ONLINE FIRST

This is a provisional PDF only. Copyedited and fully formatted version will be made available soon.

DISASTER AND EMERGENCY MEDICINE JOURNAL

Large language models in emergency medicine: potential and challenges

Authors: Natasza Blek, Bartosz Wojciech Maj, Karolina Mikołap

DOI: 10.5603/demj.101288

Article type: Letter to the Editor

Submitted: 2024-06-24

Accepted: 2024-07-08

Published online: 2024-08-02

Disaster and Emergency Medicine Journal

[Letter to the Editor]

LARGE LANGUAGE MODELS IN EMERGENCY MEDICINE: POTENTIAL AND CHALLENGES

[Short title: LLMs in emergency medicine]

Natasza Blek[1] https://orcid.org/0000-0002-3213-7330, Bartosz Wojciech Maj[2], Karolina Mikołap[1]
https://orcid.org/0009-0002-3236-7061

[1]Institute of Outcomes Research, Maria Sklodowska-Curie Medical Academy, Warsaw, Poland

[2]Research Unit, Polish Society of Disaster Medicine, Warsaw, Poland

ADDRESS FOR CORRESPONDENCE:

Natasza Blek, Institute of Outcomes Research, Maria Sklodowska-Curie Medical Academy, Warsaw, Poland, e-mail: natasza.blek@uczelniamedyczna.com.pl

[Received: 24.06.2024 Accepted: 8.07.2024 Early publication date: 2.08.2024]

DOI: 10.5603/demj.101288

KEYWORDS: large language models; LLMs; emergency medicine

To the Editor,

Large language models (LLMs), a subset of artificial intelligence (AI), are remarkably good at identifying, summarizing, translating, forecasting, and producing text and other content [1]. LLMs are very receptive to different prompts and inquiries and have a broad comprehension of the enormous volume of text data they have been trained on. With so many possible uses in the emergency department (ED), LLMs hold great promise for significant improvements in healthcare. This letter provides an overview of a future in which healthcare is more accurate, efficient, and patient-centred by pointing out places where incorporating these LLM AI developments into emergency medicine situations can be advantageous.

Large language models (LLMs) have the power to transform several emergency department (ED) processes. By examining patient data — including medical history, symptoms and test results — these powerful AI systems can deliver evidence-based recommendations for diagnosis and therapy in real-time [1]. This capacity can increase decision-making speed and accuracy in the hectic ED setting. Using main complaints and other relevant information, LLMs can create triage algorithms to rapidly determine patient acuity [2, 3]. Such algorithms may shorten waiting times and enhance results by giving priority to patients who need immediate attention. A priori LLMs should not exhibit any form of bias and therefore might be more objective in certain conditions, as it was described that some ED workers have strong prejudices that could affect patients' treatment — for example, homophobic attitudes can relate to incorrect attribution of symptoms to sexually transmitted diseases, ignoring a true cause of the complaint [4].

In one of the studies, ChatGPT could correctly assess patients in the emergency department according to their main complaints. Medical report generation and other clinical documentation management can also be automated by LLMs [1], which relieves doctors of administrative duties and saves them a great deal of time. Bradshaw [5] proposed that ED discharge instructions received by the patients may successfully be generated using ChatGPT. LLMs can also build chatbots or virtual assistants to communicate with patients, respond to their queries and offer crucial medical guidance [1]. Particularly when patients have restricted instant access to healthcare professionals, this can increase patient involvement and happiness. Further helping to identify important data quickly is LLMs' ability to extract and summarize important findings from radiological reports [6].

Although large language models (LLMs) have a lot of potential in the healthcare industry, adoption of them will need to overcome several obstacles. Addressing data privacy and security issues requires ensuring adherence to laws like Health Insurance Portability and Accountability Act (HIPAA) and safeguarding patient data anonymously [7, 8]. Considered “black boxes”, many LLMs are opaque in their decision-making processes, which can impede physicians' adoption and trust [2, 8, 9]. LLMs run the danger of sustaining prejudices in their training data, which could result in differences in treatment. Various, representative datasets and close monitoring are necessary to lessen this [1, 8, 9]. The need to stress that LLMs should support the human experience rather than take its place is essential to preserve the professional worth of healthcare practitioners. Clinical judgment and critical thinking abilities run the danger of declining if physicians become unduly dependent on LLMs, even if they can improve clinical decision-making. Using LLMs presents intricate medical and ethical issues about accountability for mistakes and guaranteeing just and moral AI application [8, 9]. Particularly in the medical industry where accuracy is essential, LLMs cannot always yield correct information [8]. LLM performance is mostly influenced by the calibre and representativeness of the training data. Untrustworthy or discriminating outcomes can occur from inaccurate or biased data [8]. What is more, for smaller or less resource-rich institutions, the financial burden of establishing and sustaining LLMs in hospital settings could be prohibitive [9].

LLMs have the potential to revolutionize emergency medicine by enhancing patient care, streamlining procedures, and improving clinical decision-making. Studies demonstrated that paramedics working in ED often experience mental overload, insomnia or burnout [10, 11]. Interventions leading to relieving their occupational burden or saving time might contribute to all, as it was proven that strategic management of human resources positively influences employees, their satisfaction and engagement. Notwithstanding, ED healthcare providers must feel secure to deliver optimal care [12]. Therefore, they must be certain that new technologies implemented in the ED are reliable and trustworthy.

There are multiple important issues to consider, such as the need for model interpretability, data privacy concerns, and the potential for bias in training data. Addressing these issues is critical as research in this area advances to guarantee the safe, efficient, and fair application of LLMs in the emergency department. Medical practitioners play a vital part in this process since their knowledge and experience are priceless in directing the integration of LLMs in emergency medicine.

Article information and declarations

Acknowledgments

None.

Author contributions

All authors participated equally in the creation of the article.

Conflict of interest

The authors declare no conflict of interest.

Funding

None.

Supplementary material

None.

REFERENCES

1. Preiksaitis C, Ashenburg N, Bunney G, et al. The role of large language models in transforming emergency medicine: scoping review. JMIR Med Inform. 2024; 12: e53787, doi: 10.2196/53787, indexed in Pubmed: 38728687.
2. Lee S, Lee J, Park J, et al. Deep learning-based natural language processing for detecting medical symptoms and histories in emergency patient triage. Am J Emerg Med. 2024; 77: 29–38, doi: 10.1016/j.ajem.2023.11.063, indexed in Pubmed: 38096637.
3. Paslı S, Şahin AS, Beşer MF, et al. Assessing the precision of artificial intelligence in ED triage decisions: Insights from a study with ChatGPT. Am J Emerg Med. 2024; 78: 170–175, doi: 10.1016/j.ajem.2024.01.037, indexed in Pubmed: 38295466.
4. Oktay M, Boğan M, Sabak M, et al. Assessment of the homophobic attitudes of the emergency department professionals: descriptive survey study. Disaster Emerg Med J. 2021; 6(3): 119–124, doi: 10.5603/demj.a2021.0018.
5. Bradshaw JC. The chatgpt era: artificial intelligence in emergency medicine. Ann Emerg Med. 2023; 81(6): 764–765, doi: 10.1016/j.annemergmed.2023.01.022, indexed in Pubmed: 37210166.
6. Infante A, Gaudino S, Orsini F, et al. Large language models (LLMs) in the evaluation of emergency radiology reports: performance of ChatGPT-4, Perplexity, and Bard. Clin Radiol. 2024; 79(2): 102–106, doi: 10.1016/j.crad.2023.11.011, indexed in Pubmed: 38087683.
7. Ashenburg N, Preiksaitis C, Dayton J, et al. 312 when AI meets the emergency department: realizing the benefits of large language models in emergency medicine. Ann Emerg Med. 2023; 82(4): S136, doi: 10.1016/j.annemergmed.2023.08.337.

8. Petrella RJ. The AI future of emergency medicine. Ann Emerg Med. 2024; 84(2): 139–153, doi: 10.1016/j.annemergmed.2024.01.031, indexed in Pubmed: 38795081.

9. Cheng R, Aggarwal A, Chakraborty A, et al. Implementation considerations for the adoption of artificial intelligence in the emergency department. Am J Emerg Med. 2024; 82: 75–81, doi: 10.1016/j.ajem.2024.05.020, indexed in Pubmed: 38820809.

10. Gonczaryk A, Chmielewski J, Strzelecka A, et al. Occupational hazards in the consciousness of the paramedic in emergency medical service. Disaster Emerg Med. 2022; 7(3): 182–190, doi: 10.5603/demj.a2022.0031.

11. Gonczaryk A, Sady N, Motyl M, et al. Prevalence of sleep disturbances among emergency response team paramedics working in shift systems. Disaster Emerg Med J. 2023; 8(1): 1–9, doi: 10.5603/demj.a2023.0009.

12. Akin GC, Olcay Z, Yildrim M, et al. Effects of occupational safety performance on work engagement of emergency workers: mediating role of job satisfaction. Disaster Emerg Med J. 2024; 9(1): 23–35, doi: 10.5603/demj.97462.