Client Alert

JANUARY 10, 2025

For more information, contact:

Lisa M. Dwyer
+1 202 626 2393
ldwyer@kslaw.com

Adam Solander
+1 202 626 5542
asolander@kslaw.com

Robert Hudock
+1 202 626 5521
rhudock@kslaw.com

Jessica Ringel
+1 202 626 9259
jringel@kslaw.com

Kyle Sampson
+1 202 626 9226
ksampson@kslaw.com

Jeffrey Shapiro
+1 202 626 2927
jshapiro@kslaw.com

Eric Henry
+1 202 661 7823
ehenry@kslaw.com

McKenzie Cato
+1 202 626 9270
mcato@kslaw.com

King & Spalding

Washington, D.C.
1700 Pennsylvania Avenue, NW
Suite 900
Washington, D.C. 20006
T. +1 202 737 0500

FDA Releases Draft Guidance on Submission Recommendations for AI-Enabled Device Software Functions

On January 7, 2025, FDA published a draft guidance titled "Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations." The draft guidance was long-anticipated; it was on FDA's "A list" for FDA to publish in fiscal year 2024 (i.e., prior to September 30, 2024) and then transferred to the fiscal year 2025 "A list" when that deadline was missed.

The draft guidance provides recommendations on the content of marketing submissions (e.g., 510(k) premarket notifications, premarket approval applications) for devices that include artificial intelligence (AI)-enabled device software functions. The draft guidance also provides recommendations on the design and development of AI-enabled devices throughout the total product life cycle (TPLC).

Throughout the draft guidance, FDA points to its Recognized Consensus Standards Database and encourages manufacturers to utilize these standards in the software development process. Although not specifically referenced in the draft guidance, we note that recognized consensus standards, such as IEC 62304 (*Medical device software – Software life cycle processes*), IEC 82304-1 (*Health software – Part 1: General requirements for product safety*), and IEC 81001-5-1 (*Health software and health IT systems safety, effectiveness and security – Part 5-1: Security – Activities in the product life cycle*), are particularly useful for managing the development process and are referenced by FDA in other guidances. In addition, the draft guidance promotes the use of the Q-Submission Program as a way for manufacturers to obtain FDA feedback on various aspects of the AI device software function elements of their submission.

SCOPE AND GENERAL PRINCIPLES

The draft guidance applies only to software functions that (i) meet the statutory definition of device in section 201(h) of the Federal Food, Drug, and Cosmetic Act (FD&C Act); and (ii) are AI-enabled device software functions (AI-DSFs). FDA defines an AI-DSF as a "device software function that implements one or more 'AI models.'"[i] A "model" is defined as a "mathematical construct that generates a reference or prediction based on new input data."[ii]

The draft guidance provides recommendations specific to the AI-DSF components of a device or combination product, but clarifies that other FDA guidance on marketing submissions, including the "Content of Premarket Submissions for Device Software Functions" (June 2023), are still applicable to AI-enabled devices. Throughout the draft guidance, FDA provides links to "additional resources" with other relevant guidance documents.

FDA further acknowledges that the draft guidance is most relevant to machine learning (specifically deep learning and neural networks). To date, FDA has not authorized any AI-DSF with generative AI (GAI), but this language in the draft guidance may open the door to future uses of GAI technology, which utilizes neural networks.

Although the draft guidance addresses post-market concerns, such as performance monitoring and data drift, FDA encourages manufacturers to consider whether a Predetermined Change Control Plan (PCCP) might also be appropriate in the submission and references the Agency's final guidance, issued last month, in December 2024, on *Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions* (see our Client Alert on the PCCP guidance).

Because AI-specific terminology is, in some respects, different than both Quality System Regulation terminology and traditional Software Development Life Cycle (SDLC) concepts, FDA clarifies in the draft guidance how it uses certain terms and points to its recently released *FDA Digital Health and Artificial Intelligence Glossary*[iii] for further clarification. Also, instead of structuring the draft guidance in alignment with the SDLC phases defined in recognized consensus standards, such as IEC 62304, the draft guidance is organized according to the sections of an eSTAR template or typical marketing submission. In each subsection (e.g., Device Description, Risk Management), FDA indicates where in the submission the recommended information should be provided. Appendix A to the draft guidance also provides a table of recommended documentation and the associated section in the marketing submission.

FDA states that the draft guidance is intended to adopt a TPLC approach to the management of AI-enabled devices by providing recommendations for their design, development, deployment, and maintenance. FDA further encourages sponsors to consider whether any of the recommended documentation for marketing submissions should exist in the sponsor's Quality System documentation, such as design controls (21 C.F.R. § 820.30(g)), design changes (21 C.F.R. § 820.30(i)), control of nonconforming product (21 C.F.R. § 820.90(a)), and corrective and preventive actions (21 C.F.R. § 820.100(a)).

DEVICE DESCRIPTION

The draft guidance provides a list of information about AI-DSFs that should be included in a marketing submission as part of the device description. This list includes: a statement that AI is used in the device, device inputs and outputs, how AI is used to achieve the device's intended use, intended users and requisite expertise, intended use environment, intended workflow, installation and maintenance procedures, and any calibration or configuration procedures. Additionally, if a device has multiple connected applications with separate interfaces, FDA recommends that the device description address all applications in the device. The draft guidance allows for the use of graphics, diagrams, illustrations, screen captured images, or video demonstrations in support of the device description. As is implied by the elements of the device description, AI is never the device in totality but a component (or "software

item" in IEC 62304 terms) of a medical device. Device descriptions therefore must incorporate an overview of the entire device and not be limited to a description of the AI components only.

USER INTERFACE AND LABELING

The draft guidance describes the user interface as "all points of interaction between the user and the device, including all elements of the device with which the user interacts" and "all sources of information transmitted by the device (including packaging and labeling), training, and all physical controls and display elements."[iv] FDA recommends that a user interface be designed such that important information is provided throughout the course of use of the device. Appendix D to the draft guidance includes details on usability evaluation considerations.

The draft guidance recommends including a description of the user interface in the software description portion of the software documentation section, which may include a graphical representation, written description, overview of the operational sequence, examples of the output format (e.g., example reports), and a demonstration of the device (e.g., through a recorded video).

The draft guidance explains that a device's user interface includes labeling, but labeling should be provided in the separate labeling section of the marketing submission. Labeling should address the following information "in a format and at a reading level that is appropriate for the intended user"[v]: statement that AI is used in the device, model inputs, model outputs, automation, model architecture, model development data, performance data, device performance metrics, performance monitoring, all known limitations, installation and use, customization, metrics or visualization to provide context to the model output, and patient/caregiver information. FDA recommends the use of a "model card" in the device labeling to communicate information about the AI-enabled device and provides an example model card in Appendix E (see below for further discussion of model cards).

RISK ASSESSMENT

As with other devices containing software, the draft guidance states that AI-enabled devices should include a risk management file that takes into account FDA's guidance on Content of Premarket Submissions for Device Software Functions and ANSI/AAMI/ISO 14971 (*Medical devices – Applications of risk management to medical devices*). FDA has also recognized AAMI CR 34971 (*Guidance on the Application of ISO 14971 to Artificial Intelligence and Machine Learning*).[vi] Risk management should consider hazards across the TPLC and should link risk controls to user interface requirements, with a particular emphasis on understanding information necessary to use or interpret the device, as appropriate.

For AI-enabled devices, FDA states that risks related to understanding information necessary to use or interpret the device are particularly important. The risk management file should include explanation of any risk controls, including elements of the user interface (e.g., labeling) that address identified risks.

DATA MANAGEMENT

The draft guidance explains that, for an AI-enabled device, the model (i.e., the algorithm and the data used to train it) is part of the device's mechanism of action. For this reason, FDA recommends a clear description of the data management practices and characterization used in the development and validation of the AI-enabled device. FDA states that reviewers will evaluate the quality, diversity, and quantity of data used to test an AI-DSF to evaluate the safety and effectiveness of the AI-enabled device. Data management is also important in identifying and mitigating bias through inclusion of representative data in training and validation datasets. Manufacturers should provide evidence of sufficient segregation of training and validation datasets to address issues of bias and overfitting.

In a submission, FDA recommends that the software description section include information on data collection (e.g., how data was collected, size of the dataset, use of synthetic data), data cleaning/processing, the reference standard (i.e., source of truth) used in device training or validation, data annotation (including qualifications and training for those responsible for annotating data), data storage, management and independence of data (including segregation of training, tuning, and validation datasets), and representativeness of the data for the intended use population (including an assessment of any non-U.S. data vs. the U.S. population). The draft guidance provides specific content recommendations in each of these data management categories.

MODEL DESCRIPTION AND DEVELOPMENT

The software description section should include a distinct model description subsection that addresses the technical aspects of the model and how it was developed, including its biases and limitations. Where multiple models are employed as part of an AI-enabled device, FDA recommends that the software description include a diagram of how model outputs combine to create the device outputs.

The software description should include an explanation of each model used as part of the AI-enabled device, including a description of model inputs and outputs, model architecture, features, feature selection process and any loss function(s) used for model design and optimization, and model parameters. Where the AI-enabled device has customizable features, a sponsor should include a description of technical elements of the model that allow for and control customization. The software description should also include a description of quality control criteria or algorithms and any methods applied to input and/or output data (e.g., pre-processing).

Additionally, the software description should include an explanation of how the model was trained (e.g., optimization methods, training paradigms), metrics and results obtained for any tuning evaluation, pre-trained models, ensemble methods, how any thresholds were determined, and any calibration of the model output.

VALIDATION

In the introduction to the draft guidance, FDA clarifies that the term "validation" is used differently by the AI community than in the context of a device marketing submission. The AI community often refers to "validation" to mean data curation or model tuning that can be combined with model training to optimize model selection. In contrast, FDA refers to "validation" as defined in 21 C.F.R. § 820.3(z), which is "confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use can be consistently fulfilled."

For an AI-enabled device, the draft guidance explains, validation includes ensuring that the device will perform its intended use safely and effectively (including whether users consistently and correctly receive, understand, interpret, and apply information related to the AI-enabled device), as well as establishing that relevant performance specifications of the device can be consistently met. FDA states that performance validation and human factors validation helps provide information on how a device will perform under real-world circumstances. Performance validation is confirmation that device specifications conform to user needs and intended uses, and that performance requirements implemented can be consistently fulfilled (discussed in Appendix C of the draft guidance). Human factors validation is confirmation that all intended users can achieve specific goals while using the device and will be able to consistently interact with the device safely and effectively (discussed in Appendix D of the draft guidance), in accordance with FDA's human factors guidance document.[vii]

The draft guidance explains that validation methods will differ depending on the intended use of the device and could come from non-clinical bench or analytical studies, pre-clinical animal studies, clinical performance studies, clinical outcome studies, or some combination thereof. The "Software testing as part of Verification and Validation" section

within the software documentation section of the submission should include study protocols with study design details and statistical analysis plans. The study results should include adequate subgroup analyses for relevant demographics and an evaluation of device repeatability and reproducibility. The draft guidance again emphasizes the importance of independent datasets for performance validation.

DEVICE PERFORMANCE MONITORING

In line with FDA's TPLC approach for AI-DSFs, a marketing submission must address the performance of AI-enabled devices in real-world environments and the risk that they will change or degrade over time (i.e., "drift"). We believe the emphasis on postmarket performance monitoring is partly in recognition that postmarket performance of AI models used in medical devices vs. the performance specifications against which these medical devices were originally approved or cleared is currently opaque. We see this issue as especially problematic when future AI models used in medical devices become more adaptive (i.e., make unattended changes to themselves in their operating environment based on real-world data).

FDA recommends that manufacturers proactively monitor, identify, and address device performance changes. The draft guidance acknowledges that FDA does not typically assess quality system compliance as part of its review of marketing submissions. FDA may, however, review details from a sponsor's quality system in a marketing submission to ensure that the device will have adequate ongoing performance. Further, the draft guidance clarifies that performance monitoring plans (although recommended for all submissions) are generally not required for 510(k) premarket notifications but may be required for De Novo requests (as a special control) and premarket approval (PMA) applications.

FDA recommends that the risk management file included in the submission include performance monitoring plans. Such plans should include a description of data collection and analysis methods for assessing changes and potential causes for changes in model performance; robust software lifecycle processes with mechanisms for monitoring the device when in use; and a plan for deploying updates, mitigations, and corrective actions that address performance changes. Some actions required to address performance may not require a new submission based on FDA's guidances on *Deciding When to Submit a 510(k) for a Change to an Existing Device* (Oct. 2017), and *Deciding When to Submit a 510(k) for a Software Change to an Existing Device* (Oct. 2017), or if the action was taken in accordance with an FDA-authorized PCCP.

CYBERSECURITY

In recent years, and in particular with the release of FDA's guidance document on premarket cybersecurity (*see* FDA, Guidance for Industry & FDA Staff, *Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions* (Sep. 2023)), FDA has been focused on cybersecurity in evaluating marketing submissions for medical devices containing software.

Sponsors of AI-enabled devices that meet the definition of a "cyber device" in section 524B(c) of the FD&C Act[viii] must include in their marketing submissions details on AI risks that can be impacted by cybersecurity threats. The draft guidance provides a list of examples of AI risks (i.e., vulnerabilities) that can be impacted by cybersecurity threats, including data poisoning, model inversion/stealing, model evasion, data leakage, overfitting, model bias, and performance drift.

In a marketing submission, FDA recommends that sponsors include in the Cybersecurity/Interoperability section information on: additional elements in the cybersecurity risk management report and other submission sections related to AI cybersecurity; how the cybersecurity testing is appropriate to address risks associated with the model; a "Security Use Case View" that covers AI-enabled considerations for the device; and a description of controls

implemented to address data vulnerability and leakage. FDA also recommends that sponsors refer to the control recommendations in Appendix 1 to the above-referenced 2023 premarket cybersecurity guidance ("*2023 Premarket Cybersecurity Guidance*"). Based on our experience with FDA deficiency letters for cyber devices, three recommendations in the *2023 Premarket Cybersecurity Guidance* (which are not mentioned in the draft guidance) are of particular importance and should be included in the Cybersecurity/Interoperability section of a marketing submission:

- A threat model (with vulnerability assessment),

- A Software Bill of Materials (SBOM), and

- The full set of security architecture views (global system view, multi-patient harm view, updateability/patchability view, and security use case view).

As is referenced in the *2023 Premarket Cybersecurity Guidance*, we also recommend that manufacturers incorporate the recognized consensus standard IEC 81001-5-1 (*Health software and health IT systems safety, effectiveness and security – Part 5-1: Security – Activities in the product life cycle*) into their software development process and quality system.

PUBLIC SUBMISSION SUMMARY

Device premarket submissions typically include a public submission summary (e.g., the 510(k) summary for a 510(k) premarket notification). The draft guidance emphasizes the importance of transparency with respect to AI-DSFs in the public summary. Specifically, the draft guidance states that the public summary should include a statement that AI is used in the device, an explanation of how AI is used as part of the device's intended use, a description of the class of model (e.g., convolutional neural network, recurrent neural network), a description of the development and validation datasets, a description of the statistical confidence level of predictions, and a description of how the model will be updated and maintained over time. Appendix B of the draft guidance discusses transparency in the design of the medical device in more detail and includes considerations related to the user interface of the AI-enabled medical device.

MODEL CARDS

One tool contributing to transparency is known as a "model card." Model cards, as defined in FDA's *Digital Health and Artificial Intelligence Glossary*, are structured reports of relevant technical characteristics of an AI model. Appendix E provides an example Model Card Template, and Appendix F provides a sample 510(k) summary with a model card. The template provided in the draft guidance includes device information, the regulatory status of the medical device, a description of the AI model, performance and limitations, risk management, and development, with detailed recommendations for each of these categories.

Although no model cards specific to medical devices have been published prior to this draft guidance, several industry groups have published model card templates for AI models used in healthcare more broadly. For example, the Health AI Partnership (HAIP) at the Duke University Institute for Health Innovation published an article in the Nature Digital Medicine journal titled "Presenting machine learning model information to clinical end users with model facts labels,"[ix] in which HAIP introduced the concept of a "nutrition label" that includes a summary, and information regarding mechanism, validation and performance, uses and directions, warnings, and so forth. The Coalition for Health AI (CHAI) published an Applied Model Card Template that includes sections for release information, summary, uses and directions, warnings, trust ingredients, key metrics (i.e., usefulness, usability, and efficacy; fairness and equity; safety and reliability), and resources.

There has been debate over the effectiveness of model cards across the wide variety of AI use cases, so we recommend that medical device manufacturers tailor their use of model cards, with appropriate rationale, to the particular use case, operating environment, and technical specifics of their AI model instead of using the model card in the draft guidance (or other published model card templates) as a mandatory checklist of elements.

<p align="center">*   *   *   *   *</p>

Although the draft guidance is still exactly that—a draft—it provides important insight into how FDA thinks about AI-DSFs and can serve as a valuable guide for the content of premarket submissions even before the guidance is finalized. FDA is accepting comments on the draft guidance until April 7, 2025, to be submitted to docket number FDA-2024-D-4488. Additionally, FDA has scheduled a public webinar on the draft guidance on February 18, 2025. If you have any questions about the draft guidance or its application, we would be happy to assist you.

---

[i] FDA, Draft Guidance for Industry & FDA Staff, *Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations*, § II (Jan. 7, 2025) [hereinafter "*Draft Guidance*"].

[ii] *Id.*

[iii] *See* FDA, *FDA Digital Health and Artificial Intelligence Glossary—Educational Resource* (updated Sep. 26, 2024).

[iv] *Draft Guidance* § VI.

[v] *Id.* § VI(B).

[vi] ISO's TC210, which maintains ISO 14971, is currently drafting a version of AAMI CR 34971 as an international guidance, which will have the identifier ISO TIR 24971-1 once published.

[vii] *See* FDA, Guidance for Industry & FDA Staff, *Applying Human Factors and Usability Engineering to Medical Devices* (Feb. 2016).

[viii] A "cyber device" is defined as a device that "(1) includes software validated, installed, or authorized by the sponsor as a device or in a device; (2) has the ability to connect to the internet; and (3) contains any such technological characteristics validated, installed, or authorized by the sponsor that could be vulnerable to cybersecurity threats." FD&C Act § 524B(c).

[ix] *See* Sendak, M.P., Gao, M., Brajer, N. *et al*. Presenting machine learning model information to clinical end users with model facts labels. *npj Digit. Med.* 3, 41 (2020). https://doi.org/10.1038/s41746-020-0253-3.