THE ASIA-PACIFIC SCHOLAR
MEDICAL AND HEALTH PROFESSIONS EDUCATION

PERSONAL VIEW

Check for updates

Submitted: 19 February 2023
Accepted: 10 July 2023
Published online: 3 October, TAPS 2023, 8(4), 50-52
https://doi.org/10.29060/TAPS.2023-8-4/PV3007

Large language models (ChatGPT) in medical education: Embrace or abjure?

Nathasha Luke[1], Reshma Taneja[1], Kenneth Ban[2], Dujeepa Samarasekera[3] & Celestial T Yap[1]

[1]Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [2]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [3]Centre for Medical Education, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

Health Professional Education has considerably evolved over the years. Traditional classroom teaching has shifted to blended learning modalities, and clinical teaching has embraced virtual reality and simulation-based learning.

Education is poised for another major change with the development of artificial intelligence (AI) models that can emulate human-like intelligent behaviour, particularly in the field of large language models (LLM) that are capable of generating text in response to user input. There has been remarkable progress in the development of these models, with each iteration having an increasing ability to generate human-like responses to user input.

In November 2022, Open AI released ChatGPT. This marked a major milestone in the ability of LLMs. This leap in performance was driven in part by the training of the model on large text datasets from various sources such as books, articles, and websites. It was combined with supervised learning and reinforcement learning to fine-tune the model based on human feedback about the quality of the output. This was further augmented by the release of GPT-4, a further advanced version in early 2023.

A study demonstrated that ChatGPT was able to pass or preform at near parsing threshold in the United States Medical Licensing Examination (USMLE) (Kung et al., 2023). Also, ChatGPT passed a law entrance examination at a level equivalent to a C+ grade (Kelly, 2023). These studies highlight the potential of modern LLMs to impact education.

Despite its impressive performance, LLMs have limitations. These caveats notwithstanding, when educators and students are aware of the capabilities and limitations, LLM tools could provide opportunities to improve the way we teach and the way students learn. More evidence is needed to depict a specific model as if how this technology could be incorporated. This article particularly focuses on capabilities and limitations of LLMs in the context of medical education with suggestions on how this technology might be used. There is a huge scope for discussion on the impact of LLMs in various dimensions of medical education. However, we limit this discussion to commonest domains pertaining undergraduate medical education.

Being widely available and accessible to educators and students around the world including resource poor settings, LLMs promote equity in medical education. Certain educational institutes have customised learning platforms to support student learning while such amenities are not accessible in resource poor settings. On the contrary, the technology of LLMs could at least partially counterbalance such shortages promoting equity.

Copyright © 2023 TAPS. All rights reserved.

In addition, Universities in resource-poor settings often find inadequate number of educators as a barrier to implement new teaching strategies and curriculum reformation, particularly with the shifting paradigms to competency based medical education (Ramanathan et al., 2022). Effective incorporation of LLM tools could ease their workload to some extent, providing more time to explore new teaching pedagogies and scope for curricular reformations.

LLM tools are being adopted in medical education, and assistive in both clinical and non-clinical settings, as discussed below. In non-clinical settings, the following are some areas where LLMs are helpful.

From an educator's perspective, LLM tools are helpful in generating summaries, quizzes, and flashcards to make the learning interesting.

From learners' perspective, LLM tools can generate customised information within a short span of time. For example, we may ask the LLM tool to answer a question 'at a level of a medical student' or 'at the level of a resident', to generate distinct responses. This will assist students in self-learning and understanding difficult concepts. LLM tools are also helpful in generating mnemonics, poems, and flashcards. Students who are not native English speakers will have the added advantage of AI tools being assistive in improving language skills.

Additionally, LLMs are assistive tools in many stages of research including design and development, implementation, literature survey and data analysis.

There is emerging literature illustrating capabilities of LLMs as useful tools in medical practice (Lee et al., 2023). Though there aren't many publications evaluating implications of current LLM technology in the domain of clinical education per -se, this technology is likely to be helpful in the development of skills such as history-taking and doctor-patient communication. Customized chatbots have been used by some medical schools to improve history-taking skills. Development of such tools are not affordable in resource-poor settings. LLM tools such as ChatGPT are not capable of 'acting' as a model patient to practice history taking. However, they can guide students to formulate relevant questions for effective history-taking in specific scenarios. Students should not be encouraged to use AI tools as the sole reference to guide the task, rather use it as an adjunct to ones' thought process. For instance, if a student encounters a patient with palpitations, he or she should brainstorm based on theoretical knowledge to formulate relevant questions to be asked in history taking. An LLM tool can be an adjunct to 'cross-check' if all salient points were captured.

LLMs could potentially be assistive in improving patient communication skills among students and junior doctors.AI systems could aid in generating facts that are comprehensible to non-medical personnel. This ability is particularly helpful for students in generating content to practice patient communication skills. Accuracy and suitability of such information should be objectively assessed, before recommending LLM tool use for this purpose.

The LLM technology is a tool that can augment the process of multi- dimensional education, encompassing competency-based approaches to education, in addition to discipline-based education. This multidimensional approach comprises knowledge along with various other skills including professionalism, communication, practice-based learning, and patient care. This approach demands more commitment from students and educators and requires more learning resources. With the emergence of more sophisticated AI technologies, harnessing of LLM capabilities could be explored as future learning resources to be developed.

Unbundling and rebundling the curriculum is a concept that emphasizes revising the existing curricula by combining various educational resources including textbooks, lectures and web-based resources (Morris et al., 2018). This is pertinent to the multi-dimensional approach discussed above. The advances in current LLM tools have the potential to become an integral component of the curriculum bundle to meet the demands of reformations in modern medical education.

Potential negative implications of LLM tools have caused anxiety among educators. Firstly, the content generated may not be accurate at all times. LLMs generate responses from language patterns learnt from the training data and not from a deeper understanding of a subject. This is also compounded by the inability of some LLMs to link to external resources to gauge the validity of the output. Another limitation of the current LLM technology is 'hallucination' to create non-existent or wrong information in a convincing manner. (Lee et al., 2023) Consequences of such information could cause huge impact particularly in patient safety in medical education.

LLMs have potential negative implications on the evaluation of students' learning. In modern-day multidimensional education, assessments have moved a long way from traditional examinations to include

Copyright © 2023 TAPS. All rights reserved.

projects, assignments, and research. Certain assignments are designed to foster the development of critical thinking and analytical skills. AI tools may direct students to take an easier path in completing these tasks, impeding the accomplishment of intended learning outcomes.

There is no consensus on how this technology should be adopted in higher education. At the start, certain educational institutes banned the use of LLMs, and software was developed to detect work done by AI. These approaches are not sustainable in the long run.

Users should have a clear understanding on potentials and imitations of current LLM technology, in order to use LLMs effectively.

LLM technology is improving rapidly, and efficacious compared to many other sources of education. However they are not yet at a standpoint to be recommended as the primary source in education, rather, they could be adjuncts to standard resources like lectures, textbooks, peer-reviewed literature, and online materials. Students should know when and when not to use it, and the content should be critically and cautiously looked into.

Educators have a crucial role in guiding the students on using AI effectively. Navigating students to experience the limitations of LLMs through practical scenarios is a potential strategy. An example would be to assign students to critically analyse a draft answer generated through an LLM platform. This will allow both the students and tutors identify the capabilities and limitations of LLMs.

In the context of evolving LLMs educators have to re look into the existing assessment modalities and implement changes to ensure the potential objectives of the assessment are met. The policies regarding LLM use for the particular task should be clearly communicated to the students on contextual basis.

Impacts of LLMs on educational development is yet another area warranting discussion. This encompasses exploring the role of LLMS in instructor, instructional and organizational development. We didn't include it within the scope of this write up.

In summary, generative AI could be harnessed to potentiate students' learning, in knowledge acquisition as well as application. Even though LLM tools may pose challenges, we foresee a larger potential for the betterment of medical education, ultimately leading to the overall goal of better patient care.

Notes on Contributors

WANVL, CTY, RT, DS and KB were involved in planning of the article.

WANVL drafted the initial version of the mauscript.

WANVL, CTY, RT, DS and KB revised and edited the initial draft manuscript and approved the final version of the manuscript for submission.

Funding

The authors received no financial support for the authorship or publication of this article.

Declaration of Interest

The authors do not have any conflicts of interest to disclose.

References

Kelly, S. M. (2023, January 26). *ChatGPT passes exams from law and business schools*. CNN Business. https://edition.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, *2*(2), Article e0000198. https://doi.org/10.1371/journal.pdig.0000198

Lee, P., Bubeck, S., & Petro, J. (2023). Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *The New England Journal of Medicine*, *388*(13), 1233–1239. https://doi.org/10.1056/NEJMsr2214184

Morris, N., Ivancheva, M., Swinnerton, B., Coop, T., & Czerniewicz, L. (2018, September 11-13). *Critical perspectives on unbundling and rebundling higher education provision online* [Research session]. [18-97]. ALT Annual Conference, Manchester, UK. https://www.youtube.com/watch?v=F2FS0n3Dr0k

Ramanathan, R., Shanmugam, J., Gopalakrishnan, S. M., Palanisamy, K. T., & Narayanan, S. (2022). Challenges in the Implementation of Competency-Based Medical Curriculum: Perspectives of Prospective Academicians. *Cureus*. https://doi.org/10.7759/cureus.32838

*Celestial T Yap
Department of Physiology,
Yong Loo Lin School of Medicine,
National University of Singapore
+6590560468
Email: phsyapc@nus.edu.sg

Copyright © 2023 TAPS. All rights reserved.