International Journal of Innovative Technologies in Social Science

e-ISSN: 2544-9435

Operating Publisher
SciFormat Publishing Inc.
ISNI: 0000 0005 1449 8214

2734 17 Avenue SW,
Calgary, Alberta, T3E0A7,
Canada
+15878858911
editorial-office@sciformat.ca

ARTICLE TITLE	THE EVOLUTION OF DIGITAL THERAPEUTICS (DTx) IN MENTAL HEALTH CARE: A SYSTEMATIC REVIEW OF SOCIAL IMPACT, ETHICAL CHALLENGES, AND PATIENT ENGAGEMENT
DOI	https://doi.org/10.31435/ijitss.1(49).2026.5192
RECEIVED	02 February 2026
ACCEPTED	25 March 2026
PUBLISHED	30 March 2026
LICENSE	The article is licensed under a **Creative Commons Attribution 4.0 International License.**

THE EVOLUTION OF DIGITAL THERAPEUTICS (DTx) IN MENTAL HEALTH CARE: A SYSTEMATIC REVIEW OF SOCIAL IMPACT, ETHICAL CHALLENGES, AND PATIENT ENGAGEMENT

Alicja Maria Mitan *(Corresponding Author, Email: alis.mitan@gmail.com)*
Fryderyk Chopin University Clinical Hospital in Rzeszów, Rzeszów, Poland
ORCID ID: 0009-0004-5617-5034

Weronika Napierała
Jerzy Popiełuszko Hospital of Bielany, Warsaw, Poland
ORCID ID: 0009-0005-1411-4046

Karolina Magda Leszczyńska
Pomeranian Medical University, Szczecin, Poland
ORCID ID: 0009-0000-6427-7269

Jeremi Leon Jasiński
Pomeranian Medical University, Szczecin, Poland
ORCID ID: 0009-0004-1690-5936

Anna Krzysztofik
Pomeranian Medical University, Szczecin, Poland
ORCID ID: 0009-0002-3130-1917

Karolina Krawczyk
Pomeranian Medical University, Szczecin, Poland
ORCID ID: 0009-0005-1191-5788

Kamila Teresa Kańska
Pomeranian Medical University, Szczecin, Poland
ORCID ID: 0009-0005-2702-7013

Aleksandra Maria Tomaszewska
University Clinical Hospital of Bialystok, Białystok, Poland
ORCID ID: 0009-0000-4684-0492

Maciej Tomasz Wieczorek
Pomeranian Medical University, Szczecin, Poland
ORCID ID: 0009-0007-4058-7650

Karolina Julia Hak
Independent Public Provincial Integrated Hospital in Szczecin, Szczecin, Poland
ORCID ID: 0009-0009-7641-0237

ABSTRACT

Background: The global landscape of mental health care has undergone a fundamental transformation driven by the emergence of Digital Therapeutics (DTx). These software-driven, evidence-based interventions are designed to deliver high-quality psychological support directly to the user. As traditional healthcare infrastructures face unprecedented strain and accessibility crises, DTx has transitioned from a niche supplementary tool to a cornerstone of modern digital psychiatry, redefining how clinical outcomes are achieved and measured in the digital age.

Objective: This systematic review critically evaluates the rapid evolution of the DTx sector between 2020 and 2026. The primary objective is to analyze three pivotal dimensions: the broader social implications of digital health democratization, the shifting ethical complexities introduced by the integration of Generative Artificial Intelligence (GenAI), and the persistent socio-technical barriers that hinder long-term patient adherence and therapeutic efficacy.

Methodology: Adhering to the PRISMA 2020 guidelines, a comprehensive systematic search was executed across PubMed, Scopus, and Web of Science. The synthesis included twenty-five high-impact studies focused on digital psychiatry, socio-technical frameworks, and algorithmic ethics. The analysis encompassed peer-reviewed literature regarding clinical validation processes, global regulatory shifts (such as the German DiGA market), and the principles of user-centered design (Abbas et al., 2025; Jiang et al., 2025; Gensorowsky et al., 2024).

Results: Findings reveal that while DTx offers a viable pathway for expanding care (Carlbring et al., 2018), significant "digital divides" persist. Socioeconomic status and digital literacy remain primary gatekeepers, preventing equitable distribution among marginalized and socioeconomically disadvantaged populations (Divatia et al., 2026; Hengst et al., 2023; Piers et al., 2023). Ethical challenges have evolved from basic data privacy toward sophisticated debates on algorithmic bias, the "black box" nature of AI diagnostics, and the critical necessity for "human-in-the-loop" oversight to mitigate clinical risks (Rahsepar Meadi et al., 2025; Shakeel et al., 2025). Furthermore, patient attrition remains the "Achilles heel" of digital interventions (Nwosu et al., 2022). Sustainable engagement is now increasingly linked to the implementation of "precision engagement" frameworks and the successful cultivation of a digital therapeutic alliance (Eiselt et al., 2025; Smith et al., 2025).

Conclusion: The evolution of DTx from 2020 to 2026 necessitates a paradigm shift toward a hybrid socio-technical ecosystem. Future success in the next decade will depend on moving beyond isolated clinical efficacy toward implementation strategies that are radically inclusive, ethically grounded, and designed for deep, human-centric engagement. Addressing the intersection of AI ethics and social equity is no longer optional but a requirement for the next generation of digital mental health care.

KEYWORDS

Digital Therapeutics, Mental Health, AI Ethics, Patient Engagement, Digital Health Equity, Socio-Technical Systems

CITATION

Alicja Maria Mitan, Weronika Napierała, Karolina Magda Leszczyńska, Jeremi Leon Jasiński, Anna Krzysztofik, Karolina Krawczyk, Kamila Teresa Kańska, Aleksandra Maria Tomaszewska, Maciej Tomasz Wieczorek, Karolina Julia Hak. (2026) The Evolution of Digital Therapeutics (DTx) in Mental Health Care: A Systematic Review of Social Impact, Ethical Challenges, and Patient Engagement. *International Journal of Innovative Technologies in Social Science*. 1(49). doi: 10.31435/ijitss.1(49).2026.5192

1. Introduction

1.1 The Paradigm Shift in Digital Mental Health

The global landscape of mental health care is currently undergoing a transformative paradigm shift, catalyzed by the rapid maturation and institutionalization of Digital Therapeutics (DTx). As we navigate the period between 2020 and 2026, the initial wave of unregulated mobile wellness applications has been superseded by a rigorous medical discipline focused on software-driven, evidence-based interventions (Friis-Healy et al., 2021). These technologies are no longer viewed as mere supplementary tools but as specialized medical devices capable of delivering high-fidelity cognitive behavioral therapy (CBT) and other clinical protocols at scale. Recent evidence suggests that these interventions can achieve clinical outcomes comparable to traditional face-to-face therapy, marking a new era in psychiatric practice (Calderone et al., 2025; Carlbring et al., 2018).

However, this rapid technological proliferation has given rise to what researchers define as the "Technological Paradox": while DTx offers the potential to bridge the treatment gap for millions, it simultaneously threatens to exacerbate existing health disparities through new forms of digital exclusion (Piers et al., 2023). The e-health revolution, intended to democratize access to care, often functions as a magnifier of social inequities. Consequently, the clinical benefits of a validated algorithm are frequently inaccessible to those residing on the "wrong side" of the digital divide—populations characterized by low socioeconomic status, limited technological infrastructure, or poor digital literacy (Fanta & Pretorius, 2023; Graham et al., 2019; Hengst et al., 2023).

1.2 The Evolution of Digital Therapeutics (2020–2026)

The evolution of the DTx sector can be categorized into three distinct chronological and functional phases. The first phase, accelerated by the global COVID-19 pandemic, was defined by rapid deployment and the necessity of crisis management to ensure continuity of care. The second phase involved the establishment of rigorous regulatory frameworks, most notably the German DiGA (Digital Health Applications) model. This framework created a global blueprint for the reimbursement and prescription of digital tools by national health systems, elevating software to the same level of clinical authority as traditional pharmaceuticals (Gensorowsky et al., 2024; Hosono et al., 2024). We are currently in the third phase: the deep integration of advanced Artificial Intelligence (AI) and Generative AI (GenAI) into therapeutic workflows, moving toward highly personalized, conversational, and adaptive interfaces (Campellone et al., 2025; Torous & Topol, 2025).

Despite these advancements, the transition from clinical trials to real-world implementation has revealed significant systemic friction. In many healthcare systems, the "clinician-as-gatekeeper" phenomenon remains a primary obstacle. Providers often exhibit implicit biases regarding a patient's technological competence, leading to a situation where innovative DTx options are disproportionately offered to younger, more affluent, and digitally literate populations. As a result, marginalized groups are frequently left with traditional, often overstretched, care models, further widening the gap in health outcomes (Piers et al., 2023; Volpe et al., 2025).

1.3 Socio-Technical Systems (STS) Theory as an Analytical Lens

To comprehensively analyze the complexities of DTx adoption, this review utilizes the Socio-Technical Systems (STS) theory. As applied in digital health research, STS theory posits that the success of a technological intervention is not determined solely by its algorithmic accuracy but by the "joint optimization" of its technical and social components (Eiselt et al., 2025; Fanta & Pretorius, 2023).

- **The Technical Subsystem:** Refers to the software architecture, the validity of the clinical algorithms, and the user interface (UI) design. In the context of 2026, this increasingly includes the interpretability of AI models (Explainable AI) and the robustness of data protection measures required for sensitive psychological assessments and identity-related interventions (Abbas et al., 2025; Shakeel et al., 2025; Srivastava & Goswami, 2026).
- **The Social Subsystem:** Encompasses the human elements, including patient-provider trust—often conceptualized as the digital therapeutic alliance—cultural attitudes toward mental health, and the organizational culture of the healthcare institutions (Calderone et al., 2025; Eiselt et al., 2025).

A failure in digital psychiatry often occurs when a high-performing technical subsystem is introduced into a fractured social subsystem. For instance, an AI-driven intervention may achieve high clinical efficacy in a controlled trial, yet fail in a community setting because it lacks cultural humility or fails to account for the specific "precision engagement" needs and literacy levels of the end-users (Graham et al., 2019; Rahsepar Meadi et al., 2025).

1.4 Objectives and Scope of the Review

This systematic review synthesizes evidence from twenty-five core studies published between 2020 and 2026 to address the critical misalignment between rapid technological innovation and the structural realities of healthcare delivery. The primary objectives are:

- **To analyze the social impact of DTx implementation:** Specifically focusing on how socioeconomic status, geographic location, and digital literacy act as primary determinants of health equity in the digital age (Carlbring et al., 2018; Piers et al., 2023).
- **To evaluate the ethical challenges of AI integration:** Examining the inherent risks of algorithmic bias, the urgent necessity for "Explainable AI" (XAI) to ensure clinical safety, and the evolving legal and global standards for data protection in psychological assessment (Kuhn et al., 2024; Rahsepar Meadi et al., 2025; Srivastava & Goswami, 2026).
- **To investigate the "Engagement Crisis":** Deconstructing the multi-faceted factors behind high attrition rates—the "Achilles heel" of digital interventions—and evaluating the critical role of "human-in-the-loop" models and precision engagement in maintaining therapeutic efficacy (Nwosu et al., 2022; Smith et al., 2025).

By applying the Socio-Technical Systems (STS) framework, this work seeks to provide a comprehensive roadmap for policy-makers and clinicians, ensuring that the future of digital mental health is built on the principles of inclusivity, ethical transparency, and sustained patient engagement (Ali et al., 2025; Torous & Topol, 2025; Volpe et al., 2025).

2. Methodology

2.1 Research Design: A Socio-Technical Synthesis

This systematic review employs a multidimensional synthesis design, moving beyond traditional clinical meta-analysis to integrate a holistic socio-technical evaluation. Given that the success of Digital Therapeutics (DTx) is determined by the complex interplay between algorithmic performance and social infrastructure, the methodology is grounded in the Socio-Technical Systems (STS) framework (Eiselt et al., 2025). This approach ensures that technical variables, such as AI precision and data security protocols, are analyzed in parallel with social determinants, including socioeconomic position, digital literacy, and the strength of the therapeutic alliance (Graham et al., 2019; Piers et al., 2023). The review strictly follows the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol to maintain the highest level of academic transparency and rigor (Torous & Topol, 2025).

2.2 Information Sources and Search Strategy

A rigorous systematic search was conducted between January 2020 and February 2026. This temporal boundary was selected to capture the critical transition from the crisis-driven telehealth deployment during the global pandemic to the current era of Generative AI and standardized, regulated digital prescribing (Fürstenau et al., 2023; Torous & Topol, 2025). Three primary academic databases were queried: PubMed/MEDLINE, Scopus, and Web of Science.

To maximize the capture of relevant literature while maintaining socio-technical specificity, the search string utilized a combination of Medical Subject Headings (MeSH) and free-text terms. Primary terms included "Digital Therapeutics," "DTx," and "Prescription Digital Therapeutics." These were intersected with socio-technical terms such as "Socioeconomic Position," "Algorithmic Bias," "Health Equity," and "Explainable AI," alongside clinical terms including "Mental Health," "Psychotherapy," and "Attrition."

2.3 Inclusion and Exclusion Criteria: The Quality Threshold

To ensure that the analysis is built on high-fidelity evidence, the following criteria were established:

Inclusion:

- Peer-reviewed research (including Randomized Controlled Trials, systematic reviews, and meta-analyses) focusing on DTx for clinically recognized psychiatric disorders (Calderone et al., 2025; Seo et al., 2024).
- Studies specifically analyzing the "Implementation Gap"—the friction between theoretical software potential and real-world clinical adoption (Gensorowsky et al., 2024; Jiang et al., 2025).
- Research addressing "Ethics-by-Design" in AI-driven mental health tools (Kuhn et al., 2024; Rahsepar Meadi et al., 2025).
- Full-text articles published in English.

Exclusion:

- Direct-to-consumer (DTC) wellness applications that lack rigorous clinical validation or formal regulatory oversight, such as non-prescriptive meditation apps (Fürstenau et al., 2023).
- Studies focusing solely on hardware without a dominant software-driven therapeutic component.
- Gray literature, conference proceedings, or non-peer-reviewed commentaries.

2.4 Data Extraction and Systematic Appraisal

The initial screening identified 342 unique records. Following the removal of duplicates and a secondary screening of titles and abstracts, 45 full-text articles were evaluated against the inclusion criteria. The final selection of 25 core studies provides a comprehensive longitudinal view of the sector's evolution.

Data extraction was conducted through a specialized matrix focused on three essential layers:

1. **The Technical Layer:** Software architecture, AI model type (e.g., rule-based vs. GenAI), and data protection standards (Carlbring et al., 2018; Torous & Topol, 2025).
2. **The Social Layer:** User demographics, socioeconomic barriers, and digital health literacy requirements (Hengst et al., 2023; Piers et al., 2023).
3. **The Clinical Layer:** Adherence rates, clinical outcomes (e.g., CBT efficacy), and the presence of "human-in-the-loop" oversight (Campellone et al., 2025; Eiselt et al., 2025; Nwosu et al., 2022).

Quality assessment was performed using the AMSTAR 2 tool for systematic reviews and the Cochrane Risk of Bias 2 (RoB 2) tool for randomized controlled trials, ensuring a robust foundation for the final synthesis (Torous & Topol, 2025).

3. Results

3.1 Systematic Synthesis and Data Extraction Matrix

The systematic evaluation of 25 core studies underscores that the evolution of Digital Therapeutics (DTx) in mental health is a multidimensional process where technical maturity frequently clashes with social infrastructure. To provide an exhaustive analytical overview, the data from these sources have been extracted into a comprehensive matrix focusing on the intersection of the technical and social subsystems.

Table 1. Comprehensive Data Extraction of DTx Evolution (2020–2026)

Author(s) & Year	Target Condition	STS Focus	Major Clinical & Social Outcomes
Abbas et al. (2025a)	General Health	Workflow	AI precision requires clinical integration to avoid failure.
Ali et al. (2025)	Psychiatry	Policy Gap	Existing frameworks provide insufficient data privacy for DTx.
Calderone et al. (2025)	Depression	Efficacy	Digital CBT achieves parity with traditional face-to-face therapy.
Campellone et al. (2025)	Psychotherapy	Trust	Human-in-the-loop oversight is the foundation of digital trust.
Carlbring et al. (2018)	Anxiety	Safety	Internet-based CBT is highly effective with strict guardrails.
Divatia et al. (2026)	Precision DTx	Inclusion	Underserved groups are largely absent from validation trials.
Eiselt et al. (2025)	Chronic Illness	Retention	The ENGAGE model reduces attrition through personalized UX.
Fanta & Pretorius (2023)	Digital Systems	Sustainability	Socio-technical factors are primary drivers of adoption.

Friis-Healy et al. (2021)	Minority Health	Equity	Technology widens gaps without explicit cultural adaptation.
Fürstenau et al. (2023)	Psychiatry	Definition	DTx defined as clinically validated, software-driven devices.
Gensorowsky et al. (2024)	Mental Health	Reimbursement	Germany's DiGA model faces physician adoption bottlenecks.
Graham et al. (2019)	mHealth Apps	Economics	Low-income status remains a major barrier to app engagement.
Hengst et al. (2023)	mHealth Adoption	Barriers	Low socioeconomic position hinders meaningful tool adoption.
Hosono et al. (2024)	Insomnia	Regulation	Trial designs are shifting toward pharmaceutical-grade standards.
Jiang et al. (2025)	Addiction	Standardization	China's market scales fast but lacks global clinical alignment.
Kuhn et al. (2024)	Global Health	Ethics	Interdisciplinary perspectives needed for global DTx equity.
Nwosu et al. (2022)	Mental Health	Attrition	Attrition is the primary bottleneck for virtual interventions.
Piers et al. (2023)	Youth Mental	Exclusion	Deprived youth face the highest risk of digital exclusion.
Rahsepar Meadi et al. (2025)	Conversational AI	Justice	Non-diverse data leads to algorithmic injustice in AI.
Seo et al. (2024)	General DTx	Trends	Validated DTx provides significant long-term clinical benefits.
Shakeel et al. (2025)	ADHD	Explainability	XAI is required for clinical trust in automated diagnostics.
Smith et al. (2025)	Mental Health	Engagement	Adherence is highly dependent on psychiatric conditions.
Srivastava & Goswami (2026)	Assessment	Legal Review	Legal standards must evolve to protect digital identities.
Torous & Topol (2025)	Psychiatry	Innovation	GenAI has empathetic potential but lacks longitudinal data.
Volpe et al. (2025)	Risk Analysis	Fairness	Risks include diminished oversight and security compromises.

3.2 The Impact of Social Determinants: The New Digital Divide

The synthesis of results demonstrates that the "digital divide" in mental health care has evolved from a simple gap in device ownership into a sophisticated socio-technical barrier (Fanta & Pretorius, 2023; Piers et al., 2023). Data from the 2020–2026 period indicates that "Meaningful Use"—the ability to consistently and effectively engage with a therapeutic platform—is heavily dictated by a patient's socioeconomic position (Hengst et al., 2023).

- **Socioeconomic Stratification:** Individuals with lower income levels report **40% higher attrition rates** in self-guided DTx compared to high-income groups. This is frequently cited as a result of lower digital health literacy and a lack of tailored onboarding (Hengst et al., 2023; Nwosu et al., 2022).

- **Geographic Barriers:** Rural populations continue to face "infrastructural exclusion," where the lack of high-speed connectivity prevents the cloud-based data synchronization required for real-time AI-driven monitoring (Abbas et al., 2025a; Jiang et al., 2025).
- **Racial Disparities in Validation:** A critical finding is that underserved and minority populations are consistently underrepresented in clinical trials used to validate DTx products (Divatia et al., 2026; Friis-Healy et al., 2021). For instance, in many high-impact studies, over **80% of participants** identify as White/Caucasian, leading to interventions that may not be culturally or linguistically relevant for marginalized groups (Friis-Healy et al., 2021; Piers et al., 2023).

3.3 Algorithmic Ethics and the Shift toward Justice

As AI-driven diagnostics become the industry standard in 2026, the ethical focus has shifted from simple data privacy to algorithmic justice (Rahsepar Meadi et al., 2025; Srivastava & Goswami, 2026).

- **Algorithmic Bias:** Systematic reviews reveal that many psychiatric AI models exhibit bias against non-majority populations due to training on non-representative datasets. This can lead to misdiagnosis or lower therapeutic efficacy for minority groups (Rahsepar Meadi et al., 2025; Volpe et al., 2025).
- **The Explainability Requirement:** The "black box" nature of earlier AI models has led to a clinical requirement for Explainable AI (XAI). Clinicians are increasingly reluctant to prescribe DTx unless the underlying logic of the therapeutic recommendation is transparent (Abbas et al., 2025a; Shakeel et al., 2025).
- **Erosion of Oversight:** Researchers warn that rapid digitalization risks diminishing clinical oversight. Automated systems may fail to detect subtle nuances in patient crises that a human practitioner would recognize, highlighting a dangerous "oversight gap" (Campellone et al., 2025; Volpe et al., 2025).

3.4 The Crisis of Engagement: Attrition as the "Achilles Heel"

Engagement remains the most significant clinical barrier to the efficacy of digital psychiatry (Nwosu et al., 2022; Smith et al., 2025).

- **The Retention Gap:** While clinical trials show high efficacy for DTx, real-world data indicates that **over 50% of users** disengage within the first 30 days of a self-guided intervention (Eiselt et al., 2025; Nwosu et al., 2022).
- **Hybrid Model Superiority:** Studies consistently show that the most effective interventions are those that maintain a "Human-in-the-Loop" (HITL) (Campellone et al., 2025; Volpe et al., 2025). Hybrid models, which combine automated dCBT with periodic check-ins from a human facilitator, achieve adherence rates up to **3x higher** than purely automated versions (Calderone et al., 2025; Eiselt et al., 2025).
- **Precision Engagement:** The emergence of the ENGAGE framework in 2025 suggests that personalized UX design, which adapts to the user's psychological state and socioeconomic constraints in real-time, is essential for reducing long-term attrition (Eiselt et al., 2025).

3.5 The Role of Explainable AI (XAI) in Clinical Validation

The results of studies conducted between 2024 and 2026 indicate that the "black-box" nature of early psychiatric algorithms was a primary driver of physician skepticism (Abbas et al., 2025; Shakeel et al., 2025). The evolution toward Explainable AI (XAI) has subsequently become a prerequisite for the clinical validation and professional acceptance of DTx (Kuhn et al., 2024).

- **Clinical Trust and Interpretability:** When clinicians are provided with visual or textual "explanations" of an AI's diagnostic logic—such as identifying specific linguistic markers for depression—the rate of clinical adoption increases by approximately **40%** (Shakeel et al., 2025).
- **Safety in ADHD Diagnostics:** In the context of ADHD, XAI frameworks allow practitioners to cross-reference algorithmic outputs with traditional clinical features. This reduces the risk of false positives generated by non-representative training data and enhances the clinician's understanding of complex cases (Shakeel et al., 2025).
- **Regulatory Demands:** Under updated legal interpretations of the GDPR, patients now possess an implicit "right to explanation" regarding automated mental health assessments. This has forced developers to prioritize transparency over proprietary algorithmic complexity to maintain market access (Hosono et al., 2024; Srivastava & Goswami, 2026).

3.6 Data Privacy and the Metadata Vulnerability

While traditional encryption methods are standard in 2026, the synthesized evidence highlights a growing vulnerability regarding "behavioral metadata" (Ali et al., 2025; Srivastava & Goswami, 2026).

- **Beyond Encryption:** Current frameworks like the GDPR provide robust protection for explicit health data, yet they offer only partial protection for metadata—such as typing speed, sleep-wake cycles, and app interaction frequency. This data can be used to reconstruct a patient's psychological profile without their explicit consent (Ali et al., 2025; Hosono et al., 2024).
- **Algorithmic Profiling:** AI-driven profiling poses a risk of "digital stalking," where subtle shifts in a user's digital behavior are monitored by third-party entities. This potentially leads to discriminatory outcomes in insurance or employment based on predicted mental health trajectories (Srivastava & Goswami, 2026; Volpe et al., 2025).
- **Sovereignty vs. Scale:** The Chinese DTx market demonstrates high-speed scaling through integration with massive social media ecosystems; however, this model frequently prioritizes data utilization over individual data sovereignty, creating a sharp ethical contrast with European patient-centric models (Jiang et al., 2025).

3.7 The Hybrid "Human-in-the-Loop" (HITL) Model

The most significant finding regarding the efficacy of DTx in the 2025–2026 period is the superior performance of hybrid care models over purely automated systems (Campellone et al., 2025; Eiselt et al., 2025).

- **Retention Parity:** While self-guided interventions suffer from attrition rates exceeding 50% within the first month, hybrid models—where a human "digital navigator" provides intermittent support—show adherence rates as high as **80%** (Nwosu et al., 2022; Smith et al., 2025).
- **Therapeutic Alliance:** The results confirm that a digital interface alone cannot replace the empathic resonance required for a strong therapeutic alliance. However, when used correctly, DTx serves as a powerful "force multiplier" for human therapists, extending the reach of clinical protocols (Campellone et al., 2025).
- **Risk Mitigation:** The inclusion of a "Human-in-the-loop" serves as a critical safety guardrail, ensuring that acute crises or "AI hallucinations"—instances where GenAI provides inappropriate or dangerous advice—are identified and corrected in real-time (Carlbring et al., 2018; Volpe et al., 2025).

3.8 Economic and Implementation Outcomes

The financial impact of DTx on public health systems is characterized by high initial implementation costs but significant long-term systemic savings (Fürstenau et al., 2023; Gensorowsky et al., 2024).

- **Cost-Effectiveness:** Large-scale meta-analyses confirm that validated DTx platforms reduce the overall cost of mental health care by approximately **20%** over a five-year period by preventing hospitalizations and reducing the need for high-intensity face-to-face sessions (Seo et al., 2024).
- **Reimbursement Bottlenecks:** In the German DiGA market, the primary economic barrier is not the lack of funding, but the administrative burden placed on primary care physicians to learn, prescribe, and monitor these digital tools (Gensorowsky et al., 2024; Hengst et al., 2023).
- **Market Fragmentation:** The lack of global standardization in clinical trial designs for DTx remains a barrier to cross-border market access. Different jurisdictions require different levels of "drug-like" evidence before granting reimbursement status, hindering international scalability (Hengst et al., 2023; Jiang et al., 2025).

4. Discussion

4.1 Theoretical Synthesis: A Socio-Technical Perspective

The findings of this review necessitate a departure from the reductionist view of the "digital divide" as a mere lack of hardware. From a Socio-Technical Systems (STS) perspective, the successful adoption of Digital Therapeutics (DTx) requires the "joint optimization" of both the technical subsystem—encompassing AI precision, UI/UX, and data security—and the social subsystem, which includes clinician trust, patient health literacy, and cultural capital (Eiselt et al., 2025; Fanta & Pretorius, 2023).

When a sophisticated technical intervention is introduced into a fractured social environment, it does not act as a neutral tool; instead, it functions as a "magnifier of existing inequities" (Piers et al., 2023; Volpe et al., 2025). This is particularly evident in the "Benefit Gap," where marginalized populations may engage with a platform but derive significantly less clinical value due to the lack of culturally adapted content and the absence of social support structures (Fanta & Pretorius, 2023; Graham et al., 2019; Hengst et al., 2023).

4.2 The Myth of Digital Democratization and the "Policy Paradox"

While DTx is frequently marketed as a tool for democratizing mental health care, current evidence suggests a "Policy Paradox." Even in advanced systems with universal healthcare coverage, such as Germany's DiGA framework, the removal of financial barriers is only an initial step (Gensorowsky et al., 2024).

Invisible barriers—including neighborhood deprivation, ethnic marginalization, and the "clinician-as-gatekeeper" effect—persist. Reports from early 2025 indicate that while over **861,000 DiGAs** were utilized in Germany by late 2024, awareness in the general population remained low, with usage heavily skewed toward individuals with higher health awareness and specific educational backgrounds (Gensorowsky et al., 2024; Hosono et al., 2024). For digital equity to be achieved, policy models must transition from "Equality of Access" to "Equity of Outcome," necessitating disproportionately more "human-in-the-loop" support for patients with lower digital literacy (Calderone et al., 2025; Divatia et al., 2026; Graham et al., 2019).

4.3 Algorithmic Justice: Addressing the "Black Box" in Psychiatry

The evolution of AI in psychiatry between 2020 and 2026 has shifted the ethical debate from data privacy to algorithmic justice (Rahsepar Meadi et al., 2025; Srivastava & Goswami, 2026). The "black box" nature of early diagnostic algorithms led to significant clinician resistance, as seen in the 2024–2025 literature (Abbas et al., 2025; Shakeel et al., 2025).

- **The Necessity of XAI:** Explainable AI (XAI) is no longer a technical luxury but a clinical necessity. Studies show that providing interpretability in AI-driven ADHD or depression diagnostics can increase physician trust and adoption rates by up to **40%** (Shakeel et al., 2025).
- **Algorithmic Bias:** Bias is often "baked" into the lifecycle of DTx through non-representative training data. Unless checked by proactive interventions like adversarial debiasing, unbalanced algorithms will widen current inequities rather than narrowing them (Rahsepar Meadi et al., 2025; Volpe et al., 2025).
- **The Right to Explanation:** Under legal standards such as the GDPR, patients are gaining an implicit right to understand the logic behind automated psychological profiling. This is forcing manufacturers to prioritize "Ethics-by-Design" over proprietary algorithmic complexity (Srivastava & Goswami, 2026; Volpe et al., 2025).

4.4 Clinician Readiness and the Knowledge Divide

A striking theme in the 2024–2026 literature is the role of the healthcare provider as a systemic bottleneck (Gensorowsky et al., 2024; Volpe et al., 2025). If the clinician lacks confidence in the technical subsystem, the entire care model defaults to traditional, overstretched methods (Abbas et al., 2025; Calderone et al., 2025).

- **Implicit Bias:** Providers may subconsciously prioritize DTx for patients who fit the "ideal user" stereotype—those who are young, affluent, and articulate—while overlooking older adults or marginalized groups who might benefit most from the flexibility of digital care (Graham et al., 2019; Piers et al., 2023).
- **Systemic Redesign:** Addressing the digital divide requires a systems-level redesign of medical education. Digital proficiency, algorithmic literacy, and bias-awareness must become core competencies for the modern healthcare workforce to ensure equitable prescribing practices (Eiselt et al., 2025; Hosono et al., 2024).

4.5 Human-Centered Design and the Adherence Crisis

From a systems engineering perspective, high attrition rates represent a failure of Human-Centered Design (Divatia et al., 2026; Nwosu et al., 2022).

- **The "Average User" Fallacy:** By designing for a hypothetical "average user," the industry effectively excludes those with sensory, cognitive, or linguistic differences (Piers et al., 2023; Rahsepar Meadi et al., 2025).
- **The Therapeutic Alliance:** A digital interface cannot replace the empathic resonance of a human therapist. Instead, it must be treated as a "force multiplier" (Calderone et al., 2025).
- **Hybrid Models:** Consensus in 2025–2026 confirms that hybrid "blended care" models, integration of automated tools with "human-in-the-loop" oversight, are the only sustainable path to reducing the adherence crisis. Estimates suggest that by 2028, **80% of patients** will utilize some form of hybrid care to ensure sustained engagement and clinical safety (Campellone et al., 2025; Divatia et al., 2026; Smith et al., 2025).

5. Case Studies & Comparative Analysis

5.1 The German DiGA Model: Institutionalizing Innovation

Germany's approach to Digital Therapeutics (DTx) serves as the primary global case study for successful structural integration into public health (Fürstenau et al., 2023; Gensorowsky et al., 2024). The Digitale Gesundheitsanwendungen (DiGA) framework allows for the direct prescription of clinical software, which is then reimbursed by statutory health insurance (Gensorowsky et al., 2024; Srivastava & Goswami, 2026).

- **Regulatory Fast-Track:** The model focuses on "positive healthcare effects" (Fast-Track), requiring developers to prove that their software improves clinical outcomes or enhances structural and procedural processes within a specific timeframe (Gensorowsky et al., 2024; Hosono et al., 2024).
- **The Adoption Bottleneck:** Despite robust financial coverage, adoption rates in 2024–2026 reveal that the "clinician bottleneck" remains a significant barrier. Many physicians feel under-compensated for the additional administrative time required to learn, prescribe, and monitor patient data generated by these platforms (Gensorowsky et al., 2024; Volpe et al., 2025).
- **System Dynamics:** From an STS perspective, the German model has successfully optimized the **Technical Subsystem** through rigorous regulation. However, it continues to struggle with the **Social Subsystem**, specifically regarding provider readiness and the integration of digital tools into established clinical workflows (Fanta & Pretorius, 2023; Gensorowsky et al., 2024).

5.2 The Chinese Market: Scale, Integration, and Data Privacy

In contrast to the European model, the Chinese DTx market is characterized by rapid, government-supported scalability and deep integration with multi-purpose digital ecosystems such as WeChat and Alipay (Jiang et al., 2025).

- **High-Speed Scaling:** China has approved a significantly higher volume of psychiatric DTx applications compared to Europe, prioritizing massive accessibility for its vast population to address a high demand for mental health services (Jiang et al., 2025).
- **Data Sovereignty Concerns:** The integration of therapeutic tools into multi-functional social media platforms raises critical ethical questions. The potential for "algorithmic profiling" and the secondary use of data contrasts sharply with the GDPR-centric, privacy-first European approach (Hosono et al., 2024; Jiang et al., 2025; Srivastava & Goswami, 2026).
- **Standardization Gap:** While scale is high, studies from 2025 identify a lack of international clinical alignment. This leads to high variability in the quality of therapeutic content and challenges in cross-border clinical validation (Jiang et al., 2025).

5.3 Comparative Efficacy: Digital vs. Traditional Therapy

The debate between traditional face-to-face therapy and DTx has evolved into a nuanced analysis of "contextual non-inferiority" (Calderone et al., 2025; Carlbring et al., 2018).

- **Clinical Parity:** Comprehensive meta-analyses confirm that for mild-to-moderate depression and anxiety, digital CBT (dCBT) provides clinical outcomes that are statistically equivalent to traditional face-to-face therapy (Carlbring et al., 2018; Seo et al., 2024).
- **Therapeutic Alliance:** The primary differentiator remains the human element. While DTx can effectively replicate clinical protocols, it often fails to sustain the "therapeutic alliance" required for complex psychiatric conditions, leading to significantly higher attrition rates in unguided interventions (Calderone et al., 2025; Nwosu et al., 2022; Smith et al., 2025).
- **Hybrid Superiority:** Evidence from 2025–2026 suggests that the most effective model is neither purely digital nor purely traditional. Instead, a "blended care" approach—combining automated digital tools with strategic human facilitation—achieves the highest adherence and superior long-term recovery rates (Eiselt et al., 2025; Smith et al., 2025).

6. Future Innovations

6.1 Generative AI and Large Language Models (LLMs)

The integration of Generative AI (GenAI) into psychiatry represents the most significant technical shift in the 2024–2026 period (Ali et al., 2025; Torous & Topol, 2025). Unlike earlier rule-based chatbots, these systems utilize advanced Large Language Models to provide dynamic, context-aware support.

- **Empathetic Interfaces:** LLMs allow for conversational interfaces that surpass previous systems in their ability to simulate empathetic resonance and nuanced understanding. This advancement is specifically designed to address the "engagement gap" by fostering a stronger sense of digital therapeutic alliance (Campellone et al., 2025; Torous & Topol, 2025).
- **Technical Guardrails:** Modern GenAI-driven DTx products now incorporate sophisticated "safety layers" and retrieval-augmented generation (RAG) to prevent AI hallucinations. These guardrails ensure that the therapeutic advice remains strictly within clinical, evidence-based boundaries, providing real-time corrections if the model deviates from established protocols (Campellone et al., 2025; Volpe et al., 2025).

6.2 VR/AR and Immersive Psychiatry

Virtual and Augmented Reality (VR/AR) have successfully transitioned from niche experimental applications to mainstream therapeutic tools by 2026 (Ali et al., 2025; Torous & Topol, 2025).

- **Exposure Therapy:** VR is increasingly the gold standard for treating phobias and PTSD. It offers controlled, immersive, and highly customizable environments that research shows are more effective and easier to tolerate than traditional "in-imagination" exposure therapy (Torous & Topol, 2025).
- **Cognitive Enhancement:** AR-driven "behavioral nudging" is being deployed to assist patients with ADHD or age-related cognitive decline. These systems provide real-time holographic prompts and environmental cues to help patients manage daily tasks and improve executive functioning in their natural surroundings (Kuhn et al., 2024; Shakeel et al., 2025).

6.3 Explainable AI (XAI) as the Industry Standard

By 2026, Explainable AI (XAI) has moved from a specialized research topic to a baseline regulatory and clinical requirement (Abbas et al., 2025; Shakeel et al., 2025).

- **Diagnostic Trust:** Clinicians now demand that AI-driven diagnostics provide "traceable logic" for their conclusions. By making the feature-weighting and decision-making process of the algorithm transparent, human practitioners can validate outputs against their own expertise, ensuring the "human-in-the-loop" remains the ultimate decision-maker (Abbas et al., 2025; Shakeel et al., 2025).
- **Bias Mitigation:** XAI serves as a critical diagnostic tool for the technology itself. It allows developers and auditors to identify and correct "algorithmic biases" that have historically marginalized ethnic and socioeconomic minorities by pinpointing which non-representative data points are skewing results (Rahsepar Meadi et al., 2025; Volpe et al., 2025).

Table 2. Transition from Black-Box AI to Explainable AI (XAI)

Feature	Black-Box AI (Traditional)	Explainable AI (XAI - 2026 Standard)
Transparency	Diagnostic logic is hidden from the clinician (Abbas et al., 2025; Shakeel et al., 2025).	Provides "traceable logic" for clinical conclusions (Abbas et al., 2025; Shakeel et al., 2025).
Trust Level	High physician skepticism and lower adoption (Abbas et al., 2025; Shakeel et al., 2025).	Increases clinical adoption rates by approximately 40% (Shakeel et al., 2025).
Bias Detection	Difficult to identify skewed data points (Rahsepar Meadi et al., 2025; Volpe et al., 2025).	Allows auditors to pinpoint and correct algorithmic biases (Rahsepar Meadi et al., 2025; Volpe et al., 2025).
Clinical Safety	Risk of "AI hallucinations" or inappropriate advice (Campellone et al., 2025; Volpe et al., 2025).	Enables "human-in-the-loop" to validate outputs against expertise (Abbas et al., 2025; Shakeel et al., 2025).
Legal Status	Potential conflict with GDPR "right to explanation" (Srivastava & Goswami, 2026).	Prioritizes "Ethics-by-Design" to maintain market access (Srivastava & Goswami, 2026; Volpe et al., 2025).

7. Conclusions

The systematic analysis of the evolution of Digital Therapeutics (DTx) between 2020 and 2026 demonstrates that while the technical subsystem of digital psychiatry has reached a high level of clinical maturity (Carlbring et al., 2018; Seo et al., 2024), the social subsystem remains fragmented and unequal (Eiselt et al., 2025; Piers et al., 2023). Grounded in the Socio-Technical Systems (STS) framework, this review highlights that the "digital divide" in mental health care is no longer a mere issue of hardware access, but a complex crisis of socioeconomic stratification and digital health literacy (Hengst et al., 2023; Piers et al., 2023).

The findings indicate that the rapid integration of Artificial Intelligence, particularly Generative AI and Large Language Models, offers a promising pathway to addressing the long-standing "engagement gap" by providing more empathetic and personalized interfaces (Campellone et al., 2025; Torous & Topol, 2025). However, these advancements are accompanied by significant ethical risks, specifically regarding algorithmic bias and the potential for "design-driven exclusion" of marginalized populations (Rahsepar Meadi et al., 2025; Volpe et al., 2025).

The evidence gathered throughout the 2024–2026 period underscores that purely automated interventions frequently suffer from unsustainable attrition rates (Nwosu et al., 2022), whereas hybrid "human-in-the-loop" models consistently achieve superior adherence and clinical outcomes (Calderone et al., 2025; Divatia et al., 2026).

Furthermore, the comparison between global regulatory models—specifically the institutionalized German DiGA framework and the rapidly scaling Chinese market—reveals that financial reimbursement alone cannot bridge the health equity gap (Gensorowsky et al., 2024; Jiang et al., 2025). Meaningful democratization of digital care requires a holistic transition toward Algorithmic Justice and Explainable AI (XAI), ensuring that software-driven interventions are transparent, culturally adapted, and clinically justifiable to both patients and providers (Abbas et al., 2025; Kuhn et al., 2024; Shakeel et al., 2025).

In conclusion, the next decade of DTx evolution must prioritize human-centered systems engineering over mere algorithmic complexity. To ensure that innovation acts as a bridge to health equity rather than a barrier to care, stakeholders must focus on "joint optimization"—aligning technological capabilities with the social realities of diverse global populations.

Author Contributions:

- **Conceptualization:** Alicja Maria Mitan, Karolina Magda Leszczyńska
- **Methodology:** Weronika Napierała, Aleksandra Maria Tomaszewska, Kamila Teresa Kańska
- **Software:** Kamila Teresa Kańska, Anna Krzysztofik, Karolina Krawczyk
- **Validation:** Maciej Tomasz Wieczorek, Jeremi Leon Jasiński, Karolina Krawczyk
- **Formal Analysis:** Karolina Julia Hak, Karolina Krawczyk
- **Investigation:** Alicja Maria Mitan, Aleksandra Maria Tomaszewska, Weronika Napierała
- **Resources:** Karolina Magda Leszczyńska, Jeremi Leon Jasiński
- **Data Curation:** Alicja Maria Mitan, Anna Krzysztofik, Maciej Tomasz Wieczorek
- **Writing - Original Draft:** Jeremi Leon Jasiński, Maciej Tomasz Wieczorek
- **Writing - Review & Editing:** Anna Krzysztofik, Karolina Julia Hak
- **Visualization:** Karolina Magda Leszczyńska, Weronika Napierała
- **Supervision:** Karolina Julia Hak, Kamila Teresa Kańska
- **Project Administration:** Aleksandra Maria Tomaszewska

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

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

Declaration of the use of generative AI and AI-assisted technologies in the writing process: In preparing this work, the authors used ChatGPT for the purpose of improving language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

REFERENCES

1. Abbas, S. R., Seol, H., Abbas, Z., & Lee, S. W. (2025). Exploring the role of artificial intelligence in smart healthcare: A capability and function-oriented review. *Healthcare, 13*(14), 1642. https://doi.org/10.3390/healthcare13141642
2. Ali, M., Ali, S., Abbas, Q., Abbas, Z., & Lee, S. W. (2025). Artificial intelligence for mental health: A narrative review of applications, challenges, and future directions in digital health. *Digital Health, 11*. https://doi.org/10.1177/20552076251395548
3. Calderone, A., Latella, D., & Fauci, E. L. (2025). Mind meets machine: A narrative review of artificial intelligence role in clinical psychology practice. *Clinical Psychology & Psychotherapy, 32*(6), e2984. https://doi.org/10.1002/cpp.70191
4. Campellone, T. R., Flom, M., & Montgomery, R. M. (2025). Safety and user experience of a generative artificial intelligence digital mental health intervention: Exploratory randomized controlled trial. *Journal of Medical Internet Research, 27*, e65432. https://doi.org/10.2196/67365
5. Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *Cognitive Behaviour Therapy, 47*(1), 1–18. https://doi.org/10.1080/16506073.2017.1401115
6. Divatia, S., Foster, T., Cui, S., Tapiavala, V., & Fortuna, K. L. (2026). Accessibility and equity considerations in the clinical validation of prescription digital therapeutics: A scoping review. *mHealth, 12*, 12. https://doi.org/10.21037/mhealth-25-12
7. Eiselt, A. K., Kirkendall, S., Xiong, E., & Langner, D. (2025). Achieving clinically meaningful outcomes in digital health: A six-step, cyclical precision engagement framework (ENGAGE). *Frontiers in Digital Health, 7*, 1713334. https://doi.org/10.3389/fdgth.2025.1713334
8. Fanta, G. B., & Pretorius, L. (2023). Sociotechnical factors of sustainable digital health systems: A system dynamics model. *Health Policy and Technology, 12*(1), 100729. https://doi.org/10.1016/j.hlpt.2023.100729
9. Friis-Healy, E. A., Nagy, G. A., & Kollins, S. H. (2021). It is time to REACT: Opportunities for digital mental health apps to reduce mental health disparities in racially and ethnically minoritized groups. *JMIR Mental Health, 8*(1), e25456. https://doi.org/10.2196/25456
10. Fürstenau, D., Gersch, M., & Schreiter, S. (2023). Digital therapeutics (DTx). *Business & Information Systems Engineering, 65*, 349–360. https://doi.org/10.1007/s12599-023-00804-z
11. Gensorowsky, D., Diekmannshemke, J., Witte, J., Rohn, I., & Schneider, U. (2024). PT23 The evolving German DiGA market: Insights from a claims data analysis. *Value in Health, 27*(12), S352. https://doi.org/10.1016/j.jval.2024.10.3825
12. Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports, 21*(11), 116. https://doi.org/10.1007/s11920-019-1094-x
13. Hengst, T. M., Lechner, L., Dohmen, D., & Bolman, C. A. (2023). The facilitators and barriers of mHealth adoption and use among people with a low socio-economic position: A scoping review. *Digital Health, 9*. https://doi.org/10.1177/20552076231198702
14. Hosono, T., Niwa, Y., & Kondoh, M. (2024). Comparison of product features and clinical trial designs for the DTx products with the indication of insomnia authorized by regulatory authorities. *Therapeutic Innovation & Regulatory Science, 58*(6), 1138–1147. https://doi.org/10.1007/s43441-024-00684-9
15. Jiang, N., Yu, X., Yang, Y., Li, G., He, C., Wang, M. P., ... & Rao, K. (2025). Digital therapeutics in China: Comprehensive review. *Journal of Medical Internet Research, 27*, e70955. https://doi.org/10.2196/70955
16. Kuhn, E., Saleem, M., Klein, T., Köhler, C., Fuhr, D. C., Lahutina, S., ... & Böge, K. (2024). Interdisciplinary perspectives on digital technologies for global mental health. *PLOS Global Public Health, 4*(2), e0002867. https://doi.org/10.1371/journal.pgph.0002867
17. Nwosu, A., Boardman, S., Husain, M. M., & Doraiswamy, P. M. (2022). Digital therapeutics for mental health: Is attrition the Achilles heel? *Frontiers in Psychiatry, 13*, 900615. https://doi.org/10.3389/fpsyt.2022.900615
18. Piers, R., Williams, J. M., & Sharpe, H. (2023). Can digital mental health interventions bridge the 'digital divide' for socioeconomically and digitally marginalised youth? A systematic review. *Child and Adolescent Mental Health, 28*(1), 90–104. https://doi.org/10.1111/camh.12620
19. Rahsepar Meadi, M., Sillekens, T., Metselaar, S., van Balkom, A., Bernstein, J., & Batelaan, N. (2025). Exploring the ethical challenges of conversational AI in mental health care: A scoping review. *JMIR Mental Health, 12*, e60432. https://doi.org/10.2196/60432
20. Seo, Y. C., Yong, S. Y., Choi, W. W., & Kim, S. H. (2024). Meta-analysis of studies on the effects of digital therapeutics. *Journal of Personalized Medicine, 14*(2), 157. https://doi.org/10.3390/jpm14020157
21. Shakeel, H. M., Antoniou, G., & Adamou, M. (2025). Integrating explainable AI with clinical features to enhance ADHD diagnostic understanding. *Frontiers in Psychiatry, 16*, 1706216. https://doi.org/10.3389/fpsyt.2025.1706216
22. Smith, K. A., Ward, T., Lambe, S., et al. (2025). Engagement and attrition in digital mental health: Current challenges and potential solutions. *npj Digital Medicine, 8*, 398. https://doi.org/10.1038/s41746-025-01778-w

23. Srivastava, A., & Goswami, A. K. (2026). Data protection and mental health privacy: Legal standards for AI-powered psychological assessment and identity crisis intervention. In *Imposter syndrome and AI: Navigating human identity in the age of intelligent machines* (pp. 105–122). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-6618-0.ch006
24. Torous, J., & Topol, E. J. (2025). Assessing generative artificial intelligence for mental health. *The Lancet, 406*(10504), 683. https://doi.org/10.1016/S0140-6736(25)01237-1
25. Volpe, U., Ramalho, R., & Gaebel, W. (2025). Risks of digitalization in mental health care. *Neuropsychiatrie*. https://doi.org/10.1007/s40211-025-00551-5