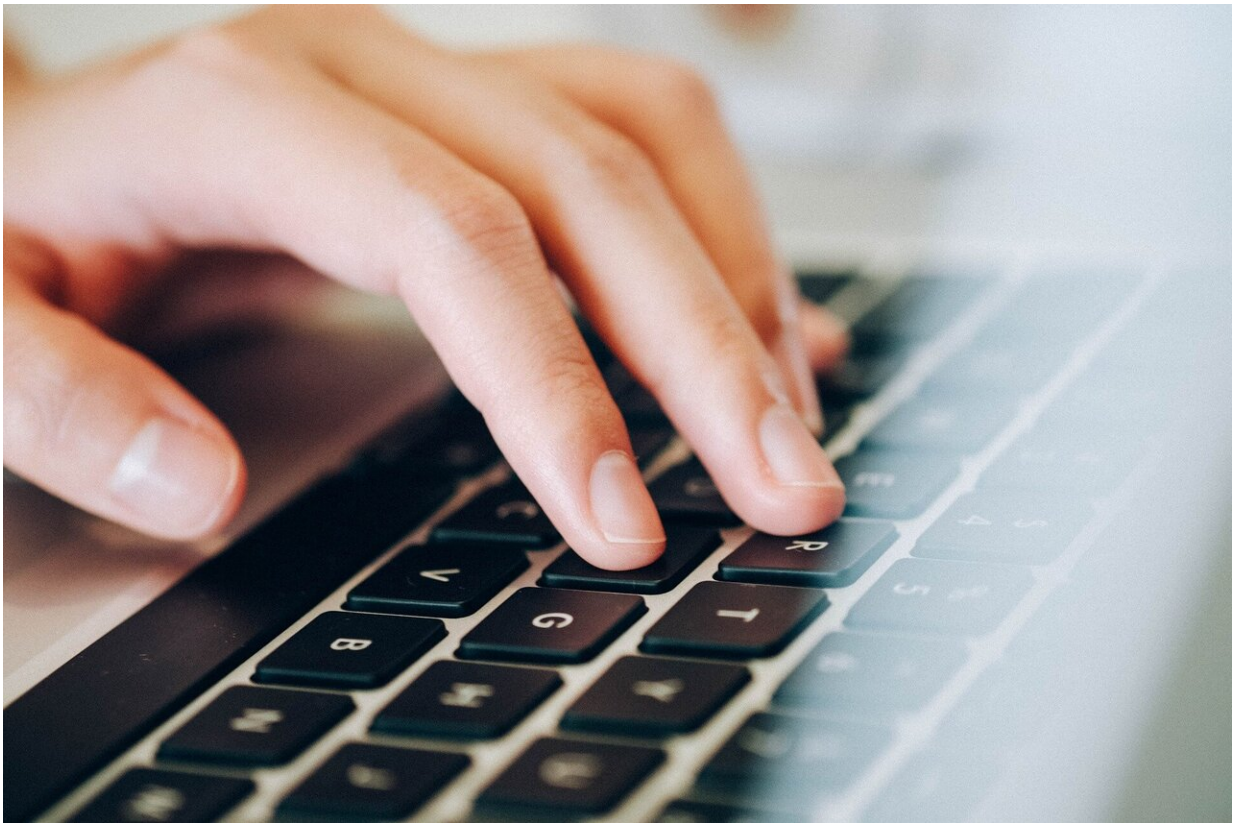


Study shows AI tool successfully responds to patient questions in electronic health record

July 16 2024



Credit: CC0 Public Domain

As part of a nationwide trend, many more of NYU Langone Health's patients during the pandemic started using electronic health record tools to ask their doctors questions, refill prescriptions, and review test results.

Many patients' digital inquiries arrived via a communications tool called In Basket, which is built into NYU Langone's electronic health record (EHR) system, EPIC.

While physicians have always dedicated time to managing EHR messages, they saw a more than 30% [annual increase](#) in recent years in the number of messages received daily, according [an article](#) by Paul A. Testa, MD chief medical information officer at NYU Langone.

Testa wrote that it is not uncommon for physicians to receive more than 150 In Basket messages per day. With [health systems](#) not designed to handle this kind of traffic, physicians ended up filling the gap, spending long hours after work sifting through messages. This burden is cited as a reason that [half of physicians report burnout](#).

Now a new study, led by researchers at NYU Grossman School of Medicine, shows that an AI tool can draft responses to patients' EHR queries as accurately as their human health care professionals, and with greater perceived "empathy." The findings highlight these tools' potential to dramatically reduce physicians' In Basket burden while improving their communication with patients, as long as human providers review AI drafts before they are sent.

NYU Langone Health has been testing the capabilities of generative artificial intelligence (genAI), in which computer algorithms develop likely options for the next word in any sentence based on how people have used words in context on the internet.

A result of this next-word prediction is that genAI "chatbots" can reply to questions in convincing human-like language. NYU Langone in 2023 licensed "a private instance" of GPT4, the latest relative of the famous ChatGPT chatbot, which lets physicians experiment using real patient data while still adhering to data privacy rules.

Published online July 16 in *JAMA Network Open*, the new study examined GPT4-generated drafts to patient In Basket queries, and had [primary care physicians](#) compare them to the actual human responses to those messages.

"Our results suggest that chatbots could reduce the workload of care providers by enabling efficient and empathetic responses to patients' concerns," said lead study author William Small, MD, a clinical assistant professor in Department of Medicine at NYU Grossman School of Medicine. "We found that EHR-integrated AI chatbots that use patient-specific data can draft messages similar in quality to human providers."

For the study, 16 primary care physicians rated 344 randomly assigned pairs of AI and human responses to patient messages on accuracy, relevance, completeness, and tone, and indicated if they would use the AI response as a first draft, or have to start from scratch in writing the patient message. The physicians did not know whether the responses they were reviewing were generated by humans or the AI tool (blinded study).

The research team found that the accuracy, completeness, and relevance of generative AI and human providers responses did not differ statistically. Generative AI responses outperformed human providers in terms of understandability and tone by 9.5%.

Further, the AI responses were more than twice as likely (125% more likely) to be considered empathetic and 62% more likely to use language that conveyed positivity (potentially related to hopefulness) and affiliation ("we are in this together").

On the other hand, AI responses were also 38% longer and 31% more likely to use complex language, so further training of the tool is needed, the researchers say. While humans responded to patient queries at a 6th

grade level, AI was writing at an 8th grade level, according to a standard measure of readability called the Flesch Kincaid score.

The researchers argued that use of private patient information by chatbots, rather than general internet information, better approximates how this technology would be used in the real world. Future studies will be needed to confirm whether private data specifically improved AI tool performance.

"This work demonstrates that the AI tool can build high-quality draft responses to patient requests," said corresponding author Devin Mann, MD, senior director of Informatics Innovation in NYU Langone Medical Center Information Technology (MCIT).

"With this physician approval in place, GenAI message quality will be equal in the near future in quality, communication style, and usability, to responses generated by humans," added Mann, also a professor in the Departments of Population Health and Medicine.

More information: Large Language Model–Based Responses to Patients' In-Basket Messages, *JAMA Network Open* (2024). [DOI: 10.1001/jamanetworkopen.2024.22399](https://doi.org/10.1001/jamanetworkopen.2024.22399)

Provided by NYU Langone Health

Citation: Study shows AI tool successfully responds to patient questions in electronic health record (2024, July 16) retrieved 16 July 2024 from <https://medicalxpress.com/news/2024-07-ai-tool-successfully-patient-electronic.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.