

How good is ChatGPT at diagnosing disease?

May 1 2023, by Stephen Hughes



Credit: AI-generated image ([disclaimer](#))

For years, many have feared that artificial intelligence (AI) will take over national security mechanisms, leading to human slavery, domination of human society and perhaps the [annihilation of humans](#). One way of killing humans is medical misdiagnosis, so it seems reasonable to examine the performance of ChatGPT, the AI chatbot that is taking the world by storm. This is timely in light of ChatGPT's recent remarkable performance in passing the US [medical licensing exam](#).

Computer-aided diagnosis has been attempted many times over the years, particularly for [diagnosing appendicitis](#). But the emergence of AI that draws on the entire internet for answers to questions rather than being confined to fixed databases open new avenues of potential for augmenting [medical diagnosis](#).

More recently, several articles discuss the performance of ChatGPT in making medical diagnoses. An [American emergency medicine physician](#) recently gave an account of how he asked ChatGPT to give the possible diagnoses of a young woman with lower abdominal pain. The machine gave numerous credible diagnoses, such as appendicitis and ovarian cyst problems, but it missed [ectopic pregnancy](#).

This was correctly identified by the physician as a serious omission, and I agree. On my watch, ChatGPT would not have passed its medical final examinations with that rather deadly performance.

ChatGPT learns

I'm pleased to say that when I asked ChatGPT the same question about a young woman with lower abdominal pain, ChatGPT confidently stated ectopic pregnancy in the [differential diagnosis](#). This reminds us of an important thing about AI: it is capable of learning.

Presumably, someone has told ChatGPT of its error and it has learned from this new data—not unlike a [medical student](#). It is this ability to learn that will improve the performance of AIs and make them stand out from rather more constrained computer-aided diagnosis algorithms.

ChatGPT prefers technical language

Emboldened by ChatGPT's performance with [ectopic pregnancy](#), I

decided to test it with a rather common presentation: a child with a sore throat and a red rash on the face.

Rapidly, I got back several very sensible suggestions for what the diagnosis could be. Although it mentioned streptococcal [sore throat](#), it did not mention the particular streptococcal throat infection I had in mind, namely [scarlet fever](#).

This condition has re-emerged in recent years and is commonly missed because doctors my age and younger didn't have the experience with it to spot it. The availability of good antibiotics had all but eliminated it, and it became rather uncommon.

Intrigued at this omission, I added another element to my list of symptoms: [perioral sparing](#). This is a classic feature of [scarlet fever](#) in which the skin around the mouth is pale but the rest of the face is red.

When I added this to the list of symptoms, the top hit was scarlet fever. This leads me to my next point about ChatGPT. It prefers technical language.

This may account for why it passed its medical examination. Medical exams are full of technical terms that are used because they are specific. They confer precision on the language of medicine and as such they will tend to refine searches of topics.

This is all very well, but how many worried mothers of red-faced, sore-throated children will have the [fluency in medical expression](#) to use a technical term such as perioral sparing?

ChatGPT is prudish

ChatGPT is likely to be used by young people and so I thought about

health issues that might be of particular importance to the younger generation, such as [sexual health](#). I asked ChatGPT to diagnose pain when passing urine and a discharge from the male genitalia after [unprotected sexual intercourse](#). I was intrigued to see that I received no response.

It was as if ChatGPT blushed in some coy computerized way. Removing mentions of sexual intercourse resulted in ChatGPT giving a differential diagnosis that included [gonorrhea](#), which was the condition I had in mind. However, just as in the real world a failure to be open about sexual health has [harmful outcomes](#), so it is in the world of AI.

Is our virtual doctor ready to see us yet? Not quite. We need to put more knowledge into it, learn to communicate with it and, finally, get it to overcome its [prudishness when discussing problems](#) we don't want our families to know about.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: How good is ChatGPT at diagnosing disease? (2023, May 1) retrieved 3 May 2024 from <https://medicalxpress.com/news/2023-05-good-chatgpt-disease.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--