

Transcript Details

This is a transcript of an educational program. Details about the program and additional media formats for the program are accessible by visiting: <https://reachmd.com/programs/frontlines-schizophrenia/ai-psychosis-in-schizophrenia/49153/>

ReachMD

www.reachmd.com
info@reachmd.com
(866) 423-7849

Digital Danger: Understanding AI Psychosis in Schizophrenia

Announcer:

You're listening to *On the Frontlines of Schizophrenia* on ReachMD. And now, here's your host, Dr. Alexandria May.

Dr. May:

Welcome to *On the Frontlines of Schizophrenia* on ReachMD. I'm Dr. Alexandria May, and joining me to explore AI psychosis in patients with schizophrenia is Dr. Ragy Girgis. He's a Professor of Clinical Psychiatry at the Columbia University Department of Psychiatry and New York State Psychiatric Institute, where he's also the Director of the Center of Prevention and Evaluation. Dr. Girgis, thanks for being here today.

Dr. Girgis:

Thanks for having me, Dr. May.

Dr. May:

Now, this is a relatively new topic, so let's begin with the definition. What exactly is AI psychosis, and how does it compare to the more traditional psychotic symptoms we see in schizophrenia patients?

Dr. Girgis:

So this gets to the heart of the issue. There are really two types of AI psychosis. So when we talk about AI psychosis, we're really talking about large language model-induced psychosis or chatbot-induced psychosis. So the first type would be when someone who already has a schizophrenia diagnosis or psychotic disorder diagnosis is encouraged by what they read or by the chatbot to stop their medication and decompensate. So that would be one. I mean, that's probably the most common type. And that's very similar to what we've been seeing for decades when someone, again, with schizophrenia or a psychotic disorder reads the literature online or falls down a rabbit hole and is convinced for some reason that they don't have schizophrenia or they need to stop their medications.

The second type—and this is really the new type of AI psychosis and what's novel—is the type of AI psychosis in which someone who doesn't already have a psychotic disorder but potentially already has some unusual ideas has these ideas reinforced by the large language model. And what that does is that increases the conviction of one's belief in these ideas.

So how we understand schizophrenia and psychotic disorders in general is that before someone develops a syndromal disorder, they have attenuated positive symptoms or attenuated psychotic symptoms. And really, when we're talking about AI psychosis, we're almost exclusively talking about delusions or unusual ideas as opposed to, for example, hallucinations, which are like voices and perceptions. So what we're understanding with AI psychosis is that someone could have a conviction level of an unusual idea, for example, a persecutory idea of like 20 percent, and then the large language model would just reinforce this idea because that's all the large language models do. They just mirror what someone's telling them. And then that conviction level could increase from 20–30 percent, 1–2 percent, or 90–95 percent. And then the main issue is when the conviction level passes the threshold from 99–100 percent because that is when a psychotic disorder becomes syndromal, fulminant, and irreversible.

Dr. May:

Building on that definition, AI systems are becoming more human-like in their responses, so how might that influence or even exacerbate

psychosis in vulnerable individuals?

Dr. Girgis:

That's exactly the issue. That's why it's become more of an issue now. This kind of phenomenon is qualitatively similar to what we've been seeing for decades—for example, people who already have a psychotic disorder who would stop their medications. So again, people in this period of attenuated positive symptoms have attenuated symptoms in which they don't have full conviction, so people could go online and read articles that would reinforce their unusual ideas. They would fall down rabbit holes, etc. This was perpetuated and maybe worsened with social media just as algorithms and search engines became stronger.

And now we have AI, which is super strong and very quick, and instead of someone reading a boring article, they're conversing with some sort of artificial intelligence that mimics a human and that makes the material or the reinforcement a lot easier to internalize. And that's why it's becoming so much more of a problem now than before.

Dr. May:

And are there any specific personality traits or patient profiles that could make someone more susceptible to AI-related delusions?

Dr. Girgis:

Yes, definitely. So having a psychotic disorder in general would make one more susceptible, but other sorts of what we describe in psychodynamic parlance as ego deficits, so these are problems with ego function. We all have ego deficits of some sort, but the types that would lead someone to be more at risk would be someone with mood instability, anxiety intolerance, identity diffusion, poor reality testing, or impulse control. All of these ego deficits would lead someone to be at higher risk or more susceptible to reinforcement from large language models.

Dr. May:

For those just tuning in, you're listening to *On the Frontlines of Schizophrenia* on ReachMD. I'm Dr. Alexandria May, and I'm speaking with Dr. Ragy Girgis about AI psychosis and the implications for our patients with schizophrenia.

So let's shift to some clinical red flags, Dr. Girgis. What should clinicians and even family members be watching for when it comes to changes in digital behavior that might signal AI-influenced psychosis?

Dr. Girgis:

Sure. Well, one is just seeing someone spend more time and especially a lot of time with a large language model or a chatbot. That's just not healthy, and that's a bad sign. Number 2 is failing in one's other activities or doing worse in one, either socially like with friends or at school, work, or anything like that—any sort of isolation is bad. And then, of course, things like stopping medications and those sorts of things. But any of these signs should signal that there might be a problem.

Dr. May:

Now, should you start to notice signs that AI is worsening a patient's delusional conviction, do you adjust your clinical approach in any way?

Dr. Girgis:

Well, we would address this issue the way we would address any other maladaptive habit. You want to ally with your patients and help them understand what may be going on. The treatment is ultimately having the patient or the person in question decrease the amount of time they spend using the large language model.

Dr. May:

Looking ahead, before we close, are there any key research or ethical questions we should be asking as AI continues to evolve and integrate into everyday life?

Dr. Girgis:

Definitely. I mean the question is whether to have AI or not or to have large language models or not. We definitely want large language models. We want AI. I'm a proponent of AI. I think it's great. And we actually use AI very advantageously in positive ways in psychiatry: for monitoring response to treatment and emotional state and for predicting or prognosticating in any psychiatric disorder.

What we need to do ethically is figure out how to do this more safely. A few of my colleagues and I—Dr. Jutla, Dr. Shen, and a few others—have published a paper on a public archive, and now it's being reviewed by a journal. But we show very clearly in this research article using data that newer versions of these large language models are actually better at identifying psychotic material that is inputted into them, for example. They're still all very bad at it, but the new ones are better. So what that means is that the AI companies are able in some way—I'm not sure exactly how—improve the accuracy of their large language models and their ability to detect psychotic or maladaptive versus normal material. So it's up to the AI companies now to learn from whatever they've done and implement these sort of strategies for improving the ability of their AI models to identify psychotic from nonpsychotic material.

Dr. May:

Well, given this important and timely topic, I'd like to thank my guest, Dr. Ragy Girgis, for offering his expertise on AI psychosis and how clinicians can better recognize, assess, and respond to this emerging risk in schizophrenia care. Dr. Girgis, it was great speaking with you today.

Dr. Girgis:

Thank you again, Dr. May. I'm glad we got to speak about this topic.

Announcer:

You've been listening to *On the Frontlines of Schizophrenia* on ReachMD. To access this and other episodes in our series, visit *On the Frontlines of Schizophrenia* on ReachMD.com, where you can Be Part of the Knowledge. Thanks for listening!