

---

# Addressing the Emotional and Ethical Risks of Using AI Chatbots in Psychotherapy

---

<sup>1</sup>Abeera Saleem Mughal, <sup>1</sup>Arwa Jabeen, <sup>1</sup>Abiha Fatima, <sup>2</sup>Hafiz Shahbaz Zahoor

<sup>1</sup>Jinnah Sindh Medical University, <sup>2</sup>Quaid-e-Azam Medical College, Pakistan

<sup>2</sup>Corresponding Author Email: [muhammadshahbaz25261@gmail.com](mailto:muhammadshahbaz25261@gmail.com)

As technology continues to transform how we live, Artificial Intelligence (AI) chatbots have entered the space of mental health therapy with promises of accessibility and innovation. These digital companions offer hope for many people who face barriers to traditional care, whether due to cost, geography, or stigma. But with hope must come caution, and with innovation, responsibility [1].

At their best, AI chatbots can guide users through structured techniques or offer reminders rooted in cognitive behavioral therapy. Yet, they fall short in the most vital areas, those that touch the human heart. A chatbot cannot truly understand suffering. It cannot feel a pause in someone's voice or respond with compassion to a moment of shared vulnerability. Empathy, trust, and ethical wisdom remain uniquely human gifts. When someone struggling with depression or trauma turns to a chatbot, the risk is not just disappointment's harm [2].

Therapists have expressed valid hesitation. Their work is built not just on words, but on presence, intuition, and deep relational understanding. While AI evolves rapidly, clinical practice must prioritize safety, dignity, and care. Technology should never rush ahead of ethical reflection [2].

Equally important is the way we set expectations for users. If someone opens up to a chatbot expecting a healing dialogue and receives a canned or contextually off response, the damage can be more than disappointment; it can be retraumatization. It's crucial that both developers and clinicians ensure users understand the chatbot's role, limitations, and boundaries before any exchange begins. Transparency isn't just good design's an ethical responsibility [3].

AI's role in mental health must be anchored in strong safeguards: explicit privacy protections, explainable algorithms, and ongoing regulatory scrutiny. Tools should be designed with empathy in mind, not just efficiency. Most importantly, they must be guided by the voices of those with lived experience, patients, therapists, and caregivers [4].

In the end, digital tools may support mental health care, but they cannot replace the healing presence of a fellow human being. A blended model where AI assists but never replaces the human connection holds the most promise. Let us shape this future not only with code and logic but also with compassion, humility, and the unshakable belief that every person deserves to be seen and heard by another human soul [4].

**KEYWORDS:** *Artificial Intelligence, Artificial Intelligence in Healthcare*

Conflict of interest: None

Funding: None

## REFERENCES

1. Jesudason D, Bacchi S, Bastiampillai T. Artificial intelligence (AI) in psychotherapy: A challenging frontier. *Australasian Psychiatry*. 2025 Jan 6:10398562251346075.
2. Pham KT, Nabizadeh A, Selek S. Artificial intelligence and chatbots in psychiatry. *Psychiatric Quarterly*. 2022 Mar;93(1):249-53.
3. Sedlakova J, Trachsel M. Conversational artificial intelligence in psychotherapy: a new therapeutic tool or agent?. *The American Journal of Bioethics*. 2023 May 4;23(5):4-13.
4. Löchner J, Carlbring P, Schuller B, Torous J, Sander L. Digital interventions in mental health: An overview and future perspectives. *Internet Interventions*. 2025 Apr 2:100824.