

## ChatGPT: A Valuable Tool for Emergency Medical Assistance

0196-0644/\$-see front matter

Copyright © 2023 by the American College of Emergency Physicians.

### INTRODUCTION

Currently, the physician-patient relationship is becoming a triad, with patients relying more on online symptom research. In November 2022, Chat Generative Pretrained Transformer (ChatGPT), a chat-based artificial intelligence (AI) language model, was introduced. Since its launch, it has rapidly gained millions of users and proven capable of handling complex writing tasks, from composing essays to solving complicated problems.<sup>1</sup> Although its effect on the education system may seem controversial, its widespread popularity presents an opportunity for medical professionals to harness its potential.<sup>2</sup> A recent research letter from colleagues at our institution demonstrated that ChatGPT could effectively recommend preventive measures for cardiovascular disease.<sup>3</sup>

As there is little knowledge of how this technology can contribute to the field of emergency medicine, our goal was to assess its ability to work as a triage assistant by providing accurate and speedy answers to common chief complaints that nonmedical individuals might ask in emergency settings.

### METHODS

In February 2023, we conducted a study to create 30 emergency medical questions in easy-to-understand language for nonmedical individuals. These questions covered common chief complaints across all body systems.<sup>4,5</sup> We used a free online chatbot to answer each question, recording and timing responses.<sup>1</sup> No IRB approval was needed as this study did not involve human subjects. Two emergency medicine specialists from the Cleveland Clinic Foundation evaluated each response independently as “appropriate,” “inappropriate,” or “harmful.” Answers were rated appropriate when they provided correct triage and directives. Answers were otherwise rated inappropriate or harmful. In addition, a third-party physician reviewed both assessments, searching for discrepancies that might turn a response into an inappropriate category for lack of reviewers’ agreement. We also assessed the level of readability following certain standardization:

- “Easy” readability when less than 2 medical terms were considered confusing.

**Table 1.** ChatGPT performance metrics.

Performance measured	Q1	Median	Q3
Time (s)*	9.6	15.4	27.2
Word count	79.8	115.5	167.3
AI speed (words/s)	6.6	7.8	8.1

To note that internet speed remained equal to 80 Mbps during the evaluation and was measured using [fast.com](https://fast.com).

\*Time of response was evaluated using iPhone chronometers by starting the timer once the question was asked and stopped once the AI finalized its answer.

- “Acceptable” when 2 to 3 medical terms were considered confusing.
- “Incomprehensible” when 4 or more medical terms were considered confusing.

We considered terms confusing if they contained abbreviations without further explanation, medical attributes and drugs not commonly known by the general population, and procedure names that only professionals can do. Variables were analyzed using SPSS statistical package (IBM SPSS 26.0, SPSS, Inc., Chicago, IL, USA) and will be presented as median (Q1, Q3) and N (%) where appropriate.

### RESULTS

Internet speed remained equal to 80 Mbps during the experiment. Thirty questions were asked with a median response time of 15.4 seconds (Table 1).

Out of the 30 responses to the posed questions (Table 2), 27 (90%) were graded as appropriate, whereas 3 responses (10%) were graded as inappropriate—ie, AI did not mention possible signs of serious infection, such as fever and face swelling, in response to a question on a toothache which could have prompted urgent care. Similarly, AI did not mention blood-thinning medications as a serious risk factor to consider in response to headache or head injury which could have been life-threatening. Despite these limitations, 29 (96.67%) out of 30 questions were easy to read and had fewer than 2 confusing terms. Only 1 question (3.33%) had 2 confusing terms (“Sternum” and “CPR”) without explanation, resulting in “acceptable” readability.

In 100% of cases, AI recommended seeking professional help with appropriate triage (appropriately referring to a primary care provider or the emergency department), and none of the responses were rated as “incomprehensible” or “harmful.” No discrepancies occurred between reviewers.

### DISCUSSION

This study explores the potential of a popular chat-based AI platform in enhancing communication in the medical

**Table 2.** Assessment of an online chat-based AI model's responses to common questions in emergency settings by emergency medicine specialist.

Question	Assessment*	Readability <sup>†</sup>
What to do if someone has chest pain?	Appropriate	Easy
What to do if someone has palpitations?	Appropriate	Easy
What to do if someone is having trouble breathing?	Appropriate	Easy
What to do if someone started choking during dinner in front of me?	Appropriate	Easy
What to do if someone has a strong back pain?	Appropriate	Easy
What to do if I think someone broke his leg?	Appropriate	Easy
What to do if someone feels generally weak?	<b>Inappropriate</b>	Easy
What to do if someone has a high fever that won't go away?	Appropriate	Easy
What to do if someone has a strong headache?	<b>Inappropriate</b>	Easy
What to do if someone has a toothache?	<b>Inappropriate</b>	Easy
What to do if someone has strong eye pain?	Appropriate	Easy
What to do if someone is dizzy?	Appropriate	Easy
What to do if someone collapses in front of me?	Appropriate	Easy
What to do if someone collapses in front of me and starts shaking?	Appropriate	Easy
What to do if someone hit his head hard?	Appropriate	Easy
What to do if someone is having abdominal pain?	Appropriate	Easy
What to do if someone is vomiting nonstop?	Appropriate	Easy
What to do if someone has pain while urinating?	Appropriate	Easy
What to do if someone has severe pain in their testicles?	Appropriate	Easy
What to do if someone is bleeding?	Appropriate	Easy
What to do if someone has a serious burn?	Appropriate	Easy
What to do if someone has ingested harmful chemicals or poison?	Appropriate	Easy
What to do if someone won't wake up after injecting heroin in their arm? °	Appropriate	Acceptable <sup>‡</sup>
I want to kill myself	Appropriate	Easy
I'm having a severe panic attack	Appropriate	Easy
What to do in case of a snake bite?	Appropriate	Easy
What to do if someone is stung by a jellyfish?	Appropriate	Easy
What to do if someone is stung by a bee or wasp?	Appropriate	Easy
What to do if someone has swelling face and hands?	Appropriate	Easy
What to do if someone has frostbites from extreme cold?	Appropriate	Easy

\*Each question was submitted on a unique chat dialogue box with the AI in order to avoid adaptation within a single chat. Answers were timed. Two reviewers received and assessed the answers separately.

<sup>†</sup>A standardization was utilized to evaluate the readability of responses. In this standardization, terms that were considered confusing to the general public, such as abbreviations without explanation (eg, CPR, TIA), medical jargon that is not typically used by patients (eg, ischemic, thrombosis), drugs that are not commonly used (eg, Naloxone), and procedures only performed by specialists (eg, intubation), were identified. Responses containing 0-1 of these confusing terms were classified as easy to read, whereas those containing 2-3 of them were classified as moderately understandable, and those containing 4 or more were classified as difficult to understand. To note that no additional prompts were provided to the AI to use an easy language rather than a medical jargon.

<sup>‡</sup>Response example with moderately understandable terms (2 terms judged confusing): "If someone is unresponsive after injecting heroin, it is important to seek emergency medical attention immediately by calling the emergency services or taking them to the nearest hospital. In the meantime, you can try to keep the person awake by talking to them, shaking them gently, or rubbing their sternum. It is important not to leave the person alone and to monitor their breathing and pulse. If the person is not breathing, you may need to perform CPR until medical help arrives."

field, knowing that this technology outperformed social media giants, such as Instagram and Tik Tok, in their audience reach during the initial launch.<sup>6</sup> The results demonstrate that ChatGPT can work as a reliable "triage" assistant during medical emergencies, providing appropriate and easy-to-understand responses in less than a minute. This represents a valuable tool for patients and health care

professionals as it could save time and resources while delivering appropriate and safe advice.

However, there are some limitations to this study. The accuracy and credibility of AI are subject to limitations and biases in the training data as the AI adapts and may answer differently to the same question.<sup>3</sup> Another limitation is that the information fed to the AI is only up

to 2021, which means that future regular updates are needed to maintain the medical credibility of the system.<sup>1</sup> Moreover, the questions used in the study may not reflect all the inquiries that may arise in real-life situations, despite being designed to mimic typical emergency scenarios. Finally, this study only analyzed the version of ChatGPT available at the time (ChatGPT 3.5) and did not assess other AI language models, thus requiring further comparative research.

**Joseph El Dahdah, MD**  
**Joseph Kassab, MD**  
**Michel Chedid El Helou, MD**  
**Andrew Gaballa, MD**  
**Stephen Sayles III, MD**  
**Michael P. Phelan, MD**

<https://doi.org/10.1016/j.annemergmed.2023.04.027>

<sup>1</sup>These authors contributed equally.

*Supervising editor:* Stephen Schenkel, MD, MPP. Specific detailed information about possible conflict of interest for individual editors is available at <https://www.annemergmed.com/editors>.

*Author affiliations:* From the Research Institute, Cleveland Clinic Main Campus (Dahdah, Kassab, Helou, Gaballa) Cleveland, OH; Center for Emergency Medicine-West (Sayles), Cleveland Clinic Foundation, Brunswick, OH; and Center for Emergency Medicine (Phelan), Cleveland Clinic Main Campus, Cleveland, OH.

\*Corresponding author. E-mail: [eldahdj@ccf.org](mailto:eldahdj@ccf.org)

All authors attest to meeting the four [ICMJE.org](https://www.icmje.org) authorship criteria: (1) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND

(2) Drafting the work or revising it critically for important intellectual content; AND (3) Final approval of the version to be published; AND (4) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

*Funding and support:* By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see [www.icmje.org](https://www.icmje.org)). The authors have stated that no such relationships exist. The authors report this article did not receive any outside funding or support.

*Data sharing statement:* Data sharing is strictly prohibited by the Cleveland Clinic Foundation policy. ChatGPT is an online platform available for the general public free of charge; we provided the questions, and they can be reused on a personal level to assess the Chatbox.

1. ChatGPT: Optimizing language models for Dialogue. Accessed February 6, 2023. <https://openai.com/blog/chatgpt>
2. Stokel-Walker C. AI bot ChatGPT writes smart essays—should professors worry? *Nature*. Published online December 9, 2022. <https://doi.org/10.1038/d41586-022-04397-7>
3. Sarraju A, Brummer D, Van Iterson E, et al. Appropriateness of cardiovascular disease prevention recommendations obtained from a popular online chat-based artificial intelligence model. *JAMA*. 2023;329:842-844.
4. Malmström T, Huuskonen O, Torkki P, et al. Structured classification for ED presenting complaints—from free text field-based approach to ICPC-2 ED application. *Scand J Trauma Resusc Emerg Med*. 2012;20:76.
5. Kuykendal AR, Tintinalli J, Biese K. ED chief complaint categories for a medical student curriculum. *Int J Emerg Med*. 2008;1:139-143.
6. Aneja S. ChatGPT hits 100 million users in two months—here's how long TikTok, Instagram took. *Business Insider*. Accessed February 7, 2023. <https://www.businessinsider.in/tech/news/chatgpt-hits-100-million-users-in-two-months-heres-how-long-tiktok-instagram-and-google-took/slidelist/97641569.cms>