

<https://doi.org/10.1038/s44184-024-00067-w>

Behavioral health and generative AI: a perspective on future of therapies and patient care

Emre Sezgin & Ian McKay

Check for updates

There have been considerable advancements in artificial intelligence (AI), specifically with generative AI (GAI) models. GAI is a class of algorithms designed to create new data, such as text, images, and audio, that resembles the data on which they have been trained. These models have been recently investigated in medicine, yet the opportunity and utility of GAI in behavioral health are relatively underexplored. In this commentary, we explore the potential uses of GAI in the field of behavioral health, specifically focusing on image generation. We propose the application of GAI for creating personalized and contextually relevant therapeutic interventions and emphasize the need to integrate human feedback into the AI-assisted therapeutics and decision-making process. We report the use of GAI with a case study of behavioral therapy on emotional recognition and management with a three-step process. We illustrate image generation-specific GAI to recognize, express, and manage emotions, featuring personalized content and interactive experiences. Furthermore, we highlighted limitations, challenges, and considerations, including the elements of human emotions, the need for human-AI collaboration, transparency and accountability, potential bias, security, privacy and ethical issues, and operational considerations. Our commentary serves as a guide for practitioners and developers to envision the future of behavioral therapies and consider the benefits and limitations of GAI in improving behavioral health practices and patient outcomes.

Rapid advancements in the fields of artificial intelligence (AI) have reshaped the healthcare and clinical practice landscape, delivering solutions that optimize processes and have the potential to enhance patient outcomes¹. These technological progress have led to a number of applications that address an extensive range of challenges encountered in healthcare. Such applications have introduced improvements in diagnostic accuracy and treatment planning^{2–4}, have personalized medical interventions^{5–7}, expedited drug discovery and development⁸, and broadened the reach of medical services through remote monitoring and telemedicine^{9,10}.

As a subset of AI, generative AI (GAI) models have gained increasing interest from the public and research communities owing to their demonstrated proficiency and computational efficiency in task execution. GAI models are principally neural networks or machine learning algorithms specialized to identify the patterns and structures within a dataset in order to generate variations or new data or content. Diffusion models¹¹, Generative adversarial networks¹², and Transformer networks¹³ are some of the widely adopted GAI models used in industry and research. These, in particular, have emerged as powerful tools in clinical research and the healthcare industry, where they serve pivotal roles in domains such as medical data processing¹⁴, genomic simulations and therapeutics¹⁵, synthetic data generation¹⁶, and image generation for diagnostics^{17,18}. The latter includes tasks such as high-resolution MRI reconstruction, image synthesis, contrast adjustment, and image enhancement, thereby broadening the scope of medical diagnostics¹⁹. Emerging at the forefront are medical foundation models, or multimodal foundation models, which combine various data types, including images, text, voice, and sensor data, to streamline clinical decision-making processes^{20,21}. Their versatility and broad applicability indicate that these models could potentially redefine healthcare practices, communication, and documentation. However, with the recent developments in image generation models including diffusion models (recent examples: [Stable Diffusion XL](#), [Midjourney v5](#)), generative adversarial networks, variational autoencoders²², and pre-trained transformers (e.g., [DALL-E 3](#)), there are new venues in healthcare to explore and leverage generated images in healthcare delivery.

Gap in behavioral health practice

Despite these advancements, the application of generative models in the behavioral health domain remains underexplored. Recent studies illustrate the application of large language models (LLMs) in augmenting empathy in online peer support platforms²³, supporting behavioral health services²⁴, building GAI integrated chatbot to support patients^{25,26}, and addressing behavioral health information seeking activities²⁷. Of the many proposed uses of GAI in behavioral health²⁸, image generation has yet to be explored²⁹. Both including and expanding these capabilities, the utilization of a

multimodal approach invites broader opportunities, especially with the inclusion of diverse data streams from personal health records, digital ecosystems, and medical records²⁰.

Meta-analyses examining psychotherapeutic interventions have found equally strong effect sizes for many common treatment modalities (e.g., CBT for Depression; 0.7 effect size)³⁰. However, when compared to control groups across several meta-analytic studies, higher variance has been found in intervention groups, suggesting heterogeneity in individual-level treatment effects. Aptitude by treatment interactions (ATIs)³¹, where an individual's unique circumstances, needs, or characteristics are prioritized for optimal outcomes, may explain this variability^{30,32}. Therefore, it is imperative to explore how AI models can be harnessed to augment treatment efficacy in evidence-based practices through the use of more personalized treatment components³⁰. In particular, generative models could support the creation of idiosyncratic and contextually relevant exercises for patients, thereby making therapeutic interventions more targeted and potentially more effective. GAI shows increased promise in conjunction with human-in-the-loop and human-AI collaboration approaches^{33,34}, which integrate human intelligence and feedback into the AI decision-making process³⁵. In line with that, GAI could be effective across a variety of therapeutic interventions to better meet the individual's unique needs and attributes^{28,36}. Within this scope, we provide a new perspective on adopting GAI in behavioral health.

Behavioral health and AI-based image generation

In clinical practice, images and visuals are often used to enrich conversation, aid in the learning process, and enhance therapeutic interventions^{37,38}. Provider (e.g., psychiatrist, psychologist, counselor, therapist, nurse, and social worker) initiated verbal conversations with visual cues in healthcare settings have already been used in a number of strategic ways, often to improve treatment engagement³⁹. Some examples include reducing emotional discomfort, improving patient education and learning, reducing healthcare disparities and stigma, increasing positive interactions between patients and providers, increasing personalization of treatment, and access to care⁴⁰.

In light of the growing evidence supporting the potential of generative models for images in various healthcare applications and integration (e.g., radiology and synthetic data¹⁹), it is necessary to investigate their utility in behavioral health services. The integration of generative models for images into existing evidence-based therapies, such as cognitive-behavioral therapy (CBT)⁴¹, dialectical behavioral therapy (DBT)⁴², and acceptance and commitment therapy (ACT)⁴³, could enhance the personalization and potential effectiveness of these interventions, ultimately leading to improved patient outcomes and patient-provider experiences. For example, an essential component of CBT includes teaching patients to modify their patterns of thinking (T), feeling (F), and behaving (B). In working with children who may struggle to understand the dynamic interaction between their thoughts, feelings, and behaviors, it may be useful to extend these interventions to include generated personalized images. These images may provide clarity and simplicity to a somewhat abstract concept, promote insight into one's maladaptive patterns, and provide strategies for change. For example, using pictures to depict one's tendency to catastrophize in social situations (T), often leads to worsening mood (F), and subsequent social isolation (B). A new "triangle" (CBT triangle showing how thoughts, feelings, and behaviors are connected) can be generated depending on the individual's goals or progress to illustrate ways new actions (e.g., approaching a social activity) may help reduce aversive emotions (e.g., sadness), and promote more balanced beliefs (e.g., "I met some new people, and they did not seem to hate me.").

As the literature is growing, and researchers are yet to study the impacts of this technology on behavioral health outcomes, we aim to provide a glimpse at the future to guide and inform practitioners and developers on how to yield the potential benefits of these innovative tools in clinical practice and patient care. As multimodal foundation models in medicine are growing with text, image, and audio⁴⁴, focusing on image generation and visual output (rather than multimodal engagement) would be a practical first step in introducing GAI in behavioral health therapies at clinics toward envisioning clinical utility and human-AI collaboration. In the following use case, we discuss the use of GAI to augment therapies with a focus on instilling emotional empowerment in the pediatric population.

Use case: emotional empowerment through the use of GAI-based images

Emotional intelligence broadly refers to a person's ability to navigate and comprehend the intricate world of emotions, both within oneself and in others. As a practical application of emotional intelligence, emotional empowerment involves the employment of emotional intelligence skills in one's daily life. Research has demonstrated the importance of fostering emotional empowerment in personal development and levels of well-being⁴⁵. Furthermore, research has demonstrated a relationship between the fostering of emotional empowerment in children and positive life outcomes, including enhanced resilience, school success, social skills, attention and concentration, decision-making abilities, and reduced long-term mental and behavioral health symptoms⁴⁶. The ability to understand and control one's emotions is crucial for children to navigate various challenges and complexities in life, and emotional empowerment can be a valuable tool for achieving this. We envision that the use of GAI models during pediatric behavioral therapies can augment the therapeutic process and contribute to promoting emotional empowerment. Figure 1 highlights "augmented" 3-step emotional empowerment therapy sessions with GAI.

Step 1: Recognizing emotions. The first step of emotional empowerment consists of developing an ability to recognize and name emotional states. The fostering of these skills in early years is vital for one's interpersonal functioning over the course of their lifespan⁴⁷, and has been inversely related to the development of psychological and behavioral challenges in both children and adults⁴⁸. These skills are also a necessary prerequisite to experiencing positive outcomes from the delivery of many therapeutic interventions⁴⁹.

Giving children opportunities to recognize a wide range of emotions in themselves and in others as well as scaling intensity may help support the acquisition of these skills early in life because children often learn better through visual cues and illustrations⁵⁰. GAI models may enhance learning through visual depictions of various facial expressions associated with different emotions. These images could serve as reference points for children to understand and recognize specific emotional states like happiness, sadness, anger, surprise, and more. This tool can be further personalized and more engaging via gamification, which has been broadly defined as the application of gamifying elements for treatment or intervention purposes^{51,52}. Prior use of these strategies to target emotional identification deficits has been related to improved emotion recognition and feedback among the pediatric population⁵³. To expand on this work, applications with image generation tools can present a series of personalized images (self or someone familiar) displaying different facial expressions, allowing children to match them with corresponding emotions (see Fig. 2). This interactive approach could make the learning process more enjoyable and engaging for children. Moreover, GAI models can be used to create personalized content that aligns with the child's cultural background, age group, or specific needs. By generating

Augmented Therapy with Generative AI

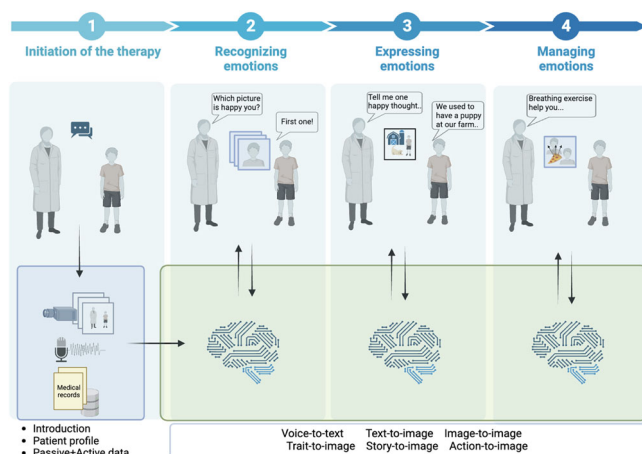


Fig. 1 | Augmented therapy with generative AI for image generation. Initiated with therapist and patient conversation, creating input including speech (prompts), visual (identifying setting, individuals, expressions), and medical records (clinician notes and patient history). Image, voice, and text inputs inform future prompts via conversations in a multi-agent relationship based on the prompt triggered (e.g., voice to text and then text to image by clinician's prompt; AI is listening to the story and using patient visual traits to create relevant images to the story). In each session, the prompt could be redefined based on the user feedback and the conversation (e.g., to manage emotion, the clinician can define actions, and those could be displayed as images).

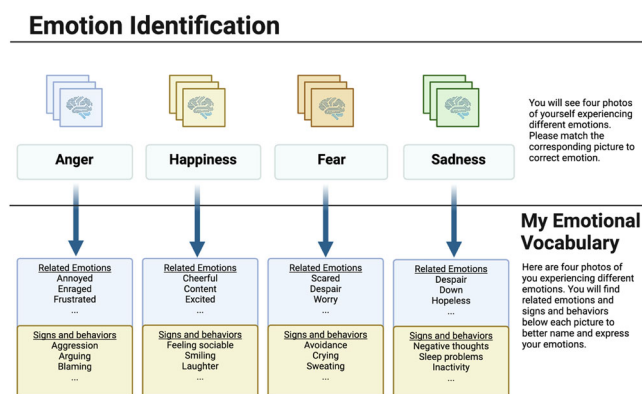


Fig. 2 | Emotion identification card. The card is including generated images of a patient reflecting anger, happiness, fear and sadness. Instructions are first asking patient to match images with emotions. Then, it guides to match images with corresponding emotional vocabulary items.

images that resemble the child's own environment, ethnicity, or familiar characters, the GAI can provide relatable visual cues that facilitate improved emotional understanding and identification. Please see <https://github.com/edsz7/supplementary>, for the sample images generated by AI (DALL-E 2 (<https://openai.com/dall-e-2>) and Stable Diffusion 1.5)⁵⁴ for research purposes only.

Step 2: Expressing emotions. Emotional self-expression, which refers to the verbal and non-verbal communication of one's emotions, is another necessary skill in building emotional empowerment in children.

Encouraging children to express their emotions in healthy ways is important for their emotional development and is a commonly used coping strategy that predicts better psychological and physical health in later life⁵⁵. There are several ways that GAI could be used to help enhance a child's ability to express their emotional state(s). For example, a GAI model could generate several images associated with specific emotions and provide corresponding labels or emotion words to help facilitate improved communication of a child's emotional state (see Fig. 2). For some individuals, expressing their feelings through words can be challenging or insufficient. Creating personalized images may provide a new layer for engagement and a non-verbal medium through which expression of feelings can be communicated and practiced. This could also help facilitate a sense of relief or catharsis by permitting individuals to externalize and express their emotions and could provide several opportunities for individuals to identify feelings in themselves and others. As a variety of creative therapeutic approaches are used to help enable individuals to better express their emotions⁵⁶, GAI may help enhance expressive therapeutic techniques through the development of more personalized and contextually relevant visual and auditory stimuli, facilitating unique and engaging avenues for emotional expression.

Step 3: Managing emotions. Lastly, GAI may prove useful in improving a child's ability to cope with various emotions through the provision of visual tools and resources to support emotional regulation. Broadly speaking, emotion regulation refers to the processes by which individuals can influence and exert control over their emotional state (i.e., which emotions they have, for how long as well as how strongly they experience and express them) and plays a vital role in young children's emotional and cognitive development and later academic achievement^{57,58}. The acquisition of emotion-regulation skills has been shown to improve the effectiveness of existing therapeutic interventions⁵⁸. GAI may help foster improved emotional regulation through the creation of personalized coping resources. More specifically, idiosyncratic emotion-related content tailored to a child's specific needs and experiences (e.g., coping cards, mood trackers, and emotion charts) could serve as visual reminders and tools to support one's experience in learning to cope with various emotions. Incorporating additional evidence-based techniques (e.g., guided imagery, relaxation practices, and mindfulness), GAI could also be used to create individualized calming and soothing visuals, such as a peaceful natural scene or comforting images (e.g., lounging on a beach chair). These can be generated by collaboration between patient, provider and GAI, and eventually help bolster existing treatment strategies aimed at reducing distress and may also provide a powerful distraction in the face of aversive emotional states.

Future of behavioral health support. Moving forward, the use of GAI-based applications may serve as a supportive tool in the delivery and/or implementation of a wide range of behavioral health treatments both during and outside of therapeutic settings. For example, GAI can enhance mindfulness-based interventions by creating visually and contextually relevant stimuli that facilitate relaxation, focus, and self-awareness. It can also be used to augment cognitive-behavioral therapies, such as exposure and response prevention through creating virtual environments or scenarios that mimic real-life situations that may trigger anxiety or phobias, allowing the patients to gradually expose themselves to specific fears in a controlled, convenient, and safe manner and learn coping skills to reduce associated distress (please see Fig. 3 and <https://github.com/edsz7/supplementary> for the sample images generated by AI for research purposes only). Broadly speaking, the personalization of various treatment

My Exposure Hierarchy

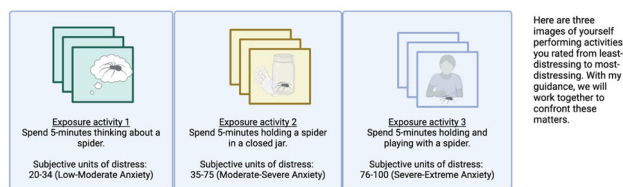


Fig. 3 | Emotion identification card with generated images. The card presents generated images of a patient reflecting anger, happiness, fear and sadness. The card shares instruction to guide patient to match images with emotions, and it guides to match images with corresponding emotional vocabulary items⁷⁹.

modalities has been found to improve the effectiveness of the treatments, reduce heterogeneity in treatment outcomes, and better meet the needs of the patients³⁰. Eventually, such personalization may help patients to be more compliant with out-of-session homework⁵⁹. In addition, the automation of therapeutic materials across various interventions may reduce the need for manual content creation by providers, thus possibly saving time and potential costs associated with developing personalized materials.

Outside of more traditional psychotherapeutic approaches, creative therapy modalities, such as art and music therapy, may also benefit from the utilization of GAI for content generation (e.g., by producing tailored stimuli to evoke desired emotional and cognitive responses) and may be woven into existing creative therapeutic strategies (i.e., narrative strengths-based storytelling). As GAI models become more sophisticated and multimodal, they may be considered for facilitating and/or supporting the development of novel therapeutic tools and approaches for behavioral health treatment, specifically for cases when it is hard to verbally communicate and express oneself.

GAI models could also assist in the provision of psychoeducation to help-seeking individuals by visually representing complicated concepts, processes, or emotions to improve understanding. Going beyond the clinical setting, the prospect of multimodal AI (e.g., image and video generation from voice inputs, text, or images) opens up new possibilities in the context of teletherapy and remote behavioral health services. Incorporated with virtual reality, extended reality, and augmented reality with LLMs, therapeutic services can also be expanded with conversational agents as personalized and guided self-therapy or self-training tools for patients²⁹.

Limitations, challenges, and considerations

Despite the benefits for healthcare service users and providers, this technology is not without potential limitations. In the context of the discussed use case, one limitation is the potential for generated images to lack authenticity and genuine emotional resonance. While GAI may be capable of producing visually and esthetically appealing images, it may struggle to capture or represent the necessary subtleties and complexities of human emotions accurately⁶⁰, and it might lack the nuanced understanding and empathy that human therapists provide⁶¹. Therefore, unsupervised GAI applications (e.g., self-generated images, text, and audio) may currently pose a challenge and risk to be used in multimodal therapy delivery. These qualities are usually exhibited by a human therapist and cannot be entirely replicated by an AI. The non-judgmental listening and rapport building, that characterize successful therapy, may be impacted by the introduction of AI in terms of practice and assessment. Therefore, it is important to keep the provider in the loop for professional guidance and control of the process. We suggest that any GAI implementation should be thoughtfully designed to

supplement rather than replace human interaction³⁵, with studies needed to assess the impacts and outcomes on the therapeutic relationship. As a visual communication tool, GAI may require to be used via iterative or chain-of-thought process to visualize closer to what is being expressed as a part of AI-human collaboration^{62,63}. A human feedback mechanism can also be used to train generative models to improve the outputs.

Building upon human-AI collaboration, inclusivity, and representativeness are necessary for equitable care delivery, and the AI-generated content should be relevant in terms of context and culture⁶⁴. Emotional expression and interpretation can vary significantly across cultures, and what may be considered empowering in one culture might not hold the same meaning in another. Cultural inclusion could be a challenging task for providers (i.e., require training in inclusivity and culture) as well as GAI (e.g., require training datasets on a variety of populations, languages, and cultures). However, GAI could also be “prompted” to produce inclusive and individualized images based on user input. Thus, it is crucial that providers are trained and familiar with how to operate GAI with manual input and deploy GAI that is tuned to take into account cultural diversity and sensitivity to avoid reinforcing stereotypes or inadvertently causing harm (see <https://github.com/edsz7/supplementary> for a prompting example).

GAI and provider collaboration, rather than the use of GAI only in self-therapy, can also reduce encountering unexpected risks regarding patients safety and quality of care³⁵. The AI-behavioral health convergence offers significant potential for care improvement. Therefore, we need to consider ethical implementations, particularly towards building privacy-preserving and trustworthy models, and supportive human interaction in therapeutic processes⁶⁵. For instance, considering the potential bias and non-representative dataset (images, pictures, illustrations, figures, and annotations) used for GAI training, human-in-the-loop quality assurance, and content validation are necessary steps before using in practice⁶⁶. Such practices may need to be defined or prescribed further by standards and guidelines towards GAI use and development, such as vetting of training datasets against biases as well as intellectual property right infringement^{36,61}, and integrating fairness and equity approaches to ensure ethical and effective use in mental and behavioral health applications⁶⁷. In line with that, an additional layer of transparency and accountability is needed in practice^{36,68}, such as adding consent in clinical workflow protocol to disclose the risks and benefits of such approaches with patients, caregivers, and families, and adherence to medicolegal practices, data protection policies and legislations (e.g., HIPAA and GDPR)⁶¹. The number of government agencies and technology providers have been releasing guidelines⁶⁹, regulations⁷⁰, and call for action^{71,72} towards ethical and trustworthy GAI use and development, which could be informative toward GAI in behavioral healthcare practices.

The current use of AI as clinical tools or support systems so far has led to only a moderate improvement in cost-effectiveness⁷³. Even though the marginal cost of AI-based decisions is minimal, its overall impact on net costs is unknown⁷⁴. Therefore, the aforementioned considerations may further require investigation of operational effectiveness, including strategizing integration and workflow adjustment, system design updates, creating infrastructure (e.g., Cloud backend architecture and EHR integration) for deployment, planning of GAI-based practice reimbursement mechanisms, patient engagement testing and treatment adherence, developing patient and provider training curriculum to effectively use and incorporate GAI in practices, and cost-effectiveness analysis⁷⁵.

In addition to the above considerations, the potential misuse of GAI (e.g., deepfake) is a significant concern⁷⁶. AI-generated images, if used unethically, can manipulate emotions, possibly causing psychological harm⁷⁷. There could be instances where these powerful tools are

used with malicious intent, such as promoting propaganda or spreading misinformation about behavioral health. It is critical to ensure that all data used by GAI is appropriately anonymized and secured following standard encryption and cybersecurity practices. However, there could be a risk of identification of protected health information, especially with the use of patient audio and image data as input. A comprehensive framework of data governance should be in place to safeguard against any unauthorized access, misuse of data and handling data breach^{61,78}. Regulatory frameworks and stringent monitoring processes need to be in place to prevent such misuse, ensuring that the power of GAI is harnessed for beneficial purposes alone.

Conclusions

In this paper, we highlight GAI impact on behavioral health, exemplifying image generation support on therapy towards recognizing, expressing, and managing emotions. The potential for integrating GAI models in therapeutic settings is open for exploration. GAI could be used for immersive therapeutic experiences leading to engaging therapy sessions, with personalized environments generated in real time based on the patient's emotional state and therapeutic needs. Given the potential of GAI to transform behavioral healthcare, it is crucial for researchers and practitioners to collaborate in the development, evaluation, governance, and implementation of these tools and provide guidelines for practices and implications. By conducting rigorous, interdisciplinary research and trials at the intersection of behavioral health and GAI, we can ensure that these innovative solutions are optimally harnessed to improve patient care and promote overall well-being.

Emre Sezgin^{1,2,4} ✉ & Ian McKay^{2,3,4}

¹The Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH, USA. ²The Ohio State University College of Medicine, Columbus, OH, USA. ³Department of Psychiatry and Behavioral Health, Nationwide Children's Hospital, Columbus, OH, USA. ⁴These authors contributed equally: Emre Sezgin, Ian McKay.

✉ e-mail: sezgin.7@osu.edu

Received: 27 November 2023; Accepted: 6 April 2024;

Published online: 07 June 2024

References

- Rajpurkar, P., Chen, E., Banerjee, O. & Topol, E. J. AI in health and medicine. *Nat. Med.* **28**, 31–38 (2022).
- Nazarian, S., Glover, B., Ashrafian, H., Darzi, A. & Teare, J. Diagnostic accuracy of artificial intelligence and computer-aided diagnosis for the detection and characterization of colorectal polyps: systematic review and meta-analysis. *J. Med. Internet Res.* **23**, e27370 (2021).
- Aggarwal, R. et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ Digit. Med.* **4**, 65 (2021).
- Wang, C., Zhu, X., Hong, J. C. & Zheng, D. Artificial intelligence in radiotherapy treatment planning: present and future. *Technol. Cancer Res. Treat.* **18**, 1533033819873922 (2019).
- Zhou, Q., Chen, Z.-H., Cao, Y.-H. & Peng, S. Clinical impact and quality of randomized controlled trials involving interventions evaluating artificial intelligence prediction tools: a systematic review. *NPJ Digit. Med.* **4**, 154 (2021).
- Ho, D. et al. Enabling technologies for personalized and precision medicine. *Trends Biotechnol.* **38**, 497–518 (2020).
- Bae, S. W. et al. Leveraging mobile phone sensors, machine learning, and explainable artificial intelligence to predict imminent same-day binge-drinking events to support just-in-time adaptive interventions: algorithm development and validation study. *JMIR Form. Res.* **7**, e39862 (2023).
- Paul, D. et al. Artificial intelligence in drug discovery and development. *Drug Discov. Today* **26**, 80–93 (2021).
- Kuziemy, C. et al. Role of Artificial Intelligence within the Telehealth Domain. *Yearb. Med. Inform.* **28**, 35–40 (2019).
- Sujith, A. V. L. N., Sajja, G. S., Mahalakshmi, V., Nuhmani, S. & Prasanalakshmi, B. Systematic review of smart health monitoring using deep learning and artificial intelligence. *Neurosci. Inform.* **2**, 100028 (2022).
- Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst.* **33**, 6840–6851 (2020).
- Goodfellow, I. et al. Generative adversarial networks. *Commun. ACM* **63**, 139–144 (2020).
- Vaswani, A. et al. Attention is all you need. In: *Advances in neural information processing systems* 5998–6008 (2017).
- Thirunavukarasu, A. J. et al. Large language models in medicine. *Nat. Med.* **29**, 1930–1940 (2023).
- Huang, K. et al. Artificial intelligence foundation for therapeutic science. *Nat. Chem. Biol.* **18**, 1033–1036 (2022).
- D'Amico, S. et al. Synthetic data generation by artificial intelligence to accelerate research and precision medicine in hematology. *JCO Clin. Cancer Inf.* **7**, e2300021 (2023).
- Zhao, J., Hou, X., Pan, M. & Zhang, H. Attention-based generative adversarial network in medical imaging: a narrative review. *Comput. Biol. Med.* **149**, 105948 (2022).
- Liu, Y. et al. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. *Med. Phys.* **47**, 2472–2483 (2020).
- Koochi-Moghadam, M. & Bae, K. T. Generative AI in medical imaging: applications, challenges, and ethics. *J. Med. Syst.* **47**, 94 (2023).
- Acosta, J. N., Falcone, G. J., Rajpurkar, P. & Topol, E. J. Multimodal biomedical AI. *Nat. Med.* **28**, 1773–1784 (2022).
- Bommasani, R. et al. On the opportunities and risks of foundation models. *arXiv* <https://arxiv.org/abs/2108.07258> (2021).
- Bond-Taylor, S., Leach, A., Long, Y. & Willcocks, C. G. Deep generative modelling: a comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 7327–7347 (2022).
- Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C. & Althoff, T. Human-AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nat. Mach. Intell.* **5**, 46–57 (2023).
- van Heerden, A. C., Pozuelo, J. R. & Kohrt, B. A. Global mental health services and the impact of artificial intelligence-powered large language models. *JAMA Psychiatry* **80**, 662–664 (2023).
- Pandey, S. & Sharma, S. A comparative study of retrieval-based and generative-based chatbots using deep learning and machine learning. *Healthc. Analytics* **3**, 100198 (2023).
- Bird, J. J. & Lotfi, A. Generative transformer chatbots for mental health support: a study on depression and anxiety. In: *Proceedings of the 16th international conference on pervasive technologies related to assistive environments*, 475–479 (Association for Computing Machinery, 2023). <https://doi.org/10.1145/3594806.3596520>.
- Sezgin, E., Chekeni, F., Lee, J. & Keim, S. Clinical accuracy of large language models and google search responses to postpartum depression questions: cross-sectional study. *J. Med. Internet Res.* **25**, e49240 (2023).
- Schuller, S. M. & Morris, R. R. Clinical science and practice in the age of large language models and generative artificial intelligence. *J. Consult. Clin. Psychol.* **91**, 559–561 (2023).
- Pataranutopon, P. et al. AI-generated characters for supporting personalized learning and well-being. *Nat. Mach. Intell.* **3**, 1013–1022 (2021).
- Nye, A., Delgadillo, J. & Barkham, M. Efficacy of personalized psychological interventions: a systematic review and meta-analysis. *J. Consult. Clin. Psychol.* **91**, 389–397 (2023).
- Cronbach, L. J. & Snow, R. E. Aptitudes and instructional methods: a handbook for research on interactions. **574**, (1977).
- Zhong, Y. et al. The issue of evidence-based medicine and artificial intelligence. *Asian J. Psychiatr.* **85**, 103627 (2023).
- Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J. & Fernández-Leal, Á. Human-in-the-loop machine learning: a state of the art. *Artif. Intell. Rev.* **56**, 3005–3054 (2023).
- Park, S. Y. et al. Identifying challenges and opportunities in human-AI collaboration in healthcare. In: *Conference companion publication of the 2019 on computer supported cooperative work and social computing*, 506–510 (Association for Computing Machinery, 2019). <https://doi.org/10.1145/3311957.3359433>.
- Sezgin, E. Artificial intelligence in healthcare: complementing, not replacing, doctors and healthcare providers. *Digit. Health* **9**, 20552076231186520 (2023).
- Zhong, Y. et al. The Artificial intelligence large language models and neuropsychiatry practice and research ethic. *Asian J. Psychiatr.* **84**, 103577 (2023).
- Holmes, E. A., Arntz, A. & Smucker, M. R. Imagery rescripting in cognitive behaviour therapy: images, treatment techniques and outcomes. *J. Behav. Ther. Exp. Psychiatry* **38**, 297–305 (2007).
- Lang, P. J. Imagery in therapy: an information processing analysis of fear. *Behav. Ther.* **8**, 862–886 (1977).
- Hafner, C., Schneider, J., Schindler, M. & Brailard, O. Visual aids in ambulatory clinical practice: experiences, perceptions and needs of patients and healthcare professionals. *PLoS One* **17**, e0263041 (2022).
- Boland, L. et al. An experimental investigation of the effects of perspective-taking on emotional discomfort, cognitive fusion and self-compassion. *J. Context. Behav. Sci.* **20**, 27–34 (2021).
- Beck, A. T. Thinking and depression. II. Theory and therapy. *Arch. Gen. Psychiatry* **10**, 561–571 (1964).
- Brasfield, C. Cognitive-behavioral treatment of borderline personality disorder. *Behav. Res. Ther.* **32**, 899 (1994).
- Hayes, S. C., Strosahl, K. D. & Wilson, K. G. Acceptance and commitment therapy: an experiential approach to behavior change. (Guilford Publications, 1999).
- Moor, M. et al. Foundation models for generalist medical artificial intelligence. *Nature* **616**, 259–265 (2023).

45. Christens, B. D., Collura, J. J. & Tahir, F. Critical hopefulness: a person-centered analysis of the intersection of cognitive and emotional empowerment. *Am. J. Community Psychol.* **52**, 170–184 (2013).
46. Kotsou, I., Mikolajczak, M., Heeren, A., Grégoire, J. & Leys, C. Improving emotional intelligence: a systematic review of existing work and future challenges. *Emot. Rev.* **11**, 151–165 (2019).
47. Lewis, G. J., Lefevre, C. E. & Young, A. W. Functional architecture of visual emotion recognition ability: a latent variable approach. *J. Exp. Psychol. Gen.* **145**, 589–602 (2016).
48. Brackett, M. A., Rivers, S. E. & Salovey, P. Emotional intelligence: implications for personal, social, academic, and workplace success. *Soc. Personal. Psychol. Compass* **5**, 88–103 (2011).
49. Morie, K. P., Crowley, M. J., Mayes, L. C. & Potenza, M. N. The process of emotion identification: considerations for psychiatric disorders. *J. Psychiatr. Res.* **148**, 264–274 (2022).
50. Bobek, E. & Tversky, B. Creating visual explanations improves learning. *Cogn. Res. Princ. Implic.* **1**, 27 (2016).
51. Cheng, V. W. S., Davenport, T., Johnson, D., Vella, K. & Hickie, I. B. Gamification in apps and technologies for improving mental health and well-being: systematic review. *JMIR Ment. Health* **6**, e13717 (2019).
52. Xie, H. A scoping review of gamification for mental health in children: uncovering its key features and impact. *Arch. Psychiatr. Nurs.* **41**, 132–143 (2022).
53. Nicolaidou, I., Aristeidis, L. & Lambrinos, L. A gamified app for supporting undergraduate students' mental health: a feasibility and usability study. *Digit Health* **8**, 20552076221109059 (2022).
54. stabilityai/stable-diffusion-2-1 · Hugging face. <https://huggingface.co/stabilityai/stable-diffusion-2-1>.
55. Stanton, A. L. & Low, C. A. Expressing emotions in stressful contexts: benefits, moderators, and mechanisms. *Curr. Dir. Psychol. Sci.* **21**, 124–128 (2012).
56. CSEFEL: center on the social and emotional foundations for early learning. <https://csefel.vanderbilt.edu/resources/family.html>.
57. Gross, J. J. The emerging field of emotion regulation: an integrative review. *Rev. Gen. Psychol.* **2**, 271–299 (1998).
58. Moltrecht, B., Deighton, J., Patalay, P. & Edbrooke-Childs, J. Effectiveness of current psychological interventions to improve emotion regulation in youth: a meta-analysis. *Eur. Child Adolesc. Psychiatry* **30**, 829–848 (2021).
59. Tang, W. & Kreindler, D. Supporting homework compliance in cognitive behavioural therapy: essential features of mobile apps. *JMIR Ment. Health* **4**, e20 (2017).
60. Azuaje, G. et al. Exploring the use of AI text-to-image generation to downregulate negative emotions in an expressive writing application. *R. Soc. Open Sci.* **10**, 220238 (2023).
61. Vial, T. & Almon, A. Artificial intelligence in mental health therapy for children and adolescents. *JAMA Pediatr.* **177**, 1251–1252 (2023).
62. Harvey, W. & Wood, F. Visual chain-of-thought diffusion models. *arXiv* <https://arxiv.org/abs/2303.16187> (2023).
63. Weisz, J. D. et al. Design principles for generative AI applications. *arXiv* <https://arxiv.org/html/2401.14484v1> (2024).
64. Cachat-Rosset, G. & Klarsfeld, A. Diversity, equity, and inclusion in artificial intelligence: an evaluation of guidelines. *Appl. Artif. Intell.* **37**, 2176618 (2023).
65. Lee, E. E. et al. Artificial intelligence for mental health care: clinical applications, barriers, facilitators, and artificial wisdom. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **6**, 856–864 (2021).
66. Sezgin, E., Sirrianni, J. & Linwood, S. L. Operationalizing and implementing pretrained, large artificial intelligence linguistic models in the US health care system: outlook of generative pretrained transformer 3 (GPT-3) as a service model. *JMIR Med. Inf.* **10**, e32875 (2022).
67. Timmons, A. C. et al. A call to action on assessing and mitigating bias in artificial intelligence applications for mental health. *Perspect. Psychol. Sci.* **18**, 1062–1096 (2023).
68. Boch, S., Sezgin, E. & Linwood, S. Ethical artificial intelligence in paediatrics. *Lancet Child Adolesc. Health* **6**, 833–835 (2022).
69. Interim guidance on government use of public generative AI tools - November 2023. <https://architecture.digital.gov.au/guidance-generative-ai>.
70. European Parliament-SpokespersonGuillot, J. D. EU AI Act: first regulation on artificial intelligence. https://www.europarl.europa.eu/pdfs/news/expert/2023/6/story/20230601STO93804/20230601STO93804_en.pdf (2023).
71. International community must urgently confront new reality of generative, artificial intelligence, speakers stress as Security Council debates risks, rewards. <https://press.un.org/en/2023/sc15359.doc.htm> (2023).
72. The White House. Executive order on the safe, secure, and trustworthy development and use of artificial intelligence. <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/> (2023).
73. Gomez Rossi, J., Rojas-Perilla, N., Krois, J. & Schwendicke, F. Cost-effectiveness of artificial intelligence as a decision-support system applied to the detection and grading of melanoma, dental caries, and diabetic retinopathy. *JAMA Netw. Open* **5**, e220269 (2022).
74. Hendrix, N., Veenstra, D. L., Cheng, M., Anderson, N. C. & Verguet, S. Assessing the economic value of clinical artificial intelligence: challenges and opportunities. *Value Health* **25**, 331–339 (2022).
75. Allen, M. R. et al. Navigating the doctor-patient-AI relationship - a mixed-methods study of physician attitudes toward artificial intelligence in primary care. *BMC Prim. Care* **25**, 42 (2024).
76. Mirsky, Y. & Lee, W. The creation and detection of deepfakes: a survey. *ACM Comput. Surv.* **54**, 1–41 (2021).
77. Brundage, M. et al. The malicious use of artificial intelligence: forecasting, prevention, and mitigation. *arXiv* <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf> (2018).
78. Janssen, M., Brous, P., Estevez, E., Barbosa, L. S. & Janowski, T. Data governance: organizing data for trustworthy artificial intelligence. *Gov. Inf. Q.* **37**, 101493 (2020).
79. Tanner, B. A. Validity of global physical and emotional SUDS. *Appl. Psychophysiol. Biof.* **37**, 31–34 (2012).

Acknowledgements

The authors thank John Ackerman for his valuable comments and feedback on the manuscript. Figures are created with BioRender.com.

Author contributions

E.S. conceived the idea. E.S. and I.M. contributed to ideation and planning; contributed equally in drafting and finalizing the manuscript; reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Emre Sezgin.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024, corrected publication 2024