ChatGPT and other Chatbots in Psychiatry

Goran Arbanas[1,2]

[1]Department for Forensic Psychiatry, University Psychiatric Hospital Vrapče, Zagreb, Croatia, [2]Medical School Rijeka, University of Rijeka, Rijeka, Croatia

Abstract

Aim: Artificial intelligence (AI) is making significant inroads into the field of psychiatry, offering new tools and applications. ChatGPT, a specific chatbot, is at the forefront of this digital revolution. AI's use in psychiatry ranges from identifying psychiatric symptoms, predicting treatment responses, and improving medication adherence to assisting in patient education, monitoring, and bridging gaps in mental health care. **Materials and Methods:** This review used a literature study method. **Results:** ChatGPT functions as a clinical decision support tool. It can analyse patient data and provide diagnostic insights, recommend evidence-based treatments, and offer drug information. It has demonstrated proficiency in generating summaries from medical records, saving clinicians time and enabling them to focus on patient care. Additionally, chatbots like ChatGPT serve as therapist assistants, offering emotional support between therapy sessions and potentially conducting psychotherapy. Studies have shown positive outcomes, with chatbots reducing depression, anxiety symptoms, and providing 24/7 availability for crisis situations. Users find them non-judgmental and comfortable for discussing sensitive issues. Despite their potential, chatbots have limitations, such as the risk of incorrect or biased information due to their training data. They lack genuine understanding, creativity, and the ability to clarify user input. Ethical considerations regarding responsibility and data usage are paramount. **Conclusion:** AI, particularly ChatGPT, holds substantial promise in modern psychiatry, enhancing diagnostics, patient education, monitoring, and therapeutic support. Its integration into everyday psychiatric practice requires careful use, continuous oversight, and ethical considerations. Psychiatrists must become more familiar with AI tools to leverage their benefits in patient care.

Introduction

Artificial intelligence in psychiatry

There have been four industrial revolutions in the history of humankind: mechanical revolution (with the invention of a steam engine), electrical (with electrification of industry and everyday life), Internet (which has become widespread) and digital (characterized by a fusion of technology types) [1]. Artificial intelligence (AI) is one of the forms of a digital revolution. It was originally recognized in 1956, and the term AI was coined by computer scientist John McCarthy, who defined it as „the science and engineering of making intelligent machines" [2]. The word „intelligent" is used when the technology has the ability to learn. There are several ways in which a machine/computer is able to learn. Ma-

Correspondence to:
Goran Arbanas
Klinika za psihijatriju Vrapče,
Bolnička cesta 32, 10 000 Zagreb
Phone: +385 98 304405
E-mail: goran.arbanas@bolnica-vrapce.hr

chine learning (ML) is an approach to enabling an algorithm to learn [3]. There are three types of ML: supervised ML (data are pre-labelled and the algorithm learns to associate input features to predict the labels; it learns from a large amount of labelled training data), unsupervised ML (it uses clustering techniques to sort raw, unlabelled data) and deep learning (by employing artificial neuronal networks /ANNs/) [3]. ANNs are computerized models that imitate the construction and operation of biological neural networks in the brain [4]. They are composed of interlinked artificial neurons classified into three layers: input layer (which accepts the data), hidden layers (which perform calculations, but invisibly to the customer) and output layer (generates the results) [4]. The "neuron" can be either activated or inactive, based on the total weighted input. During the training process it is exactly these weights and biases that make the adjustment possible.

The first use of AI in psychiatry dates back to 1960s with computer programmes used to support diagnostic decisions and create treatment plans [5]. Today's use of AI in contemporary psychiatry is wide: it can identify pre-determined psychiatric symptoms or suicide ideation from health data and can predict depression and suicidality on unstructured texts, social media posts, twitter data or Instagram posts; it can predict which patients will respond to citalopram based on clinical history, sociodemographics and mood; which clusters of symptoms of depression (i.e. core symptoms, sleep symptoms and atypical symptoms) will respond to different antidepressants, which group of patients will better respond to placebo (those with higher education) [1,6-8]. Apart from these mainly statistical methods, an AI platform on mobile devices was able to improve medication adherence, compared to directly observed therapy in persons with schizophrenia [9]. Accuracies of these different methods ranged from 62 % (smartphone data and information from social media) to 98% (predictions from physical functions and sociodemographic data) [1]. It is important to interpret the results we obtain from AI wisely. For example, depressive posts recognized by AI in social media may suggest depressive mood and not a diagnosis of depression.

Chatbots are specially designed AI platforms that generate human-like conversations; therefore, some call them conversational AI [10,11]. They have been widely used in service industries: by hotels, phone providers, etc (in any domain where it is expected that many people will ask similar questions – e.g. about prices, cancellation policies, terms and conditions). Chatbots often start the conversation by stating „My name is Marina, how can I help you?" or similar greetings (mimicking human conversation starters), giving the illusion the customer is speaking to a real human being.

ChatGPT is a specific chatbot. It stands for Chat Generative Pre-trained Transformer. It is a large language model (LLM) trained on a massive amount of culled textual data and can replicate human discourse [12]. It was fed by a huge amount of text found on the Internet. The word generative means that it can create new contents [13]. Responses are generated via an autoregressive statistical model which outputs a word based on the probabilities of different words following the previous ones [11]. This means that the response is not intelligent, ChatGPT does not understand why this is the response, but simply 'calculates' the best match. In other words, it generates text based on the prompts provided, without understanding the prompts' and the responses' meaning [4]. It is a probabilistic text generator [10]. It is a simulation of intelligence and not intelligence itself [14]. Nevertheless, it emulates human speech and interacts with its user in a dialogue way [15]. ChatGPT was launched on November 30, 2022 by San Francisco-based Open AI [12,16].

Still, the responses can be optimized using reinforcement learning with human feedback and therefore, the contemporary ChatGPT provides better fitting answers, compared to the earlier versions of it. Furthermore, it can be optimized with specific feedback (e.g. from professional texts or on a specific topic) and in such cases specific chatbots are produced that are better suited for a specific task [4,10]. This is what P (pre-trained) stands for in GPT. It is important to mention that ChatGPT was not primarily introduced to be a medical help and it was fed by different, both professional and lay texts, and therefore may produce answers that are based on all the texts provided and not only on texts that are medically appropriate.

The aim of this review was to gather and discuss the contemporary knowledge about the use of chatbots (and especially ChatGPT) in clinical psychiatric work.

Materials and Methods

This review used a literature study method. The literature search was conducted on PubMed, Google Scholar and ResearchGate platforms, using keywords "ChatGPT", "chatbot" and "psychiatry". Furthermore, all the references from the searched studies were carefully considered and included if appropriate.

Potentials of ChatGPT and other chatbots in psychiatry

ChatGPT can be used as a clinical decision support tool to analyse patient data and assist a psychiatrist in the clinical assessment and management. The user can feed it with information about their patient and ask ChatGPT to analyse the data and provide ideas on a diagnosis for the given symptoms, or to

provide differential diagnostic options and appropriate work-up to either confirm them or dispute them. Usually, it was able to identify a medical condition based on initial symptoms and was able to provide a list of differentials and an incomplete list of assessments [17]. Nevertheless, with additional provocation it gave a complete list (and this underlines the fact that there is always the need for a medical professional to approve information) [12,13]. It was the least successful in diagnosing personality disorders [18].

ChatGPT may provide evidence-based treatment recommendations tailored to patient's needs and preferences and a holistic plan of the management, including pharmacological and non-pharmacological approaches [12]. Studies prove that ChatGPT gives suggestions on treatment based on generally accepted guidelines. It can also be used as a drug information retriever providing details. A clinician can ask ChatGPT to give him details about possible side-effects and adverse reactions, laboratory changes or interactions with other medications. Furthermore, it gives information about possible indications and dosages [12,15].

It is worth noting that ChatGPT has passed United States Medical Licencing Examination (USMLE), steps 1-3 and has been proved to give differential diagnoses from clinical vignettes by the 93.3 % accuracy [19]. Of course, all the details should be checked by the physician and not followed blindly. With a sensible use, psychiatrists can make more accurate and informed diagnoses, design more effective treatments and personalized care for individual patients, but also make the access to mental health care more approachable for patients [15]. It is important to note that 78 % of participants were willing to use ChatGPT for self-diagnosis [20]. Participants in studies have shown that they are overall willing to receive health advice from chatbots (especially on logistical issues and preventative measures) [10,20,21]. When given both answers from ChatGPT and human professionals, participants were only weakly able to distinguish who was the author of the responses (ability to identify the author ranged from 49 to 86 %, depending on a question, with the average of 65 %) [10]. Yet, ChatGPT was not trained on medical data specifically, and there are medically trained chatbots, such as Med-PaLM, which might produce even better results [10].

There is an additional function of trained chatbots. It has been shown that they can identify the content of speech that includes hallucinations and delusions, as well as other pre-defined sets of symptoms [15]. It can also recognize main mood from the text and can explain its reasoning. Thus, it can help gather information about symptoms and assist in making a diagnosis [5]. Similarly, we can use ChatGPT as a tool to help educate our patients. It can give information about the disease, possible risk factors, but also about life style changes or about medications. It can save time since it is available 24/7 and can answer patient's questions repeatedly, without tiring [15].

Furthermore, it can be used for patient monitoring. This can be specially the case in children, adolescents and young adults [16]. There are studies conforming that it can give support be-tween therapy sessions [4]. It can help in regulation of emotions and behaviours and support patients through meditation, mindfulness or helping patients setting their goals [5].

Adolescents and young adults are more prone to use digital technologies, because they have been born with this technology and are available on-line all the time. Furthermore, lack of child and adolescent mental health professionals is even more pronounced than in adult psychiatry. Chatbots can be one of the possibilities of bridging this gap [16]. ChatGPT can help clinicians to generate summaries from medical records and put them in admission notes or discharge summaries and thus save time and clinicians can focus on patient care [13,22]. Another function of ChatGPT is to assist medical education. It is easy to create simulated patient examples (simply by asking it to produce an example of a patient with particular symptoms and particular aethyology; by additional instructions ChatGPT can produce specific details). In such a way we can produce examples for students or residents without the risk of harmful behaviour or any ethical issues and without responsibility [23]. ChatGPT can create case vignettes on certain topics, even in other languages then English [24]. This might be important, because some research show that people more often use terms in indigenous languages than in English, when searching for health information [25]. Further, chatbots can help with data collection and analysis by researchers. It can be used to polish academic writing for readability [13]. Due to extended use of ChatGPT by some researchers and due to ethical pressures, World Association of Medical Editors have reached some recommendations regarding the use of AI in publishing results: ChatGPT cannot be listed as a co-author, its use should be acknowledged and it should be clearly stated what it was used for [13,14].

Chatbots as psychiatrists/therapists

Science fiction author Frederik Pohl described a computerized version of a psychiatrist in his novel Gateway, 35 years ago [26]. He even gave paratext examples of a computer programme standing behind this AI therapist (if we use contemporary terms). It seems that his predictions came true.

Eshghie and Eshgie showed that it is possible to give prompts to ChatGPT to instruct it how to talk with their patients in between sessions [4]. They showed that ChatGPT is capable of holding the conversation in a positive, non-judgmental and supportive way, to ask clarifying questions and provide feedback that validated patients' feelings and experiences. It also suggested potential coping strategies and encouraged patients to explore the underlying triggers. Thus, ChatGPT can be used as a therapist assistant for emotional support in between therapy sessions.

The final frontier is use of chatbots as therapist substitutes. This is an important issue, knowing that there is an insufficient number of mental health professionals available [11]. There have been some studies showing that chatbots can be used for psychotherapy sessions, allowing patients to interact with a virtual therapist through an online platform [5]. Even

ChatGPT proved to be a good listener, can be employed 24/7, and is more cost effective. Furthermore, it is not biased and judgemental and there are patients who are unwilling to open in front of a human therapist (fearing they might judge them) but feel talking about intimate issues in front of a non-human machine less frightening [11]. Some consider sharing personal history with AI to be less stigmatizing and would prefer to talk to an AI and not to a human psychiatrist because they feel more comfortable with self-disclosure to a computer [23]. Furthermore, ChatGPT is able to provide companionship at any time [16]. Because the responses of ChatGPT are human-like, they can provide company, support and therapy [27].

We already have some results on therapeutic use of chatbots. There are some positive results in reducing depression among college students with a conversational chatbot (Woebot), after a two-week cognitive behavioural based therapy and the users considered it to be empathic [16,28]. Also, ChatGPT can mimic therapists for people with social anxiety disorder [29]. On the other hand, depression-focused chatbot called Tess proved positive effect on the reduction of anxiety symptoms, but not depressive symptoms [30]. Abd-Alrazaq et al found weak evidence that chatbots can help in managing depression, stress, and distress but had no effect on subjective psychological wellbeing [31]. Furthermore, due to its 24/7 availability it can be valuable in crises situations, to help control suicidal ideation and panic attacks [16]. Others showed it can help in regulation of emotions and behaviours and support patients through meditation, mindfulness or helping patients setting their goals [5].

Some authors even studied human-chatbot relationship development (on 18 subjects). Skjuve et al found that the relationship was of a superficial character at the outset, but evolved to a substantial affective exploration and engagement and was rewarding for the users. They concluded there are three stages to the relationship: explorative, affective and stable [32]. When chatbot was being seen as accepting, understanding and nonjudgmental, the relationship was facilitated. Meng and Dai also confirmed that chatbots are as effective in providing emotional support as human partners in reducing stress and worry. The reduction in symptoms happened only when chatbots offered emotional support; those chatbots that were solely self-disclosing did not provide any reduction in stress and worry [33].

Would attaching friendly avatars help as well as combination of language analysis with physiological measurements achieved by wearable devices? [13]. Such idea has been already described in Pohl's Gateway. He suggested that a patient should be connected to instruments measuring physiological parameters (e.g. pulse, blood pressure, sweating intensity) while talking to an AI psychiatrist SF [26].

Chatbots weaknesses

Similarly to humans, chatbots can provide false answers. There are three types of false answers [11]. The first type is due to the data chatbots were fed with. ChatGPT was fed with a large amount of data from Internet. Yet, these data were not only specific or professional, authentic data, but anything available, from both professional and lay sources. Since ChatGPT produces answers by calculating the most probable response, based on the given information, it can give answers that result from the large dataset which was not specific, professional, nor scientific [16]. Therefore, it can produce wrong information and inappropriate advice [27]. Because the data it depends upon were biased, it will give biased answers. The data chatbots are based on were created with overrepresentation of certain groups, e.g. European and North American datasets, and therefore, chatbots' answers are biased towards these groups [10]. The second type of false answers is the result of the fact that ChatGPT does not broaden its information base. It was launched with the dataset finalized in September 2021 and it does not know anything that was added to the Internet during the last two years (because its pretraining ended in September 2021) [13]. Finally, the third type of false answers are just fabricated data that it sometimes gives. If you ask ChatGPT to produce literature on a specific psychiatric topic it will give you references that are technically correct (the correct order of authors and punctuation), but with non-existent articles. Computer scientists call these incorrect, non-existent outputs that were generated by the model and which even appear in the new version, GPT4 - hallucinations [13].

Chatbots do not understand what they produce, they just provide calculated answers from the information they were fed with. Words for them do not have meanings. They lack common sense and make reasoning errors [13,23]. They do not have creativity, nor can they produce personal, individual responses. They are not tuned to individual differences and do not identify a person who needs more assistance [16]. If a user, for example, puts an incorrect decimal point or misspells something, ChatGPT cannot recognize that and it will give inaccurate output. It does not have the ability to ask users to clarify input [12]. One of the answers to these risks is a co-creation and collaboration between psychiatrists and AI to provide best possible care. The combination of the expertise of psychiatrists and the analytic power of AI provides a synergistic effect [5]. For all these reasons, professionals should always verify and revise the content and routine monitoring of the system should be mandatory [13].

It should be added that ChatGPT does not recall conversations from previous sessions if a new thread is started [4]. There is also an ethical problem of whose responsibility it is when something goes wrong – is it the responsibility of a psychiatrist using these data, of a patient (if they gave their consent to a medical doctor to use AI), of a company who developed the specific AI or the system itself? Maybe we should think of asking our patients to give their consent that we can use chatbots or other AI applications in our practice. Or to give their consent to provide their social media content for assessment by GPT-based medical apps [13].

For all of these reasons American Psychiatric Association has formulated a digital psychiatry task force to evaluate and monitor AI and mental health related applications [27]. Where are Croatian psychiatrists and professional associations in this respect?

Use of ChatGPT in everyday psychiatric practice

So far, it seems that psychiatrists are less involved into application of AI, including ChatGPT, into their everyday work, and some believe it is because of lack of training of psychiatrists on skills of using software for clinical applications and computer science and AI [23]. Today, use of AI is not included into medical education and medical curriculum in major Medical Schools [23].

One possible application of ChatGPT in everyday work is to use patient data when diagnosing or creating treatment plans and write it in a style of a case vignette and ask ChatGPT for differential diagnoses and proper investigation, treatment plan, side effect profile and standards of care. AI will give detailed list of all the abovementioned elements. The user needs to keep in mind that the result will depend on the input (more specific the input, more specific the output). If we are not fully satisfied with the answer, there is a „regenerate" button, which will provide a new answer to the same question and this button can be applied as many times as we want. Another useful phrase is to ask „Can you expend on this " or „Can you tell me more" and it will provide more specific details. It is possible to give new data, new details and continue the conversation with ChatGPT on some specific aspects of the case.

It is also possible to ask ChatGPT to provide an example of whatever case, diagnosis, symptom or problem we are dealing with [12]. ChatGPT (if we ask it) can make it in a form of an educational material for the patients, in a form of a lecture or in a form of a presentation for students or colleagues. The text it provides can be shorter or longer (we can instruct it to produce a certain number of sentences) and in any level of technical language (e.g. let it sound as a lay person made it or let it sound as a scientist made it). Yet, do never copy the answers from chatbots directly, without checking it because the results can be biased, incorrect, or just false ("hallucinations").

The main limitation for this review is that ChatGPT and other chatbots is a technology that has been changing and improving constantly and published articles cannot produce contemporary knowledge but depicts their use several months ago. Furthermore, majority of studies are professional opinions on the use of chatbots and there has been published only a few original research into the topic.

Conclusion

If we understand how ChatGPT and other chatbots work, we can use them knowledgeably and prudently, primarily for the analysis, organisation, and review of large amounts of data, allowing us more time to focus on patient care.

References

1. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, et al. Artificial intelligence for mental health and mental illnesses: an overview. Curr Psychiatry Rep. 2019;21:116.

2. Simon HA. Artificial intelligence: where has it been, and where is it going? IEEE Knowl Data Eng. 1991;3:128-36.

3. Bzdok D, Meyer-Lindenberg A. Machine learning for precision psychiatry: opportunities and challenges. Biol Psychiatry Cogn Neurosci Neuroimaging, 2018;3:223-30.

4. Eshghie M, Eshghie M. ChatGPT as a therapy assistant: a suitability study. SSRN Electronic. 2304.09873v1.

5. van Dellen E. Artificial intelligence in psychiatry: co-creation of human and ChatGPT. Tijdschri Psychiatr. 2023;65:241-3.

6. Choi SB, Lee W, Yoon JH, Won JU, Kim DW. Ten-year prediction of suicide death using Cox regression and machine learning in a nationwide retrospective cohort study in South Korea. J Affect Disord. 2018;231:8-14.

7. Sau A, Bhakta I. Artificial neural network (ANN) model to predict depression among geriatric population at a slum in Kolkata, India. J Clin Diagnost Res. 2017;11:VC01-4.

8. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. Lancet Psychiatry. 2016;3:243-50.

9. Bain EE, Shafner L, Walling DP, Othman AA, Chuang-Stein C, Hinkle J, et al. Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a phase 2 clinical trial in subjects with schizophrenia. JMIR Mhealth Uhealth. 2017;5:e18.

10. Nov O, Singh N, Mann D. Putting ChatGPT's medical advice to the (Turing) test: survey study. JMIR Med Educ. 2023;9:e46939.

11. Amram B, Klempner U, Shturman S, Greenbaum D. Therapists or replicants? Ethical, legal, and social considerations for using ChatGPT in therapy. Am J Bioeth. 2023;23:40-2.

12. Galido PV, Butala S, Chakerian M, Agustines D. A case study demonstrating applications of ChatGPT in the clinical management of treatment resistant schizophrenia. Cureus. 2023;15:e38166.

13. Cheng SW, Chang CW, Chang WJ, Wang HW, Liang CS, Kishimoto T, et al. The now and future of ChatGPT and GPT in psychiatry. Psychiatry Clin Neurosci. 2023;77:592-6.

14. Kahambing JG. ChatGPT, polypsychic artificial intelligence and psychiatry in museums. Asian J Psychiatr. 2023;83:103548.

15. Çaliyurt O. AI and psychiatry: the ChatGPT perspective. Alpha Psychiatry. 2023;24:41-2.

16. Imran N, Hashmi A, Imran A. Chat-GPT: opportunities and challenges in child mental healthcare. Pak J Med Sci. 2023;39:1191-3.

17. Hirosawa T, Harada Y, Yokose M, Sakamoto T, Kawamura R, Shimizu T. Diagnostic accuracy of differential-diagnosis lists generated by generative pretrained transformer 3 chatbot for clinical vignettes with common chief complaints: a pilot study. Int J Environ Res Public Health. 2023;20:3378.

18. Amin MM, Cambria E, Schuller B. Will affective computing emerge from foundation models and general AI? A first evaluation on ChatGPT. IEEE Iltel Syst. 2023;38:15-23.

19. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health. 2023;2:e0000198.

20. Shahsavar Y, Choudhury A. User intentions to use ChatGPT for self-diagnosis and health-related purposes: cross-sectional survey study. JMIR Hum Factors. 2023;17:e47564.

21. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Health. 2019;21:2055207619871808.

22. Patel SB, Lam K. ChatGPT: the future of discharge summaries? Lancet Digital Health. 2023;5:e107-8.

23. Ventriglio A, Ricci F. ChatGPT use in teaching social psychiatry. Int J Soc Psychiatry. 2023;207640231178481.

24. Torales J, O'Higgins M. ChatGPT and social psychiatry: A commentary on the article Old dog, new tricks? Exploring the potential functionalities of ChatGPT in supporting educational methods in social psychiatry. Int J Soc Psychiatry. 2023;207640231178488.

25. Singh PM, Wight CA, Sercinoglu O, Wilson DC, Boytsov A, Raizada MN. Language preferences on websites and in Google searches for human health and food information. J Med Internet Res. 2007;9:e18.

26. Pohl F. Gateway. London (GB): Gollancz; 2006.

27. Singh OP. Artificial intelligence in the era of ChatGPT - opportunities and challenges in mental health care. Indian J Psychiatry. 2023;65:297-8.

28. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health. 2017;4:e19.

29. Hall J, Jordan S, van Ommeren M, Au T, Sway RA, Crawford J, et al. Sustainable technology for adolescents and youth to reduce stress (STARS): a WHO transdiagnostic chatbot for distressed youth. World Psychiatry. 2022;21:156-7.

30. Klos MC, Escoredo M, Joerin A, Lemos VN, Rauws M, Bunge EL. Artificial intelligence-based chatbot for anxiety and depression in university students: pilot randomized controlled trial. JMIR Form Res. 2021;5:e20678.

31. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res. 2020;22:e16021.

32. Skjuve M, Følstad A, Fostervold KI, Brandtzaeg PB. My chatbot companion – a study of human-chatbot relationships. Int J Human-Comput Stud. 2021;149:102601.

33. Meng J, Dai YN. Emotional support from AI chatbots: should a supportive partner self-disclose or not? JCMC. 2021;26:207-22.