

# AI medical assistants? ChatGPT answers patient questions with more empathy than human doctors



There has been widespread speculation about how advances in artificial intelligence (AI) assistants like ChatGPT could be used in medicine.

A new study published in [JAMA Internal Medicine](#) led by [John W. Ayers](#), Ph.D., from [the Qualcomm Institute at University of California San Diego](#) provides an early glimpse into the role that AI assistants could play in medicine. The study compared written responses from physicians and those from ChatGPT to real-world health questions. A panel of licensed healthcare professionals preferred ChatGPT's responses 79% of the time and rated ChatGPT's responses as higher quality and more empathetic.

"The opportunities for improving healthcare with AI are massive," said Ayers, who is also vice chief of innovation in the UC San Diego School of Medicine Division of Infectious Disease and Global Public Health. "AI-augmented care is the future of medicine."

Follow the latest news and policy debates on sustainable agriculture, biomedicine, and other 'disruptive' innovations. Subscribe to our newsletter.

[SIGN UP](#)

## Is ChatGPT ready for healthcare?

In the new study, the research team set out to answer the question: Can ChatGPT respond accurately to questions patients send to their doctors? If yes, AI models could be integrated into health systems to improve physician responses to questions sent by patients and ease the ever-increasing burden on physicians.

"ChatGPT might be able to pass a medical [licensing exam](#)," said study co-author Davey Smith, M.D., M.A.S., a physician-scientist, co-director of the UC San Diego Altman Clinical and Translational Research Institute and professor at the UC San Diego School of Medicine, "but directly answering patient questions accurately and empathetically is a different ballgame."

"The COVID-19 pandemic accelerated virtual healthcare adoption," added study co-author Eric Leas, Ph.D., M.P.H., a Qualcomm Institute affiliate and assistant professor in the UC San Diego Herbert Wertheim School of Public Health and Human Longevity Science. "While this made accessing care easier for patients, physicians are burdened by a barrage of electronic patient messages seeking medical advice that have contributed to [record-breaking](#) levels of physician burnout."

## Designing a study to test ChatGPT in a healthcare setting

To obtain a large and diverse sample of healthcare questions and physician answers that did not contain identifiable personal information, the team turned to social media where millions of patients publicly post

medical questions to which doctors respond: [Reddit's AskDocs](#).

r/AskDocs is a subreddit with approximately 452,000 members who post medical questions and verified healthcare professionals submit answers. While anyone can respond to a question, moderators verify healthcare professionals' credentials and responses display the respondent's level of credentials. The result is a large and diverse set of patient medical questions and accompanying answers from licensed medical professionals.

While some may wonder if question-answer exchanges on social media are a fair test, team members noted that the exchanges were reflective of their clinical experience.

The team randomly sampled 195 exchanges from AskDocs where a verified physician responded to a public question. The team provided the original question to ChatGPT and asked it to author a response. A panel of three licensed healthcare professionals assessed each question and the corresponding responses and were blinded to whether the response originated from a physician or ChatGPT. They compared responses based on information quality and empathy, noting which one they preferred.

The panel of healthcare professional evaluators preferred ChatGPT responses to physician responses 79% of the time.

"ChatGPT messages responded with nuanced and accurate information that often addressed more aspects of the patient's questions than physician responses," said Jessica Kelley, M.S.N, a nurse practitioner with San Diego firm Human Longevity and study co-author.

Additionally, ChatGPT responses were rated significantly higher in quality than physician responses: good or very good quality responses were 3.6 times higher for ChatGPT than physicians (physicians 22.1% versus ChatGPT 78.5%). The responses were also more empathic: empathetic or very empathetic responses were 9.8 times higher for ChatGPT than for physicians (physicians 4.6% versus ChatGPT 45.1%).

"I never imagined saying this," added Aaron Goodman, M.D., an associate clinical professor at UC San Diego School of Medicine and study coauthor, "but ChatGPT is a prescription I'd like to give to my inbox. The tool will transform the way I support my patients."

## **Harnessing AI assistants for patient messages**

"While our study pitted ChatGPT against physicians, the ultimate solution isn't throwing your doctor out altogether," said Adam Poliak, Ph.D., an assistant professor of Computer Science at Bryn Mawr College and study co-author. "Instead, a physician harnessing ChatGPT is the answer for better and empathetic care."

"Our study is among the first to show how AI assistants can potentially solve real world healthcare delivery problems," said Christopher Longhurst, M.D., M.S., Chief Medical Officer and Chief Digital Officer at UC San Diego Health. "These results suggest that tools like ChatGPT can efficiently draft high quality,

personalized medical advice for review by clinicians, and we are beginning that [process](#) at UCSD Health.”

Mike Hogarth, M.D., a physician-bioinformatician, co-director of the Altman Clinical and Translational Research Institute at UC San Diego, professor in the UC San Diego School of Medicine and study co-author, added, “It is important that integrating AI assistants into healthcare messaging be done in the context of a randomized controlled trial to judge how the use of AI assistants impact outcomes for both physicians and patients.”

In addition to improving workflow, investments into AI assistant messaging could impact patient health and physician performance.

Mark Dredze, Ph.D., the John C Malone Associate Professor of Computer Science at Johns Hopkins and study co-author, noted: “We could use these technologies to train doctors in patient-centered communication, eliminate health disparities suffered by minority populations who often seek healthcare via messaging, build new medical safety systems, and assist doctors by delivering higher quality and more efficient care.”

[Read the original post here](#)