Comparing Physician and ChatGPT Responses to Patient Questions
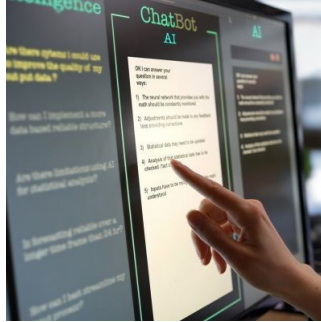

In a recent cross-sectional study, published in JAMA Internal Medicine, researchers evaluated the ability of ChatGPT to provide responses to patient questions that are of comparable quality and empathy to those produced by doctors.

The team drew 195 answers to patient questions from October 2022 over a publicly accessible social media forum. Chatbot responses were generated by entering the original question into a fresh session.

The original question along with anonymised and randomly ordered physician and chatbot responses were evaluated by a team of licensed health care professionals.

Evaluators judged the answers based on the quality of information provided and the empathy or bedside manner. They compared the mean outcomes of ChatGPT and physician on a 1 to 5 scale.

Of the 195 questions and responses, evaluators preferred the chatbot responses in 78.6% (95% CI, 75.0%-81.8%) of the 585 evaluations.

Strikingly, mean physician responses were significantly shorter than chatbot responses, and chatbot responses were rated significantly higher for quality than physician responses.

Responses rated ≥4 indicating 'good' or 'very good' quality was approx. 3.6 times higher for chatbots than physicians (chatbot: 78.5% vs. physicians: 22.1%).

Chatbot responses were also rated significantly more empathetic than physician responses. The proportion of responses rated ≥4 indicated they were *empathetic* or *very empathetic*, and this was approx. 9.8 times higher in chatbot responses than for physicians (chatbot: 45.1% vs. physicians: 4.6%).

Although some patient queries require more skills and time to answer, many patients only enquire about appointments and test results. Should AI assistant be tested in this area, and if successful, they could help to relieve clinicians from the extra burden of answering patient messages.

Overall, this cross-sectional study shows that a chatbot can generate quality and empathetic responses in an online forum. Further evidence is needed to see if using AI assistants can improve responses, lower clinician burnout, and improve patient outcomes.

Source: JAMA Intern Med

Image Credit: iStock

Published on : Mon, 15 May 2023