

Performance evaluation of ChatGPT in detecting diagnostic errors and their contributing factors: an analysis of 545 case reports of diagnostic errors.

June 19, 2024

Harada Y, Suzuki T, Harada T, et al. Performance evaluation of ChatGPT in detecting diagnostic errors and their contributing factors: an analysis of 545 case reports of diagnostic errors. *BMJ Open Qual.* 2024;13(2):e002654. doi:10.1136/bmjog-2023-002654.

<https://psnet.ahrq.gov/issue/performance-evaluation-chatgpt-detecting-diagnostic-errors-and-their-contributing-factors>

ChatGPT and other large language models are being studied for their ability to provide [correct diagnoses](#) in numerous clinical areas. This study sought to evaluate ChatGPT's [ability](#) to identify diagnostic errors in case studies. ChatGPT was able to identify diagnostic errors in 95% of the 545 case reports. The results suggest ChatGPT could be used to screen for possible diagnostic errors in electronic health records, but significant research and testing is still [required](#).