SSM - Mental Health 6 (2024) 100373

Contents lists available at ScienceDirect

SSM - Mental Health

journal homepage: www.journals.elsevier.com/ssm-mental-health

Review

Exploring the societal implications of digital mental health technologies: A critical review

Olivia A. Stein *, Audrey Prost

Institute for Global Health, University College London, London, United Kingdom

ARTICLE INFO

Handling Editor: E Mendenhall

ABSTRACT

Introduction: Digital mental health technologies are services that rely significantly on big data and artificial intelligence and are widely championed as possible solutions to global mental healthcare shortages. Services include prediction and detection of symptoms, personalized treatment, chatbot therapy, and both personal and population-level monitoring. Existing research has focused on describing the functionality, acceptability, and efficacy of these technologies, as well as data governance challenges. This critical review explores the societal implications of digital mental health technologies in terms of its impacts on mental healthcare, population-based monitoring of mental health, and commodification of mental health data.

Methods: Searched six databases for literature on digital mental health technologies published between 2014 and 2023 following PRISMA-ScR. Conducted qualitative data analysis of 53 records using the Framework method, bringing into conversation wider literature on mental healthcare, ethics, health equity, and data capitalism.

Results: The literature on digital mental health technologies highlights three main areas of ethical concern. First, these technologies could affect treatment and management through changes in accessibility, quality and resource availability of mental healthcare in either positive or negative ways, depending on linkages with clinical services. In addition, these technologies may have ramifications due to the objectification or dehumanization of mental healthcare, the medicalization of poor mental health, and the prominence of self-management. Second, the implications of novel clinical and population-based monitoring are explored, including algorithm-triggered mental health interventions and surveillance. Third, the literature brings forth reservations about the commodification of mental health data through the practice of data capitalism.

Conclusion: This critical review suggests an urgent need for comprehensive regulation of digital mental health technologies and scholarly collaboration to curb adverse effects on mental healthcare systems and society, while remaining optimistic regarding the potential benefits of these services if implemented in collaboration with clinicians and communities who experience mental illness.

1. Introduction

Over the past few decades, reported rates of mental illness have risen dramatically worldwide, with the global number of disability-adjusted life years attributed to mental disorders increasing by 55% between 1990 and 2019 (Ferrari et al., 2022). This alarming change in burden of disease denotes a worldwide crisis in mental wellbeing, one that health systems are inadequately equipped to address due to widespread lack of funding for and access to mental healthcare, particularly in low-income countries (World Health Organization, 2022). Semi- and fully-autonomous digital mental health technologies have emerged in recent years with a strong emphasis on big data and artificial intelligence (AI)-related services. Such services are often championed as potential solutions to mental healthcare accessibility and quality issues (Ahmed et al., 2022; D'Alfonso, 2020). Prominent services include wellbeing or mental illness specific mobile apps, chatbots and mental health monitoring or risk prediction algorithms (D'Alfonso, 2020; Hahn et al., 2017). These technologies serve a range of different purposes which encompasses detection of mental health symptoms or self-harm risks, diagnostics, personalized treatment including therapy, as well as self, clinical or population-based monitoring (Rocheteau, 2023; Shatte et al., 2019). In this review, we use the term "digital mental health technologies" or simply "technologies" or "services" to refer to this diverse domain.

* Corresponding author. Institute for Global Health, University College London, 30 Guilford Street, London, WC1N 1EH, United Kingdom.
E-mail address: olivia.stein.22@alumni.ucl.ac.uk (O.A. Stein).

https://doi.org/10.1016/j.ssmmh.2024.100373
Received 7 August 2024; Received in revised form 12 November 2024; Accepted 15 November 2024
Available online 16 November 2024

Digital mental health technologies conduct active and/or passive data collection and analysis to inform their operations (Onnela, 2021). Active data consist of information intentionally entered by users for mental health purposes (e.g. wellbeing reports documented in a designated app), whereas passive data are information collected continuously through devices without a user's explicit input and without a specific mental health objective (e.g. general mobile phone, internet and social media usage, voice/keystroke characteristics, sleep patterns, location information) (D'Alfonso, 2020; Hindley et al., 2022). The aggregation of such data can form a "digital phenotype" that can be used to infer new health information such as an illness risk prediction, diagnosis, or personalized treatment (Liang et al., 2019). These tools can be designed to be used clinically with support from mental health professionals, autonomously by individuals, or for public health aims (Torous and Haim, 2018). Public and private institutions ranging from small businesses and universities to large internet companies such as Facebook or Google are actively involved in developing digital mental health technologies (Gooding, 2019). The age of artificial intelligence and big data therapeutics will undoubtedly alter mental healthcare globally and may ultimately impact societal conceptions of illness and wellbeing, as well as the prevalence and distribution of inequity (Moncrieff, 2022; Rubeis, 2023).

Existing reviews on digital mental health technologies comprehensively describe their functionality and user acceptability (Ahmed et al., 2022; Milne-Ives et al., 2022; Sequeira et al., 2020; Shatte et al., 2019). Efficacy issues have also been extensively addressed, with studies highlighting the lack of empirical evidence and equity concerns regarding AI objectivity, algorithm bias, the opaqueness of algorithmic inferences and issues of false positives (Birk and Samuel, 2022; Dwyer and Koutsouleris, 2022; Lagen et al. 2020; Malhotra and Jindal, 2022). Data-related ethical concerns have been widely raised in relation to privacy breaches, protection of personal mental health data and lack of regulation in the field globally (Lagen et al. 2020; Liu et al., 2022; Terra et al., 2023). Due to the thorough coverage of these issues, this review does not contend with these topics but instead focuses on outlining the diverse possible societal implications of these technologies, including their potential impacts on mental healthcare and the effects of population-based monitoring and commodification of mental health data.

This critical review uses both a database search of the literature on digital mental health technologies and wider critical sources to answer the question: what are the potential societal implications of digital mental health technologies? Following an overview of methodology, this review discusses three key themes identified in the literature: (1) changes to mental healthcare; (2) the effect of population-based monitoring including algorithm-triggered interventions and surveillance; and (3) commodification. This discussion is followed by a summary of the societal implications of these emerging services through critical literature beyond that identified through the literature search. In particular, we address ethical concerns regarding prominent framings of mental disorders promoted by these technologies and discuss the wider social influences that may determine effects on patient outcomes.

2. Methods

This work is a critical review, which is a form of review that "aim[s] to critically analyze the extant literature on a broad topic to reveal weaknesses, contradictions, controversies, or inconsistencies" (Paré et al., 2015, p.189). The methodology follows the PRISMA-ScR checklist (Tricco et al., 2018) as well as the Framework method for analyzing qualitative data (Ritchie and Spencer, 1994) and was conducted by the lead author OS under supervision of AP. This first encompassed initial reading, identifying keywords, choosing databases and running a search. Subsequently, OS followed a four-step process that consisted of combining the search results, removing duplicates, screening titles/abstracts and assessing full texts for inclusion. This process, the specification of the research question and the subsequent qualitative analysis of the literature included in the review are described in detail below.

2.1. Keywords

To identify keywords, OS reviewed a survey of literature relevant to digital mental health technology and its ethical implications derived from Google Scholar and extracted key terms that were categorized into "mental health", "technology" or "concept". The final iteration of the search terms includes the four conceptual categories of "mental health", "function", "technology" and "concept" and are presented in Table 2 in Appendix A, along with an example of an exact database search in Table 3.

2.2. Search strategy

OS searched six databases to cover multiple relevant subject areas. The biomedical and life science resources included to source mental illness publications are PubMed, Web of Science (contains Core Collection, Biosis & Medline) and Embase 1974–2023 (OVID version). Global Health 1973-Present (OVID version) was added to provide a public health-specific database and ProQuest Social Science to incorporate anthropology and bioethics perspectives. Scopus was then included to ensure a large multi-disciplinary repository for any remaining resources. Database searches were conducted in July 2023 and yielded a total of 671 results. Table 4 in Appendix A provides an overview of the initial search results. Finally, four supplementary sources obtained through OS's initial readings were added, bringing the total number of records to 675.

2.3. Refining search results

OS scanned the 675 results for duplicates and removed 272. The 403 remaining records were assessed for inclusion or exclusion based on title/abstract, following the first-round criteria described in Table 5 in Appendix A. The inclusion criteria specified pertaining to psychotic or emotional disorders and a mental health technology intervention. The outcome of this step was the exclusion of 245 records and 158 remaining for full-text review.

At this time, 26 reviews were identified within the 158 remaining results, then screened and coded into seven observed themes (see Table 6 in Appendix A). This inquiry offered an overview of the existing research in this field. We noted that common topics were applications of digital technologies, research on effectiveness, user acceptability and data-related ethical concerns. This investigation highlighted a shortage of reviews addressing the potential impact of these technologies on existing mental healthcare systems, the role of private companies in this field, and concerns regarding patient agency and the framing of mental health. These findings informed the final inclusion/exclusion criteria.

Using the second-round inclusion and exclusion criteria presented in Table 5 in Appendix A, OS conducted a full-text review of the 158 sources and excluded 105 records, leaving 53 remaining for inclusion. The inclusion criteria comprised relating to the mental healthcare system, follow-up/action generated from mental health technology, commercial aspects of mental health technology, self-management of mental health or non-data related ethical concerns (i.e. consent, coercion, surveillance, human rights). Fig. 1 offers a flow diagram describing the search screening process and results. A brief summary of the content and relevance of each record selected as well as characteristics, including subject of interest, institutional affiliation location, article type and overall tone, can be found in Table 7, Table 9 and Table 10 in the appendices.

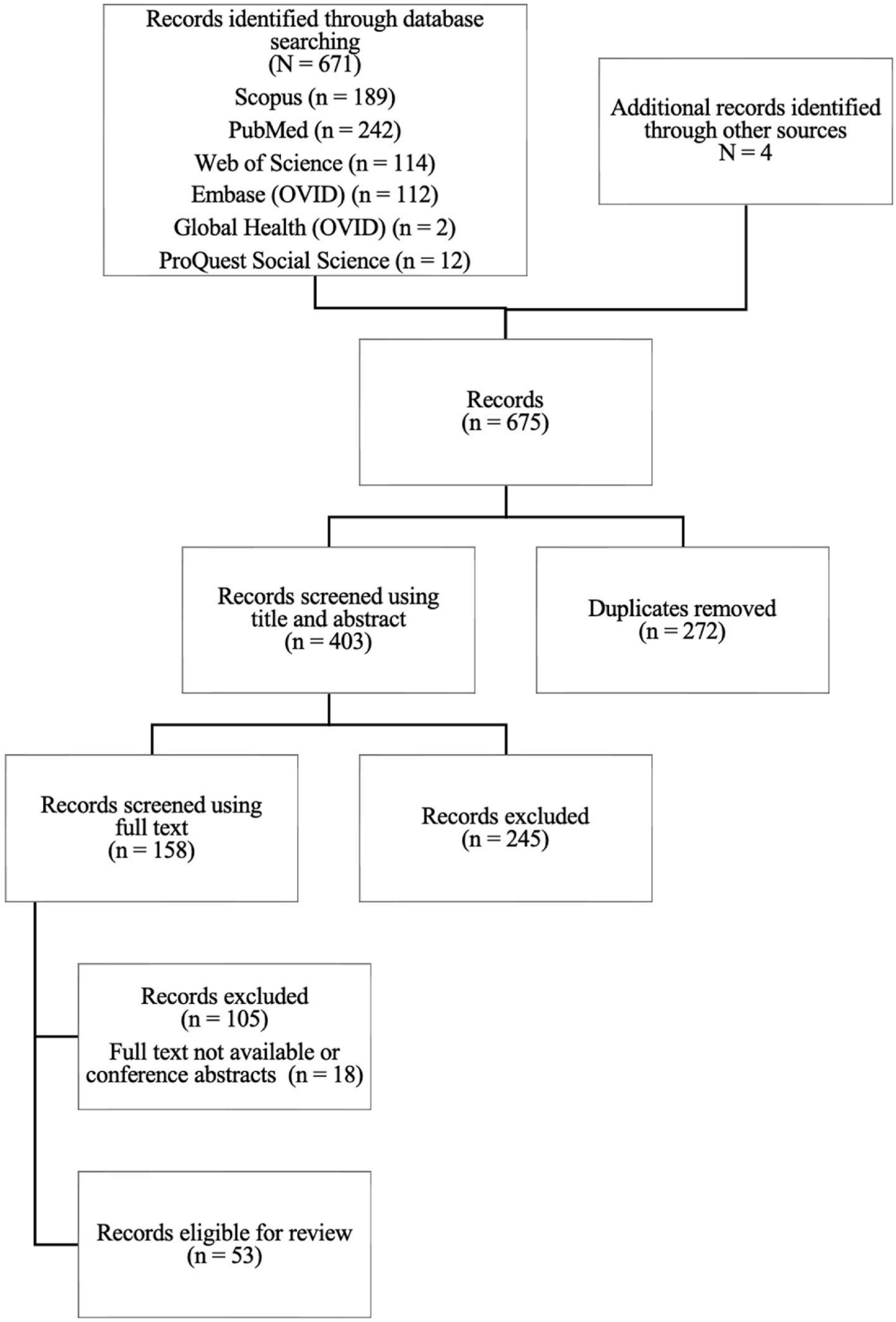


Fig. 1. Flow diagram of search screening process and results.

2.4. Qualitative data analysis

The Framework method (Ritchie and Spencer, 1994) was followed by OS to independently chart the data. Based on the research question and familiarization with the literature, OS first identified the following five major themes to constitute the thematic framework: mental healthcare system, monitoring, self-management, commodification and ethics. To analyze the records included in this review, OS thoroughly read through each source, pulled out significant text portions into a separate document and indexed the major theme(s) and sub-categories. After all sources had been coded, 5 documents were created to separate the data by major theme. Each major theme document underwent charting whereby the textual data was carefully synthesized into a new file that delineated, under each sub-theme and further intricate sub-headings, the ideas mentioned along with a citation of the corresponding source author(s) and page number(s). Once completed, OS created a list of all the sub-categories mentioned in the records, which are detailed in Table 1 below. These sub-categories were then used to produce a mind map (see Fig. 2 below), which informed the approach to analyze the key themes elucidated from the literature and describe their interrelationships, as well as served to identify the topics of mental health framing and ethical considerations to be addressed in the discussion. A summary of the content and coding of each record is presented in Tables 7 in Appendix A, along with the frequency each theme and sub-theme in Table 8. The most common coded major theme was mental healthcare system at 92.5% of records, followed by ethics at 67.9%, monitoring at 60.4%, commodification at 43.4% and finally self-management at 37.7%.

Table 1
Thematic coding themes and sub-categories.

Major Coding Theme	Sub-Categories
Mental Healthcare System	Accessibility; Quality of care; Undermines/supports healthcare system; Dehumanizing/objectifying care; Therapeutic relationship; Overmedicalization; Clinical tool
Self-Management	Beneficial; Harmful
Monitoring	Lack/vague intervention; Soft intervention; Hard intervention; Surveillance
Commodification	Profit incentive; Transparency; Gamification; Data capitalism
Ethics	Agency; Consent; Coercion; Inequity/equity; Framing of mental illness

2.5. Positionality

Critical reviews are influenced by an elevated degree of subjectivity (Paré et al., 2015). Both the lead author (OS) and second author (AP) acknowledge their privilege and position as educated White women. Both have experienced mental illnesses, and OS further shares a personal diagnosis of neurodiversity as well as experiences with loved ones with severe mental illness. Despite an active practice in reflexivity, our views on mental health are deeply influenced by our lived experiences and a Western conception of mental disorders.

3. Results

3.1. Implications for the treatment and management of mental health conditions

3.1.1. Accessibility

Proponents of digital mental health technologies emphasize their potential to increase the accessibility of mental healthcare in numerous ways. Since these services are anticipated to partially or entirely complete the tasks of mental healthcare providers, their presence is predicted to increase the availability of care (Balcombe and De Leo, 2021; Stern et al., 2022). Treatment or monitoring by technology is not limited by time or location constraints, which is a major barrier for people living in resource-poor settings or for whom conventional care structures are unsuitable (Fiske et al., 2019; Wasil et al., 2022). Digital technologies overcome financial barriers to access for mental healthcare by either offering free access or paid usage at fraction of the cost of professional care (Koh et al., 2022; Kolenik and Gams, 2021). Yet the extent of this affordability will depend on the ability of technologies to replace a patient's need for a clinician or therapist (Eagle et al., 2022; Uma et al., 2023), and the true costs of such services may be personal data (Baumgartner, 2021). Access may additionally be increased indirectly through greater acceptability of technological mental healthcare to users, predominantly due to discretion (ibid).

Opponents point out that accessibility is limited to digital accessibility. The "digital divide" marks the phenomenon of partial or complete lack of access to technology or internet for many people due to financial, regional or generational limitations (Rickard et al., 2022; Wies et al., 2021). Many individuals otherwise have personal aversions to digital mental healthcare either due to unfamiliarity or to security reasons (Bauer et al., 2020; Omarov et al., 2023). Doubt has moreover been expressed about whether digital healthcare can achieve its promises of equity or will instead become a premium product (Feldman and Perret, 2023).

3.1.2. Quality of care

In addition to providing *more* mental healthcare, advocates point out that digital mental health technology may provide *better* mental healthcare through novel treatment and monitoring features. Firstly, digital phenotyping technologies passively track mental illness progression through mobile and internet user data and claim to be able to detect early symptoms and accordingly trigger preventative interventions (Rosenfeld et al., 2021; Uusitalo et al., 2021). These services

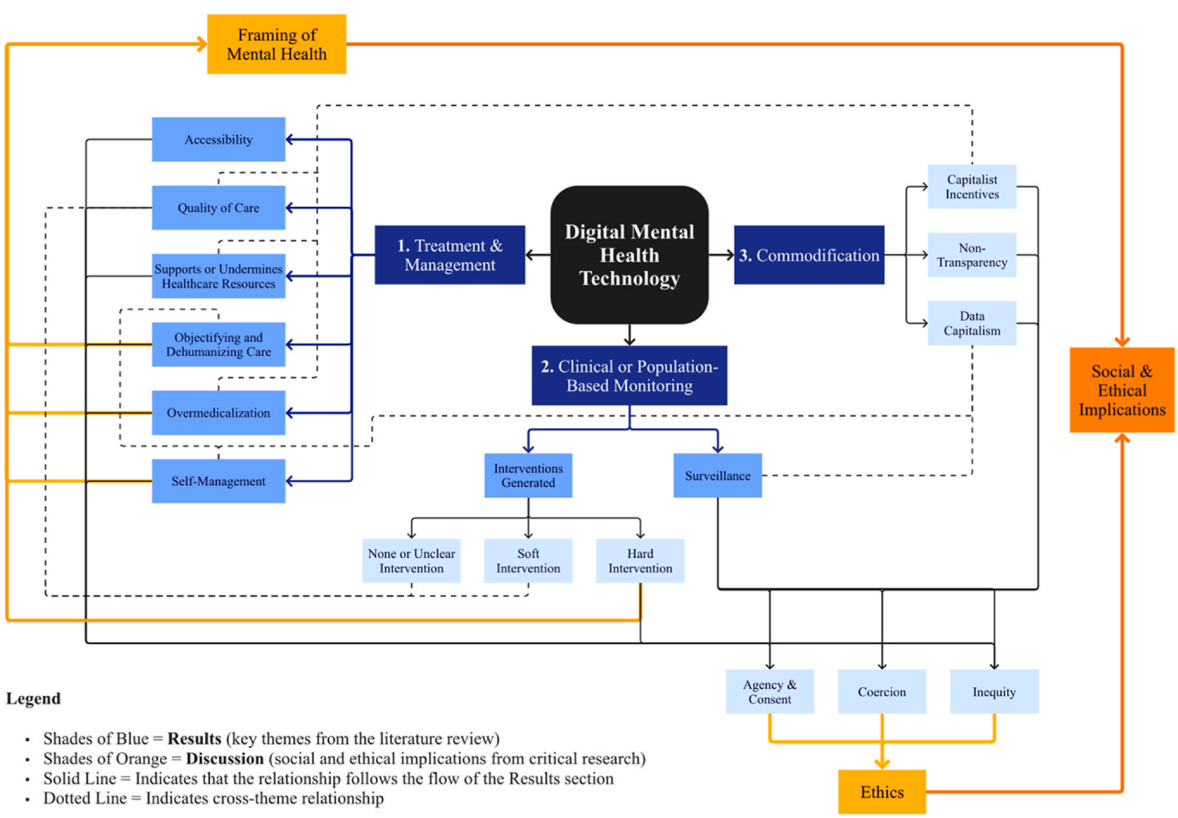


Fig. 2. Thematic mind map.

could potentially be used to provide personalized care by predicting what treatment will be most successful to a given individual and by monitoring their ongoing health or pharmaceutical adherence (Feldman and Perret, 2023; Graham et al., 2019; Woodward et al., 2022). Second, mental health apps can provide psychoeducation, therapeutic exercises, guided meditations, journaling functions, goal setting, daily planning support, reminders and AI chatbot therapy (Berry and Lai, 2014; Denecke et al., 2022; Rauseo-Ricupero et al., 2021; Wasil et al., 2022). Third, many technologies collect and aggregate the active data inputted by users regarding their symptoms to provide perspective on mental states over time (Denecke et al., 2021). Fourth, digital mental healthcare, especially chatbots, is unique because it provides a human-free form of care that is seen as completely non-judgemental and anonymous, which increases comfort for patients who experience external or self-stigma regarding seeking mental healthcare (Koh et al., 2022; Stern et al., 2022). Finally, AI systems can also continuously improve due to a feedback loop and allow for the standardization of the quality of therapy (Rosenfeld et al., 2021).

Despite these potential benefits, critics argue that most digital mental health technologies are not yet clinically validated and tend to provide vague and substandard services that do not take into account the complexity of mental disorders (Abdelrahman, 2023; Balcombe and De Leo, 2022; Rubeis, 2022). This runs the risk of exposing people to inappropriate care, potentially leading to negative patient outcomes (Williams and Pykett, 2022). Moreover, the trend of limited use and poor retention of users for mental health apps speaks to substandard quality (Bauer et al., 2020; Koh et al., 2022).

3.1.3. *Effects on resources for mental healthcare*

The impending adoption of data-driven technology will undoubtedly affect the provision of professional mental healthcare. Some scholars argue that such implementation will increase efficiency, reduce healthcare costs and relieve clinicians of time-consuming administrative tasks, thus allowing more time to be devoted to patients (Bidargaddi et al., 2021; Oakey-Neate et al., 2020; Yu et al., 2022). System stress could be relieved if technological interventions could provide treatment for non-severe mental health cases in order to allow clinicians more time to attend to severe cases that require higher expertise (Fiske et al., 2019).

On the other hand, observers fear that the advent of these services will divert resources away from existing mental healthcare systems (Berry and Lai, 2014; Starke et al., 2021). Should this happen, it would become increasingly difficult to see a human mental healthcare provider, and it is feared that disparities in seeking care based on socio-economic status would widen (Omarov et al., 2023). Nonetheless, proponents point out that having access to some care is better than having no access at all (Fiske et al., 2019). As many mental healthcare technologies are delivered by private companies, undermining human mental healthcare resources may furthermore create dependence on these for-profit businesses and potentially compromise quality, access and patient trust (Gültekin, 2022; Tekin, 2021).

3.1.4. *Objectifying and dehumanizing mental healthcare*

Several scholars claim that AI systems will become experts in mental illness and be able to conduct complex decision making and patient diagnosis (Atlam et al., 2022; Luxton, 2014; Starke et al., 2021). Digital phenotyping is championed to be able to re-define mental illness diagnostic categories and reveal risk factors in an objective manner (Van Assche et al., 2022). The argument contends that algorithms transcend clinicians' skillsets because they are unbiased, and that data provides a more reliable account than patient testimony of mental wellbeing, which may be impacted by warped perception, recall error, or social desirability bias (Luxton et al., 2016; Rosenfeld et al., 2021; Schmidt and D'Alfonso, 2023). Through this practice, AI algorithms define thresholds of normal and abnormal mental health (Williams and Pykett, 2022). Conversely, many call into question the ability of AI to determine objective and clinically relevant diagnostic categories due to the complexity of mental disorders and algorithmic bias (Tekin, 2021; Uusitalo et al., 2021). Ethicists caution that objectifying mental healthcare may in fact dehumanize care by undermining patient testimony of their own experiences and needs (Ma et al., 2023). Requiring outside validation of suffering can arguably impede human agency and be an abuse of epistemic power (McCradden et al., 2022). This practice would likely be aggravated for patients from vulnerable groups who already experience systemic delegitimization and coercion at the hands of mental healthcare personnel and may further erode trust towards seeking support (Carr, 2020; McCradden et al., 2022).

Mental health professionals often stress that the patient-clinician therapeutic relationship is central to providing effective care (Rosenfeld et al., 2021; Rubeis, 2022). Various researchers have found it is possible for AI chatbots to create analogous empathetic therapeutic relationship with patients, especially if individuals are more comfortable confiding in a non-human as discussed above (Gültekin, 2022; Omarov et al., 2023). Nonetheless, worry is expressed that developing such connections with technological systems will lead to overreliance, addiction, erosion of human relationships and possibly manipulation (Denecke et al., 2021; Fiske et al., 2019). Skeptics question the depth of these therapeutic relationships and argue that AI is incapable of replacing human empathy, wisdom and connection (Tekin, 2021; Uusitalo et al., 2021). Others add that removing humans from mental healthcare may lead to people feeling abandoned and to the dehumanization of patients (Abdelrahman, 2023; Baumgartner, 2021).

3.1.5. *Overmedicalization*

Mental healthcare apps commonly suggest that poor mental health is ubiquitous (Parker et al., 2018). This mindset, along with the development of granular AI diagnostic algorithms, incites alarm among scholars of overmedicalization (Bauer et al., 2020). Difficult emotions, such as responses to stress or grief, may become more easily attributed to mental disorders and therefore deemed unnatural (Abdelrahman, 2023; Cosgrove et al., 2020b). Such services could furthermore promote a singular pathology of mental illness that does not account for cultural variation or diversity in experiences (Tekin, 2021). While many people experience poor mental health and need support, some scholars argue that the framing of all instances of poor mental health as chronic neurophysiological disability that requires prescription medication is deeply problematic (Parker et al., 2018). Evidence suggests that unnecessary diagnosis and treatment can promote poor mental health symptoms due to the looping effect and create undue stress on healthcare systems (Cosgrove et al., 2020b; Uusitalo et al., 2021).

Authors repeatedly point out that digital phenotyping and AI initiatives to find "objective" diagnostic categories typically do not account for situational and socio-economic factors of poor mental health such as discrimination, poverty and abuse (Baumgartner, 2021; Cosgrove et al., 2020a; Tekin, 2021). Therefore those who experience oppression would be most likely to be misdiagnosed by algorithms and possibly coerced into treatment, following the history of injustice in mental healthcare (Cosgrove et al., 2020a). If mental illness is seen as purely biological in origin, individuals may be portrayed as genetically responsible for their mental health outcomes (Uusitalo et al., 2021). The onus would then be removed from governments to dismantle systemic forms of oppression and establish supportive social systems (Ma et al., 2023).

3.1.6. *Self-management*

A major facet of digital mental health technologies is their autonomous nature. Self-management encompasses individuals independently using resources to treat and monitor their mental health symptoms (Rubeis, 2022). Many authors stress that these technologies allow individuals to live more effectively with their condition(s) (Denecke et al., 2022; Kolenik and Gams, 2021; Wies et al., 2021). This is achieved through tracking symptoms over time to discern potential triggers, using educational resources to gain understanding of one's experiences and exercises to manage symptoms as well as promoting accountable behaviour change (ibid). Patients may furthermore gain a sense of control, autonomy and empowerment in the face of chronic mental illness (Rubeis, 2022). Self-management is highlighted as a common goal of therapy practices such as cognitive behavioural therapy (Berry and Lai, 2014). Many applications furthermore intimate that mental illness is undemanding to manage with their services (Parker et al., 2018; Radovic et al., 2016).

Concerned parties highlight that it may be particularly difficult for some people with mental disorders to manage their health alone, such as in instances where disease symptoms impair judgement and/or executive functioning (Rubeis, 2022). Certain individuals may furthermore be unaware of their mental health needs and would not obtain any relief without clinical guidance (Tekin, 2021). A dominant narrative that goes hand in hand with self-management technology is that individuals are morally responsible for their mental health (Abdelrahman, 2023; Parker et al., 2018). Those with poor mental health may be seen as failing to manage their illness and as economically burdensome to society (Abdelrahman, 2023). This line of thinking risks unjustly ignoring the difficult reality of living with mental illness and the structural causes of poor mental health, and may consequentially perpetuate mental illness stigma (Williams and Pykett, 2022).

3.2. Implications of clinical and population-based monitoring

Beyond self-monitoring, advocates emphasize the benefits of technology for monitoring patients and the general public for mental health concerns. Clinical monitoring via mental health technology may consist of a mental health professional keeping track of symptom data inputted by patients into a specific server, passive digital phenotype data shared by patients or in the most extreme example by tracking psychiatric medication adherence through indigestible markers attached to pharmaceuticals (Cosgrove et al., 2020a; Schmidt and D'Alfonso, 2023). Population-based monitoring may be conducted through mental health apps, social media networks, internet usage or through passive data collected from mobile biometric or utilization tracking systems (D'Hotman and Loh, 2020; Nebeker et al., 2019). AI developers believe that algorithms can predict changes in mental health, suicide risk or risk of violence from a variety of data ranging from social media posts and google searches to typing dynamics and voice characteristics from mobile phones and in-home assistants, a claim that is heavily debated (Feldman and Perret, 2023; Ma et al., 2023; Neuman, 2016). What then are the possible consequences of this monitoring, and what are the ethical responsibilities associated with tracking mental health risk information?

3.2.1. Interventions generated from monitoring algorithms

Considering the possibility of mental health monitoring systems prospering, it is important to ask what tangible action would be taken in response to an alert generated by an algorithm. Currently, developers have largely left this question unanswered. Many services state an inability to respond to a mental health crisis due to limited functionality or predefined content (Denecke et al., 2022; Parrish et al., 2022), while others mention real-time support without fleshed-out explanations (Atlam et al., 2022; Robinson et al., 2020; Van Assche et al., 2022).

When technologies address this problem, most opt to provide soft interventions. A soft action from a non-clinical system includes providing individuals with automated messages, grounding exercises, local crisis hotlines, online counselling or other resources (D'Hotman and Loh, 2020; Wasil et al., 2022). Key to these techniques is that they are non-invasive and thus can either be accepted or rejected at an individual's discretion (Feldman and Perret, 2023). One problem is that these techniques may not be particularly effective, especially if someone is in imminent danger (D'Hotman et al., 2021). In the case of a clinical monitoring system, a possible strategy is for a doctor or case manager to reach out to a patient directly to assess their wellbeing (Betthauser et al., 2020; Bidargaddi et al., 2021). Another hypothetical scenario discussed in the literature is for individuals to choose a pre-identified emergency contact or guardian whom they would want notified if the algorithm detected a significant risk (Neuman, 2016; Poulin et al., 2016).

In some cases, digital mental health technologies have triggered hard/coercive interventions. In the United States for example, Facebook can and has directed emergency services to people's homes when an acute suicide risk is detected (D'Hotman and Loh, 2020). This form of action is controversial and viewed as a profound violation of individual privacy and agency (Cosgrove et al., 2020a). There also remains a risk of false positives whereby someone may be mistakenly identified as suicidal and could conceivably have their narrative ignored and be forcibly detained or treated (D'Hotman et al., 2021; McCradden et al., 2022).

3.2.2. Surveillance

Scholars recurrently voice unease regarding mental health monitoring being akin to mass surveillance, a practice which infringes upon autonomy and denies people privacy in their own minds (Abdelrahman, 2023; Rubeis, 2022). Data gathered may be perversely used to control, manipulate and/or discriminate against people for social, commercial or political ends (Baumgartner, 2021; Cosgrove et al., 2020a; Feldman and Perret, 2023). Categorizing people based on perceived mental health could proliferate sigma and conceivably lead to profiling for insurance, employment and crime (Cosgrove et al., 2020a; Ma et al., 2023). This practice can be defined as disability surveillance, which has problematic historical links to ableism and eugenics (Ma et al., 2023).

3.2.3. Consent, paternalism and responsibility

Population-based monitoring brings about ethical concerns related to consent and paternalism, particularly when considering the potential use for surveillance or to enact hard interventions. As it stands, private companies do not require true informed consent to collect and process peoples' data, beyond perhaps agreeing to dense terms and services, and often do not allow opting-out (Eagle et al., 2022). Therefore most individuals are unaware to what extent their data is being collected and the ways in which their information is being used (Uusitalo et al., 2021). Insofar as people have not explicitly consented to monitoring, ethicists argue that their rights to agency and privacy are being violated (Cosgrove et al., 2020a). Others contend that the benefit of saving people from harm outweighs the invasion of their liberties (Neuman, 2016; Schmidt and D'Alfonso, 2023). One author adds that comprehensive surveillance is necessary to prevent instances of public violence (Neuman, 2016). Conversely, scholars warn against abuse of power and justifying the problematic tradition of the law ignoring the rights of mentally ill people, particularly those from disadvantaged backgrounds, on the assumption they "lack capacity" (Carr, 2020; Cosgrove et al., 2020a). However, an opt-out system would foreseeably exclude someone that would benefit from a life-saving mental health intervention (D'Hotman et al., 2021). This illustrates the common moral challenge of balancing respect for autonomy and benevolence. Given that mental health risk monitoring algorithms exist and are being used, a question remains unanswered as to the ethical responsibility intertwined with the possession of such information (Williams and Pykett, 2022).

Of note, the discussion of ethical obstacles within the literature surveyed for this critical review shows stark biases alongside industry lines, with high techno-optimism and justification of rights infringements on one side and pessimism and heavy critique on the other (Abdelrahman, 2023; Omarov et al., 2023). This indicates deficient dialogue and collaboration between the academic and technical fields.

3.3. Implications of commodification of mental health data

Numerous scholars discuss digital mental health technologies to be used in a clinical setting with doctors' and patients' involvement (Graham et al., 2019; Rauseo-Ricupero et al., 2021). In this case, technologies would mainly be used as an adjunctive treatment (Stern et al., 2022). This type of implementation would follow local medical regulatory guidelines and patient ethical codes, such as informed consent and respect for autonomy (Schmidt and D'Alfonso, 2023). Unfortunately, adoption of digital mental health services in this area is likely be slow and limited due to insufficient empirical evidence regarding safety and efficacy, and the substantial cost of uptake and continued use alongside clinicians, which seem impossible for already strained health systems to implement (Nogueira-Leite et al., 2023; Williams and Pykett, 2022).

Simultaneously, a large number of digital mental health technologies are being developed by private companies to be used independently of health professionals (Williams and Pykett, 2022). Private implementation is considered more feasible short-term due to lower cost and current weak regulatory barriers (Nebeker et al., 2019; Woodward et al., 2022). Many of these products rely on internal claims of efficacy and have not undergone external scientific testing, in part due to proprietary licencing laws (D'Hotman et al., 2021; Nebeker et al., 2019). Scholars criticize insufficient transparency regarding business practices, user privacy and informed consent (D'Hotman and Loh, 2020). There is a conflict of interest here: companies can inflate the quality of their technologies for profit and carry out unethical practices without scientific or regulatory scrutiny (Balcombe and De Leo, 2021; Cosgrove et al., 2020b). Authors add that many businesses are not properly equipped to provide quality and long-term mental healthcare and may oversimplify their product(s) to appeal to a large audience, resulting in a service that cannot support the complexities of mental health (Abdelrahman, 2023; Eagle et al., 2022; Ma et al., 2023). To increase and maintain user engagement, technologies may additionally integrate gamification which can lead to addiction (Denecke et al., 2022).

Seemingly free digital mental health services in fact require 'payment' in personal data, a prized resource that has generated billions of dollars (Nebeker et al., 2019). Through the practice of data capitalism, user data is routinely sold to corporate marketing teams, insurance companies and governments in order to sell products and manipulate individuals, without user's knowledge or consent (Cosgrove et al., 2020a). This commodification is considered an abhorrent violation of privacy and humanity since the data of commercial interest is mental fragility (Baumgartner, 2021).

4. Discussion

Of the potential outcomes of digital mental health technology, those of most concern relate to services developed by private companies for non-clinical use, both at an individual and population level. This discussion brings elements from the literature summarised above into conversation with wider literature by analyzing four prominent framings of mental health and addressing two ethical themes.

4.1. Framing of mental health

Many mental health applications convey the idea that poor mental health constitutes mental illness. This spreads the idea that suffering is likely due to a medical abnormality such as mental illness, rather than a potentially normal facet of the human emotional experience (Cosgrove and Karter, 2018; Doblytė, 2020). Normal is thus constrained to positive and productive mental states, creating the impression that human beings should be constantly happy (Sweet and Decoteaua, 2017). Such psychiatrization ignores the fact that normality is defined by culture and life experience and may lead to the imposition of largely Western ideals as universal standards of being (Beeker et al., 2021; Fusar-Poli et al., 2020). Since suffering is not legitimized if it is not mental illness, people seek out diagnoses and pharmaceutical treatment in order to validate their suffering, and moreover taking prescription drugs becomes the norm for managing mental health (Doblytė, 2020; Esposito and Perez, 2014). Medicalizing average distress as chronic mental health issues can reinforce negative feelings in individuals, prevent resilience and detract from the gravity of severe mental disorders (Beeker et al., 2021). Alternatively, a wider breadth of what is considered mental illness, in line with the broadening of the DSM, may reflect a growing sensitivity to suffering in society, a lowering of barriers to care, and a dismantling of stigma (ibid). It should be noted that due to the cultural nature of mental illness categorization, there are many historical instances of beneficial psychiatrization and de-psychiatrization obtained through patient advocacy, such as the recognition of attention-deficit/hyperactivity disorder in women and removing homosexuality from the DSM (ibid). While the threshold of what defines a mental illness remains unclear and must be continually scrutinized, society must concurrently allow for the validation of non-medicalized suffering.

Many emerging technologies, particularly AI, promote that the idea that mental health outcomes are a consequence of one's biology and can be objectively measured, rather than a composite of biological, situational and societal forces. This belief persists notwithstanding substantial evidence refuting theories of chemical imbalance and genetic causes of mental illness (Gorjup and Makovec, 2021; Moncrieff, 2022). Rejection of mental illness subjectivity enforces the idea that patients are unreliable and unknowing of their needs (Doblytė, 2020; Esposito and Perez, 2014; Russo, 2023). This approach leads to the dehumanization of mental healthcare in which treatment is geared towards a disease rather than a human being (Sweet and Decoteaua, 2017). Whereas proponents of objective mental health detection emphasize that to frame mental illness as biological provides understanding and alleviates notions of fault (Lebowitz, 2019). However this framework can equally increase stigma from health providers and cause hopelessness in mentally ill individuals (ibid). Furthermore, by not acknowledging the role of socio-economic status and discriminatory structures in many people's poor mental health, governments may absolve themselves of responsibility to enact systemic change that would impart an enormous positive impact on mental wellbeing (Doblytė, 2020; Zeira, 2022).

Another dominant framing within semi- and fully autonomous digital mental health technologies is the notion that one's mental health can be self-managed. This may be experienced as empowering and motivating for individuals to exert control over their mental healthcare (McAndrew et al., 2018). Alternatively, self-management can be akin to abandonment, especially if therapeutic relationships are undermined and for those that lack support systems (Esposito and Perez, 2014; Wang et al., 2018). This conception creates expectations that all mental health issues are manageable and that one is morally responsible for their mental state (Sweet and Decoteaua, 2017). This follows a neoliberal health framework common in Western countries in which poor health is seen as an individual failure, rather than a composite of individual choice, biology and societal factors (Esposito and Perez, 2014). Supporting such a mentality reinforces stigma and complies with a capitalist ideology that holds productivity above wellbeing (Moncrieff, 2022; Zeira, 2022). Paradoxically, this perspective simultaneously holds that individuals are responsible for taking care of their health independently, but that they are not trustworthy enough to not require overarching surveillance (Cosgrove and Karter, 2018). The idea of mental illness as self-manageable promotes a world in which people must constantly self-regulate against predicted risks from monitoring algorithms to ensure a level of perfect mental health that is unattainable (Rubeis, 2023).

Digital mental health technologies recurrently rely on and justify privacy, consent and agency rights infringements, from the use of population-based monitoring to medication adherence tracking and hard interventions such as sending authorities to user's homes. Potential consequences of these undertakings include proliferating stigma, forced treatment, traumatic experiences with law enforcement or medical staff, and destroying individuals' trust in their healthcare system (Nazroo et al., 2019). The discriminatory rhetoric underscoring these initiatives is that coercion is justified due to the undermined rationality of those with mental disorders (Russo, 2023). In response, mental health professionals have combatted the advent of indigestible markers on psychiatric pharmaceuticals, arguing that it is unjustified paternalism against those with severe mental illness and that there are many valid reasons people discontinue these notoriously problematic drugs (Cosgrove and Karter, 2018). Additionally, they emphasize restricting autonomy only if someone is a danger to themselves and to do so in the least intrusive way possible, with the highest amount of empathy and patient collaboration (Hem et al., 2018). Framing mental illness as a justification for curtailing rights lays the groundwork for human rights abuses, profiling and coercive diagnosis/treatment, which following historical and societal trends would unjustly and disproportionately be inflicted upon marginalized people (Nazroo et al., 2019; Russo, 2023).

4.2. Ethical themes

The effects of impending digital mental health technologies will largely be predicted by an individual's illness severity, access to in-person mental healthcare, socio-economic status, support system strength, racial and gender identity and geographical location. In other words, it will be largely predicted by privilege. The literature suggests those who will gain the most from these technologies will likely have less complex mental health issues and a sufficient degree of executive functioning to allow for self-management, and/or access to in-person mental healthcare, meaning these services would be used as a clinical tool or adjunctive treatment. Under such circumstances, it is quite probable that these technologies will be effective, though it is still possible that distress may be overmedicalized and result in unnecessary treatment that may harm patients (Rothstein, 2021). Ultimately, the evidence suggests that the people most likely to benefit are not the individuals most in need of care, for whom these technologies claim to increase accessibility and quality of care. Furthermore, the expansion of the digital mental health technology market may undermine already feeble existing services. Thus, people in less privileged positions may either solely rely on mental health technologies due to lack of access to traditional care or face no access due to digital divide constraints (Skorburg and Yam, 2021). Realistically, most of these services are not equipped to help those with the complexities of severe mental disorders or those whose poor mental health is in part due to their socio-economic, societal or life circumstances. Within the non-clinical applications of mental health technologies, the main treatment options available will either be relatively simple hands-off care or coercive care. Thus it is foreseeable that a large proportion of people with mental disorders will be unable to obtain the efficient and effective healthcare they need. Based on trends of discrimination and abuse in society, those who will suffer coercive diagnosis, treatment and profiling/discrimination will unjustly be those who are racialized and/or have severe mental illness (Russo, 2023). The way digital mental health technologies are poised to be implemented runs the risk of magnifying individual privilege or vulnerability; this is antithetical to ethical mental healthcare.

The framings of mental health encouraged throughout digital mental health technologies purposely coincide with capitalist incentives, as they are rooted in medical neoliberalism (Cosgrove and Karter, 2018). First, focusing on biological causes of mental illness and self-management allows for a simple and inexpensive product that places the onus of care on the patient (Esposito and Perez, 2014; Moncrieff, 2022). Second, encouraging the medicalization of distress works to expand the consumer base for both mental healthcare and pharmaceuticals, through the practice of disease mongering (Gorjup and Makovec, 2021; Kaczmarek, 2019). Finally, by justifying the need for self-directed or population-based mental health monitoring, businesses have created a perfect model to widely collect mental health data. In line with the theory of surveillance capitalism, procuring this immense amount of personal data may in fact be one of the main drivers for private companies to develop digital mental health technologies (Rubeis, 2023; Zuboff, 2015). Surveillance capitalism contends that the purpose of collecting data is to sell it to third parties, who use this information to modify people's behaviour for profit and control (Zuboff, 2015). Potential outcomes of this practice applied to mental health data include pharmaceutical companies spurring further overmedicalization in order to sell drugs, insurance companies enacting discriminatory pricing and discrimination by employers or governments based on calculated mental illness status (Doblytė, 2020; Monteith and Glenn, 2016). The fallout is the creation of "common sense" around mental health that serves to promote Eurocentric neoliberalism, individualism, and health consumerism (Esposito and Perez, 2014). The central goal of private capitalist businesses is to make profit and retain customers, which is contradictory to the aims of ethical and effective healthcare (ibid). Productive mental healthcare must address the complexity of the mental health problems, conserve therapeutic relationships and prioritize patient agency and privacy. This should include the integration of technology in a regulated manner in with extensive collaboration across affected communities, clinical services, areas of scholarship and public-private partnerships (Hahn et al., 2017; Kozelka et al., 2021). General messaging is furthermore needed that it is normal to experience poor mental health, and that people are deserving of care even if their suffering is not due to a medical condition (Fergusson et al., 2023). This requires societies to implement deep structural changes to alleviate inequity and allocate appropriate funding to mental healthcare and wellbeing care.

4.3. Limitations

This critical review is both strengthened and weakened by its broad scope, which provides an overview of societal implications of digital mental health technologies but does not address all issues presented exhaustively. A significant limitation is the subjectivity introduced due to the fact that the record screening process and qualitative analysis was conducted solely by one author (OS) with support from a second (AP). Furthermore, a major limitation of this research is that most of the review's records are from high-income Western countries. Only 2 of 51 sources pertained specifically to a non-Western or non-high-income country subject of interest and 83% of sources originated from institutions located in geographically Western countries and 92% from institutions in high-income countries, although this does not account for diversity in author backgrounds. This bias results in a Eurocentric worldview within the sources, which is reflected in this review. It is significant to note that the review does not extend focus to ideas based on popularity, but rather presents the breadth of thinking on the themes discussed in an effort to highlight diversity in perspectives. Additionally, there is a shortage of empirical data (26% of review records) as many of these technologies are still being developed and evaluated. A strength of this work is its examination of diverse perspectives ranging from techno-optimistic to heavily techno-skeptical, which provides this research with a more balanced, nuanced and comprehensive view. Record characteristics are presented in Appendix B.

4.4. Steps forward

Further research in this area is required to address how the adoption of digital mental health technologies will affect mental healthcare in low- and middle-income countries, and how diverse cultures will interact with these technologies and their links to neoliberal capitalist ideologies. Scholarship must be devoted to examining the potential impacts on vulnerable communities, especially in regard to coercion and surveillance features.

Public health programs and research that critically examine and challenge detrimental mental health framings while creating supportive structures for wellbeing are essential. This may include specific initiatives to increase literacy related to mental health technologies for both patients and providers (Suh et al., 2024). Initiatives are additionally needed to breakdown silos within this field to allow for cross-disciplinary collaboration between academics, digital technology developers and those with lived-experience of mental illness (ibid). Policy advocacy and participatory research are crucial to inform and establish international and national regulations for these services that follow ethics of care, to enforce user safety, efficacy standards, informed consent, safeguard individual agency and protect mental health data against nefarious use (Tavory, 2024). Finally, we believe the considerations presented in this critical review can provide vital insight to developers of mental health technologies to achieve the thoughtful and beneficial products they strive for.

5. Conclusion

Global development of semi- and fully autonomous digital mental health technologies, with applications of artificial intelligence and big data, are poised to revolutionize mental healthcare. These innovations provide great promise to improve the accessibility, quality and efficiency of mental healthcare, particularly if used in conjunction with clinical care. However, current versions of these technologies spark warnings of overmedicalization and dehumanization and may undermine already feeble traditional systems. These services shift mental healthcare predominantly towards self-management alongside population-based monitoring systems, prompting concerns about patient abandonment, disability surveillance, coercion, and discriminatory profiling. The presence of numerous private actors in this domain creates circumstances ripe for the commodification of mental healthcare, which may compromise quality of care as well as patient privacy and agency through the practice of selling user health data to third parties. Prominent narratives of mental illness promoted by technologies emerging from the review include framing 'poor mental health as mental illness', mental illness as 'objective', 'self-manageable' and 'as justification for rights infringement'. Analysis of inequity in relation to potential outcomes from digital mental health technologies cautions that these services may amplify vulnerability or privilege in individuals. Surveillance capitalism and neoliberalism furthermore are argued to underpin the development of many of these technologies for private financial gain. This review suggests an urgent need for comprehensive regulation and scholarly collaboration in this field to curb adverse effects, while remaining optimistic regarding the beneficial applications of many technologies if used in collaboration with clinicians and communities of people who experience mental disorders. Addressing the global mental health crisis requires a transformation of current systems including allocating adequate funding and dismantling oppressive neoliberal structures; and while digital mental health technologies live unregulated within these foundations, they will inevitably magnify existing injustice.

CRediT authorship contribution statement

Olivia A. Stein: Writing – review & editing, Writing – original draft, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Audrey Prost:** Writing – review & editing, Supervision, Conceptualization.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

OS would like to thank Michael, Martha and Samuel Stein for their everlasting support, guidance and vivid spirit.

Abbreviations/Terms

Digital mental health technologies *referred to as* "technologies" or "services"
DSM Diagnostic and Statistical Manual of Mental Disorders

Appendix A. Additional Information Regarding Methodology

Table 2
Final Search Terms

Category N^{o}	Category Title	Keyword/Synonym
1	Mental Health	Mental health
		Mental* ill*
		Mental disorder*
2	Function	Treatment
		Monitoring
3	Technology	Artificial intelligence
		AI
		Predictive analytics
		Digital phenotyping
		Mental health app*
		Data-driven technolog*
4	Concept	Surveillance capitalism
		Data capitalism
		Capitalism
		Big data
		Privacy
		Ethic*
	Complete search	1 AND 2 AND 3 AND 4 *Keywords/synonyms combined with OR*

Table 3
Exact Database Search Example

Database	Exact Search
PubMed	("Mental health" OR "Mental* ill*" OR "Mental disorder*") AND ("Treatment" OR "Monitoring") AND ("Artificial intelligence" OR "AI" OR "Predictive analytics" OR "Digital phenotyping" OR "Mental health app*" OR "Data-driven technolog*") AND ("Surveillance capitalism" OR "Surveillance" OR "Data capitalism" OR "Capitalism" OR "Big data" OR "Privacy" OR "Ethic*")

Table 4
Initial Search Results

	Criterion Restriction	Number of Results
Database Searches		
PubMed	“title/abstract”	242
Scopus	“title-abs-key”	189
Web of Science	“topic”	114
Embase 1974–2023 (OVID version)	“keyword”	112
ProQuest Social Science	“anywhere except full text (noft)”	12
Global Health (OVID version)	“keyword”	2
Other Sources		
Initial Readings	n/a	4
		Total = 675 results

Table 5
Inclusion and Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria
Round 1	- Psychotic or emotional disorders - Mental health technology intervention (i.e. mental health app, AI, big data)	- Not specifically pertaining to mental health - Non-psychotic or emotional disorders (i.e. ADHD, autism, dementia, Parkinson’s, anorexia, OCD, addiction) - Mental health in relation to another illness (i.e. HIV, tuberculosis, other viruses) - Intervention not related to technology - Online therapy (i.e. therapy delivered by a human being via video conferencing) - Resource not available in English
Round 2	- Relation to the mental healthcare system - Follow-up/action generated from mental health technology - Commercial aspects of mental health technology - Self-management of mental health - Non-data related ethical issues (i.e. consent, coercion, surveillance, agency, privacy, human rights)	- Description of mental health technology functioning or services - Tool validity & objectivity - User acceptability/feasibility of mental health apps - Data related ethical issues (i.e. data privacy/protection, data bias, algorithmic inferences) - Study protocols - Conference abstracts - Full text not available

Table 6
Coding of Reviews from Search Results at Full-Text Review Stage (n = 26) Existing reviews from search results at the full-text review stage =

1	Digital Phenotyping for Mental Health: Reviewing the Challenges of Using Data to Monitor and Predict Mental Health Problems (Birk and Samuel, 2022)
2	Review of Use and Integration of Mobile Apps into Psychiatric Treatments (Chan et al., 2017)
3	AI enabled suicide prediction tools: a qualitative narrative review (D’Hotman and Loh, 2020)
4	Smartphone Ownership, Smartphone Utilization, and Interest in Using Mental Health Apps to Address Substance Use Disorders: Literature Review and Cross-sectional Survey Study Across Two Sites (Hsu et al., 2022)
5	A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining (Islam et al., 2018)
6	Potential and Pitfalls of Mobile Mental Health Apps in Traditional Treatment: An Umbrella Review (Koh et al., 2022)
7	Digital health developments and drawbacks: a review and analysis of top-returned apps for bipolar disorder (Lagen et al. 2020)
8	Influencing factors, prediction and prevention of depression in college students: A literature review (Liu et al., 2022)
9	Deep learning techniques for suicide and depression detection from online social media: A scoping review (Malhotra and Jindal, 2022)
10	Mobile intervention for individuals with psychosis, dual disorders, and their common comorbidities: A literature review (Pennou et al., 2019)
11	Mobile and wearable technology for monitoring depressive symptoms in children and adolescents: A scoping review (Sequeira et al., 2020)
12	Machine learning in mental health: a scoping review of methods and applications (Shatte et al., 2019)

(*continued on next page*)

Table 6 (*continued*)

13	Process and Outcome Evaluations of Smartphone Apps for Bipolar Disorder: Scoping Review (Tatham et al., 2022)
14	Machine Learning in Mental Health: A systematic review of the HCI literature to support the development of effective and implementable ML Systems (Thieme et al., 2020)
15	Smartphones, Sensors, and Machine Learning to Advance Real-Time Prediction and Interventions for Suicide Prevention: a Review of Current Progress and Next Steps (Torous et al. 2018)
16	Is There an App for That? A Review of Popular Apps for Depression, Anxiety, and Well-Being (Wasil et al., 2022)
17	Digital Mental Health for Young People: A Scoping Review of Ethical Promises and Challenges (Wies et al., 2021)
18	Mental health monitoring apps for depression and anxiety in children and young people: A scoping review and critical ecological analysis (Williams and Pykett, 2022)
19	The Emerging Role of Technology in Cognitive-Behavioral Therapy for Anxious Youth: A Review (Berry and Lai, 2014)
20	Implementation of Cognitive Behavioral Therapy in e-Mental Health Apps: Literature Review (Denecke et al., 2022)
21	Annual Research Review: Translational machine learning for child and adolescent psychiatry (Dwyer and Koutsouleris, 2022)
22	Could Robots Empatize? A Review on The Employment of Social Robots in Mental Healthcare (Gültekin, 2022)
23	Artificial intelligence (AI) and machine learning (ML) based decision support systems in mental health: An integrative review (Higgins et al., 2023)
24	Artificial Intelligence-Enabled Chatbots in Mental Health: A Systematic Review (Omarov et al., 2023)
25	How Can Digital Mental Health Enhance Psychiatry? (Stern et al., 2022)
26	Opportunities, applications, challenges and ethical implications of artificial intelligence in psychiatry: a narrative review (Terra et al., 2023)

Coding Theme	Reviews
Broad applications for mental healthcare	Chan et al., 2017; Islam et al., 2018; Shatte et al., (2019); Thieme et al., 2020; Dwyer and Koutsouleris (2022); Higgins et al., 2023; Stern et al., (2022); Terra et al. (2023)
Technology for predicting/detecting mental illness or suicide risk	D'Hotman and Loh (2020); Liu et al., (2022); Malhotra and Jindal (2022); Torous et al. 2018
Technology for monitoring or treating mental health	Pennou et al., 2019; Sequeira et al., (2020); Williams and Pykett (2022)
AI chatbot therapy technology	Berry and Lai (2014); Denecke et al., (2022); Gültekin (2022); Omarov et al. (2023)
Efficacy	Tatham et al., 2022; Torous et al. 2018; Wies et al., (2021); Williams and Pykett (2022); Dwyer and Koutsouleris (2022)
User acceptability, utility and accessibility	Chan et al., 2017; Hsu et al., 2022; Lagen et al. 2020; Wasil et al. (2022)
Data & ethical concerns	Birk and Samuel (2022); Koh et al., (2022); Malhotra and Jindal (2022); Thieme et al., 2020; Wies et al., (2021); Williams and Pykett (2022); Denecke et al., (2022); Terra et al. (2023)

Table 7
Description and Thematic Coding of Records Included in the Review (n = 53) Please note that sub-categories are not necessarily exhaustive, but seek to denote major and minor themes within the literature as relevant to this review

Author (Year)	Summary & Relevance	Thematic Coding Categories (Sub-Categories)
Records Identified through Database Searching		
Abdelrahman (2023)	Critique of mental illness/trauma self-management apps for refugees	Mental healthcare (MHC) system (quality of care; undermine system; overmedicalization; dehumanizing); Self-management (harmful); Ethics (framing of mental illness); Commodification (gamification; data capitalism); Monitoring (surveillance)
Atlam et al. (2022)	AI for emotional recognition; action & role in MHC system	Monitoring (vague intervention); MHC system (objectifying; support system)

(*continued on next page*)

Table 7 (*continued*)

Author (Year)	Summary & Relevance	Thematic Coding Categories (Sub-Categories)
Balcombe and De Leo (2021)	General overview of technology for mental health; challenges; efficacy & effect on MHC system	MHC system (accessibility, quality of care)
Balcombe and De Leo (2022)	Presents concerns and promise of digital mental health tools, encourages human-computer interaction investment	MHC system (quality of care; accessibility; overmedicalization); Ethics (framing of mental illness; inequity)
Baumgartner (2021)	Digital phenotyping for mental illness; ethical issues with surveillance, conceptualization of mental health & data capitalism	MHC system (undermine system; overmedicalization); Ethics (coercion); Commodification (data capitalism); Self-management (harmful); Monitoring (surveillance)
Berry and Lai (2014)	Technology-based CBT or mental health apps for anxious youth; talks about self-monitoring;	Self-management (beneficial); MHC system (accessibility; quality of care; undermine system); Ethics (equity)
Betthauser et al. (2020)	User acceptability of the Cogito Companion app for veterans; self-monitoring & passive data collection; app may prompt clinician follow-up	Self-management (beneficial); Monitoring (soft intervention); MHC system (accessibility; quality of care; therapeutic relationship; clinical tool); Commodification (gamification)
Bidargaddi et al. (2021)	Study; AI2 software that tracks HC data for medication adherence; clinical use to trigger f/u	Monitoring (soft intervention); MHC system (quality of care; support system; clinical tool); Ethics (agency)
Cosgrove et al. (2020a)	Human rights & ethics around digital phenotyping; digital drug for compliance	Ethics (coercion; agency; consent; inequity); Commodification (data capitalism; profit incentive; transparency); MHC system (overmedicalization); Monitoring (hard intervention; surveillance)
D'Hotman and Loh (2020)	Medical & social suicide prediction tools; seen as positive for suicide prediction; many applications for data to predict mental illness, private company and government involvement	Monitoring (soft intervention; hard intervention); Commodification; MHC system (clinical tool)
D'Hotman et al. (2021)	Overview of digital mental health technology for medical professionals specifically; addresses human rights concerns	Monitoring (soft intervention; hard intervention); Ethics (agency; consent; inequity); Commodification (profit incentive; transparency; data capitalism)
Denecke et al. (2021)	Benefits and challenges of chatbots for mental health therapy; ethical concerns	MHC system (accessibility; quality of care; therapeutic relationship; undermine system; support system); Commodification (data capitalism); Monitoring (lack intervention); Ethics (agency; inequity)
Denecke et al. (2022)	CBT techniques implementation into mental health apps; potential benefits and ethical concerns	Self-management (beneficial); MHC system (accessibility; quality of care; support system); Commodification (gamification); Monitoring (lack intervention); Ethics (agency)
Eagle et al. (2022)	Negative potential outcomes of freemium mental health apps; suggestions for regulation	Commodification (profit incentive; gamification); Ethics (consent; agency; inequity); Self-management (harmful)
Feldman and Perret (2023)	Digital mental health specifically for post-partum depression; focus on connection to treatment, accessibility, agency and ease of access, ethical cons	MHC system (accessibility; quality of care; clinical tool); Self-management (beneficial); Monitoring (surveillance; soft intervention); Ethics (agency; inequity)
Fiske et al. (2019)	Ethics of AI for mental health care and effect on the mental healthcare system	Ethics (equity; agency; coercion; inequity); MHC system (accessibility; quality of care; support system; undermine system; therapeutic relationship); Monitoring (lack intervention
Graham et al. (2019)	Overview of AI for mental health; objective diagnosis; identify illness; personalized treatment; also goes into AI redefining mental illness & healthcare system; ethics	MHC system (quality of care; support system; clinical tool); Ethics (consent)
Gültekin (2022)	Employment of social robots with a physical body in mental healthcare	MHC system (therapeutic relationship; quality of care; objectifying; dehumanizing; undermine system)
Koh et al. (2022)	Overview of potential benefits and pitfalls of clinical adjunct use of mental health apps	MHC system (accessibility; quality of care; clinical use); Monitoring (lack intervention); Ethics (consent; inequity)
Luxton (2014)	Various usages of AI in the mental health field; ideations around the future of AI and superintelligence	MHC system (quality of care; objectifying; dehumanizing; undermine system)
Luxton et al. (2016): *Book entitled "Artificial Intelligence in Behavioral and Mental Health Care" that contains the following 3 chapters that met the inclusion criteria (considered separate sources throughout this review):*		
Poulin et al. *(2016)*	Study; opt-in suicide prediction for veterans using Facebook	Monitoring (soft intervention); MHC system (clinical use); Commodification (data capitalism; transparency); Ethics (consent)
Neuman (2016)	Monitoring for depression or tendencies towards violence	Monitoring (soft intervention; surveillance); Self-management; Ethics (framing of mental illness; agency)
Luxton et al. *(2016)*	Therapeutic relationship with robots; regulation needed; ethics around intelligent autonomous care providers collecting data and surveillance; AI system morality	MHC system (therapeutic relationship; quality of care; objectifying; dehumanizing); Ethics (consent; agency); Monitoring (surveillance); Commodification (data capitalism)
Ma et al. (2023)	Ethical criticism of Sonde (depression detection technology via vocal biomarkers); criticizes objectivity, universality and eugenic; critical disability studies	Ethics (inequity; agency; framing of mental illness); MHC system (dehumanizing; undermine system; overmedicalization); Commodification (profit incentive); Monitoring (surveillance)
McCradden et al. (2022)	Data objectivity & epistemic justice; AI vs patient testimonies	Ethics (agency; inequity; framing of mental illness); MHC system (dehumanizing); Monitoring (surveillance)
Nebeker et al. (2019)	Ethical concerns about direct-to-consumer mental health apps and need for industry regulation	Ethics (consent); Commodification (transparency); MHC system (quality of care)
Nogueira-Leite et al. (2023)	Mental health professional's attitudes regarding using mental health apps in clinical practice	MHC system (accessibility; quality of care; support system; clinical tool)
Oakey-Neate et al. (2020)	Study protocol: AI^2 software to nudge patients or clinicians based on med adherence	Monitoring (soft intervention); MHC system (clinical tool; support system)
Omarov et al. (2023)	Overview of AI chatbot functionality and usage; presents some ethical considerations	MHC system (accessibility; quality of care; support system; undermine system; therapeutic relationship); Ethics (inequity; consent)
Parker et al. (2018)	Problematic notion of mental health apps & over-medicalization of mental states	MHC system (overmedicalization); Ethics (framing of mental illness); Self-management (harmful)
Parrish et al. (2022)	Study evaluates mental health app crisis resources for suicidal ideation/ behaviour	Monitoring (lack intervention); MHC system (quality of care)
Radovic et al. (2016)	Outlines & critiques features and rhetoric of common mental health apps	Ethics (framing of mental illness); MHC system (quality of care); Self-management (harmful; beneficial)
Rauseo-Ricupero et al. (2021)	MindLAMP app used by patients as a clinical tool; provides journaling, CBT activities and biomarker tracking	MHC system (clinical tool; quality of care; accessibility); Monitoring (lack intervention)

(*continued on next page*)

Table 7 (*continued*)

Author (Year)	Summary & Relevance	Thematic Coding Categories (Sub-Categories)
Rickard et al. (2022)	Study evaluates quality of mental health apps targeted at mood disorders and compares to app store visibility	MHC system (accessibility; quality of care); Self-management (beneficial)
Robinson et al. (2020)	Self-harm monitoring technology based on emergency dept. health records; data used to inform interventions	Monitoring (vague intervention); MHC system (quality of care)
Rosenfeld et al. (2021)	Limitations of technology for mental health due to lack of biomarkers & objective diagnosis; talks about treatment delivery through AI therapist/ mindfulness; use of monitoring data	MHC system (therapeutic relationship; quality of care; accessibility, support system; objectifying; clinical tool); Monitoring (lack intervention; surveillance)
Rubeis (2022)	Explores potential benefits and ethical concerns of self-monitoring, ecological momentary assessment and data mining in AI/big data mental health technologies	Self-management (beneficial; harmful); MHC system (quality of care; therapeutic relationship; dehumanization); Ethics (agency; consent); Monitoring (surveillance); Commodification (data capitalism)
Schmidt & D'Alfonso (2023)	Ethical themes in involved digital phenotyping in mental health care; data informing treatment	Ethics (agency, consent); MHC system (clinical tool; quality of care; objectifying); Monitoring (surveillance); Commodification (data capitalism)
Starke et al. (2021)	Machine learning for schizophrenia diagnosis via MRIs; pros & cons; concerns of redefining diagnosis, patient consent & safeguards	MHC system (quality of care; overmedicalization; undermine system); Ethics (inequity; framing of mental illness); Commodification (profit incentive)
Stern et al. (2022)	Overview of digital technology pros & cons in psychiatry to assist clinical care, app evaluation; data for research	MHC system (accessibility; quality of care)
Tekin (2021)	Criticism of smartphone psychotherapy chatbots on epistemic and ethical grounds	MHC system (dehumanizing; therapeutic relationship; undermine system; overmedicalization); Ethics (inequity; framing of mental illness); Commodification (data capitalism); Self-management (harmful)
Uma et al. (2023)	Benefits of AI mental health technology for healthcare system	MHC system (quality of care); Self-management (beneficial)
Uusitalo et al. (2021)	Ethical concerns regarding AI in clinical practice; focus on AI redefining diagnosis and impact on mental health framing	Ethics (framing of mental illness; consent; agency); MHC system (quality of care; dehumanizing; undermine system; therapeutic relationship; accessibility; overmedicalization); Monitoring (surveillance)
Van Assche et al. (2022)	Outlines various applications of digital phenotyping for depression; presents ethical considerations	Ethics (consent); MHC system (quality of care; objectifying; clinical tool; support system; accessibility); Self-management (harmful); Monitoring (vague intervention)
Wasil et al. (2022)	Description of content of popular mental health/wellness apps for mindfulness, self-monitoring or AI chatbots	Self-management (beneficial); MHC system (accessibility); Monitoring (soft intervention); Commodification (transparency); Ethics (framing of mental illness)
Wies et al. (2021)	Outlines broad overview of pros & cons of technology for mental health for ages 0–25; ethics focused	MHC system (accessibility; quality of care); Ethics (equity; agency; consent; inequity; framing of mental illness); Self-management (beneficial; harmful); Commodification (gamification)
Williams and Pykett (2022)	Mental health monitoring apps: self-management interventions; standardization of mental health symptoms & effect on the societal framing of mental illness	Self-management (beneficial; harmful); Ethics (framing of mental illness); MHC system (overmedicalization; quality of care); Commodification (profit incentive); Monitoring (lack intervention)
Woodward et al. (2022)	Overview of technology to sense mental health states or provide mental health treatment; discusses ethical and logistical challenges	MHC system (quality of care; accessibility; objectifying); Commodification (data capitalism); Monitoring; Ethics (framing of mental illness; consent; inequity)
Yu et al. (2022)	Healthcare needs and stigma barriers for people with severe mental illness; AI increasing access and efficiency	MHC system (support system; accessibility; quality of care)
Records Identified through Other Sources		
Bauer et al. (2020)	Mental health app and self-monitoring unpopularity; ethical concerns regarding passive monitoring and data capitalism	Self-management (harmful); Ethics (consent); MHC system (overmedicalization; quality of care); Monitoring (lack intervention); Commodification (data capitalism)
Carr (2020)	Power and mental illness; consent to use mental health monitoring data and coercion to treatment	Ethics (consent, agency, coercion, inequity)
Cosgrove et al. (2020b)	Surveillance capitalism, overmedicalization of mental health; ethics of consent, business conflict of interest and treatment coercion	Ethics (consent; agency; coercion); MHC system (overmedicalization); Commodification (data capitalism; profit incentive); Monitoring (hard intervention)
Kolenik and Gams (2021)	Self-management through behaviour change; benefits for mental healthcare system; accessibility to marginalized groups; dismantling stigma	Self-management (beneficial); MHC system (accessibility; support system)

Table 8

Frequency of Themes and Sub-Themes (n = 53) The frequency signifies the number of individual review records that significantly mention a theme/sub-theme at least 1 time and percentage is the frequency out of the 53 total records.

Theme	Sub-Theme	Frequency	Percentage
Mental Healthcare System		**49**	**92.5%**
	Accessibility	22	41.5%
	Quality of care	36	67.9%
	Undermines/supports system	22	41.5%
	Dehumanizing/objectifying	14	26.4%
	Therapeutic relationship	10	18.9%
	Overmedicalization	12	22.6%
	Clinical tool	11	20.8%
Self-Management		**20**	**37.7%**
	Beneficial	12	22.6%
	Harmful	11	20.8%
Monitoring		**32**	**60.4%**
	Lack or vague intervention	12	22.6%
	Soft intervention	9	17.0%
	Hard intervention	4	7.5%

(*continued on next page*)

Table 8 (*continued*)

Theme	Sub-Theme	Frequency	Percentage
	Surveillance	12	22.6%
Commodification		**23**	**43.4%**
	Profit incentive	7	13.2%
	Transparency	5	9.4%
	Gamification	5	9.4%
	Data capitalism	13	24.5%
Ethics		**36**	**67.9%**
	Agency	18	34.0%
	Consent	18	34.0%
	Coercion	5	9.4%
	Inequity/equity	17	32.1%
	Framing of mental illness	14	26.4%

Appendix B. Additional Information Relevant to Limitations

Table 9
Characteristics of Records Included in the Review (n = 53) Note: Empirical vs non-empirical criterion follows the characteristics outlined by Dan (2017). Overall tone is determined by the interpretation of OS.

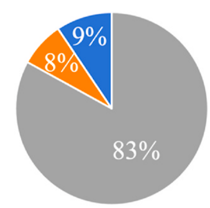
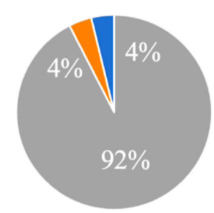
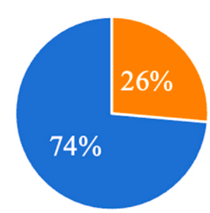
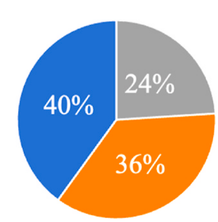
Author (Year)	Subject of Interest (if specified)	Affiliated Institution(s) Location(s)	Article Type; Dominant Methods	Overall Tone
Records Identified through Database Searching				
Abdelrahman (2023)	Refugees	UK	Non-Empirical	Pessimistic
Atlam et al. (2022)		UK, Egypt	Non-Empirical	Optimistic
Balcombe and De Leo (2021)	Australia	Australia	Non-Empirical	Optimistic
Balcombe and De Leo (2022)		Australia	Review; Non-Empirical	Mixed
Baumgartner (2021)		Germany	Research; Non-Empirical	Pessimistic
Berry and Lai (2014)	Youth; Anxiety	USA	Review; Non-Empirical	Optimistic
Betthauser et al. (2020)	USA; Veterans	USA	Mixed Methods Research; Empirical	Optimistic
Bidargaddi et al. (2021)	Australia	Australia	Quantitative Research; Empirical	Optimistic
Cosgrove et al. (2020a)		USA	Perspective; Non-Empirical	Pessimistic
D'Hotman and Loh (2020)		UK, Australia	Review; Non-Empirical	Mixed
D'Hotman et al. (2021)		UK, Australia	Commentary; Non-Empirical	Mixed
Denecke et al. (2021)		Switzerland, Qatar	Non-Empirical	Mixed
Denecke et al. (2022)		Switzerland	Review; Non-Empirical	Optimistic
Eagle et al. (2022)	USA	USA	Mixed Methods Research; Empirical	Pessimistic
Feldman and Perret (2023)	Post-partum depression	USA	Comment; Non-Empirical	Mixed
Fiske et al. (2019)		Germany	Research; Non-Empirical	Mixed
Graham et al. (2019)		USA	Review; Non-Empirical	Optimistic
Gültekin (2022)		Turkey	Review; Non-Empirical	Mixed
Koh et al. (2022)		USA, Singapore	Review; Non-Empirical	Mixed
Luxton (2014)		USA	Non-Empirical	Mixed
Luxton et al. (2016): *Book entitled "Artificial Intelligence in Behavioral and Mental Health Care" that contains the following 3 chapters that met the inclusion criteria (considered separate sources throughout this review):*				
Poulin et al. *(2016)*	USA; Veterans	USA	Non-Empirical	Mixed
Neuman (2016)	USA; Mass shootings	Israel	Non-Empirical	Mixed
Luxton et al. *(2016)*		USA	Non-Empirical	Mixed
Ma et al. (2023)		Canada	Non-Empirical	Pessimistic
McCradden et al. (2022)		Canada	Non-Empirical	Pessimistic
Nebeker et al. (2019)		USA	Review; Non-Empirical	Mixed
Nogueira-Leite et al. (2023)	Portugal	Australia	Mixed Methods Research; Empirical	Optimistic
Oakey-Neate et al. (2020)		Portugal	Protocol; Non-Empirical	Optimistic
Omarov et al. (2023)		Kazakhstan	Review; Non-Empirical	Optimistic
Parker et al. (2018)		Australia	Qualitative Research; Empirical	Pessimistic
Parrish et al. (2022)		USA	Mixed Methods Research; Empirical	Pessimistic
Radovic et al. (2016)		USA	Mixed Methods Research; Empirical	Mixed
Rauseo-Ricupero et al. (2021)		USA	Mixed Methods Case Reports Research; Empirical	Optimistic
Rickard et al. (2022)		Australia	Systemic Assessment; Empirical	Mixed
Robinson et al. (2020)	Australia	Australia	Protocol; Non-Empirical	Optimistic
Rosenfeld et al. (2021)		Canada, Israel	Non-Empirical	Optimistic
Rubeis (2022)		Austria	Non-Empirical	Mixed
Schmidt & D'Alfonso (2023)		Australia	Qualitative Research; Empirical	Mixed
Starke et al. (2021)	Schizophrenia	France, Canada	Editorial; Non-Empirical	Mixed
Stern et al. (2022)		France, Canada	Review; Non-Empirical	Optimistic

(*continued on next page*)

Table 9 (*continued*)

Author (Year)	Subject of Interest (if specified)	Affiliated Institution(s) Location(s)	Article Type; Dominant Methods	Overall Tone
Tekin (2021)		USA	Research; Non-Empirical	Pessimistic
Uma et al. (2023)		India, Saudi Arabia	Qualitative Secondary Research; Empirical	Optimistic
Uusitalo et al. (2021)		Finland	Non-Empirical	Pessimistic
Van Assche et al. (2022)	Depression	Germany, Spain, UK, Switzerland, Australia, the Netherlands, Estonia	Review; Non-Empirical	Optimistic
Wasil et al. (2022)		USA	Qualitative Research; Empirical	Optimistic
Wies et al. (2021)	Youth aged 0-25	Switzerland	Review; Non-Empirical	Mixed
Williams and Pykett (2022)	UK; Youth	UK	Review; Non-Empirical	Mixed
Woodward et al. (2022)	UK	UK	Qualitative Research; Empirical	Mixed
Yu et al. (2022)	China; Serious mental illness	China	Non-Empirical	Optimistic
Records Identified through Other Sources				
Bauer et al. (2020)	UK, USA, Australia, Canada, EU, Singapore, New Zealand	Germany, USA, UK, Canada, Denmark	Review; Empirical	Pessimistic
Carr (2020)		UK	Editorial; Non-Empirical	Pessimistic
Cosgrove et al. (2020b)		USA	Non-Empirical	Pessimistic
Kolenik and Gams (2021)		Slovenia	Non-Empirical	Optimistic

Table 10
Distribution of Characteristics of Records Included in the Review (n = 53)

	Characteristic	Percentage	Graph
Affiliated Institution(s) Location(s) – Geographic Region ('Western' includes North America, Europe and Australia)	Only Geographically Western Countries Both Geographically Western and Non-Western Countries Only Geographically Non-Western Countries	83% 8% 9%	
Affiliated Institution(s) Location(s) – Income Classification based on World Bank (2024)	Only High-Income Countries Both High-Income Countries & Non-High-Income Countries Only Non-High-Income Countries	92% 4% 4%	
Dominant Methods	Empirical Non-Empirical	26% 74%	
Overall Tone	Optimistic Mixed Pessimistic	36% 40% 24%	

References

Abdelrahman, M., 2023. Trauma apps and the making of the 'smart' refugee. EPD: Society and Space 0 (0), 1–16. https://doi.org/10.1177/02637758231173416.

Ahmed, A., Agus, M., Alzubaidi, M., Aziz, S., Abd-Alrazaq, A., Giannicchi, A., Househ, M., 2022. Overview of the role of big data in mental health: a scoping review. Computer Methods and Programs in Biomedicine Update 2, 100076. https://doi.org/10.1016/j.cmpbup.2022.100076, 1-11.

Atlam, H.F., Shafik, M., Kurugollu, F., Elkelany, Y., 2022. Emotions in mental healthcare and psychological interventions: towards an inventive emotions recognition framework using AI. In: Shafik, M., Case, K. (Eds.), Advances in Manufacturing Technology XXXV. IOS Press, pp. 317–322. https://doi.org/10.3233/ATDE220609.

Balcombe, L., De Leo, D., 2021. Digital mental health challenges and the horizon ahead for solutions. JMIR Mental Health 8 (3). https://doi.org/10.2196/26811.

Balcombe, L., De Leo, D., 2022. Human-computer interaction in digital mental health. Informatics 9 (14), e26811. https://doi.org/10.3390/informatics9010014, 1-18.

Bauer, M., Glenn, T., Geddes, J., Gitlin, M., Grof, P., Kessing, L.V., Monteith, S., et al., 2020. Smartphones in mental health: a critical review of background issues, current status and future concerns. Int. J. Behav. Dev. 8 (2), 1–19. https://doi.org/10.1186/s40345-019-0164-x.

Baumgartner, R., 2021. Precision medicine and digital phenotyping: digital medicine's way from more data to better health. Big Data & Society 1–12. https://doi.org/10.1177/20539517211066452.

Beeker, T., Mills, C., Bhugra, D., te Meerman, S., Thoma, S., Heinze, M., von Peter, S., 2021. Psychiatrization of society: a conceptual framework and call for transdisciplinary research. Front. Psychiatr. 12, 1–11. https://doi.org/10.3389/fpsyt.2021.645556.

Berry, R.R., Lai, B., 2014. The emerging role of technology in cognitive-behavioral therapy for anxious youth: a review. J. Ration. Emot. Cogn. Behav. Ther. 32, 57–66. https://doi.org/10.1007/s10942-014-0184-5.

Betthauser, L.M., Stearns-Yoder, K.A., McGarity, S., Smith, V., Place, S., Brenner, L.A., 2020. Mobile app for mental health monitoring and clinical outreach in veterans: mixed methods feasibility and acceptability study. J. Med. Internet Res. 22 (8), 1–14. https://doi.org/10.2196/15506.

Bidargaddi, N., Schrader, G., Myles, H., Schubert, K.O., van Kasteren, Y., Zhang, T., Bastiampillai, T., et al., 2021. Demonstration of automated non-adherence and service disengagement risk monitoring with active follow-up for severe mental illness. Aust. N. Z. J. Psychiatr. 55 (10), 976–982. https://doi.org/10.1177/0004867421998800.

Birk, R.H., Samuel, G., 2022. Digital phenotyping for mental health: reviewing the challenges of using data to monitor and predict mental health problems. Curr. Psychiatr. Rep. 24, 523–528. https://doi.org/10.1007/s11920-022-01358-9.

Carr, S., 2020. AI gone mental': engagement and ethics in data- driven technology for mental health. J. Ment. Health 29 (2), 12–130. https://doi.org/10.1080/09638237.2020.1714011.

Cosgrove, L., Karter, J.M., 2018. The poison in the cure: neoliberalism and contemporary movements in mental health. Theor. Psychol. 28 (5), 669–683. https://doi.org/10.1177/0959354318796307.

Cosgrove, L., Karter, J.M., McGinley, M., Morrill, Z., 2020a. Digital phenotyping and digital psychotropic drugs: mental health surveillance tools that threaten human rights. Health and Human Rights Journal 22 (2), 33–39.

Cosgrove, L., Karter, J.M., Morrill, Z., McGinley, M., 2020b. Psychology and surveillance capitalism: the risk of pushing mental health apps during the COVID-19 pandemic. J. Humanist. Psychol. 60 (5), 611–625. https://doi.org/10.1177/0022167820937498.

D'Hotman, D., Loh, E., 2020. AI enabled suicide prediction tools: a qualitative narrative review. BMJ Health & Care Informatics 27, e100175. https://doi.org/10.1136/bmjhci-2020-100175, 1-10.

D'Hotman, D., Loh, E., Savulescu, J., 2021. AI-enabled suicide prediction tools: ethical considerations for medical leaders. BMJ Leader 5, 102–107. https://doi.org/10.1136/leader-2020-000275.

D'Alfonso, S., 2020. AI in mental health. Current Opinion in Psychology 36, 112–117. https://doi.org/10.1016/j.copsyc.2020.04.005.

Dan, V., 2017. Empirical and nonempirical methods. In: Matthes, J., Davis, C.S., Potter, R.F. (Eds.), The International Encyclopedia of Communication Research Methods. John Wiley & Sons. https://doi.org/10.1002/9781118901731.iecrm0083.

Denecke, K., Abd-Alrazaq, A., Househ, M., 2021. Artificial intelligence for chatbots in mental health: opportunities and challenges. In: Househ, M., Borycki, E., Kushniruk, A. (Eds.), *Multiple Perspectives On Artificial Intelligence In Healthcare*. Lecture Notes in Bioengineering. Springer. https://doi.org/10.1007/978-3-030-67303-1_10.

Denecke, K., Schmid, N., Nussli, S., 2022. Implementation of cognitive behavioral therapy in e-mental health apps: literature review. J. Med. Internet Res. 24 (3), 1–13. https://doi.org/10.2196/27791.

Doblytė, S., 2020. Under- or overtreatment of mental distress? Practices, consequences, and resistance in the field of mental health care. Qual. Health Res. 30 (10), 1503–1516. https://doi.org/10.1177/1049732320918531.

Dwyer, D., Koutsouleris, N., 2022. Annual Research Review: translational machine learning for child and adolescent psychiatry. JCPP (J. Child Psychol. Psychiatry) 63 (4), 421–443. https://doi.org/10.1111/jcpp.13545.

Eagle, T., Mehrotra, A., Sharma, A., Zuniga, A., Whittaker, S., 2022. Money doesn't buy you happiness": negative consequences of using the freemium model for mental health apps. Proceedings of the ACM on Human-Computer Interaction 265 (CSCW2), 1–38. https://doi.org/10.1145/3555155, 6.

Esposito, L., Perez, F.M., 2014. Neoliberalism and the commodification of mental health. Humanity Soc. 38 (4), 414–442. https://doi.org/10.1177/0160597614544958.

Feldman, N., Perret, S., 2023. Digital mental health for postpartum women: perils, pitfalls, and promise. npj Digital Medicine 6 (11), 1–4. https://doi.org/10.1038/s41746-023-00756-4.

Fergusson, E., Reed-Purvis, S., Foulkes, L., 2023. Overmedicalization of young people's distress is undermining and disempowering families. Nature Mental Health 1, 381–382. https://doi.org/10.1038/s44220-023-00071-7.

Ferrari, A.J., Santomauro, D.F., Herrera, A.M.M., Shadid, J., Ashbaugh, C., Erskine, H.E., Charlson, F.J., et al., 2022. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatr. 9, 137–150. https://doi.org/10.1016/S2215-0366(21)00395-3.

Fiske, A., Henningsen, P., Buyx, A., 2019. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. J. Med. Internet Res. 21 (5), 1–12. https://doi.org/10.2196/13216.

Fusar-Poli, P., de Pablo, G.S., De Micheli, A., Nieman, D.H., Correll, C.U., Kessing, L.V., Pfennig, A., et al., 2020. Eur. Neuropsychopharmacol 31, 33–46. https://doi.org/10.1016/j.euroneuro.2019.12.105.

Gooding, P., 2019. Mapping the rise of digital mental health technologies: emerging issues for law and society. Int. J. Law Psychiatr. 67, 101498. https://doi.org/10.1016/j.ijlp.2019.101498, 1-11.

Gorjup, R., Makovec, T., 2021. Chemical imbalance and mental health. Journal of Clinical Review & Case Reports 6 (9), 737–740. https://doi.org/10.33140/JCRC.

Graham, S., Depp, C., Lee, E.E., Nebeker, C., Tu, X., Kim, H., Jeste, D.V., 2019. Artificial intelligence for mental health and mental illnesses: an overview. Curr. Psychiatr. Rep. 21 (116), 1–18. https://doi.org/10.1007/s11920-019-1094-0.

Gültekin, M., 2022. Could robots empatize? A review on the employment of social robots in mental healthcare. Türk Psikolojik Danışma ve Rehberlik Dergisi 12 (67), 593–618. https://doi.org/10.17066/tpdrd.12250405.

Hahn, T., Nierenberg, A.A., Whitfield-Gabrieli, S., 2017. Predictive analytics in mental health: applications, guidelines, challenges and perspectives. Mol. Psychiatr. 22, 37–43. https://doi.org/10.1038/mp.2016.201.

Hem, M.H., Gjerberg, E., Husum, T.L., Pedersen, R., 2018. Ethical challenges when using coercion in mental healthcare: a systematic literature review. Nurs. Ethics 25 (1), 92–110. https://doi.org/10.1177/0969733016629770.

Hindley, G., Smeland, O.B., Frei, O., Andreassen, O.A., 2022. Big data and the goal of personalized health interventions. In: Stein, D.J., Fineberg, N.A., Chamberlain, S.R. (Eds.), Mental Health in a Digital World. Academic Press, pp. 41–61. https://doi.org/10.1016/B978-0-12-822201-0.00021-6.

Kaczmarek, E., 2019. How to distinguish medicalization from over-medicalization? Med. Healthc. Philos. 22, 119–128. https://doi.org/10.1007/s11019-018-9850-1.

Koh, J., Tng, G.Y.Q., Hartanto, A., 2022. Potential and pitfalls of mobile mental health apps in traditional treatment: an umbrella review. J. Personalized Med. 12 (9), 1376. https://doi.org/10.3390/jpm12091376.

Kolenik, T., Gams, M., 2021. Persuasive technology for mental health: one step closer to (mental health care) equality? IEEE Technol. Soc. Mag. 40 (1), 80–86. https://doi.org/10.1109/MTS.2021.3056288. March.

Kozelka, E.E., Jenkins, J.H., Carpenter-Song, E., 2021. 'Advancing health equity in digital mental health: lessons from medical anthropology for global mental health', JMIR Mental Health, 8(8), e28555, pp.1-7. doi: 10.2196/28555.

Lagan, S., Ramakrishnan, A., Lamont, E., Ramakrishnan, A., Frye, M., Torous, J., 2020. Digital health developments and drawbacks: a review and analysis of top-returned apps for bipolar disorder. Int. J. Behav. Dev. 8 (39), 1–8. https://doi.org/10.1186/s40345-020-00202-4.

Lebowitz, M.S., 2019. Implications of genetic and other biological explanations for thinking about mental disorders. Hastings Cent. Rep. 49 (Suppl. 1), S82–S87. https://doi.org/10.1002/hast.1020.

Liang, Y., Zheng, X., Zeng, D.D., 2019. A survey on big data-driven digital phenotyping of mental health. Inf. Fusion 52, 290–307. https://doi.org/10.1016/j.inffus.2019.04.001.

Liu, X., Guo, Y., Zhang, W., Gao, W., 2022. Influencing factors, prediction and prevention of depression in college students: a literature review. World J. Psychiatr. 12 (7), 860–873. https://doi.org/10.5498/wjp.v12.i7.860.

Luxton, D.D., 2014. Artificial intelligence in psychological practice: current and future applications and implications. Prof. Psychol. Res. Pract. 45 (5), 332–339. https://doi.org/10.1037/a0034559.

Luxton, D.D., Anderson, S.L., Anderson, M., 2016. Chapter 11 - ethical issues and artificial intelligence technologies in behavioral and mental health care. In: Luxton, D.D. (Ed.), Artificial Intelligence in Behavioral and Mental Health Care. Academic Press, pp. 255–276. https://doi.org/10.1016/B978-0-12-420248-1.00011-8.

Ma, A., Patitsas, E., Sterne, J., 2023. You sound depressed: a case study on sonde health's diagnostic use of voice analysis AI. In: FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, pp. 639–650. https://doi.org/10.1145/3593013.3594032.

Malhotra, A., Jindal, R., 2022. Deep learning techniques for suicide and depression detection from online social media: a scoping review. Appl. Soft Comput. 130, 109713. https://doi.org/10.1016/j.asoc.2022.109713, 1-37.

McAndrew, L.M., Martin, J.L., Friedlander, M., Shaffer, K., Breland, J., Slotkin, S., Leventhal, H., 2018. The common sense of counseling psychology: introducing the Common-Sense Model of self- regulation. Counsell. Psychol. Q. 31 (4), 497–512. https://doi.org/10.1080/09515070.2017.1336076.

McCradden, M., Hui, K., Buchman, D.Z., 2022. Evidence, ethics and the promise of artificial intelligence in psychiatry. J. Med. Ethics 49 (8), 573–579. https://doi.org/10.1136/jme-2022-108447.

Milne-Ives, M., Selby, E., Inkster, B., Lam, C., Meinert, E., 2022. Artificial intelligence and machine learning in mobile apps for mental health: a scoping review. PLOS Digit Health 1 (8), e0000079. https://doi.org/10.1371/journal.pdig.0000079, 1-13.
Moncrieff, J., 2022. The political economy of the mental health system: a marxist analysis. Frontiers in Sociology 6, 1–11. https://doi.org/10.3389/fsoc.2021.771875.
Monteith, S., Glenn, T., 2016. Automated decision-making and Big Data: concerns for people with mental illness. Curr. Psych. Rep. 18 (12), 112. https://doi.org/10.1007/s11920-016-0746-6.
Nazroo, J.Y., Bhui, K.S., Rhodes, J., 2019. Where next for understanding race/ethnic inequalities in severe mental illness? Structural, interpersonal and institutional racism. Sociol. Health Illness 42 (2), 262–276. https://doi.org/10.1111/1467-9566.13001.
Nebeker, C., Torous, J., Ellis, R.J.B., 2019. Building the case for actionable ethics in digital health research supported by artificial intelligence. BMC Med. 17 (137), 1–7. https://doi.org/10.1186/s12916-019-1377-7.
Neuman, Y., 2016. Chapter 9 - public health surveillance: predictive analytics and big data. In: Luxton, D.D. (Ed.), Artificial Intelligence in Behavioral and Mental Health Care. Academic Press, pp. 255–276. https://doi.org/10.1016/B978-0-12-420248-1.00011-8.
Nogueira-Leite, D., Diniz, J.M., Cruz-Correia, R., 2023. Mental health professionals' attitudes toward digital mental health apps and implications for adoption in Portugal: mixed methods study. JMIR Human Factors 10, e45949. https://doi.org/10.2196/45949.
Oakey-Neate, L., Schrader, G., Strobel, J., Bastiampillai, T., van Kasteren, Y., Bidargaddi, N., 2020. Using algorithms to initiate needs-based interventions for people on antipsychotic medication: implementation protocol. BMJ Health & Care Informatics 27, e100084. https://doi.org/10.1136/bmjhci-2019-100084.
Omarov, B., Narynov, S., Zhumanov, Z., 2023. Artificial intelligence-enabled chatbots in mental health: a systematic review. Comput. Mater. Continua (CMC) 74 (3), 5105–5122. https://doi.org/10.32604/cmc.2023.034655.
Onnela, J., 2021. Opportunities and challenges in the collection and analysis of digital phenotyping data. Neuropsychopharmacology 46 (1), 45–54. https://doi.org/10.1038/s41386-020-0771-3.
Paré, G., Trudel, M., Jaana, M., Kitsiou, S., 2015. Synthesizing information systems knowledge: a typology of literature reviews. Inf. Manag. 52 (2), 183–199. https://doi.org/10.1016/j.im.2014.08.008.
Parker, L., Bero, L., Gillies, D., Raven, M., Mintzes, B., Jureidini, J., Grundy, Q., 2018. Mental health messages in prominent mental health apps. Ann. Fam. Med. 16 (4), 338–342. https://doi.org/10.1370/afm.2260.
Parrish, E.M., Filip, T.F., Torous, J., Nebeker, C., Moore, R.C., Depp, C.A., 2022. Are mental health apps adequately equipped to handle users in crisis? Crisis 43 (4), 289–298. https://doi.org/10.1027/0227-5910/a000785.
Poulin, C., Thompson, P., Bryan, C., 2016. Chapter 10 - artificial intelligence in public health surveillance and research. In: Luxton, D.D. (Ed.), Artificial Intelligence in Behavioral and Mental Health Care. Academic Press, pp. 205–230. https://doi.org/10.1016/B978-0-12-420248-1.00011-8.
Radovic, A., Vona, P.L., Santostefano, A.M., Ciaravino, S., Miller, E., Stein, B.D., 2016. Smartphone applications for mental health. Cyberpsychol., Behav. Soc. Netw. 19 (7), 465–470. https://doi.org/10.1089/cyber.2015.0619.
Rauseo-Ricupero, N., Henson, P., Agate-Mays, M., Torous, J., 2021. Case studies from the digital clinic: integrating digital phenotyping and clinical practice into today's world. Int. Rev. Psychiatr. 33 (4), 394–403. https://doi.org/10.1080/09540261.2020.1859465.
Rickard, N.S., Kurt, P., Meade, T., 2022. Systematic assessment of the quality and integrity of popular mental health smartphone apps using the American Psychiatric Association's app evaluation model. Frontiers in Digital Health 29 (4), 1003181. https://doi.org/10.3389/fdgth.2022.1003181, 1-13.
Ritchie, J., Spencer, L., 1994. Chapter 9 - qualitative data analysis for applied policy research. In: Bryman, A., Burgess, R.G. (Eds.), Analyzing Qualitative Data. Routledge, pp. 173–194.
Robinson, J., Witt, K., Lamblin, M., Spittal, M.J., Carter, G., Verspoor, K., Page, A., et al., 2020. Development of a self-harm monitoring system for victoria. Int. J. Environ. Res. Publ. Health 17 (24), 1–12. https://doi.org/10.3390/ijerph17249385, 9385.
Rocheteau, E., 2023. On the role of artificial intelligence in psychiatry. Br. J. Psychiatr. 222, 54–57. https://doi.org/10.1192/bjp.2022.132.
Rosenfeld, A., Benrimoh, D., Armstrong, C., Mirchi, N., Langlois-Therrien, T., Rollins, C., Tanguay-Sela, M., et al., 2021. Big Data analytics and artificial intelligence in mental healthcare. In: Khanna, A., Gupta, D., Dey, N. (Eds.), Applications of Big Data in Healthcare. Academic Press, pp. 137–171. https://doi.org/10.1016/B978-0-12-820203-6.00001-1.
Rothstein, M.A., 2021. Big data, surveillance capitalism, and precision medicine: challenges for privacy. J. Law Med. Ethics 49 (4), 666–676. https://doi.org/10.1017/jme.2021.91.
Rubeis, G., 2022. iHealth: the ethics of artificial intelligence and big data in mental healthcare. Internet Interventions 28, 100518. https://doi.org/10.1016/j.invent.2022.100518, 1-6.
Rubeis, G., 2023. Liquid Health. Medicine in the age of surveillance capitalism. Soc. Sci. Med. 322 (115810), 1–8. https://doi.org/10.1016/j.socscimed.2023.115810.
Russo, J., 2023. Psychiatrization, assertions of epistemic justice, and the question of agency. Frontiers in Sociology 8, 1–6. https://doi.org/10.3389/fsoc.2023.1092298.
Schmidt, S., D'Alfonso, S., 2023. Clinician perspectives on how digital phenotyping can inform client treatment. Acta Psychol. 235, 103886. https://doi.org/10.1016/j.actpsy.2023.103886, 1-12.
Sequeira, L., Perrottaa, S., LaGrassaa, J., Merikangasc, K., Kreindlerd, D., Kundurf, D., Courtney, D., et al., 2020. Mobile and wearable technology for monitoring depressive symptoms in children and adolescents: a scoping review. J. Affect. Disord. 265, 314–324. https://doi.org/10.1016/j.jad.2019.11.156.
Shatte, A.B.R., Hutchinson, D.M., Teague, S.J., 2019. Machine learning in mental health: a scoping review of methods and applications. Psychol. Med. 49, 1426–1448. https://doi.org/10.1017/S0033291719000151.
Skorburg, J.A., Yam, J., 2021. Is there an app for that?: ethical issues in the digital mental health response to COVID-19. AJOB Neuroscience 13 (3), 177–190. https://doi.org/10.1080/21507740.2021.1918284.
Starke, G., De Clercq, E., Borgwardt, S., Elger, B.S., 2021. Computing schizophrenia: ethical challenges for machine learning in psychiatry. Psychol. Med. 51 (15), 2515–2521. https://doi.org/10.1017/S0033291720001683.
Stern, E., Franchi, J.M., Dumas, G., Moreira, J., Mouchabac, S., Maruani, J., Philip, P., et al., 2022. How can digital mental health enhance psychiatry? Neuroscientist 1–13. https://doi.org/10.1177/10738584221098603.
Suh, J., Pendse, S.R., Lewis, R., Howe, E., Saha, K., Okoli, E., et al., 2024. Rethinking technology innovation for mental health: framework for multi-sectoral collaboration. Nature Mental Health 2, 4780488. https://doi.org/10.1038/s44220-024-00232-2.
Sweet, P.L., Decoteaua, C.L., 2017. Contesting normal: the DSM-5 and psychiatric subjectivation. BioSocieties 13, 103–122. https://doi.org/10.1057/s41292-017-0056-1.
Tavory, T., 2024. Regulating AI in mental health: ethics of care perspective. JMIR Mental Health 11, e58493. https://doi.org/10.2196/58493.
Tekin, Ş., 2021. Is big data the new stethoscope? Perils of digital phenotyping to address mental illness. Philosophy & Technology 34, 447–461. https://doi.org/10.1007/s13347-020-00395-7.
Terra, M., Baklola, M., Ali, S., El-Bastawisy, K., 2023. Opportunities, applications, challenges and ethical implications of artificial intelligence in psychiatry: a narrative review. The Egyptian Journal of Neurology, Psychiatry and Neurosurgery 59 (80), 1–10. https://doi.org/10.1186/s41983-023-00681-z.
Torous, J., Haim, A., 2018. Dichotomies in the development and implementation of digital mental health tools. Psychiatr. Serv. 69 (12), 1204–1206. https://doi.org/10.1176/appi.ps.201800193.
Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., et al., 2018. PRISMA extension for scoping reviews (PRISMAScR): checklist and explanation. Ann. Intern. Med. 169, 467–473. https://doi.org/10.7326/M18-0850.
Uma, V., Banappagoudar, S.B., Nivethitha, S.S., Thangam, M.M.N., 2023. The implications of AI-based systems in the psychological management of post chronic stress disorder. Journal for ReAttach Therapy and Developmental Diversities 6 (5s), 278–284.
Uusitalo, S., Tuominen, J., Arstila, V., 2021. Mapping out the philosophical questions of AI and clinical practice in diagnosing and treating mental disorders. J. Eval. Clin. Pract. 27 (3), 478–484. https://doi.org/10.1111/jep.13485.
Van Assche, E.V., Ramos-Quiroga, J.A., Pariante, C.M., Sforzini, L., Young, A.H., Flossbach, Y., Gold, S.M., et al., 2022. Digital tools for the assessment of pharmacological treatment for depressive disorder: state of the art. Eur. Neuropsychopharmacol 60, 100–116. https://doi.org/10.1016/j.euroneuro.2022.05.007.
Wang, J., Mann, F., Lloyd-Evans, B., Ma, R., Johnson, S., 2018. Associations between loneliness and perceived social support and outcomes of mental health problems: a systematic review. BMC Psychiatr. 18 (156), 1–16. https://doi.org/10.1186/s12888-018-1736-5.
Wasil, A.R., Palermo, E.H., Lorenzo-Luaces, L., DeRubeis, R.J., 2022. Is there an app for that? A review of popular apps for depression, anxiety, and well-being. Cognit. Behav. Pract. 29 (4), 883–901. https://doi.org/10.1016/j.cbpra.2021.07.001.
Wies, B., Landers, C., Ienca, M., 2021. Digital mental health for young people: a scoping review of ethical promises and challenges. Frontiers in Digital Health 3, 1–11. https://doi.org/10.3389/fdgth.2021.697072.
Williams, J.E., Pykett, J., 2022. Mental health monitoring apps for depression and anxiety in children and young people: a scoping review and critical ecological analysis. Soc. Sci. Med. 297, 114802. https://doi.org/10.1016/j.socscimed.2022.114802.
Woodward, 2022. Beyond mobile apps: a survey of technologies for mental well-being. IEEE Transactions on Affective Computing 13 (3), 1216–1235. https://doi.org/10.1109/TAFFC.2020.3015018 et al.
World Bank, 2024. World Bank Country and Lending Groups. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. (Accessed 8 November 2024).
World Health Organization (WHO), 2022. World mental health report: transforming mental health for all. Available at: https://www.who.int/publications/i/item/9789240049338. (Accessed 29 October 2023).
Yu, T., Deng, K., Yao, Q., 2022. Application of artificial intelligence in the process of rehabilitation of mentally ill patients returning to society. In: ISAIMS '22: Proceedings of the 3rd International Symposium on Artificial Intelligence for Medicine Sciences, pp. 457–462. https://doi.org/10.1145/3570773.3570803.
Zeira, A., 2022. Mental health challenges related to neoliberal capitalism in the United States. Community Ment. Health J. 58 (2), 205–212. https://doi.org/10.1007/s10597-021-00840-7.
Zuboff, S., 2015. Big other: surveillance capitalism and the prospects of an information civilization. J. Inf. Technol. 30 (1), 75–89. https://doi.org/10.1057/jit.2015.5.