3 June 2019

To: Food and Drug Administration, HHS

Re: Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)

The International Society for CNS Clinical Trials and Methodology (ISCTM) welcomes this opportunity to respond to the FDA request for feedback: Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD).

The International Society for CNS Clinical Trials and Methodology (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 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 Adam Butler and Larry Alphs, to review and provide comments on behalf of the Society. Authors (in alphabetical order):

Lawrence Adler, MD, *Clinical Insights, Inc*
Larry Alphs, MD, PhD, *Newron* (Co-chair)
Ariana E. Anderson, PhD, *University of California, Los Angeles*
Adam Butler (Co-chair)
Daniel DeBonis, *CRF Bracket*
Miguel Garcia, MS, *Boehringer Ingelheim*
Hugo Geerts, PhD, BachMed, PharmaMBA, *In Silico Biosciences*
Joseph Geraci, PhD, Queen's University, *NetraMark Corp*
Amir Kalali, MD
Timothy Mariano, MD, PhD, MSc, *Harvard Medical School, Sage Therapeutics*
Kemi Olugemo, MD, *Parexel*
Jill Rasmussen, MD, *psi-napse*
Vikas Mohan Sharma, MBBS, MD, *Boehringer Ingelheim*
Stephanie Sommer, Dr. rer. nat., *Boehringer Ingelheim*
Glen Wunderlich, PhD, *Boehringer Ingelheim*
Silvia Zaragoza-Domingo, PhD, *Neuropsynchro*

COMMENTS ON THE PROPOSED REGULATORY FRAMEWORK FOR MODIFICATIONS TO ARTIFICIAL INTELLIGENCE/MACHINE LEARNING (AI/ML)-BASED SOFTWARE AS A MEDICAL DEVICE (SAMD):

General Comments

ISCTM is encouraged by this proactive and detailed document regarding some of the issues related to how AI and ML will be utilized in clinical trials. Clarification on how AI and ML tools will interact with existing SaMD guidance are helpful and should serve to expand their use in clinical trials. Many drug developers, and SaMD developers are cautious about how to use machine learning and algorithms in research programs, especially regarding clinical decision making, and having clear expectations about regulatory pathways will serve to support their appropriate and meaningful use.

II. Background

The risk categorization matrix on page 5 of the document is useful in terms of identifying scenarios in which AI and ML may be utilized. This matrix will be most useful when these risk categories are combined with additional criteria for evidence generation or validation. For example, will a lower threshold for validation of a machine learning algorithm be applied if it is for an SaMD to be used only for clinical site selection for a clinical trial of a non-serious medical condition?

III. Types of AI/ML-based SaMD Modifications

Question 1 "Do these categories of AI/ML-Software as a Medical Device (SaMD) modifications align with the modifications that would typically be encountered in software development that could require premarket submission?"

- In general, these categories seem consistent with existing software development and validation paradigms. Allowing for prespecified input that draws on existing biological knowledge and domain expertise are important. As the goal of AI/ML is to improve the performance of an algorithm, a performance improvement in a certain defined frame (parameters) should be possible without pre-market submission.
- There are some risks associated with changes in ownership or design of predicate data sets that are used to train AI/ML algorithms. A modification plan should take into account how these changes are accounted for, and to ensure that the statistical validity of any training done on one data set is replicable if those data sets are changed or replaced.

Question 2 "What additional categories, if any, of AI/ML-SaMD modifications should be considered in this proposed approach?"

- Any change in clinical guidelines or patient classification; for example, the introduction of new staging criteria for Alzheimer's Disease (as discussed in the FDA Draft Guidance "Early Alzheimer's Disease: Developing Drugs for Treatment") may require a significant modification of an AI/ML tool designed to enrich patients into a clinical trial.
- Strictly technical software changes, such as new versions of a statistical software package or an update to a standardized library such as TensorFlow, should probably be explicitly excluded from modification requirements.

Question 3 "Would the proposed framework for addressing modifications and modification types assist the development AI/ML software?"

- Establishing clear boundaries, with examples, on what would require modification will be a hugely important regulatory task and will be important for developers in understanding what both their development and post-marketing plans need to look like. To this end, being clear regarding the population to which the software applies will be important. In addition, many fields of medicine are developing rapidly, and the definitions of populations are

changing or are being refined. Providing guidance on how these anticipated changes will be managed is important.

IV. A Total Product Lifecycle Regulatory Approach for AI/ML-Based SaMD

Section 1 "Quality Systems and Good Machine Learning Practices (GMLP)"

Question 1 "What additional considerations exist for Good Machine Learning Practice?"

- GMLP should include a definition of "obsolescence" to allow for a pre-defined criteria that would determine an algorithm was no longer generating current and clinically relevant outcomes.
- Complimentary to this would be specification of how generation and collection of new data relevant to the baseline training dataset can be accommodated. It is likely that the most powerful applications of AI/ML will be to assist clinical, human, decision making based on very large data sets. GMLP should define how continuous data collection and analysis can be addressed in algorithms and whether or not reapproval is necessary.
- GMLP should define what role human decision making may have, in both the design of the algorithms and analysis during the design phase, but in the ongoing evaluation and updates to the algorithms.

Question 2 "How can the FDA support development of Good Machine Learning Practice?"

- A clear guidance document, with a transparent and efficient mechanism for submitting SaMD for review, is the most effective way to support this area.
- Addressing many of the ethical and other GCP considerations raised by AI/ML. These include;
    - Clear definitions of what protections are available for subjects whose data is included in datasets used for AI/ML, including how informed consent should be considered and defined.
    - Guidelines on how the FDA would consider "ownership" of data, both the underlying datasets used for training AI/ML, and the outcomes of those algorithms.
    - Clarification on how existing privacy and data ownership laws should be considered when using AI/ML.
    - Definitions on quality, source and complexity of datasets that are being used for AI/ML.
    - Similar to CDISC data standards, are their standards that can be leveraged for how datasets are coded and defined when being used for AI/ML?
- Providing clarification on the duration of a regulatory approval for a device to ensure that advances in technology, developments in medicine, etc are adequately addressed.
- Ensuring that the FDA's position on key matters are as consistent as possible around the globe.

Section 2 "Initial Premarket Assurance of Safety and Effectiveness"

Question 2 "What are the appropriate elements for the Algorithms Change Protocol (ACP) to support SPS?"

- The FDA may consider recommendations for change protocols for AI/ML software/devices not currently meeting SaMD criteria, but likely to function as such in a future version.
- If the Algorithm is approved in a range of indications / populations / age-ranges / gender should the algorithm change be verified in all or only some? Defining how this criteria should be evaluated in regards to an ACP would be helpful.

Section 3 "Approach for modifications after initial review with an established SPS and ACP"

Question 3 "What content should be included in a "focused review?"

- FDA should consider how to deal with potential abuses of the modification process, especially in regard to possible erroneous or malicious attempts to modify algorithms or underlying datasets to manipulate outcomes.

Section 4 "Transparency and real-world performance monitoring of AI/ML-based SaMD"

Question 1 "In what ways can a manufacturer demonstrate transparency about AI/ML-SaMD algorithm updates, performance improvements, or labeling changes, etc?"

- There should be some transparency from the FDA regarding how evaluation and monitoring is done.
- Adverse Events generated by SaMD should be tracked or managed in a manner similar to how the existing FAERS system operates.

Question 2 "What role can real-world evidence play in supporting transparency for AI/ML-SaMD?"

- There are opportunities to leverage many existing datasets in the development of AI/ML. For example, testing outcomes of a predictive model on an existing RWE dataset and compare predictions with actual outcomes *(see Kadra et al 2018, Predicting parkinsonism side effects of antipsychotic polypharmacy prescribed in secondary mental healthcare, Journal of Psychopharmacology, 32(11):1191-1196).*

VII. Questions and Feedback

1. A major concern with many of these machine-learning approaches—especially if at all trained with crowdsourced data—is the ability to bias the algorithm with deliberately malicious input data. This could be used, however, to result in an algorithm that provides deliberately biased outputs. This risk should be addressed specifically and rigorously, with appropriate input from bioethicists.
2. Given the product development timelines and rapid proliferation of software manufacturers, there will likely be a requirement for short turnaround times and quicker decision-making than is typical for FDA. A concern with which to grapple will be how to meet this expectation. Are there any considerations for conditional approval of SaMD? What would those conditional requirements look like?