

American Society of Echocardiography

October 16, 2026

Digital Health Center of Excellence
Center for Devices and Radiological Health
U.S. Food and Drug Administration
Submitted electronically via Regulations.gov

Re: Docket FDA-2026-N-7874

Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback

Dear CDRH and Digital Health Center of Excellence Leadership:

On behalf of the American Society of Echocardiography (ASE), we appreciate the opportunity to comment on FDA's discussion paper concerning generative artificial intelligence (GenAI)-enabled medical devices. ASE supports a risk-proportionate approach that enables useful innovation while requiring evidence that a device is safe and effective for its intended clinical use. We urge FDA to incorporate evaluation specific to cardiovascular imaging, meaningful clinician oversight, transparent change management, and manufacturer-led monitoring into any future framework.

ASE represents cardiovascular ultrasound professionals, including physicians and sonographers, and advances patient care through education, research, clinical standards, and advocacy. Our AI Policy Statement supports responsible innovation, transparent validation, and recognition of clinician expertise, with distinctions among assistive, augmentative, and autonomous functions.[1] GenAI could support image acquisition, synthesize imaging findings and prior studies, draft reports, and assist cardiovascular care decisions. These are potential applications, not assertions that any particular product is authorized or that every such function is a medical device.

We recognize that this paper seeks early stakeholder input and does not establish draft or final guidance.[2] The recommendations below address the discussion questions most relevant to echocardiography and cardiology practice. They are intended to help FDA develop a predictable framework without imposing the same evidentiary or operational requirements on every GenAI function.

I. Preserve clear boundaries for device regulation

Section III and Questions 1 and 2

FDA should provide cardiovascular imaging examples that distinguish regulated device functions from functions outside device oversight and explain how mixed-function products will be evaluated. A tool that

[1] ASE. AI Policy Statement. Approved October 25, 2023. Policy Statements. Link

[2] FDA. Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback. Sections III–VII and Appendix A. 2026. Link

formats a cardiologist's completed report presents a different regulatory question from a tool that interprets images or generates a patient-specific management recommendation. Classification should turn on the actual function, intended use, and applicable statutory criteria. Neither use of GenAI nor a clinician review statement should determine device status by itself.

Future policies should distinguish generative functions from conventional predictive AI, including image segmentation, automated quantification, and disease detection. Existing devices should not automatically acquire new GenAI requirements because they use AI or share a platform with a generative component. FDA should explain when interactions between functions warrant evaluation, particularly when a generated report or recommendation incorporates automated measurements. Common terminology should describe the clinician's role and the degree of device autonomy without substituting for FDA's device classification analysis.

II. Apply risk assessment to the actual cardiovascular workflow

Questions 1 through 6

ASE supports using device activity and the consequences of an incorrect output as organizing dimensions for risk assessment. FDA should supplement them with explicit consideration of the likelihood and detectability of error, reversibility, urgency, and the opportunity for effective intervention. In cardiovascular imaging, a draft report available for deliberate review has a different risk profile from intraprocedural guidance during structural heart intervention or an automated instruction that changes care before review is possible.

Risk assessment should address clinically important omissions as well as incorrect statements. A generated interpretation could mislead care if it overlooks limited image quality, incomplete Doppler acquisition, discordant valve measurements, or changes from a prior study. A plausible numerical output may be difficult to verify when the underlying image or signal has been altered. Evaluation should examine whether the device identifies these limitations, distinguishes missing data from normal findings, and defers conclusions when essential inputs are unavailable. FDA should show how these risks and safeguards affect the evidence expected.

A clinician's presence should reduce assessed risk only when the workflow allows meaningful review. Testing should demonstrate that intended users can inspect source images and measurements, recognize limitations, and correct an output before harm occurs. Validation should reflect the distinct roles of sonographers, interpreting cardiologists, and clinicians using focused or point-of-care ultrasound. For patient-facing tools, comprehension, access to a human contact, and clear escalation instructions are more useful safeguards than a generic disclaimer. Multi-turn testing should include evolving chest pain, dyspnea, or postintervention symptoms. Both unnecessary escalation and false reassurance should be measured, with acceptance limits reflecting their different consequences.

III. Validate the complete device with specialty relevant benchmarks

Questions 7 through 10 and 17

ASE supports benchmarking followed by clinical confirmation, provided the evidence applies to the complete device as deployed. Testing a foundation model alone cannot establish the performance of a product whose retrieval sources, prompts, interfaces, or connected tools alter its behavior. FDA should

expect evaluation of the configuration intended for clinical use, including foreseeable missing inputs, conflicting information, repeated runs, and failures of connected systems.

Benchmarks should use realistic cardiovascular cases and criteria developed with qualified imaging specialists and sonographers. Depending on intended use, cases could test ventricular function, valvular disease, hemodynamic assessment, congenital anatomy, and guidance of structural heart procedures. Testing should assess consistency between images, measurements, narrative conclusions, and recommendations. Clinically consequential omissions and fabricated findings should have distinct acceptance criteria; a high average score should not conceal an unacceptable safety failure rate. Sources supporting recommendations should be identifiable and current.

Sponsors should demonstrate separation of development and evaluation data and use independent, sequestered cases where feasible. Sponsor-developed tests may fill gaps but should include a justification of clinical relevance and independent review. For multimodal products, evaluation should verify that images, cine loops, Doppler data, electrocardiographic timing, and clinical records are correctly linked to the same patient and examination. When integration of CT, cardiac MRI, or other modalities is claimed, sponsors should validate that integration and show safe handling of discordant findings. Evidence for one modality should not establish performance in another.

IV. Protect image integrity and quantitative reliability

Questions 1, 9, 12, 14, and 17

FDA should explicitly address generative image reconstruction, enhancement, and synthesis when used for diagnosis or measurement. Testing should establish that processing does not introduce or conceal clinically important anatomy or change measured function and flow. In echocardiography, assessment should examine preservation of temporal relationships across cardiac cycles, structural detail, and Doppler information. A visually persuasive image is not sufficient evidence of diagnostic fidelity.
Generated or materially altered images should be clearly identified, with source data accessible for clinical review where applicable. Clinicians should be able to distinguish measured values from inferred or generated values and trace conclusions to the relevant image, contour, or Doppler envelope. Quantitative validation should characterize bias, agreement, and repeatability across the intended range, including performance near thresholds that influence clinical decisions. Correlation alone should not establish interchangeability with a clinical reference method.

For longitudinal imaging, evaluation should determine whether changes in a model, acquisition system, or processing method can create an apparent change in cardiac function or disease severity. The device should disclose when studies cannot be compared reliably. Training and workflow assessment should preserve the sonographer's ability to obtain additional views and the interpreting physician's ability to review and revise measurements. These safeguards should be evaluated alongside any claimed gains in speed or access.

V. Match clinical confirmation and performance standards to risk

Questions 11, 12, 14, and 15

FDA should permit clinical confirmation methods proportionate to the intended use and risk, including retrospective evaluation, clinician adjudication, and shadow deployment when appropriate. However, retrospective agreement alone is unlikely to resolve questions about user reliance or autonomous action.

Devices that materially influence urgent cardiac triage, treatment selection, or intraprocedural decisions should generally require prospective evidence in representative workflows sufficient to address those risks. A randomized trial should be required when needed to answer the relevant safety or effectiveness question, rather than imposed uniformly.

For specialist functions, the comparator should reflect applicable specialty standards and qualified expert assessment, rather than assume that a median clinician's performance is an adequate safety threshold. Expert panels should manage conflicts of interest, document disagreements, and recognize acceptable alternative responses. When a device is intended to assist clinicians, the principal evaluation should compare the clinician-device team with an appropriate usual-care comparator and examine automation bias, correction burden, and missed errors. Autonomous use requires separate evidence; results from supervised use should not be extrapolated automatically.

Acceptance criteria should be justified for the clinical task. Potential measures include diagnostic concordance, clinically significant measurement error, correct severity classification, appropriate follow-up recommendations, and time to escalation. Sample sizes should support confidence intervals for critical error rates, with clustering by patient, clinician, or site addressed where relevant. Efficiency gains should be assessed alongside safety. Improved access may be an important benefit, but comparison with delayed or unavailable specialist care should not excuse failure to meet an appropriate minimum clinical safety standard.

VI. Establish generalizability using representative patient data

Questions 12 and 13

Evaluation should reflect the populations, equipment, and settings included in the intended use. Relevant variation includes age, sex, body habitus, rhythm, image quality, disease severity, and altered anatomy. Testing should address scanner vendors, acquisition protocols, and user experience, including community laboratories and point-of-care settings when intended. Performance in adult transthoracic echocardiography should not be extrapolated automatically to transesophageal, stress, pediatric, or fetal echocardiography. Congenital heart disease and postintervention anatomy warrant specific assessment when included. Multilingual patient outputs should be validated for comprehension. Sponsors should disclose subgroup limitations and pursue additional evaluation where needed.

Synthetic data can supplement testing of unusual scenarios and stress conditions, but should not replace real patient evidence needed to establish clinical performance. In particular, data generated by a related model may reproduce the assumptions or blind spots of the device under evaluation. Synthetic and real-patient results should be reported separately unless pooling has a defensible statistical and clinical basis. Real-world and international data should be usable when sponsors demonstrate relevance to U.S. practice, including clinical terminology, guideline differences, workflow, and patient characteristics.

VII. Use independent expertise while preserving access to evaluation

Question 16

Qualified independent third parties can help maintain evaluation datasets, adjudicate outputs, and develop reproducible testing methods. FDA should include cardiovascular imaging specialists, sonographers, and patient representatives in developing methods for imaging-related functions. Adjudicators should have expertise appropriate to the claimed task, with documented independence from the device sponsor and

underlying model developer. An AI system used for scoring should itself be validated against qualified human review, especially for safety-critical judgments.

Participation should remain practical for smaller developers and should not depend on a single proprietary benchmark provider. FDA should publish transparent qualification criteria, permit justified alternative methods, and retain responsibility for regulatory determinations. ASE recommends that specialty societies contribute to clinical criteria and education while maintaining appropriate conflict-of-interest safeguards.

VIII. Keep postmarket monitoring accountable and feasible

Questions 18 through 21

Postmarket monitoring should complement evidence sufficient for authorization. ASE would support accepting additional residual uncertainty only when the potential harm is limited and reversible, adequate baseline performance has been established, and an operational monitoring and response plan can detect and contain emerging problems promptly. Monitoring cannot reliably protect a patient from a serious error that causes harm before detection. Greater reliance on future evidence is therefore inappropriate when unresolved uncertainty concerns potentially irreversible harm from autonomous action or delayed diagnosis or treatment.

Monitoring plans should specify the deployed version, sampling strategy, metrics, review cadence, and thresholds for investigation, notification, restriction, or suspension. They should capture clinically important omissions, near misses, inappropriate escalation, and performance differences across settings or populations, rather than depend solely on adverse-event reports. Triggered reassessment should follow changes likely to affect clinical behavior, including important source-guideline updates and changes in patient populations, model components, or connected systems.

Manufacturers should retain responsibility for device performance monitoring, investigation, and corrective action. Practices should have simple ways to report problems and obtain timely responses; they should not be expected to build bespoke surveillance infrastructure or provide routine, uncompensated expert adjudication for a commercial product. Monitoring obligations should be achievable in office practices, community echocardiography laboratories, and hospitals. Device integration should support existing quality processes rather than require duplicate manual data entry. Consistent with ASE's AI policy, responsibilities should not place a disproportionate burden on frontline clinicians.[1] Data collection should protect confidentiality and minimize unnecessary retention of identifiable records. Machine-based supervisory agents may support detection, but need independent validation, testing for correlated failures, and a pathway to human review.

IX. Control updates and foundation model dependencies

Questions 22 through 25

ASE supports predictable change management, including use of FDA-authorized predetermined change control plans where appropriate.[3] FDA should distinguish changes manageable within a manufacturer's quality system, changes covered by an authorized plan, and changes requiring additional premarket review. The distinction should reflect the expected effect on clinical behavior. A seemingly small change to

[3] FDA. Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions. Final guidance. August 2025. Link

prompts, retrieval sources, or refusal behavior may matter more clinically than its technical description suggests.

Updates should undergo regression testing of affected functions and critical safety behaviors before deployment, with clinical confirmation when the change could materially alter performance in practice. Clinicians should receive understandable notice of changes to intended use, limitations, or recommendations. Device records should identify the version used for a clinical output. Manufacturers should maintain rollback or safe suspension procedures and prevent unvalidated third-party updates from reaching clinical use through version control, monitoring, and appropriate agreements with model providers.

A voluntary Foundation Model Device Master File could reduce duplicative review if it contains usable information about supported versions, known failure modes, evaluation results, and update processes. FDA should preserve alternative evidentiary routes when a model developer does not participate. A master file should neither imply authorization of the model for medical use nor substitute for the sponsor's device-specific evidence. Confidential submissions should be complemented by user-facing information sufficient to understand device capabilities, limitations, and clinically meaningful changes.

X. Require explicit safeguards for autonomous clinical actions

Question 26

For agentic devices, FDA should evaluate complete sequences of actions and interactions with external systems, including wrong-patient errors, incorrect tool outputs, prompt injection, partial completion, and failures during handoffs. A device that drafts an imaging interpretation differs from one that finalizes a diagnosis, executes a clinical order, controls image acquisition, or directs another device during a cardiac intervention. Authorization and labeling should identify which actions are permitted and what human oversight is required.

High-consequence or irreversible actions should require explicit clinician authorization unless autonomous performance is specifically supported by evidence and authorized for that use. Safeguards should include appropriately limited permissions, confirmation of patient identity, effective stop controls, and an auditable record of inputs and actions. Testing should establish safe behavior during loss of connectivity or tool failure, including return to a defined clinical workflow. Passing individual component tests should not substitute for evaluating the safety of the complete action sequence.

ASE urges FDA to engage specialty societies and patients as it develops any subsequent guidance, with clear echocardiography and cardiology examples and an opportunity for further public comment. A workable framework should establish credible clinical evidence, maintain manufacturer accountability, and allow clinicians and patients to benefit from useful innovation without creating impractical obligations for the practices delivering care.

Thank you for considering these comments.

Sincerely,



Cynthia C. Taub, MD, MBA, FASE
President, American Society of Echocardiography