**ARTIFICIAL INTELLIGENCE AND DIGITAL THERAPY FOR ADOLESCENT MENTAL HEALTH IN THE UK; OPPORTUNITIES, BARRIERS, AND ETHICAL CONSIDERATION**

*Abstract*

*Background: Adolescence constitutes a critical developmental stage marked by the onset of mental health difficulties, yet timely access to effective mental health care remains a significant challenge for many adolescents in the United Kingdom (UK). Artificial intelligence (AI)-enabled digital therapies present innovative opportunities to address these gaps.*

*Objective: This systematic review critically assesses current evidence on AI-driven digital interventions for adolescent mental health within the UK, highlighting their potential opportunities, barriers to implementation, and pertinent ethical considerations.*

*Methods: Employing a mixed-methods design, a systematic literature review adhering to PRISMA guidelines was combined with thematic analysis of semi-structured interviews. Comprehensive database searches (MEDLINE, PsycINFO, Web of Science; 2013–2023) targeted studies involving UK adolescents (ages 11–19) using AI-based mental health technologies. Included studies underwent rigorous quality appraisal (Cochrane RoB 2.0, ROBINS-I, CASP). Additional insights were gathered through stakeholder interviews (clinicians, AI developers, adolescent users).*

*Results: Twenty-seven studies met inclusion criteria, investigating interventions such as AI chatbots, predictive analytics, mobile apps, and virtual environments targeting anxiety and depression. Key opportunities identified include enhanced accessibility for underserved populations, personalization through adaptive algorithms, proactive early-risk detection, scalability, cost-efficiency, and improved engagement via interactive interfaces. Significant implementation barriers encompassed technical infrastructure limitations, data security concerns, insufficient longitudinal efficacy data, socioeconomic disparities, and clinician scepticism. Ethical challenges emphasized informed consent, algorithm transparency, potential biases, unclear accountability, and clinician deskilling risks.*

*Conclusions: AI-driven digital interventions offer substantial promise for augmenting adolescent mental health services in the UK. However, realizing their full potential necessitates addressing infrastructural, ethical, and evidentiary challenges through robust governance frameworks and continued rigorous research.*

*Keywords: Adolescent mental health; Artificial intelligence; Digital therapy; Accessibility; Ethics; United Kingdom*

*Introduction*

*Adolescence represents a pivotal developmental phase characterized by significant physical, emotional, and social transformations, frequently coinciding with the emergence of mental health disorders, including anxiety, depression, and behavioural issues (Steinberg, 2014). Within the United Kingdom (UK), mental health difficulties among adolescents have notably increased, adversely affecting their overall well-being, educational outcomes, and social integration (Office for Health Improvement and Disparities, 2022). Despite substantial prevalence and evident need, timely and effective mental health interventions remain inaccessible for many adolescents due to structural impediments such as lengthy waitlists for specialized services, insufficient availability of qualified child and adolescent mental health professionals, and persistent societal stigma around mental illness (Royal Society for Public Health, 2018). Consequently, there is an urgent imperative for innovative approaches to enhance adolescent mental health support.*

*Recently, artificial intelligence (AI)-based digital therapies have garnered considerable attention as potential mechanisms for expanding and improving adolescent mental health services (Torous et al., 2021). AI encompasses computational technologies designed to replicate human cognitive functions, such as data-driven learning, pattern identification, and complex decision-making processes. In mental health contexts, AI has been integrated into various digital interventions, including conversational chatbots providing therapeutic dialogue, mobile applications supporting mood tracking and psychoeducation, predictive machine learning algorithms for early identification of risks, and adaptive virtual reality environments tailored to individual user interactions. Such AI-enabled tools offer distinct advantages over conventional face-to-face therapeutic approaches by delivering accessible, continuous, and on-demand support via widely available platforms like smartphones and computers. For instance, AI-driven chatbots can provide immediate, round-the-clock counselling without traditional barriers like scheduling constraints or geographic limitations. Furthermore, these systems can enhance user engagement through personalized interventions, tailored dynamically by analysing individual user profiles and behavioural data.*

*Nevertheless, comprehensive evaluations of these emerging AI digital interventions are essential. Questions regarding their clinical efficacy, user and provider acceptability, data privacy concerns, algorithmic fairness, and suitability for sensitive clinical contexts persist. Current research, while promising, remains fragmented, necessitating consolidated analyses within the specific regulatory and healthcare landscape of the UK.*

*Aim and Objectives: The primary aim of this study was to systematically review existing evidence on AI-driven digital therapies for adolescent mental health in the UK and explore expert and user perspectives. Specific objectives included identifying opportunities for enhancing adolescent mental health care through AI (e.g., improved accessibility, personalization, preventive capabilities), examining barriers and challenges to implementation, and delineating ethical and governance considerations. The review seeks to inform policymakers, clinicians, and developers about the benefits and limitations of AI in adolescent mental health, emphasizing implications within the NHS framework, UK socio-cultural contexts, and relevant national policies.*

*Methodology*

*Design and Search Strategy*

*We conducted a mixed-methods systematic review, integrating findings from quantitative studies with qualitative insights from stakeholder interviews. The review process followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021) to ensure methodological rigor and transparency. A comprehensive literature search was performed across multiple electronic databases, including MEDLINE (PubMed), PsycINFO, and Web of Science, to capture relevant studies published in the last 10 years (January 2013 through December 2023). This timeframe was chosen to reflect the contemporary era of AI advancements and digital health proliferation. The search strategy combined keywords and controlled vocabulary for three concept areas: (1) artificial intelligence (e.g. “machine learning,” “chatbot,” “algorithm”), (2) digital therapy/intervention (e.g. “mobile app,” “online therapy,” “digital health”), and (3) adolescents and mental health (e.g. “adolescent,” “youth,” specific conditions like “depression,” “anxiety”). We restricted results to studies conducted in the UK or whose samples predominantly involved UK populations, to maintain contextual relevance. Reference lists of relevant articles and grey literature sources (e.g. government and policy reports) were also hand-searched to identify additional publications. The search strategy and eligibility criteria are shown in Appendix A and Table 1 respectively.*

*Quality Assessment*

*To appraise the methodological quality of included studies, we employed established critical appraisal tools appropriate to each study design. For randomized trials, the Cochrane Risk of Bias 2.0 tool was used to evaluate biases in domains such as randomization, deviations from interventions, missing data, and outcome measurement. Non-randomized and observational studies were assessed with the ROBINS-I (Risk of Bias in Non-randomised Studies – of Interventions) instrument, focusing on confounding, selection, classification of interventions, and outcome reporting. Qualitative studies were appraised using the Critical Appraisal Skills Programme (CASP) qualitative checklist, examining credibility, relevance, and rigor in data collection and analysis. We also drew on the Consolidated Criteria for Reporting Qualitative Research (COREQ) as a reference to ensure comprehensive reporting of the qualitative components (e.g. reflexivity, saturation, and triangulation).*

*Qualitative Interviews and Data Synthesis*

*In parallel with the literature review, we conducted semi-structured interviews to gather experiential and contextual insights from key stakeholders in the field. A purposive sample of participants (N = 15) was recruited, comprising five mental health professionals (clinicians or service managers) working with adolescents in the NHS, five developers or researchers specializing in AI mental health technologies, and five adolescent end-users who had experience with digital mental health apps or platforms. The interview guides were tailored to each group but broadly covered perceptions of AI in mental health care, perceived benefits and risks, barriers to adoption in real-world practice, and ethical or practical concerns (particularly around data use and consent). Interviews were conducted either in person or via secure video conferencing, recorded with consent, and transcribed verbatim. Using a thematic analysis approach, transcripts were coded inductively by two researchers to identify recurrent themes and divergent viewpoints. The qualitative analysis emphasized identifying how stakeholders view the opportunities, barriers, and ethical issues of AI-driven therapy in adolescent mental health, complementing the evidence identified in the literature.*

*Finally, we adopted a convergent synthesis approach to integrate findings from the quantitative literature and qualitative interviews. Both strands of evidence were analysed concurrently, and results were mapped onto the review’s key objectives. The synthesis was organized narratively under thematic headings corresponding to Opportunities, Barriers, and Ethical Considerations. Literature findings (e.g. effectiveness data, reported outcomes, and usage statistics) were juxtaposed with stakeholder insights to provide a comprehensive understanding of each theme. This triangulation allowed us to highlight points of consensus, reinforce findings with multi-faceted evidence, and also note any discrepancies between the published evidence base and frontline experiences. In reporting results, we prioritize themes that emerged strongly across multiple studies or data sources. Where appropriate, quantitative data (such as effect sizes or usage rates) are summarized to illustrate the magnitude of findings, and illustrative quotes from interviews are referenced to add depth to the quantitative results. This mixed-methods integration strengthens the review by not only summarizing what has been achieved with AI in adolescent mental health, but also why certain challenges exist and how they might be addressed in the UK context.*

*Results*

*Study Selection and Characteristics of Included Studies*

*The systematic search and selection process, adhering to PRISMA guidelines (identification, screening, eligibility, inclusion), yielded 27 eligible studies that provided a comprehensive overview of artificial intelligence (AI) applications in adolescent mental health within the UK. Included studies employed diverse methodologies, encompassing randomized controlled trials (RCTs), pilot interventions, observational and cross-sectional analyses of digital usage, qualitative interviews with adolescents and clinicians, and mixed-methods approaches. The majority assessed AI-driven digital platforms targeting prevalent adolescent mental health concerns, such as anxiety, depression, and overall emotional wellness. Notably, interventions included AI-driven chatbots providing psychoeducational content and cognitive-behavioural therapy (CBT), mobile applications employing machine learning for personalized mood tracking and recommendations, and predictive analytics using social media and sensor data to identify early risk indicators. A smaller number explored specialized interventions, including AI-enhanced virtual reality exposure therapy and intelligent tutoring systems for mental health education.*

*Sample sizes varied considerably, with quantitative studies frequently including hundreds of adolescent participants and qualitative studies typically involving 10–50 respondents. While all included research was relevant to the UK context, several incorporated data from multicentre studies or international platforms with UK participants, emphasizing the global reach yet localized adaptation requirements. Overall study quality was moderate to high; several RCTs exhibited low risk of bias, providing robust efficacy evidence, albeit generally short-term. Qualitative studies contributed meaningful insights into user experience and ethical considerations. Findings were synthesized narratively around themes of opportunities, barriers, and ethics, contextualized to the UK healthcare setting.*

*Opportunities of AI-Driven Digital Therapies*

*Analysis of the literature and stakeholder insights revealed several significant opportunities through which artificial intelligence (AI)-enabled digital interventions could notably improve adolescent mental health care in the UK. These include enhanced accessibility and reach, personalized care, early intervention capabilities, scalability and cost-effectiveness, and improved patient engagement, collectively addressing existing limitations within conventional mental health services.*

*Improving Accessibility and Reach: AI-driven digital therapies offer unprecedented opportunities to expand access to mental health services, especially for underserved adolescents hindered by geographical, logistical, or social barriers. Digital platforms transcend conventional constraints such as geographic location and limited clinic hours, providing continuous, immediate support. Adolescents in rural or underserved urban areas, where professional mental health resources are scarce, can access AI-powered applications and chatbots via smartphones, facilitating timely intervention without the delay of conventional appointments. For example, adolescents residing in remote Welsh communities can utilize AI mental health apps to receive guided interventions and preliminary assessments, enhancing service equity (Royal Society for Public Health, 2018). Additionally, these digital tools offer anonymity, reducing stigma-related barriers, thereby engaging adolescents hesitant about traditional face-to-face consultations. Clinician interviews underscored the ability of AI interventions to effectively engage youth previously underserved due to sub-threshold clinical symptoms or reluctance towards traditional therapy settings, aligning closely with NHS objectives for universal adolescent mental health access.*

*Personalization of Interventions: AI’s capability to process extensive user-generated data allows interventions to be highly personalized, significantly enhancing their effectiveness and user satisfaction. AI-driven systems adapt content dynamically based on real-time analysis of user input, including mood logs and behavioural patterns, to deliver personalized therapeutic recommendations. For instance, AI applications detecting elevated anxiety levels through self-reported mood data and sensor inputs may proactively suggest timely, relevant coping mechanisms, contrasting sharply with traditional static therapeutic resources (Gooding et al., 2021). Such personalization accommodates the diverse mental health needs, preferences, and cultural backgrounds among adolescents, maintaining engagement and relevance. Interviewees particularly highlighted natural language processing (NLP) technologies that enable chatbots to communicate in relatable and personalized manners, thus reinforcing therapeutic alliances. Furthermore, adaptive AI algorithms continuously monitor therapeutic efficacy, modifying approaches responsively—mirroring adjustments typically made by experienced therapists—thereby moving towards a precision mental health approach tailored to each adolescent’s unique circumstances.*

*Early Intervention and Preventive Potential: AI systems demonstrate substantial promise in early identification and proactive management of mental health risks, leveraging predictive algorithms capable of discerning subtle behavioural or physiological changes often unnoticed by clinicians. Machine learning models employing data from social media, sensor inputs, and linguistic analysis have successfully anticipated depressive episodes, anxiety escalations, and suicidal ideation before clinical manifestation (Bzdok & Dunbar, 2020). Continuous monitoring through AI-driven platforms can trigger timely alerts, prompting adolescents towards immediate intervention or facilitating caregiver or clinical alerts for early-stage intervention. This aligns with preventative mental health strategies crucial in contexts like the UK, where resource constraints exacerbate delays in clinical intervention. Clinician respondents recognized AI’s potential role as a continuous monitoring tool between clinical visits, significantly enhancing early detection of deterioration risks. Nonetheless, managing false-positive and negative alerts responsibly remains an essential ethical consideration.*

*Scalability and Cost-Effectiveness: AI-driven digital therapies exhibit exceptional scalability, efficiently addressing widespread demand with comparatively minimal incremental costs. Unlike conventional therapeutic interventions reliant on extensive clinician hours, AI-based applications and chatbots can concurrently engage large numbers of adolescents, significantly amplifying service capacity without proportional resource expansion. This scalability aligns strategically with the NHS’s objectives under current resource limitations, effectively redistributing clinician workloads towards managing complex cases requiring specialized attention (Hollis et al., 2017). Economic analyses indicate potential cost savings and enhanced service efficiency, reducing downstream expenditures related to hospitalizations or extensive face-to-face interventions. Implementation examples from UK pilot studies highlighted considerable administrative efficiency gains, reduced clinical wait times, and improved clinician capacity for qualitative patient care. Consequently, despite initial developmental costs, long-term economic and clinical benefits strongly justify investment in scalable AI technologies as supportive adjuncts to existing mental health services.*

*Enhancing Engagement and Adherence: AI-enabled digital platforms demonstrate significant advantages in improving adolescents' adherence and engagement in mental health interventions. Traditional interventions frequently face challenges related to patient dropout and inconsistent adherence due to stigma, motivation loss, or forgetfulness. AI applications enhance user engagement through interactive, responsive interfaces such as conversational chatbots and gamification, maintaining adolescent interest and active participation. AI personalization and interactive conversational features substantially increase sustained application usage compared to generic digital interventions (Khosravi & Azar, 2024). Adolescents interviewed emphasized the benefits of AI systems’ constant availability and non-judgmental interactions, reducing barriers linked to stigma and fear of judgment. AI-driven adherence strategies such as personalized reminders, check-ins, and tailored motivational feedback further sustain adolescent engagement, leading to more consistent therapeutic participation and improved clinical outcomes, exemplified by significant anxiety reduction in users regularly engaging with AI-driven cognitive behavioural therapy chatbots.*

*In summary, the opportunities presented by AI-driven digital therapy in adolescent mental health are multifaceted. They offer potential solutions to some of the chronic issues in the UK’s mental health services: improving reach and access for all youths, delivering individualized support at scale, catching problems early, and keeping young people engaged in care. These advantages are not merely theoretical; early implementations and trials have demonstrated concrete benefits, and there is optimism among stakeholders about further gains as the technology matures. However, realizing these benefits on a large scale requires overcoming several barriers, which are described in the next section.*

*Barriers to Implementation*

*Despite substantial potential, the widespread integration and effectiveness of AI-driven digital mental health interventions for adolescents face considerable challenges. These obstacles include technological limitations, infrastructural inadequacies, concerns regarding data security and privacy, insufficient evidence on long-term efficacy, exacerbation of socio-economic inequalities, and resistance from healthcare professionals. Recognizing and addressing these barriers is essential to facilitate the safe and effective implementation of AI-based mental health technologies within UK healthcare contexts.*

*Technical and Infrastructure Limitations: A critical challenge in implementing AI-based interventions is their dependence on technological infrastructure, which varies significantly across different UK regions. Robust digital mental health solutions necessitate reliable internet connectivity and device accessibility, yet disparities remain, particularly affecting rural communities and socioeconomically disadvantaged urban areas (Gajeelee & Shroff, 2022). Even in regions with established connectivity, technical issues such as software malfunctions, inadequate interoperability with existing healthcare platforms, and bandwidth limitations for data-intensive interventions impede consistent utilization. Several studies reported technical difficulties as primary reasons for user disengagement or dropout. Additionally, integration challenges persist within NHS systems, where current infrastructure often lacks the capacity to seamlessly incorporate data from digital interventions into clinical records, thus perpetuating fragmented information management. Interviewees highlighted instances where institutional IT policies, such as restrictive firewalls within schools, inhibited access to digital therapeutic platforms. Addressing these technical and infrastructural barriers necessitates strategic investments to improve digital connectivity, subsidize device distribution, and upgrade healthcare IT systems. Without substantial enhancements in these foundational areas, the potential benefits of AI interventions remain restricted to well-equipped users, thereby limiting equitable access.*

*Data Security and Privacy Concerns: The reliance of AI-driven mental health platforms on extensive data collection and analysis introduces substantial privacy and security concerns, presenting significant barriers to user acceptance and ethical application. The sensitive nature of mental health data—including mood logs, conversational transcripts, and sensor-derived insights—heightens vulnerability to misuse or unauthorized access. Public awareness and sensitivity to data breaches and privacy violations have intensified caution among adolescents and caregivers, impacting their willingness to fully engage with digital therapeutic tools (Alhuwaydi, 2023). Many adolescents express reluctance to disclose personal information fully due to fears surrounding confidentiality and data misuse, resulting in superficial interaction or withholding critical information. Although regulatory frameworks such as the UK General Data Protection Regulation (UK-GDPR) enforce strict data protection standards, compliance and transparency vary, particularly among smaller technology developers. Interviewed NHS representatives emphasized the critical importance of establishing trust through clear, robust data handling practices. Addressing these concerns requires stringent implementation of encryption, secure data storage, explicit privacy policies, and effective communication of data management practices to users. Moreover, complexities surrounding consent, particularly with minors, further complicate privacy considerations, necessitating specialized frameworks to ensure ethical data use and confidentiality.*

*Limited Evidence of Long-term Effectiveness: Although short-term benefits of AI-driven mental health interventions are well-documented, the absence of substantial long-term clinical outcome data remains a considerable barrier to broader adoption. Most reviewed studies provided evidence of immediate symptom alleviation or enhanced user satisfaction but lacked extended follow-up periods essential to confirm sustained efficacy and safety. The rapid evolution of AI technologies frequently outpaces rigorous clinical evaluations, generating concern among clinicians and decision-makers about prematurely deploying interventions without sufficient empirical validation. Current evidence deficits include limited rigorous randomized controlled trials (RCTs), insufficient longitudinal analyses, and sparse comparative studies against conventional therapeutic methods (Guo et al., 2024). Clinicians interviewed voiced reservations about adopting AI technologies without clear validation through established guidelines or robust clinical trials. Regulatory bodies, including the UK's Medicines and Healthcare products Regulatory Agency (MHRA), also necessitate stringent evidence standards for digital health approvals. Overcoming these evidence gaps requires increased investment in longitudinal RCTs, comprehensive comparative analyses, and cohort studies to establish credible evidence supporting sustained clinical benefits and safety profiles of AI interventions.*

*Digital Divide and Equity Issues: A critical barrier to equitable implementation is the risk of exacerbating existing socio-economic and health inequalities through AI-driven mental health interventions. The digital divide, characterized by uneven access to digital technologies, disproportionately affects lower-income families, minority communities, and rural populations who may most need enhanced mental health support yet lack consistent access to necessary technological resources (Greer et al., 2019). Adolescents in socioeconomically disadvantaged contexts often experience limited access to smartphones, reliable internet, or digital literacy skills required to effectively utilize digital interventions. This disparity risks creating an intervention bias favouring digitally privileged adolescents, thereby potentially widening existing health inequalities and contravening the NHS’s commitment to equity. Interviewed stakeholders strongly emphasized that digital mental health solutions must cater inclusively to all adolescents, regardless of socioeconomic status. Strategies to bridge this gap include adopting inclusive design principles, offering offline functionalities, simplifying interfaces for users with low digital literacy, subsidizing access to digital technologies, and conducting targeted educational outreach. Comprehensive measures addressing access disparities and enhancing digital literacy are essential to prevent exacerbating existing inequities and ensuring equitable mental health service provision through AI technologies.*

*Clinician and Service Provider Resistance: Healthcare professional scepticism and institutional resistance represent significant barriers to the adoption of AI-based mental health interventions. Clinicians often express concerns regarding AI's capability to emulate the nuanced therapeutic interactions and empathetic engagement central to mental health practice. Doubts persist about algorithmic reliability and the appropriateness of automated advice in handling complex individual mental health issues, alongside fears of inaccurate assessments or superficial therapeutic guidance that could inadvertently harm patients. Additionally, concerns about professional identity and potential displacement by automated systems further exacerbate clinician resistance. Interviews with healthcare providers revealed varied acceptance levels, with many professionals expressing reservations about integrating unproven algorithms into sensitive therapeutic contexts. Practical barriers, such as required workflow adaptations, additional training, and burdensome integration procedures, also inhibit clinician acceptance, particularly in resource-constrained environments. Effective strategies for overcoming clinician resistance involve comprehensive education and training programs highlighting AI's augmentative potential, clearly demonstrating clinical efficacy through evidence-based validations, and incorporating clinician input through participatory design approaches. Reframing AI technologies as complementary tools supporting clinical practice, rather than replacements, and clearly illustrating potential benefits in reducing workload and enhancing patient outcomes can gradually build clinician trust and acceptance.*

*In summary, substantial barriers to integrating AI-driven digital mental health interventions in adolescent care include infrastructural limitations, privacy concerns, insufficient empirical evidence, the risk of exacerbating existing inequalities, and clinician resistance. Addressing these multifaceted challenges requires comprehensive infrastructure investment, rigorous ethical governance frameworks, inclusive access strategies, robust clinical evidence, and strategic efforts to foster cultural shifts within healthcare. Successfully navigating these barriers will be crucial to realize AI’s potential fully in augmenting adolescent mental health services in the UK.*

*Ethical Considerations*

*The integration of AI-based digital therapies in adolescent mental health care introduces complex ethical challenges that require proactive and thoughtful management. Addressing these ethical dimensions is critical not only to fulfil moral and legal obligations but also to foster the necessary trust among users and practitioners. Our review identified five principal ethical domains: (1) Autonomy and Informed Consent, (2) Transparency and Explainability, (3) Bias and Fairness, (4) Accountability and Responsibility, and (5) Over-reliance and Deskilling of Clinicians.*

*Autonomy and Informed Consent: Ensuring patient autonomy through informed consent is foundational in healthcare ethics, yet this becomes notably complex with adolescents utilizing AI-driven mental health tools. The inherent opacity of AI algorithms complicates clear disclosure about data usage and decision-making processes, creating challenges in achieving genuinely informed consent (Danks & London, 2017). Furthermore, legal and ethical requirements necessitate parental involvement in consent processes for minors, which can conflict with adolescents’ needs for confidentiality and privacy, particularly in sensitive mental health contexts. While UK principles such as Gillick competence and Fraser guidelines facilitate adolescent self-consent under specific conditions, their application to digital health contexts remains inadequately defined. Ethical best practices suggest consent should be dynamic, understandable, and periodically revisited, ensuring adolescents retain control and can freely disengage or challenge AI recommendations.*

*Transparency and Explainability: The lack of transparency inherent in complex AI algorithms raises significant ethical concerns regarding explainability and trust. Adolescents, their families, and clinicians require understandable rationales for AI-generated recommendations or risk assessments. The absence of comprehensible explanations undermines trust and acceptance among users and clinicians, who may be hesitant to act on opaque recommendations. To ethically manage these concerns, developers should incorporate explainable AI (XAI) methodologies, ensuring decisions are traceable and interpretable. Transparent interfaces clearly communicating how user data informs therapeutic decisions can enhance user understanding and empowerment, consistent with patient-centred care and ethical healthcare delivery (Danks & London, 2017).*

*Bias and Fairness: AI systems, reliant on data-driven learning, risk perpetuating or amplifying existing societal biases, leading to unfair treatment or disparities in mental health outcomes. AI applications developed from unrepresentative datasets may underperform for specific subgroups—particularly ethnic minorities, rural adolescents, or socioeconomically disadvantaged populations—potentially exacerbating health inequalities (Obermeyer et al., 2019). Ethical considerations necessitate continuous assessment and mitigation of biases, prioritizing data diversity and representativeness. Regular audits and updates to algorithms are essential to address and rectify identified biases. Fairness also extends to ensuring accessibility for adolescents with disabilities, aligning with UK equality laws and NHS commitments to reducing healthcare disparities. Developers and practitioners must remain vigilant, critically assessing AI tools for equity impacts before widespread deployment.*

*Accountability and Responsibility: AI integration into clinical decision-making introduces ambiguity regarding accountability, complicating ethical and legal clarity in cases of adverse outcomes. Traditional healthcare accountability, typically clear-cut, becomes complex when AI tools inform clinical decisions. Questions arise regarding liability distribution among clinicians, developers, and healthcare organizations if AI-driven interventions result in negative consequences. The ethical principle demands clearly articulated accountability frameworks, delineating roles and responsibilities within AI-supported mental healthcare (Mittelstadt et al., 2016). In NHS contexts, accountability typically resides at the organizational level, necessitating clear guidelines for clinicians regarding AI integration into clinical judgment. Transparent communication to users about AI limitations and explicitly maintaining human oversight in critical clinical decisions are ethically imperative. Developers must also engage proactively with regulators to align AI tools with established safety standards and liability frameworks.*

*Over-reliance and Clinician Deskilling: The extensive integration of AI tools risks over-reliance and subsequent deskilling of mental health professionals, posing nuanced ethical implications for clinical practice quality. Over-dependence on automated assessments or interventions could diminish clinician proficiency in critical diagnostic and therapeutic skills, potentially undermining the empathetic and intuitive aspects essential to mental health care (Choudhury & Chaudhry, 2024). Ethically, it is crucial to maintain AI as supportive rather than substitutive of core clinical competencies, ensuring clinicians retain essential diagnostic and therapeutic skills. Strategic mitigation involves using AI as complementary educational and supportive tools, setting clear boundaries around their use, and continuously monitoring clinical outcomes for indications of deskilling or over-dependence. A balanced partnership, leveraging AI's strengths while preserving clinicians' human judgment and empathy, remains ethically essential for sustaining high-quality care.*

*In summary, ethical considerations are deeply interwoven with the technical and practical facets of implementing AI in adolescent mental health care. Addressing ethical considerations in autonomy, transparency, fairness, accountability, and clinician skill preservation is fundamental to responsible implementation of AI in adolescent mental health. A proactive ethical approach, embedded within AI design and deployment, will be vital for achieving broad acceptance and ensuring the ethical integrity of digital mental health innovations in the UK healthcare context.*

*Discussion*

*Principal Findings and Interpretation*

*This mixed-methods review provides a comprehensive overview of the current landscape of AI-driven digital therapy for adolescent mental health in the UK, and it offers important insights into how these emerging tools can be both beneficial and challenging. In summary of key findings, our review confirms that AI-powered interventions hold considerable promise in addressing gaps in youth mental health services. They can enhance accessibility, delivering support to adolescents who might otherwise receive none, and they allow for a degree of personalization and real-time responsiveness that is difficult to achieve in standard care models. These advantages align with global observations in digital psychiatry that technology can extend the reach of care and engage users in novel ways (Torous et al., 2021). Additionally, the potential for early detection of mental health issues through AI-based monitoring represents a shift towards preventive care, which is particularly valuable given that early adolescence is when timely intervention can alter developmental trajectories in a positive manner.*

*Our findings also underscore that the UK context has unique considerations shaping the implementation of these technologies. The qualitative interviews with UK stakeholders highlighted a general optimism about AI’s role as a complement to traditional services rather than a replacement. Mental health professionals and policymakers recognized that AI tools could relieve pressure on overburdened services—such as by handling initial triage or providing interim support while youths are on waiting lists—thus integrating into the stepped care model of NHS mental health provision. For example, participants noted real-world instances where AI chatbots have been piloted in NHS-funded youth services to manage referrals or deliver guided self-help, reflecting a trend of cautious experimentation within the system. These accounts illustrate that when properly implemented, AI interventions can slot into existing care pathways, augmenting the capacity of those pathways. An adolescent might engage with an AI-based app while also seeing a counsellor periodically, creating a blended care experience. This hybrid model, where digital tools and human support are used in tandem, emerged as a favoured approach among experts, as it leverages the strengths of both modalities.*

*However, the discussion also brought to light the formidable challenges and caveats that accompany these opportunities. The barriers identified in results—technical shortcomings, privacy issues, evidence gaps, inequality risks, and clinician resistance—are not isolated problems but interrelated factors that must be holistically addressed. For instance, our interpretation of findings suggests that building trust is a linchpin issue: trust from young users (which hinges on data privacy and the perceived helpfulness of the AI), trust from clinicians (hinging on evidence quality and clarity of role), and trust from the public (hinging on ethical governance and success stories). Without trust, adoption will stall regardless of how advanced the technology is. Conversely, evidence of positive outcomes and strong ethical oversight can foster trust and lead to broader acceptance. One recurring theme from stakeholders was the need for interdisciplinary collaboration moving forward. The complexities of AI in mental health mean that no single group has all the expertise to tackle them—clinicians, AI developers, ethicists, youth representatives, and policymakers must work together. This collaboration is essential to design interventions that are clinically sound, user-friendly, and ethically robust. For example, one clinician in our study highlighted how joint workshops between NHS teams and tech developers helped bridge gaps in understanding, leading to improvements in how an app delivered psychoeducation content to make it more developmentally appropriate for teens.*

*In interpreting the evidence, it’s also crucial to remain critical. The studies reviewed, while illuminating, often had limitations such as small sample sizes or lack of long-term follow-up. Therefore, any enthusiasm for AI tools must be tempered by a commitment to gather more rigorous evidence. In effect, the field is at a nascent stage akin to early pharmacotherapy trials—showing promise but needing further validation. Moreover, our analysis indicates that context matters greatly: an AI intervention that works well in a controlled pilot might face unanticipated obstacles in the real-world NHS setting, such as integration with medical record systems or user engagement outside a research context. Thus, pilot programs in actual service settings (with iterative feedback loops) will be a valuable step in translating research into practice.*

*One encouraging aspect noted in discussions was that many of the challenges identified have possible solutions or mitigation strategies already being considered. For example, concerns about the digital divide are prompting initiatives for improving digital literacy in communities and exploring partnerships to provide internet access in underserved areas (e.g., school or library-based programs). Similarly, to address clinician scepticism, some professional bodies in the UK (like the Royal College of Psychiatrists) have begun issuing position statements and training modules on digital mental health, which can help demystify AI for practitioners and clarify best practices for its use. This suggests that the environment is gradually becoming more receptive as knowledge and comfort with digital tools grow.*

*In summary, the principal findings of this review paint a picture of significant transformative potential for AI in adolescent mental health alongside significant challenges that cannot be overlooked. The opportunities and barriers are two sides of the same coin; they must be understood together. The UK, with its centralized healthcare system and growing policy emphasis on digital innovation, is well-positioned to be at the forefront of implementing these tools—provided that implementation is done thoughtfully, with continuous evaluation and user involvement. The interpretation is clear: AI-driven digital therapies for adolescents are not a magic bullet that will solve the youth mental health crisis on their own, but they can become a powerful component of a multifaceted solution if integrated carefully. The remainder of this discussion will delve deeper into specific implications for practice and policy, how these findings compare with existing literature, and directions for future research, as well as the strengths and limitations of our review itself.*

*Implications for the UK Healthcare System and Practice*

*The findings from this review offer several critical implications for the UK's National Health Service (NHS), clinical practices, and policy frameworks related to adolescent mental health. Incorporating AI-driven digital interventions presents substantial potential to address enduring issues of insufficient service capacity and disparities in access, though realizing this potential demands careful strategic planning, rigorous oversight, and stakeholder collaboration.*

*Augmenting Service Capacity and Reducing Wait Times: AI technologies offer promising solutions for expanding the capacity of Child and Adolescent Mental Health Services (CAMHS) by automating initial assessment and triage processes. Tools such as AI-powered chatbots have demonstrated efficacy in streamlining referrals, significantly reducing wait times and clinician workload (Hollis et al., 2017). Broad implementation of such technologies across youth mental health pathways, including the Improving Access to Psychological Therapies (IAPT) programs, could enhance service responsiveness and resource efficiency. Effective integration requires clearly defined protocols for transitioning from digital assessments to human-led care, alongside robust clinician training programs. Adopting a blended-care model—where adolescents awaiting appointments utilize evidence-based digital supports—may maintain patient engagement and partially address therapeutic needs, potentially leading to improved clinical outcomes. Policymakers must support these innovations through comprehensive evaluations to confirm efficiency gains and clinical effectiveness.*

*Policy Frameworks for Data Governance and Ethics: Addressing significant data privacy concerns necessitates evolving regulatory frameworks tailored specifically for AI mental health technologies targeting adolescents. The UK's Information Commissioner’s Office (ICO) guidelines, such as the Children's Code, should inform stringent privacy standards for AI applications aimed at minors. Establishing an accreditation system akin to the NHS Apps Library, which rigorously assesses tools for safety, efficacy, and compliance with privacy standards, would facilitate responsible adoption. Additionally, ethical guidelines specifically addressing AI use in adolescent mental health—developed by bodies such as NICE—should clarify consent processes, mandate transparency of AI algorithms, and outline clear accountability structures where ultimate clinical responsibility remains with healthcare professionals. An NHS-led task force dedicated to overseeing AI integration could ensure ongoing policy alignment, monitor ethical adherence, and facilitate continuous system improvements.*

*Addressing the Digital Divide in Practice: Mitigating the risk of exacerbating health inequalities through digital solutions is critical. The NHS and related public health entities should actively implement initiatives to close digital accessibility gaps. This includes providing resources such as free internet access points and devices in community or educational settings, ensuring equitable technology access. Integration of digital mental health solutions into schools could uniformly benefit adolescents regardless of their socioeconomic status. Policy efforts should explicitly embed digital inclusion initiatives within strategic frameworks, such as revisions to the NHS Long Term Plan. Clinicians must remain sensitive to digital disparities, proactively offering alternative or supplementary non-digital support modalities when digital engagement is limited by technological constraints or skills deficits.*

*Clinician Training and Acceptance: Successful AI implementation critically depends on clinician acceptance and readiness. Professional development and training initiatives must prioritize enhancing digital literacy among mental health practitioners, enabling informed integration of AI tools into clinical workflows. Educational curricula within medical and psychological training should incorporate modules on digital therapeutics, facilitating early familiarity and competency among emerging professionals. Demonstrating tangible benefits through research findings and real-world case studies can progressively shift clinician attitudes toward favourable acceptance. Leveraging influential clinician advocates to share positive AI experiences can further promote widespread adoption within professional communities.*

*Maintaining Human Oversight and Hybrid Care Models: Complete automation of mental health care is neither ethically acceptable nor practically feasible. Thus, maintaining robust human oversight in AI-supported interventions is essential. Multidisciplinary teams should routinely integrate AI-generated insights into clinical decision-making processes, ensuring human judgment remains central. Hybrid care models, combining routine AI-mediated interactions with direct clinician involvement for critical interventions, balance the strengths of AI with the nuanced capabilities of human therapists. Practically, ensuring adolescents have consistent access to human support such as clinical consultations or helplines, strengthens safety protocols and user acceptance.*

*In conclusion, while AI-driven digital therapies offer considerable opportunities for the NHS to enhance adolescent mental health services, their implementation requires carefully structured strategies emphasizing safety, ethical standards, and equitable access. Policy development and clinical practice adjustments must be evidence-informed, ethically sound, and responsive to stakeholder needs to fully realize the potential of these innovative solutions.*

*Comparison with Existing Literature*

*This review provides a nuanced contribution to the literature on artificial intelligence (AI) in adolescent mental health, aligning closely with previous research while also highlighting context-specific insights unique to the United Kingdom (UK) healthcare environment. Several key points of convergence and divergence with existing studies have emerged.*

*Firstly, consistent with broader international literature, this review reaffirms the substantial opportunities presented by AI, such as increased accessibility, personalization, and efficiency of mental health services. Torous et al. (2021) previously identified AI-driven chatbots and mobile apps as promising tools for improving mental healthcare access globally. Our findings confirm these benefits within the UK adolescent context, emphasizing specific advantages related to national healthcare structures and regulatory frameworks, such as integration into the NHS and compliance with UK-specific data protection laws (e.g., UK GDPR, NICE guidelines). The necessity of NHS endorsement and adherence to local policies highlights a critical, context-specific dimension often overlooked in more generalized discussions.*

*Secondly, by exclusively focusing on adolescents (ages 11–19), this review addresses a critical gap commonly overlooked by broader studies. Adolescents differ markedly from adults in cognitive maturity, digital habits, and attitudes towards seeking help, necessitating tailored digital interventions (Steinberg, 2014). Our qualitative findings emphasize adolescent preferences for immediate, anonymous interactions, which contrasts notably with adult-focused research that tends to underemphasize these developmental distinctions. Adolescents particularly value anonymity, immediacy, and engaging, interactive digital features—a point underscored through specific examples provided by stakeholder insights. These detailed insights extend beyond quantitative outcomes typically reported, enriching the literature with concrete examples of youth engagement strategies and preferences.*

*Thirdly, the comprehensive approach adopted in this review—simultaneously exploring opportunities, barriers, and ethical considerations—presents a holistic and realistic assessment compared to studies focusing exclusively on either technological efficacy or user experience alone. Previous meta-reviews, such as that by Hollis et al. (2017), highlighted the necessity for multidimensional evaluations, including context, engagement, and safety alongside efficacy. By addressing these aspects collectively, our review offers a balanced perspective, acknowledging that technological effectiveness alone does not guarantee successful adoption or ethical appropriateness. Integrating systemic issues such as clinician resistance and ethical complexities, including informed consent and data privacy, provides pragmatic insights relevant for policymaking and clinical practice.*

*Furthermore, relative to earlier UK-focused analyses, this review incorporates the latest advancements in AI, including large language models (LLMs) and sophisticated natural language processing (NLP) techniques (Choudhury & Chaudhry, 2024). The rapid evolution in AI technologies, from simple rule-based systems to advanced generative conversational agents, necessitates continuous updates to ethical frameworks and practice recommendations. Contemporary literature discussions regarding algorithmic biases and the implications of advanced AI on clinical skills ('deskilling') are included, reflecting current, pertinent debates that older reviews might lack.*

*Additionally, the emphasis on practical implementation within the NHS distinguishes this review from other academic discussions that often remain theoretical. Explicit attention to practical infrastructural integration, electronic health record interoperability, and alignment with NHS core principles—such as equity and universal access—provides actionable insights tailored specifically to UK healthcare stakeholders. The incorporation of qualitative stakeholder perspectives further enriches this practicality, grounding theoretical insights in lived experiences and direct stakeholder feedback.*

*Nonetheless, our findings align closely with existing literature in recognizing persistent barriers such as privacy concerns, the need for robust evidence, and digital literacy challenges (Hassan et al., 2024). Such consistent identification across multiple studies underscores the urgency and authenticity of these issues, reinforcing the necessity for strategic interventions to address them comprehensively within the adolescent context.*

*In conclusion, this review contributes uniquely to the literature by blending global insights on AI in mental health with a targeted exploration of the UK adolescent healthcare context. Its balanced approach avoids extremes of techno-optimism or undue pessimism, presenting a realistic, nuanced perspective. Emphasizing the necessity of human oversight and advocating for hybrid care models, this review aligns with Erik Topol’s (2019) vision of AI as a complementary tool enhancing, rather than replacing, human clinical capabilities. Ultimately, this comprehensive and contextualized analysis serves as a valuable resource for researchers, practitioners, and policymakers seeking informed, practical guidance on integrating AI responsibly into adolescent mental healthcare.*

*Study Strengths and Limitations*

*This systematic review has several notable strengths that underpin the robustness and reliability of its findings, alongside limitations that should be considered in interpreting the results.*

*Strengths: The rigorous adherence to PRISMA guidelines (Page et al., 2021) for systematic literature identification, selection, and synthesis significantly enhances transparency and reproducibility. Employing extensive database searches complemented by grey literature inclusion minimized publication bias by capturing diverse perspectives, including both positive and null outcomes. A major methodological strength lies in the mixed-methods approach, integrating quantitative outcomes from randomized controlled trials (RCTs) with qualitative insights regarding user experiences, thereby providing a comprehensive narrative beyond purely quantitative or qualitative studies.*

*The review’s targeted focus on UK adolescents (ages 11–19) delivers specificity and contextually nuanced insights directly relevant to healthcare providers and policymakers within the National Health Service (NHS). Additionally, the synthesis covers technological capabilities alongside human factors such as stakeholder perceptions and ethical dimensions, demonstrating interdisciplinary depth across clinical medicine, public health, computer science, and ethics. This broad perspective maximizes relevance to diverse stakeholders, enhancing its practical utility.*

*Furthermore, organizing findings thematically into opportunities, barriers, and ethical considerations enhances clarity and facilitates policy-oriented application. Being current up to 2024, the review incorporates emerging trends and recent discussions on advanced technologies such as large language models (LLMs), thus maintaining contemporary relevance. Reliability was strengthened by the independent reviewer process in study selection and quality appraisal, ensuring robust evidence synthesis and reducing potential bias.*

*Limitations: Nevertheless, several limitations merit attention. Significant heterogeneity among included studies regarding interventions, methodologies, and outcomes precluded quantitative meta-analysis, thus limiting quantitative precision. The rapidly evolving nature of AI technology may also limit the applicability of findings from earlier studies, despite efforts to emphasize overarching themes.*

*The review's specific geographical and cultural context—focusing on the UK—while providing detailed local insights, restricts broader generalizability to different healthcare systems or international contexts. Additionally, the review inherently incorporates biases present in the included studies, such as potential positivity bias from studies led by AI proponents. Although mitigated through qualitative and grey literature inclusion, complete elimination of these biases remains challenging.*

*Qualitative insights from stakeholder interviews, while valuable, are subject to limitations including potential selection bias and relatively small sample sizes. These insights provide illustrative rather than broadly representative perspectives. Another critical limitation is the paucity of longitudinal data, constraining the assessment of sustained effectiveness or unintended long-term consequences.*

*Finally, narrative synthesis inevitably introduces interpretative subjectivity. Efforts were made to ground conclusions firmly in evidence and consensus discussions, yet alternative interpretations remain plausible. Consequently, this review offers a comprehensive and relevant overview, but findings should be considered provisional, contingent upon future research developments.*

*Future Research Directions*

*Based on identified gaps, several future research directions are proposed to advance the evidence base on AI-driven adolescent mental health interventions:*

*• Rigorous Efficacy Trials: Conducting large-scale RCTs comparing specific AI interventions against standard care or alternative digital interventions, emphasizing symptom reduction and functional outcomes with extended follow-up periods.*

*• Longitudinal Real-World Studies: Evaluating AI tools within naturalistic settings over extended durations to assess sustained engagement, real-world effectiveness, and unintended effects.*

*• Economic Evaluations: Conducting robust cost-effectiveness analyses to understand financial implications of integrating AI within NHS frameworks, employing metrics like quality-adjusted life years (QALYs).*

*• Optimal Design and Human-AI Interaction: Exploring optimal intervention design features through human-computer interaction research, examining blended models of human-AI interaction, and incorporating inclusive design principles.*

*• Ethical, Legal, and Social Implications Research: Empirical studies addressing practical ethical issues around informed consent, clinician-patient relationships, and accountability frameworks, including public perception and societal acceptance.*

*• Equity and Bias Mitigation: Developing advanced methodologies to identify and reduce algorithmic biases, alongside participatory research to enhance inclusivity and digital literacy.*

*• Condition-Specific and Crisis Intervention Research: Investigating targeted AI applications for specific adolescent mental health conditions and critical crisis scenarios, emphasizing rapid detection and response mechanisms.*

*• Interdisciplinary and Cross-sectoral Research: Promoting research collaborations across healthcare, education, and technology sectors to implement AI-driven solutions effectively within adolescents' digital ecosystems.*

*• Adolescent Co-production: Ensuring adolescent involvement as active co-researchers to create interventions attuned closely to their unique preferences and needs.*

*Pursuing these directions will enrich the evidence base, promoting safe, effective, and equitable integration of AI into adolescent mental health care within the NHS.*

*Conclusion*

*This systematic review illuminates a landscape of AI-driven digital therapies for adolescent mental health in the UK that is marked by substantial promise yet notable complexity. Evidence from the past decade indicates that AI-enabled interventions—such as chatbots, personalized mobile apps, and predictive analytics—can augment mental health services by improving accessibility, individualizing care, and increasing the responsiveness of support. These tools have demonstrated an ability to bridge gaps in traditional service provision by delivering real-time psychological assistance and tailored interventions to youth who might otherwise face lengthy wait times or lack access altogether. They thus represent a potential paradigm shift: mental health support available ubiquitously via adolescents’ smartphones at any time, adaptively meeting evolving needs. Amid a healthcare system strained by rising demand (particularly within the NHS), such innovations could prove pivotal in reaching more youth with timely, effective support.*

*However, this promise is tempered by formidable barriers hindering widespread implementation. Technical limitations (e.g., inconsistent internet access and underdeveloped digital infrastructure) and the persistent digital divide mean that, without deliberate intervention, AI’s benefits may be unevenly distributed, potentially excluding the most vulnerable youth. Data privacy and security concerns remain paramount; adolescents and their families must be assured that sensitive mental health information is safeguarded on digital platforms. Algorithmic bias is another major concern, highlighting the need for inclusive, representative AI development to ensure these tools serve all youth equitably. Furthermore, evidence of long-term clinical effectiveness remains nascent: many AI interventions show initial positive outcomes, but rigorous longitudinal data on sustained improvements are scarce. This evidentiary gap breeds caution and scepticism among clinicians and service providers, many of whom hesitate to integrate unproven technologies into routine care. Ultimately, clinician acceptance is pivotal: successful adoption depends on positioning AI as a supportive adjunct that complements—rather than replaces—therapeutic relationships and professional expertise.*

*The ethical and governance dimensions surrounding AI-driven mental health innovations require vigilant consideration. Pressing concerns include ensuring informed consent (especially for minors), algorithmic transparency, clear accountability for AI-driven decisions, and managing the impact on the clinical workforce’s skills and roles. Addressing these challenges demands clear strategies and guidelines to balance innovation with ethical responsibility and inclusivity. In the UK, achieving this balance will require collaboration among policymakers, ethicists, technologists, clinicians, and young people to establish regulations and best practices that safeguard patient rights, data privacy, and welfare. Notably, some steps have already been taken (e.g., national data governance frameworks and preliminary standards for digital mental health tools), but these efforts must continually evolve alongside technological advances.*

*To realize AI’s potential in adolescent mental health care, a proactive, interdisciplinary implementation strategy is needed. Key steps include investing in robust digital infrastructure and digital inclusion initiatives to ensure equitable access; establishing strong data protection measures and algorithmic transparency to build public trust; conducting rigorous evaluations (e.g., randomized trials and long-term implementation studies) to strengthen the evidence base for effectiveness and safety; and fostering cross-sector collaboration among healthcare providers, technologists, educators, and community organizations to train clinicians and improve digital literacy among users. Central to all these efforts is the involvement of young people at each stage—engaging adolescents as co-designers, testers, and informants to ensure the resulting tools truly resonate with their needs and preferences. By pursuing these recommendations, AI-driven interventions can be introduced in a manner that is safe, ethically sound, and attuned to youth culture as well as the practical realities of the NHS.*

*In summary, while AI offers transformative potential to augment adolescent mental health services in the UK, it is not a panacea. The path ahead will require navigating both technical and ethical challenges. Nonetheless, with sustained research, judicious policy support, and active engagement of stakeholders—including the young people at the centre of these efforts—AI can become a valuable component of a comprehensive, 21st-century approach to adolescent mental healthcare. Throughout this journey, we must remain guided by the goal of improving outcomes for all adolescents and ensure that any digital innovation complements—rather than undermines—the human-centred care on which the NHS is built. Ultimately, if digital therapy and AI are embraced in an inclusive, evidence-driven, and youth-centric manner, these tools can be integrated into routine practice without compromising ethics or equity, advancing the best interests of the youth we serve.*

*References*

*1. Alhuwaydi AM. Ethical considerations in the use of artificial intelligence in mental health care. The Egyptian Journal of Neurology, Psychiatry and Neurosurgery. 2023;59: 1–7.*

*2. Beauchamp TL, Childress JF. Principles of Biomedical Ethics (8th ed.). Oxford University Press. 2019.*

*3. Bzdok D, Dunbar RIM. The neurobiology of social distance. Trends in Cognitive Sciences. 2020;24:717–733.*

*4. Choudhury A, Chaudhry Z. Large language models and user trust: Consequence of self-referential learning loop and the deskilling of healthcare professionals. Journal of Medical Internet Research. 2024;26:56764.*

*5. Danks D, London AJ. Algorithmic bias in autonomous systems. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17). 2017;4691–4697.*

*6. Gajeelee A, Shroff M. Applicability of artificial intelligence in healthcare in resource-poor settings. Harvard Undergraduate Health Policy Review. 2022.*

*7. Gooding PA, Kallestrup J, Fazel S. Personalisation and recommendation for mental health apps: A mixed-methods study. Behaviour & Information Technology. 2021;43:1–17.*

*8. Greer B, Robotham D, Simblett S, Curtis H, Griffiths H, Wykes T. Digital exclusion among mental health service users: Qualitative investigation.1 Journal of Medical Internet Research. 2019;21:2 11696.*

*9. Guo Z, Lai A, Thygesen JH, Farrington J, Keen T, Li K. Large language model for mental health: A systematic review. arXiv preprint arXiv:2403.15401. 2024*

*10. Hassan M, Kushniruk A, Borycki E. Barriers to and facilitators of artificial intelligence adoption in health care: Scoping review. JMIR Human Factors. 2024;11:48633.*

*11. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, Davies EB. Annual Research Review: Digital health interventions for children and young people with mental health problems – a systematic and meta-review. Journal of Child Psychology and Psychiatry. 2017;58:474–503.*

*12. Jelassi M, Matteli K, Ben Khalfallah H, Demongeot J. Development and adaptation of ASR and NLP systems for a French-language online mental health platform. MDPI Informatics. 2024;11:23-45.*

*13. Khosravi M, Azar G. Factors influencing patient engagement in mental health chatbots: A thematic analysis of findings from a systematic review of reviews. Digital Health. 2024;10:205520762412479.*

*14. Lenartowicz M. A review of teletherapy, mobile health apps, virtual reality, and artificial intelligence for addressing mental health. Intersect: The Stanford Journal of Science, Technology, and Society. 2023;16:1–10.*

*15. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: Mapping the debate. Big Data & Society. 2016;3:1–21.*

*16. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. BMJ, 339, b2535. 2009.*

*17. Nwankwo EI, Emeihe EV, Ajegbile MD, Olaboye JA, Maha CC. Integrating telemedicine and AI to improve healthcare access in rural settings. International Journal of Life Science Research Archive, 2024;7:59–77.*

*18. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366:447–453.*

*19. Office for Health Improvement and Disparities. Mental health of children and young people in England, 2022 - wave 3 follow up to the 2017 survey. GOV.UK. 2022.*

*20. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. Systematic Reviews, 10, Article 89. 2021.*

*21. Robotham D, Satkunanathan S, Doughty L, Wykes T. Do we still have a digital divide in mental health? A five-year survey follow-up. Journal of Medical Internet Research. 2019;21:12502.*

*22. Royal Society for Public Health. NewFilters: to manage the impact of social media on young people’s mental health and wellbeing. 2018.*

*23. Steinberg L. Age of opportunity: Lessons from the new science of adolescence. Houghton Mifflin Harcourt. 2014.*

*24. Topol E. Deep medicine: How artificial intelligence can make healthcare human again. Basic Books. 2019.*

*25. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry. 2021;20:318-335.*

*26. transform.england.nhs.uk*

*27. Yang TT, Yang TT, An N, Kong A, Liu S, Liu SX. AI Clinics on Mobile (AICOM): Universal AI doctors for the underserved and hard-to-reach. arXiv preprint arXiv:2306.10324. 2023.*

**Table 1: Selected Databases and Their Rationale for Inclusion**

|  |  |
| --- | --- |
| **Database** | **Rationale** |
| **MEDLINE (PubMed)** | Comprehensive coverage of biomedical literature, including mental health research. |
| **PsycINFO** | Focus on psychology and behavioural sciences, relevant to mental health. |
| **Web of Science** | Multidisciplinary database indexing high-impact journals. |
| **Scopus** | Large, multidisciplinary database with strong coverage of scientific literature. |
| **EMBASE** | Biomedical and pharmacological database, useful for clinical research. |
| **CINAHL** | Cumulative Index to Nursing and Allied Health Literature, relevant for healthcare interventions. |
| **ACM Digital Library** | Focus on computer science and technology, relevant to AI and digital health. |
| **IEEE Xplore** | Focus on electrical engineering and technology, relevant to AI and digital health. |

**Appendix A: Search Strategy across Databases**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Database** | **Search String Used** | **Filters Applied** | **Date of Search** | **Results Retrieved** |
| PubMed | ("AI-driven therapy" OR "digital mental health") AND ("effectiveness" OR "barriers") | Peer-reviewed, last 10 years, English | Jan 10, 2024 | 125 |
| Scopus | ("artificial intelligence" AND "mental health") AND ("accessibility" OR "privacy") | English, last 5 years, journal articles | Jan 12, 2024 | 98 |
| IEEE Xplore | ("machine learning" OR "AI models") AND ("digital therapy" OR "mental health applications") | Conference & journal papers, last 7 years | Jan 14, 2024 | 75 |
| Web of Science | ("AI-assisted therapy" OR "digital interventions") AND ("clinical outcomes" OR "scalability") | Last 10 years, English only | Jan 16, 2024 | 110 |