How Does AI Improve Differential Diagnosis Generation?

Rasit Dinc

Abstract

Artificial intelligence (AI) is rapidly transforming the healthcare landscape, and one of its most promising applications is in improving the generation of d...

How Does AI Improve Differential Diagnosis Generation?

Author: Rasit Dinc

Introduction

Artificial intelligence (AI) is rapidly transforming the healthcare landscape, and one of its most promising applications is in improving the generation of differential diagnoses. The ability of AI, particularly large language models (LLMs), to analyze vast amounts of data and identify patterns that may be missed by human clinicians has the potential to significantly enhance diagnostic accuracy and efficiency. A comprehensive differential diagnosis is a cornerstone of medical care, and AI is proving to be a powerful tool in this critical process. This article explores how AI is being used to improve differential diagnosis generation, drawing on recent academic research to provide an overview of the current state of the field.

The Role of Large Language Models in Diagnostic Reasoning

Recent studies have demonstrated the potential of LLMs to assist clinicians in diagnostic reasoning. For example, a 2025 study published in *Nature* introduced the Articulate Medical Intelligence Explorer (AMIE), an LLM optimized for diagnostic reasoning. The study found that AMIE's standalone performance exceeded that of unassisted clinicians in generating a differential diagnosis for challenging real-world medical cases. Furthermore, clinicians assisted by AMIE produced more comprehensive and accurate differential diagnoses than those who relied on standard medical resources alone [1]. This highlights the potential of LLMs to not only provide decision support but also

to augment the diagnostic capabilities of clinicians.

AI-Powered Tools for Differential Diagnosis

Several AI-powered tools have been developed to assist with differential diagnosis. These tools leverage machine learning algorithms to analyze patient data, including symptoms, medical history, and lab results, to generate a list of potential diagnoses. For instance, DxGPT, an AI-assisted medical diagnosis platform, uses advanced language models to provide rapid differential analysis. While these tools are not intended to replace the clinical judgment of physicians, they can serve as valuable aids in the diagnostic process, helping to broaden the scope of considered diagnoses and reduce the risk of diagnostic errors. The integration of such tools into clinical workflows has the potential to streamline the diagnostic process and improve patient outcomes.

Comparing AI and Human Diagnostic Accuracy

A 2025 study published in *JAMA Network Open* compared the diagnostic accuracy of a dedicated AI expert system with that of a generative AI with a large language model. The study found that the dedicated AI diagnostic decision support system listed the correct diagnosis more often and higher up in its differential diagnosis list than the generative AI [2]. This suggests that specialized AI systems may offer advantages in diagnostic accuracy over more general-purpose models. However, both types of AI systems have the potential to improve upon human diagnostic accuracy, particularly in complex cases.

Another study, published in *JMIR Medical Informatics* in 2023, evaluated the accuracy of differential diagnosis lists generated by ChatGPT-3.5 and ChatGPT-4 for complex clinical vignettes. The study found that ChatGPT-4's performance was comparable to that of physicians, with the correct diagnosis appearing in the top 10 differential diagnoses in 83% of cases [3]. These findings underscore the potential of LLMs as a supplementary tool for physicians, particularly in the context of general internal medicine.

Challenges and Future Directions

Despite the promising potential of AI in differential diagnosis, there are several challenges that need to be addressed. These include the need for large, high-quality datasets to train AI models, the risk of bias in AI algorithms, and the importance of ensuring that AI tools are integrated into clinical workflows in a way that supports, rather than hinders, the work of clinicians. The "black box" nature of some AI models can also be a barrier to their adoption in clinical practice, as clinicians may be hesitant to trust the recommendations of a system whose reasoning they cannot understand. Future research will need to focus on addressing these challenges, developing more transparent and explainable AI models, and conducting rigorous real-world evaluations of their impact on diagnostic accuracy and patient outcomes.

Conclusion

AI is poised to revolutionize the process of differential diagnosis generation. By leveraging the power of machine learning and large language models, AI-powered tools can assist clinicians in making more accurate and timely diagnoses. While there are still challenges to be overcome, the continued development and refinement of these technologies hold great promise for the future of healthcare. As AI becomes more integrated into clinical practice, it has the potential to not only improve diagnostic accuracy but also to free up clinicians' time, allowing them to focus on the human aspects of patient care.

References

[1] McDuff, D., Schaekermann, M., Tu, T. et al. Towards accurate differential diagnosis with large language models. Nature 642, 451–457 (2025). https://doi.org/10.1038/s41586-025-08869-4

[2] Feldman, M. J., Hoffer, E. P., & Barnett, G. O. (2025). Dedicated AI Expert System vs Generative AI With Large Language Model for Clinical Diagnoses. JAMA Network Open, 8(5), e2512994. https://doi.org/10.1001/jamanetworkopen.2025.12994

[3] Hirosawa, T., Kawamura, R., Harada, Y., Mizuta, K., Tokumasu, K., Kaji, Y., Suzuki, T., & Shimizu, T. (2023). ChatGPT-Generated Differential Diagnosis Lists for Complex Case–Derived Clinical Vignettes: Diagnostic Accuracy Evaluation. JMIR Medical Informatics, 11, e48808. https://doi.org/10.2196/48808