BMJ Health & Care Informatics

Effectiveness of chatbot-based interventions on mental well-being of the general population in Asia: protocol for a systematic review and meta-analysis of randomised controlled trials

Wilson Leung ,[1,2] Simon Ching Lam ,[1,2] Fowie Ng,[2,3] Calvin Chi Kong Yip,[2,4] Chi-Keung Chan[2,5]

To cite: Leung W, Lam SC, Ng F, *et al*. Effectiveness of chatbot-based interventions on mental well-being of the general population in Asia: protocol for a systematic review and meta-analysis of randomised controlled trials. *BMJ Health Care Inform* 2024;**31**:e101148. doi:10.1136/bmjhci-2024-101148

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjhci-2024-101148).

Received 23 May 2024
Accepted 22 November 2024

For numbered affiliations see end of article.

Correspondence to
Professor Simon Ching Lam; simonlam@twc.edu.hk; simlc@alumni.cuhk.net

ABSTRACT

Introduction In Asian countries, stigma against psychiatric disorders and shortage of manpower are the two major challenges that hinder people from receiving treatments. Chatbots can surely help people surpass the stigmatising and manpower shortage challenges. Since a comprehensive review in the Asian context is lacking, this systematic review will evaluate the effects of chatbot interventions on the mental well-being of the general population in Asia.

Methods and analysis Four electronic databases (PubMed, CINAHL, PsycINFO and MEDLINE) will be searched until December 2024. Randomised controlled trials with English/Chinese full text available will be included. Random-effect models will be used for meta-analyses. The risk of bias (RoB) and certainty of evidence across studies will be assessed using the Cochrane RoB2 and Grading of Recommendation Assessment, Development and Evaluation tools, respectively.

Ethics and dissemination This study will not require ethical approval. The findings will be disseminated through peer-reviewed publications.

Funding School Research Grant of the Tung Wah College (2023-04-52-SRG230401)

PROSPERO registration number CRD42024546316.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ It was estimated that 1 in every 8 people worldwide, equating to 970 million people, were living with a mental disorder.
⇒ Conventional mental health services usually have limited accessibility (eg, limited opening hours) and manpower shortage (eg, 0.1–90 psychiatrists per 1 000 000 people).
⇒ In line with the WHO Special Initiative for Mental Health (2019–2023) aimed at increasing access to quality and affordable care for mental health conditions, digital health due to its ubiquity and wide applicability can be considered an adjuvant treatment option in favour of access to quality mental health services.
⇒ Chatbot-based interventions have broad applicability (eg, health education, panic management and physical activity promotion) across multiple settings and populations.
⇒ Chatbot-related mental services are relatively flexible, readily accessible and 24/7 available compared with traditional mental health services.

INTRODUCTION

A chatbot, also known as a virtual conversational agent, is a computational system that simulates real-life human conversations.[1] Rule-based chatbots adopt a set of predefined rules to generate responses whereas artificial intelligence (AI)-powered chatbots employ more advanced algorithms (eg, machine learning) to allow deeper, adaptable communications by comprehending, generating and manipulating human language as it is written or spoken. During the conversations, people can interact with the agent using spoken and written language and the agent can respond verbally and non-verbally via spoken, written and visual (eg, body movements and facial expressions) language.[2] Recent systematic reviews showed that chatbots were a feasible, effective and safe adjunct modality for improving mental health (eg, depression, anxiety, distress, stress, acrophobia, subjective psychological well-being and affect).[2–4] In particular, chatbot interventions were clinically beneficial for alleviating depression (95% CI: −0.87 to −0.23).[2]

Since only a paucity of studies (n=2 to 4) were included for meta-analysis[2] and many reviews even did not quantitatively evaluate the effects of chatbot-based interventions on mental health outcomes,[4] a comprehensive review involving more updated evidence is required to uncover the potential efficacy of

WHAT THIS STUDY ADDS

⇒ This systematic review will provide evidence about the effects of varying chatbot interventions on mental health outcomes among the general population in Asia.
⇒ This study will examine the strength of the causal association between chatbot interventions and mental health outcomes among the general population in Asia.
⇒ This study will provide evidence about the feasibility of chatbot interventions in mental health services in respect of recruitment rates, completion rates, safety and adverse effects.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Apart from conventional mental health services, mental health chatbots can provide patients with more therapeutic options for mental illnesses.
⇒ Given that stigma against psychiatric illnesses (eg, dangerous, aggressive and weak) is very common and serious in Asian societies, advocating the use of chatbots in mental health services can help Asian people surpass the stigmatising challenges because they are more receptive and open to talk with a humanised machine than a real person.
⇒ An improved understanding of the chatbots' effectiveness within the context of mental health could surely better shed light on their adjunctive roles to already available interventions in healthcare and more importantly, enhance the comprehensiveness of services by offering more treatment options that are highly acceptable and easily accessible to people having mental health service needs.

chatbots on individuals who may have multifaceted needs (eg, emotional support) that are unresolved by conventional therapies. Moreover, previous reviews fell short in addressing the effects of different cultural backgrounds across countries on people's chatbot adoption behaviour given that culture is a strong predictor for the way people choose to cope with everyday problems and adversities.[5 6] Importantly, stigma against psychiatric illnesses (eg, dangerous, aggressive and weak) is very common and more serious in Asian societies.[7] For instance, patients' family members can be blamed for poor family lineage and hence their marriage and economic prospects will be adversely affected. Advocating the use of chatbots in mental health services not only provides patients with more therapeutic options but also helps them surpass stigmatising challenges because they are more receptive to share sensitive information with a humanised machine than a real person.[4]

This systematic review aims at systematically examining the effects of chatbot interventions on the mental well-being of the general population in Asia. Evidence about the feasibility of chatbot interventions in respect of recruitment, completion and safety will also be synthesised.

METHODS AND ANALYSIS

Reporting of the review findings will be in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 updated guidance and exemplars for reporting systematic reviews.[8] The protocol was registered in the PROSPERO registry. One change has been made to the project title from 'artificial intelligence-driven conversational agents' to 'chatbot-based interventions' which generally consider all kinds of chatbots to ensure a more comprehensive analysis of their mental health impacts by comparing different chatbot-based interventions. However, all other processes including the eligibility criteria for study inclusion, data extraction and synthesis and reporting bias assessment remain unchanged.

Eligibility criteria

The Population, Intervention, Comparison, Outcomes, Study Design framework will be employed to guide the inclusion of eligible studies.[8] The eligibility criteria are shown as follows:

- This systematic review will only include randomised controlled trials (RCTs) among Asian people of any age whom both healthy people and people at risk of mental health conditions (eg, socially disadvantaged groups including lesbian, gay, bisexual, transgender or people having chronic conditions/substance misuse) will be considered. If the study populations consist of both Asian and non-Asian people, only data derived from the Asian participants will be extracted for data analyses. Parallel, cross-over and cluster RCTs will be considered. For cross-over RCTs, only data before the cross-over period will be used for data analyses. Only RCTs with English/Chinese full text available will be included. Studies such as conference abstracts/proceedings and case studies on one or a few subjects or other than original research including grey literature, reviews, meta-analyses, study protocols, editorials and commentaries or retracted articles will be excluded.
- Standalone chatbot-related interventions using either mobile application-based or online web-based platforms have to be in at least one of the interventional arms within the trials. Studies involving chatbot interventions combined with other types of mental health services in one single experimental arm will be excluded because the therapeutic effects of the chatbot intervention may be caused by other interventions. Meanwhile, chatbots controlled by human operators will not be considered.
- The effects of the chatbot-based interventions will be compared with either no interventions including 'usual care', 'no treatment' and 'wait-list control' or compared with other interventions.
- The RCTs aimed at examining the effects of AI-driven chatbots on improving mental health outcomes including depression, anxiety, stress, affect, distress, sleep and subjective psychological well-being[9 10].

Information sources

Studies published in the databases of PubMed, CINAHL, PsycINFO and MEDLINE will be searched from their inception through December 2024. The search queries

for every electronic database are summarised in online supplemental information. The studies that were conducted in Asian countries/regions will be manually identified.

Search strategy

The initial screening (titles and abstracts) and full-text assessment of the searched studies will be conducted by the first author (WL) and subsequently verified by an independent reviewer (SCL). Since subject recruitment can be conducted through internet-based methods (eg, social media) which can target populations of different countries,[11] identification of whether the studies were conducted in Asian populations will be carried out through the full-text assessment (except relevant information explicitly specified in the abstracts or titles). To avoid missing any potentially eligible studies, the reference lists of the searched relevant systematic reviews[2–4 12–20] and all the included articles will also be screened for additional potential studies. Disagreements between the two reviewers (WL and SCL) will be resolved by discussion involving all review authors. The proposed PRISMA 2020 flow diagram for the study selection process is presented in online supplemental figure S1.

Data extraction

Information including publication details (authors, publication year and countries/regions), participant characteristics, intervention and control groups, training volume, recruitment and completion rates, safety and adverse effects and main findings will be extracted and summarised in an evidence table by the first author. The recruitment rate will be defined as the percentage of participants giving consent after being assessed as eligible while the completion rate will be defined as the percentage of participants remaining in the study until the last assessment visits.[4] To guarantee the robustness of evidence, all authors will verify and edit every entry. The template data collection form is shown in online supplemental table S1.

Risk of bias assessment

The Cochrane Risk of Bias 2 (RoB 2) will be employed to examine the RoB of the included RCTs.[21] The tool is structured into five domains covering all possible biases that may potentially affect the results of the included trials, namely bias arising from the randomisation process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in the measurement of the outcome and bias in the selection of the reported result. An overall RoB judgement for every study (low risk, some concerns, high risk) will be drawn based on the results of each bias domain. The RoB assessments will be first conducted by the first author and independently verified by another review author. Disagreements between the two reviewers will be resolved through discussion.

Certainty of evidence assessment

The Grading of Recommendation Assessment, Development and Evaluation approach will be employed to examine the certainty of evidence (CoE) across included studies for clinical endpoints (eg, depression and anxiety) that were most reported. Evidence about the clinical endpoints will be downgraded one place if (1) the RoB is high (ie, most included trials have some concerns or high RoB), (2) there is evidence of imprecision (ie, 95% CI fell within the minimal clinically important difference (MCID) or the effect estimate is derived from only one or two small studies), (3) there is evidence of unexplained inconsistency or high heterogeneity (ie, I^2 >50%), (4) the directness of evidence is questionable in respect of the participants studied, intervention of interest, measurements of the outcome (eg, surrogate vs intended markers) and the method of the trial of a candidate intervention (ie, depending on whether direct comparisons are available or not), (5) publication bias is detected if asymmetry of the funnel plot is observed. Overall, the CoE for the clinical endpoints will be classified as very low, low, moderate and high. The first author will assess the CoE which will hence be verified by an independent reviewer. Disagreements between the two reviewers will be resolved through discussion.

Data synthesis

Meta-analysis will be conducted for mental health outcomes (eg, depression and anxiety)[9] that were reported in at least two included studies.[2 4] For controlled trials, pairwise meta-analyses of postintervention data will be performed using random-effects models.[22] The overall effect will be examined using either standardised mean difference or weighted mean difference depending on whether the outcomes were assessed using different scales or the same scale, respectively. Subgroup analyses will also be conducted for potential moderators (eg, age, sex, ethnicity, medical history, types of chatbots and comparison controls). Boundaries of an MCID for the outcomes will be set as half of a SD from the mean value in the control arm(s) of the included studies at baseline.[2]

The degree of heterogeneity across study findings will be assessed using the Higgins I^2 statistics. Percentages of the I^2 statistics in 0–25%, 25–50% and>50% represent low, medium and high heterogeneity, respectively. For the meta-analyses that will be identified as having high heterogeneity, sensitivity analyses will be carried out by excluding studies with a high RoB. If there is missing information, we will contact the study authors directly via email. All meta-analyses in the present proposed study will be performed using Review Manager (RevMan V.5.4) software.

Assessment of publication bias

Funnel plots will be constructed to visually examine the bias if there are 10 or more trials included in the meta-analyses or subgroup analyses.[8] Asymmetry of the funnel plots indicates possible publication bias.

ETHICAL AND DISSEMINATION

No formal ethical approval will be required because the present review will only consider published articles and all data available among included studies should be anonymous with no concerns about participant privacy or confidentiality. The findings of this systematic review will be disseminated through peer-reviewed publications.

Author affiliations
[1]School of Nursing, Tung Wah College, Kowloon, Hong Kong
[2]Translational Research Centre for Digital Mental Health, Tung Wah College, Hong Kong SAR, Hong Kong
[3]School of Management, Tung Wah College, Kowloon, Hong Kong
[4]School of Medical and Health Sciences, Tung Wah College, Kowloon, Hong Kong
[5]School of Arts and Humanities, Tung Wah College, Kowloon, Hong Kong

Contributors WL, SCL, FN, CCKY and ACKC conceived and designed the study. This protocol was drafted by WL and then edited by SCL, FN, CCKY and ACKC. WL designed the search strategies while WL and SCL will conduct the search, data extraction, risk of bias assessment independently. WL and SCL will analyse and interpret the data. FN, CCKY and ACKC will resolve any disagreements during the review. All authors have approved the final version of this study protocol. WL acted as guarantor.

Funding This work is being supported by the School Research Grant (2023-04-52-SRG230401), School of Nursing, Tung Wah College, Hong Kong SAR, China.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Wilson Leung http://orcid.org/0000-0002-6196-4984
Simon Ching Lam http://orcid.org/0000-0002-2982-9192

REFERENCES

1 Nadarzynski T, Miles O, Cowie A, *et al*. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Dig Health* 2019;5:2055207619871808.
2 Abd-Alrazaq AA, Rababeh A, Alajlani M, *et al*. Effectiveness and Safety of Using Chatbots to Improve Mental Health: Systematic Review and Meta-Analysis. *J Med Internet Res* 2020;22:e16021.
3 Chou Y-H, Lin C, Lee S-H, *et al*. Potential Mobile Health Applications for Improving the Mental Health of the Elderly: A Systematic Review. *Clin Interv Aging* 2023;18:1523–34.
4 Li Y, Liang S, Zhu B, *et al*. Feasibility and effectiveness of artificial intelligence-driven conversational agents in healthcare interventions: A systematic review of randomized controlled trials. *Int J Nurs Stud* 2023;143:104494.
5 Manson SM. Extending the boundaries, bridging the gaps: crafting Mental Health: Culture, Race, and Ethnicity, a Supplement to the Surgeon General's Report on Mental Health. *Cult Med Psychiatry* 2003;27:395–408.
6 Chang DF. An introduction to the politics of science: culture, race, ethnicity, and the Supplement to the Surgeon General's Report on Mental Health. *Cult Med Psychiatry* 2003;27:373–83.
7 Zhang Z, Sun K, Jatchavala C, *et al*. Overview of Stigma against Psychiatric Illnesses and Advancements of Anti-Stigma Activities in Six Asian Societies. *Int J Environ Res Public Health* 2019;17:280.
8 Page MJ, Moher D, Bossuyt PM, *et al*. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;372:n160:160:.
9 Cheung T, Cheng CPW, Fong TKH, *et al*. Psychological impact on healthcare workers, general population and affected individuals of SARS and COVID-19: A systematic review and meta-analysis. *Front Public Health* 2022;10:1004558:1004558:.
10 Cheung DST, Yu BY-M, Lam SC, *et al*. A longitudinal study on the change in sleep across three waves of the COVID-19 outbreaks in Hong Kong. *Sleep Biol Rhythms* 2024;22:93–102.
11 Ollier J, Suryapalli P, Fleisch E, *et al*. Can digital health researchers make a difference during the pandemic? Results of the single-arm, chatbot-led Elena+: Care for COVID-19 interventional study. *Front Public Health* 2023;11:1185702.
12 Bérubé C, Schachner T, Keller R, *et al*. Voice-Based Conversational Agents for the Prevention and Management of Chronic and Mental Health Conditions: Systematic Literature Review. *J Med Internet Res* 2021;23:e25933.
13 Gaffney H, Mansell W, Tai S. Conversational Agents in the Treatment of Mental Health Problems: Mixed-Method Systematic Review. *JMIR Ment Health* 2019;6:e14166.
14 He Y, Yang L, Qian C, *et al*. Conversational Agent Interventions for Mental Health Problems: Systematic Review and Meta-analysis of Randomized Controlled Trials. *J Med Internet Res* 2023;25:e43862.
15 Hoermann S, McCabe KL, Milne DN, *et al*. Application of Synchronous Text-Based Dialogue Systems in Mental Health Interventions: Systematic Review. *J Med Internet Res* 2017;19:e267.
16 Li H, Zhang R, Lee Y-C, *et al*. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *NPJ Digit Med* 2023;6:236.
17 Lim SM, Shiau CWC, Cheng LJ, *et al*. Chatbot-Delivered Psychotherapy for Adults With Depressive and Anxiety Symptoms: A Systematic Review and Meta-Regression. *Behav Ther* 2022;53:334–47.
18 Ogilvie L, Prescott J, Carson J. The Use of Chatbots as Supportive Agents for People Seeking Help with Substance Use Disorder: A Systematic Review. *Eur Addict Res* 2022;28:405–18.
19 Sin J, Spain D, Furuta M, *et al*. Psychological interventions for post-traumatic stress disorder (PTSD) in people with severe mental illness. *Cochrane Database Syst Rev* 2017;1:CD011464.
20 Vaidyam AN, Wisniewski H, Halamka JD, *et al*. Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *Can J Psychiatry* 2019;64:456–64.
21 CochraneMethods. Risk of bias 2 (rob 2) tool 2024. n.d. Available: https://methods.cochrane.org/risk-bias-2
22 Borenstein M, Hedges LV, Higgins JPT, *et al*. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* 2010;1:97–111.