratic* or *Wolfram Alpha*) will provide you with an answer. Some companies decided to skip the innocent-sounding image of a "study tool" and instead they directly market their AI tools as "cheating aids" (e.g. *CheatMate*).

Exams and tests assessing qualifications are an essential part of ensuring competence and safety, and not only in the

medical profession. Undoubtedly there has been cheating as long as there have been tests, but the difference is that gen AI not only makes cheating effortless, but also more sophisticated and organized. Cheating during a test does matter because as soon as one test-taker decides to cheat, it reduces the value of the test for the rest, who are performing honestly. Exams are supposed to be tools for evaluating knowledge, but currently there is a risk that they instead evaluate the gen AI's "knowledge" and the exam taker's ability to use that technology. This raises another question: do we know how to accurately and fairly evaluate students' knowledge (and by extension the effectiveness of teaching) in the AI era?

To be fair, some AI tools are also used to prevent and detect cheating, e.g. plagiarism checkers and real time lockdown browsers. Attempts have also been made using advanced algorithms to alert faculty when open-ended answers (or essays) do not match the student's word choice in previous written work or if recent test results deviate significantly from previous academic achievements [9]. Hence, gen AI has equipped both sides in the generations-long cat & mouse game involving educators and students.

Data from the United States suggest that the percentage of high school students who cheat is high (~ 60-70%) but remains unchanged, despite technological advancements brought by AI [10-11]. Unfortunately, there is little new and publicly available data specific to cheating during medical studies [12-15]. Frankly, I am not surprised that medical schools are not exactly in a hurry to publish data on this topic. However, as we all know it is not the quantity but the quality that matters and AI-supported cheating generates higher test results (pun definitely intended). Consequently, more students receiving credit for medical knowledge and clinical reasoning skills they do not have, leads to score/grade inflation and more difficult questions on future editions of the tests, both of which directly increase the pressure on the honest test-takers.

At the core of all university-level learning are tasks such as reading dozens upon dozens of pages of text, interpreting that information (and what was said in class) in order to understand it, identifying key points, summarizing information into notes or flashcards and solving problems by reasoning and/or creativity. Gen AI takes away from us each of those processes every single time we ask it to summarize a lecture we never listened to or a book chapter we never read. It seems that a handful of high-tech gen AI companies succeeded in convincing millions of people that the above-mentioned mental tasks that lead to our learning are not only "a waste of time" but that they should be outsourced to internet-powered machines. Is it indeed true that such "efficiency" is more valuable than our learning?

Looking from a broader perspective, perhaps we shouldn't be so surprised. After all, each of the previous changes in

the learning tradition has stirred concern, e.g. the shift from the storytelling and discussions to writing. Currently we are seeing paper books competing with their electronic (and audio-only) versions along with the rest of the internet and gen AI. Kwame Anthony Appiah described this in more detail (and more elegantly) as the process of “de-skilling” [16]. It is well-known that commercially available gen AI tools tend to provide incomplete, partially correct answers and sometimes even the so-called hallucinations (confabulated information, lies) [17]. As medical students continue to “de-skill” themselves by outsourcing critical reading and thinking to AI tools, how will they differentiate a summary of medical facts from a summary of medical-sounding AI-generated lies? How will they handle advising patients who are confused by suggestions generated by their gen-AI “medical advisors”?

Researchers found that even when ChatGPT answered MCQs correctly, the explanations it generated for its answers were shallow at best and lacked pathophysiological insight [18]. Another study concluded that AI-generated answers were often biased, e.g. radiological imaging was more often recommended for white than for black patients, despite the same presentations of symptoms [19]. In other words, if we use AI for learning, the biases of gen AI can easily become our biases. Rather than using gen AI as a study tool or incorporating gen AI technologies in medical education, we seem to desperately need to incorporate into our lives some education about AI usage.

There are some pioneers, e.g. Harvard Medical School or Örebro University (Sweden), which have already implemented into their curricula courses about AI in health care [20-21]. Educating future doctors about gen AI is arguably much more important than implementing AI tools in their training itself, because without understanding AI technology's uses and limitations, they will not be able to safely use it when making evidence-based clinical decisions. After all, we don't leave medical students in the classroom by themselves with a box of peripheral and central intravenous catheters and hope that sooner or later they will “figure out” how to safely and correctly use them. We don't do that because that would be irresponsible and potentially harmful to their health (not to mention the legal liability). Why then are we, the medical community, so irresponsible with AI tools?

Among the discussions regarding what should (or should not be) considered cheating or if and how to use AI, stands the patient. There are a plethora of studies about the ethical concerns about implementing AI, but what does seeking advice from an AI-educated doctor actually mean to the patient? Will we lose what Hippocrates once described as “the art of medicine”? A 2025 study showed that participants preferred the AI-generated responses to medical questions and considered them “more compassionate” than those written by the human doctors [22]. Even if gen AI can be convincing

in sounding empathetic, people tend to be more receptive to answers/advice when they think it is provided by a fellow human, even if in fact it was generated by AI [22]. Another team of researchers arrived at more alarming conclusions: patients are not able to distinguish the accuracy of medical advice provided by gen-AI and place far too much trust in it, leading to potentially harmful decisions [23]. Beneath the many layers of this “AI in medicine” dilemma, more question inevitably arises: what makes a good doctor in today's society? Is it the doctor that has the most information readily available “in his/her head” or is it the one who can quickly use technology to find the information that is needed at the moment? My elementary school math teacher adamantly insisted that we calculate as much arithmetic as possible without using calculators. In the current AI-era I thank her more than ever for teaching that skill.

Multiple data show that starting in the mid-2000s education measures such as basic literacy, numeracy, problem solving, creativity, memory, attention span and general cognitive performance either plateaued or declined among adolescents in many Western nations, despite the fact that they attended more classroom hours than their peers in the previous generations [24-28]. Of note, the mid-2000s was the time of the rapid and wide-scale digital transformation in the classrooms around the world. About 16 years ago my country's government (Sweden) widely introduced digital tools such as tablets into classrooms, including even some preschools. Since 2023 that same government has been investing significant money and effort into bringing paper books and handwriting back to those same classrooms in order to reverse the students' decline in attention span and reading skills [29]. Recently, nearly 1400 AI engineers, researchers and co-founders (definitely not neo-Luddites) from the leading high-tech companies in the USA and Europe signed an open letter calling for an international and institutional effort to control the pace and direction of AI development [30]. Just a few weeks ago, public school authorities in New York City issued a moratorium on the use of all gen AI tools in their kindergarten and elementary school classrooms [31].

One does not need to be an expert in education to see these signs and trends. Yet we appear so passive and assume that the technologies that diminish the learning outcomes of elementary and high students do not have the same effects on the brains of medical students. It seems that we fell in love with digital technologies in the classrooms a bit too quickly. But what does that mean for medical schools: do we need a return to the roots, e.g. paper-pencil tests and oral exams?

I still don't have the answer to the question I asked in the title of this editorial.

I also see that we, the medical community, are not answering it either.

References

1. Good IJ. Speculations Concerning the First Ultraintelligent Machine. In 1966. p. 31–88. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0065245808604180>
2. Forskaren utnyttjad i bluffhärra: "Det är livsfarligt" [Internet]. Sweden: Nyheterna TV4; 2025. Available from: <https://www.tv4.se/artikel/7zcuB7MisQaEGsOkQDBw6N/prisade-professorn-utnyttjad-i-bluffhaerva-det-aer-livsfarligt>
3. Park PS, Goldstein S, O'Gara A, Chen M, Hendrycks D. AI deception: A survey of examples, risks, and potential solutions. *Patterns* [Internet]. 2024;5(5):100988. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S266638992400103X>
4. Kaiser C, Kaiser J, Schallner R, Schneider S. A New Era of Online Search? A Large-Scale Study of User Behavior and Personal Preferences during Practical Search Tasks with Generative AI versus Traditional Search Engines. In: *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* [Internet]. New York, NY, USA: ACM; 2025. p. 1–7. Available from: <https://dl.acm.org/doi/10.1145/3706599.3720123>
5. Carter C. AI search vs. Google. AI search traffic, users, and market share compared to Google, as of Q1 2026 [Internet]. Wix Studio. 2026 [cited 2026 Oct 14]. Available from: <https://www.wix.com/studio/ai-search-lab/research/ai-search-vs-google>
6. Shahsavari Y, Choudhury A. User Intentions to Use ChatGPT for Self-Diagnosis and Health-Related Purposes: Cross-sectional Survey Study. *JMIR Hum Factors* [Internet]. 2023;10:e47564. Available from: <https://humanfactors.jmir.org/2023/1/e47564>
7. Rodigin A. Is medicine ready for ChatGPT – why not just ask ChatGPT? *Eur J Transl Clin Med* [Internet]. 2023;6(1):5–8. Available from: <https://ejtcm.gumed.edu.pl/articles/162647>
8. Avanzo M, Stancanelli J, Pirrone G, Drigo A, Retico A. The Evolution of Artificial Intelligence in Medical Imaging: From Computer Science to Machine and Deep Learning. *Cancers (Basel)* [Internet]. 2024 Nov 1;16(21):3702. Available from: <https://www.mdpi.com/2072-6694/16/21/3702>
9. Coldwell W. I received a first but it felt tainted and undeserved': inside the university AI cheating crisis. *Guard* [Internet]. 2024; Available from: <https://www.theguardian.com/technology/2024/dec/15/i-received-a-first-but-it-felt-tainted-and-undeserved-inside-the-university-ai-cheating-crisis>
10. Spector C. What do AI chatbots really mean for students and cheating? [Internet]. Stanford. Graduate School of Education. 2023 [cited 2026 Oct 14]. Available from: <https://ed.stanford.edu/news/what-do-ai-chatbots-really-mean-students-and-cheating>
11. Lee VR, Pope D, Miles S, Zárate RC. Cheating in the age of generative AI: A high school survey study of cheating behaviors before and after the release of ChatGPT. *Comput Educ Artif Intell* [Internet]. 2024;7:100253. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2666920X24000560>
12. Kusnoor A V., Falik R. Cheating in Medical School. *South Med J* [Internet]. 2013 Aug;106(8):479–83. Available from: <http://sma.org/southern-medical-journal/article/cheating-in-medical-school-the-unacknowledged-ailment>
13. Abdulghani HM, Haque S, Almusalam YA, Alanezi SL, Alsulaiman YA, Irshad M, et al. Self-reported cheating among medical students: An alarming finding in a cross-sectional study from Saudi Arabia. *Manalo E, editor. PLoS One* [Internet]. 2018 Mar 29;13(3):e0194963. Available from: <https://dx.plos.org/10.1371/journal.pone.0194963>
14. Monteiro J, Silva-Pereira F, Severo M. Investigating the existence of social networks in cheating behaviors in medical students. *BMC Med Educ* [Internet]. 2018;18(1):193. Available from: <https://bmcmmeduc.biomedcentral.com/articles/10.1186/s12909-018-1299-7>
15. Dar UF, Khan YS. Self-Reported Academic Misconduct among Medical Students: Perception and Prevalence. *Martins CHG, editor. Sci World J* [Internet]. 2021 Aug 23;2021:1–8. Available from: <https://www.hindawi.com/journals/tswj/2021/5580797/>
16. Appiah KA. The Age of De-Skilling Will AI stretch our minds—or stunt them? [Internet]. *The Atlantic*. 2025 [cited 2026 Oct 14]. Available from: <https://www.theatlantic.com/ideas/archive/2025/10/ai-deskilling-automation-technology/684669/>
17. Saeidnia HR, Jahani S, Ghiasi N, Keshavarz H. Generative AI and health misinformation: production, propagation, and mitigation—a systematic review. *BMC Public Health* [Internet]. 2026;26(1):693. Available from: <https://link.springer.com/10.1186/s12889-025-26148-9>
18. Furfaro D, Celi LA, Schwartzstein RM. Artificial Intelligence in Medical Education. *Chest* [Internet]. 2024;165(4):771–4. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0012369223058208>
19. Zack T, Lehman E, Suzgun M, Rodriguez JA, Celi LA, Gichoya J, et al. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. *Lancet Digit Heal* [Internet]. 2024;6(1):e12–22. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S258975002300225X>

20. Gehrman E. How Generative AI Is Transforming Medical Education [Internet]. Harvard Medicine. The Magazine of Harvard Medical School. 2024 [cited 2026 Oct 14]. Available from: <https://magazine.hms.harvard.edu/articles/how-generative-ai-transforming-medical-education>
21. Krantz M. AI to be included in medical students' curriculum [Internet]. Örebro University. 2025 [cited 2026 Oct 14]. Available from: <https://www.oru.se/english/news/news-archive/news-archive-2025/ai-to-be-included-in-medical-students-curriculum/>
22. Franke Föyen L, Zapel E, Lekander M, Hedman-Lagerlöf E, Lindsäter E. Artificial intelligence vs. human expert: Licensed mental health clinicians' blinded evaluation of AI-generated and expert psychological advice on quality, empathy, and perceived authorship. *Internet Interv* [Internet]. 2025;41:100841. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2214782925000429>
23. Shekar S, Pataranutaporn P, Sarabu C, Cecchi GA, Maes P. People Overtrust AI-Generated Medical Advice despite Low Accuracy. *NEJM AI* [Internet]. 2025;2(6). Available from: <https://ai.nejm.org/doi/10.1056/Aloa2300015>
24. Mullis I, von Davier M, Foy P, Fishbein B, Reynolds K, Wry E. PIRLS 2021 International Results in Reading [Internet]. 2023. Available from: <http://pirls2021.org/results>
25. PISA 2022 Results (Volume I) [Internet]. OECD Publishing; 2023. (PISA). Available from: https://www.oecd.org/en/publications/pisa-2022-results-volume-i_53f23881-en.html
26. Andrzejewski D, Zeilinger EL, Pietschnig J. Is there a Flynn effect for attention? Cross-temporal meta-analytical evidence for better test performance (1990–2021). *Pers Individ Dif* [Internet]. 2024;216:112417. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0191886923003409>
27. Kim KH. Creativity Crisis Update: America Follows Asia in Pursuing High Test Scores Over Learning. *Roeper Rev* [Internet]. 2021;43(1):21–41. Available from: <https://www.tandfonline.com/doi/full/10.1080/02783193.2020.1840464>
28. Dutton E, van der Linden D, Lynn R. The negative Flynn Effect: A systematic literature review. *Intelligence* [Internet]. 2016;59:163–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0160289616300198>
29. Foster B. Sweden confirms it: screens are directly linked to falling student performance [Internet]. Glass Almanac. 2025 [cited 2026 Oct 14]. Available from: <https://glassalmanac.com/sweden-confirms-it-screens-are-directly-linked-to-falling-student-performance/>
30. Pacing the Frontier. A statement from 1386 employees of frontier AI companies. [Internet]. 2026 [cited 2026 Oct 14]. Available from: <https://www.pacingthefrontier.com/>
31. Samuels J. Mayor Mamdani and Chancellor Samuels Put Students First with Nation's Broadest Generative AI Moratorium in Schools [Internet]. NYC. Office of the Mayor. 2026 [cited 2026 Oct 14]. Available from: <https://www.nyc.gov/mayors-office/news/2026/09/mayor-mamdani-and-chancellor-samuels-put-students-first-with-nat>