



ChatGPT and Generative Artificial Intelligence for Medical Education: Potential Impact and Opportunity

Christy K. Boscardin, PhD, Brian Gin, MD, PhD, Polo Black Golde, and Karen E. Hauer, MD, PhD

Abstract

ChatGPT has ushered in a new era of artificial intelligence (AI) that already has significant consequences for many industries, including health care and education. Generative AI tools, such as ChatGPT, refer to AI that is designed to create or generate new content, such as text, images, or music, from their trained parameters. With free access online and an easy-to-use conversational interface, ChatGPT quickly accumulated more than 100 million users within the first few months of its launch. Recent headlines in

the popular press have ignited concerns relevant to medical education over the possible implications of cheating and plagiarism in assessments as well as excitement over new opportunities for learning, assessment, and research. In this Scholarly Perspective, the authors offer insights and recommendations about generative AI for medical educators based on literature review, including the AI literacy framework. The authors provide a definition of generative AI, introduce an AI literacy framework and competencies, and

offer considerations for potential impacts and opportunities to optimize integration of generative AI for admissions, learning, assessment, and medical education research to help medical educators navigate and start planning for this new environment. As generative AI tools continue to expand, educators need to increase their AI literacy through education and vigilance around new advances in the technology and serve as stewards of AI literacy to foster social responsibility and ethical awareness around the use of AI.

Recent headlines in the popular press and social media have created palpable excitement and worry over the advent of generative artificial intelligence (AI) technologies, such as ChatGPT, Bard, LLaMA, DALL-E Bing, and Stable Diffusion, in many domains, including health care and education. For educators, the availability of ChatGPT has raised alarm because of implications for assessment and the potential for a new mechanism of cheating and plagiarism. ChatGPT can generate responses to written prompts and accurately answer multiple-choice questions.¹ In a recent Op-Ed in *Inside Higher Ed*, Weissman² declared that “what winter of 2020 was for COVID-19, winter of 2023 is for ChatGPT, and higher education will never

be the same.” Some large school districts across the United States, including New York City’s Department of Education and Los Angeles Unified School Districts, temporarily blocked access to ChatGPT and the OpenAI website to give some space and time for educators to adapt to this new environment. After initial concern around the negative implications of ChatGPT, there is some cautious optimism for its potential utility and resignation that ChatGPT is here to stay.

As we embark on this new era of technological breakthroughs, in this Scholarly Perspective, we provide a definition of generative AI, introduce an AI literacy framework and competencies, and offer considerations for potential impacts and opportunities to optimize integration of generative AI, such as ChatGPT, for admissions, learning, assessment, and medical education research to help medical educators navigate and start planning for this new environment. Although we focus on these 4 domains with specific examples for illustrative purposes, framing this discussion around the immediate impact and looking ahead for opportunities for optimization should extrapolate across various other activities within medical education.

What Are ChatGPT, Chatbots, and Generative AI?

ChatGPT (Chat Generative Pre-Trained Transformer) is an AI tool released November 20, 2022, by OpenAI as a web-based chatbot. Chatbots are AI-based software designed to mimic human conversation through text or audio, providing natural language responses to human input in a conversational format. Although the language-predictive technology behind ChatGPT has been available since at least 2018, ChatGPT provided an easily accessible and simple-to-use interface leveraging existing AI technology. The underlying AI algorithms behind ChatGPT, OpenAI’s GPT-3/4 models, are large language models (LLMs) capable of generating text responses based on patterns learned from text data and more recently from images (with GPT-4, released March 14, 2023). These patterns are encoded within the LLM’s underlying *deep learning neural network*—an electronically modeled collection of artificial neurons with multiple layers that mimic the cortex of the human brain. Training has the effect of tuning the interactions between artificial neurons, and these interactions are stored as numerical values also known as *parameters*. *Generative AI* is a term used to describe models (such as LLMs) that generate new content, such as text, images,

Please see the end of this article for information about the authors.

Correspondence should be addressed to Christy K. Boscardin, Departments of Medicine and Anesthesia, UCSF Office of Medical Education, Box 0710, 521 Parnassus Ave., Mail Room 0104, San Francisco, CA 94143-0710; telephone: (415) 519-3570; email: christy.boscardin@ucsf.edu.

Acad Med. 2024;99:22–27.

First published online August 31, 2023

doi: 10.1097/ACM.0000000000005439

Copyright © 2023 the Association of American Medical Colleges.

An Academic Medicine Podcast episode featuring this article is available wherever you get your podcasts.

or music, from their trained parameters. Such models can also be updated based on user feedback but in general do not self-update or record prior conversations as long-term memory. Generative AI, such as ChatGPT, is trained across Internet resources, published texts, and images to generate responses to user-provided prompts in a variety of conversational or academic styles, including journal papers, academic presentations, and various other formats (e.g., poems, blogs, programming codes, and scripts). ChatGPT is available free of charge and garnered more than 100 million users within the first few months of its launch.³ Other LLMs available at the time of this publication include Google's BERT (one of the first LLMs) and PaLM 2 (used in its Bard chatbot), and Facebook's LLaMA (which can be used to create chatbots on local computers without requiring a cloud connection).

As an illustration, ChatGPT generated this text in response to the prompt, "What are the potential benefits and drawbacks to ChatGPT for medical education?"

Potential benefits of using ChatGPT for medical education include its ability to provide personalized and interactive learning experiences, as well as its ability to simulate real-world scenarios and patient interactions. Additionally, ChatGPT can help students practice and improve their communication skills, which are essential for being an effective health care provider. Potential drawbacks of using ChatGPT for medical education include the limitations of its current knowledge base, which may not be up-to-date with the latest medical research and best practices. Additionally, ChatGPT is not able to provide hands-on experience or physical examination, which are important aspects of medical education. Lastly, there is a risk that students may rely too heavily on ChatGPT and not develop the critical thinking skills necessary to make accurate diagnosis and treatment decisions.

This text generated by ChatGPT constituted a reasonable and plausible response to the prompt. However, evaluation of the accuracy, credibility, and adequacy of the response requires human judgment. A notable drawback highlighted in this response is the acknowledgment that ChatGPT may not be up-to-date and accurate in current research and medical practice. Educators also need to be aware that ChatGPT's other significant limitations include the potential for bias in the sources and data that it draws from, leading to generation of

inaccurate or insufficient information. The accuracy and validity of information generated depend on the accuracy of the data that the model is trained on.⁴ Despite these limitations, with thoughtful planning, generative AI can be a potentially useful resource for medical education. Therefore, educators need to understand the limitations in addition to possible opportunities to minimize potential negative impact and maximize benefits. Grounded in an AI literacy framework and AI competencies for educators, this article raises awareness of implications of generative AI for medical education and provides clear actionable steps for educators.

AI Literacy Framework and AI Competencies

Expanding on the idea of digital literacy, AI literacy has emerged in recent years as an essential skill within multiple disciplines and industries.⁵ From the perspective of health professions education, AI literacy includes (1) understanding the capabilities of AI, including tools to enhance health care; (2) integrating AI tools into teaching; and (3) ensuring inclusion, equity, and ethically responsible use of AI for societal good. A recent study by Gong et al⁶ in China found that, despite being exposed to AI as part of the core curriculum, most students lacked understanding of basic AI ethics or consideration for intellectual property associated with AI, highlighting the current curricular gap around AI ethics. More recently, Russell et al⁷ also highlighted social and ethical implications as part of 6 core competencies for the use of AI-based clinical tools for health care professionals. The other 4 competencies include basic knowledge of AI, workflow analysis for AI-based tools, AI-enhanced clinical encounters, and evidence-based evaluation of AI-based tools. These competencies and the literature supporting the integration of AI as part of medical education curricula emphasize the importance of knowing the basics of AI and recognizing the risks around data security and accuracy. Being able to evaluate the appropriate use of AI is a new core competency for health care professionals.^{8,9}

Drawing on these AI literacy and AI competency frameworks, we first provide a brief description of AI capabilities and the immediate impact on medical

education, including potential ethical issues. We then discuss opportunities afforded by generative AI in 4 areas of medical education that represent the major programmatic phases: (1) selection into medical school and training programs, (2) learning, (3) assessment, and (4) research to illustrate the potential for diverse and broad implications throughout medical education. For each of the 4 areas in which we foresee the most immediate impact, we identified 1 specific example activity for illustrative purposes.

To generate these insights and recommendations, the authors engaged in discussions among themselves and at curriculum committee meetings with other educators. C.K.B. is a research methodologist with expertise in AI and educational statistics. B.G. is a pediatric hospitalist and doctoral candidate in health professions education with expertise in AI and machine learning. P.B.G. is a director of data and analytics and leads a technical team overseeing the implementation of AI for a school of education. K.E.H. is an associate dean for assessment, an experienced medical education researcher, and a practicing general internist.

Selection Into Medical School and Training Programs

Aim

Despite some conflicting evidence around the predictive validity of application essays for future performance,¹⁰ personal statements are used widely as qualitative information during medical student and graduate medical education selection. Admissions and selection committees often use personal statements or responses to other open-ended questions to discern something distinguishing about each applicant as a person and ensure that applicants' writing and analytic thinking skills are sufficient for success in the program.¹¹

Potential impact

Shortly after the release of ChatGPT, one article¹² reported that ChatGPT was able to generate 2 different plausible college essays to 2 different application essay prompts in approximately 20 minutes. Responses to this report cited clear limitations of ChatGPT-generated essays as being "impersonal, standard, and 'soulless.'"¹³ Guo et al¹⁴ also suggested that

the major differences between human- and ChatGPT-generated texts are the lack of subjective expressions and the more formal, objective language in texts generated by ChatGPT. Unlike responses to an exam prompt requiring an academic style of writing, essay prompts for admissions personal statements invite more expressiveness and may require texts that are linguistically different from exam responses. However, with continuous updates and additional training of the current model, we anticipate rapid improvements in generative AI such that users can increasingly refine the style of narrative response they seek, and differences in linguistic style may become less apparent.

Opportunity

Generative AI tools can provide general templates for writing essays and written responses to prompts, assist as editors for grammatical checks, and help organize ideas. For learners with less confidence or who need some guidance, ChatGPT and similar writing tools (e.g., Jasper) can be used for creating initial outlines and drafts. Thus, generative AI can potentially level the playing field across students with varying degrees of writing and language proficiency and access to resources such as tutors. Despite the concerns raised around the authenticity of personal statements created by ChatGPT, potential for use as a free editing service is undeniable.¹⁵ This capability of generative AI should compel authors of essay prompts to be thoughtful, deliberate, and purposeful to ensure that their prompts maximize the potential for personalization in responses and minimize excessive reliance on AI assistance. For example, an essay prompt to promote a personal narrative might state, "Describe a significant personal challenge or a unique experience that has prepared you for medical school." This type of prompt would require creative thinking beyond ChatGPT's ability to merely summarize facts and information.

Learning

Aim

As part of their development of lifelong learning skills and self-regulated learning, students engage in independent study for licensing exam preparation and other learning opportunities during and after clerkships. Most students draw on multiple resources for their independent study, including textbooks, primary

literature, question banks, board review and test preparation, notes from coursework, and online cases and resources.¹⁶ The practice of clinical medicine today relies on the ready availability of such information technology, and training has consequently placed less emphasis on memorization and more on information access, review, synthesis, and critical appraisal.¹⁷

Potential impact

We contemplated how ChatGPT (and other generative AI) may or may not be different from other information technology resources used to enhance learning in classroom and clinical settings. When prompted by the user, ChatGPT can provide general references, pointing the user to textbooks and online resources, such as MedlinePlus and UpToDate. However, it remains unable to provide specific details regarding the references used to generate responses.

In medical education (and medicine in general), validity of information is paramount to rendering patient care. Such validity is critically dependent on being able to trace evidence to its origins and to understand contexts that may limit its generalizability. ChatGPT's inability to provide references used to generate its responses thus severely limits its usability as a learning tool. Although the natural language processing chatbot offered by Microsoft Bing can list citations for its responses to clinical questions, it (as well as ChatGPT) can at times assert misinformation as factual.¹⁸ This limitation is not surprising because the underlying AI technology behind language prediction is based on probabilistic models in which responses are chosen as being most likely but are not guaranteed to be accurate. As such, these tools in their current form cannot be depended on as reliable learning tools for evidence-based medicine (Bard, Google's chatbot, initially avoided providing summaries of medical information altogether).

Opportunity

Despite the potential for inaccuracy, these tools may appear enticing to learners to answer specific queries and readily synthesize a broad base of information. Educators should anticipate that learners will use generative AI tools as a quick reference and information synthesis tool.

We see this as a potential opportunity for educators to continue emphasizing critical appraisal skills in assessing information, whether it be from chatbots or other sources. The risks of misinformation and the lack of context and generalizability posed by today's chatbots are real, and like other online information sources, independent verification is essential. It remains just as important for learners to develop skills in tracing evidence to its primary sources from assertions provided by chatbots compared with summaries and opinions provided by human experts.

Ongoing improvements in generative AI tools can also aid learners in assessing the validity and applicability of AI-generated responses. Although the first wave of these tools appears to have emphasized the novelty of being able to mimic human responses, to improve credibility, the next generation of such tools needs to focus on the following: (1) citing resources used to generate a particular response (such as references used to generate a differential diagnosis), (2) providing contextual information for the response (such as information about the population for which a differential diagnosis is applicable), and (3) providing transparency about the certainty of any given response (such as providing a percentage that indicates how strongly the AI algorithm ascertains the statement as being factual). Generative AI can provide opportunities for learners to exercise critical appraisal and inquiry skills and help identify potential gaps in the literature or current knowledge. Future opportunities will emerge for trainees and practicing physicians to incorporate AI in generating diagnostic considerations and differential diagnoses for real patients, highlighting the urgency for educators to develop plans for teaching these appraisal and inquiry skills.¹⁹

Assessments

Aim

According to Epstein,²⁰ the 3 aims of assessment in medical education are to (1) motivate and provide future direction in learning (feedback), (2) ensure physician competence (social accountability), and (3) serve as a basis for advancement (competency-based medical education). Similarly, programmatic assessment serves 2 primary purposes: to drive learning through feedback and to make judgments about readiness for progression.²¹ We envisioned how

generative AI may undermine or facilitate these goals of assessment.

Potential impact

Recent headlines and scholarly publications have signaled the potential for generative AI to change the landscape of high-stakes knowledge assessment. ChatGPT can generate answers to essay or open-ended prompts by drawing on its training from a wide range of online information sources. Headlines that ChatGPT can pass high-stakes graduate exams and produce narratives indistinguishable from learner-generated content immediately generated alarm among educators.^{22–24} In medical education, ChatGPT performed at or near the passing threshold for all 3 of the United States Medical Licensing Examinations.¹ These findings rightfully raise concerns about cheating and raise challenges for academic institutions in detecting dishonesty. Academic integrity is essential to achieve valid determinations about learners' readiness for advancement. Although exam security problems are not new in education, generative AI tools such as ChatGPT broaden the nature of security concerns. Cheating has been a longstanding problem with online exams.²⁵ ChatGPT is a new facilitator of information access in the context of online exams. Examinees who access ChatGPT during open-book exams or while taking exams using computers that are not locked down or carefully proctored can use ChatGPT to generate draft responses. With open-ended question exams, ChatGPT generates text in response to prompts that may be indistinguishable from student-generated answers and undetectable by plagiarism checkers.²⁶ Therefore, educators would not be able to discern a student's true work or level of understanding. Mechanisms to detect language generated by tools such as ChatGPT may become available but are not currently in routine use. A common response to these concerns is for educators to propose strategies to prevent students from accessing information from ChatGPT during exams. Some approaches already in place may need updating to reflect the current AI environment. Locking down computers during exams is a useful strategy to avoiding plagiarism, reading online resources during exams, and using generative AI tools. Strengthening proctoring protocols further limits students' ability to access resources during exams. Avoiding the use

of take-home exams similarly reduces access to information aids. Policies, honor codes, and student instructions about exam security and cheating should address permitted and prohibited uses of AI to ensure transparency.

Opportunity

Educators can move from a defensive, security-focused approach to creative alternatives for assessment that steer learners away from opportunities to use generative AI. For example, group projects, multimedia work products, and presentations or oral exams represent novel assessment techniques in many settings. Another strategy is to incorporate use of generative AI tools into the learning and assessment process. For example, to engage and assess students' critical appraisal skills, an educator could invite students to generate responses using AI or review AI-generated material to critique or revise it.

Formative uses of generative AI can guide learning.²⁷ For example, AI can help identify areas of learner misconceptions or knowledge gaps based on thematic text analysis of assessment data. AI-generated assessment tools are adaptive; they can provide individualized, timely, and continuous feedback based on a learner's response to a prompt to guide the student as a tutor would do, as illustrated by the new tool being tested by Khan Academy called Khamingo. Given the importance of timely and frequent feedback as integral to formative assessment, generative AI can increase learners' access to feedback with less expenditure of faculty time and can potentially strengthen the depth or quality of feedback shared.²⁷

For summative assessments that ensure competence and readiness for advancement, AI can facilitate scoring and interpretation of performance data. For written assessments (e.g., open-ended constructed responses) scored with rubrics, generative AI can automate part of the scoring and suggest narrative feedback using smart phrases to ease the burden on raters. Although the reliability of automated scoring by ChatGPT is still low compared with human raters,²⁸ with more training and improvements in the algorithm for detecting semantic nuances, generative AI can make scoring of responses to written assessments more efficient. AI tools could facilitate scoring learner patient notes, with educator

attention to the confidentiality of imputed patient information.

AI tools can signal learning trajectories, identify patterns of growth, or flag learners below thresholds of competence to make assessment inferences more continuous rather than periodic (only at the end of a learning period). For educators or committees synthesizing learner performance information, AI can generate summary conclusions. For example, a generative AI tool could generate draft student assessment summary paragraphs for the medical student performance evaluation using supervisor narratives or other sources of data based on certain criteria, such as the clerkship competencies, or compare resident performance to specialty milestones. These summaries would need to be manually reviewed to determine whether they accurately represent the original assessment data. However, such use on open AI platforms violates current privacy requirements for individual student data protected by the Family Educational Rights and Privacy Act, and any use of assessment data in generative AI tools must comply with local policies and laws.

Research

Aim

The primary aim of medical education research is to advance the science and evidence within our field of inquiry. Research manuscripts rely on adhering to the accepted conventions and methods of scientific inquiry, including critical appraisal of the literature, accurate presentation of data, and appropriate citation and references for source information. As researchers start using ChatGPT in various phases of the research process, the importance of adhering to these existing scientific protocols is paramount.

Potential impact

There are already some recommendations for how to use ChatGPT to formulate outlines for research papers.²⁹ ChatGPT can speed the process of identifying and reviewing potential lists of literature review sources and help paraphrase texts. Several publications and preprints crediting generative AI with authorship have prompted discussion among journal editors and researchers on the appropriateness of citing ChatGPT as an author on publications. Some journal editors, including those for *Nature* and

Science, have opposed including AI as a coauthor because tools such as ChatGPT do not meet current authorship guidelines, cannot take responsibility for a study, and do not have the capacity to agree to coauthorship.³⁰ As discussed earlier, potential pitfalls of relying on tools such as ChatGPT in research are the lack of clear references and citations for sources of information. Although some information sources can be extracted from ChatGPT when prompted, the current iteration of the tool makes verifying and linking specific citations to generated responses challenging and limits the usefulness of ChatGPT as a literature search tool.

Opportunity

Authorships aside, there are some potential uses of generative AI in the

research process, including initiating the writing process, formulating early research questions, and identifying gaps in current knowledge. Researchers remain accountable for the transparency and accuracy of the information they report with appropriate citations and references. As we anticipate improvements in generative AI tools, encouraging full disclosure about the role of generative AI during the research process will be important. The JAMA Network recently published changes to their authorship guidelines stating that the use of AI to create content or assist in writing should be included in the acknowledgment and/or methods sections clarifying the qualification criteria for authorship.³¹ Another opportunity to leverage AI tools for research is during the literature review process. By accessing the wider scope of

literature afforded through AI tools beyond the traditional databases, we can bring new perspectives and ideas to the field that were previously untapped or overlooked.³²

Conclusions

Generative AI has the potential to be a valuable tool for educators, learners, and institutions. However, concerns around the accuracy of generated information, potential for bias, and the ethical use of the tool need to be considered carefully for successful adoption within medical education. In Table 1, we provide a list of recommendations and resources for medical educators and institutions to increase AI literacy for optimal adoption of generative AI tools such as ChatGPT. We anticipate more significant changes to

Table 1
Recommendations for Integrating Generative AI in Medical Education

Recommendations	Helpful resources
Recommendations for educators	
Increase AI knowledge	<ul style="list-style-type: none"> Atewell S. How Generative AI Is Impacting the Way We Work. National Center for AI in Tertiary Education; July 6, 2023. Accessed July 16, 2023. https://nationalcentreforai.jiscinvolve.org/wp/. Santa Fe Community College Library. Faculty help: ChatGPT comprehensive resource guide: writing assignments. Accessed July 16, 2023. https://libraryhelp.sfcc.edu/ChatGPT/revised-assignments.
Understand the current landscape of AI use in medical education	<ul style="list-style-type: none"> Tolsgaard MG, Pusic MV, Sebok-Syer SS, et al. The fundamentals of artificial intelligence in medical education research: AMEE Guide No. 156. <i>Med Teach</i>. 2023;45(6):565–573. Russell RG, Novak LL, Patel M et al. Competencies for the use of artificial intelligence-based tools by health care professionals. <i>Acad Med</i>. 2023;98(3):348–356.
Review strategies for successful AI integration into education	<ul style="list-style-type: none"> Holmes W, Tuomi I. State of the art and practice in AI in education. <i>Eur J Educ</i>. 2022;57(4):542–570. ISTE. AI exploration for educators: putting educators and students in the driver's seat. Accessed July 16, 2023. https://www.iste.org/areas-of-focus/AI-in-education. McCoy LG, Nagaraj S, Morgado F, Harish V, Das S, Celi LE. What do medical students actually need to know about artificial intelligence? <i>NPJ Dig Med</i>. 2020;3:86.
Become stewards of ethical use of AI	<ul style="list-style-type: none"> TeachOnlineCA. About us. Accessed July 16, 2023. https://teachonline.ca/about-us. Katznelson G, Gerke S. The need for health AI ethics in medical school education. <i>Adv Health Sci Educa</i>. 2021;26:1447–1458.
Recommendations for institutions	
Review and revise school policies (and create new policies as needed) regarding use of generative AI	<ul style="list-style-type: none"> National Centre for AI. Considerations on wording when creating advice or policy on AI use. Accessed July 16, 2023. https://nationalcentreforai.jiscinvolve.org/wp/2023/02/14/considerations-on-wording-ai-advice/. University of California. University of California Learning Data Privacy Principles. https://www.ets.berkeley.edu/sites/default/files/general/uc_learning_data_principles_final03_05_2018.pdf. Accessed July 16, 2023.
Support faculty development about AI and provide resources for teaching	<ul style="list-style-type: none"> Texas State University. Artificial intelligence (AI) in academia: resources for faculty. Accessed July 16, 2023. https://www.facdv.txst.edu/Teaching-Learning/ai-resources-faculty.html. ISTE. AI exploration for educators: putting educators and students in the driver's seat. Accessed July 16, 2023. https://www.iste.org/areas-of-focus/AI-in-education.
Offer information-checking tools for originality and plagiarism to faculty	<ul style="list-style-type: none"> OpenAI. AI text classifier. Accessed July 16, 2023. https://platform.openai.com/ai-text-classifier. CopyLeaks. AI content detector. Accessed July 16, 2023. https://copyleaks.com/ai-content-detector. Beebom. 15 Best AI plagiarism checkers to detect ChatGPT-generated content. Accessed July 16, 2023. https://beebom.com/best-ai-plagiarism-checkers/.

Abbreviation: AI, artificial intelligence.

medical education beyond what we have outlined in this article as the use of generative AI expands and tools improve. As educators, we need to increase our AI literacy through education and vigilance around new advances in technology. Additionally, educators need to be stewards of AI literacy to foster social responsibility and ethical awareness around the use of AI with our learners.^{5,33}

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

C. Boscardin is professor, Department of Medicine and Department of Anesthesia and Perioperative Care, and director of student assessment, School of Medicine, University of California, San Francisco, San Francisco, California.

B. Gin is associate professor, Department of Pediatrics, University of California, San Francisco, San Francisco, California.

P.B. Golde is director of data and analytics, School of Medicine, University of California, San Francisco, San Francisco, California.

K.E. Hauer is associate dean for competency assessment and professional standards and professor of medicine, University of California, San Francisco, San Francisco, California.

References

- 1 Kung TH, Cheatham M, Medinilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. medRxiv. Preprint posted online December 20, 2022. doi:10.1101/2022.12.19.22283643
- 2 Weissman J. ChatGPT is a plague upon education [opinion]. Accessed March 8, 2023. <https://www.insidehighered.com/views/2023/02/09/chatgpt-plague-upon-education-opinion>.
- 3 Reuters. ChatGPT sets record for fastest-growing user base—Analyst note. Accessed March 22, 2023. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>.
- 4 Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56.
- 5 Ng DTK, Leung JKL, Chu SKW, Qiao MS. Conceptualizing AI literacy: an exploratory review. *Comput Educ Artif Intell*. 2021;2:100041.
- 6 Gong X, Tang Y, Liu X, et al. K-9 Artificial intelligence education in Qingdao: issues, challenges and suggestions. In: Proceedings of the 2020 IEEE International Conference on Networking, Sensing and Control (ICNSC). IEEE; 2020:1–6.
- 7 Russell RG, Lovett Novak L, Patel M, et al. Competencies for the use of artificial intelligence-based tools by health care professionals. *Acad Med*. 2023;98(3):348–356.
- 8 McCoy LG, Nagaraj S, Morgado F, Harish V, Das S, Celi LA. What do medical students actually need to know about artificial intelligence? *NPJ Digit Med*. 2020;3:86.
- 9 Lomis K, Jeffries P, Palatta A, et al. Artificial intelligence for health professions educators. *NAM Perspect*. 2021;2021. doi:10.31478/202109a.
- 10 Dong T, Kay A, Artino AR, et al. Application essays and future performance in medical school: are they related? *Teach Learn Med*. 2013;25(1):55–58.
- 11 Cowen VS, Kaufman D, Schoenherr L. A review of creative and expressive writing as a pedagogical tool in medical education. *Med Educ*. 2016;50(3):311–319.
- 12 Here's How Forbes Got the ChatGPT AI to Write 2 College Essays in 20 Minutes. Accessed March 8, 2023. <https://www.forbes.com/sites/emmahwhitford/2022/12/09/heres-how-forbes-got-the-chatgpt-ai-to-write-2-college-essays-in-20-minutes/?sh=6dea74a756ad>.
- 13 Will ChatGPT Eliminate the College Admissions Essay? Accessed March 8, 2023. <https://www.saraharberson.com/blog/chatgpt-college-admissions-essay>.
- 14 Guo B, Zhang X, Wang Z, et al. How close is ChatGPT to human experts? Comparison corpus, evaluation, and detection. arXiv. Preprint posted online January 18, 2023. doi:10.48550/arxiv.2301.07597
- 15 Biswas S. ChatGPT and the future of medical writing. *Radiology*. 2023;307(2):e223312.
- 16 Kind T, Olvet DM, Farina G, et al. Reading and study habits of medical students on clerkships and performance outcomes: a multi-institutional study. *Med Sci Educ*. 2021;31(6):1957–1966.
- 17 Cutrer WB, Spickard WA 3rd, Triola MM, et al. Exploiting the power of information in medical education. *Med Teach*. 2021;43(suppl 2):S17–S24.
- 18 Bing's ChatGPT-powered search engine is making stuff up and throwing tantrums. TechRadar. Accessed March 22, 2023. <https://www.techradar.com/news/bings-chatgpt-powered-search-engine-is-making-stuff-up-and-throwing-tantrums>.
- 19 Liang H, Tsui BY, Ni H, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. *Nat Med*. 2019;25(3):433–438.
- 20 Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356(4):387–396.
- 21 van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205–214.
- 22 Terwiesch T. Would Chat GPT get a Wharton MBA? White Paper. Accessed March 22, 2023. <https://mackinstitute.wharton.upenn.edu/2023/would-chat-gpt3-get-a-wharton-mba-new-white-paper-by-christian-terwiesch/>.
- 23 Choi JH, Hickman KE, Monahan A, Schwarcz DB. ChatGPT goes to law school. *Soc Sci Res Netw J*. Published online January 23, 2023. doi:10.2139/ssrn.4335905
- 24 Hung AJ, Chen J, Che Z, et al. Utilizing machine learning and automated performance metrics to evaluate robot-assisted radical prostatectomy performance and predict outcomes. *J Endourol*. 2018;32(5):438–444.
- 25 Noorbahani F, Mohammadi A, Aminazadeh M. A systematic review of research on cheating in online exams from 2010 to 2021. *Educ Inf Technol*. 2022;27(6):8413–8460.
- 26 Cotton D, Cotton P, Shipway JR. Chatting and cheating: ensuring academic integrity in the era of ChatGPT. EdArXiv. Preprint posted online January 10, 2023. doi:10.35542/osf.io/mrz8h.
- 27 Cardona MA, Rodríguez RJ, Ishmael K. Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations. Office of Educational Technology, US Department of Education; May 25, 2023.
- 28 Khademi A. Can ChatGPT and Bard generate aligned assessment items? A reliability analysis against human performance. *J Appl Learn Teach*. 2023;6(1):1–39.
- 29 Turban S. How will ChatGPT change research paper writing? Accessed March 8, 2023. <https://www.lumiere-education.com/post/how-will-chatgpt-change-research-paper-writing>.
- 30 Stokel-Walker C. ChatGPT listed as author on research papers: many scientists disapprove. *Nature*. 2023;613(7945):620–621.
- 31 Flanagan A, Bibbins-Domingo K, Berkwits M, Christiansen SL. Nonhuman “authors” and implications for the integrity of scientific publication and medical knowledge. *JAMA*. 2023;329(8):637–639.
- 32 Tolsgaard MG, Boscardin CK, Park YS, Cuddy MM, Sebok-Syer SS. The role of data science and machine learning in health professions education: practical applications, theoretical contributions, and epistemic beliefs. *Adv Health Sci Educ Theory Pract*. 2020;25(5):1057–1086.
- 33 Laupichler MC, Aster A, Schirch J, Raupach T. Artificial intelligence literacy in higher and adult education: a scoping literature review. *Comput Educ Artif Intell*. 2022;3:100101.