AFMN BIOMEDICINE

REVIEW ARTICLE

EXPLORING USABILITY, SAFETY, AND PRELIMINARY EFFICACY OF NOVEL DIGITAL INTERVENTIONS IN ADOLESCENT PSYCHOTHERAPY: A NARRATIVE REVIEW

Miodrag Stanković[1,2] **Aleksandra Stojanović**[2]

[1]Department of Psychiatry, University of Niš Faculty of Medicine, Niš, Serbia [2]Center for Mental Health Protection, University Clinical Center Niš, Niš, Serbia

Artificial intelligence (AI) is increasingly recognized as a novel tool for enhancing psychotherapy, particularly for adolescents. The integration of AI-based platforms into mental health care promises improved accessibility, personalized interventions, and support for therapeutic processes. We aimed to evaluate the usability, safety, and preliminary effectiveness of a novel AI-assisted psychotherapy intervention in adolescent populations, with a focus on ethical considerations, user profiles, and limitations. We conducted a theoretical analysis of AI integration in psychotherapy, focusing on its potential application among adolescents, current ethical debates, and user patterns, particularly in post-pandemic contexts. Adolescents aged 16–25 years, who are highly immersed in digital environments, appear to be most open to using AI platforms for psychological support. The use of AI may enhance therapeutic access; however, limitations include a lack of emotional intelligence, reduced therapist involvement, and vulnerability to unethical data usage. Concerns also arise regarding the commodification of mental health through commercially driven AI applications. AI-supported psychotherapy for adolescents holds great potential but must remain adjunctive and ethically grounded. Further empirical research is necessary to ensure the safety, therapeutic efficacy, and ethical integrity of AI technologies in psychiatric care.

Keywords: artificial intelligence, digital psychotherapy, adolescents, mental health, ethics

Submitted: April 24, 2025 **Accepted:** August 11, 2025
Published online: December 17, 2025

Correspondence to:
Miodrag Stanković
Department of Psychiatry
University of Niš Faculty of Medicine
Bulevar dr Zorana Đinđića 81, Niš, Serbia
E-mail: miodrag.stankovic@medfak.ni.ac.rs

INTRODUCTION

Psychotherapy remains a primary psychological therapy aimed at assisting individuals with a wide range of mental states and disorders. As digital health technologies evolve, artificial intelligence (AI) has introduced new avenues for enhancing the accessibility, personalization, and scalability of psychotherapeutic care. Unlike traditional methods that rely on in-person sessions, AI-based solutions have shown promise in delivering structured, evidence-based interventions through digital interfaces, including mobile apps and conversational agents (1).

Role of AI in modern psychotherapy

Recent studies have demonstrated that AI-enabled tools can be particularly effective in managing symptoms of depression and anxiety by replicating key therapeutic mechanisms, such as cognitive restructuring, behavioral activation, and emotional support, through automated digital platforms (1-3). For instance, chat agents like Woebot and Tess have yielded promising outcomes in reducing psychological distress among young adults, providing consistent, immediate, and stigma-free support without requiring therapist involvement (1,3).

Despite the broader acceptance of AI in general medicine, its integration into mental healthcare remains in its early development stages. Many systems are still undergoing evaluation for clinical efficacy, ethical compliance, and long-term engagement (4-6). A key advantage of AI in psychotherapy is its ability to provide low-barrier interventions in underserved populations, particularly in low- and middle-income countries where mental health professionals are scarce (7).

Nevertheless, significant ethical considerations must be addressed before AI systems can be fully integrated into routine clinical practice, particularly in psychotherapy. These include concerns regarding data privacy, algorithmic bias, obtaining informed consent, and the preservation of patient autonomy and dignity (8,6,9,10). Furthermore, because AI lacks human empathy and nuanced clinical judgment, current guidelines recommend using it as a supportive tool rather than a replacement for human therapists (4,11).

Key milestones in the integration of AI into psychotherapy

The necessity for digital mental health interventions became particularly evident during the COVID-19 pandemic and associated global lockdowns. Movement restrictions, combined with a surge in demand for psychotherapeutic services and a simultaneous shortage of trained professionals, resulted in a mental health emergency on a global scale (12). Even high-income countries found themselves underprepared for the rising need. The integration of AI into mental health treatment, particularly through digital psychotherapy applications, could emerge as a promising solution for low- and middle-income countries, offering a more accessible and cost-effective method to deliver psychological support (13). Digital psychotherapy using AI can be more cost-effective, accessible, and available than traditional methods, providing a significant advantage in developing countries where access to licensed psychotherapists is often limited or expensive. The use of AI in mental health could enable the automation of therapy sessions, such as those employing acceptance and commitment therapy, which has shown efficacy in reducing symptoms of depression and anxiety, particularly among adolescents (13). Despite these promising prospects, the long-term efficacy of these digital therapies remains uncertain, especially in terms of maintaining therapeutic engagement and achieving sustained therapeutic outcomes. While early studies suggest improvements in treating depression and anxiety disorders, it is important to consider that these systems cannot replace human interaction and the personal empathy of a therapist (14).

AI and the diagnosis and treatment of mental disorders

The diagnostic process in psychiatry primarily relies on clinician-patient communication and behavioral observation. Psychiatric diagnoses are inherently interpretative, based on symptom clusters and consensus-driven diagnostic criteria. AI offers the potential to significantly improve diagnostic accuracy by objectively analyzing real-time data, including verbal and non-verbal communication, affective expression, and cognitive functioning (15, 16). Furthermore, AI systems may provide more consistent assessments of suicide risk (17).

On the therapeutic front, psychiatric treatment generally involves facilitating motivation, building adaptive coping strategies, modifying maladaptive habits, fostering psychological insight, and managing dysfunctional biological patterns, often through pharmacotherapy. The therapeutic relationship remains the primary agent of change, with the clinician's empathy and judgment being central to treatment success (18). However, AI-enhanced

systems may improve treatment outcomes, particularly in depression and anxiety disorders, by simulating therapeutic engagement and facilitating meaningful interaction (19). Recent studies suggest that AI tools like Wysa show promising results in enhancing digital mental well-being through conversational empathy. Nevertheless, it remains unclear whether such effects are sustainable over time, as this area of research is still in its early stages (20).

User profiles, ethical challenges, and evidence of AI usability

The rapid and irreversible advancement of artificial intelligence (AI) in mental health care raises complex ethical considerations. Core principles such as patient autonomy, equitable access to care, protection from discrimination, preservation of dignity, and data privacy must be thoroughly addressed before AI systems are integrated into clinical practice. Warrier et al. highlight the significance of these ethical aspects, emphasizing the need for privacy, impartiality, transparency, responsibility, and the physician–patient relationship in the context of AI in mental health (21). Alfano et al. emphasize the ethical implications of AI in psychotherapy for adolescents, underscoring the delicate balance between technological intervention and the preservation of the therapeutic alliance (22). Additionally, Yan et al. discuss the challenges of accurately recognizing mental disorders through AI applications, highlighting that AI's current limitations pose significant obstacles in clinical settings (23). Ciliberti et al. further explore the ethical dilemmas surrounding AI's role in caring relationships, emphasizing the importance of maintaining empathy and trust even in digital therapy environments (24).

Preliminary and still limited data suggest that adolescents and emerging adults represent the most frequent users of AI-based mental health services, particularly those aged between 16 and 25 years (25,26). This demographic cohort, having grown up in a technologically saturated environment, naturally incorporates digital tools into nearly all forms of communication. Feelings of physical isolation among these individuals are often mitigated by superficial but readily available virtual connections via smartphones and video conferencing. Furthermore, this generation spent a significant portion of its formative years under lockdown during the COVID-19 pandemic, relying heavily on remote and digital education systems (27). As a result, a behavioral pattern has developed in which young people exhibit a normalized comfort with digital platforms, even in contexts that traditionally rely on in-person interaction, such as psychotherapy.

This shift has led to an almost intuitive acceptance of AI-based mental health tools, where therapeutic encounters occur without direct physical presence. In contrast, adults, particularly those less fluent in emerging technologies, may struggle to communicate effectively with this cohort, increasing the risk of a generational disconnect. Nevertheless, active involvement of adults and professionals in guiding adolescents toward evidence-based digital interventions could yield long-term benefits (22). Encouraging the use of clinically validated AI tools may serve as a protective measure against reliance on unregulated, potentially harmful content that is often more accessible to young people. Introducing AI-assisted psychotherapeutic tools in schools, under the supervision of school psychologists, could help educational institutions provide much-needed mental health support to a generation that is already digitally literate. This approach may increase access to care while addressing the shortage of qualified professionals (28).

Ensuring that AI-based mental health systems comply with data protection regulations and patient privacy laws is an ethical imperative. Open questions remain regarding the storage and access of sensitive data, including session recordings: Where is the data stored? Who manages it? Who has access? Can third parties utilize the data, and under what conditions? Client privacy is especially vulnerable in digital ecosystems where personal data is often stored in cloud-based environments. These circumstances raise legitimate concerns about the trustworthiness of technology providers and their third-party collaborators. Furthermore, safeguarding client confidentiality requires robust encryption protocols and transparent governance mechanisms to ensure that sensitive data is protected and cannot be misused (28,29).

Online psychotherapy and the application of AI in psychotherapeutic practice

The advent of digital psychotherapy has significantly disrupted the traditional therapeutic "setting" based on face-to-face encounters bound by physical space and scheduled sessions. In digital formats, in-person attendance is no longer a prerequisite, and therapy becomes more flexible in terms of time and location. Therapists can now accommodate clients across various time zones, ensuring continuity of care regardless of geographical distance or personal travel. One of the most significant breakthroughs is

the decoupling of patients from their immediate local mental health resources, thus removing geographic limitations and expanding access to treatment (30,31).

Digital psychotherapy platforms enable individuals, especially those in remote or underserved areas, to engage in therapy that would otherwise be inaccessible due to professional shortages or logistical barriers. These innovations contribute to the democratization of mental health care, aligning with principles of justice and equity, and providing essential services to vulnerable populations who are typically underrepresented in traditional care systems (32,33).

For adolescents, digital therapy is often perceived as less stigmatizing and more approachable than traditional in-person models. AI interfaces do not evoke concerns about hidden agendas and can help reduce shame associated with discussing sensitive topics (34). Several AI-based applications designed for adolescents with depressive symptoms incorporate algorithms capable of identifying linguistic markers of suicidality based on user interactions with chatbots. These systems can recognize risk patterns through language use, historical user data, or clinician-provided metadata (35,36).

In some instances, AI has demonstrated a superior capacity to detect high-risk scenarios compared to even the most experienced therapists, due to its ability to process and analyze data instantaneously (37). Nonetheless, therapeutic responsibility must remain within the scope of human professionals, given ethical, legal, and clinical accountability (38,39). This highlights the importance of a hybrid model in which AI assists with data processing and risk assessment, while human therapists provide nuanced clinical judgment, empathy, and ethical oversight (40).

AI can serve as an auxiliary tool, flagging urgent situations in real time and alerting the supervising therapist, who remains actively involved in treatment decisions (41). The absence of human oversight may result in misinterpretation, misdiagnosis, or clinical harm, analogous to iatrogenic consequences (42). Patients must be fully informed about the role of AI in their treatment and its implications. Informed consent and patient autonomy must be upheld throughout the therapeutic process. Ensuring transparency about the scope and limitations of AI involvement is essential to maintaining therapeutic trust (43).

Limitations and critiques of AI use in psychotherapy

The application of artificial intelligence (AI) in psychotherapy has drawn considerable and warranted criticism. Firstly, many AI-based applications are developed within commercial frameworks, raising concerns about profit-driven motives. Suppose such platforms are designed to generate misleading claims or offer false promises of mental health improvement without empirical support. In that case, they may cause significant harm to users and hinder the future credibility of AI-assisted interventions (44).

Even when platforms operate with the highest ethical standards, therapy conducted with minimal clinician oversight still poses substantial risks. At present, AI cannot independently monitor patients or make clinical decisions. While algorithms are practical in pattern recognition, they cannot replace human intelligence. One of the most significant limitations of AI is its inability to perceive and process the emotional, symbolic, relational, and anthropological dimensions of human communication. AI operates strictly within empirical frameworks and lacks the capacity for empathy, a core element of therapeutic interaction (45).

For AI, the concept of empathy holds no evoked emotional significance. Words that may appear neutral to an algorithm can evoke strong emotional responses in humans, underscoring the importance of context and emotional resonance. In psychotherapy, language often carries meanings far beyond the literal, and such nuances are not interpretable by AI in their current form (46). Additional concerns include the lack of robust evidence supporting the diagnostic and therapeutic effectiveness of AI tools, as well as their failure to embody humanistic qualities necessary for meaningful engagement, such as self-reflection, professionalism, reliability, and the ability to recognize when a patient may be withholding or distorting the truth. These are all fundamental attributes of the psychotherapist, essential for ensuring ethical and effective mental health care (44). The absence of humanistic qualities increases the risk of oversimplification and misdiagnosis, particularly when clients engage in self-diagnosis without professional input (47).

The integration of artificial intelligence into adolescent psychotherapy represents a promising yet complex advancement in mental health care. While AI-assisted tools offer potential benefits in accessibility, scalability, and support for therapeutic interventions, they also introduce significant ethical, professional, and clinical concerns. The digital format may particularly appeal to younger populations accustomed to technological interaction; however, it cannot replace the empathic, symbolic, and relational depth characteristic of human therapeutic encounters. Adolescents, as digital natives, may benefit from supervised AI applications when these are integrated into existing mental health systems and guided by qualified professionals. Nonetheless, AI systems must be transparent, ethically designed, and subject to strict data protection standards. Without sustained clinical oversight and empirical validation of efficacy, AI-driven psychotherapy remains an adjunctive, rather than a substitutive tool in mental health care. Future research must address long-term outcomes, ethical standards, and the psychological integrity of digital interventions before they are fully integrated into psychiatric practice.

Acknowledgements

This study was not supported by any sponsor or funder.

Competing Interest

The authors declare no relevant conflicts of interest.

REFERENCES

1. 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, 6;4(2):e19. [CrossRef]

2. Kolding S, Lundin RM, Hansen L, Østergaard SD. Use of generative artificial intelligence (AI) in psychiatry and mental health care: a systematic review. Acta Neuropsychiatr 2024 11;37:e37. [CrossRef]

3. Fulmer R, Joerin A, Gentile B, et al. Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. JMIR Ment Health. 2018;5(4):e64. [CrossRef]

4. Vaidyam AN, Wisniewski H, Halamka JD, et al. Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. Can J Psychiatry 2019;64(7):456-464. [CrossRef]

5. Dehbozorgi R, Zangeneh S, Khooshab E, et al. The application of artificial intelligence in the field of mental health: a systematic review. BMC Psychiatry 2025;25(1):132. [CrossRef]

6. Inkster B, Sarda S, Subramanian V. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation. JMIR Mhealth Uhealth 2018;6(11):e12106. [CrossRef]

7. Mudiyanselage KWW, De Santis KK, Jörg F, et al. The effectiveness of mental health interventions involving non-specialists and digital technology in low-and middle-income countries - a systematic review. BMC Public Health. 2024;24(1):77. [CrossRef]

8. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. Artif Intell Med 2014;62(1):1-10. [CrossRef]

9. Rahsepar Meadi M, Sillekens T, Metselaar S, et al. Exploring the Ethical Challenges of Conversational AI in Mental Health Care: Scoping Review. JMIR Ment Health. 2025 Feb 21;12:e60432. [CrossRef]

10. American Psychiatric Association. Ethical Principles of Psychologists and Code of Conduct. 2017. https://www.apa.org/ethics/code

11. World Health Organization. Ethics and Governance of Artificial Intelligence for Health. Geneva: WHO; 2021.

12. Kola L, Kohrt BA, Hanlon C, et al. COVID-19 mental health impact and responses in low-income and middle-income countries: reimagining global mental health. Lancet Psychiatry. 2021;8(6):535-550. [CrossRef]

13. Vertsberger A, Shani L, Naor N, et al. Digital mental health

interventions for adolescents: a systematic review of applications using artificial intelligence. J Med Internet Res. 2022;24(1):e36018.

14. Cross S, Bell I, Nicholas J, et al. Use of AI in Mental Health Care: Community and Mental Health Professionals Survey. JMIR Ment Health. 2024 11:e60589. [CrossRef]

15. Gooding P. Mapping the rise of digital mental health technologies: Emerging issues for law and society. Int J Law Psychiatry 2019;67:101498. [CrossRef]

16. Bzdok D, Meyer-Lindenberg A. Machine learning for precision psychiatry: Opportunities and challenges. Biol Psychiatry Cogn Neurosci Neuroimaging 2018;3(3):223-30. [CrossRef]

17. Shatte ABR, Hutchinson DM, Teague SJ. Machine learning in mental health: A scoping review of methods and applications. Psychol Med 2019;49(9):1426-48. [CrossRef]

18. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. J Med Internet Res 2019;21(5):e13216. [CrossRef]

19. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation. JMIR Mhealth Uhealth 2018;6(11):e12106. [CrossRef]

20. Topol EJ. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. Basic Books; 2019.

21. Warrier U, Warrier A, Khandelwal K. Ethical considerations in the use of artificial intelligence in mental health. Egypt J Neurol Psychiatry Neurosurg 2023;59:139. [CrossRef]

22. Alfano L, Malcotti I, Ciliberti R. Psychotherapy, artificial intelligence and adolescents: ethical aspects. J Prev Med Hyg. 2024 Jan 1;64(4):E438-E442. [CrossRef]

23. Yan WJ, Ruan QN, Jiang K. Challenges for Artificial Intelligence in Recognizing Mental Disorders. Diagnostics (Basel) 2022;13:2. [CrossRef]

24. Schiavone V, Alfano L, Ciliberti R. Ethics and humanization of care: Reflections in the teaching of French institutional psychotherapy experience. J Prev Med Hyg 2023;64(2):E198-E203. [CrossRef]

25. Beg MJ, Verma M, Chanthar KM M, Verma MK. Artificial Intelligence for Psychotherapy: A Review of the Current State and Future Directions. Indian J. Psychol. Med. 2024;47(4):314–325. [CrossRef]

26. Tavory T. Regulating AI in Mental Health: Ethics of Care Perspective. JMIR Ment Health. 2024;11:e58493. [CrossRef]

27. Vertsberger D, Naor N, Winsberg M. Adolescents' Well-being While Using a Mobile Artificial Intelligence-Powered Acceptance Commitment Therapy Tool: Evidence From a Longitudinal Study. JMIR AI 2022;1(1):e38171. [CrossRef]

28. D'Alfonso S, Santesteban-Echarri O, Rice S, et al. Artificial Intelligence-Assisted Online Social Therapy for Youth Mental Health. Front Psychol 2017;8:796. [CrossRef]

29. Ienca M, Vayena E. Mental data protection and the GDPR. J Law Biosci 2022;9(1):lsac001. [CrossRef]

30. McGorry PD, Mei C, Dalal N, et al. The Lancet Psychiatry Commission on youth mental health. Lancet Psychiatry 2024;11(9):731-74. [CrossRef]

31. Alhuwaydi AM. Exploring the Role of Artificial Intelligence in Mental Healthcare: Current Trends and Future Directions - A Narrative Review for a Comprehensive Insight. Risk Manag Healthc Policy. 2024;17:1339-1348. [CrossRef]

32. Lehtimaki S, Martic J, Wahl B, et al. Evidence on Digital Mental Health Interventions for Adolescents and Young People: Systematic Overview. JMIR Ment Health. 2021;8(4):e25847. [CrossRef]

33. Spytska L. The use of artificial intelligence in psychotherapy: development of intelligent therapeutic systems. BMC Psychol 2025;13:175. [CrossRef]

34. Rogan J, Bucci S, Firth J. Health Care Professionals' Views on the Use of Passive Sensing, AI, and Machine Learning in Mental Health Care: Systematic Review With Meta-Synthesis. JMIR Ment Health. 2024;11:e49577. [CrossRef]

35. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry. 2021;20(3):318-335. [CrossRef]

36. Gkintoni E, Vassilopoulos SP, Nikolaou G, Boutsinas B. Digital and AI-Enhanced Cognitive Behavioral Therapy for Insomnia: Neurocognitive Mechanisms and Clinical Outcomes. J Clin Med. 2025;14(7):2265. [CrossRef]

37. Chen J, Yuan D, Dong R, et al. Artificial intelligence significantly facilitates development in the mental health of college students: a bibliometric analysis. Front Psychol 2024;15:1375294. [CrossRef]

38. Harishbhai Tilala M, Kumar Chenchala P, Choppadandi A, et al. Ethical Considerations in the Use of Artificial Intelligence and Machine Learning in Health Care: A Comprehensive Review. Cureus. 2024;16(6):e62443. [CrossRef]

39. Marshall NJ, Loades ME, Jacobs C, et al. Integrating Artificial Intelligence in Youth Mental Health Care: Advances, Challenges, and Future Directions. Curr Treat Options Psych. 2025;12(1):11. [CrossRef]

40. Vial T, Almon A. Artificial Intelligence in Mental Health Therapy for Children and Adolescents. JAMA Pediatr. 2023;177(12):1251-2. [CrossRef]

41. Dobbs MF, McGowan A, Selloni A, et al. Linguistic correlates of suicidal ideation in youth at clinical high-risk for psychosis. Schizophr Res 2023;259:20-27. [CrossRef]

42. Bloch-Atefi A. Balancing ethics and opportunities: The role of AI in psychotherapy and counselling. Psychother Couns J Aust. 2025;13(1). [CrossRef]

43. Saeidnia HR, Hashemi Fotami SG, Lund B, Ghiasi N. Ethical Considerations in Artificial Intelligence Interventions for Mental Health and Well-Being: Ensuring Responsible Implementation and Impact. Social Sciences 2024; 13(7):381. [CrossRef]

44. Wykes T, Lipshitz J, Schueller SM. Towards the design of ethical standards related to digital mental health and all its applications. World Psychiatry 2023;22(1):110-11. [CrossRef]

45. Salles A, Evers K, Farisco M. Anthropomorphizing AI: Ethical and societal implications. AI & Society. 2020;35:611-622.

46. Høeg TB, Gammon D, Strand M. Empathy in digital therapy: What can artificial intelligence really understand? Journal of Medical Internet Research Mental Health 2021;8(3):e27878.

47. Kretzschmar K, Tyroll H, Pavarini G, et al.; NeurOx Young People's Advisory Group. Can Your Phone Be Your Therapist? Young People's Ethical Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health Support. Biomed Inform Insights 2019;11:1178222619829083. [CrossRef]