



June 3, 2019

Submitted Electronically

Dr. Norman Sharpless, M.D.
Acting Commissioner
U.S. Department of Health and Human Services
Food and Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993-0002

Re: Comments of the ACM U.S. Technology Policy Committee on Proposed FDA Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning Based Software as a Medical Device Discussion Paper (Docket No. FDA-2019-N-1185)

Dear Commissioner Sharpless:

In response to the FDA's call for public comment on the above-referenced Discussion Paper, I am pleased to submit the following observations and recommendations of the U.S. Technology Policy Committee ("Committee") of the Association for Computing Machinery ("ACM"). ACM is the longest-established and, with more than 100,000 global members, the largest association of individual professionals engaged in all aspects of computing in the U.S. and the world. A non-lobbying and otherwise wholly apolitical organization, ACM's mission includes providing unbiased, expert technical advice to policymakers on matters of our members' wide-ranging expertise. That work is accomplished in the United States by and through the Committee, to which these comments should be attributed.

First, the Committee respectfully urges the FDA in future stages of this proceeding to separately identify and request input on two distinct developments in the design of medical software ("devices") currently commingled in the proposed Framework. This concern, elaborated upon below, relates to the: 1) need and opportunity for software-based devices to evolve, and risks associated with such evolution; and 2) use of artificial intelligence to make decisions.

* The Committee recognizes the members of its ad hoc working group responsible for organizing and generating this submission: James Hendler, Harry Hochheiser, Juan Miguel de Joya, Lorraine Kisselburgh, Andy Oram, Arnon S. Rosenthal, and Shahid N. Shah.

Evolution of Software-Based Devices and Associated Risks

The FDA is properly focused in this docket on software-based devices capable, when connected to networks, of being improved, corrected, and upgraded to enhance security. Such connectivity also introduces (and potentially heightens) the risk of error in FDA-approved devices. It also introduces new opportunities for devices to fail and dramatically increases their exposure to attack. We also note that modern and increasingly agile software development methods permit the ever more rapid update and iteration of such products. FDA testing and validation thus need to keep pace with such accelerated schedules.

We recommend, therefore, that the FDA redesign its processes to specify what changes (whether human- or AI-generated) may be made to a device when the manufacturer applies a previously-approved testing process, but without FDA review. Examples include emergency security patches and adjusting a setting within bounds that have been tested. In doing so, it may be productive to permit or require a manufacturer to share information collected automatically from its version control, development, and test processes documenting quality measures.¹

Use of Artificial Intelligence to Make Decisions

The success of FDA regulatory models for devices that change their function based on AI input will be a function of how AI is used in practice in specific applications. Specifically, no regulatory change is likely to be needed if a medical device manufacturer collects data from the field (through its devices or otherwise) and uses AI purely to identify design features and parameters during its standard development process.

For example, if AI analysis of data collected reveals that a device should activate at a lower measurement for a vital sign, the change should continue to undergo an FDA-approved validation process, as it does now. Such use of AI thus is simply one of many ways to improve the design of devices. All the usual quality checks and regulatory requirements for development and testing should still have the desired effects of guaranteeing quality.

¹ Given that validation may require manufacturers to supply large amounts of data about a device and its testing, the FDA will need to define quantitative and automatable processes for specifying, collecting, and validating such data. Otherwise, both the manufacturer and the FDA will be overwhelmed by the amount of manual validation involved. Such manual validation, of course, would undermine the goals of accelerated development. A productive first step toward the necessary automation thus could be to extend or clarify the FDA's definition of "intended use" to fully embrace a more detailed, modern, measurable, metrics-oriented, and machine-computable definition of what a device is *supposed* to do.

By contrast, new regulatory models will be essential if a manufacturer proposes to use AI to dynamically change a device's behavior in the field without being subject to a regulated development and testing process. Removing the human from the “loop,” coupled with bypassing testing, would increase the risk that the AI employed will have unintended effects on data accuracy and device safety.² There are many possible regulatory responses to these important statistical and physical issues.³ These include:

- Banning such dynamic, significant AI-dictated updates to device behavior because the results cannot be guaranteed accurate and safe given our current knowledge of AI;
- Requiring a manufacturer to install “governor” software in its device which will reliably assure that the device’s behavior cannot degrade beyond articulable and enforceable limits;⁴ and
- Mandating or encouraging data sharing among health care providers and device manufacturers to maximize the amount, and assure the validity, of large quantities of training data.

Second, the Committee also urges the FDA, in devising its new Artificial Intelligence/ Machine Learning testing and certification protocols, to:

- Employ and require outcomes-driven, automated and (where possible) deterministically reproducible testing outside of the vendors’ own development laboratories;⁵

² Before AI, devices behaved deterministically. In other words, a device could reliably be expected to behave the same way under the same conditions. In the use case above, however, AI would render devices non-deterministic; their behavior would be unpredictably variable under unchanged conditions.

³ The Committee notes that the use of AI/ML in SaMDs inevitably also raises important *ethical* issues outside the scope of the FDA’s current Discussion Paper and thus the immediate docket. However, such matters – particularly with respect to the use and operation of potentially autonomous implanted devices – must not remain outside the scope of agency and legislative determinations about whether and how such technologies should be utilized in human beings, and the degree of input into and feedback about their operation that humans (especially those implanted with such devices and their medical supervisors) must continue to be assured.

⁴ This fear is not hypothetical. Infamously and instructively, in the mid- to late-1980s, Therac-25 machines were misconfigured and delivered radiation doses hundreds of times greater than justified. *See, e.g., “An Investigation of the Therac-25 Accidents,” Computer*, Vol. 26 No. 7 (July 1993), pp. 18-41 available online at: web.stanford.edu/class/cs240/old/sp2014/readings/therac-25.pdf

⁵ At the first stage of such modern test-first and test-driven evaluation, the focus can and should be primarily on the quality and rigor of the tests themselves. If the intended use statements are well defined, measurable -- and testing is automated and reproducible -- the internal workings of a device are less important at this early juncture.

- Require manufacturers to create a common pool of data for input to AI analyses, including both real-world deidentified data and synthetic data; and
- Foster the development of a common pool of varied simulation and other test environments using the deidentified and synthetic data endorsed above.⁶

Conclusion

AI and its forms of data processing have produced huge changes in industrial decision-making and promise at least equivalently profound revolution in the relationship between human and machine. This is true in no sphere more so than that of human health.

The Committee commends the FDA for its efforts to both foster and regulate the transformative power of Artificial Intelligence and Machine Learning by striving to understand when the use of AI calls for regulatory changes in device approval and review, and when current oversight models will suffice. We hope that the foregoing comments will be useful in that critical work.

Thank you for the opportunity to participate in this critical effort. Should you or your staff have any questions regarding these Comments, or seek further expert analysis or information our members may provide, please email Adam Eisgrau, ACM's Washington-based Director of Global Policy & Public Affairs, at the address below or reach him at 202-580-6555.

Sincerely,



James A. Hendler, Chair

⁶ Just as no aircraft would be test flown (and certainly would not be put into public service) before being rigorously challenged virtually, SaMD offerings also should be fully tested in simulators and/or by other appropriate means before being approved for use.