

53

Ethical, Technical, and Global Equity Challenges in AI-Powered Mental Health A Systematic Review

Sara Taher

An Najah National University, Nablus, Palestine

Abstract

The integration of artificial intelligence (AI) into mental health care marks a transformative shift in diagnostic and therapeutic paradigms, yet its ethical and technical implications demand rigorous scrutiny. The development, validation, and deployment of AI systems have shown persistent disparities that stem from biases in training datasets which include over representation of Western populations and exclusion of nonbinary identities. Technical challenges, including variable diagnostic accuracy (21%–100%) and opaque algorithmic decision-making, hinder clinical adoption. The ethical concerns such as privacy risks, data ownership ambiguities, and accountability gaps remain unresolved across global contexts. The review synthesizes peer-reviewed studies from 2020 to 2025 to critically analyze methodological limitations, mitigation strategies (e.g. Strategies like SHAP explanations and adversarial debiasing are proposed to mitigate bias), and interdisciplinary approaches to ensure equitable access. The analysis highlights the urgent need for standardized benchmarks, culturally inclusive datasets, and human-AI collaboration models to address enduring inequities in mental health technology