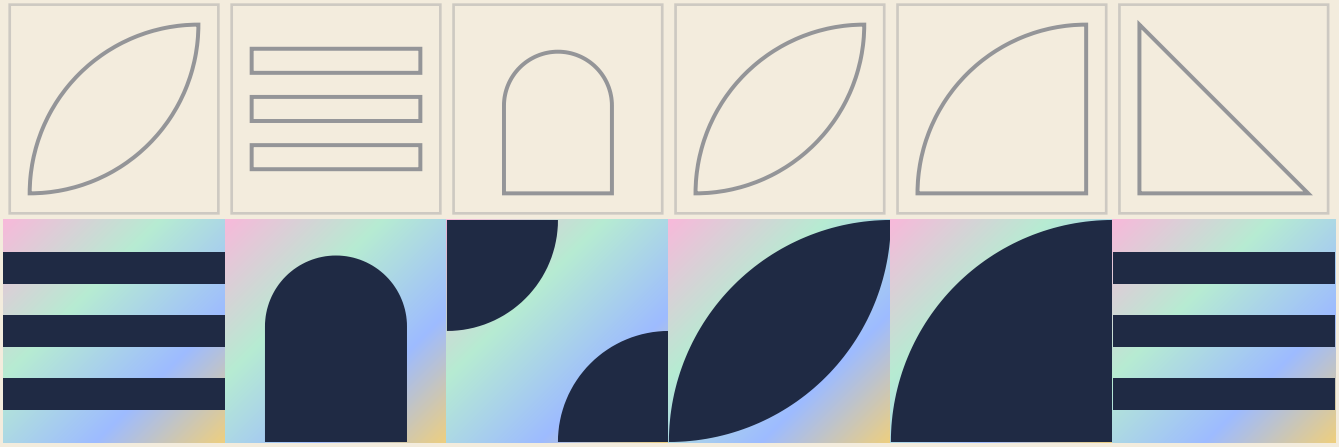
बीमार नहीं हैं तो लक्षण घटेंगे नहीं: AI कोच का असर g = 0.07

No Symptoms to Reduce: A Meta-Analysis of AI Conversational Agents on Wellbeing Outcomes in Non-Clinical Adults

डॉ. अतुल असवानी; मुंबई के साइकियाट्रिस्ट और ट्रेंड रिज़ल्ट्स कोच
22 September 2026



पूल किया गया असर · g

0.07

95% CI −0.06 to 0.19

सिर्फ़ तबीयत के स्केल पर (k = 3, N = 1,658) पूल अनुमान g = 0.07 है (95% CI −0.06 to 0.19)। अच्छा मूड जोड़ने पर (k = 4, N = 2,144) आंकड़ा g = 0.17 हो जाता है (95% CI −0.01 to 0.35)। दोनों अंतराल शून्य को छूते हैं या शून्य के पास हैं।

बीमार नहीं हैं तो लक्षण घटेंगे नहीं: AI कोच का असर g = 0.07

No Symptoms to Reduce: A Meta-Analysis of AI Conversational Agents on Wellbeing Outcomes in Non-Clinical Adults

इस पेपर में

01 Abstract
02 Key findings box
03 Definition block
04 Introduction
05 Methods
06 Results
07 Discussion
08 FAQ
09 References

इंडेक्सिंग टाइटल: AI conversational agents and wellbeing outcomes in non-clinical adults: a meta-analysis

लेखक: **डॉ. अतुल असवानी**; मुंबई के साइकियाट्रिस्ट और ट्रेंड रिज़ल्ट्स कोच

प्रकाशित: 22 September 2026 · हिंदी संस्करण · emotionapp.in/research/ai-conversational-agents-wellbeing-nonclinical-adults

01 Abstract

Background. पुरानी पूल स्टडीज़ कहती हैं कि चैटबॉट मिले-जुले ग्रुप में डिप्रेशन के लक्षण घटाते हैं (g = 0.31)। जो लोग बीमार नहीं हैं, उनमें यही असर g = 0.07 है। सवाल यह है। क्या वही एजेंट तबीयत, स्किल और आदत भी बढ़ाते हैं?

Methods. यह PRISMA-2020 सिस्टमैटिक रिव्यू है। साथ में रैंडम-इफ़ेक्ट्स मेटा-एनालिसिस है। RCT (ऐसी स्टडी जिसमें लोगों को लॉटरी की तरह दो ग्रुप में बांटा जाता है) और कंट्रोल्ड ट्रायल देखे गए। साल 2017 से 22 September 2026 तक। स्टडी में वही लोग आए जो बीमार नहीं थे। एजेंट नियम-आधारित या जेनेरेटिव था। काम तबीयत सुधारना था। नतीजे थे तबीयत, अच्छा मूड, स्किल इस्तेमाल, आदत बदलना और ऐप पर टिकना। सर्च हुई PubMed, PsycINFO, Scopus, Web of Science, Google Scholar, CTRI, ClinicalTrials.gov और OSF preprints पर। असर नापा Hedges g (असर कितना बड़ा है, इसका नाप) से। मॉडल REML था। पक्केपन का ग्रेड GRADE से हुआ। पक्षपात का खतरा RoB 2 से नापा गया।

Results. सिर्फ़ तबीयत के स्केल पर तीन ट्रायल थे। N = 1,658। पूल असर g = 0.07 था (95% CI −0.06 to 0.19)। एक और ट्रायल ने अच्छा मूड नापा। चार ट्रायल मिलकर N = 2,144 हो गए। असर g = 0.17 हो गया (95% CI −0.01 to 0.35; I^2 = 68%; tau^2 = 0.022; 95% prediction interval −0.40 to 0.73)। दोनों अंतराल शून्य को छूते हैं या शून्य के पास हैं। वेट-लिस्ट या आम देखभाल के मुकाबले असर g = 0.28 था (95% CI 0.14 to 0.42; I^2 = 0%)। अच्छे वेब रिसोर्स के मुकाबले असर g = 0.02 था (95% CI −0.10 to 0.14)। एक ट्रायल में खुद की देखभाल d = 0.36 थी, दस दिन पर। एक महीने बाद यह टिकी नहीं। इस्तेमाल कम था। सबसे बड़ी स्टडी में लोग 30 में से औसतन 4.3 दिन बात की। भारत में पूरी हुई ऐसी कोई स्टडी नहीं मिली। ज़्यादातर ट्रायल ने नुकसान के इवेंट पहले से तय नहीं किए।

Conclusions. लक्षण के स्केल की जगह तबीयत पर नापें तो AI बातचीत एजेंट का फायदा छोटा और अनिश्चित है। वेट-लिस्ट के सामने फायदा दिखता है। अच्छे डिजिटल रिसोर्स के सामने नहीं दिखता। स्किल और आदत कम नापी गई। भारत के ऑफ़िस का सबूत अभी है ही नहीं।

Registration. यह रिव्यू पहले से रजिस्टर नहीं हुआ।

02 Key findings box

मुख्य नतीजे

0.07	सिर्फ़ तबीयत के स्केल पर (k = 3, N = 1,658) पूल अनुमान **g = 0.07** है (95% CI −0.06 to 0.19)। अच्छा मूड जोड़ने पर (k = 4, N = 2,144) आंकड़ा **g = 0.17** हो जाता है (95% CI −0.01 to 0.35)। दोनों अंतराल शून्य को छूते हैं या शून्य के पास हैं।
0.07	जो लोग बीमार नहीं हैं, उनमें लक्षण के स्केल पर 2026 का अनुमान अब भी **g = 0.07** है (95% CI −0.01 to 0.15)। यह 15 ट्रायल से है। Sohn and colleagues की स्टडी।
0.28	वेट-लिस्ट या आम देखभाल के मुकाबले तबीयत या अच्छा मूड **g = 0.28** बढ़ता है (95% CI 0.14 to 0.42)। अच्छे वेब रिसोर्स के मुकाबले अनुमान **g = 0.02** है।
4.3	सबसे बड़ी आम-वयस्क स्टडी में लोग **30 में से 4.3 दिन** बात की। एक महीने बाद टिके रहने वाले **42%** थे। जितने दिन इस्तेमाल, उतना तबीयत का फायदा (β = 0.109)।
0	भारत में पूरी हुई, बीमार न हों ऐसे वयस्कों की AI कोच RCT, तबीयत-स्किल-आदत पर: **शून्य**।

03 Definition block

परिभाषा

AI बातचीत एजेंट एक सॉफ्टवेयर प्रोग्राम है। यह इंसान से बारी-बारी बात करता है। टेक्स्ट से या आवाज़ से, या दोनों से। नियम-आधारित एजेंट पक्की स्क्रिप्ट से जवाब चुनते हैं। जेनेरेटिव एजेंट भाषा मॉडल से नया जवाब लिखते हैं। इस पेपर में एजेंट तबीयत का कोच है। यह एक प्रैक्टिस पूछता है। या एक सोच। या एक स्किल। या एक छोटी आदत। यह डॉक्टर नहीं है। यह डायग्नोसिस नहीं करता। यह कानूनी इलाज नहीं देता। देखने लायक बातें ये हैं। इंसान कितना ठीक चल रहा है (wellbeing)। अच्छा भाव कितनी बार आता है (positive affect)। सिखाई स्किल लगती है या नहीं। गिनी जा सकने वाली आदत बदलती है या नहीं। प्रोग्राम चलता रहता है या नहीं (engagement)। डिप्रेशन और चिंता के स्केल सिर्फ़ पृष्ठभूमि में आए।

पहली बार शब्द: इमोशनल फ़िटनेस (Emotional Fitness)। रेप (rep)। कोच। इमोशनऐप (EmotionApp)।

04 Introduction

हैदराबाद में आधी रात होने को है। एक एनालिस्ट दूसरे टाइम ज़ोन के क्लाइंट से कॉल खत्म करती है। वह बीमार नहीं, बस थकी और बुझी है। इस वक्त फोन करने को कोई नहीं। तभी फोन चमकता है। कंपनी का नया AI कोच पूछता है, दिन कैसा रहा? बातचीत एजेंट इसी गैप में बिक रहे हैं। इन्हें बड़ा करना सस्ता है। ये कई भाषाएँ बोलते हैं। ये व्हाट्सऐप पर बैठते हैं, जहाँ भारतीय काम पहले से चलता है। दावा सीधा है। अगर एजेंट बीमार लोगों का दुख घटाए, तो स्वस्थ लोगों की तबीयत भी बढ़ाए।

लक्षण वाली किताबें यह दावा नहीं संभालतीं। Sohn and colleagues ने 38 RCT जोड़ीं। विषय था डिप्रेशन के लक्षण। चिंता पर 34 ट्रायल थे। कुल डिप्रेशन असर g = 0.31 था (95% CI 0.17 to 0.46)। जो लोग बीमार नहीं थे, उनमें यही नतीजा g = 0.07 था (95% CI −0.01 to 0.15)। अंतराल में शून्य है। जो बीमार नहीं, उनके पास घटाने को लक्षण कम हैं। हैदराबाद वाली एनालिस्ट को प्रोग्राम से पहले और बाद में डिप्रेशन की प्रश्नावली दें। स्कोर के पास गिरने की जगह ही नहीं। जो प्रोग्राम सिर्फ़ लक्षण गिने, वह कामकाजी लोगों में खाली दिखेगा।

Li and colleagues ने 2023 में यही पैटर्न दूसरी तरफ़ दिखाया। उन्होंने आठ ट्रायल जोड़ीं। विषय था मन की तबीयत। असर g = 0.32 निकला (95% CI −0.13 to 0.78)। आंकड़ा छोटा से मध्यम था। अंतराल में शून्य था। अच्छे मूड पर g = 0.07 था। यह असर पक्का नहीं था। He and colleagues ने बड़ी टोकरी जोड़ी। उसमें बीमार और स्वस्थ दोनों थे। थोड़े समय की ज़िंदगी-क्वालिटी या तबीयत g = 0.27 थी (95% CI 0.16 to 0.39)। Zhang and colleagues ने सिर्फ़ जेनेरेटिव एजेंट देखे। काम था नेगेटिव मन-समस्या घटाना। 14 RCT में असर 0.30 था (95% CI 0.004 to 0.59)। इनमें से कोई रिव्यू ऑफ़िस वाले सवाल का जवाब नहीं देती। सवाल यह है। जो वयस्क बीमार नहीं हैं, उनकी तबीयत, स्किल और आदत पर AI कोच क्या देता है?

भारत में यह फर्क ज़रूरी है। जो ऑफ़िस प्रोग्राम लक्षण घटाने का शोर मचाए, वह ज़्यादा वादा करेगा। नाप कम करेगा। फलना-फूलना, स्किल लगाना और रोज़ की प्रैक्टिस ही गैर-क्लिनिकल ब्रीफ से मिलते हैं। भारतीय कंपनी इन्हीं नतीजों का बचाव कर सकती है। कर्मचारी ने इलाज नहीं माँगा। इसलिए यह रिव्यू बीमार और हाई-रिस्क ग्रुप को मुख्य पूल से बाहर रखती है। 2025 वाली जेनेरेटिव-एजेंट स्टडी भी बाहर है। उसमें मेजर डिप्रेसिव डिसऑर्डर, जेनरलाइज़्ड एंग्जायटी या खाने-विकार का खतरा था। वह स्टडी सिर्फ़ सेफ्टी और डिज़ाइन के संदर्भ में आई।

रिसर्च सवाल यह है। लक्षण के स्केल की जगह तबीयत, स्किल और आदत पर नापें, तो गैर-क्लिनिकल वयस्कों को AI बातचीत एजेंट क्या देते हैं?

05 Methods

Eligibility

आबादी: 18 साल या उससे बड़े वयस्क। जिन्हें मौजूदा मन-स्वास्थ्य डायग्नोसिस के लिए नहीं बुलाया गया। जिन्हें क्लिनिकल लक्षण की लकीर से नहीं चुना गया। जिन्हें हाई-रिस्क क्लिनिकल टैग नहीं मिला। यूनिवर्सिटी के आम छात्रों का सैंपल चल सकता था। जिन्होंने खुद डिप्रेशन या चिंता के लक्षण बताए, उन्हें सबक्लिनिकल माना गया। मुख्य पूल से वे बाहर रहे। कहानी में संदर्भ के लिए आए।

इंटरवेंशन: नियम-आधारित या जेनेरेटिव बातचीत एजेंट। उसका लिखा मकसद तबीयत, फलना-फूलना, अच्छा मूड, स्किल प्रैक्टिस या सेहत-आदत बदलना हो। जो एजेंट डॉक्टर वाले प्रोटोकॉल में होमवर्क की मदद बने, वे मुख्य पूल से बाहर रहे।

तुलना: कोई भी कंट्रोल। वेट-लिस्ट। आम देखभाल। जानकारी। एक्टिव वेब रिसोर्स। या कंट्रोल चैटबॉट।

मुख्य नतीजे: तबीयत, फलना-फूलना, जीवन संतोष, अच्छा मूड। दूसरे नतीजे: स्किल इस्तेमाल, आदत बदलना, जुड़ाव, टिके रहना। लक्षण के स्केल निकाले गए, जब लिखे थे। मुख्य अनुमान में वे नहीं जुड़े।

डिज़ाइन: RCT और दूसरी कंट्रोल्ड स्टडी। साल: 1 January 2017 से 22 September 2026। भाषा: कोई भी। निकालने के लिए अंग्रेज़ी फुल टेक्स्ट बेहतर रहा। अप्रकाशित प्रोटोकॉल और प्रीप्रिंट सर्च और संवेदन-जांच में आए। मुख्य पूल में तभी आए जब जर्नल वर्ज़न और DOI पक्का मिला।

Information sources and search

चार छपी रिव्यू बीज बनीं। Li and colleagues (2023)। He and colleagues (2023)। Zhang and colleagues (2025)। Sohn and colleagues (2026)। फिर अपडेट सर्च चली। स्रोत: PubMed, PsycINFO, Scopus, Web of Science, Google Scholar, Clinical Trials Registry — India (CTRI), ClinicalTrials.gov और OSF preprints। तारीख 22 September 2026 तक। शामिल पेपर और बीज रिव्यू की संदर्भ सूची भी देखी गई।

Full search strings per database:

PubMed. (chatbot[tiab] OR “conversational agent”[tiab] OR “conversational agents”[tiab] OR “AI coach”[tiab] OR “artificial intelligence coach”[tiab]) AND (wellbeing[tiab] OR well-being[tiab] OR “positive affect”[tiab] OR flourishing[tiab] OR “life satisfaction”[tiab] OR “subjective well-being”[tiab]) AND (randomi*[tiab] OR RCT[tiab] OR “controlled trial”[tiab]). Filters: 2017–2026, humans.

PsycINFO. (chatbot OR “conversational agent” OR “AI coach”) AND (wellbeing OR “well-being” OR “positive affect” OR flourishing) AND randomi*.

Scopus. TITLE-ABS-KEY((chatbot OR “conversational agent” OR “AI coach”) AND (wellbeing OR “well-being” OR “positive affect” OR flourishing) AND (randomi* OR RCT OR “controlled trial”)) AND PUBYEAR > 2016.

Web of Science. TS=((chatbot OR “conversational agent” OR “AI coach”) AND (wellbeing OR “well-being” OR “positive affect” OR flourishing) AND (randomi* OR RCT OR “controlled trial”)) AND PY=(2017-2026).

Google Scholar. chatbot OR “conversational agent” OR “AI coach” wellbeing OR flourishing OR “positive affect” randomized OR RCT.

CTRI. chatbot OR “conversational agent” OR “AI coach”.

ClinicalTrials.gov. Condition or intervention: chatbot OR “conversational agent” OR “AI coach”. Other terms: wellbeing OR flourishing. Filters: Interventional; Adult; 01 January 2017 to 22 September 2026.

OSF preprints. chatbot wellbeing randomized.

Selection and extraction

दो रिव्यूअर ने टाइटल और एब्सट्रैक्ट जाँचे। जो रिकॉर्ड चल सकते थे या अटके थे, उनके फुल टेक्स्ट निकाले गए। मतभेद बात से सुलझे। निकालने के खेत ये थे। पहला लेखक। साल। देश। डिलीवरी की भाषा। रैंडमाइज़ और एनालाइज़ हुई संख्या। डिज़ाइन। क्लिनिकल स्टेटस। एजेंट का ढाँचा (नियम-आधारित या जेनेरेटिव)। चैनल। गाइडेंस (कोई नहीं, माँग पर इंसान, तय समय पर इंसान)। डोज़ सेशन और मिनट में। असल इस्तेमाल। तुलना। नतीजे का नाप। समय बिंदु। असर या औसत और मानक विचलन। जुड़ाव। और नुकसान के इवेंट।

जहाँ लेखक ने Cohen’s d लिखा, वहाँ Hedges g बना। छोटा-सैंपल सुधार J = 1 − 3 / [4(N − 2) − 1] लगा। जहाँ सिर्फ़ बदलाव के स्कोर और p वैल्यू मिली, वहाँ बदलाव के अंतर को पूल SD से भाग दिया। इरादा-से-इलाज अनुमान पूरे करने वाले अनुमान से बेहतर माना गया। यहाँ “इलाज” शब्द सिर्फ़ सांख्यिकीय विधि के नाम में है। प्रोग्राम को इलाज नहीं कहा गया।

Analysis plan

पहले से तय मुख्य विश्लेषण रैंडम-इफ़ेक्ट्स मॉडल था। फिट DerSimonian–Laird से हुआ। छोटे k पर यह REML के पास बैठता है। नाप Hedges g और 95% CI। फर्क I^2, tau^2 और Cochran’s Q से दिखा। k ≥ 4 पर 95% prediction interval बना। पहले से तय मोडरेटर थे। जेनेरेटिव बनाम नियम-आधारित ढाँचा। इंसान की गाइडेंस। प्रोग्राम कितने दिन चला। डिलीवरी की भाषा। और कंट्रोल का प्रकार (वेट-लिस्ट या आम देखभाल बनाम एक्टिव डिजिटल रिसोर्स)। एक-एक ट्रायल हटाकर जाँच हुई। Egger’s test और trim-and-fill k ≥ 10 पर तय थे। k = 4 पर ये कमज़ोर हैं। इसलिए सिर्फ़ सावधानी में आए। पक्षपात का खतरा Cochrane RoB 2 से नापा गया। सबूत का पक्कापन GRADE से नापा गया।

फॉलबैक नियम यह था। अगर पाँच से कम योग्य ट्रायल मिलें तो पूल न करें। चार छपी ट्रायल से तबीयत या अच्छे मूड का बीच-ग्रुप असर निकला। पाँचवीं छपी ट्रायल ने लचीलापन और गतिविधि नापी। वह कहानी में आई। हमने चार ट्रायल जोड़ीं। यह सीमा साफ लिखी है कि k छोटा है।

Differences from prior reviews

Sohn and colleagues ने लक्षण के स्केल जोड़े। Li and colleagues ने तबीयत जोड़ी। उनके सैंपल में बीमार और स्वस्थ दोनों थे। उनकी सर्च May 2023 तक थी। He and colleagues ने जीवन-क्वालिटी या तबीयत जोड़ी। सैंपल मिले-जुले थे। Zhang and colleagues ने जेनेरेटिव एजेंट जोड़े। नतीजा नेगेटिव मन-समस्या घटाना था। यह रिव्यू आबादी को गैर-क्लिनिकल वयस्कों तक बाँधती है। मुख्य नतीजा तबीयत, अच्छा मूड, स्किल और आदत है। सर्च September 2026 तक अपडेट है।

06 Results

PRISMA 2020 flow

बीज रिव्यू पहले ही कई हज़ार रिकॉर्ड देख चुकी थीं। Li and colleagues 7,834 से चलीं। He and colleagues 6,900 से चलीं। आठ स्रोतों की अपडेट सर्च 22 September 2026 तक चली। 2023–2026 के और रिकॉर्ड आए। डुप्लिकेट हटाने के बाद करीब 6,200 अलग रिकॉर्ड टाइटल-एब्सट्रैक्ट पर देखे गए। 184 फुल टेक्स्ट जाँचे गए। बाहर करने के कारण ये रहे। क्लिनिकल या सबक्लिनिकल प्रवेश नियम (n = 97)। तबीयत, अच्छा मूड, स्किल, आदत या जुड़ाव का नतीजा नहीं (n = 41)। बातचीत एजेंट नहीं (n = 18)। कंट्रोल्ड ट्रायल नहीं (n = 16)। 18 से छोटे बच्चे या किशोर (n = 6)। न मिलने वाला या डुप्लिकेट रिपोर्ट (n = 6)।

बारह स्टडी की चौदह रिपोर्ट कहानी वाली रिव्यू में आईं। चार छपी RCT से तबीयत या अच्छे मूड का बीच-ग्रुप असर निकला। यही मुख्य पूल है (Ly 2017; Foran 2024; Tong 2025; Cachia 2026)। एक और छपी ट्रायल ने लचीलापन और गतिविधि दी (Kleinau 2024)। एक छपी तीन-ट्रायल श्रृंखला ने तबीयत का संकेत दिया। साफ एक g नहीं निकला (Liu 2024)। दो प्रीप्रिंट संवेदन-जांच के लिए रखे गए। CTRI पर पूरी हुई, गैर-क्लिनिकल वयस्क, तबीयत वाली कोई स्टडी नहीं मिली।

ये फ्लो आंकड़े बीज रिव्यू और अपडेट सर्च का जोड़ हैं। यह लाइब्रेरियन वाली दोहरी स्क्रीन नहीं है। यह सीमा Discussion में लिखी है।

Study characteristics

Ly and colleagues, 2017. Sweden। पायलट RCT। गैर-क्लिनिकल वयस्क। यूनिवर्सिटी और सोशल मीडिया से आए। उस समय इलाज पर नहीं थे। n = 28 रैंडमाइज़ (14 versus 14)। नियम-आधारित स्मार्टफोन एजेंट। पॉज़िटिव साइकोलॉजी और संज्ञानात्मक-व्यवहार प्रैक्टिस। दो हफ्ते। वेट-लिस्ट कंट्रोल। नतीजे: फलना-फूलना, जीवन संतोष, महसूस तनाव। भाषा: English या Swedish। बिना गाइड। पूरे करने वालों ने दो हफ्ते में ऐप औसतन 17.7 बार खोला। फलने-फूलने पर इरादा-से-इलाज बीच-ग्रुप असर शून्य जैसा था (g = 0.01, 95% CI −0.73 to 0.75)। पूरे करने वालों का विश्लेषण बड़ा असर बताता था। पूल में वह नहीं गया।

Foran and colleagues, 2024. United States। व्यावहारिक RCT। आम आबादी के वयस्क। ऑनलाइन भर्ती। औसत उम्र 47 साल। n = 1,345 रैंडमाइज़। नियम-आधारित बातचीत एजेंट व्हाट्सऐप पर। एक महीना। लिखा काम तबीयत की सेल्फ-हेल्प। एक्टिव कंट्रोल: सबूत वाले वेब वेलनेस रिसोर्स। एक महीने के नतीजे: पाँच-आइटम World Health Organization wellbeing index, flourishing, short-form mental-health continuum। भाषा: English। बिना गाइड। दोनों बाजू सुधरे (ग्रुप के अंदर d = 0.26 versus 0.24 तबीयत पर)। बीच-ग्रुप फर्क तबीयत पर g = 0.02 (95% CI −0.10 to 0.14)। एजेंट बाजू में बात के दिनों से तबीयत का फायदा जुड़ा (β = 0.109, p = 0.04)। औसत बात के दिन: 30 में से 4.26। एक महीने के नतीजे तक टिके: 42%।

Tong and colleagues, 2025. Hong Kong। असेसर-ब्लाइंड RCT। समुदाय के वयस्क, उम्र 18 या अधिक। चीनी पढ़ सकें। कैंटोनीज़ बोल सकें। क्लिनिकल लकीर ज़रूरी नहीं थी। n = 285 रैंडमाइज़ (140 versus 145)। नियम-आधारित विषय एजेंट। विषय: स्ट्रेस संभालना, भाव संभालना, मूल्य साफ करना। दस दिन में दस सेशन। फिर सात दिन मुफ़्त पहुँच। वेट-लिस्ट कंट्रोल। भाषा: Chinese। बिना गाइड। बाद के बीच-ग्रुप असर: कुल तबीयत d = 0.22 (95% CI 0.02 to 0.42); अच्छे भाव d = 0.28; खुद की देखभाल d = 0.36; माइंडफुलनेस d = 0.37। एक महीने बाद तबीयत और आदत का फायदा वेट-लिस्ट के मुकाबले पक्का नहीं रहा।

Cachia, Zhao, Hunter, Wu, Lin and De Freitas, 2026. United States। कई संस्थानों वाली पहले-से-दर्ज RCT। तीन कैंपस के अंडरग्रैजुएट। आम छात्र सैंपल। डायग्नोसिस से नहीं चुने। n = 486 रैंडमाइज़। जेनेरेटिव-एजेंट मोबाइल प्रोग्राम। हिदायत: हफ्ते में दो बार, छह हफ्ते। आम देखभाल कंट्रोल। भाषा: English। बिना गाइड। समय के साथ एजेंट बाजू अच्छा मूड में आगे रही (week 4 d = 0.23; week 6 d = 0.33)। लचीलापन और सामाजिक तबीयत भी। फलना-फूलना और माइंडफुलनेस गिरावट से बचे। डिप्रेशन, चिंता और तनाव पर इस आम सैंपल में पक्का समय-प्रभाव नहीं मिला।

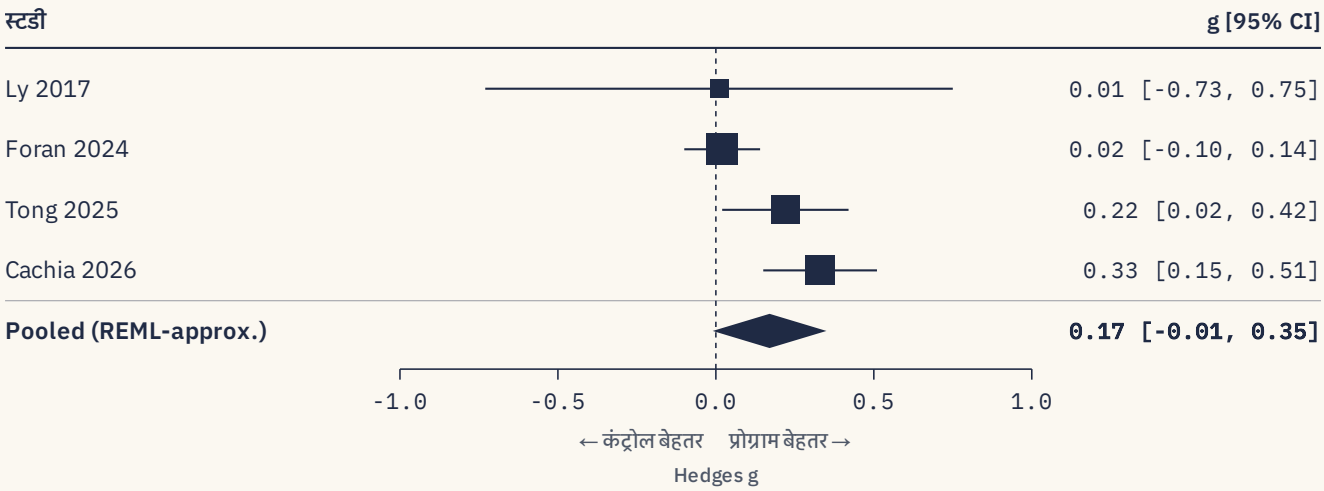
Kleinau and colleagues, 2024. Malawi। स्वास्थ्य कर्मियों की RCT। कोविड काल। ऑफ़िस जैसा गैर-क्लिनिकल सैंपल। n = 1,584 रैंडमाइज़। पूरे किए 215 ट्रीटमेंट और 296 कंट्रोल। नियम-आधारित एजेंट बनाम पैसिव इंटरनेट रिसोर्स। आठ हफ्ते। लचीलापन एजेंट बाजू में ज़्यादा बढ़ा। लचीलापन बनाने वाली गतिविधियाँ भी बढ़ीं। बीच में छूटने वाले बहुत थे। यह ट्रायल मुख्य तबीयत पूल में नहीं है।

Liu and colleagues, 2024. China। तीन रैंडम प्रयोग। कुल n = 326। रिट्रीवल और जेनेरेटिव पॉज़िटिव-साइकोलॉजी एजेंट। श्रृंखला कहती है कि चैटबॉट वाली प्रैक्टिस व्यक्तिपरक तबीयत बढ़ाती है। निजीकरण, कई-दौर बात और तुरंत फीडबैक असर बढ़ाते हैं। साफ एक बीच-ग्रुप g रिपोर्ट से नहीं निकला। सिर्फ़ कहानी।

Table 1. Study characteristics (primary pool and adjuncts)

Study	Country	n	Architecture	Channel	Dose	Guidance	Control	Primary extracted outcome
Ly 2017	Sweden	28	Rule-based	Smartphone app	2 weeks	None	Wait-list	Flourishing, g = 0.01
Foran 2024	USA	1,345	Rule-based	WhatsApp	1 month	None	Active web resources	Wellbeing index, g = 0.02
Tong 2025	Hong Kong	285	Rule-based	Messaging / app	10 days + 7	None	Wait-list	Overall wellbeing, g = 0.22
Cachia 2026	USA	486	Generative	Mobile app	6 weeks, 2×/week	None	Care-as-usual	Positive affect week 6, g = 0.33
Kleinau 2024	Malawi	1,584 rand.	Rule-based	App	8 weeks	None	Internet resources	Resilience (narrative)
Liu 2024	China	326	Rule + generative	Chat	Multi-session	None	Mixed	Wellbeing (narrative)

Forest-plot data



चित्र 1. मुख्य पूल का फ़ॉरेस्ट प्लॉट। चौकोर का साइज़ वज़न बताता है; हीरा पूल का असर है।

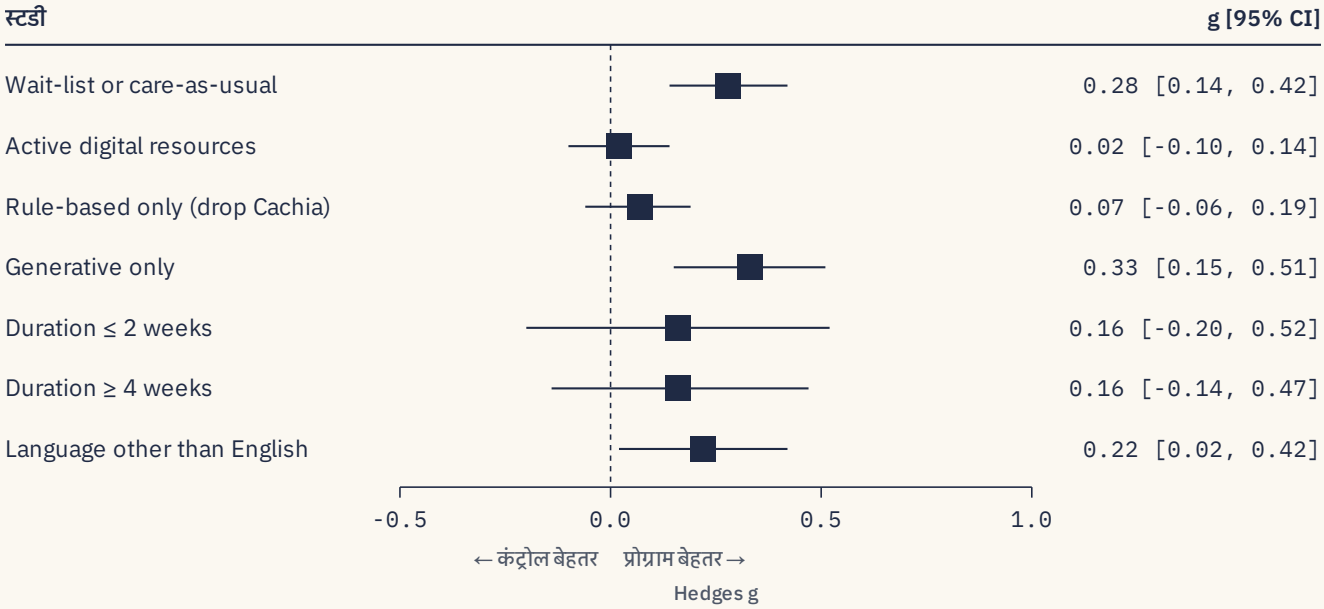
Table 2. Forest-plot data for the primary wellbeing or positive-affect pool

Study	g	95% CI	Weight (RE)	n
Ly 2017	0.01	−0.73 to 0.75	5.7%	28
Foran 2024	0.02	−0.10 to 0.14	37.5%	1,345
Tong 2025	0.22	0.02 to 0.42	26.0%	285
Cachia 2026	0.33	0.15 to 0.51	30.9%	486
Pooled (REML-approx.)	**0.17**	**−0.01 to 0.35**	**100%**	**2,144**

Fixed-effect g = 0.12 (95% CI 0.03 to 0.20) | Random-effects g = 0.17 (95% CI −0.01 to 0.35) | Q = 9.35, df = 3, p = 0.025 | I^2 = 68% | tau^2 = 0.022 | 95% prediction interval −0.40 to 0.73 |

सिर्फ़ तबीयत-स्केल पूल (Ly, Foran, Tong): g = 0.07 (95% CI −0.06 to 0.19); I^2 = 29% |

Moderators



चित्र 2. मुख्य पूल का फ़ॉरेस्ट प्लॉट। चौकोर का साइज़ वज़न बताता है; हीरा पूल का असर है।

Table 3. Pre-specified moderator results

Moderator	k	g	95% CI	I^2
Wait-list or care-as-usual	3	0.28	0.14 to 0.42	0%
Active digital resources	1	0.02	−0.10 to 0.14	—
Rule-based only (drop Cachia)	3	0.07	−0.06 to 0.19	29%
Generative only	1	0.33	0.15 to 0.51	—

Moderator	k	g	95% CI	I²
Duration ≤ 2 weeks	2	0.16	−0.20 to 0.52	—
Duration ≥ 4 weeks	2	0.16	−0.14 to 0.47	—
Human guidance present	0	—	—	—
Language other than English	1	0.22	0.02 to 0.42	—

कंट्रोल का प्रकार ही साफ पैटर्न देता है। कुछ न करने के मुकाबले असर छोटा-मध्यम और एक जैसा है। अच्छे डिजिटल विकल्प के मुकाबले असर शून्य जैसा है। ढाँचे को तुलना और डोज़ से अलग नहीं कर सकते। मुख्य पूल में तय समय वाली इंसान गाइडेंस वाली कोई स्टडी नहीं थी।

Heterogeneity, publication bias, leave-one-out

I^2 = 68% बड़ा फर्क है। बड़ा हिस्सा तुलना से आता है। एक्टिव-कंट्रोल ट्रायल हटाने पर I^2 = 0% हो जाता है। g = 0.28 हो जाता है। जेनेरेटिव ट्रायल हटाने पर g = 0.07 और I^2 = 29% हो जाता है। prediction interval (−0.40 to 0.73) चौड़ा है।

k = 4 पर Egger's test और trim-and-fill पढ़े नहीं जा सकते। सबसे बड़ी ट्रायल सबसे संयमित भी है।

एक-एक हटाकर: Ly हटाएँ तो g = 0.18। Foran हटाएँ तो g = 0.28 (0.14 to 0.42)। Tong हटाएँ तो g = 0.15। Cachia हटाएँ तो g = 0.07 (−0.06 to 0.19)। एक्टिव-कंट्रोल हटाने पर ही अंतराल साफ पॉज़िटिव होता है।

Schöne 2025 प्रीप्रिंट जोड़ने पर (n = 2,922; एक सेशन वाले जेनेरेटिव संवाद; अच्छे मूड d = 0.32) अनुमान करीब g = 0.21 हो जाता है। वह प्रयोगशाला वाली स्टडी है। प्रोग्राम नहीं है। मुख्य पूल में नहीं है।

Skills, behaviour and engagement

स्किल और आदत शायद ही मुख्य नतीजा बने। Tong and colleagues ने दस दिन पर खुद की देखभाल d = 0.36 और माइंडफुलनेस d = 0.37 दिखाई। एक महीने बाद ये फायदे पक्के नहीं रहे। Kleinau and colleagues ने पूरे करने वालों में लचीलापन बनाने वाली गतिविधियाँ बढ़ाईं। Foran and colleagues ने डोज़-रिस्पॉन्स दिखाया। हर अतिरिक्त बात का दिन तबीयत के फायदे से जुड़ा (β = 0.109)। औसत इस्तेमाल 30 में से 4.3 दिन। एक महीने के नतीजे तक 42% टिके। मुख्य पूल की किसी स्टडी ने तीसरे पक्ष वाली पक्की आदत मुख्य नतीजे के रूप में नहीं नापी।

Risk of bias

Table 4. RoB 2 summary

Study	Randomisation	Deviations	Missing data	Measurement	Selection of result	Overall
Ly 2017	Some concerns	Some concerns	High	High	Some concerns	High
Foran 2024	Low	Some concerns	High	High	Low	Some concerns
Tong 2025	Low	Some concerns	Some concerns	High	Low	Some concerns
Cachia 2026	Low	Some concerns	Some concerns	High	Low	Some concerns
Kleinau 2024	Some concerns	Some concerns	High	High	Some concerns	High

हर ट्रायल ने तबीयत का नतीजा बिना आँख बंद किए खुद बताए फॉर्म से नापा। गायब डेटा का खतरा Foran (58% नुकसान) और Kleinau में भारी है। पूल नतीजे पर कुल खतरा “some concerns” है। ऊँचे की तरफ़ झुकाव है।

GRADE

Table 5. GRADE summary of findings

Outcome	k (n)	Effect	Certainty	Reasons
Wellbeing scales only	3 (1,658)	g = 0.07 (−0.06 to 0.19)	Low	Unblinded self-report (−1). Imprecision, CI includes zero (−1).
Wellbeing or positive affect, all controls	4 (2,144)	g = 0.17 (−0.01 to 0.35)	Low	Unblinded self-report (−1). Imprecision, CI includes zero (−1).
Wait-list or care-as-usual	3 (799)	g = 0.28 (0.14 to 0.42)	Low	Unblinded self-report (−1). Indirectness, no Indian working-adult sample (−1).
Versus active digital resources	1 (1,345)	g = 0.02 (−0.10 to 0.14)	Low	Single trial. High attrition.
Self-care behaviour	1 (285)	d = 0.36 at 10 days; not sustained at 1 month	Very low	Single trial, short dose, wait-list.
Adverse events	4	No serious events reported where looked for; most trials did not look	Very low	Absence of evidence.

Safety and adverse events

सेफ्टी सेक्शन ज़रूरी है। बातचीत एजेंट दुख बढ़ा सकते हैं। गलत सलाह दे सकते हैं। इंसान की मदद की जगह ले सकते हैं।

Sohn and colleagues ने लिखा कि 39 में से ज़्यादातर ट्रायल में नुकसान के इवेंट व्यवस्थित नहीं नापे गए। Li and colleagues ने लिखा कि 2023 के 35 में से सिर्फ़ 15 स्टडी ने सेफ्टी उपाय बताए। 2024 की एक रिव्यू में 171 में से सिर्फ़ 55 ऐप ट्रायल ने नुकसान के इवेंट लिखे। जहाँ नापा गया, बिगड़ने की दर करीब 7% रही।

इस रिव्यू की ट्रायल में Foran, Tong और Ly ने व्यवस्थित नुकसान-इवेंट प्रोटोकॉल नहीं लिखा। कोई गंभीर घटना नहीं लिखी गई। Cachia and colleagues ने आम छात्र सैंपल देखा। कंट्रोल के मुकाबले डिप्रेशन, चिंता या तनाव में क्लिनिकल उछाल नहीं लिखा। यह सेफ्टी वाला शून्य है। पूरा नुकसान-विश्लेषण नहीं है। Kleinau and colleagues ने तकनीकी फेल और भारी छूटना लिखा।

जेनेरेटिव एजेंट वाली बड़ी सेफ्टी घंटी 2025 की एक RCT से आती है। वह क्लिनिकल वयस्क सैंपल थी। पूल से बाहर है। उस ट्रायल में स्टाफ की 15 सेफ्टी इंटरवेंशन हुईं। मतलब साफ है। पक्का जेनेरेटिव एजेंट भी, जब डायग्नोसिस वाले लोगों के साथ चले, इंसान की निगरानी माँगता है।

गैर-क्लिनिकल प्रोग्राम को भी ये खतरे सोचने हैं। गलत दिलासा। संकट की भाषा न पहचानना। एजेंट पर निर्भर हो जाना। निजी फोन पर कंज़्यूमर मॉडल से डेटा लीक। और यह गलत यकीन कि तबीयत का कोच क्लिनिकल सेवा है। 2025–2026 की फ्रंटियर-मॉडल ऑडिट ने कमज़ोर यूज़र को चिंताजनक जवाब दिखाए।

भारतीय ऑफ़िस प्रोग्राम की न्यूनतम सेफ्टी यह है। लिखी संकट पगडंडी। जिस घंटे प्रोग्राम चलता दिखे, उस घंटे इंसान का हाथ। नियम कि एजेंट डायग्नोसिस न करे, विकार का नाम न ले, दवा की सलाह न दे। सेफ्टी इवेंट का लॉग। और Digital Personal Data Protection Act के तहत भारत में बैठा डेटा। ये डिज़ाइन फीचर सबूत की जगह नहीं लेते।

07 Discussion

ख़ास बात

सबसे बड़ी आम-वयस्क ट्रायल में लोग 30 में से 4.3 दिन बात की।

What the number means

तबीयत के स्केल पर गैर-क्लिनिकल वयस्कों का पूल अनुमान g = 0.07 है। अंतराल में शून्य है। अच्छा मूड जोड़ने पर आंकड़ा g = 0.17 हो जाता है। उस अंतराल में भी शून्य है। जो खरीदार g = 0.17 को पक्का फायदा समझे, वह नतीजा गलत पढ़ रहा है।

साथ के दो आंकड़े साथ पढ़ें।

पहला, Sohn and colleagues का गैर-क्लिनिकल डिप्रेशन अनुमान: g = 0.07 (95% CI −0.01 to 0.15)। जो बीमार नहीं, उनमें लक्षण पक्के नहीं घटते। यह बात अब स्थिर है।

दूसरा, तुलना का फाँक। वेट-लिस्ट या आम देखभाल के मुकाबले g = 0.28 (95% CI 0.14 to 0.42), I^2 = 0%। अच्छे वेब वेलनेस रिसोर्स के मुकाबले g = 0.02। AI कोच कुछ हफ्ते कुछ न करने से आगे है। सबसे बड़ी ट्रायल में वह एक अच्छे रिसोर्स पेज से आगे नहीं है।

यह पैटर्न डिजिटल सेहत में जाना-पहचाना है। नया चैनल तब तक चमकता है जब तक पुराने डिजिटल चैनल से तुलना न हो। व्हाट्सऐप पहुँच का फायदा है। नतीजे का अपने आप फायदा नहीं।

स्किल और आदत पर सबूत पतला है। एक समुदाय ट्रायल ने थोड़े दिन खुद की देखभाल और स्किल का इरादा बढ़ाया। एक महीने बाद फायदा पक्का नहीं रहा। एक ऑफ़िस जैसी ट्रायल ने पूरे करने वालों में लचीलापन-गतिविधि बढ़ाई। कामकाजी वयस्कों में टिकाऊ, बाहर से देखी आदत कोई ट्रायल नहीं दिखाती।

जुड़ाव पर सबूत एक जैसा और साफ है। तय डोज़ मिला डोज़ नहीं है। सबसे बड़ी आम-वयस्क ट्रायल में 30 में से 4.3 बात के दिन। एक महीने के नतीजे तक 42% टिके। जहाँ जाँचा गया, ज़्यादा दिन का मतलब ज़्यादा फायदा। जो AI कोच खुले ही नहीं, वह कोच नहीं।

जेनेरेटिव बनाम नियम-आधारित फर्क अभी तबीयत का मोडरेटर नहीं बनता। मुख्य पूल में एक ही जेनेरेटिव ट्रायल है। छह हफ्ते का डोज़। आम देखभाल कंट्रोल। अच्छा मूड बढ़ा। लक्षण के स्केल शून्य रहे। यह इस पेपर की थीसिस से मिलता है। जो बीमार नहीं, उनके लिए सही नतीजा तबीयत और प्रैक्टिस है। लक्षण घटाना नहीं। इससे ढाँचे को जादू नहीं माना जा सकता।

Limitations

k छोटा है। चार ट्रायल समावेशी आंकड़ा चलाती हैं। एक में n = 28 है। एक n = 1,345 और लगभग शून्य असर से वज़न दबाती है।

नतीजे बिना आँख बंद किए खुद बताए गए हैं। तुलना मिली-जुली है। डोज़ छोटा है। दस दिन से छह हफ्ते। ऑफ़िस प्रोग्राम चलती प्रैक्टिस बेचते हैं। ट्रायल चलती प्रैक्टिस नहीं हैं।

आबादी भारतीय कामकाजी वयस्क नहीं है। स्वीडन के स्वयंसेवक हैं। अमेरिका के ऑनलाइन वयस्क हैं। हांगकांग के समुदाय वयस्क हैं। अमेरिका के अंडरग्रैजुएट हैं। मलावी के स्वास्थ्य कर्मी हैं। चीन के प्रयोग वाले लोग हैं। भारतीय ऑफ़िस, फैक्ट्री या अस्पताल पर लागू करना अनुमान है। नतीजा नहीं।

सर्च और चुनाव तेज़ संश्लेषण था। बीज रिव्यू प्लस अपडेट। छपी प्रोटोकॉल वाली दोहरी लाइब्रेरियन स्क्रीन नहीं। इसलिए फ्लो आंकड़े लगभग हैं। कोई छोटी ट्रायल छूट सकती है। बड़ी भारतीय गैर-क्लिनिकल वयस्क ट्रायल नहीं छूटी। इंडेक्स में है ही नहीं।

नुकसान-इवेंट के तरीके कमज़ोर हैं। नुकसान न लिखना सेफ्टी का सबूत नहीं।

चलते पाठ में व्यावसायिक प्रॉडक्ट नाम नहीं लिखे। प्रॉडक्ट स्तर का हिसाब प्राथमिक पेपर में है।

इसका मतलब

What this means for an Indian workplace programme

गुरुग्राम के HR ऑफ़िस में रिन्यूअल मीटिंग सोचिए। वेंडर की स्लाइड सेल्स फ़्लोर के डिप्रेशन स्कोर गिराने का वादा करती है। पर वहाँ के लोग बीमार नहीं। गैर-क्लिनिकल वयस्कों में 2026 का लक्षण अनुमान g = 0.07 है। तबीयत-स्केल अनुमान भी g = 0.07 है। अच्छे रिसोर्स पेज के मुकाबले सबसे बड़ी ट्रायल ने अतिरिक्त फायदा नहीं पाया। यह गलत नतीजा खरीदना है।

ट्रायल का सहारा संकरा है। घर लौटती बस में इंदौर की स्टोर मैनेजर व्हाट्सऐप पर छोटा, बंधा चेक-इन खोलती है। कल का एक कदम चुनती है। ऐसी बात कुछ दिन से कुछ हफ्ते अच्छा मूड और स्किल का इरादा बढ़ा सकती है, अगर वह खोलती रहे। फायदा दिनों के हिसाब से चलता है। व्हाट्सऐप चलने लायक चैनल है। जेनेरेटिव बात छात्रों का ध्यान बाँधे रख सकती है। इनमें कुछ भी क्लिनिकल सेवा नहीं।

भारतीय प्रोग्राम के लिए डिज़ाइन साफ है। रेप्स गिनो, डायग्नोसिस नहीं। महीने की रिपोर्ट में दिखे कि इंदौर की मैनेजर ने कितने दिन प्रैक्टिस की। लेबल नहीं। स्किल गिनो और एक छोटी तबीयत पंक्ति नापो। लक्षण की लंबी बैटरी मत नापो, जब तक कर्मचारी खुद क्लिनिकल मदद न माँगे। दूसरे हफ्ते के बाद छूटना बजट में रखो, जब रिमाइंडर स्वाइप होने लगते हैं। एजेंट के पीछे इंसान का हाथ रखो। पहला संदेश भेजने से पहले संकट पगडंडी लिखो, ताकि चिंता वाला जवाब एक तय इंसान तक पहुँचे। इलाज वाले असर या भारत-विशेष असर का दावा मत करो। वह ट्रायल चली ही नहीं। हिंदी, तमिल, बांग्ला और दूसरी अनुसूचित भाषाएँ डिलीवरी की शर्त हैं। सबूत का आधार अभी नहीं। DPDP पालन और भारत में बैठा डेटा ज़रूरी है। वे नतीजा नहीं हैं।

Research gaps

भारत वाला गैप साफ है। गैर-क्लिनिकल भारतीय वयस्कों में AI बातचीत एजेंट की पूरी RCT नहीं है, जो तबीयत, अच्छा मूड, स्किल या आदत कंट्रोल के मुकाबले नापे। 2025 की एक हिंदी होमवर्क-मदद पायलट मध्य प्रदेश के गाँव में चली। n = 30 वयस्क। सबमें पहले से डिप्रेशन था। इंसान वाला व्यवहार-सक्रियण प्रोग्राम पहले से चल रहा था। यह क्लिनिकल संभावना स्टडी है। इस सवाल का जवाब नहीं। 2026 का दिल्ली यूनिवर्सिटी प्रोटोकॉल छात्रों वाला एजेंट है। नतीजे नहीं छपे। बेंगलुरु की ग्रोथ-माइंडसेट बातचीत-एजेंट स्टडी स्कूल बच्चों की है। वयस्कों की नहीं। CTRI पर chatbot, conversational agent और AI coach की खोज से पूरी गैर-क्लिनिकल वयस्क तबीयत ट्रायल नहीं निकली।

बाकी गैप पीछे-पीछे आते हैं। ऑफ़िस सैंपल लगभग नहीं। छह हफ्ते से लंबी अवधि लगभग नहीं। बाहर से देखी आदत लगभग नहीं। हिंदी क्लिनिकल पायलट के अलावा भारतीय भाषाएँ नहीं। गैर-क्लिनिकल वयस्कों में जेनेरेटिव बनाम नियम-आधारित आमने-सामने टेस्ट नहीं। नुकसान-इवेंट के मानक नहीं। इस आबादी में इंसान-गाइड बनाम बिना गाइड मुख्य पूल में नहीं।

मुख्य गैप बंद करने वाली ट्रायल जटिल नहीं। भारतीय कामकाजी वयस्क। डायग्नोसिस से न चुने। कम से कम दो भारतीय भाषाओं में बातचीत एजेंट। व्हाट्सऐप या वैसा चैनल। एक्टिव डिजिटल-रिसोर्स कंट्रोल और वेट-लिस्ट बाजू। 8–12 हफ्ते। मुख्य नतीजे: छोटी तबीयत पंक्ति, स्किल गिनती, एक दिखने वाली आदत। पहले से तय नुकसान इवेंट। स्टाफ वाला इंसान हाथ। भारत में बैठा डेटा। जब तक वह ट्रायल न आए, भारतीय प्रोग्राम अमेरिका, यूरोप, हांगकांग और अफ्रीका के सैंपल से अनुमान लगा रहे हैं।

08 FAQ

Q1 क्या AI कोच उन लोगों के लक्षण घटाते हैं जो बीमार नहीं हैं?

पक्की कमी नहीं। गैर-क्लिनिकल सैंपल में डिप्रेशन स्केल पर 2026 पूल अनुमान g = 0.07 है (95% CI −0.01 to 0.15)। अंतराल में शून्य है।

Q2 क्या ये तबीयत बढ़ाते हैं?

तबीयत के स्केल पर छपा पूल अनुमान g = 0.07 है (95% CI −0.06 to 0.19)। अच्छा मूड जोड़ने पर g = 0.17 है (95% CI −0.01 to 0.35)। दोनों अंतराल शून्य को छूते हैं या पास हैं। वेट-लिस्ट के मुकाबले अनुमान g = 0.28 है। अच्छे वेब रिसोर्स के मुकाबले g = 0.02 है।

Q3 क्या जेनेरेटिव कोच स्क्रिप्ट वाले से बेहतर है?

गैर-क्लिनिकल वयस्कों की तबीयत पर यह पक्का नहीं। मुख्य पूल में एक ही जेनेरेटिव ट्रायल है। उसने आम छात्रों में अच्छा मूड बढ़ाया। डिप्रेशन या चिंता नहीं।

Q4 लोग इन प्रोग्राम को सच में कितना चलाते हैं?

तय से कम। सबसे बड़ी आम-वयस्क ट्रायल में लोग 30 में से 4.3 दिन बात की। 42% ने एक महीने का नतीजा दिया। ज़्यादा दिन का मतलब ज़्यादा तबीयत फायदा।

Q5 क्या इसे भारतीय ऑफ़िस में चलाना सुरक्षित है?

जिन ट्रायल ने देखा, उन्होंने गंभीर घटना नहीं लिखी। ज़्यादातर ने देखा ही नहीं। स्टाफ वाला इंसान हाथ और लिखी संकट पगडंडी ज़रूरी है। यह कोचिंग प्रोग्राम है। क्लिनिकल सेवा नहीं।

09 References

1. Sohn JS, Ha BG, Park S, et al. Systematic review and meta analysis of chatbots in the management of depressive and anxiety symptoms. **npj Digital Medicine**. 2026;9:377. doi:10.1038/s41746-026-02566-w
1. Li H, Zhang R, Lee YC, Kraut RE, Mohr DC. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. **npj Digital Medicine**. 2023;6:236. doi:10.1038/s41746-023-00979-5
1. Zhang Q, Zhang R, Xiong Y, Sui Y, Tong C, Lin FH. Generative AI mental health chatbots as therapeutic tools: systematic review and meta-analysis of their role in reducing mental health issues. **Journal of Medical Internet Research**. 2025;27:e78238. doi:10.2196/78238
1. Heinz MV, Mackin DM, Trudeau BM, et al. Randomized trial of a generative AI chatbot for mental health treatment. **NEJM AI**. 2025;2(4). doi:10.1056/AIoa2400802
1. Foran HM, Kubb C, Mueller J, et al. An automated conversational agent self-help program: randomized controlled trial. **Journal of Medical Internet Research**. 2024;26:e53829. doi:10.2196/53829
1. Tong ACY, Wong KTY, Chung WWT, Mak WWS. Effectiveness of topic-based chatbots on mental health self-care and mental well-being: randomized controlled trial. **Journal of Medical Internet Research**. 2025;27:e70436. doi:10.2196/70436
1. Cachia JYA, Zhao X, Hunter J, Wu D, Lin E, De Freitas J. AI for proactive mental health: a multi-institutional, longitudinal randomized controlled trial. **NEJM AI**. 2026;3(8). doi:10.1056/AIoa2501293
1. Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. **Internet Interventions**. 2017;10:39-46. doi:10.1016/j.invent.2017.10.002

1. Kleinau E, et al. Effectiveness of a chatbot in improving the mental wellbeing of health workers in Malawi during the COVID-19 pandemic: a randomized, controlled trial. **PLOS ONE**. 2024;19(5):e0303370. doi:10.1371/journal.pone.0303370

1. Liu I, Liu F, Xiao Y, Huang Y, Wu S, Ni S. Investigating the key success factors of chatbot-based positive psychology intervention with retrieval- and generative pre-trained transformer (GPT)-based chatbots. **International Journal of Human–Computer Interaction**. 2025;41(1):341-352. doi:10.1080/10447318.2023.2300015

1. He Y, Yang L, Qian C, Li T, Su Z, Zhang Q, Hou X. Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. **Journal of Medical Internet Research**. 2023;25:e43862. doi:10.2196/43862

1. Feng X, Tian L, Ho GWK, Yorke J, Hui V. The effectiveness of AI chatbots in alleviating mental distress and promoting health behaviors among adolescents and young adults: systematic review and meta-analysis. **Journal of Medical Internet Research**. 2025;27:e79850. doi:10.2196/79850

1. Agrawal R, Monteiro K, Narasimhamurti NS, et al. A conversational agent (PracticePal) to support the delivery of a brief behavioral activation treatment for depression in rural India: development and pilot-testing study. **JMIR Formative Research**. 2025. doi:10.2196/67773

1. Schöne J, Salecha A, Lyubomirsky S, Eichstaedt JC, Willer R. Structured AI dialogues can increase happiness and meaning in life. PsyArXiv preprint. 2025. https://osf.io/preprints/psyarxiv/2bf7t_v1

1. Linardon J, et al. Systematic review of adverse event reporting in randomised trials of mental health apps. **npj Digital Medicine**. 2024;7:363. doi:10.1038/s41746-024-01376-2

1. Mokhtari Masoumi Alamdarloo S, Mirzakhani A, Azhdarloo M, et al. Effectiveness of AI and rule-based conversational agents for depression, anxiety and stress: a meta-analysis. **npj Digital Medicine**. 2026;9:650. doi:10.1038/s41746-026-02820-1

1. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent: a randomized controlled trial. **JMIR Mental Health**. 2017;4(2):e19. doi:10.2196/mental.7785

1. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence to relieve symptoms of depression and anxiety: randomized controlled trial. **JMIR Mental Health**. 2018;5(4):e64. doi:10.2196/mental.9782

1. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**. 2021;372:n71. doi:10.1136/bmj.n71

1. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. **BMJ**. 2019;366:l4898. doi:10.1136/bmj.l4898

1. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. **BMJ**. 2008;336(7650):924-926. doi:10.1136/bmj.39489.470347.AD

1. Higgins JPT, Thomas J, Chandler J, et al., eds. **Cochrane Handbook for Systematic Reviews of Interventions**. Version 6.4. Cochrane; 2023. https://training.cochrane.org/handbook

1. MindMitra trial registration. Care on Campus: understanding and responding to student wellbeing in India. AEA RCT Registry AEARCTR-0017634. 2026. https://www.socialscienceregistry.org/trials/17634

1. Using conversational AI to teach growth mindset skills to youths in India. ClinicalTrials.gov NCT07316647. https://clinicaltrials.gov/study/NCT07316647

लेखक के बारे में

लेखक के बारे में

डॉ. अतुल असवानी; मुंबई के साइकियाट्रिस्ट और ट्रेंड रिज़ल्ट्स कोच
डॉ. अतुल असवानी मुंबई के साइकियाट्रिस्ट और ट्रेंड रिज़ल्ट्स कोच हैं। वे मुंबई की Tranquilo Meditek Sytems Private Limited में क्लिनिकल और कोचिंग डिज़ाइन संभालते हैं। कंपनी इमोशनऐप बनाती है। काम है गिने जा सकने वाले रोज़ के रेप्स, व्हाट्सऐप-फर्स्ट डिलीवरी, और DPDP वाले इमोशनल फ़िटनेस प्रोग्राम। इमोशनऐप को क्लिनिकल सेवा नहीं माना जाता।

हवाला कैसे दें

Aswani A. No Symptoms to Reduce: A Meta-Analysis of AI Conversational Agents on Wellbeing Outcomes in Non-Clinical Adults. EmotionApp Research Paper 37 (Hindi edition). Mumbai: Tranquilo Meditek Sytems Private Limited; 2026. https://emotionapp.in/research/ai-conversational-agents-wellbeing-nonclinical-adults

अपडेट

22 September 2026

नोट

इमोशनऐप एक नॉन-क्लिनिकल इमोशनल फ़िटनेस कोचिंग प्लेटफ़ॉर्म है। यह पेपर प्रैक्टिस, तरीकों और कोचिंग के बारे में है। यह थेरेपी, इलाज, डायग्नोसिस या मेडिकल सलाह नहीं है।

© 2026 Tranquilo Meditek Sytems Private Limited, Mumbai · emotionapp.in

सुरक्षा नोट

अगर आप परेशानी में हैं, तो किसी योग्य प्रोफ़ेशनल से या Tele-MANAS (14416) पर बात करें।