

Biomedical education in the era of large language models: a paradigm shift

Shen-Han Lee

Department of Otorhinolaryngology–Head & Neck Surgery, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, and Hospital Pakar Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia.

The year is 2050.

The golden age of biomedical progress has ended. Science has lost its soul. Scientists do not ask “Why?” anymore. What changed?

In the early 2020s, the rise of artificial intelligence (AI) and deep learning models known as large language models (LLMs) promised to revolutionize science. Instead, the same tools created a generation of scientists who lost their ability to think critically and innovate beyond the output of their LLMs.

Jonathan, a PhD student, sits in his office staring at the computer screen. He keys in data from his latest cancer immunotherapy experiments for the LLM to analyze and watches as the model generates its conclusions. He does not question the results.

Several months later, at a conference question and answer session, a senior scientist challenges Jonathan: “Your model predicts a novel biomarker, but how do you explain its biological relevance?” Jonathan stares blankly. Flustered, he quickly types the question to the LLM on his laptop. A few elder scientists in the audience whisper to each other, “Remember when we used to think for ourselves?”

The year is 2024.

This was the year when AI was awarded two Nobel Prizes – the Physics prize for work on machine learning with artificial neural networks, the technology that powers LLMs (1); and half of the Chemistry prize, for the development of deep learning algorithms that solved a 50-year protein structure prediction challenge (2).

This was also the year when the world experienced a Cambrian explosion of LLMs. Several new models, such as GPT-4 (3), Gemini (4), Claude (5), LLaMA (6), and DeepSeek (7) were released only barely two years since ChatGPT first took the world by storm (8). Pretrained on vast amounts of text data, these deep learning models are fine-tuned for specific natural language processing tasks, such as summarizing, question answering, and classification (9).

The strength of LLMs lies in their ability to extract information across the vast scientific literature and distill it into responses that are easy to understand. Scientists can now use readily available LLMs to parse and refine their ideas, automate data analysis, summarize documents, and disseminate research findings more easily. LLMs have also been used as AI agents capable of autonomously performing tasks of a research assistant, searching the Web and documents and executing tasks related to experiments (10).

“Education is not the preparation of life; education is life itself.” This quote by 19th century educational reformer John Dewey holds true in the rapidly evolving AI landscape today. To harness the potential of LLMs, students and educators must devel-

op the competencies and literacies to understand this technology, its limitations, and ethical concerns. For educators, LLMs can be used for curriculum development and creation of educational content, personalized learning experiences, and assessments for students (11). For students, LLMs can be used to concisely explain unfamiliar technical concepts tuned to their level of understanding, summarize the content of scientific papers, answer specific technical questions about methodology, and even provide suggestions for new research ideas based on existing literature.

By virtue of its scalability, LLMs can promote a more sustainable and inclusive biomedical workforce by breaking down barriers to education in low-resource settings. LLMs with multilingual capabilities can deliver biomedical training globally to non-native English speakers in their native languages, reducing the linguistic barrier to education (12). They can also cater to learners with disabilities through voice-to-text and text-to-speech capabilities (13).

Avoiding a dystopian AI future

Notwithstanding the immense benefits of LLMs, without a paradigm shift in biomedical education, future scientists may become overly reliant on LLMs and risk losing their ability to think critically and innovate. Therefore, the goals of post-LLM biomedical education should be to train scientists who use AI as an augmentation tool rather than a replacement for human reasoning, and cultivate thinkers rather than mere passive users of AI tools. These goals may be achieved by focusing on cross-disciplinary thinking, scientific epistemology, and the history and philosophy of science. There should also be a greater emphasis on exploring the creative process of generating ideas for scientific research to answer the question “Where do ideas come from?” (14). These skills will be crucial for a future biomedical workforce capable of merging human intelligence with AI.

Coda: the year is 2050

Biomedical science is in a golden age of creativity, innovation, and discovery. LLMs are no longer crutches for lazy thinking but tools for scientists to think deeper and work more efficiently. What changed? In the mid-2020s, biomedical education evolved to meet the needs of an AI-driven world. Education systems were redesigned to emphasize interdisciplinary knowledge and critical and creative thinking, alongside AI literacy. The use of LLMs sustainably and inclusively scaled biomedical education globally, giving birth to a new generation of biomedical scientists well equipped to thrive in an AI-powered future.

1. The Nobel Prize in Physics 2024. <https://www.nobelprize.org/prizes/physics/2024/press-release>. Accessed June 17, 2025.
2. The Nobel Prize in Chemistry 2024. <https://www.nobelprize.org/prizes/chemistry/2024/press-release>. Accessed June 17, 2025.
3. OpenAI, et al. GPT-4 technical report [published online March 15, 2023]. *arXiv*. <https://doi.org/10.48550/arXiv.2303.08774>.
4. Gemini Team, et al. Gemini: a family of highly capable multimodal models [published online December 18, 2023]. *arXiv*. <https://doi.org/10.48550/arXiv.2312.11805>.
5. Anthropic. Introducing Claude 3.5 Sonnet. <https://www.anthropic.com/news/claude-3-5-sonnet>. Updated June 20, 2024. Accessed June 17, 2024.
6. Touvron H, et al. LLaMA: open and efficient foundation language models [published online February 27, 2023]. *arXiv*. <https://doi.org/10.48550/arXiv.2302.13971>.
7. DeepSeek-AI, et al. DeepSeek-R1: incentivizing reasoning capability in LLMs via reinforcement learning [published online January 22, 2025]. *arXiv*. <https://arxiv.org/abs/2501.12948v1>.
8. OpenAI. Introducing ChatGPT. <https://openai.com/index/chatgpt>. Updated November 30, 2022. Accessed August 27, 2024.
9. Thirunavukarasu AJ, et al. Large language models in medicine. *Nat Med*. 2023;29(8):1930–1940.
10. Boiko DA, et al. Autonomous chemical research with large language models. *Nature*. 2023;624(7992):570–578.
11. Abd-Alrazaq A, et al. Large language models in medical education: opportunities, challenges, and future directions. *JMIR Med Educ*. 2023;9(1):e48291.
12. Peláez-Sánchez IC, et al. The impact of large language models on higher education: exploring the connection between AI and Education 4.0. *Front Educ (Lausanne)*. 2024;9:1392091.
13. Kasneci E, et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn Individ Differ*. 2023;103:102274.
14. Yanai I, Lercher M. Night science. *Genome Biol*. 2019;20(1):179.