

BHB+

As AI Adoption Accelerates, Liability Becomes the Uneven Ground Beneath Providers' Feet

By [Ashleigh Hollowell](#) | December 8, 2025

While savvy behavioral health providers are considering AI liability from the outset of implementation, the technology's rapid evolution makes long-term planning difficult.

Just as quickly as the health care industry has embraced AI for its demonstrated benefits, incidents of the technology's limits and liabilities have drawn criticism and uncertainty among mental health professionals.

Many of the controversial incidents have been specifically tied to generative AI, which uses large language models (LLMs) to understand and analyze natural language and generate a response.

For example, AI chatbots, which are on the rise and increasingly under scrutiny for the harms they can incite, use generative AI. But any AI tool, used for front-end therapy or on the back-end of business for analysis or claims, comes with a level of risk.

"Any kind of AI assistance is going to have risk to it," Dr. Rachel Wood, a cyberpsychology researcher and the founder of the AI Mental Health Collective, told Behavioral Health Business. "But when you are moving over to a non HIPAA-compliant platform – like ChatGPT – that trains on your data, you really don't want to be using this in any kind of professional capacity with client or patient information."

The AI Mental Health Collective is a group of behavioral health leaders from various disciplines that aims to foster dialogue about the use of AI within the mental health field.

One [survey](#) found that OpenAI's [ChatGPT](#), one of the most widely used generative AI chat tools, may be the largest provider of mental health support in the U.S., with 49% of respondents reporting they have used it for psychological support within the past year. Notably, around 64% said these tools improved their mental health to some degree and only 9% reported encountering harmful responses.

However, as the user base for tools like ChatGPT has grown to more than 700 million weekly users, so have reports of [AI-induced psychosis](#), where the technology has reinforced delusions, reports of AI [encouraging suicide](#) and other harmful mental health behaviors when a user is in crisis. All of which has led to [multiple lawsuits](#) against the company, even as it has worked to develop better [mental health guardrails](#) and redirect users to resources when they're experiencing a crisis.

AI's controversy goes beyond just ChatGPT. [Preprint research](#) released Sept. 13 concluded that "across 1,536 simulated conversation turns, all LLMs demonstrated psychogenic potential, showing a strong tendency to perpetuate rather than challenge delusions."

It's something to watch, especially as more companies, particularly in the digital mental health realm, aim to use AI and to tailor a chatbot for their own platform. Many also use generative AI to record and analyze

therapy sessions or to summarize notes from a patient's case, and the liabilities associated with doing so are mounting.

“Understand the limitations and the liabilities that are present,” Wood said, “Things to be looking for are: How do they store your data? Are they training the model on your data? Are they using deidentified information? All of these things are really important. Then also knowing how their model works and what kind of feedback loops of recourse are present. For instance, if you pull up an output from an AI and there are some gross errors in it, do you have access to let the developers know directly?”

HIPAA-compliant AI tools do exist, but regardless of which AI tool a mental health professional considers adopting, it's important to be aware that liabilities exist to a degree with all of them. Researching the tool, speaking with colleagues and becoming aware of both the risks and benefits of its use are just some steps Wood recommends taking prior to AI adoption.

AI is unique in that it autonomously trains on data and evolves. Many of its capabilities and liabilities have yet to be fully realized, making it challenging for providers and insurers to anticipate future risks.

What's at stake

A major challenge with evaluating the risks vs. benefits of AI is that the technology is so new. Safeguards, legislation and legal cases involving its use within the mental health space are still in early stages.

California, which often paves the way in technology and health practices, just passed an [inaugural law](#) in October that establishes some of the first guardrails for AI chatbots in the nation, particularly their use with minors. The state has also passed other legislation that mandates transparency with patients about the use of any kind of clinical AI communication.

New York followed suit and [passed its own law](#) in November, establishing mandates and safety protocols for AI systems designed to simulate human companionship. Both California and New York laws are slated to take effect in January 2026.

The Food and Drug Administration (FDA) is also engaged in ongoing review and regulation of generative AI-powered mental health tools. A Nov. 6, 2025, meeting of the agency's Digital Health Advisory Committee focused on AI-enabled mental health medical devices and determined that “generative AI and some LLMs, to date, have demonstrated vulnerabilities in some of the areas where human therapy excels,” a meeting [document](#) states.

“As generative AI advances and therapeutic roles for generative AI continue to be explored, the necessity of developing effective guardrails becomes important to balance unintended consequences and mitigate adverse effects in using digital mental health medical devices as a replacement to human therapy,” the FDA committee wrote.

It was the first of what are likely to be several meetings the agency will hold as it works to develop policy on AI's use in the behavioral health field.

There are thousands of AI-powered tools and platforms with different uses, data sets and capabilities – which can make regulation hard to establish, especially as the technology has rapidly taken off. But providers also need to be careful about what AI adoption could mean for malpractice liabilities.

“It is a little bit hard to generalize over what is potentially a very broad suite of tools,” Michelle Mello, a professor of law and health policy who specializes in AI research at Stanford University, told BHB. “There’s a lot of uncertainty about how malpractice liability would be divided between the end user and the organization that decided to deploy the AI system, such as a physician practice or hospital and the developer itself. There’s just a lot that’s unsettled in the law around when developers can be held liable.”

Historically, Mello said, it has been challenging to hold product developers liable for malpractice-related injuries in the health sector. Clinicians tend to be “at the pointy end of the stick” and become the party in cases that “suck up that residual liability,” she said.

Right now, there are a lot of conversations about AI tool liability, but no cases yet that show how courts will evaluate these tools or what litigation might look like.

“I think in the near term, at least, it will be challenging for plaintiffs to prevail in these types of claims when they do start bringing them,” Mello said, adding that the burden of proof ordinarily required in these types of cases may be even higher.

But what concerns Mello most is another liability. As AI advances, much like with other technologies that are used and trusted today, “our desire to retentively review their output and make sure it’s right will lessen, as it does for all kinds of computer decision support,” she said.

AI is imperfect. The technology can hallucinate, produce analyses and answers that are flawed and exhibit bias from the data it was trained on. The tools can be dangerous without a human in the loop to review and correct errors, or if relaxed approaches are taken in review.

Right now, it’s common for mental health C-suite professionals to reiterate that AI should not replace human expertise in care, but in practice, the “human nature” of technology reliance and relaxed oversight might be a challenge that bleeds into more liabilities, Mello said.

“It’s really the same as any other piece of technology or any other non-AI decision support tool,” Mello said. “If you follow decision support for a patient and it gives you the wrong answer, you are potentially liable for the injury to the patient that results unless you can show that your reliance on it was reasonable. There’s nothing really new about AI in that sense.”

Preparing for what’s ahead

AI is also changing how payers evaluate risk, with implications for practitioners.

A spokesperson for the Blue Cross Blue Shield Association (BCBSA), a national federation of independent Blue Cross and Blue Shield companies, reiterated Mello’s concern: that the differing outcomes and levels of oversight are still evolving — and as a result, risk is hard to predict.

Independent Blue Cross Blue Shield companies conduct their own independent evaluation of AI tools, therapies and treatments to inform coverage decisions, but there is still much to monitor.

“The science on AI-enabled tools is still emerging and there is still a lack of sufficient data on appropriate use cases and outcomes, and levels of human oversight vary,” the BCBSA spokesperson told BHB. “These factors can have profound implications on whether the tools have a positive or negative impact on a patient’s mental

health and their ability to get the right care at the right time. BCBS companies continue to monitor new research and updates in the regulatory environment to ensure AI is deployed safely and responsibly.”

The America’s Health Insurance Plans (AHIP) has already seen health plans across the U.S. embrace AI for its benefits, but noted that a commitment to “appropriate governance, monitoring and safeguards to prevent unintended consequences,” is critical to preparing for what is ahead, a spokesperson told BHB.

Providers evaluating AI tools for risk and liabilities should anticipate harms before they happen, Dr. Nikhil Nadkarni, chief medical officer at Brightline, a pediatric mental health practice with locations across New York, explained.

“I spend a lot of time thinking about not just the promise of AI and how it might be really exciting, and what are the things that it can be useful for, but what are the harms that I expect?” Nadkarni told BHB. “What are the harms that I’m seeing immediately in the day-to-day? And what are the ways that those harms will impact various stakeholders: companies, children, families – all of the above?”

As the technology evolves, one of the key priorities that Gijo Mathew, chief product officer at Spring Health, said he will be engaged in is [VERA-MH](#) – an open-source evaluation tool for validating AI uses across the mental health field.

The acronym stands for “Validation of Ethical and Responsible AI in Mental Health” and is a partnership with a coalition of leaders across healthcare, technology and ethics.

The framework is designed to create clear evaluation criteria to help determine if an AI tool can recognize and appropriately respond to signs of mental health crises like suicidal ideation and ensure escalation to a human clinician when needed.

“Its goal is to create a common standard for safety and performance so innovation in AI for mental health can move forward responsibly and with confidence,” Mathew told BHB. “It is critical for doctors and therapists to continuously review AI outputs and conduct their work with human oversight and decision-making while practicing with the technology.”

Spring Health is a New York-based digital mental health provider that partners with employers and health plans to offer a range of mental health support services.

Mathew said that as AI evolves, this framework will help address concerns about inaccurate outputs, clinician liabilities and malpractice issues. Since the tool is open source, it can be freely utilized, adapted and modified across specific companies and practices to fit their own needs.

“Every AI mental health tool should meet the highest global compliance standards, including

HIPAA, SOC2, and HITRUST,” Mathew said. “These data security standards set the bar even higher than typical HIPAA compliance requirements and can determine whether the tool is built safely. ... Ethical use means maintaining human oversight, informed consent and full transparency about what the AI is doing and what it’s not. That’s why frameworks like VERA-MH are important in ensuring AI in mental health platforms meet rigorous clinical and ethical standards.”