

Uncomposed, edited manuscript published online ahead of print.

This published ahead-of-print manuscript is not the final version of this article, but it may be cited and shared publicly.

Author: Hashimoto Daniel A. MD, MS; Johnson Kevin B. MD, MS

Title: The Use of Artificial Intelligence Tools to Prepare Medical School Applications

DOI: 10.1097/ACM.00000000005309

Academic Medicine

DOI: 10.1097/ACM.000000000005309

The Use of Artificial Intelligence Tools to Prepare Medical School Applications

Daniel A. Hashimoto, MD, MS, and Kevin B. Johnson, MD, MS

D.A. Hashimoto is assistant professor of surgery and computer and information science, and affiliated faculty, General Robotics, Automation, Sensing, and Perception Laboratory, University of Pennsylvania, Philadelphia, Pennsylvania; ORCID: http://orcid.org/0000-0003-4725-3104.

K.B. Johnson is the David L. Cohen University Professor of Pediatrics, Biomedical Informatics,

and Science Communication, University of Pennsylvania, Philadelphia, Pennsylvania.

Correspondence should be addressed to Daniel A. Hashimoto, 3400 Spruce Street, 4 Silverstein Pavilion, Philadelphia, PA 19104; telephone: 215-662-4503; email:

daniel.hashimoto@pennmedicine.upenn.edu; Twitter: @laparoscopes.

Acknowledgments: The authors thank Lyle Ungar, PhD, for insights on the evolution of large language models and Lilach Mollick, PhD, for insights on how business schools are using large language models in their curricula.

Funding/Support: None reported.

Other disclosures: D. Hashimoto is vice chair of the board of directors of the Global Surgical AI Collaborative, a nonprofit institution that facilitates surgical data sharing for researchers. The authors utilized Grammarly for word choice and grammar and GPT-4 (OpenAI) to generate the text in Box 1 to illustrate the type of response a large language model can generate. *Ethical approval:* Reported as not applicable.

Abstract

Advances in artificial intelligence (AI) have been changing the landscape in daily life and the practice of medicine. As these tools have evolved to become consumer-friendly, AI has become more accessible to many individuals, including applicants to medical school. With the rise of AI models capable of generating complex passages of text, questions have arisen regarding the appropriateness of using such tools to assist in the preparation of medical school applications. In this commentary, the authors offer a brief history of AI tools in medicine and describe large language models, a form of AI capable of generating natural language text passages. They question whether AI assistance should be considered inappropriate in preparing applications and compare it with the assistance some applicants receive from family, physician friends, or consultants. They call for clearer guidelines on what forms of assistance-human and technological-are permitted in the preparation of medical school applications. They recommend that medical schools steer away from blanket bans on AI tools in medical education and instead consider mechanisms for knowledge sharing about AI between students and faculty members, incorporation of AI tools into assignments, and the development of curricula to teach the use of AI tools as a competency.

In the 2022–2023 U.S. medical school admissions cycle, 41% of the more than 55,000 applicants were accepted.¹ In 2012, there were approximately 45,000 applicants.² Facing increasing competition for medical school seats, applicants have found themselves in somewhat of an academic arms race to optimize their test scores, grade point averages, extracurricular research and clinical experiences, letters of recommendation, and personal statements. For some applicants, preparing their application involves taking a test preparation course, consulting with physician family members or friends, or perhaps even securing the services of an admissions consultant. Such assistance may be readily attained by the 75% of U.S. medical students whose families are in the top 2 income quintiles,³ but may be inequitably accessible to a large number of applicants.

Recent advances in consumer-friendly artificial intelligence (AI) technologies—such as mobile applications for flashcards and study quizzes, software services for grammar assistance in essays, and services for language generation—have dramatically enhanced the low- and no-cost tools available to students to improve their applications. Fields ranging from law to business to higher education and medicine are grappling with whether and how to use AI tools in the admissions process and in daily practice. While some industries are exploring how to rapidly incorporate these technologies to augment or replace their workforce, others have blocked their use in an attempt to first better understand their implications. AI tools can introduce risks—for example, perpetuation of bias from datasets, de-skilling of the workforce, and opportunities for intellectual dishonesty—but they also offer the potential to improve or augment human performance on various tasks, including the preparation of written material.

Copyright © by the Association of American Medical Colleges. Unauthorized reproduction of this article is prohibited

In this commentary, we provide a brief history of the evolution of AI tools in medicine, describe current AI tools that are capable of generating passages of text, and call for guidance for medical school admissions committees regarding whether and how to manage applications that leverage these tools. We emphasize that these tools offer opportunities for medical school applicants to improve their applications as these resources may be accessible to a broader pool of applicants than more costly resources like admissions consultants. We also explore the need for safeguards on the use of AI in medical education beyond the admissions process.

A Brief History of Artificial Intelligence

In 1956, a group of academics spanning computer science, mathematics, linguistics, psychology, and other fields came together to discuss a set of ideas under the umbrella of a field they coined as "artificial intelligence."⁴ Today, AI is broadly considered to represent the field of study concerned with developing machines that can reason and perform cognitive functions that approximate those of humans.⁵

As hardware and software capabilities have advanced, the public's expectations of AI capabilities have evolved. While we are not close to general AI agents that act as human surrogates, such as the fictional HAL9000 from *2001: A Space Odyssey* or C-3P0 from *Star Wars*, narrow AI tools have enabled semi-automated driving, image generation from text prompts, autonomous factory robots, and a host of functions on mobile phones.

As AI tools have become ubiquitous and have subtly begun to permeate many aspects of our daily lives in ways both seen and unseen, the goalposts for what individuals consider as AI have shifted. Early AI programs used rule-based systems that required expert input. The first medical system that exhibited what many would call AI capabilities was MYCIN, a system for treating blood infections developed by Shortliffe and colleagues at Stanford University.⁶ INTERNIST-I,

a rule-based computer diagnostic system developed in the 1970s based on the expertise of a single clinician at the University of Pittsburgh, was another notable advance.⁷ Medical AI tools have since evolved to incorporate more complex methods, such as deep learning, that have yielded automated preliminary reads in diagnostic radiology and pathology as well as advanced surgical guidance systems.^{8,9}

This evolution, both inside and outside medicine, has altered our perception of what "counts" as AI. For example, Japanese-style rice cookers, use rule-based systems like fuzzy logic¹⁰ (developed in the 1960s) to execute their tasks. However, few people today would look at a rice cooker and consider it to be an AI-enabled machine when compared to smartphones and driverless cars.

Modern perceptions and expectations of AI have largely been influenced by advances in deep learning techniques over the last decade.¹¹ With these advances, there has been a proliferation of tools, including AI voice assistants (e.g., Apple's Siri and Amazon's Alexa), natural language processing assistants (e.g., Grammarly), and speech recognition tools (e.g., Nuance's Dragon). As datasets have grown and processors have become more powerful, scientists have become more adept and creative at leveraging them to create more impressive AI tools.

Large Language Models

Perhaps the most popular advance in AI as of 2023 has been in the space of large language models (LLMs). LLMs are deep learning algorithms that recognize, predict, and generate text, images, and other content.¹² Two tools using this approach are BERT (Bidirectional Encoder Representations from Transformers)¹² and GPT (Generative Pre-trained Transformer)¹³ first introduced by Google and OpenAI, respectively, in 2018, with ChatGPT released by OpenAI to the lay public via a chat interface in 2022. These AI tools leverage enormous amounts of training

data from datasets scraped from the internet and use this training to generate responses to queries or prompts posed in natural language (as opposed to requests made in computer code). They generate responses to queries by considering their training data and generating statistically likely sequences of words that would fulfill the query. While the specific architectures of the most popular LLMs are proprietary, the predominant training technique for these models is known as reinforcement learning from human feedback (RLHF).¹⁴ Reinforcement learning is a machine learning technique that is akin to operant conditioning, wherein an algorithm receives mathematical positive or negative rewards based on its performance of a task. With RLHF, a human provides feedback to the algorithm by ranking its responses, thus introducing human preferences into the algorithm's reward structure.

This training process allows LLMs to generate text answers that are complex and may seem to have been written by a human. Box 1 demonstrates the type of answer an LLM may give in response to a human prompt—in this case, our prompt asking for assistance with planning a medical school admissions essay. One of the limitations of the current generation of LLMs is that small changes in prompt structure can yield vastly different responses. To refine the responses to achieve the desired output, users can engage in prompt engineering, which refers to the process of crafting prompts to optimize the performance of LLMs in generating accurate, contextually relevant responses.

LLM-generated text is based on the statistical associations of patterns of words to the patterns seen in training data and prompts. Early LLM-generated text was plagued by elements such as unnecessary and unnatural repetition, lack of clarity or specificity, out-of-context content, inconsistent phrasing and arguments (particularly across long essays), incorrect word selection,

and occasional incorrect statements presented as facts—called "hallucinations" in the computer science literature.¹⁵ Some of these issues persist today but have become less perceptible. Despite these issues, LLMs and transformer-based tools have performed astonishingly well on tests such as the Law School Admission Test and Medical College Admission Test (MCAT).¹⁶ Further, a recent study by Ayers and colleagues¹⁷ compared physician and AI chatbot responses to patient questions and found that investigators preferred the chatbot responses and rated them as having higher quality and empathy than the physician responses. This study is the harbinger of a tsunami of research and technology based on new AI methods.

Using AI to Prepare Medical School Applications

Medical school applications are intended to help schools select students who are likely to excel in medicine as clinicians, researchers, and/or educators. While a large component of the application is composed of scores (i.e., grade point average and transcript, MCAT score), much of the application consists of the applicant's descriptions of their research, volunteer, and other activities and their personal statement. Together, these text elements of the application can contribute to holistic review of an applicant,¹⁸ which the AAMC defines as considering "experiences, attributes, and academic metrics as well as the value an applicant would contribute to learning, practice, and teaching."¹⁹

With the applicant's activity descriptions and personal statement comprising 50% of the primary application (and more of a secondary application), we must consider whether AI tools like LLMs are acceptable for use by students to assist in preparing their applications and whether the type of assistance provided should affect the perception of how AI was used. Faculty members would be likely to agree that using AI tools to generate an entire personal statement or to fabricate a description of a volunteer clinical experience would be unethical and unacceptable. However,

would all faculty members be opposed to the use of AI tools to clean up grammar and phrasing? Grammar and spell-check features are embedded into common word-processing programs. Would faculty members be opposed to the use of a search engine to find articles that offer tips on how to construct an effective personal statement? Content in such articles rarely differs from the content suggested by an LLM in the "Body Paragraphs" section of our example (Box 1), except that an LLM can generate highly tailored responses based on the human-provided prompt. These personalized suggestions likely resemble those that an admissions consultant, premedical advisor, or physician family member might offer an applicant. Using AI tools could, therefore, provide more equitable access to tailored assistance in preparing each element of a medical school application (Table 1).²⁰

Incidentally, concurrent with the release of AI tools (e.g., GPT), tools are being released to distinguish between AI-written and human-written text.²¹ It is reasonable to assume that these detection tools will improve, as will the AI tools that aim to evade detection.

Clear Rules Needed for the Use of AI in Applications

Content augmentation of written statements using technology (including AI) is not new; therefore, banning the use of AI in medical school applications seems unreasonable. Medical school applicants have learned to utilize a wide range of tools (e.g., search engines, mobile applications, presentation software) to succeed in their scholastic endeavors, and AI is arguably another tool in their armamentarium.

Scientific journals have policies outlining how and when LLMs can be used to contribute to scientific writing.²² LLMs cannot coauthor publications as they cannot meet the criteria for authorship set forth by the International Committee of Medical Journal Editors (namely that AI tools cannot be accountable for the text they generate).²³ Journals such as *Academic Medicine*

have published their guidance on how use of AI tools should be disclosed and reported in manuscripts.²⁴

Medical schools, perhaps with facilitation by the Association of American Medical Colleges, should establish clear rules and guidance on the use of AI in preparing medical school applications. The 2024 American Medical College Application Service (AMCAS) application does provide certification statements to which applicants must attest, including a statement about AI and authorship: "I certify that all written passages, such as the personal statement ... are my own and have not been written, in part or in whole, by another author and are not the product of artificial intelligence"²⁵ (emphasis added). Although this attestation is important and seemingly straightforward, it may be insufficient given the ambiguity in the definition of a "product of artificial intelligence" as noted above. It is unclear whether this includes editorial review of passages, advice on structuring passages, or suggestions on types of content to include. Specific guidance on use of AI would provide applicants (and admissions committees) with a transparent understanding of what is and is not allowed. The AAMC could consider asking students to indicate on their AMCAS application if and how AI was used to assist them in preparing their materials. However, if such a step is taken, we suggest students should similarly be asked if and how other humans assisted in preparing their application. Asking only about AI assistance could lead to bias against applicants who use widely available AI tools due to lack of access to assistance from consultants, parents, or physicians in their social circle.

Safeguards for the Use of AI in Medical Education

AI tools are embedded in nearly every facet of life whether we recognize or define them explicitly as "AI." Every popular search engine and web browser uses AI, social media and entertainment applications use AI, our cellular telephones use AI, and our electronic medical records systems use AI. Therefore, it would be prudent for medical educators to consider safeguards to encourage appropriate, safe use of AI tools—and to avoid their improper use—beyond the admissions process.

As we have noted, transparent rules for the use of AI are critical to responsible use. Blanket bans are impractical given the ubiquity of AI tools and would inadvertently capture common software (e.g., Microsoft Word, Google Search, Apple iOS, Epic). Educators must thus think critically about when and how AI tools can be used in coursework and on the wards and clearly communicate their expectations to trainees. The wide range of AI tools for medical education includes flashcard applications, conversational agents, question banks, and intelligent tutoring systems.

Medical schools should consider mechanisms to allow individuals to share their knowledge of and experience with the use of AI tools. Such knowledge should be shared among faculty members and between faculty members and learners. Medical students are motivated to maximize their learning and performance, and they often share tips with one another about helpful applications, books, and other resources for a given course or rotation. They will assuredly do the same with AI tools, and faculty members should encourage their students to share helpful AI tools with them as well. The faculty should then vet these tools. Business schools are beginning to incorporate the use of LLMs in their coursework and assignments.²⁶ Medical schools should consider doing the same in their curricula to support students' learning. Practically, the proficient use of AI tools will become a competency that benefits trainees, faculty, and patients alike. Medical schools should consider developing curricula that address this competency.²⁷ We believe the clearest safeguard against inappropriate and unethical use of

AI tools is to empower medical professionals and trainees with an accurate and realistic understanding of AI, its capabilities, and its potential for both harm and good.

Conclusion

The use of AI tools may become inevitable in many aspects of our personal and professional lives if it has not already. In the medical school admissions process, clear rules on the transparent use of assistance (AI or human) would help draw a line between demonstrating students' prowess and resourcefulness in their use of assistance versus hiding their deficiencies and weaknesses. The sharing of knowledge about AI, the integration of AI tools into existing curricula, and the creation of specialized curricula that emphasize AI utilization as a competency warrant careful consideration as clinicians and patients alike would benefit from a greater understanding of AI's benefits and risks.

References

- 2022 FACTS: Applicants and Matriculants Data.. Association of American Medical Colleges. Accessed June 16, 2023. https://www.aamc.org/data-reports/studentsresidents/data/2022-facts-applicants-and-matriculants-data
- FACTS. U.S. Medical School Total Applicants, 2002-2017. In: 2017 Applicant and Matriculant Data Tables. Association of American Medical Colleges. Published December 2017. Accessed June 16, 2023. https://www.aamc.org/system/files/d/1/5-2017_applicant_and_matriculant_data_tables.pdf
- Williams C, Perez MA, Vapiwala N, Shea JA. The impact of socioeconomic factors on medical school acceptance rates. Acad Med. 2021;96(11 Suppl):S219-S220.
- Hashimoto DA, Rosman G, Meireles OR, eds. Artificial Intelligence in Surgery: Understanding the Role of AI in Surgical Practice. New York, NY: McGraw-Hill Education; 2021.
- Bellman R. An Introduction to Artificial Intelligence: Can Computers Think? San Francisco, CA: Thomson Course Technology; 1978.
- Shortliffe EH, Axline SG, Buchanan BG, Merigan TC, Cohen SN. An artificial intelligence program to advise physicians regarding antimicrobial therapy. Comput Biomed Res. 1973;6(6):544-560.
- Miller RA, Pople HE Jr, Myers JD. Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. N Engl J Med. 1982;307(8):468-476.
- Mascagni P, Alapatt D, Sestini L, et al. Computer vision in surgery: From potential to clinical value. NPJ Digit Med. 2022;5(1):163.

- Cohen TA, Patel VL, Shortliffe EH, eds. Intelligent Systems in Medicine and Health: The Role of AI. Cham, Switzerland: Springer Cham; 2022.
- 10. Zadeh LA. Fuzzy sets. Information and Control. 1965;8:338–353.
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, eds. Advances in Neural Information Processing Systems. 2012;25:1097-1105.
- Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional Transformers for language understanding. arXiv [csCL]. Preprint posted online October 10, 2018. doi:10.48550/arXiv.1810.04805
- 13. Radford A, Narasimhan K, Salimans T, Sutskever I. Improving Language Understanding by Generative Pre-Training. OpenAI.com. Preprint posted 2018. Accessed June 16, 2023. Available from: https://s3-us-west-2.amazonaws.com/openai-assets/researchcovers/language-unsupervised/language_understanding_paper.pdf
- 14. Stiennon N, Ouyang L, Wu J, et al. Learning to summarize with human feedback. In: Larochelle H, Ranzato M, Hadsell R, Balcan MF, Lin H, eds. Advances in Neural Information Processing Systems. 2020;33:3008-3021. https://proceedings.neurips.cc/paper_files/paper/2020/file/1f89885d556929e98d3ef9b86448 f951-Paper.pdf
- Ji Z, Lee N, Frieske R, et al. Survey of hallucination in natural language generation. ACM Comput Surv. 2023;55:1-38.
- OpenAI. GPT-4 technical report. arXiv [csCL]. Preprint posted online March 15, 2023. doi:10.48550/arXiv.2303.08774

Copyright © by the Association of American Medical Colleges. Unauthorized reproduction of this article is prohibited

- Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med. 2023;183:589-596.
- Conrad SS, Addams AN, Young GH. Holistic review in medical school admissions and selection: A strategic, mission-driven response to shifting societal needs. Acad Med. 2016;91(11):1472-1474.
- Holistic Review. Association of American Medical Colleges. Accessed June 9, 2023. https://www.aamc.org/services/member-capacity-building/holistic-review
- 20. The Parts of Your Medical School Application. Students & Residents. Association of American Medical Colleges. Accessed June 9, 2023. https://students-residents.aamc.org/real-stories-demonstrating-core-competencies/parts-your-medical-school-application
- New AI classifier for indicating AI-written text. Open AI. Accessed June 9, 2023. https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text
- 22. [No authors listed.] Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature. 2023;613(7945):612.
- 23. Defining the Role of Authors and Contributors. International Committee of Medical Journal Editors. Accessed June 9, 2023. https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html
- DeVilbiss MB, Roberts LW. Artificial intelligence tools in scholarly publishing: Guidance for academic medicine authors. Acad Med. Published online April 28, 2023. doi:10.1097/ACM.000000000005261

- 25. Application and Acceptance Protocols for Applicants. Students & Residents. Association of American Medical Colleges. Accessed June 9, 2023. https://students-residents.aamc.org/applying-medical-school-amcas/application-and-acceptance-protocols-applicants
- Mollick ER, Mollick L. Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts. SSRN. Published online March 24, 2023. doi:10.2139/ssrn.4391243
- 27. 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.

Table 1

Roles for AI Augmentation in Medical School Applications Without Replacing An Applicant's Own Perspective and Work

AMCAS application	
component ²⁰	Role of AI augmentation
Description of institutional	Create a more sympathetic or understandable explanation
action	for egregious actions
Description of felonies and	Create a more sympathetic or understandable explanation
misdemeanors	for egregious actions
Work and activities	Help select most common or least common activities from
	a list of activities
	Draft text explaining why these activities might be
	significant
	Shorten text to stay within the allotted character limits
Letter of evaluation	Draft multiple different letters from a single prompt, or
	create versions of a letter for different authors
	Generate list of recommended points to highlight in a letter
	from a CV/résumé in a letter
Essays	Proofread essay
	Shorten text to stay within the allotted character limits
	Use school-specific applicant instructions to adjust a draft
	essay or create different essay versions to adhere to criteria
	set forth by a given school
	Assist in brainstorming ideas for essay topics based on an
	applicant's CV and personal background
Assessment of English	Translate text into English at any level of readability
proficiency	
Interview(s)	Simulate an interview to allow applicants to test their
	answers to interview questions
Social media/internet search	Screen applicants' potential social media posts prior to
	posting to ensure appropriateness
Abbraviation: AMCAS Amorican Madice	College Application Service: AL artificial intelligence: CV curriculum

Abbreviation: AMCAS, American Medical College Application Service; AI, artificial intelligence; CV, curriculum vitae.

Box 1

Example of AI Assistance in Planning a Medical School Admissions Essay Based on a Human Prompt^a

Prompt: I am writing my medical school admissions essay. I have trouble coming up with writing ideas and could use help. I grew up in an immigrant family, and my parent was a foreign medical graduate who had to restart training. I never really wanted to pursue a career in medicine but grew to really value what doctors give to others. How can I start my essay?

LLM response: Growing up in an immigrant family, I witnessed firsthand the resilience and determination of my parent, a foreign medical graduate who had to restart their medical training from scratch upon arriving in a new country. Through their eyes, I was exposed to the myriad challenges and triumphs of a career in medicine. As a child, I was initially reluctant to follow in their footsteps, but over time, I found myself captivated by the compassion and impact that doctors have on people's lives. In this essay, I will share how my unique upbringing and the transformative experiences that followed have ultimately fueled my desire to become a physician and contribute to the betterment of the medical community.

Body Paragraphs:

- Narrate your upbringing and the challenges your parent faced as a foreign medical graduate, drawing connections to the qualities you admire in them (e.g., perseverance, adaptability, empathy) and how these qualities have influenced your personal growth.
- Describe a key experience or event that sparked your initial interest in medicine, focusing on the emotions and realizations you encountered. This could be a personal or family health issue, a volunteer experience, or an interaction with a medical professional.

- Reflect on any additional experiences or activities that have further strengthened your passion for medicine, such as shadowing doctors, participating in research, or engaging in community service.
- 4. Discuss how your background as the child of an immigrant physician has shaped your understanding of healthcare disparities and cultural competency, and how you plan to utilize this perspective to better serve diverse patient populations.
- Outline the personal qualities and skills you possess that you believe will make you a successful medical student and future physician. Emphasize your commitment to lifelong learning, teamwork, and empathy.

Conclusion: In conclusion, my journey towards a career in medicine has been profoundly influenced by my parent's unwavering perseverance as a foreign medical graduate and the invaluable lessons I have gleaned from their experiences. My unique upbringing, coupled with my passion for helping others and dedication to continuous growth, has prepared me to take on the challenges and rewards of medical school. I look forward to contributing to the medical community and using my skills and insights to make a lasting impact on the lives of my future patients.

Abbreviation: AI, artificial intelligence.

^aThe prompt was written by the authors and the response was generated by a large learning model, GPT-4 (March 23 version, Open AI), on April 1, 2023.