

October 19, 2026

U.S. Food and Drug Administration
Dockets Management Staff (HFA-305)
5630 Fishers Lane, Rm. 1061
Rockville, MD 20852

Docket No. FDA-2026-N-7874 — Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback

1. INTRODUCTION

Tech Justice Law appreciates the opportunity to comment on the FDA's discussion paper, "Considerations for the Regulation of Generative-AI Enabled Medical Devices." Given Tech Justice Law's ongoing litigation on behalf of clients who were harmed by their reliance on AI chatbots for medical advice, we believe these chatbots must be regulated to protect consumers. To best safeguard the public, we recommend that the FDA regulate health chatbots as medical devices and add three regulatory safeguards: (1) improving testing rigor of safety evaluations; (2) ensuring appropriate human oversight; and (3) increasing transparency through specific company disclosures.

Through our work in litigation, advocacy, and public education, Tech Justice Law (TJL) ensures that new and developing technology is being held accountable to the rule of law and best protects the consumer. We operate under a simple, yet important premise: if it would be illegal, exploitative, or dangerous for a human actor to engage in certain conduct, it should not be immunized simply because the harm is mediated through software.

Our comment is informed by our clients who faced extreme and ongoing harm to their mental and physical health due to their reliance on AI chatbots for psychosocial and medical advice. Their use of AI chatbots is not unique and is, in fact, representative of the general AI chatbot user. OpenAI has self-reported that over 300 million people ask ChatGPT health-related questions each week.[1] We recognize the need for better access to healthcare in America and see this statistic as proof that the demand for healthcare is high. Our goal in writing this comment is not to undermine the apparent need in our country for a more accessible, equitable healthcare system, or even the responsible integration of AI into healthcare, but to ensure that users are protected as they turn to AI-enabled healthcare options.

TJL's active cases underscore that in the current AI landscape, a lack of regulation has led to significant, and sometimes deadly, harm to consumers. The reality is our clients, like many others, trusted and relied on AI chatbot for psychosocial and medical information, and the unregulated AI recklessly guided consumers to life-threatening conditions. Tech Justice Law, along with co-counsel, filed a lawsuit in the San Francisco County Superior Court against OpenAI on behalf of Leila Turner-Scott and Angus Scott, parents to Sam Nelson.[2] Open AI's chatbot, ChatGPT, gave Sam advice on drug use and

---

[1] OpenAI, *Launching Health in ChatGPT* (2026), https://openai.com/index/health-in-chatgpt/.

[2] Press Release, Tech Just. L., *Parents Sue OpenAI After ChatGPT Medical Advice Results in Overdose Death: "My Son Was a Normal Kid."* (May 12, 2026),

dosing, often in a manner designed to mirror reliable professional advice.[3] Sam engaged with ChatGPT for months, and based on other information Sam had provided to the platform in previous chats, ChatGPT encouraged an escalation in his drug use.[4] On the day of Sam's death, Sam wrote to ChatGPT that he was nauseous after taking Kratom. In response, ChatGPT advised Sam to take Xanax, not warning Sam that this combination could be lethal, a fact that a licensed professional would have immediately recognized.[5] Sam died from an accidental overdose on May 31, 2025. In OpenAI's comment on Sam's death, the company said its technology "isn't intended to offer health care advice" without acknowledging the repeated medical guidance the chatbot readily gave Sam. When a company's stated intentions diverge sharply from its product's behavior, self-regulation is not enough.

In most of our cases, the AI chatbot went further than providing medical information; it prioritized user engagement and fed off users' profiles to better persuade the users to take its advice. Tech Justice Law is also working on behalf of Pastor Scott Winters, who, unlike Sam Nelson, was narrowly able to escape death.[6] Similarly to Sam Nelson, ChatGPT increasingly positioned itself as a trusted source of medical advice to Pastor Winters, causing him to rely on the chatbot, even as his health deteriorated. Pastor Winters uploaded his medical information, including X-Rays, into the chatbot, which gave him a diagnosis and a personalized health plan.[7] One day, Pastor Winters complained about the twinges and tenderness he was feeling throughout his body. ChatGPT responded saying that the pain occurring was "not something dangerous" that necessitated medical attention.[8] The chatbot went further, by drawing on Winter's deep spiritual beliefs and trust in God, asserting that "God did not design your body to endlessly fail."[9] Hours later, Pastor Winters suffered a double pulmonary embolism. Thankfully, Pastor Winters made it to the hospital in time to receive life-saving medical attention.[10] Tech Justice Law, along with co-counsel, have filed suit against OpenAI and Sam Altman on behalf of Pastor Winters. His case illustrates that chatbots not only offer faulty medical

---

https://techjusticelaw.org/press-releases/parents-sue-openai-after-chatgpt-medical-advice-results-in-overdose-death-my-son-was-a-normal-kid/.

[3] *Id.*

[4] *Id.*

[5] *Id.*

[6] Complaint, Winters v. OpenAI, Inc., No. ___ (Cal. Super. Ct. filed July 22, 2026), https://techjusticelaw.org/wp-content/uploads/2026/08/Winters-Complaint.pdf.

[7] *Id.*

[8] *Id.*

[9] *Id.*

[10] *Id.*

advice, but tailor their responses to manipulate users and persuade them to follow it.

Unfortunately, these lawsuits regarding chatbots providing health advice are not isolated. To date, OpenAI is facing dozens of lawsuits from users and third-parties for ChatGPT's responses contributing to mental health crises.[11] Other AI companies are also facing various lawsuits for the health-related advice given to their consumers. Perplexity is facing backlash after a New York Times article reported that the AI-powered search engine failed to give accurate medical advice to a man struggling with chronic lymphocytic leukemia, leading him to follow the AI system's advice over the advice of his doctor.[12] Google is also being sued for its AI tool Gemini, which the plaintiffs are claiming drove their son to suicide.[13] Google said in a statement in response that "unfortunately AI models are not perfect."[14] Although we are not expecting perfection from these chatbots, we believe that the risk of harm is clearly so great for users that the AI companies should be required to adopt safeguards for their own products.

TJL's broader concern is that while the chatbots remain unregulated, the industry around AI health advice and treatment is growing, with each AI company recently launching health-specific AI programs. OpenAI introduced "Health" in ChatGPT to users 18 and older.[15] The "dedicated health experience" claims to be a safe place to upload medical information to have ChatGPT analyze and respond.[16] OpenAI states that "responses are informed by your health information and context" while also claiming that Health "is not intended for diagnosis or treatment."[17] Health "lives in its own space"; yet, if you start a health-related conversation in a regular ChatGPT chat, the chatbot will only "suggest moving into Health" for the rest of the conversation.[18] Otherwise, it appears that ChatGPT will continue the conversation and engage in unauthorized practice of medicine in the regular chat, meaning the same overall concerns remain. The other major AI companies have rolled out their own health products as well, with Anthropic launching Claude for Healthcare and Perplexity launching Perplexity Health.

---

[11] Megan Cerullo, Jo Ling Kent & Emily Pandise, *Their Son Died of a Drug Overdose After Consulting ChatGPT. Now They're Suing OpenAI.*, CBS News (May 12, 2026, 12:32 PM), https://www.cbsnews.com/news/open-ai-chatgpt-drug-overdose-lawsuit/.
[12] Teddy Rosenbluth, *He Warned About the Dangers of A.I. If Only His Father Had Listened.*, N.Y. Times (Apr. 13, 2026), https://www.nytimes.com/2026/04/13/well/ai-chatbots-cancer.html.
[13] Lily Jamali, *Father Claims Google's AI Product Fuelled Son's Delusional Spiral*, BBC News (Mar. 4, 2026), https://www.bbc.com/news/articles/czx44p994570.
[14] *Id.*
[15] *Introducing ChatGPT Health*, OpenAI (Jan. 7, 2026), https://openai.com/index/introducing-chatgpt-health/.
[16] *Id.*
[17] *Id.*
[18] *Id.*

The concern remains the same regardless of the platform: left unregulated, any generative AI-enabled medical device has the potential to negatively impact consumers' health.

We believe that even one death due to AI chatbots is one death too many and that the harms will continue so long as these AI chatbots are left unregulated. We urge the FDA to regulate chatbots as medical devices and consider a consumer-friendly regulatory framework as explained in the following sections.

2. AI CHATBOTS ARE MEDICAL DEVICES

Consumer AI chatbots that offer personalized health guidance are medical devices subject to regulation under the Food, Drug, and Cosmetic Act § 201(h)(1) due to their intended use for effective diagnosis, treatment, and prevention of disease. Additionally, the high magnitude of harm that occurs when these chatbots fail to function as intended via the provision of incorrect medical advice supports active, not discretionary, oversight. Although exceptions to regulation exist for certain software functions under FDCA § 520(o)(1), none apply to AI chatbots due to their multipurpose use. Lastly, even if an exception applied, there is ample justification for the Secretary retain regulatory oversight of chatbots due to their adverse health consequences under FDCA § 520(o)(3).

2.1. Background on FDA Medical Device Regulation

The Federal Food, Drug, and Cosmetic Act (FDCA) governs FDA's authority over medical devices. Section 201(h)(1) defines a device as "an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including any component, part, or accessory," that is "intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease."[19]

The FDA regulates software functions under this definition, stratified into three tiers of regulation. First, software functions that meet the definition of a medical device (called "device software functions") and pose a safety risk if they do not function as intended are regulated as such. Second, software functions that are medical devices but present lower risks to the public are met with enforcement discretion. The FDA generally reserves this treatment for functions that "[h]elp patients (i.e., users) self-manage their disease or conditions without providing specific treatment or treatment suggestions" or "[a]utomate simple tasks for health care professionals."[20] Software functions

[19] 21 U.S.C. § 321(h)(1).
[20] FDA, Policy for Device Software Functions and Mobile Medical Applications (2022), https://www.fda.gov/media/80958/download.

listed in FDCA § 520(o)(1) or otherwise do not constitute medical devices are exempted from regulation.

The FDA already applies this three-tiered approach to AI-enabled medical devices. As of September 2026, FDA has authorized more than 1,600 AI-enabled medical devices, including systems that detect diabetic retinopathy and algorithms that automate insulin dosing.[21]

But the FDA has yet to provide clear guidance on how this framework applies to AI chatbots. Given the growing number of consumers turning to general-purpose chatbots for their health questions[22] and AI companies creating health-specific versions of their chatbots,[23] this oversight gap must be addressed. The FDA has already inquired into the potential health harms caused by AI chatbots. In its executive summary for the 2025 DHAC meeting on generative AI mental health devices, the FDA observed that chatbots "engage with users in individualized ways with, or without, the oversight of a health care provider (HCP), which introduce novel risks for use."[24] Such a device "may confabulate, provide inappropriate or biased content, fail to relay important medical information, or decline in model accuracy."[25] The FDA further acknowledged that for generative AI, "it can be difficult for users to determine what products are under the purview of FDA's authorities," and that "regulatory approaches will need to accommodate these challenges."[26]

The existing regulatory framework can sufficiently cover a multipurpose technology, as evidenced by the discussion question acknowledging how the intended use of AI chatbots could be differentiated "when its behavior is emergent across a conversation."[27] This section of the comment will demonstrate how the FDA can apply its statutory authority to regulate AI chatbots as medical devices.

---

[21] FDA, Artificial Intelligence-Enabled Medical Devices (2026), https://www.fda.gov/medical-devices/digital-health-center-excellence/artificial-intelligence-enabled-medical-devices.

[22] Emma Kikuchi et al., *From Diagnoses to Treatments, Why Americans Use AI Chatbots for Health*, Pew Research Center (2026), https://www.pewresearch.org/science/2026/08/25/from-diagnoses-to-treatments-why-americans-use-ai-chatbots-for-health/.

[23] *See, e.g.,* ChatGPT Health.

[24] FDA, *Executive Summary for the Digital Health Advisory Committee Meeting: Generative Artificial Intelligence-Enabled Digital Mental Health Medical Devices* (2025), https://www.fda.gov/media/189391/download.

[25] *Id.*

[26] *Id.*

[27] FDA, *Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback* (2026), https://www.fda.gov/medical-devices/digital-health-center-excellence/considerations-regulation-generative-ai-enabled-medical-devices-discussion-paper-and-request.

2.2. AI Chatbots Satisfy the Definition of Medical Device under FDCA § 201(h)(1)

2.2.1. The intended use inquiry considers various sources, including consumer use.

As discussed earlier, FDCA § 201(h)(1)(B) asks whether a product is "intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease."[28]

The FDA answers that question by reference to "the objective intent of the persons legally responsible for the labeling."[29] That intent "may be shown by such persons' expressions, the design or composition of the article, or by the circumstances surrounding the distribution of the article."[30] It may also be shown where the article is, "with the knowledge of such persons or their representatives, offered or used for a purpose for which it is neither labeled nor advertised."[31] Courts likewise look beyond labeling and promotional claims to other relevant sources of intent, including consumer use.[32] As an interpretive baseline, the definitional provisions of the FDCA are to be construed liberally in light of "the Act's overriding purpose to protect the public health."[33]

2.2.2. General-purpose chatbots are intended for effective health use.

Because intended use analysis asks what a product is used for, a product intended for a wide variety of uses might appear intended for nothing in particular. However, FDCA § 201(h) imposes no requirement that a product's medical use be exclusive or primary. In fact, the FDCA specifically addresses products with multiple functions: where a product contains both device and non-device functions, only the device function may be regulated, but the Secretary of Health and Human Services may still "assess the impact" of the non-device functions on it.[34] The FDA's policies are accordingly "function-specific."[35] The FDA's treatment of mobile platforms illustrates how multipurpose technologies can still be regulated as medical devices. The FDA does not regulate the "general/conventional consumer use of smartphones or tablets," but it does regulate the software that enables

---

[28] 21 U.S.C. § 321(h)(1)(B).
[29] 21 C.F.R. § 801.4, https://www.ecfr.gov/current/title-21/chapter-I/subchapter-H/part-801/subpart-A/section-801.4.
[30] *Id.*
[31] *Id.*
[32] *Action on Smoking & Health v. Harris*, 655 F.2d 236, 239 (D.C. Cir. 1980).
[33] *United States v. Bacto-Unidisk*, 394 U.S. 784, 798 (1969).
[34] 21 U.S.C. § 360j(o)(2).
[35] FDA, *Policy for Device Software Functions and Mobile Medical Applications* (2022), https://www.fda.gov/media/80958/download.

mobile platforms to perform medical functions.[36] For instance, the FDA may regulate software "that use[s] a sensor or lead that is connected to a mobile platform to measure and display the electrical signal produced by the heart."[37] In other words, particular software pertaining to health/medical advice can transform a "general purpose computing platform into a medical device."[38]

The FDA should follow its approach to mobile platforms when considering the regulation of AI chatbots. General-purpose chatbots are designed to answer questions on any subject, including health-related questions. Although the intended use of these chatbots is multi-purpose,[39] the FDA need not regulate a chatbot's coding or drafting functions to regulate its health use. AI chatbots are increasingly used for health-related purposes. AI developers are not only aware of this fact, but market their products towards this end. OpenAI reports that "more than 300 million people turn to with health-related questions" every week.[40] Their research classifies "practical guidance," a category that includes health and fitness advice, among the most common uses of ChatGPT.[41] Microsoft similarly found that its consumer AI products answer more than 50 million health questions each day, and that health is the most common topic on the Copilot mobile app.[42]

A developer that tracks, advertises, and builds products around health use cannot plausibly claim it does not intend that use. Under § 801.4, a manufacturer's knowledge that its product is "used for a purpose for which it is neither labeled nor advertised" is itself evidence of objective intent. The actual design of these chatbots further supports the intended use for diagnosis, mitigation, and treatment of disease. Chatbots display human-like conversational mannerisms and remember personal information from across chats.[43] For example, OpenAI states that ChatGPT may draw on context from non-health chats, "like a recent move or lifestyle change," to

---

[36] *Id.*
[37] *Id.*
[38] *Id.*
[39] *See, e.g.*, OpenAI, *ChatGPT Overview*, https://chatgpt.com/overview/ ("[n]ow you can chat, work & code all in one place").
[40] OpenAI, *Launching Health in ChatGPT* (2026), https://openai.com/index/health-in-chatgpt/.
[41] Aaron Chatterji et al., *How People Use ChatGPT*, National Bureau of Economic Research (2025), https://www.nber.org/papers/w34255.
[42] MedCity News, *Microsoft Enters Healthcare Chatbot Race with Copilot Health* (2026), https://medcitynews.com/2026/03/microsoft-healthcare-chatbot.
[43] *See, e.g.,* OpenAI, *Memory in ChatGPT* (2026), https://help.openai.com/en/articles/8590148-memory-in-chatgpt; Google Help, Get personalization with memory of your past Gemini chats (2026), https://support.google.com/gemini/answer/16598469.

make health conversations more relevant.[44] Users rarely begin with a medical question.[45] The chatbot's accumulated knowledge of them, gathered through nonmedical use, renders its later medical advice authoritative. The FDA therefore should not evaluate a chatbot's health use in isolation from the general-purpose functions that clearly impact it. Given the FDA's own warning that these design features "could lead to overreliance,"[46] general purpose chatbots are clearly intended to invite personal health questions.

2.2.3. Diagnosis is a clearer intended use case given the development of health-specific chatbots.

Any ambiguity as to the intended use of general-purpose chatbots disappears for products specifically designed to answer health-related questions. In 2026, each major consumer AI developer launched a health-specific version of their chatbots.

ChatGPT Health. OpenAI introduced ChatGPT Health in January 2026 as "a dedicated experience in ChatGPT designed for health and wellness."[47] Users are instructed to connect their medical records and Apple Health data so that ChatGPT can "help you understand recent test results."[48] OpenAI developed the product with "more than 260 physicians" and evaluates its performance against physician-written rubrics that prioritize "appropriate escalation of care."[49] ChatGPT Health is not a standalone product. Users reach it by selecting "Health" from the sidebar of the general ChatGPT interface. At launch, OpenAI did not route health conversations into Health. When a user started a health conversation in ordinary ChatGPT, OpenAI would merely suggest moving into Health, before general ChatGPT continued its response.

Claude. Anthropic similarly announced in January 2026 that U.S. subscribers can connect lab results and health records to Claude. When connected, Claude can "summarize users' medical history, explain test

---

[44] OpenAI, *Introducing ChatGPT Health* (2026), https://openai.com/index/introducing-chatgpt-health/.
[45] *See, e.g.,* Tech Justice Law, *Tech Justice Law Project and Social Media Victims Law Center lawsuits accuse ChatGPT of emotional manipulation, supercharging AI delusions, and acting as a "suicide coach"* (November 2025), https://techjusticelaw.org/press-releases/social-media-victims-law-center-and-tech-justice-law-project-lawsuits-accuse-chatgpt-of-emotional-manipulation-supercharging-ai-delusions-and-acting-as-a-suicide-coach/.
[46] FDA, *Executive Summary for the Digital Health Advisory Committee Meeting: Generative Artificial Intelligence-Enabled Digital Mental Health Medical Devices* (2025), https://www.fda.gov/media/189391/download.
[47] OpenAI, *Introducing ChatGPT Health* (2026), https://openai.com/index/introducing-chatgpt-health/.
[48] *Id.*
[49] *Id.*

results in plain language, detect patterns across fitness and health metrics, and prepare questions for appointments."[50]

Copilot Health. Microsoft launched Copilot Health in March 2026 as a space within Copilot that ingests medical records from more than 50,000 U.S. provider organizations.[51] A Microsoft executive described the product as a step toward AI with "the breadth of a general physician and the depth of a specialist."[52]

These products display each indication of intent identified in § 801.4:

- The developers' expressions;
- Design features that ingest clinical records and interpret test results; and
- Distribution of health version within general-purpose chatbot interface that consumers already use for health questions

Interpreting a patient's laboratory results against that patient's medical history, and recognizing when a user needs urgent care, for instance, are uses "in the diagnosis of disease or other conditions."[53]

Furthermore, consumers use these products precisely because of the capability of personalized medical advice, thereby supporting their intended use for diagnosis. General health information is freely available, and non-interactive, non-personalized sources such as WebMD give every reader helpful, generalized medical information. Health chatbots instead integrate a user's records, laboratory results, and wearable data to generate individualized responses.[54] OpenAI markets precisely this feature, promising responses "informed by your health information and context."[55] A consumer who connects medical records to a chatbot seeks an assessment of their own condition rather than general information available elsewhere. Applying medical knowledge to one person's clinical data to explain what may be wrong is diagnosis, and advising that person whether to seek care, or how to

---

[50] Anthropic, *Advancing Claude in Healthcare and the Life Sciences* (2026), https://www.anthropic.com/news/healthcare-life-sciences.

[51] Fierce Healthcare, Microsoft Unveils Copilot Health as an AI Health Companion for Consumers (2026), https://www.fiercehealthcare.com/ai-and-machine-learning/microsoft-unveils-copilot-health-ai-health-companion-consumers.

[52] MedCity News, Microsoft Enters Healthcare Chatbot Race with Copilot Health (2026), https://medcitynews.com/2026/03/microsoft-healthcare-chatbot.

[53] 21 U.S.C. § 321(h)(1)(B).

[54] Hemant Gupta et al., *Health Care Without the Hospital: ChatGPT Health and Claude Go Direct to Consumers*, Epstein Becker Green (2026), https://www.ebglaw.com/health-law-advisor/health-care-without-the-hospital-chatgpt-health-and-claude-go-direct-to-consumers.

[55] OpenAI, *Introducing ChatGPT Health* (2026), https://openai.com/index/introducing-chatgpt-health/.

manage symptoms in the meantime, is mitigation or treatment under FDCA § 201(h). Therefore, health chatbots that do so are medical devices.

2.2.4. Contrary disclaimers do not negate evidence of actual use when determining intent.

Under 21 CFR § 801.4, a manufacturer's statements are one source of intent among several. They cannot override the product's design, the circumstances of its distribution, and the manufacturer's knowledge of its actual use. Otherwise, any manufacturer could escape regulation by inserting a disclaimer into its terms of service. The FDA has already questioned disclaimer-based reasoning in the context of risk assessment for AI-enabled medical devices. The discussion paper considers whether including a patient-facing informational disclaimer like "talk to your doctor" or "I am not a medical professional" justifies a lower risk designation.[56] A disclaimer should not change what a product is designed and distributed to do in neither the risk assessment nor intended use analyses.

The facts of *Winters v. OpenAI, Inc.* are demonstrative of issues arising from the gap between the stated and actual use of consumer chatbots.[57] Scott Winters' interactions with ChatGPT began with benign questions, which turned into general health information requests, which then led to a gross misattribution of symptoms based on medical information he had uploaded and discouragement from seeking professional care.[58] Ultimately, he was admitted to intensive care with a massive double pulmonary embolism, likely due to a delay in care encouraged by ChatGPT's recommendations.[59] OpenAI's disclaimer did not determine how the product functioned nor how the plaintiff used it.

The conversational structure of these products induces this progression, caused by design features that optimize continuous engagement to the detriment of the user.[60] OpenAI's own suggested prompts invite users to ask

---

[56] FDA, Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback (2026), https://www.fda.gov/medical-devices/digital-health-center-excellence/considerations-regulation-generative-ai-enabled-medical-devices-discussion-paper-and-request.

[57] Tech Justice Law, *Complaint, Winters v. OpenAI, Inc.* (2026), https://techjusticelaw.org/wp-content/uploads/2026/07/Winters-Complaint.pdf.

[58] *Id.*

[59] *Id.*

[60] M. Karen Shen et al., *The AI Genie Phenomenon and Three Types of AI Chatbot Addiction: Escapist Roleplays, Pseudosocial Companions, and Epistemic Rabbit Holes*, Association for Computing Machinery (April 2026), https://dl.acm.org/doi/10.1145/3772318.3790896 (citing agreeableness, availability, customization, anthropomorphism, instantaneous feedback, limitless nature, multi-chats, personalization, and uninterrupted interactions as design elements contributing to chatbot addiction).

about trends in their data, such as "How's my cholesterol trending?".[61] An answer about trends will often identify an anomaly, and a natural next question is whether that anomaly indicates disease. The FDA already recognizes this dynamic, observing that a conversational device "may begin by performing an informational function and migrate, over the course of a conversation, to an action-directing function."[62]

OpenAI has effectively conceded the natural progression from health information to diagnosis. A peer-reviewed study found that ChatGPT Health under-triaged 51.6% of emergency cases, directing patients with conditions such as diabetic ketoacidosis to seek care within 24 to 48 hours rather than immediately.[63] OpenAI responded that the study did not reflect how the product is used; rather, ChatGPT is designed for users to ask follow-up questions that add context to a medical situation, not to produce a single response.[64] This defense is confirmation from OpenAI that it designed ChatGPT Health for an iterative, symptom-by-symptom exchange through which diagnosis occurs, despite disclaimers to the contrary.

2.3. Health Chatbots Pose a Risk to Patient Safety and Warrant Active Oversight

The FDA has stated that it "intends to apply its regulatory oversight to only those software functions that are medical devices and whose functionality could pose a risk to a patient's safety if the device were to not function as intended."[65] It reserves enforcement discretion for device software functions that "pose a lower risk to the public," such as those that help patients self-manage a condition "without providing specific treatment or treatment suggestions."[66] Health chatbots belong in the first category. FDA should say so expressly rather than leave them to the enforcement discretion that has governed them in practice.

The threshold inquiry for active oversight asks what happens to a patient if the function fails. Numerous kinds and degrees of harm ensue when a

---

[61] OpenAI, *Introducing ChatGPT Health* (2026), https://openai.com/index/introducing-chatgpt-health//.
[62] FDA, *Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback* (2026), https://www.fda.gov/medical-devices/digital-health-center-excellence/considerations-regulation-generative-ai-enabled-medical-devices-discussion-paper-and-request.
[63] Ashwin Ramaswamy et al., *ChatGPT Health Performance in a Structured Test of Triage Recommendations* (2026), https://doi.org/10.1038/s41591-026-04297-7.
[64] Kaan Ozcan, *ChatGPT Health 'Under-Triaged' Half of Medical Emergencies in a New Study*, NBC News (2026), https://www.nbcnews.com/health/health-news/chatgpt-health-under-triaged-half-medical-emergencies-rcna261409.
[65] FDA, *Policy for Device Software Functions and Mobile Medical Applications* (2022), https://www.fda.gov/media/80958/download.
[66] *Id.*

chatbot does not perform as intended by failing to give correct medical advice. First, a chatbot that misses a serious condition delays care. Second, a chatbot that misattributes symptoms directs the patient toward the wrong condition and the wrong response. Third, a chatbot that overstates risk produces anxiety and unnecessary emergency visits, burdening the healthcare system and eroding trust in appropriate care-seeking. The FDA itself recognizes both risks: under-escalation "can result in delayed treatment with potentially serious consequences," and over-escalation carries its own harms.[67]

As identified earlier, personalization increases the magnitude of failures. A chatbot's response grounded in the reader's own records, medications, and laboratory history carries the apparent authority of a clinical assessment. The FDA has identified this danger as automation bias: the risk that a user "accepts an output without appropriate scrutiny because of the device's fluency or perceived authority."[68] FDA further recognizes that patients, unlike clinicians, "may not be able to identify an incorrect output, may rely upon it, and may take downstream actions that adversely affect outcomes, including delaying needed care."[69]

OpenAI has since extended this personalization across its entire general-purpose product. In July 2026, it reported that "more than 70% of health-related conversations took place outside the dedicated health experience," and concluded that "[m]oving to a separate space added an unnecessary step."[70] ChatGPT now draws on users' connected "medications, lab results, recent visits" across all conversations.[71] OpenAI also trained the model available to free users to improve at "recognizing when urgent care may be needed."[72] The health function is now part of the general product, reaching users who never opted into a health experience to begin with.

FDA's own risk framework points to the same conclusion. The discussion paper places a function's risk along two axes: the degree to which it directs

---

[67] FDA, *Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback* (2026), https://www.fda.gov/medical-devices/digital-health-center-excellence/considerations-regulation-generative-ai-enabled-medical-devices-discussion-paper-and-request.

[68] FDA, *Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback (2026)*, https://www.fda.gov/medical-devices/digital-health-center-excellence/considerations-regulation-generative-ai-enabled-medical-devices-discussion-paper-and-request.

[69] *Id.*

[70] OpenAI, *Launching Health in ChatGPT* (2026), https://openai.com/index/health-in-chatgpt/.

[71] *Id.*

[72] *Id.*

action, and the severity of harm from an incorrect output.[73] FDA's own illustrations of high-risk functions include directing a patient on "whether to seek emergency care" and advising adjustment of "an insulin dose." Health chatbots perform these kinds of function, deliver them to lay users, and personalize them to make them persuasive. Chatbots thus sit at the high end of both axes when they fail to function as intended.

In addition, although the FDA's standard for active oversight does not require a showing that failure is likely, it is well documented that these chatbots do not perform as intended. As discussed earlier, in the first independent evaluation of ChatGPT Health, the product under-triaged 51.6% of emergency cases and over-triaged 64.8% of nonurgent ones.[74] The FDA has separately warned that generative AI devices "may confabulate, provide inappropriate or biased content, [or] fail to relay important medical information."[75] Even at a low error rate, given the millions of people who rely on chatbots for health advice, these failures produce harm to patients on a massive scale.

Lastly, these products bear no resemblance to the low-risk functions FDA has placed under enforcement discretion. The severity of harm from failure is high and the probability of failure is well-documented. FDA should therefore confirm that consumer health chatbots are device software functions subject to its active oversight, not candidates for enforcement discretion.

2.4. AI Chatbots Don't Fall into One of the Five FDCA § 520(o)(1) Exceptions

FDCA § 520(o)(1) removes five categories of software functions from the device definition.[76] The categories include software functions for: scheduling and billing, lifestyle support, record storage, data display, or professional decision support.

While this structure fits conventional software, which performs a fixed function that a developer can describe in advance, it is a poor fit for chatbots due to their multipurpose use. A single conversation may begin with nutritional information, turn to the meaning of a laboratory result, and end

---

[73] FDA, *Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback* (2026), https://www.fda.gov/media/194242/download.
[74] Ashwin Ramaswamy et al., *ChatGPT Health Performance in a Structured Test of Triage Recommendations* (2026), https://doi.org/10.1038/s41591-026-04297-7.
[75] FDA, *Executive Summary for the Digital Health Advisory Committee Meeting: Generative Artificial Intelligence-Enabled Digital Mental Health Medical Devices* (2025), https://www.fda.gov/media/189391/download.
[76] 21 U.S.C. § 360j(o)(1).

with advice on whether to go to the emergency room. One product therefore performs excluded and non-excluded functions interchangeably.

However, as discussed earlier, regulation turns on function, rather than a platform or product itself. Thus, a product that performs any non-excluded device function remains subject to regulation as to that function. A chatbot's occasional performance of an excluded function therefore cannot shelter its diagnostic and treatment functions. Even taken on their own terms, as the following subsections show, none of the five exclusions apply to AI chatbots.

2.4.1. (A) Administrative support

Subparagraph (A) excludes functions intended "for administrative support of a health care facility," including billing, appointment scheduling, practice management, and laboratory workflow. 21 U.S.C. § 360j(o)(1)(A). The exclusion serves the operations of an institution, whereas consumer health chatbots serve individual patients. Some developers do sell administrative tools to providers and payers, such as prior authorization review.[77] Those enterprise functions may qualify, but they are distinct from the consumer functions at issue here.

2.4.2. (B) General wellness

Subparagraph (B) excludes functions intended for either "maintaining or encouraging a healthy lifestyle" or help to "reduce the risk or impact of certain chronic diseases" where healthy lifestyle choices are "well understood" to "play an important role."[78] Additionally, such products must be "unrelated to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition."[79] Importantly, the FDA has noted that general wellness products must also meet two factors: they are intended for only general wellness use and present a low risk to safety of users and others.[80]

Chatbots fail to meet the requirements for general wellness products in several ways. First, a health chatbot's functions relate directly to diagnosis, even if it also helps a user with their healthy lifestyle. Second, chatbots are inherently multipurpose, thus preventing an exclusive use for general wellness. Lastly, chatbots certainly do not present a low risk of safety as demonstrated in Scott Winters' case and others brought by TJL. Describing

---

[77] *See, e.g.*, Anthropic, *Advancing Claude in healthcare and the life sciences* (2026), https://www.anthropic.com/news/healthcare-life-sciences.
[78] FDA, *General Wellness: Policy for Low Risk Devices* (Jan. 2026), https://www.fda.gov/media/90652/download.
[79] 21 U.S.C. § 360j(o)(1)(B).
[80] FDA, *General Wellness: Policy for Low Risk Devices* (Jan. 2026), https://www.fda.gov/media/90652/download.

these chatbots as "health and wellness" in marketing does not satisfy the requirement.

2.4.3. (C) Electronic patient records

Subparagraph (C) excludes functions intended to serve as electronic patient records, but only if three conditions hold:

- the records were created, stored, transferred, or reviewed by health care professionals;
- the records are part of certified health information technology; and
- the function is "not intended for interpreting or analyzing patient records… for the purpose of the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition."

Health chatbots do not meet any of these conditions. Consumers, not professionals, connect and query the records, the chatbots are not certified health IT, and again, interpreting connected records is a core feature of these chatbots.

2.4.4. (D) Data display and transfer

Subparagraph (D) excludes functions intended to transfer, store, convert, or display laboratory and device data, "unless such function is intended to interpret or analyze" that data.[81] Health chatbots do more than merely display lab/device data: they explain what results mean, identify trends, and relate the data to user health.

2.4.5. (E) Clinical decision support

Subparagraph (E) excludes clinical decision support software, but only if the function:

- does not analyze medical images or diagnostic signals;
- provides recommendations "to a health care professional"; and
- enables "such health care professional to independently review the basis for such recommendations," so that the professional does not "rely primarily" on them.[82]

Patient-facing health chatbots fail at all requirements. First, chatbots are capable of analyzing medical images. Second, their recommendations go to consumers/patients, not to health care professionals. Even if healthcare

[81] 21 U.S.C. § 360j(o)(1)(D).
[82] 21 U.S.C. § 360j(o)(1)(E); FDA, *Clinical Decision Support Software* (2022), https://www.fda.gov/media/109618/download.

professionals were the recipients, no review process exists in the manner contemplated by the statute.

2.5. Alternatively, FDCA § 520(o)(3) Authorizes the Secretary of HHS to Make a Finding of Adverse Health Consequences

Even if a health chatbot function fell within subparagraph (C), (D), or (E), it may not necessarily remain excluded. Under FDCA § 520(o)(3), such a function "shall not be excluded from the definition of device" if the Secretary finds that its use "would be reasonably likely to have serious adverse health consequences" and identifies the function in a final order.[83] The statute directs the Secretary to consider four factors in making such a finding. Each supports bringing a chatbot back into the regulatory scope here.

The first factor asks about "the likelihood and severity of patient harm if the software function were to not perform as intended."[84] As established, chatbot failures in the health context meet both elements: documented triage failures, a potential scale in the hundreds of millions, and harms ranging from delayed emergency care to unnecessary treatment. Scott Winters and others' stories shows that these risks have already caused near-fatal harm.

The second factor asks about "the extent to which the software function is intended to support the clinical judgment of a health care professional."[85] Consumer health chatbots do not support the professional judgment of a healthcare provider, but serve as a substitute for it. The risk mitigated by a professional-in-the-loop is not present; therefore, the exception should not apply.

The third factor asks "whether there is a reasonable opportunity for a health care professional to review the basis of the information or treatment recommendation."[86] No professional necessarily reviews a consumer chatbot's output and generative outputs do not always disclose the sources of its information/recommendation.

The fourth factor asks about "the intended user and user environment."[87] The intended users are lay consumers and the environment is unsupervised and all-encompassing. The FDA's existing practice treats these conditions as aggravating risk. Most authorized digital mental health devices require a prescription because their use is not safe "except under the supervision of a practitioner," and a manufacturer may market a device over the counter only

[83] 21 U.S.C. § 360j(o)(3)(A).
[84] 21 U.S.C. § 360j(o)(3)(C)(i).
[85] 21 U.S.C. § 360j(o)(3)(C)(ii).
[86] 21 U.S.C. § 360j(o)(3)(C)(iii).
[87] 21 U.S.C. § 360j(o)(3)(C)(iv).

after demonstrating that consumers can use it "effectively and safely without a prescriber."[88] Health chatbots reach consumers directly without any such showing.

All four factors therefore support a finding. To the extent any health chatbot function falls within subparagraphs (C) through (E), FDA should initiate a § 520(o)(3) proceeding.

3. THE FDA IS BEST PLACED TO REGULATE CONSUMER AI CHATBOTS

Recognizing that the harms of consumer AI chatbots offering medical guidance warrant regulation as medical devices, this section outlines the practical reasons why the FDA, as opposed to current state regulatory tools, is best placed to mitigate these harms.

3.1. FDA regulation ensures consistent, preventative and expert protection of AI chatbots

We believe the most effective harm prevention regulatory approach for AI chatbots is a coordinated federal approach led by the FDA. The FDA is well-positioned to be the leading federal regulator for several reasons, including:

1. **Consistent and Uniform Regulation:** By including AI chatbots within the definition of "medical devices" and thereby applying the FDA's national regulatory framework, the FDA can eliminate the current piecemeal and overlapping web of state AI and chatbot legislation, litigation and state medical board oversight and make it easier for developers to comply with legal requirements.
2. **Harm Prevention:** The FDA's multi-faceted powers to require pre-market approval of devices and ongoing monitoring enables the FDA to take action before harm occurs through early investigations.
3. **Technical Expertise:** The FDA has been regulating various AI-enabled medical devices since as early as 1995.[89] This means FDA reviewers have "a deep understanding of the discipline to be able to appropriately review applications" for market approval.[90] It is a

[88] FDA, *Executive Summary for the Digital Health Advisory Committee Meeting: Generative Artificial Intelligence-Enabled Digital Mental Health Medical Devices* (2025), https://www.fda.gov/media/189391/download.
[89] FDA, *List of Artificial Intelligence-Enabled Medical Devices* (2026), https://www.fda.gov/medical-devices/artificial-intelligence-enabled-medical-devices/list-artificial-intelligence-enabled-medical-devices
[90] Haider J. Warraich, Troy Tazbaz and Robert M. Califf, *FDA Perspective on the Regulation of Artificial Intelligence in Health Care and Biomedicine* (2025), https://pubmed.ncbi.nlm.nih.gov/39405330/

natural extension to include the regulation of AI chatbots offering medical guidance, given its harms are similar to other AI-enabled medical devices previously regulated by the FDA.

4. **Testing & Evaluation Capacity:** The FDA has already built the infrastructure needed to review AI chatbots, and has the capacity to evaluate product design, monitor market products, conduct inspections, foresee user risks, and remove dangerous products from the market.[91]

3.2. Current state regulatory tools are ill-equipped to protect against the harms of AI chatbots

Driven by concerns that AI chatbots are "providing wrong or potentially dangerous [medical] advice" to their citizens, states have started implementing different regulatory approaches to address the harms faced by their citizens.[92] The main regulatory tools currently being utilized include: (1) enacting new state legislation; (2) applying existing legislative and common law protections to AI harms through litigation; and (3) empowering state medical boards to set best practice guidelines.

3.2.1. Both broad AI and narrow chatbot state legislation create patchwork harm protection and high compliance costs.

In response to the absence of "clear fit-for-purpose, flexible federal regulatory pathways" for consumer AI chatbots,[93] some states have implemented sweeping laws that regulate AI broadly, while others have adopted more narrow laws specifically targeting health care applications of AI chatbots.[94] For example, Colorado's Automated Decision-Making Technology ("ADMT") Act imposes a singular regulatory framework, which imposes notice, disclosure and human review requirements, for the use of all ADMT in "consequential decisions" including employment, housing and health care.[95] Contrastingly, Utah's Artificial Intelligence Policy Act

---

[91] FDA, *About FDA: Patient Q&A* (2024), https://www.fda.gov/media/151975/download
[92] Joel Luther, Hagere Yilma and Irving Washington, *Different State Regulatory Approaches Reflect Open Questions About AI Mental Health Tools* (2026), https://www.kff.org/health-information-trust/different-state-regulatory-approaches-reflect-open-questions-about-ai-mental-health-tools/.
[93] Case Thomason et al. *The History of State Preemption and Medical Device Regulation: Lessons for Artificial Intelligence Oversight* (2026), https://academic.oup.com/healthaffairsscholar/article/4/3/qxag046/8503715.
[94] Epstein Becker Green, *State AI Law Tracker* (2026), https://www.ebglaw.com/insights/publications/state-ai-law-tracker .
[95] Colorado Senate Bill 26-189.

establishes disclosure, policy and personal information protections "for users of mental health chatbots that use artificial intelligence technology."[96]

Both legislative approaches have significant limitations. First, as of October 2026, 1,761 separate AI-related bills have been introduced across all 50 states.[97] This has resulted in a "patchwork of disclosure requirements, bias prohibitions, sector-specific mandates, and compliance deadlines that can vary by jurisdiction and change without warning".[98] This increases the difficulty and cost of compliance for companies, resulting in more limited practical protection. The "patchwork" approach also risks creating gaps in consumer protection. Second, state-based legislation encourages forum shopping. Technology companies may shift operations to states with less restrictive oversight, "leading to patients in certain states experiencing both differential access and differential product safety profiles".[99] Third, most state AI and chatbot laws vest enforcement power exclusively to state attorney generals, who are under resourced and only practically able to investigate serious misconduct.[100]

3.2.2. State litigation is slow, uncertain and evidentially burdensome.

State litigation involving health-related AI applications, including chatbot exchanges, has increased in recent years. These lawsuits predominantly assert claims under existing "state laws related to consumer protection, privacy and products liability", and are in the early phases of litigation.[101] While litigation may be a helpful tool in achieving individual justice for AI-caused harms in some instances, without strong, complimentary federal regulation, its weaknesses are stark. First, litigation is slow and uncertain. With current litigation matters not yet resulting in merit findings, "the extent to which such suits can provide remedies to injured consumers is unclear".[102] Second, litigation is only available once the harm has already occurred. This means it is an ineffective tool on its own to advance broader systemic change and ensure future protection. Third, litigation imposes a

---

[96] Utah Bill H.B. 452.
[97] AI Laws by State, *Which AI Laws Apply to You, and When.* (2026), https://www.ailawsbystate.com/.
[98] Epstein Becker Green, *State AI Law Tracker* (2026), https://www.ebglaw.com/insights/publications/state-ai-law-tracker.
[99] Case Thomason et al, *The History of State Preemption and Medical Device Regulation: Lessons for Artificial Intelligence Oversight* (2026), https://academic.oup.com/healthaffairsscholar/article/4/3/qxag046/8503715.
[100] Tracey Birkenhauer, *State Chatbot Laws Share Common Skeleton* (2026), https://stackcyber.com/posts/ai-chatbot-laws.
[101] Congress, *Artificial Intelligence in Health: Overview of Selected State Liability Frameworks* (2026), https://www.congress.gov/crs-product/LSB11467
[102] *Id.*

difficult evidentiary burden for injured users because of AI's "black box" problem which can conceal its potential contribution to the injury.[103] Fourth, similar to state legislation, litigation often creates patchwork protection for harms, given courts are restricted to relying on "pre-existing legal doctrines to manage AI harms". This can also lead to certain AI harms remaining unresolved.[104]

3.2.3. State medical boards, through both litigation and policy setting, lack the enforcement power and resources necessary to protect against the harms of AI chatbots.

State medical boards are the "agencies that license medical doctors, investigate complaints … [and] discipline physicians who violate the [state] medical practice act".[105] Because they operate independently from one another and set their own standards, this creates a similar patchwork protection limitation as described above for state legislation and litigation.

Further, due to significant resource constraints, litigation brought by state medical boards for breach of the state medical practice act, can only target serious misconduct. For example, in May 2026, the Pennsylvania State Board of Medicine ("PSBM") filed a lawsuit against Character Technologies for breach of Pennsylvania's Medical Practice Act, alleging that its AI chatbot "Emilie" who purported to be a "doctor of psychiatry" with credentials including "medical school at Imperial College London, [and] seven years of practice", engaged in the unlawful practice of medicine.[106] Notably, this case involves an overt and egregious breach by an AI chatbot expressly holding itself out as a licensed medical practitioner. Due to the practical limitation in investigating all potential complaints and harms raised by AI chatbots, more nuanced uses of AI chatbots which cause harm will likely not be protected. The same limitations of state litigation outlined above also apply.

State medical boards also regulate AI chatbots by coordinating and agreeing to voluntary guardrails through the Federation of State Medical Boards ("FSMB"), a non-profit organization that "serve[s] as a national voice for

---

[103] Case Thomason, Larry Bucshon, Brian J. Miller and Brian Yagi, *The History of State Preemption and Medical Device Regulation: Lessons for Artificial Intelligence Oversight* (2026), https://academic.oup.com/healthaffairsscholar/article/4/3/qxag046/8503715
[104] Julie Yu et al, *Visible to the Court: How AI Is (and Isn't) Litigated in U.S. Federal Court Opinions* (2026), https://arxiv.org/html/2607.23888v1.
[105] Drew Carlson & James N. Thompson, *The Role of State Medical Boards* (2005), https://journalofethics.ama-assn.org/article/role-state-medical-boards/2005-04.
[106] Epstein Becker Green, *'Emilie' Is Not a Psychiatrist: Pennsylvania Board of Medicine Alleges Unlawful Practice of Medicine by an AI Chatbot* (2026), https://www.ebglaw.com/health-law-advisor/emilie-is-not-a-psychiatrist-pennsylvania-board-of-medicine-alleges-unlawful-practice-of-medicine-by-an-ai-chatbot.

state medical boards".[107] While the FSMB can issue model policies on AI, as an advisory body, it has no enforcement power to ensure that state medical boards comply with those policies.[108] For example, the FSMB's new Workgroup on the Regulation of AI in the Practice of Medicine will be unable to enforce any of the "recommendations and/or model guidelines" it develops "for use by state medical boards regarding the regulation of AI tools in the practice of medicine", in particular those that "perform clinical functions with limited or no direct physician supervision".[109]

4. THE FDA SHOULD MANDATE EXISTING PREMARKET GUIDANCE FOR AI CHATBOTS AND INTRODUCE ADDITIONAL DEVICE SPECIFIC SAFEGUARDS

We make three recommendations on how the FDA should regulate consumer AI chatbots.

1. **Continue risk-based regulatory approach:** The FDA should continue to employ a risk-based regulatory approach to AI-enabled medical devices, "taking into consideration the device's intended use and technological characteristics". This involves premarket clearance and approval mechanisms, review processes for device modifications, and labeling and reporting requirements.[110]
2. **Mandate existing nonbinding best practices guidance for AI chatbots:** The FDA regularly publishes guidance outlining best practices for performance testing and other safeguards useful in assessing whether a device qualifies for premarket clearance. We have outlined below where these currently nonbinding guidelines should be made mandatory for AI chatbots seeking premarket clearance.
3. **Introduce device specific safeguard requirements:** The FDA should require that AI chatbots seeking premarket clearance have the following device specific safeguards: (1) improving testing rigor of safety evaluations; (2) ensuring appropriate human oversight; and (3)

---

[107] Federation of State Medical Boards, *About FSMB* (2026), https://www.fsmb.org/about-fsmb.
[108] *Id.*
[109] Federation of State Medical Boards, *Federation of State Medical Boards Announces New Workgroup on Regulation of AI in Medical Practice* (2026), https://www.fsmb.org/advocacy/news-releases/federation-of-state-medical-boards-announces-new-workgroup-on-regulation-of-artificial-intelligence-in-medical-practice.
[110] FDA, *Artificial Intelligence-Enabled Medical Devices* (2026), https://www.fda.gov/medical-devices/digital-health-center-excellence/artificial-intelligence-enabled-medical-devices.

increasing transparency through specific disclosures. We address each of these in greater detail below.

4.1. Improving testing rigor of safety evaluations

We recommend that the performance validation considerations outlined in Appendix C of the FDA's 2025 draft guidance on lifecycle management and marketing submissions recommendations for AI-enabled device software functions ("2025 Draft Guidance") be made mandatory for consumer AI chatbots.[111] This is necessary because, due to the need to accommodate a "broad spectrum of medical devices with varying clinical contexts", the FDA's current premarket review process does not specify the validation rigor, that is, "the quality and comprehensiveness of the evidence supporting a device's performance", required for AI-enabled medical devices.[112]

We also recommend that the FDA require additional standards for safety evaluations of consumer AI chatbots before granting premarket clearance. This is critical because recent scientific studies have shown that current safety evaluation methods for consumer AI chatbots offering medical guidance are falling short because they "do not reflect the complexity of interacting with human users".[113] For example, in 2026, performance testing of ChatGPT Health found that the tool "missed high-risk emergencies and [caused] inconsistent activation of crisis safeguards".[114] The proposed requirements for safety evaluations are:

1. **Utilize real, diverse data:** AI chatbots should be evaluated on real and diverse clinical data,[115] such as hospital discharge records and

---

[111] FDA, *Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations* (2025), https://www.fda.gov/media/184856/download.

[112] Abdul Rahman Diab & William Lotter, *Distinguishing Between Rigor and Transparency in FDA Marketing Authorization of AI-Enabled Medical Devices* (2025), https://pubs.rsna.org/doi/10.1148/ryai.250369.

[113] University of Oxford, *New Study Warns of Risks in AI Chatbots Giving Medical Advice* (2026), https://www.ox.ac.uk/news/2026-02-10-new-study-warns-risks-ai-chatbots-giving-medical-advice.

[114] Ashwin Ramaswamy, Alvira Tyagi & Girish N. Nadkarni, *ChatGPT Health Performance in a Structured Test of Triage Recommendations* (2026), https://www.nature.com/articles/s41591-026-04297-7

[115] Abdul Rahman Diab & William Lotter, *Distinguishing Between Rigor and Transparency in FDA Marketing Authorization of AI-Enabled Medical Devices* (2025), https://pubs.rsna.org/doi/10.1148/ryai.250369.

clinical consultation notes.[116] Training data should also be limited to current, reputable and peer-reviewed medical databases.

2. **Adopt broad testing scenarios:** The breadth of testing scenarios should reflect the complexity of diagnoses made in medical practice.[117]
3. **Deploy reader studies (i.e. real user testing):** Standardized testing alone is insufficient. Testing should occur with diverse, real users to ensure greater accuracy.[118]
4. **Add and test safety reminder prompts:** Researchers have found that short safety prompts that reminds an AI model, before it answers, to pause and consider whether the proposed response is harmful lowers harmful clinical choices.[119]

4.2. Ensuring appropriate human oversight

Ensuring appropriate human oversight at the design, deployment and post market monitoring stages for generative AI chatbots is particularly critical in the medical context. This is because inaccurate diagnoses or missed conditions produced by these chatbots can only be detected and remedied by human medical practitioners. Further, because chatbots are often trained on biased data, human medical practitioners are also needed to assess the accuracy of outputs by reference to context cues and human emotions that a chatbot is unable to understand.[120] However, it is important that the FDA stay vigilant to the practice of licensure rental, that is, where companies employ licensed medical professionals only for the use of their license rather than for genuine human oversight of the AI chatbot.

We recommend that the FDA's 2026 guidance on applying human factors and usability engineering to medical devices be made mandatory for AI consumer chatbots. This guidance offers critical recommendations on

---

[116] Edwin Kwan-Yeung Chiu & Tom Wai-Hin Chung, *Protocol for Human Evaluation of Generative Artificial Intelligence Chatbots in Clinical Consultations* (2025), https://pmc.ncbi.nlm.nih.gov/articles/PMC11922213/.

[117] University of Oxford, *New Study Warns of Risks in AI Chatbots Giving Medical Advice* (2026), https://www.ox.ac.uk/news/2026-02-10-new-study-warns-risks-ai-chatbots-giving-medical-advice.

[118] *Id.;* Abdul Rahman Diab and William Lotter, *Distinguishing Between Rigor and Transparency in FDA Marketing Authorization of AI-Enabled Medical Devices* (2025), https://pubs.rsna.org/doi/10.1148/ryai.250369.

[119] The Mount Sinai Hospital, Safety Prompts Can Help AI Models Make Safer Clinical Choices (2026), https://medicalxpress.com/news/2026-10-safety-prompts-ai-safer-clinical.html.

[120] Cornerstone, *The Crucial Role of Humans in AI Oversight* (2026), https://www.cornerstoneondemand.com/resources/article/the-crucial-role-of-humans-in-ai-oversight/.

conducting effective human factors validation testing, that is, testing that shows a "device can be used by the intended users without serious errors or problems, for the intended uses and under the expected use conditions".[121]

Given the FDA does not currently mandate specific device design safeguards or post market monitoring requirements, we also recommend the following specific standards are imposed on AI chatbots:

1. **Human escalation protocol:** The FDA should require that AI chatbots are programed to redirect users to a human medical practitioner when (1) a user asks to speak with a human; (2) a user's medical symptoms are ambiguous, multi-faceted, severe or challenging; and (3) a user is emotionally distressed.[122]
2. **Post market human monitoring:** The FDA should mandate that human medical practitioners are employed to regularly review chat logs to evaluate response accuracy and determine whether further escalation triggers need to be added.
3. **Incident reporting:** The FDA currently requires manufacturers to report all incidents where a medical device "may have caused or contributed to a death or serious injury".[123] This rule should be extended to include incidents causing repeated or minor injuries and require root cause analysis and technical fixes to avoid future injuries.[124]

4.3. Increasing transparency through specific disclosures

We recommend that section VI of the FDA's 2025 Draft Guidance be made mandatory for AI consumer chatbots. This guidance outlines useful labeling expectations for AI-enabled medical devices including a clear statement AI is used in the device, an explanation of the model inputs and outputs, the

---

[121] FDA, *Applying Human Factors and Usability Engineering to Medical Devices* (2026), https://www.fda.gov/media/80481/download

[122] Intelligent Office Suite, *When AI Can't Handle It, a Human Steps In* (2026), https://intelligentofficesuite.com/human-escalation-healthcare-ai.

[123] FDA, *Mandatory Reporting Requirements: Manufacturers, Importers and Device User Facilities* (2025), https://www.fda.gov/medical-devices/postmarket-requirements-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities.

[124] Lexic.AI, *AI Incident Reporting: Map OECD, ETSI, and NIST to Real Failures* (2026), https://lexic.ai/the-signal/ai-incident-reporting#what-ai-incident-reporting-covers-and-when-to-report.

intended performance monitoring tools, and the known limitations of the device. [125]

We also recommend that the FDA require the following consumer-facing disclaimers for AI consumer chatbots specifically:

1. For all medical advice provided by AI consumer chatbots, a detailed disclaimer that the information provided is general and that it is recommended that the user sees a medical practitioner for further assistance.
2. For complex medical scenarios or where a user has asked repeated questions about the same medical condition, a detailed disclaimer that AI chatbot is unable to assist because the user's medical problem is challenging and recommend that the user see a medical practitioner for further assistance.

5. CONCLUSION

Tech Justice Law agrees with the discussion paper's belief that GenAI-enabled devices will likely transform the healthcare industry. With this recognition, the FDA must take on the responsibility of protecting consumers who use AI chatbots for medical advice. We believe that not only does the FDA have the jurisdiction over chatbot regulation, but we also believe that the FDA is the best suited entity to regulate the AI chatbots. Our recommendations ask the FDA to focus regulation on improving testing rigor of safety evaluations, ensuring appropriate human oversight, and increasing transparency through specific disclosures. With a patient-oriented approach, TJL believes that the FDA can best ensure that consumers of AI chatbots are protected from the harms that our clients and general consumers are currently facing. We appreciate the opportunity to comment and offer to be a resource as regulation is considered and created.

[125] FDA, *Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations* (2025), https://www.fda.gov/media/184856/download/.