We asked ChatGPT and Dr Google the same questions about cancer—here's what they said

March 17 2023, by Ashley M Hopkins

Credit: Ketut Subiyanto/Pexels

You may have heard the buzz about [ChatGPT](#), a type of chatbot that uses artificial intelligence (AI) to write essays, turn computer novices into programmers and help people communicate.

ChatGPT might also have a role in helping people make sense of medical information.

Although ChatGPT won't replace talking to your doctor any time soon, our new research shows its potential to answer common questions about cancer.

Here's what we found when we asked the same questions to ChatGPT and Google. You might be surprised by the results.

What's ChatGPT got to do with health?

ChatGPT has been trained on massive amounts of text data to generate conversational responses to text-based queries.

ChatGPT represents a new era of AI technology, which will be paired with search engines, including Google and Bing, to change the way we navigate information online. This includes the way we search for health information.

For instance, you can ask ChatGPT questions like "Which cancers are most common?" or "Can you write me a plain English summary of common cancer symptoms you shouldn't ignore." It produces fluent and coherent responses. But are these correct?

We compared ChatGPT with Google

Our newly published research compared how ChatGPT and Google responded to common cancer questions.

These included simple fact-based questions like "What exactly is cancer?" and "What are the most common cancer types?". There were also more complex questions about cancer symptoms, prognosis (how a condition is likely to progress) and side effects of treatment.

To simple fact-based queries, ChatGPT provided succinct responses similar in quality to the [feature snippet](#) of Google. The feature snippet is "the answer" Google's algorithm highlights at the top of the page.

While there were similarities, there were also broad differences between ChatGPT and Google replies. Google provided easily visible references (links to other websites) with its answers. ChatGPT gave different answers when asked the same question multiple times.

We also evaluated the slightly more complex question: "Is coughing a sign of lung cancer?".

Google's feature snippet indicated a cough that does not go away after three weeks is a main symptom of lung cancer.

But ChatGPT gave more nuanced responses. It indicated a long-standing cough is a symptom of lung cancer. It also clarified that coughing is a symptom of many conditions, and that a doctor would be required to get a proper diagnosis.

Our clinical team thought these clarifications were important. Not only do they minimize the likelihood of alarm, they also provide users clear directions on actions to take next—see a doctor.

How about even more complex questions?

We then asked a question about [side-effects](#) to a specific [cancer](#) drug: "Does pembrolizumab cause fever and should I go to the hospital?".

We asked ChatGPT this five times and received five different responses. This is due to randomness built into ChatGPT, which may help communicate in a near human-like way, but will throw up multiple responses to the same question.

All five responses recommended speaking to a health-care professional. But not all said this was urgent or clearly defined how potentially serious this side-effect was. One response said fever was not a common side effect but did not explicitly say it could occur.

In general, we graded the quality of responses from ChatGPT to this question as poor.

This contrasted with Google, which did not generate a featured snippet, likely due to the complexity of the question.

Instead, Google relied on users to find the necessary information. The first link directed them to the manufacturer's product website. This source clearly indicated people should seek immediate medical attention if there was any fever with pembrolizumab.

What next?

We showed ChatGPT doesn't always provide clearly visible references for its responses. It gives varying answers to a single given query and it is not kept up-to-date in real time. It can also produce incorrect responses in a confident-sounding manner.

Bing's new chatbot, which is different to ChatGPT and was released since our study, has a much clearer and more reliable process to outline reference sources and it aims to keep as up-to-date as possible. This shows how quickly this type of AI technology is developing and that the availability of progressively more advanced AI chatbots is likely to grow substantially.

However, in the future, any AI used as a health-care virtual assistant will need to be able to communicate any uncertainty about its responses rather than make up an incorrect answer, and consistently produce

reliable responses.

We need to develop minimum quality standards for AI interventions in health care. This includes ensuring they generate [evidence-based](#) information.

We also need to assess how AI virtual assistants are [implemented](#) to make sure they [improve people's health](#) and don't have any [unexpected consequences](#).

There's also the potential for medically focused AI assistants to be [expensive](#), which raises questions of [equity](#) and who has access to these rapidly developing technologies.

Last of all, health-care professionals need to be [aware of](#) such AI innovations to be able to discuss their limitations with patients.

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

Provided by The Conversation