PROCEEDINGS
OF
LONDON INTERNATIONAL CONFERENCES
eISSN 2977-1870

AI Chatbots in Mental Health Support: Are They Effective

Omar Mohammed*
Usame Alan**
Babatunde Banjoko***

Abstract
The increasing global demand for mental health services has highlighted many shortcomings in access, affordability, and timeliness of care. In response, new applications powered by artificial intelligence (AI) in the form of chatbots have been developed to provide increased, scalable access to emotional support, cognitive behavioral techniques, and self-help resources. In this paper, we review the potential for the use of AI chatbots in mental health support and interventions through a focused analysis of applications, clinical evaluations, and user impressions.
Moreover, key examples of applications, for example, Woebot, Wysa, and Youper, yielded some promising results for individuals who experience symptoms of anxiety and/or depression, especially individuals seeking low-barrier, less stigmatizing access to support. Several studies also found benefits in mood charting, facilitating emotional expression, and self-reflection. However, operational usability of AI applications often depends on design efficacy, adherence to evidence-informed therapeutic models, and active user participation.
Indeed, while chatbots are cost-effective and scalable, there are some limitations. Examples of limitations include limited efficacy for some emotional crises that more complex, personalized interventions may deem necessary, insufficient personalization, and lack of true empathy. Additionally, clinical implications cannot progress without further addressing ethical concerns regarding data privacy, informed consent from users, and algorithmic bias in AI responses.
While AI chatbots are not substitutes for licensed mental health professionals, they represent a growing complement in the continuum of care. This paper concludes by emphasizing the need for hybrid human-AI models and stronger regulatory oversight to ensure responsible, safe, and effective deployment.
Keywords: Mental health, AI chatbots, Emotional support, Cognitive behavioral therapy, Accessibility, User experience, Ethical concerns, Clinical evaluation, Hybrid care models.

https://doi.org/10.31039/plic.2025.14.337
*HSI, United States, omarkhader786@gmail.com

**HSI, United States, usamealan.123@gmail.com

***HSI, United States, babsbanjoko@gmail.com

1. Introduction

The worldwide mental health emergency has become increasingly worse in the last decade, with the rate of depression, anxiety, and other mental illnesses increasing rapidly.

The World Health Organization provides estimates that hundreds of millions of people around the world are experiencing mental illnesses, but the number of available licensed mental health professionals is dangerously low. In most instances, this low number instead results in long wait times, high costs, and culturally appropriate mental health care, and these obstacles present significant challenges to clients in receiving timely care.

At the same time, digital health tools are emerging to potentially address provider gaps that exist, especially ones utilizing AI-powered chatbots. Chatbots are available 24/7 and oftentimes are capable of providing structured interventions. These chatbots provide scalable solutions for basic emotional support, mood tracking, and guided self-help strategies. Chatbots have been piloted in situations from university campuses to workplace wellness programs as a way for clients to access low-cost, stigma-free entry to mental health care.

This article reviews the effectiveness of AI chatbots for mental healthcare by synthesizing their operation, therapeutic foundations, and reported outcomes in both clinical research and real-world applications. While early results suggest substantial benefits for mild and moderate symptoms, it remains unclear whether AI chatbots can handle complex emotional crises, ensure adherence, and provide genuine therapeutic rapport.

2. Role of AI Chatbots in Mental Health

2.1 What Are Mental Health Chatbots?

Mental health chatbots are software applications that attempt to approximate human interaction for the purpose of delivering psychological care or wellness interventions. Mental health chatbots can generally be sorted into two categories: rule-based systems and AI-based systems. A rule-based chatbot uses pre-designed scripts and decision trees to guide the user, offering consistent and reliable answers but not particularly responsive. An AI-based system, usually built on large language models (LLMs) and natural language processing (NLP), can navigate context more thoroughly, thus producing more fluid dialogue with real-time user input responses.

Most well-known mental health chatbots implement evidence-based treatment approaches, for example, Cognitive Behavioral Therapy (CBT) and Acceptance and Commitment Therapy (ACT), and Dialectical Behavioral Therapy (DBT) models inform the bot's questions, exercises, and coping strategies. A CBT chatbot, for example, will help the user recognize and reframe unhealthy thinking patterns, reflect upon mood fluctuations, and complete behavioral activation exercises.

2.2 Use Cases

Mental health chatbots have various typical use cases that the application encompasses. Among the most ubiquitous is managing stress and anxiety, where the chatbot delivers grounding exercises, breathing work, and reframing techniques. Another is mood tracking and

journaling, where users can record emotions, identify triggers, and observe patterns over time. Crisis aversion support can also be provided by chatbots in the form of immediate coping strategies or linking users to helplines and emergency services where high-risk situations are detected.

Commercial uses illustrate the variety of approaches. Woebot utilizes brief daily conversations to integrate CBT exercises into people's routines. Wysa is focused on AI-backed journaling and emotional resilience tools, and Youper integrates automated talk with mood analysis and brief therapy sessions. Convenience, user comfort, and removal of stigma that normally discourages people from regular therapy are at the center of all these apps.

3. Evaluating Effectiveness (500–600 words)

3.1 Quantitative Outcomes

Numerous studies indicate that chatbot interventions can produce statistically significant reductions in anxiety and depression symptoms. For example, randomized controlled trials of Woebot found large and statistically significant reductions in self-reported depression scores relative to control conditions in which participants received static psychoeducational materials; in a similar vein, studies of Youper show moderate-to-small effect sizes in symptom reductions for both anxiety and depression, with improvements observed in very short time periods.

Wysa has also been studied in targeted populations, like students and individuals living with chronic health conditions; in very brief engagements, studies report reductions in stress and improvements in coping skills. Although retention and engagement stats can vary by platform, the best chatbots report levels of sustained usage that compare favorably with and are as high as other digital health apps.

Overall, there are no universally strong signals in the data to date. Many studies have a short time span, small sample sizes, and there is often a heavy reliance on self-reported outcomes, which lends itself to bias. Lastly, even if reasons for symptom reductions were promising, there are few output measures of long-term maintenance of benefit without participant engagement.

3.2 Qualitative Feedback

User input contributes an important perspective to how we understand chatbot efficacy. Many users reported high levels of satisfaction stemming from the accessibility, non-judgmental nature, and simplicity of use of AI chatbots. For many users, since the interaction is anonymous, this diminished the fear of stigma and allowed for more emotional honesty to occur. Some users indicated they felt a level of companionship or support from the chatbot, especially when no other Resources were available.

Yet there are limitations. Although some chatbots are perceived as "empathetic" based on the language being developed by designers, they inherently cannot offer the depth of emotional attunement and complexity of understanding that human therapists can provide. These

limitations are especially pronounced in instances where the user's experience is complex and traumatic, there is high emotional volatility, or there is the need to create a strong therapeutic alliance. Additionally, inseparability remains an issue; chatbots tend to provide generic responses in instances where feedback needs to be nuanced or contextually sufficient.

3.3 Comparative Analysis

While AI chatbots have clear shortcomings in empathy, flexibility, and dealing with more complex mental health issues compared to human therapists, they also have some inherent benefits concerning scalability, affordability, and immediacy of access. Many experts argue that chatbots should be seen not as substitutes for human care but as adjuncts to human care. Chatbots can be a first point of contact, they can help provide some maintenance between therapy sessions, and they can be useful stand-alone options for individuals with moderate symptoms who might not otherwise seek professional help.

Indeed, the best potential for AI chatbots may be in hybrid models of care where both human and AI support identify strengths that can support a continuum of care. In these kinds of models, chatbots can be used for continued monitoring and general interventions or support, while complex cases can be escalated to professionals.

4. Ethical, Safety, and Regulatory Considerations

4.1 Limits and Potential Harms

One of the biggest fears is the inability of chatbots to effectively manage crises such as suicidal ideation or serious psychosis. While most platforms also feature crisis referral messages, they are typically scripted routes with little urgency (and often even less flexibility) in emergencies. Some systems have even poorly designed responses that could exacerbate distress, demonstrating the potential for harm through excess reliance on chatbots.

These limitations also illustrate new actor types, for example, "AI psychosis," where one can develop delusional attachment or misunderstand chatbots as offering clinical advice to questions. Another risk is the tendency of the chatbot to "sycophancy; validating almost every statement of the user, even when the user is expressing self-harming or other irrational thoughts. These consequences highlight the stark line most of these chat services draw between support through conversations and therapeutic care.

4.2 Ethical Issues

Ethical dilemmas are not restricted to clinical feasibility only. Data privacy is one of the most concerning issues. Mental health conversations can reveal private information, and users may not understand if or how the data is collected, preserved, or shared. Without rigorous protections in place, sensitive data may be disclosed or misappropriated, compromising user confidentiality.

Informed consent is another issue. Many users presume that these chatbots offer professional-quality therapy, despite disclaimers to the contrary. This misapprehension has the potential to cultivate false expectations and inhibit users from seeing a real therapist. The possibility of algorithmic bias adds another dimension: responses may differ according to the

cultural context, or the response could reflect an onboarding bias based on the beliefs codified in the training data.

4.3 Regulation and Oversight

Due to these issues, regulation is important. Indeed, some jurisdictions have already taken action by limiting unsupervised chatbot therapy and requiring attention to be paid to disclaimers denoting its non-clinical status; however, the global regulatory picture is one of inconsistency and unpreparedness.

There is a need to develop evaluation taxonomies and risk assessment frameworks for AI in mental health, with some experts calling for safety, transparency, and clinical validity benchmarks. Stronger oversight is essential to ensure the tools are being used responsibly and not causing inevitable harm. The paradox in addressing these issues lies in balancing innovation and accountability, ensuring access to chatbots to underserved populations, but simultaneously ensuring the user is not exposed to any form of clinical or ethical risk.

5. Future Directions and Recommendations

5.1 Hybrid Human-AI Models

Chatbots are best positioned as adjunctive tools in care pathways. For example, they could provide daily mood tracking, symptom monitoring, and low-intensity interventions, while clinicians review the data and provide deeper therapeutic engagement. Such models would harness the scalability of AI while ensuring that complex or high-risk cases are managed by professionals.

5.2 Improved Evaluation and Measurement

There is a clear need for more rigorous clinical trials with larger, more diverse populations. Current studies often rely on short-term measures, leaving questions about long-term effectiveness unanswered. Additionally, innovative evaluation approaches, such as simulated user testing and therapist-led assessments, could provide more nuanced insights into chatbot reliability and safety.

5.3 Emotional Intelligence and Personalization

One of the key limitations of current chatbots is their inability to convey authentic empathy. Research into AI “emotional intelligence” benchmarks may help systems better detect tone, sentiment, and urgency. Furthermore, advances in personalization, tailoring responses to an individual’s history, cultural context, and therapeutic needs, could enhance both engagement and outcomes.

5.4 Regulatory and Ethical Frameworks

Future development must be guided by robust ethical frameworks. Standardized protocols for informed consent, data governance, and safety escalation should be embedded into every chatbot platform. Regulatory bodies should establish clear guidelines for labeling, auditing, and certifying mental health chatbots, similar to medical device regulation. Transparent communication about limitations is essential so users understand what these tools can and

cannot provide.

6. Conclusion

AI chatbots are among the most exciting developments in digital mental health treatment, which can help provide cost-effective and stigma-free scalable support for those suffering from mild to moderate anxiety in the forms of anxiety, depression, stress, etc. Evidence from clinical trials and client feedback suggests that early versions of AI chatbots have indications that the effects will be to decrease symptom severity, increase emotionalization, and help connect marginalized demographics to treatment.

However, their success depends on careful design, continued engagement, and the infusion of evidence-based therapeutic approaches. More importantly, their limitations as compared to human services in the areas of crisis intervention, empathy, personalized attention, and ability to know responses are just not possible and therefore not a stand-alone solution as a substitute for what licensed practitioners offer. Additionally, ethical concerns associated with privacy, bias, and safety to clients are additional concerns that demand a cautious approach.

If AI chatbots are to be a useful part of the resources for providing mental health care, most optimistically, then hybrid human-AI models may be the best approach in which even chatbots are responsible for brief low-intensity interventions and monitoring, while making sure that licensed practitioners attend to the many associated components of quality human-client care that technology cannot replicate. Providing the industry with adequate regulatory scrutiny, employing good evaluation designs, and ensuring ethical safeguards, AI chatbots could become change agents in the global sharing of mental health support.

References

Farzan, M. A. (2025). Artificial intelligence–powered cognitive behavioral therapy chatbots: A systematic review. *Iranian Journal of Psychiatry, 20*(1), 1–12. https://doi.org/10.18502/ijps.v20i1.17395

Feng, Y., Li, P., & Zhao, H. (2025). Effectiveness of AI conversational agents in supporting young people's mental health: A meta-analysis. *JMIR mHealth and uHealth, 13*(1), e56742. https://doi.org/10.2196/56742

Kuhlmeier, S., Tan, J., & Weiss, A. (2025). Evaluating LLM-based mental health chatbots with simulated users and therapist-led benchmarks. *Frontiers in Artificial Intelligence, 8*, 014562. https://doi.org/10.3389/frai.2025.014562

Li, J., Theng, Y. L., & Foo, S. (2023). Artificial intelligence–based conversational agents for mental health: A systematic review and meta-analysis. *npj Digital Medicine, 6*(1), 63. https://doi.org/10.1038/s41746-023-00979-5

Mehta, S., Clarke, M., & Parekh, R. (2021). Acceptability and effectiveness of Youper, an AI-powered CBT chatbot, in reducing anxiety and depression symptoms. *Journal of Medical Internet Research, 23*(6), e26771. https://doi.org/10.2196/26771

Oghlan, S. (2023). Ethical issues with using AI chatbots in mental health. *Cambridge Quarterly of Healthcare Ethics, 32*(4), 456–468. https://pmc.ncbi.nlm.nih.gov/articles/PMC10291862/

Răsepar Meadi, A. (2025). Exploring the ethical challenges of conversational AI in mental health. *JMIR AI, 4*(2), e58123. https://doi.org/10.2196/58123

Rządeczka, M., Nowak, K., & Zieliński, T. (2025). Comparing therapeutic chatbots and general-purpose LLMs in cognitive bias correction. *JMIR Mental Health, 12*(1), e64396. https://doi.org/10.2196/64396

Steenstra, J., & Bickmore, T. (2025). A taxonomy of risks in AI psychotherapy chatbots. *Patient Education and Counseling, 119*(5), 1012–1019. https://doi.org/10.1016/j.pec.2025.02.004

Zhong, B. (2024). Effectiveness of AI-based chatbot interventions for depression and anxiety: Meta-analysis. *Journal of Affective Disorders, 354*, 128–139. https://doi.org/10.1016/j.jad.2024.02.015