Alliance for AI in Healthcare (AAIH)
1340 Smith Ave, Suite 400
Baltimore, MD 21209
Tele: (410) 779-1245

June 3rd, 2019

Subject: Proposed Regulatory Framework for Modifications to AI/ML-based SaMD

Dear FDA,

AAIH is an **international multi-stakeholder advocacy organization** based in Washington, D.C. that **promotes scientific, legislative, and regulatory initiatives** necessary to facilitate the development of, access to, and implementation of **artificial intelligence (AI) powered healthcare solutions**. AAIH is comprised of over 25 organizations that utilize AI in **biomedical R&D** and **clinical applications** including growth phase start-ups, biopharma, **diagnostics and device manufacturers**, and research institutions. AAIH **takes the lead** on the sector's most **pressing and significant issues**, fostering **research, development, investment**, and **commercialization** of **transformational treatments and cures** for **patients worldwide.**

It is out of that **dedication to our mission** that we submit our comments today:

After review of the document in total, we have compiled a **combination of general and specific comments** on FDA's proposed Regulatory Framework Discussion Paper. From a **general perspective**, we see **many positive features** of this **proposal and applaud FDA** on taking a **proactive approach in developing AI/ML regulatory frameworks**. We look forward to working with FDA on the continued development of this framework, and to clarifying key aspects of the framework. Due to the **varied nature of these general comments**, and **requests for clarity**, we have supplied **feedback at both a high level** as well as with **specific responses to the 18 posed questions**, which are tabulated in *Appendix 1, Specific Comments on FDA Discussion Paper.*

At a high level:

- We believe this is the **right approach** from the FDA and it is in **line with the technical evolution** in the field toward continuous product improvements and continuous life-cycle management
- We note that some **elements of this framework** seem to borrow from the **well-established QbD principles**. We encourage **FDA to utilize existing, proven concepts as part of building** this ***new*** **regulatory framework**, but caution that devices and AI products at large are not biologics, and so QbD principles need to be carefully considered and modified to match the technology in question before being integrated into a new regulatory framework
- We have **several questions regarding the SPS and ACP** approach, which we feel may be helpful in the **continued development** of this **regulatory framework**, and would be pleased to **work with FDA on a workshop or other educational opportunity** to further explore these concepts:
    - What are the **appropriate elements** for the **SPS**?
    - What are the **appropriate elements** for the **ACP to support the SPS**?
    - What **potential formats** do you suggest for **appropriately describing a SPS and an ACP** in the premarket review submission or application?

Respectfully Submitted,

Annastasiah Mudiwa Mhaka, PhD (Co-Founder) and Adam Roose (Strategic Advisor)

info@theaaih.org www.theaaih.org

Appendix 1 – Specific comments on FDA Discussion Paper:

Line Number/Question	Comment and Rationale
1. Do these categories of AI/ML-SaMD modifications align with the modifications that would typically be encountered in software development that could require premarket submission?	These are typical software development modifications that manufacturers encounter: - Algorithm improvements (retraining, performance, input/output parameters, robustness, etc.) - Adoption of a new computing platform (hardware or software) and upgrade or replacement of embedded frameworks - Principle changes in software architecture, data, and/or methods of training - Changes in user interface (graphic, acoustic, etc.), localization, internationalization, human-device interfaces, form-factor - Integration with other devices or data systems including new means of communication and life-cycle management - Cyber-security and data privacy updates including changes in managing secrets (keys, certificates) - Corrections to defects and specification mismatch The suggested categories of modifications in the document are sufficiently broad to cover major directions of software and algorithm modifications mentioned above.
2. What additional categories, if any, of AI/ML-SaMD modifications should be considered in this proposed approach?	AI/ML model, architecture (ex. CNN model) and training methods are not explicitly addressed in the existing three modifications and so should be addressed. We suggest adding categories for AI/ML model, architecture changes as well as output changes. This would include a ***principle change in the software, user interface, and/or model architecture, and/or training set***. It is theoretically possible, although unlikely, that such change does not affect the system's performance. The change may not affect intended use either. Yet, such change constitutes a major modification to the established model and software characteristics and shall trigger a review. Modifications in the degree of decision autonomy may also fall under this category.
3. Would the proposed framework for addressing modifications and modification types assist the development AI/ML software?	Yes, we believe it would

4. What additional considerations exist for GMLP?	In the development of "GMLP" as a principle we encourage FDA to also consider: - Standards and regulations (ethical, trustworthiness, etc.) - Data management practices including applicable controls (for data integrity) - Documentation of training methods and validation plans - Quality of data sets used in training, testing, and validation - Algorithm's robustness and reliability through the life-cycle - Algorithm's performance metrics - Bias or 'overfit' of data - Auditability of field operations and process transparency - Management of patient health information
5. How can FDA support development of GMLP?	We encourage FDA to continue its methods, and expand its thinking to facilitate development of GMLP principles through the following approaches: - Support development and standardization of Real World Data collections and interfaces for training, testing, and validating algorithms - Build upon existing efforts, such as the Pre-Cert program, and engage both academics and industry to develop guidance to help codify operation of existing and incipient programs - Work closely with NIST, NIH, IEEE, ISO, and other agencies and organizations on data and process standards, and in particular the development of communal resources for AI/ML product development
6. How do manufacturers and software developers incorporate GMLP in their organization?	We believe that this answer differs based on the company and its methodology, and so the answer would be very specific to any individual organization. Regardless, we believe that GMLP should be incorporated into all AI/ML SaMD development processes.
7. What are the appropriate elements for the SPS?	Examples of expected changes, which can be captured by SPS: - Changes in SaMD analytical characteristics (e.g. accuracy, specificity, sensitivity, etc.) - Changes in SaMD clinical characteristics (e.g. intended use, target population, intended use environment, etc.) - Changes in SaMD outputs - Changes in data inputs - Changes in interoperability and integration with other clinical systems

8. What are the appropriate elements for the ACP to support the SPS?	As we understand the proposal, the ACP is meant to be the procedures used to implement changes outlined by SPS. Under these assumptions we believe the higher-level objectives are: - ACP has to be repeatable and reproducible - Clear rationale for the change and its intended effect on clinical outcome should be documented It would be helpful if FDA clarifies what is required for the assessment of change to be fully automated, so the continuous learning/improvement becomes possible. It is unclear if FDA suggests that retraining can be done at run-time or if it is meant to be bundled in software releases. This is a very important difference. As it is written, the SPS/ACP approach could be interpreted as pertaining to. run-time ML opportunity only. However, when the document mentions V&V, does it mean it assumes the full qualification and controlled releases? We would appreciate more clarity on this point.
9. What potential formats do you suggest for appropriately describing a SPS and an ACP in the premarket review submission or application?	**For the SPS:** Structure the format according to the possible areas of change: - Changes in SaMD analytical characteristics (e.g. accuracy, specificity, sensitivity, etc.) - Changes in SaMD clinical characteristics (e.g. intended use, target population, intended use environment, etc.) - Changes in SaMD outputs - Changes in data inputs - Changes in interoperability and integration with other clinical systems For each group, answer the following elements: - Which parameters/components/ functions may change? - Which factors/methods lead to the change? - What are the foreseeable ranges/ categories of change? - What are the potential established risk factors that can be affected by the change? Are there new risks introduced by the change? - How often/regularly is the change expected to take place? **For ACP:** The Figure 4 (ACP components) is adequate.

10. How should FDA handle changes outside of the "agreed upon SPS and ACP"?	Along with the process suggested by FDA and illustrated by Figure 5, we suggest to give the manufacturer an opportunity to revise the SPS and ACP, and ask for a focused review of the revisions. FDA (and/or manufacturer) may come to a conclusion that the scope of the changes, risk profile modifications, advancement in performance and intended use/indications are too significant to be accepted as a simple revision of the SPS/ACP. Then a premarket submission might be required in this case. However, if, based on submitted evidence and rationale, FDA accepts manufacturer's proposal for revising the SPS/ACP documents within the existing approval, then the process returns to the previously approved state with the appropriate revisions to the SPA and ACP. We perceive that there remain several uncertainties: - If ACP is a plan, we are assuming that the expectation of FDA is to see Algorithm Change Report (ACR) filed with DHF. - Is it conceivable that the algorithm may have a built-in continuous life-cycle update management with an automated V&V. Can such generated artifacts be filed as proof of compliance with the approved ACP? - In the process of reviewing SPS/ACP, how far does FDA want to go in evaluating the appropriateness of the data sources when standardized testing sets are not available and/or mandated?
11. What additional mechanisms could achieve a "focused review" of an SPS and ACP?	A special 510 (k) would be appropriate
12. What content should be included in a "focused review"?	We would ask for a Special 510(k) focused review which does not require clinical data (only risk analysis and mechanisms to mitigate risks - based on QS) We ask that a focused review be non-data driven and is totally based on the controlled design documents.
13. In what ways can a manufacturer demonstrate transparency about AI/ML-SaMD algorithm updates, performance improvements, or labeling changes, to name a few?	Already existing design controls contain such elements as design reviews, risk analysis, V&V, complaints. New suggested instruments (SPS/ACP) should be sufficient to support the transparency. With the increasing frequency of the substantial SaMD updates, we suggest preparing mandatory software release notes containing a few major points: - Brief description of the major changes in the software system. - Brief description of the major changes in the algorithmic core of the SaMD.

	- Definition of data sources you are adding into the training set and describe how they may affect the algorithm - A concise summary of the expected effect of the change on algorithm performance, risk profile, intended use, clinical workflow and/or outcomes, user experience - Reference (trace to) DHF documents providing more detailed report on the outlined changes FDA may want to create a registry where this briefs/release notes can be submitted for transparency and traceability. We may also suggest the following 'rule of thumb': - If changes are within the approved SPS and APC, then there is no need for an update to FDA outside of existing regulatory requirements such as complaint handling - If changes fall outside the bounds of an approved SPS / ACP then a new premarket review is expected, in which the triggering change and all changes since 510k clearance would be included Proposals for a fixed interval update to FDA o This should depend on the nature of the review process Ex. For a PMA, you submit an annual report o There is no such process for 510(k) or De Novo o If we want there to be a similar mechanism for 510(k) in this space, FDA may need to create something new as described in the Pre-cert program
14. What role can real-world evidence play in supporting transparency for AI/ML-SaMD?	RWE can have a substantial impact in AI/ML transparency through: - Introduction to new challenges, don't know what your algorithm is encountering out in the field, but RWE may assist in determination - Introduction of representative sample size - Cause the algorithm to experience new features that were not in the dataset, may provide transparency about possible degradation - While more data is better, not all data is equal so RWE may have costly effects on algorithm training - If we can use RWE to produce smaller yet effective sample sizes and increased transparency, this is a win - RWE is really a safeguard, having datasets with RWE might make your algorithm more effective - Risk – it can open up models to malicious data hacking
15. What additional mechanisms exist for real-world performance monitoring of AI/ML-SaMD?	We recommend that FDA harmonize with principles in the Pre-Cert program, and to not 'reinvent the wheel'. However, we would also recommend FDA consider reconciling subjective quantification of information vs the potential for objective degradation of RWE in the

	process. Considerations must be made for how computerizing of human biases could affect AI.

16. What additional mechanisms might be needed for real-world performance monitoring of AI/ML- SaMD?	- Existing complaint handling (how to quantify complaints when used for RWE) We suggest the relevant elements of the Pre-Cert program be implemented here as well
	- Conduct periodic algorithm performance review with users and stakeholders (healthcare professionals and associations, patient focus groups and associations, distributors, etc.):
	  o Does the device meet performance expectations?
	  o Are there specific cases when the device does not perform?
	  o Are there known cases of off-label use? What kind?
	  o Are there any suggestions for possible device improvements?
	- Are the complaints and/or adverse events being reported as suggested?
	- If algorithmic bias and/or incompleteness of the data sources is important to consider, the manufacturer shall address this in the ACP. Appropriate addressing of the issue should include analytical reasoning regarding the origin of the bias and proposals on how to mitigate for its presence. The manufacturer should demonstrate good understanding of the problem dimensionality and possible error in the data.

17. Are there additional components for inclusion in the ACP that should be specified?	None other than those mentioned above
18. What additional level of detail would you add for the described components of an ACP?	We see this proposal as having good detail and appreciate that each of the four main categories is also broken into sub-categories with specific elements. However, additional detail would be appreciated in particular as it pertains to device-specific, intended use-specific considerations.