

# Accelerating the impact of artificial intelligence in mental healthcare through implementation science

Implementation Research and Practice  
Volume 3: Jan-Dec 2022 1–10

© The Author(s) 2022

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/26334895221112033

journals.sagepub.com/home/irp



Per Nilsen<sup>1</sup>, Petra Svedberg<sup>2</sup>, Jens Nygren<sup>2</sup>, Micael Frideros<sup>1</sup> , Jan Johansson<sup>3</sup>  
and Stephen Schueller<sup>4</sup>

## Abstract

### Background

The implementation of artificial intelligence (AI) in mental healthcare offers a potential solution to some of the problems associated with the availability, attractiveness, and accessibility of mental healthcare services. However, there are many knowledge gaps regarding how to implement and best use AI to add value to mental healthcare services, providers, and consumers. The aim of this paper is to identify challenges and opportunities for AI use in mental healthcare and to describe key insights from implementation science of potential relevance to understand and facilitate AI implementation in mental healthcare.

### Methods

The paper is based on a selective review of articles concerning AI in mental healthcare and implementation science.

### Results

Research in implementation science has established the importance of considering and planning for implementation from the start, the progression of implementation through different stages, and the appreciation of determinants at multiple levels. Determinant frameworks and implementation theories have been developed to understand and explain how different determinants impact on implementation. AI research should explore the relevance of these determinants for AI implementation. Implementation strategies to support AI implementation must address determinants specific to AI implementation in mental health. There might also be a need to develop new theoretical approaches or augment and recontextualize existing ones. Implementation outcomes may have to be adapted to be relevant in an AI implementation context.

### Conclusion

Knowledge derived from implementation science could provide an important starting point for research on implementation of AI in mental healthcare. This field has generated many insights and provides a broad range of theories, frameworks, and concepts that are likely relevant for this research. However, when taking advantage of the existing knowledge basis, it is important to also be explorative and study AI implementation in health and mental healthcare as a new phenomenon in its own right since implementing AI may differ in various ways from implementing evidence-based practices in terms of what implementation determinants, strategies, and outcomes are most relevant.

**Plain Language Summary:** The implementation of artificial intelligence (AI) in mental healthcare offers a potential solution to some of the problems associated with the availability, attractiveness, and accessibility of mental healthcare

<sup>1</sup>Linköping University, Linköping, Sweden

<sup>2</sup>Halmstad University School of Health and Welfare, Halmstad University, Halmstad, Sweden

<sup>3</sup>Region Halland, Halmstad, Sweden

<sup>4</sup>Psychological Science, University of California Irvine, Irvine, CA, USA

### Corresponding author:

Per Nilsen, Linköping University, Uvebergsv 47, Linköping 581 83, Sweden.

Email: per.nilsen@liu.se



services. However, there are many knowledge gaps concerning how to implement and best use AI to add value to mental healthcare services, providers, and consumers. This paper is based on a selective review of articles concerning AI in mental healthcare and implementation science, with the aim to identify challenges and opportunities for the use of AI in mental healthcare and describe key insights from implementation science of potential relevance to understand and facilitate AI implementation in mental healthcare. AI offers opportunities for identifying the patients most in need of care or the interventions that might be most appropriate for a given population or individual. AI also offers opportunities for supporting a more reliable diagnosis of psychiatric disorders and ongoing monitoring and tailoring during the course of treatment. However, AI implementation challenges exist at organizational/policy, individual, and technical levels, making it relevant to draw on implementation science knowledge for understanding and facilitating implementation of AI in mental healthcare. Knowledge derived from implementation science could provide an important starting point for research on AI implementation in mental healthcare. This field has generated many insights and provides a broad range of theories, frameworks, and concepts that are likely relevant for this research.

## Keywords

mental health services, implementation, artificial intelligence

Mental health disorders are prevalent in society and are a serious public health problem. In 2001, the World Health Organization estimated that 450 million people worldwide had such problems (World Health Organization, 2001). Approximately 75% of mental disorders emerge before the age of 25 years (Jones, 2013), and it is expected that 29% of people will experience a mental disorder at least once in their lifetime (Barrett et al., 2017). Mental illness restricts individuals' abilities to function, engage in daily activities, and maintain social relationships, causing significant suffering to individuals and their families (Fricchione et al., 2012; Rosenfeld et al., 2021).

Despite the prevalence and severity of mental illness, there are many barriers to achieving optimal detection, prevention, treatment, and monitoring of mental health disorders (Schueller & Torous, 2020). Availability of mental healthcare remains limited. Many people with mental illnesses receive no mental healthcare, and even those who do rarely receive evidence-based mental healthcare (Alegría et al., 2018; Kilbourne et al., 2018). There is a shortage of psychiatrists, psychologists, nurses, and social workers (Fricchione et al., 2012; Liu et al., 2017); almost half of the world's population live in countries where there is less than one psychiatrist per 100,000 people (World Health Organization, 2017). This shortage of mental healthcare professionals restricts care opportunities, resulting in long waiting lists, delays, and multiple, diverse contacts before appropriate care is obtained (MacDonald et al., 2018; Westberg et al., 2020).

Even when providers exist, studies have documented accessibility barriers, including concerns about the cost, transportation or inconvenience, problems recognizing mental illness symptoms (i.e., poor mental health literacy), and lack of awareness of care options (Gulliver et al., 2010; Pedersen & Paves, 2014). Attitudinal barriers also prevent people from seeking mental healthcare, including the belief that it will not help and a preference for self-reliance.

Public stigma against mental health disorders and those who receive care is a common concern that also functions as a barrier (Gulliver et al., 2010; Wainberg et al., 2017). As such, even where care is hypothetically available, few in need of such care actually receive it.

The implementation of artificial intelligence (AI) in mental healthcare offers a potential solution to some of the problems with the availability, attractiveness, and accessibility of mental healthcare services (Bickman, 2020; European Institute of Innovation and Technology, 2020; Lee, 2019; Topol, 2019). AI generally refers to a computerized system (hardware or software) that is able to perform tasks or reasoning processes that we usually associate with intelligence in a human being (European Union, 2018). AI applied to mental healthcare includes techniques to improve the ability to detect and predict various conditions, with implications for screening, assessment, and clinical decision-making (Graham et al., 2019). AI can also be used to improve treatments by identifying treatments or treatment elements most likely to provide benefit, thus promoting personalized mental healthcare (Bickman, 2020), or can be incorporated into digital interventions, such as smartphone apps, to create novel interventions, thus reducing reliance on professionals and increasing scalability (D'Alfonso, 2020).

While AI applications have become increasingly commonplace in many areas of business and society, the potential of using AI in mental healthcare has not been realized to date. Still, there are examples of AI applications that have been implemented and demonstrated the promise of AI in mental health practice, such as leveraging existing data streams or creating new interventions (Lee et al., 2021) and platforms to enable therapists to receive data-led insights about each patient during the assessment stage of treatment (Ieso Health, 2022). An example is the implementation of a suicide prevention AI application, Recovery Engagement, and Coordination for Health-Veterans

Enhanced Treatment (REACH VET) intended to identify individuals in the Veterans Health Administration electronic health records most likely to die from suicide. Using an interactive dashboard, program coordinators receive information and communicate with clinicians who re-evaluate treatment strategies and conduct outreach to initiate care enhancements (McCarthy et al., 2021). There are also numerous mental health self-assessment AI applications available, e.g., the Ellie digital avatar that was initially developed to help war veterans struggling with depression and Post Traumatic Stress Syndrome, the BioBase app that leverages AI to interpret sensor data from a wearable as well as apps like Woebot and Elomia that target individuals with mild symptoms of anxiety or depression (Lavrentyeva, 2021; Raibagi, 2020). AI mental health applications such as these have produced considerable enthusiasm for development and implementation, including a substantial investment in mental health startups using AI (Shah & Berry, 2021) and pilots in healthcare systems to consider opportunities (Universal Health Services, 2020). Despite examples of AI use in mental healthcare practice, Davenport and Kalakota (2019) argue that the greatest challenge to AI “is not whether the technologies will be capable enough to be useful, but rather ensuring their adoption in daily clinical practice.” Seneviratne et al. (2019) suggest that AI implementation in healthcare is “the elephant in the room.”

Despite the promise of AI applications in mental healthcare, considerable knowledge gaps still exist, including how to best implement and use these innovations to add value to healthcare services, providers, and consumers. Implementation science knowledge may be useful for understanding and addressing the challenges of implementing AI in mental healthcare. The aim of this paper is to review the extant literature to identify challenges and opportunities for the use of AI in mental healthcare and describe key insights from implementation science of potential relevance to understand and facilitate AI implementation in mental healthcare. To this end, we conducted a selective literature review of articles concerning AI in mental healthcare and implementation science.

## AI in Mental Healthcare

Mental healthcare differs in many ways from medicine, which contributes to the need to consider unique challenges and opportunities related to its implementation. Unlike the diagnosis of a physical condition that can be based on laboratory tests or physiologic measurements, diagnoses of mental illness typically rely on patients’ self-reported information and mental health professionals’ judgment (Rosenfeld et al., 2021; Su et al., 2020). Mental health professionals emphasize the importance of the relational and observational aspects in therapeutic practice, such as forming a therapeutic alliance with the patient and directly observing patient behaviors and emotions

(Graham et al., 2019). A multi-country survey by Doraiswamy et al. (2020) showed that mental health professionals believed that AI can be used for documenting and synthesizing information, but they did not think it can replace the interaction between patient and professional. However, there is considerable optimism among AI and mental health researchers that mental healthcare can benefit greatly from the use of AI (Barrett et al., 2017; D’Alfonso, 2020; Graham et al., 2019; Mohr et al., 2017; Rosenfeld et al., 2021; Su et al., 2020; Topol, 2019).

AI has a potentially important role in improving our understanding of mental illness to achieve improved detection, prevention, treatment, and monitoring of mental health disorders (Graham et al., 2019; Lee et al., 2021; Su et al., 2020). In these areas, the application of AI in mental health can broadly be classified into opportunities for selection and for assessment.

## Opportunities for Selection

Opportunities for selection involve identifying the patients most in need of care or the interventions that might be most appropriate for a given population or individual. In the service of selection, AI applications based on data from large populations can help identify previously unidentified correlations. Such models can help detect unobservable mental health states, analyze complex problems, and monitor trends and thresholds of key mental health indicators. As examples, AI applied to social media data can detect individuals who are at higher risk for suicide (Coppersmith et al., 2018). AI has also been used to determine who might be most likely to benefit from cognitive behavioral or psychodynamic therapy in routine healthcare settings (Schwartz et al., 2021).

## Opportunities for Assessment

Opportunities for assessment include supporting a more reliable diagnosis of psychiatric disorders and ongoing monitoring and tailoring during the course of treatment to maximize opportunities for response. In the service of assessment, AI offers opportunities to combine bio-psycho-social data to create novel assessment streams such as the determination of mental health states through passive smartphone data (Mohr et al., 2017). Applied to evaluation of interventions, AI can support understanding of what treatment elements most have an impact on improvement (Ewbank et al., 2020) and might assist in monitoring elements during therapy, such as emotion (Tanana et al., 2021) or symptom change that might facilitate better measurement-based care (Huckvale et al., 2019) or fidelity to evidence-based practices in ways that can help clinicians improve or adapt their techniques (Flemotomos et al., 2021). In the service of these multiple areas, AI offers the opportunity to use novel or larger data streams which can facilitate a more reliable diagnosis of psychiatric

disorders and support patients and providers in making more informed shared decisions as part of a more person-centered care (Aafjes-van Doorn et al., 2021; Graham et al., 2019; Su et al., 2020).

## Implementation of AI in Mental Healthcare

Much AI research and development (R&D) has followed a generic R&D pipeline, from ideas and hypotheses to AI applications used in routine practice, with progress varying between areas of AI use. The pipeline begins with hypotheses and in silico experiments (i.e., performed on a computer or via computer simulation), progressing to proof-of-concept (prototype) projects to document the feasibility of the AI application, before studies are carried out to evaluate the efficacy and effectiveness of the application when it is deployed in practice. The final step of this pipeline is implementation of the application as standard practice in routine health and mental healthcare (Stead, 2018).

It is firmly established in implementation and innovation research that novel technologies are often resisted by healthcare professionals, contributing to a slow and variable uptake (Safi et al., 2018; Whitelaw et al., 2020). Several types of implementation challenges (i.e., barriers and facilitators) at different levels have been documented in AI research. Here, we provide an overview of challenges at different levels: organizational and policy, providers, patients, and technical. The challenges are likely interdependent, making it difficult to assess the relative importance of various challenges or identify which challenges might be most difficult to overcome. However, dividing the challenges into these levels provides an opportunity for implementation frameworks to support understanding and addressing them given the considerations of such levels in many such frameworks.

### Organizational and Policy Challenges

There are many organization- and policy-level challenges to AI implementation in healthcare. Regulatory concerns include governance of autonomous AI systems, responsibility and accountability issues, paucity of industry standards for use of AI use and assessment of safety, and inadequate or non-existing clinical and economic impact measurement (Buch et al., 2018; Esmailzadeh, 2020; Horgan et al., 2019; Liyanage et al., 2019). The literature does not expound on appropriate strategies to address various barriers, but He et al. (2019) recommend the use of specific task force committees in healthcare organizations to work specifically with AI issues.

A major hurdle with regard to the organization and policy levels is the lack of clear and consistent guidelines regarding the application of AI in healthcare generally,

and mental healthcare consistently. In the United States, regulatory issues with regard to AI in healthcare are still being defined by the U.S. Food and Drug Administration. The European Union has developed Ethical Guidelines for Trustworthy AI and is now working on turning these into law. In the proposed legislation, most healthcare applications will likely be considered high-risk applications, with significant regulation associated (European Union, 2019). In other countries, standards for the use of digital mental health products (Brown et al., 2021) have not addressed the additional complications that AI applications bring. Therefore, given the current lack of regulation, new shifts in regulatory guidelines could have considerable impact on implementation and organization- and policy-level challenges.

### Challenges for Providers

At the individual level, professionals must have access to education and training to learn new skills as AI users. They need tools to be able to train AI systems to set them up to perform specified tasks, and they must be able to understand and interpret AI outputs (Buch et al., 2018; He et al., 2019; Mistry, 2019). Integrating AI into clinical practice will require professionals to adapt their practice and actively engage with AI to achieve “a partnership where the machine predicts, and the human explains and decides on action” (Verghese et al., 2018). However, research has documented considerable skepticism among healthcare professionals on using AI (Lee, 2019; Liyanage et al., 2019). Reflecting on the AI implementation challenges, Kelly et al. (2019) concluded that “human barriers to adoption [of AI] are substantial.”

### Challenges for Patients

There may also be mistrust among patients in mental healthcare, with concerns about technical aspects (e.g., communication barriers), ethical issues (e.g., privacy), and regulatory aspects (e.g., unregulated standards) (Esmailzadeh, 2020). These concerns relate to AI in general, but they might be even more pronounced in mental healthcare. For example, psychiatric diagnoses are defined by criteria related to daily life and are formulated in ordinary language that is often easy for patients to comprehend. However, it may be more difficult to understand and accept that signs indicating, for example, depression, suicidal ideation, or social anxiety, are identified by AI through analysis of data contained in a patient’s medical record (Esmailzadeh, 2020), smartphone (Gong et al., 2019) or on social media (Lekkas et al., 2021), potentially leading to communication barriers between patient and mental healthcare professionals. Patients might also experience various privacy concerns because the use of AI often requires substantial datasets to provide insights (Uusitalo et al., 2021). Patients might be more willing to provide

mental health data based on the different proposed uses or recipients of that data (Nicholas et al., 2019), but in many applications of AI, the ultimate use of that data may not be known when it is first collected.

### Technical Challenges

Technical challenges to AI implementation include the lack of transparency of data and algorithms, safety of AI recommendations, and complexities in interpreting outputs and dealing with responsibilities for the outcomes of decisions based on AI-generated information. Concerns have been raised about data being too fragmented and siloed in many systems that can be difficult to bring together. Data may also be unusable due to a lack of data that meets precise requirements or is sufficiently labeled. Thus, there may paradoxically be both too much data (i.e., datasets are large but lack meaningful information or sufficient labels) and too little data (i.e., data that are not appropriate or sufficient for specified purposes). Just because a dataset is large does not mean it will be applicable to a given problem, especially as too much similar data might not provide much additional value. Furthermore, when Electronic Health Record data are used, mental health data tend to be less structured and therefore more messy (Dawoodbhoy et al., 2021). When datasets are small, they may not be representative and fail to generalize when applied more broadly. This might introduce algorithm bias and models that fail to predict in specific populations (Brodwin & Ross, 2021; Coley et al., 2021).

Furthermore, software must be continuously updated and evaluated with regard to applicability for datasets that might vary from the data on which models have been trained and validated. Also, sufficient hardware to support any provider interface with models is necessary (Buch et al., 2018; Horgan et al., 2019; Keane & Topol, 2018; Kelly et al., 2019). Given that mental healthcare settings are under resourced, sufficient technical infrastructure may not be present.

Methods used to capture the data need to be sufficiently standardized and robust to prevent negative effects from handling or data management (Geirhos et al., 2020). Another technical challenge is that the data used to train an AI application must be relevant for the actual application. For example, if an AI system is used to interpret facial expressions, it needs to be trained on examples representative for the actual target group because social communication might differ depending on factors such as age, culture, or region (Kim et al., 2019). Some of the technical challenges may be particularly relevant for mental healthcare. AI algorithms based on neural networks require many different examples, hundreds or thousands, to learn a particular diagnosis or treatment (Cho et al., 2016). This means that an AI system based on neural networks will often struggle when confronted with a rare, albeit

distinct and well described, diagnosis. For a well-trained human, on the other hand, one rich example described in research might suffice.

### Using Implementation Science to Understand Implementation of AI

In recognition of the challenges in implementing AI, it may be important to draw on implementation science knowledge to understand and facilitate implementation of AI in mental healthcare. Research in implementation science has generated many insights concerning implementation in various healthcare settings, establishing the importance of considering and planning for implementation from the start, the progression of implementation through different stages, and the appreciation of influences from multiple levels of the healthcare system and beyond (Nilsen, 2015). So-called hybrid designs are used in implementation science to accelerate the research-to-practice process. This design allows researchers to investigate the effects of both a clinical (patient-level) intervention and the implementation of this intervention (Curran et al., 2012). This study design may be relevant to investigate whether or the extent to which various AI applications warrant implementation in mental health.

Numerous determinant frameworks and implementation theories have been developed in implementation science to understand and explain how different barriers and facilitators impact on implementation (Nilsen, 2015), e.g., Consolidated Framework of Implementation Research (Damschroder et al., 2009), Promoting Action on Research in Health Services (PARIHS) and integrated-PARIHS (Harvey & Kitson, 2016), Exploration Preparation Implementation Sustainability (Aarons et al., 2011), Theoretical Domains Framework (Michie et al., 2014), Capability Opportunity Motivation Behaviour (COM-B) (Michie et al., 2011), and Normalization Process Theory (May & Finch, 2009), all of which are widely used in the field. Knowledge about determinants of implementation outcomes provides input for selecting the most effective strategies to overcome barriers and/or harness facilitators. Together these frameworks and theories describe five interdependent domains of influence on implementation: (1) effectiveness of the strategy used to support implementation; (2) attributes of the implemented practice (i.e. an intervention, programme, or service), e.g., its perceived complexity and compatibility with previous practices; (3) features of the adopters, e.g., the providers' attitudes, beliefs, and motivation concerning the implemented practice; (4) features of the patients or recipients of the implemented practice, e.g., their preferences; and (5) contextual influences, e.g., the culture, leadership, and other collective-level influences on the adopters (Nilsen, 2015).

AI implementation research should explore the extent to which the barriers and facilitators to AI implementation overlap with those identified in studies of implementation

of evidence-based practices. Furthermore, because AI is not one single technology but rather many different ones (Davenport & Kalakota, 2019; Shaw et al., 2019), it is also important to investigate different AI applications with regard to these barriers and facilitators. Studies of different applications will allow assessment of the extent to which there are certain determinants that are more important than others. It is also important to investigate the extent to which there are more application-specific barriers and facilitators and whether there are determinants that are more broadly generalizable. Few of the frameworks used in implementation science have yet been applied or “tested” in AI research with regard to their applicability. Determinant frameworks and implementation theories are multi-level, which is important because the implementation of AI in mental healthcare is a socio-technical system that requires conceptualizing the unique, and sometimes asymmetric, contributions of human and technology elements (Holton & Boyd, 2021). Furthermore, applications of AI in mental healthcare can often circumvent traditional workflow and care delivery pathways (Hermes et al., 2019), requiring consideration of how this disruption in traditional care processes necessitates systems-level thinking with regard to implementation.

Implementation science, thus far, has most often focused on implementation of various individual evidence-based practices, typically health interventions with empirical support for their efficacy or effectiveness (Nilsen & Birken, 2020). The implementation and routine use of such practices may require smaller, more incremental changes to professionals’ existing practices than when implementing and deploying a new AI application. AI, on the other hand, is a discontinuous and more disruptive form of change (Scott et al., 2000), likely requiring considerable professional and organizational learning to achieve optimal ways of working.

Implementation strategies to support AI implementation must address barriers specific to implementation in mental health (Graham et al., 2019). Implementation strategies constitute the “how-to” component of changing practice, being methods and techniques that are used to facilitate implementation of evidence-based practices (Curran et al., 2012). Implementation science has documented limited effectiveness for many strategies, which has been attributed to a paucity of the use of theory in the field. However, it could also be that many strategies have been based on inappropriate assumptions about how to achieve practice change. After all, theories are assumptions about a phenomenon, which means that the explanatory power of a given theory depends on the extent to which the assumptions underpinning the theory provide an accurate or plausible explanation of how current practice can be changed (Moore & Evans, 2017). Research is needed to investigate whether the current conceptualization of implementation strategies is sufficient to develop insights and provide the recommendations necessary to inform AI

implementation. More careful definition of strategies could be needed, including not only their name, but also the actor and action targets (Leeman et al., 2017). In practice, many implementations involve not one implementation strategy but a bundle of implementation strategies. Given the socio-technical systems involved in AI implementation, strategy bundles are often necessary. Thus, our thinking on implementation strategies needs to adjust to mirror the complexity of AI implementation in mental healthcare.

Implementation outcomes used in implementation science have been defined as the effects of deliberate and purposive actions to implement evidence-based practices (Powell et al., 2012). A range of outcomes has been described: acceptability (is the practice agreeable or satisfactory?); appropriateness (is the practice compatible with the setting, providers or patients?); feasibility (can the practice be used or carried out within the setting?); fidelity (was the practice implemented as intended?); adoption (was there an initial decision or action to try or employ the practice?); sustainability (was the practice maintained or institutionalized within the setting?); penetration (was the practice integrated within the setting?); and cost (what was the cost of implementing the practice?) (Proctor et al., 2010). The concept of outcomes in implementation research is distinct from service system outcomes and clinical patient outcomes. The relevance of various outcomes for AI implementation research needs to be explored (Hermes et al., 2019).

There might also be a need to develop new theoretical approaches or augment and re-contextualize existing ones from implementation science. Greenhalgh et al. (2017) recognized that the present theoretical approaches in implementation science were inadequate to provide satisfactory explanations of the challenges involved in going from small-scale proof-of-concept projects to implementation of new technologies in healthcare. They developed the Non-adoption, Abandonment, Scale-up, Spread, Sustainability framework, which acknowledges the complexity of the environment in which technologies are introduced and accounts for many collective-level influences, including characteristics of the adopter system (e.g., changes in staff roles), the organization’s readiness and capacity to innovate, and the wider context.

Despite the obvious potential of using knowledge generated within the field of implementation science to guide AI implementation, a scoping review by Gama et al. (2021) found only seven studies that described various aspects of AI implementation in different healthcare settings. Most of the 2,541 unique articles identified in the searches concerned studies that were not conducted in healthcare settings and/or focused on the earlier steps of the R&D pipeline, e.g., algorithm development, proof-of-concept projects, and AI testing in the laboratory. The authors concluded that “understanding of how to implement AI technology in healthcare practice is still at its early stages of development” (Gama et al., 2021).

In summary, knowledge derived from implementation science could provide an important starting point for research on AI implementation in mental healthcare. This field has generated many insights and provides a broad range of theories, frameworks, and concepts that are likely relevant for the study of AI implementation in mental healthcare. However, when taking advantage of the existing knowledge basis, it is important to also be explorative and study AI implementation in health and mental healthcare as a new phenomenon in its own right since implementing AI may differ in various ways from implementing evidence-based practices in terms of what implementation determinants, strategies, and outcomes are most relevant. Implementation researchers are used to interdisciplinary work and will benefit from collaborating with healthcare professionals, AI developers, and recipients of mental healthcare to generate conceptual, instrumental, and strategic knowledge that can contribute toward realizing the expectations of AI in mental healthcare.

### Acknowledgments

The authors thank Fredrik Heintz for valuable comments on the technology of AI.


### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iD

Micael Frideros  <https://orcid.org/0000-0003-4241-0270>

### References

- Aafjes-van Doorn, K., Kamsteeg, C., Bate, J., & Aafjes, M. (2021). A scoping review of machine learning in psychotherapy research. *Psychotherapy Research*, 31(1), 92–116. <https://doi.org/10.1080/10503307.2020.1808729>
- Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research*, 38, 4–23. <https://doi.org/10.1007/s10488-010-0327-7>
- Alegria, M., Nakash, O., & NeMoyer, A. (2018). Increasing equity in access to mental health care: A critical first step in improving service quality. *World Psychiatry*, 17(1), 43–44. <https://doi.org/10.1002/wps.20486>
- Barrett, P. M., Steinhubl, S. R., Muse, E. D., & Topol, E. (2017). Digital medicine: Digitising the mind. *The Lancet*, 389, 1877. [https://doi.org/10.1016/S0140-6736\(17\)31218-7](https://doi.org/10.1016/S0140-6736(17)31218-7)
- Bickman, L. (2020). Improving mental health services: A 50-year journey from randomized experiments to artificial intelligence and precision mental health. *Administration and Policy in Mental Health*, 47(5), 795–843. <https://doi.org/10.1007/s10488-020-01065-8>
- Brodwin, E., & Ross, C. (2021). Promise and peril: How artificial intelligence is transforming health care. *STAT*. [https://www.statnews.com/wp-content/uploads/2021/04/STAT\\_Promise\\_and\\_Peril\\_2021\\_Report.pdf](https://www.statnews.com/wp-content/uploads/2021/04/STAT_Promise_and_Peril_2021_Report.pdf)
- Brown, P., Prest, B., Miles, P., & Rossi, V. (2021). The development of national safety and quality digital mental health standards. *Australasian Psychiatry*, 10398562211042361. Advance online publication. <https://doi.org/10.1177/10398562211042361>
- Buch, V. H., Ahmed, I., & Maruthappu, M. (2018). Artificial intelligence in medicine: Current trends and future possibilities. *British Journal of General Practice*, 68(668), 143–144. <https://doi.org/10.3399/bjgp18X695213>
- Cho, J., Lee, K., Shin, E., Choy, G., & Do, S. (2016). *How much data is needed to train a medical image deep learning system to achieve necessary high accuracy?* arXiv. <https://arxiv.org/pdf/1511.06348.pdf>
- Coley, R. Y., Johnson, E., Simon, G. E., Cruz, M., & Shortreed, S. M. (2021). Racial/ethnic disparities in the performance of prediction models for death by suicide after mental health visits. *JAMA Psychiatry*, 78(7), 726–734. <https://doi.org/10.1001/jamapsychiatry.2021.0493>
- Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2018). Natural language processing of social media as screening for suicide risk. *Biomedical Informatics Insights*, 10, 1178222618792860. <https://doi.org/10.1177/1178222618792860>
- Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical Care*, 50, 217–226. <https://doi.org/10.1097/MLR.0b013e3182408812>
- 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>
- Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, 4, 50. <https://doi.org/10.1186/1748-5908-4-50>
- Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2), 94–98. <https://doi.org/10.7861/futurehosp.6-2-94>
- Dawoodbhoy, F. M., Delaney, J., Cecula, P., Yu, J., Peacock, I., Tan, J., & Cox, B. (2021). AI in patient flow: Applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units. *Heliyon*, 7(5), e06993. <https://doi.org/10.1016/j.heliyon.2021.e06993>
- Doraiswamy, P. M., Blease, C., & Bodner, K. (2020). Artificial intelligence and the future of psychiatry: Insights from a global physician survey. *Artificial Intelligence in Medicine*, 102, 101753. <https://doi.org/10.1016/j.artmed.2019.101753>
- Esmailzadeh, P. (2020). Use of AI-based tools for healthcare purposes: A survey study from consumers' perspectives. *BMC Medical Informatics and Decision Making*, 20(1), 170. <https://doi.org/10.1186/s12911-020-01191-1>
- European Institute of Innovation and Technology (2020). *Transforming healthcare with AI*. EIT Health and McKinsey & Company. <https://eithealth.eu/wp-content/uploads/2020/>

- 03/EIT-Health-and-McKinsey\_Transforming-Healthcare-with-AI.pdf
- European Union (2018). *A definition of AI: Main capabilities and scientific disciplines*. [https://ec.europa.eu/futurium/en/system/files/ged/ai\\_hleg\\_definition\\_of\\_ai\\_18\\_december\\_1.pdf](https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf)
- European Union (2019). *Ethics guidelines for trustworthy AI*. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Ewbank, M. P., Cummins, R., Tablan, V., Bateup, S., Catarino, A., Martin, A. J., & Blackwell, A. D. (2020). Quantifying the association between psychotherapy content and clinical outcomes using deep learning. *JAMA Psychiatry*, 77(1), 35–43. <https://doi.org/10.1001/jamapsychiatry.2019.2664>
- Flemotomos, N., Martinez, V. R., Chen, Z., Creed, T. A., Atkins, D. C., & Narayanan, S. (2021). Automated quality assessment of cognitive behavioral therapy sessions through highly contextualized language representations. *PloS one*, 16(10), e0258639. <https://doi.org/10.1371/journal.pone.0258639>
- Fricchione, G. L., Borba, C. P., Alem, A., Shibre, T., Carney, J. R., & Henderson, D. C. (2012). Capacity building in global mental health: Professional training. *Harvard Review of Psychiatry*, 20(1), 47–57. <https://doi.org/10.3109/10673229.2012.655211>
- Gama, F., Tyskbo, D., Nygren, J., Barlow, J., Reed, J., & Svedberg, P. (2021). *Implementation frameworks for AI translation into healthcare practice: A scoping review*. <https://doi.org/10.2196/preprints.32215>
- Geirhos, R., Medina Temme, C. R., Rauber, J., Schütt, H. H., Bethge, M., & Wichmann, F. A. (2020). *Generalisation in humans and deep neural networks*. arXiv. <https://arxiv.org/pdf/1808.08750.pdf>
- Gong, J., Huang, Y., Chow, P. I., Fua, K., Gerber, M. S., Teachman, B. A., & Barnes, L. E. (2019). Understanding behavioral dynamics of social anxiety among college students through smartphone sensors. *Information Fusion*, 49, 57–68. <https://doi.org/10.1016/j.inffus.2018.09.002>
- Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports*, 21(11), 116. <https://doi.org/10.1007/s11920-019-1094-0>
- Greenhalgh, T., Wherton, J., Papoutsis, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research*, 19(1), e367. <https://doi.org/10.2196/jmir.8775>
- Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry*, 10, 113. <https://doi.org/10.1186/1471-244X-10-113>
- Harvey, G., & Kitson, A. (2016). PARIHS revisited: From heuristic to integrated framework for the successful implementation of knowledge into practice. *Implementation Science*, 11, 33. <https://doi.org/10.1186/s13012-016-0398-2>
- He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, 25(1), 30–36. <https://doi.org/10.1038/s41591-018-0307-0>
- Hermes, E. D., Lyon, A. R., Schueller, S. M., & Glass, J. E. (2019). Measuring the implementation of behavioral intervention technologies: Recharacterization of established outcomes. *Journal of Medical Internet Research*, 21(1), e11752. <https://doi.org/10.2196/11752>
- Holton, R., & Boyd, R. (2021). ‘Where are the people? What are they doing? Why are they doing it?’ (Mindell) situating artificial intelligence within a socio-technical framework. *Journal of Sociology*, 57(2), 179–195. <https://doi.org/10.1177/02F1440783319873046>
- Horgan, D., Romao, M., Morré, S. A., & Kalra, D. (2019). Artificial intelligence: Power for civilisation - and for better healthcare. *Public Health Genomics*, 22(5–6), 145–161. <https://doi.org/10.1159/000504785>
- Huckvale, K., Venkatesh, S., & Christensen, H. (2019). Toward clinical digital phenotyping: A timely opportunity to consider purpose, quality, and safety. *NPJ Digital Medicine*, 2, 88. <https://doi.org/10.1038/s41746-019-0166-1>
- Ieso Health (2022). *World's first AI-enabled mental health treatment platform goes live*. Retrieved March 14, 2022, from <https://www.iesohealth.com/en-gb/news/world-s-first-ai-enabled-mental-health-treatment-platform-goes-live>
- Jones, P. B. (2013). Adult mental health disorders and their age at onset. *British Journal of Psychiatry Supplement*, 54, s5–s10. <https://doi.org/10.1192/bjp.bp.112.119164>
- Keane, P. A., & Topol, E. J. (2018). With an eye to AI and autonomous diagnosis. *NPJ Digital Medicine*, 1, 40. <https://doi.org/10.1038/s41746-018-0048-y>
- Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17, 195. <https://doi.org/10.1186/s12916-019-1426-2>
- Kilbourne, A. M., Beck, K., Spaeth-Rublee, B., Ramanuj, P., O'Brien, R. W., Tomoyasu, N., & Pincus, H. A. (2018). Measuring and improving the quality of mental health care: A global perspective. *World Psychiatry*, 17(1), 30–38. <https://doi.org/10.1002/wps.20482>
- Kim, J. H., Kim, B. G., Roy, P. P., & Jeong, D. M. (2019). Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. *IEEE Access*, 7, 41273–41285. <https://doi.org/10.1109/ACCESS.2019.2907327>
- Lavrentyeva, Y. (2021). *The big promise AI holds for mental health*. Retrieved May 20, 2022, from <https://itrexgroup.com/blog/ai-mental-health-examples-trends/#header>
- Lee, J. C. (2019). The perils of artificial intelligence in healthcare: Disease diagnosis and treatment. *Journal of Computational Biology and Bioinformatics Research*, 9(1), 1–6. <https://doi.org/10.5897/JCBBR2019.0122>
- Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H. C., Paulus, M. P., Krystal, J. H., & Jeste, D. V. (2021). Artificial intelligence for mental health care: Clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry, Cognitive Neuroscience and Neuroimaging*, 6(9), 856–864. <https://doi.org/10.1016/j.bpsc.2021.02.001>
- Leeman, J., Birken, S. A., Powell, B. J., Rohweder, C., & Shea, C. M. (2017). Beyond “implementation strategies”: Classifying the full range of strategies used in implementation science and practice. *Implementation Science*, 12(1), 125. <https://doi.org/10.1186/s13012-017-0657-x>
- Lekkas, D., Klein, R. J., & Jacobson, N. C. (2021). Predicting acute suicidal ideation on Instagram using ensemble



- machine learning models. *Internet Interventions*, 25, 100424. <https://doi.org/10.1016/j.invent.2021.100424>
- Liu, J. X., Goryakin, Y., Maeda, A., Bruckner, T., & Scheffler, R. (2017). Global health workforce labor market projections for 2030. *Human Resources for Health*, 15(1), 11. <https://doi.org/10.1186/s12960-017-0187-2>
- Liyanage, H., Liaw, S. T., Jonnagaddala, J., Schreiber, R., Kuziemy, C., Terry, A. L., & de Lusignan, S. (2019). Artificial intelligence in primary health care: Perceptions, issues, and challenges. *Yearbook of Medical Informatics*, 28(1), 41–46. <https://doi.org/10.1055/s-0039-1677901>
- MacDonald, K., Fainman-Adelman, N., Anderson, K. K., & Iyer, S. N. (2018). Pathways to mental health services for young people: A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, 53(10), 1005–1038. <https://doi.org/10.1007/s00127-018-1578-y>
- May, C., & Finch, T. (2009). Implementing, embedding, and integrating practices: An outline of normalization process theory. *Sociology*, 43, 535–554. <https://doi.org/10.1177/0038038509103208>
- McCarthy, J. F., Cooper, S. A., Dent, K. R., Eagan, A. E., Matarazzo, B. B., Hannemann, C. M., Reger, M. A., Landes, S. J., Trafton, J. A., & Schoenbaum, M. (2021). Evaluation of the recovery engagement and coordination for health-veterans enhanced treatment suicide risk modeling clinical program in the veterans health administration. *JAMA Network Open*, 4(10), e2129900. <https://doi.org/10.1001/jamanetworkopen.2021.29900>
- Michie, S., Atkins, L., & West, R. (2014). *The behaviour change wheel: A guide to designing interventions*. Silverback.
- Michie, S., Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6, 42. <https://doi.org/10.1186/1748-5908-6-42>
- Mistry, P. (2019). Artificial intelligence in primary care. *British Journal of General Practice*, 69(686), 422–423. <https://doi.org/10.3399/bjgp19X705137>
- Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, 13, 23–47. <https://doi.org/10.1146/annurev-clinpsy-032816-044949>
- Moore, G. F., & Evans, R. E. (2017). What theory, for whom and in which context? Reflections on the application of theory in the development and evaluation of complex population health interventions. *SSM - Population Health*, 3, 132–135. <https://doi.org/10.1016/j.ssmph.2016.12.005>
- Nicholas, J., Shilton, K., Schueller, S. M., Gray, E. L., Kwasny, M. J., & Mohr, D. C. (2019). The role of data type and recipient in individuals' perspectives on sharing passively collected smartphone data for mental health: Cross-sectional questionnaire study. *JMIR mHealth and UHealth*, 7(4), e12578. <https://doi.org/10.2196/12578>
- Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science*, 10, 53. <https://doi.org/10.1186/s13012-015-0242-0>
- Nilsen, P., & Birken, S. A. (2020). Prologue. In P. Nilsen, & S. A. Birken (Eds.), *Handbook on implementation science* (pp. 1–6). Edward Elgar.
- Pedersen, E. R., & Paves, A. P. (2014). Comparing perceived public stigma and personal stigma of mental health treatment seeking in a young adult sample. *Psychiatry Research*, 219(1), 143–150. <https://doi.org/10.1016/j.psychres.2014.05.017>
- Powell, B. J., McMillen, J. C., Proctor, E. K., Carpenter, C. R., Griffey, R. T., Bunger, A. C., Glass, J. E., & York, J. L. (2012). A compilation of strategies for implementing clinical innovations in health and mental health. *Medical Care Research and Review*, 69, 123–157. <https://doi.org/10.1177/1077558711430690>
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, 38, 65–76. <https://doi.org/10.1007/s10488-010-0319-7>
- Raibagi, K. (2020, December 20). *Top AI-based Mental Health Apps of 2020*. Opinions. <https://analyticsindiamag.com/top-ai-based-mental-health-apps-of-2020/>
- Rosenfeld, A., Benrimoh, D., Armstrong, C., Mirchi, N., Langlois-Therrien, T., Rollins, C., Tanguay-Sela, M., Mehlretter, J., Fratila, R., Israel, S., Snook, E., Perlman, K., Kleinerman, A., Saab, B., Thoburn, M., Gabbay, C., & Yaniv-Rosenfeld, A. (2021). Big data analytics and artificial intelligence in mental healthcare. *Applications of Big Data in Healthcare Theory and Practice*, 137–171. <https://doi.org/10.1016/B978-0-12-820203-6.00001-1>
- Safi, S., Thiessen, T., & Schmailzl, K. J. (2018). Acceptance and resistance of new digital technologies in medicine: Qualitative study. *JMIR Research Protocols*, 7(12), e11072. <https://doi.org/10.2196/11072>
- Schueller, S. M., & Torous, J. (2020). Scaling evidence-based treatments through digital mental health. *American Psychologist*, 75(8), 1093–1104. <https://doi.org/10.1037/amp0000654>
- Schwartz, B., Cohen, Z. D., Rubel, J. A., Zimmermann, D., Wittmann, W. W., & Lutz, W. (2021). Personalized treatment selection in routine care: Integrating machine learning and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy. *Psychotherapy Research*, 31(1), 33–51. <https://doi.org/10.1080/10503307.2020.1769219>
- Scott, W. R., Ruef, M., Mendel, P. J., & Caronna, C. A. (2000). *Institutional change and healthcare organizations*. The University of Chicago Press.
- Seneviratne, M. G., Shah, N., & Chu, L. (2019). Bridging the implementation gap of machine learning in healthcare. *BMJ Innovations*, 6(2), 45–47. <https://doi.org/10.1136/bmjinnov-2019-000359>
- Shah, R. N., & Berry, O. O. (2021). The rise of venture capital investing in mental health. *JAMA psychiatry*, 78(4), 351–352. <https://doi.org/10.1001/jamapsychiatry.2020.2847>
- Shaw, J., Rudzicz, F., Jamieson, T., & Goldfarb, A. (2019). Artificial intelligence and the implementation challenge. *Journal of Medical Internet Research*, 21(7), e13659. <https://doi.org/10.2196/13659>
- Stead, W. W. (2018). Clinical implications and challenges of artificial intelligence and deep learning. *JAMA*, 320(11), 1107–1108. <https://doi.org/10.1001/jama.2018.11029>
- Su, C., Xu, Z., Pathak, J., & Wang, F. (2020). Deep learning in mental health outcome research: A scoping review. *Translational Psychiatry*, 10, 116. <https://doi.org/10.1038/s41398-020-0780-3>
- Tanana, M. J., Soma, C. S., Kuo, P. B., Bertagnolli, N. M., Dembe, A., Pace, B. T., Srikumar, V., Atkins, D. C., &

- Imel, Z. E. (2021). How do you feel? Using natural language processing to automatically rate emotion in psychotherapy. *Behavior Research Methods*, 53(5), 2069–2082. <https://doi.org/10.3758/s13428-020-01531-z>
- Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
- Universal Health Services (2020). *Advancing Behavioral Health Outcomes Using Innovative Artificial Intelligence (AI)*. Retrieved March 15, 2022, from <https://uhs.com/advancing-behavioral-health-outcomes-using-innovative-artificial-intelligence-ai>.
- Uusitalo, S., Tuominen, J., & Arstila, V. (2021). Mapping out the philosophical questions of AI and clinical practice in diagnosing and treating mental disorders. *Journal of Evaluation in Clinical Practice*, 27(3), 478–484. <https://doi.org/10.1111/jep.13485>
- Verghese, A., Shah, N. H., & Harrington, R. A. (2018). What this computer needs is a physician: Humanism and artificial intelligence. *JAMA*, 319(1), 19–20. <https://doi.org/10.1001/jama.2017.19198>
- Wainberg, M. L., Scorza, P., Shultz, J. M., Helpman, L., Mootz, J. J., Johnson, K. A., Neria, Y., Bradford, J. E., Oquendo, M. A., & Arbuckle, M. R. (2017). Challenges and opportunities in global mental health: A research-to-practice perspective. *Current Psychiatry Reports*, 19(5), 28. <https://doi.org/10.1007/s11920-017-0780-z>
- Westberg, K. H., Nygren, J. M., Nyholm, M., Carlsson, I. M., & Svedberg, P. (2020). Lost in space - an exploration of help-seeking among young people with mental health problems: A constructivist grounded theory study. *Archives of Public Health*, 78, 93. <https://doi.org/10.1186/s13690-020-00471-6>
- Whitelaw, S., Mamas, M. A., Topol, E., & Van Spall, H. G. C. (2020). Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digital Health*, 2(8), e435–e440. [https://doi.org/10.1016/S2589-7500\(20\)30142-4](https://doi.org/10.1016/S2589-7500(20)30142-4)
- World Health Organization (2001). *The world health report 2001. Mental health: New understanding, new hope*. <https://apps.who.int/iris/handle/10665/42390>
- World Health Organization (2017). *Depression and other common mental disorders. Global health estimates*. <https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf>