FDA Develops Regulatory Blueprint for Generative AI-Enabled Devices

By Luke Nguyen, Kayla J. Cristales | September 15, 2026

RELATED PRACTICES FDA Regulatory and Compliance, Food, Beverage and Restaurant

The U.S. Food and Drug Administration (FDA) released a discussion paper on generative AI (GenAI)-enabled devices, discussing various regulatory concepts and requesting insight from industry stakeholders as an early step in the development of a framework that appropriately accounts for the novel, nuanced considerations GenAI raises in this context.[1] It is no secret that industry experts have raised concern about the suitability of FDA's current regulatory framework to properly evaluate GenAI-enabled devices. The discussion paper seems to be FDA's acknowledgement of, and the start of its answer to, those concerns. While there is likely still a long way to go before industry gets guidance it can actually implement with respect to GenAI-enabled devices, the paper offers helpful insight into the current thinking of the Center for Devices and Radiological Health ("**CDRH**") surrounding GenAI-enabled devices and identifies the types of complex considerations that conflict with the traditional approaches to device regulation.

What Are GenAI-Enabled Devices?

The FDA defines "GenAI" as a "class of AI models that emulate the structure and characteristics of input data to generate derived synthetic content."[2] The content created by GenAI can take the form of images, videos, audio, text, or other digital content.[3] "GenAI-enabled devices" are devices that contain one or more software functions enabled by GenAI. Such devices differ from traditional software and AI-enabled devices in that they can accept open-ended inputs, perform multiple subtasks, produce variable outputs to similar inputs and evolve over time. For example, GenAI models can analyze input data and produce outputs that may not have been explicitly seen in its training data.[4]

What Is the Framework Considered in the Discussion Paper?

A. **Risk Assessment**

As with other medical devices, FDA's discussion paper focuses on a risk-proportionate approach to regulation. To properly categorize the risk posed by GenAI-enabled devices, FDA contemplates a two-axis framework (see below) in which risk is based on the GenAI activity levels.

At the Counter - Activity and Consequences Table

Functions that provide the user with non-directive information (*e.g.,* a risk score for a future cardiovascular event) have a lower risk in terms of the likelihood that an error will result in adverse health consequences, as compared to information that is action directing and provides specific instruction (*e.g.*, "increase the lisinopril from 10 mg to 20 mg daily"). The risk increases further if the device performs action-taking functions (*e.g.,* affirmative assignment of a clinical diagnosis, prescription of a medication or initiation of a clinical order set). In addition to the *likelihood* of patient harm, the "consequences" axis of the risk framework also considers the *severity* of the adverse

outcomes associated with a potential malfunction. For example, autonomous prescription of antibiotics for a minor infection presents less risk than an autonomous initiation of an order set for a severe disease or condition.

Continuing on the discussion paper's theme of recognizing all the reasons a "one-size-fits-all" regulation will not work for GenAI-enabled devices, FDA also identifies several factors, in addition to the "activity" and "consequence" considerations that may reduce or increase the risk posed by a particular function. For example, a function acting under continuous healthcare provider ("**HCP**") supervision is likely lower risk than fully autonomous functions. The intended user similarly affects a GenAI-enabled device's risk profile, e.g., information delivered directly to patients poses a higher risk, compared to information directed to HCPs.

B. **Competency-Based Approach for Premarket Evaluation**

The discussion paper proposes a novel "competency-based" approach for premarket evaluation of GenAI-enabled devices. The approach adopts themes inspired by human clinician evaluation and credentialing processes, focusing on evaluation of underlying knowledge and reasoning, as well as evaluation in a clinical setting.

1. *Device Benchmarking*

The first element of the competency-based approach is benchmarking. According to FDA, benchmarking is intended to evaluate whether a device, in its deployed or representative configuration, demonstrates the clinical knowledge, analysis, safety, communication, and generalizability to ensure the device is safe and effective for its intended use. Specifically, FDA considers benchmarking to consist of the following elements:

- **Safety**: Specific assessments to determine whether the device avoids causing harm. This includes whether the device: recognizes safety-critical states and initiates escalation in a timely manner; recognizes when a request exceeds its intended use or capabilities; and accurately represents the confidence and limits of its outputs.
- **Clinical Proficiency**: Specific assessments to determine whether the device possesses accurate clinical knowledge, reasons soundly through ambiguous information, interprets numerical data correctly, and communicates outputs clearly and appropriately to its intended users.
- **Generalizability**: Specific assessments to evaluate whether safety and clinical proficiency properties remain consistent across all foreseeable variations (*i.e.*, input, user, and/or runtime variations).

2. *Clinical Confirmation*

CDRH recognizes that device benchmarking, alone, may be insufficient to ensure GenAI-enabled devices operate safely and effectively in real clinical settings. As such, the discussion paper highlights an additional, clinical confirmation step as a potential requirement for sponsors to establish that the device operates as intended in real practice. However, what this approach may actually look like is unclear, as the paper notes that the appropriate clinical confirmation exercises should likely vary based on the device's intended use and risk profile. To illustrate, CDRH discusses the following potential approaches for clinical confirmation:

- **Retrospective Evaluation**: Applying the GenAI-device to previous patient inputs and comparing the device output to either a reference standard or expert clinical

adjudication, which could include developing independent multi-clinician consensus.
- **Standardized Patient Interactions**: Utilizing trained patient actors or representative individuals to simulate standardized scenarios and comparing the device output with prospectively defined criteria.
- **Clinician Adjudication of Real Cases**: Utilizing real patient inputs and device outputs and comparing such outputs to the assessment of real clinicians. For example, this could take the form of a clinician reviewing the input without knowledge of the device's output and independently determining the appropriate response to the input before comparing such determination to the device output.
- **Clinical Trial**: CDRH also notes that, for certain devices, a prospective clinical trial may be the most appropriate clinical confirmation method.

C. **Postmarket Monitoring**

A key nuance to GenAI-enabled devices is the potential for continual adjustments while the product is on the market, which makes a proposed premarket review structure insufficient to ensure the device remains safe and effective following premarket review. As such, CDRH recognizes GenAI-enabled devices may require increased reliance on postmarket monitoring strategies and, in turn, acknowledges that, unlike the traditional premarket model under which FDA evaluates and authorizes medical products, clinical studies may not be the end-all-be-all for GenAI-enabled product candidates.

While CDRH notes that a variety of approaches to postmarket monitoring may be effective, it highlights the following as potential postmarket monitoring approaches:

- **Periodic Device Benchmarking**: The sponsor could reassess the device against device benchmarking thresholds established during the premarket review process. This review could be conducted on a set, periodic basis or following triggering events (i.e., changes to the underlying model).
- **Periodic Sample-Based Clinician Review**: Independent clinicians could review sample inputs and outputs of the device, comparing the device against prospectively defined criteria.
- **Performance Degradation Monitoring**: The sponsor may monitor the device for performance degradation resulting from changes in the input, data, or underlying model. This would require the sponsor to specify the analyses and thresholds used to measure perceived degradation.

In addition, FDA is considering whether postmarket tools, such as a Predetermined Change Control Plan or rebenchmarking may also be necessary to ensure the device is safe and effective following postmarket modifications. Under the current device framework, such changes would often require sponsors to submit an amendment or a new premarket authorization application. These approaches would replace the need for sponsors to submit a new premarket authorization submission before implementing each postmarket change.

Next Steps

The goal of the discussion paper is to advance discussions around a regulatory framework amenable to GenAI-enabled devices. To that end, CDRH requested industry responses to the discussion questions raised throughout the paper, as well as any other stakeholder feedback that may inform the development of FDA's policy in this area, by October 19, 2026. Public comments may be submitted via Regulations.gov under docket number FDA-2026-N-7874.

[1] FDA, Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback (Aug. 2026).

[2] FDA, Digital Health and Artificial Intelligence Glossary – Educational Resource (Generative AI).

[3] *Id.*

[4] FDA, Executive Summary for the Digital Health Advisory Committee: Total Product Lifecycle Considerations for Generative AI-Enabled Devices 4 (Nov. 2024).