To: Food and Drug Administration, HHS

Re: Docket: FDA-2024-D-4689

The International Society for CNS Clinical Trials and Methodology (ISCTM) welcomes this opportunity to respond to the FDA request for comment: *Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products*

The ISCTM offers these comments for consideration based on our experience and expertise in human CNS research. The ISCTM is an independent organization focused on advancing the development of improved treatments for CNS disorders. No member of this Working Group, comprised of scientists, clinicians, trialists, statisticians, and drug developers from both industry and academia, received compensation for comments provided. Comments represent personal opinions and not that of the institution, agency, or company affiliation of group members.

The ISCTM formed a group, led by Marc Aafjes and Erica Smith, to review, and provide comments on behalf of the Society. Authors (in alphabetical order):

Marc Aafjes, MSM, MSc, *Deliberate AI (co-chair)*
Deepa Arora, MD, *Clinexel Life Sciences*
Michael Avissar, MD, PhD, *Alto Neuroscience*
Adam Butler, *Independent*
Nicholas DeMartinis, MD, *Independent*
Franco DiCesare, MD, *Loeben Research Aurora*
Martin Dunbar, PhD, *Alkermes*
George Garibaldi, MD, *Garibaldi Consulting*
Ni Khin, MD, *Neurocrine Biosciences*
Hardik Kothare, PhD, *Modality.AI*
Atul R. Mahableshwarkar, MD, *Cybin*
Raeanne Moore, PhD, *University of California San Diego*
Alvaro Pereira, PhD, *Cognivia*
Maria Louise Petersen, MD, *H. Lundbeck A/S*
Jan Sedway, PhD, *WCG Clinical Endpoint Solutions*
Beatrice Setnik, PhD, *Altasciences*
Leif Simmatis, PhD, *University of Toronto*
Erica Smith, PhD, *WCG Clinical Endpoint Solutions (co-chair)*
Josh Spiegel, CFA, *NetraMark*

Qing Wang, PhD, *Neumarker*
Lynn Webster, MD, *Dr. Vince Clinical Research*
Scott Woods, MD, *Yale University*

COMMENTS ON CONSIDERATIONS FOR THE USE OF ARTIFICIAL INTELLIGENCE TO SUPPORT REGULATORY DECISION-MAKING FOR DRUG AND BIOLOGICAL PRODUCTS, GUIDANCE FOR INDUSTRY:

General Comments

The ISCTM appreciates the FDA's timely release of this draft guidance, addressing critical questions regarding integrating artificial intelligence (AI) into regulatory decision-making. We recognize the importance of a structured, risk-based approach to evaluating AI models' credibility and support the FDA's initiative to define clear guidelines for their use in clinical development and manufacturing.

The ISCTM looks forward to the evolution of guidance as AI gets more used in regulatory decision making and is prepared to, and would readily participate in, further public debate to achieve further integration of AI in drug development.

Throughout this document, additional suggestions or modifications for text are inserted in italicized text after regular font draft guidance text.

Specific Comments

II. Scope

In addition to exclusions already mentioned in line 47-51, ISCTM proposes to clarify that this guidance also does not cover recruitment-related activities:

"(2) when used for operational efficiencies (e.g., internal workflows, resource allocation, drafting/writing a regulatory submission) that do not impact patient safety, drug quality, or the reliability of results from a nonclinical or clinical study*—this includes the use of computer-based algorithms and automation that are used to identify, highlight or screen patients.*"

III. Background

Line 17: The ISCTM suggests clarifying the definition of 'credibility', which is not often used, to reflect the government's discourse around the concept "Trustworthy AI." A clear definition of credibility should explicitly reference aspects such as explainability, interpretability, and validation.

> "For the purposes of this guidance, credibility *incorporates the principles of Trustworthy AI—as defined in the Executive Order 13960 on Promoting Trustworthy AI (2020), HHS Trustworthy AI (TAI) Playbook (2021), NIST AI Risk Management Framework (RMF 1.0, 2023), FDA's AI/ML-Based Software as a Medical Device Action Plan (2023—and is established through the systematic collection of evidence demonstrating the reliability, safety, fairness, transparency, and accountability of an AI model's performance within its intended context of use (COU).*"

IV. A Risk-Based Credibility Assessment Framework

Step 3: Assess the AI Model Risk. Lines 208-257: The ISCTM finds Figure 1 useful but recommends enhancing it with additional examples or details to improve clarity and practical applicability for various contexts. A more detailed matrix indicating clear delineations between low, medium, and high-risk scenarios would significantly enhance usability.

Step 4: Develop a Plan to Establish AI Model Credibility. Starting at line 293, the agency suggests detailed descriptions of AI models, which the ISCTM supports. However, we recommend explicitly mentioning the transparency and interpretability of AI models, particularly for high-risk applications. We suggest either clarifying this here, or adding the following text to line 314 Model Features:

> *"For high-risk models, sponsors should provide detailed information on model interpretability methods (e.g., SHAP, LIME, etc.) to facilitate understanding of decision processes by regulators and other stakeholders."*

Regarding line 310, we propose strengthening this to:

> "Model inputs and outputs, *including preprocessing and feature extraction steps if these are integral to model performance.*"

Note 27 (Line 314): The ISCTM suggests expanding the description of "features" to explicitly acknowledge a broader range of data types:

> "Features can *encompass a wide range of data types, including clinical measurements, demographics, imaging data, physiological signals, and other behavioral or biometric attributes.*"

Line 367, we propose strengthening this to:

> "Describe whether the development data *is stored and processed in a centralized manner or distributed across multiple sites (e.g., using federated learning), and explain how this approach ensures data privacy, security, regulatory compliance, and model reproducibility."*

Regarding lines 381 & 451, we suggest broadening these terms beyond classification models as other types of models are likely to be encountered more in regulatory decision-making for drug and biological products than what be more typically seen in diagnostic medical devices (e.g. regression models or unsupervised clustering)

> *"Performance metrics used to evaluate the model should align with its specific learning task. For classification models, relevant metrics include the area under the receiver operating characteristic (ROC) curve, recall (sensitivity), specificity, positive/negative predictive values (PPV/NPV), true/false positive and true/false negative counts (e.g., in a confusion matrix), positive/negative diagnostic likelihood ratios (PLR/NLR), precision, and F1 scores. For regression models, appropriate metrics include root mean squared error (RMSE), mean absolute error (MAE), Pearson correlation (R^2), and concordance correlation coefficients. For unsupervised learning models, evaluation should consider metrics such as silhouette score, Davies-Bouldin index, or anomaly detection performance where applicable. All performance estimates should be provided with confidence intervals or uncertainty quantification where feasible."*

Line 466. We propose to strengthen this with the kind of analysis the agency would expect.

> "Describe limitations of the modeling approach, including potential biases *using, for example, subgroup analysis*."

IV.C Early Engagement. Lines 570-588 discuss early engagement options. The ISCTM strongly supports the FDA's encouragement of early engagement and proposes extending examples of successful regulatory interactions related to AI model usage. Providing examples would help guide sponsors more effectively.

Conclusion

The ISCTM is grateful for the opportunity to comment and looks forward to ongoing dialogue with the FDA to advance robust and reliable AI methodologies that enhance clinical and regulatory decision-making.