ACRO Innovate Advocate Collaborate

April 4, 2025

Tala Fakhouri, Center for Drug Evaluation and Research
James Myers, Center for Biologics Evaluation and Research
Sonja Fulmer, Center for Devices and Radiological Health
Food and Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993-0002

RE: ACRO Comment on *Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products* [FDA-2024-D-4689-0001]

Dear Dr. Fakhouri, Mr. Myers, and Dr. Fulmer,

The Association of Clinical Research Organizations (ACRO) represents the world's leading clinical research and clinical technology organizations. Our member companies provide a wide range of specialized services across the entire spectrum of development for new drugs, biologics and medical devices, from pre-clinical, proof of concept and first-in-human studies through post-approval, pharmacovigilance and health data research. ACRO member companies manage or otherwise support a majority of all biopharmaceutical sponsored clinical investigations worldwide and advance clinical outsourcing to improve the quality, efficiency and safety of biomedical research.

ACRO thanks the Agency for the valuable recommendations in the draft guidance and the opportunity to provide feedback. We offer general reactions followed by section-specific comments.

For years, the FDA has been strategically supporting the integration of AI/ML across medical product regulation. While the Center for Devices and Radiological Health (CDRH) has led with specific action plans and guidance for medical devices, the FDA is now establishing a framework for drugs and biologics. This draft guidance represents a key milestone, building upon internal collaborative efforts and working groups across all medical product centers. In 2023, the Center for Drug Evaluation and Research (CDER) initiated a dialogue with stakeholders through discussion papers on AI/ML in drug development and manufacturing. These papers underscored the need for a clear regulatory path, outlining potential applications from drug discovery to post-market surveillance. To further accelerate progress, CDER is actively driving innovation through programs like the Quantitative Medicine Center of Excellence (QMCoE), the Center for Clinical Trial Innovation (C3TI), and the Emerging Drug Safety Technology Program (EDSTP). These initiatives, coupled with the AI Council and cross-center guidance, are building momentum for the responsible and effective implementation of AI/ML throughout the entire medical product lifecycle. This guidance is a valuable next step in the FDA's ongoing efforts to ensure the safe and effective use of AI/ML in drug development. We support the agency's approach of using a risk-based framework to guide AI model evaluation.

Regulators around the globe are grappling with the complex challenges of designing regulatory regimes for AI that foster patient safety, data integrity, and high evidentiary standards – without stifling innovation. As with previous draft guidances on timely issues (e.g., decentralized trials, integrating randomized controlled trials into routine care, and risk-based quality management), the Agency's scientific expertise and forward leadership have enabled it to blaze a globally influential path by promptly issuing much-needed draft guidance that embraces a risk-based, flexible, and pragmatic approach to AI oversight.

Growing experience with AI/ML in regulatory science will likely result in the evolution of approaches and strategies. As regulatory concepts and review processes mature, we ask that core principles, including definitions and terminology, remain aligned across FDA's medical product centers. Alignment of this draft guidance with CDRH's AI/ML framework for medical devices will help ensure consistency across regulatory pathways.

I. General Comments

There are numerous instances throughout the draft guidance where it would be valuable for the final guidance to provide further elaboration, to clarify what "good" looks like. Stakeholders would benefit from insight into AI models that have successfully met regulatory expectations. To clarify expectations and promote best practices, we ask the Agency to consider sharing in the final guidance anonymized examples (case studies) of successful ("Dos") and unsuccessful ("Don'ts") AI models it has had the opportunity to evaluate in regulatory submissions. We have identified three sections of the draft guidance where these anonymized examples/case studies from actual regulatory submissions would be particularly helpful.

- *Model Risk as a Function of Model Influence and Decision Consequence (Section IV.A, 208-253)*
  Step 3 explains how to understand model risk as a function of model influence and decision consequence. We thank FDA for the examples provided in the draft guidance. It would be helpful to see additional anonymized case examples of how to think about model risk from previous submissions FDA has received.

- *Human-in-the-Loop Models (Section IV.A, 446-449)*
  The draft guidance emphasizes that AI models may require human oversight but does not provide examples of how this oversight might be structured. It would be particularly illuminating for industry to see illustrations of how FDA's AI risk framework has been applied and "human-in-the-loop" models/approaches have been successfully and effectively implemented in regulatory submissions.

- *Risk-Based Approach for Lifecycle Monitoring (Section IV.B, 538-567)*
  The draft guidance recommends a risk-based approach to lifecycle monitoring. We ask the Agency to consider including examples of successfully and effectively implemented lifecycle monitoring programs in the final guidance, by sharing anonymized case studies of lifecycle monitoring approaches and plans it has had the opportunity to evaluate in regulatory submissions.

Additionally, we recommend that the Agency explicitly states in the final guidance that "transparency" does necessitate disclosure of input-output relationships, data, and methodology, but does not require either "model interpretability" or "model explainability." Moreover, we ask the Agency to consider including specific language acknowledging the inherent tensions in the multiple objectives of any AI model along the lines of: "AI models contain inherent tensions among competing objectives. Sponsors should use good judgment and provide a rationale for adjudicating the trade-offs and balance among model explainability, transparency requirements, model performance, and model complexity."

II. Section-Specific Comments

Section III – "Background"

Lines 88-91:
The draft guidance notes that data used to develop AI models should be "fit for use" – particularly with respect to the need for data to be representative. ACRO asks the Agency to consider including further discussion and elaboration in the final guidance of how to achieve fit for use, representative data in the not uncommon situation where absence of bias and representativeness is, in practice, difficult to demonstrate.

Lines 92-95:
The draft guidance mentions that "model transparency may be necessary" but does not define what constitutes transparency in this context. ACRO recommends that the Agency clarify the meaning of transparency in the final guidance so that it aligns with the principle of Occam's Razor. This principle of parsimony means that the simplest explanation is usually the best, favoring simpler models over more complex ones.

Section IV.A – "A Risk-Based Credibility Assessment Framework"

Omission in Steps 1-3 (Lines 150-257):
We recommend that the final guidance includes additional, clarifying discussion of multi-model risk considerations. AI models may operate in isolation, but they may also be used in combination – for example, in the case of two (or more) linked AI models where one model's output is used as input for another. The draft guidance does not specify whether risk should be assessed for individual models, across linked models, or both. We ask that the final guidance emphasize sponsor flexibility in using the AI risk aggregation methodologies for multi-agent AI systems that they see as most appropriate for that use case and provide a rationale for why they decided to assess risk for individual models, across linked models, or both.

Omission in Step 3 (Lines 208-257):
Step 3 discusses how to understand model risk as a function of two factors: model influence and decision consequence. We would like to ask the Agency to consider including a reference to a framework introduced by Kuemmel[1] which is a relevant source for the framework on which this draft guidance rests.

Section IV.B – "Special Consideration: Life Cycle Maintenance of the Credibility of AI Model Outputs in Certain Contexts of Use"

Lines 528-536:
In the discussion of self-evolving models, the draft guidance states that "due to the evolving nature of AI models," certain risk-mitigation measures are necessary. This phrasing implies that all AI models are dynamic and evolve, which is not the case. We recommend that the final guidance clarifies that self-evolving models are a subset of AI models and also clarifies that sponsors may use their best judgment and have flexibility to use differing risk-mitigation measures in static versus dynamic (evolving) AI systems.

[1] Kuemmel C et al; *Consideration of a credibility assessment framework in model-informed drug development: potential application to physiologically-based pharmacokinetic modelling and simulation*. CPT Pharmacometrics Syst. Phamacol. (2020), 21-28; doi: 10.1002/psp4.12479

ACRO is grateful for the opportunity to provide input on this draft guidance. To foster a more robust understanding of AI in regulatory contexts, we believe the final guidance would benefit from concrete examples, such as anonymized case studies, that provide guidance on critical regulatory concepts and offer stakeholders tangible illustrations of both exemplary and potentially suboptimal development and implementation of credible AI models. These case studies could also provide more detailed guidance around "transparency" and “explainability” requirements. It will be important for stakeholders to understand the application of these concepts -- particularly when there are multi-modal risk considerations; when data representativeness is paramount; and when humans are in the loop to verify model output. Addressing these points will provide greater clarity and strengthen the framework for the responsible integration and use of AI in medical product regulation. Please contact ACRO (knoonan@acrohealth.org) if we can answer any questions or provide additional information.

Respectfully submitted,

Karen Noonan

Karen Noonan
Senior Vice President, Global Regulatory Policy