

ChatGPT for Mental Health Support: A Systematic Scoping Review of Human–Computer Interaction Implications

Health Education & Behavior

1–11

© 2026 Society for Public

Health Education

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/10901981261435479

journals.sagepub.com/home/heb



Thseen Nazir, PhD¹ 

Abstract

Large language models (LLMs) such as ChatGPT are increasingly used around mental health, yet design-oriented syntheses remain limited. We conducted a PRISMA-aligned systematic scoping review focused specifically on ChatGPT (GPT-3.5/4+) in mental health-related use including “in the wild” adoption by laypeople, training applications, and clinical-adjacent pilots. Searches of five databases (November 2022 to August 2025), with citation tracking and gray-literature screening, yielded 34 studies spanning randomized and non-randomized experiments, pilot trials, surveys/interviews, simulations, digital ethnography, and structured editorials. Evidence supports adjunct, not replacement, roles. In education/supervision, one randomized trial and a comparative supervision study show skill gains when practice is scaffolded with rubrics and human oversight; expert ratings judged trainee case conceptualizations acceptable. In clinical/adjacent contexts, signals include quality-of-life improvement (small inpatient pilot), short-term anxiety reduction when the model provides empathetic feedback, and a clinical RCT (outside psychiatry) showing reduced anxiety/depression with a ChatGPT adjunct. Studies of public/self-help use document appropriation of ChatGPT as a “digital therapist” with identified risks including privacy concerns, boundary violations, and over-reliance. Safety-critical tasks remain unreliable (e.g., under-identification of suicide risk, degradation with complexity, and cultural-fit gaps). We derive human–computer interaction requirements: clear scope-of-use messaging, prompt scaffolding, human-in-the-loop, privacy-preserving defaults, and explicit escalation/hand-off pathways.

Keywords

large language models, mental health, human–computer interaction, counselor education and supervision

Introduction

Global demand for mental health support continues to outstrip available services. Cost, geography, long waits, and stigma keep many people outside of formal care, sustaining a large “treatment gap” across settings and age groups. In this vacuum, general-purpose, conversational tools including large-language-model (LLM) chatbots like ChatGPT are increasingly used as just-in-time sources of coping tips, psychoeducation, and “someone to talk to” when human care is unavailable or delayed (Jung et al., 2025; Song et al., 2025). From a human–computer interaction (HCI) perspective, this moment is consequential: a general-purpose, always-on dialog system has become part of the help-seeking journey for millions of people, yet we lack a consolidated, design-oriented synthesis of what users actually do with ChatGPT for mental health, what benefits and risks have been observed, and how these map onto interaction design choices (e.g., prompts, safety rails, empathy displays, and hand-off pathways).

ChatGPT reached mass adoption at unprecedented speed following its late-2022 release, catalyzing both enthusiasm and concern in mental health communities. Reports in counseling and counselor-education literature describe it as the “fastest-growing” platform, with more than 100 million users shortly after launch and a scale that virtually guarantees spillover into wellness and mental health use cases (Ayo-Ajibola et al., 2024; Cross et al., 2024; Siddals et al., 2024). Early commentaries in psychiatry likewise argue that LLM chatbots may help narrow treatment-access gaps while flagging unresolved issues around safety, privacy, and clinical reliability (Cheng et al., 2023; Singh, 2019). Empirically, multiple strands of research already document how people

¹Ibn Haldun University, Istanbul, Turkey

Corresponding Author:

Thseen Nazir, Ibn Haldun University, Başak Mah. Ordu Cad. No. 3 P.K., 34480 Başakşehir/İstanbul, Turkey.
Email: thseen.nazir@ihu.edu.tr

use ChatGPT around mental health. Social-media ethnography shows organically emerging practices in which individuals “shape” the model into a “digital therapist” for emotion regulation, cognitive reframing, and reflective writing, while negotiating boundaries, expectations, and trust (Collins et al., 2025). Survey work and classroom studies in counseling and mental health education find high perceived usefulness for psychoeducation and skill practice (e.g., empathic reflection rehearsal), tempered by concerns about accuracy, privacy, and over-reliance (Ajrlouni et al., 2023; Maurya, 2024).

Clinical-adjacent evaluations are mixed: in a small inpatient pilot, brief guided ChatGPT sessions were associated with improved WHOQOL-BREF scores and high satisfaction (Melo et al., 2024), whereas simulated-patient studies caution that performance degrades as clinical complexity rises, producing potentially unsafe recommendations and revealing limits in information-gathering and judgment (Dergaa et al., 2024). Beyond outcome signals, HCI-relevant work interrogates how the interaction feels and what the system seems to “do” for users. Qualitative and experimental studies point to perceived empathy, non-judgmental tone, and 24/7 availability as key drivers of engagement, even when users explicitly recognize the model’s non-human status (Luo et al., 2025; Raile, 2024).

At the same time, human raters still tend to outperform ChatGPT on nuanced emotional-reflection tasks, and expectations of “therapeutic alliance” regularly collide with the model’s limited situational awareness and inability to perceive non-verbal cues or reliably detect risk (Eryilmaz & Başal, 2024; Imran et al., 2023). Conceptual work in bioethics frames these tensions as an emerging “artificial third” in the therapeutic space neither therapist nor tool alone requiring new norms around boundaries, agency, and informed use (Tal et al., 2023).

Safety, ethics, and governance recur across this literature and are central to HCI design obligations. Content accuracy, hallucinations, and domain drift remain problematic in uncurated health contexts; privacy risks are acute because help-seeking conversations often include sensitive identifiers; and crisis detection/response is fragile without robust escalation pathways (Kalam et al., 2024; Singh, 2019). Pediatric and adolescent uses amplify these concerns and have prompted professional bodies to prototype evaluative frameworks (e.g., APA’s App Advisor/Digital Psychiatry Task Force) that call for transparent data handling, clear scope-of-use messaging, and human oversight principles that translate directly into HCI requirements for onboarding, consent, and hand-off design (Imran & Almusharraf, 2023; Singh, 2019).

As the boundary between “consumer” and “clinical” contexts blurs, adjacent findings map promising indirect roles for ChatGPT: as a psychoeducational coach, reflection partner, and practice simulator for counselors-in-training functions that can be engineered to avoid diagnosis or treatment claims while still supporting mental health literacy and skills acquisition (Cioffi et al., 2025; Maurya & Cavanaugh, 2025).

Scope and Contribution

This review focuses specifically on ChatGPT (and direct ChatGPT-accessing interfaces), synthesizing empirical and conceptual studies that examine its use by laypeople and trainees for self-help, psychoeducation, coping support, and early crisis messaging. Studies centered exclusively on purpose-built mental health chatbots (e.g., Woebot) were excluded unless used as comparators, and ChatGPT was not primarily evaluated as a diagnostic or treatment device, though some included studies examined outputs in diagnostic-adjacent contexts, and these are noted where relevant. Using PRISMA-aligned methods, the review integrates quantitative and qualitative findings into an evidence matrix and translates them into actionable HCI implications for safety, usability, trust, and equity. In doing so, the review complements prior position pieces and single-setting evaluations with a domain-specific, interaction-focused synthesis (Cheng et al., 2023; Kalam et al., 2024; Raile, 2024).

A defining and distinguishing feature of this review is its focus on “in the wild” use, meaning organic, largely unguided engagement with a general-purpose LLM by people who were not enrolled in a formal treatment program and were not using a clinically validated, purpose-built tool. This scope is intentional and reflects an underexamined gap in the literature. Most prior syntheses have focused either on purpose-built conversational agents (e.g., Woebot and Wysa) designed with narrow therapeutic scripts, hard safety rails, and clinical oversight built into the product, or on formal clinical evaluations of AI-assisted therapy delivered under controlled conditions. That work has yielded meaningful insights within its bounded scope. Yet those findings do not speak to what happens when a general-purpose conversational model, carrying no such architecture, is appropriated for the same functions by users acting outside any formal program. ChatGPT occupies a fundamentally different niche, as it is a freely accessible, general-purpose system that users can and do steer toward therapeutic functions entirely of their own volition, without clinical framing, structured protocols, or professional monitoring.

When a layperson “shapes” ChatGPT into a digital therapist through iterative prompting (Luo et al., 2025), the resulting interaction is governed by neither clinical training nor product design, but by the user’s own expectations, mental models, and help-seeking behavior. Understanding what happens in these unstructured encounters, including what benefits emerge, where failures cluster, and what interaction design choices could mitigate harm, is therefore a distinct and pressing research question that purpose-built chatbot evaluations do not address.

The “in the wild” framing also encompasses two related use contexts that share its key characteristic of operating outside formal clinical oversight: training and education applications, where counselors-in-training use ChatGPT informally or in semi-structured classroom settings; and clinical-adjacent pilots, where clinicians or researchers have introduced ChatGPT as a supplementary resource alongside, though not

Table 1. Eligibility Criteria (PICOS) for Inclusion of Studies.

Criterion	Inclusion	Exclusion
Population	Lay users, patients, students/trainees, or professionals engaging with ChatGPT for mental health purposes	Non-ChatGPT chatbots (e.g., Woebot and Replika), unless used for comparison
Intervention	Explicit use of ChatGPT (GPT-3.5, GPT-4, or later)	Purely technical NLP studies with no mental health implications
Comparison	Human counselors, baseline datasets, or no comparator	—
Outcomes	Benefits (e.g., knowledge, satisfaction, and coping), risks (e.g., misinformation, privacy, and safety), and HCI-relevant features (e.g., empathy, usability, and boundaries)	Opinion pieces with no analytic component
Study design	Empirical studies (qualitative, quantitative, mixed methods, and pilot trials), comparative evaluations, ethnographies, reviews/commentaries with structured analysis	—

replacing, formal care. These contexts are included because they similarly involve general-purpose LLM use without the guardrails of a purpose-built clinical product, and because their HCI implications are closely related. By synthesizing across all three contexts, this review aims to clarify what is happening in the wild, what seems to help and for whom, where failures cluster, and how HCI design can concretely improve safety and utility in LLM-mediated mental health interactions.

Research Questions

Guided by gaps identified above, we organize the review around the following questions:

- RQ1.** In what contexts and for what purposes are people using ChatGPT for mental health help (e.g., emotion regulation, psychoeducation, coping, and between-session support)?
- RQ2.** What benefits (perceived and/or measured) have been reported when ChatGPT is used for mental health support?
- RQ3.** What types of supportive functions does ChatGPT provide (e.g., psychoeducation, cognitive restructuring prompts, empathic reflection, and behavioral planning), and how are these elicited by interaction design?
- RQ4.** What expectations do users bring to ChatGPT as helper, and what recurrent usability, safety, ethical, and privacy issues arise?
- RQ5.** How do risks and benefits vary across populations (e.g., youth and low-resource settings), and what inclusive design considerations emerge?
- RQ6.** What roles for human oversight (clinicians, peers, and educators) are supported by evidence, and how should interfaces handle escalation and hand-offs?

Method

Review Design and Registration

A systematic scoping review was conducted in accordance with PRISMA 2020. The protocol was not preregistered; all methods were determined a priori and applied consistently. This approach was chosen to map a rapidly emerging, heterogeneous body of evidence while allowing structured synthesis across both empirical and conceptual studies.

Eligibility Criteria (PICOS Framework)

We applied the PICOS framework (Population, Intervention, Comparison, Outcomes, Study design) to determine inclusion and exclusion (Table 1).

Information Sources and Search Strategy

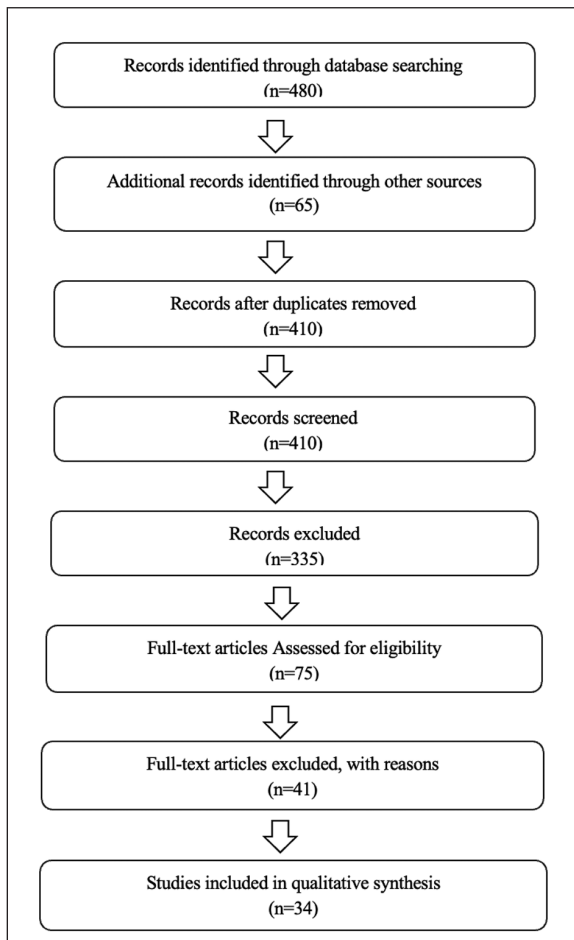
Systematic searches were conducted in PubMed, PsycINFO, Web of Science, Scopus, and ACM Digital Library. Searches covered November 2022 (ChatGPT public release) through August 2025. No language restrictions were applied at the search stage; non-English records were screened for eligibility using automated translation where needed. Additional sources included hand searches of reference lists, gray literature from conference proceedings, and citation tracking.

Study Selection

The search across 5 databases yielded 480 records, with an additional 65 identified through other sources (e.g., reference lists, gray literature, and citation tracking). After removal of duplicates, 410 unique records remained for screening. Two reviewers independently screened titles and abstracts, excluding 335 that clearly did not meet eligibility criteria. The full texts of 75 articles were assessed, of which 41 were excluded with documented reasons are summarized in Table 2 (e.g., focus on

Table 2. Reasons for Full-Text Exclusion ($n = 41$).

Reason for exclusion	Examples	Number of studies (n)
Not ChatGPT-focused	Studies on Woebot, Replika, or general AI chatbots without ChatGPT	12
No mental health focus	AI in education, language learning, and medical writing without the MH link	10
Technical-only NLP work	Benchmarking language models, dataset creation, and sentiment analysis	9
Editorial/commentary without empirical or MH focus	General AI perspectives without relevance to MH practice	6
Duplicate/overlap	Conference paper + journal article of the same study	4
Total excluded		41

**Figure 1.** PRISMA 2020 Flow Diagram.

non-ChatGPT chatbots, technical NLP evaluations, or no mental health relevance). Ultimately, 34 studies met the inclusion criteria and were synthesized in the review. The study selection process is summarized in Figure 1.

Data Extraction and Synthesis

Study metadata were extracted, including authors, year, country, study design, population, setting, ChatGPT version or

interface, mental health context, outcomes, and key findings. Both quantitative and qualitative outcomes were tabulated. A convergent synthesis approach was applied: numerical results were summarized descriptively, and qualitative findings were coded thematically.

Risk of Bias and Quality Appraisal

Methodological quality was assessed using tools mapped to study designs (Table 3), and study-level appraisals are provided in Table S1 (Supplementary Materials). Randomized controlled trials were evaluated with the Cochrane Risk of Bias 2 (RoB 2) tool. Qualitative studies were appraised with the Critical Appraisal Skills Programme (CASP) Qualitative Checklist, and cross-sectional surveys were evaluated using the AXIS critical appraisal tool. Commentaries and editorials were appraised narratively, with attention to the rigor of argumentation and the disclosure of potential conflicts of interest.

All risk-of-bias assessments were conducted independently by two reviewers, and discrepancies were resolved through discussion until consensus was reached. Appraisal tools were mapped to study designs as shown in Table 3.

Reporting

A PRISMA flow diagram summarizes the screening and selection process (Figure 1). The evidence matrix (Table 4) provides an overview of included studies.

Results

Study Characteristics

Our searches yielded 34 eligible studies published between November 2022 and August 2025. These comprised a mix of empirical work ($n = 21$) including surveys, interviews, pilot trials, ethnographies, and experimental simulations and conceptual/commentary articles ($n = 13$). Study populations were diverse, spanning psychiatric inpatients (Melo et al., 2024), counseling trainees and educators (Ajlouni et al., 2023; Hsieh et al., 2024; Maurya, 2024), outpatients in community

Table 3. Quality Appraisal Tools by Study Design.

Study (year)	Design	Appraisal tool	Overall judgment	Key concerns (one-line)
Melo et al. (2024)	Pilot before–after	Narrative appraisal	High	Very small sample ($n = 12$); no control; self-report outcomes
Ajlouni et al. (2023)	Cross-sectional survey	AXIS	Low–moderate	Purposive sample; self-report; clear measures
Eryılmaz & Başal (2024)	Mixed-methods experiment	Narrative + AXIS/CASP alignment	Moderate	Scenario-based; convenience raters; limited external validity
Dergaa et al. (2024)	Simulation	Narrative appraisal	High	Complex cases; safety concerns; limited validation

Note: AXIS = appraisal tool for cross-sectional studies; CASP = Critical Appraisal Skills Programme.

clinics (Alanezi, 2024), and lay users posting on social media (Giray, 2025; Luo et al., 2025). Geographically, work originated from Asia, the Middle East, Europe, and North America, with several cross-national or online settings.

Roles of ChatGPT/LLMs

Education, Training, and Supervision. Across education settings, ChatGPT was deployed as a feedback provider, practice partner, and case-conceptualization aid. Evidence from one randomized trial shows skill gains when AI is embedded into structured practice: novice counselors who engaged an LLM “patient” and received rubric-guided feedback demonstrated improved reflective and questioning skills versus practice-alone controls (Louie et al., 2025). A comparative supervision study with Gestalt trainees reported that AI feedback supported supervision goals while still requiring human oversight for calibration and boundary setting (Cioffi et al., 2025). In an expert-rated evaluation, trainee case conceptualizations produced with ChatGPT support were judged acceptable, indicating feasibility for early-stage training tasks (Hsieh et al., 2024). Student and educator reports similarly described perceived usefulness alongside a need for AI literacy and clear scope-of-use guidance (Ajlouni et al., 2023; Maurya, 2024; Maurya & Cavanaugh, 2025; Rehman & Sajjad, 2025). Notably, a mixed-methods comparison found that humans outperformed ChatGPT on complex emotional reflection, suggesting LLMs are better suited to structured, well-scaffolded learning activities than to advanced, nuance-heavy competencies (Eryılmaz & Başal, 2024).

Clinical and Clinical-Adjacent Adjuncts. Evidence for direct clinical impact is emerging and remains preliminary. In a small inpatient before–after pilot, brief guided ChatGPT sessions were associated with improved quality of life and high satisfaction, but the absence of a control group and the small sample constrain inference (Melo et al., 2024). A single-session, non-randomized experiment found that anxiety decreased only when ChatGPT provided empathetic feedback during a self-compassion task, suggesting short-term affective benefits contingent on the chatbot’s feedback function (Kishimoto et al., 2025). Outside psychiatry, a clinical

RCT reported reduced anxiety and depression for patients receiving a ChatGPT-based counseling adjunct relative to usual care, reinforcing the plausibility of adjunct effects on mental health endpoints (Akdogan et al., 2025).

In psychodynamic work, expert psychiatrists judged ChatGPT’s formulations appropriate when prompts embedded psychodynamic concepts, underscoring the importance of prompt scaffolding and domain framing (Hwang et al., 2024). A qualitative case analysis with therapists and patients described practical value between sessions, while cautioning about CBT-skew and the risk of over-attribution if the tool is treated as a substitute for therapy (Raile, 2024). Outpatient interviews highlighted psychoeducation/CBT-like benefits alongside cultural considerations and the need for clinician oversight (Alanezi, 2024). In relationship-support settings, users perceived empathy and usability yet noted weak risk-checking and limited collaborative problem solving, indicating gaps for safety and shared decision making (Vowels et al., 2024).

Public/Self-Help and Social-Media Use. Across public settings, people appropriated ChatGPT as a “digital therapist” for just-in-time support, most commonly for empathy, reframing, journaling, and coping tips. Digital-ethnographic and content-analytic studies on Reddit described pervasive themes of constant availability, non-judgmental tone, and an easy on-ramp for disclosing distress (Luo et al., 2025). In parallel, single-session relationship-support evaluations showed high perceived empathy and usability, but highlighted weak risk-checking and limited collaborative problem solving, aligning with the broader pattern that LLMs help users “talk through” problems more than they help decide on safe next steps (Vowels et al., 2024). Across studies, participants and observers flagged privacy, boundary, and over-reliance risks in open communities, underscoring the need for clear scope-of-use messaging, escalation pathways, and community-aware guardrails (e.g., guidance on disclaimers, data visibility, and signaling when to hand off to humans).

Safety-Critical Tasks and Capability Limits. Across high-risk contexts, three converging evidence streams caution against

Table 4. Summarizes the Characteristics of the 34 Included Studies, Including Design, Populations, ChatGPT Roles, and Key Findings.

Author (year)	Design	Population/setting	ChatGPT role	Key findings
Cioffi et al. (2025)	Comparative supervision	Gestalt therapy trainees	Feedback provider	AI feedback supports supervision goals (career guidance/skills); it still requires human oversight.
Louie et al. (2025)	Randomized study	Novice counselors (n = 90)	AI-simulated practice + feedback	Practice with feedback improved reflective/questioning skills vs. practice alone.
Giray (2025)	Narrative case study	Reddit users	Self-help and grief support	Comfort/companionship reported; boundaries can blur.
Maurya and Cavanaugh (2025)	Self-study (education)	Counselor education programs	Training tool	Students value practice but need AI literacy; fit as adjunct.
Maurya (2024)	Content analysis	ChatGPT responses to prompts	Psychoeducational coach	Generally accurate, empathic answers on structured prompts; watch for limitations.
Rehman & Sajjad (2025)	Qualitative interviews	Counselors + students (n = 40)	Adjunct to counseling	Perceived as accessible, flexible, and low-cost; not a replacement.
Luo et al. (2025)	Digital ethnography	Reddit posts ("ChatGPT" + "therapy")	"Digital therapist" (user-shaped)	Users co-create therapy spaces; empathy/availability with privacy/over-reliance risks.
Akdogan et al. (2025)	RCT (clinical adjunct)	Chemo-naïve cancer pts	Digital counseling adjunct	HADS-Anxiety/Depression ↓ vs. control; majority of outputs appropriate; still adjunct.
Kishimoto et al. (2025)	Non-randomized 3-arm (PSM)	Adults (n = 98)	Empathetic feedback coach	Self-compassion ↑ for writing groups; anxiety ↓ only with ChatGPT feedback (single-session).
Hsieh et al. (2024)	Expert evaluation (training)	Counseling trainees	Case-conceptualization support	Experts rated trainee conceptualizations as acceptable with ChatGPT-supported tasks.
Raile (2024)	Qualitative case analysis	Psychotherapists and patients	Adjunct between sessions	Helpful as stop-gap; CBT bias noted; not a substitute for therapy.
Melo et al. (2024)	Pilot before-after	Psychiatric inpatients (n = 12)	Support sessions	WHOQOL-BREF ↑; high satisfaction; no control. small n → preliminary.
Alanezi (2024)	Qualitative study	Outpatients (Saudi Arabia)	Psychoeducation + CBT-like	Useful adjunct; cultural considerations and clinician oversight needed.
Hwang et al. (2024)	Exploratory expert-rated	Psych case vignette	Psychodynamic formulation assistant	Psychiatrists rated outputs appropriate when psychodynamic prompts used; prompt scaffolding matters.
Vowels et al. (2024)	Qualitative + technical eval	Relationship-help users	Single-session support	High perceived empathy/usability; weak on risk-checking/collaborative planning.
Eryilmaz & Başal (2024)	Mixed-methods comparison	Counselors vs. ChatGPT	Emotional reflection	Humans outperform in nuanced/complex reflection; AI useful for structured elements.
Naher (2024)	Comparative content analysis	ChatGPT vs. datasets	Response quality	ChatGPT more empathic-sounding than baselines; empathy ≠ accuracy.
Xu et al. (2024)	Reddit content analysis	r/ChatGPT and related subs	Public attitudes to MH use	Perceived potential for MH support with misinformation/accuracy concerns.
Dergaa et al. (2024)	Simulation	Complex psychiatric cases	Therapeutic advice (simulated)	Adequate in simple cases; unsafe as complexity rises → guardrails required.
Aleem et al. (2024)	Prompt-testing	Multicultural scenarios	Cultural adaptivity	Gaps in cultural humility/adaptation; tuning and context matter.
Kalam et al. (2024)	Perspective	MH policy	Governance	Calls for regulation, safety standards, and AI literacy.
Zhang & Wang (2024)	Opinion	–	AI in therapy	AI cannot replace therapists; adjunct/triage roles only.
Ajlouti et al. (2023)	Survey	Students (n = 210)	Education/training	82% positive; perceived skill development & usefulness.
Imran et al. (2023)	Perspective (child MH)	Child/adolescent psychiatry	Clinician support	Could assist diagnosis/treatment; risks/ethics emphasized.
Blease & Torous (2023)	Editorial/Perspective	MH systems	LLM chatbots in psychiatry	Benefits vs harms; misinformation, privacy, and clinician oversight.
Cheng et al. (2023)	Review	Psychiatry overview	LLM roles	Integration pathways; current limits and future potential.
Tal et al. (2023)	Bioethics commentary	Clinicians + clients	"Artificial third"	Conceptualizes chatbot as new relational entity in treatment triad.
Elyoseph & Levkovich (2023)	Empirical comparison	Suicide-risk vignettes	Risk-assessment assistant	Under-identified suicide risk vs. psychiatrists → escalate to humans.
Levkovich & Elyoseph (2023)	Empirical evaluation	Suicide-risk vignettes	Risk-assessment assistant	Limited/inconsistent risk stratification; needs human oversight and escalation.
Ghanadian et al. (2023)	Technical evaluation	Reddit suicidality dataset	Risk screening on social media	Reasonable accuracy but below fine-tuned baselines; tuning/temperature matter; adjunct only.
Hadar-Shoval et al. (2023)	Empirical study	Personality-style vignettes	Mentalizing assistant	Style-tuning to personality contexts possible; reliability unproven.
van Schaikwyk (2023)	Perspective/Narrative	Child/adolescent	Psychoeducation and self-help	AI can aid youth MH information/coping; safeguards needed.
Carlbring et al. (2023)	Editorial	Internet interventions (ICBT)	AI tutor/coach	Frames ChatGPT as adjunct to ICBT with need for guardrails/clinician control.
Monteith et al. (2024)	Editorial	MH practice	Ethics and misinformation	Generative AI can produce plausible but dangerous advice; it mandates mitigation and oversight.

unsupervised use. First, in suicide-risk vignettes, ChatGPT under-identified risk and produced inconsistent stratification relative to psychiatrists, indicating the need for human oversight and explicit escalation protocols (Levkovich & Elyoseph, 2023). Technical work on social-media risk screening similarly showed that zero/few-shot LLMs trail fine-tuned baselines and are sensitive to temperature/tuning, supporting adjunct-only deployment (Ghanadian et al., 2023). Second, simulation and prompt-stress testing demonstrated that output degrades with case complexity (Dergaa et al., 2024) and that cultural humility/adaptation remains a persistent weakness, particularly in multicultural scenarios (Aleem et al., 2024). Third, although responses often sound empathic (Naher, 2024; Vowels et al., 2024), empathy cues do not guarantee accuracy or safety, and may increase user over-trust without adequate risk-checking.

Outcomes and Acceptability

Across studies, ChatGPT/LLMs were generally acceptable to end users and trainees. In education, skill gains were observed when practice was scaffolded with rubrics or supervision (Cioffi et al., 2025; Louie et al., 2025), and expert raters judged trainee case conceptualizations as acceptable in controlled tasks (Hsieh et al., 2024). In clinical/adjacent contexts, users frequently reported empathy, availability, and convenience (Luo et al., 2025; Raile, 2024; Vowels et al., 2024). Early signals of benefit included quality-of-life improvement in a small inpatient pilot and short-term anxiety reduction when the chatbot provided empathetic feedback in a single-session experiment, alongside a clinical RCT outside psychiatry showing reduced anxiety/depression with a ChatGPT adjunct (Akdogan et al., 2025; Kishimoto et al., 2025; Melo et al., 2024). At the same time, risk detection, complex reasoning, and cultural fit were limited, and privacy/boundary issues were common in open communities (Dergaa et al., 2024; Luo et al., 2025).

Risk of Bias and Evidence Maturity

Beyond two randomized trials, most evidence derives from small pilots, non-randomized experiments, simulations, or observational social-media analyses. Training gains frequently depended on prompt scaffolding and human feedback; clinical pilots lacked control groups; and public/self-help findings are vulnerable to self-selection and platform effects. Safety-critical tasks (e.g., suicide-risk appraisal) repeatedly showed under-identification of risk or inconsistent stratification relative to clinicians, and performance degraded with case complexity. Overall, the body of evidence supports adjunct roles in education and low-risk psychoeducation, and indicates that deployment in high-risk pathways requires human oversight, clear escalation, and larger controlled trials.

Discussion

Principal Findings

This scoping review synthesized 34 studies examining ChatGPT/LLMs in mental health contexts across education/supervision, clinical and clinical-adjacent pilots, and public/self-help use. Taken together, the literature indicates feasibility and acceptability for adjunctive use rather than replacement of human intervention. In training settings, the most consistent signals appear where interaction is scaffolded: one randomized counselor-training trial reported gains in reflective/questioning skills when novices practiced with an AI “patient” and received rubric-guided feedback (Louie et al., 2025), and a comparative supervision study with Gestalt trainees found AI feedback useful while underscoring the need for human oversight (Cioffi et al., 2025). Expert raters also judged trainee case conceptualizations generated with ChatGPT support to be acceptable, suggesting viability for early-stage skill development (Hsieh et al., 2024).

In clinical and clinical-adjacent domains, early outcome signals include quality-of-life improvement in a small inpatient pilot (Melo et al., 2024) and short-term anxiety reduction in a single-session study when the chatbot provided empathetic feedback (Kishimoto et al., 2025). Outside psychiatry, a clinical RCT showed reduced anxiety and depression with a ChatGPT-based counseling adjunct compared with usual care (Akdogan et al., 2025). In public/self-help settings, digital ethnography and content analyses describe users informally shaping ChatGPT into a “digital therapist,” citing constant availability, non-judgmental tone, and a low threshold for disclosure (Giray, 2025; Luo et al., 2025); complementary user studies in relationship support report high perceived empathy and usability but weak risk-checking and limited collaborative problem solving (Vowels et al., 2024). Across these streams, three limitations recur and temper conclusions about autonomous use: models tend to under-identify suicide risk or stratify it inconsistently relative to clinicians (Elyoseph & Levkovich, 2023; Ghanadian et al., 2023; Levkovich & Elyoseph, 2023), performance degrades with case complexity (Dergaa et al., 2024), and cultural humility/adaptation is uneven (Aleem et al., 2024). The net picture is therefore one of adjunct value, particularly for scaffolded education and low-risk psychoeducation paired with a consistent requirement for human oversight, clear scope boundaries, and escalation mechanisms for higher-risk situations.

Implications for Human-Computer Interaction

From an HCI standpoint, the present findings call for interfaces that calibrate trust without dampening legitimate benefits. Users frequently experience interactions as empathic and helpful (Luo et al., 2025; Vowels et al., 2024), yet this same empathic fluency can mask gaps in accuracy, cultural fit, and risk detection documented in simulations and risk

vignettes (Aleem et al., 2024; Dergaa et al., 2024; Elyoseph & Levkovich, 2023). Interfaces should therefore open with transparent scope-of-use messaging and plain-language statements about capabilities and limits; they should make handoff to humans obvious and actionable at the moment of need (Blease & Torous, 2023; Carlbring et al., 2023). Because training and formulation benefits in the evidence base are strongest under prompt scaffolding and rubric-guided feedback (Cioffi et al., 2025; Hsieh et al., 2024; Louie et al., 2025), the interaction layer should embed that scaffolding, offering templates, exemplars, and reflective prompts that keep tasks within competence boundaries while enabling instructors or clinicians to review outputs. In clinical-adjacent and public contexts, the same accessibility features that enable disclosure (24/7 availability, non-judgmental tone) create privacy and boundary risks (Luo et al., 2025); privacy-preserving defaults, visible data-handling notices, and gentle friction before sharing sensitive details can mitigate over-disclosure (Kalam et al., 2024; Monteith et al., 2024). For safety-critical tasks, risk cues must trigger escalation affordances and conservative behaviors (e.g., de-escalation language, verified hotline options, and clear signals that the system is not performing an assessment), consistent with ethical analyses that frame LLMs as an “artificial third” reshaping therapeutic dynamics and responsibilities (Tal et al., 2023).

Comparison With Prior Literature

Earlier work on purpose-built conversational agents for mental health (e.g., CBT-module systems) emphasized narrow scripting and hard safety rails, which delivered predictable protocols but often struggled with engagement and personalization. ChatGPT differs as a general-purpose LLM that users can steer into therapist-like roles across unregulated contexts. The flexibility documented here helps explain higher perceived empathy, availability, and user-driven shaping (Luo et al., 2025; Vowels et al., 2024), while also explaining why safety and reliability are more variable: unlike legacy systems, ChatGPT’s performance depends heavily on prompting, task framing, and human oversight (Cioffi et al., 2025; Dergaa et al., 2024; Louie et al., 2025). The contrast clarifies the current design challenge for HCI: to retain the engagement advantages of open-ended conversation while reconstructing, at the interface level, the guardrails, escalation paths, and scope cues that purpose-built tools obtained by design (Blease & Torous, 2023; Carlbring et al., 2023). The present synthesis also refines claims from early viewpoint pieces by anchoring them in empirical patterns: improvements are most credible in scaffolded training and low-risk psychoeducation (Akdogan et al., 2025; Hsieh et al., 2024; Kishimoto et al., 2025; Melo et al., 2024), whereas safety-critical decision support remains inappropriate without human review (Elyoseph & Levkovich, 2023; Ghanadian et al., 2023; Levkovich & Elyoseph, 2023).

Equity Considerations

Equity concerns surfaced across populations and settings, though they remain insufficiently evaluated in controlled studies. Pediatric and adolescent perspectives call for age-appropriate consent, privacy protections, and conservative defaults, reflecting the higher-risk sensitivity in youth mental health contexts (Imran et al., 2023; van Schalkwyk, 2023). Cross-cultural prompt-testing and qualitative reports point to deficits in cultural humility and adaptation, with models sometimes misreading idioms or norms; at the interface level, this argues for localized content, opportunities for users to correct cultural assumptions without penalty, and transparent acknowledgment when the model’s fit is uncertain (Alanezi, 2024; Aleem et al., 2024). Studies from resource-constrained settings describe potential to narrow access gaps where human services are scarce, but also warn that uneven infrastructure and AI-literacy disparities could widen inequities if adoption proceeds without support (Alanezi, 2024). In counselor education, widespread informal use risks amplifying skill gaps unless curricula provide scaffolded AI literacy and explicit criteria for appropriate vs. inappropriate use (Ajilouni et al., 2023; Maurya & Cavanaugh, 2025; Maurya, 2024; Rehman & Sajjad, 2025). An equity-by-design approach combining localization, multilingual support, privacy-first defaults, and embedded handoff to human intervention follows directly from the mixed benefits and risks documented in the “Results” section.

Strength of Evidence and Limitations of the Literature

The evidence base remains early-stage. Beyond two randomized trials (one in counselor training and one clinical adjunct outside psychiatry) and a non-randomized single-session experiment (Akdogan et al., 2025; Kishimoto et al., 2025; Louie et al., 2025), most studies are small pilots, simulations, observational social-media analyses, or expert-rating/qualitative designs (e.g., Dergaa et al., 2024; Luo et al., 2025; Melo et al., 2024). Training benefits often rely on prompt scaffolding and human feedback (Cioffi et al., 2025; Louie et al., 2025); clinical pilots lack controls (Melo et al., 2024); and public-use findings are vulnerable to self-selection and platform effects (Xu et al., 2024). At the same time, multiple independent strands of suicide-risk vignettes, simulations under complexity, and multicultural prompt-testing converge on the conclusion that unsupervised use is unsafe in high-risk contexts (Aleem et al., 2024; Dergaa et al., 2024; Elyoseph & Levkovich, 2023). Future work should therefore prioritize controlled evaluations in clinical populations, prospective safety testing of escalation mechanisms, cross-cultural studies with representative samples, and interface-level experiments that test whether trust-calibrating designs (e.g., uncertainty cues, scope prompts, supervised modes) reduce error, over-reliance, and harm.

Future Directions

The trajectory of this evidence base points toward several priorities for research and design. First, the most promising applications appear to be in structured, adjunctive roles, such as psychoeducation, skills rehearsal, and reflective writing. Future studies should therefore move beyond exploratory pilots and systematically evaluate ChatGPT as an adjunct to, not a replacement for, human intervention. Controlled trials in these bounded contexts, combined with qualitative follow-up, would help clarify both efficacy and risks.

Second, there is a need to examine the long-term dynamics of use, particularly around dependency, coping, and expectations of therapeutic alliance. Most available studies are short-term, yet our synthesis shows that users often anthropomorphize ChatGPT and develop patterns of repeated engagement. Understanding whether these interactions support resilience or foster over-reliance will require longitudinal research.

Third, the HCI community is well-positioned to investigate design interventions that can calibrate trust and support appropriate reliance. This includes testing mechanisms such as real-time disclaimers, cultural adaptation prompts, privacy-preserving defaults, and crisis escalation pathways. Such features could help transform ChatGPT's perceived empathy and accessibility into safe and sustainable affordances.

Finally, equity considerations demand sustained attention. Evidence from pediatric and cross-cultural contexts shows that risks are not evenly distributed: young users, low-resource populations, and those from non-Western cultural backgrounds face distinctive challenges. Future work should therefore adopt equity-by-design principles, exploring how multilingual models, localized training data, and scaffolded AI literacy programs can ensure that ChatGPT's benefits are broadly shared rather than stratified.

Acknowledgments

This research was supported by the Ibn Haldun University for their valuable support in making this study possible.

ORCID iD

Thseen Nazir  <https://orcid.org/0000-0002-5541-7749>

Ethical Considerations

This study is a literature review and does not involve human participants or animals. Therefore, ethical approval from an institutional ethics committee was not required. All sources used in this study have been appropriately cited and referenced in accordance with academic integrity standards.

Consent to Participate

This study does not involve human participants; therefore, informed consent was not required.

Consent for Publication

Not applicable.

Author Contributions

TN is solely responsible for all aspects of this research, including the conceptualization and design of the study, data collection, analysis, and interpretation. The author also conducted the literature review, developed the research methodology, and wrote the entire manuscript. All revisions and final approvals were completed by the author, ensuring the integrity and accuracy of the research presented.

Funding

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

Declaration of Conflicting Interests

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

Data Availability Statement

The data supporting this study's findings are available on request from the corresponding author, TN.

Declaration of Generative AI in Scientific Writing

During the preparation of this work, the authors used Grammarly AI to improve fluency and clarity. The authors reviewed and edited the content and take full responsibility for the final version.

Supplemental Material

Supplemental material for this article is available online.

References

- Ajlouni, A. O., Almahaireh, A. S., & Whaba, F. A.-A. (2023). Students' perception of using ChatGPT in counseling and mental health education: The benefits and challenges. *International Journal of Emerging Technologies in Learning (IJET)*, 18(20), 199–218. <https://doi.org/10.3991/ijet.v18i20.42075>
- Akdogan, O., Uyar, G. C., Yesilbas, E., Baskurt, K., Malkoc, N. A., Ozdemir, N., Yazici, O., Oksuzoglu, B., Uner, A., Ozet, A., & Sutcuoglu, O. (2025). Effect of a ChatGPT-based digital counseling intervention on anxiety and depression in patients with cancer: A prospective, randomized trial. *European Journal of Cancer*, 221, Article 115408. <https://doi.org/10.1016/j.ejca.2025.115408>
- Alanezi, F. (2024). Assessing the effectiveness of ChatGPT in delivering mental health support: A qualitative study. *Journal of Multidisciplinary Healthcare*, 17, 461–471. <https://doi.org/10.2147/JMDH.S447368>
- Aleem, M., Zahoor, I., & Naseem, M. (2024). Towards culturally adaptive large language models in mental health: Using ChatGPT as a case study. In *Proceedings of the ACM conference on computer supported cooperative work, CSCW* (pp. 240–247). Association for Computing Machinery. <https://doi.org/10.1145/3678884.3681858>
- Ayo-Ajibola, O., Davis, R. J., Lin, M. E., Riddell, J., & Kravitz, R. L. (2024). Characterizing the adoption and experiences of users of artificial intelligence-generated health information in the United States: Cross-sectional questionnaire study. *Journal of Medical Internet Research*, 26, Article e55138. <https://doi.org/10.2196/55138>

- Blease, C., & Torous, J. (2023). ChatGPT and mental health-care: Balancing benefits with risks of harms. *BMJ Mental Health, 26*(1), Article e300884. <https://doi.org/10.1136/bmjment-2023-300884>
- Carlbring, P., Hadjistavropoulos, H., Kleiboer, A., & Andersson, G. (2023). A new era in internet interventions: The advent of Chat-GPT and AI-assisted therapist guidance. In *Internet Interventions, 32*, Article 100621. <https://doi.org/10.1016/j.invent.2023.100621>
- Cheng, S., Chang, C., Chang, W., Wang, H., Liang, C., Kishimoto, T., Chang, J. P., Kuo, J. S., & Su, K. (2023). The now and future of ChatGPT and GPT in psychiatry. *Psychiatry and Clinical Neurosciences, 77*(11), 592–596. <https://doi.org/10.1111/pcn.13588>
- Cioffi, V., Ragozzino, O., Mosca, L. L., Moretto, E., Tortora, E., Acocella, A., Montanari, C., Ferrara, A., Crispino, S., Gigante, E., Lommatzsch, A., Pizzimenti, M., Temporin, E., Barlacchi, V., Billi, C., Salonia, G., & Sperandio, R. (2025). Can AI technologies support clinical supervision? Assessing the potential of ChatGPT. *Informatics, 12*(1), Article 29. <https://doi.org/10.3390/informatics12010029>
- Collins, A. C., Lekkas, D., Heinz, M. V., Annor, J., Ruan, F., & Jacobson, N. C. (2025). ChatGPT as therapy: A qualitative and network-based thematic profiling of shared experiences, attitudes, and beliefs on Reddit. *Journal of Psychiatric Research, 191*, 277–288. <https://doi.org/10.1016/j.jpsychires.2025.09.057>
- Cross, S., Bell, I., Nicholas, J., Valentine, L., Mangelsdorf, S., Baker, S., Titov, N., & Alvarez-Jimenez, M. (2024). Use of AI in mental health care: Community and mental health professionals survey. *JMIR Mental Health, 11*, Article e60589. <https://doi.org/10.2196/60589>
- Dergaa, I., Fekih-Romdhane, F., Hallit, S., Loch, A. A., Glenn, J. M., Fessi, M. S., Ben Aissa, M., Souissi, N., Guelmami, N., Swed, S., El Omri, A., Bragazzi, N. L., & Ben Saad, H. (2024). ChatGPT is not ready yet for use in providing mental health assessment and interventions. *Frontiers in Psychiatry, 14*, Article 1277756. <https://doi.org/10.3389/fpsy.2023.1277756>
- Elyoseph, Z., & Levkovich, I. (2023). Beyond human expertise: The promise and limitations of ChatGPT in suicide risk assessment. *Frontiers in Psychiatry, 14*, Article 1213141. <https://doi.org/10.3389/fpsy.2023.1213141>
- Eryilmaz, A., & Bařal, A. (2024). Rational AIs with emotional deficits: ChatGPT vs. counselors in providing emotional reflections. *Current Psychology, 43*(45), 34962–34977. <https://doi.org/10.1007/s12144-024-06947-w>
- Ghanadian, H., Nejadgholi, I., & Al Osman, H. (2023). ChatGPT for suicide risk assessment on social media: Quantitative evaluation of model performance, potentials and limitations. In *Proceedings of the 13th workshop on computational approaches to subjectivity, sentiment, & social media analysis* (pp. 172–183). <https://doi.org/10.18653/v1/2023.wassa-1.16>
- Giray, L. (2025). Cases of using ChatGPT as a mental health and psychological support tool. *Journal of Consumer Health and the Internet, 29*(1), 29–48. <https://doi.org/10.1080/15398285.2024.2442374>
- Hadar-Shoval, D., Elyoseph, Z., & Lvovsky, M. (2023). The plasticity of ChatGPT's mentalizing abilities: Personalization for personality structures. *Frontiers in Psychiatry, 14*, 1234397. <https://doi.org/10.3389/fpsy.2023.1234397>
- 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), Article 100083. <https://doi.org/10.1016/j.chbah.2024.100083>
- Hwang, G., Lee, D. Y., Seol, S., Jung, J., Choi, Y., Her, E. S., An, M. H., & Park, R. W. (2024). Assessing the potential of ChatGPT for psychodynamic formulations in psychiatry: An exploratory study. *Psychiatry Research, 331*, Article 115655. <https://doi.org/10.1016/j.psychres.2023.115655>
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology, 15*(4), Article ep464. <https://doi.org/10.30935/cedtech/13605>
- Imran, N., Hashmi, A., & Imran, A. (2023). Chat-GPT: Opportunities and challenges in child mental healthcare. *Pakistan Journal of Medical Sciences, 39*(4), 1191–1193. <https://doi.org/10.12669/PJMS.39.4.8118>
- Jung, K., Lee, G., Huang, Y., & Chen, Y. (2025). "I've talked to ChatGPT about my issues last night": Examining Mental Health Conversations with Large Language Models through Reddit Analysis. In *Proceedings of the ACM on human-computer interaction* (pp. 1–25). Association for Computing Machinery. <https://doi.org/10.1145/3757537>
- Kalam, K. T., Rahman, J. M., Islam Md, R., & Dewan, S. M. R. (2024). ChatGPT and mental health: Friends or foes? *Health Science Reports, 7*(2), Article 1912. <https://doi.org/10.1002/hsr2.1912>
- Kishimoto, T., Hao, X., Chang, T., & Luo, Z. (2025). Single online self-compassion writing intervention reduces anxiety: With the feedback of ChatGPT. *Internet Interventions, 39*, Article 100810. <https://doi.org/10.1016/j.invent.2025.100810>
- Levkovich, I., & Elyoseph, Z. (2023). Suicide risk assessments through the eyes of ChatGPT-3.5 versus ChatGPT-4: Vignette study. *JMIR Mental Health, 10*, Article e51232. <https://doi.org/10.2196/51232>
- Louie, R., Orney, I. H., Pacheco, J. P., Shah, R. S., Brunskill, E., & Yang, D. (2025). Can LLM-simulated practice and feedback upskill human counselors? A randomized study with 90+ novice counselors. *arXiv*. <https://doi.org/10.48550/arXiv.2505.02428>
- Luo, X., Ghosh, S., Tilley, J. L., Besada, P., Wang, J., & Xiang, Y. (2025). "Shaping ChatGPT into my digital Therapist": A thematic analysis of social media discourse on using generative artificial intelligence for mental health. *Digital Health, 11*, Article 1351088. <https://doi.org/10.1177/20552076251351088>
- Maurya, K., & Cavanaugh, T. (2025). Counselor education in the era of ChatGPT and other artificial intelligence-based Chatbots. *Journal of Asia Pacific Counseling, 15*(1), 1–33. <https://doi.org/10.18401/2025.15.1.1>
- Maurya, R. K. (2024). Using AI based chatbot ChatGPT for practicing counseling skills through role-play. *Journal of Creativity in Mental Health, 19*(4), 513–528. <https://doi.org/10.1080/15401383.2023.2297857>
- Melo, A., Silva, I., & Lopes, J. (2024). ChatGPT: A pilot study on a promising tool for mental health support in psychiatric inpatient care. *International Journal of Psychiatric Trainees, 2*(2), Article 92367. <https://doi.org/10.55922/001c.92367>
- Monteith, S., Glenn, T., Geddes, J. R., Whybrow, P. C., Achtyes, E., & Bauer, M. (2024). Artificial intelligence and increasing

- misinformation. *The British Journal of Psychiatry*, 224(2), 33–35. <https://doi.org/10.1192/bjp.2023.136>
- Naher, J. (2024). Can ChatGPT provide a better support: A comparative analysis of ChatGPT and dataset responses in mental health dialogues. *Current Psychology*, 43(28), 23837–23845. <https://doi.org/10.1007/s12144-024-06140-z>
- Raile, P. (2024). The usefulness of ChatGPT for psychotherapists and patients. *Humanities and Social Sciences Communications*, 11(1), Article 47. <https://doi.org/10.1057/s41599-023-02567-0>
- Rehman, F., & Sajjad, S. (2025). Bridging technology and therapy: Exploring AI in mental health services through counselors' and students' perspectives. *Online Media and Society*, 6(1), 31–44. <https://doi.org/10.71016/oms/hn8an983>
- 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), Article 48. <https://doi.org/10.1038/s44184-024-00097-4>
- Singh, O. (2019). Chatbots in psychiatry: Can treatment gap be lessened for psychiatric disorders in India. *Indian Journal of Psychiatry*, 61(3), Article 225. <https://doi.org/10.4103/0019-5545.258323>
- Song, I., Pendse, S. R., Kumar, N., & De Choudhury, M. (2025). The typing cure: Experiences with large language model chatbots for mental health support. *Proceedings of the ACM on Human-Computer Interaction*, 9(7), 1–29. <https://doi.org/10.1145/3757430>
- Tal, A., Elyoseph, Z., Haber, Y., Angert, T., Gur, T., Simon, T., & Asman, O. (2023). The artificial third: Utilizing ChatGPT in mental health. *The American Journal of Bioethics*, 23(10), 74–77. <https://doi.org/10.1080/15265161.2023.2250297>
- van Schalkwyk, G. (2023). Artificial intelligence in pediatric behavioral health. *Child and Adolescent Psychiatry and Mental Health*, 17(1), Article 38. <https://doi.org/10.1186/s13034-023-00586-y>
- Vowels, L. M., Francois-Walcott, R. R. R., & Darwiche, J. (2024). AI in relationship counselling: Evaluating ChatGPT's therapeutic capabilities in providing relationship advice. *Computers in Human Behavior: Artificial Humans*, 2(2), Article 100078. <https://doi.org/10.1016/j.chbah.2024.100078>
- Xu, Z., Fang, Q., Huang, Y., & Xie, M. (2024). The public attitude towards ChatGPT on reddit: A study based on unsupervised learning from sentiment analysis and topic modeling. *PLOS ONE*, 19(5), Article e0302502. <https://doi.org/10.1371/journal.pone.0302502>
- Zhang, Z., & Wang, J. (2024). Can AI replace psychotherapists? Exploring the future of mental health care. *Frontiers in Psychiatry*, 15, 1444382. <https://doi.org/10.3389/fpsy.2024.1444382>