Current AI Should Extend (Not Replace) Human Care in Mental Health

Nick Kabrel[1,2*], Elizabeth C. Stade[3,4], Jaan Aru[5], and Johannes C. Eichstaedt[4,6*]

[1]Department of Psychology, University of Zurich, ZH, Switzerland
[2]Digital Society Initiative, University of Zurich, ZH, Switzerland
[3]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA
[4]Institute for Human-Centered Artificial Intelligence, Stanford University, Stanford, CA, USA
[5]Institute of Computer Science, University of Tartu, Tartu, Estonia
[6]Department of Psychology, Stanford University, Stanford, CA, USA

nick.kabrel@gmail.com johannes.stanford@gmail.com

*Corresponding authors

Abstract

Conversational AI (CAI) is being deployed in mental healthcare settings, including in chatbots designed to deliver therapy directly to patients. However, in light of many unresolved questions, this focus on replacing human therapists may be premature, while overlooking safer, immediate opportunities for CAI to extend the existing systems of care in areas where human support and resources are minimal or absent. Specifically, we identify opportunities for CAI to facilitate (1) prevention and screening, (2) assist while patients are on waitlists, (3) support between sessions, and (4) provide follow-up care after human therapy. Rather than focusing on replacing clinicians in therapy, CAI should be strategically integrated into the broader care infrastructure to make the care system more integrated, less burdensome, and increase its capacity, thereby enhancing its effectiveness. We review preliminary studies that suggest that this is plausible. Across stages of healthcare delivery, we highlight specific opportunities and candidate mechanisms for effective and ethical implementation of such blended human-AI care.

Keywords: Psychotherapy, Conversational AI, Artificial Intelligence, Mental Health Care.

1. Introduction

Conversational Artificial Intelligence (CAI) has been increasingly applied in mental healthcare settings since its advent in November 2022. Several randomized controlled trials (RCTs) have demonstrated the efficacy of elements of psychotherapy delivered by CAI-based chatbots (reviewed in He et al., 2023; Li et al., 2023; Zhong et al., 2024; see also Heinz et al., 2025). For example, a recent first-of-its-kind RCT of a fully generative chatbot showed significant improvements in symptoms of Major Depressive Disorder (8 weeks: d=0.9), Generalized Anxiety Disorder (8 weeks: d=0.8), and weight concerns (8 weeks: d=0.6) (Heinz et al., 2025). The qualitative study by Siddals et al. (2024) has also found positive impacts of a CAI-based chatbot, including high patient engagement and reduction of symptoms related to trauma and loss.

However, while these preliminary results are promising, there is insufficient evidence to suggest that direct therapy delivery is effective and safe in real-world mental healthcare systems, and that patient interactions with stand-alone CAI can be sustained long enough to be effective (Mohr et al., 2017, 2018). Despite that, numerous companies and research groups worldwide are working toward building chatbots designed to act as therapists, often envisioning a future in which CAI can fully substitute for human clinicians. Yet, many open questions remain about what it would take for a CAI system to approximate the real-world effectiveness of human-level therapy (Stade et al., 2025).

Instead of focusing on substituting for human care, we suggest that CAI should be deployed where it has the greatest relative potential to improve the status quo within mental healthcare systems. Following a recent saying in the tech industry that "AI is not better than the best human, but often better than the best available human," there are many stages of the mental health care pathway where, currently, "the best available human" is nobody. For example, real-world care is often limited in prevention, early detection, while patients are on waitlists, between sessions, and post-treatment.

This current status quo leaves healthcare systems as a whole less effective. Specifically, in most systems, universal screening and prevention efforts are often insufficiently implemented, resulting in undetected cases and unsupported individuals, despite the demonstrated value of screening (e.g., Dodge, 2018; Ridenour et al., 2022). Once a referral is made, patients often face long waitlists. For instance, in the U.S., wait times commonly range from 2–3 weeks to several months (Steinman et al., 2015; Blech et al., 2017) with an average delay of around seven weeks, with millions awaiting care (Peipert et al., 2022). Moreover, therapy sessions are often spaced two to three weeks apart due to insufficient clinicians. Between-session support is typically minimal or inconsistent, compounding the gap in care (Orlinsky et al., 1993). Likewise, when patients complete psychotherapy, relapse prevention, and aftercare programs are often minimal (Henneman et al., 2018), even though the risk of relapse is high, particularly for those with chronic conditions (Hölzel et al., 2011).

2. Considering the entire mental healthcare pathway

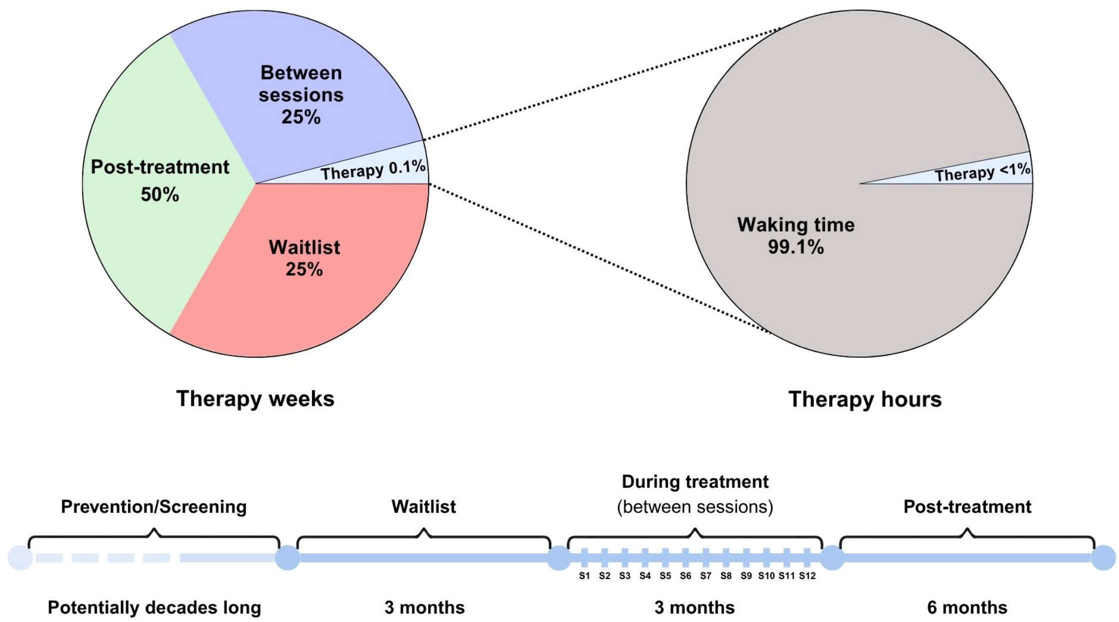
Experts have long emphasized that a focus on the patient in isolation is insufficient to reduce the overall population burden of mental disorders, pointing instead to systems-level interventions such as universal screening and prevention (Chater & Loewenstein, 2022; Dodge et al., 2024). In these overlooked spaces, even modest AI-based interventions could produce meaningful benefits. We argue that they are a promising target for CAI development efforts in the near term.

Although psychotherapy is often seen as the central to care, meaningful change can begin well before treatment starts and must often be sustained long after it ends (see **Fig. 1**). Indeed, empirical evidence shows that mental health disorders can be prevented before they emerge (Dodge et al., 2024; even though the number needed to treat remains substantial, Cuijpers et al., 2021) and that aftercare programs help reduce relapses after release from treatment (Krijnen-de Bruin, 2022; Chen et al., 2022; Zhou et al., 2023).

Therefore, we propose a "figure ground reversal": to focus on using CAI in mental healthcare to support mental healthcare pathways outside of psychotherapy (which may very well be best continued to be delivered by human clinicians). In this paper, we offer four entry points for integrating CAI into the mental healthcare pathway: (1) for prevention – before there is an explicit need for therapy, (2) waitlist – when there is a need for psychotherapy but psychotherapy has not begun, (3) intersession processes – between therapy sessions when therapy is ongoing, and (4) sustainment – after formal psychotherapy has concluded (see **Fig. 2**).

The pathways have in common that they always have "a therapist in the loop" to supervise and mitigate risk (following Stade et al., 2024). There have been numerous reported cases in which CAI chatbots gave inappropriate, unethical, or dangerous advice (Sedlakova & Trachsel, 2023; Blease & Rodman,

Figure 1: Proportional timeline of the mental healthcare pathway.



Note. This figure represents an illustrative, typical one-year timeline of the mental healthcare pathway, highlighting the relative proportions of time spent in each stage. Psychotherapy itself typically occupies only a small fraction of the overall process, with much more time spent on prevention, waiting for treatment, between sessions, and in aftercare. We relied on published estimates to estimate relative time use: 3 months (25% of the year) for the waitlist period (based on European healthcare data showing average wait times of 2–5 months, Civio, 2021); for the therapy phase, we assumed 3 months of weekly psychotherapy (0.1% of the year spent in session, based on 12 one-hour sessions, and 24.9% in intersession periods, Flückiger et al., 2020). The remaining 6 months (50%) were allocated to aftercare, which matches durations reported for follow-up and maintenance programs (e.g., APA, 2018). For the right pie chart, we assumed one hour of weekly psychotherapy and 16 hours of waking activity per day.

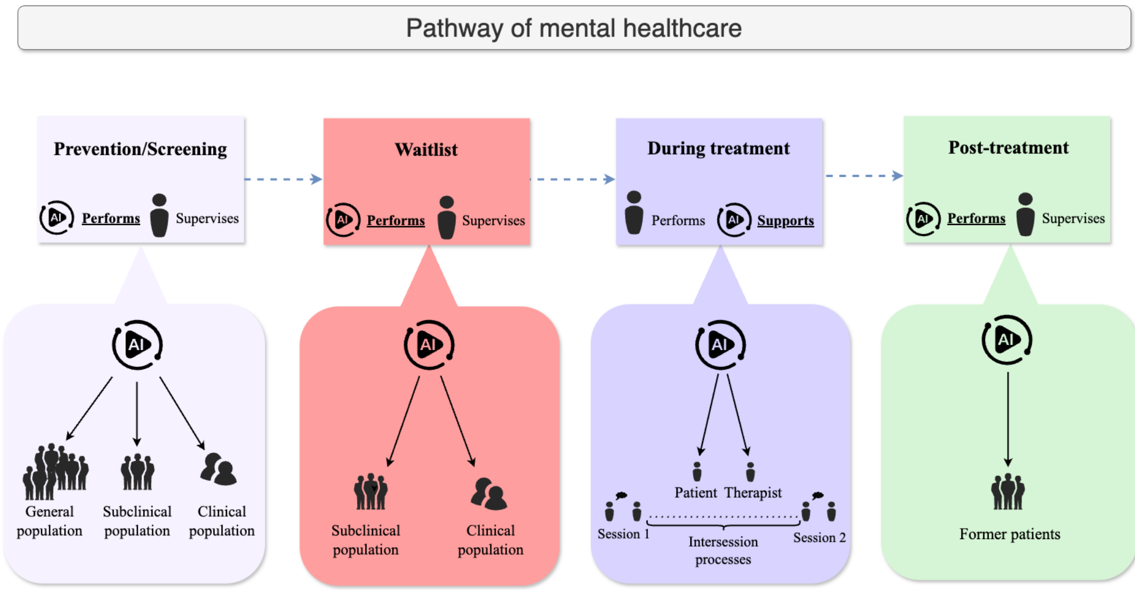
2025). When a CAI tool interacts directly with patients without therapist oversight, it creates a critical safety gap. This not only poses risks to the client's safety but can also disrupt the clinician's therapeutic plans. By contrast, a model in which the clinician reviews and selectively uses the content generated by CAI offers a safer and more controlled alternative. Of course, such implementations require adherence to healthcare privacy standards, but most major CAI systems (including ChatGPT) are available in HIPAA-compliant versions, and such systems can also be run locally on university or hospital servers.

2.1. Prevention

There are three types of prevention models: (1) universal prevention, which is aimed at targeting the overall population irrespective of the risk factors, (2) selective prevention, which targets subpopulations at risk of developing a disorder based on bio-psycho-social risk factors, and (3) indicated prevention, which targets individuals with subthreshold symptoms of a mental disorder who do not yet qualify for a diagnosis (Mrazek & Haggerty, 1994; National Research Council, 2009).

Despite being effective and cost-effective (Cuijpers et al., 2021; Le et al., 2021; Sander et al., 2016), mental health prevention strategies do not receive sufficient governmental funding,

Figure 2: Pathway of care augmented by CAI.



Note. There are at least four states in mental health care that can be supported by CAI, either with CAI performing a task under human supervision or a human performing a task with CAI support (Miner et al., 2019). The prevention stage can target three distinct populations: the general population, individuals with subclinical conditions, and patients with diagnosed clinical conditions. In the screening stage, individuals exhibiting clinical symptoms are evaluated and may be placed on a waitlist. During the waitlist period, prospective patients may receive support in preparation for psychotherapy, or CAI may conduct symptom assessment/intake. During the treatment stage, both the patient and therapist may be supported between therapy sessions. Finally, the post-treatment stage provides patients with ongoing support to maintain therapeutic outcomes, reinforce acquired skills, monitor for potential relapse, and facilitate timely readmission, if necessary.

as they require long periods of engagement and are costly (Jacka & Reavley, 2014). Using CAI for prevention is promising since it can address some of these limitations. First, CAI technology can be implemented on all three levels of prevention classification: it can target the general population as well as (sub)clinical populations (Fig. 2).

To illustrate, for universal prevention targeting the general population, CAI can be embedded into existing digital ecosystems, such as public health apps, workplace wellness platforms, or educational settings, as optional, user-initiated tools offering psychoeducation, stress-reduction practices, or mood tracking. Also, as noted by Dodge et al. (2024), embedding services in public institutions like schools, primary pediatric care, and other health care centers, churches, and neighborhood centers could increase participation. CAI can be offered as part of routine screening in primary care, for example, during annual check-ups or wellness consultations where patients already expect health-related discussions. For selective and indicated prevention, CAI can support individuals at elevated risk or with subclinical symptoms through tailored interventions, such as through integration with electronic health systems or referrals from school counselors or general practitioners.

Such a strategy is relatively affordable since it does not require intensive human involvement. Furthermore, studies have shown that digital prevention programs administered through web applications and websites have a small but significant effect in preventing mental disorders (Sander et al., 2016). Since it is known that more engaging web-based programs lead to better effectiveness (Karekla et al., 2019), CAI warrants investigation for prevention as it may enhance adherence through greater flexibility and personalization, and thus, engagement (e.g., Heinz et al., 2025).

Extensive research has identified numerous risk and protective factors associated with the development of mental disorders (e.g., Hölzel et al., 2011; Vink et al., 2008; Cairns et al., 2014; Shortt & Spence, 2006). CAI could help individuals strengthen protective factors (e.g., emotion regulation or social connectedness) and detect early signs of risk. Such tools could also empower individuals to activate existing psychological and social resources, such as reaching out to trusted peers or engaging in meaningful activities, thereby aligning with a strength-based approach to mental health prevention (Tse et al., 2016).

Effective psychoeducation alone might help increase awareness about mental health, address common biases, and reduce stigma. The psychoeducation skills of ChatGPT have been recently evaluated by psychotherapists and found to be adequate (Maurya et al., 2025).

Future research should systematically investigate how CAI can be most effectively tailored and deployed across the different levels of prevention. This includes identifying which conversational strategies, content formats, and delivery contexts are most beneficial for diverse populations. Controlled trials comparing CAI-based interventions with existing digital and human-delivered preventive approaches are needed to evaluate their effectiveness, adherence, and long-term impact.

2.2. Screening

Prior work has found that general practitioners do not have the time or resources to conduct even short routine screenings for mental disorders (Parker et al., 2020; Zhang et al., 2019). CAI could support assessment, for example, by engaging patients in short conversations at the beginning or end of a medical visit to flag potential mental health concerns. This could both enhance prevention and reduce the burden on already overloaded general practitioners. Similarly, a brief, structured check-in with a chatbot during school or workplace health assessments could deliver programs that reduce risk factors for anxiety and depression (Schleider & Weisz, 2016).

Although it is perhaps too early to speak about CAI diagnosing a condition and severity in real-world settings, some studies suggest that it may become a feasible option in the future (Johri et al., 2025). For example, studies show that chatbots might be able to detect mental issues by asking sets of questions based on screening questionnaires, similar to clinicians (Johri et al., 2025; Levkovich & Elyoseph, 2023; Vaidyam et al., 2019; Kjell et al., 2023). Research should explore how to integrate CAI ethically and equitably into real-world care infrastructures—such as schools, workplaces, or primary care settings—without overwhelming users or providers.

2.3. Waitlist periods

Long waitlist time is a formidable challenge for mental healthcare systems across the world (Mohr et al., 2006). For instance, the general waitlist time for psychotherapy is up to 3 months in the US (Steinman et al., 2015; Blech et al., 2017) and approximately 4-5 months in Germany (Civio, 2021). There's emerging acceptance of the idea that waitlists are harmful to people's mental health: in general, the condition of prospective therapy patients deteriorates while being on a waitlist (Steinert et al., 2017; Cristea, 2019; van Dijk et al., 2023; Reichert & Jacobs, 2018). Building on previous developments in computerized CBT programs (Huang et al., 2024), using CAI-based tools to support individuals during the wait for treatment represents a promising and scalable solution during a time when they would otherwise see deterioration in their symptoms.

For example, Rollwage et al. (2023) recently tested a CAI-based self-referral tool. This conversational chatbot is integrated into the service's website and assists patients by collecting the necessary intake information (eligibility criteria, demographic information, etc). This additional information includes free-text input regarding the patient's presenting symptoms as well as standardized, clinically validated routine outcome measures and screening questions. The outcomes of this program showed that the use of CAI increased clinical efficiency by reducing the time clinicians spend on preliminary mental health assessments and reducing wait times for patients, reducing dropout rates, and increasing recovery rates (Rollwage et al., 2023). The study found the sharing of clinically relevant information obtained through CAI with service providers to be particularly important.

Another crucial aspect is that many patients require a substantial number of therapy sessions before experiencing meaningful improvement; given that 9–13 sessions can be needed for meaningful clinical change (Hansen et al., 2002; Harper et al., 2024). One promising strategy to reduce this delay is to offer low-intensity interventions prior to therapy. For example, single-session interventions have been shown to produce measurable improvements for individuals awaiting treatment (Schleider & Weisz, 2016). Likewise, single-session consultations (SSC), which prompt patients to identify and implement achievable behavior changes while on waitlists, have demonstrated feasibility and symptom reduction in both in-person and telehealth formats (Schleider et al., 2021; Sung et al., 2023).

In addition, there is strong evidence that role induction—providing patients with information about what to expect from therapy and how to engage effectively—improves early engagement and outcomes (Swift et al., 2023). Similarly, studies indicate that motivational interviewing (MI) before treatment can enhance retention and outcomes (Westra et al., 2006; 2009; Carroll et al., 2006). Together, these findings suggest that evidence-based pre-therapy interventions can accelerate therapeutic gains and reduce dropout. CAI could provide a scalable way to deliver such interventions, improving readiness for therapy while helping to alleviate systemic bottlenecks in access to care.

That said, future research could focus on proof-of-concept studies aimed at testing whether chatbot interventions can activate specific psychological mechanisms known to support treatment outcomes. For instance, a brief chatbot conversation conducted before the first therapy session might increase a client's positive expectations toward treatment, which is a well-established predictor of therapeutic success (Constantino et al., 2011). In medicine, proof-of-concept studies demonstrate that an intervention engages its intended mechanism (e.g., a drug reaches the bloodstream or alters a biomarker) (Cartwright et al., 2010). Similarly, in mental health, we need preliminary studies to assess whether key change mechanisms can be effectively targeted by digital tools before therapy.

2.4. Between therapy sessions

Psychotherapy sessions typically occur once a week for about one hour, which is a tiny fraction of the waking or working hours in a given week. While much attention has been given to in-session factors that contribute to change (e.g., McAleavey & Castonguay, 2014), the time between sessions also plays a significant role (Orlinsky et al., 1993). Assuming the person is awake for 112 hours per week (based on 16 waking hours per day), 99%+ of those hours are spent in between sessions. Events that occur to the patient between sessions are called intersession processes (Orlinsky et al., 1993), and they typically refer to patients' reflections on the previous sessions, their internalized image of psychotherapy, and thoughts about their therapist, along with different homework assignments. Since these processes occur in patients' daily lives and it is challenging to track them, supporting these processes with CAI appears to be a fruitful target.

What can CAI offer to support intersession processes? CAI can help patients practice and consolidate skills learned during therapy – a powerful mediator of lasting change (e.g., Lane et al., 2015). Meta-analytic evidence indicates that both the quantity and quality of homework compliance predict therapy success across cognitive-behavioral interventions (Kazantzis et al., 2016; Mausbach et al., 2010). Despite this, much of the time between sessions is underutilized: patients often do not fully engage with exercises or apply skills in daily life. CAI-based applications could fill this gap by guiding patients through homework assignments, providing reminders, answering questions about exercises, or helping interpret experiences (e.g., distinguishing thoughts from emotions), thereby reinforcing skill acquisition and supporting behavioral change (Stade et al., 2024; Hartford & Stein, 2024). Supporting this view, a recent real-world study found that patients using an AI-enabled therapy support tool alongside group CBT demonstrated higher session attendance, lower dropout, and greater clinical improvement compared to those using standard workbooks, with benefits linked to engagement and personalization (Habicht et al., 2025). Similarly, Sharma et al. (2024) have found that self-guided LLM-based exercises might facilitate adaptive cognitive restructuring.

Beyond homework support, CAI may also facilitate additional processes such as reflection and insight generation. For example, chatbots can help organize patient reflections, summarize intersession experiences, or track cognitive and emotional patterns, which could then be summarized by CAI and reviewed by the therapist in subsequent sessions. Multi-modal CAI systems with increasingly seamless speech-to-text capabilities can make these processes more convenient, allowing patients to log experiences as voice notes (Haag et al., 2024). Summarization is one of the strongest use cases for CAI (Lee et al., 2024); thus, daily or weekly summaries with key points could be created based on such recordings, which patients can share with their therapist. Also, these systems can extract the main cognitive patterns, e.g., patients' negative thoughts and certain emotional reactions, and illuminate the connections between them, for which self-insight might be limited.

Future research should examine how CAI can most effectively support intersession processes, importantly, without undermining the therapeutic alliance or replacing essential human judgment. Studies could explore which types of prompts, reflection tools, and modes of interaction (e.g., text, voice, or summaries) are most beneficial for different patient groups. In particular, longitudinal trials are needed to assess whether CAI-enhanced intersession engagement leads to better therapeutic outcomes, such as increased insight, adherence to homework, and therapeutic session-to-session continuity.

2.5. Post-treatment

Mental disorders are, in general, highly recurrent and are prone to relapse and chronification (Paykel et al., 2005; Yonkers et al., 2003; Hölzel et al., 2011), with depression being a prime example, characterized by a particularly high recurrence (Hölzel et al., 2011). While it is known that evidence-based psychotherapy is generally effective (Campbell, 2013; Cuijpers et al., 2023; 2025), these effects often tend to decrease after psychotherapy is finished (Steinert et al., 2014). Hence, follow-up aftercare programs, including outpatient psychotherapy and monitoring, are recommended for people with depressive disorders (Krijnen-de Bruin, 2022). The main aim of this period is to consolidate treatment outcomes and ensure their longevity (Hennemann et al., 2018).

Studies show that aftercare can reduce rates of relapse, as well as increase rates of appropriate treatment following relapse due to increased monitoring (Dennis et al., 2003; McKay, 2021). This might allow clinicians to reduce the risk of relapse. However, one of the main barriers to administering effective aftercare programs is the limited resources of healthcare systems, given high costs (Henemann et al., 2018). To address these issues, in recent years, innovative low-cost and low-threshold aftercare models have emerged, such as web- or telephone-based follow-up monitoring programs (Hennemann et al., 2016; Jacobi et al., 2017; reviewed in Petre et al., 2024).

Building on these developments, CAI could offer structured, interactive modules tailored to post-therapy needs. For example, a CAI system could guide users through periodic relapse monitoring using validated mood and symptom questionnaires, prompting reflection on triggers and early warning signs, and offering personalized relapse prevention, prompting the patient to make use of the skills that worked the best for them in therapy. It could also provide personalized psychoeducation refreshers, facilitate structured behavioral activation exercises, and offer motivational dialogue to sustain therapeutic habits. Importantly, the CAI system could escalate cases by notifying clinicians when risk thresholds are exceeded.

Future studies should investigate the efficacy and acceptability of CAI-based aftercare interventions in sustaining treatment gains and preventing relapse, particularly among individuals with recurrent conditions like depression or psychosis. RCTs could compare CAI-supported aftercare with traditional follow-up models and evaluate outcomes such as relapse rates, user engagement, and cost-effectiveness. Research should also explore which components (e.g., mood tracking, motivational dialogue, or elements of CBT) are most impactful and how to adapt these for varying clinical profiles.

3. CAI support for therapists

In addition to filling the gaps for therapy patients, CAI can also assist therapists outside of therapy proper. Based on the transcripts of psychotherapy sessions, CAI can perform a number of actions at different levels of analytic depth. On the basic level, it can summarize the session, find key points discussed in the session, and identify areas for progress, akin to meeting "co-pilots" that are being routinely implemented in video conferencing solutions, like Zoom or Google Teams. Indeed, one of the most immediate opportunities lies in supporting the non-patient-facing administrative tasks that often consume a substantial portion of therapists' time. CAI systems can generate structured summaries of therapy sessions (Asgari et al., 2025), thereby reducing the burden of routine documentation. Automatic CAI-based services to deliver such documentation have experienced rapid adoption in other parts of medicine (Tierney et al., 2024). Beyond session notes, they can assist in drafting referral requests, preparing statements for insurance providers, and maintaining up-to-date clinical records (Tierney et al., 2024). By taking on these administrative roles rather than the therapeutic relationship itself, CAI can improve efficiency, reduce clinician cognitive load, and allow therapists to concentrate on core therapeutic tasks and higher-order decision-making (Das & Ghoshal, 2023). Supporting the superordinate administrative structure of healthcare is, therefore, a meaningful avenue for supporting mental health care with CAI.

Beyond these applications of summarizations, on a deeper level, CAI that is grounded in particular evidence-based treatment manuals could act as supervisors for psychotherapists (e.g., Hsieh et al., 2024; Stade et al., 2024) and generally provide actionable feedback. For instance, it can help with identifying the effective and ineffective aspects of the therapist's behavior throughout the session. It can also analyze the behavior of the client in the session, for example: were there some moments when the client avoided some questions, or the therapist ignored points the patient brought up?

In addition, CAI systems could be used to monitor established mechanisms of change within evidence-based treatments, offering therapists actionable information between sessions. For example, in cognitive therapy for PTSD, changes in trauma-related beliefs are central to recovery and strongly predict treatment outcome (Kleim et al., 2013). Similarly, in exposure-based interventions, therapeutic progress is often mediated by expectancy violation processes (patients learning that feared outcomes do not occur), which can be assessed and tracked (Craske et al., 2022). A CAI-based system could analyze session transcripts or patient reflections to detect such shifts, flagging them for therapist review. The therapist then decides whether and how to address, reinforce, or further explore these mechanisms with the client.

Future research should investigate how CAI-based supervision tools can be reliably integrated into clinical training without undermining the development of therapists' critical thinking, intuition, and interpersonal sensitivity. Empirical studies could explore the effectiveness of CAI feedback in improving specific therapist competencies across different therapeutic modalities and training levels. Additionally, research is needed to determine how therapists interact with, interpret, and act on CAI-generated insights, especially in complex or ambiguous clinical cases.

4. Concluding remarks

Since the advent of advanced conversational AI (CAI), the dominant view among technologists has been to use it to directly provide psychotherapy. It may be appropriate to expand our perspective on CAI's role in mental healthcare. Rather than positioning it as a substitute for (human) therapy, we can think of CAI as a new technological affordance (like the internet) that can be used to transform real-world healthcare systems to make them more integrated, less burdensome, improve their reach, and thus, their impact on public health. To start, the biggest opportunities may lie in areas in which currently little or no human support is available to deliver care. Future research and private sector investment could focus on these gap areas, spanning prevention, waitlist periods, intersession processes, and post-treatment stages. To summarize, then, the question is not "Is CAI better at therapy than humans?" but "Will a health care system with CAI solutions be better than a system without?" Based on much preliminary work cited in this article, we feel the answer is demonstrably yes. Now is the time for private sector and academic innovators to "think bigger" beyond therapy and about the system as a whole.

Acknowledgements

We are thankful to Markus Wolf and Birgit Watzke for fruitful discussions that inspired certain aspects of this work.

Authors' contributions

N.K. developed the concept and drafted the manuscript. E.C.S., J.A., and J.C.E. critically revised the manuscript.

Funding

N.K. was supported by the University of Zurich and the Digital Society Initiative Excellence Fellowship. J.A. was supported by the Estonian Research Council grant PSG728 and Tem-TA 120, and the Estonian Centre of Excellence in Artificial Intelligence (EXAI), funded by the Estonian Ministry of Education and Research. J.C.E. and E.C.S. are supported by Stanford's Institute for Human-Centered AI and Center for Artificial Intelligence in Medicine and Imaging, as well as through the NIMH P50 grant #MH139450.

Conflict of interest

N.K. and J.A. report no conflict of interest. E.C.S. reports paid advising work for Sonar Mental Health. J.C.E. reports equity and consulting income from Sonar Mental Health and Jimini Health, LLC.

References

American Psychological Association. (2018). How Long Will It Take for Treatment to Work? https://www.apa.org/ptsd-guideline/patients-and-families/length-treatment

Asgari, E., Montaña-Brown, N., Dubois, M., Khalil, S., Balloch, J., Yeung, J. A., & Pimenta, D. (2025). A framework to assess clinical safety and hallucination rates of LLMs for medical text summarisation. npj Digital Medicine, 8(1), 1-15.

Blease, C., & Rodman, A. (2025). Generative Artificial Intelligence in Mental Healthcare: An Ethical Evaluation. Current Treatment Options in Psychiatry, 12(1), 1-9.

Blech B, West JC, Yang Z, Barber KD, Wang P, Coyle C. Availability of network psychiatrists among the largest health insurance carriers in Washington, D.C. Psychiatr Serv. 2017 Oct 01;68(9):962–965. doi: 10.1176/appi.ps.201600454.

Campbell, L. F., Norcross, J. C., Vasquez, M. J., & Kaslow, N. J. (2013). Recognition of psychotherapy effectiveness: the APA resolution. Psychotherapy (Chicago, Ill.), 50(1), 98–101. https://doi.org/10.1037/a0031817

Carroll, K. M., Ball, S. A., Nich, C., Martino, S., Frankforter, T. L., Farentinos, C., Kunkel, L. E., Mikulich-Gilbertson, S. K., Morgenstern, J., Obert, J. L., Polcin, D., Snead, N., Woody, G. E., & National Institute on Drug Abuse Clinical Trials Network (2006). Motivational interviewing to improve treatment engagement and outcome in individuals seeking treatment for substance abuse: a multisite effectiveness study. Drug and alcohol dependence, 81(3), 301–312. https://doi.org/10.1016/j.drugalcdep.2005.08.002

Cartwright, M. E., Cohen, S., Fleishaker, J. C., Madani, S., McLeod, J. F., Musser, B., & Williams, S. A. (2010). Proof of concept: a PhRMA position paper with recommendations for best practice. Clinical Pharmacology & Therapeutics, 87(3), 278-285.

Chater, N., & Loewenstein, G. (2022). The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. Behavioral and Brain Sciences, 46, Article e147. https://doi.org/10.1017/S0140525X22002023

Chen, H., He, Q., Wang, M., Wang, X., Pu, C., Li, S., & Li, M. (2022). Effectiveness of CBT and its modifications for prevention of relapse/recurrence in depression: A systematic review and meta-analysis of randomized controlled trials. Journal of Affective Disorders, 319, 469-481.

Civio. (2021, March 9). Access to mental health in Europe. Civio. https://civio.es/sanidad/2021/03/09/access-to-mental-health-in-europe/

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. Journal of clinical psychology, 67(2), 184-192.

Craske, M. G., Sandman, C. F., & Stein, M. B. (2022). How can neurobiology of fear extinction inform treatment?. Neuroscience & Biobehavioral Reviews, 143, 104923.

Cristea, I. A. (2019). The waiting list is an inadequate benchmark for estimating the effectiveness of psychotherapy for depression. Epidemiology and Psychiatric Sciences, 28(3), 278-279.

Cuijpers, P., Harrer, M., Miguel, C., Ciharova, M., & Karyotaki, E. (2023). Five decades of research on psychological treatments of depression: A historical and meta-analytic overview. American psychologist.

Cuijpers, P., Harrer, M., Miguel, C., Ciharova, M., Papola, D., Basic, D., ... & Furukawa, T. A. (2025). Cognitive behavior therapy for mental disorders in adults: A unified series of meta-analyses. JAMA psychiatry, 82(6), 563-571.

Cuijpers, P., Pineda, B. S., Quero, S., Karyotaki, E., Struijs, S. Y., Figueroa, C. A., ... & Muñoz, R. F. (2021). Psychological interventions to prevent the onset of depressive disorders: a meta-analysis of randomized controlled trials. Clinical Psychology Review, 83, 101955.

Das, S., & Ghoshal, A. (2023). Can Artificial Intelligence Ever Develop the Human Touch and Replace a Psychiatrist?-A letter to the editor of the Journal of Medical Systems: Regarding "Artificial Intelligence in Medicine & ChatGPT: De-Tether the Physician". Journal of Medical Systems, 47(1), 72.

Dennis, M., Scott, C. K., & Funk, R. (2003). An experimental evaluation of recovery management checkups (RMC) for people with chronic substance use disorders. Evaluation and program planning, 26(3), 339-352.

Dodge, K. A., Prinstein, M. J., Evans, A. C., Ahuvia, I. L., Alvarez, K., Beidas, R. S., Brown, A. J., Cuijpers, P., Denton, E., Hoagwood, K. E., Johnson, C., Kazdin, A. E., McDanal, R., Metzger, I. W., Rowley, S. N., Schleider, J., & Shaw, D. S. (2024). Population mental health science: Guiding principles and initial agenda. American Psychologist, youth. https://doi.org/10.1037/amp0001334

Dodge, K. A. (2018). Toward population impact from early childhood psychological interventions. American Psychologist, 73(9), 1117–1129. https://doi.org/10.1037/amp0000393

Eubanks, C. F., Muran, J. C., & Safran, J. D. (2018). Alliance rupture repair: A meta-analysis. Psychotherapy, 55(4), 508.

Flückiger, C., Wampold, B. E., Delgadillo, J., Rubel, J., Vîslă, A., & Lutz, W. (2020). Is There an Evidence-Based Number of Sessions in Outpatient Psychotherapy? - A Comparison of Naturalistic Conditions across Countries. Psychotherapy and psychosomatics, 89(5), 333–335. https://doi.org/10.1159/000507793

Haag, C., Steinemann, N., Ajdacic-Gross, V., Schlomberg, J. T. T., Ineichen, B. V., Stanikić, M., ... & von Wyl, V. (2024). Natural language processing analysis of the theories of people with multiple sclerosis about causes of their disease. Communications Medicine, 4(1), 122.

Habicht, J., Dina, L. M., McFadyen, J., Stylianou, M., Harper, R., Hauser, T. U., & Rollwage, M. (2025). Generative AI–enabled therapy support tool for improved clinical outcomes and patient engagement in group therapy: real-world observational study. Journal of medical Internet research, 27, e60435.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. Clinical Psychology: Science and Practice.

Harper, K. L., Lee, D. J., Moshier, S., Zweig, I., Keane, T. M., & Marx, B. P. (2024). Is adequate dose adequate? An examination of the impact of psychotherapy on posttraumatic stress disorder symptoms utilizing Veterans Health Administration medical records. Psychological Services. https://doi.org/10.1037/ser0000830

Hartford, A., & Stein, D. J. (2024). The Machine Speaks: Conversational AI and the Importance of Effort to Relationships of Meaning. JMIR Mental Health, 11, e53203.

He, Y., Yang, L., Qian, C., Li, T., Su, Z., Zhang, Q., & Hou, X. (2023). Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. Journal of medical Internet research, 25, e43862.

Heinz, M. V., Mackin, D. M., Trudeau, B. M., Bhattacharya, S., Wang, Y., Banta, H. A., ... & Jacobson, N. C. (2025). Randomized trial of a generative ai chatbot for mental health treatment. NEJM AI, 2(4), AIoa2400802.

Hennemann, S., Beutel, M. E., & Zwerenz, R. (2016). Drivers and barriers to acceptance of web-based aftercare of patients in inpatient routine care: a cross-sectional survey. Journal of medical Internet research, 18(12), e337.

Hennemann, S., Farnsteiner, S., & Sander, L. (2018). Internet-and mobile-based aftercare and relapse prevention in mental disorders: a systematic review and recommendations for future research. Internet interventions, 14, 1-17.

Hölzel, L., Härter, M., Reese, C., & Kriston, L. (2011). Risk factors for chronic depression—a systematic review. Journal of affective disorders, 129(1-3), 1-13.

Hsieh, L. H., Liao, W. C., & Liu, E. Y. (2024). Feasibility assessment of using ChatGPT for training case conceptualization skills in psychological counseling. Computers in Human Behavior: Artificial Humans, 2(2), 100083.

Huang, S., Wang, Y., Li, G., Hall, B. J., & Nyman, T. J. (2024). Digital mental health interventions for alleviating depression and anxiety during psychotherapy waiting lists: systematic review. JMIR mental health, 11(1), e56650.

Jacka, F. N., & Reavley, N. J. (2014). Prevention of mental disorders: evidence, challenges and opportunities. BMC medicine, 12, 1-3.

Jacobi, C., Beintner, I., Fittig, E., Trockel, M., Braks, K., Schade-Brittinger, C., & Dempfle, A. (2017). Web-based aftercare for women with bulimia nervosa following inpatient treatment: randomized controlled efficacy trial. Journal of medical Internet research, 19(9), e321.

Johri, S., Jeong, J., Tran, B. A., Schlessinger, D. I., Wongvibulsin, S., Barnes, L. A., ... & Rajpurkar, P. (2025). An evaluation framework for clinical use of large language models in patient interaction tasks. Nature Medicine, 1-10.

Karekla, M., Kasinopoulos, O., Neto, D. D., Ebert, D. D., Van Daele, T., Nordgreen, T., ... & Jensen, K. L. (2019). Best practices and recommendations for digital interventions to improve engagement and adherence in chronic illness sufferers. European Psychologist.

Kjell, O. N., Kjell, K., & Schwartz, H. A. (2024). Beyond rating scales: With targeted evaluation, large language models are poised for psychological assessment. Psychiatry Research, 333, 115667.

Kleim, B., Grey, N., Wild, J., Nussbeck, F. W., Stott, R., Hackmann, A., ... & Ehlers, A. (2013). Cognitive change predicts symptom reduction with cognitive therapy for posttraumatic stress disorder. Journal of consulting and clinical psychology, 81(3), 383.

Krijnen-de Bruin, E., Scholten, W., Muntingh, A., Maarsingh, O., van Meijel, B., van Straten, A., & Batelaan, N. (2022). Psychological interventions to prevent relapse in anxiety and depression: A systematic review and meta-analysis. PLoS One, 17(8), e0272200.

Le, L. K. D., Esturas, A. C., Mihalopoulos, C., Chiotelis, O., Bucholc, J., Chatterton, M. L., & Engel, L. (2021). Cost-effectiveness evidence of mental health prevention and promotion interventions: A systematic review of economic evaluations. PLoS medicine, 18(5), e1003606.

Lee, C., Vogt, K. A., & Kumar, S. (2024). Prospects for AI clinical summarization to reduce the burden of patient chart review. Frontiers in Digital Health, 6, 1475092.

Levkovich, I., & Elyoseph, Z. (2023). Identifying depression and its determinants upon initiating treatment: ChatGPT versus primary care physicians. Family medicine and community health, 11(4), e002391. https://doi.org/10.1136/fmch-2023-002391

Li, H., Zhang, R., Lee, Y. C., Kraut, R. E., & Mohr, D. C. (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. NPJ Digital Medicine, 6(1), 236.

Maurya, R. K., Montesinos, S., Bogomaz, M., & DeDiego, A. C. (2025). Assessing the use of ChatGPT as a psychoeducational tool for mental health practice. Counselling and Psychotherapy Research, 25(1), e12759.

McAleavey, A. A., & Castonguay, L. G. (2014). The process of change in psychotherapy: Common and unique factors. In Psychotherapy research: Foundations, process, and outcome (pp. 293-310). Vienna: Springer Vienna.

McKay J. R. (2021). Impact of Continuing Care on Recovery From Substance Use Disorder. Alcohol research : current reviews, 41(1), 01. https://doi.org/10.35946/arcr.v41.1.01

Miner, A. S., Shah, N., Bullock, K. D., Arnow, B. A., Bailenson, J., & Hancock, J. (2019). Key considerations for incorporating conversational AI in psychotherapy. Frontiers in psychiatry, 10, 746.

Mohr, D. C., Hart, S. L., Howard, I., Julian, L., Vella, L., Catledge, C., & Feldman, M. D. (2006). Barriers to psychotherapy among depressed and nondepressed primary care patients. Annals of Behavioral Medicine, 32(3), 254-258.

Mohr, D. C., Riper, H., & Schueller, S. M. (2018). A solution-focused research approach to achieve an implementable revolution in digital mental health. JAMA psychiatry, 75(2), 113-114.

Mohr, D. C., Weingardt, K. R., Reddy, M., & Schueller, S. M. (2017). Three problems with current digital mental health research... and three things we can do about them. Psychiatric services, 68(5), 427-429.

Mrazek P., Haggerty R. (1994) Reducing risks for mental disorders: frontiers for preventive intervention research. National Academy Press, Washington, DC

National Research Council, Institute of Medicine (2009) Preventing mental, emotional, and behavioral disorders among young people: progress and possibilities. The National Academies Press, Washington, DC

Orlinsky, D. E., Geller, J. D., Tarragona, M., & Farber, B. (1993). Patients' representations of psychotherapy: A new focus for psychodynamic research. Journal of Consulting and Clinical Psychology, 61(4), 596.

Parker, D., Byng, R., Dickens, C., & McCabe, R. (2020). "Every structure we're taught goes out the window": General practitioners' experiences of providing help for patients with emotional concerns'. Health & social care in the community, 28(1), 260–269. https://doi.org/10.1111/hsc.12860

Paykel, E. S., Scott, J., Cornwall, P. L., Abbott, R., Crane, C., Pope, M., & Johnson, A. L. (2005). Duration of relapse prevention after cognitive therapy in residual depression: follow-up of controlled trial. Psychological Medicine, 35(1), 59-68.

Peipert, A., Krendl, A. C., & Lorenzo-Luaces, L. (2022). Waiting Lists for Psychotherapy and Provider Attitudes Toward Low-Intensity Treatments as Potential Interventions: Survey Study. JMIR formative research, 6(9), e39787. https://doi.org/10.2196/39787

Petre, L. M., Piepiora, P. A., Gemescu, M., & Gheorghe, D. A. (2024). Internet-and mobile-based aftercare and relapse prevention interventions for anxiety and depressive disorders: a systematic review. Frontiers in Psychology, 15, 1474016.

Reichert, A., & Jacobs, R. (2018). The impact of waiting time on patient outcomes: Evidence from early intervention in psychosis services in England. Health economics, 27(11), 1772–1787. https://doi.org/10.1002/hec.3800

Ridenour, T. A., Murray, D. W., Hinde, J., Glasheen, C., Wilkinson, A., Rackers, H., & Coyne-Beasley, T. (2022). Addressing barriers to primary care screening and referral to prevention for youth risky health behaviors: Evidence regarding potential cost-savings and provider concerns. Prevention Science, 23(2), 212–223. https://doi.org/10.1007/ s11121-021-01321-9

Rollwage, M., Habicht, J., Juechems, K., Carrington, B., Viswanathan, S., Stylianou, M., ... & Harper, R. (2023). Using conversational AI to facilitate mental health assessments and improve clinical efficiency within psychotherapy services: real-world observational study. JMIR AI, 2(1), e44358.

Sander, L., Rausch, L., & Baumeister, H. (2016). Effectiveness of internet-based interventions for the prevention of mental disorders: a systematic review and meta-analysis. JMIR mental health, 3(3), e6061.

Sander, L., Rausch, L., & Baumeister, H. (2016). Effectiveness of internet-based interventions for the prevention of mental disorders: a systematic review and meta-analysis. JMIR mental health, 3(3), e6061.

Schleider, J. L., & Weisz, J. R. (2016). Reducing risk for anxiety and depression in adolescents: Effects of a single-session intervention teaching that personality can change. Behaviour research and therapy, 87, 170-181.

Schleider, J. L., Mullarkey, M. C., Fox, K., Dobias, M., Shroff, A., Hart, E., & Roulston, C. A. (2021). Single-session interventions for adolescent depression in the context of COVID-19: A nationwide randomized-controlled trial. Nature Human Behaviour, 6, 258-268.

Sedlakova, J., & Trachsel, M. (2023). Conversational artificial intelligence in psychotherapy: a new therapeutic tool or agent?. The American Journal of Bioethics, 23(5), 4-13.

Sharma, A., Rushton, K., Lin, I. W., Nguyen, T., & Althoff, T. (2024, May). Facilitating self-guided mental health interventions through human-language model interaction: A case study of cognitive restructuring. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (pp. 1-29).

Shortt, A. L., & Spence, S. H. (2006). Risk and protective factors for depression in youth. Behaviour Change, 23(1), 1-30.

Siddals, S., Torous, J., & Coxon, A. (2024). "It happened to be the perfect thing": experiences of generative AI chatbots for mental health. npj Mental Health Research, 3(1), 48.

Stade, E., Tait, Z., Campione, S., & Stirman, S. (2025). Current Real-World Use of Large Language Models for Mental Health.

Stade, E. C., Stirman, S. W., Ungar, L. H., Boland, C. L., Schwartz, H. A., Yaden, D. B., ... & Eichstaedt, J. C. (2024). Large language models could change the future of behavioral healthcare: a proposal for responsible development and evaluation. NPJ Mental Health Research, 3(1), 12.

Steinert, C., Hofmann, M., Kruse, J., & Leichsenring, F. (2014). Relapse rates after psychotherapy for depression–stable long-term effects? A meta-analysis. Journal of affective disorders, 168, 107-118.

Steinert, C., Stadter, K., Stark, R., & Leichsenring, F. (2017). The effects of waiting for treatment: A meta-analysis of waitlist control groups in randomized controlled trials for social anxiety disorder. Clinical Psychology & Psychotherapy, 24(3), 649-660.

Steinman KJ, Shoben AB, Dembe AE, Kelleher KJ. How long do adolescents wait for psychiatry appointments? Community Ment Health J. 2015 Oct 25;51(7):782–9. doi: 10.1007/s10597-015-9897-x.10.1007/s10597-015-9897-x

Sung, J. Y., Bugatti, M., Vivian, D., & Schleider, J. L. (2023). Evaluating a telehealth single-session consultation service for clients on psychotherapy wait-lists. Practice innovations, 8(2), 141.

Swift, J. K., Penix, E. A., & Li, A. (2023). A meta-analysis of the effects of role induction in psychotherapy. Psychotherapy, 60(3), 342.

Tierney, A. A., Gayre, G., Hoberman, B., Mattern, B., Ballesca, M., Kipnis, P., ... & Lee, K. (2024). Ambient artificial intelligence scribes to alleviate the burden of clinical documentation. NEJM Catalyst Innovations in Care Delivery, 5(3), CAT-23.

Tse, S., Tsoi, E. W., Hamilton, B., O'Hagan, M., Shepherd, G., Slade, M., ... & Petrakis, M. (2016). Uses of strength-based interventions for people with serious mental illness: A critical review. International Journal of Social Psychiatry, 62(3), 281-291.

Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and conversational agents in mental health: a review of the psychiatric landscape. The Canadian Journal of Psychiatry, 64(7), 456-464.

Van Dijk, D. A., Meijer, R. M., Van Den Boogaard, T. M., Spijker, J., Ruhé, H. G., & Peeters, F. P. M. L. (2023). Worse off by waiting for treatment? The impact of waiting time on clinical course and treatment outcome for depression in routine care. Journal of affective disorders, 322, 205-211.

Vink, D., Aartsen, M. J., & Schoevers, R. A. (2008). Risk factors for anxiety and depression in the elderly: a review. Journal of affective disorders, 106(1-2), 29-44.

Westra, H. A., & Dozois, D. J. (2006). Preparing clients for cognitive behavioral therapy: A randomized pilot study of motivational interviewing for anxiety. Cognitive Therapy and Research, 30, 481-498.

Westra, H. A., Arkowitz, H., & Dozois, D. J. (2009). Adding a motivational interviewing pretreatment to cognitive behavioral therapy for generalized anxiety disorder: A preliminary randomized controlled trial. Journal of anxiety disorders, 23(8), 1106-1117.

Yonkers, K. A., Bruce, S. E., Dyck, I. R., & Keller, M. B. (2003). Chronicity, relapse, and illness—course of panic disorder, social phobia, and generalized anxiety disorder: findings in men and women from 8 years of follow-up. Depression and anxiety, 17(3), 173-179.

Zhang, H., Yu, D., Wang, Z., Shi, J., & Qian, J. (2019). What impedes general practitioners' Identification of mental disorders at outpatient departments? A qualitative study in Shanghai, China. Annals of Global Health, 85(1).

Zhong, W., Luo, J., & Zhang, H. (2024). The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. Journal of Affective Disorders. https://doi.org/10.1016/j.jad.2024.04.057

Zhou, Y., Zhao, D., Zhu, X., Liu, L., Meng, M., Shao, X., ... & Zhu, G. (2023). Psychological interventions for the prevention of depression relapse: systematic review and network meta-analysis. Translational Psychiatry, 13(1), 300.