Congress Must Update FDA Regulations for Medical AI

Scott Gottlieb, MD

When the US Food and Drug Administration (FDA) cleared the Apple Watch in 2018 for use in detecting irregular heart rhythms, many observers were worried that the agency would become mired by its need to dismantle the device to affirm its functionality. Instead, when Apple came to the FDA with a proverbial mountain of data proving its reliability in real-world settings, it bolstered a regulatory concept that was already forming among the agency’s policymakers: taking a firm-based approach to the regulation of lower-risk medical devices. This approach shifted the regulatory focus from scrutinizing the product itself to the results it generated and the developer’s overall approach to confirming the device’s reliability. The FDA cleared the Apple Watch based on the careful validation it had undergone, not by trying to dissect its hardware. The FDA shifted the nexus of its regulation from the product’s mechanism to the firm’s overall development process.1,2

The FDA’s firm-based approach to regulating medical devices means that the regulatory oversight focuses on the quality systems and processes of the manufacturing firm, rather than just on the mechanics of the individual products. This approach emphasizes the importance of a company’s quality management system to ensure consistent production and control of medical devices, and involves routine inspections and audits to verify the firm’s procedures and operational controls. By ensuring that the manufacturer has robust systems in place, the FDA aims to enhance the overall safety and effectiveness of the medical devices produced, thereby providing better protection for public health.

This same concept is uniquely suited to the regulation of artificial intelligence (AI) medical devices that can augment patient care. Legislation currently before Congress (Senate Bill 2209 and House Bill 4128), the Verifying Accurate Leading-edge IVCT Development Act (VALID Act), codifies this firm-based approach to regulation. It enables the FDA to oversee the methods used to develop a technology and validate its reliability, rather than trying to decouple the product’s construction.

The VALID Act was crafted to modernize the FDA’s oversight of in vitro diagnostics and strike the right balance for the agency’s regulation of laboratory-developed tests. However, much like advanced diagnostics and genomic tests, the AI used in medical devices is designed to analyze vast amounts of data to generate clinical insights. The challenges Congress is attempting to solve with the VALID Act are similar to those confronted by the FDA’s evaluation of medical AI. One of the challenges involves the rapid cycles of innovation inherent to these products because they undergo constant modification as new information becomes available.3

The AI models have evolved to a point where they can directly manage certain patient interactions. For instance, ChatGPT (OpenAI) has demonstrated substantial semantic medical knowledge and the ability to perform medical reasoning.4 The AI medical devices can interface with patients, handle routine follow-up care, and defer to clinicians for complex cases. These devices have the potential to offer patients specific information relevant to their health status and assist them in managing their care alongside clinicians. However, large language models trained on vast datasets embody the ultimate black box in the realm of FDA regulation. These models involve complex layers of computations and extensive data processing, resulting in predictions with typically opaque reasoning pathways. Their nonlinear and high-dimensional nature make it challenging to trace specific inputs to outputs—even their developers cannot fully elucidate how these models draw their conclusions.

Open Access. This is an open access article distributed under the terms of the CC-BY License.
One risk, however, is that these systems can return wrong answers, especially when the devices are trained on unreliable datasets. Just as regulators could not be expected to fully deconstruct the Apple Watch to understand every aspect of its functionality, the FDA cannot fully comprehend how AI medical devices generate every possible answer. These systems meet the definition of being a medical device and fall squarely inside the FDA’s purview. However, lacking the ability to take a firm-based approach to regulating these devices, the FDA will struggle with how to efficiently oversee these advanced tools.

Product developers are equally unsure about how the FDA will approve these AI medical devices. To navigate this uncertainty, many manufacturers are opting to classify them as non-device clinical decision support software, which is a type of health information technology designed to assist clinicians in making decisions for patient care. Clinical decision support software analyzes data from various sources (such as patient records and medical databases) to contextualize patient information and provide evidence-based recommendations to clinicians. Because the clinician remains involved in the decision-making and can verify the system’s outputs, these devices generally avoid the need for premarket FDA clearance.

Much like their approach taken in creating electronic medical records, in which developers intentionally limited functionality to avoid classifying these software tools as medical devices, developers seeking to apply AI to clinical care are intentionally curbing the utilities of these tools. Product developers are apprehensive that if the regulatory process remains uncertain, the FDA will become entangled in the boundless task of deciphering the product’s underlying code.

Under a firm-based approach to regulation, innovators could bring certain new products to market and update their existing ones without undergoing the same premarket review in every case. This is a modern approach to regulation that would make the introduction of new innovations far more efficient. It is well suited for products like AI medical devices, for which the goal is to undergo near-constant evolution of the model to continuously improve its performance.

Additional reforms could establish a framework for third-party certification of some lower-risk AI devices. Independent organizations meeting criteria set by the FDA would assist in evaluating and certifying these devices. In the complex realm of AI (an area where the FDA might struggle to recruit and retain top experts), this flexibility to rely on certified evaluations would enable the agency to tap into external expertise. Such an approach could streamline the FDA’s review process while enhancing the depth of expert oversight the agency can provide.

Because the functionality of these AI devices is heavily dependent on access to reliable health care data, Congress can also facilitate the creation of large datasets for AI development, training, validation, and postmarket monitoring. Private organizations already help the FDA aggregate data from electronic health records and payer claims to support postmarket surveillance of medical products. Similar entities could be empowered to create a federated network of data required to adhere to FDA standards for accuracy and representativeness, ensuring that the data meet a regulatory grade for inclusion in the development and training of AI medical devices. As Newman-Toker and Sharfstein noted in proposing a similar idea, the quality of the information in data training sets for AI systems is essential for certain uses (such as diagnostic accuracy). In turn, models that are trained or validated on this regulatory grade data could undergo less premarket scrutiny because there would be greater assurance that the models were developed using reliable datasets.

To unlock the potential of these tools for patients, regulatory policies need to be crafted to address their unique characteristics. Just as the FDA could not dismantle the inner workings of the Apple Watch, it cannot unravel the programming behind large language models. The FDA’s traditional regulatory approach, which depends on the agency’s capacity to meticulously examine a product’s construction, might prove infeasible in this context. Furthermore, the FDA cannot develop a firm-based approach solely through its own guidance and regulation. Congress must enact new laws to codify these modern concepts that are essential for regulating this new technology.