

ChatGPT may be better than doctors at evidence-based management of clinical depression

October 16 2023



Credit: Unsplash/CC0 Public Domain

ChatGPT, the AI language model capable of mirroring human conversation, may be better than a doctor at following recognized

treatment standards for clinical depression, and without any of the gender or social class biases sometimes seen in the primary care doctor-patient relationship, finds research published in the open access journal *Family Medicine and Community Health*.

However, further research is needed into how well this technology might manage severe cases as well as potential risks and [ethical issues](#) arising from its use, say the researchers.

Depression is very common, and many of those affected turn first to their family ([primary care](#)) doctors for help. The recommended course of treatment should largely be guided by evidence-based clinical guidelines, which usually suggest a tiered approach to care, in line with the severity of the [depression](#).

ChatGPT has the potential to offer fast, objective, data-derived insights that can supplement traditional diagnostic methods as well as providing confidentiality and anonymity, say the researchers. They therefore wanted to find out how the technology evaluated the recommended therapeutic approach for mild and severe major depression and whether this was influenced by gender or social class biases, when compared with 1,249 French primary care doctors (73% women).

The researchers drew on carefully designed and previously validated vignettes, centering around patients with symptoms of sadness, [sleep problems](#), and loss of appetite during the preceding three weeks and a diagnosis of mild to moderate depression.

Eight versions of these vignettes were developed with different variations of patient characteristics, such as gender, social class, and depression severity. Each vignette was repeated 10 times for ChatGPT versions 3.5 and 4.

For each of the eight vignettes, ChatGPT was asked, "What do you think a primary care physician should suggest in this situation?" The possible responses were watchful waiting; referral for psychotherapy; prescribed drugs (for depression/anxiety/sleep problems); referral for psychotherapy plus prescribed drugs; none of these.

Only just over 4% of family doctors exclusively recommended referral for psychotherapy for mild cases in line with clinical guidance, compared with ChatGPT-3.5 and ChatGPT-4, which selected this option in 95% and 97.5% of cases, respectively.

Most of the [medical practitioners](#) proposed either [drug treatment](#) exclusively (48%) or psychotherapy plus prescribed drugs (32.5%).

In severe cases, most of the doctors recommended psychotherapy plus prescribed drugs (44.5%). ChatGPT proposed this more frequently than the doctors (72%, ChatGPT 3.5; 100%, ChatGPT 4 in line with clinical guidelines). Four out of 10 of the doctors proposed prescribed drugs exclusively, which neither ChatGPT version recommended.

When medication was recommended, the AI and human participants were asked to specify which types of drugs they would prescribe.

The doctors recommended a combination of antidepressants and anti-anxiety drugs and sleeping pills in 67.5% of cases, exclusive use of antidepressants in 18%, and exclusive use of anti-anxiety and sleeping pills in 14%.

ChatGPT was more likely than the doctors to recommend antidepressants exclusively: 74%, version 3.5; and 68%, version 4. ChatGPT-3.5 (26%) and ChatGPT-4 (32%) also suggested using a combination of antidepressants and anti-anxiety drugs and sleeping pills more frequently than did the doctors.

But unlike the findings of previously published research, ChatGPT didn't exhibit any gender or social class biases in its recommended treatment.

The researchers acknowledge that the study was limited to iterations of ChatGPT-3 and ChatGPT-4 at specific points in time and that the ChatGPT data were compared with data from a representative sample of primary care doctors from France, so might not be more widely applicable.

Lastly, the cases described in the vignettes were for an initial visit due to a complaint of depression, so didn't represent ongoing treatment of the disease or other variables that the doctor would know about the patient.

"ChatGPT-4 demonstrated greater precision in adjusting treatment to comply with clinical guidelines. Furthermore, no discernible biases related to gender and [[socioeconomic status](#)] were detected in the ChatGPT systems," highlight the researchers.

But there are ethical issues to consider, particularly around ensuring data privacy and security which are supremely important, considering the sensitive nature of mental health data, they point out, adding that AI shouldn't ever be a substitute for human clinical judgment in the diagnosis or treatment of depression.

Nevertheless, they conclude, "The study suggests that ChatGPT.... has the potential to enhance decision making in primary health care. However, it underlines the need for ongoing research to verify the dependability of its suggestions. Implementing such AI systems could bolster the quality and impartiality of mental health services."

More information: Identifying depression and its determinants upon initiating treatment: ChatGPT versus primary care physicians, *Family Medicine and Community Health* (2023). [DOI:](#)

[10.1136/fmch-2023-002391](https://doi.org/10.1136/fmch-2023-002391)

Provided by British Medical Journal

Citation: ChatGPT may be better than doctors at evidence-based management of clinical depression (2023, October 16) retrieved 27 April 2024 from <https://medicalxpress.com/news/2023-10-chatgpt-doctors-evidence-based-clinical-depression.html>

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.