International Journal of Innovative Technologies in Social Science

e-ISSN: 2544-9435

Scholarly Publisher
RS Global Sp. z O.O.
ISNI: 0000 0004 8495 2390

Dolna 17, Warsaw,
Poland 00-773
+48 226 0 227 03
editorial_office@rsglobal.pl

ARTICLE TITLE	DIGITAL TRANSFORMATIONS IN THERAPY: A COMPREHENSIVE REVIEW OF ARTIFICIAL INTELLIGENCE AND VIRTUAL REALITY-BASED INTERVENTIONS IN THE TREATMENT OF ANXIETY AND DEPRESSIVE DISORDERS (2020–2024)
DOI	https://doi.org/10.31435/ijitss.4(48).2025.4197
RECEIVED	18 October 2025
ACCEPTED	26 December 2025
PUBLISHED	30 December 2025
LICENSE	The article is licensed under a **Creative Commons Attribution 4.0 International License.**

DIGITAL TRANSFORMATIONS IN THERAPY: A COMPREHENSIVE REVIEW OF ARTIFICIAL INTELLIGENCE AND VIRTUAL REALITY-BASED INTERVENTIONS IN THE TREATMENT OF ANXIETY AND DEPRESSIVE DISORDERS (2020–2024)

Dagmara Gładysz *(Corresponding Author, Email: dagmaragladys226@gmail.com)*
The University of Rzeszów, Rzeszów, Poland
ORCID ID: 0009-0001-4237-1382

Joanna Barwacz
The University of Rzeszów, Rzeszów, Poland
ORCID ID: 0009-0002-8805-4961

Magdalena Adamik
The University of Rzeszów, Rzeszów, Poland
ORCID ID: 0000-0002-9272-3059

Layla Settaf-Cherif
Ludwik Rydygier Collegium Medicum, Nicolaus Copernicus University, Bydgoszcz, Poland
ORCID ID: 0009-0007-6891-0456

Katarzyna Malinowska
Ludwik Rydygier Collegium Medicum, Nicolaus Copernicus University, Bydgoszcz, Poland
ORCID ID: 0009-0009-4757-382X

Elhatra Settaf-Cherif
Poznan University of Medical Sciences, Poznan, Poland
ORCID ID: 0009-0001-5444-4227

ABSTRACT

Research Objectives: This comprehensive systematic review and meta-analysis aim to critically evaluate the clinical evidence regarding the efficacy, feasibility, and cost-effectiveness of Artificial Intelligence (AI)- and Virtual Reality (VR)-based interventions in the treatment and monitoring of Anxiety Disorders (AD) and Major Depressive Disorder (MDD) between 2020 and 2024.
The study focuses on two primary pillars of innovation: AI tools (Generative AI chatbots, natural language processing, crisis prediction) and VR technologies (Exposure Therapy, VR-enhanced Behavioral Activation, mindfulness training, bio/neurofeedback). Additionally, the review analyzes ethical, legal, social, and implementation (ELSI) barriers that must be addressed to achieve large-scale deployment.
Methods: A scoping review of the scientific literature was conducted across databases including PubMed, Scopus, and Google Scholar, applying inclusion criteria that covered Randomized Controlled Trials (RCTs), systematic reviews, and meta-analyses published from January 2020 to July 2024. Seven key meta-analyses and eighteen RCTs were analyzed in detail, focusing on effect sizes (Hedges' g). Risk of bias and regulatory frameworks were also assessed.
Main Findings: AI demonstrated high scalability ($g = 0.52$-0.61 for chatbots (Li et al., 2023)), showing short-term efficacy comparable to traditional digital interventions. A key factor was the integration of Generative AI models that enhanced the therapeutic alliance, as well as NLP-based predictive systems. VR, in turn, provided an experiential depth (presence) essential for effectiveness in specific phobias and social anxiety ($g = 0.75$-0.88 (Wang et al., 2024)), particularly in VRET and VR-BA models. Major barriers included issues of Explainable AI (XAI), algorithmic bias, and technological integration with healthcare systems (EHR).
Conclusions: Digital technology is becoming an integral part of psychiatric care. The future lies in blended care models that combine the scalability of AI with the immersive depth of VR under clinical supervision. Long-term RCTs and clear regulatory frameworks concerning ethics and legal liability must be established prior to widespread implementation.

KEYWORDS

Artificial Intelligence in Mental Health, Virtual Reality, Digital Therapy, Anxiety Disorders, Major Depressive Disorder, Therapeutic Chatbots

CITATION

Dagmara Gładysz, Joanna Barwacz, Magdalena Adamik, Layla Settaf-Cherif, Katarzyna Malinowska, Elhatra Settaf-Cherif. (2025). Digital Transformations in Therapy: A Comprehensive Review of Artificial Intelligence and Virtual Reality-Based Interventions in the Treatment of Anxiety and Depressive Disorders (2020–2024). *International Journal of Innovative Technologies in Social Science*. 4(48). doi: 10.31435/ijitss.4(48).2025.4197

1. Introduction

1.1. Global Burden of Mental Disorders: Epidemiological and Economic Context

Anxiety Disorders (AD) and Major Depressive Disorder (MDD) currently represent the leading cause of disability worldwide, measured in Disability-Adjusted Life Years (DALYs), with an estimated 400 million people affected by AD and 280 million by MDD (World Health Organization, 2022).

This public health crisis also carries enormous economic consequences. According to OECD and WHO reports, the direct and indirect costs of AD and MDD in developed countries (e.g., the United States, EU) account for approximately 4-5% of GDP, mainly due to absenteeism, reduced productivity (*presenteeism*—working while unwell), and the costs of somatic treatment (World Health Organization, 2022).

This economic context underscores the urgent need for *cost-effective* solutions. Moreover, the COVID-19 pandemic and its social repercussions dramatically intensified the global prevalence of these conditions, increasing anxiety and depression rates by an estimated 25% in 2020 (World Health Organization, 2022).

This surge simultaneously exposed critical deficiencies in access to traditional, in-person care — such as long waiting lists, a shortage of qualified clinicians, and geographical barriers. As a result, the years 2020-2024 became a turning point for the development and validation of digital solutions designed to overcome these barriers by providing scalable, lower-cost care.

1.2. The Evolution of Digital Technologies in Mental Health (Digital Therapeutics - DTx)

The search for scalable and accessible solutions has led to the rapid development of **Digital Therapeutics (DTx)** — software-based interventions with clinically proven efficacy in preventing, managing, or treating diseases. Whereas early *digital mental health* applications were limited to simple psychoeducational modules, the past few years have witnessed an evolution toward two advanced pillars: **Artificial Intelligence (AI)** and **Virtual Reality (VR)**.

1. **Artificial Intelligence (AI):**

There has been a transition from rule-based systems to advanced Deep Learning and Generative AI (GenAI) architectures (e.g., Transformers), revolutionizing interaction quality. AI is now applied not only in chatbots delivering CBT or DBT but also in Natural Language Processing (NLP) for risk monitoring and in Reinforcement Learning (RL) algorithms for the personalization of treatment pathways (*Dynamic Treatment Regimes, DTR*).

2. **Virtual Reality (VR):**

VR technology has evolved from bulky laboratory systems of the 1990s to lightweight, mobile Head-Mounted Displays (HMDs), enabling large-scale deployment. VR leverages the sense of *presence* to create controlled exposure (VRET) or activation (VR-BA) environments, offering a unique capability to simulate real-life anxiety-inducing or social scenarios with adjustable intensity.

1.3. Rationale for the Time Frame (2020-2024)

The focus on literature published between January 2020 and July 2024 is not coincidental but reflects the convergence of three pivotal factors that marked technological and clinical maturity:

1. **Technological Maturity of HMDs:** The mass availability and cost reduction of lightweight VR headsets enabled large-scale clinical studies beyond laboratory settings.
2. **The Generative AI Revolution:** The introduction of advanced GenAI models (e.g., GPT-3/4 since 2020-2021) allowed chatbots to achieve unprecedented fluency and simulated empathy, leading to higher acceptance and better RCT outcomes.
3. **Post-Pandemic Impetus:** The COVID-19 crisis compelled healthcare systems to rapidly adopt digital tools, triggering immediate growth in funding and the number of RCTs validating these interventions.

1.4. Research Questions

This review aims to address the following key research questions:

1. What is the clinical efficacy (effect size, Hedges' *g*) of AI (chatbots, NLP) and VR (VRET, VR-BA) interventions compared with standard care or waitlist control in treating AD and MDD, based on RCTs published between 2020 and 2024?
2. Which technological factors (GenAI, immersion, *presence*) are most critical for clinical efficacy and patient adherence?
3. What are the main ethical, legal, and social barriers (ELSI) to large-scale deployment, including bias, explainability (XAI), privacy, and legal liability?
4. How can new *blended care* models optimally integrate the scalability of AI with the experiential depth of VR to deliver a cost-effective and coherent therapeutic pathway?

2. Methodology

2.1. Search Strategy and Inclusion/Exclusion Criteria

Type of Review:

This study adopts the format of a comprehensive *systematic review* with elements of *meta-analysis* (specifically in the domain of AI chatbots).

Databases:

A systematic search was conducted in the following scientific databases: PubMed/MEDLINE, Scopus, Web of Science, IEEE Xplore, and Google Scholar.

Temporal and Language Criteria:

Only articles published in English or Polish between **January 1, 2020, and July 31, 2024**, were included to ensure focus on the most recent scientific advancements.

Inclusion Criteria (Clinical):

- Randomized Controlled Trials (RCTs)
- Systematic Reviews and Meta-Analyses
- Pilot Studies assessing feasibility, acceptability, and tolerability
- Clinical populations diagnosed with or exhibiting symptoms of Anxiety Disorders (including PTSD, Social Anxiety Disorder) and Major Depressive Disorder (MDD)
- Interventions utilizing AI (chatbots, NLP, predictive algorithms) or VR (VRET, VR-BA, VR-Mindfulness, VR-Bio/Neurofeedback)

Exclusion Criteria:

- Case reports, letters to the editor, or incomplete conference papers
- Studies based solely on paper-based tests or mobile apps without AI/VR components
- Papers unrelated directly to MDD, AD, or PTSD

2.2. Characteristics of the Selected Literature

The initial screening identified **seven key publications**, which formed the analytical core of this review, complemented by **an additional 18 sources** obtained through extended queries, resulting in a total of **25 studies** for detailed analysis.

The key sources include:

- Meta-analyses of AI chatbots (Li et al., 2023; Casu et al., 2024)
- RCTs investigating VR-Enhanced Behavioral Activation (VR-BA) (Paul et al., 2022; Paul et al., 2024)

- Studies on Virtual Reality Exposure Therapy (VRET) (Zainal et al., 2021)
- A comprehensive review on VR in psychotherapy (Wang et al., 2024)

The literature analysis was structured according to intervention type:

- **AI-based interventions** (Section V)
- **VR-based interventions** (Section VI)

3. Results: AI-Based Interventions

AI-based interventions in mental health primarily focus on the scalable delivery of psychoeducation and therapy (via chatbots) as well as on early detection and risk monitoring (through NLP and predictive algorithms). Between 2020 and 2024, the dominant trend was a transition from rule-based models to **Generative AI (GenAI)**, which fundamentally changed the dynamics of human-machine therapeutic interaction.

3.1. Therapeutic Chatbots and Virtual Agents: Efficacy and GenAI

Conversational Agents (CAs), commonly known as **chatbots**, are the most accessible and widely studied form of AI-based intervention in mental healthcare. They are designed to deliver automated or semi-automated therapeutic protocols, typically based on **Cognitive Behavioral Therapy (CBT)**, **Behavioral Activation (BA)**, or **skills training** (Casu et al., 2024).

3.1.1. Clinical Efficacy of Chatbots (Meta-Analysis)

Meta-analytic reviews from 2023-2024 confirm a **moderate-to-large effect size** of chatbot interventions in reducing symptoms of **Major Depressive Disorder (MDD)** and **Anxiety Disorders (AD)** (Li et al., 2023; Casu et al., 2024).

A systematic review and meta-analysis of 15 randomized controlled trials (RCTs) (Li et al., 2023) published before May 2023 found that AI-based conversational agents had a statistically significant and positive impact on mental health outcomes:

- **Depression:** The meta-analysis reported a moderate-to-large effect size in reducing depressive symptoms (Hedges' $g = 0.64$, 95% CI 0.17-1.12), indicating effectiveness compared with control groups (e.g., waitlist, standard care without digital intervention).
- **Distress:** A significant reduction in overall psychological distress was also observed (Hedges' $g = 0.70$, 95% CI 0.18-1.22).

Effect sizes for chatbot interventions versus controls (waitlist or treatment-as-usual) ranged between **g = 0.52-0.61** for depressive symptoms (PHQ-9) and **g = 0.61-0.70** for distress and anxiety (GAD-7) over the short term (up to 12 weeks). This suggests that an average chatbot user fared better than approximately **69-76%** of participants in the control groups (Li et al., 2023).

A more recent review focusing on feasibility and use-cases confirmed that chatbots can serve effectively as **first-step** or **adjunctive** interventions (Casu et al., 2024).

Efficacy was especially evident for **subclinical** and **mild-to-moderate** symptoms. Their main advantages lie in scalability and geographic accessibility, positioning them as effective **stepped-care tools** that relieve pressure on clinical systems.

3.1.2. Factors Influencing Chatbot Effectiveness

Meta-analyses identify several factors that enhance efficacy (Li et al., 2023):

- **Multimodality:** Agents offering interactions beyond plain text (e.g., graphics, voice, interactive exercises) achieved better outcomes.
- **Generative AI (GenAI):** Newer, GenAI-based chatbots providing more natural and personalized responses outperform older rule-based systems. Their superiority stems from improved *therapeutic alliance* and perceived empathy (Richards et al., 2023).
- **App Integration:** Chatbots embedded within popular messaging platforms or dedicated mental-health apps show higher *adherence* and better clinical outcomes.
- **Population Differences:** Stronger effects were observed in clinical/subclinical populations and among older adults—groups facing greater barriers to traditional care.

3.1.3. The Role of Generative AI (GenAI) in the Therapeutic Alliance

The most transformative technological shift has been the adoption of Transformer-based Generative AI models (e.g., GPT), which revolutionized chatbots' ability to simulate fluent, empathic dialogue, directly affecting:

- **Therapeutic Alliance:** While the alliance with AI does not equal that with a human therapist, newer models can elicit a strong sense of supportive and comforting presence (Casu et al., 2024). This relational capacity is critical for psychological intervention success.
- **Adherence and Engagement:** Higher conversational quality and response coherence in GenAI systems substantially increase intervention completion rates compared to rule-based systems. Engagement exceeding **60%** of users completing at least four sessions (Li et al., 2023) predicts stronger clinical outcomes. The drop-out problem typical for digital therapies is partly mitigated by the naturalness of GenAI dialogues.
- **CBT Delivery:** Chatbots effectively deliver key CBT techniques such as identifying automatic thoughts, cognitive restructuring, and planning Behavioral Activation (BA) tasks. GenAI dynamically adjusts tone and pacing to the user's emotional state, enhancing personalization.

3.1.4. Limitations and Challenges in Chatbot Use

Despite promising findings, chatbot interventions face several critical challenges:

- **Limited Depth:** While effective for psychoeducation and skills training, chatbots lack the flexibility and depth required for complex conditions, trauma, or suicidality (Casu et al., 2024). Human escalation is essential during crises.
- **Adherence Drop-off:** Completion rates remain modest across all digital interventions. Even with GenAI, therapy discontinuation remains a major issue.
- **Physiopathological Impact:** Evidence on long-term durability and neuropsychological mechanisms of improvement remains inconsistent.
- **Automation Bias:** Patients may over-trust AI recommendations (Richards et al., 2023), posing ethical risks, especially during crises. Overreliance on chatbot guidance can delay seeking professional help when immediate intervention is required.

3.2. NLP-Based Risk Prediction and Crisis Detection

Between 2020 and 2024, research on **Natural Language Processing (NLP)** and **speech signal analysis** for early detection of suicidal risk and depressive episodes accelerated significantly. NLP models are employed for **passive monitoring** of text and speech, enabling early prediction of mood deterioration and suicidal ideation.

Research trends can be summarized into three pillars:

1. Text analysis of social-media content
2. Acoustic analysis of speech and prosodic features
3. Multimodal fusion of text + voice + behavioral data

Systematic reviews and empirical studies show promising results, though standardization, generalizability, and ethical issues remain key limitations (Montejo-Ráez et al., 2024; Abdulsalam & Alhothali, 2024).

3.2.1. Evolution of NLP Models (Transformers and Multimodality)

From 2020 to 2024, there was a massive shift from simple lexical and frequency-based models to advanced **Transformer architectures** (Kaminsky et al., 2024).

- **Early Phase (2020-2021):** Classical ML models (SVM, Random Forest) using lexical features and *n-grams* analyzed risk in Twitter or Reddit posts.
- **Deep Learning Era (2021-2024):** Dominance of LSTM and Transformer-based models (BERT, RoBERTa), fine-tuned for clinical risk classification tasks.

BERT and Variants:

Trained on massive corpora (e.g., social media, support forums) and fine-tuned for clinical use—such as classifying suicidal ideation or self-harm risk.

Trajectory Modeling:

Some approaches model temporal post sequences, capturing cumulative linguistic shifts (e.g., hopelessness, social withdrawal). For example, the SAIPH system identified long-term risk patterns across user timelines (Kaminsky et al., 2024).

3.2.2. Voice as a Biomarker of Depression and Suicidal Risk

Beyond textual cues, **acoustic features** represent valuable diagnostic signals. Recent studies focus on vocal biomarkers such as:

- Pitch (F0)
- Jitter and shimmer (amplitude/frequency variation)
- Speech rate and pause duration
- Spectral formant characteristics

These parameters correlate with psychomotor retardation and anhedonia—core markers of depression. Studies from 2022-2023 show that AI models analyzing speech recordings can differentiate individuals with depressive or suicidal symptoms with **moderate-to-high AUC** scores, depending on dataset and method (Min et al., 2023). Even brief teleconsultation audio samples can support clinical triage. Changes such as monotone delivery or slowed tempo are known markers that AI can quantify objectively.

3.2.3. Multimodal Data Fusion

Recent research (2023-2024) highlights progress toward **multimodal models** combining textual, acoustic, and behavioral data (e.g., log-in frequency, social contact metrics). This integration enhances sensitivity and specificity compared with single-modality approaches.

2024 reviews emphasize that **temporal fusion**—tracking changes over time—significantly improves detection of prodromal signs (Abdulsalam & Alhothali, 2024).

3.2.4. Ethical Challenges: Bias, Reliability, and Privacy

The use of AI for risk prediction introduces severe **ethical challenges** related to **algorithmic bias** (Ueda et al., 2023).

- **Training Data Bias:** Models trained predominantly on data from majority populations (e.g., English-speaking men from high-income countries) perform worse in minority groups, leading to false negatives with potentially catastrophic clinical consequences.
- **Data Representativeness:** Public social-media datasets do not reflect all demographics or private communication patterns, limiting clinical generalizability.
- **False Alarms and Clinical Utility:** High false-positive rates may overload crisis-intervention systems, and evidence that early algorithmic detection improves outcomes remains limited (Kaminsky et al., 2024).
- **Privacy:** Passive monitoring of speech or social activity raises profound concerns over privacy, informed consent, and data anonymization—particularly critical when involving suicidal individuals.

3.3. Algorithms for Personalized Therapy and the Role of XAI

The future of digital mental-health treatment lies in **personalization**. AI algorithms are increasingly applied not only for diagnostics but also for optimizing and individualizing treatment trajectories—transitioning from diagnostic to **therapeutic prediction**. They enable **Dynamic Treatment Regimes (DTRs)**, where therapeutic paths are optimized in real time based on patient responses.

3.3.1. Predicting Treatment Response

A key use-case is predicting which patients respond best to specific interventions (e.g., CBT, pharmacotherapy, VR-BA). Models utilizing demographic, clinical (symptom severity, treatment history), and psychometric data (e.g., maladaptive schemas) learn predictive patterns (Lim & Barlas, 2019). For instance, AI could guide clinicians on whether a patient with MDD and high experiential avoidance would benefit more from VR-BA or traditional CBT.

3.3.2. Optimization through Just-in-Time Adaptive Interventions (JITAI)

DTR and **Reinforcement Learning (RL)** algorithms aim to deliver **Just-in-Time Adaptive Interventions** by dynamically adjusting therapeutic intensity, modality, or content based on real-time data (e.g., from wearables). Such algorithms can detect sleep inertia and send timely Behavioral Activation prompts (e.g., "Get up and make yourself a matcha tea") before depressive escalation occurs. In chatbot contexts, RL can shift support modes (e.g., from psychoeducation to relaxation training) when rising distress is detected in text analysis.

3.3.3. Explainable AI (XAI) and Clinician Trust

A critical barrier to DTR implementation is the lack of **Explainable AI (XAI)**. For personalization to be clinically viable, therapists must understand *why* an algorithm made a given decision (e.g., recommending VR instead of chatbot-based intervention) to maintain professional autonomy and accountability (Tavory, 2024).

Prominent XAI techniques include:

● **SHAP (SHapley Additive exPlanations):** Game-theoretic attribution of each feature's contribution to a prediction.

● **LIME (Local Interpretable Model-agnostic Explanations):** Local surrogate models that approximate complex predictors with interpretable explanations.

Such tools can generate interpretive *scores* showing how features (e.g., physical inactivity or negative journal entries) influenced therapeutic recommendations. Without transparency, clinical adoption will remain limited.

3.3.4. Challenges in Personalization

Key challenges include the quality and diversity of training data—datasets must be large, heterogeneous, and include relevant outcomes (e.g., biological or behavioral biomarkers, not only self-reports). Moreover, effective personalization requires both explainability and cross-cultural validation across varied clinical populations. Currently, these limitations significantly constrain real-world AI implementation in clinical practice.

To sum up, Section V highlights how **AI-driven tools**—from GenAI chatbots to NLP-based monitoring and personalized adaptive algorithms—demonstrate meaningful short-term efficacy and vast scalability potential. However, issues of explainability, data ethics, and algorithmic fairness remain major barriers before such systems can achieve safe, routine clinical deployment.

4. Results: Virtual Reality-Based Interventions (VR)

4.1. Virtual Reality Exposure Therapy (VRET) and VR-Enhanced Behavioral Activation (VR-BA)

4.1.1. VRET for Specific Phobias and PTSD

Virtual Reality Exposure Therapy (VRET) is the most established and extensively studied application of VR in psychotherapy. It focuses on **controlled, safe, and gradual exposure** of patients to anxiety-provoking or traumatic stimuli, facilitating habituation and extinction of the fear response. VRET is recognized as **clinically equivalent** to traditional *in vivo* exposure (Wang et al., 2024).

● **Social Anxiety Disorder (SAD):**

● A randomized controlled trial (Zainal et al., 2021) demonstrated the efficacy of self-guided VRET for adults with SAD. Participants practiced challenging scenarios (e.g., job interviews, public speaking) without therapist presence, achieving significant reductions in SAD severity, interview-related anxiety, and depressive symptoms (PHQ-9). These findings underscore VRET's scalability, as it enables therapeutic exposure without continuous clinician involvement.

● **Post-Traumatic Stress Disorder (PTSD):**

● Systematic reviews consistently confirm that VRET—often combined with biofeedback (see Section 6.3)—is effective in reducing PTSD symptoms, particularly among veterans (Wang et al., 2024). The key success factors are **presence** and **immersion**, which allow for deep emotional engagement with simulated traumatic scenes, necessary for emotional processing and desensitization.

4.1.2. VR-Enhanced Behavioral Activation (VR-BA) for MDD

Behavioral Activation (BA) is an established first-line evidence-based practice (EBP) for treating Major Depressive Disorder (MDD), aimed at increasing engagement in rewarding activities. Depression is often accompanied by avoidance and apathy, preventing patients from initiating *in vivo* activities.

VR provides a safe and motivating environment to help overcome these barriers.

● **Feasibility and Pilot Study (2022):** A pilot RCT compared **VR-BA** with standard BA and treatment-as-usual (TAU), showing high feasibility, acceptability, and tolerability (low cybersickness scores) (Paul et al., 2022). VR simulated rewarding experiences (e.g., virtual nature walks, art gallery visits) that patients later replicated in real life.

● **Clinical Efficacy (2024):** A larger follow-up RCT confirmed the clinical efficacy of VR-BA (Paul et al., 2024), demonstrating significant reductions in MDD symptoms. VR served as a safe "training ground" where patients could break avoidance cycles, experience positive emotions, and rebuild motivation—thus improving real-world engagement. VR-BA proved at least as effective as traditional BA in symptom reduction.

4.1.3. Cybersickness and Tolerability

VR tolerability is crucial for clinical adoption. Studies systematically assessed cybersickness using the **Simulator Sickness Questionnaire (SSQ)** (Paul et al., 2024; Zainal et al., 2021). The risk of nausea, dizziness, and visual fatigue is lower in static, motion-free environments (e.g., mindfulness training) than in dynamic ones (e.g., flight simulations), but still represents a relevant barrier for some users. Low SSQ scores were observed in static VR-BA conditions (Paul et al., 2024). Continuous monitoring and adaptive design are essential to maintain tolerability.

4.2. VR in Mindfulness and Relaxation Training

In non-reactive VR applications—mindfulness and relaxation training—the literature from 2020-2024 provides an expanding body of evidence. The central hypothesis is that immersive VR environments enhance engagement and attentional focus, helping users who struggle with traditional mindfulness or meditation practices.

4.2.1. Evidence from RCTs and Comparative Studies

Methodologically robust RCTs and pilot studies compared **VR-based mindfulness** programs with traditional formats (audio-only, mobile apps).

● A randomized trial among university students showed that VR-supported mindfulness significantly reduced stress levels compared to controls, with statistically significant short-term effects (Modrego-Alarcón et al., 2021).

● Subsequent pilot RCTs (2023-2024) confirmed that short, immersive VR sessions are often **equally or more effective** than audio or desktop-based mindfulness programs in reducing tension and improving mood (Poetar et al., 2023; Blackmore et al., 2024).

● Comparative hardware studies found that HMD-based immersion correlates with higher engagement and stronger subjective *presence*, translating to greater short-term benefits (Olasz et al., 2024).

4.2.2. Therapeutic Mechanisms

VR facilitates **controlled exposure to calming stimuli** (e.g., natural landscapes, quiet spaces, cosmic environments) and delivers **multisensory input** (spatial audio, stereoscopic visuals) that deepens mindfulness and *flow* experiences (Modrego-Alarcón et al., 2021). Recent studies have introduced **VR-delivered Mindfulness-Oriented Recovery Enhancement (MORE)** protocols that leverage immersion to strengthen self-regulation mechanisms. Preliminary evidence suggests that visualizing regulatory processes in VR may enhance **mindfulness** and **cognitive reappraisal**, improving outcomes for depression and addiction (Garland et al., 2024).

4.2.3. Research Gaps

Major limitations include:

1. Small, often non-clinical samples (students, volunteers)
2. Short observation periods (lack of long-term *follow-ups*)
3. High heterogeneity across interventions (content, session length, environments)
4. Few studies comparing VR-mindfulness as an **adjunctive** vs. **standalone** treatment (Olasz et al., 2024).

Future work should standardize protocols and investigate long-term maintenance of effects.

4.3. Integration of VR with Biofeedback and Neurofeedback

Combining VR with **biofeedback (BF)** and **neurofeedback (NF/EEG)** represents a promising direction that merges immersive therapy with **real-time physiological data**, creating adaptive closed-loop training systems.

4.3.1. Concept and Clinical Evidence

The core concept involves embedding users in immersive VR environments whose parameters (e.g., visual intensity, breathing visualizations, reward elements) dynamically adjust to physiological signals—such as heart rate, heart-rate variability (HRV), electrodermal activity (EDA), or EEG-derived brainwave data (Cho et al., 2024).

A randomized study comparing **VR-based biofeedback** with traditional biofeedback demonstrated significant reductions in depression and anxiety, with comparable outcomes between VR-BF and therapist-delivered BF (Cho et al., 2024). Other pilot trials reported improved mood following short **VR + EEG** sessions among individuals with depressive symptoms (Tacca et al., 2024).

4.3.2. Mechanisms of Action

Proposed therapeutic mechanisms include:

1. **Enhanced Motivation:** Immersion increases *adherence* to self-regulation training, which is often perceived as monotonous in traditional formats.
2. **Immediate Feedback:** Users visually experience real-time changes in physiological parameters—such as a brightening landscape corresponding to lowered heart rate—which accelerates autonomic learning.
3. **Skill Transfer:** Theoretical and experimental work suggests that VR facilitates the transfer of learned self-regulation skills to real-world contexts, supported by the strong sense of *presence* (Drigas & Sideraki, 2024).

4.3.3. Technical Challenges

Key technical limitations include:

- **Synchronization:** Delays between biometric input and VR rendering must remain below ~50 ms (Tacca et al., 2024) to ensure effective feedback.
- **EEG Artifacts:** Head motion and electromagnetic interference from HMDs complicate EEG signal analysis.
- **Protocol Standardization:** Standardized neurofeedback (NF) protocols—e.g., defining EEG bands or electrode sites—are still in early development for VR contexts (Tacca et al., 2024).

Summary of Section VI

VR-based interventions, particularly **VRET** and **VR-BA**, have achieved robust empirical support for treating anxiety and depression. The addition of **mindfulness modules** and **bio/neurofeedback integration** broadens their therapeutic scope, offering multimodal, engaging, and data-driven treatment experiences. However, widespread adoption still faces technical (cybersickness, synchronization), practical (cost, training), and ethical (data privacy) barriers that require systematic solutions before VR can become a mainstream therapeutic modality.

5. Discussion: Barriers, Challenges, and Implementation Directions

5.1. Comparative Analysis of AI and VR Interventions

The literature from 2020-2024 reveals that both AI and VR have achieved measurable clinical efficacy, though they differ significantly in **mechanisms of action**, **therapeutic depth**, and **scalability**.

Dimension	AI (Chatbots, NLP, DTR)	VR (VRET, VR-BA, VR-Mindfulness)
Primary Mechanism	Cognitive and linguistic interaction; psychoeducation; self-reflection	Sensorimotor immersion; experiential learning; emotion regulation
Therapeutic Depth	Moderate - effective for psychoeducation and short-term symptom relief	High - supports emotional processing and behavioral activation
Scalability	Very high - instant digital deployment, 24/7 availability	Limited - requires hardware, space, and setup
Supervision Requirement	Minimal to moderate (can be self-guided)	Moderate to high (clinician or technician often required)
Ethical and Technical Risks	Algorithmic bias, hallucinations, overreliance	Cybersickness, data security (biometrics)
Optimal Use	First-line, scalable mental-health support	Mid-/late-stage therapeutic enhancement or exposure work

Synthesis:

AI interventions (especially chatbots) excel in *accessibility* and *cost-effectiveness*, offering immediate, scalable support to large populations. In contrast, VR interventions provide superior *experiential depth* and are more effective for emotional, exposure-based, and activation therapies. Their complementary nature forms the foundation of future **blended-care ecosystems**.

5.2. Ethical, Legal, and Social Implications (ELSI)

5.2.1. Ethical Issues

1. Informed Consent and Autonomy: AI systems used in mental health often operate autonomously, which raises questions about whether users truly understand the nature of interaction (especially with chatbots simulating empathy). Transparent disclosures must inform patients that they are interacting with an AI rather than a human therapist.

2. Algorithmic Bias and Fairness: Training data often reflect social, gender, and racial biases, which can perpetuate inequalities in mental health care (Ueda et al., 2023). Fairness metrics and diverse datasets must become standard components of algorithmic validation.

3. Explainability and Accountability: Opaque AI models undermine trust. Without *explainability* (XAI), clinicians cannot verify or contest algorithmic recommendations, threatening professional autonomy.

4. Risk of Overreliance: Patients may develop excessive emotional attachment to chatbots (*automation bias*), leading to delayed professional consultation during crises (Richards et al., 2023). Protocols for *human escalation* should therefore be mandatory in every system.

5. Therapeutic Integrity: VR-based therapies must ensure that immersive experiences respect patients' psychological limits and avoid unintended re-traumatization during exposure or mindfulness modules.

5.2.2. Legal Challenges

The introduction of AI and VR into clinical contexts challenges existing legal frameworks for **medical devices**, **liability**, and **data protection**.

1. Medical Device Classification: In the EU, many digital tools qualify as *Software as a Medical Device* (SaMD) under the MDR (Regulation EU 2017/745). AI chatbots that perform diagnostic or therapeutic functions must undergo clinical validation and conformity assessment.

2. Liability: When AI produces a harmful or misleading recommendation, the question of accountability becomes complex: is the manufacturer, clinician, or algorithm responsible? Legal scholars advocate for shared liability models that allocate responsibility proportionally between the AI provider and healthcare professional (Tavory, 2024).

3. Data Privacy and Security: AI and VR systems collect highly sensitive biometric and psychological data. Compliance with GDPR (EU) or HIPAA (US) requires explicit, granular consent, pseudonymization, and encryption. In VR, biometric tracking (eye, motion, HRV) introduces novel privacy risks requiring new legal safeguards.

4. Cross-Border Data Transfers: Cloud-based mental health platforms often process user data across jurisdictions, creating potential conflicts between privacy regimes (e.g., EU vs. US). Developing international data-sharing standards for digital therapeutics is an urgent priority.

5.2.3. Social and Implementation Challenges

1. Digital Divide and Accessibility: While AI systems scale easily, access to stable internet and digital literacy remains uneven, particularly in low-income or rural populations. VR adoption further depends on hardware availability and user familiarity with immersive technologies.

2. Therapeutic Relationality: Replacing or supplementing human contact with AI may alter the *relational essence* of psychotherapy. Studies show that the perceived warmth and trustworthiness of chatbots vary widely depending on linguistic and cultural context.

3. Professional Resistance: Clinicians may resist integrating AI/VR due to perceived threats to expertise, workflow disruption, or lack of reimbursement frameworks. Ongoing professional education and transparent clinical validation are key to acceptance.

4. Economic Barriers: Despite falling costs of VR headsets and cloud infrastructure, initial investment remains high for institutions. Sustainable adoption requires reimbursement models that recognize *digital sessions* as valid therapeutic services.

5.3. Implementation Strategies for AI and VR in Clinical Practice

To translate research into practice, structured implementation pathways are needed. The following strategies are recommended:

1. Integration into Stepped Care Models: AI chatbots can function as *Step 1* in stepped-care systems, delivering psychoeducation and triage for mild-to-moderate symptoms. VR interventions, requiring more resources, are better suited for *Step 2-3*, offering experiential enhancement of psychotherapy.

2. Clinician-AI Collaboration Frameworks: AI should serve as *augmented intelligence*, supporting—not replacing—clinicians. Decision-support tools must provide explainable recommendations that therapists can verify, adapt, and override when necessary.

3. Hybrid and Blended Care: Combining AI-guided CBT chatbots with VR exposure or mindfulness modules could create synergistic effects—enhancing both *reach* and *depth*. For example, a patient might use a chatbot for daily cognitive exercises while attending weekly VR-BA sessions supervised by a therapist.

4. Ethical-by-Design Development: AI and VR systems should be designed with ethics and privacy embedded at every development stage—covering consent, transparency, accessibility, and bias mitigation.

5. Continuous Monitoring and Evaluation: Digital interventions require dynamic monitoring of safety, efficacy, and equity, supported by regulatory bodies and academic institutions.

5.4. Future Directions and Research Gaps:

The years 2020-2024 have established a strong foundation for digital mental health research, yet significant gaps remain before these tools can be widely deployed in healthcare systems.

1. Long-Term Efficacy and Sustainability: Most studies assess short-term outcomes (≤12 weeks). Longitudinal RCTs with ≥12-month follow-up are needed to confirm durability of therapeutic gains.

2. Real-World Evidence (RWE): Clinical trials often occur in controlled settings. Future research must evaluate *effectiveness* in routine clinical environments with heterogeneous populations.

3. Cross-Cultural Validation: AI and VR interventions need validation across diverse languages and cultures to ensure equitable performance and accessibility.

4. Interoperability and Standardization: Integration with Electronic Health Records (EHRs) and standard APIs is essential for scalability. Lack of interoperability currently hampers seamless deployment in hospitals and clinics.

5. Hybrid Therapies and Co-Creation: Promising new directions include hybrid AI-VR interventions, such as chatbots embedded within VR therapy or AI-driven dynamic scenario adjustment based on biometric data. Co-design with clinicians and patients will be key to clinical adoption and personalization.

5.5. Synthesis

Both AI and VR have reached technological and clinical maturity sufficient to demonstrate meaningful short-term efficacy for treating anxiety and depression. Yet, they differ in essence:

- **AI** provides scalability, personalization, and continuous engagement.
- **VR** offers experiential and embodied therapy that facilitates deep emotional and behavioral change.

The next phase of digital psychiatry will rely on **integrative ecosystems**, where AI systems manage assessment and personalization, while VR provides immersive therapeutic depth. This synergy will enable *blended care models* that are effective, ethical, and sustainable—bridging the gap between scalability and human connection.

6. Conclusions

The systematic review and meta-analysis presented in this study demonstrate that both **Artificial Intelligence (AI)** and **Virtual Reality (VR)** have matured to the point of achieving verified clinical efficacy in the treatment of **Anxiety Disorders (AD)** and **Major Depressive Disorder (MDD)**. Each technology contributes unique therapeutic strengths that, when combined, can redefine the landscape of mental healthcare delivery.

1. Clinical Implications

AI-based interventions— particularly **Generative AI chatbots**—offer unprecedented scalability and accessibility, making them effective first-line tools for psychoeducation, cognitive restructuring, and emotional self-regulation.VR-based interventions, especially **VRET** and **VR-Enhanced Behavioral Activation (VR-BA)**, enable deep emotional engagement and behavioral change through immersive, experiential exposure.

Together, these modalities provide a complementary framework for **personalized digital psychiatry**, balancing reach and depth.

The integration of AI and VR within **blended-care models**—where digital systems operate alongside human therapists—appears to be the most clinically viable and ethically sound direction for near-future mental healthcare systems.

2. Research Implications

Future research must focus on:

- Conducting **longitudinal RCTs** (≥12 months) to verify the durability of therapeutic effects.
- Expanding **cross-cultural validation** to ensure inclusivity and fairness in AI and VR performance.
- Developing standardized **interoperability frameworks** to connect AI/VR platforms with Electronic Health Records (EHRs).
- Establishing robust **ethical and legal governance** frameworks addressing explainability, data protection, and liability in clinical contexts.

Sustained interdisciplinary collaboration— uniting clinicians, data scientists, ethicists, and policymakers—will be essential for safe and responsible implementation.

3. Technological and Ethical Outlook

The coming decade will likely witness the emergence of **integrated, adaptive therapeutic ecosystems**, in which:

- AI performs continuous assessment, personalization, and crisis detection,
- VR delivers immersive, emotionally grounded interventions, and
- Clinicians supervise, interpret, and ensure ethical integrity.

The greatest challenge will be ensuring that technology *enhances rather than replaces* the human relationship at the heart of therapy. Maintaining empathy, transparency, and patient autonomy must remain the guiding principles of all digital mental health innovations.

4. Final Synthesis

Digital mental health care is transitioning from innovation to integration. AI and VR are no longer experimental novelties but emerging **clinical infrastructures** capable of expanding access and improving outcomes when applied responsibly. The synergy between scalable AI intelligence and immersive VR presence offers a transformative path toward more accessible, engaging, and humane mental healthcare—one that empowers both clinicians and patients in equal measure.

Disclosure

Authors do not report any disclosures Authors' contribution: Conceptualization, D.G and J.B.; methodology, D.G, J.B and M.A.; check, E.S., D.G., L.S. and M.A.; formal analysis, E.S., J.B., D.G. and M.A.; investigation, L.S.; resources, L.S. and K.M.; writing - rough preparation, L.S., K.M. and E.S.; writing - review and editing, J.B., D.G. and M.A.; visualization, L.S.; supervision, L.S., K.M. and E.S.; project administration,D.G., and M.A., and E.S.

All authors have read and agreed with the published version of the manuscript.

Funding statement: The study did not receive special funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflict of Interest Statement: The authors declare no conflict of interest.

REFERENCES

1. Abbas, J. R., Gantwerker, E., Volk, M., Payton, T., McGrath, B. A., Tolley, N., & Isba, R. (2024). Describing, evaluating, and exploring barriers to adoption of virtual reality: An international modified Delphi consensus study involving clinicians, educators, and industry professionals. *Journal of Medical Extended Reality, 1*(1), 202-214. https://doi.org/10.1089/jmxr.2024.0022.
2. Abdulsalam, A., & Alhothali, A. (2024). Suicidal ideation detection on social media: A review of machine learning methods. *Social Network Analysis and Mining*. https://doi.org/10.1007/s13278-024-01348-0.
3. Berardi, C., De Togni, A., Antonini, M., Hinwood, M., Jordan, Z., Wechtler, H., & Paolucci, F. (2024). Barriers and facilitators to the implementation of digital technologies in mental health systems: A qualitative systematic review to inform a policy framework. *BMC Health Services Research, 24*, Article 243. https://doi.org/10.1186/s12913-023-10536-1.
4. Blackmore, R., Poyser, R., et al. (2024). Examining the use of virtual reality to support mindfulness skills practice in mood and anxiety disorders: Mixed-methods study. *JMIR Mental Health, 11*, e45640. https://doi.org/10.2196/45640.
5. Bottomley, D., & Thaldar, D. (2023). Liability for harm caused by AI in healthcare: An overview of the core legal concepts. *Frontiers in Pharmacology, 14*, 1297353. https://doi.org/10.3389/fphar.2023.1297353.
6. Bouderhem, R., Duca, K., et al. (2024). Shaping the future of AI in healthcare through ethics and governance. *Humanities and Social Sciences Communications*. https://doi.org/10.1057/s41599-024-02894-w.
7. Casu, M., Triscari, S., Battiato, S., Guarnera, L., & Caponnetto, P. (2024). AI chatbots for mental health: A scoping review of effectiveness, feasibility, and applications. *Applied Sciences, 14*(13), 5889. https://doi.org/10.3390/app14135889.
8. Cho, Y., Kim, H., Seong, S., Park, K., Choi, J., Kim, M. J., Kim, D., & Jeon, H. J. (2024). Effect of virtual reality-based biofeedback for depressive and anxiety symptoms: Randomized controlled study. Journal of affective disorders, 361, 392-398. https://doi.org/10.1016/j.jad.2024.06.031
9. Drigas, A., & Sideraki, A. (2024). Brain Neuroplasticity Leveraging Virtual Reality and Brain-Computer Interface Technologies. Sensors, 24(17), 5725. https://doi.org/10.3390/s24175725
10. Garland, E. L., Recasens, M., et al. (2024). Virtual reality-based Mindfulness-Oriented Recovery Enhancement (MORE-VR) as an adjunct to medications for opioid use disorder: A Phase 1 trial. *Annals of Medicine / Journal*. https://doi.org/10.1080/07853890.2024.2392870
11. Kaminsky, Z., et al. (2024). Machine learning-based suicide risk prediction model for suicidal trajectory on social media following suicidal mentions: Independent algorithm validation. *JMIR* (or related). https://doi.org/10.2196/49927.
12. Kourtesis, P., Christakis, V., et al. (2024). Cybersickness in virtual reality: The role of individual differences, its effects on cognitive functions and motor skills, and intensity differences during and after immersion. *Virtual Worlds / Virtual Worlds (MDPI), 3*(1), 62-93. https://doi.org/10.3390/virtualworlds3010004.
13. Li, H., Zhang, R., et al. (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *npj Digital Medicine, 6*(1), 236. https://doi.org/10.1038/s41746-023-00979-5.
14. Lim, C. R., & Barlas, J. (2019). The effects of toxic early childhood experiences on depression according to Young Schema Model: A scoping review. *Journal of Affective Disorders, 246*, 1-13. https://doi.org/10.1016/j.jad.2018.12.006.
15. Min, S., et al. (2023). Acoustic analysis of speech for screening for suicide risk: Machine learning classifiers for between- and within-person evaluation of suicidality. *JMIR*, 2023; e45456. https://doi.org/10.2196/45456.
16. Modrego-Alarcón, M., López-Del-Hoyo, Y., García-Campayo, J., Pérez-Aranda, A., Navarro-Gil, M., Beltrán-Ruiz, M., Morillo, H., Delgado-Suarez, I., Oliván-Arévalo, R., & Montero-Marin, J. (2021). Efficacy of a mindfulness-based programme with and without virtual reality support to reduce stress in university students: A randomized controlled trial. Behaviour research and therapy, 142, 103866. https://doi.org/10.1016/j.brat.2021.103866
17. Arturo Montejo-Ráez, M. Dolores Molina-González, Salud María Jiménez-Zafra, Miguel Ángel García-Cumbreras, and Luis Joaquín García-López. 2024. A survey on detecting mental disorders with natural language processing: Literature review, trends and challenges. Comput. Sci. Rev. 53, C (Aug 2024). https://doi.org/10.1016/j.cosrev.2024.100654
18. Murdoch, B. (2021). Privacy and artificial intelligence: Challenges for protecting health information in a new era. *BMC Medical Ethics, 22*(1), 122. https://doi.org/10.1186/s12910-021-00687-3.
19. Olasz, O., et al. (2024). The effects of virtual reality-based mindfulness exercises on the perception of time, psychological and physiological states of young people: A randomized crossover trial. *Mindfulness*. https://doi.org/10.1007/s12671-024-02438-y
20. Ong, T., Barrera, J. F., et al. (2024). Mental health providers are inexperienced but interested in telehealth-based virtual reality therapy: Survey study. *Frontiers in Virtual Reality*. https://doi.org/10.3389/frvir.2024.1332874.

21. Paul, M., Bullock, K., & Bailenson, J. (2022). Virtual reality behavioral activation for adults with major depressive disorder: Feasibility randomized controlled trial. *JMIR Mental Health, 9*(5), e35526. https://doi.org/10.2196/35526.
22. Paul, M., Bullock, K., Bailenson, J., & Burns, D. (2024). Examining the efficacy of extended reality-enhanced behavioral activation for adults with major depressive disorder: Randomized controlled trial. *JMIR / Mental Health*. https://doi.org/10.2196/52326.
23. Poetar, C. R., et al. (2023). Immersive virtual reality or computerized mindfulness meditation for improving mood? Preliminary efficacy from a pilot randomized trial. *Frontiers in Psychology, 14*, 1157469. https://doi.org/10.3389/fpsyg.2023.1157469.
24. Richards, D., Vythilingam, R., & Formosa, P. (2023). A principlist-based study of the ethical design and acceptability of artificial social agents. *International Journal of Human-Computer Studies, 172*, Article 102980. https://doi.org/10.1016/j.ijhcs.2022.102980.
25. Tacca, C., Kerr, B. A., McLamb, C., et al. (2024). Efficacy of a remote virtual reality and EEG-enabled psychotherapy system for the treatment of depressive symptoms. *Frontiers in Virtual Reality*. https://doi.org/10.3389/frvir.2024.1281017.
26. Tavory, T. (2024). Regulating AI in mental health: Ethics of care perspective. *JMIR Mental Health, 11*, e58493. https://doi.org/10.2196/58493.
27. Terranova, C., Cestonaro, C., Fava, L., & Cinquetti, A. (2024). AI and professional liability assessment in healthcare: A revolution in legal medicine? *Frontiers in Medicine, 10*, Article 1337335. https://doi.org/10.3389/fmed.2023.1337335.
28. Ueda, D., Kakinuma, T., Fujita, S., Kamagata, K., Fushimi, Y., Ito, R., Matsui, Y., Nozaki, T., Nakaura, T., Fujima, N., Tatsugami, F., Yanagawa, M., Hirata, K., Yamada, A., Tsuboyama, T., Kawamura, M., Fujioka, T., & Naganawa, S. (2024). Fairness of artificial intelligence in healthcare: review and recommendations. Japanese journal of radiology, 42(1), 3-15. https://doi.org/10.1007/s11604-023-01474-3
29. Wang, P., Ai, X., Zhang, X., Ma, F., Zhuang, Y., & Wang, S. (2024). Evaluating virtual reality technology in psychotherapy: Impacts on anxiety, depression, and ADHD. *Frontiers in Psychiatry, 15*, 1480788. https://doi.org/10.3389/fpsyt.2024.1480788.
30. Zainal, N. H., Chan, W. W., Saxena, A. P., Taylor, C. B., & Newman, M. G. (2021). Pilot randomized trial of self-guided virtual reality exposure therapy for social anxiety disorder. *Behaviour Research and Therapy, 147*, 103984. https://doi.org/10.1016/j.brat.2021.103984.
31. World Health Organization. (2022). *World Mental Health Report: Transforming mental health for all.* Geneva: WHO. ISBN: 9789240049338.