October 17, 2025

FDA Digital Health Advisory Committee
Center for Devices and Radiological Health (CDRH)
U.S. Food and Drug Administration
Silver Spring, MD 20993

Re: Generative AI-Enabled Digital Mental Health Devices (Docket No. FDA-2025-N-2338)

Dear Members of the Digital Health Advisory Committee:

SonderMind is a comprehensive mental health provider dedicated to improving access to high-quality, continuous care through technology and clinical expertise. We appreciate the opportunity to comment on the FDA's examination of generative AI-enabled digital mental health devices. As a stakeholder committed to safe and effective mental health solutions, we support a regulatory framework that protects patients and promotes equity while also encouraging innovation in this emerging field.

In the following comments, we respectfully recommend a tiered, risk-based regulatory approach, support transparency and flexibility over rigid premarket controls, and highlight safety, equity, and privacy as guiding principles. Our goal is to help the FDA balance health innovation with public and provider trust in AI mental health tools.

**Tiered, Risk-Based Approach to Mental Health AI Tools**

We urge the FDA to adopt a tiered, risk-based regulatory approach for AI-enabled mental health tools. Not all such applications carry the same level of risk, and oversight should be calibrated accordingly. In particular, general wellness and coaching tools should be clearly distinguished from higher-risk clinical interventions and excluded from medical device regulation. Congress and the FDA have already recognized that software intended solely for general wellness – to maintain or encourage a healthy lifestyle without diagnosing or treating a condition – is not a medical device.[1] Consistent with this policy, AI-powered mental wellness apps (for stress reduction, meditation, coaching, etc., without specific disease claims) should remain outside FDA device oversight. This clarity ensures low-risk tools that benefit consumers (e.g. mood trackers or meditation apps) are not stifled by unnecessary regulation, allowing the FDA to focus on higher-risk AI applications (such as diagnostic chatbots or clinician decision-support systems).

[1] *General Wellness: Policy for Low Risk Devices*, Guidance for Industry and Food and Drug Administration Staff, U.S. Food & Drug Administration (September 2019), sec. 520(o)(1)(B) (citing amendments under section 3060(a) of the 21st Century Cures Act), https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-wellness-policy-low-risk-devices.

A tiered framework would assign different levels of regulatory scrutiny based on each tool's risk. For example:

- Low-Risk Tools (Wellness/Coaching): Apps or chatbots that provide general wellness support or coaching without medical claims. These would be subject to minimal requirements (e.g. basic safety standards and consumer disclosures) and not treated as medical devices.

- Moderate to High-Risk Tools (Clinical Use): AI intended to be used by a clinician to diagnose, treat, or guide management of mental health conditions. These would require FDA review with evidence of safety and effectiveness commensurate with their risk level.

By carving out low-risk products and concentrating oversight on tools with genuine clinical risk, FDA can promote innovation and access without compromising safety. A risk-based model provides clarity for developers and directs regulatory attention where it is most needed.

**Emphasizing Transparency and Alternatives to Premarket Controls**

We encourage the FDA to favor transparency and accountability in lieu of overly prescriptive premarket requirements for AI mental health tools. Given the rapid evolution of AI, a flexible approach that emphasizes ongoing transparency and post-market oversight is preferable to rigid prescriptive rules.

Transparent documentation and user disclosures should play a central role. For example, developers can provide model cards or similar documentation summarizing an AI tool's intended use, design, performance, and limitations.[2] Notably, FDA's own draft guidance suggests that model cards help communicate key details of AI devices and can "increase user trust and understanding".[3] Additionally, developers should publish or submit evaluation reports detailing how the tool was validated, including accuracy metrics and known limitations or biases. Equally important, end users – both patients and clinicians – should receive clear plain-language disclosures when using AI mental health tools. For instance, an app could notify users that it is an AI-driven assistant, describe its scope, and advise on seeking human help for crises.

---

[2] U.S. Food & Drug Administration, *Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations*, Draft Guidance (Jan. 7, 2025), at Appendix E (Example Model Card), https://www.fda.gov/media/184856/download.

[3] U.S. Food & Drug Administration, *Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations* (Draft Guidance, Jan. 7, 2025), at 1512, https://www.fda.gov/media/184856/download.

Rather than mandating identical premarket tests for every AI tool, the FDA can set baseline expectations for transparency and monitoring to ensure accountability. For example, the agency could require robust labeling so users understand an AI tool's limits, encourage

sponsors to include model documentation in their submissions, and enforce post-market monitoring so developers track real-world use and address emerging risks. This transparency-first approach holds developers responsible for demonstrating quality and informing users, without unduly hindering iterative improvements. By working with industry to develop clear guidelines on documentation and disclosures, FDA can foster an environment where mental health AI tools are trusted and used appropriately while innovation thrives.

**Flexibility in Validation Methods and Protecting Proprietary Data**

We recommend that the FDA maintain flexibility in how companies validate AI-driven mental health devices, and exercise caution about requiring over-disclosure of proprietary data. The field of mental health AI is diverse–ranging from self-help chatbots to clinical decision support–so a rigid, one-size-fits-all validation approach could inadvertently hamper beneficial technologies.

- Flexible Validation: FDA should allow a variety of evidence to demonstrate an AI tool's safety and effectiveness. In some cases traditional trials may be feasible; in others, real-world evidence or observational studies might suffice. The key is to focus on meaningful outcomes and performance standards, rather than prescribing one uniform testing method for all products. FDA can provide guidance on best practices, but developers should have leeway to choose validation methods suited to their tool and patient population, as long as the evidence is sound.

- Proprietary Information: We also urge caution on any requirements to publicly disclose sensitive proprietary details, such as source code or algorithms. Regulators should have access to detailed information as needed for evaluation, but this can be done confidentially. Forcing companies to reveal trade secrets or extensive training data is not necessary for accountability and would discourage innovation. We support transparency about performance–for example sharing summary results, intended use, and limitations–without requiring companies to expose their intellectual property. This balanced approach to keeping proprietary details confidential while being transparent with performance data protects both patients and the incentives to develop new mental health technologies.

**Prioritizing Safety, Equity, and Privacy as Core Principles**

Finally, safety, equity, and privacy must remain the bedrock principles for overseeing AI mental health tools. We urge the FDA to embed these values in its framework:

- Patient Safety**:** AI tools for mental health must be designed and tested to do no harm. This means building in safeguards to prevent unsafe or inappropriate outputs and ensuring the AI knows when to defer to human help. For example, a mental health chatbot should be programmed to recognize crisis cues like suicidal ideation

  and immediately prompt human intervention rather than attempt to handle a crisis alone. Clear limits on the system's role should be defined–for example, it should not diagnose conditions or give advice outside its scope–and continuous monitoring for harmful behavior is essential. Higher-risk applications may warrant human-in-the-loop oversight or on-call clinician backup.

- Equity and Bias Mitigation: It is crucial that AI mental health devices work for diverse populations and do not worsen health disparities. We recommend requiring regular bias audits to evaluate performance across demographic groups, and require developers to address any significant disparities. This aligns with the medical consensus that AI systems should "identify and take steps to address bias and avoid… exacerbating health care disparities."[4] The FDA should make such equity assessments part of its evaluation criteria, requiring evidence that AI tools have been tested for bias. Prioritizing fairness will help ensure these technologies reduce gaps in access to care rather than reinforce them.

- Privacy and Data Protection: Mental health data is highly sensitive, so strong data protection is imperative. AI tools should adhere to robust privacy standards – for example, using encryption, following applicable health privacy laws, and being transparent about data use. Users deserve to know what personal information (if any) an AI application collects, how it is stored and used, and who might access it. We echo that safeguarding privacy is as critical as addressing bias in health AI.[5] Clear regulatory expectations for privacy-by-design and secure data handling will protect users and maintain public trust in digital health innovations.

- User Education and Informed Use: Both consumers and providers need to understand the appropriate role of AI in mental health care. Regulators should encourage efforts to educate users on each tool's purpose and limitations. For example, an AI self-help app might include an onboarding tutorial explaining that it provides coping exercises and psychoeducation–not a medical diagnosis–and advising when to seek professional help. Similarly, clinicians using an AI decision-support tool should be trained on its proper use and limitations. By promoting informed use of these technologies, we can prevent misuse or

[4] Robeznieks, "FDA Must Guard Against Bias in AI, Focus on Patient Outcomes," *American Medical Association*, June 24, 2014, https://www.ama-assn.org/practice-management/digital/fda-must-guard-against-bias-ai-focus-patient-outcomes.
[5] *Ibid.*

overreliance and ensure that AI serves as a supplement to, not a replacement for, professional care.

SonderMind strongly supports the FDA's efforts to develop a thoughtful regulatory approach for AI-enabled mental health devices and is committed to working with all stakeholders to ensure that digital mental health innovation proceeds responsibly. We believe a balanced framework, one that is risk-based, transparency-oriented, and grounded in safety, equity, and privacy, will best serve patients. Such an approach can protect people from harm and bias, preserve trust, and still allow promising AI tools to improve access to care and outcomes. We appreciate the Agency's consideration and welcome continued dialogue.

Thank you for the opportunity to share our perspective.

Sincerely,

Carmen Feldman

Carmen Feldman
VP, Head of Legal & GRC

Miguel Alvarado

Miguel Alvarado
VP of Engineering, AI Engineering