

REVIEW ARTICLE **OPEN ACCESS**

Chatbot-Delivered Interventions for Improving Mental Health Among Young People: A Systematic Review and Meta-Analysis

Jiaying Li  | Yan Li  | Yule Hu  | Dennis Chak Fai Ma  | Xiaoxiao Mei  | Engle Angela Chan  | Janelle Yorke 

School of Nursing, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Correspondence: Xiaoxiao Mei (xiaoxiao.mei@connect.polyu.hk)

Received: 21 December 2024 | **Revised:** 12 February 2025 | **Accepted:** 28 May 2025

Funding: The authors received no specific funding for this work.

Keywords: chatbot | conversational agent | mental health | psychological distress | psychological well-being | young people

ABSTRACT

Background: The characteristics, application, and effectiveness of chatbots in improving the mental health of young people have yet to be confirmed through systematic review and meta-analysis.

Aim: This systematic review aims to evaluate the effectiveness of chatbot-delivered interventions for improving mental health among young people, identify factors influencing effectiveness, and examine feasibility and acceptability.

Methods: To identify eligible interventional studies, we systematically searched 11 databases and search engines covering a publication period of January 2014 to September 2024. Meta-analyses and subgroup analyses were performed on randomized controlled trials to investigate the effectiveness of chatbot-delivered interventions and potential influencing factors. Narrative syntheses were conducted to summarize the feasibility and acceptability of these interventions in all the included studies.

Results: We identified 29 eligible interventional studies, 13 of which were randomized controlled trials. The meta-analysis indicated that chatbot-delivered interventions significantly reduced distress (Hedge's $g = -0.28$, 95% CI $[-0.46, -0.10]$), but did not have a significant effect on psychological well-being (Hedge's $g = 0.13$, 95% CI $[-0.16, 0.41]$). The observed treatment effects were influenced by factors including sample type, delivery platform, interaction mode, and response generation approach. Overall, this review demonstrates that chatbot-delivered interventions were feasible and acceptable.

Linking Evidence to Action: This review demonstrated that chatbot-delivered interventions had positive effects on psychological distress among young people. Chatbot-delivered interventions have the potential to supplement existing mental health services provided by multidisciplinary healthcare professionals. Future recommendations include using instant messenger platforms for delivery, enhancing chatbots with multiple communication methods to improve interaction quality, and refining language processing, accuracy, privacy, and security measures.

1 | Introduction

In 2024, the global population of young people aged 10–24 years (i.e., adolescents and young adults) was approximately 1.95 billion, accounting for 24% of the total population (United

Nations Population Fund 2024). Recent trends indicate an increasing prevalence of mental health problems, such as symptoms of depression and anxiety, among this age group (Lipson et al. 2019; Li, Bressington, et al. 2021). Moreover, 62.5% of mental disorders emerge during this period, resulting in

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *Worldviews on Evidence-Based Nursing* published by Wiley Periodicals LLC on behalf of Sigma Theta Tau International.

lifelong effects on individuals' well-being and social functioning (Solmi et al. 2022). Despite the critical importance of intervention and preventive measures, a significant proportion of young people in need do not receive adequate mental health services (Islam et al. 2022; Yan, et al. 2023). The most prominent barriers to accessing these services are attributed to the scarcity of mental health resources and the pervasive stigma surrounding psychological issues (Aguirre Velasco et al. 2020).

As an innovative subset of digital technology, chatbots have been applied to mental health interventions (Abd-Alrazaq et al. 2019). Chatbots are autonomous systems that simulate and process conversations across various modalities, including written, spoken, and visual languages, enabling synchronous interactions between humans and digital devices (McTear et al. 2016). Chatbots, being conducted on platforms independent of time and place, offer enhanced accessibility compared to traditional face-to-face treatments (Lim et al. 2022). By providing users with anonymity, chatbots effectively address concerns related to discussing sensitive topics and combat the stigma associated with seeking help (Li, Lee, et al. 2024). The evolution of chatbots for mental health issues has transitioned from rule-based bots with predefined interactions to artificial intelligence (AI) chatbots that employ algorithms or neural networks to process natural language (Abd-Alrazaq et al. 2019).

The rapid development of chatbot technology presents significant potential in supporting the mental health of young people, which is further reinforced by the increasing demand for digital tools in this context (Pretorius and Coyle 2021). Chatbots have demonstrated diverse applications in the field of youth mental health, including screening, preventive psychoeducation, and therapy (Balan et al. 2024). Furthermore, there is a growing body of evidence supporting the positive effects of chatbot-delivered interventions on the mental health of young people (Li, Chung, et al. 2024; Vertsberger et al. 2022).

Several systematic reviews have synthesized evidence on the effectiveness of chatbots in promoting mental health. However, these studies have primarily focused on the adult or general population (Abd-Alrazaq et al. 2020; He et al. 2023; Li, Zhang, et al. 2023; Lim et al. 2022), with limited attention given to young people. Nevertheless, considering the unique social, academic, and career challenges faced by young people, their distinct needs and preferences for chatbot features and patterns of usage may differ from those of other age groups (Brandtzaeg et al. 2021). Furthermore, their higher proficiency and acceptance of digital technology may contribute to a broader demand for chatbots (van Doorn et al. 2021). Given the increasing prevalence of mental health issues among young people, it is crucial to explore the characteristics and effectiveness of chatbots tailored specifically to this population. Indeed, a scoping review has provided an overview of the characteristics and application of chatbots in improving the mental health of the young population (Balan et al. 2024), the conclusion regarding their effectiveness has yet to be confirmed due to the absence of rigorous quality assessment and meta-analysis of outcomes.

1.1 | Aims

To address these knowledge gaps, this systematic review aims to (1) evaluate the effectiveness of chatbot-delivered interventions on mental health (including psychological distress and well-being) in young people and explore factors that may influence the magnitude of these treatment effects and (2) summarize the feasibility and acceptability of such interventions.

2 | Methods

This review was reported in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement (Page et al. 2021). The protocol was registered with PROSPERO (CRD42024529404).

2.1 | Eligibility Criteria

The eligibility criteria were developed based on the PICOS framework:

1. *Population*: The target population comprised young people aged between 10 and 24 years as end users (United Nations Department of Economic and Social Affairs 2016). Studies with a broader age range were considered eligible if either the mean age or over 50% of the participants fell within the predefined age range.
2. *Intervention*: Studies incorporating autonomous chatbots involving synchronous two-way interactions with users were included. Additionally, chatbots that are part of virtual reality and robots were also considered.
3. *Control*: Chatbot-delivered interventions were compared to control groups receiving various types of comparisons, ranging from active (e.g., psychoeducation) to inactive (e.g., usual care, waitlist control), or those without a direct control (e.g., uncontrolled pre-post evaluation).
4. *Outcome*: Any outcomes related to psychological distress or well-being were considered, including anxiety, depression, stress, mood, self-esteem, well-being, and coping, as defined by the Cochrane Consumers and Communication Group (Consumers and Communication Group 2012). Included studies were required to measure these outcomes using validated questionnaires or objective assessments (e.g., cortisol measurement).
5. *Design*: Any interventional study design was considered eligible for inclusion.
6. *Publication type and language*: The review included peer-reviewed articles, dissertations, conference proceedings, and reports published in English and Chinese. Exclusion criteria included reviews, conference abstracts, proposals, editorials, and letters.

2.2 | Search Strategy

A comprehensive search was conducted covering the publication period from January 2014 to September 2024. This 10-year

timeframe was selected based on the recognition that the rise of chatbot technology may be traced back to 2014 (Grudin and Jacques 2019), and most of the relevant research has been carried out during this period (Balan et al. 2024; Li, Zhang, et al. 2023). Nine bibliographic databases were searched: Embase, PubMed, Scopus, The Cochrane Library, PsycINFO, IEEE Xplore, ACM Digital Library, China National Knowledge Infrastructure, and Wanfang Data. In addition, gray literature and relevant records were further explored using “Bielefeld Academic search engine” and “Google Scholar.” Due to the substantial number of studies retrieved by Google Scholar, the screening was limited to the first 100 hits ranked by relevance to the search topic. The search terms were constructed using keywords related to the chatbot, the target population, and the desired outcomes (refer to Table S1 for a complete sample of the search strategy). Furthermore, the reference lists of included publications were manually checked to identify any additional relevant articles.

2.3 | Study Selection

After removing duplicate records, the study selection process involved two steps. Initially, two reviewers (Authors 1 and 3) independently screened the titles and abstracts based on the pre-determined eligibility criteria. Subsequently, the full texts of the studies meeting the criteria from the first step were assessed by the same reviewers. Any discrepancies were resolved through discussion or consultation with a third reviewer (Author 2). Inter-rater reliability was evaluated using Cohen's κ , resulting in values of 0.88 and 0.90 for the first and second steps, respectively, indicating an excellent level of agreement (McHugh 2012).

2.4 | Data Extraction

A comprehensive data extraction form were developed and underwent pilot testing on a randomly selected subset of five full-text studies. The extracted information encompassed the following aspects: (1) general study information (author, publication year, country, setting, publication type, methodology, and study design); (2) participant characteristics (sample type, sample size, age, and gender); (3) chatbot-delivered intervention details (interventional aspects: purpose, theory basis, use of co-design, deployment, and duration; technical aspects (platform, response generation, interaction mode, embodiment, and safety measures); (4) controls; (5) mental health outcomes and measurement tools; and (6) summary of results (main findings, feasibility, and acceptability for each study). Table S2 provides the detailed definitions for each item. Data extraction was performed by two independent reviewers Authors 1 and 3), and any discrepancies were resolved through discussion with a third reviewer (Author 2).

2.5 | Quality Appraisal and Certainty of Evidence

The risk of bias in randomized controlled trials (RCTs) was assessed using the revised Cochrane risk-of-bias tool (Sterne et al. 2019). Two reviewers (authors 1 and 5) independently evaluated bias across five domains: randomization, deviations

from intended interventions, missing outcome data, measurement of the outcome, and selective reporting. Each domain was assessed based on specific questions (Yes/Probably yes/No/Probably no/No information), and an overall risk of bias level (“High”/“Low”/“Some concerns”) was classified to each outcome. Considering the difficulty of blinding participants in chatbot-delivered interventions, a “high” risk of bias was assigned in the measurement of the outcome domain when a clear difference existed between the intervention and comparison groups (i.e., usual care and waitlist control). The risk of bias was rated as “some concerns” when participants received alternative treatments (i.e., psychoeducation). Regarding the quasi-experimental studies, the risk of bias was evaluated using the Mixed Methods Appraisal Tool (MMAT) (Hong et al. 2018). Five criteria were assessed to determine bias: representativeness of the sample, appropriateness of measurements, completeness of outcome data, treatment of confounding factors, and whether the intervention operated as intended. Each criterion was rated as “Yes,” “No,” or “Can't tell” to assess the level of bias.

The quality of the meta-analysis results was appraised using the Grading of Recommendation, Assessment, Development, and Evaluation guideline (Guyatt et al. 2011). Two independent reviewers (authors 1 and 5) assessed the level of certainty by considering five domains: risk of bias, indirectness, inconsistency, imprecision, and publication bias. Subsequently, the evidence was categorized into four levels: “high”, “moderate”, “low”, and “very low”, based on the assessments conducted.

2.6 | Data Synthesis

The meta-analyses of RCTs were performed using the metafor package in R software (version 4.3.2). Hedges' g , which represents the standard mean differences, along with their corresponding 95% confidence intervals (CI), were computed for each study. The effect sizes were categorized as small ($g=0.2$), medium ($g=0.5$), or large ($g=0.8$) (Hedges and Olkin 2014). A random-effects model was employed for all analyses to account for variations across studies. The determination of Hedges' g utilized post-intervention data (mean, sample size, and standard deviation). In cases where studies had multiple groups, relevant intervention or control groups were merged to establish a single pairwise comparison (Higgins et al. 2023). Missing data were obtained by contacting corresponding authors, while studies lacking essential data (mean, sample size, standard deviation) were excluded from the meta-analysis.

Given that all the included RCTs provided data on measures of psychological distress (i.e., depression, anxiety, stress, and negative affect) (Sartore et al. 2021) and/or psychological well-being (i.e., well-being, positive affect, coping self-efficacy, and mental health self-efficacy) (Li, Zhang, et al. 2023), we conducted two distinct meta-analyses to calculate the pooled effect sizes for these two overarching mental health outcomes. However, since most included studies contributed multiple effect sizes when evaluating these two categories, combining these correlated effect sizes could lead to an overestimated overall effect (Van den Noortgate et al. 2015). To account for multiple effect sizes within individual trials and optimize statistical power, we utilized two three-level random-effects meta-analytical models that handle

dependencies between effect sizes from the same study (Assink and Wibbelink 2016). Heterogeneity was assessed using I^2 and Cochrane's Q statistics, with I^2 values categorized as “might not be important” (0%–40%), “moderate” (30%–60%), “substantial” (50%–90%), or “considerable” (75%–100%) heterogeneity (Deeks et al. 2023). The assessment of publication bias was conducted using Egger's regression test. Sensitivity analyses were conducted using the leave-one-out method to assess the impact of excluding individual trials and identify potential sources of heterogeneity. The meta-analysis included four major subgroup analyses to explore variations in efficacy among sample type (clinical/subclinical vs. nonclinical), platform (instant messenger vs. mobile application vs. web-based), interaction mode (text-based vs. multimodal), and response generation (rule-based vs. AI-based). Only outcome data from RCTs were retrieved for meta-analysis. Narrative syntheses were performed on all included studies to summarize the feasibility and acceptability of chatbot-delivered interventions.

3 | Results

3.1 | Study Selection

A total of 4031 records were obtained from 11 databases and search engines ($n=4011$ records) and reference lists ($n=20$). After removing duplicate records ($n=1415$) and screening titles and abstracts ($n=2469$), 147 relevant articles were selected for full-text screening. Ultimately, 29 interventional studies met the eligibility criteria for inclusion in this review. Among them, 16 studies were not RCTs, and 2 studies did not provide sufficient outcome data, resulting in 11 studies being included in the meta-analysis. Figure 1 presents the PRISMA flow chart illustrating the search and selection process.

3.2 | Study Characteristics

3.2.1 | General Study Information

As shown in Table 1, the methodological approach most frequently employed was mixed methods ($n=20$). 55.2% of the studies ($n=16$) adopted a quasi-experimental design, while the remaining studies were RCTs ($n=13$). Most studies ($n=21$) were published as peer-reviewed journal articles. The studies encompassed three distinct income groups, with high-income countries (Hamadeh et al. 2023) accounting for 62.1% of the total studies ($n=18$). Specifically, the research was conducted across 12 countries, with the United States ($n=9$), China ($n=5$), and New Zealand ($n=3$) being the most frequently represented locations (Table S3). Interventions were primarily conducted in educational settings ($n=19$), followed by community settings ($n=7$) and hospitals ($n=3$). Only one study was reported in Chinese (Liu 2022), and the remaining studies were reported in English.

3.2.2 | Characteristics of Participants, Controls, and Outcomes

The age range of participants in the included studies varied from 8 to 33 years, with male participants making up an average of 38.3%. Twenty studies recruited nonclinical populations, five studies enrolled participants with psychological symptoms, and the remaining four studies focused on young people diagnosed with physical or mental disorders. In 15 studies, chatbot-delivered interventions were compared with a control condition. The studies encompassed a wide range of mental health outcomes, with the severity of depression ($n=14$) and anxiety ($n=14$) most commonly assessed. A summary of study

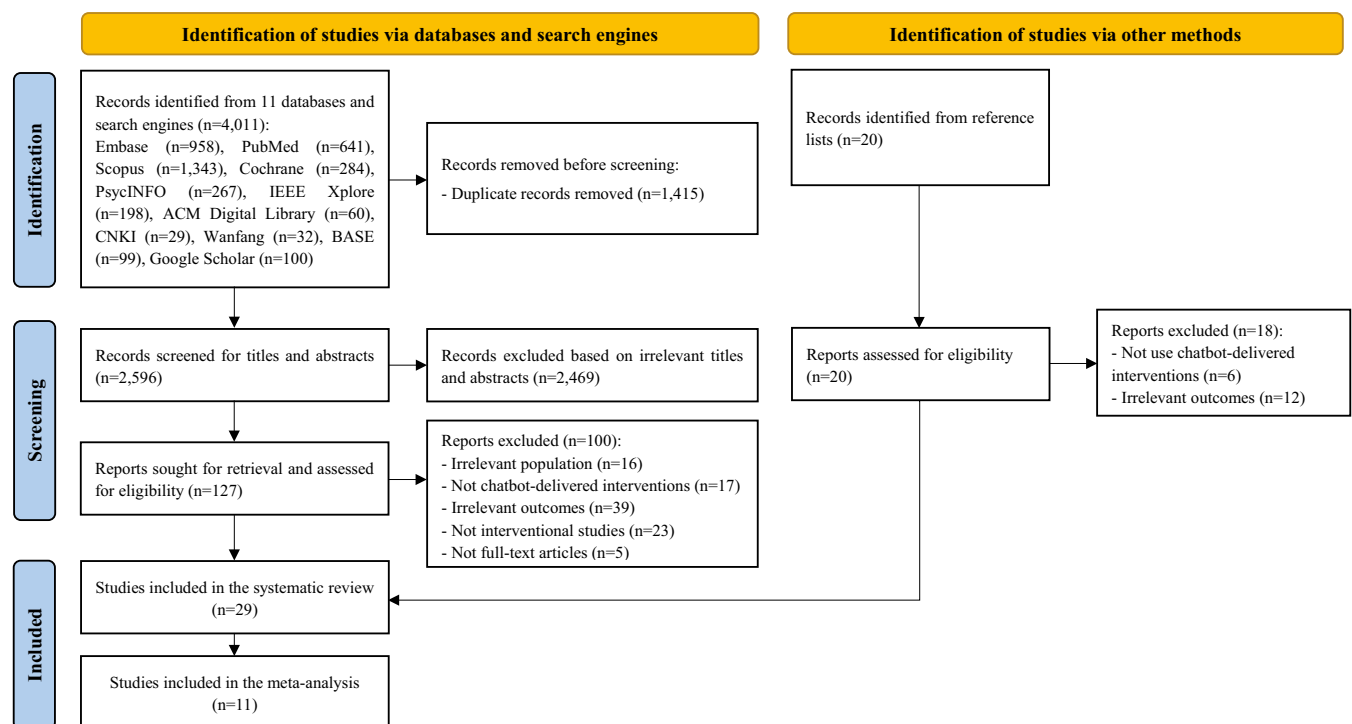


FIGURE 1 | PRISMA flow diagram of study selection.

TABLE 1 | Summary of characteristics of included studies ($N = 29$).

Characteristics	No. of studies (%)
Methodology	
Mixed methods	20 (69.0)
Quantitative	9 (31.0)
Study design	
Quasi-experimental	16 (55.2)
Randomized controlled trial	13 (44.8)
Type of publication	
Journal article	21 (72.4)
Conference proceeding	7 (24.1)
Thesis	1 (3.5)
Income group	
High	18 (62.1)
Upper-middle	8 (27.6)
Lower-middle	3 (10.3)
Setting	
Educational	19 (65.5)
Community	7 (24.1)
Hospital	3 (10.4)
Sample size	
< 100	18 (62.1)
100–200	6 (20.7)
> 200	5 (17.2)
Gender (male%, reported in 26 studies)	38.3
Sample type	
Nonclinical	20 (69.0)
Subclinical	5 (17.2)
Clinical	4 (13.8)
Control	
Psychoeducation	8 (27.6)
Therapist-delivered control	1 (3.5)
Usual care	3 (10.3)
Waitlist control	3 (10.3)
No control group	14 (48.3)
Measured outcomes (five most reported indicators)	
Severity of depression	14 (48.3)
Severity of anxiety	14 (48.3)
Positive and negative affect	8 (17.2)

(Continues)

TABLE 1 | (Continued)

Characteristics	No. of studies (%)
Stress	8 (17.2)
Psychological well-being	5 (17.2)

characteristics can be found in Table 1, with detailed information available in Table S3.

3.2.3 | Characteristics of Interventions

The characteristics of chatbots have been summarized in Table 2 (further details in Table S4). As for the interventional features, the predominant use of chatbots in the included studies ($n = 28$) was for the delivery of psychotherapy and/or education. Additionally, chatbots were utilized for specific purposes, such as counseling (Bray et al. 2020; Trappey et al. 2022) and self-management (Peuters et al. 2024). Among the 26 studies that incorporated theoretical frameworks, an integrative approach ($n = 12$) and cognitive behavioral therapy ($n = 7$) were the most frequently utilized therapeutic approaches (Table S4). Eleven studies explicitly mentioned employing a co-design approach, involving end users in the development of chatbot-delivered interventions. In terms of implementation, most of the studies ($n = 25$) deployed chatbots as independent, standalone systems, while the remaining four studies integrated chatbots as components within mental health applications or digital platforms. The duration of the studies varied significantly, ranging from a single session lasting 15min to interventions spanning over 4 months.

In terms of the technical features, chatbots are primarily implemented through instant messengers ($n = 14$) and mobile applications ($n = 10$), followed by web-based platforms ($n = 2$), robots ($n = 2$), and virtual reality ($n = 1$). In most studies ($n = 20$), AI-based chatbots were utilized to comprehend user responses and generate corresponding replies. In contrast, other studies employed rule-based approaches, where responses were determined by predefined rules or decision trees. Interactions between users and chatbots encompassed various modalities, including only written language via text ($n = 21$), only spoken language via voice ($n = 1$), a combination of written and spoken languages ($n = 4$), a combination of spoken and visual languages ($n = 3$), and one study did not specify the modality. In eight studies, chatbots were equipped with embodiments (e.g., robot, virtual human, and avatar). Despite the rising importance of addressing safety issues concerning chatbots in mental health, only 14 studies integrated safeguarding measures, encompassing emergency assistance ($n = 12$), crisis identification ($n = 6$), and professional accompaniment ($n = 2$).

3.3 | Risk of Bias of Included Studies and Quality of Evidence of Meta-Analysis Results

The risk of bias within the RCTs focusing on psychological distress and well-being outcomes produced similar findings

TABLE 2 | Summary of interventional and technical characteristics of chatbots ($N=29$).

Characteristics	No. of studies (%)
Chatbot purpose ^a	
Therapy and/or education	28 (96.6)
Counseling	2 (6.9)
Self-management	1 (3.4)
Theory basis	
Yes	26 (89.7)
Not reported	3 (10.3)
Use of co-design	
Yes	11 (37.9)
No/Not reported	18 (62.1)
Deployment	
Stand alone	25 (86.2)
Component	4 (13.8)
Duration (weeks)	
< 1	4 (13.8)
2–4	20 (69.0)
> 4	5 (17.2)
Platform	
Instant messenger	14 (48.3)
Mobile application	10 (34.5)
Web-based	2 (6.9)
Robot	2 (6.9)
Virtual reality	1 (3.4)
Response generation ^b	
AI-based	20 (69.0)
Rule-based	10 (34.5)
Interaction mode ^b	
Written only	21 (72.4)
Spoken only	1 (3.4)
Written and spoken	4 (13.8)
Spoken and visual	3 (10.3)
Not reported	1 (3.4)
Embodiment ^b	
Yes	8 (27.6)
No	22 (78.6)
Safety measures	
Yes	14 (48.3)
No	15 (51.7)

^aThe total number exceeds 29 because most chatbots served more than one purpose.

^bThe total number exceeds 29 because one study (Karhiy et al. 2023) used two types of chatbots.

(Figures S1 and S2). Notably, the most significant source of bias stemmed from outcome measurements, specifically the reliance on self-reported instruments without assessor-blinding. Regarding the selection of reported results, 21 outcomes across six studies raised “some concerns” due to insufficient information regarding prespecified plans and protocols. Additionally, 12 included outcomes in five studies exhibited biased aspects in deviations from the intended intervention, including inadequate analyses and procedural deviations. The domain of missing outcome data indicated a certain level of bias in nine outcomes, influenced by high attrition rates and inadequate bias correction methods. Apart from four outcomes in two studies that showed “some concerns” in the randomization process, the remaining outcomes were deemed “low” risk. Table S5 summarizes the quality appraisal of the included quasi-experimental studies against MMAT criteria. Two studies fulfilled all the criteria set (5/5), while five, four, and five studies achieved four, three, and two, respectively.

The certainty of evidence regarding psychological distress was assessed as “moderate” due to significant heterogeneity among the included outcomes. For psychological well-being, the evidence was downgraded by two levels to “low” because of the presence of bias and inconsistency (Table S6).

3.4 | Effectiveness of Chatbot-Delivered Interventions

The meta-analysis consisted of 11 studies, excluding two RCTs due to insufficient data reported (Fulmer et al. 2018; Romanovsky et al. 2021). In the case of three-arm trials, multiple groups were combined to make a single pairwise comparison (He et al. 2022; Karhiy et al. 2023).

In the meta-analysis, 11 trials assessed psychological distress outcomes (Figure 2), while seven trials evaluated psychological well-being (Figure 3). The Egger's regression test results indicated no statistically significant publication bias (Table S7). When comparing chatbot-delivered interventions to various control conditions, participants who engaged with chatbots experienced a statistically significant reduction in psychological distress, with an effect size of $g = -0.28$ (95% CI $[-0.46, -0.10]$). The robustness of this result was confirmed through “leave-one-out” sensitivity analyses, which yielded estimated effect sizes ranging from -0.23 to -0.31 (Table S8). Although participants engaging with chatbots demonstrated positive changes in psychological well-being, these benefits were not statistically significant ($g = 0.13$, 95% CI $[-0.16, 0.41]$). The sensitivity analyses further supported the robustness of these findings (Table S8).

Both psychological distress ($I^2 = 63.94\%$, $p < 0.01$) and psychological well-being ($I^2 = 66.90\%$, $p < 0.01$) syntheses exhibited substantial heterogeneity. Predefined subgroup analyses were conducted to investigate variations in heterogeneity and efficacy. Regarding psychological distress, a statistically significant reduction was observed in the clinical/subclinical population ($g = -0.34$, 95% CI $[-0.57, -0.10]$, $n = 5$), while the effect was not significant in the nonclinical population ($g = -0.19$, 95% CI $[-0.48, 0.10]$, $n = 6$). Platform subgroup

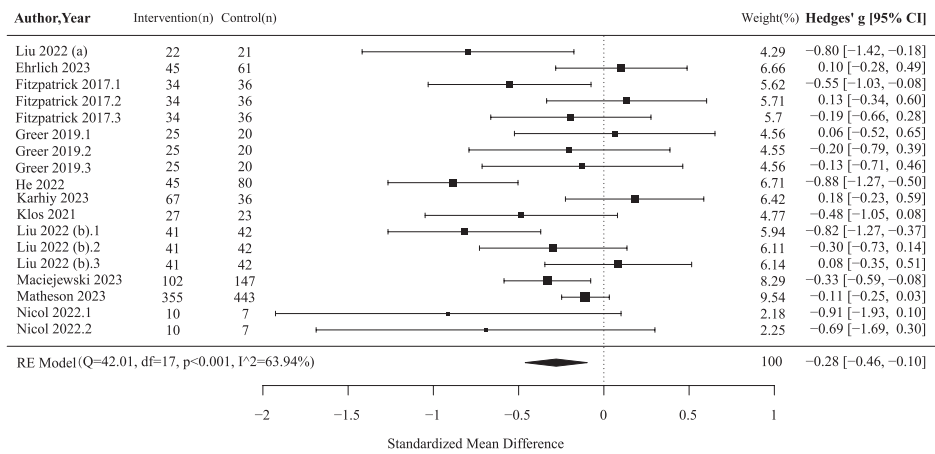


FIGURE 2 | Effects of chatbot-delivered interventions on psychological distress. *Negative effect sizes indicate a more favorable outcome for the intervention group.

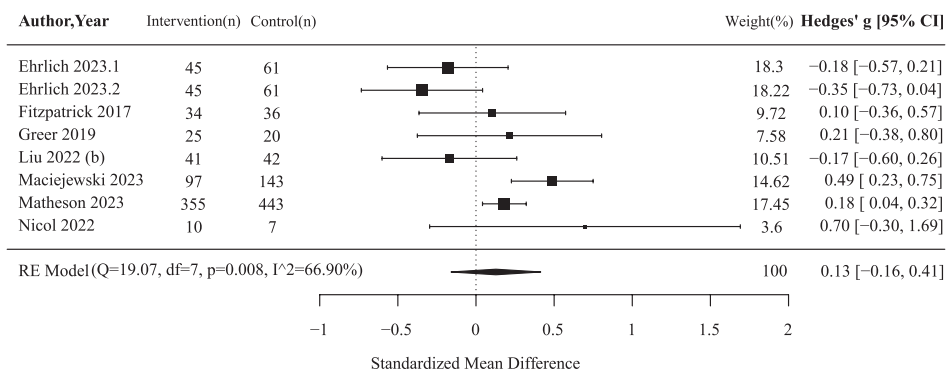


FIGURE 3 | Effects of chatbot-delivered interventions on psychological well-being. *Positive effect sizes indicate a more favorable outcome for the intervention group.

analysis revealed that chatbots delivered via instant messenger ($g = -0.29$, 95% CI $[-0.50, -0.08]$, $n = 7$) were beneficial in improving psychological distress. Furthermore, chatbots employing multimodal interaction mode ($g = -0.39$, 95% CI $[-0.67, -0.11]$, $n = 4$) demonstrated a more pronounced ameliorative impact on psychological distress compared to text-based chatbots ($g = -0.19$, 95% CI $[-0.41, 0.02]$, $n = 8$), while AI-based chatbots ($g = -0.36$, 95% CI $[-0.58, -0.14]$, $n = 8$) exhibited a stronger effect compared to rule-based chatbots ($g = -0.09$, 95% CI $[-0.39, 0.20]$, $n = 4$). In terms of psychological well-being, it was shown that only rule-based chatbots had a statistically significant impact on enhancing well-being ($g = 0.29$, 95% CI $[0.01, 0.57]$, $n = 3$). The results of subgroup analyses are presented in Tables S9 and S10.

3.5 | Narrative Synthesis of Feasibility and Acceptability

The recruitment rate reported in the studies ranged from 24% to 100% (Median = 46.9%; $n = 18$), while the attrition rate varied from 0% to 89.7% (Median = 23.4%; $n = 28$). Out of the 29 studies, 19 provided detailed measures regarding engagement, including metrics such as the frequency and duration of chatbot usage ($n = 13$) and the interactions/conversations ($n = 9$). The acceptability of chatbot-delivered interventions was

examined in 25 studies, employing various measurement tools such as scales (e.g., System Usability Scale), questionnaires (e.g., single-item Likert questionnaire), and interviews (e.g., focus group interviews). Despite variations in the selection of measures and reporting of quantitative outcomes, participants consistently expressed positive satisfaction and acceptance towards the chatbot-delivered interventions. Nineteen studies gathered qualitative feedback from end users regarding their interactions with chatbots. Participants reported positive aspects such as empathic and interactive communication, accessibility, practicality, non-judgment, personalization, alliance relationship, educational content, and pleasant chatbot design. However, some studies also identified challenges, including input misunderstanding, lack of targeted output content, technical issues, unnatural and impersonal interactions, and repetitiveness (details regarding feasibility and acceptability are presented in Table S11).

4 | Discussion

This is the first systematic review and meta-analysis that examines the effectiveness of chatbot-delivered interventions for mental health in young people. While the review covered a 10-year timeframe, most of the included studies (79.3%) were published between 2021 and 2024, indicating a recent surge

in research development within this field. Despite being a relatively new development, most studies (69.0%) employed mixed methods designs to comprehensively evaluate the effectiveness and acceptability of chatbot-delivered interventions.

4.1 | Effectiveness on Mental Health Outcomes and Key Influencing Characteristics

Our meta-analysis revealed a small-to-moderate effect of chatbot-delivered interventions in reducing psychological distress among young people ($g = -0.46$ to -0.10), echoing a previous meta-analysis conducted in the adult population ($g = 0.24$ – 0.47) (He et al. 2023). Notably, this effect size was lower than another review that reported in a subgroup focusing on adolescents and young adults ($g = 0.64$) (Li, Zhang, et al. 2023). This difference may be attributed to variations in population and intervention criteria, as Li, Zhang, et al. (2023) defined young people as 13–40 years old and focused solely on the effectiveness of AI-based chatbots.

However, this review did not find statistically significant improvements in psychological well-being outcomes, which aligns with another study (Li, Zhang, et al. 2023). One possible reason could be the smaller number of studies evaluating well-being ($n = 7$) compared to those assessing psychological distress ($n = 11$), leading to reduced statistical power. Furthermore, most interventions included in the well-being analyses were based on therapies aimed at symptom reduction (e.g., CBT) and lacked integration of positive psychology elements. According to the dual-continua model of mental health, well-being and distress are not the extreme ends of the same spectrum, indicating that interventions effectively alleviating distress may not necessarily enhance well-being (Iasiello and van Agteren 2020). Hence, to gain a comprehensive understanding of the effectiveness of chatbot-delivered interventions in mental health, future research could consider incorporating positive psychology techniques into interventions and simultaneously assessing their impact on psychological well-being.

Chatbot-delivered interventions were more effective in reducing psychological distress in clinical and subclinical populations ($g = -0.34$, $p < 0.01$) compared to nonclinical populations ($p > 0.05$). This supports prior research indicating that mental health interventions targeting young people with subclinical symptoms and higher-risk conditions yield better outcomes (Werner-Seidler et al. 2017). Moreover, interventions delivered through instant messenger platforms, such as Facebook and WeChat, yielded a greater effect in reducing psychological distress ($g = -0.29$, $p < 0.01$). Leveraging these readily accessible platforms can remove entry barriers and improve the availability of mental health support (Naslund et al. 2020), potentially leading to higher adherence and better outcomes.

Our subgroup analyses indicated that interacting with chatbots employing multiple modalities had greater effects ($g = -0.39$, $p < 0.01$) on reducing psychological distress than those using written language alone ($p > 0.05$). This can be attributed to the combination of various modes (e.g., speech, text, facial

expressions, or gestures), which may enhance interaction quality and user engagement (Provoost et al. 2017). The subgroup analyses yielded contradictory results for the response generation approach. AI-based chatbots significantly reduced psychological distress ($g = -0.36$, $p < 0.01$), whereas rule-based chatbots had a better effect on enhancing well-being ($g = 0.29$, $p < 0.05$). While AI-based chatbots excel in autonomous learning and handling complex conversations, providing personalized and human-like interactions, they also entail risks including misunderstandings, technical glitches, and privacy concerns (Li, Liang, et al. 2023). To create a more comprehensive and effective mental health support system, integrating the strengths of AI-based chatbots and rule-based chatbots appears promising.

While chatbots have made advancements in generating responses, the integration of embodiment has been limited, with only a quarter of reviewed chatbots incorporating this feature. Evidence suggests that human-like conversational behaviors in embodied chatbots foster rapport with users (Loveys et al. 2020), which is reinforced by one included comparative study demonstrating higher user engagement and satisfaction in embodied chatbots (Karhiy et al. 2023). However, excessive pursuit of human-like interactions can yield negative outcomes based on the uncanny valley theory (Rapp et al. 2021). Consequently, future research should strive for a balanced level of anthropomorphism. Moreover, the fully automated nature of chatbots in mental health care may raise safety concerns, such as limited capacity to monitor and address negative emotional reactions during interventions (Viduan et al. 2023). Considering that only half of the included studies have implemented safeguarding measures, future chatbots should prioritize the integration of robust safety protocols to ensure the well-being of users.

4.2 | Feasibility and Acceptability

In contrast to the attrition rate of 36% in technology-delivered interventions for children and adolescents with depression and anxiety (Grist et al. 2019), most reviewed studies ($n = 18$) reported attrition rates below 30%, suggesting the general feasibility of chatbot-delivered interventions in youth mental health. However, several of the included studies showed high attrition rates and early discontinuation, which could be attributed to the fully remote and self-help nature of the interventions (Matheson et al. 2023; Vertsberger et al. 2022). Integrating human support elements alongside chatbot virtual assistance may improve young people's adherence (Struthers et al. 2015). However, only three included studies explicitly mentioned integrating human assistance into the chatbot and showed lower attrition rates (Bray et al. 2020; Nicol et al. 2022; Russell et al. 2021). Therefore, further investigation is needed to explore the effects of integrating human support and determine the types and levels of support that can yield optimal results. Additionally, the included studies generally reported positive acceptability but also highlighted challenges that could reduce adherence and engagement. To enhance interaction fluency and quality, techniques such as deep learning algorithms can be used (Casheekar et al. 2024). Furthermore, involving end users and relevant stakeholders through a

co-design approach from the early development stages can enhance chatbots' ability to meet youth-specific needs, resulting in increased engagement and satisfaction (Gemert-Pijnen et al. 2018; Lu, et al. 2025).

4.3 | Limitations

In light of the emerging nature of research in this field, we conducted a comprehensive search, including gray literature, to mitigate publication bias and enhance the comprehensiveness of the evidence. However, there are several limitations to this review. First, most studies were conducted in nonclinical populations within educational settings and developed countries, potentially limiting the generalizability of the findings. Second, caution is warranted when interpreting the meta-analysis findings due to the risk of bias and heterogeneity (e.g., differences in participants' characteristics and control groups) among the included studies. Third, the limited number of studies prevented us from examining the long-term effects of chatbot-delivered interventions and identifying other potential factors for successful outcomes (e.g., the presence of embodiment and human support). Lastly, this review only included publications in English and Chinese, potentially overlooking studies published in other languages.

4.4 | Implications to Research and Practice

To further improve accessibility and convenience, future chatbot-delivered interventions could explore delivering interventions through instant messenger platforms. With advancements in technology, equipping multiple modalities of communication is expected to improve the quality of interactions with chatbots. Alongside improving language processing capabilities and accuracy, future chatbots must prioritize privacy enhancements and strengthen security measures. Involving end users and stakeholders, such as therapists and nurses, early in the development process ensures that the final product is customized to meet the needs of young people and effectively complements existing mental health services. This co-design approach can inform decisions such as the degree of anthropomorphism of the chatbot and the inclusion of human support in the intervention.

This review revealed a lack of studies conducted in low-income countries, highlighting the need to implement chatbot-delivered interventions in these mental health resource-limited regions. Additionally, future research should prioritize clinical and sub-clinical populations, considering their heightened mental health needs Bressington, (Li, Chan, et al. 2024; Li, Li, et al. 2024) as most interventions analyzed in this review focused on young people in nonclinical settings. Longer-term follow-up is needed to confirm the lasting effectiveness of chatbot-delivered interventions. Future trials should rigorously assess the impact of interventions on psychological well-being outcomes, considering the current scarcity of studies and the low certainty of existing evidence. To enhance the methodological quality of future research in this field, it is recommended to use active controls, pre-register research proposals, account for missing data, and conduct intention-to-treat analyses.

4.5 | Linking Evidence to Action

- Chatbot-delivered interventions show great potential for supporting the mental health of young people, complementing existing mental health services provided by multidisciplinary healthcare professionals.
- Instant messenger platforms are recommended for carrying out chatbot-delivered interventions.
- Multiple modalities of communication can be adopted to improve the interaction quality with chatbots.
- Alongside improving language processing capabilities and accuracy, future chatbots must prioritize privacy enhancements and strengthen security measures.
- Future research should emphasize the impact of chatbot-delivered interventions in low-income countries and among young people with physical or psychological symptoms, with extended follow-up periods.

5 | Conclusions

In conclusion, chatbot-delivered interventions for young people's mental health have shown rapid growth, with diverse interventional and technical features. Preliminary evidence from RCTs supports their effectiveness in reducing psychological distress and identifies influential factors. While feasibility and acceptability have been demonstrated, there is still room for improvement in engagement, safety, and interaction quality. Our review highlights the potential of chatbot-delivered interventions as an easily accessible and effective solution to support the mental health of young people and complement existing mental health services. Given the heterogeneity and risk of bias of the included studies, the current evidence could be strengthened with more well-designed trials that examine the effectiveness and mechanism of action of chatbot-delivered interventions for young people.

Acknowledgments

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Abd-Alrazaq, A. A., M. Alajlani, A. A. Alalwan, B. M. Bewick, P. Gardner, and M. Househ. 2019. "An Overview of the Features of Chatbots in Mental Health: A Scoping Review." *International Journal of Medical Informatics* 132: 103978. <https://doi.org/10.1016/j.ijmedinf.2019.103978>.
- Abd-Alrazaq, A. A., A. Rababeh, M. Alajlani, B. M. Bewick, and M. Househ. 2020. "Effectiveness and Safety of Using Chatbots to Improve Mental Health: Systematic Review and Meta-Analysis." *Journal of*

- Medical Internet Research 22, no. 7: e16021. <https://doi.org/10.2196/16021>.
- Aguirre Velasco, A., I. S. S. Cruz, J. Billings, M. Jimenez, and S. Rowe. 2020. "What Are the Barriers, Facilitators and Interventions Targeting Help-Seeking Behaviours for Common Mental Health Problems in Adolescents? A Systematic Review." *BMC Psychiatry* 20, no. 1: 293. <https://doi.org/10.1186/s12888-020-02659-0>.
- Assink, M., and C. Wibbelink. 2016. "Fitting Three-Level Meta-Analytic Models in R: A Step-By-Step Tutorial." *Quantitative Methods for Psychology* 12: 154–174. <https://doi.org/10.20982/tqmp.12.3.p154>.
- Balan, R., A. Dobrea, and C. R. Poetar. 2024. "Use of Automated Conversational Agents in Improving Young Population Mental Health: A Scoping Review." *npj Digital Medicine* 7, no. 1: 75. <https://doi.org/10.1038/s41746-024-01072-1>.
- Brandtzæg, P. B. B., M. Skjuve, K. K. K. Dysthe, and A. Følstad. 2021. "When the Social Becomes Non-Human: Young People's Perception of Social Support in Chatbots Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems." Yokohama, Japan. <https://doi.org/10.1145/3411764.3445318>.
- Bray, L., A. Sharpe, P. Gichuru, P.-M. Fortune, L. Blake, and V. Appleton. 2020. "The Acceptability and Impact of the Xploro Digital Therapeutic Platform to Inform and Prepare Children for Planned Procedures in a Hospital: Before and After Evaluation Study." *Journal of Medical Internet Research* 22, no. 8: e17367. <https://doi.org/10.2196/17367>.
- Casheekar, A., A. Lahiri, K. Rath, K. S. Prabhakar, and K. Srinivasan. 2024. "A Contemporary Review on Chatbots, AI-Powered Virtual Conversational Agents, ChatGPT: Applications, Open Challenges and Future Research Directions." *Computer Science Review* 52: 100632. <https://doi.org/10.1016/j.cosrev.2024.100632>.
- Consumers and Communication Group. 2012. "Outcomes of Interest to the Cochrane Consumers & Communication Review Group." <https://cccr.cochrane.org/sites/cccr.cochrane.org/files/uploads/Outcomes.pdf>.
- Deeks, J. J., J. P. T. Higgins, and D. G. Altman. 2023. "Chapter 10: Analysing Data and Undertaking Meta-Analyses." In *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (Updated August 2023)*. Cochrane. www.training.cochrane.org/handbook.
- Fulmer, R., A. Joerin, B. Gentile, L. Lakerink, and M. Rauws. 2018. "Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial." *JMIR Mental Health* 5, no. 4: e64. <https://doi.org/10.2196/mental.9782>.
- Gemert-Pijnen, J., S. Kelders, H. Kip, and R. Sanderman. 2018. *eHealth Research, Theory and Development: A Multi-Disciplinary Approach*. Routledge.
- Grist, R., A. Croker, M. Denne, and P. Stallard. 2019. "Technology Delivered Interventions for Depression and Anxiety in Children and Adolescents: A Systematic Review and Meta-Analysis." *Clinical Child and Family Psychology Review* 22, no. 2: 147–171. <https://doi.org/10.1007/s10567-018-0271-8>.
- Grudin, J., and R. Jacques. 2019. "Chatbots, Humbots, and the Quest for Artificial General Intelligence Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems." Glasgow, Scotland UK. <https://doi.org/10.1145/3290605.3300439>.
- Guyatt, G. H., A. D. Oxman, H. J. Schünemann, P. Tugwell, and A. Knottnerus. 2011. "GRADE Guidelines: A New Series of Articles in the Journal of Clinical Epidemiology." *Journal of Clinical Epidemiology* 64, no. 4: 380–382. <https://doi.org/10.1016/j.jclinepi.2010.09.011>.
- Hamadeh, N., C. Van Rompaey, and E. Metreau. 2023. "World Bank Group Country Classifications by Income Level for FY24 (July 1, 2023–June 30, 2024)." World Bank Blogs. <https://blogs.worldbank.org/en/pendata/new-world-bank-group-country-classifications-income-level-fy24>.
- He, Y., L. Yang, C. Qian, et al. 2023. "Conversational Agent Interventions for Mental Health Problems: Systematic Review and Meta-Analysis of Randomized Controlled Trials." *Journal of Medical Internet Research* 25: e43862. <https://doi.org/10.2196/43862>.
- He, Y., L. Yang, X. Zhu, et al. 2022. "Mental Health Chatbot for Young Adults With Depressive Symptoms During the COVID-19 Pandemic: Single-Blind, Three-Arm Randomized Controlled Trial." *Journal of Medical Internet Research* 24, no. 11: e40719. <https://doi.org/10.2196/40719>.
- Hedges, L. V., and I. Olkin. 2014. *Statistical Methods for Meta-Analysis*. Academic press.
- Higgins, J. P. T., S. Eldridge, and T. Li. 2023. "Chapter 23: Including Variants on Randomized Trials." In *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4 (Updated August 2023)*. Cochrane. www.training.cochrane.org/handbook.
- Hong, Q. N., A. Gonzalez-Reyes, and P. Pluye. 2018. "Improving the Usefulness of a Tool for Appraising the Quality of Qualitative, Quantitative and Mixed Methods Studies, the Mixed Methods Appraisal Tool (MMAT)." *Journal of Evaluation in Clinical Practice* 24, no. 3: 459–467. <https://doi.org/10.1111/jep.12884>.
- Iasiello, M., and J. van Agteren. 2020. "Mental Health and/or Mental Illness: A Scoping Review of the Evidence and Implications of the Dual-Continua Model of Mental Health." *Evidence Base* 1, no. 1: 1–45.
- Islam, M. I., F. M. Yunus, S. N. Isha, E. Kabir, R. Khanam, and A. Martiniuk. 2022. "The Gap Between Perceived Mental Health Needs and Actual Service Utilization in Australian Adolescents." *Scientific Reports* 12, no. 1: 5430. <https://doi.org/10.1038/s41598-022-09352-0>.
- Karhiy, M., M. Sagar, M. Antoni, K. Loveys, and E. Broadbent. 2023. "Mindfulness Based Stress Reduction: A Randomised Trial of a Virtual Human, Teletherapy, and a Chatbot." 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), Cambridge, MA, USA.
- Li, Y., D. Bressington, S. Wang, S. F. Leung, and Y. W. Mak. 2021. "Relationship Between Parental Psychological Control and Optimism Among Hong Kong Adolescents: The Mediating Role of Self-Mastery." *Current Psychology* 42, no. 12: 10115–10122. <https://doi.org/10.1007/s12144-021-02312-3>.
- Li, J., E. A. Chan, M. Li, et al. 2024. "Am I Different?" Coping and Mental Health Among Teenagers With Adolescent Idiopathic Scoliosis: A Qualitative Study." *Journal of Pediatric Nursing* 75: e135–e141. <https://doi.org/10.1016/j.pedn.2024.01.004>.
- Li, Y., T. Y. Chung, W. Lu, et al. 2024. "Chatbot-Based Mindfulness-Based Stress Reduction Program for University Students With Depressive Symptoms: Intervention Development and Pilot Evaluation." *Journal of the American Psychiatric Nurses Association*. <https://doi.org/10.1177/10783903241302092>.
- Li, J., Y. Li, X. Mei, et al. 2024. "Effectiveness of Cognitive-Based Interventions on Psychological Distress in Adolescents With Physical Disabilities: A Systematic Review and Meta-Analysis." *International Journal of Mental Health Nursing* 33, no. 6: 1729–1742. <https://doi.org/10.1111/inm.13391>.
- Li, Y., S. Liang, B. Zhu, et al. 2023. "Feasibility and Effectiveness of Artificial Intelligence-Driven Conversational Agents in Healthcare Interventions: A Systematic Review of Randomized Controlled Trials." *International Journal of Nursing Studies* 143: 104494. <https://doi.org/10.1016/j.ijnurstu.2023.104494>.
- Li, Y., K.-C. Lee, D. Bressington, et al. 2024. "A Theory and Evidence-Based Artificial Intelligence-Driven Motivational Digital Assistant to Decrease Vaccine Hesitancy: Intervention Development and Validation." *Vaccine* 12, no. 7: 708. <https://doi.org/10.3390/vaccines12070708>.
- Li, H., R. Zhang, Y. C. Lee, R. E. Kraut, and D. C. Mohr. 2023. "Systematic Review and Meta-Analysis of AI-Based Conversational Agents for Promoting Mental Health and Well-Being." *npj Digital Medicine* 6, no. 1: 236. <https://doi.org/10.1038/s41746-023-00979-5>.

- Lim, S. M., C. W. C. Shiao, L. J. Cheng, and Y. Lau. 2022. "Chatbot-Delivered Psychotherapy for Adults With Depressive and Anxiety Symptoms: A Systematic Review and Meta-Regression." *Behavior Therapy* 53, no. 2: 334–347. <https://doi.org/10.1016/j.beth.2021.09.007>.
- Lipson, S. K., E. G. Lattie, and D. Eisenberg. 2019. "Increased Rates of Mental Health Service Utilization by U.S. College Students: 10-Year Population-Level Trends (2007–2017)." *Psychiatric Services* 70, no. 1: 60–63. <https://doi.org/10.1176/appi.ps.201800332>.
- Liu, H. 2022. "Research on the Design of an Intelligent Chat Agent for University Students' Depression Psychotherapy." Master, South China University of Technology.
- Loveys, K., G. Sebaratnam, M. Sagar, and E. Broadbent. 2020. "The Effect of Design Features on Relationship Quality With Embodied Conversational Agents: A Systematic Review." *International Journal of Social Robotics* 12, no. 6: 1293–1312. <https://doi.org/10.1007/s12369-020-00680-7>.
- Lu, W., Y. Li, J. Montayre, et al. 2025. "A Bibliometric Analysis of Healthcare Intervention-Related Studies Reporting Patient and Public Involvement and Engagement." *Health* 13, no. 3: 305. <https://doi.org/10.3390/healthcare13030305>.
- Matheson, E. L., H. G. Smith, A. C. S. Amaral, et al. 2023. "Using Chatbot Technology to Improve Brazilian Adolescents' Body Image and Mental Health at Scale: Randomized Controlled Trial." *JMIR mHealth and uHealth* 11: e39934. <https://doi.org/10.2196/39934>.
- McHugh, M. L. 2012. "Interrater Reliability: The Kappa Statistic." *Biochemia Medica* 22, no. 3: 276–282.
- McTear, M., Z. Callejas, and D. Griol. 2016. *The Conversational Interface: Talking to Smart Devices*. Springer.
- Naslund, J. A., A. Bondre, J. Torous, and K. A. Aschbrenner. 2020. "Social Media and Mental Health: Benefits, Risks, and Opportunities for Research and Practice." *Journal of Technology in Behavioral Science* 5, no. 3: 245–257. <https://doi.org/10.1007/s41347-020-00134-x>.
- Nicol, G., R. Wang, S. Graham, S. Dodd, and J. Garbutt. 2022. "Chatbot-Delivered Cognitive Behavioral Therapy in Adolescents With Depression and Anxiety During the COVID-19 Pandemic: Feasibility and Acceptability Study." *JMIR Formative Research* 6, no. 11: e40242. <https://doi.org/10.2196/40242>.
- Page, M. J., J. E. McKenzie, P. M. Bossuyt, et al. 2021. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." *BMJ* 372: n71. <https://doi.org/10.1136/bmj.n71>.
- Peuters, C., L. Maenhout, G. Cardon, et al. 2024. "A Mobile Healthy Lifestyle Intervention to Promote Mental Health in Adolescence: A Mixed-Methods Evaluation." *BMC Public Health* 24, no. 1: 44. <https://doi.org/10.1186/s12889-023-17260-9>.
- Pretorius, C., and D. Coyle. 2021. "Young People's Use of Digital Tools to Support Their Mental Health During Covid-19 Restrictions." *Frontiers in Digital Health* 3: 763876. <https://doi.org/10.3389/fdgth.2021.763876>.
- Provoost, S., H. M. Lau, J. Ruwaard, and H. Riper. 2017. "Embodied Conversational Agents in Clinical Psychology: A Scoping Review." *Journal of Medical Internet Research* 19, no. 5: e151. <https://doi.org/10.2196/jmir.6553>.
- Rapp, A., L. Curti, and A. Boldi. 2021. "The Human Side of Human-Chatbot Interaction: A Systematic Literature Review of Ten Years of Research on Text-Based Chatbots." *International Journal of Human-Computer Studies* 151: 102630. <https://doi.org/10.1016/j.ijhcs.2021.102630>.
- Romanovskiy, O., N. Pidbutska, and A. Knysh. 2021. "Elomia Chatbot: The Effectiveness of Artificial Intelligence in the Fight for Mental Health. 5th International Conference on Computational Linguistics and Intelligent Systems, Ukraine."
- Russell, J. K., E. Strodl, and D. Kavanagh. 2021. "Use of a Social Robot in the Implementation of a Narrative Intervention for Young People With Cystic Fibrosis: A Feasibility Study." *International Journal of Social Robotics* 13, no. 7: 1787–1801. <https://doi.org/10.1007/s12369-021-00765-x>.
- Sartore, G. M., A. Pourliakas, and V. Lagioia. 2021. "Peer Support Interventions for Parents and Carers of Children With Complex Needs." *Cochrane Database of Systematic Reviews* 12, no. 12: Cd010618. <https://doi.org/10.1002/14651858.CD010618.pub2>.
- Solmi, M., J. Radua, M. Olivola, et al. 2022. "Age at Onset of Mental Disorders Worldwide: Large-Scale Meta-Analysis of 192 Epidemiological Studies." *Molecular Psychiatry* 27, no. 1: 281–295. <https://doi.org/10.1038/s41380-021-01161-7>.
- Sterne, J. A. C., J. Savović, M. J. Page, et al. 2019. "RoB 2: A Revised Tool for Assessing Risk of Bias in Randomised Trials." *BMJ* 366: 14898. <https://doi.org/10.1136/bmj.14898>.
- Struthers, A., C. Charette, S. B. Bapuji, et al. 2015. "The Acceptability of E-Mental Health Services for Children, Adolescents, and Young Adults: A Systematic Search and Review." *Canadian Journal of Community Mental Health* 34, no. 2: 1–21. <https://doi.org/10.7870/cjcmh-2015-006>.
- Trappey, A. J. C., A. P. C. Lin, K. Y. K. Hsu, C. V. Trappey, and K. L. K. Tu. 2022. "Development of an Empathy-Centric Counseling Chatbot System Capable of Sentimental Dialogue Analysis." *PRO* 10, no. 5: 930.
- United Nations Department of Economic and Social Affairs. 2016. "Definition of Youth." <https://www.un.org/esa/socdev/documents/youth/fact-sheets/youth-definition.pdf>.
- United Nations Population Fund. 2024. "World Population Dashboard." <https://www.unfpa.org/data/world-population-dashboard>.
- Van den Noortgate, W., J. A. López-López, F. Marín-Martínez, and J. Sánchez-Meca. 2015. "Meta-Analysis of Multiple Outcomes: A Multilevel Approach." *Behavior Research Methods* 47, no. 4: 1274–1294. <https://doi.org/10.3758/s13428-014-0527-2>.
- van Doorn, M., L. A. Nijhuis, M. D. Egeler, et al. 2021. "Online Indicated Preventive Mental Health Interventions for Youth: A Scoping Review." *Frontiers in Psychiatry* 12: 580843. <https://doi.org/10.3389/fpsy.2021.580843>.
- Vertsberger, D., N. Naor, and M. Winsberg. 2022. "Adolescents' Well-Being While Using a Mobile Artificial Intelligence–Powered Acceptance Commitment Therapy Tool: Evidence From a Longitudinal Study." *JMIR Ai* 1, no. 1: e38171. <https://doi.org/10.2196/38171>.
- Viduani, A., V. Cosenza, R. M. Araújo, and C. Kieling. 2023. "Chatbots in the Field of Mental Health: Challenges and Opportunities." In *Digital Mental Health: A Practitioner's Guide*, edited by I. C. Passos, F. D. Rabelo-da-Ponte, and F. Kapczinski, 133–148. Springer International Publishing. https://doi.org/10.1007/978-3-031-10698-9_8.
- Werner-Seidler, A., Y. Perry, A. L. Caele, J. M. Newby, and H. Christensen. 2017. "School-Based Depression and Anxiety Prevention Programs for Young People: A Systematic Review and Meta-Analysis." *Clinical Psychology Review* 51: 30–47. <https://doi.org/10.1016/j.cpr.2016.10.005>.
- Yan, L. I., A. Y. Wong, J. P. Cheung, et al. 2023. "Psychosocial Interventions for Teenagers With Adolescent Idiopathic Scoliosis: A Systematic Literature Review." *Journal of Pediatric Nursing* 73: e586–e593. <https://doi.org/10.1016/j.pedn.2023.10.037>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.