

LAY ABSTRACT

TITLE: Accuracy and Reliability of Chatbot Responses to Physician Questions

JOURNAL: JAMA Network Open. 2023. 6:10

AUTHORS: Rachel S Goodman¹, J Randall Patrinely², Cosby A Stone Jr³, Eli Zimmerman⁴, Rebecca R Donald⁵, Sam S Chang⁶, Sean T Berkowitz⁷, Avni P Finn⁷, Eiman Jahangir⁸, Elizabeth A Scoville⁹, Tyler S Reese¹⁰, Debra L Friedman¹¹, Julie A Bastarache³, Yuri F van der Heijden¹², Jordan J Wright¹³, Fei Ye¹⁴, Nicholas Carter¹⁵, Matthew R Alexander¹⁶, Jennifer H Choe¹⁷, Cody A Chastain¹², John A Zic², Sara N Horst⁹, Isik Turker¹⁸, Rajiv Agarwal¹⁷, Evan Osmundson¹⁹, Kamran Idrees²⁰, Colleen M Kiernan²⁰, Chandrasekhar Padmanabhan²⁰, Christina E Bailey²⁰, Cameron E Schlegel²⁰, Lola B Chambless²¹, Michael K Gibson¹⁷, Travis J Osterman²², Lee E Wheless^{2,*}, Douglas B Johnson^{17,*}

* Corresponding authors

INSTITUTIONS:

1 Vanderbilt University School of Medicine, Nashville, Tennessee.

2 Department of Dermatology, Vanderbilt University Medical Center, Nashville, Tennessee.

3 Department of Allergy, Pulmonology, and Critical Care, Vanderbilt University Medical Center, Nashville, Tennessee.

4 Department of Neurology, Vanderbilt University Medical Center, Nashville, Tennessee.

5 Department of Anesthesiology, Vanderbilt University Medical Center, Nashville, Tennessee.

6 Department of Urology, Vanderbilt University Medical Center, Nashville, Tennessee.

7 Vanderbilt Eye Institute, Department of Ophthalmology, Vanderbilt University Medical, Nashville, Tennessee.

8 Department of Cardiovascular Medicine, Vanderbilt University Medical Center, Nashville, Tennessee.

9 Department of Gastroenterology, Hepatology, and Nutrition, Vanderbilt University Medical Center, Nashville, Tennessee.

10 Department of Rheumatology and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee.

11 Department of Pediatric Hematology/Oncology, Vanderbilt University Medical Center, Nashville, Tennessee.

12 Department of Infectious Disease, Vanderbilt University Medical Center, Nashville, Tennessee.

13 Department of Diabetes, Endocrinology, and Metabolism, Vanderbilt University Medical Center, Nashville, Tennessee.

14 Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee.

15 Division of Trauma and Surgical Critical Care, University of Miami Miller School of Medicine, Miami, Florida.

16 Department of Cardiovascular Medicine and Clinical Pharmacology, Vanderbilt University Medical Center, Nashville, Tennessee.

17 Department of Hematology/Oncology, Vanderbilt University Medical Center, Nashville, Tennessee.

18 Department of Cardiology, Washington University School of Medicine in St Louis, St Louis, Missouri.

19 Department of Radiation Oncology, Vanderbilt University Medical Center, Nashville, Tennessee.

20 Department of Surgical Oncology & Endocrine Surgery, Vanderbilt University Medical Center, Nashville, Tennessee.

21 Department of Neurological Surgery, Vanderbilt University Medical Center, Nashville, Tennessee.

22 Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee.

This is attributed to the CEECR grant: UG3 CA265846

LAY ABSTRACT

ChatGPT is an online robot that can answer medical questions. This paper looked at how accurate and complete ChatGPT is. In 2023, 33 doctors from 17 types of clinics came up with 284 medical questions. The questions were easy, medium, or hard. Answers to the questions could have been “yes or no” or longer responses.

The doctors graded the accuracy of the answers ChatGPT gave to the questions on a 1-6 scale (1 being wrong and 6 being right). Doctors also graded how complete answers were on a 1-3 scale (1 being incomplete and 3 being better than complete).

For all the questions, the average score for accuracy was 4.8. The average score for completeness was 2.5. ChatGPT did worse as the questions got harder. Accuracy scores for the “yes or no” and longer answers were similar. ChatGPT was asked to answer the questions a second time that it scored the worse on. These answers scored better when the questions were asked a second time.

Overall, ChatGPT gives better answers to medical questions over time. Most answers were fairly accurate. More studies can better train the ChatGPT to be even more accurate.