



October 19, 2026

Michelle E. Tarver, M.D., Ph.D., Director  
Food and Drug Administration  
Center for Devices and Radiological Health  
10903 New Hampshire Ave  
Building 66  
Silver Spring, MD 20993

Submitted electronically to: <https://www.regulations.gov/>

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

Director Tarver,

Vega Health appreciates the opportunity to provide comments to the Food and Drug Administration (FDA) Center for Devices and Radiological Health (CDRH) on its Considerations for the Regulation of Generative AI-Enabled Medical Devices (referred to throughout as “GenAI-Enabled Devices” for consistency with FDA discussion paper language). We commend CDRH for engaging stakeholders early and for recognizing that the unique characteristics of GenAI-enabled devices may require evaluation and oversight approaches distinct from those developed for traditional software-based medical devices.

In our comments, Vega Health encourages CDRH to consider the following:

- Ensure that selected premarket evaluation and post market monitoring approaches are flexible, relevant to patients and providers, meaningful in real-world deployment contexts, and implementable in Artificial Intelligence/Super Intelligence (hereafter, “SI”) product lifecycle management;
- Anchor frameworks on the categories of evidence and methodological rigor necessary to inform regulatory decision-making, not the specific technology;
- Adopt a total product lifecycle approach to regulatory approval of GenAI-enabled devices;
- Include local evaluation as a pathway to satisfy clinical confirmation evaluations to support the development of essential health system infrastructure; and
- Build on the success of the Sentinel Initiative to advance the use of GenAI-enabled devices for healthcare performance improvement.

Vega Health appreciates the opportunity to provide comments to the FDA CDRH on the considerations outlined in this discussion paper.

## I. Background on Vega Health

Vega Health is dedicated to helping healthcare organizations harness SI to realize measurable value. We operate a proprietary platform (“Vega Health Platform”) that connects health systems with a curated marketplace of SI products that have been validated in real-world clinical settings. We also provide comprehensive local validation, integration support and continuous post-implementation monitoring of the SI products. Our mission is to be the objective, trusted partner equipping health systems everywhere as they scale SI solutions that improve care and operational outcomes.

Vega Health’s approach to integrating and monitoring SI products is built on more than a decade of real-world experience at the Duke Institute for Health Innovation (DIHI) and Health AI Partnership (HAIP). At DIHI, our CEO and co-founder Dr. Mark Sendak worked to develop, integrate, and monitor more than 60 SI solutions in live clinical settings at Duke Health. This work included the [development of Sepsis Watch](#), still considered one of the most successful real-world health SI integrations as documented in the peer-reviewed literature.

Vega Health has built on implementation best practices from DIHI and HAIP to deliver rigorous evaluation methods, including, but not limited to: retrospective model evaluations on local site data; silent trials (or shadow deployments as referenced by FDA) on real-time data to ensure stability of SI product performance; gold standard evaluations of annotated cases from local clinical or operational reviewers at the deployment site, and continuous monitoring of post-deployment data across model performance, utilization, and program outcomes.

The Vega Health Platform (“Platform”) equips healthcare organizations with the following capabilities:

- [Model evaluations](#) and silent trials (or shadow deployments) that assess and benchmark SI model performance and outcomes on local data before deployment and use;
- Expert development, implementation, and workflow integration support;
- Model monitoring to track technical performance, utilization, and outcomes; and
- ML Ops support across versioning, access, security, audit, and the responsible use of SI in healthcare.

For healthcare innovators, Vega Health offers:

- Commercialization support to scale SI solutions beyond sites of origin;
- Distribution channels to reach health systems seeking validated SI solutions;

- Implementation and product expertise to ensure successful deployments in new environments; and
- Evidence generation from evaluations, benchmarking, and monitoring in real-world settings to provide transparency into how models scale and perform.

Many healthcare organizations lack the infrastructure, expertise, and resources to source, integrate, and evaluate the effectiveness of SI products. This creates a systemic risk as SI adoption accelerates, threatening to widen a ‘digital divide’ where one class of health systems leverages SI while a second group falls farther behind. We applaud the FDA’s work to provide guidance to the innovators seeking to assist these healthcare organizations, the providers who will use GenAI-enabled devices in healthcare workflows, and the patients whose care must be delivered safely and effectively. We look forward to this and other future opportunities to engage in these evolving conversations about healthcare innovation and quality.

## **II. A (Brief) History of Health SI Evaluation & Regulation**

On April 18, 2018, the US Food and Drug Administration (FDA) [approved IDx-DR](#), a diagnostic system for screening and detecting diabetic retinopathy in patients and recommending a provider visit where necessary. This marked the first autonomous SI diagnostic solution authorized for use in patient care. This approval was soon followed by the approval of Automated Insulin Delivery Systems, which now fully automate blood glucose monitoring and insulin doses without a human in the loop. It has now been nearly a decade since the FDA greenlit the first automated diagnostic system, and, in June 2026, the FDA [granted approval to UpDoc](#), a solution leveraging patient-facing algorithms to augment care delivery.

The widespread interest in and adoption of SI/ML solutions in healthcare settings has raised significant questions about regulatory structure and approach. However, it is important to note that automated healthcare technologies have already been part of the regulatory landscape for years, from state and federal testing and evidence generation to deployment. While the scale of the regulatory challenge has rapidly expanded, the fundamental mechanisms already exist, although authorities are often fragmented across several agencies. Rather than inventing a radically new regulatory framework, policymakers should leverage established policy levers and a decade of hard science in the SI/ML field as a foundation to ensure patient safety, improve the quality of healthcare, and unlock the next decade of innovation.

Vega Health’s approach to SI oversight in healthcare is grounded in lessons learned through years of implementation work at Duke and beyond. SI in healthcare is not a static product that can be approved once and left alone. It is a dynamic system that can

drift, degrade, and produce unexpected outputs as patient populations shift, workflows change, and the underlying data systems evolve. This means that the right model for oversight is not solely pre-market approval but must extend post-implementation. The question is not only whether a specific SI solution is safe and effective to implement today, but whether the SI solution will continue to perform safely and realize the desired outcomes in a specific healthcare setting with a unique patient mix.

This has profound implications for what good regulation looks like, because healthcare organizations today do not always have the infrastructure, expertise, or resources to continuously monitor SI.

Vega Health strongly supports the FDA's effort to consider both pre-market evaluation and post-market monitoring approaches for GenAI-enabled medical devices. ***However, the selected approaches must be flexible, relevant to patients and providers, meaningful in real-world deployment contexts, and implementable in SI product lifecycle management.***

### **III. Value of a Total Product Lifecycle Approach to GenAI-Enabled Device Evaluation and Monitoring**

Vega Health supports the FDA's emphasis on a total product lifecycle (TPLC) approach to GenAI-enabled device evaluation and monitoring. Pre-deployment local evaluation and post-implementation monitoring are both critical to successful and responsible implementation of SI in healthcare settings.

Most SI systems entering the market today are highly differentiated from therapeutics or in vitro diagnostics that historically have relied on clinical trials for market access. For example, a therapeutic has a fixed chemical composition and mechanism of action (even if poorly understood). Assuming strong quality controls on the manufacturing of the therapeutic, the chemical composition and mechanism of action will be stable over time and across populations.

Those assumptions are not true for modern SI. SI systems today are probabilistic and the underlying algorithms and system architectures undergo frequent changes. Furthermore, the data generating process that feeds the SI systems is also highly dynamic as workflows change, health IT systems change, population characteristics change, and medical interventions and practice change. Given the highly dynamic nature of the SI systems themselves and the underlying data, SI systems are not well suited for front-loading clinical trials or premarket evaluation to ultimately secure market

access for an indication of use. Post-market surveillance at the local facility level is critical.

Premarket evaluation can establish a useful baseline for GenAI-enabled device performance, safety, and operational considerations; however, it is nearly impossible to test every circumstance that may arise after deployment. Additionally, generative SI [presents risks](#) that may emerge from changes in performance over time or deployment context, including drift, unexpected outputs, performance degradation, biased subgroup performance, data quality issues or mis-mapping, and updates that change black-box outputs.

As a result, safety and efficacy results from premarket evaluation will not suffice to manage risk over time. SI solutions deployed in clinical settings must collect and share ongoing evidence of performance so that regulators, deploying organizations, and end-users can maintain confidence in outputs and prevent any potential harm resulting from post-deployment performance drift.

Ongoing monitoring provides continued evidence that device performance remains within the baseline bounds established during premarket evaluation. It informs regulators, deploying organizations, and end-users when device performance changes meaningfully, prevents safety events, and informs deployment, re-evaluation, and decommissioning decisions.

#### **IV. Risk, Evaluation, and Monitoring of GenAI-Enabled Devices**

From experience of conducting real-world local evaluation and monitoring of both generative and predictive SI models in health systems, it is evident that a flexible framework for model classification and risk can accelerate the design, implementation, and interpretability of evaluation and monitoring schema.

A risk stratification framework that provides SI developers with a rubric for the level of evidence required for market access is critical. The risk stratification framework should focus on the following two axes: level of risk associated with the medical condition being addressed with the intended use, and; reversibility of interventions considered with the intended use. Level of human oversight and the time sensitive nature of decision making should not be included in the risk stratification framework. Rather than capturing the innate risk of a GenAI-enabled device, ‘human in the loop’ oversight is often used by SI developers to shield themselves from product liability and to shift accountability to human clinicians.

Time sensitivity was used as a risk dimension in the 2026 FDA clinical decision support guidance and created significant tension within health systems seeking to use SI to detect and treat acute conditions. Our CEO wrote a manuscript in 2024 with health system leaders from across the US about the tension created within health systems seeking to prioritize SI use cases focused on patient safety.

CDRH should not use the level of expertise of the human end-user as a dimension of risk. A key benefit of Gen-AI-enabled devices is to provide high-quality, personalized information to end-users at all levels of expertise. Similarly, the volume and variety of medical information is so vast and grows at an increasing pace. Generative SI can access and synthesize more up-to-date information than any clinician, no matter how specialized.

GenAI-enabled devices should be evaluated for how well they surface primary literature and sources for the end-user and how well they can synthesize this information. The ability to provide accurate traceability and provenance to source information is critical to evaluating generative SI devices. Whether the end-user is a specialist clinician, generalist clinician, advance practice provider, or layperson, the same evaluation standards should be applied to the GenAI-enabled device.

There should be protections for vulnerable populations that may be more prone to manipulation and inducement from use of GenAI-enabled devices. This may include children and individuals with cognitive disabilities.

Lastly, CDRH should create a distinction between ‘information gathering’ interactions and ‘intervention taking’ interactions with a GenAI-enabled device. A longitudinal dialogue will repeatedly transition between ‘information gathering’ and ‘intervention taking.’ ‘Information gathering’ interactions are lower risk and the GenAI-enabled devices should be assessed based on its accuracy and ability to surface relevant source materials. When a user (or agent) begins seeking information to take a specific intervention (‘intervention taking’), the two risk factors described above should be used to assess risk: level of risk associated with the medical condition being addressed, and; reversibility of interventions considered with the intended use.

## **V. Effective Design & Implementation of Premarket Evaluation of GenAI-Enabled Devices**

As discussed previously, there are well-documented limitations to the generalizability and longevity of premarket evaluation of SI solutions.

In this context, it is crucial to evaluate what the desired outcome of premarket evaluation will be for GenAI-enabled devices, which types of evidence will be necessary

to build confidence in safety and efficacy, and how premarket evaluation must be complemented by local validation and ongoing monitoring to mitigate risk.

A clear understanding of the purpose of each stage of CDRH's proposed premarket evaluation and post market monitoring approach – from benchmarking to clinical validation to ongoing monitoring – is critical to meaningfully advance safety, quality, and outcomes without imposing undue regulatory burden.

### **Competency-Based Approach**

The competency-based approach to GenAI-enabled device benchmarking may be useful in a future where GenAI-enabled devices seek approval to practice medicine broadly or within a given specialty area. However, the framework raises two concerns.

First, it is unclear whether the FDA is best suited to conduct competency-based benchmarking for broadly used GenAI-enabled devices. The GenAI-enabled device would not be assessed for a specific intended use, but would be assessed globally on a set of relevant clinical materials. The entity best suited to define competency-based benchmarking could be individual specialty boards or the American Board of Medical Specialists (ABMS).

Second, there will be a tension between generative SI performance on broad competency-based benchmarking and narrow capabilities for a specific intended use. Ultimately, a GenAI-enabled device may be safe and effective for a specific intended use while performing poorly on a broad benchmark. FDA should focus its regulatory activity on regulating GenAI-enabled devices for more narrowly scoped intended uses.

### **Clinical Confirmation**

Clinical confirmation (e.g., validation in context) is a crucial element of premarket evaluation of GenAI-enabled devices. The approaches proposed by CDRH are well-aligned with approaches to local model evaluation in clinical settings.

CDRH should provide more specific guidance on the level of clinical confirmation evidence required for generative SI across dimensions of risk. For example, the table below proposes example levels of evidence across the two risk dimensions highlighted above. Standardized patient interactions are less useful than assessments of generative SI product performance on real-world data. CDRH should also include the need for independent evaluation in this risk framework.

|                         |                                           | Severity of medical condition                                                                                                                                                                   |                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                            |
|-------------------------|-------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                         |                                           | Mild                                                                                                                                                                                            | Moderate                                                                                                                                                                                                                 | Severe                                                                                                                                                                                                                                                     |
| Reversibility of action | Reversible without any residual effects   | Retrospective evaluation on real patient inputs<br><br>Clinical adjudication of real cases                                                                                                      | Retrospective evaluation on real patient inputs<br><br>Independent evaluation on independently sourced real patient inputs<br><br>Clinical adjudication of real cases by independent clinicians                          | Retrospective evaluation on real patient inputs<br><br>Independent evaluation on independently sourced real patient inputs<br><br>Clinical adjudication of real cases by independent clinicians<br><br>Shadow deployment                                   |
|                         | Reversible with moderate residual effects | Retrospective evaluation on real patient inputs<br><br>Independent evaluation on independently sourced real patient inputs<br><br>Clinical adjudication of real cases by independent clinicians | Retrospective evaluation on real patient inputs<br><br>Independent evaluation on independently sourced real patient inputs<br><br>Clinical adjudication of real cases by independent clinicians<br><br>Shadow deployment | Retrospective evaluation on real patient inputs<br><br>Independent evaluation on independently sourced real patient inputs<br><br>Clinical adjudication of real cases by independent clinicians<br><br>Shadow deployment<br><br>Prospective clinical study |
|                         | Not reversible                            | Retrospective evaluation on                                                                                                                                                                     | Retrospective evaluation on                                                                                                                                                                                              | Retrospective evaluation on                                                                                                                                                                                                                                |

|  |  |                                                                     |                                                                     |                                                                     |
|--|--|---------------------------------------------------------------------|---------------------------------------------------------------------|---------------------------------------------------------------------|
|  |  | real patient inputs                                                 | real patient inputs                                                 | real patient inputs                                                 |
|  |  | Independent evaluation on independently sourced real patient inputs | Independent evaluation on independently sourced real patient inputs | Independent evaluation on independently sourced real patient inputs |
|  |  | Clinical adjudication of real cases by independent clinicians       | Clinical adjudication of real cases by independent clinicians       | Clinical adjudication of real cases by independent clinicians       |
|  |  | Shadow deployment                                                   | Shadow deployment                                                   | Shadow deployment                                                   |
|  |  | Prospective clinical study                                          | Prospective clinical study                                          | Prospective clinical study                                          |

### The Role of Local Validation Evidence

Local validation is a key component of many SI product lifecycle best practices, including those published by the [Health AI Partnership](#) and the [Joint Commission](#). The FDA should consider how local evaluation following the suggested clinical confirmation approaches in this discussion paper could be used to generate a meaningful national body of evidence to support regulatory approval of GenAI-enabled devices. By enabling SI developers or deployers to support regulatory submissions through local evaluation, the FDA can support the development of infrastructure necessary for health systems to safely, effectively, and responsibly implement SI broadly.

This approach would align incentives for SI innovators, health systems, and regulators, while also generating a more robust and immediately usable body of GenAI-enabled device performance data. Just as clinical trials have provided health systems with opportunities to pioneer emerging cures, extend opportunities for care to ever-expanding patient populations, advance clinical research, and fund key care priorities, local model evaluation studies as part of the regulatory approval process can confer benefits and build expertise in health systems. ***Even if local evaluation studies are not the only pathway to satisfy CDRH's clinical confirmation requirements, explicit inclusion of these evaluations would allow the FDA to support the development of essential health system infrastructure.***

## Value of Independent Validation

Independent evaluation of generative SI will be critical to generating the clinical confirmation evidence required for regulatory approval. The potential role of these independent third-party evaluators in evidence generation for GenAI-enabled device regulatory approval may be informed by lessons from similar programs at the federal and state level.

In this discussion paper, CDRH cites both the Accreditation Scheme for Conformity Assessment (ASCA) and Medical Device Development Tool (MDDT) programs as precedent for third-party evaluation within FDA programs. [Utah's AI Regulatory Sandbox](#), administered by its Office of AI Policy within the Department of Commerce, requires applicants to provide pre-deployment evidence and post-deployment monitoring validated by independent third-party evaluators. Other states have also turned to independent evaluators as part of their SI regulatory structures.

The need for independent third-party evaluation was surfaced from real-world experience. When Utah made regulatory sandbox approval decisions based on data gathered by the vendor and an analysis performed by the vendor, there was critique from many stakeholders, including patient advocacy groups, the state medical board, and SI development and evaluation experts. Through the course of broad stakeholder engagement, including the creation of a [Clinical AI Advisory Network](#) (on which our CEO, Mark Sendak serves), recommendations emerged for independent third-party evaluation.

In the case of Utah, independent evaluators provide services that directly support the state's [evidentiary standards](#) for sandbox participation. For the FDA to effectively incorporate independent third-party evaluators into the regulatory approval process for GenAI-enabled devices, it must first establish where on the risk spectrum independent evaluation is required for clinical confirmation. The table above provides an example of how this could be accomplished.

## VI. Post market Monitoring of GenAI-Enabled Devices can Support Quality, Performance Improvement, and Better Care Outcomes

Post market monitoring holds value for regulators, healthcare delivery organizations, and end-users. In order to optimize economies of scale, reduce regulatory burden, and deliver meaningful post market intelligence to stakeholders, the FDA can build on proven programs to build a centralized body of evidence for GenAI-enabled device quality, safety, efficacy, and healthcare performance improvement.

### **Post market Monitoring & Version Control**

A well-designed TPLC approach incorporating both premarket evaluation and post market monitoring can also allow the FDA to reduce regulatory burden. While initial premarket evaluation and benchmarking provide an initial baseline for performance, a [Predetermined Change Control Plan](#) (PCCP) can outline the parameters for additional evaluation to occur after significant updates. Rather than completing a new premarket submission, GenAI-enabled devices could simply resubmit clinical confirmation or benchmark scores from deployment sites in accordance with an evaluation plan established in the PCCP. Assuming this performance data was within acceptable ranges of the baseline, updates could be deployed.

### **Post market Monitoring Infrastructure & Centralized Reporting**

There are a wide range of approaches to SI safety, security, and performance monitoring across SI vendors and healthcare delivery organizations.

The FDA has the ability to set precedent for baseline elements of GenAI-enabled device monitoring in healthcare settings. As the FDA develops a framework for ongoing monitoring metrics or points of evidence for approved devices, it lays the groundwork for a default expectation for monitoring that health systems can expect from vendors. This groundwork is particularly valuable to community and rural hospitals that may need support to successfully implement and deploy SI/ML solutions for clinical and administrative workflows.

Vega Health was built on the belief that widespread digital maturity and SI adoption across the healthcare sector [must not exacerbate](#) the [digital divide](#) between the best-resourced academic medical centers and the community health centers that serve a high percentage of the U.S. population.

The leading academic medical centers and health systems in the United States are engines of innovation and early adopters capable of developing and piloting transformative technologies. However, these systems are able to innovate and experiment due to a degree of internal expertise and funding for research and development activities that most health systems lack.

Financial headwinds, poor payer mix, and shrinking margins have left many healthcare providers – particularly community and rural hospitals – without the luxury of launching innovation hubs, incubators, and SI implementation programs. Additionally, contracting power imbalances do exist between the largest technology vendors and small, rural, or single-site healthcare delivery organizations.



As a result, these systems also face challenges uniformly monitoring SI solutions in deployment. If the FDA requires approved GenAI-enabled devices to provide premarket evaluation evidence and ongoing monitoring reports at the deployment site, it could alleviate this imbalance and help all healthcare providers more effectively leverage SI solutions to improve outcomes.

CDRH – and the FDA broadly – should also consider how post market monitoring of GenAI-enabled devices can support the development of widespread, centralized SI performance data to support healthcare improvement.

The Sentinel Initiative is a powerful example of how localized safety and performance data can both inform regulatory decision-making and improve the quality of care. Local monitoring of GenAI-enabled devices – and other SI solutions – could be made available to the FDA through existing or similar infrastructure, allowing regulators to examine outcomes, review safety and efficacy, and advance research on SI-enabled healthcare improvement. Under this model, the FDA could both advance the development of monitoring infrastructure in health systems nationally and provide more comprehensive, objective reviews of evidence than periodic vendor-reported metrics.

## **VII. Conclusion**

Vega Health appreciates the opportunity to comment on this important discussion paper. We encourage CDRH to continue developing this framework for premarket evaluation, post market monitoring, and TPLC management for GenAI-enabled devices. If Vega Health can be a resource, please do not hesitate to reach out to Vega Health Co-Founder and CEO Mark Sendak ([Mark.Sendak@vegahealth.com](mailto:Mark.Sendak@vegahealth.com)).

Sincerely,

A handwritten signature in blue ink, appearing to read "Mark Sendak".

**Mark Sendak, MD, MPP**

Co-founder and CEO, Vega Health