



## OPEN ACCESS

## EDITED BY

Uffe Kock Wil, University of Southern Denmark, Denmark

## REVIEWED BY

Luke Balcombe, Griffith University, Australia  
Sonia Israel, McGill University, Canada

## \*CORRESPONDENCE

Katherine C. Kellogg  
kkellogg@mit.edu

## SPECIALTY SECTION

This article was submitted to Digital Mental Health, a section of the journal Frontiers in Psychiatry

RECEIVED 09 July 2022

ACCEPTED 19 August 2022

PUBLISHED 06 September 2022

## CITATION

Kellogg KC and Sadeh-Sharvit S (2022) Pragmatic AI-augmentation in mental healthcare: Key technologies, potential benefits, and real-world challenges and solutions for frontline clinicians. *Front. Psychiatry* 13:990370. doi: 10.3389/fpsy.2022.990370

## COPYRIGHT

© 2022 Kellogg and Sadeh-Sharvit. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Pragmatic AI-augmentation in mental healthcare: Key technologies, potential benefits, and real-world challenges and solutions for frontline clinicians

Katherine C. Kellogg<sup>1\*</sup> and Shiri Sadeh-Sharvit<sup>2,3</sup>

<sup>1</sup>Department of Work and Organization Studies, MIT Sloan School of Management, Cambridge, MA, United States, <sup>2</sup>Eleos Health, Cambridge, MA, United States, <sup>3</sup>Center for M<sup>2</sup>Health, Palo Alto University, Palo Alto, CA, United States

The integration of artificial intelligence (AI) technologies into mental health holds the promise of increasing patient access, engagement, and quality of care, and of improving clinician quality of work life. However, to date, studies of AI technologies in mental health have focused primarily on challenges that policymakers, clinical leaders, and data and computer scientists face, rather than on challenges that frontline mental health clinicians are likely to face as they attempt to integrate AI-based technologies into their everyday clinical practice. In this Perspective, we describe a framework for “pragmatic AI-augmentation” that addresses these issues by describing three categories of emerging AI-based mental health technologies which frontline clinicians can leverage in their clinical practice—automation, engagement, and clinical decision support technologies. We elaborate the potential benefits offered by these technologies, the likely day-to-day challenges they may raise for mental health clinicians, and some solutions that clinical leaders and technology developers can use to address these challenges, based on emerging experience with the integration of AI technologies into clinician daily practice in other healthcare disciplines.

## KEYWORDS

artificial intelligence, AI-augmentation, automation technologies, clinical practice, decision support technologies, engagement technologies, mental healthcare

## Introduction

Artificial intelligence (AI) technologies—computer systems that perform human-like physical and cognitive tasks such as sensing, perceiving, problem solving, and learning subtle patterns of language and behavior (1)—may both enable mental health (MH) clinicians to focus on the human aspects of medicine that can only be achieved through the clinician–patient relationship, and at the same time make therapies more effective. In the health domain, the automation, engagement, and decisions empowered by AI are commonly reviewed by domain experts before they can be implemented in a patient’s treatment plan, a process called “augmented intelligence”

or “intelligence amplification.” By utilizing the capabilities of AI systems to support clinicians and patients, they transform mental healthcare (MHC) through improving patient experience and retention and the work life of healthcare professionals, and reducing costs (2). AI-Augmentation incorporates a pragmatic approach, led by the frontline clinician, wherein AI technology informs and augments, rather than replaces, clinician experience and cognition (3).

To date, much of the literature on implementing AI technologies in MHC has highlighted the potential clinical and economic value of AI (1), as well as the technical, ethical, and regulatory challenges associated with its effective implementation (4). A research agenda and initiatives for addressing these challenges has also been proposed, such as redesigning elements of the technical and regulatory infrastructure of organizations (4), using robust data and science practices in developing AI technologies (5, 6), studying AI efficacy and fairness (7), and providing formal training to medical professionals (7).

To add to this emerging understanding of effective AI development and use in MHC, we provide a complementary perspective of “Pragmatic AI-Augmentation.” We argue that, in order to realize the promise of AI technologies in MHC, it is necessary to answer three questions: (1) What are some specific AI technologies that frontline MH clinicians can leverage to inform and augment their human intelligence in clinical practice?; (2) What challenges are likely to arise for MH clinicians as they attempt to use these AI technologies in their daily work?; and (3) What solutions can clinical leaders and technology developers use to help address MH clinicians’ AI implementation challenges?

Three categories of AI technologies that are of particular interest to MH clinicians are automation technologies, engagement technologies, and clinical decision support technologies (8, 9). AI-based automation technologies allow for the automation of healthcare management processes to optimize healthcare delivery or reduce administrative cost, often via computer vision and machine learning (ML) systems. AI-based engagement technologies allow for engagement with patients using natural language processing (NLP) chatbots and intelligent agents. And AI-based clinical decision-making support technologies allow for early detection and diagnosis through extensive analysis of structured data, most often using ML algorithms and neural networks. While we do not cover all types of technologies in this paper, we provide a framework for patients, clinicians, and stakeholders to think about AI-augmentation in their practice (10). Table 1 illustrates key characteristics, applications, and potential benefits of AI for MH clinicians in each of these three domains. Table 2 details related challenges and solutions required for pragmatic AI-augmentation.

## AI-based automation technologies for mental healthcare

### Key technologies for AI-based automation

AI-based automation technologies enable the automation of structured or semi-structured tasks, often *via* computer vision, ML, and NLP. They can perform a wider variety of tasks than traditional automation technologies, such as recognizing conversations and emotions, and therefore can automate or semi-automate certain aspects of care, such as screening, diagnosis, and treatment recommendations, or psychosocial therapy. For example, computer vision in combination with ML offers early autism screening by assessing the eye gaze patterns of children while they watch a short movie on an iPad (11). Platforms using NLP in combination with ML can provide chatbot-based screening as well as assist with routine or mundane tasks that clinicians dislike or are not reimbursed for. For instance, AI-based automation technologies can support measurement-based care by automatically collecting and analyzing data pre-, during, and post-session, administering digital surveys, and generating insights from a patient’s talking patterns, tone, and word choice (12).

### Potential benefits of AI-based automation

AI-based automation technologies facilitate the automation of cognitive and physical tasks that are often repetitive and cumbersome, thereby facilitating early screening, increasing patient access, and improving clinician quality of work life. For example, while measurement-based care is perceived to advance MHC and make it more data-informed (13), administering standardized assessments is a laborious task that can be easily computerized. Digital screening tools for autism, eating disorders, or other psychosocial concerns, can be complementary to other assessment approaches and ultimately increase the accuracy, exportability, accessibility, and scalability of screening and intervention (14–16). More generally, these technologies can be used to extend, rather than replace, the skills and experience of frontline clinicians, so could help address some of the acute workforce shortages in MHC (17). This augmented AI methodology could also free up staff time to focus on the human aspects of care. In a blended-care approach, AI-driven chatbots, apps, and online interventions can provide data on a person’s responsiveness to them, and quickly identify cases where improvement criteria are not met, prioritizing a higher level of care as needed (18).

TABLE 1 Key AI technologies, example applications in mental healthcare, and potential benefits.

	Automation technologies	Patient engagement technologies	Clinical decision support technologies
Key AI technologies	<ul style="list-style-type: none"> <li>• Computer vision</li> <li>• Machine learning</li> <li>• Natural language processing</li> </ul>	<ul style="list-style-type: none"> <li>• Conversational Agents</li> <li>• Chatbots</li> </ul>	<ul style="list-style-type: none"> <li>• Machine learning</li> <li>• Neural networks</li> <li>• Deep learning</li> </ul>
Example applications in MH	<ul style="list-style-type: none"> <li>• Screening for autism, eating disorders</li> <li>• Administering digital surveys</li> <li>• Generating insights from a patient's talking patterns, tone, word choice</li> </ul>	<ul style="list-style-type: none"> <li>• Coaching for smoking cessation, exercise, nutrition</li> <li>• Scripted CBT, MI, dialectical behavioral therapy</li> </ul>	<ul style="list-style-type: none"> <li>• Depression and anxiety prediction</li> <li>• Stroke prediction</li> <li>• Suicide ideation prediction</li> </ul>
Potential benefits	<ul style="list-style-type: none"> <li>• Increase early screening</li> <li>• Increase patient access</li> <li>• Improve clinician quality of work life</li> </ul>	<ul style="list-style-type: none"> <li>• Increase patient access</li> <li>• Increase patient engagement</li> <li>• Extend services provided by organizations adopting a hybrid therapy model</li> </ul>	<ul style="list-style-type: none"> <li>• Help clinicians identify mental illnesses at an earlier stage when interventions may be more effective</li> <li>• Help clinicians customize treatments based on a patient's characteristics</li> </ul>
AI-augmentation in MH practice	<ul style="list-style-type: none"> <li>• Automation technologies could assist alleviate some of the severe staffing shortages in mental healthcare by extending rather than replacing the knowledge and expertise of human clinicians</li> </ul>	<ul style="list-style-type: none"> <li>• AI engagement tools could complement the therapist's interventions in a blended therapy model. To support skill practice between sessions, for example, the therapist could assign a specific GSH component, extending therapy beyond the meeting</li> </ul>	<ul style="list-style-type: none"> <li>• To increase clinicians' readiness to consider using AI tools to help decision-making, there has to be more openness about how algorithms are developed, the data utilized for their creation, and the engagement of mental health practitioners and service users in their evaluation and improvement</li> </ul>

AI, artificial intelligence; MH, mental health; GSH, guided self-help.

TABLE 2 Challenges and potential solutions related to pragmatic AI-augmentation.

	Automation technologies	Patient engagement technologies	Clinical decision support technologies
<b>Potential challenges</b>			
• Work practices	<ul style="list-style-type: none"> <li>• Automation complacency</li> </ul>	<ul style="list-style-type: none"> <li>• Required workflow changes</li> </ul>	<ul style="list-style-type: none"> <li>• Lack of interpretability of AI recommendations</li> </ul>
• Beliefs and identity	<ul style="list-style-type: none"> <li>• Fear of obsolescence</li> </ul>	<ul style="list-style-type: none"> <li>• Fear of negative impact on clinician-patient therapeutic alliance</li> </ul>	<ul style="list-style-type: none"> <li>• Concerns about model accuracy and bias</li> <li>• Lack of evidence-based evaluation studies</li> </ul>
• Tasks and roles	<ul style="list-style-type: none"> <li>• Clinician-AI task allocation</li> </ul>	<ul style="list-style-type: none"> <li>• Concerns about patient overreliance on technology</li> </ul>	<ul style="list-style-type: none"> <li>• Clinician surveillance concerns</li> </ul>
<b>Potential solutions</b>			
• Work practices	<ul style="list-style-type: none"> <li>• Provide proactive risk assessment</li> <li>• Expose clinicians to potential automation failures</li> </ul>	<ul style="list-style-type: none"> <li>• Engage clinicians in iterative technology development and implementation process</li> </ul>	<ul style="list-style-type: none"> <li>• Train clinicians in computational thinking</li> <li>• Use explainable AI (XAI) techniques</li> </ul>
• Beliefs and identity	<ul style="list-style-type: none"> <li>• Provide more accurate portrayals of capabilities of AI technologies</li> </ul>	<ul style="list-style-type: none"> <li>• Discuss potential risks related to the patient becoming too emotionally attached</li> </ul>	<ul style="list-style-type: none"> <li>• Use human-in-the-loop development</li> <li>• Evaluate efficacy vs. current models</li> </ul>
• Tasks and roles	<ul style="list-style-type: none"> <li>• Add organizational structures and roles to govern AI projects</li> </ul>	<ul style="list-style-type: none"> <li>• Automate detection of threat, and provide recommended action</li> </ul>	<ul style="list-style-type: none"> <li>• Configure AI technologies so that predictions do not infringe upon clinicians' core tasks</li> </ul>

AI, artificial intelligence.

## Potential challenges of AI-based automation

With these new potential benefits come new potential challenges. First, regarding clinician daily work practices, these technologies raise the potential for automation complacency. This can occur, for example with an “auto-complete” function for drug names after entering the first few letters. Because clinicians are faced with complex tasks, multitasking, heavier workloads, and increasing time pressures, they may choose the route requiring the least amount of cognitive effort, which may lead to them letting technology direct their path (19). Additionally, clinicians may overestimate the performance of these technologies because they believe that technology has better analytical capabilities than humans, and they might put less effort or responsibility into carrying out a task (20).

Second, regarding beliefs and identities, narratives surrounding AI (21) might influence how frontline clinicians perceive and approach the technology. Because AI-based automation enables the automation of well-defined cognitive tasks that can be captured by sets of patterns and rules, this may lead to clinician concerns about obsolescence. Clinicians may worry that AI will replace the need for their expert work (9, 22).

Third, regarding new tasks, clinicians may be concerned about task allocation between clinicians and AI technologies. They may wonder how the delegation of tasks among interdependent humans and machines should be determined (23). This uncertainty may, in turn, undermine confidence and prevent the widespread use of these technologies.

## Potential solutions for AI-based automation

Regarding potential automation complacency, this issue is not unique to AI technologies, and has been handled effectively in healthcare settings using two strategies. First, before implementing new technology, clinical leaders can first undertake a proactive risk assessment to identify any unexpected vulnerabilities and remediate them (24). Second, leaders can allow clinicians to experience automation failures during training, such as technology failure to issue an important alert or “auto-fill” errors, to encourage critical thinking when using automated systems, and increase the likelihood of recognizing these failures during daily work (25).

Regarding fear of obsolescence, although in the popular imagination software often takes on mythic qualities of near-omnipotence (26), AI will likely augment the importance of human judgement and knowledge (27–29). Clinical leaders could highlight the essential role played by humans at each phase of the development and implementation of an AI technology (30). For example, leaders could note that radiologists may

relinquish the preliminary scanning of medical images to the AI technology in order to focus on the more complicated cases, or those that the algorithm finds ambiguous (31); technology developers can include feedback during screening that recommends reaching out for in-person evaluation to verify the clinical recommendation (32).

Regarding concerns about task allocation, the design of AI systems—like the design of other sociotechnical systems—involves decisions about how to integrate augmented intelligence and delegate sub-tasks among humans and non-humans. Clinical leaders at healthcare organizations using AI have put in place a wide range of internal organizational structures and roles to manage and govern AI projects (33). Governance committees should offer a balanced view of challenges and opportunities associated with implementing AI technologies (34).

## AI-enabled engagement technologies for mental healthcare

### Key technologies for AI-enabled engagement

AI-enabled voice-and text-based engagement technologies allow for patient engagement, ranging from the handling of repetitive patient queries to the undertaking of more complex tasks that involve greater interaction, conversation, reasoning, prediction, accuracy, and emotional display. Conversational agents (CAs) such as chatbots use NLP, speech recognition, and synthesis technologies to conduct limited text conversations with human users. Intelligent virtual agents (IVAs) add more realism to conversational emulation through computer-generated, artificially intelligent virtual characters.

IVAs and CAs have been used as coaches to support behavior change such as smoking cessation, starting exercise programs, nutrition, and to provide limited treatment functions, such as elements of cognitive behavioral therapy (CBT), dialectical behavior therapy, or motivational interviewing, in areas where highly scripted protocols are applicable (35). For instance, Woebot is a conversational agent designed to help reduce depression and anxiety. It provides CBT tools, learning materials, and videos as well to guide users through “self-directed” therapy (36). Further, digital guided self-help (GSH) programs can improve access to evidence-based therapy and offer patients standardized interventions (37).

### Potential benefits of AI-enabled engagement

While these technologies cannot function at the level of an experienced, empathic human clinician, they may play a special

role in MHC prevention and intervention, both supportive of and distinct from the functions of human healthcare providers. AI-enabled engagement technologies have the potential to increase patient access because they offer 24/7 availability in remote areas where access to MH professionals may be limited (38). They can also increase the dissemination of evidence-based practices. For instance, “Elizabeth,” a virtual nurse, assists patients in understanding information about hospital discharge, such as follow-up care and prescription needs (38). AI can also identify program participants’ disengagement and alert the assigned case manager or clinician, using hundreds of datapoints that humans do not always have access to or cannot process in a timely manner (39).

In addition, these technologies may provide between-session support through facilitating homework and skills practice. In a blended therapy model, AI tools for engagement augment clinicians’ interventions. Similarly, AI-augmented engagement can extend the services provided by organizations adopting a hybrid therapy model—integrating both in-person and technology-based modalities (40). For instance, clinicians can assign a specific GSH module, thereby extending therapy beyond the meeting (37). Further, a study on the use of CAs in inner city outpatient clinics found that patients with chronic pain and depression who interacted with a CA reported a high degree of compliance with the CAs’ recommendations for stress reduction and healthy eating (37).

## Potential challenges of AI-enabled engagement

Just as AI-enabled engagement technologies may afford new benefits for MH clinicians, so they may raise a new set of new challenges. First, regarding clinician daily work practices, these technologies will require workflow changes that enable clinicians to collaborate with chatbots and patients, and there are no codified practices that clinicians can use to implement chatbots in a way that meets patients’ needs, goals, and lifestyles while facilitating trust in AI to improve MH outcomes (41).

Second, regarding beliefs and identities associated with AI-Enabled Engagement, recent research on CAs and GSH have found that program participants report forming some sort of a therapeutic alliance with automated MH platforms. For instance, the users of a smoking cessation chatbot reported that their perceived therapeutic alliance with the chatbot increased over time (42). Patients forming a relationship with the chatbot may lead to clinician concerns about a negative impact on the therapeutic alliance between clinicians and patients (43). Further, the literature has noted that the therapeutic alliance with a human being and with a therapist may be distinct concepts that need to be assessed using different tools (44).

Third, regarding concerns about task allocation, clinicians may be concerned that their patients may assume that the technology is adequately addressing their MH needs, and may over-rely on technology (7). Additionally, there is a potential risk of harm to patients if a system is unable to appropriately address circumstances in which a user needs rapid crisis care or if other safety-related action is necessary (43). Risk assessment is complicated enough for providers; if we delegate it entirely to machines, patients’ unique circumstances may be overlooked (45–47).

## Potential solutions for AI-enabled engagement

Regarding workflow changes required by AI-enabled engagement technologies, research on AI technology implementation in other healthcare domains has shown that clinical leaders can facilitate the creation of new workflows by supporting an iterative process between AI technology’s developers and clinicians—through which both the clinicians’ daily activities and the technology itself are transformed (48).

Regarding a potential negative impact on the therapeutic alliance, clinical leaders could encourage the responsible party to discuss with the patient any potential risks related to the patient relying on the agent more than on human caregiver or clinician. Technology developers could design the technology to detect if the user is becoming too emotionally dependent, broach this topic with them, and discuss possible risks (7).

Third, regarding clinician concerns about task allocation, technology developers should include an explicit message to the patient outlining circumstances under which patients should seek clinician help. There is a need for blending AI insights with clinician-led engagement to incorporate the nuances, risk, and support resources an individual has. Developers could design agents to identify types of risks and take appropriate action, such as detecting suicide risk and immediately directing the patient to contact a suicide prevention hotline, a clinician, or caregiver (7).

## AI-enabled decision support technologies for mental healthcare

### Key technologies for AI-enabled decision support

AI-enabled decision support technologies use ML algorithms to make predictions based on past data. New large-scale advances, such as deep or reinforcement learning (49, 50), have enabled AI to make inroads into much more complex decision-making settings, including those that involve audio and speech recognition. These improved decision-making processes stem from big datasets now available in healthcare,

that are enhanced by computational analyses (51). For example, supervised ML has been shown to predict attention-deficit and hyperactivity disorder (ADHD), autism, schizophrenia, and Tourette syndrome and suicidal ideation. ML-based models have been shown to more accurately predict the severity of depression or anxiety (52), and stroke (53) and allow clinicians to identify which interventions will be most effective for which patient populations.

## Potential benefits of AI-enabled decision support

AI-enabled decision support technologies can enable clinicians to detect mental illnesses at an earlier stage when interventions may be more effective, and offer customized treatments (1). These technologies, as opposed to clinicians who are human, never grow weary, can integrate new knowledge into their models, and can process large datasets in a matter of minutes. However, algorithms do not have the empathy and capacity for nuance of a human being. Therefore, platforms that integrate data from multiple resources such as assessments, the treatment alliance, content from treatment sessions, and the patient's perception of therapy, are poised to augment, rather than replace, clinician decision making (54).

## Potential challenges of AI-enabled decision support

With these new potential benefits comes a set of new challenges. First, regarding clinician daily work practices, ML model outputs are often complex and uninterpretable to humans—the so called 'black box' problem. The precise reason for an ML model's recommendation often cannot be easily pinpointed, even by data scientists (31). In addition, technology-based approaches are pattern-based, while traditional research methods are hypothesis-driven, and most clinicians do not have the training required to engage in the computational thinking associated with these interpretations (17, 37).

Second, regarding beliefs and identities, clinicians may not trust that ML-based algorithms yield accurate predictions and that recommendations can be integrated into clinical practice securely and efficiently for the benefit of patients. There is currently a low quality and quantity of evaluated evidence-based studies for these technologies (55). In the healthcare context, lives may depend on assuring that model predictions not only are accurate in the original context, but also function robustly in other contexts, and that the results of predictive models are not biased (17).

Third, regarding concerns about task allocation, AI decision-support tools are by their nature prescriptive in that

they recommend actions for the clinician to take. Clinicians may be concerned that the availability of AI-enabled decision support technologies may allow others from outside their domain to attempt to scrutinize and dictate frontline clinicians' diagnosis and treatment decisions (56).

Fourth, most studies on AI-augmentation for personalizing therapy have been carried out on clients with depression and anxiety (57). AI-based clinical decision-support technologies have been understudied in severe mental illnesses (SMI), including schizophrenia, bipolar disorder, severe depressive disorder and psychotic disorders (58).

## Potential solutions for AI-enabled decision support

To address the black box problem and misperceptions of ML models, clinicians should be informed about the advantages, limitations, and risks of AI-based systems. Such knowledge and abilities are required so that clinicians feel confident in using AI and communicating the results effectively with patients and carers. The more technology developers incorporate explainable Artificial Intelligence (XAI), i.e., explanatory techniques that clearly state why the AI has made a specific suggestion, the more they can increase the credibility, accountability, and trust in AI in mission-critical fields like MHC (59). Developers should also consider the audience when determining what explanations to provide; for example, clinicians may be most interested in explanations that make the model's functioning easy to understand, and in those that help them trust the model itself (60).

Second, regarding therapists' concerns about how AI-enabled decision support technologies can be incorporated effectively and safely into clinical practice, developers need to foster greater involvement of MH providers and service users in AI models evaluation and refinement. Developers can use a human-in-the-loop approach to data collection and model creation that enlists MH practitioners in all stages of development and use (61). There also needs to be more separation between studies that measure degree of use of AI technologies, and those that evaluate the efficacy of these technologies; in order to gain the trust and understanding of clinicians, developers need to demonstrate that these technologies work better than existing service delivery (17, 62). For example, in the case of stroke risk prediction and prevention, this means that novel ML-based approaches need to compete against established models to win clinicians' and patients' trust (53).

Third, regarding task allocation, research on AI technology implementation in other healthcare domains has shown that clinician concerns regarding autonomy and surveillance can be addressed by clinical leaders facilitating greater clinician

involvement in developing these technologies and by developers configuring the AI tools so that their predictions do not infringe upon clinicians' core tasks, but instead assist them with important tasks less valued by the clinicians such as monitoring the completion of follow-up items (56).

Fourth, in regard to the claim that AI-based decision support tools should be expanded to SMI, since these conditions may be relatively complicated to diagnose, we suggest that efforts to develop technologies for AI decision support adopt human-in-the-loop methods to develop, assess, refine, and test the tools before they are deployed in practice. In addition, technology developers should give more attention to the impacts of using these tools to tailor and personalize treatment (58).

## Conclusion

While much has been written about the potential for AI technologies to transform the delivery of MHC, the predominant focus to date has been on the technical, ethical, and regulatory challenges associated with these technologies. What is missing is a pragmatic approach to AI-augmentation that outlines challenges and potential solutions related to the incorporation of AI-based technologies into clinicians' everyday work. Such an approach can help frontline clinicians use AI-based technologies to transform how MH is understood, accessed, treated, and integrated, as clinicians make decisions, day in and day out, to care for their patients in need.

## References

- Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim H-C, et al. Artificial intelligence for mental health and mental illnesses: an overview. *Curr Psychiatry Rep.* (2019) 2111:1–18. doi: 10.1007/s11920-019-1094-0
- Crigger E, Reinbold K, Hanson C, Kao A, Blake K, Irons M. Trustworthy augmented intelligence in health care. *J Med Syst.* (2022) 462:1–11. doi: 10.1007/s10916-021-01790-z
- Bazoukis G, Hall J, Loscalzo J, Antman EM, Fuster V, Armoundas AA. The inclusion of augmented intelligence in medicine: a framework for successful implementation. *Cell Rep Med.* (2022) 31:100485–100485. doi: 10.1016/j.xcrm.2021.100485
- Jotterand F, Ienca M. Artificial intelligence in brain and mental health: philosophical, ethical and policy issues. *Adv Neuroethics.* (2021). 2021:1–7. doi: 10.1007/978-3-030-74188-4\_1
- Plis SM, Hjelm DR, Salakhutdinov R, Allen EA, Bockholt HJ, Long JD, et al. Deep learning for neuroimaging: a validation study. *Front Neurosci.* (2014) 8:229. doi: 10.3389/fnins.2014.00229
- Xiao T, Albert MV. *Big Data in Medical AI: How Larger Data Sets Lead to Robust, Automated Learning for Medicine. Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues.* Cham, Switzerland: Springer (2021). p. 11–25. doi: 10.1007/978-3-030-74188-4\_2
- Luxton DD, Hudlicka E. *Intelligent Virtual Agents in Behavioral and Mental Healthcare: Ethics and Application Considerations. Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues.* Cham, Switzerland: Springer (2021). p. 41–55. doi: 10.1007/978-3-030-74188-4\_4
- Benbya H, Pachidi S, Jarvenpaa S. Special issue editorial: artificial intelligence in organizations: implications for information systems research. *J Assoc Inform Syst.* (2021) 222:10. doi: 10.17705/1jais.00662
- Benbya H, Nan N, Tanriverdi H, Yoo Y. Complexity and information systems research in the emerging digital world. *Mis Q.* (2020) 441:1–17. doi: 10.25300/MISQ/2020/13304
- Bickman L. Improving mental health services: a 50-year journey from randomized experiments to artificial intelligence and precision mental health. *Adm Policy Ment Health.* (2020) 475:795–843.
- Chang Z, Di Martino JM, Aiello R, Baker J, Carpenter K, Compton S, et al. Computational methods to measure patterns of gaze in toddlers with autism spectrum disorder. *JAMA Pediatr.* (2021) 175:827–36. doi: 10.1001/jamapediatrics.2021.0530
- Sadeh-Sharvit S, Hollon SD. Leveraging the power of nondisruptive technologies to optimize mental health treatment: case study. *JMIR Mental Health.* (2020) 711:e20646. doi: 10.2196/preprints.20646
- Connors EH, Douglas S, Jensen-Doss A, Landes SJ, Lewis CC, McLeod BD, et al. What gets measured gets done: how mental health agencies can leverage measurement-based care for better patient care, clinician supports, and organizational goals. *Adm Policy Ment Health Mental Health Serv Res.* (2021) 482:250–65. doi: 10.1007/s10488-020-01063-w
- Shahamiri SR, Thabtah F. Autism AI: a new autism screening system based on artificial intelligence. *Cogn Comput.* (2020) 124:766–77. doi: 10.1007/s12559-020-09743-3

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## Author contributions

KK and SS-S: conceptualization and writing—review and editing. All authors contributed to the article and approved the submitted version.

## Conflict of interest

Author SS-S is the Chief Clinical Officer of Eleos Health.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

15. Sadeh-Sharvit S, Fitzsimmons-Craft EE, Taylor CB, Yom-Tov E. Predicting eating disorders from Internet activity. *Int J Eat Disord.* (2020) 539:1526–33. doi: 10.1002/eat.23338
16. Smrke U, Mlakar I, Lin S, Musil B, Plohl N. Language, speech, and facial expression features for artificial intelligence-based detection of cancer survivors' depression: scoping meta-review. *JMIR Mental Health.* (2021) 812:e30439. doi: 10.2196/preprints.30439
17. Balcombe L, De Leo D. Digital mental health challenges and the horizon ahead for solutions. *JMIR Mental Health.* (2021) 83:e26811. doi: 10.2196/26811
18. Lim HM, Teo CH, Ng CJ, Chiew TK, Ng WL, Abdullah A, et al. An automated patient self-monitoring system to reduce health care system burden during the Covid-19 pandemic in Malaysia: development and implementation study. *JMIR Med Inform.* (2021) 92:e23427. doi: 10.2196/23427
19. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc.* (2012) 191:121–7. doi: 10.1136/amiajnl-2011-000089
20. Parasuraman R, Manzey DH. Complacency and bias in human use of automation: an attentional integration. *Hum Factors.* (2010) 523:381–410. doi: 10.1177/0018720810376055
21. Sartori L, Bocca G. Minding the gap (s): public perceptions of AI and socio-technical imaginaries. *AI Soc.* (2022) 2022:1–16. doi: 10.1007/s00146-022-01422-1
22. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res.* (2019) 215:e13216. doi: 10.2196/13216
23. Martinengo L, Van Galen L, Lum E, Kowalski M, Subramaniam M, Car J. Suicide prevention and depression apps' suicide risk assessment and management: a systematic assessment of adherence to clinical guidelines. *BMC Med.* (2019) 171:1–12. doi: 10.1186/s12916-019-1461-z
24. Colman N, Stone K, Arnold J, Doughty C, Reid J, Younker S, et al. Prevent safety threats in new construction through integration of simulation and FMEA. *Pediatr Qual Saf.* (2019) 44:e189. doi: 10.1097/pq9.000000000000189
25. Grissinger M. Understanding human over-reliance on technology. *Pharm Therap.* (2019) 446:320.
26. Ziewitz M. Governing algorithms: myth, mess, and methods. *Sci Technol Human Values.* (2016) 411:3–16. doi: 10.1177/0162243915608948
27. Ekbia HR, Nardi BA. Heteromation, and other stories of computing and capitalism. Cambridge, MA: MIT Press. (2017). doi: 10.7551/mitpress/10767.001.0001
28. Autor D. Why are there still so many jobs? The history and future of workplace automation. *J Econ Perspect.* (2015) 29:3–30. doi: 10.1257/jep.29.3.3
29. Autor DH. *The Paradox of Abundance: Automation Anxiety Returns. Performance and Progress: Essays on Capitalism, Business, and Society.* New York, NY: Oxford University Press (2015). p. 237–60. doi: 10.1093/acprof:oso/9780198744283.003.0017
30. Johnson DG, Verdicchio M. Reframing AI discourse. *Minds Mach.* (2017) 274:575–90. doi: 10.1007/s11023-017-9417-6
31. Lebovitz S, Lifshitz-Assaf H, Levina N. To engage or not to engage with AI for critical judgments: how professionals deal with opacity when using AI for medical diagnosis. *Organ Sci.* (2022) 331:126–48. doi: 10.1287/orsc.2021.1549
32. Fitzsimmons-Craft EE, Balantekin KN, Eichen DM, Graham AK, Monterubio GE, Sadeh-Sharvit S, et al. Screening and offering online programs for eating disorders: reach, pathology, and differences across eating disorder status groups at 28 US universities. *Int J Eat Disord.* (2019) 5210:1125–36. doi: 10.1002/eat.23134
33. Cruz Rivera S, Liu X, Chan A-W, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med.* (2020) 269:1351–63. doi: 10.1136/bmj.m3210
34. Wilson A, Saeed H, Pringle C, Eleftheriou I, Bromiley PA, Brass A. Artificial intelligence projects in healthcare: 10 practical tips for success in a clinical environment. *BMJ Health Care Inform.* (2021) 281:e100323. doi: 10.1136/bmjhci-2021-100323
35. Hudlicka E. *Virtual Affective Agents and Therapeutic Games. Artificial Intelligence in Behavioral and Mental Health Care.* San Diego, CA: Elsevier (2016). p. 81–115. doi: 10.1016/B978-0-12-420248-1.00004-0
36. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health.* (2017) 42:e7785. doi: 10.2196/mental.7785
37. Taylor CB, Graham AK, Flatt RE, Waldherr K, Fitzsimmons-Craft EE. Current state of scientific evidence on internet-based interventions for the treatment of depression, anxiety, eating disorders and substance abuse: an overview of systematic reviews and meta-analyses. *Eur J Public Health.* (2021) 31:i3–i10. doi: 10.1093/eurpub/ckz208
38. Torous J, Myrick KJ, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. *JMIR Mental Health.* (2020) 73:e18848. doi: 10.2196/18848
39. Funk B, Sadeh-Sharvit S, Fitzsimmons-Craft EE, Trockel MT, Monterubio GE, Goel NJ, et al. A framework for applying natural language processing in digital health interventions. *J Med Internet Res.* (2020) 222:e13855. doi: 10.2196/13855
40. Singh S, Germine L. Technology meets tradition: a hybrid model for implementing digital tools in neuropsychology. *Int Rev Psychiatry.* (2021) 334:382–93. doi: 10.1080/09540261.2020.1835839
41. Boucher EM, Harake NR, Ward HE, Stoeckl SE, Vargas J, Minkel J, et al. Artificially intelligent chatbots in digital mental health interventions: a review. *Expert Rev Med Devices.* (2021) 18:37–49. doi: 10.1080/17434440.2021.2013200
42. He L, Basar E, Wiers RW, Antheunis ML, Kraemer E. Can chatbots help to motivate smoking cessation? A study on the effectiveness of motivational interviewing on engagement and therapeutic alliance. *BMC Public Health.* (2022) 221:1–14. doi: 10.1186/s12889-022-13115-x
43. Miner AS, Milstein A, Schueller S, Hegde R, Mangurian C, Linos E. Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Intern Med.* (2016) 1765:619–25. doi: 10.1001/jamainternmed.2016.0400
44. D'Alfonso S, Lederman R, Bucci S, Berry K. The digital therapeutic alliance and human-computer interaction. *JMIR Mental Health.* (2020) 712:e21895. doi: 10.2196/preprints.21895
45. D'Hotman D, Loh E. AI enabled suicide prediction tools: a qualitative narrative review. *BMJ Health Care Inform.* (2020) 273:e100715. doi: 10.1136/bmjhci-2020-100175
46. Mulder R, Newton-Howes G, Coid JW. The utility of risk prediction in psychiatry. *Br J Psychiatry.* (2016) 2094:271–2. doi: 10.1192/bjp.bp.116.184960
47. Chan MK, Bhatti H, Meader N, Stockton S, Evans J, O'Connor RC, et al. Predicting suicide following self-harm: systematic review of risk factors and risk scales. *Br J Psychiatry.* (2016) 2094:277–83. doi: 10.1192/bjp.bp.115.170050
48. Singer SJ, Kellogg KC, Galper AB, Viola D. Enhancing the value to users of machine learning-based clinical decision support tools: a framework for iterative, collaborative development and implementation. *Health Care Manage Rev.* (2022) 472:E21–31. doi: 10.1097/HMR.0000000000000324
49. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* (2015) 521:436–44. doi: 10.1038/nature14539
50. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction.* Massachusetts: MIT Press (2018).
51. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer.* (2018) 188:500–10. doi: 10.1038/s41568-018-0016-5
52. Fletcher S, Spittal MJ, Chondros P, Palmer VJ, Chatterton ML, Densley K, et al. Clinical efficacy of a Decision Support Tool (Link-me) to guide intensity of mental health care in primary practice: a pragmatic stratified randomised controlled trial. *Lancet Psychiatry.* (2021) 83:202–14. doi: 10.1016/S2215-0366(20)30517-4
53. Amann J. *Machine Learning in Stroke Medicine: Opportunities and Challenges for Risk Prediction and Prevention. Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues.* Cham, Switzerland: Springer (2021). p. 57–71. doi: 10.1007/978-3-030-74188-4\_5
54. Rundo L, Pirrone R, Vitabile S, Sala E, Gambino O. Recent advances of HCI in decision-making tasks for optimized clinical workflows and precision medicine. *J Biomed Inform.* (2020) 108:103479. doi: 10.1016/j.jbi.2020.103479
55. Balcombe L, De Leo D, editors. Human-computer interaction in digital mental health. *Informatics.* (2022) 9:14. doi: 10.3390/informatics9010014
56. Kellogg KC, Sendak M, Balu S. AI on the front lines. *MIT Sloan Manag Rev.* (2022) 634:44–50.
57. Cohen ZD, DeRubeis RJ. Treatment selection in depression. *Annu Rev Clin Psychol.* (2018) 141:209–36. doi: 10.1146/annurev-clinpsy-050817-084746
58. Dawoodbhoj FM, Delaney J, Cecula P, Yu J, Peacock I, Tan J, et al. AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units. *Heliyon.* (2021) 75:e06993. doi: 10.1016/j.heliyon.2021.e06993
59. Ammar N, Shaban-Nejad A. Explainable Artificial Intelligence recommendation system by leveraging the semantics of adverse



childhood experiences: proof-of-concept prototype development. *JMIR Med Inform.* (2020) 811:e18752. doi: 10.2196/preprints.18752

60. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inform Fusion.* (2020) 58:82–115. doi: 10.1016/j.inffus.2019.12.012

61. Chandler C, Foltz PW, Elvevåg B. Improving the applicability of AI for psychiatric applications through human-in-the-loop methodologies. *Schizophr Bull.* (2022) sbac038. doi: 10.1093/schbul/sbac038

62. Balcombe L, De Leo D. The potential impact of adjunct digital tools and technology to help distressed and suicidal men: an integrative review. *Front Psychol.* (2021) 12:796371. doi: 10.3389/fpsyg.2021.796371