ConnectedHealthInitiative

October 19, 2026

Dockets Management Staff (HFA-305)
Food and Drug Administration
Department of Health and Human Services
5630 Fishers Lane, Rm. 1061
Rockville, Maryland 20852

RE: Comments of the Connected Health Initiative on *Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback* [Docket No. FDA-2026-N-7874]

The Connected Health Initiative (CHI) submits these comments in response to the discussion paper issued by the Food and Drug Administration (FDA) Center for Devices and Radiological Health (CDRH) on the regulation of generative artificial intelligence (GenAI)-enabled medical devices.[1] The discussion paper asks for input on how FDA should assess the risk of GenAI-enabled devices, what evidence should support their marketing authorization, how they should be monitored and modified after they reach the market, and how FDA should approach foundation models and agentic AI systems. CHI appreciates that FDA has chosen to gather stakeholder views before proposing guidance, and these comments begin with CHI's general views and then respond to the paper's sections and discussion questions in order.

I. Introduction and Statement of Interest

CHI is the leading multistakeholder policy and legal advocacy effort dedicated to improving health outcomes while reducing costs through digital health innovation. CHI aims to realize an environment in which Americans see improvements in their health through policies that allow connected health technologies to advance outcomes and reduce costs. Our members develop and deploy connected health technologies across a wide range of use cases, including telehealth platforms, remote physiologic and remote therapeutic monitoring, software as a medical device (SaMD), prescription digital therapeutics, clinical decision support, and AI-enabled diagnostic and workflow tools. We advocate before Congress, federal agencies, and state legislatures and agencies on coverage and payment, privacy and security, health data interoperability, FDA regulation of digital health, and the role of AI in care delivery. For more information, see https://connectedhi.com.

CHI has engaged with FDA on the regulation of software and AI throughout the development of its digital health policies, including its recent comments on the real-world performance of AI-enabled

[1] U.S. Food and Drug Administration, Center for Devices and Radiological Health, Digital Health Center of Excellence, *Considerations for the Regulation of Generative AI-Enabled Medical Devices: Discussion Paper and Request for Feedback* (Aug. 18, 2026), https://www.fda.gov/media/194242/download. See also U.S. Food and Drug Administration, FDA Seeks Public Feedback to Inform Regulatory Approach for Generative AI-Enabled Medical Devices (Aug. 18, 2026), https://www.fda.gov/news-events/press-announcements/fda-seeks-public-feedback-inform-regulatory-approach-generative-ai-enabled-medical-devices.

devices, the Digital Health Advisory Committee's review of GenAI-enabled digital mental health devices, the International Medical Device Regulators Forum (IMDRF) technical framework for AI life cycle management, the Medical Device User Fee Amendments (MDUFA VI) reauthorization, and FDA's report on non-device software functions.[2] CHI is also an appointed member of the American Medical Association's (AMA) Digital Medicine Payment Advisory Group, which brings together experts to identify barriers to digital medicine adoption and propose solutions on coding, payment, and coverage. Many CHI members are developing products that use GenAI, either within a device software function or in tools that support clinicians, patients, and administrative staff outside of FDA oversight. These developers need a regulatory approach that is predictable and proportionate to risk and that does not change from one agency to the next, and CHI offers these comments to help FDA build one.

II. General Views

a. CHI's Health AI Principles and Health AI Roles and Interdependency Framework

CHI urges FDA to draw on two resources that CHI developed with clinicians, developers, health systems, and patient organizations, both of which CHI released in updated form in September 2026. The first is CHI's Health AI Principles for State Policymakers, the most recent statement of CHI's health AI policy principles, which sets out how policymakers can support the safe and effective use of AI in healthcare.[3] Those principles call for quality assurance and oversight that is risk-based and calibrated to intended use, clinical significance, the consequence of an error, the degree of autonomy, and the human oversight in place. They distinguish FDA-regulated devices from non-device clinical decision support, administrative tools, and patient-facing tools, and they treat FDA's oversight of AI-enabled devices, together with the National Institute of Standards and Technology (NIST) AI Risk Management Framework (AI RMF), as the base that other policymakers should build on without duplicating. The principles also ask that disclosure expectations be calibrated to risk and audience and, consistent with FDA labeling, that bias be addressed through

[2] See, e.g., Comments of the Connected Health Initiative on Measuring and Evaluating Artificial Intelligence-Enabled Medical Device Performance in the Real World (Dec. 1, 2025) (Docket No. FDA-2025-N-4203), https://connectedhi.com/wp-content/uploads/2025/12/CHI-Comments-re-FDA-Measuring-and-Evaluating-AI-Med-Device-Performance-in-the-Real-World-1-Dec-2025-w-appendices.pdf; Comments of the Connected Health Initiative to the Digital Health Advisory Committee on Generative Artificial Intelligence-Enabled Digital Mental Health Medical Devices (Oct. 17, 2025) (Docket No. FDA-2025-N-2338), https://connectedhi.com/wp-content/uploads/2025/10/CHI-Comments-re-FDA-GAI-Digital-Mental-Health-Medical-Devices-17-Oct-2025-w-appendicies.pdf; Comments of the Connected Health Initiative on IMDRF Draft Technical Framework for AI Life Cycle Management (N93) (July 10, 2026), https://connectedhi.com/wp-content/uploads/2026/07/CHI-Comments-IMDRF-N93-AI-Life-Cycle-10-Jul-2026.pdf; Comments of the Connected Health Initiative on the Medical Device User Fee Amendments (MDUFA VI) Reauthorization Public Meeting (Aug. 7, 2026) (Docket No. FDA-2026-N-6655), https://connectedhi.com/wp-content/uploads/2026/08/CHI-Comments-re-MDUFA-VI-Letter-7-Aug-2026.pdf; Comments of the Connected Health Initiative on the Report on Non-Device Software Functions Under Section 3060 of the 21st Century Cures Act (Aug. 13, 2026) (Docket No. FDA-2018-N-1910), https://connectedhi.com/wp-content/uploads/2026/08/CHI-Comments-FDA-NonDevice-Software-13-Aug-2026.pdf.

[3] Connected Health Initiative, Health AI Principles for State Policymakers (Sept. 2026), https://connectedhi.com/wp-content/uploads/2026/09/Health-AI-Principles-for-State-Policymakers-1.pdf.

testing across subpopulations and monitoring for drift, and that compliance requirements not undercut intellectual property and trade secrets.

The second is CHI's Health AI Roles and Interdependency Framework, which defines the stakeholders in the health AI value chain, from foundation model developers through AI platform developers, health AI platform developers, digital health solution developers, deploying organizations, clinicians and administrative users, payers, patients, standard-setting organizations, certification bodies and test beds, and accrediting and licensing bodies.[4] For each group, the framework recommends roles that support safety, efficacy, and transparency, and maps those roles to the functions in the NIST AI RMF. The framework rests on the view that responsibility for AI safety, efficacy, and transparency is shared across the value chain. The level of review and oversight should be proportionate to potential harms, so that validation and monitoring scale with the consequences of an error and one standard does not apply to every tool. Responsibility for a given risk should fall on the party with the knowledge, ability, and incentive to minimize it.

The discussion paper raises the same questions the framework was built to answer, including what a foundation model developer should tell the companies that build on its model, how a device manufacturer should account for an upstream model update, and what deploying organizations and clinicians can contribute to monitoring without diluting the accountability of the manufacturer. The framework recommends, for example, that foundation model developers document known bias and safety issues and their mitigations in transparency documentation such as system cards, give guidance on adapting the model to foreseeable uses, and provide version information that explains how each release differs from the last so that downstream parties can decide whether to revalidate. It recommends that digital health solution developers specify appropriate uses and design interfaces that let users act safely on outputs, and that deploying organizations monitor performance over time and revalidate after significant changes in proportion to the harm an error could cause. CHI refers to both resources throughout these comments and encourages FDA to use them as it develops its approach.

b. Building on Existing Authorities and a Risk-Based, Least Burdensome Approach

CHI supports the paper's statement that FDA does not regulate GenAI as such and that its oversight attaches to products that meet the statutory definition of a device. CHI also supports FDA's focus on the GenAI-enabled device software function and not on every function in a product, consistent with FDA's policy on multiple function device products.[5] Many GenAI tools now in use, including tools that draft clinical documentation, summarize records, support scheduling and billing, or provide general health and wellness information, are not devices, and others fall within the clinical decision support exclusion in section 520(o)(1)(E) of the Federal Food, Drug, and Cosmetic Act (FD&C Act) or within FDA's enforcement discretion policies, and FDA's January 2026 updates to its clinical decision support and general wellness guidances clarified where those lines

[4] Connected Health Initiative, Health AI Roles & Interdependency Framework (rev. Sept. 2026), https://connectedhi.com/wp-content/uploads/2026/09/CHI-Health-AI-Roles-Interdependencies-FW-rev-Sept-2026.pdf.

[5] U.S. Food and Drug Administration, Multiple Function Device Products: Policy and Considerations, Guidance for Industry and Food and Drug Administration Staff (July 2020), https://www.fda.gov/regulatory-information/search-fda-guidance-documents/multiple-function-device-products-policy-and-considerations.

fall.[6] CHI asks FDA to state in any future guidance that the use of GenAI does not by itself bring a software function within device oversight, and that the risk framework in Section IV of the paper applies only after FDA has determined that a function is a device that FDA intends to oversee. For GenAI-enabled functions that give users informational, non-directive output where an error has limited consequences, such as a feature that helps a user understand the user's own wearable data and suggests consulting a clinician, FDA should continue the enforcement discretion it already applies to symptom checkers and to tools that help patients self-manage their conditions, since the use of a generative model rather than a decision tree does not change the function or its risk.[7] Regulatory status should depend on a function's intended use and risk, not on the type of company or technology that offers it, so that similar functions with similar intended uses receive the same treatment. CHI also asks FDA to explain in any future guidance how the framework interacts with its general wellness policy, and to confirm that a non-directive reference to a disease or condition, made only to give context on a user's own wearable data and paired with a suggestion to consult a clinician, does not by itself make a wellness product a device.

CHI also encourages FDA to rely on the tools it already has, including the total product life cycle (TPLC) approach, predetermined change control plans (PCCPs), the Breakthrough Devices Program, FDA-recognized consensus standards, the Accreditation Scheme for Conformity Assessment (ASCA), and the Medical Device Development Tools (MDDT) program, and not create a separate regulatory track for GenAI. The paper notes that FDA intends to employ least burdensome principles, and CHI asks that this commitment be applied to each element FDA eventually adopts, with particular attention to the costs that new testing and documentation expectations impose on developers. FDA should also continue to work through the IMDRF so that its approach to GenAI-enabled devices is consistent with the approaches of other regulators, which allows developers to rely on a single body of evidence across markets. Developers that follow international standards for AI management systems, such as ISO/IEC 42001, should be able to draw on that work in their device-level documentation.[8] CHI also supports transparency that is calibrated to risk and to the audience that needs it. The labeling of a GenAI-enabled device should give clinicians and, where relevant, patients the information they need to use it appropriately, including its intended use and users, its known limitations, its performance across relevant subpopulations, a summary of the characteristics of the data used to develop and test it, and how and when it will be updated. This device-level information can be provided without requiring the disclosure of trade secrets, model weights, or the proprietary details of an underlying foundation model. Finally, CHI urges FDA to coordinate with other parts of the Department of Health and Human Services (HHS), including the Assistant Secretary for Technology Policy/Office of the National Coordinator for Health Information

[6] U.S. Food and Drug Administration, Clinical Decision Support Software, Guidance for Industry and Food and Drug Administration Staff (Jan. 2026), https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software; U.S. Food and Drug Administration, General Wellness: Policy for Low Risk Devices (Jan. 2026), https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-wellness-policy-low-risk-devices.

[7] U.S. Food and Drug Administration, *Policy for Device Software Functions and Mobile Medical Applications*, Guidance for Industry and Food and Drug Administration Staff (Sept. 28, 2022), https://www.fda.gov/media/80958/download (Section V.B and Appendix B).

[8] International Organization for Standardization and International Electrotechnical Commission, ISO/IEC 42001:2023, Information technology, Artificial intelligence, Management system (Dec. 2023), https://www.iso.org/standard/42001.

Technology (ASTP/ONC) and the Centers for Medicare & Medicaid Services (CMS), so that developers do not face conflicting transparency and evidence expectations for the same product.

III. Considerations for the Assessment of Risk (Section IV, Questions 1 to 6)

a. The Two-Axis Framework (Questions 1 and 2)

CHI supports the paper's two-axis framework, which assesses risk according to the degree and independence of the device's activity and the consequence of relying on an incorrect output. The framework is consistent with CHI's view that oversight should scale with potential harm, and it closely tracks the IMDRF SaMD risk categorization, which considers the significance of the information a software function provides to a healthcare decision and the state of the healthcare situation or condition.[9] CHI recommends that FDA explain how the two approaches relate so that developers who already classify their products under the IMDRF categories can apply FDA's framework without starting over. The activity axis also corresponds in large part to the taxonomy in Current Procedural Terminology (CPT) Appendix S, which describes the work performed by AI as assistive, augmentative, or autonomous, and FDA should consider how its terms can match those used in coding so that labeling and coding decisions describe the same product in the same way.[10]

On Question 1, CHI recommends that FDA treat the additional factors it lists, such as reversibility, time pressure, and traceability, as considerations that move a device along one of the two existing axes. In particular, the safeguards built into a device and the clinical setting in which it is used, including whether a licensed clinician meaningfully reviews the output before it is acted on, where the clinician knows that AI was used and has the information and time needed to evaluate the output, should lower the assessed risk. Adding more independent dimensions would make the framework harder for developers and FDA reviewers to apply consistently.

On Question 2, CHI agrees that directiveness is a continuum that depends on substance and context. The paper notes that a statement directing the user to consult a physician does not by itself reduce directiveness, and CHI agrees that FDA should look to what the output actually conveys. For output that is otherwise informational and non-directive, however, encouraging the user to consult a clinician is an appropriate safeguard, and FDA's general wellness guidance already allows a wellness product, in some circumstances, to tell users that evaluation by a healthcare professional may be helpful.[11] CHI asks FDA to make clear that this continuum operates within the statutory framework, so that a function providing information to a healthcare professional who can independently review its basis remains outside device oversight under section 520(o)(1)(E) when it meets the statutory criteria, and so that general health information that

[9] International Medical Device Regulators Forum, "Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations, IMDRF/SaMD WG/N12FINAL:2014 (Sept. 18, 2014), https://imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf.

[10] American Medical Association, CPT Appendix S: Taxonomy for Artificial Intelligence in Medical Services & Procedures, https://www.ama-assn.org/practice-management/cpt/cpt-appendix-s-ai-taxonomy-medical-services-procedures.

[11] U.S. Food and Drug Administration, *General Wellness: Policy for Low Risk Devices*, supra note 6.

does not relate to the diagnosis or treatment of a particular patient's disease or condition is not treated as a device function.

b. Patient-Facing Functions and Intended Users (Questions 3 and 4)

CHI agrees with the paper that FDA should not underestimate patients, and patient-facing tools can extend access to reliable information and timely care between visits, particularly for patients in rural and underserved areas and for those managing chronic conditions. CHI agrees that relying on an incorrect output can carry greater consequences when a patient acts on it without a clinician's involvement, and FDA's framework can account for that on the consequence axis. CHI recommends that FDA assess patient-facing functions on the same two axes that apply to other functions, and that FDA give weight to safeguards that are designed for lay users, including plain-language communication, clear statements of what the tool can and cannot do, and reliable escalation to a clinician or to emergency care. Escalation to a licensed clinician, and the involvement of the patient's physician where the intended use calls for it, are among the most important of these safeguards, and the evidence for a patient-facing device should show that it recognizes when a patient needs clinical care and directs the patient there.

Many patient-facing tools are offered as part of an established care relationship. They draw on the patient's own medical record with the patient's consent, they route the patient to the licensed clinicians who provide the patient's care when clinical judgment is needed, and their conversations are monitored as they occur by automated tools and by clinicians. A tool designed in this way supports the patient's relationship with the care team and is not intended to diagnose or treat without a clinician's involvement, and its risk is lower than that of a standalone tool that gives the same information with no clinician available. CHI asks FDA to account for this integration when it assesses the risk of a patient-facing function, and to give weight to testing before deployment that shows the tool performs at least as well as clinicians on decisions that are critical to safety, such as recognizing an emergency. On Question 4, the device's labeling should state its intended user, whether a generalist, a specialist, or a patient, and its benchmarking and clinical confirmation should reflect that user, which addresses the differences among users without adding a separate risk factor.

c. Multi-Turn Interactions and Escalation (Questions 5 and 6)

CHI recognizes that a conversation that begins with general information may move toward specific advice. Intended use should continue to be defined by the manufacturer's labeling, design, and reasonably foreseeable use, and the evidence that a device stays within that intended use should come from the scope maintenance and boundary adherence testing the paper describes in Appendix A. CHI recommends that FDA not define a device's intended use by the most directive output a user could elicit through persistent or adversarial prompting, since that approach would make it difficult for any developer to know in advance which pathway applies. Where testing shows that a device drifts beyond its intended use, the developer should address the drift through design changes and guardrails, and FDA should evaluate the device on whether those controls work.

On Question 6, CHI agrees that both under-escalation and over-escalation are risks. Under-escalation can delay needed care, and over-escalation can send patients to emergency departments unnecessarily and add to clinician workload. CHI recommends that acceptance criteria for escalation be set according to the clinical context and intended population, with a lower

tolerance for missed escalation where the consequences of delay are severe, and that FDA compare escalation performance to the performance of the care pathway the patient would otherwise use, such as a nurse advice line or a patient portal message.

IV. Premarket Evaluation (Section V, Questions 7 to 17)

a. The Competency-Based Approach and Level of Evidence (Questions 7 and 8)

CHI supports the paper's competency-based approach, which combines high-throughput benchmarking of the final, user-facing device with clinical confirmation of its performance in real or clinically representative settings. The approach should work through FDA's existing intended-use framework, so that an authorization covers the device as configured and deployed for the uses described in its labeling and does not imply clinical competence beyond those uses. This approach suits devices whose outputs are open-ended and whose performance depends on the whole system, including the model, prompts, retrieval, guardrails, and user interface. CHI agrees with the paper that a prospective clinical study is not always necessary, and CHI recommends that FDA tie the expected evidence directly to the device's position on the two axes. Devices that provide non-directive information with low consequences of error may be supported mainly by benchmarking and a retrospective or shadow-deployment confirmation, while devices that direct or take action in high-consequence settings may warrant prospective evidence. Publishing this relationship, with examples across several device types, would give developers the predictability they need to plan their studies.

b. Benchmarks and Construct Validity (Questions 9 and 10)

The benchmarking elements in Appendix A, which address safety, clinical proficiency, generalizability, and agentic competencies, provide a useful structure. CHI recommends that FDA treat them as a menu from which sponsors select the elements relevant to the device's intended use and risk, with a justification for any element that is omitted, and should not expect every device to address every element. Health AI benchmarks are developing quickly, and FDA should recognize consensus benchmarks and test methods through its standards recognition program as they mature. Until then, sponsor-developed benchmarks should be acceptable where the sponsor shows that the benchmark reflects the intended use and population, uses prespecified methods and acceptance criteria, keeps test data separate from training and tuning data, and is adjudicated by qualified clinical experts. CHI recommends that sponsors be able to support construct validity by showing that benchmark results are consistent with results from the clinical confirmation stage, and that FDA allow this showing to build over time through postmarket data where the device's risk permits.

c. Clinical Confirmation, Statistical Approaches, and Synthetic Data (Questions 11 to 13)

The range of clinical confirmation approaches in the paper, from retrospective evaluation and shadow deployment to standardized patient interactions and prospective studies, gives sponsors appropriate flexibility. CHI recommends that the sponsor select the approach based on the device's risk and intended use and explain the choice in a presubmission, and that FDA use the Q-Submission program to reach early agreement on the approach. Decentralized and remote study elements can make clinical confirmation more representative and less costly, and FDA should

encourage them. CHI also supports the use of synthetic data to supplement real-world data, especially to test rare but serious scenarios and to evaluate performance in subgroups with limited real data, where the sponsor explains how the synthetic data were generated, shows that they are representative of the deployment setting, and does not rely on the same model class to generate both the test data and the outputs being tested. Because the most serious failures of a GenAI-enabled device may be rare, CHI also recommends that statistical approaches look beyond average performance to the frequency and severity of high-consequence errors, including the omission of clinically important information such as a contraindication or a red-flag symptom, using targeted test sets that are large enough to estimate those rates with confidence.

d. Comparators and Counterfactual Care (Questions 14 and 15)

CHI recommends that FDA select comparators based on the device's intended use and the care it is designed to support. For a device intended to be used under the review of a clinician, the performance of the clinician and the device together is usually the most meaningful measure, and FDA should not require the device alone to match a specialist panel when the labeling calls for clinician review. CHI supports FDA's consideration of the counterfactual, because many GenAI-enabled devices are designed for settings where a specialist is not available, and comparing them only to an idealized standard of care would understate their benefit and could keep useful tools from patients who would otherwise have no comparable resource. FDA should allow sponsors to show benefit relative to the care realistically available to the intended population, while still requiring that the device meet safety thresholds appropriate to the consequences of error.

e. Independent Third Parties (Question 16)

Consistent with CHI's framework, which recommends that certification bodies and test beds offer voluntary processes for verifying conformity with consensus standards, CHI supports a role for independent third parties that sponsors may choose to use. FDA's existing programs, including ASCA and MDDT, provide a sound basis for that role because they rest on FDA-recognized standards and FDA qualification. CHI urges FDA not to require sponsors to use a third party or to designate one organization, or a small group of them, as gatekeepers. FDA should also avoid relying on sequestered datasets that only one provider can offer. Those approaches would raise costs and cause delays for sponsors, and they would give some organizations an advantage unrelated to the quality of their products. Any third party should have no financial ties to the sponsors it evaluates or to their competitors, and it should follow published methods and protect the confidential information it receives.

f. Multimodal Models (Question 17)

The competency-based approach should apply to multimodal vision-language models, with benchmarking elements adapted to the input types the device accepts, such as image quality and acquisition variability, and with clinical confirmation that reflects the settings where those inputs are captured. CHI recommends that FDA address multimodal devices through the same framework, so that developers can apply one set of expectations as their products add capabilities.

V. Postmarket Monitoring (Section VI, Questions 18 to 24)

a. Balancing Premarket and Postmarket Evidence (Questions 18 to 20)

CHI supports FDA's view that postmarket evidence can complement premarket evidence for GenAI-enabled devices, and that stronger postmarket monitoring can justify a reduced premarket burden for appropriate devices. That tradeoff is best suited to devices where the consequences of error are lower and where monitoring can detect problems before they cause serious harm, and it should not replace clinical confirmation for devices that direct or take action in high-consequence settings. This approach is consistent with the TPLC model and with CHI's longstanding support for the use of real-world evidence. CHI recommends that FDA describe the conditions under which it would accept this tradeoff, including a monitoring plan with defined performance metrics and triggers for action, and that it confirm that the tradeoff will not be used to add postmarket obligations on top of an unchanged premarket burden. The monitoring approaches in the paper, including periodic re-benchmarking, sample-based clinician review, and performance degradation monitoring, are all reasonable options, and the sponsor should select among them based on the device's risk, with the frequency of reassessment tied to the consequence of error and to the rate at which the device or its inputs change. Monitoring results are more useful when they are reported by care setting and relevant subpopulation. CHI also encourages FDA to work with stakeholders on a common, voluntary taxonomy for AI-related events, so that reports from manufacturers, deploying organizations, and clinicians can be compared and acted on. Machine-based supervisory agents can help monitor large volumes of outputs, including by reviewing conversations as they occur and referring those that need attention to a clinician, provided that the sponsor validates their performance in proportion to the role they play and keeps clinician review in the loop for high-consequence findings.

b. Shared Ecosystem Responsibility (Question 21)

CHI agrees with the paper that the manufacturer is primarily responsible for postmarket monitoring of its device and that other parties add to that responsibility without replacing it. As CHI's framework describes, developers inform deployers and users about the data a tool requires, its intended population and use cases, and its known limitations, and they monitor and retest their models. Deploying organizations set governance policies, train their staff, monitor performance in their own setting, and revalidate after significant changes to their clinical or data infrastructure. Clinicians review outputs before acting on them where they know AI was used, and report errors, biases, and harms to the developer. Physicians and other clinicians also bring the clinical expertise that benchmark adjudication, sample-based review, and postmarket evaluation depend on, and CHI encourages FDA to involve them in developing these methods. A clinician who uses a device as labeled should be able to rely on the manufacturer's evidence of its performance and should not be expected to answer for performance problems the clinician had no reason to know about. Standard-setting organizations and certification bodies develop standards and testing resources, and medical specialty societies issue professional guidance. CHI encourages FDA to point to these roles in its guidance as a description of good practice, and to use its authority over manufacturers, including labeling, to support them. FDA should not, however, impose obligations on clinicians or deploying organizations that would regulate the practice of medicine, which section 1006 of the FD&C Act preserves, and it should continue to leave professional standards and scopes of practice to the licensing bodies and specialty societies that set them.

c. Modifications and Predetermined Change Control Plans (Questions 22 and 23)

CHI supports the paper's proposal to use the premarket benchmark as a baseline against which modified versions of a device are re-benchmarked, and to scale the level of oversight to the impact of a change. CHI recommends that FDA identify categories of changes that a manufacturer may implement under its quality management system with documentation and re-benchmarking, such as changes to user interface text, prompt refinements, and guardrail updates that do not change the intended use and that meet the original acceptance criteria. Changes that expand the intended use, add a new patient population, or move the device to a higher position on either axis would warrant FDA review.

CHI has supported PCCPs as one of the most effective tools FDA has for AI-enabled devices.[12] For GenAI-enabled devices, a PCCP that lists each specific future modification will often be impractical. CHI recommends that FDA allow a PCCP to describe classes of modifications, together with the modification protocol, the benchmark suite and acceptance criteria that each modified version must meet, the subgroups that must be tested, and the circumstances that would require a new submission. A modified device that meets those criteria would be within the authorized PCCP. This approach keeps FDA's oversight focused on the performance of the device while allowing the frequent updates that keep GenAI-enabled devices safe and current.

d. Changes Originating with Third-Party Foundation Models (Question 24)

Device manufacturers that build on third-party foundation models can manage upstream changes with tools they already use, including selecting a specific model version, controlling when a new version is adopted, running regression testing against the device benchmark before deployment, and monitoring for changes in output after deployment. CHI's framework recommends that foundation model developers provide version information that explains how each release differs from the last, so that downstream developers can decide whether revalidation is needed, and it recommends that downstream developers specify appropriate uses so that their products do not amplify issues inherited from underlying models. The terms on which this information is shared are best set in the commercial relationship between the parties. CHI recommends that FDA describe the outcomes it expects the device manufacturer to achieve, such as knowing which model version the device uses and confirming that a new version meets the device's acceptance criteria before it reaches users, and that FDA allow a PCCP to cover the adoption of new upstream model versions on that basis. The manufacturer should also tell deploying organizations when a model change affects how the device performs or should be used. FDA should not expect a device manufacturer to obtain or disclose proprietary information about a third-party model, such as training data, that the manufacturer does not hold. Where less information about an upstream model is available, the manufacturer can address that uncertainty through more rigorous device-level validation and monitoring. Because the device manufacturer is the party that makes the device claim and controls whether a new model version is put into use, it should remain the party accountable to FDA for the

[12] U.S. Food and Drug Administration, Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions, Guidance for Industry and Food and Drug Administration Staff (Aug. 2025), https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence.

device, and FDA should not seek to regulate a foundation model developer that makes no device claims.

VI. Other Topics (Section VII, Questions 25 and 26)

a. Foundation Model Master Files (Question 25)

CHI supports FDA's proposal to keep Foundation Model Master Files (MAFs) voluntary and its statement that holding a MAF does not authorize a model for any intended use. A MAF could reduce duplication where many sponsors build on the same model, since the model developer could share confidential information with FDA one time, without separate arrangements for each sponsor. CHI recommends that FDA design the program so that its contents build on the transparency documentation that foundation model developers already prepare, such as model cards and system cards, and do not require new disclosures of training data, model weights, or other trade secrets. FDA should also make clear that a sponsor's submission will not be viewed less favorably because the foundation model it uses is not the subject of a MAF, so long as the sponsor shows through its own testing and documentation that the device is safe and effective. This alternative, in which the sponsor relies on its own device-level evidence and the documentation it obtains from the model developer, should remain fully available.

b. Agentic AI Systems (Question 26)

CHI agrees with the paper that many of the likely uses of agentic AI in healthcare, including care coordination, clinical documentation, patient outreach, and workflow support, may not be functions on which FDA focuses its device oversight. CHI asks FDA to confirm this in any future guidance, along with its view that using an agentic system to carry out these tasks does not by itself bring them within device oversight, since these administrative and operational uses are some of the most immediate ways to reduce clinician burden. For agentic functions that do meet the device definition, such as a sequence of actions that controls another medical device, FDA should apply the same two-axis framework and competency-based approach that apply to other devices. The agentic competencies in Appendix A, including oversight checkpoints before high-consequence actions and resistance to prompt injection, are a reasonable starting point, and the acceptance criteria for each should scale with the consequences of the actions the device can take and the degree of human review built into its design.

VII. Conclusion

CHI appreciates the opportunity to comment on FDA's discussion paper and urges FDA to consider the recommendations above as it develops its approach to GenAI-enabled medical devices. We look forward to working further with FDA and other stakeholders and would be pleased to discuss any of these recommendations in greater detail.

Sincerely,



Brian Scarpelli
Executive Director

Chapin Gregor
Policy Counsel

Connected Health Initiative
1401 K St NW (Ste 501)
Washington, DC 20005
p: +1 517-507-1446
e: bscarpelli@healthismobile.org