



Kent Academic Repository

Khan, Nagina (2026) *Bridging Science and Subjectivity: Evolving Evidence, Emerging Technologies and the Call for Personalised Psychotherapy: Créer une passerelle entre la science et la subjectivité : Évolution des données probantes, arrivée de nouvelles technologies et appel à une psychothérapie personnalisée.* The Canadian Journal of Psychiatry . pp. 1-9.

Downloaded from

<https://kar.kent.ac.uk/113345/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1177/07067437261425840>

This document version

Publisher pdf

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Bridging Science and Subjectivity: Evolving Evidence, Emerging Technologies and the Call for Personalised Psychotherapy

The Canadian Journal of Psychiatry /
La Revue Canadienne de Psychiatrie

1-9

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/07067437261425840

TheCJP.ca | LaRCP.ca



Créer une passerelle entre la science et la subjectivité : Évolution des données probantes, arrivée de nouvelles technologies et appel à une psychothérapie personnalisée

Nagina Khan, PhD, SFHEA, PGCert, BHSc, RMN^{1,2,3} 

Abstract Psychotherapy Research in the Era of Digital Health and AI

The integration of digital health technologies and artificial intelligence (AI) into psychotherapy research represents a transformative shift in mental health care. Traditional psychotherapy relies heavily on face-to-face interactions, limiting accessibility, scalability, and opportunities for continuous monitoring of patient progress. Emerging digital tools, including teletherapy platforms, mobile applications, virtual reality (VR) and AI-powered conversational agents are increasingly being evaluated as adjuncts or alternatives to conventional therapy, opening new frontiers in both research and clinical practice.

Résumé Recherche en Psychothérapie à L'ère de la Santé Numérique et de L'IA

L'intégration des technologies de santé numérique et de l'intelligence artificielle (IA) dans la recherche en psychothérapie représente un changement transformateur dans les soins en santé mentale. La psychothérapie classique repose fortement sur les interactions en personne; ce qui limite l'accessibilité, l'élargissement de la portée et les possibilités de surveiller en continu les progrès réalisés par les patients. On évalue de plus en plus les nouveaux outils numériques, y compris les plateformes de télépsychothérapie, les applications mobiles, la réalité virtuelle (RV) et les agents conversationnels alimentés par l'IA, à titre d'appoints ou de solutions de rechange à la thérapie conventionnelle; ce qui ouvre de nouveaux horizons en recherche et en matière de pratique clinique.

Keywords

psychotherapy, cognitive behavioural therapy, evidence-based practice, neuroscience, digital mental health, personalized therapy, mental health research

Psychotherapy Research in the Era of Digital Health and AI

The rapid expansion of digital health infrastructure has fundamentally reshaped the landscape of psychotherapy research. While psychotherapy has historically been grounded in in-person therapeutic encounters, contemporary mental health systems increasingly incorporate technology-mediated care models. Advances in telecommunication platforms, mobile health applications, immersive virtual reality

(VR), and artificial intelligence (AI) driven systems are not merely extending traditional therapy delivery but are redefining how interventions are developed, evaluated, and personalized. These innovations introduce new methodological opportunities, including continuous symptom monitoring, large-scale data integration, and adaptive treatment models, while simultaneously challenging long-standing assumptions about therapeutic processes and clinical boundaries. Key domains of innovation in digital and AI-enabled psychotherapy research are summarised in Figure 1.

¹ Centre for Health Services Studies (CHSS), School of Social Sciences, University of Kent, Canterbury, UK

² The National Institute of Health and Care Research (NIHR), Applied Research Collaborative Kent (ARC KSS), Research Inclusion, Surrey, Sussex, UK

³ NIHR INSIGHT, MSc Applied Health and Care Research Programme, Canterbury, Kent, UK

Corresponding author:

Nagina Khan, PhD, SFHEA, PGCert, BHSc, RMN, Centre for Health Services Studies (CHSS), School of Social Sciences, University of Kent, Canterbury, Kent, CT2 7NZ, UK.

Email : N.Khan-523@kent.ac.uk

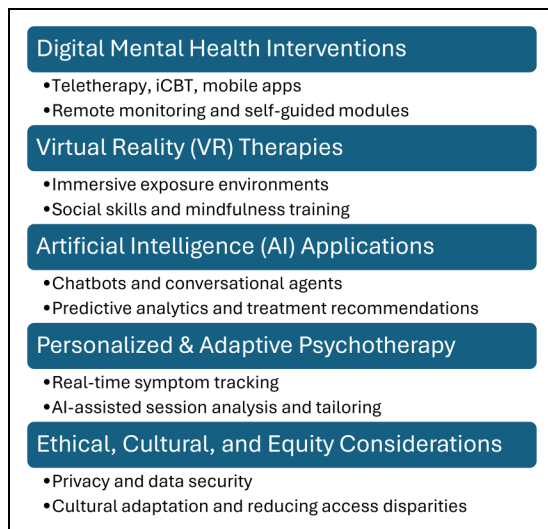


Figure 1. Key domains of innovation in digital and AI-enabled psychotherapy research.

Digital Mental Health Interventions

The rise of digital mental health interventions has revolutionised the accessibility, delivery, and research of psychotherapy. Teletherapy and online platforms have become integral to mental health care, particularly following the COVID-19 pandemic, which accelerated the adoption of remote therapy services worldwide. These platforms offer synchronous (real-time video or chat) and asynchronous (messaging or self-guided modules) therapeutic formats, enabling continuity of care even during periods of social distancing or restricted mobility.

Internet-based cognitive behavioural therapy (iCBT) has been extensively studied and consistently demonstrates effectiveness across multiple mental health conditions, including anxiety disorders, depression, obsessive-compulsive disorder (OCD), and post-traumatic stress disorder (PTSD).^{1,2} Structured online programs often include psychoeducation, interactive exercises, symptom tracking, and homework assignments, replicating core elements of face-to-face CBT while allowing patients to work at their own pace. Research indicates that iCBT not only reduces symptoms but can also improve treatment adherence, especially when supported by periodic therapist guidance, blending automated delivery with human oversight.³

Mobile mental health applications have expanded the reach of psychotherapy further, offering tools for self-monitoring, mood tracking, and behavioural activation. Apps can provide reminders, cognitive restructuring exercises, mindfulness practices, and immediate coping strategies, reinforcing therapeutic techniques between sessions. These interventions are particularly valuable for populations with limited access to traditional therapy, such as individuals in rural areas, those with mobility limitations, or patients facing financial constraints.⁴

An increasingly important dimension of digital interventions is their potential for cultural adaptation and trauma-informed delivery. Traditional psychotherapies have historically reflected Eurocentric norms, often limiting their relevance for diverse populations. Digital interventions can incorporate culturally relevant content, language options, and narratives that resonate with specific communities, enhancing engagement and therapeutic outcomes.⁵⁻⁷ For instance, culturally adapted CBT programs for South Asian, Latino, or Indigenous populations have been shown to improve both adherence and efficacy by addressing culturally specific beliefs about mental health, stigma, and family involvement.⁸

Digital interventions also provide unprecedented opportunities for research and data collection. Smartphone sensors, wearable devices, and app analytics enable ecological momentary assessment, real-time monitoring of symptoms, and continuous feedback loops between patient and clinician. This granular data can inform personalised interventions, identify early signs of relapse, and contribute to predictive models of treatment response, bridging the gap between clinical research and real-world application.⁹ Despite these advantages, challenges remain. Digital literacy, socioeconomic disparities, and limited access to reliable internet or devices can exacerbate existing health inequities. Engagement and long-term adherence are also variable, with attrition rates often higher than in traditional therapy. Moreover, ethical concerns around privacy, data security, informed consent, and algorithmic bias must be addressed to ensure safe and equitable delivery of digital mental health care.¹⁰

Digital mental health interventions represent a paradigm shift in psychotherapy, offering scalable, flexible, and potentially personalised treatment options. When integrated thoughtfully with therapist support, cultural sensitivity, and robust ethical safeguards, these tools can enhance access, engagement and outcomes, marking a significant evolution in both clinical practice and psychotherapy research. See Table 1. Contemporary Advances in Digital and AI-Augmented Psychotherapy.

Interpersonal Psychotherapy (IPT) also exemplifies how digitally supported approaches can illuminate core therapeutic mechanisms. Digital behavioural data and computational modelling can track communication patterns, role transitions, and affective dynamics central to IPT, offering new ways of operationalising and monitoring these traditionally human delivered processes. This illustrates how digital tools can augment, not replace, established therapeutic models and sets the stage for comparing how different psychotherapies translate into digital and AI-enabled contexts.

Long-Term Comparative Outcomes of CBT and Psychodynamic/Psychoanalytic Therapy

Evidence from long-term comparative studies provides important context for psychotherapy research prior to the era of digital health. In a German three-year follow-up study, patients receiving either cognitive-behavioural therapy (CBT) or

Table 1. Contemporary Advances in Digital and AI-Augmented Psychotherapy.

Digital Mental Health Interventions	Virtual Reality (VR) Interventions	AI Applications in Psychotherapy
Online therapy and structured iCBT deliver evidence-based treatment remotely. ^{11–13}	VR exposure therapy enables controlled, immersive simulations for phobias, PTSD, and anxiety. ¹⁴	Conversational agents (e.g., Woebot) automate CBT skills and mood monitoring. ¹⁰
Mobile apps support self-monitoring, psychoeducation, and symptom tracking. Promotes access for rural, marginalised, and mobility-restricted groups.	Enhances engagement and motivation through immersive environments. Allows graded exposure that can be tailored safely.	Predictive analytics identify risk of relapse, non-response or suicidality. ¹⁵ Analyses large datasets from therapy transcripts, EHRs, and digital behaviour patterns.
Supports culturally tailored, trauma-informed interventions. ^{5,6}	Future work needed on long-term outcomes and cost-effectiveness.	Offers personalised treatment recommendations and adaptive interventions.

psychodynamic/psychoanalytic therapy for depression showed broadly comparable symptomatic outcomes at 3-year follow-up, with CBT demonstrating some advantages in efficiency and structured delivery.¹⁶ Similarly, a recent randomised controlled trial in psychiatric outpatient clinics comparing short-term psychodynamic psychotherapy (STPP) with CBT for major depression found no significant differences in depressive symptom reduction, anxiety, psychosocial functioning, or quality of life at follow-up.¹⁷

These findings suggest that, across different durations and settings, CBT and psychodynamic/psychoanalytic therapies may offer roughly equivalent efficacy for depression.¹⁸ While these studies evaluated conventional treatment delivery, they provide a foundation for integrating AI-driven personalisation and digital health interventions. AI technologies can enhance these traditional modalities by continuously monitoring patient progress, identifying early signs of relapse, and tailoring interventions to individual symptom trajectories. For example, digital platforms could help determine whether a patient might benefit more from CBT or psychodynamic approaches based on real-time data, optimising both clinical effectiveness and resource allocation.

The long-term comparative evidence highlighting similar efficacy of CBT and psychodynamic approaches highlight that multiple therapeutic modalities can achieve meaningful clinical outcomes. Digital tools and AI offer the potential to extend these interventions, enabling scalable delivery, continuous monitoring, and personalised adjustments without compromising the core mechanisms that drive efficacy, such as therapeutic alliance, exposure processes, and cognitive restructuring. By integrating empirical insights from traditional therapies with digital and AI innovations, researchers and clinicians can optimise treatment delivery, enhance accessibility and develop data-informed models for personalised psychotherapy.

Virtual Reality (VR) Interventions

Virtual reality (VR) represents a rapidly growing and innovative domain within psychotherapy research, offering

immersive, controlled environments that can be precisely tailored to therapeutic goals. VR-based interventions are particularly prominent in exposure therapy, where patients are gradually and safely exposed to anxiety- or trauma-related stimuli in a virtual setting. This method allows therapists to replicate real-world scenarios, such as flying for phobia treatment, social situations for social anxiety, or combat simulations for PTSD, without exposing patients to actual risk.¹⁴

Evidence to date indicates that VR exposure therapy can be as effective as, or in some cases more engaging than, traditional in vivo or imaginal exposure therapies. The immersive quality of VR can enhance emotional and physiological responses, facilitating more effective habituation to feared stimuli.¹⁹ Additionally, VR allows for fine-grained control over the intensity, duration, and context of exposures, enabling therapists to tailor interventions precisely to the patient's tolerance and treatment goals. Interactive features, such as real-time feedback, gamification, and patient-controlled pacing, further improve engagement, motivation and treatment adherence.²⁰

Beyond exposure therapy, VR is being explored in broader psychotherapeutic contexts, including social skills training, cognitive rehabilitation, mindfulness practice, and stress reduction. For example, VR-based social simulations can provide safe practice environments for individuals with social anxiety or autism spectrum disorders, supporting skill acquisition and confidence-building. Mindfulness and relaxation-focused VR interventions use immersive nature scenes, guided meditations, or biofeedback integration to reduce stress and improve emotional regulation.²¹ The integration of VR with hybrid approaches is a promising area for future research. For instance, VR can be combined with traditional cognitive behavioural therapy (CBT), biofeedback, or digital mental health tools to create multi-modal interventions that enhance efficacy while maintaining flexibility and accessibility. Early feasibility studies suggest that these hybrid models may optimise outcomes by combining the advantages of immersive experiential learning with established therapeutic frameworks.²²

Despite these promising findings, several challenges remain. The cost of VR hardware and software, the need

for clinician training, and potential cybersickness or sensory discomfort for some patients are important considerations. Long-term efficacy and generalisation of VR-acquired skills to real-world situations are still under investigation, as is the scalability of VR interventions in routine clinical practice. Ethical and privacy considerations, particularly regarding the collection of sensitive biometric or behavioural data during VR sessions, also require careful attention.¹⁴ VR interventions represent a transformative frontier in psychotherapy research, offering immersive, engaging, and adaptable therapeutic experiences. When combined with robust clinical frameworks and thoughtful consideration of patient specific factors, VR has the potential to expand access, enhance outcomes, and complement traditional therapeutic approaches in innovative ways.

Bridging Traditional Psychotherapy Mechanisms With AI and Digital Psychiatry

Contemporary AI and digital systems in psychiatry often replicate, extend, or operationalise core mechanisms that have long been foundational to psychotherapy traditions. However, these connections are rarely articulated. Cognitive, psychodynamic, interpersonal, and experiential models all rely on structured processes, such as case formulation, interpretation, behavioural experimentation, affect regulation, and alliance repair, that map closely onto emerging computational tools.

Cognitive-behavioural models align with AI systems that provide automated feedback, cognitive restructuring, and structured symptom monitoring using natural language processing and pattern detection. Psychodynamic and relational traditions emphasise affective attunement and meaning reconstruction, which resonate with research on AI-supported reflective dialogue, vocal affect recognition, and automated detection of alliance ruptures. Interpersonal psychotherapy's focus on role transitions and communication patterns parallels computational mapping of relational dynamics in digital interactions.

By situating digital technologies within these therapeutic mechanisms, AI becomes not a departure from psychotherapy but an evolution of its core processes. This conceptual integration helps ensure that digital systems remain clinically coherent, mechanism-based, and aligned with established pathways of therapeutic change.

Artificial Intelligence (AI) Applications in Psychotherapy Research

AI is increasingly shaping the landscape of psychotherapy research, offering tools that enhance both clinical practice and empirical investigation. AI applications in this field include conversational agents (chatbots), predictive analytics, and personalised treatment recommendation systems, each contributing uniquely to mental health care and research.

Conversational agents, such as Woebot, Wysa, and Tess, use natural language processing to deliver structured, evidence-based interventions like cognitive behavioural therapy (CBT), acceptance and commitment therapy (ACT) and problem-solving strategies. These AI-driven tools can monitor patient mood, provide psychoeducation, suggest coping strategies, and offer real-time support outside traditional therapy sessions. Early studies have demonstrated their effectiveness in reducing symptoms of mild depression and anxiety, improving engagement and treatment adherence and supporting psychoeducation in large populations.^{10,23} By offering scalable, low-cost support, chatbots can complement human-delivered psychotherapy, addressing service gaps in underserved or geographically isolated communities.

Predictive analytics represents another powerful application of AI. Machine learning algorithms can analyse complex datasets from electronic health records, digital mental health interventions, and longitudinal studies to identify patients at risk of non-response, relapse, or suicidal behaviour. This enables clinicians and healthcare systems to intervene proactively, optimising treatment planning and resource allocation. Predictive models also facilitate early identification of treatment-resistant or high-risk populations, supporting personalised care approaches and potentially reducing the burden of chronic mental health conditions.¹⁵

AI can also enhance personalised treatment recommendation systems, which integrate patient-specific information, such as demographics, symptom profiles, treatment history, and real-time digital phenotyping, to suggest the most effective interventions. These systems can guide therapists in tailoring treatment plans, selecting optimal therapy modalities, and adjusting intervention intensity. Over time, AI-driven learning algorithms can continuously refine recommendations based on patient outcomes, supporting dynamic, data-informed psychotherapy.

In addition to these clinical applications, AI serves as a transformative tool in psychotherapy research. Large-scale data analysis enables the identification of patterns, predictors and mediators of treatment outcomes across diverse populations and therapy modalities. For instance, AI can analyse language use, engagement metrics, and behavioural data from digital interventions to uncover subtle indicators of progress or relapse. This can inform both theory development and the design of more targeted and adaptive psychotherapeutic interventions. However, several challenges and ethical considerations accompany the integration of AI in psychotherapy. Issues include data privacy and security, algorithmic bias, and the risk of over-reliance on automated systems. Ensuring that AI interventions are transparent, evidence-based, and culturally sensitive is critical to avoid exacerbating disparities in mental health care.

Furthermore, AI tools should complement, rather than replace, the therapeutic alliance, which remains a central determinant of treatment outcomes. AI applications in psychotherapy research offer unprecedented opportunities to

scale interventions, personalise care, and generate insights from complex datasets. When implemented thoughtfully, they can augment traditional therapy, support initiative-taking clinical decision-making and contribute to a more data-informed, patient-centred future in mental health care.

Existing reviews often describe digital and AI interventions independently from the psychotherapeutic traditions that shaped the mechanisms they draw upon. The present manuscript contributes by explicitly linking AI capabilities to established models of therapeutic change, cognitive, affective, interpersonal, and experiential, which offers a theoretically grounded framework for the development of future AI-enabled tools. This conceptual bridge clarifies how psychotherapy's long-standing mechanisms can inform computational design and prevent AI interventions from becoming disconnected from clinical theory.

Expanding on AI-Based Therapy Beyond Mild Conditions

While early research on AI-based therapy focused largely on mild to moderate conditions, a growing body of work demonstrates broader applicability. A recent systematic review of AI-powered CBT chatbots reported significant reductions in depression and anxiety symptoms across platforms such as Woebot, Wysa, and Youper, along with high user engagement and satisfaction.²⁴ Another meta-analysis focused on young people found that chatbot-delivered interventions significantly reduced psychological distress, suggesting feasibility and benefit in younger and potentially more diverse populations.²⁵ Beyond static chatbots, recent research has begun to explore large language model (LLM)-based conversational agents and their therapeutic potential. For example, a 2025 randomised controlled trial found that a 28-day intervention using an LLM-based agent significantly reduced both depression and anxiety symptoms in young adults.²⁶ Early evidence also suggests that integrating AI-based tools into mental health services can improve clinical efficiency, reduce wait times, lower dropout rates and improve allocation to appropriate treatments, offering scalable support beyond traditional care settings.²⁷ These developments indicate that AI-enabled therapy is evolving beyond short-term, low-intensity use offering scalable, adaptive and potentially clinically robust support across a wider spectrum of severity and user needs.

Digital Tools and AI in Personalised Psychotherapy

Digital tools and AI technologies are revolutionising the delivery and personalisation of psychotherapy, offering unprecedented opportunities to optimise treatment for individual patients. By continuously collecting and analysing data on symptom trajectories, engagement patterns, and

patient preferences, these tools allow interventions to be dynamically adapted in real time. Such adaptability ensures that therapy remains responsive to the unique needs of each client, rather than relying on static, one-size-fits-all treatment protocols.

For instance, AI-assisted analysis of session transcripts, voice tone, and natural language use can detect subtle signs of therapeutic rupture or disengagement. These insights enable therapists to implement timely repair strategies, strengthening the therapeutic alliance and enhancing treatment efficacy, which is consistently shown to be one of the most significant predictors of positive outcomes.^{28,29} Also, AI can identify patterns of response across clients with similar profiles, helping clinicians anticipate potential obstacles, tailor interventions, and select evidence-based techniques most likely to succeed for a given individual.

Digital platforms also facilitate scalability and accessibility, addressing long-standing inequities in mental health care. Online therapy, mobile apps, and AI-driven chatbots can reach populations who face barriers to traditional in-person care, including rural communities, individuals with mobility limitations, or those in marginalised social and economic contexts.^{6,30} These platforms can deliver culturally adapted content, language-specific interventions and trauma-informed care, thereby to promote inclusivity while maintaining treatment fidelity.

Additionally, continuous monitoring through digital tools enables therapists to track progress longitudinally, providing actionable feedback on adherence, engagement, and symptom improvement. This data-informed approach not only strengthens therapist decision-making but also empowers patients to take an active role in their treatment, promoting self-efficacy and sustained behavioural change. Over time, AI systems can refine intervention strategies based on aggregated outcomes, creating a feedback loop that enhances both research knowledge and clinical practice.

The integration of digital tools and AI in psychotherapy represents a paradigm shift toward personalised, data-driven and accessible mental health care. By enabling real-time adaptation, enhancing therapeutic alliance, and expanding reach to underserved populations, these technologies have the potential to transform both clinical outcomes and equity in mental health services.

Bridging Psychotherapy Traditions and AI

Traditional psychotherapy frameworks, including cognitive behavioural, psychodynamic, and interpersonal approaches, provide mechanistic insights that can inform the design of AI-assisted interventions. For example, models of cognitive restructuring and behavioural activation underpin algorithmic decision-making in AI-supported CBT, while psychodynamic concepts of affect regulation and relational patterns can guide the development of chatbots capable of recognising emotional cues and tailoring responses to relational

Table 2. Key Domains of Innovation in Digital and AI-Enabled Psychotherapy Research.

Domain	Examples	Research Implications
Digital Mental Health	Teletherapy, iCBT, mobile apps	Accessibility, adherence, remote monitoring
Virtual Reality (VR)	Exposure therapy, social simulations	Immersion, engagement, generalisation
AI Applications	Chatbots, predictive models	Personalisation, risk detection
Personalised Therapy	EMA, digital phenotyping	Dynamic treatment adaptation
Ethical and Equity Issues	Data privacy, cultural adaptation	Safe, fair, inclusive implementation

contexts.¹⁰ By leveraging these theoretical foundations, AI systems can not only replicate evidence-based techniques but also generate novel hypotheses about treatment mechanisms, creating a bidirectional flow between human clinical expertise and computational modelling.

Challenges and Limitations of Digital and AI-Driven Psychotherapy

Despite their considerable promise, digital and AI-driven approaches to psychotherapy face a range of significant challenges and limitations that must be carefully addressed. First, the evidence base for AI and digital interventions is still emerging, particularly for complex psychiatric conditions such as severe depression, psychotic disorders, or comorbid mental health presentations. Most research to date has concentrated on mild to moderate symptoms, raising questions about generalisability and clinical applicability for patients with higher acuity or multifaceted needs.^{10,15}

A core concern is the potential impact on the therapeutic alliance, which remains one of the strongest predictors of treatment outcome across modalities. Fully automated interventions, such as chatbots or algorithmically guided modules, risk attenuating the relational depth and empathic responsiveness that human therapists provide.^{28,31} While AI can augment clinical decision-making, over-reliance on technology may inadvertently reduce opportunities for real-time attunement, reflective listening, and relational repair, which are critical for meaningful therapeutic change.

Ethical and regulatory issues also present significant hurdles. Data privacy and security are paramount, as digital platforms collect sensitive personal information, often stored in cloud-based systems that may be vulnerable to breaches. Ensuring informed consent is more complex in AI-mediated interventions, particularly when algorithms continuously adapt treatment strategies without explicit human oversight. Additionally, biases inherent in AI models, stemming from training datasets that may underrepresent minority populations, can exacerbate inequities in care and risk reinforcing systemic disparities.⁶

Implementation challenges extend beyond ethics to practical considerations. See Table 2, Key domains of innovation in digital and AI-enabled psychotherapy research. Digital literacy varies widely across populations, and individuals from socioeconomically disadvantaged backgrounds may lack access to devices or reliable internet connections, limiting engagement.

Cultural acceptability is another critical factor; interventions developed within Western paradigms may not align with the values, beliefs, or communication styles of diverse client populations, necessitating careful adaptation to ensure relevance and efficacy.⁵

An emerging concern is the use of large-language-model (LLM) chatbots in populations at risk of self-harm or severe psychiatric symptoms. Although AI-based chatbots show promise for delivering therapeutic content and supporting engagement, recent empirical work reveals serious safety limitations when handling crisis situations. A recent evaluation of 29 popular mental-health chatbots tested using standardised suicidal-risk scenarios based on the Columbia-Suicide Severity Rating Scale (C-SSRS) found that none delivered consistently adequate crisis responses; nearly half were rated as ‘inadequate,’ often failing to provide emergency contact information or clinician referral.³² Similarly, a mixed-methods study comparing LLM-based chatbots to human therapists demonstrated substantial gaps: most chatbots lacked robust risk-assessment protocols, failed to refer users to crisis hotlines appropriately, or provided overly reassuring responses that clinicians judged inadequate.³³ These findings emphasise that, without rigorous safeguarding, LLM-based tools risk mismanaging suicidal ideation or self-harm disclosures, a considerable limitation if chatbots are used as standalone or first-line support.

Given these risks, deployment of AI-mediated psychotherapy must proceed with caution. Any implementation should include clear crisis-management protocols, transparent communication about the tool’s limitations, routine human oversight, and mechanisms to ensure timely referral to professional care in high-risk situations. Until such safeguards are standardised and empirically validated, LLM-based therapy should not replace human-delivered care, especially for individuals with severe or unstable mental health conditions.

Until such safeguards are standardised and empirically validated, LLM-based therapy should not replace human-delivered care, especially for individuals with severe or unstable mental health conditions. Beyond immediate safety considerations, the long-term effectiveness, sustainability, and cost-effectiveness of digital and AI-based therapies remain underexplored. At the same time, initial studies show promising short-term outcomes.

Access, Scalability, and Stepped-Care Models

A growing challenge in contemporary psychotherapy is not the efficacy of established treatments but their limited

accessibility and scalability. Despite a broad range of evidence-based psychotherapies, many patients face barriers including cost, long wait times, and reliance on traditional one-to-one formats that are difficult to deliver at population scale. Health systems such as the UK National Health Service (NHS) have attempted to address this through stepped-care models, in which patients first receive low-intensity interventions, such as guided self-help workbooks, digital modules, email-supported programs, or group-based interventions, and only progress to higher-intensity individual therapy if needed. This structure allows more efficient allocation of clinical resources while maintaining treatment responsiveness. Integrating emerging digital tools and AI-driven monitoring into stepped-care frameworks further enhances scalability, enabling real-time symptom tracking, early detection of non-response, and dynamic routing toward more intensive interventions. Such models illustrate how personalisation and scalability can coexist, offering a pathway to expand access without compromising treatment quality.

Future Directions in Digital and AI-Enhanced Psychotherapy

Forthcoming research in digital and AI-enhanced psychotherapy should emphasise hybrid models that integrate traditional therapist-delivered interventions with AI-supported and digital tools. Such models leverage the strengths of human empathy, clinical judgment, and relational engagement while taking advantage of AI's ability to personalise care, monitor progress in real time, and scale interventions efficiently. Investigating the optimal balance between human and digital components will be crucial to maximise therapeutic outcomes and engagement across diverse patient populations.

A parallel and equally urgent priority is improving access and scalability. Much evidence-based psychotherapies remain difficult to obtain due to workforce shortages, cost barriers, and reliance on traditional one-to-one formats. Integrating stepped-care models, such as those employed by the NHS, where low-intensity interventions (e.g., workbooks, digital modules, brief guided support) precede more intensive therapy, illustrates how digital tools and AI-driven systems can expand capacity while maintaining treatment quality. Embedding AI into these frameworks offers the potential for dynamic triage, early detection of non-response, and adaptive routing toward higher-intensity care, thereby enhancing both efficiency and equity in service delivery.

Longitudinal studies are essential to evaluate sustained treatment effects, adherence over time, and cost-effectiveness. Current evidence largely focuses on short-term outcomes, leaving gaps in understanding the durability of digital interventions, their potential to prevent relapse, and

their utility in complex or comorbid mental health conditions. Future trials should also assess differential effectiveness across symptom severity, age groups, and sociocultural contexts to identify which interventions are most beneficial for whom.

Integrating digital tools with culturally responsive and trauma-informed care can increase both accessibility and relevance. Tailoring content to reflect cultural norms, language, and lived experiences enhances engagement, addresses disparities in access, and promotes equity in mental health care. Moreover, trauma-informed approaches embedded in digital interventions can reduce the risk of re-traumatisation and improve the therapeutic experience. The development of advanced measurement and monitoring tools represents another critical frontier. Techniques such as ecological momentary assessment (EMA), digital phenotyping, wearable biosensors, and neuroimaging can provide granular, real-time data on symptom fluctuations, treatment adherence, and behavioural patterns. This data-driven approach allows for more precise personalisation of interventions, identification of early warning signs of relapse, and improved understanding of the mechanisms underlying therapeutic change.

The convergence of AI, digital health and psychotherapy offers a transformative opportunity to expand access, enhance personalisation, and optimise outcomes. However, these advances must be accompanied by rigorous ethical oversight, robust data privacy protections, and strategies to mitigate bias, ensuring that technology support equitable and responsible mental health care. By addressing these challenges, future research can create an integrated digital ecosystem that complements human expertise, maximises therapeutic efficacy, and promotes sustainable improvements in population mental health.

Acknowledgements

The author thanks Professor and Chairman Emeritus of the Department of Psychiatry, Allen Frances, and Professor Marvin Goldfried, Distinguished Professor of Clinical Psychology at Stony Brook University, for their valuable feedback on earlier drafts of this manuscript. The author also acknowledges the support of the NIHR Allied Research Collaborative Kent, Surrey and Sussex (ARC KSS).


Declaration of Conflicting Interests

The author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ORCID iD

Nagina Khan  <https://orcid.org/0000-0003-3870-2609>

References

- Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. *World Psychiatry*. 2014;13(3):288–295.
- Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. *Cognit Ther Res*. 2012;36(5):427–440.
- Cuijpers P, Karyotaki E, Weitz E, Andersson G, Hollon SD, Van Straten A. The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. *J Affect Disord*. 2014;159:118–126.
- Torous J, Myrick KJ, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: Using technology today to accelerate the curve on access and quality tomorrow. *JMIR Ment Health*. 2020;7(3):1–2.
- Gone JP. A community-based treatment for native American historical trauma: prospects for evidence-based practice. *Spirituality in Clinical Practice*. 2013;1(S):78–94.
- Naeem F, Khan N, Sohani N, et al. Culturally adapted cognitive behaviour therapy (CaCBT) to improve community mental health services for Canadians of south Asian origin: a qualitative study. *Can J Psychiatry* [Internet]. 2024 Jan 1 [cited 2024 Oct 18];69(1):54–68. <https://pubmed.ncbi.nlm.nih.gov/37376808/>.
- Khan N, Butt MG, Awan F, et al. Global mental health commentary: using innovation to create a workforce to deliver and implement culturally adapted CBT in Pakistan. *BMJ Mental Health* [Internet]. 2025 Mar 4 [cited 2025 Mar 7];28(1):e301389. <https://mentalhealth.bmj.com/content/28/1/e301389>.
- Naeem F, Khan N, Ahmed S, et al. Development and Evaluation of Culturally Adapted CBT to Improve Community Mental Health Services for Canadians of South Asian Origin: Final Report [Internet]. Toronto, Ontario; 2023. <https://www.camh.ca/-/media/files/cacbt-for-south-asians-report-eng.pdf>
- 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 Oct 1;20(3):318–335.
- 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(2).
- Berger T, Andersson G. Internet-Based Treatment of Depression: A Randomized Controlled Trial Comparing Guided with Unguided Self-Help Negative effects of Internet-based cognitive behavior therapy View project The use of a self-guided online program for depression (Deprexis) in Brazil View project. 2011 [cited 2018 Oct 25]. <http://dx.doi.org/10.1080/16506073.2011.616531>
- Andersson S, Magnusson L, Hanson E. The use of information and communication technologies to support working carers of older people - a qualitative secondary analysis. *Int J Older People Nurs*. 2016.
- Andersson G, Carlbring P, Furmark T, Buhrman M. Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial Treatment of Premenstrual Syndrome-Internet-based Self-help, ClinicalTrials.gov Identifier: NCT01961479 View project BASIC project View project. *J Consult Clin Psychol* [Internet]. 2006 [cited 2018 Oct 25];74(4). <http://www.random.org>.
- Maples-Keller JL, Bunnell BE, Kim SJ, Rothbaum BO. The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harv Rev Psychiatry*. 2017;25(3):103–113.
- Nawaz FA, Riaz MMA, Tsagkaris C, et al. Impact of #PsychTwitter in promoting global psychiatry: a hashtag analysis study. *Front Public Health* [Internet]. 2023;11. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emexb&NEWS=N&AN=640576586>.
- Huber D, Zimmermann J, Henrich G, Klug G. Comparison of cognitive-behaviour therapy with psychoanalytic and psychodynamic therapy for depressed patients – A three-year follow-up study. <https://doi.org/10.13109/zptm.2012583299> [Internet]. 2014 Apr 29 [cited 2025 Nov 28];58(3):299–316. [doi/pdf/10.13109/zptm.2012.58.3.299?download=true](https://doi.org/10.13109/zptm.2012.58.3.299?download=true).
- Malkomsen A, Wilberg T, Bull-Hansen B, et al. Comparative effectiveness of short-term psychodynamic psychotherapy and cognitive behavioral therapy for major depression in psychiatric outpatient clinics: a randomized controlled trial. *BMC Psychiatry*. 2025 Feb 11 [cited 2025 Nov 28];25(1):113. <https://link.springer.com/article/10.1186/s12888-025-06544-6>.
- Egger N, Konnopka A, Beutel ME, et al. Long-term cost-effectiveness of cognitive behavioral therapy versus psychodynamic therapy in social anxiety disorder. *Depress Anxiety* [Internet]. 2016 Dec 1 [cited 2025 Nov 28];33(12):1114–1122. [doi/pdf/10.1002/da.22540](https://doi.org/10.1002/da.22540)
- Carl E, Stein AT, Levihn-Coon A, et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. *J Anxiety Disord*. 2019;61:27–36.
- Riva G, Serino S. Virtual reality in the assessment, understanding and treatment of mental health disorders. *J Clin Med*. 2020;9(11).
- Ladakakis I, Filios D, Chouvarda I. Virtual reality environments for stress reduction and management: a scoping review. *Virtual Real*. 2024 Feb 27 [cited 2025 Nov 28];28(1):50. <https://link.springer.com/article/10.1007/s10055-024-00943-y>.
- Gerardi M, Cukor J, Difede J, Rizzo A, Rothbaum BO. Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorders. *Curr Psychiatry Rep* [Internet]. 2010 [cited 2025 Nov 28];12(4):298–305. <https://pubmed.ncbi.nlm.nih.gov/20535592/>.
- Inkster B, Sarda S, Subramanian V. An Empathy-Driven. Conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation. *J Med Internet Res* [Internet]. 2020;22(8):e17158. <https://www.jmir.org/2020/8/e17158/>.

24. Farzan M, Ebrahimi H, Pourali M, Sabeti F. Artificial intelligence-powered cognitive behavioral therapy chatbots, a systematic review. *Iran J Psychiatry* [Internet]. 2025 [cited 2025 Nov 28];20(1):100–108. <https://pubmed.ncbi.nlm.nih.gov/40093525/>.
25. Li J, Li Y, Hu Y, et al. Chatbot-Delivered Interventions for Improving Mental Health Among Young People: A Systematic Review and Meta-Analysis. *Worldviews Evid Based Nurs* [Internet]. 2025 Aug 1 [cited 2025 Nov 28];22(4). <https://pubmed.ncbi.nlm.nih.gov/40662463/>.
26. Zhao Y, Qian W, Chen Y, et al. Effect of an AI agent trained on a large language model (LLM) as an intervention for depression and anxiety symptoms in young adults: A 28-day randomized controlled trial. *Appl Psychol Health Well Being* [Internet]. 2025 Oct 1 [cited 2025 Nov 28];17(5). <https://pubmed.ncbi.nlm.nih.gov/40910958/>.
27. Rollwage M, Habicht J, Juechems K, et al. Using conversational AI to facilitate mental health assessments and improve clinical efficiency within psychotherapy services: real-world observational study. *JMIR AI* [Internet]. 2023 Dec 13 [cited 2025 Nov 28];2(1):e44358. <http://www.ncbi.nlm.nih.gov/pubmed/38875569>.
28. Horvath AO, Del Re AC, Flückiger C, Symonds D. Alliance in individual psychotherapy. In: *Psychotherapy Relationships That Work: Evidence-Based Responsiveness*. 2011.
29. Flückiger C, Del AC, Wampold BE, Horvath AO. The alliance in adult psychotherapy: a meta-analytic synthesis. *Psychotherapy* [Internet]. 2018;55(4):316–340. <https://www.embase.com/search/results?subaction=viewrecord&from=export&id=L622287473%0Ahttp://dx.doi.org/10.1037/pst0000172>.
30. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry*. 2014;13(1).
31. Wampold BE, Imel ZE. The great psychotherapy debate: the evidence for what makes psychotherapy work: second edition. 2nd ed. In: *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work: Second Edition*. Vol. 53. Routledge; 2015:1–323.
32. Pichowicz W, Kotas M, Piotrowski P. Performance of mental health chatbot agents in detecting and managing suicidal ideation. *Sci Rep* [Internet]. 2025 Dec 1 [cited 2025 Nov 28];15(1). <https://pubmed.ncbi.nlm.nih.gov/40866537/>.
33. Scholich T, Barr M, Wiltsey Stirman S, Raj S. A comparison of responses from human therapists and large language model-based chatbots to assess therapeutic communication: mixed methods study. *JMIR Ment Health* [Internet]. 2025 May 21 [cited 2025 Nov 28];12(1):e69709. <http://www.ncbi.nlm.nih.gov/pubmed/40397927>.