

Should AI Companions Be Included in Treatment Plans for Mental Health Issues With Associated Loneliness?

Team B: Kobby Hanson WooJung Kim Shilpi Pasricha Ender Ricart

Last Updated: Dec 7, 2024

Presentation: [12/7] Should AI Companions Be Included in Treatment Plans for Mental He...

Overview

Problem Statement

Loneliness is a widespread mental health concern, worsened by COVID-19-related social isolation¹. It often co-occurs with conditions like depression, anxiety, and PTSD and is linked to a 26% higher risk of premature death. Despite its significant impact, patients may not recognize loneliness as a factor, presenting instead with symptoms like low mood or social withdrawal. Mental health professionals play a vital role in identifying and addressing loneliness alongside other mental health issues.

Traditional approaches, including Cognitive Behavioral Therapy (CBT), mindfulness, and fostering community connections, are effective but often hindered by resource limitations and challenges in patient engagement². AI companions like Woebot³ and Replika⁴ offer a scalable solution, simulating therapeutic techniques such as guided mindfulness or role-playing social scenarios. The emergence of conversational GenAI tools like ChatGPT and Gemini holds additional promise when fine-tuned for mental health applications and integrated with professional care.

However, the effectiveness and ethical integration of GenAI into therapy remain underexplored⁵. This study seeks to evaluate whether AI companions can improve outcomes for mental health patients experiencing loneliness while addressing risks such as emotional overdependence and delusional attachments.

¹ N. Morrish, S. Choudhury, and A. Medina-Lara, "What Works in Interventions Targeting Loneliness: A Systematic Review of Intervention Characteristics," *BMC Public Health* 23 (2023): 2214, <https://doi.org/10.1186/s12889-023-17097-2>.

² Morrish, Choudhury, and Medina-Lara, "What Works in Interventions Targeting Loneliness," Sec. 3, <https://doi.org/10.1186/s12889-023-17097-2>.

³ Woebot is explicitly a "mental health chatbot" with approximately 1.5 million users as of 2023. With a smaller user base there isn't much objective research available yet.

⁴ Classified as an Intelligent Social Agent, Replika has over 25 million users. Replika asks a series of informational questions of its new users upon downloading the App and then personalizes its generative AI model to the individual.

⁵ J. Linardon et al., "The Efficacy of App-Supported Smartphone Interventions for Mental Health Problems: A Meta-Analysis of Randomized Controlled Trials," *World Psychiatry* 18 (2019): 325–336.

Intended Audience

Our intended audience is two-fold: 1) mental-health professionals specializing in mental health concerns with co-occurring feelings of loneliness, and 2) tech-companies considering the role GenAI could play in improving social and psychological wellbeing.

Existing Literature

Barriers to Receiving Mental-Health Care

Effective treatments like medication, psychotherapy, and CBT can reduce loneliness and mental illness symptoms, but significant barriers prevent access for many. Research indicates that two-thirds of individuals with mental health disorders never receive outpatient care due to stigma, discrimination, and a lack of affordable options⁶. Since the COVID-19 pandemic, demand for mental health professionals has surged, often resulting in long waitlists or reliance on costly private practices, making care inaccessible for many. Conversational AI and AI companions offer a promising interim solution, providing immediate support to those waiting for licensed practitioners and complementing care for those already in treatment.

Risks of Digitally Mediated Social Interactions

While tools like conversational AI hold promise, concerns persist about overdependence and "problematic internet usage" (PIU)⁷, where excessive reliance on digital tools may worsen loneliness and limit real-world connections. Studies like those by Moretta and Buodo (2020) highlight these risks⁸. However, with thoughtful development and intentional application, such technologies can mitigate loneliness and support mental health effectively, balancing potential drawbacks with their benefits.

The Potential of Conversational GenAI

Recent research demonstrates the promise of generative AI in mental health. For instance, De Freitas et al. (2023) found AI companions effectively reduce loneliness, offering support comparable to human interactions. These tools can help address barriers to care and improve mental health outcomes. While oversight from licensed practitioners remains essential, conversational GenAI provides valuable support for

⁶ Bethanie Maples, Merve Cerit, Aditya Vishwanath, and Roy Pea, "Loneliness and Suicide Mitigation for Students Using GPT-3-Enabled Chatbots," *NPJ Mental Health Research* 3 (January 22, 2024): 4, <https://doi.org/10.1038/s44184-023-00047-6>

⁷ Richard A. Davis, "A Cognitive-Behavioral Model of Pathological Internet Use," *Computers in Human Behavior* 17 (2001): 187–195, [https://doi.org/10.1016/S0747-5632\(00\)00041-8](https://doi.org/10.1016/S0747-5632(00)00041-8).

⁸ Teresa Moretta, Gabriele Buodo, Zsolt Demetrovics, and Marc N. Potenza, "Tracing 20 Years of Research on Problematic Use of the Internet and Social Media: Theoretical Models, Assessment Tools, and an Agenda for Future Work," *Comprehensive Psychiatry* 112 (2022): 152286, <https://doi.org/10.1016/j.comppsy.2022.152286>.

those waiting for or supplementing professional care. This study aims to explore the effectiveness of conversational GenAI in reducing loneliness, improving mental health, and identifying best practices for its application⁹.

Anticipated Impact

- Evaluate the effectiveness of conversational AI in reducing loneliness and enhancing mental health alongside professional care.
- Develop guidelines for using conversational GenAI in mental health support and loneliness interventions.
- Foster collaboration between GenAI developers and mental health professionals in digital care.
- Update the mHONcode to address the rise of generative AI in mental health applications.

Research Questions

Main Question

Can integrating AI companions into professional mental health care improve therapeutic outcomes for patients experiencing loneliness and mental illness?

Sub-Questions:

- How do loneliness and mental health outcomes *differ* between patients receiving traditional therapy and those using AI companions (beyond improvement)?
- Which AI companion features (e.g., empathy, goal-setting, availability) most effectively reduce loneliness?
- How might therapeutic approaches for loneliness and mental illness evolve with the addition of AI companions?
- What risks (e.g., emotional overdependence, problematic internet use) arise from integrating AI companions, and how can they be mitigated?

Definitions

Term	Definition
Loneliness	A subjective feeling of social or emotional isolation.

⁹ See the [Selection of AI Companion Criteria](#) section below about inclusion criteria and the mHONcode (Medical Health On the Net Code).

mHONcode	Ethical standards adapted from the Health On the Net (HON) Code, including Authority, Complementarity, Confidentiality, and Validity for mental health tools.
Problematic Internet Use (PIU)	Excessive or poorly controlled online behavior that disrupts daily life, potentially worsening loneliness or depression.
Digital Literacy	The ability to use digital tools effectively, influencing how users interact with AI companions in therapeutic contexts.
Mental-Health Professionals	Licensed professionals including LPCs, Psychiatrists, Psychologists, LCSWs, PMHNPs, and Licensed School Counselors.
Conversational GenAI	AI companions that generate human-like, contextually relevant responses, tailored for emotional support and mental health applications.
UCLA Loneliness Scale	A 20-item questionnaire measuring loneliness and isolation, with scores ranging from 20 to 80. Higher scores indicate greater loneliness.

Study Design

This mixed-methods study will evaluate whether AI companions, integrated into professional-guided interventions, improve loneliness outcomes in outpatient care compared to traditional therapy alone. The design includes randomized assignment to assess causal inference, combining qualitative and quantitative data from patients and professionals to address the study's main and sub-questions.

Recruitment Criteria

We will recruit 20 licensed mental-health professionals from mental-health networking platforms - Lyra, BetterHealth, and Talkspace. Each professional will invite 24-26 patients with moderate to severe loneliness. Half of the patients will be randomly assigned to the control or treatment groups to minimize bias. Eligible participants must score 35-80 on the UCLA Loneliness Scale and have access to a smartphone or computer for AI interactions. Exclusion criteria include severe psychiatric conditions or language barriers preventing engagement.

- **Control Group:** Traditional therapy alone
- **Treatment Group:** Traditional therapy + AI companion support

Randomization

Participants will be randomly assigned to the groups using a random number generator, ensuring balanced therapist factors across groups. All therapists will receive standardized training for consistent therapy delivery and effective AI companion integration, including guidance on monitoring engagement and managing risks like emotional overdependence.

Intervention

The intervention will control for individual differences by considering baseline UCLA Loneliness Scale scores, demographics (age, gender, socioeconomic status), and comfort with digital tools. Pre-study surveys will assess attitudes toward AI and digital interventions. For the treatment group, data from app/web usage will track interaction patterns and engagement with the AI companion, supplemented by monthly self-reports on perceived emotional support. Conversation transcripts will be analyzed for tone and empathy, while ensuring participant privacy. This multi-faceted approach will provide a comprehensive evaluation of the intervention's impact on emotional well-being and loneliness reduction.

Selection Criteria for AI Companions

We will evaluate AI companions based on ethical standards from the Mental Health on the Net (mHON) Code, including Authority, Complementarity, Confidentiality, Validity, and new principles like Usability and Intervention¹⁰. While not yet adopted by conversational AI tools, mHON certification could apply. Candidates must demonstrate advanced natural language processing (NLP) capabilities, such as context awareness and tone detection, and meet criteria including over one million downloads or daily active users, a proven operational track record of at least one year, and accessibility for diverse digital literacy levels. Finalists will be reviewed by participating mental health professionals, and the top-rated companion will be selected for the study. Using a single AI companion minimizes variability, simplifying the research and ensuring accurate assessment of its integration into therapist-guided interventions.

¹⁰ Ahmed E. Siddiqi, Alla Sikorskii, Charles W. Given, and Barbara Given, "Early Participant Attrition from Clinical Trials: Role of Trial Design and Logistics," *Clinical Trials* 5, no. 4 (2008): 328–35, <https://doi.org/10.1177/1740774508094406>.

Data

Quantitative Data:

- (Treatment & Control) Weekly UCLA Loneliness Scale to track pre- and post-intervention changes over time.
- (Treatment) Monthly collection of data on frequency, duration, and types of AI companion interactions (e.g., goal setting, emotional support).
- (Treatment & Control) Monthly surveys for care professionals to assess patient engagement, progress, and adherence.
- (Treatment & Control) Monthly self-reported outcomes on emotional support and therapy satisfaction.

Qualitative Data:

- (Treatment & Control) Bi-monthly open-ended survey feedback from patients on their experience.
- (Treatment & Control) Structured interviews with therapists before and after the study.
- (Treatment) Content analysis of AI companion conversations to identify patient-led topics, emotional expression, and interaction patterns, alongside therapeutic techniques and companion behaviors that enhance engagement.

Data Privacy and Ethics

All patient and therapist data will be anonymized, encrypted, and stored on secure servers with access limited to authorized researchers only. Participants will receive an informed consent form introducing the study's purpose, data collection, storage, and duration. To prevent bias, control and treatment group details will not be disclosed to participants. Mental health professionals will have full transparency about the study's objectives and variables. All participants, including professionals and patients, may withdraw at any time or decline to answer specific questions.

Sample

Our sample size of 522 participants was calculated using power analysis to detect meaningful differences between the control and treatment groups. Prior research suggests a control group mean loneliness score of 42 (SD = 10) on the 20-item UCLA Loneliness Scale¹¹. The intervention is expected to reduce loneliness by 2.5 points,

¹¹ Kate L. Lapane, Emily Lim, Emily McPhillips, Adrita Barooah, Yiyang Yuan, and Catherine E. Dube, "Health Effects of Loneliness and Social Isolation in Older Adults Living in Congregate Long-Term Care

resulting in an anticipated mean score of 39.5 for the treatment group. These figures are based on similar studies involving individuals with cognitive impairment or depression¹².

Power analysis conducted with ClinCalc determined that 251 participants per group (502 total) are needed to achieve 80% power (beta = 0.20) at an alpha level of 5%¹². Given the average 10% attrition rate reported in healthcare studies¹³, we have added 20 additional participants (4%) to account for potential attrition, bringing the total recruitment to 522.

Hypotheses: Review Theory

H1: Impact of AI Companions Patients using AI companions under professional guidance will report lower loneliness and higher mental well-being than the control group, with an expected 2.5 point reduction in UCLA Loneliness Scale scores.

H2: Risk Mitigation Professional oversight will mitigate risks like emotional overdependence and problematic internet use, ensuring balanced and supportive AI tool usage in the treatment group

H3: Long-Term Benefits Treatment group patients will sustain reduced loneliness in long-term follow-ups, indicating lasting benefits of the AI intervention.

Concept	Control Group	Treatment Group
Reduction in loneliness	Low	Medium - High
Sustained reduction in loneliness	Low	Medium
Improved mental health	Varied	Medium - High
Sustained improved mental health	Low	Medium
Risks of overdependence	NA	Low-Medium

Variables

Independent variable

Our independent variable is the addition of an AI companion in patient care plans. The control group consists of participants receiving traditional therapy without any AI companion interaction. The treatment group consists of participants receiving traditional

Settings: A Systematic Review of Quantitative and Qualitative Evidence," *Archives of Gerontology and Geriatrics* 102 (2022): 104728, <https://doi.org/10.1016/j.archger.2022.104728>.

¹² "Sample Size Calculator," *ClinCalc*, accessed November 23, 2024, <https://clincalc.com/stats/samplesize.aspx>.

¹³ Ahmed E. Siddiqi, Alla Sikorskii, Charles W. Given, and Barbara Given, "Early Participant Attrition from Clinical Trials: Role of Trial Design and Logistics," *Clinical Trials* 5, no. 4 (2008): 328–35, <https://doi.org/10.1177/1740774508094406>.

therapy in addition to interacting with an AI companion. This variable is manipulated to study how AI companionship impacts the outcomes of the therapy.

Dependent variable

The dependent variable in this study is loneliness levels with an associated focus on mental wellbeing. This will be measured by the UCLA Loneliness Scale, a 20-item questionnaire with a range of 20 to 80. On this scale, higher scores indicate greater loneliness.

Covariates

This study will account for key covariates to better isolate the effects of the independent variable on the dependent outcomes, minimizing confounding factors. These include:

- Baseline loneliness scores to control for initial differences.
- Demographics such as age, education, socioeconomic status, and access to technology.
- Digital literacy and comfort using conversational GenAI for mental health and loneliness treatment.
- Duration of prior therapy to account for familiarity and comfort with the therapeutic setting.

Duration of the Study

The experimentation phase of this study will last six months, with a long-term follow-up conducted six months after completion, at the one year mark. Prior to experimentation begins, four weeks will involve identifying and screening mental health practitioners. Another four weeks will focus on onboarding these practitioners to the study objectives, training them to integrate AI companions into treatment, and recruiting qualified patients. Recruitment will ensure diversity and account for key covariates, aiming to meet the target of 522 participants. This timeline aligns with prior studies on web-based interventions for depression, ensuring sufficient duration to observe the intervention's impact.

Statistical Methods

Statistical analyses will focus on evaluating pre- and post-intervention differences between control and treatment groups, as well as exploring interactions between covariates and outcomes.

Primary Analysis

- A paired t-test will assess changes in loneliness levels (measured by the UCLA Loneliness Scale) within each group over time.
- An independent t-test will compare mean changes in loneliness scores between the treatment and control groups.
- A two-way ANOVA will examine how demographic factors (such as age group or level of digital literacy) might change the way the AI companion affects loneliness. This analysis will test whether certain types of participants benefit more than others from having an AI companion.
- Multiple regression modeling will control for covariates such as baseline loneliness, age, digital literacy, and socioeconomic status. By doing this, we can better understand how much of the improvement in loneliness and mental health is truly linked to using the AI companion, separate from these other influences.

Secondary Analysis

- Repeated measures ANOVA will analyze trends in weekly UCLA Loneliness Scale scores over the course of the study, helping us understand how the scores evolve over time. By assessing whether improvements are maintained, grow, or diminish from one point to the next, we can determine not just if the AI companion leads to initial benefit (H1), but also whether these benefits continue throughout the intervention period and beyond (H3).

All statistical tests will be conducted at a 5% significance level ($\alpha=0.05$), with adjustments for multiple comparisons as needed (e.g., Bonferroni correction). Statistical power of 80% ($1-\beta=0.80$) has been ensured through sample size calculations, allowing for detection of an effect size of 2.5 points on the UCLA Loneliness Scale.

Potential Risks

Integrating AI companions into therapeutic settings introduces several potential risks, which this study seeks to identify, manage, and mitigate:

1. Emotional Overdependence

Patients may form excessive emotional reliance on AI companions, potentially reducing engagement with real-world relationships. Mitigation includes:

- Monitoring by mental health professionals.
- Educating patients on the complementary role of AI in therapy.

2. **Problematic Internet Use (PIU)**

Increased interaction with AI tools could lead to unhealthy usage patterns.

Strategies to address this risk include:

- Tracking app usage data to identify signs of overuse.
- Including digital literacy training as part of the intervention.

3. **Data Privacy Concerns**

Sensitive patient data may be compromised during AI interactions. Safeguards include:

- Anonymizing data and using secure, encrypted servers.
- Transparent communication of data handling practices in consent forms.

4. **Unintended Emotional Responses**

AI interactions could trigger negative emotions if poorly calibrated. Mitigation includes:

- Pre-testing AI companions for tone and appropriateness.
- Preparing patients with warning of the AI's inconsistency and possible inappropriate responses.

By systematically addressing these risks, this study aims to establish ethical guidelines for the safe integration of AI companions into mental health care.

Deliverables

This study will produce three key deliverables:

1. **Efficacy Analysis:** Evaluate whether AI companions reduce loneliness and improve mental health outcomes when paired with professional therapeutic care, providing data-driven insights on their role in addressing these challenges.
2. **Guidelines Development:** Create ethical principles for integrating conversational GenAI into therapy, addressing risks like emotional overdependence, data privacy, and effective use, to support safe and effective implementation by mental health professionals and developers.
3. **mHONcode Updates:** Propose updates to the Medical Health on the Net Code (mHONcode) to address the unique needs of conversational GenAI in mental health, enhancing its relevance for digital care tools.

These deliverables aim to foster collaboration between technology companies and mental health professionals, promoting ethical innovation in AI-assisted therapeutic care.

References

- Ahmed, Arfan, Nashva Ali, Sarah Aziz, Alaa A. Abd-alrazaq, Asmaa Hassan, Mohamed Khalifa, Bushra Elhusein, Maram Ahmed, Mohamed Ali Siddig Ahmed, and Mowafa Househ. "A Review of Mobile Chatbot Apps for Anxiety and Depression and Their Self-Care Features." *Computer Methods and Programs in Biomedicine Update* 1 (2021): 100012.
- Baptista, C., Silva, A. R., Lima, M. P., and Afonso, R. M. "Evaluating the Cognitive Effects of Interventions to Reduce Social Isolation and Loneliness in Older Adults: A Systematic Review." *Aging & Mental Health* (2024): 1–10.
- Barjaková, M., Garnero, A., and d'Hombres, B. "Risk Factors for Loneliness: A Literature Review." *Social Science & Medicine* 334 (October 2023): 116163.
<https://doi.org/10.1016/j.socscimed.2023.116163>.
- ClinCalc.com. "Sample Size Calculator." *ClinCalc*. Accessed November 23, 2024.
<https://clincalc.com/stats/samplesize.aspx>.
- Creswell, John W., and J. David Creswell. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: SAGE Publications, Inc., 2023.
- Davis, Richard A. "A Cognitive-Behavioral Model of Pathological Internet Use." *Computers in Human Behavior* 17 (2001): 187–195. [https://doi.org/10.1016/S0747-5632\(00\)00041-8](https://doi.org/10.1016/S0747-5632(00)00041-8).
- De Freitas, J., A. K. Uguralp, Z. O. Uguralp, and P. Stefano. "AI Companions Reduce Loneliness." *arXiv preprint*, arXiv:2407.19096, 2024. <https://arxiv.org/abs/2407.19096>.
- Donkin, Liesje, Ian Hickie, Helen Christensen, Sharon Naismith, Bruce Neal, Neil Cockayne, and Nick Glozier. "Rethinking the Dose-Response Relationship between Usage and Outcome in an Online Intervention for Depression: Randomized Controlled Trial." *Journal of Medical Internet Research* 15, no. 10 (2013): e231. <https://doi.org/10.2196/jmir.2771>.
- Garcia, Patricia. "Alone Together: Why We Expect More from Technology and Less from Each Other by Sherry Turkle." *InterActions: UCLA Journal of Education and Information Studies* 8, no. 1 (2012).
- Harvard T.H. Chan School of Public Health. "From Loneliness to Social Connection: Lessons from Research and a Global Pandemic." Last modified October 11, 2023.
<https://www.hsph.harvard.edu/health-happiness/2023/10/11/from-loneliness-to-social-connection-lessons-from-research-and-a-global-pandemic/>.
- Hickin, Natasha, Annika Käll, Roz Shafran, Sarah Sutcliffe, Giovanni Manzotti, and Daniel Langan. "The Effectiveness of Psychological Interventions for Loneliness: A Systematic Review and Meta-Analysis." *Clinical Psychology Review* 88 (2021): 102066.

De Jong Gierveld, Jenny, and Theo Van Tilburg. "The De Jong Gierveld Short Scales for Emotional and Social Loneliness: Tested on Data from 7 Countries in the UN Generations and Gender Surveys." *European Journal of Ageing* 7, no. 2 (June 2010): 121–30.

<https://doi.org/10.1007/s10433-010-0144-6>.

Klein, Jan Philipp, Christian Späth, Jörg Schröder, Beate Meyer, Wolfgang Greiner, Martin Hautzinger, Wolfgang Lutz, et al. "Time to Remission from Mild to Moderate Depressive Symptoms: One-Year Results from the EVIDENT Study, an RCT of an Internet Intervention for Depression." *Behaviour Research and Therapy* 97 (2017): 154–62.

<https://doi.org/10.1016/j.brat.2017.07.013>.

Lapane, Kate L., Emily Lim, Emily McPhillips, Adrita Barooah, Yiyang Yuan, and Catherine E. Dube. "Health Effects of Loneliness and Social Isolation in Older Adults Living in Congregate Long-Term Care Settings: A Systematic Review of Quantitative and Qualitative Evidence." *Archives of Gerontology and Geriatrics* 102 (2022): 104728.

<https://doi.org/10.1016/j.archger.2022.104728>.

Linardon, Jason, et al. "The Efficacy of App-Supported Smartphone Interventions for Mental Health Problems: A Meta-Analysis of Randomized Controlled Trials." *World Psychiatry* 18 (2019): 325–36.

Maples, Bethanie, Merve Cerit, Aditya Vishwanath, and Roy Pea. "Loneliness and Suicide Mitigation for Students Using GPT-3-Enabled Chatbots." *NPJ Mental Health Research* 3 (January 22, 2024): 4. <https://doi.org/10.1038/s44184-023-00047-6>.

Moretta, T., G. Buodo, Z. Demetrovics, and M. N. Potenza. "Tracing 20 Years of Research on Problematic Use of the Internet and Social Media: Theoretical Models, Assessment Tools, and an Agenda for Future Work." *Comprehensive Psychiatry* 112 (2022): 152286.

Morrish, Nicola, Srishti Choudhury, and Anibal Medina-Lara. "What Works in Interventions Targeting Loneliness: A Systematic Review of Intervention Characteristics." *BMC Public Health* 23 (2023): 2214. <https://doi.org/10.1186/s12889-023-17097-2>.

Reed, Phil, Angharad Davies, Katie Evans, and Lisa A. Osborne. "Longitudinal Relationship between Problematic Internet Use with Loneliness during and after COVID-19 Social Restrictions." *Psychiatry Research* 323 (2023): 115148.

Russell, Daniel, Letitia A. Peplau, and Carolyn E. Ferguson. "Developing a Measure of Loneliness." *Journal of Personality Assessment* 42, no. 3 (1978): 290–94.

https://doi.org/10.1207/s15327752jpa4203_11.

Siddiqi, Ahmed E., Alla Sikorskii, Charles W. Given, and Barbara Given. "Early Participant Attrition from Clinical Trials: Role of Trial Design and Logistics." *Clinical Trials* 5, no. 4 (2008): 328–35. <https://doi.org/10.1177/1740774508094406>.