

Title: FDA Seeks Input on How to Regulate Generative AI-Enabled Medical Devices to Shape Upcoming Guidance

Authors:

[Tina Hu-Rodgers](#)

202 452 7308

tina.hu@bipc.com

[Natalie P. Crow](#)

202 452 5477

natalie.crow@bipc.com

Full Article: The U.S. Food and Drug Administration (FDA) is considering how its existing medical device framework should apply to products that use generative artificial intelligence (GenAI). In a recent discussion paper, FDA describes the unique characteristics of GenAI-enabled medical devices and asks the industry for input on how these products should be evaluated, monitored and modified over time.

Notably, while the discussion paper identifies several areas where FDA believes its traditional approaches to evaluating medical devices may be difficult to apply to GenAI, it does not establish new requirements or propose a new regulatory framework. FDA states that the paper is simply intended to generate discussion and early stakeholder input, not to communicate proposed regulatory expectations. However, FDA has indicated that the discussion paper is a step toward formal policy, as the agency's goal is to translate stakeholder input into policy. For companies awaiting regulatory clarity, this means the discussion paper is a meaningful signal of FDA's direction, but formal guidance documents are still to come. This was confirmed by Rick Abramson, director of FDA's Digital Health Center of Excellence, who this week stated in an interview with STAT that the agency's goal is "formal policy guidance," and that the ecosystem should expect both broad guidance on GenAI generally and narrower guidance addressing specific GenAI topics of particular interest or complexity.

FDA Has Been Engaging With GenAI Developers

In 2026, the Agency granted Breakthrough Device designation to at least two publicly announced GenAI-enabled devices. The Breakthrough Device Program is designed to facilitate development and FDA review of certain innovative devices that have the potential to provide more effective diagnosis or treatment of serious or life-threatening conditions. Breakthrough designation does not mean that FDA has cleared or approved the device for marketing.

In March, FDA granted Breakthrough Device Designation to Cognita Chest X-Ray, a generative vision-language model designed to assist radiologists in interpreting chest X-rays and drafting preliminary findings for physician review. Subsequently in June, Aidoc announced that FDA had granted Breakthrough Device Designation to First Read, a GenAI tool designed to analyze chest X-rays and produce preliminary radiology report text for review by a radiologist.

These products are notable because they move beyond traditional AI systems that identify a particular finding or provide a defined output. They use GenAI to produce substantive clinical content (i.e., preliminary radiology reports) for a physician to review.

The designations, therefore, provide an interesting backdrop to FDA's discussion paper, as FDA is actively working with companies developing GenAI-enabled medical devices while simultaneously considering whether its existing regulatory tools are adequate for these products.

FDA Is Not Creating a New Risk-Based Framework

As previously noted, one point is important to understand about the discussion paper: FDA is not abandoning or replacing its existing risk-based approach.

FDA expressly states that it expects to apply its existing risk-based approach to GenAI-enabled devices, considering the product's intended use and technological characteristics. FDA also continues to view total product lifecycle management as important for these devices.

This is consistent with FDA's broader approach to software. FDA already evaluates software functions based on factors such as what the software does, its intended use, the risks associated with its use and the extent to which a healthcare professional can independently evaluate and act on its output. FDA's January 2026 final guidance on Clinical Decision Support Software, for example, continues to distinguish non-device clinical decision support from device functions in part based on whether a healthcare provider can independently review the basis for the software's recommendation rather than primarily relying on it.

FDA's discussion paper builds on these concepts rather than replacing them. Therefore, the more significant question is how FDA can apply those concepts when the technology itself behaves differently from traditional software.

What Is Different About GenAI?

FDA identifies several characteristics of GenAI that can create additional regulatory challenges: GenAI-enabled devices may accept open-ended inputs, perform multiple tasks, generate different outputs in response to similar inputs, interact with users over multiple turns, change as underlying models or components change and rely on third-party foundation models that the manufacturer may not fully control.

These characteristics can make it difficult to determine in advance exactly how a device will behave, and create challenges in demonstrating that it will perform safely and consistently across the wide range of inputs and outputs it may encounter.

FDA specifically identifies risks such as hallucinations or "confabulations," uncertainty about the boundaries of intended use, limited visibility into third-party foundation models and performance degradation across different testing environments and real-world settings.

FDA Is Exploring How Existing Risk Concepts Should Apply to GenAI

Rather than creating a new risk-based framework, FDA is instead asking whether the factors traditionally used to assess risk need to be adapted for GenAI.

The paper proposes, for discussion purposes, a two-axis framework:

the form and independence of the device's activity; and

the potential consequences of relying on an incorrect output.

This is not a fundamentally new concept for FDA. The agency applied a similar framework to Software as a Medical Device.

However, this approach offers a useful model for how existing risk principles might extend to GenAI. For example, FDA distinguishes between non-directive information and a recommended specific action, and between a system under continuous provider supervision and one acting autonomously. The agency also asks whether factors such as wording, specificity and personalization of an output should affect risk classification and whether a disclaimer like "talk to your doctor" should meaningfully reduce the risk of an otherwise action-directing output — a

question of particular relevance to conversational GenAI systems that may drift from general information toward individualized, action-directing advice over the course of a conversation.

The Bigger Challenge: How Do You Test a GenAI Device?

Traditional software can often be tested by establishing representative inputs and determining whether the software produces the expected outputs. FDA recognizes that this approach may not work as well for GenAI because the range of possible inputs and outputs can be extremely large. It may simply be impractical to test every conceivable interaction.

FDA is therefore considering a competency-based approach to premarket evaluation, meaning that companies could demonstrate that the system has certain capabilities and performs them safely and reliably within a defined scope. Depending on the product, those competencies could include clinical knowledge and task performance, appropriate reasoning, knowing when to refuse to answer, communication, reliability and reproducibility, performance across different patient populations, and, for more autonomous systems, the ability to safely plan and execute multiple steps.

FDA May Also Put More Weight on Real-World Performance

The paper also raises the possibility that premarket testing may need to be supplemented by clinical confirmation and post-market monitoring, given that some GenAI devices may be difficult to fully evaluate before marketing because of their broad capabilities and potentially changing behavior. FDA discusses various approaches, including retrospective evaluation using real patient data, testing in clinical workflows, evaluation by clinicians, prospective clinical studies and ongoing monitoring. FDA is also considering whether post-market monitoring could identify performance degradation or "drift" resulting from changes in patient populations, data environments, or underlying models.

This does not necessarily mean FDA is proposing to lower the premarket bar. Rather, the paper raises the possibility of a more lifecycle-oriented evidence strategy, in which premarket and post-market evidence work together.

Model Changes Could Become a Major Regulatory Issue

In situations where a company is relying on a third-party foundation model modified by another company for their device, GenAI can also create a more complicated change-control problem. FDA specifically identifies several types of changes that could affect a GenAI device, including

changes to the underlying model, prompts, retrieval strategies, guardrails, orchestration logic and user interfaces.

FDA is considering how manufacturers should manage these changes and how tools such as Predetermined Change Control Plans (PCCPs) might be used. This builds on FDA's existing approach to AI-enabled devices. FDA's 2025 PCCP guidance provides a framework for allowing certain planned modifications to AI-enabled devices while maintaining reasonable assurance of safety and effectiveness.

For GenAI, however, the challenge may be greater where a manufacturer depends on a third-party foundation model and does not control when or how the underlying model changes. Contracts may need to address notice of model changes, access to information needed for validation, performance monitoring, version control and responsibility for problems caused by changes to the underlying model.

Human Oversight Remains Important, but May Not Solve Everything

FDA also recognizes the potential for automation bias, where users may place too much confidence in GenAI output because it sounds authoritative. The appropriate level of human oversight may therefore depend on who is using the product and what the product does.

FDA specifically questions whether patient-facing GenAI functions should sometimes be treated as presenting greater risk than similar healthcare provider-facing functions because patients may be less able to identify an incorrect output. At the same time, FDA recognizes that GenAI could increase patient access to information and support patient engagement. This intersects with FDA's existing clinical decision framework: human review remains relevant, but FDA is questioning how much weight it should receive when the reviewer may lack the expertise or information to identify an incorrect GenAI output.

What Does This Mean for Companies?

The discussion paper does not require companies to change their products or regulatory strategies today. FDA has indicated that formal guidance will follow, including guidance addressing particular GenAI topics of special interest or complexity, but has not committed to a timeline. In the meantime, the paper provides several useful signals for companies developing GenAI-enabled medical devices

1. Think carefully about what the product actually does.

Companies should look beyond broad descriptions such as "clinical decision support" or "AI assistant." The regulatory analysis may depend on whether the product provides general information, gives individualized recommendations, directs a specific action or takes action itself.

2. Consider the output, not just the underlying technology.

For GenAI products, the regulatory risk may depend heavily on how the product communicates information to the user. Companies should consider the degree of personalization, specificity and directiveness of outputs and whether the product can move from general information to individualized recommendations during a conversation.

3. Design validation around the product's capabilities.

Traditional accuracy testing may not tell the whole story for GenAI. Companies should consider whether testing addresses issues such as inappropriate responses, refusal behavior, consistency, robustness, different patient populations and the system's ability to remain within its intended scope.

4. Plan for change from the beginning.

Companies should identify which parts of the system may change, who controls those changes, how changes will be evaluated and when a change could affect the device's safety or effectiveness.

5. Treat third-party AI providers as part of the regulatory strategy.

If a device relies on an outside foundation model, the manufacturer's relationship with that provider may become a critical regulatory issue. Companies should consider whether their

agreements provide sufficient transparency, notice of changes, access to relevant information and the ability to monitor the model's performance.

6. Build post-market monitoring into the product strategy

Demonstrating performance before launch of a GenAI product may not be the end of the regulatory exercise. Companies should consider how they will identify changes in performance once the product is being used in the real world and how they will respond if the system begins behaving differently.

Bottom Line

FDA's discussion paper should not be read as FDA abandoning its existing risk-based framework for medical devices. In fact, FDA expressly says the opposite. The more significant issue is that GenAI may make some of FDA's traditional methods for applying that framework harder to use.

FDA is therefore exploring whether new approaches are needed to evaluate products whose outputs can be open-ended and variable, whose behavior can change over time and whose underlying technology may be controlled partly by third parties. FDA has stated that formal policy guidance, including both broad and topic-specific guidance, will follow the comment period, though no timeline has been announced.

FDA's recent Breakthrough Device designations also suggest that the Agency is willing to work with developers of promising GenAI technologies while the framework is being worked out. The regulatory landscape is therefore likely to evolve alongside the technology, making early engagement with FDA particularly important.

Stakeholders wishing to provide feedback on the paper can submit comments under the docket FDA-2026-N-7874 by October 19, 2026.