International Journal of Innovative Technologies in Social Science

e-ISSN: 2544-9435

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

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

ARTICLE TITLE	ARTIFICIAL INTELLIGENCE IN MENTAL HEALTH CARE: OPPORTUNITIES, CHALLENGES, AND ETHICAL DILEMMAS
ARTICLE INFO	Piotr Rzyczniok, Mateusz Kopczyński, Aneta Rasińska, Justyna Matusik, Justyna Jachimczak, Paulina Bala. (2025) Artificial Intelligence in Mental Health Care: Opportunities, Challenges, and Ethical Dilemmas. *International Journal of Innovative Technologies in Social Science*. 3(47). doi: 10.31435/ijitss.3(47).2025.3529
DOI	https://doi.org/10.31435/ijitss.3(47).2025.3529
RECEIVED	28 June 2025
ACCEPTED	05 August 2025
PUBLISHED	12 August 2025
LICENSE	The article is licensed under a **Creative Commons Attribution 4.0 International License.**

ARTIFICIAL INTELLIGENCE IN MENTAL HEALTH CARE: OPPORTUNITIES, CHALLENGES, AND ETHICAL DILEMMAS

Piotr Rzyczniok *(Corresponding Author, Email: piorzy@gmail.com)*
Private Healthcare Institution ProCordi Ltd., 44-335 Jastrzębie Zdrój, ul. Wyspiańskiego 8, Poland
ORCID ID: 0009-0004-5899-2345

Mateusz Kopczyński
Municipal Hospital Complex in Częstochowa, Częstochowa 42-200, ul. Mirowska 15, Poland
ORCID ID: 0009-0009-6592-3061

Aneta Rasińska
Municipal Hospital No. 4 in Gliwice Ltd. Zygmunta Starego 20, 44-100, Poland
ORCID ID: 0009-0007-2872-8000

Justyna Matusik
Academy of Silesia in Katowice, 40-555 Katowice, ul. Rolna 43, Poland
ORCID ID: 0009-0009-7907-9690

Justyna Jachimczak
Our Lady of Perpetual Help Hospital in Wołomin, 05-200 Wołomin, ul. Gdyńska 1/3, Poland
ORCID ID: 0009-0005-7614-9500

Paulina Bala
Our Lady of Perpetual Help Hospital in Wołomin, 05-200 Wołomin, ul. Gdyńska 1/3, Poland
ORCID ID: 0009-0007-0604-8549

ABSTRACT

Introduction and Objective: The increasing global burden of mental health disorders, exacerbated by the COVID-19 pandemic and the limitations of traditional mental health systems, has accelerated interest in digital health solutions. Artificial intelligence (AI) has emerged as a transformative force in mental health care, offering tools for diagnosis, intervention, and patient monitoring. This review aims to explore current applications, opportunities, and ethical challenges of AI-based tools in mental health, with an emphasis on responsible and equitable deployment.
Review Methods: A narrative literature review was conducted using PubMed, Scopus, Web of Science, and Google Scholar. Peer-reviewed articles published between 2014 and 2022 were considered, with a focus on interdisciplinary sources covering clinical psychology, digital health technologies, AI development, and medical ethics. Key themes were synthesized across domains to provide a holistic understanding.
State of Knowledge: AI technologies, including chatbots, machine learning algorithms, and predictive analytics, are increasingly integrated into mental health services. They offer scalable solutions for screening, personalized intervention, and early risk detection. However, concerns remain about algorithmic bias, privacy, transparency, and the digital divide. The current body of evidence supports AI's potential to complement—rather than replace—human care, particularly when integrated responsibly within clinical frameworks.
Conclusion: AI holds significant promise in improving access, personalization, and efficiency in mental health care. To harness its benefits, interdisciplinary collaboration, robust ethical oversight, and patient-centered design are essential. Further research is needed to evaluate long-term outcomes and ensure AI systems uphold clinical integrity, equity, and trust.

KEYWORDS
Artificial Intelligence (AI), Mental Health, Digital Health Tools, Ethical Challenges, Telepsychiatry, Health Equity

CITATION
Piotr Rzyczniok, Mateusz Kopczyński, Aneta Rasińska, Justyna Matusik, Justyna Jachimczak, Paulina Bala. (2025) Artificial Intelligence in Mental Health Care: Opportunities, Challenges, and Ethical Dilemmas. *International Journal of Innovative Technologies in Social Science*. 3(47). doi: 10.31435/ijitss.3(47).2025.3529

Introduction.

Mental health disorders are among the leading causes of disability worldwide, affecting over 970 million people as of recent estimates (World Health Organization [WHO], 2022). Conditions such as depression, anxiety, and post-traumatic stress disorder (PTSD) have not only increased in prevalence over the past decades but have also gained renewed attention due to the global impact of the COVID-19 pandemic. The pandemic highlighted significant shortcomings in mental health service availability, particularly in underserved areas and populations with limited access to psychiatric care (Pfefferbaum & North, 2020).

Parallel to the rising demand for mental health services is the rapid evolution of digital health technologies. Among these, artificial intelligence (AI) has emerged as a promising tool for transforming mental health care delivery. AI-driven applications, including mobile health (mHealth) platforms, conversational agents, predictive analytics, and emotion recognition algorithms, are increasingly being developed to enhance screening, diagnosis, and treatment of mental illnesses (Torous et al., 2021). These innovations offer the potential for scalable, cost-effective, and personalized care—addressing critical gaps in traditional mental health systems.

However, the integration of AI into mental health care also raises complex questions regarding ethical use, clinical validation, data privacy, and equitable access. Understanding the implications of AI's growing role in this space requires an interdisciplinary approach that considers technological capabilities, clinical evidence, public health needs, and societal impact.

This review aims to explore the current landscape of AI applications in mental health care, highlighting both the opportunities and the ethical dilemmas that come with these technologies. It also discusses the future directions needed to responsibly integrate AI into mental health systems, especially in preparation for future global health crises.

Methodology

This narrative review was conducted to explore the role, potential, and challenges of AI in mental health care, with a particular focus on post-2010 developments in digital psychiatry, machine learning applications, and ethical frameworks. The narrative format was chosen to allow for a broader interpretative synthesis of interdisciplinary sources across clinical psychology, digital health, bioethics, and information science.

A comprehensive literature search was performed using databases such as PubMed, Scopus, Web of Science, and Google Scholar. The search strategy incorporated keywords including *"artificial intelligence"*, *"mental health"*, *"digital psychiatry"*, *"machine learning"*, *"chatbots in therapy"*, *"predictive analytics in psychiatry"*, *"ethical AI"*, and *"algorithmic bias"*. Preference was given to peer-reviewed journal articles published between 2014 and 2022.

Articles were selected based on their contribution to one or more of the following domains: effectiveness of AI-based interventions in mental health care, ethical considerations surrounding AI implementation, issues of equity and access, and the impact of AI on clinical decision-making. Literature from both medical and technological fields was included to reflect the interdisciplinary nature of the topic.

This review does not follow the systematic review protocol (e.g., PRISMA) but instead adopts a critical and thematic approach to integrate findings, identify trends, and suggest implications for future research and practice.

Overview of AI Technologies in Mental Health Applications

AI encompasses a wide array of computational techniques, including machine learning (ML), natural language processing (NLP), and deep learning, which enable machines to mimic cognitive functions such as learning and reasoning. In mental health care, these technologies are increasingly leveraged to support early detection, diagnosis, monitoring, and treatment of psychiatric conditions.

One of the most prominent applications of AI in this domain is the use of predictive analytics to identify individuals at risk of developing mental illnesses. Machine learning algorithms can analyze diverse data sources—such as electronic health records (EHRs), social media activity, and wearable device data—to detect behavioral patterns indicative of conditions like depression, anxiety, or suicidal ideation (Inkster et al., 2018; Shatte et al., 2019). For example, NLP models can process large volumes of unstructured text, such as therapy session transcripts or patient-reported outcomes, to extract clinically relevant insights.

Conversational agents or chatbots—such as Woebot, Wysa, and Tess—use AI to provide real-time psychological support and cognitive behavioral therapy (CBT)-based interventions. These tools are particularly valuable in resource-limited settings, where human mental health professionals may be scarce (Fitzpatrick et al., 2017). Moreover, emotion recognition systems utilizing voice and facial expression analysis are being developed to enhance diagnostic precision and patient engagement, although their real-world effectiveness remains under evaluation.

Despite the promise, many AI tools in mental health are still in their nascent stages, requiring robust clinical validation and regulatory oversight. Ensuring their reliability, transparency, and generalizability across populations is a key challenge as these technologies become more integrated into care delivery.

Benefits and Opportunities

AI is increasingly recognized as a game-changing tool in the mental health landscape, offering innovative solutions to longstanding challenges in diagnosis, treatment, and service delivery. One of the most significant benefits of AI is its potential to enhance accessibility to mental health care. Globally, millions of people live in areas where mental health professionals are scarce or altogether unavailable (World Health Organization [WHO], 2021). AI-powered tools such as mental health chatbots, virtual therapists, and self-guided cognitive behavioral therapy (CBT) applications are helping to bridge this treatment gap. For example, apps like Woebot and Wysa employ natural language processing to provide empathetic, evidence-based support, and are available 24/7, offering immediate assistance without the long waiting times typical of traditional services (Fitzpatrick et al., 2017; Inkster et al., 2018).

AI also contributes significantly to early identification and intervention. By analyzing data from smartphones, wearables, and social media activity, AI algorithms can detect early warning signs of mental health conditions such as depression, anxiety, or suicidal ideation—often before a person becomes consciously aware of their symptoms (Barnett et al., 2018). These systems utilize machine learning to detect subtle changes in behavior, speech patterns, sleep, or physical activity, enabling real-time risk assessments and allowing for preventive measures before crises occur (Mohr et al., 2017).

Furthermore, AI fosters personalization in mental health treatment. Traditional therapeutic models often rely on trial-and-error to find effective interventions, but AI can synthesize vast datasets to recommend tailored treatments based on individual history, genetic information, comorbidities, and even real-time mood data. Platforms like Quartet or IBM Watson Health are designed to optimize treatment matching, potentially improving patient adherence and outcomes (Chekroud et al., 2016; Price et al., 2019).

In clinical settings, AI also reduces the administrative burden on healthcare professionals by automating routine but time-consuming tasks such as patient intake documentation, scheduling, transcriptions, and progress tracking. This efficiency not only improves service delivery but allows clinicians to spend more time engaging directly with patients (Jiang et al., 2017). Additionally, advanced AI tools are being developed to assist clinicians in their diagnostic reasoning, offering decision support that integrates clinical guidelines with real-time patient data.

AI-driven sentiment analysis and voice emotion recognition are even being used during psychotherapy sessions to provide therapists with objective feedback on emotional tone, speech hesitations, and rapport quality, opening new avenues for therapist training and supervision (Miner et al., 2016). These insights can support clinical judgment, especially for novice practitioners or in telehealth settings where visual cues may be limited.

Overall, the integration of AI into mental health care represents a paradigm shift toward more scalable, personalized, and preventive services. While not a substitute for human empathy and clinical expertise, AI acts as a powerful complement that can augment existing services and expand care to populations previously left behind.

Challenges and Ethical Considerations

Despite its promising benefits, the integration of AI into mental health care presents a complex set of challenges and ethical concerns that require careful scrutiny. One of the foremost issues is data privacy and security. AI systems rely heavily on large volumes of personal and often sensitive data—including mood logs, biometric signals, and online behavioral patterns. If mishandled or breached, such data could expose users to stigmatization, discrimination, or even psychological harm. Unlike general medical data, mental health information often carries a deeper layer of vulnerability, and yet many mental health apps lack clear, enforceable privacy policies (Huckvale et al., 2019; Martinez-Martin et al., 2018).

Another critical concern is algorithmic bias. AI systems trained on datasets that do not adequately represent diverse populations—by race, gender, age, socioeconomic status, or geography—may perpetuate or even exacerbate existing disparities in mental health care. For instance, predictive algorithms may fail to identify depressive symptoms in non-Western populations or misinterpret cultural expressions of distress, leading to misdiagnoses or exclusion from care pathways (Obermeyer et al., 2019). This raises important questions about fairness and equity, particularly for marginalized or underrepresented communities.

Transparency and explainability are also major limitations in the current AI models. Many advanced AI systems, particularly those utilizing deep learning, function as "black boxes"—producing outcomes without clear, interpretable rationales. In the context of mental health, where therapeutic relationships are built on trust and understanding, it is ethically problematic for clinicians or patients to rely on recommendations that cannot be meaningfully explained (Samek et al., 2017). The lack of transparency undermines informed consent and erodes patient autonomy, especially in high-stakes contexts like suicide risk assessment or psychosis detection.

Additionally, over-reliance on AI tools may lead to a dehumanization of care. While AI can support clinical decision-making or supplement therapy, it cannot replicate human empathy, contextual judgment, or the nuanced understanding of a trained mental health professional. There is a growing concern that commercial pressures or health system efficiencies may lead to replacing rather than augmenting clinicians, potentially reducing the quality of care, particularly for complex cases (Blease et al., 2019).

The use of chatbots and digital agents in crisis intervention also raises critical safety concerns. Although AI-driven mental health bots can provide initial support, they are not equipped to handle acute psychiatric emergencies or unpredictable user behavior. Instances have been documented where bots gave inappropriate or even harmful responses to disclosures of suicidal ideation or trauma (Miner et al., 2016). The absence of regulatory standards means that many of these tools are deployed with minimal oversight, leading to a “wild west” landscape of digital mental health technologies.

Lastly, legal and regulatory frameworks have not kept pace with the rapid development of AI technologies. Jurisdictions differ widely in how they regulate medical devices, digital health tools, and data protection. This patchwork of standards creates uncertainty around liability, accountability, and user rights. Who is responsible if an AI system fails to flag suicidal ideation? How should clinicians incorporate AI suggestions into medical records? These unresolved questions require international collaboration and interdisciplinary governance (Vayena et al., 2018).

Taken together, these challenges underscore the necessity of developing AI in mental health with a strong ethical foundation—one that prioritizes equity, privacy, transparency, and human oversight. Without these safeguards, the very technologies designed to democratize and enhance mental health care risk reinforcing the structural inequities they aim to solve.

Future Directions

The future of AI in mental health care is poised to be transformative, provided that innovations are developed in tandem with evidence-based practices and stakeholder collaboration. One promising direction is the integration of multimodal data sources—such as speech patterns, facial expressions, smartphone usage, and physiological signals—to create more holistic and accurate assessments of mental health status (Cummins et al., 2018). These advances aim to move beyond static diagnostic categories and enable real-time monitoring of psychological well-being. Moreover, developments in natural language processing (NLP) and affective computing offer the potential to identify subtle linguistic and emotional cues associated with depression, anxiety, or suicidal ideation, which could support earlier interventions (Calvo et al., 2017).

Another key area is the use of personalized AI-driven interventions, where mental health support is dynamically tailored to an individual’s preferences, needs, and context. Adaptive digital platforms could adjust content delivery based on user engagement or emotional state, making therapy more responsive and effective. Additionally, cross-disciplinary collaborations—between technologists, clinicians, ethicists, and patient groups—are essential for designing tools that are not only technologically advanced but also user-centered and clinically meaningful.

Emerging trends also include global scalability of mental health services through AI, particularly in resource-limited settings. AI-powered tools can bridge gaps in access by delivering basic psychological support via smartphones or chatbots in underserved regions. Finally, the incorporation of explainable AI (XAI) is gaining attention as a means of making algorithmic decisions more transparent and acceptable in clinical environments (Gunning & Aha, 2019). Such transparency is critical for integrating AI into mental health systems where trust and accountability are paramount.

Recommendations for Responsible AI Use in Mental Health

To ensure that AI technologies are deployed ethically and effectively in mental health care, a series of structured, responsible practices must be established. First and foremost, regulatory oversight and standardized evaluation protocols should be implemented before any AI-based tools are introduced into clinical settings. These standards must assess clinical validity, user safety, and psychological impact, much like pharmaceutical trials (Topol, 2019). Without such regulation, there is a risk that ineffective or harmful tools could be marketed directly to vulnerable users.

Second, developers and researchers should adhere to principles of fairness, inclusivity, and cultural competence. This includes ensuring diverse training datasets and involving users from different socioeconomic and cultural backgrounds in both design and testing phases. Addressing systemic bias proactively is essential to prevent the amplification of existing health disparities (Jobin et al., 2019).

Third, AI tools must be designed with human-in-the-loop frameworks, where clinicians retain decision-making authority and AI serves as an assistive tool rather than a replacement. This not only preserves the therapeutic alliance between provider and patient but also helps ensure that AI outputs are interpreted within the appropriate clinical context. The importance of informed consent and digital literacy cannot be overstated; users must clearly understand how their data is collected, processed, and used, and be given meaningful choices about their engagement with AI systems.

Lastly, ongoing monitoring and post-deployment auditing of AI tools should be institutionalized to assess real-world effectiveness, unintended consequences, and user satisfaction over time. Feedback loops involving clinicians and patients can drive iterative improvements, while ethical review boards can oversee compliance with emerging guidelines and public expectations (Morley et al., 2020).

Discussion

The integration of AI into mental health care marks a profound shift in how psychological conditions are detected, monitored, and managed. While earlier sections of this review have outlined the technological and clinical promise of AI, it is essential to critically reflect on both its opportunities and unresolved challenges. AI applications, including chatbots, sentiment analysis, and predictive diagnostics, have shown effectiveness in enhancing accessibility and personalization of care (Fitzpatrick et al., 2017; Inkster et al., 2018). However, their real-world utility depends on a complex interplay of technological robustness, ethical safeguards, user trust, and systemic integration.

One key issue is the risk of over-reliance on AI tools in emotionally nuanced domains such as mental health. Algorithms lack the empathetic depth and contextual sensitivity that human clinicians provide, making them prone to misinterpretation of emotional cues, especially among diverse populations or atypical presentations (Luxton, 2014). Additionally, without adequate regulatory frameworks and clinical validation, there is a danger that poorly tested applications may enter the market, potentially causing harm or fostering false reassurance among users (He et al., 2019).

Moreover, the ethical tensions surrounding data privacy, autonomy, and consent remain central to the discourse. Mental health data is particularly sensitive, and many AI applications rely on continuous data collection through smartphones or wearable devices. Despite anonymization efforts, risks of data breaches or misuse persist, raising concerns about surveillance and stigmatization (Whittaker et al., 2019). Public confidence in AI will depend not only on technological transparency but also on developers' ability to demonstrate responsible stewardship of user data.

Another layer of complexity arises from algorithmic bias. AI systems trained on non-representative data may yield inaccurate or discriminatory outcomes, particularly for marginalized communities. For instance, racial or gender disparities in data sources can result in AI tools that systematically underdiagnose or overlook key symptoms in certain groups (Vokinger et al., 2021). This underscores the importance of inclusive design and interdisciplinary oversight in developing equitable mental health technologies.

Finally, while AI may offer cost-effective solutions to global mental health shortages, its success hinges on integration into broader systems of care. This includes aligning AI outputs with clinical pathways, training professionals to interpret algorithmic recommendations, and ensuring interoperability with existing health infrastructure. The role of blended care models—which combine digital tools with traditional therapy—appears particularly promising, offering scalability without sacrificing human support (Torous et al., 2020).

Conclusions

As mental health challenges continue to rise globally, artificial intelligence offers a powerful, though still evolving, set of tools for improving mental health care delivery. From conversational agents to predictive analytics, AI has the capacity to address long-standing gaps in access, personalization, and early intervention. However, its deployment must be carefully navigated, keeping in mind the ethical, cultural, and practical complexities inherent in mental health care.

AI's role should not be viewed as a replacement for human therapists or clinical judgment but rather as a complement that can enhance the scalability, responsiveness, and consistency of support. The future of mental health care lies in hybrid models where human empathy is augmented—not substituted—by technological efficiency. Realizing this vision will require rigorous research, inclusive algorithm design, transparent policy frameworks, and sustained collaboration between clinicians, technologists, and communities.

Ultimately, to ensure that AI truly contributes to health equity, its development and implementation must center the dignity, safety, and autonomy of all individuals—particularly the most vulnerable. Responsible innovation in this space holds transformative potential, but only if guided by a commitment to ethical practice, social inclusion, and continuous evaluation.

Disclosure

Authors do not report any disclosures.

Authors' contributions

Conceptualization: P. Bala, A. Rasińska;
Methodology: P. Rzyczniok;
Software: n/a; check: A. Rasińska;
Formal analysis: M. Matusik;
Investigation: A. Rasińska;
Resources: J. Jachimczak, M. Kopczyński;
Data curation: M. Kopczyński, J. Jachimczak;
Writing -rough preparation: A. Rasińska, P. Bala;
Writing -review and editing: A. Rzyczniok;
Visualization, P. Rzyczniok;
Supervision: A. Rasińska, M. Matusik;
Project administration: P. Rzyczniok;
Receiving funding: n/a.

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

Funding statement: This research received no external funding.
Institutional Review: Board Statement Not applicable.
Informed Consent Statement: Not applicable.
Data availability statement: Not applicable.
Acknowledgments: The authors declare that there are no acknowledgments for this study.
Conflict of Interest Statement: The authors declare no conflict of interest.

REFERENCES

1. Barnett, I., Torous, J., Staples, P., Sandoval, L., Keshavan, M., & Onnela, J. P. (2018). Relapse prediction in schizophrenia through digital phenotyping: A pilot study. *Neuropsychopharmacology, 43*(8), 1660–1666. https://doi.org/10.1038/s41386-018-0030-z
2. Blease, C., Kaptchuk, T. J., Bernstein, M. H., Mandl, K. D., Halamka, J. D., & DesRoches, C. M. (2019). Artificial intelligence and the future of psychiatry: Insights from a global physician survey. *NPJ Digital Medicine, 2*, 1–6. https://doi.org/10.1038/s41746-019-0182-7
3. Calvo, R. A., Milne, D. N., Hussain, M. S., & Christensen, H. (2017). Natural language processing in mental health applications using non-clinical texts. *Natural Language Engineering, 23*(5), 649–685. https://doi.org/10.1017/S1351324916000383
4. Chekroud, A. M., Zotti, R. J., Shehzad, Z., et al. (2016). Cross-trial prediction of treatment outcome in depression: A machine learning approach. *The Lancet Psychiatry, 3*(3), 243–250. https://doi.org/10.1016/S2215-0366(15)00471-X
5. Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2018). A review of depression and suicide risk assessment using speech analysis. *Speech Communication, 71*, 10–49. https://doi.org/10.1016/j.specom.2015.03.004
6. Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health, 4*(2), e19. https://doi.org/10.2196/mental.7785
7. Gunning, D., & Aha, D. W. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine, 40*(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850
8. He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine, 25*(1), 30–36. https://doi.org/10.1038/s41591-018-0307-0
9. Huckvale, K., Torous, J., & Larsen, M. E. (2019). Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. *JAMA Network Open, 2*(4), e192542. https://doi.org/10.1001/jamanetworkopen.2019.2542
10. Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation. *JMIR mHealth and uHealth, 6*(11), e12106. https://doi.org/10.2196/12106
11. Inkster, B., Stillwell, D., Kosinski, M., & Jones, P. B. (2018). A decade into Facebook: Where is psychiatry in the digital age? *The Lancet Psychiatry, 5*(11), 900–902. https://doi.org/10.1016/S2215-0366(18)30358-2
12. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101
13. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2
14. Luxton, D. D. (2014). Recommendations for the ethical use and design of artificial intelligent care providers. *Artificial Intelligence in Medicine, 62*(1), 1–10. https://doi.org/10.1016/j.artmed.2014.06.004
15. Martinez-Martin, N., Wieten, S., Magnus, D., & Cho, M. K. (2018). "Data mining for health": Ethical responsibilities for clinical practice and research. *The Hastings Center Report, 48*(4), 22–31. https://doi.org/10.1002/hast.875
16. Miner, A. S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C., & Linos, E. (2016). Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Internal Medicine, 176*(5), 619–625. https://doi.org/10.1001/jamainternmed.2016.0400
17. Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology, 13*, 23–47. https://doi.org/10.1146/annurev-clinpsy-032816-044949
18. Morley, J., Machado, C. C., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *Social Science & Medicine, 260*, 113172. https://doi.org/10.1016/j.socscimed.2020.113172
19. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342
20. Pfefferbaum, B., & North, C. S. (2020). Mental health and the COVID-19 pandemic. *The New England Journal of Medicine, 383*(6), 510–512. https://doi.org/10.1056/NEJMp2008017
21. Price, M., Yuen, E. K., Goetter, E. M., Herbert, J. D., Forman, E. M., Acierno, R., & Ruggiero, K. J. (2019). mHealth: A mechanism to deliver more accessible, more effective mental health care. *Clinical Psychology & Psychotherapy, 26*(3), 232–240. https://doi.org/10.1002/cpp.2330
22. Samek, W., Wiegand, T., & Müller, K.-R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv preprint*. https://arxiv.org/abs/1708.08296

23. Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine, 49*(9), 1426–1448. https://doi.org/10.1017/S0033291719000151
24. Torous, J., Larsen, M. E., Depp, C., Cosco, T. D., Barnett, I., Nock, M. K., & Firth, J. (2021). Smartphones, sensors, and machine learning to advance real-time prediction and interventions for suicide prevention: A review of current progress and next steps. *Current Psychiatry Reports, 23*(7), 51. https://doi.org/10.1007/s11920-021-01245-2
25. Torous, J., Lipschitz, J., Ng, M., & Firth, J. (2020). Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. *Journal of Affective Disorders, 263*, 413–419. https://doi.org/10.1016/j.jad.2019.11.167
26. Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.
27. Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLoS Medicine, 15*(11), e1002689. https://doi.org/10.1371/journal.pmed.1002689
28. Vokinger, K. N., Feuerriegel, S., & Kesselheim, A. S. (2021). Mitigating bias in machine learning for medicine. *Communications Medicine, 1*(1), 1–3. https://doi.org/10.1038/s43856-021-00028-4
29. Whittaker, R., McRobbie, H., Bullen, C., Rodgers, A., & Gu, Y. (2019). Mobile phone-based interventions for smoking cessation. *Cochrane Database of Systematic Reviews*, (10), CD006611. https://doi.org/10.1002/14651858.CD006611.pub5
30. World Health Organization. (2021). *Mental health atlas 2020*. https://www.who.int/publications/i/item/9789240036703
31. World Health Organization. (2022). *Mental health*. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response