



A PODCAST
BY PHYSICIANS
FOR PHYSICIANS

HOSTED BY DR. DAVID GRATZER

November 1, 2023

ChatGPT and mental health care with Dr. John Torous

Edited for grammar and clarity by CAMH

[Musical intro]

Running time: 26:30

David Gratzter: It's the most downloaded app in history. ChatGPT has caused a stir. People use it to write their resumes, plan dinner and help with college term papers. But what are the implications for mental health? Will AI change our work? Should we be retraining?

My name is Dr. David Gratzter. I'm a psychiatrist here at CAMH and welcome to *Quick Takes*, a podcast series by Physicians for Physicians. Joining us today is a returning guest to weigh in on AI and ChatGPT. Dr. John Torous is Director of Digital Psychiatry Division at Beth Israel Deaconess Medical Centre. Dr. Torous is very active in research and is the author of some 300 or more papers, as well as five book chapters. He serves as editor in chief for the journal *JMIR Mental Health* and is the Web editor for *JAMA Psychiatry*. Welcome, John.

John Torous: Thank you so much for having me back.

David Gratzter: Well, we're always delighted to hear your thoughts. And now we're picking something very timely and topical. ChatGPT. Have you played with ChatGPT?

John Torous: So I have played with ChatGPT and I'll disclose that I am red teaming ChatGPT for OpenAI. I'm not compensated for it, I'm not paid, and we actually haven't started it, but our team is going to try to test it and we have to sign a non-disclosure that whatever we test we can't talk about. But I think we'll at least try to see if we can see what happens when we play a bit on a testing model.

David Gratzter: Without going into that work, which you can't talk about anyway, what are some of the things you've tried to do with ChatGPT, and what's been your experience?

John Torous: Yeah, our team actually bought one of the early licenses and subscriptions to ChatGPT so we could play with it, so we could learn about it. We've asked it to solve different DSM cases. The DSM actually publishes a case series book and we like to put the cases in and see what it responds. We've asked it to do drug-drug interactions. We've asked it for personal things to do to feel better. We've asked it for sleep advice. We've certainly played with it a lot. We have not put patient data into it ever. One should not put patient data into it at this time. There's even American Psychiatric Association directive basically saying, please don't put any PHI in, of course. ChatGPT, when you open it up, says, do not put patient information, do not ask me for medical advice. But overall, it's been pretty impressive in how we've played with it. It's been, overall, I think – sometimes I'm known for having a cynical take on things – but overall, I think it's been very exciting to play with and learn.

David Gratzer: Lots to process in your last comment. Let's break it down. The case series. So just to remind our listeners, the DSM series includes some cases that they publish, and then they include diagnosis and some comment. How did ChatGPT do with diagnosis?

John Torous: Pretty well. I think, overall it's able to spit out a diagnosis. I think, of course, that doesn't mean anyone's going to lose their job to ChatGPT. Wikipedia also has a lot of information. You can probably Google a lot of things. And if you say: "I'm feeling tired and more sleepy and low appetite and a negative mood", you can probably even Google search it, and it's going to come out with depression. So I don't think it means... it doesn't transform the practice of medicine to get a diagnosis right. A lot of what clinical care about is helping people manage diagnosis and be on the path towards recovery. Often patients come in and say, "Doctor Torous, I'm depressed." And after an assessment I will agree. Or they will I say: "You know, I'm really anxious, I have panic attacks." And again, sometimes there's other reasons. We do an exam, but the diagnosis is often relatively sometimes the easier part. Management is much harder.

David Gratzer: Have you explored ChatGPT's ability to give advice to patients, or 'hypothetical patients,' I suppose?

John Torous: So. We've certainly put in use cases of kind of "what would you think, or recommend, or do?" But a lot of what I've noticed actually is a lot of times it says: "You cannot rely on me for medical advice." It's actually changed, I think since Christmas. We've been playing a bit now and have used a version that often says: "That sounds like something that you should get clinical advice with, or you should seek a professional". So, I do think they're actively updating it to kind of have it not give as much information, especially when you go for management. Even when you kind of put in a case example around diagnosis, it very clearly frames it and says, "I'm a program. I don't know a lot of things. If this is about clinical – first, don't do it – and actually talk to someone who knows what they're talking about."

David Gratzer: Drug-drug interactions, though I suspect it would comment on.

John Torous: It actually gives you a lot of information, right? Because you just think about all those different receptors we learned about. Some patients are on really complex medication regimens. And you can imagine there must be a table, right, where you could kind of say, well, it hits this receptor, this receptor. So I think cases like that, it certainly seems to give information that would be actionable and useful, especially if you could verify it. It could quickly kind of look up all those tables and figure out what could be a possible interaction. I mean that had Epic has it too, right? Medical records have this also when you prescribe, depending on what electronic medical record you use. So it's not the first time a computer can help us kind of sort through complex information.

David Gratzer: True, though the ability to access this would be different, right? So, if you've got a patient portal in the organisation that is helping you with your health care subscribes to, that sort of service perhaps, but many patients wouldn't have that. Whereas ChatGPT after you ask ChatGPT for tips on what to make for dinner based on the contents of your kitchen, you might ask them for tips on what you can pair with your SSRI. How impressed are you?

John Torous: I think the ease of getting information is very impressive, right? The ease of that modality and the format of chat I think makes it very intuitive and easy to access. I think not only for psychiatrists, I think for patients to look up information. It's a really nice way to personalise content for people. As we know, it's really only trained up to 2021 data and it has not been trained on explicitly medical sources or information at this point. So, I think we have to take into account that it's not working the latest data. It hasn't really trained on a lot of PubMed papers. It hasn't read your newsletter about what are the recent papers coming out in the world. And so I think the fact that it can give answers that are relevant when it hasn't really been given explicit psychiatry training – I don't think it's read the DSM – I think that's what makes it very impressive. You say in newer iterations, what would be possible when it kind of focuses especially on health and behavioural, health

and psychiatry? What could that be like? And I think we are seeing especially like Google, right, is working on different ones like Med-PaLM. So there are kind of specialised of these large language models that will hopefully learn about health and train on health. And of course, they have to prove themselves just because you build one, you have to show that it works and it's reliable, it's valid. But I think we are seeing as a general one, ChatGPT is it has utility. It can be impressive and I think the hopefully more specialised ones will continue to do that as well.

David Gratzer: What's the biggest surprise in terms of what your team has experienced playing with ChatGPT over these last number of months?

John Torous: I think one big surprise, though, is you don't always get the same answer back when you put something in. Which I think is a little bit concerning. We can see our search history and we can ask the same prompt over and over again. And sometimes we get back different information. Which probably means that the model is changing, underlying it, it's learnt something new. We have seen, I think 'hallucinations' is not a good term, being a psychiatrist, for when ChatGPT gets it wrong, we should say it makes factual errors or it 'lies' to us about what's happening. So sometimes for fun we put in "what is mindLAMP?", which is again the open-source app that we built. And it sometimes it gets it right. Sometimes it's a piece of software. It was built by different people. It was built in a different decade. Sometimes it doesn't know what mindLAMP is. So that's kind of our litmus test. We just say, "What is mindLAMP?" and sometimes it's spot on, but it's not quite reliable, if that makes sense. It's always entertaining though, for issues like that. So again, but again, mindLAMP is a more specialised thing. It doesn't have a large body of literature. Not everyone's writing about it like depression. So, I think as you get to more narrow topics, I guess what I'm saying is we begin to see it being less reliable.

David Gratzer: And I think that's a fair comment. Scott Patten, who is the former editor in chief of the *Canadian Journal of Psychiatry* and a professor at the University of Calgary, had emailed us a few months ago about getting ChatGPT to do a short bio, and he talked about the biographical information. Basic things were wrong, like the medical school he went to. I too played with it and was surprised that when I was in medical school I had been inducted into the Canadian Medical Hall of Fame. Great thing to read about, unfortunately, completely made up. So we're not going to use the word 'hallucination', but 'nonsense' might be our substitute word. How would that influence the patient experience?

John Torous: I think we know just from early digital work, right, when you're predicting something, giving information in health care, we have to get it right. And if you're going to get it wrong, even one out of 100 times, it's not enough. Right? We want things to work well. And I think we set the bar on purpose at a high level, and I think the technology will evolve to meet that. I think we certainly don't have to settle, just like we've talked in the past, we don't have to settle for apps we don't like. We don't have to settle for these language models, large language models, that are not ready yet. And I think it's great because, again, the technology companies will meet the challenge. I think they'll make the models better. But I think all of us as clinicians should be excited. We should be learning about them and then saying, here's what the threshold is, when we think they'll be useful and encourage the companies to get there. Because, again, there's potential benefits to it. But certainly, as we're recording this, they're certainly entertaining, they're exciting, but they're by no means ready for clinical use. At all.

David Gratzer: Fair. Early days. Though I mean have some evidence that these models might help us. As an example, as you know, there was the *JAMA Internal Medicine* paper where they looked at responses on Reddit to basic health questions. So, off the top of my head, the scenario is: I was biking, I hit my head, I have a headache now. Should I see a doctor? And then they took doctors' responses and compared it to ChatGPT, and the answers that ChatGPT put forward were more fulsome, were more reliable. And here's the really unsettling part, more empathetic as rated by blinded responders. You're familiar with the paper. What did you think about it?

John Torous: I mean, I think it's an impressive result. And again, I think if we use our psychiatry jargon, this is psychoeducation, right? This is information. And I think things like offering basic information, we want everyone to have access to information and to know what the standards are and how to follow it. And those are things I think that, again, you could likely find on different websites or different medical societies. That is good public accessible information. I think ChatGPT is making it easier to understand and more empathetic. That's not delivery of medicine or that's not the delivery of care, right. And so I think it's a good first step. And I think psychoeducation is a wonderful thing to probably begin to use these things for. But if you went to a doctor and all they could give you was psychoeducation, you would quickly leave that doctor. You would find one that does more. So, I think it's a good start.

David Gratzer: How far do you think we could go with AI? So, psychoeducation, you're right, it's low hanging fruit. Diagnosis and psychiatry also isn't usually so complicated sometimes. So that's low hanging fruit. What do you think we could see over the next 5 to 10 years?

John Torous: I think what will be exciting is we know, and you've covered, we have psychiatric genetics, right. It's never really made a direct impact in care. There's some evidence you could use it in the past to pick a medication, but that's never borne out. But what if these systems could integrate genetic data, your personal data, your medical record data, your environmental data. If it could really put together these different data streams that it's impossible for one person to synthesise today, it would really give you a more differential diagnosis based on a comprehensive, holistic picture – that would be exciting, right? That's something truly novel. You can't do it today. That's not predicting I have depression because I tell you, I'm depressed. I think then if it could give us these kind of more specific pictures of illness or kind of what it could be, maybe we could use those to personalise treatment? We've always wanted to do preventive psychiatry. We've always wanted to do personalised psychiatry. These have become these buzzwords that are a little hollow, right? It doesn't often translate into things we can do. But maybe if we really understood from a holistic picture because a source, a program like this can integrate so much diverse information, we really would understand changes and could guide those types of treatments. It's a little bit of a pipe dream at this point, but I think we have a lot of data about our health, and we never quite put it together and maybe this is a great way to do that in the future.

David Gratzer: There are issues around access to care in your country, the United States and my country, Canada, and right across the West. Goodness, it's not better in low income and middle-income countries either. Do you think there would be a role for therapy or some aspect of therapy?

John Torous:

Mental health has not gotten better, in a macro picture, in the last 20 years. You know, we've had more and more ways to make therapy accessible to people. So, I think in some ways it's a good time for us to pause and say, "this will be a new tool and mechanism to deliver therapy. What can we do to make sure it's different than apps or it's different than web-based programs?" Which again have not hurt, but have they had that kind of transformative, paradigm shifting impact? Somehow, not yet. And I think we have seen that. Well, again, we saw, unfortunately, some unethical studies. There was a company called Koko that did not tell people it was getting responses from ChatGPT when people asked for basic emotional support. And then later on, the company said, actually, that was ChatGPT, not a person. So that involved deception. It did not get ethics. Slightly, at least to me personally, terrifying that that kind of people would conduct unethical research on people without consent. But I think we'll see new efforts probably come and show that, especially for the basic parts of therapy, especially the psychoeducation, which is part of therapy, right, teaching people about the condition. So maybe we'll have ChatGPT do those parts and maybe with time it will do more.

David Gratzer: You bring up ethics. One concern always with digital solutions is the implications for privacy. So, you're talking about, as an example, loading up lots of personal information and it could personalise care as a result. But that would be a lot of personal data. How concerning is this?

John Torous: I think extremely concerning, right? If we look back to things we've talked about on prior episodes of *Quick Takes*, right. The privacy of mental health apps, which are gathering just a little bit of information has been abysmal, and it continues to be abysmal. So, I don't think we have a good track record of technology companies respecting people's personal mental health information. And I think in some ways it's so abysmal that the Mozilla Foundation released an update to its report of 'how creepy is your mental health app?' Again, we haven't done this for large language models because they're newer. But I think that we're seeing a point where we may actually need to have some type of external regulation. I think we all want innovation to thrive, and we know if we overregulate things, that's not good. But in some ways, as we've talked about in the past, if we don't have trust, we can't put information into these models. We're not going to get out those magical results that we want. And it'll be interesting to see do we see kind of new efforts or commitments towards trust that we can have faith in? Hard to know. In spring of this year, a large kind of US online telehealth provider Betterhelp was sued by the Federal Trade Commission – or not sued, they settled for \$8 million – for privacy breaches. And, you know, things are probably not good if the US federal government takes time out of itself to go after you as a mental health provider. So, I think that's probably signalling that there's broader issues happening across the whole industry that we have to be careful of. And I think we don't need to thus give a benefit of the doubt. They need to prove to us that they're going to be private.

David Gratzer: Doctors and other health care workers do many things that are tedious, mundane, frankly, things like discharge summaries, doctors actually do kind of a crummy job because we we're thinking about the next patient, not the one we just discharged. There's a good paper in *Digital Health* speculating, actually, that ChatGPT could be used for those sorts of mundane tasks. What are your thoughts?

John Torous: Yeah, I've seen some estimates on the potential cost savings of letting things like ChatGPT do more administrative work. And I think, given that the financial case makes sense, it's something that doctors would like to do less of, patients would like to have more time. I think you have a lot of different stakeholders probably agreeing that's a use case that could make sense. I think as soon as the privacy issues are solved, which again can be solved, technology can be very private if it does what it wants to do. Most of our banking information is relatively private and secure. There's always some risk. It's minimal. But if we could get over that privacy, I think it would make a lot of sense. There was a news article I read, I think even Epic is beginning to integrate some of these large language models, so we're already seeing the electronic medical record providers thinking about how to integrate this and how to put this in. And I wonder if we look at the history of innovation, sometimes it's not the snazziest thing, right? It's not the chatbot playing doctor and making a miraculous diagnosis. It's the chatbot filling out a discharge summary. It's not as glamorous, but the reality is that's going to probably be where we start and it's going to make a difference. And I think if it goes well in the discharge summary, I think it can probably move up. But if we're being realistic, it's going to be that kind of mundane but very important work that has to be done and this can help.

David Gratzer: So computers don't get bored and computers don't find things mundane. Could that be the future of discharge summaries and maybe eventually admission notes and day to day records drawing, perhaps even on voice records?

John Torous: Yeah. And I mean, I've talked with a Charlotte Blease is a researcher in Finland, and she's even said, well, maybe ChatGPT could make the notes accessible to patients with not English as a first language. It could put in their education level so I could write a note and it could really translate into a way that's very understandable for different patients. So, we're getting back to that kind of education and understanding, but it could actually make the notes better. Dare say, who's ever said let's make notes better? It's a rare thing. So, I think there's potential there too.

David Gratzer: You and I have gone back and forth on a few studies. We haven't really talked about how the sort of journal you edit might be changed by ChatGPT, but already some of the big journals have put out statements about authorship and what's allowed and not allowed. What are your thoughts?

John Torous: I mean, I think it goes back in part to the privacy issues, right? When you put something into these chatbots, you really don't know what's happening. So when you're putting your research data, your own personal ideas, figures, it makes sense that we have issues there. I think, though, I know a lot of people that are using it to help proof it. They may put a paragraph in. It'll certainly give something back and you can edit it. And I think, again, most of us have used spellcheck. I'm very impressed. If you're a listener who doesn't need to use spellcheck! And I think some people have used different kind of grammar programs, so I think a lot of people would be using it for kind of editing, proofreading, I imagine, at this point in time. I do think it doesn't always write, if you ask it to write kind of more conceptual things or to think about kind of arguments, it's not quite there yet. I think as an editor, actually only one person we've seen submitted a paper that clearly was written by ChatGPT, but it doesn't take long to spot it, at least at this point. Again, maybe next year. Someone listening to this in 2024 there'll be more issues, but it's not that sophisticated for let's call it psychiatry journal writing at this point in time. And I think honestly, I think most authors are doing a very good job respecting those boundaries because also it really wouldn't make sense. Or again, if you and I tried to write a paper in ten minutes based on ChatGPT, I think everyone would spot it very quickly and it would not go far in the peer review process. So I think it's not it may become more of an issue, I think right now we're seeing the journals preparing when it becomes more sophisticated and subtle.

David Gratzer: We do have AI, including with apps. What's a nightmare scenario we should think about?

John Torous: Yeah. So I think we don't want to cause harm. And I think we did see a case a couple of months ago with the Tessa chatbot that was rolled out by the National Association of Eating Disorders, where it was meant to help people with eating disorders, it seemed to be giving incorrect advice for unclear reasons, giving people dangerous and unhelpful information. Thankfully, it was quickly pulled from the Internet, but I think we learned about how even a well-tested chatbot can go off the rails and we have to be very careful. And the fact that we've had one of these cases already in 2023 is telling that again, we want to really be cautious because these are complex pieces of software, and they can cause harm. We have tangible cases of that harm.

David Gratzer: Do no harm is good advice for human doctors and ChatGPT helpers, perhaps.

Dr. Torous. It is a tradition here at *Quick Takes* that we do a rapid-fire minute. You've done a few of them already. Are you ready today?

John Torous: They're very stressful, but I'm willing to give it a shot!

David Gratzer: We keep inviting you back. Clearly, they're going well, sir. Let's put a minute on the clock. First question. Are you optimistic about ChatGPT and clinical work?

John Torous: Yes.

David Gratzer: Are you hesitant?

John Torous: Yes, one can hold two contradictory positions at once, I suppose.

David Gratzer: What excites you the most about AI in these early experiments?

John Torous: I'm not being quick on this one. I think what excites me the most is really making it easier for patients to understand complex medical information.

David Gratzer: Now it's possible to use ChatGPT for non-medical things. Have you used it in your personal life?

John Torous: Yes, it has. Very. Yes, we'll say.

David Gratzer: What have you done?

John Torous: As a tourist in a different city, it can actually give you a lot of great things to do that are pretty well, pretty good. And you don't have to browse through all those wild travel sites with pop up ads.

David Gratzer: Your last trip was partly a ChatGPT inspired trip?

John Torous: Yes.

David Gratzer: Okay. Now, at the buzzer, you've told us about your travels and your enthusiasm, at the buzzer here's the last question. Should I be retraining?

John Torous: No.

David Gratzer: That was pretty darn succinct! Yeah. Do you want to expand further on your eloquence?

John Torous: So, I think the best way to think of ChatGPT is a new modality to put together and share information, just like Wikipedia was a new modality to put together and share information. No one lost their job to Wikipedia. If anything, people were excited. Wikipedia made it easier for people to look up information. There's always risks of misinformation on Wikipedia. People know that. But overall, it helped kind of elevate everyone. And I think if we think of ChatGPT like a tool that's going to spread, it's a new conduit and it's exciting to have new conduits in vehicles, to share information, to put data together for psychiatry.

David Gratzer: A thoughtful comment and a thoughtful interview. Dr. Torous, as always, we appreciate your time and your insights.

John Torous: Thank you all. Bye.

David Gratzer: Thank you, sir. Cheers.

[Outro:] *Quick Takes* is a production of the Center for Addiction and Mental Health. You can find links to the relevant content mentioned in the show and accessible transcripts of all the episodes we produce online at CAMH.ca/professional/podcasts.

If you like what we're doing here, please subscribe. Until next time.