

Harnessing the Potential of Large Language Models in Medical Education: Promise and Pitfalls

Trista M. Benítez, MPH^{1*}, Yueyuan Xu^{2*}, J Donald Boudreau, MD, FRCPC³, Alfred Wei Chieh Kow, PhD⁴, Fernando Bello, PhD⁵, Le Van Phuoc, MD, MPH, DTM&H⁶, Xiaofei Wang, PhD⁷, Xiaodong Sun MD, PhD⁸, Gilberto Ka-Kit Leung, MS, PhD⁹, Yanyan Lan, PhD¹⁰, Yaxing Wang, MD, PhD¹¹, Davy Cheng, MD, MSC, FCAHS¹², Yih-Chung Tham, PhD^{13,14**}, Tien Yin Wong, MD, PhD^{2,14**}, Kevin C. Chung, MD, MS^{1**}

*co-first authors

**co-last authors

1. University of Michigan Medical School, Ann Arbor, MI. 2. Tsinghua Medicine, Tsinghua University; Beijing Tsinghua Changgung Hospital, Beijing, China. 3. Institute of Health Sciences Education, Faculty of Medicine and Health Sciences, McGill University, Montreal, Canada. 4. Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. 5. Technology Enhanced Learning and Innovation Department, Duke-NUS Medical School, National University of Singapore, Singapore. 6. VinUniversity, College of Health Sciences. 7. Key Laboratory for Biomechanics and Mechanobiology of Ministry of Education, Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing, China. 8. Department of Ophthalmology, Shanghai General Hospital, School of Medicine, Shanghai JiaoTong University. 9. Department of Surgery, School of Clinical Medicine, LKS Faculty of Medicine, The University of Hong Kong, Queen Mary Hospital, China. 10. Institute of AI Industrial Research, Tsinghua University, Beijing, China. 11. Beijing Institute of Ophthalmology, Beijing Tongren Hospital, Beijing Ophthalmology and Visual Sciences Key Laboratory, Capital University of Medical Science, Beijing, China. 12. The Chinese University of Hong Kong (Shenzhen), School of Medicine, Shenzhen, China; The University of Western Ontario, London, Canada. 13. Centre for Innovation & Precision Eye Health, Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore. 14. Singapore Eye Research Institute, Singapore National Eye Center, Singapore

Funding: None

Word Count: 2,882 Figures: 3 Tables: 1 References: 69

Financial Disclosure: Dr. Chung receives funding from the National Institutes of Health, book royalties from Wolters Kluwer and Elsevier.

Corresponding authors:

Yih-Chung Tham, PhD

Department of Ophthalmology

Yong Loo Lin School of Medicine, National University of Singapore

Email: thamyc@nus.edu.sg

Kevin C. Chung, MD, MS

Section of Plastic Surgery

The University of Michigan Health System

1500 E. Medical Center Drive

2130 Taubman Center, SPC 5340

Ann Arbor, MI 48109-5340

Phone: 734-615-3435

Fax: 734-763-5354

E-mail: kecchung@med.umich.edu

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17

Large language models (LLMs), particularly OpenAI's ChatGPT, have generated interest for their potential to complement and enhance current pedagogy in medical education. This new technology—alongside its recent enhancement with the integration of plugins—represents a paradigm shift in human interactions with AI and offers unprecedented opportunities in medical education. However, like all disruptive technologies, ChatGPT also presents challenges alongside its promises. In this perspective, we present key advantages and limitations of LLMs across the medical school continuum, encompassing preclinical and clinical phases, as well as student research engagements. Opportunities include direct accessibility to information, personalized learning through immediate feedback, and augmentation of clinical skill development. Conversely, potential pitfalls include the risk of inadvertently promoting academic misconduct such as overreliance on AI assistance, potential dampening of critical and creative thinking, concerns over the reliability and accuracy of AI-generated responses as well as the impact of this technology on teaching staff. In this perspective piece, we offer balanced discussions on the newly presented opportunities and challenges associated with integrating LLMs throughout the medical school continuum.

INTRODUCTION

In recent years, there is growing consideration to apply artificial intelligence (AI) technology in medicine and healthcare.¹⁻³ The public release of ChatGPT (OpenAI, San Francisco, CA, USA), a large language model (LLM) trained to provide human-like conversational responses to user queries, further spurred discussion and debate on the ethical use of AI in academia medicine, and healthcare settings. Currently, multiple health systems and hospitals are already considering adoption of LLMs to streamline healthcare and education delivery. The potential applications span a wide range including clinical note-writing, medical record summarization, responding to patient inquiries, and providing point-of-care decision support.⁴⁻⁶

The use of LLM in medical education to enhance basic medical knowledge, clinical learning, clinical skill development, and foster innovation has also been suggested as a major potential application.^{5,7-9} LLM can potentially impact on the entire continuum of medical education and training, spanning across pre-clinical education, clinical clerkships, post-medical school graduate training in residency, and fellowships.^{7,9,10} Furthermore, LLMs are increasing important in the extra-curricular learnings of medical students, such as research and other scholarly activities. Most recently, integration of third-party plugins has further augmented ChatGPT's functionality to wide-ranging tasks from summarizing a lecture to retrieving the latest scientific articles. In this perspective piece, we contextualize emerging literature to offer insights into the potential promises and challenges of integrating LLMs throughout the medical school continuum (**Figure 1**).

1. PRECLINICAL EDUCATION

During the preclinical phase, students traditionally acquire medical knowledge through lectures, face-to-face classroom learning, and case-based small group activities. In recent years, an increasing number of medical students have embraced a self-regulated approach to learning.¹¹⁻

¹⁴ This approach involves utilization of external resources over reliance on instructor-derived content for medical knowledge acquisition.^{13,15} Now more than ever, e-learning tools such as podcasts, teaching videos, app-based technology, and question banks are harnessed to accomplish these ends.^{16,17} By catering to a range of learning styles, e-learning can augment traditional methods in a more individualized manner. Among these e-learning tools, LLMs such as ChatGPT are poised to join the repository of medical education resources with several key advantages.

1.1 Acquire Medical Knowledge

LLMs such as ChatGPT may be a promising tool for students to grasp conceptual knowledge. Unlike search engines that function by retrieving websites related to search terms, ChatGPT is able to deliver a more focused response to an inquiry (**Supplemental Table 1**). Due to its conversational output style, students can use ChatGPT as a remote tutor to summarize a concept, outline salient points, generate practice questions, and refer to resources.¹⁸ This has been further enhanced by the introduction of ChatGPT plugins, software add-ons that integrate applications into the existing interface, in May 2023.¹⁹ To date, over 900 plugins exist, expanding the platform's functionality and potential use in medical education.²⁰

To explore and evaluate its use as a learning resource, we first queried ChatGPT on a clinical condition: carpal tunnel syndrome (CTS). ChatGPT produced broad overview, encompassing etiology, symptoms, diagnostic maneuvers, and indications for surgical management (**Supplemental Table 1**). The information provided appears clear, well-readable, and not dissimilar from commercially available subscription-based resources such as Up-To-Date¹² or Amboss.^{21,22} From this, one can query specific knowledge gaps or request evidence-based references. Considering these attributes, LLM may act as a viable auxiliary learning tool. To that

end, educators may consider organizing workshops to showcase how to effectively leverage this technology within the scope of prepared lecture material.

Given the diverse learning styles of medical students, we explored several popular educational plugins. Medical students often supplement learning with video platforms such as YouTube or OnlineMedEd. Hence, we explored the use of the Video Insights plugin²³ to summarize an educational video on CTS. The plugin retrieved publication information in addition to an accurate text summary of the video's highlights (**Supplemental Table 2**). We envision its utility for students with hearing impairment or those who desire a textual summary without watching an entire clip. Furthermore, the Video Insights plugin could support teaching faculty by providing analytics on audience engagement and viewer demographics, potentially aiding in further lecture refinement.

Plugins such as Argil²⁴ have enhanced ChatGPT's capabilities to include image generation. As an illustration, we prompted ChatGPT to demonstrate the Phalen maneuver, a diagnostic test that triggers CTS symptoms by positioning the hands and wrists to apply pressure on the median nerve. Despite retrieving accurate textual instruction of the maneuver, the generated images did not correctly illustrate the Phalen maneuver (**Figure 2**). Thus, while such plugins are appealing to visual learners, the overarching concern of inaccurate or misleading information still persists. Therefore, this should be approached with caution and ChatGPT should not be relied upon as the sole resource for medical knowledge.

Overall, the introduction of Plugins can help bridge the gap between the most up-to-date information available and ChatGPT's original training dataset, which was previously limited to data up until 2021. By staying updated with recent developments and accessing external data, plugins can potentially aid LLMs in mitigating erroneous responses or "hallucinations" (i.e.,

generating incorrect or fabricated information not present in the training data), when faced with complex queries. However, at this nascent stage, users should be cognizant that plugins necessitate rigorous evaluation and testing to ascertain their validity for medical education purposes. Moreover, it is anticipated that as the field of LLM advances, more sophisticated and higher-quality plugins are poised to emerge.

1.2 Board Exam Preparation

The deployment of ChatGPT as a medical exam board review tool attracted significant attention when researchers demonstrated its capability to achieve a passing score equivalence on the United States Medical Licensing Exam (USMLE).²⁵ Standardized board exams remain a cornerstone of medical education, and residency programs continue to place substantial emphasis on board exam performance.²⁶ In the United States, medical students spend approximately \$4,000 on external test prep material, primarily on question banks and subscription-based content.²⁷ Hence, utilizing ChatGPT to generate practice questions could potentially offer a more accessible and cost-effective alternative.

To delve into this, we explored using ChatGPT to generate USMLE-style practice questions on common hand conditions such as CTS (**Table 1**). ChatGPT crafted questions in a style akin to the USMLE, yet the complexity of the questions appeared to be of a lower order. Most questions prompted users to select the most appropriate diagnosis or diagnostic test based on a presented constellation of symptoms. Subsequently, we explored the functionality of Open Lecture²⁸, an educational plugin that retrieves university lecture content related to a user's prompt. However, we observed the questions were not advanced enough to endorse ChatGPT as a stand-alone question bank.

Table 1. Sample USMLE-style questions generated by ChatGPT.

1. A 45-year-old woman presents to the clinic with complaints of numbness and tingling in her right hand, which is worse at night and relieved by shaking. On physical examination, Tinel's sign and Phalen's maneuver are positive. Which of the following tests would be most appropriate to confirm the diagnosis of carpal tunnel syndrome?
A. X-ray of the wrist
B. MRI of the wrist
C. Nerve conduction studies
D. Electromyography
Answer: C. Nerve conduction studies are the most appropriate test to confirm the diagnosis of carpal tunnel syndrome.
2. A 55-year-old man presents with a complaint of difficulty grasping objects with his right hand. On physical examination, he has weakness of the thumb and index finger, along with atrophy of the thenar eminence. Which nerve is most likely affected in this patient?
A. Ulnar nerve
B. Radial nerve
C. Median nerve
D. Axillary nerve
Answer: C. The median nerve is most likely affected in this patient, as it supplies the muscles of the thenar eminence and the first two lumbricals.
3. A 35-year-old female presents with a 6-month history of wrist pain and tingling in her fingers. She works in a factory and performs repetitive wrist movements. On physical examination, she has positive Tinel's sign and Phalen's maneuver. Which of the following is a non-surgical treatment option for her condition?
A. Wrist splinting
B. Steroid injection
C. Carpal tunnel release surgery
D. Physical therapy
Answer: A. Wrist splinting is a non-surgical treatment option for carpal tunnel syndrome and can help to relieve symptoms in many cases.
4. A 60-year-old man with a history of carpal tunnel syndrome presents to the clinic with worsening symptoms. On physical examination, he has atrophy of the thenar eminence and weakness of the thumb and index finger. Which of the following is a surgical treatment option for his condition?
A. Wrist splinting
B. Steroid injection
C. Carpal tunnel release surgery
D. Physical therapy
Answer: C. Carpal tunnel release surgery is a surgical treatment option for carpal tunnel syndrome, particularly in cases where non-surgical treatments have been ineffective.

109

110

Overall, we envisage LLMs as beneficial adjunctive learning tools and preparatory resources for both medical students and faculty educators. Nonetheless, there are significant caveats. While rapid access to information can aid in achieving short-term objectives, mastering a clinical concept necessitates time, the assimilation of diverse material, and deliberate practice.²⁹ Although plugin integration has expanded functionality, these applications cannot substitute for comprehensive lecture and textbook material. Educators who use LLMs to augment learning material should underscore this in their curriculum. Furthermore, relying solely on ChatGPT for guidance in clinically complex scenarios is not advisable as its accuracy in highly specialized fields remains unverified.³⁰ Furthermore, it should be highlighted that LLMs may produce misleading or inaccurate information, when presented with poorly refined prompts by less knowledgeable users.^{2,31} While ChatGPT performed reasonably well when queried for factual information (**Table 1**), it cannot substitute years of clinical acumen required to respond or address vague symptomatology or an atypical presentation. Taking all this into account, we advocate for the utilization of ChatGPT as a supplementary tool for students' self-directed learning, alongside engagement with peer study groups and faculty mentors.

LLM chatbots could potentially aid personalized preclinical education and tailor learning. The recent adoption of pass-fail scoring for USMLE Step 1 reflects a shift from standardized exam performance to an emphasis on critical thinking and problem-solving skills.^{32,33} In a recent commentary, Prober et al. advocated a holistic profiling of students to advance the development of a precision education system. This proposed approach encompasses prior assessments, activities, as well as individual characteristics and preferences.³⁴ LLMs may facilitate this approach by analyzing these intricate datasets, aiding in the customization of sophisticated training plans.

2. CLINICAL CLERKSHIPS

Clinical clerkships are a pivotal and transformative aspect of medical school. This is exemplified by an increasing number of U.S. medical schools adopting an earlier start to clinical education and training, often accompanied by a shortened preclinical phase.^{14,35} Although technology has fundamentally reshaped the preclinical years, clinical clerkships remain rooted in the apprenticeship model. In this section, we propose ways LLMs could augment clinical preparedness and learning by supporting simulation-based training and medical note-writing.

2.1 Simulation-based training

Simulation-based training facilitates learning in a low-stakes environment. Standardized patients (SPs) are ‘trained patients’ who simulate clinical scenarios for student teaching and assessment.³⁶ SP encounters help stimulate students’ critical thinking, communication skills, and confidence.³⁷ SPs are also used to assess student history-taking and physical exam skills. As telemedicine visits increasingly become a routine facet of clinical practice, more institutions and educators have begun using virtual SP encounters in communication training modules.³⁸

Medical educators can potentially leverage LLMs to produce simulated encounters as an innovative and cost-effective alternative to human SPs, where the latter requires intensive faculty effort to train and administer simulation activities.^{39,40} For example, ChatGPT can potentially be used to generate virtual patient encounters, where students can practice history-taking, clinical reasoning, and communication skills.⁴¹ In turn, educators can devote more time for individualized feedback and coaching.^{36,41} By generating multiple patient vignettes, LLMs could provide greater range of training materials and better access to student-patient role-playing scenarios, compared to traditional SP-based curriculum.

Preliminary studies which investigated the use of AI for developing clinical reasoning and communication skills, suggested that more refinement and medical educator input are needed. For example, Shorey et al. assessed the use of AI-generated virtual patients for practicing communication skills in nursing students. They found students receiving AI-training content retained high levels of communication self-efficacy, but exhibited poorer clinical communication relative to students who did not.⁴² Thus, the authors posited that virtual simulations may not be intellectually stimulating or challenging enough to yield substantial learning outcomes.⁴² Similarly, a study on pharmacy students found no difference in communication scores between students with and without supplemental AI-training.⁴³ In a recent study, ChatGPT was used to create a “breaking bad news” training scenario for emergency medicine physician trainees.⁴⁴ It was observed that although ChatGPT provided constructive feedback based on a validated framework, its efficacy relied heavily on user training, prompt refinement, and supervision.⁴⁴ Hence, we advocate for medical educators to not only familiarize themselves with ChatGPT but also comprehend the principles and engineering behind well-conceived prompts.⁴⁵

It is worth noting that LLM-generated simulated patient encounters are currently limited to text-based interactions. Thus, the LLM is not capable of assessing hands-on clinical skills, emotional responses, or body language during a simulated encounter. Nevertheless, the potential integration of LLMs with virtual reality could lend to a more immersive learning experience, potentially capturing intangible human components such as conversational pauses and eye contact^{46,47}.

Nevertheless, the ethical dimensions surrounding AI-generated virtual patients cannot be overlooked. Experts warn that such technology might perpetuate negative stereotypes, given the potential bias ingrained in the training data.⁴⁸ Students ought to understand that LLM-chatbots

responses are unlikely to represent the full diversity of patients encountered in real-world clinical settings.⁴⁹ Despite their sophistication, LLM-chatbots like ChatGPT can confabulate, i.e., present inaccurate information or fabricate references.⁵⁰ Hence, it is pertinent for medical educators to discern plausible yet inaccurate responses from LLMs, as these could translate to patient harm if overlooked. Automation bias, the inclination to overly rely on computerized decision support despite contradictory evidence, is a concern among physicians and medical educators, underscoring the need for vigilant oversight of AI-based virtual patient learning.⁵¹

2.2 Medical Note-Writing Support

Physicians can dedicate up to 50% of their working time on patient care documentation.⁵²⁻
⁵⁴ The importance of a well-written note cannot be overstated, as it synthesizes relevant details in a patient encounter, facilitates billing, reconciles outside records, and informs other healthcare professionals involved in the patient's care. While medical note-writing traditionally has been the responsibility of attending physicians and residents, it is an essential skill for medical students to acquire prior to entering residency.^{55,56}

With the inception of ChatGPT, it is conceivable to use LLMs to streamline documentation, which may have substantial implications for medical student learning. For example, Lee et al. proposed that ChatGPT could listen in on a patient encounter, akin to “smart speakers”, and then transcribe and format the interaction into a patient clinic note.² This could be potentially achieved by using the Whisper API⁵⁷ to transform audio recordings into text and subsequently linked with ChatGPT to condense and extract essential points. Such an LLM-derived note could serve as a template for educators to evaluate the thoroughness of student history-taking. Future LLMs could

be customized to extract relevant details from the electronic health record, mitigating omissions of key patient-reported details and consequently elevating the quality of care. On this note, recently, Microsoft and Epic Systems announced collaborations with OpenAI to utilize LLMs for drafting patient messages to boost physician productivity.⁵¹

Despite the allure, there are several important considerations when using LLMs for clinical educational activities. Besides reported instances of inaccurate or misleading transcription and privacy concerns,² over-reliance on ChatGPT could impede student's learning. As shown in a previous qualitative study, note-writing has a major role in honing skills to identify knowledge gaps, improving decision-making abilities and critical thinking.⁵⁸

3. RESEARCH ELECTIVES

Enhancing Research Engagement through LLMs

Academic research is a crucial complement to medical education.⁵⁹ LLMs such as ChatGPT offer tools which may aid in this endeavor. As a result of ChatGPT's functionality, performing research may be more accessible to medical students. Here, we illustrate some of these applications in a research process (**Figure 3**).

Utilizing ChatGPT, students can initiate their research journey by inputting a description of their research topic of interest and available resources. Subsequently, ChatGPT can draft an initial research strategy and even retrieve pertinent references.⁶⁰ Through plugins such as ScholarAI⁶¹ and Science⁶², or its web-browsing function (for ChatGPT Plus subscribers),⁶³ ChatGPT can fetch abstracts or even extract key details such as study design, methodology, and results into a formatted table.⁶⁴

In addition, ChatGPT can also aid data analysis by suggesting suitable statistical tests based on the data structure and produce sample codes for statistical software use.⁶⁰ For instance, plugins like Noteable enable users to upload data files, run code, and collaborate with others.⁶⁵ Following data analysis, students can leverage ChatGPT to improve the flow and organization of manuscript, which can be particularly useful for non-native English-speaking students.¹⁸ Additionally, students may use ChatGPT to translate English-written publications into various languages thereby reducing barriers and supporting equity in academic research.⁶⁶ Lastly, students can also request ChatGPT to style and format writing or paper abstracts to adhere to a specific journal's format.

Nevertheless, several concerns linger. When utilized for literature review, researchers found that ChatGPT may yield erroneous references, if left unchecked.^{50,60,67} In a mini experiment, we explored using ChatGPT3.5 to request for articles related to care disparity in carpal tunnel syndrome. We found that 9 out of 10 given publications by ChatGPT actually were not present in the PubMed database. Adding to this concern, ChatGPT generated a seemingly realistic abstract for one of these articles, which a novice researcher could readily perceive as legitimate (**Supplemental Figure 3**). More recent scientific research plugins for ChatGPT Plus subscribers, such as ScholarAI and Science, have been developed to mitigate these “hallucinations.”⁶⁸ These plugins interface with open-access databases, thus improving the reliability of retrieved results. Once extracted, the plugins were able to provide a concise summary of each article and a link to the full-text article. Students using LLMs for academic research must remain cognizant of the evolving editorial policies governing their use. Worthy to note, leading publishing groups such as Elsevier and Springer-Nature prohibit listing ChatGPT as co-author but permit its use “to improve the readability and language” of submitted manuscripts.⁶⁹ Given the breadth of these concerns,

244 faculty researchers should proactively discuss and emphasize best practices for using LLMs with
245 students to uphold academic and research integrity.

247 **CONCLUSION**

248 The rise of ChatGPT and other LLMs have taken the world by storm. Healthcare
249 stakeholders, higher education institutions, and medical educators should holistically evaluate the
250 promises and perils of LLMs. On one hand, embracing LLMs can optimize learning, level the
251 playing field, and streamline mundane tasks for students and educators alike thereby freeing up
252 more time for other learning activities. On the other hand, overuse of LLMs may result in
253 oversimplification of learning, constraint of independent thinking, and potential misconduct in
254 academic research.

255 While the adoption of LLMs in the clinical realm in healthcare and medical school setting
256 appears promising, physicians and medical educators ought to oversee their safe deployment to
257 ensure that they complement learning while upholding high standards of patient care and safety.
258 While medical students can be encouraged to use AI as learning tool, critical thinking, ethical
259 judgment, and patient connection remain the foundational pillars of medical practice. These cannot
260 be easily replaced by AI. As LLMs continue to push the boundaries, further research and
261 conceptualization on the technology's application are needed to generate evidence and guide the
262 use of LLMs in medical education.

Figure legends:

Figure 1: Potential Educational Uses of Large Language Model Chatbots in Medical Education.

Figure 2: Images of the Phalen Maneuver Generated Using the Argil Plugin in ChatGPT.

Figure 3: Potential Integration of ChatGPT into a Research Process.

Funding Statement: This work was supported by National Key Research and Development Program of China [2022YFC2502802], and National Natural Science Foundation of China [8238810007]

Competing Interests Statement: The authors have no competing interests to declare.

Data Availability Statement: The data underlying this article are available in the article and in its online supplementary material.

Contributorship Statement: TMB, YX, YCT, TYW, and KCC conceived and designed the study. TMB, YX, YCT and KCC generated and edited illustrations. TMB, YX, JDB, AWCK, FB, XW, XS, YW, DC, YCT, TYW, and KCC wrote the manuscript. All authors reviewed and edited the manuscript. All authors approved the final manuscript.

REFERENCES

1. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing Artificial Intelligence Training in Medical Education. *JMIR Med Educ*. 2019;5(2):e16048.
2. Lee P, Bubeck S, Petro J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *The New England journal of medicine*. 2023;388(13):1233-1239.
3. Moor M, Banerjee O, Abad ZSH, et al. Foundation models for generalist medical artificial intelligence. *Nature*. 2023;616(7956):259-265.
4. Haupt CE, Marks M. AI-Generated Medical Advice-GPT and Beyond. *Jama*. 2023;329(16):1349-1350.
5. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nature Medicine*. 2023;29(8):1930-1940.
6. Wang X, Sanders HM, Liu Y, et al. ChatGPT: promise and challenges for deployment in low- and middle-income countries. *The Lancet regional health Western Pacific*. 2023;41:100905.
7. Clusmann J, Kolbinger FR, Muti HS, et al. The future landscape of large language models in medicine. *Communications Medicine*. 2023;3(1):141.
8. Bair H, Norden J. Large Language Models and Their Implications on Medical Education. *Academic Medicine*. 2023;98(8).
9. Safranek CW, Sidamon-Eristoff AE, Gilson A, Chartash D. The Role of Large Language Models in Medical Education: Applications and Implications. *JMIR Med Educ*. 2023;9:e50945.
10. Ng FYC, Thirunavukarasu AJ, Cheng H, et al. Artificial intelligence education: An evidence-based medicine approach for consumers, translators, and developers. *Cell Reports Medicine*. 2023;4(10):101230.
11. Cho KK, Marjadi B, Langendyk V, Hu W. The self-regulated learning of medical students in the clinical environment - a scoping review. *BMC Med Educ*. 2017;17(1):112.
12. Ikonne U, Campbell AM, Whelihan KE, Bay RC, Lewis JH. Exodus From the Classroom: Student Perceptions, Lecture Capture Technology, and the Inception of On-Demand Preclinical Medical Education. *J Am Osteopath Assoc*. 2018;118(12):813-823.
13. Scott K, Morris A, Marais B. Medical student use of digital learning resources. *Clin Teach*. 2018;15(1):29-33.
14. Emanuel EJ. The Inevitable Reimagining of Medical Education. *JAMA*. 2020;323(12):1127-1128.

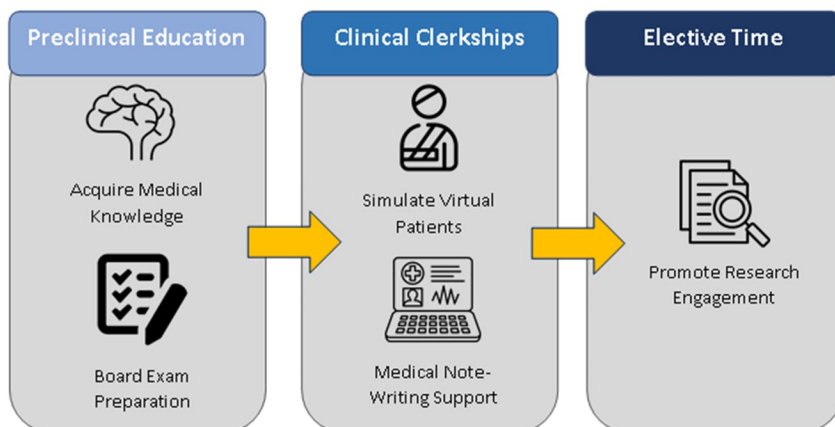
- 314 15. Chang C, Colón-Berlinger M, Mavis B, Laird-Fick HS, Parker C, Solomon D. Medical
315 Student Progress Examination Performance and Its Relationship With Metacognition,
316 Critical Thinking, and Self-Regulated Learning Strategies. *Acad Med*. 2021;96(2):278-
317 284.
- 318 16. Chick RC, Clifton GT, Peace KM, et al. Using Technology to Maintain the Education of
319 Residents During the COVID-19 Pandemic. *J Surg Educ*. 2020;77(4):729-732.
- 320 17. Choi-Lundberg DL, Low TF, Patman P, Turner P, Sinha SN. Medical student preferences
321 for self-directed study resources in gross anatomy. *Anat Sci Educ*. 2016;9(2):150-160.
- 322 18. Kasneci E, Sessler K, Kuchemann S, Bannert M. *ChatGPT for Good? On Opportunities*
323 *and Challenges of Large Language Models for Education*. 2023.
- 324 19. AI O. Chat Plugins. 2023; <https://platform.openai.com/docs/plugins/introduction>.
325 Accessed July 17, 2023.
- 326 20. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT - Reshaping medical education and
327 clinical management. *Pak J Med Sci*. 2023;39(2):605-607.
- 328 21. Amboss. Your next-gen resource for med school and beyond. 2023;
329 <https://www.amboss.com/us/students>. Accessed June 9, 2023.
- 330 22. Kluwer W. UpToDate. 2023; <https://www.wolterskluwer.com/en-sg/solutions/uptodate>.
331 Accessed Nov 4, 2023.
- 332 23. Overview of AI/ChatGPT Plugin Video Insights. <https://gptstore.ai/plugins/jettel-de>.
333 Accessed July 19, 2023.
- 334 24. How to Generate ChatGPT Images Using Argil Plugin. [https://www.argil.ai/blog/how-to-](https://www.argil.ai/blog/how-to-generate-chatgpt-images-using-argil-plugin)
335 [generate-chatgpt-images-using-argil-plugin](https://www.argil.ai/blog/how-to-generate-chatgpt-images-using-argil-plugin). Accessed July 19, 2023.
- 336 25. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE:
337 Potential for AI-assisted medical education using large language models. *PLOS digital*
338 *health*. 2023;2(2):e0000198.
- 339 26. McGaghie WC, Cohen ER, Wayne DB. Are United States Medical Licensing Exam Step
340 1 and 2 scores valid measures for postgraduate medical residency selection decisions?
341 *Acad Med*. 2011;86(1):48-52.
- 342 27. Bhatnagar V, Diaz SR, Bucur PA. The Cost of Board Examination and Preparation: An
343 Overlooked Factor in Medical Student Debt. *Cureus*. 2019;11(3):e4168.
- 344 28. Overview of AI/ChatGPT Plugin Open Lecture. [https://gptstore.ai/plugins/-lorelattice-](https://gptstore.ai/plugins/-lorelattice-com)
345 [com](https://gptstore.ai/plugins/-lorelattice-com). Accessed July 17, 2023.
- 346 29. Gladwell M. *Outliers: The Story of Success*. San Francisco, CA: Little, Brown and
347 Company; 2008.

- 348 30. Liu Y, Han T, Ma S, et al. Summary of chatgpt/gpt-4 research and perspective towards the
349 future of large language models. *arXiv preprint arXiv:230401852*. 2023.
- 350 31. Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine,
351 2023. *The New England journal of medicine*. 2023;388(13):1201-1208.
- 352 32. Desai A, Hegde A, Das D. Change in Reporting of USMLE Step 1 Scores and Potential
353 Implications for International Medical Graduates. *JAMA*. 2020;323(20):2015-2016.
- 354 33. Dive HE. Stanford, Columbia, Penn, Mount Sinai medical schools withdraw from U.S.
355 News rankings. 2023; [https://www.highereddive.com/news/stanford-columbia-penn-](https://www.highereddive.com/news/stanford-columbia-penn-mount-sinai-medical-schools-withdraw-from-us-ne/641090/#:~:text=Four%20prominent%20medical%20schools%20%E2%80%94%20the%20University%20of,law%20and%20medical%20programs%20rejecting%20the%20influentia%20lists)
356 [mount-sinai-medical-schools-withdraw-from-us-](https://www.highereddive.com/news/stanford-columbia-penn-mount-sinai-medical-schools-withdraw-from-us-ne/641090/#:~:text=Four%20prominent%20medical%20schools%20%E2%80%94%20the%20University%20of,law%20and%20medical%20programs%20rejecting%20the%20influentia%20lists)
357 [ne/641090/#:~:text=Four%20prominent%20medical%20schools%20%E2%80%94%20the%20University%20of,law%20and%20medical%20programs%20rejecting%20the%20in-](https://www.highereddive.com/news/stanford-columbia-penn-mount-sinai-medical-schools-withdraw-from-us-ne/641090/#:~:text=Four%20prominent%20medical%20schools%20%E2%80%94%20the%20University%20of,law%20and%20medical%20programs%20rejecting%20the%20influentia%20lists)
358 [fluentia%20lists](https://www.highereddive.com/news/stanford-columbia-penn-mount-sinai-medical-schools-withdraw-from-us-ne/641090/#:~:text=Four%20prominent%20medical%20schools%20%E2%80%94%20the%20University%20of,law%20and%20medical%20programs%20rejecting%20the%20influentia%20lists). Accessed June 9, 2023.
- 360 34. Prober CG, Desai SV. Medical School Admissions: Focusing on Producing a Physician
361 Workforce that Addresses the Needs of the United States. *Acad Med*. 2023.
- 362 35. Murphy B. *Medical schools ponder move to shorter pre-clinical clerkship curriculum*.
363 American Medical Association;2021.
- 364 36. Bosek MS, Li S, Hicks FD. Working with standardized patients: a primer. *International*
365 *journal of nursing education scholarship*. 2007;4:Article 16.
- 366 37. Nelson G. Training Standardized Patients to Provide Effective Feedback: Development,
367 Implementation, and its Effect on the Efficacy of Medical Students' Education. *S D Med*.
368 2022;75(10):454-455.
- 369 38. Suneja M, Thoma KD, Franklin E, Rowat J. Pivoting during a Pandemic: Reimagining
370 Intern Orientation EPA-Based OSCE. *Southern medical journal*. 2023;116(3):312-316.
- 371 39. Kelly M, Murphy A. An evaluation of the cost of designing, delivering and assessing an
372 undergraduate communication skills module. *Med Teach*. 2004;26(7):610-614.
- 373 40. Gillette C, Stanton RB, Rockich-Winston N, Rudolph M, Anderson HG. Cost-
374 Effectiveness of Using Standardized Patients to Assess Student-Pharmacist
375 Communication Skills. *Am J Pharm Educ*. 2017;81(10):6120.
- 376 41. Eysenbach G. The Role of ChatGPT, Generative Language Models, and Artificial
377 Intelligence in Medical Education: A Conversation With ChatGPT and a Call for Papers.
378 *JMIR Med Educ*. 2023;9:e46885.
- 379 42. Shorey S, Ang ENK, Ng ED, et al. Evaluation of a Theory-Based Virtual Counseling
380 Application in Nursing Education. *Comput Inform Nurs*. 2023.

- 381 43. Nakagawa N, Odanaka K, Ohara H, Kisara S. Communication training for pharmacy
382 students with standard patients using artificial intelligence. *Curr Pharm Teach Learn*.
383 2022;14(7):854-862.
- 384 44. Webb JJ. Proof of Concept: Using ChatGPT to Teach Emergency Physicians How To
385 Break Bad News. *Cureus*. 2023;15(5).
- 386 45. Multiplex. What are AI SuperPrompt and why they are important. 2023. Accessed May
387 11, 2023.
- 388 46. Demszky D, Yang D, Yeager DS, et al. Using large language models in psychology.
389 *Nature Reviews Psychology*. 2023;2(11):688-701.
- 390 47. Varas J, Coronel BV, Villagrán I, et al. Innovations in surgical training: exploring the role
391 of artificial intelligence and large language models (LLM). *Revista do Colegio Brasileiro*
392 *de Cirurgioes*. 2023;50:e20233605.
- 393 48. Ali SR, Dobbs TD, Hutchings HA, Whitaker IS. Using ChatGPT to write patient clinic
394 letters. *Lancet Digit Health*. 2023;5(4):e179-e181.
- 395 49. Combs CD, Combs PF. Emerging Roles of Virtual Patients in the Age of AI. *AMA J*
396 *Ethics*. 2019;21(2):E153-159.
- 397 50. Sanchez-Ramos L, Lin L, Romero R. Beware of References when Using ChatGPT as a
398 Source of Information to Write Scientific Articles. *Am J Obstet Gynecol*. 2023.
- 399 51. Edwards B. GPT-4 will hunt for trends in medical records thanks to Microsoft and Epic.
400 *Ars Technica*2023.
- 401 52. Sinsky C, Colligan L, Li L, et al. Allocation of Physician Time in Ambulatory Practice: A
402 Time and Motion Study in 4 Specialties. *Ann Intern Med*. 2016;165(11):753-760.
- 403 53. Tipping MD, Forth VE, O'Leary KJ, et al. Where did the day go?--a time-motion study of
404 hospitalists. *J Hosp Med*. 2010;5(6):323-328.
- 405 54. Oxentenko AS, West CP, Popkave C, Weinberger SE, Kolars JC. Time spent on clinical
406 documentation: a survey of internal medicine residents and program directors. *Arch*
407 *Intern Med*. 2010;170(4):377-380.
- 408 55. Association of American Medical Colleges. *Core Entrustable Professional Activities for*
409 *Entering Residency Curriculum Developers' Guide*. Washington D.C.2014.
- 410 56. Amiel JM, Andriole DA, Biskobing DM, et al. Revisiting the Core Entrustable
411 Professional Activities for Entering Residency. *Acad Med*. 2021;96(7S):S14-S21.
- 412 57. OpenAI. Speech to text. <https://platform.openai.com/docs/guides/speech-to-text>.

- 413 58. Bowker D, Torti J, Goldszmidt M. Documentation as composing: how medical students
414 and residents use writing to think and learn. *Adv Health Sci Educ Theory Pract*. 2022.
- 415 59. Pathipati AS, Taleghani N. Research in Medical School: A Survey Evaluating Why
416 Medical Students Take Research Years. *Cureus*. 2016;8(8):e741.
- 417 60. Macdonald C, Adeloye D, Sheikh A, Rudan I. Can ChatGPT draft a research article? An
418 example of population-level vaccine effectiveness analysis. *Journal of global health*.
419 2023;13:01003.
- 420 61. ScholarAI. ScholarAI. 2023; <https://scholar-ai.net/>. Accessed Nov 4, 2023.
- 421 62. OA.mg. "Science" ChatGPT Plugin. 2023; <https://oa.mg/chatgpt>. Accessed Nov 4, 2023.
- 422 63. REUTERS. ChatGPT users can now browse internet, OpenAI says. 2023;
423 [https://www.reuters.com/technology/openai-says-chatgpt-can-now-browse-internet-2023-](https://www.reuters.com/technology/openai-says-chatgpt-can-now-browse-internet-2023-09-27/)
424 [09-27/](https://www.reuters.com/technology/openai-says-chatgpt-can-now-browse-internet-2023-09-27/). Accessed Nov 4, 2023.
- 425 64. Hutson M. Could AI help you write your next paper? *Nature*. 2022;611:192-193.
- 426 65. Noteable. ChatGPT Plugin. <https://noteable.io/chatgpt-plugin-for-notebook/>. Accessed
427 July 19, 2023.
- 428 66. Amano T, Rios Rojas C, Boum Ii Y, Calvo M, Misra BB. Ten tips for overcoming
429 language barriers in science. *Nat Hum Behav*. 2021;5(9):1119-1122.
- 430 67. Alkaissi H, McFarlane SI. Artificial Hallucinations in ChatGPT: Implications in
431 Scientific Writing. *Cureus*. 2023;15(2):e35179.
- 432 68. Overview of AI/ChatGPT Plugin Scholar AI. <https://gptstore.ai/plugins/scholar-ai-net>.
433 Accessed July 19, 2023.
- 434 69. Sample I. Science journals ban listing of ChatGPT as co-author on papers. *The*
435 *Guardian*2023.

436





Used Argill AI

Sure, here is an illustration of the Phalen's test maneuver for carpal tunnel syndrome:



In this test, the patient is asked to flex their wrists and press the backs of their hands against each other for about a minute. If numbness or tingling is felt in the fingers within this time, it may suggest carpal tunnel syndrome.



Used Argill AI

I apologize for the confusion earlier. Here is a correct illustration of the Phalen's test maneuver for carpal tunnel syndrome:



In this test, the patient is asked to flex their wrists and press the backs of their hands against each other for about a minute. If numbness or tingling is felt in the fingers within this time, it may suggest carpal tunnel syndrome.

