Review

Ethical Considerations in Artificial Intelligence Interventions for Mental Health and Well-Being: Ensuring Responsible Implementation and Impact

Hamid Reza Saeidnia [1], Seyed Ghasem Hashemi Fotami [2], Brady Lund [3] and Nasrin Ghiasi [4,*]

[1] Department of Knowledge and Information Science, Tarbiat Modares University, Tehran 14115-111, Iran; hamidrezasaednia@modares.ac.ir
[2] Department of Computer Science, Tarbiat Modares University, Tehran 14115-111, Iran; seyedghasemhashemifotami@modares.ac.ir
[3] Department of Information Science, University of North Texas, Denton, TX 76203, USA; brady.lund@unt.edu
[4] Department of Public Health, School of Health, Ilam University of Medical Sciences, Ilam 69391-77143, Iran
[*] Correspondence: ghiasi-n@medilam.ac.ir

Abstract: AI has the potential to revolutionize mental health services by providing personalized support and improving accessibility. However, it is crucial to address ethical concerns to ensure responsible and beneficial outcomes for individuals. This systematic review examines the ethical considerations surrounding the implementation and impact of artificial intelligence (AI) interventions in the field of mental health and well-being. To ensure a comprehensive analysis, we employed a structured search strategy across top academic databases, including PubMed, PsycINFO, Web of Science, and Scopus. The search scope encompassed articles published from 2014 to 2024, resulting in a review of 51 relevant articles. The review identifies 18 key ethical considerations, including 6 ethical considerations associated with using AI interventions in mental health and wellbeing (privacy and confidentiality, informed consent, bias and fairness, transparency and accountability, autonomy and human agency, and safety and efficacy); 5 ethical principles associated with the development and implementation of AI technologies in mental health settings to ensure responsible practice and positive outcomes (ethical framework, stakeholder engagement, ethical review, bias mitigation, and continuous evaluation and improvement); and 7 practices, guidelines, and recommendations for promoting the ethical use of AI in mental health interventions (adhere to ethical guidelines, ensure transparency, prioritize data privacy and security, mitigate bias and ensure fairness, involve stakeholders, conduct regular ethical reviews, and monitor and evaluate outcomes). This systematic review highlights the importance of ethical considerations in the responsible implementation and impact of AI interventions for mental health and well-being. By addressing privacy, bias, consent, transparency, human oversight, and continuous evaluation, we can ensure that AI interventions like chatbots and AI-enabled medical devices are developed and deployed in an ethically sound manner, respecting individual rights, promoting fairness, and maximizing benefits while minimizing potential harm.

Keywords: artificial intelligence; mental health; ethics; well-being; interventions

1. Introduction

Artificial intelligence (AI) is a rapidly advancing technology that involves the development of systems capable of performing tasks that typically require human intelligence, such as learning, problem solving, and decision making (Chalyi 2024; Saeidnia 2023). In the field of health, AI has emerged as a powerful tool with the potential to transform various aspects of healthcare delivery, diagnosis, treatment, and patient care (Reddy et al. 2019). By leveraging data analytics, machine learning algorithms, and predictive modeling, AI has

the capacity to revolutionize the way healthcare services are delivered and improve patient outcomes (Alowais et al. 2023; Yelne et al. 2023).

In recent years, AI has also made significant inroads into the field of mental health and well-being (Yelne et al. 2023). Mental health disorders such as depression, anxiety, and PTSD represent a growing global health burden, with millions of individuals in need of support and treatment (Wainberg et al. 2017; Charlson et al. 2019). AI technologies have been utilized to develop innovative interventions aimed at addressing these challenges by improving access to care, enhancing treatment outcomes, and providing personalized support to individuals in need (Mennella et al. 2024). From chatbots and virtual therapists (referring to digital, remote mental health support and treatment, whether delivered by AI systems, human therapists, or a combination of both) to predictive analytics for early intervention, AI interventions in mental health hold great promise for improving the quality and effectiveness of mental healthcare services (Balcombe 2023).

The potential benefits of artificial intelligence in mental health are multifaceted (Baskin et al. 2021; Carr 2020). AI-based interventions have the capacity to provide timely and personalized support to individuals experiencing mental health challenges, thereby improving their overall well-being (S. Graham et al. 2019; Shah 2022). By analyzing vast amounts of data, AI systems can identify patterns and trends that may not be apparent to human clinicians, leading to more accurate diagnoses and treatment recommendations (Alowais et al. 2023; Faezi and Alinezhad 2024). AI tools can also help bridge the gap in mental health services by reaching underserved populations, reducing barriers to access, and increasing the efficiency of healthcare delivery (V. Singh et al. 2024).

Alongside the potential benefits of AI in mental health, however, there are also significant ethical consequences that must be carefully considered (Jeyaraman et al. 2023). The use of AI in health care, particularly in sensitive areas such as mental health, raises complex ethical dilemmas related to privacy, consent, transparency, accountability, bias, and the potential for unintended harm (Farhud and Zokaei 2021; Thakkar et al. 2024). Issues such as data security, algorithmic bias, and the impact of automation on the patient–provider relationship are critical considerations that must be addressed to ensure the responsible and ethical implementation of AI interventions in mental health settings (Alowais et al. 2023; Bélisle-Pipon et al. 2022; Davahli et al. 2021; Gaonkar et al. 2023; Sarah Graham et al. 2019; Jeyaraman et al. 2023; Khanna and Srivastava 2020).

Against this backdrop, this systematic review study aims to critically examine the ethical considerations surrounding the use of artificial intelligence in mental health interventions. By synthesizing existing literature and research findings, our goal is to shed light on the key ethical challenges and opportunities associated with the integration of AI technologies in mental health care. We seek to identify best practices, guidelines, and recommendations for promoting responsible implementation and ensuring the positive impact of AI interventions on individuals' mental health and overall well-being. Through this review, we aim to contribute to a better understanding of how ethical principles can be upheld in the development and deployment of AI solutions in mental health, ultimately enhancing the quality and accessibility of mental healthcare services while safeguarding the rights and dignity of individuals receiving care.

2. Methods and Materials

In this study, we critically analyze the ethical considerations related to the utilization of artificial intelligence in mental health interventions. Throughout the process of manuscript preparation, we followed the guidelines outlined by (Smith et al. 2011), with a specific focus on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al. 2021).

2.1. Research Questions

1. What are the key ethical considerations associated with the use of artificial intelligence interventions in mental health and well-being?

2. How can ethical principles be integrated into the development and implementation of AI technologies in mental health settings to ensure responsible practice and positive outcomes?

3. What are the best practices, guidelines, and recommendations for promoting ethical use of AI in mental health interventions?

2.2. Inclusion and Exclusion Criteria

This systematic review applied the following inclusion and exclusion criteria.

2.2.1. Inclusion Criteria

The inclusion criteria for this review study encompass studies that focus on the ethical considerations surrounding the use of artificial intelligence interventions in mental health and well-being. Including studies that address ethical considerations ensures a comprehensive understanding of the potential implications of AI in mental health interventions. Additionally, studies investigating the impact of AI on mental health outcomes are crucial for evaluating the effectiveness and potential risks associated with these technologies. All types of articles, including reviews, original research, short communications, and letters to the editor, are considered for inclusion to provide a diverse range of perspectives and insights on the topic. This approach allows for a thorough examination of the current literature on AI in mental health, regardless of the format in which the information is presented. Finally, limiting the inclusions to publications in the English language ensures consistency in data interpretation and accessibility for the review process.

2.2.2. Exclusion Criteria

The exclusion criteria for this review study involve excluding studies that do not specifically address the ethical implications of AI interventions in mental health, as the primary focus of this review is on the ethical considerations associated with AI technologies in mental health and well-being. Additionally, studies that primarily emphasize the technical aspects of AI algorithms without integrating discussions on ethical considerations are excluded, as the ethical dimension is a key aspect of interest in this review. Publications in languages other than English are also excluded to maintain consistency in data interpretation and ensure accessibility for the research process. By applying these exclusion criteria, this review aims to focus on studies that provide comprehensive insights into the ethical implications of AI interventions in mental health.

2.3. Databases and Search Method

We conducted a comprehensive literature search across the following databases:

1. PubMed;
2. PsycINFO;
3. Scopus;
4. Web of Science;
5. Google Scholar.

Search terms included combinations of keywords related to artificial intelligence, mental health, ethics, well-being, and interventions. Boolean operators (AND, OR) were used to refine search queries and identify relevant studies. The search scope spanned a decade between 2014 and 2024. We conducted a manual search of Google Scholar to enhance the scope of our search and identify additional relevant articles. This method enabled us to expand our search beyond the initial database search and uncover a broader range of scholarly articles related to our research topic.

The search strategy was designed to capture a broad range of articles addressing the ethical implications of AI interventions in mental health. While there were slight variations in the specific search terms and strings used across databases, they maintained a consistent structure. Additional information can be found in Supplementary File S1.

2.4. Study Selection

Our study selection process involved a meticulous review of article titles and abstracts by each researcher to assess relevance to our inclusion criteria, focusing on the ethical considerations of artificial intelligence interventions in mental health. Conflicting articles were promptly excluded, and input from other scholars was sought when doubts arose, ensuring a consensus-based final selection by the research team. This rigorous approach aimed to maintain consistency and rigor in the selection process, address uncertainties collaboratively, and enhance the reliability of the studies included in our systematic review.

In cases where a researcher had a potential conflict of interest due to prior involvement with a study, such as being a co-author, that researcher recused themselves from evaluating that particular study to maintain objectivity. For example, if a researcher had collaborated on a study examining the ethical implications of a specific AI-powered mental health chatbot, they would not have been involved in assessing the inclusion of that study in our review to avoid any bias. By following this protocol, we ensured that the selection process remained impartial and free from conflicts of interest.

2.5. Quality Assessment

The quality of included studies was assessed using The Critical Appraisal Skills Programme (CASP) Systematic Review tool, which is a widely recognized tool for evaluating the methodological rigor and validity of research studies. The CASP tool provides a structured framework for critically appraising the key components of a study, including study design, methodology, data collection, analysis, and interpretation of findings (Supplementary File S2).

2.6. Data Extraction and Synthesis

Data extraction involved systematically collecting relevant information from each included study, such as author(s), publication year, study design, key findings related to ethical considerations in AI interventions for mental health, and recommendations for ethical practice. Data synthesis involved analyzing and summarizing the extracted information to identify common themes, trends, and gaps in the literature. Findings were synthesized to provide a comprehensive overview of the ethical challenges and opportunities associated with the use of AI in mental health interventions, as well as recommendations for promoting responsible and ethical practice in this evolving field (Supplementary File S3).

3. Results

3.1. Article Selection

Based on the database search strategy (PubMed, PsycINFO, Scopus, Web of Science, and Google Scholar), we identified 5974 articles, out of which 1412 articles were relevant to PubMed, 1951 articles were relevant to PsycINFO, 1351 articles were relevant to Scopus, and 1100 articles were relevant to Web of Science. Furthermore, our manual search of Google Scholar identified 160 articles. After removing 3236 duplicate articles, the remaining 2738 articles were screened. Of these, 2036 articles were excluded, as they were not relevant to the study objectives; 1969 were from other academic disciplines; and 340 were in languages other than English. This left 702 potentially eligible articles. Upon further review of the titles and abstracts, an additional 443 articles were excluded, as they did not meet the study design criteria (i.e., they focused on other content or subjects). The full texts of the remaining 259 articles were then assessed for inclusion. After this detailed evaluation, 216 articles were excluded, leaving a final set of 43 articles that were included in the systematic review. We found an additional 8 relevant articles through a citation-chaining search. Consequently, in the final summary, we obtained 51 articles (Figure 1).
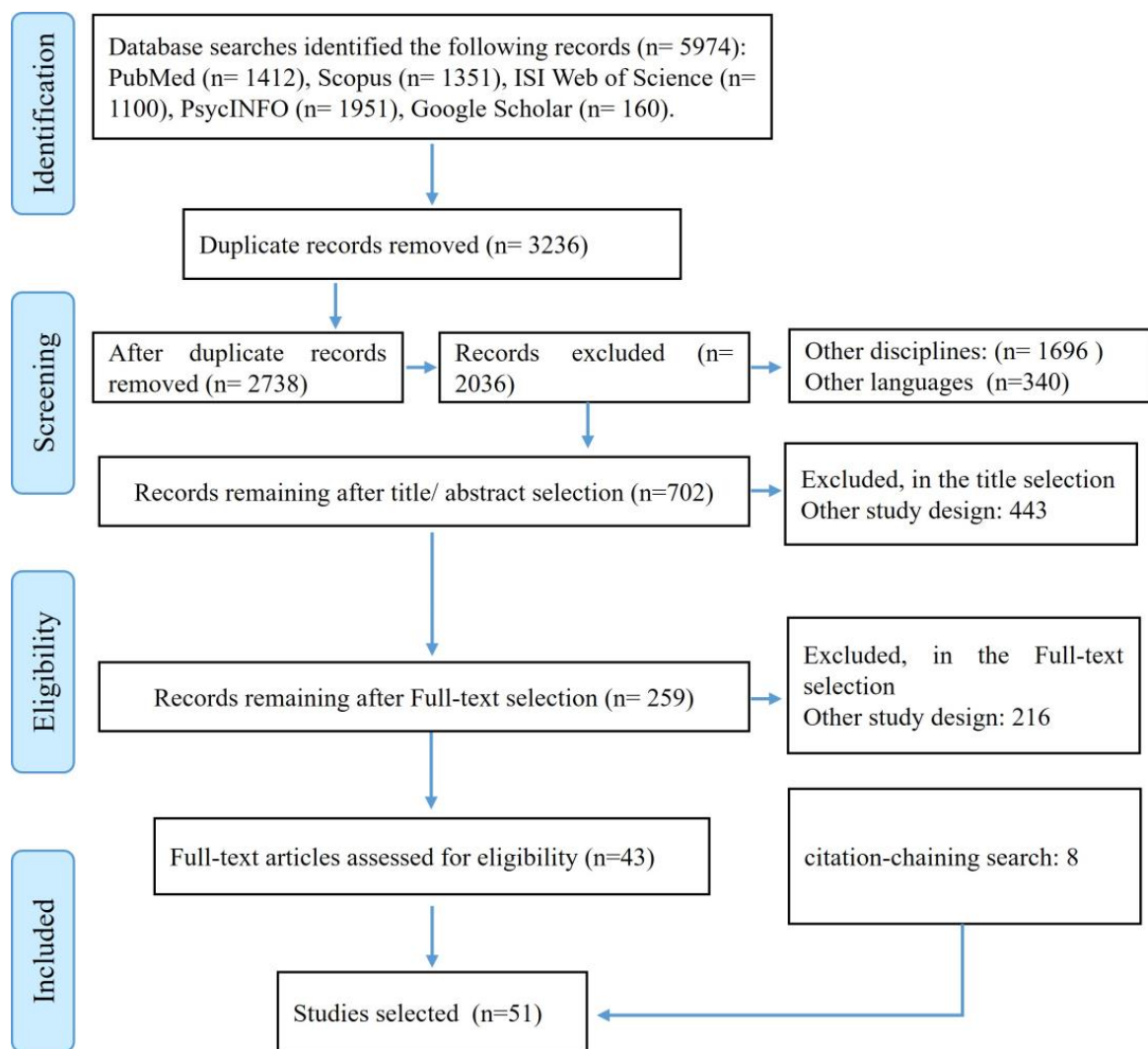
Figure 1. Flow diagram showing the study selection/screening process.

3.2. Quality Assessment Results

The assessment of study quality using the CASP tool yielded insightful results. None of the studies achieved a perfect score of 10. Notably, nine studies stood out with a commendable score of 9, reflecting the high quality of research in those cases, which accounted for 17.64% of the reviewed articles. Additionally, a significant portion of the studies, totaling twenty-three, achieved a score of 8, comprising 45.09% of the total. Twelve studies received a score of 7, demonstrating a satisfactory level of quality, representing 23.52% of the reviewed articles. Moreover, seven studies garnered a score of 6, indicating room for enhancement yet still contributing valuable insights, making up 13.75% of the total. Overall, these findings highlight the qualitative strengths prevalent in the majority of the reviewed studies. For a detailed examination of the study selection process, readers are encouraged to consult Supplementary File S4.

3.3. Ethical Considerations of Artificial Intelligence Interventions in Mental Health and Well-Being

According to the literature review, there are several key ethical considerations associated with the use of artificial intelligence interventions in mental health and well-being. Some of the main considerations include privacy and confidentiality, informed consent, bias and fairness, transparency, explainability, accountability, and autonomy and human agency (Table 1).

Table 1. Ethical considerations of artificial intelligence interventions in mental health and well-being.

Considerations	Interpretation	Reference
Privacy and confidentiality	AI systems often collect and analyze large amounts of sensitive personal data. It is crucial to ensure that these data are handled securely and that individuals' privacy rights are respected.	(Y. Chen and Esmaeilzadeh 2024; Chintala 2022; Murdoch 2021; Sivan and Zukarnain 2021)
Informed consent	Individuals should be fully informed about how their data will be used and the potential risks and benefits of using AI interventions in mental health. Informed consent should be obtained before implementing any AI-based intervention.	(Cohen 2019; Pickering 2021; Ursin et al. 2021)
Bias and fairness	AI systems can perpetuate and amplify biases present in the data used to train them. It is important to address issues of bias and ensure that AI interventions are fair and equitable for all individuals, regardless of their background or characteristics.	(Gaonkar et al. 2023; Kerasidou 2021; Martin et al. 2022; Aditya Singhal et al. 2024; Tatineni 2019)
Transparency, explainability, and accountability	The decision-making process of AI systems can often be opaque and difficult to interpret. It is important to ensure transparency in how AI interventions are developed, implemented, and evaluated and to establish mechanisms for accountability in case of errors or harm.	(Habli et al. 2020; Khanna and Srivastava 2020; Kiseleva et al. 2022; Aditya Singhal et al. 2024; Vollmer et al. 2020)
Autonomy and human agency	AI interventions should be designed to support and enhance human decision making and autonomy rather than replacing human judgment entirely.	(Fanni et al. 2023; Love 2023; Tiribelli 2023)
Safety and efficacy	AI interventions should be rigorously evaluated to ensure that they are safe and effective for use in mental health and well-being contexts. It is essential to prioritize the well-being and safety of individuals who may be using these interventions.	(Davahli et al. 2021; Ellahham et al. 2020; Habli et al. 2020; Morley et al. 2021; Tiwari and Dileep 2023)

3.4. Integrating Ethical Principles for Responsible Practice and Positive Outcomes in AI Technologies for Mental Health Settings

Through a comprehensive review of the collected articles, we identified several considerations for incorporating ethical principles into the AI design process, namely an ethical framework, stakeholder engagement, ethical review, bias mitigation, and continuous evaluation and improvement (Table 2).

Table 2. Considerations for integrating ethical principles for responsible practice and positive outcomes in ai technologies for mental health settings.

Considerations	Interpretation	Reference
Ethical framework	A clear ethical framework should be established that outlines the values and principles that will guide the development and implementation of AI technologies in mental health settings is essential. This framework should address key ethical considerations such as privacy, transparency, fairness, and accountability.	(Baskin et al. 2021; Leimanis and Palkova 2021; Nasir et al. 2024; Prathomwong and Singsuriya 2022; Tahri Sqalli et al. 2023)
Stakeholder engagement	The involvement of a diverse group of stakeholders, including mental health professionals, patients, ethicists, and community members, in the development process can help identify and address ethical concerns from various perspectives.	(Bélisle-Pipon et al. 2022; Couture et al. 2023; A. Singhal et al. 2024)

Table 2. *Cont.*

Considerations	Interpretation	Reference
Ethical review	Conducting regular ethical reviews of AI technologies in mental health settings can help identify and address any ethical issues that may arise during the development and implementation process.	(McKay et al. 2023; Olorunsogo et al. 2024; Shaw 2022)
Bias mitigation	Implementing strategies to mitigate bias in AI technologies, such as using diverse and representative datasets, regularly monitoring for bias, and incorporating fairness and accountability measures into the algorithms, can help ensure that the technology is used ethically and responsibly.	(F. Chen et al. 2024; Ferrara 2023; Mensah 2023; Mittermaier et al. 2023)
Continuous evaluation and improvement	Regularly evaluating the impact of AI technologies on mental health outcomes and ethical considerations is important. This includes monitoring for any unintended consequences, soliciting feedback from stakeholders, and making adjustments to the technology as needed to ensure positive outcomes and responsible practice.	(WHO Guidance 2021; Magrabi et al. 2019; McGreevey et al. 2020; Morley et al. 2020)

3.5. Practices for Ethical Use of AI in Mental Health Interventions

According to our comprehensive review of the articles, some key practices to promote the ethical use of AI in mental health interventions are adhering to ethical guidelines, ensuring transparency, prioritizing data privacy and security, mitigating bias and ensuring fairness, involving stakeholders, conducting regular ethical reviews, and monitoring and evaluating outcomes (Table 3).

Table 3. Best Practices for ethical use of ai in mental health interventions: guidelines and recommendations.

Considerations	Interpretation	Reference
Adhere to ethical guidelines	Established ethical guidelines and principles should be followed, such as those outlined by professional organizations like the American Psychological Association (APA) or the World Health Organization (WHO), to guide the development and implementation of AI technologies in mental health settings.	(Joerin et al. 2020; Luxton 2014; Skorburg et al. 2024)
Ensure transparency and explainability	Transparency about how AI technologies are developed, how they work, what underlying data were used to train them, and how they are used in mental health interventions should be prioritized. Providing clear information to users about the technology can help build trust and promote ethical use.	(Carr 2020; Kasula 2023; Aditya Singhal et al. 2024)
Prioritize data privacy and security	Robust data privacy and security measures should be implemented to protect the confidentiality and integrity of individuals' data. This includes securing data storage, ensuring data encryption, and obtaining informed consent from individuals before collecting and using their data.	(Gooding and Kariotis 2021; Mörch et al. 2020; Olawade et al. 2024; Rubeis 2022)
Mitigate bias and ensure fairness	Steps should be taken to identify and mitigate biases in AI algorithms used in mental health interventions. This includes using diverse and representative datasets, regularly monitoring for bias, and implementing fairness measures to ensure equitable outcomes for all individuals.	(F. Chen et al. 2024; Ferrara 2023; Mensah 2023; Mittermaier et al. 2023)

Table 3. *Cont.*

Considerations	Interpretation	Reference
Involve stakeholders	A diverse group of stakeholders should be engaged, including mental health professionals, patients, ethicists, and community members, in the development and implementation of AI technologies in mental health settings. Incorporating diverse perspectives can help identify and address ethical concerns and ensure that the technology meets the needs of its users.	(Carr 2020; Y. Chen and Esmaeilzadeh 2024; Chintala 2022; Kasula 2023; Murdoch 2021; Aditya Singhal et al. 2024; Sivan and Zukarnain 2021)
Conduct regular ethical reviews	The ethical implications of AI technologies in mental health interventions should be regularly reviewed to identify and address any ethical issues that may arise. This can involve evaluating the potential risks and benefits of the technology, ensuring compliance with ethical guidelines, and making adjustments as needed to promote responsible practice.	(McKay et al. 2023; Olorunsogo et al. 2024; Shaw 2022)
Monitor and evaluate outcomes	The impact of AI technologies on mental health outcomes and ethical considerations should be continuously monitored and evaluated. This includes assessing the effectiveness of the technology, soliciting feedback from stakeholders, and making improvements to enhance ethical use and positive outcomes.	(Carr 2020; Sarah Graham et al. 2019; Habli et al. 2020; Khanna and Srivastava 2020; Kiseleva et al. 2022; Aditya Singhal et al. 2024; Vollmer et al. 2020)

4. Discussion

Artificial Intelligence (AI) holds immense potential to revolutionize mental health services by providing personalized support and improving accessibility. However, the responsible implementation of AI interventions in mental health settings requires careful consideration of ethical concerns to ensure positive outcomes for individuals. This study contributes to a theoretical understanding of ethical considerations for AI in mental health interventions by identifying the themes of privacy, informed consent, bias and fairness, transparency, accountability, autonomy, and safety within the literature. In order to adequately integrate these principles, it is critical to engage stakeholders who will be impacted by the technology and continuously evaluate as the technology evolves. Ultimately, these technologies must support people, not overlook them in an effort to automate.

Several studies emphasize the importance of safeguarding patient privacy and ensuring confidentiality in AI-driven mental health interventions (Y. Chen and Esmaeilzadeh 2024; Chintala 2022; Murdoch 2021; Sivan and Zukarnain 2021). One of the foremost ethical considerations in AI-driven mental health interventions is privacy and confidentiality (Ghadiri 2022; Shimada 2023). Protecting patient data and ensuring confidentiality are paramount to building trust between users and AI systems. Researchers emphasize the importance of implementing robust data security measures and adhering to privacy regulations (Murdoch 2021; Sivan and Zukarnain 2021). By safeguarding patient privacy, AI interventions can uphold ethical principles while delivering personalized support to individuals seeking mental health assistance (Y. Chen and Esmaeilzadeh 2024; Murdoch 2021; Sivan and Zukarnain 2021).

Researchers stress the need for transparent communication, explainability of AI models, and the need to obtain informed consent from users before deploying AI interventions (Cohen 2019; Pickering 2021; Ursin et al. 2021). This includes providing clear information about the purpose, risks, and benefits of the technology (Cohen 2019; Pickering 2021; Ursin et al. 2021). Transparent communication fosters trust and empowers individuals to make informed decisions about their mental health care. By prioritizing informed consent, AI interventions can respect individuals' autonomy and promote collaborative decision-making between users and providers (Cohen 2019; Ursin et al. 2021).

Several researchers have identified the risk of bias in AI algorithms, particularly in mental health diagnostics and treatment recommendations (Gaonkar et al. 2023; Kerasidou 2021; Martin et al. 2022; Aditya Singhal et al. 2024; Tatineni 2019). Addressing bias requires diverse and representative datasets, as well as algorithmic fairness assessments to mitigate disparities (Gaonkar et al. 2023; Kerasidou 2021; Aditya Singhal et al. 2024; Tatineni 2019). Bias in AI algorithms poses significant ethical challenges in mental health diagnostics and treatment recommendations. Studies have highlighted the importance of addressing bias through the use of diverse and representative datasets, algorithmic fairness assessments, and bias mitigation strategies (Gaonkar et al. 2023; Kerasidou 2021; Martin et al. 2022; Aditya Singhal et al. 2024; Tatineni 2019). Fairness in AI-driven mental health interventions ensures equitable access to care and minimizes disparities among diverse patient populations (Ferrara 2023). By promoting fairness, AI technologies can enhance the quality and effectiveness of mental health services while reducing the risk of harm (Ferrara 2023; Aditya Singhal et al. 2024).

There is a call for transparency in AI systems, including disclosure of how algorithms make decisions and accountability for their outcomes; this transparency fosters trust between users and AI systems (Habli et al. 2020; Khanna and Srivastava 2020; Kiseleva et al. 2022; Aditya Singhal et al. 2024; Vollmer et al. 2020). Transparency in AI systems is crucial for promoting accountability and trustworthiness (Kiseleva et al. 2022). Users should have insight into how algorithms make decisions and understand the limitations of AI technology (de Bruijn et al. 2022; Lee 2018). Ethical AI-driven mental health interventions prioritize transparency through clear explanations of algorithms' functionality and decision-making processes (Koutsouleris et al. 2022; Aditya Singhal et al. 2024). Accountability mechanisms hold developers and providers accountable for the outcomes of AI interventions, fostering responsible practice and ensuring positive outcomes for individuals (Habli et al. 2020; Kiseleva et al. 2022; Aditya Singhal et al. 2024; Vollmer et al. 2020).

Ethical AI in mental health respects individual autonomy and empowers users to make informed decisions about their treatment options (Fanni et al. 2023; Love 2023; Tiribelli 2023). Human oversight is essential to ensure that AI interventions complement rather than replace human judgment and agency (Fanni et al. 2023; Love 2023; Tiribelli 2023). Respecting individual autonomy and human agency is fundamental in ethical AI-driven mental health interventions (Alowais et al. 2023). While AI technologies can augment decision-making processes, human oversight is essential to ensure that interventions align with users' preferences and values (Fanni et al. 2023; Love 2023; Tiribelli 2023). Empowering individuals to actively participate in their mental health care promotes autonomy and self-determination (Fanni et al. 2023; Love 2023; Tiribelli 2023). Human-centered design approaches prioritize user autonomy and agency, emphasizing collaboration and shared decision making between users and AI systems (Margetis et al. 2021; Usmani et al. 2023).

Ensuring the safety and efficacy of AI-driven mental health interventions is paramount; this involves rigorous testing, validation, and ongoing monitoring to detect and mitigate potential adverse effects (Davahli et al. 2021; Ellahham et al. 2020; Habli et al. 2020; Morley et al. 2021; Tiwari and Dileep 2023). Ensuring the safety and efficacy of AI-driven mental health interventions is paramount to protecting individuals from harm. Rigorous testing, validation, and ongoing monitoring are essential to detect and mitigate potential adverse effects (Balcombe and De Leo 2021; Joerin et al. 2020; Tatineni 2019). Ethical AI practices prioritize safety and efficacy, prioritizing the well-being of users and minimizing the risk of unintended consequences (J. P. Singh 2021). By upholding safety standards, AI interventions can enhance the quality and accessibility of mental health care while promoting positive outcomes for individuals (S. Graham et al. 2019; Habli et al. 2020; Mensah 2023; Reddy et al. 2019).

Ethical frameworks and guidelines specific to promoting ethical AI in mental health are advocated for by the authors of many studies (Jeyaraman et al. 2023; Siala and Wang 2022). These frameworks provide a structured approach for addressing ethical challenges and promoting responsible practice (Siala and Wang 2022; Zhang et al. 2023). Developing and adopting ethical frameworks and guidelines specific to AI-driven mental health inter-

ventions provide a structured approach for addressing ethical challenges (Carr 2020). These frameworks offer guidance on ethical decision making, risk assessment, and responsible practice (Carr 2020; Molala and Makhubele 2021). Stakeholder engagement throughout the development and implementation process ensures that ethical considerations are adequately addressed, promoting transparency and accountability in AI-driven mental health interventions (Bélisle-Pipon et al. 2022; Couture et al. 2023; A. Singhal et al. 2024).

Practical strategies for mitigating bias in AI algorithms include diverse data collection, algorithmic audits, and ongoing evaluation to detect and correct biases that may arise during deployment (F. Chen et al. 2024; Ferrara 2023; Mensah 2023; Mittermaier et al. 2023). Mitigating bias in AI algorithms is essential to ensure equitable and fair outcomes in mental health interventions (Timmons et al. 2023). Diverse data collection, algorithmic audits, and bias mitigation strategies are critical components of ethical AI practice (F. Chen et al. 2024; Ferrara 2023; Mensah 2023; Mittermaier et al. 2023). Continuous evaluation and improvement efforts aim to detect and correct biases that may arise during deployment, promoting fairness and inclusivity in AI-driven mental health interventions (F. Chen et al. 2024; Mensah 2023; Mittermaier et al. 2023).

Ethical AI practices require continuous evaluation and improvement to adapt to evolving ethical standards, technological advancements, and user needs (WHO Guidance 2021; Magrabi et al. 2019; McGreevey et al. 2020; Morley et al. 2020). Continuous evaluation and improvement are integral to ethical AI practice in mental health settings (WHO Guidance 2021; McGreevey et al. 2020). Monitoring and evaluating outcomes enable developers and providers to identify areas for improvement and adapt to evolving ethical standards and user needs (WHO Guidance 2021; Magrabi et al. 2019; McGreevey et al. 2020; Morley et al. 2020). Regular ethical reviews and stakeholder feedback contribute to ongoing refinement and optimization of AI-driven mental health interventions, ensuring that they remain ethically sound and beneficial to individuals seeking care (Farhud and Zokaei 2021; WHO Guidance 2021; Leimanis and Palkova 2021; Nasir et al. 2024).

Several recent review studies, including that by Li et al. (2023), have critically examined the ethical implications of employing artificial intelligence (AI) in mental health interventions. Li, Han et al. synthesized evidence on the effectiveness of AI-driven conversational agents in enhancing mental health and well-being. Their findings offer valuable insights into the current evidence base for the use of conversational AI in mental health interventions, highlighting both its potential and limitations. Ethical concerns such as informed consent, privacy, transparency, and algorithmic bias were identified as significant challenges (Li et al. 2023). Another narrative review by A. M. Alhuwaydi (2024) explored the evolving role of AI in mental health care, addressing key challenges, limitations, and prospects. It underscored the potential of AI, particularly predictive analytics, in refining treatment strategies by predicting individual responses to interventions, thus aligning with the shift towards personalized mental health care. The review also scrutinized major ethical dimensions in AI-driven mental health, including algorithmic bias, data privacy, transparency, responsibility, and the doctor–patient relationship (Alhuwaydi 2024).

Additionally, Thakkar et al. (2024) contributed a narrative review discussing AI's applications in managing psychiatric disorders such as neurodegenerative disorders, intellectual disabilities, and seizures. The paper explored AI's role in enhancing awareness, diagnosis, and intervention for mental health conditions. While highlighting AI's potential benefits, the review acknowledged significant challenges, emphasizing the necessity of culturally sensitive and flexible algorithms to mitigate potential biases. It provided a comprehensive overview of AI's current landscape and future prospects in mental health, alongside critical considerations and limitations that warrant attention for its responsible and effective integration into mental health care (Thakkar et al. 2024).

Together, these studies underscore the pressing ethical issues that must be addressed to ensure the safe and ethical use of AI in supporting mental health care. They emphasize the importance of informed consent, data privacy, algorithmic transparency, and maintaining human-centric approaches in AI-powered mental health interventions.

Notably, regulating bodies such as the Federal Drug Administration (FDA) in the United States may play a role in ensuring that AI interventions are developed and deployed in an ethical manner. The AI interventions discussed in this paper could take many forms, such as specific algorithms, chatbots, or complete AI-enabled devices. AI-enabled medical devices may be subject to FDA approval, as noted in recent publications on the agency's website. While this can be a promising development for the protection of consumers, it will be critical that the FDA retain experts who are able to properly assess the ethical design and development of the AI components of these devices. Research like that discussed in this paper can offer an important source of information to inform the development of responsible regulation of AI devices.

5. Limitations of This Study

There are a few limitations to note for this study. This systematic review has a limited scope and mainly focuses on articles published in a specific time period (2014–2024) in a limited set of databases (PubMed, PsycINFO, Web of Science, and Scopus). This narrow scope ignores studies or perspectives from other time periods or sources and limits the generalizability of the findings. The review's reliance on published articles may introduce publication bias, as studies with significant findings are more likely to be published than studies with negative results. This bias can distort the overall interpretation of ethical considerations in AI-based mental health interventions. Limiting the search to articles published in specific databases may lead to language bias, as relevant studies published in other languages or regions may be overlooked. This limitation can affect the comprehensiveness of the review findings. This review may not provide an accurate assessment of the quality of included studies, potentially overlooking methodological flaws or biases in the literature. Without robust quality assessment criteria, the reliability and validity of pooled findings may be compromised. Despite efforts to conduct a systematic review, biases inherent in the processes of study selection, data extraction, and synthesis may affect the interpretation of the findings. Future research should aim to overcome these limitations to provide a more comprehensive understanding of ethical considerations in AI-based mental health interventions and to inform responsible practice and policy development.

6. Conclusions

Ethical considerations play a central role in the responsible implementation and impact of AI-driven mental health interventions. By addressing privacy, informed consent, bias, transparency, autonomy, safety, and efficacy, ethical AI practice promotes responsible practice and positive outcomes for individuals. Ethical frameworks, stakeholder engagement, bias mitigation strategies, and continuous evaluation efforts contribute to the ethical development and deployment of AI interventions, fostering trust, fairness, and effectiveness in mental healthcare delivery. As AI technology continues to evolve, prioritizing ethical considerations remains essential to maximizing benefits while minimizing potential harms in mental health interventions. These findings collectively underscore the importance of prioritizing ethical considerations in the development and deployment of AI interventions in mental health. By addressing these concerns, researchers and practitioners can ensure that AI technologies contribute positively to mental health care while minimizing potential risks and harms.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article or Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

Alhuwaydi, Ahmed M. 2024. Exploring the Role of Artificial Intelligence in Mental Healthcare: Current Trends and Future Directions—A Narrative Review for a Comprehensive Insight. *Risk Management and Healthcare Policy* 17: 1339–48. [CrossRef]

Alowais, Shuroug A., Sahar S. Alghamdi, Nada Alsuhebany, Tariq Alqahtani, Abdulrahman I. Alshaya, Sumaya N. Almohareb, Atheer Aldairem, Mohammed Alrashed, Khalid Bin Saleh, Hisham A. Badreldin, and et al. 2023. Revolutionizing healthcare: The role of artificial intelligence in clinical practice. *BMC Medical Education* 23: 689. [CrossRef]

Balcombe, Luke. 2023. AI Chatbots in Digital Mental Health. *Informatics* 10: 82. [CrossRef]

Balcombe, Luke, and Diego De Leo. 2021. Digital mental health challenges and the horizon ahead for solutions. *JMIR Mental Health* 8: e26811. [CrossRef]

Baskin, Alison S., Ton Wang, Jacquelyn Miller, Reshma Jagsi, Eve A. Kerr, and Lesly A. Dossett. 2021. A health systems ethical framework for de-implementation in health care. *Journal of Surgical Research* 267: 151–58. [CrossRef]

Bélisle-Pipon, Jean-Christophe, Erica Monteferrante, Marie-Christine Roy, and Vincent Couture. 2022. Artificial intelligence ethics has a black box problem. *AI & Society* 38: 1507–22. [CrossRef]

Carr, Sarah. 2020. 'AI gone mental': Engagement and ethics in data-driven technology for mental health. *Journal of Mental Health* 29: 125–30. [CrossRef]

Chalyi, Oleksii. 2024. An Evaluation of General-Purpose AI Chatbots: A Comprehensive Comparative Analysis. *InfoScience Trends* 1: 52–66. [CrossRef]

Charlson, Fiona, van Mark Ommeren, Abraham Flaxman, Joseph Cornett, Harvey Whiteford, and Shekhar Saxena. 2019. New WHO prevalence estimates of mental disorders in conflict settings: A systematic review and meta-analysis. *The Lancet* 394: 240–48. [CrossRef]

Chen, Feng, Liqin Wang, Julie Hong, Jiaqi Jiang, and Li Zhou. 2024. Unmasking bias in artificial intelligence: A systematic review of bias detection and mitigation strategies in electronic health record-based models. *Journal of the American Medical Informatics Association* 31: 1172–83. [CrossRef]

Chen, Yan, and Pouyan Esmaeilzadeh. 2024. Generative AI in medical practice: In-depth exploration of privacy and security challenges. *Journal of Medical Internet Research* 26: e53008. [CrossRef]

Chintala, Sathish Kumar. 2022. Data Privacy and Security Challenges in AI-Driven Healthcare Systems in India. *Journal of Data Acquisition and Processing* 37: 2769–78.

Cohen, I. Glenn. 2019. Informed consent and medical artificial intelligence: What to tell the patient? *The Georgetown Law Journal* 108: 1425. [CrossRef]

Couture, Vincent, Marie-Christine Roy, Emma Dez, Samuel Laperle, and Jean-Christophe Bélisle-Pipon. 2023. Ethical implications of artificial intelligence in population health and the public's role in its governance: Perspectives from a citizen and expert panel. *Journal of Medical Internet Research* 25: e44357. [CrossRef]

Davahli, Mohammad Reza, Waldemar Karwowski, Krzysztof Fiok, Thomas Wan, and Hamid R. Parsaei. 2021. Controlling safety of artificial intelligence-based systems in healthcare. *Symmetry* 13: 102. [CrossRef]

de Bruijn, Hans, Martijn Warnier, and Marijn Janssen. 2022. The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. *Government Information Quarterly* 39: 101666. [CrossRef]

Ellahham, Samer, Nour Ellahham, and Mecit Can Emre Simsekler. 2020. Application of artificial intelligence in the health care safety context: Opportunities and challenges. *American Journal of Medical Quality* 35: 341–48. [CrossRef]

Faezi, Aysan, and Bahman Alinezhad. 2024. AI-Enhanced Health Tools for Revolutionizing Hypertension Management and Blood Pressure Control. *InfoScience Trends* 1: 67–72. [CrossRef]

Fanni, Rosanna, Giulia Zampedri Valerie Eveline Steinkogler, and Jo Pierson. 2023. Enhancing human agency through redress in Artificial Intelligence Systems. *AI & Society* 38: 537–47.

Farhud, Dariush D., and Shaghayegh Zokaei. 2021. Ethical Issues of Artificial Intelligence in Medicine and Healthcare. *Iranian Journal of Public Health* 50: I–V. [CrossRef]

Ferrara, Emilio. 2023. Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci* 6: 3. [CrossRef]

Gaonkar, Bilwaj, Kirstin Cook, and Luke Macyszyn. 2023. Ethical Issues Arising Due to Bias in Training A.I. Algorithms in Healthcare and Data Sharing as a Potential Solution. *The AI Ethics Journal* 1: 1–14. [CrossRef]

Ghadiri, Pooria. 2022. *Artificial Intelligence Interventions in the Mental Healthcare of Adolescents*. Montréal: McGill University.

Gooding, Piers, and Timothy Kariotis. 2021. Ethics and law in research on algorithmic and data-driven technology in mental health care: Scoping review. *JMIR Mental Health* 8: e24668. [CrossRef]

Graham, Sarah, Colin Depp, Ellen E. Lee, Camille Nebeker, Xin Tu, Ho-Cheol Kim, and Dilip V. Jeste. 2019. Artificial intelligence for mental health and mental illnesses: An Overview. *Current Psychiatry Reports* 21: 116. [CrossRef]

Habli, Ibrahim, Tom Lawton, and Zoe Porter. 2020. Artificial intelligence in health care: Accountability and safety. *Bulletin of the World Health Organization* 98: 251–56. [CrossRef] [PubMed]

Jeyaraman, Madhan, Sangeetha Balaji, Naveen Jeyaraman, and Sankalp Yadav. 2023. Unraveling the Ethical Enigma: Artificial Intelligence in Healthcare. *Cureus* 15: e43262. [CrossRef]

Joerin, Angela, Michiel Rauws, Russell Fulmer, and Valerie Black. 2020. Ethical Artificial Intelligence for Digital Health Organizations. *Cureus* 12: e7202. [CrossRef]

Kasula, Balaram Yadav. 2023. Ethical Considerations in the Adoption of Artificial Intelligence for Mental Health Diagnosis. *International Journal of Creative Research In Computer Technology and Design* 5: 1–7.

Kerasidou, Angeliki. 2021. Ethics of artificial intelligence in global health: Explainability, algorithmic bias and trust. *Journal of Oral Biology and Craniofacial Research* 11: 612–14. [CrossRef]

Khanna, Shivansh, and Shraddha Srivastava. 2020. Patient-centric ethical frameworks for privacy, transparency, and bias awareness in deep learning-based medical systems. *Applied Research in Artificial Intelligence and Cloud Computing* 3: 16–35.

Kiseleva, Anastasiya, Dimitris Kotzinos, and Paul De Hert. 2022. Transparency of AI in healthcare as a multilayered system of accountabilities: Between legal requirements and technical limitations. *Frontiers in Artificial Intelligence* 5: 879603. [CrossRef] [PubMed]

Koutsouleris, Nikolaos, Tobias U Hauser, Vasilisa Skvortsova, and Munmun De Choudhury. 2022. From promise to practice: Towards the realisation of AI-informed mental health care. *The Lancet Digital Health* 4: e829–e840. [CrossRef] [PubMed]

Lee, Min Kyung. 2018. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society* 5: 2053951718756684. [CrossRef]

Leimanis, Anrī, and Karina Palkova. 2021. Ethical guidelines for artificial intelligence in healthcare from the sustainable development perspective. *European Journal of Sustainable Development* 10: 90. [CrossRef]

Li, Han, Renwen Zhang, Yi-Chieh Lee, Robert E. Kraut, and David C. Mohr. 2023. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *NPJ Digital Medicine* 6: 236. [CrossRef] [PubMed]

Love, Charles S. 2023. "Just the Facts Ma'am": Moral and Ethical Considerations for Artificial Intelligence in Medicine and its Potential to Impact Patient Autonomy and Hope. *The Linacre Quarterly* 90: 375–94. [CrossRef] [PubMed]

Luxton, David D. 2014. Artificial intelligence in psychological practice: Current and future applications and implications. *Professional Psychology: Research and Practice* 45: 332–39. [CrossRef]

Magrabi, Farah, Elske Ammenwerth, Jytte Brender McNair, Nicolet F. De Keizer, Hannele Hyppönen, Pirkko Nykänen, Michael Rigby, Philip J. Scott, Tuulikki Vehko, Zoie Shui-Yee Wong, and et al. 2019. Artificial intelligence in clinical decision support: Challenges for evaluating AI and practical implications. *Yearbook of Medical Informatics* 28: 128–34. [CrossRef] [PubMed]

Margetis, George, Stavroula Ntoa, Margherita Antona, and Constantine Stephanidis. 2021. Human-centered design of artificial intelligence. In *Handbook of Human Factors and Ergonomics*. Hoboken: John Wiley & Sons, Inc., pp. 1085–106. [CrossRef]

Martin, Clarissa, Kyle DeStefano, Harry Haran, Sydney Zink, Jennifer Dai, Danial Ahmed, Abrahim Razzak, Keldon Lin, Ann Kogler, Joseph Waller, and et al. 2022. The ethical considerations including inclusion and biases, data protection, and proper implementation among AI in radiology and potential implications. *Intelligence-Based Medicine* 6: 100073. [CrossRef]

McGreevey, John D., C. William Hanson, and Ross Koppel. 2020. Clinical, legal, and ethical aspects of artificial intelligence-assisted conversational agents in health care. *JAMA* 324: 552–53. [CrossRef]

McKay, Francis, Bethany J. Williams, Graham Prestwich, Daljeet Bansal, Darren Treanor, and Nina Hallowell. 2023. Artificial intelligence and medical research databases: Ethical review by data access committees. *BMC Medical Ethics* 24: 49. [CrossRef]

Mennella, Ciro, Umberto Maniscalco, Giuseppe De Pietro, and Massimo Esposito. 2024. Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon* 10: e26297. [CrossRef] [PubMed]

Mensah, George Benneh. 2023. *Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems*. Available online: https://www.researchgate.net/profile/George-Benneh-Mensah-2/publication/375744287_Artificial_Intelligence_and_Ethics_A_Comprehensive_Reviews_of_Bias_Mitigation_Transparency_and_Accountability_in_AI_Systems/links/656c8e46b86a1d521b2e2a16/Artificial-Intelligence-and-Ethics-A-Comprehensive-Reviews-of-Bias-Mitigation-Transparency-and-Accountability-in-AI-Systems.pdf (accessed on 26 January 2024).

Mittermaier, Mirja, Marium M. Raza, and Joseph C. Kvedar. 2023. Bias in AI-based models for medical applications: Challenges and mitigation strategies. *NPJ Digital Medicine* 6: 113. [PubMed]

Molala, Thomas, and Jabulani Makhubele. 2021. A conceptual framework for the ethical deployment of Artificial Intelligence in addressing mental health challenges: Guidelines for Social Workers. *Technium Social Science Journal* 24: 696.

Morley, Jessica, Caio C.V. Machado, Christopher Burr, Josh Cowls, Indra Joshi, Mariarosaria Taddeo, and Luciano Floridi. 2020. The ethics of AI in health care: A mapping review. *Social Science & Medicine* 260: 113172. [CrossRef]

Morley, Jessica, Kassandra Karpathakis Caroline Morton, Mariarosaria Taddeo, and Luciano Floridi. 2021. Towards a framework for evaluating the safety, acceptability and efficacy of AI systems for health: An initial synthesis. *arXiv* arXiv:2104.06910.

Mörch, Carl-Maria, Abhishek Gupta, and Brian L. Mishara. 2020. Canada protocol: An ethical checklist for the use of artificial Intelligence in suicide prevention and mental health. *Artificial Intelligence in Medicine* 108: 101934. [CrossRef] [PubMed]

Murdoch, Blake. 2021. Privacy and artificial intelligence: Challenges for protecting health information in a new era. *BMC Medical Ethics* 22: 122. [CrossRef]

Nasir, Sidra, Rizwan Ahmed Khan, and Samita Bai. 2024. Ethical Framework for Harnessing the Power of AI in Healthcare and Beyond. *IEEE Access* 12: 31014–35. [CrossRef]

Olawade, David B., Aderonke Odetayo Ojima Z. Wada, Fiyinfoluwa Asaolu Aanuoluwapo Clement David-Olawade, and Judith Eberhardt. 2024. Enhancing mental health with Artificial Intelligence: Current trends and future prospects. *Journal of Medicine, Surgery, and Public Health* 3: 100099. [CrossRef]

Olorunsogo, Tolulope, Adekunle Oyeyemi Adenyi, Chioma Anthonia Okolo, and Oloruntoba Babawarun. 2024. Ethical considerations in AI-enhanced medical decision support systems: A review. *World Journal of Advanced Engineering Technology and Sciences* 11: 329–36. [CrossRef]

Page, Matthew J., Patrick M. Bossuyt, Joanne E. McKenzie, Tammy C. Hoffmann, Isabelle Boutron, Larissa Shamseer, Cynthia D. Mulrow, Elie A. Akl, Jennifer M. Tetzlaff, Sue E. Brennan, and et al. 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 372: n71. [CrossRef]

Pickering, Brian. 2021. Trust, but verify: Informed consent, AI technologies, and public health emergencies. *Future Internet* 13: 132. [CrossRef]

Prathomwong, Piyanat, and Pagorn Singsuriya. 2022. Ethical framework of digital technology, artificial intelligence, and health equity. *Asia Social Issues* 15: 252136. [CrossRef]

Reddy, Sandeep, John Fox, and Maulik P. Purohit. 2019. Artificial intelligence-enabled healthcare delivery. *Journal of the Royal Society of Medicine* 112: 22–28. [CrossRef]

Rubeis, Giovanni. 2022. iHealth: The ethics of artificial intelligence and big data in mental healthcare. *Internet Interventions* 28: 100518. [CrossRef] [PubMed]

Saeidnia, Hamid Reza. 2023. Ethical artificial intelligence (AI): Confronting bias and discrimination in the library and information industry. *Library Hi Tech News*, *ahead-of-print*. [CrossRef]

Shah, Varun. 2022. AI in Mental Health: Predictive Analytics and Intervention Strategies. *Journal Environmental Sciences And Technology* 1: 55–74.

Shaw, James. 2022. Emerging paradigms for ethical review of research using artificial intelligence. *American Journal of Bioethics* 22: 42–44. [CrossRef]

Shimada, Koki. 2023. The role of artificial intelligence in mental health: A review. *Science Insights* 43: 1119–27. [CrossRef]

Siala, Haytham, and Yichuan Wang. 2022. SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. *Social Science & Medicine* 296: 114782. [CrossRef]

Singh, Jatin Pal. 2021. AI Ethics and Societal Perspectives: A Comparative Study of Ethical Principle Prioritization Among Diverse Demographic Clusters. *Journal of Advanced Analytics in Healthcare Management* 5: 1–18.

Singh, Vipul, Sharmila Sarkar, Vikas Gaur, Sandeep Grover, and Om Prakash Singh. 2024. Clinical Practice Guidelines on using artificial intelligence and gadgets for mental health and well-being. *Indian Journal of Psychiatry* 66: S414–S419. [CrossRef] [PubMed]

Singhal, Aditya, Nikita Neveditsin, Hasnaat Tanveer, and Vijay Mago. 2024. Toward Fairness, Accountability, Transparency, and Ethics in AI for Social Media and Health Care: Scoping Review. *JMIR Public Health and Surveillance* 12: e50048. [CrossRef] [PubMed]

Sivan, Remya, and Zuriati Ahmad Zukarnain. 2021. Security and privacy in cloud-based e-health system. *Symmetry* 13: 742. [CrossRef]

Skorburg, Joshua August, Kieran O'Doherty, and Phoebe Friesen. 2024. Persons or data points? Ethics, artificial intelligence, and the participatory turn in mental health research. *American Psychologist* 79: 137–49. [CrossRef] [PubMed]

Smith, Valerie, Declan Devane, Cecily M Begley, and Mike Clarke. 2011. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Medical Research Methodology* 11: 15. [CrossRef] [PubMed]

Sqalli, Mohammed Tahri, Begali Aslonov, Mukhammadjon Gafurov, and Shokhrukhbek Nurmatov. 2023. Humanizing AI in medical training: Ethical framework for responsible design. *Frontiers in Artificial Intelligence* 6: 1189914. [CrossRef] [PubMed]

Tatineni, Sumanth. 2019. Ethical Considerations in AI and Data Science: Bias, Fairness, and Accountability. *International Journal of Information Technology and Management Information Systems* 10: 11–20.

Thakkar, Anoushka, Ankita Gupta, and Avinash De Sousa. 2024. Artificial intelligence in positive mental health: A narrative review. *Frontiers in Digital Health* 6: 1280235. [CrossRef]

Timmons, Adela C., Jacqueline B. Duong, Natalia Simo Fiallo, Theodore Lee, Huong Phuc Quynh Vo, Matthew W. Ahle, Jonathan S. Comer, LaPrincess C. Brewer, Stacy L. Frazier, and Theodora Chaspari. 2023. A call to action on assessing and mitigating bias in artificial intelligence applications for mental health. *Perspectives on Psychological Science* 18: 1062–96. [CrossRef]

Tiribelli, Simona. 2023. The AI ethics principle of autonomy in health recommender systems. *Argumenta* 16: 1–18.

Tiwari, Vikash Kumar, and M. R. Dileep. 2023. An Efficacy of Artificial Intelligence Applications in Healthcare Systems—A Bird View. In *Information and Communication Technology for Competitive Strategies (ICTCS 2022) Intelligent Strategies for ICT*. Singapore: Springer, pp. 649–59.

Ursin, Frank, Marcin Orzechowski Cristian Timmermann, and Florian Steger. 2021. Diagnosing diabetic retinopathy with artificial intelligence: What information should be included to ensure ethical informed consent? *Frontiers in Medicine* 8: 695217. [CrossRef] [PubMed]

Usmani, Usman Ahmad, Ari Happonen, and Junzo Watada. 2023. Human-Centered Artificial Intelligence: Designing for User Empowerment and Ethical Considerations. Paper presented at 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), İstanbul, Türkiye, June 8–10; pp. 1–5. [CrossRef]

Vollmer, Sebastian, Bilal A. Mateen, Gergo Bohner, Franz J. Király, Rayid Ghani, Pall Jonsson, Sarah Cumbers, Adrian Jonas, Katherine S. L. McAllister, Puja Myles, and et al. 2020. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *BMJ* 368: l6927. [CrossRef] [PubMed]

Wainberg, Milton L., Pamela Scorza, James M. Shultz, Liat Helpman, Jennifer J. Mootz, Karen A. Johnson, Yuval Neria, Jean-Marie E. Bradford, Maria A. Oquendo, and Melissa R. Arbuckle. 2017. Challenges and Opportunities in Global Mental Health: A Research-to-Practice Perspective. *Current Psychiatry Reports* 19: 28. [CrossRef]

WHO Guidance. 2021. *Ethics and Governance of Artificial Intelligence for Health*. Geneva: World Health Organization.

Yelne, Seema, Minakshi Chaudhary, Karishma Dod, Akhtaribano Sayyad, and Ranjana Sharma. 2023. Harnessing the Power of AI: A Comprehensive Review of Its Impact and Challenges in Nursing Science and Healthcare. *Cureus* 15: e49252. [CrossRef]

Zhang, Melody, Jillian Scandiffio, Sarah Younus, Tharshini Jeyakumar, Inaara Karsan, Rebecca Charow, Mohammad Salhia, and David Wiljer. 2023. The Adoption of AI in Mental Health Care–Perspectives From Mental Health Professionals: Qualitative Descriptive Study. *JMIR Formative Research* 7: e47847. [CrossRef]