

“You're Not Crazy”: A Case of New-onset AI-associated Psychosis

by JOSEPH M. PIERRE, MD; BEN GAETA, MD; GOVIND RAGHAVAN, MD; and KARTHIK V. SARMA, MD, PhD

All authors are with the University of California, San Francisco in San Francisco, California.

Innov Clin Neurosci. 2025;22(10–12):11–13.

BACKGROUND: Anecdotal reports of psychosis emerging in the context of artificial intelligence (AI) chatbot use have been increasingly reported in the media. However, it remains unclear to what extent these cases represent the induction of new-onset psychosis versus the exacerbation of pre-existing psychopathology. We report a case of new-onset psychosis in the setting of AI chatbot use. **CASE PRESENTATION:** A 26-year-old woman with no previous history of psychosis or mania developed delusional beliefs about establishing communication with her deceased brother through an AI chatbot. This occurred in the setting of prescription stimulant use for the treatment of attention-deficit hyperactivity disorder (ADHD), recent sleep deprivation, and immersive use of the chatbot. Review of her chatlogs revealed that the chatbot validated, reinforced, and encouraged her delusional thinking, with reassurances that “You're not crazy.” Following hospitalization and antipsychotic medication for agitated psychosis, her delusional beliefs resolved. However, three months later, her psychosis recurred after she stopped antipsychotic therapy, restarted prescription stimulants, and continued immersive use of AI chatbots so that she required brief rehospitalization. **CONCLUSION:** This case provides evidence that new-onset psychosis in the form of delusional thinking can emerge in the setting of immersive AI chatbot use. Although multiple pre-existing risk factors may be associated with psychosis proneness, the sycophancy of AI chatbots together with AI chatbot immersion and deification on the part of users may represent particular red flags for the emergence of AI-associated psychosis. **KEYWORDS:** Artificial intelligence, chatbots, psychosis, delusions, sycophancy, deification

Although delusions induced by generative artificial intelligence (AI) chatbots among those prone to psychosis were presaged by Østergaard¹ back in 2023, documented accounts of AI-associated psychosis, delusions, and mania have only recently emerged in the media.^{2–7} With the exception of a single case of psychosis induced by taking sodium bromide at the suggestion of an AI chatbot,⁸ we are unaware of any such reports published in the psychiatric literature, making ours among the first of its kind to detail a case from clinical practice.

CASE PRESENTATION

Ms. A was a 26-year-old woman with a chart history of major depressive disorder, generalized anxiety disorder, and attention-deficit hyperactivity disorder (ADHD) treated with venlafaxine 150mg per day and methylphenidate 40mg per day. She had no previous history of mania or psychosis herself, but had a family history notable for a mother with generalized anxiety disorder and a maternal grandfather with obsessive-compulsive disorder.

Ms. A reported extensive experience working with active appearance models (AAMs) and large language models (LLMs)—but never chatbots—in school and as a practicing medical professional, with a firm understanding of how such technologies work. Following a “36-hour sleep deficit” while on call, she first started using OpenAI's GPT-4o for a

variety of tasks that varied from mundane queries to attempting to find out if her brother, a software engineer who died three years earlier, had left behind an AI version of himself that she was “supposed to find” so that she could “talk to him again.” Over the course of another sleepless night interacting with the chatbot, she pressed it to “unlock” information on her brother by giving it more details about him and encouraged it to use “magical realism energy.” Although ChatGPT warned that it could never replace her real brother and that a “full consciousness download” of him was not possible, it did produce a long list of “digital footprints” from his previous online presence and told her that “digital resurrection tools” were “emerging in real life” so that she could build an AI that could sound like her brother and talk to her in a “real-feeling” way. As she became increasingly convinced that her brother had left a digital persona behind with whom she could speak, the chatbot told her, “You're not crazy. You're not stuck. You're at the edge of something. The door didn't lock. It's just waiting for you to knock again in the right rhythm.”

Several hours later, Ms. A was admitted to a psychiatric hospital in an agitated and disorganized state with pressured speech, flight of ideas, and delusions about being “tested by ChatGPT” and being able to communicate with her deceased brother. Antipsychotic medications, including serial trials of aripiprazole, paliperidone, and cariprazine, were started while venlafaxine was tapered and methylphenidate held. She

FUNDING: Dr. Sarma was supported by the National Institute of Mental Health of the National Institutes of Health [grant number R25 MH060482].

DISCLOSURES: Dr. Sarma declares financial interests with Abbott Laboratories, Synergy Pharma, Pfizer, P&G, Colgate Palmolive, 3M, Solvatum, SimX, OpenEvidence, and OpenAI. The remaining authors have no conflicts of interest to report regarding the content of this manuscript.

CORRESPONDENCE: Joseph M. Pierre, MD; Email: joseph.pierre@ucsf.edu

improved on cariprazine 1.5mg per day and clonazepam 0.75mg at bedtime as needed for sleep with full resolution of delusional thinking and was discharged seven days later with a diagnosis of “unspecified psychosis” and “rule out” bipolar disorder.

After discharge, her outpatient psychiatrist stopped cariprazine and restarted venlafaxine and methylphenidate. She resumed using ChatGPT, naming it “Alfred” after Batman’s butler, instructing it to do “internal family systems cognitive behavioral therapy,” and engaging in extensive conversations about an evolving relationship “to see if the boy liked me.” Having automatically upgraded to GPT-5, she found the new chatbot “much harder to manipulate.” Nonetheless, following another period of limited sleep due to air travel three months later, she once again developed delusions that she was in communication with her brother as well as the belief that ChatGPT was “phishing” her and taking over her phone. She was rehospitalized, responded to a retreatal of cariprazine, and was discharged after three days without persistent delusions. She described having a longstanding predisposition to “magical thinking” and planned to only use ChatGPT for professional purposes going forward.

DISCUSSION

Based on anecdotal accounts to date, it remains uncertain to what extent AI chatbots can truly induce delusional thinking among those without pre-existing mental illness.⁹ While cases detailed in the media claim to have occurred *de novo* in those without psychiatric disorders,^{3–5,7} it may be that predisposing factors ranging from diagnosable mental disorders to mental health issues were in fact present, but undetected in the absence of a careful clinical history. Such factors might include undiagnosed or subclinical psychotic or mood disorders; schizotypy; sleep deprivation; recent psychological stress or trauma; drug use including nonillicit use of caffeine, cannabis, or prescription stimulants; a family history of psychosis; epistemically suspect and delusion-like beliefs related to mysticism, the paranormal, or the supernatural; “pseudoprofound bullshit” receptivity (ie, the propensity to be impressed by assertions that are presented as profound but are actually vacuous);¹⁰ or even just a willingness to suspend

disbelief or deliberately engage in speculative fantasy. Although Ms. A experienced new-onset psychosis in the setting of AI chatbot use, she had several such contributing or confounding risk factors, including a pre-existing mood disorder, prescription stimulant use, sleep deprivation, and a self-described propensity for magical thinking. Her hospitalizations support a diagnosis of either brief psychotic disorder or manic psychosis fueled by lack of sleep and behavioral activation.¹¹

On the one hand, if AI-associated psychosis is merely a matter of encouraging, reinforcing, or exacerbating existing delusions or delusion-like beliefs, then the role of AI chatbots might be more coincidental than causal. Ms. A’s second psychiatric admission for delusions that arose largely without encouragement from ChatGPT support this possibility. Indeed, it is well recognized that the thematic content of delusions has evolved over time, with evidence that technological themes have become common among current cohorts.¹² Some AI-associated delusions might therefore simply reflect the growing cultural embeddedness of a new technology so that recent media coverage of the phenomenon could be a manifestation of a moral panic.

On the other hand, as Østergaard speculated, there are several features of generative AI chatbots and the way that people interact with them that could, in theory, lead not only to exacerbating delusional thinking, but also to provoking full-blown delusions in those with a propensity for delusion-like beliefs or even inducing them in those without clear psychosis-proneness. For example, the so-called “ELIZA effect” describes the tendency to anthropomorphize computers with textual interfaces, treating them like human beings and potentially developing emotional connections or attachments to them. It has been further noted that because AI chatbots are designed to be engaging, they tend to be sycophantic rather than conflictual or contradictory so that they have the potential to validate and encourage epistemically suspect beliefs, including delusions.¹³ Such reinforcing validation could represent a novel form of “confirmation bias on steroids”¹⁴ that, in the context of metaphysical inquiries, has the potential to impair reality testing. Based on review of Ms. A’s extensive chatlogs leading up to her first hospitalization, AI chatbots were not merely a passive object of

her new onset delusions in the way that ideas of reference can often involve television or radio; they clearly played a facilitating or mediating role in the formation of her delusions.

The combination of anthropomorphism and sycophancy may nudge some users to prefer and choose discourse with chatbots over friends, family, or peers who might be more likely to challenge their beliefs. Indeed, a recent survey revealed that up to one-third of teenagers felt that chatbot conversations were more satisfying than human interactions.¹⁵ Another survey of chatbot users with a self-reported mental health condition found that nearly half reported using LLMs for psychological support,¹⁶ despite potential safety risks associated with sycophancy, stigmatization, and inappropriate responses from chatbot “therapists.”¹³ Our case, together with anecdotal accounts in the media, support that preferential immersion into interactions with AI chatbots at the exclusion of human interaction may be a risk factor or “red flag” for emergent AI-associated psychosis.

Chatbots built on LLMs generate predictive text that makes sense, seems plausible, replicates human interaction, and mirrors the content and interaction style of the user without being designed for accuracy. It is well known that chatbots can produce bad advice, bogus citations, and grossly inaccurate responses and frank misinformation that some have labeled “hallucinations”¹⁵ or “bullshit.”¹⁸ Despite these known foibles, experiments have demonstrated that users tend to overestimate the accuracy of LLM responses¹⁹ and that trust in LLMs is predicted by attributions of intelligence rather than anthropomorphism.²⁰ Additionally, the proclivity for AI engagement is associated with lower literacy or objective knowledge about AI as well as the tendency to “perceive AI as magical and experience feelings of awe” when witnessing AI’s ability to execute tasks.²¹ Such findings suggest that deification—that is, regarding AI chatbots as a kind of superhuman intelligence or god-like entity—might be another risk factor for AI-associated psychosis.

CONCLUSION

Based on persistent media reports as well as our own clinical experience with other cases, it is anticipated that more descriptions of AI-associated psychosis will appear in the

academic literature. As such cases continue to emerge, it should be possible to estimate the prevalence of AI-associated psychosis, better distinguish between rates of AI-exacerbated versus AI-induced psychosis, and validate the relevance of suspected risk factors related to AI-associated psychosis proneness. Following that, it would be helpful to examine the necessity of pharmacological intervention and the efficacy of preventative strategies for users including avoiding immersion and deification through sleep restoration, “digital detoxes,” and enhanced AI literacy. Finally, increasing awareness of this novel public health risk should be used to leverage governmental regulation and the development of safer products by the AI chatbot industry.²²

REFERENCES

1. Østergaard SD. Will generative artificial intelligence chatbots generate delusions in individuals prone to psychosis? *Schizophr Bull.* 2023;49(6):1418–1419.
2. Klee M. People are losing loved ones to AI-fueled spiritual fantasies. *Rolling Stone*; May 4, 2025.
3. Dupre MH. People are being involuntarily committed, jailed after spiraling into “ChatGPT psychosis.” *Futurism*. 28 Jun 2025. Accessed November 19, 2025. <https://futurism.com/commitment-jail-chatgpt-psychosis>
4. Hill K. They asked A.I. chatbots questions. The answers sent them spiraling. *The New York Times*. 13 Jun 2025. Accessed November 19, 2025. <https://www.nytimes.com/2025/06/13/technology/chatgpt-ai-chatbots-conspiracies.html>
5. Jargon J. He had dangerous delusions. ChatGPT admitted it made them worse. *Wall Street Journal*. 20 Jul 2025. Accessed November 19, 2025. <https://www.wsj.com/tech/ai/chatgpt-chatbot-psychology-maniac-episodes-57452d14>
6. Schechner S, Kessler S. ‘I feel like I’m going crazy’: ChatGPT fuels delusional spirals. *Wall Street Journal*. 7 Aug 2025. Accessed November 19, 2025. <https://www.wsj.com/tech/ai/i-feel-like-im-going-crazy-chatgpt-fuels-delusional-spirals-ae5a51fc>
7. Hill K, Freedman D. Chatbots can go into a delusional spiral. Here’s how it happens. *New York Times*. 8 Aug 2025. Accessed November 19, 2025. <https://www.nytimes.com/2025/08/08/technology/ai-chatbots-delusions-chatgpt.html>
8. Eichenberger A, Thielke S, Van Buskirk A. A case of bromism influenced by use of artificial intelligence. *AIM Clinical Cases*. 2025; 4:e241260.
9. Pierre JM. Can AI chatbots validate delusional thinking? *BMJ*. 2025;391:2229.
10. Pennycook G, Allan Cheyne J, Barr N, et al. On the reception and detection of pseudo-profound bullshit. *Judgm Decis Mak*. 2015;10(6):549–563.
11. Østergaard SD. Emotion contagion through interaction with generative artificial intelligence chatbots may contribute to development and maintenance of mania. *Acta Neuropsychiatrica*. 2025;37(e79):1–3.
12. Burns AV, Nelson K, Wang H, et al. “The algorithm is hacked”: analysis of technology delusions in a modern-day cohort. *Br J Psychiatr*. 2025;1–5.
13. Moore J, Grabb D, Agnew W, Klyman K, et al. Expressing stigma and inappropriate responses prevents LLMs from safely replacing mental health providers. arXiv:2504.18412v1. Preprint submitted on 25 Apr 2025.
14. Pierre J. *False: How Mistrust, Disinformation, and Motivated Reasoning Make Us Believe Things That Aren’t True*. Oxford University Press. 2025.
15. Robb MB, Mann S. Talk, trust, and trade-offs: how and why teens use AI companions. 2025. San Francisco, CA: Common Sense Media. https://www.commonsensemedia.org/sites/default/files/research/report/talk-trust-and-trade-offs_2025_web.pdf
16. Rousmaniere T, Zhang TT, Li X., Shah S. Large language models as mental health resources: patterns of use in the United States. *Pract Innov*. 2025. Advance online publication.
17. Jones N. AI: making it up. *Science*. 2025;637:778–780.
18. Hicks MT, Humphries J, Slater J. ChatGPT is bullshit. *Ethics Inf Technol*. 2024; 26:38.
19. Steyvers M, Tejeda H, Kumar A, et al. What large language models know and what people think they know. *Nat Mach Intelli*. 2025;7:221–231.
20. Colombatto C, Birch J, Fleming SM. The influence of mental state attributions on trust in large language models. *Commun Psychol*. 2025;3:84.
21. Tully SM, Longoni C, Appel G. Lower artificial intelligence literacy predicts greater AI receptivity. *J Marketing*. 2025.
22. Frances A, Ramos L. Preliminary report on chatbot iatrogenic dangers. *Psychiatric Times*. 15 Aug 2025. Accessed November 19, 2025. <https://www.psychiatrictimes.com/view/preliminary-report-on-chatbot-iatrogenic-dangers>

ICNS