The future of large language models in fighting emerging outbreaks: lights and shadows

We read with great interest a comment published in *The Lancet Microbe* by Albert Jan van Hoek and colleagues.[1] The authors highlighted the potential role of large language models (LLMs), particularly specialised LLM-based agents, in assisting scientists and policy makers in real-time analytics pipelines for emerging outbreaks.[1] We believe that LLMs, such as the newly updated GPT-4o from OpenAI, have enhanced vision and audio capacities that could directly assist physicians in diagnosing emergent or re-emergent infectious diseases.

Van Hoek and colleagues emphasised that real-time evidence-based responses could restrict disease spread and save lives, particularly in the context of mpox and COVID-19 outbreaks.

Machine learning-based artificial intelligence systems are currently being deployed for mpox diagnosis.[2] Similarly, trained LLMs, with sophisticated vision and data analysis capabilities, might serve as effective tools in evaluating and diagnosing unknown or poorly understood skin or mucosal lesions caused by emerging or re-emerging pathogens, such as the monkeypox virus. Specifically, LLMs capable of simultaneously analysing clinical-anamnestic data and medical images could support health-care professionals in preliminary evaluations. Additionally, making these LLMs accessible through smartphones or other low-demand devices could be invaluable in remote or low-resource settings.

In the same way, preliminary studies have indicated that machine learning can detect subtle changes in the voice features of individuals with COVID-19, caused by alterations in body structures involved in vocal tasks.[3] Advanced LLMs could be used for voice assessment in diagnosing COVID-19 or other respiratory infections, serving as a far more accurate and scalable mass-screening tool than traditional PCR or rapid diagnostic tests. Integrating outbreak analytical workflows of LLMs and specialised LLM-based agents, following a careful evaluation of their abilities to analyse and perform dedicated tasks, might represent a major advancement in responding to an outbreak. Furthermore, the vision and audio capacities of advanced LLMs could be beneficial in situations requiring timely and accurate information, such as massive screening initiatives or in low-resource settings.

In conclusion, the adoption of LLMs requires a thorough evaluation of the models themselves. Literature shows that current evaluations could overestimate the performance of these models without fully considering data leakage and benchmark contamination.[4,5] These issues, although relevant, are particularly concerning when LLM models are used for diagnosis in the medical field.

We declare no competing interests.

Alberto Rizzo, Enrico Mensa, Andrea Giacomelli
rizzo.alberto@asst-fbf-sacco.it

Laboratory of Clinical Microbiology, Virology and Bioemergencies (AR) and III Infectious Diseases Unit (AG), ASST Fatebenefratelli Sacco, Luigi Sacco Hospital, Milan 20157, Italy; Department of Computer Science, University of Turin, Turin, Italy (EM); Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan, Italy (AG)

1 van Hoek AJ, Funk S, Flasche S, et al. Importance of investing time and money in integrating large language model-based agents into outbreak analytics pipelines. *Lancet Microbe* 2024; **5:** 100881.
2 Thieme AH, Zheng Y, Machiraju G, et al. A deep-learning algorithm to classify skin lesions from mpox virus infection. *Nat Med* 2023; **29:** 738–47.
3 Robotti C, Costantini G, Saggio G, et al. Machine learning-based voice assessment for the detection of positive and recovered COVID-19 patients. *J Voice* 2024; **38:** 796.e1–13.
4 Li C, Flanigan J. Task contamination: language models may not be few-shot anymore. *AAAI* 2024; **38:** 18471–80.
5 Balloccu S, Schmidtová P, Lango M, Dusek O. Leak, cheat, repeat: data contamination and evaluation malpractices in closed-source LLMs. *arXiv* 2024; published online Feb 22. https://doi.org/10.48550/arXiv.2402.03927 (preprint).

Lancet Microbe 2024;5: 100954
Published Online July 30, 2024 https://doi.org/10.1016/j.lanmic.2024.07.017
See Comment https://doi.org/10.1016/S2666-5247(24)00104-6