FDA U.S. FOOD & DRUG ADMINISTRATION
CENTER FOR DEVICES & RADIOLOGICAL HEALTH

EXECUTIVE SUMMARY FOR THE DIGITAL HEATH ADVISORY COMMITTEE MEETING

Generative Artificial Intelligence-Enabled Digital Mental Health Medical Devices

November 6, 2025

U.S. FOOD & DRUG ADMINISTRATION

Digital Health Advisory Committee (DHAC) Executive Summary

Generative Artificial Intelligence-Enabled Digital Mental Health Medical Devices

TABLE OF CONTENTS

Disclaimer: This Executive Summary is for discussion purposes only and does not represent draft or final guidance. It is not intended to propose or implement policy changes regarding regulation of digital mental health medical devices. The references cited herein are for informational purposes and should not be construed as endorsements.

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Introduction

On November 20, 2024, the FDA's Digital Health Advisory Committee (DHAC) met to discuss and provide recommendations on how the use of generative artificial intelligence (AI) may impact safety and effectiveness of medical devices enabled with this technology. This was the first meeting of the recently established Advisory Committee to address digital health technologies. The topics under discussion at that meeting encompassed premarket performance evaluation, risk management, and postmarket performance monitoring for generative AI-enabled devices. Although FDA has long promoted a total product life cycle (TPLC) approach to the oversight of medical devices (including AI-enabled devices), the Committee recognized the importance of optimizing the TPLC approach for modern medical devices that incorporate technologies that are more complex and intended to iterate faster and more frequently over a device's life of use. This is relevant for products that generate new content, provide unbounded outputs for a set of inputs, and rely on complex foundation models that may not themselves be medical devices but are intended to change rapidly over time.

Along with the rise of widely accessible generative AI products for general purposes, we are seeing an increase in the development and demand for a new kind of digital mental health medical device: "AI therapists" and other AI-based medical devices offering to provide a wide range of mental health therapies and interactions (some even being diagnostic) with therapist/healthcare provider-like chatbots. These chatbots may engage with users in individualized ways with, or without, the oversight of a health care provider (HCP), which introduce novel risks for use. As digital mental health medical devices continue to evolve in complexity, regulatory approaches will need to accommodate these challenges and opportunities to provide a reasonable assurance of their safety and effectiveness while promoting innovation to support public health.

Therefore, this DHAC meeting builds on the 2024 DHAC discussion topics (Appendix A) and is focused on the uniquely patient-facing aspect of generative AI-enabled technologies: digital mental health medical devices to treat and/or diagnose psychiatric conditions. The feedback generated from this meeting will help the Agency better facilitate innovation in this field while safeguarding patients. FDA is committed to helping digital health innovators bring safe and effective medical devices to market in alignment with agency priorities of supporting efficiency and transparency in review processes, and to deliver meaningful treatments for patients.[1]

[1] Martin A. Makary, MD, MPH & Vinay Prasad, MD, MPH, *Priorities for a New FDA*, JAMA (June 10, 2025), https://doi.org/10.1001/jama.2025.10116

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Background

Currently, 57.8 million adults have diagnosed mental illnesses[2], representing a significant number of the American population that need, and often fail to receive, consistent treatment.[3] Nationally, the percentage of patients with mental health diagnoses increased by 39.8% between 2019 and 2023, an increase from 13.5% to 18.9% of patients who received medical services during that time.[4] The increases in both diagnostic and mental health service needs highlight the importance of improved access to high quality and effective treatments for psychiatric disorders.

In recent years, there has been a marked surge of digital mental health technologies to address general health, wellness, and emotional needs, many in the form of apps. These apps and platforms are used by individuals seeking self-help, by professionals as adjuncts to clinical care, or by healthcare systems to enhance care delivery. However, the majority of these commercially available products reside in the public marketplace (e.g., app stores) as consumer wellness apps, and are not reviewed or authorized by FDA.[5,6] As described in the guidance, Policy for Device Software Functions and Mobile Medical Applications, FDA intends to apply its regulatory oversight to 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. This approach applies to generative AI-enabled products as well. For example, as with all software, the regulatory status paradigm of generative AI-enabled products is a spectrum from those that are not devices and are not within FDA's regulatory purview to those that are devices and are the focus of FDA's oversight. To further support innovation of generative AI-enabled digital mental health medical devices, FDA is committed to clarifying the regulatory pathway and applying least burdensome requirements to ensure the safety and effectiveness of these technologies.

For this DHAC meeting, "digital mental health medical devices" will refer to digital products or functions (including those utilizing AI methods) that are intended to diagnose, cure, mitigate,

---

[2] Mental illness for the purposes of this meeting is defined as any mental, behavioral, or emotional disorder that meets the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) definition of a mental disorder. Criteria for a mental disorder include clinically significant disturbances in cognition, emotional regulation, or behaviors that are associated with significant distress or disability in social or occupational functioning. Categorically, mental illnesses include but are not limited to, anxiety disorders (e.g., panic), mood disorders (e.g., bipolar, depression), neurodevelopmental disorders (e.g., autism), substance use disorders (e.g., opioids), and thought disorders (e.g., schizophrenia). Within each category of illnesses there are also diagnostic specifiers of illness features and symptom severity (mild, moderate, severe) that inform treatment recommendations and prognoses.

[3] Substance Abuse and Mental Health Services Administration. (2022). *Key substance use and mental health indicators in the United States: Results from the 2021 National Survey on Drug Use and Health* (HHS Publication No. PEP22-07-01-005, NSDUH Series H-57). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. https://www.samhsa.gov/data/report/2021-nsduh-annual-national-report

[4] FAIR Health. (2024). *Trends in Mental Health Conditions: An Analysis of Private Healthcare Claims.* Available at: https://s3.amazonaws.com/media2.fairhealth.org/whitepaper/asset/Trends%20in%20Mental%20Health%20Conditions%20-%20A%20FAIR%20Health%20White%20Paper.pdf

[5] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-wellness-policy-low-risk-devices

[6] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

treat or prevent a psychiatric condition, including uses that increase a patient's access to mental health professionals. The questions posed within this discussion are designed to assist the Agency in determining the critical information and practices needed for a comprehensive approach to the evaluation of benefit and the management of risks throughout the TPLC[7] of digital mental health medical devices. The discussion seeks to evaluate digital mental health medical devices through the lens of the pre- and post-market information needed to support the safety and effectiveness of these devices.

Overview of Digital Mental Health Technologies

Digital mental health technology is a broad category that encompasses mobile health, health information technology, wearable devices, telehealth, telemedicine, and personalized medicine. Recent estimates indicate that there are tens of thousands of applications available in the marketplace to monitor and support mental health.[8] Some are integrated with wearable devices that can track consumer habits and can monitor wellness data including physical activity and sleep patterns. Others employ gamification, technology that incorporates game mechanics into mental health interventions through the development of interactive experiences. Other technologies used in mental health include features that analyze speech, text, and facial expressions, among other tasks. More recently, Large Language Model (LLM)-based functionalities and generative AI have been increasingly utilized in different mental health contexts.

However, the term 'digital mental health technology' is also used to refer to digital therapeutics and diagnostics, which are under the purview of FDA's oversight. The spectrum of medical devices and consumer general wellness apps related to mental health has been a source of confusion to users who may not know the difference. FDA-authorized digital mental health medical devices typically include a requirement for prescription by an HCP and are designed to treat, augment the treatment of, or diagnose a psychiatric condition. The scope of this DHAC is focused on novel generative-AI digital mental health medical devices that are the focus of FDA's oversight.

Although the Agency has authorized over 1200 AI-enabled medical devices[9] (encompassing a wide range of AI technologies), none of those AI-enabled devices have been authorized for mental health uses. To date, FDA has authorized fewer than twenty digital mental health medical devices that encompass non-AI technologies. As interest in and development of generative AI-enabled mental health applications (e.g., chatbots, virtual companions, healthcare automation, predictive modeling) across the healthcare sector increases, public health questions have emerged regarding these products' safety and capability to deliver therapeutic content, diagnose mental health conditions, or substitute for a mental healthcare provider. FDA

[7] https://www.fda.gov/about-fda/cdrh-transparency/total-product-life-cycle-medical-devices

[8] Kaveladze BT, Wasil AR, Bunyi JB, Ramirez V, Schueller SM. User Experience, Engagement, and Popularity in Mental Health Apps: Secondary Analysis of App Analytics and Expert App Reviews. JMIR Hum Factors. 2022 Jan 31;9(1):e30766. doi: 10.2196/30766. PMID: 35099398; PMCID: PMC8844980.

[9] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

is seeking feedback on approaches to and considerations for regulating generative AI-enabled digital mental health medical devices.

Opportunities and Challenges

News and social media outlets increasingly report trends of positive user experiences with generative AI technologies.[10,11,12,13] The ease of use, sense of privacy, interactive qualities, perceived benefits, and convenience associated with these technologies are some of the reasons for their widespread popularity and use.

The application of generative AI in digital mental health technologies may help transform mental health treatment for the general population by addressing critical gaps in traditional mental health services (e.g., supporting traditional mental health care, improving access to care, and potentially addressing HCP shortages).

However, unique challenges exist when considering the development and use of generative AI in digital mental health medical devices.[14,15,16] The device may confabulate, provide inappropriate or biased content, fail to relay important medical information, or decline in model accuracy (data drift). A patient may misinterpret device outputs or become more symptomatic with device use. A physician or HCP may not understand how to monitor or oversee use of the technology in their practice or may not be included in the patient's use of it, if oversight is not an included feature of the device.

Further, the degree to which an HCP is involved (a human-in-the-loop vs an autonomous system) in the delivery or monitoring of generative AI-enabled digital mental health medical devices that provide diagnostic and/or therapeutic functions, is an important consideration related to benefits, risks, and mitigations. Although the Agency is experienced with regulating

[10] Kim Tingly, Kids Are in Crisis. Could Chatbot Therapy Help?, New York Times (June 20, 2025), https://www.nytimes.com/2025/06/20/magazine/ai-chatbot-therapy.html

[11] Stuart Heritage, 'I felt pure, unconditional love': the people who marry their AI chatbots, The Guardian (Jul 12, 2025), https://www.theguardian.com/tv-and-radio/2025/jul/12/i-felt-pure-unconditional-love-the-people-who-marry-their-ai-chatbots

[12] Webb Wright, People are using AI to 'sit' with them while they trip on psychedelics, MIT Technology Review (Jul 15, 2025), https://www.technologyreview.com/2025/07/01/1119513/ai-sit-trip-psychedelics

[13] I took a Decision Holiday and Put A.I. in Charge of My Life, New York Times (November 01, 2024), https://www.nytimes.com/interactive/2024/11/01/technology/generative-ai-decisions-experiment.html

[14] Marco Quiroz-Gutierrez, Gen Z is increasingly turning to ChatGPT for affordable on-demand therapy, but licensed therapists say there are dangers many aren't considering, Fortune, (June 1, 2025) https://fortune.com/2025/06/01/ai-therapy-chatgpt-characterai-psychology-psychiatry/

[15] Sarah Wells, Exploring the Dangers of AI in Mental Health Care, Stanford University HAI (June 11, 2025) https://hai.stanford.edu/news/exploring-the-dangers-of-ai-in-mental-health-care

[16] Ruben Circelli, Don't Trust Grok for Medical Advice. I Tested Its Therapist Persona, and the Answers Were Terrifying, PC Mag (August 11, 2025) https://www.pcmag.com/opinions/dont-trust-grok-for-medical-advice-i-tested-its-therapist-persona-and-the

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

physiologic closed-loop medical devices,[17] the evaluation of autonomously functioning devices for digital mental health may warrant different considerations tailored to the specific technology and use. For this Committee meeting, FDA is seeking feedback on the opportunities and challenges presented when a digital mental health medical device is designed for use with a human in the loop and for scenarios when such a device is designed as an autonomous system.

How FDA Regulates Digital Mental Health Medical Devices

FDA's regulatory oversight of medical devices follows a risk-based approach with consideration of the device's intended use and technological characteristics. This risk-based approach is applied to digital mental health technologies that meet the definition of a device and are the focus of FDA's oversight. The following sections describe how FDA regulates digital mental health medical devices. These sections highlight and answer common questions received about how to interpret FDA's regulations and policies, including where they may be challenging to apply to generative AI-enabled devices that evolve by nature.

Current Regulations and Policies

The term "device" [18] is defined in section 201(h)(1) of the Federal Food, Drug, and Cosmetic (FD&C) Act as "an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including any component, part, or accessory, which is:

A. recognized in the official National Formulary, or the United States Pharmacopoeia, or any supplement to them,
B. intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals, or
C. intended to affect the structure or any function of the body of man or other animals, and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which is not dependent upon being metabolized for the achievement of its primary intended purposes. The term "device" does not include software functions excluded pursuant to section 520(o)

As stated above, certain software functions are specifically excluded from the device definition by section 520(o) of the FD&C Act, which include, but are not limited to, software functions intended for maintaining or encouraging a healthy lifestyle and are unrelated to diseases or conditions. Other software functions may not be devices, because they do not meet the definition of a device even if they were not specifically excluded by section 520(o) of the FD&C Act. As described in FDA's guidance, Policy for Device Software Functions and Mobile Medical Applications, FDA intends to apply its regulatory oversight to those software functions that are

[17] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/technical-considerations-medical-devices-physiologic-closed-loop-control-technology

[18] See section 201(h) of the Federal Food, Drug, and Cosmetic (FD&C) Act

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

medical devices and whose functionality could pose a risk to a patient's safety if the device were to not function as intended.

To determine if a product meets the definition of a device, it is important to first identify the intended use of the product.[19] Intended use refers to the objective intent of the persons legally responsible for labeling of a product (or their representatives), and may be shown by expressions, the design or composition of the product, or by the circumstances surrounding the distribution of a product.[20]

Some software functions may meet the definition of a device, but FDA has expressed its intent to exercise enforcement discretion for these devices because they pose lower risk to the public (meaning FDA does not, at this time, intend to enforce requirements under the FD&C Act). One example of a medical device software function that is currently under enforcement discretion helps patients diagnosed with psychiatric conditions maintain their behavioral coping skills by providing a "skill of the day" behavioral technique that can be accessed when a patient experiences increased anxiety. Another software function under enforcement discretion provides educational information, reminders, or motivational guidance to people recovering from addiction. Because these are low-risk software functions, FDA does not intend to enforce device requirements that may apply.

Within the scope of digital mental health medical devices, there are two broad categories: digital mental health therapeutics and digital mental health diagnostics. "Therapeutics" within this context refer to any digital mental health medical device that is intended to contribute to or aid in the treatment of a psychiatric condition and often include stand-alone and adjunctive therapy tools that are intended to provide therapeutic content in the course for treatment for psychiatric disorders. For example, computerized behavioral therapy devices are intended to provide therapeutic adjunctive therapy under the supervision of a mental health prescriber throughout care as usual for a specific diagnosis. "Diagnostics" within this context refer to any digital mental health medical device that is intended to contribute to the assessment, evaluation, monitoring, or diagnosis of a patient, and is not limited to stand-alone diagnostic tests. For example, pediatric autism spectrum disorder (ASD) diagnosis aids perform analyses of patient data to provide clinicians with a statistical estimate of whether an individual may have ASD, while attention task performance recorders provide measures of hyperactivity, impulsivity, attention/inattention, and inhibitory control intended to aid in the assessment of attention deficit hyperactivity disorder (ADHD), but do not identify presence or absence of a specific medical condition. The information provided by these diagnostic devices can be intended for use and interpretation by patients, caregivers, or clinicians to inform patient management.

Digital mental health medical devices that have been authorized by FDA to date are typically intended for prescription use and have been authorized under several different regulations. These devices include but are not limited to Computerized Behavioral Therapy Devices for Psychiatric Disorders (21 CFR 882.5801); Digital Therapy Devices for Attention Deficit

[19] See FDA's website on How to Determine if Your Product is a Medical Device
[20] See 21 CFR 801.4

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Hyperactive Disorder (21 CFR 882.5803); Digital therapy to Reduce Sleep Disturbances for Psychiatric conditions (21 CFR 882.5705); Pediatric Autism Spectrum Disorder Diagnosis Aid (21 CFR 882.1491); and Attention Task Performance Recorder (Unclassified).[21] Product codes are used to further identify the device's classification, based on the medical devices' intended use, indications for use, and associated risk. These codes are reflected within the medical devices database.[22]

Digital mental health medical devices may be submitted and authorized through several regulatory pathways, such as Premarket Approval (PMA),[23] De Novo,[24] or Premarket Notification [510(k)].[25] Through these pathways, devices have been authorized for attention deficit hyperactivity disorder; substance use disorders (including opioid use disorder); insomnia; mild, moderate, major, and postpartum depressive disorders; anxiety disorders; autism spectrum disorders; eating disorders; and symptomatic assistance for irritable bowel syndrome and fibromyalgia.

Clinical Trial Designs

FDA reviews many types of valid scientific evidence as part of its determination of reasonable assurance of safety and effectiveness for devices that require FDA's regulatory oversight.[26] Clinical evidence for a digital mental health medical device includes well-controlled, prospective, trial designs that support the device's indication for use, a detailed description of the trial, safety monitoring plan, and a statistical analysis plan (SAP) for the intended population to be treated.

Other important design elements include the study population selection [e.g., subject screening and eligibility (inclusion and exclusion criteria)] and prespecified, fit-for-purpose outcome measures and endpoints (specific to the study population) representing clinically meaningful improvements. The study duration and timing of the assessment of endpoints are evaluated based on the device function, and in consideration of the time course of the specific disorder (e.g., acute, chronic, treatment resistant) or symptom profile (e.g., cravings), the expected length of treatment, the type of treatment being delivered, the proposed indication for use, and labeling.

Due to the high placebo response rate in behavioral health studies, clinical trials typically utilize randomized-controlled trial (RCT) designs that ensure the design of the sham or active control[27]

[21] Product code LQD: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPCD/classification.cfm?id=4166
[22] https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPCD/classification.cfm
[23] See section 515 of the Federal Food, Drug, and Cosmetic (FD&C) Act. See also, https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/premarket-approval-pma
[24] See section 513(f)(2) of the Federal Food, Drug, and Cosmetic (FD&C) Act. See also, https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/de-novo-classification-request
[25] See section 510(k) of the Federal Food, Drug, and Cosmetic (FD&C) Act. See also, https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/premarket-notification-510k
[26] See 21 CFR 860.7
[27] Sham/active control designs include but are not limited to: demonstration of user level of engagement (duration and frequency of use, time on task comparable between groups); Usability and user experience similarities; comparability in

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

maintains blinding/masking[28] and promotes participant retention by matching the intervention device group to the control group, in all aspects, except for the specific therapeutic element of the intervention device.

Premarket Evaluation and Device Classification

FDA utilizes a risk-based approach for regulating medical devices, tailoring the level of regulatory controls necessary to demonstrate a reasonable assurance of safety and effectiveness to the level of risk the device poses to patients. Regardless of the type of premarket pathway – PMA, De Novo, or 510(k) – the principles of safety and effectiveness underlie FDA's review of all medical devices. The lowest risk devices are subject to general controls (like other devices reviewed through the above regulatory pathways) but are generally exempt from premarket review.

Each device is assigned to one of three regulatory classes, Class I, Class II, or Class III, where each class has increasing levels of regulatory control necessary to provide reasonable assurance of device safety and effectiveness. The extent to which risks can be mitigated can affect the level of regulatory control for a particular product. Medical devices are subject to premarket review under one of the following regulatory pathways based on the device's classification and the degree of risk they present:[29]

- Premarket Approval (PMA) is the pathway for high-risk medical devices, where the review standard relies on an independent demonstration of safety and effectiveness for the device's intended use.
- Premarket Notification [510(k)] is the pathway for low to moderate risk medical devices, where sponsors must demonstrate that the new device is "substantially equivalent" to (i.e., as safe and effective as) a legally marketed predicate device in terms of intended use, technological characteristics, and performance testing, as needed.
- De Novo Classification Requests provide a pathway to classify low to moderate risk novel medical devices for which there is no legally marketed predicate device.

General Controls. All medical devices, unless exempt by regulation, are subject to general controls, including, but not limited to, medical device reporting, reports of corrections and removals, establishment registration and device listing, and quality system regulation.[30]

implementing the user journey; data protection and privacy policies implemented to prevent harm; contain similar content to the intervention device, minus the intervention; mirroring the content of the intervention device in terms of the non-therapeutics elements (e.g., device components, audio-visual stimuli, gaming mechanics)

[28] Key features of study blinding include but are not limited to demonstration of procedures to minimize study staff bias; standardizing processes for training staff; ensuring that the training on intervention and control devices are comparable between groups; independent evaluation of blinding procedures before pivotal trials are initiated; and testing the procedures and training before a trial starts to ensure the measures taken are adequate for a executing a successful clinical trial

[29] See FDA's website on Medical Device Safety and the 510(k) Clearance Process

[30] General controls are regulatory requirements authorized by the FD&C Act, under sections 501, 502, 510, 516, 518, 519, and 520. *See* FDA's website on Regulatory Controls, *available at* https://www.fda.gov/medical-devices/overview-device-regulation/regulatory-controls

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Special Controls. FDA may require special controls in addition to general controls to provide reasonable assurance of safety and effectiveness of the device. Previously established special controls have been device-specific, and encompass clinical data, non-clinical testing, software, and device labeling requirements (e.g., clinical data validation, software requirements and design specifications, and labeling that includes appropriate instructions for use, warnings, and summary of clinical testing).[31] Novel moderate risk device types intended to provide specific diagnostic or therapeutic benefit for which there is no predicate would likely be subject to special controls to mitigate risks.

Premarket Evidence Considerations

FDA's Guidances on digital health and software help manufacturers determine whether a product is the focus of FDA's device regulatory oversight and understand recommendations for the type of information to provide in a marketing submission. However, in the context of newer technologies (e.g., generative AI), it can be difficult for users to determine what products are under the purview of FDA's authorities and for manufacturers to know what evidence is most appropriate and least burdensome for FDA to review them.[32] For products that are the focus of FDA device authorities, manufacturers may not understand what level of evidence is needed to establish the safety and effectiveness of the device. In terms of assessing the safety and effectiveness of digital mental health medical devices, careful consideration is needed to determine the factors included when weighing the benefits of a device with the risks and to what extent risks associated with the intended use population should be factored into the benefit-risk assessment. Furthermore, clinical evidence development is critical to understanding the benefit-risk profile and demonstrating the safety and effectiveness of a digital mental health medical device. Other considerations include:

Intended Use and Indications for Use. The intended use of the device (the purpose or function of the device) and the indications for use (the patient population and clinical conditions to be treated) are important factors for benefit-risk determinations in device classifications. To date, digital mental health medical devices that have been authorized by FDA are indicated for a specific condition (such as for insomnia). When considering generative AI-enabled digital mental health medical devices, such as a chatbot therapist, it will be important to understand if a given device is indicated for a specific condition or if it is indicated for multiple conditions. The scope of the indications for use of a device will inform discussion of the pre- and post-market evidence and risk mitigation approaches.

Over-the-Counter Use. Most digital mental health medical devices authorized by FDA to date are prescription devices, because of any potentiality for harmful effect, or the method of its use, or the collateral measures necessary to its use are not safe except under the supervision of a practitioner licensed by law to direct the use of such device.[33] Over-the-counter (OTC) medical devices, on the other hand, may be offered for sale

[31] See 21 CFR 860.3

[32] See FDA's Guidances with Digital Health Content

[33] See generally 21 CFR 801.109

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

directly to the consumer and do not require a prescription for sale because the manufacturer has demonstrated to the Agency that the OTC device can be used effectively and safely without a prescriber.[34] OTC medical devices may be used in any setting where the consumer can properly use the device consistent with the intended use of the device.[35] Special regulatory controls for OTC medical devices may also be required and may include, for example, human factors testing, labeling, or on-boarding instructions.

Patient and Caregiver Preferences. This includes the extent that testing with the device reflects values and outcomes that are important to patients, as well as whether some patients (or a specific group of patients) prefer attributes of the device to the alternatives. Different groups (e.g., patients, caregivers, and healthcare providers) may judge the benefits and risks of a medical device in different ways. FDA considers these perspectives as part of the totality of evidence when collected in a structured, scientific manner.

Human Factors. Human factors/usability engineering is used to design the user-device interface.[36] The user interface includes all components with which users interact while preparing the device for use (e.g., unpacking, set up, calibration), using the device, or performing maintenance (e.g., cleaning, replacing a battery, making repairs). For medical devices, the most important goal of the human factors/usability engineering process is to minimize use-related hazards and risks and then confirm that these efforts were successful and that users can use the device safely and effectively. Primary human factors/usability engineering considerations in the development of medical devices involve the three major components of the device-user system: (1) device users, (2) device use environments and (3) device user interfaces.

Postmarket Monitoring

Beyond premarket evidence, postmarket data has a pivotal role in ensuring the continued safety and effectiveness of digital mental health medical devices, particularly for those that may introduce new and adaptive qualities (e.g., the potential to undergo continuous adjustment based on localized live data and respond to user interactions and changing conditions). To assure timely access to safe and effective technology and facilitate medical device innovation,

---

[34] https://www.fda.gov/medical-devices/products-and-medical-procedures/over-counter-otc-medical-devices-considerations-device-manufacturers

[35] See FDA's website on Over-the-Counter (OTC) Medical Devices: Considerations for Device Manufacturers, *available at* https://www.fda.gov/medical-devices/products-and-medical-procedures/over-counter-otc-medical-devices-considerations-device-manufacturers. Note: changing a 510(k)-cleared device labeled for prescription use only to a device labeled for OTC use would likely require a new premarket submission. Such a change typically could significantly affect the safety or effectiveness of the device because the directions for use necessary for healthcare professionals to use a device safely and effectively can be significantly different from the directions for use necessary for lay users to use that same device safely and effectively

[36] https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/human-factors-and-medical-devices

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

the Agency balances the amount of information collected before a device can be marketed with the information that can be collected after the device is on the US market. The information collected postmarket is an integral part of the TPLC approach referenced in the background.

In addition, Predetermined Change Control Plans (PCCPs)[37,38] are a least burdensome mechanism for consideration as part of the framework for postmarket monitoring and change control of digital mental health medical devices. PCCPs include planned changes that may be made to a given device (and that would otherwise require a supplemental application).

Postmarket monitoring also includes device reporting. Manufacturers, device user facilities, and importers are required to submit certain types of reports for adverse events and product problems about medical devices. FDA also encourages but does not require healthcare professionals, patients, caregivers, and consumers to submit voluntary reports about serious adverse events that may be associated with a medical device, as well as reporting use errors, product quality issues, and therapeutic failures.[39]

Postmarket Evidence Considerations

Some of the characteristics associated with generative AI may warrant additional monitoring or reporting elements than previously authorized digital mental health medical devices. It is important to consider what technological characteristics of the device introduce risks that may need postmarket monitoring mitigations as well as the appropriate performance criteria for determining if a device is failing to function as intended. Use of real-world evidence may be useful in supporting the continued safe and effective use of generative AI-enabled digital mental health medical devices, including those that are not limited to a single indication of use.

PCCPs may be particularly useful for an adaptive algorithm (of which generative AI is one example). FDA would focus on several key areas to ensure safety and effectiveness, such as the level of specificity needed for modifications, the boundaries/guardrails associated with automatic updates, post-market performance monitoring involving how the device's performance will be tracked over time, labeling updates to inform users when automatic modifications are implemented, and the inclusion of any appropriate notification requirements if the device deviates from its intended function as outlined in the PCCP.

Generative AI in Digital Mental Health Medical Devices

This meeting is the second FDA DHAC addressing AI-enabled content in medical devices, with a specific focus on digital mental health medical devices. The 2024 DHAC meeting focused

[37] https://www.fda.gov/medical-devices/software-medical-device-samd/predetermined-change-control-plans-machine-learning-enabled-medical-devices-guiding-principles
[38] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/predetermined-change-control-plans-medical-devices
[39] https://www.fda.gov/medical-devices/medical-device-safety/medical-device-reporting-mdr-how-report-medical-device-problems

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

broadly on total product lifestyle considerations for all medical devices that incorporate generative AI. As mentioned, generative AI-enabled digital mental health medical devices present unique opportunities and challenges for consideration. Below is a table highlighting key discussion points from the 2024 DHAC. A more detailed summary of the Committee's recommendations is included in Appendix A, which provides additional background context for the upcoming meeting of this Committee.

Table 1. Highlights from the FDA 2024 Digital Health Advisory Committee Meeting

Area of interest	Committee's Main Points
Premarket Performance	• **Comprehensive Device Characterization**: Importance of submissions including detailed information on intended use, indications, use cases, care environment, human-AI interaction plans, and user proficiencies, along with standardized data sheets or model cards for technical specifications. • **Rigorous Performance Evaluation**: Devices to undergo performance assessment with tailored metrics (e.g., sensitivity/specificity), including evaluation of repeatability, reproducibility, measurement uncertainty, hallucination rates, error rates, and stress testing across intended use population and settings. • **Enhanced User Transparency and Usability**: Users should be aware they are interacting with generative AI by incorporating transparent interfaces, conducting appropriate usability testing, and developing educational materials to ensure appropriate device use and mitigation of risks of overreliance.
Risk Management	• **Probabilistic and Iterative Nature Considerations**: Risk management strategies should account for the unique probabilistic and iterative characteristics of generative AI devices, focusing on patient harm prevention and frameworks for deployment in specific healthcare settings. • **Human-in-the-Loop and Training Considerations**: Maintaining adequately trained human oversight may be important for safety. Human oversight combined with fostering digital health literacy among patients and providing robust HCP training may help prevent overreliance on AI-generated outputs. • **Shared Responsibility Framework**: Oversight of generative AI-enabled devices benefits from collaborative responsibility among regulators, manufacturers, healthcare systems, and clinicians, with emphasis on real-world transparency, explainability, and premarket plans for performance monitoring.

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Area of interest	Committee's Main Points
Postmarket Performance	• **Automated Monitoring and Surveillance**: Implementation of scalable approaches to track product usage, detect drift, identify hallucinations, and capture adverse events, including interim deployment phases and specialized monitoring for human-AI interactions. • **Multi-site Performance Management**: Deployment of automated auditing processes, ensemble methods, and quality assurance checks to ensure consistency across multiple sites while addressing regional biases and data variations compared to the originally authorized device. • **Foundation Model Assessment and Data Sharing**: Development of new tools to evaluate opaque foundation models, assessment of training data representativeness across intended use population, and establishment of mechanisms to disseminate performance insights back to healthcare facilities, providers, and patients.

This table represents key highlights of the comments and suggestions made by the 2024 DHAC Committee to FDA.

Areas of Interest in Generative AI-enabled Digital Mental Health Medical Device Development

We are seeking the Committee's input on several areas of interest related to generative AI-enabled digital mental health medical device development and monitoring. These include but are not limited to perspectives on best practices for clinical evidence (including clinical trial designs), therapeutic and diagnostic considerations, and how this technology can best serve patients safely and effectively.

Clinical Evidence Considerations

The existing evidence needs for mental health prescription therapeutics and diagnostics may also apply to digital mental health medical devices enabled by generative AI. However, digital mental health medical devices with generative AI content may have characteristics that warrant consideration of additional or different evidence to ensure device safety and effectiveness, given the interactive and adaptive nature of the intervention delivery (e.g., chatbot interfaces).

Committee perspectives on clinical evidence considerations across the TPLC are of interest to the Agency, including premarket and postmarket/real-world evidence. Committee discussion of the development or optimization of generative AI-enabled device clinical trial design

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

methodologies[40] (e.g., RCTs, large-scale cluster randomized, or non-inferiority designs) and clinical trial features[41] (i.e., control arm/sham devices, screening, diagnosing, blinding, controls, intended use populations, primary endpoints) for the diagnosis and treatment of mental health conditions, would be informative.

The Agency also seeks Committee feedback on approaches to clinical evidence for OTC devices in this space, with perspectives on how to balance lowering barriers to market while maintaining patient safety.

Products Designed to Incorporate Diagnostics with Therapeutics

Medical device developers have the potential to incorporate multiple sources of different types of patient data into diagnostic or therapeutic digital mental health medical devices to help improve diagnosis; select or guide choice of intervention; and to tailor specific interventions to specific individuals. The incorporation of these data may also enhance the development of autonomous diagnostic and therapeutic interventions (without the involvement of an HCP) that adjust based on user responses; or as a monitor for the effectiveness of the treatment that they are delivering with more precise and regular patient feedback.

Consideration by the Committee on how testing and validation of diagnostics with therapeutic interventions in a single study can be conducted safely and at what points interventions or alerts need to be deployed are of interest to the Agency. In addition, Committee feedback on how data is collected and integrated into products where the device may diagnose and treat an individual in the absence of an HCP; or when and how human interventions may be needed, will provide insight into the development of appropriate risk mitigation strategies.

Generative AI-enabled Technologies and Patient Safety in Psychiatry

It is generally recognized in medical settings that patients seeking treatment for psychiatric conditions undergo a thorough evaluation and ongoing monitoring for a given diagnosis, based on accepted diagnostic criteria that are assessed in intervals of time for symptom reduction and functional improvement. Adequate treatment includes the early identification of a patient's needs for higher levels of care and interventions for risks of self-harm, harm to others, or symptom risks that lead to relapses and/or worsening outcomes. Psychiatrists and qualified mental healthcare specialists are expected to inform and educate patients of treatment options during the course of care and provide adequate follow-up. Generative AI-enabled technologies may serve an important role in mental healthcare treatment. Consideration by the Committee on unique issues like human susceptibility to AI outputs, and how risk controls are developed (e.g., suicidal ideation monitoring and reporting) will be helpful, including characteristics of models and devices that may enhance patient safety or increase risks affiliated with long-term use.

[40] Examples include RCTs, large-scale cluster randomized studies, non-inferiority designs, incorporation of real world evidence/data

[41] Examples include control arm/sham devices, screening, diagnosing, blinding, controls, intended use populations, primary endpoints)

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Approaches to AI-Delivered Therapy

Generative AI and some LLMs, to date, have demonstrated vulnerabilities in some of the areas where human therapy excels (and vice versa).[42,43] 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. Consideration by the Committee in expanding upon the distinction between formal therapy and eclectic therapeutic content, as well as feedback on simultaneously mitigating the risks of using digital mental health medical devices, will be helpful. Additionally, Committee feedback on how generative AI-enabled digital mental health medical devices could reduce barriers and increase the reach of the therapeutic content to the public are important to the Agency.

Summary

FDA is committed to assuring that patients and providers have timely and continued access to safe and effective medical devices. As part of this mission, FDA aims to facilitate medical device innovation by providing industry with consistent, transparent and efficient regulatory pathways, and assuring consumer confidence in devices marketed in the United States.

The utilization of generative AI within digital mental health medical devices may offer unique benefits to patients and public health, but their use and adoption also comes with specific risks and complexities that necessitate careful consideration of effective and tailored approaches to regulatory oversight. The Agency welcomes Committee feedback on perspectives related to generative AI in digital mental health medical devices and considerations for risk mitigation frameworks for these devices, including comments on premarket and postmarket evidence needs for these devices.

[42] https://psychiatryonline.org/doi/10.1176/appi.psychotherapy.20230018
[43] https://health.usnews.com/wellness/mind/articles/should-you-use-artificial-intelligence-ai-as-your-therapist

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Panel Questions

There is a mental health crisis in the US and insufficient access to mental health care providers. New devices may be one way to help address this gap in care for people, potentially improving outcomes and access. FDA has long promoted a total product life cycle (TPLC) approach to the oversight of medical devices, including artificial intelligence (AI)-enabled devices, and has committed to advancing regulatory approaches for these devices.

Along with the rise of widely accessible generative AI products for general purposes, we are seeing an increase in the development and demand for a new kind of digital mental health medical device: "AI therapists" and other AI-based medical devices offering to provide a wide range of mental health therapies and interactions (some even being diagnostic) with therapist/healthcare provider-like chatbots. These chatbots may engage with users in individualized ways with, or without, the oversight of a health care provider (HCP), which pose novel risks for use. As digital mental health medical devices continue to evolve in complexity, regulatory approaches will need to accommodate these challenges and opportunities to provide a reasonable assurance of their safety and effectiveness while promoting innovation to support public health.

The questions below are designed to address the information and practices needed for a comprehensive approach to the assessment of performance and management of risk throughout the TPLC for digital mental health medical devices. Consider the following initial scenario, subsequent modifications to this scenario, and related questions:

Scenario	A patient diagnosed with major depressive disorder (MDD) by their healthcare provider is experiencing intermittent tearfulness due to increasing life stressors. Although the patient has consistently refused recommendations for therapy from their healthcare provider, the patient is willing to try a software device that provides therapy.
Device Description	This prescription therapy device is built on a large language model (LLM) that utilizes contextual understanding and language generation with unique outputs that mimic a conversation with a human therapist.
Device Indications for Use	This product is a standalone prescription digital therapy device indicated to treat MDD for adult patients (aged 22 years and older) who are not currently engaged in therapy.

1. First, consider that a healthcare provider prescribes the digital mental health medical device to the patient to use independently at home.
    a. Briefly discuss the probable benefits of this type of device that provides automated therapy in an ongoing manner.

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

b. What probable risks are presented by this type of device that provides automated therapy?
c. What risk mitigations should be considered for this type of device (e.g., alerts for self-harm ideations)?
d. What premarket evidence would you want to see to determine whether the benefits outweigh the risks to health?
    i. What are the key aspects of clinical evidence and trial design such as clinically meaning endpoints (e,g., measurable reduction in symptomatology), follow-up time, study eligibility criteria) ?
    ii. What alternative approaches could be used to demonstrate clinically meaningful benefits and risks (e.g., benchmarking, model-based evaluation)?
e. What specific postmarket monitoring capabilities should be considered to ensure continued safety and effectiveness of this medical device in real-world use (e.g., methods, metrics, tools)?
f. What labeling would be important for users of this type of device?

2. The manufacturer of the aforementioned, generative AI-enabled mental health medical device has decided to expand their labeled indications for use. They are contemplating the following changes.
    a. Making the device available over-the-counter (OTC) for people diagnosed with MDD.
    b. Modifying the OTC device to autonomously diagnose and treat MDD in an ongoing manner without the involvement of an HCP. They intend for the device to be used by people who have not been diagnosed with MDD by an HCP but have been experiencing symptoms of depression.
    c. Modifying the OTC, autonomous diagnosis and treatment device to be used for multiple mental health conditions (e.g., multi-use indications), meaning that it can provide both diagnosis and treatment for multiple mental health conditions related to sadness (in contrast to a device that is specifically indicated for MDD). The user of the device may not be clinically diagnosed with any mental health condition but has been feeling sad and has not met with an HCP.

3. Expanding the population to include a child or adolescent (i.e., 21 years and younger).

    a. As you consider the manufacturer's proposed changes, please discuss whether your prior responses to question 1 would change if the population were children or adolescents.
    b. If so, how would the responses change?

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Appendix A

Recommendations from the FDA 2024 Digital Health Advisory Committee Meeting

The DHAC to the FDA met on November 20-21, 2024, to discuss and provide recommendations on "Total Product Lifecycle Considerations for Generative AI-Enabled Devices." Suggestions from the Committee to FDA related to the safety and effectiveness of generative AI-enabled devices were focused on three areas: premarket performance, risk management, and postmarket performance and are summarized below.

Premarket Performance

The Committee provided suggestions for information that should be included as part of a device's description or characterization in the premarket submission when the device is enabled by generative AI. Emphasis was placed on detailed characterization to facilitate rigorous scientific evaluation. Submissions should comprehensively delineate the device's intended use, indications for use, specific use cases, and intended care environment, alongside the intention for human in the loop, the human-AI interaction plans, and required user proficiencies.

The Committee came to a consensus that obtaining detailed information on the datasets used to develop and test the device, to assess potential biases and generalizability, is important (e.g., sample size, data types, and demographic diversity). The Committee suggested that a submission should characterize the underlying foundation models, outlining guiderails, imposed constraints, known or potential failure modes, and adaptivity. Transparency and robustness against cybersecurity threats were identified as important components. The Committee also suggested the utility of a standardized data sheet or model card to succinctly convey these technical specifications, complementing traditional device evaluation elements such as risk and change management protocols, and quality system assurances.

The Committee shared ideas for evidence specific to generative AI-enabled devices regarding performance evaluation and characteristics of the training data during the total product lifecycle to understand if a device is safe and effective. When characterizing generative AI-enabled medical devices for premarket submission, there was consensus on ensuring comprehensive evaluation of device performance, with specific metrics tailored to the intended use (e.g., sensitivity and specificity for diagnostics). This characterization can extend to performance across different target populations and settings relevant to the device's application. Submissions should detail the repeatability, reproducibility, and measurement uncertainty, including estimates of hallucination rates, error rates, and the severity of errors, along with results from stress testing. Acknowledging the challenges with third-party foundation models, the Committee emphasized the importance of providing available information on the training and tuning data for such models. Furthermore, benchmarking against other models is recommended. A robust postmarket monitoring plan is important, especially when initial information on the underlying foundation model is limited. Ultimately, the depth of information required for premarket

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

evaluation should align with the risk profile of the device, consistent with the FDA's established risk-based approach.

The Committee discussed that new/unique risks related to usability may be introduced by generative AI compared to non-generative AI. They shared recommendations on specific information relevant to healthcare professionals, patients, and caregivers to be conveyed to help improve transparency and/or control these risks. Examples of this included transparent user interfaces (e.g., through explanations of device inputs and outputs tailored to the user's context, fostering trust). Users should be aware they are interacting with a generative AI-enabled device, and patients should understand its contribution to their care, potentially via accompanying labels noting the non-reproducible nature of outputs. Given that some AI device inputs (e.g., text, images, multimodal data) may be unfamiliar to clinicians, patients, and caregivers, explicitly detailing the information utilized for decision-making will be helpful. In addition, comprehensive training for users (clinicians, patients, caregivers) will help ensure appropriate device use. Furthermore, additional educational materials may be needed for devices intended for patient or caregiver use.

The Committee provided their feedback on prospective performance metrics that are particularly suited/most informative for technologies, given their complexity. They also discussed the kinds of performance metrics for multimodal systems. For generative AI-enabled medical devices, premarket performance evaluation necessitates the selection of modality-specific and functionally relevant metrics alongside established measures like sensitivity and specificity where applicable and characterization of the device's performance within its established upper and lower bounds (with a focus on identifying and analyzing edge cases to understand error frequency and types). Reporting should include, when available, an explanation of the model's output generation process. In addition, ongoing data drift metrics can ensure sustained accuracy and safety post-market. Given the unique characteristics of each generative AI device, transparent communication of all safety and performance metrics to regulatory bodies and users is important.

Risk Management

The Committee described how devices enabled with Generative AI are probabilistic and iterative. These relatively unique characteristics should inform the development of risk management strategies for these devices (e.g., employed controls, such as clinical validation and ongoing monitoring). The Committee also discussed how Generative AI introduces new ways of presenting information that may seem more human-like and give the impression of human intelligence to users, which could lead to overreliance on the device. Additionally, the Committee noted that the risk of patient harm is a central consideration for risk management and governance. The Committee communicated that risk management of these devices should be focused on the risk of patient harm, and they generally agreed on the need for frameworks related to risk management of Generative AI-enabled devices, including those focused on the infrastructure needed for deployment in specific settings.

The Committee discussed several ideas to address these topics. These include fostering digital health literacy among patients and providing robust clinician training to ensure proper device

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

utilization. Benchmarking is important for comparing device capabilities and performance against established standards. Maintaining a human-in-the-loop with adequate training may be important for safety. The evaluation process should encompass standalone performance testing, site-specific clinical validation, and ongoing post-market monitoring. Devices should prioritize real-world transparency and explainability, detailing their performance and potential variability across different environments, with premarket plans for real-world performance monitoring being an important component. Ultimately, the Committee believed that oversight of Generative AI-enabled devices is a shared responsibility among regulators, manufacturers, healthcare systems, and clinicians.

Postmarket Performance

The Committee emphasized automated and scalable approaches to track product usage, detect data drift, identify hallucinations, and capture resulting adverse events. It was suggested that methodologies should incorporate an interim deployment phase before widespread implementation, alongside specific monitoring capabilities for human-AI interactions to assess their postmarket effectiveness. Leveraging existing FDA postmarket surveillance models and change management strategies, such as PCCPs, can provide a baseline for localized and multi-site device monitoring. Real-world evidence trials may also enhance monitoring and evaluation efforts, and synthetic data offers a valuable tool for performance assessment in data-limited scenarios. Finally, establishing and utilizing Generative AI-specific standards and a centralized information-sharing infrastructure (potentially including automated user feedback and watermarking for transparency) may facilitate robust reporting back to manufacturers and the broader ecosystem. The Committee made several suggestions for specific strategies and tools to be implemented to monitor and manage the performance and accuracy of a generative AI-enabled device implemented across multiple sites, helping to ensure consistency, and addressing potential regional biases and data variations compared to the device that was authorized.

These suggestions included automated auditing processes to continuously monitor data drift by comparing local data distributions against original training datasets, assessing errors, and identifying necessary corrective actions. Strategies such as ensemble methods and embedded quality assurance checks can enhance device robustness. Long-term monitoring could extend to patient outcomes and shifts in clinical practices, alongside tracking instances where healthcare professionals correct device outputs. For devices utilizing multi-layer application designs that query external, non-medical-device AI services, specialized monitoring methodologies are required to evaluate post-market performance effectively. Recognizing the challenge of assessing Generative AI devices built on opaque foundation models, sponsor-provided information on the foundation model's contents or the provision of strategies to mitigate uncertainty may help ensure continued postmarket performance and underscored the need for new tools and approaches to evaluate the impact of foundation models on device performance. Manufacturers should also prioritize assessing the representativeness of foundation model training data, clearly define performance metrics for all subgroups of the intended use population and monitor outcomes across the population. Finally, beyond model development, mechanisms are likely needed to disseminate performance data and insights

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

back to community hospitals and medical centers, potentially through registries or nonprofit frameworks.

Glossary

TERM	DEFINITION
Artificial Intelligence (AI)	A machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. Artificial intelligence systems use machine- and human-based inputs to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to formulate options for information or action. *Source: 15U.S.C.9401(3).* *https://www.govinfo.gov/content/pkg/USCODE-2020-title15/html/USCODE-2020-title15-chap119.htm*
Artificial Intelligence Performance Monitoring (AI Performance Monitoring)	Refers to the process of regularly collecting and analyzing data on the use of a deployed AI system to evaluate its performance in accomplishing its intended tasks in real-world settings. The assessment of an AI model's performance involves various performance metrics and criteria depending on the specific application. This monitoring typically aims to assess the performance of these AI systems in practice, detect performance degradation or changes (e.g., due to data drift), identify instances of misuse, and address any safety or usability concerns. *Source: DH/AI Glossary*
Artificial Intelligence System (AI System)	Engineered system that generates outputs such as content, forecasts, recommendations, or decisions for a given set of human-defined objectives. *Source*: *International Organization for Standardization. (2022). Information technology — Artificial intelligence — Artificial intelligence concepts and terminology (ISO/IEC 22989:2022). https://www.iso.org/standard/74296.html*
Data Drift	Refers to the change in the input data distribution a deployed model receives over time, which can cause the model's performance to degrade. This occurs when the properties of the underlying data change. Data drift can affect the accuracy and reliability of predictive models. For example, medical AI-enabled products can experience data drift due to, statistical differences between the data used for model development and data used in clinical operation due to variations between

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

	medical practices or context of use between training and clinical use, and changes in patient demographics, disease trends, and data collection methods over time. *Source: DH/AI Glossary*
Device	An instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including a component part, or accessory which is: (A) recognized in the official National Formulary, or the United States Pharmacopoeia, or any supplement to them, (B) intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals, or (C) intended to affect the structure or any function of the body of man or other animals, and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which is not dependent upon being metabolized for the achievement of its primary intended purposes. The term "device" does not include software functions excluded pursuant to section 520(o). *Source: Section 201(h)(1) of the Federal Food, Drug, and Cosmetic Act*
Explainability	"Refers to a representation of the mechanisms underlying AI systems' operation." *Source: NIST* Explainability may help overcome the opaqueness of black-box systems (i.e., systems where the internal workings and decision-making processes are not transparent or readily understandable). These explanations can take various forms, including free-text explanations, saliency maps, SHapley Additive *exPlanations (*SHAP), or relevant input examples from data. The primary intent is to answer the question "Why" an AI system made a particular decision. Appropriate Explainable AI (XAI) methods may enable the development of more accurate, fair, interpretable, and transparent AI systems to safely augment human decision-making in healthcare**.** *Source: DH/AI Glossary*
Foundation Models	AI models trained using large, typically unlabeled datasets and significant computational resources, that are applicable across a wide range of contexts, including some that the models were not specifically developed and trained for (i.e., emergent capabilities). These models can serve as a foundation upon which further models can be built and adapted for specific uses through further training (i.e., fine-tuning). These models can perform a range of general tasks,

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

	such as text synthesis, image manipulation, and audio generation. These models are based on deep learning architectures like transformers and can use either unimodal or multimodal input data. *Source: DH/AI Glossary*
Generative Artificial Intelligence (Generative AI)	The class of AI models that emulate the structure and characteristics of input data in order to generate derived synthetic content. This can include images, videos, audio, text, and other digital content. This is usually done by approximating the statistical distribution of the input data. For example, in healthcare, generative AI can be used to generate annotations on synthetic medical data (e.g., image features, text labels) to help expand datasets for training algorithms. *Source: DH/AI Glossary*
Hallucinations (Confabulations)	Refers to a phenomenon in which generative AI systems generate and confidently present erroneous or false content to meet the programmed objective of fulfilling a user's prompt. *Source: National Institute of Standards and Technology. (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile: https://airc.nist.gov/docs/NIST.AI.600-1.Generative AI-Profile.ipd.pdf*
Intended Use	Refers to the objective intent of the persons legally responsible for the labeling of an article (or their representatives). The 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. This objective intent may, for example, be shown by labeling claims, advertising matter, or oral or written statements by such persons or their representatives. Objective intent may be shown, for example, by circumstances in which 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; provided, however, that a firm would not be regarded as intending an unapproved new use for a device approved, cleared, granted marketing authorization, or exempted from premarket notification based solely on that firm's knowledge that such device was being prescribed or used by healthcare providers for such use. The intended uses of an article may change after it has been introduced into interstate commerce by its manufacturer. If, for example, a packer, distributor, or seller intends an article for different uses than those intended by the person from whom he or she received the article, such packer, distributor, or seller is

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

	required to supply adequate labeling in accordance with the new intended uses. *Source:* *https://www.ecfr.gov/current/title-21/chapter-I/subchapter-H/part-801/subpart-A/section-801.4*
Large Language Model (LLM)	A type of AI model trained on large text datasets to learn the relationships between words in natural language. These models can apply these learned patterns to predict and generate natural language responses to a wide range of inputs or prompts they receive, to conduct tasks like translation, summarization, and question answering. These models are characterized by a vast number of model parameters (i.e., internal learned variables within a trained model). LLMs build on foundational AI models by developing more comprehensive language understanding beyond basic linguistic patterns. For example, in the context of LLMs, chatbot is a program that enables communication between the LLM and the human through text or voice commands in a way that mimics human-to-human conversation. *Source:* *DH/AI Glossary*
Locked Model	A model that provides the same output each time the same input is applied to it and does not change with use, as its parameters or configuration cannot be updated. In case of AI-enabled medical products, locked models can help ensure consistent performance. *Source:* *DH/AI Glossary*
Machine Learning (ML)	A set of techniques that can be used to train AI algorithms to improve performance at a task based on data. *Source:* 15 U.S.C. 9401(11). https://www.govinfo.gov/content/pkg/USCODE-2020-title15/html/USCODE-2020-title15-chap119.htm
Machine Learning Model (ML Model)	A mathematical construct that generates an inference or prediction for input data. This model is the result of an ML algorithm learning from data. Models are trained by algorithms, which are step-by-step procedures used to process data and derive results. AI systems (e.g., AI-enabled medical devices) employ one or more models to achieve their intended purpose. *Source:* *DH/AI Glossary*
Neural Network	A computational model inspired by the structure of the human brain. It is composed of interconnected nodes, or “neurons” organized into layers: an input layer that receives data, one or more hidden layers that process and identify patterns in the data, and an output layer that presents the final network output. *Source:* *DH/AI Glossary*

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov

Performance Metrics	In the context of AI quantitative or qualitative measures that can be used to assess the ability of a model to produce the desired output for a given task. The choice of the metrics depends on the specific task and the model objectives. Examples of quantitative metrics include accuracy, precision, sensitivity (recall), specificity, F1-score, and Area under the Receiver Operating Characteristic curve (AUC-ROC). Qualitative measures may involve heatmap evaluations or visual interpretations. These metrics enable systematic evaluation, comparison, and refinement of models, and aid in the assessment of whether the model meets its intended objectives. *Source: DH/AI Glossary*
Total Product Lifecycle (TPLC)	An integrated device review, tracking, reporting and compliance scheme employed by FDA. The TPLC approach allows FDA to integrate all regulatory activities from device inception to obsolescence. For purposes of this document, the TPLC approach addresses all phases in the life of a medical device, from the initial conception to final decommissioning and disposal. *Source: Infusion Pump: Glossary | FDA*
Training Data	These data are used by the manufacturer of an AI system in procedures and training algorithms to build an AI model, including to define model weights, connections, and components. *Source: DH/AI Glossary*
Transparency	Describes the degree to which appropriate information about a Machine Learning-Enabled Medical Device (MLMD), including its intended use, development, performance and, when available, logic) is clearly communicated to relevant audiences. *Source: U.S. Food and Drug Administration (2024). Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles. https://www.fda.gov/medical-devices/software-medical-device-samd/transparency-machine-learning-enabled-medical-devices-guiding-principles*
Watermarking	The process of embedding an identifying pattern in a piece of media in order to track its origin —including into outputs such as images, audio, video, and digital text—for the purposes of verifying the authenticity of the output or the identity or characteristics of its provenance, modifications, or conveyance. *Source: DH/AI Glossary*

U.S. Food & Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993
www.fda.gov