# Use and Limitations of ChatGPT in Mental Health Disorders

Jovana Petrović[†]
Clinic of Psychiatry
University Clinical center
Nis, Serbia
zdravkovicj91@gmail.com

Iva Binić
Clinic of Psychiatry
University Clinical center
Nis, Serbia
ivabinic@gmail.com

Maša Vacev
Clinic of Psychiatry
University Clinical center
Nis, Serbia
vacevmasa@gmail.com

Stevo Lukić[†]
Clinic of Neurology,
University Clinical centre
Niš, Serbia
slukic@medfak.ni.ac.rs

## Abstract / Povzetek

ChatGPT is one of the most advanced and rapidly evolving large language model-based chatbots. It excels in everything from handling simple questions to performing complex medical examinations. While current technology cannot replace the expertise and judgment of skilled psychiatrists, it can assist in early detection of mental problems, patient evaluations, differential diagnoses, psychotherapy and in planning and conducting medical research. Ensuring privacy and adhering to professional, ethical, and legal standards is crucial when processing training data. This is especially important in mental health settings, where disclosing sensitive personal information increases the risk of data misuse and the potential for harmful advice. Current uses of ChatGPT in mental health care are constrained by its design as a general chatbot, rather than a specialized psychiatric tool. Despite this, the model proves useful for handling routine psychiatric and administrative tasks. As GPT technology evolves, it holds significant promise for psychiatry, including integration into diagnostics, psychotherapy, and early detection of mental health issues. To deploy these advancements responsibly and effectively, it is crucial to develop and refine professional ethical standards and practice guidelines.

## Keywords / Ključne besede

Keywords mental health disorders, large language models, deep learning, ChatGPT

## Introduction

ChatGPT has emerged as one of the most advanced and rapidly evolving large language model-based chatbot systems. Its extensive capabilities, ranging from responding to basic inquiries to performing well in complex medical examinations, have garnered significant attention from the global scientific and research communities, prompting ongoing discourse regarding its potential applications across diverse domain [1]

The discourse surrounding the potential applications of ChatGPT in mental health disorders remains relatively underexplored. This work seeks to offer a quick overview of the current state of ChatGPT implementations within the mental health domain, while also projecting future advancements in digital mental health care through the integration and development of ChatGPT technology.

## Current use of ChatGPT for mental health disorders

In managing mental health disorders, human contact is especially crucial compared to other medical fields, as it is accompanied by understanding and empathy. This is why, in the most vital aspects of psychiatric work, human relationships remain central [2]. While ChatGPT's practical applications in the field of mental health are limited because it specializes in language generation, it can still support certain routine tasks within the field. Although evaluations, diagnoses, psychotherapy, and patient assessments are mainly conducted by human therapists, ongoing trials are exploring how ChatGPT might be utilized in mental health services. Several platforms, like ChatBeacon, or Koko, are already available on the market, claiming to offer mental health assistance powered by ChatGPT [3,4]. For instance, Koko is a peer-support platform that provides crowdsourced cognitive therapy. It's experimenting with using GPT-3 to generate bot-written responses to users while they wait for peer responses. Koko is an online mental health intervention that has reached nearly two million people, mostly adolescents. The platform started as a clinical trial at MIT and is based on the concept of crowdsourced cognitive therapy. Users are taught to help each other think more hopefully about the world. Unlike traditional peer support platforms, all interactions on our service are supported and augmented by AI.

**Applied to the classification of psychiatric disorders**.

Recent advancements in deep learning, the foundational algorithm of GPT, have significantly impacted the field of mental health disorders. This technology has been applied to classify psychiatric disorders using neuroimaging data [5], develop models based on electroencephalograms [6], and utilize a range of patient characteristics for diagnosing and predicting mental disorders [7]. These deep learning models have shown good diagnostic accuracy (AUC 0.74- 0.81) suggesting the possibility of combining genetics and registry data to predict both mental disorder diagnosis and disorder progression in a clinically relevant, cross-diagnostic setting prior to clinical

* All authors contributed equally to this research
[†]Corresponding authors


*Information Society 2024, 7–11 October 2024, Ljubljana, Slovenia*


**https://doi.org/10.70314/is.2024.chtm.6**

assessment. The main limitation of this model is that it is restricted to learning from historical data and should be continuously assessed and evaluated by trained clinicians and never stand alone in the decision-making

**Utilized to alleviate the burdens associated with clinical documentation, communication, and research tasks**.
Also, new technologies can assist clinicians by allowing them to focus more on direct patient care and alleviate the high clinical workload and bureaucratic tasks- such as handling admissions and managing paperwork- that have been linked to burnout in earlier research [8]. ChatGPT can assist in processing clinical case transcripts, generating summaries, completing medical record documentation, and efficiently preparing discharge summaries. It can also help facilitate communication between clinicians of different specialties when consultations are needed, aiding in an integrative approach [9].
While current technology cannot replace the expertise and judgment of skilled psychiatrists, it can assist in generating differential diagnoses based on relevant signs and symptoms. ChatGPT is recognized for its substantial potential to assist experts with clinical and laboratory diagnoses, as well as in planning and conducting medical research [10].

**Applied in psychotherapy to enhance therapeutic processes**
Chatbots can be also beneficial in psychotherapy. The therapist's emotions and the emotional alignment between therapist and client are crucial factors influencing the process and outcomes of therapy [11]. A study conducted during the COVID-19 pandemic showed that technology can offer an effective method, providing at least a first level counseling support structure [12]. This implies that GPT models may potentially develop cognitive empathy over time, making it possible for ChatGPT to achieve a notable level of accuracy in identifying users' emotions [13], nevertheless it is important to make systematic testing to ensure a non-superficial comparison between human and artificial intelligences [14]. Namely, ChatGPT currently lacks the capability to accurately assess personality traits, a task that demands extensive training and expertise from psychiatrists. AI researchers are diligently pursuing technical advancements to improve the precision of personality detection [15].
A study designed to assess the accuracy and appropriateness of psychodynamic formulations generated by ChatGPT found that the model produced suitable results even without additional psychoanalytic information. It also demonstrated a strong ability to generate psychodynamic formulations consistent with various psychoanalytic theories when given appropriate instructions. [16].
The research underscores that ChatGPT is not intended to substitute psychiatrists or psychologists but rather to function as an initial resource and a first line of support for those dealing with mental distress. When used wisely and within appropriate limits, ChatGPT can be an effective tool in supporting mental health services. [17]. However, it is essential to process training data in a manner that ensures privacy protection and adheres to all professional, ethical, and legal standards. Given that individuals may be at increased risk of data misuse when disclosing sensitive personal and family information during mental health treatment [18].

**Employed for the prevention and early detection of mental health issues**
The role of AI in the prevention and early detection of mental problems can also be very significant. Patients frequently turn to ChatGPT to seek information about their symptoms, possible diagnoses, and treatment options. ChatGPT performs better than Google Search in delivering general medical knowledge but scores lower when it comes to providing medical recommendations [19]. A recent study highlighted early success for an AI model that can detect cognitive distortions in text messages with accuracy comparable to that of clinically trained human raters [20]. GPT's ability to recognize mental health warning signs in routine conversations or text-based telemedicine interactions has the potential to improve early and effective intervention strategies when necessary.

## Risks and limitations ChatGPT use for mental health disorders

Despite its strengths and potential, the use of AI technologies in psychiatric clinical practice carries several risks. A significant concern is the phenomenon of "artificial hallucinations" where the conversational model may confidently produce text that is factually incorrect, nonsensical, or misleading [18].

Recent systematic review that included 118 articles identified some limitations regarding the potential of ChatGPT in patient care and medical research, noted that the solutions provided by ChatGPT are often insufficient and contradictory, raising concerns about their originality, privacy, accuracy, and legality [10]. It is well established that ChatGPT may generate inaccurate facts and references when summarizing previous research, and the quality of its responses often hinges on how the questions are phrased [21].

Even with the advanced GPT-4 model, there is still a risk of providing harmful advice. The absence of clinical reasoning and experience in ChatGPT can lead to the omission of important clinical details in patient summaries and medical records. Thus, the most prudent approach is to employ AI systems as supplementary tools for mental health professionals, ensuring they are used under close supervision to uphold the safety and quality of patient care. [22]

## Conclusion

The recent introduction of GPT-4 has significantly enhanced the capabilities of the GPT system. Current implementations of ChatGPT within mental health care are limited by its inherent design as a chatbot, rather than as a specialized AI tool specifically tailored for psychiatric use. Nonetheless, this sophisticated language model demonstrates significant utility in addressing various routine psychiatric and administrative functions.
As this technology evolves and advances, we anticipate substantial potential for future applications of GPT technology in psychiatry, including its integration into diagnostic processes, the provision of psychotherapy within clinical environments, and the rapid identification of early warning signs for mental health disorders. Crucially, the development and refinement of professional ethical standards and practice guidelines are imperative for the responsible and effective

deployment of these transformative GPT technologies in the mental health sector.

## Acknowledgments / Zahvala

The work presented in this paper was partially funded in part by the European Union under Grand 101159214- ChatMED.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. Insert paragraph text here.

## References / Literatura

[1] Franco D'Souza R, Amanullah S, Mathew M, Surapaneni KM. (2023) Appraising the performance of ChatGPT in psychiatry using 100 clinical case vignettes. Asian J Psychiatr. 89:103770.
[2] Spencer L, Broome M. (2023) The epistemic harms of empathy in phenomenological psychopathology. Phenom Cogn Sci doi: https://doi.org/10.1007/s11097-023-09930-1.
[3] ChatBeacon. Mental health assistant powered by ChatGPT. (2024) URL: https://www.chatbeacon.io/industrychatgpt/mentalhealth (Accessed 24-Aug-2024)
[4] Koko. GPT-3 mental health intervention. (2024) URL: https://gpt3demo.com/apps/koko-ai (Accessed 24-Aug-2024).
[5] Quaak M, van de Mortel L, Thomas RM, van Wingen G. (2021) Deep learning applications for the classification of psychiatric disorders using neuroim- aging data: Systematic review and meta-analysis. Neuroimage. Clin. 30:102584.
[6] de Bardeci M, Ip CT, Olbrich S. (2021) Deep learning applied to electroencephalogram data in mental disorders: A systematic review. Biol. Psychol. 162:108117.
[7] Allesoe RL, Thompson WK, Bybjerg-Grauholm J et al. (2023) Deep learning for crossdiagnostic prediction of mental disorder diagnosis and prognosis using Danish Nationwide register and genetic data. JAMA Psychiatry 80:146-155
[8] Kumar S. (2007) Burnout in psychiatrists. World Psychiatry 6:186-189.
[9] Ali SR, Dobbs TD, Hutchings HA, Whitaker IS. (2023) Using ChatGPT to write patient clinic letters. Lancet Digit. Health 5:e179-e181.
[10] Garg RK, Urs VL, Agarwal AA, Chaudhary SK, Paliwal V, Kar SK. (2023) Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: A systematic review. Health Promot Perspect. 13(3):183-191.
[11] Chui H, Li X, Luk S. (2022) Therapist emotion and emotional change with clients: Effects on perceived empathy and session quality. Psychotherapy (Chic). 59(4):594-605.
[12] Mawani A, Nderu L. (2020) Towards an online empathy assisted counselling web application. EAI Endorsed Transactions on Context-aware Systems and Applications 7(22):167792.
[13] Kosinski M. (2023) Theory of mind may have spontaneously emerged in large language models. ArXiv abs/2302.02083.
[14] Strachan J, Albergo D, Borghini G. et al. (2024) Testing theory of mind in large language models and humans. Nat Hum Behav 8:1285-1295.5
[15] El-Demerdash K, El-Khoribi RA, Ismail Shoman MA, Abdou S. (2022) Deep learning based fusion strategies for personality prediction. Egypt Inform J 23:47-53.
[16] Hwang G, Lee DY, Seol S. et al. (2024). Assessing the potential of ChatGPT for psychodynamic formulations in psychiatry: An exploratory study. Psychiatry Res. 331:115655
[17] Arjanto P, Senduk FFW, Nahdiyah U, Utami MS. (2024) AI and ethics in mental health: exploring the controversy over the use of ChatGPT. J Public Health (Oxf). 46(2):e340-e341.
[18] Wei Y, Guo L, Lian C, Chen J. (2023) ChatGPT: Opportunities, risks and priorities for psychiatry. Asian J Psychiatr. 90:103808
[19] Ayoub NF, Lee YJ, Grimm D, Divi V. (2024) Head-to-Head Comparison of ChatGPT Versus Google Search for Medical Knowledge Acquisition. Otolaryngol Head Neck Surg. 170(6):1484-1491.
[20] Tauscher JS, Lybarger K, Ding X et al. (2023) Automated detection of cognitive distortions in text exchanges between clinicians and people with serious mental illness. Psychiatr. Serv. 74: 407-410.
[21] The Lancet Digital Health. (2023) ChatGPT: friend or foe? Lancet Digit Health. 5(3):e102.
[22] Jo E, Song S, Kim JH, et al. (2024) Assessing GPT-4's Performance in delivering medical advice: Comparative analysis with human experts. JMIR Med Educ. 10:e51282.