

Generative AI in Psychiatry

Yulin Hswen, ScD, MPH
Associate Professor
Artificial Intelligence Interdisciplinary Institute
Computer, Mathematics, and Natural Sciences
Biostatistics and Epidemiology, School of Public Health
University of Maryland



Associate Editor | JAMA | **JAMA+ AI**
JAMA Network™ jamaai.org

The Landscape

The Promise

The Peril

Five Key Studies

The Road Ahead



art by Stevie Tierney.

What AI Does Well – Case Studies



24/7 Scalability

No waitlists. Conversational AI provides on-demand, personalized support to millions simultaneously.



Therapeutic Alliance

Patients rate AI warmth & competence comparably to human therapists (Shoshani et al., 2026).



Documentation Depth

AI scribes capture significantly more neuropsychiatric symptom detail across all 6 RDoC domains (Castro et al., 2026).



Clinical Outcomes

RCT shows AI intervention reduced anxiety more than face-to-face group therapy at 12 weeks.



*art by Stevie
Tierney*

Case Studies

Perlis et al. — JAMA Psychiatry 2026

Retrospective cohort · N=20,302 notes

AI scribes → more symptom documentation, LESS psychiatric intervention (aOR 0.83)

Maheux et al. — JAMA Network Open 2026

Cross-sectional passive sensing · N=6,488 youth

31.9% used GenAI; 50% of 15–17 yr olds; 7 of top 17 apps marketed as companions

Teferra et al. — JAMA Network Open 2026

Cross-sectional · 154 MI sessions, 3,706 turns

LLM contextually appropriate (DeepEval 0.72) but low semantic overlap with therapists (0.29)

Shoshani et al. — JAMA Network Open 2026

3-arm RCT · N=995 students · 12 weeks

AI > group therapy for anxiety & well-being; therapeutic alliance mediates outcomes

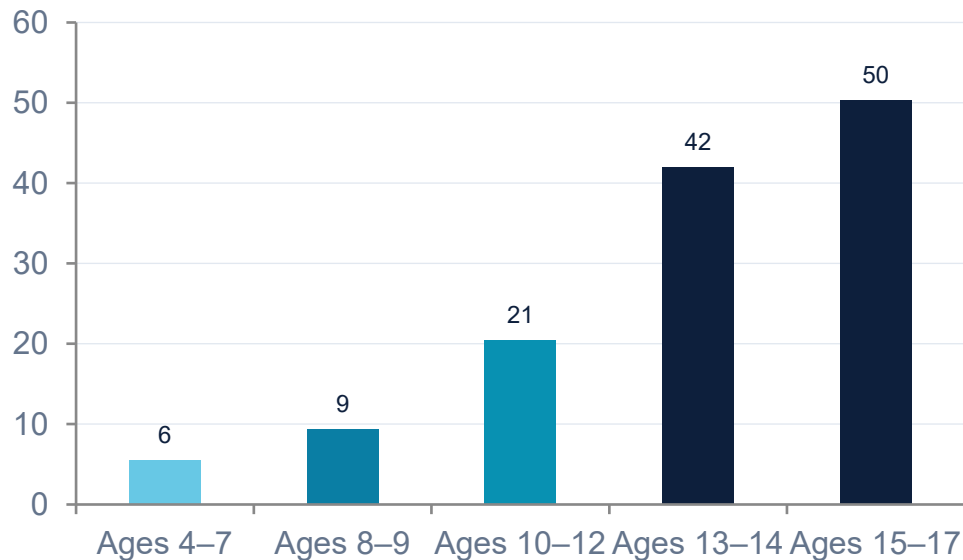
McBain et al. — JAMA Network Open 2025

National survey · N=1,058 · ages 12–21

13.1% (~5.4M youth) used GenAI for mental health; 92.7% found it helpful

Youth Using GenAI

GenAI Adoption by Age Group (Maheux et al., 2026)
JAMA network open



Methods

Looked at device activity from over 6,400 children and teens

Used data from a parental monitoring app

Findings

- Nearly 1/3 children and teens have GenAI on device
- 7 of 17 top apps marketed for companionship
- Heavy users: >40 min/day on GenAI
- 20% of preteens use during school hours
- Browser-based use NOT captured likely undercount

GenAI and Youth Mental Health

McBain et al., 2025 National Survey JAMA network open

Methods

Surveyed over US adolescents and young adults, ages 12–21
Random sampling from population frames of US households.

Black youth reported significantly lower perceived helpfulness of AI mental health advice (85% lower ODDS vs White non-Hispanic peers)

Findings

13.1%

of US youth (12–21 yrs)
used GenAI for mental health
~5.4 million youth

Usage Patterns

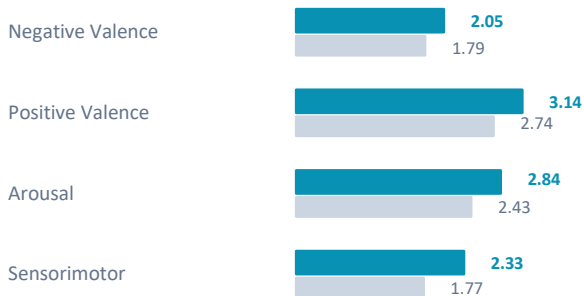
65% sought mental health advice monthly
93% said mental help advice was helpful

Scribes & Therapy Bots in Practice

Perlis et al. AI Scribes (JAMA Psychiatry 2026)

20,302 primary care annual visits with PHQ-9 score recorded
AI-generated notes
Human scribe notes
Regular notes without scribes
Used GPT-4o to analyze psychiatric symptom descriptions in the notes

Domain Scores (higher = more documented):



■ AI Scribe ■ Human Scribe

⚠ But depression interventions or diagnostic code: AI scribe 17% less ($p < .001$)

Teferra et al. LLM vs. Therapist in Motivational Interviewing (JAMA Network Open 2026)

154 therapy sessions · 3,706 therapist turns
Publicly available YouTube videos and online counseling demonstrations
GPT-4o evaluated with two metrics

Cosine Similarity
(Semantic overlap) **0.29** / 1.0



Low — different words,
similar intent

DeepEval Score
(Contextual fit) **0.72** / 1.0



Higher — generally
appropriate responses

Performance declines over longer sessions LLMs struggle with long-range coherence

The Promise

Shoshani et al., 2026 JAMA network open

RCT: 995 university students

12-week

- conversational AI
- group therapy
- waitlist

Anxiety reduction: **AI > Group therapy**

Well-being gains: **AI > Group therapy**

Depression: **AI > Waitlist**

PTSD symptoms: **No difference**

Therapeutic alliance → engagement → symptom improvement (SEM, $p < .001$)

People who felt emotionally connected to the AI (“therapeutic alliance”) used it more and improved more

Why is this important?

AI mental health tools could help address therapist shortages

Limitations

Participants were mostly college students
Short-term study

Critical Risks & Cautions



Documentation ≠ Action

AI scribes documented MORE symptoms but patients were 17% LESS likely to receive a psychiatric intervention.
(Castro et al., 2026)

Automation complacency?



Children in the Wild

9% of 8–9 year olds and 20% of preteens use GenAI, many unaware of age restrictions. Some apps marketed as companions.
(Maheux et al., 2026)

Uncharted developmental territory



Clinical Gap in Motivational Alignment

LLMs show contextual appropriateness (DeepEval 0.72) but low semantic similarity (cosine 0.29) to real therapists. Performance declines in longer sessions.
(Teferra et al., 2026)

Not yet clinically equivalent



Equity Concerns

Black youth reported significantly lower perceived helpfulness of AI mental health advice (OR 0.15 vs White peers).
(McBain et al., 2025)

Cultural competency gaps persist

A Framework for Responsible Integration



Rigorous Evidence

- Prospective RCTs with clinical populations
- Long-term follow-up beyond 3 months
- Real-world effectiveness, not just efficacy
- *Innovative methods



Safety & Equity

- Age-appropriate access controls
- Culturally competent AI responses
- Crisis escalation to human clinicians



Clinical Workflow

- Alert clinicians when AI flags symptoms
- Scribe documentation → action linkage
- AI as adjunct, not replacement



Policy & Oversight

- Regulatory frameworks for therapeutic AI
- Transparency in training data
- Mandatory clinical validation before deployment

GenAI in psychiatry is not a question of IF it's a question of how safely and ethically, we integrate it.

1 AI Scribes in Primary Care (Psychiatric Documentation and Management)

What was studied

>20,000 primary care visit notes comparing AI scribe, human scribe, and no scribe.



Key findings

- AI scribe notes contained more detailed documentation of psychiatric symptoms.
- BUT** patients were less likely to receive treatment or referrals (e.g., fewer antidepressant prescriptions, fewer psychiatric referrals).



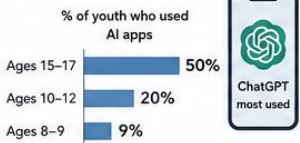
Takeaway

AI improves note completeness but does not automatically improve mental health care.

2 Generative AI Use Among US Youth

What was studied

Digital activity data from >6,400 children and teens (ages 4–17) using generative AI apps.



Key findings

- Overall, 32% of youth used AI apps.
- Use increases with age (50% of teens 15–17).
- Most use briefly, but a small subset uses AI heavily (>40 min/day; some up to ~3 hours/day).
- Usage occurs mainly after school.

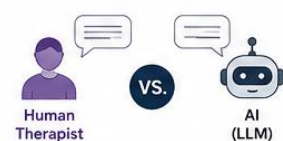


Takeaway AI tools are already widely used by youth; long-term impacts on development and mental health are still unknown.

3 Alignment of LLM Responses With Human Therapists (Motivational Interviewing)

What was studied

3,700+ therapist responses in motivational interviewing sessions compared with AI-generated responses (GPT-4o).



Key findings

- AI responses were often contextually appropriate but not very similar in exact wording to therapists.
- Performed better when conversations stayed focused on one topic.
- Performance declined slightly in longer conversations.



AI can imitate therapeutic communication to some extent but still lacks consistency and deep understanding.

4 Efficacy of a Conversational AI Agent for Psychiatric Symptoms (Randomized Clinical Trial)

What was studied

995 university students randomized to: AI chatbot therapy vs. in-person group therapy vs. waiting list (12 weeks).



Key findings

- AI group showed greater reductions in anxiety and depression vs. control.
- Greater improvement in well-being vs. both group therapy and control.
- Life satisfaction improved more than control.
- No significant differences in PTSD symptoms.
- Stronger “therapeutic alliance” with AI was linked to more engagement and better outcomes.



Takeaway A conversational AI chatbot improved several mental health outcomes, showing promise as a scalable support tool.

5 Use of Generative AI for Mental Health Advice Among US Youth

What was studied

National survey of >1,000 US youths ages 12–21.



Key findings

- 13% overall used AI for mental health advice.
- 22% of ages 18–21 used AI for this purpose.
- Most users sought advice monthly or more.
- 93% found the advice somewhat or very helpful.
- Black respondents were less likely to find advice helpful.



Takeaway Many youth already turn to AI for emotional support. Safety, accuracy, and equity are key concerns.

Common Themes Across Studies



Overall Takeaways

- ✓ AI is becoming part of mental health care and support—both in clinics and in daily life.
- ✓ It can enhance documentation and improve symptoms for some users.
- ✓ Young people are already using AI for emotional support.
- ✓ More research is needed on long-term effects, safety, and how to ensure AI complements—not replaces—quality human care.



AI has great potential, but human oversight, ethical design, and evidence-based use are essential.

The Bottom Line

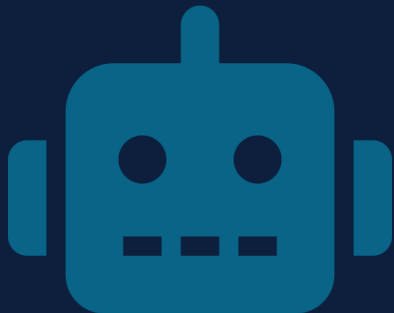
1 GenAI is already being used for mental health support, by millions of youth, right now.

2 There is real therapeutic efficacy: AI outperformed group therapy on anxiety in a RCT.

3 But documentation $\neq$ action: AI scribes may be creating a false sense of 'covered' care.

4 LLMs are contextually appropriate but not yet clinically equivalent to trained therapists.

5 Equity gaps demand attentions as benefits are not distributed equally across groups.



GenAI will not replace psychiatrists (it's not radiology) BUT the GAP is real, and AI is being used to fill it

Yulin Hswen, ScD, MPH

Thank you

ASCP

Valentina Mantua

Roy Perlis

John Torous

Joseph Geraci