

‘Dr. Google’ Meets Its Match: Dr. ChatGPT

Researchers are cautious about recommending that patients put their full trust in the current state of AI.

September 27, 2023 By Andrew Leonard and KFF Health News

As a fourth-year ophthalmology resident at Emory University School of Medicine, Riley Lyons’ biggest responsibilities include triage: When a patient comes in with an eye-related complaint, Lyons must make an immediate assessment of its urgency.

He often finds patients have already turned to “Dr. Google.” Online, Lyons said, they are likely to find that “any number of terrible things could be going on based on the symptoms that they’re experiencing.”

So, when two of Lyons’ fellow ophthalmologists at Emory came to him and suggested evaluating the accuracy of the AI chatbot ChatGPT in diagnosing eye-related complaints, he jumped at the chance.

In June, Lyons and his colleagues reported in medRxiv, an online publisher of health science preprints, that ChatGPT [compared quite well](#) to human doctors who reviewed the same symptoms — and performed vastly better than the symptom checker on the popular health website WebMD. And despite the much-publicized “hallucination” problem known to afflict ChatGPT — its habit of occasionally making outright false statements — the Emory study reported that the most recent version of ChatGPT made zero “grossly inaccurate” statements when presented with a standard set of eye complaints.

The relative proficiency of ChatGPT, which debuted in November 2022, was a surprise to Lyons and his co-authors. The artificial intelligence engine “is definitely an improvement over just putting something into a Google search bar and seeing what you find,” said co-author Nieraj Jain, an assistant professor at the Emory Eye Center who specializes in vitreoretinal surgery and disease.

But the findings underscore a challenge facing the health care industry as it assesses the promise and pitfalls of generative AI, the type of artificial intelligence used by ChatGPT: The accuracy of chatbot-delivered medical information may represent an improvement over Dr. Google, but there are still many questions about how to integrate this new technology into health care systems with the same safeguards historically applied to the introduction of new drugs or medical devices.

The smooth syntax, authoritative tone, and dexterity of generative AI have drawn extraordinary attention from all sectors of society, with some comparing its future impact to that of the internet itself. In health care, companies are working feverishly to implement generative AI in areas such as radiology and medical records.

When it comes to consumer chatbots, though, there is still caution, even though the technology is already widely available — and better than many alternatives. Many doctors believe AI-based medical tools should undergo an approval process similar to the FDA's regime for drugs, but that would be years away. It's unclear how such a regime might apply to general-purpose AIs like ChatGPT.

"There's no question we have issues with access to care, and whether or not it is a good idea to deploy ChatGPT to cover the holes or fill the gaps in access, it's going to happen and it's happening already," said Jain. "People have already discovered its utility. So, we need to understand the potential advantages and the pitfalls."

The Emory study is not alone in ratifying the relative accuracy of the new generation of AI chatbots. A [report published in Nature](#) in early July by a group led by Google computer scientists said answers generated by Med-PaLM, an AI chatbot the company built specifically for medical use, "compare favorably with answers given by clinicians."

AI may also have better bedside manner. Another study, published in April by researchers from the University of California-San Diego and other institutions, even noted that health care professionals rated ChatGPT answers as [more empathetic](#) than responses from human doctors.

Indeed, a number of companies are exploring how chatbots could be used for mental health therapy, and some investors in the companies are betting that healthy people might also enjoy chatting and even bonding with an AI "friend." The company behind Replika, one of the most advanced of that genre, markets its chatbot as, "The AI companion who cares. Always here to listen and talk. Always on your side."

"We need physicians to start realizing that these new tools are here to stay and they're offering new capabilities both to physicians and patients," said James Benoit, an AI consultant. While a postdoctoral fellow in nursing at the University of Alberta in Canada, he published a study in February reporting that ChatGPT [significantly outperformed](#) online symptom checkers in evaluating a set of medical scenarios. "They are accurate enough at this point to start meriting some consideration," he said.

Still, even the researchers who have demonstrated ChatGPT's relative reliability are cautious about recommending that patients put their full trust in the current state of AI. For many medical professionals, AI chatbots are an invitation to trouble: They cite a host of issues relating to privacy, safety, bias, liability, transparency, and the current absence of regulatory oversight.

The proposition that AI should be embraced because it represents a marginal improvement over Dr. Google is unconvincing, these critics say.

“That’s a little bit of a disappointing bar to set, isn’t it?” said Mason Marks, a professor and MD who specializes in health law at Florida State University. He recently wrote an [opinion piece](#) on AI chatbots and privacy in the Journal of the American Medical Association. “I don’t know how helpful it is to say, ‘Well, let’s just throw this conversational AI on as a band-aid to make up for these deeper systemic issues,’” he said to KFF Health News.

The biggest danger, in his view, is the likelihood that market incentives will result in AI interfaces designed to steer patients to particular drugs or medical services. “Companies might want to push a particular product over another,” said Marks. “The potential for exploitation of people and the commercialization of data is unprecedented.”

OpenAI, the company that developed ChatGPT, also urged caution.

“OpenAI’s models are not fine-tuned to provide medical information,” a company spokesperson said. “You should never use our models to provide diagnostic or treatment services for serious medical conditions.”

John Ayers, a computational epidemiologist who was the lead author of the UCSD study, said that as with other medical interventions, the focus should be on patient outcomes.

“If regulators came out and said that if you want to provide patient services using a chatbot, you have to demonstrate that chatbots improve patient outcomes, then randomized controlled trials would be registered tomorrow for a host of outcomes,” Ayers said.

He would like to see a more urgent stance from regulators.

“One hundred million people have ChatGPT on their phone,” said Ayers, “and are asking questions right now. People are going to use chatbots with or without us.”

At present, though, there are few signs that rigorous testing of AIs for safety and effectiveness is imminent. [In May](#), Robert Califf, the commissioner of the FDA, described “the regulation of large language models as critical to our future,” but aside from recommending that regulators be “nimble” in their approach, he offered few details.

In the meantime, the race is on. In July, The Wall Street Journal reported that the Mayo Clinic was [partnering with Google](#) to integrate the Med-PaLM 2 chatbot into its system. In June, [WebMD announced](#) it was partnering with a Pasadena, California-based startup, HIA Technologies Inc., to provide interactive “digital health assistants.” And the ongoing integration of AI into both Microsoft’s Bing and Google Search suggests that Dr. Google is already well on its way to being replaced by Dr. Chatbot.

This article was produced by [KFF Health News](#), which publishes [California Healthline](#), an editorially independent service of the [California Health Care Foundation](#).

[This story](#) was published by KFF Health News on September 12, 2023. It is republished with

permission.