

Comment of Skovos on FDA-2026-N-7874: Considerations for the Regulation of Generative AI-Enabled Medical Devices

September 29, 2026

Submitted via regulations.gov

Skovos builds governance tooling that hospitals and clinics use to run AI tools and AI agents after they go live. Our vantage point is the day after deployment: what the institution needs in place to know what an AI system is doing, to limit what it can do, and to stop it. We appreciate the Center's decision to seek early input, and we agree with the paper's premise that generative AI devices, and especially agentic ones, will need more of the safety case to be carried after market than before it. Our comments respond to Questions 18 through 22, 24, and 26, which concern postmarket monitoring, change control, and agentic devices. We do not offer views on the premarket evidence questions beyond how they connect to post-deployment controls.

Question 18: conditions for relying more on postmarket monitoring

Greater reliance on postmarket monitoring is reasonable only if the monitoring can actually happen at the site of use. In practice that depends on four things being present at the deploying institution, whatever the manufacturer's own program looks like.

First, an inventory. The institution can name every generative AI function in use, the version in production, an accountable owner, and the clinical settings it is authorized for. Without this, a signal from monitoring cannot be traced to a deployment, and a recall cannot be executed.

Second, written permissions. What the device is allowed to read, write, order, or send is recorded before use and checked when the action occurs, not reconstructed afterward. For a chat or summarization function this is simple. For an agentic function it is the entire safety case.

Third, an audit trail that the institution owns. A complete, tamper-evident record of inputs, outputs, and actions, held in the institution's custody rather than only in the manufacturer's systems. Manufacturer telemetry is valuable, but a hospital cannot investigate an event, respond to a patient, or answer a regulator from logs it has to request.

Fourth, a way to stop the device. The institution can suspend or revoke a device's ability to act immediately, on its own authority, and the stop is recorded with who did it and why.

We suggest the Center state these as outcomes a monitoring program must support (the deployer can produce the inventory, the log, and proof of the stop) rather than as prescribed technology. Where a device type cannot support them, for example because it acts through channels the institution cannot observe or interrupt, reduced premarket evidence would not be appropriate.

Question 19: approaches to postmarket performance evaluation

Periodic re-benchmarking, sample-based clinician review, and degradation monitoring all depend on the deployer being able to select records from a complete log. We recommend that the Center treat log completeness and custody as a precondition for any of the three approaches. Two additions would help. Triggering events for reassessment should include changes on the deployer's side, such as a new EHR version, a new clinical workflow, or a new patient population, not only changes to the model. And sample-based review should include a

sample of denied or out-of-scope actions, since the pattern of what a device tried to do and was prevented from doing is often the earliest signal of drift.

Question 20: machine-based supervisory agents

Supervisory agents can help, with two cautions. A supervisor built and controlled by the same manufacturer as the device it supervises does not give the institution independent assurance. And a supervisor that can only observe, but not enforce, adds reporting without adding safety. We suggest the Center consider supervision that sits at the point where the device acts on institutional systems, is configured by the institution, and can deny an out-of-scope action rather than only record it. The supervisor itself should be inventoried, permissioned, logged, and stoppable in the same way as the device.

Question 21: roles of healthcare institutions without diffusing manufacturer accountability

The cleanest division is by what each party can see. The manufacturer is accountable for the device's performance against its intended use and for acting on signals. The institution is accountable for the conditions of use: which functions are enabled, for whom, with what permissions, and with what record. Institutions should not be asked to validate models. They should be expected to keep the inventory, the permissions, the log, and the stop control, and to share log extracts with the manufacturer on request. Stating these institutional duties plainly would make manufacturer accountability clearer, not weaker, because it removes the argument that a failure was invisible to everyone.

Questions 22 and 24: modifications and third-party foundation model changes

A deployer cannot manage modification risk it does not know about. We recommend that any change-control approach require the manufacturer to notify deployers of model, prompt, or tool changes before they take effect in production, with a version identifier the deployer can record in its inventory. For foundation model changes initiated by a third party, the same notification obligation should flow through the manufacturer, and the manufacturer should be able to pin or roll back the model version on the deployer's behalf. Contract terms are helpful, but the practical control is technical: the deployer's ability to see the version that is running and to stop it if the version changed without notice.

Question 26: agentic devices

For agentic functions, the reduced opportunity for human review should be offset by controls on action rather than only by evidence about outputs. We suggest acceptance criteria and oversight for agentic devices include: a declared set of actions and data scopes the agent may use, with everything else denied by default; a check of each action against that set at the time it occurs; a record of every action taken and every action denied; and an institution-controlled stop that takes effect on the next action. These controls are familiar from how institutions already govern human and system accounts, and they can be verified without access to the model.

Summary

Whatever the premarket path, a generative AI device is safe in use only where the institution can list it, limit it, log it, and stop it. We encourage the Center to name those four capabilities as expectations for postmarket monitoring of generative AI devices, and to require that manufacturers make them possible for deployers. We would be glad to discuss any of this further.

Will Haver

Founder, Skovos

will@skovos.ai