NBER WORKING PAPER SERIES

THE REGULATION OF MEDICAL AI: POLICY APPROACHES, DATA, AND INNOVATION INCENTIVES

Ariel Dora Stern

Working Paper 30639
http://www.nber.org/papers/w30639

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2022

The author is grateful to Melissa Ouellet for excellent research and programming support. Boris Babic, Anna Goldenberg, Nicholson Price, Daniel Yong, and participants at the Fall 2022 Economics of Artificial Intelligence Conference provided thoughtful comments and feedback. The author reports sitting on the Scientific Advisory Board of the German Society of Digital Medicine and the Strategic Advisory Board of HumanFirst. From 2020-2021 she consulted to the German Federal Ministry of Health on topics that included the regulation and reimbursement of digital medical devices. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

The Regulation of Medical AI: Policy Approaches, Data, and Innovation Incentives
Ariel Dora Stern
NBER Working Paper No. 30639
December 2022
JEL No. I11,I18,K2,K32,O31,O32,O33

ABSTRACT

For those who follow health and technology news, it is difficult to go more than a few days without reading about a compelling new application of Artificial Intelligence (AI) to health care. AI has myriad applications in medicine and its adjacent industries, with AI-driven tools already in use in basic science, translational medicine, and numerous corners of health care delivery, including administrative work, diagnosis, and treatment. In diagnosis and treatment, a large and growing number of AI tools meet the statutory definition of a medical device or that of an in-vitro diagnostic. Those that do are subject to regulation by local authorities, resulting in both practical and strategic implications for manufacturers, along with a more complex set of innovation incentives. This chapter presents background on medical device regulation—especially as it relates to software products—and quantitatively describes the emergence of AI among FDA-regulated products. The empirical section of this chapter explores characteristics of AI-supported/driven medical devices ("AI devices") in the United States. It presents data on their origins (by firm type and country), their safety profiles (as measured by associated adverse events and recalls), and concludes with a discussion of the implications of regulation for innovation incentives in medical AI.

Ariel Dora Stern
Harvard Business School
Morgan Hall 433
Boston, MA 02163
astern@hbs.edu

1. INTRODUCTION

For those who follow health and technology news, it is difficult to go more than a few days without reading about a compelling new application of Artificial Intelligence (AI) to health care. Applications range from basic science (e.g., understanding protein folding), to translational science (e.g., supporting drug discovery), to improving existing digital offerings (e.g., machine learning (ML) algorithms to adjust for missing data in whole genome sequencing software), to tools that promise to improve health care delivery in myriad ways. Recent work has highlighted and categorized the applications of AI to health care delivery and emphasized how contemporary deep learning approaches are likely to transform health care (Hinton 2022). The overwhelming majority of applications of AI fall into one of three broad categories: (1) administrative work, (2) diagnosis, and (3) treatment. Table 1 provides examples of the types of AI tools that would fit into each of these categories.

Table 1: Applications of AI to health care delivery: examples

Category	**Examples**
(1) Administrative Work	Provider documentation
	Order, prescription, coding entry for providers Data entry
	Scheduling
	Triaging
(2) Diagnosis	Imaging / pathology review
	Diagnostic models / symptom analysis
	Phenotyping
	Incorporating non-traditional data sources
(3) Treatment	Surgical assistance
	Individualized / personalized medicine
	Adherence and health coaching
	Generating treatment recommendations
	Digital therapeutics

Source: Adapted from Sanders et al. 2019

To make their way into routine health care delivery, AI tools for administrative work will need to cater to provider preferences, workflows, and other site-specific norms (Sanders et al. 2019). Beyond these practical and design challenges, AI-driven administrative support tools need to comply with data privacy regulations in the

jurisdictions in which they are used—most notably, the HIPAA Privacy Rule[1] in the United States and the GDPR[2] in Europe. AI tools to perform or support administrative work in health care hold great promise to improve the efficiency of health care delivery by aiding in clinician notetaking and documentation, scheduling, triaging, ordering medications, and avoiding medication errors – including foreseeable negative interactions. Importantly, however, administrative support tools rarely qualify as regulated medical products and, conditional on compliance with appliable privacy laws, therefore rarely fall under the jurisdiction of medical product—chiefly, medical device—regulations.

In diagnosis and treatment, however, a large and growing number of AI tools meet the definition of a medical device or that of an in-vitro diagnostic. Those that do are subject to regulation by local authorities, with implications for manufacturers and a more complex set of innovation incentives. This chapter provides brief background on medical device regulation in the United States and Europe and discusses a few emergent regulatory approaches that are designed to address some of the unique challenges of regulating software as a medical device. It then takes a closer look at regulated AI devices in the United States by identifying such devices in the FDA's databases.

The empirical section of this chapter explores regulated, software-based, AI-supported/driven medical devices ("AI devices") in the United States. By taking advantage of publicly available information about all medical device clearances and associated product summaries, this section uses text analysis to identify AI devices and compare these to other devices in the same medical product areas – including the subset of comparator devices that are themselves software-driven. In particular, we characterize AI devices based on the types of firms they originate in and the countries in which they are developed. The chapter also presents summary data on the safety profiles of AI devices, as measured by mandatory adverse event reports and product recalls. Building on descriptive statistics from the U.S. data, the fourth section of this chapter discusses how regulation is likely to shape innovation incentives for (certain types of) AI devices and how both regulatory innovation and regulatory transparency may play a role in the future of regulated AI devices. The chapter concludes with a brief discussion of a forward-looking research agenda.

[1] The U.S. Health Insurance Portability and Accountability Act (HIPAA) dates back to 1996. The HIPAA Privacy Rule "establishes national standards to protect individuals' medical records and other individually identifiable health information (collectively defined as "protected health information" [PHI]) and applies to health plans, health care clearinghouses, and those health care providers that conduct certain health care transactions electronically." In addition to requiring appropriate safeguards" to protect PHI, the rule limits how data can be used/reused without an individual's authorization and gives individuals the right to obtain and examine copies of their own health records. (https://www.hhs.gov/hipaa/for-professionals/privacy/index.html)

[2] The European Union's General Data Protection Regulation (GDPR) "lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data." The regulation further "protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data" and governs the movement of such data within the EU. (Art. 1 GDPR, https://gdpr-info.eu/art-1-gdpr). The GDPR specifically recognizes "data concerning health" as its own category and provides specific definitions for health data for the purposes of data protection under the GDPR (https://edps.europa.eu/data-protection/our-work/subjects/health_en).

2. BACKGROUND AND POLICY APPROACHES

Medical device regulation in the United States

US medical product regulation can be traced back over a century to the Pure Food and Drug Act of 1906. However, modern medical *device* regulation began with the 1976 "Medical Device Amendments" (MDA) to the 1938 Federal Food, Drug, and Cosmetic Act. The MDA created federal oversight of medical devices for the first time (previously they had been regulated by the states) and established the framework for how medical devices are regulated today. Section 201(h) of the Food, Drug, and Cosmetic Act defines a medical device as:

> *An instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including a component part, or accessory which is:*
>
> 1. *recognized in the official National Formulary, or the United States Pharmacopoeia, or any supplement to them,*
> 2. *intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals, or*
> 3. *intended to affect the structure or any function of the body of man or other animals, and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and*
>
> *which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which is not dependent upon being metabolized for the achievement of its primary intended purposes. The term "device" does not include software functions excluded pursuant to section 520(o).*[3]

In short, a medical device is a tool for the diagnosis or treatment of disease, which is not a metabolized (biological or pharmaceutical) product.

Current medical device regulations are focused on providing users (including patients, clinicians, provider organizations, and caregivers) reasonable assurance regarding the safety and effectiveness of medical devices. Medical devices are regulated by the FDA's Center for Devices and Radiological Health (CDRH), which uses a risk-based (3-tier) classification system for all devices:[4]

- Devices of the lowest risk (Class I) are subject to only general manufacturing controls and typically exempt from needing a premarketing submission/application. These include products such as tongue depressors, condoms, latex gloves, bandages, and surgical masks.
- Moderate risk (Class II) devices are typically regulated through a process called "Premarket Notification" or, more often, the "510(k) process." This process requires a device to demonstrate "substantial equivalence" with one or more already legally marketed devices.[5] Class II devices that do not have a legally marketed "predicate" device can also use a "De Novo" Classification request to

[3] https://www.fda.gov/medical-devices/classify-your-medical-device/how-determine-if-your-product-medical-device
[4] https://www.fda.gov/medical-devices/overview-device-regulation/classify-your-medical-device
[5] A device is considered substantially equivalent if, in comparison to a predicate it has the same intended, the same technological characteristics, or the same intended use and has different technological characteristics and does not raise different questions of safety and effectiveness; and the information submitted to FDA demonstrates that the device is as safe and effective as the legally marketed device (https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/premarket-notification-510k)

come to market if the manufacturer is able to provide reasonable assurance of safety and effectiveness of the device for the intended use. Subsequent devices can then use a device that came to market through the De Novo process as a predicate in their own ensuing 510(k) applications.

- Finally, devices of the highest risk (Class III) are those that are implantable and/or life sustaining and, as such, require significant evidence of safety and effectiveness to be approved for marketing. With a few exceptions for devices that pre-date the MDA, Class III devices are regulated through a process called "Premarket Approval" or the "PMA process," which is the most rigorous of all premarket submissions and typically requires evidence from clinical studies. The PMA process is significantly more onerous than the 510(k) process and has been associated with longer periods of regulatory approval for first movers in new medical device product codes (Stern 2017).

In addition to establishing these regulatory pathways for new medical devices, the MDA created a regulatory pathway for new investigational devices to be studied in human patients, the Investigational Device Exemption (IDE), and established several postmarket requirements and processes—including adverse event reporting requirements—and Good Manufacturing Practices (GMPs). The Safe Medical Devices Act of 1990 filled in additional policy gaps in medical device regulation by authorizing the FDA to order device recalls, to impose civil penalties for violations, and improved postmarket surveillance by requiring both manufacturers as well as user facilities (hospitals, clinics, nursing homes, etc.) to report adverse events associated with the use of specific medical devices.[6] Further, the FDA's Breakthrough Devices Program was created in 2018 to provide patients and health care professionals with more timely access to devices that "provide for more effective treatment or diagnosis of life-threatening or irreversibly debilitating diseases or conditions" and as of the end of Q3, 2022 a total of 56 devices with the Breakthrough Device Designation had received marketing authorization.[7]

Regardless of the regulatory pathway used, all devices are categorized into 3-letter product codes that describe the device's generic category of use. Within a product code, devices are thus very good to excellent substitutes for one another. For example, unique product codes exist for *Coronary Drug-Eluting Stent* (NIQ), *Catheter, Balloon For Retinal Reattachment* (LOG), *Oximeter, Fetal Pulse* (MMA), and *Infusion Safety Management Software* (PHC).[8] Regulation happens at the level of 18 panels of the CDRH Advisory Committee, which are organized by medical specialty (e.g., the "Circulatory System Devices Panel" reviews cardiovascular devices, while the "Radiological Devices Panel" reviews radiology devices).

For devices that are cleared via the 510(k) or De Novo processes or approved via the PMA process, several public documents are published online at the time that a device receives a positive regulatory decision. These include a device "Summary" for 510(k)-track devices, which "includes a description of the device such as might

[6] https://www.fda.gov/medical-devices/overview-device-regulation/history-medical-device-regulation-oversight-united-states
[7] https://www.fda.gov/medical-devices/how-study-and-market-your-device/breakthrough-devices-program
[8] https://www.fda.gov/medical-devices/classify-your-medical-device/product-code-classification-database

be found in the labeling or promotional material for the device, including an explanation of how the device functions, the scientific concepts that form the basis for the device" as well as information on "the significant physical and performance characteristics of the device, such as device design, material used, and physical properties," making this document an excellent source of information on a device's key technological characteristics. The PMA process also requires a product-specific summary document, which is made publicly available at the time the device is approved. PMA summary documents also contain information on indications for a device's use and a detailed device description, including "how the device functions, the basic scientific concepts that form the basis for the device, and the significant physical and performance characteristics of the device" in addition to other requirements.[9] The text analysis that follows takes advantage of publicly available product summaries in order to understand and categorize devices' technical content and functionality at scale.

Regulation of Software-Driven Medical Devices

Because U.S. medical device regulation is grounded in legislation from 1976 (namely, the MDA), provisions for thinking about the regulation of software-driven products were not codified for the first several decades. This means that until recently, medical device regulations were woefully mismatched to the special needs and nuances of software products. Specifically, any significant updates to medical devices have historically required new applications to regulators. For a moderate risk device, there are no regulatory provisions for amending or changing an existing 510(k) clearance – that is "if it is determined the modification is not covered by the current 510(k) a new 510(k) must be submitted."[10] PMA-track devices can only be modified through a "PMA supplement," a "submission required for a change affecting the safety or effectiveness of the device for which the applicant has an approved PMA."[11] For example, a software change "that significantly affects clinical functionality or performance specifications" would require a new premarket submission (FDA 2019).

This policy rigidity and the discreteness of product updates in this context contrast starkly with products in consumer technology settings, where software programs are regularly, if not constantly, being improved upon and modified by developers. Nevertheless, a strict interpretation of U.S. regulatory policies would require a new regulatory submission in the event of a modification to an existing software program that improves the accuracy of a diagnosis being made or information being conveyed to clinicians. While some exceptions were made previously for the addressing of safety and security issues associated with medical device software, it was not until 2017 that

[9] https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?FR=814.20
[10] https://www.fda.gov/medical-devices/premarket-notification-510k/new-510k-required-modification-device
[11] https://www.fda.gov/medical-devices/premarket-approval-pma/pma-supplements-and-amendments

the FDA published formal regulatory guidance addressing when a manufacturer should submit a 510(k) for a software change to an existing medical device.[12]

There are two primary ways in which software can be included in a medical device: (1) the medical device may be software-driven in that it is a physical device that is powered by software that is inextricable from the device's functionality, sometimes called "Software in a Medical Device" or "SiMD." An example of SiMD would be the software that powers a CT scanner: the hardware device does not work without the software. (2) Alternatively, a medical device may *entirely* software-based—that is, the software itself meets the definition of a medical device (including in-vitro diagnostics). This second category is termed "Software as a Medical Device" or "SaMD" by the International Medical Device Regulators Forum, which defines SaMD as "software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device"—that is, stand-alone software (IMDRF 2013).

In the United States, SaMD products are typically classified as Class II devices and are regulated via the 510(k) or De Novo pathways. However, low-risk products that meet the definition of SaMD often qualify for "enforcement discretion" meaning that the FDA will not enforce regulatory requirements for these software products.[13] [14] In light of the complexity of regulating SaMD products, there have been recent calls for "innovation in regulatory approaches" to further address the unique needs of SaMD (Torous et al. 2022).

One proposed approach to the regulation of SaMD products was considered in the FDA's Digital Health Software Pre-Certification Program (Pre-Cert), a pilot program that was first initiated in 2017 and ran until September of 2022.[15] The pilot phase of the program was intended to "help inform the development of a future regulatory model that [would] provide more streamlined and efficient regulatory oversight of software-based medical devices" and included an outline of how FDA might evaluate SaMD products more responsively by focusing on reviewing an "Excellence Appraisal" of the manufacturer's software development practices, as well as review pathway determination, streamlined review, and real-world performance data collection and evaluation.[16] While the program held the promise of being more dynamic—an approach that surely makes sense for medical software—it was ultimately concluded abruptly in September of 2022 after "FDA encountered challenges with implementing the proposed approach under [its] current statutory authorities." Among other things, FDA reported

[12] https://www.fda.gov/media/99785/download
[13] https://www.fda.gov/media/80958/download
[14] Other sometimes-regulated products include "smart wearables" such as devices manufactured by Fitbit, AliveCor, Gramin, and Apple, which have both medical and consumer applications. Felber and Maciorowski (2023) provide an overview of how such "smart wearables" are used and regulated in the United States along with relevant risks and public purpose considerations.
[15] https://www.fda.gov/media/161815/download
[16] https://www.fda.gov/media/106331/download

that it "was simultaneously unable to pilot the program approaches with a broad sample of devices while also being unable to limit the scope of any resulting device classifications" It noted that certain types of information gathering were hampered by the fact that it "could not require pilot participants to provide information under the pilot that was not otherwise already required under existing statute," however many participants helped voluntarily.[17] Other challenges with the Pre-Cert approach were foreshadowed even before the conclusion of the pilot. For example, researchers found it difficult "to identify a standard measure that differentiated apps requiring regulatory review from those that would not" using publicly available information, including product descriptions (Alon et al. 2020). Thus, while it is almost certainly the case that novel regulatory approaches are needed for SaMD in general and for AI devices in particular, it will likely be difficult to implement such approaches in the absence of formal updates to regulators' statutory authority that take into account the specific and dynamic needs associated with software products.

Medical Device Regulation in the European Union

The regulation of medical devices in European Union (EU) member states bears several similarities to U.S. medical device regulation (including a risk-based framework governing device classification and regulation), but also has many important differences. The EU Medical Device Regulation 2017/745 (MDR)[18] was adopted in 2017 and fully implemented—thereby replacing previous directives and regulations—in May of 2021. The MDR defines a medical device as:

any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes:

- *diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease,*
- *diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability,*
- *investigation, replacement or modification of the anatomy or of a physiological or pathological process or state,*
- *providing information by means of in vitro examination of specimens derived from the human body, including organ, blood and tissue donations,*

and which does not achieve its principal intended action by pharmacological, immunological or metabolic means, in or on the human body, but which may be assisted in its function by such means.

In the EU, medical devices undergo a conformity assessment, leading to a "CE-Mark," from a notified body (an organization designated by an EU country to assess the conformity of certain products before being placed

[17] https://www.fda.gov/media/161815/download
[18] https://eur-lex.europa.eu/eli/reg/2017/745/2020-04-24

on the market[19]) and must comply with the EU's General Data Protection Regulation ((EU) 2016/679, GDPR)[20], which "protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data" and presents a general and binding framework for processing the personal data of any person within the EU and by any data processor (party) in the EU. Brönneke et al. (2021) provide an overview of regulatory, legal, and market aspects facing digital products, including a more detailed discussion of the application of the MDR and GDPR to digital medical devices.

3. DATA

The empirical section of this chapter explores regulated, software-based, AI-supported & AI-driven medical devices (henceforward "AI devices")[21] in the United States. By taking advantage of detailed, publicly available information about medical device clearances and associated product summaries, this section uses text analysis to identify AI devices and compare these to other devices in the same medical product areas – including the subset of those comparator devices that are themselves software-driven.

One of the first studies to survey FDA-regulated AI devices was Benjamens et al. 2020, which claimed to publish "the first comprehensive and open access database of strictly AI/ML-based medical technologies that have been approved by the FDA." The database, hosted by The Medical Futurist Institute (TMF) included 79 devices, as of mid-2022 (but had not been updated since mid-2021). [22] While the data used in this chapter were collected independently, the TMF database provided early clues to inform how to best identify AI devices and which medical specialties are likely to be most relevant for AI applications at present. For example, just two of the devices in the database were regulated through the PMA process for devices of the highest risk, with the remainder were brought to market via the 510(k) or De Novo pathways. This is consistent with most SaMD products coming to market through these pathways, as envisioned by the Pre-Cert program. Further, among the devices in the TMF database, over 80% are either radiology or cardiology devices, pointing to the outsized representation of these two medical specialties among AI devices. Combined, these facts strongly suggest that focusing just on a) devices regulated via the 510(k) pathway and b) the most common regulatory medical specialties including cardiology and radiology, should allow us to limit the scope of data collection, while still likely capturing the vast majority of FDA-regulated AI devices. A subsequently published list of AI devices released by the FDA and

[19] https://single-market-economy.ec.europa.eu/single-market/goods/building-blocks/notified-bodies_en
[20] https://gdpr-info.eu/
[21] This study does not distinguish between AI-supported vs. (entirely) AI-driven medical devices. For example, a piece of radiology equipment that uses AI to improve image quality (AI-supported) would qualify as an AI device, as would a SaMD product in which the algorithm itself constitutes the entirety of the medical device (fully AI-driven).
[22] List pulled on June 7, 2022 from https://medicalfuturist.com/fda-approved-ai-based-algorithms

updated in October 2022 confirmed similar patterns, with just two out of hundreds of AI devices that came to market since 2010 having done so via the PMA pathway.[23] The data assembly strategy and empirical analyses that follow are based on this data-informed approach.

As the basis for our analysis, we download the full 510(k) database for the years 2010 through 2022Q3.[24] We focused on the eight largest medical specialties (as defined by their respective FDA Advisory Committee Panels), which are (1) Clinical Chemistry and Toxicology Devices; (2) Cardiovascular Devices; (3) Dental Products; (4) Gastroenterology-Urology Devices; (5) General Hospital and Personal Use Devices; (6) Orthopedic and Rehabilitation Devices (7) Radiological Devices; and (8) General and Plastic Surgery Devices. This resulted in a total of 38,812 unique device clearances over 13 calendar years, as presented in Table 2. The 510(k) database includes information on the device type (product code), data about the applicant firm (device manufacturer), such as its name and filing address, the dates on which each application was submitted to regulators, the dates on which each device was cleared by the FDA, and the medical specialty (as assigned to one of the FDA Medical Device Advisory Committees[25]) associated with the device. We further flag devices that were part of the "Breakthrough Devices Program" (see section 2 for more detail).

We merge onto this database two additional sources of information on device outcomes: (1) data on adverse events associated with medical devices, as collected in the FDA's medical device adverse event reporting database (the "Manufacturer and User Facility Device Experience" or MAUDE database)[26]; and (2) data on medical device recalls—a more definitive indication of a systematic problem with a product—from the FDA's recall database.[27]

The MAUDE database includes all reported adverse events involving medical devices. Notably, the FDA cautions against causal interpretation: because a device is involved in an adverse event, it does not mean that the device *caused* the adverse event. For example, the FDA's website explains:

> "The FDA reviews all medical device reports (MDRs) received. The FDA's analysis of MDRs evaluates the totality of information provided in the initial MDR as well as any MDR supplemental reports subsequently provided. The submission of an MDR itself is not evidence that the device caused or contributed to the adverse outcome or event. For example, in certain MDRs, the text of the report may include the word "death" or a related term. However, the MDR would not, and should not, be classified as death

[23] For the most recent list of AI/ML devices from the FDA see https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices
[24] Data downloaded on October 4, 2002 from https://www.fda.gov/medical-devices/510k-clearances/downloadable-510k-files
[25] https://www.fda.gov/advisory-committees/medical-devices/medical-devices-advisory-committee
[26] Data downloaded on March 2, 2022 from https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities/manufacturer-and-user-facility-device-experience-database-maude
[27] Manually downloaded and scraped recall database from FDA in May 2021 https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfres/res.cfm

unless the reporter believes the patient's cause of death was or may have been attributed to the device or the device was or may have been a factor in the death." [28]

Nevertheless, the MAUDE database is useful as a surveillance tool, and it provides regulators and researchers with a quantitative and qualitative overview of potential safety issues associated with devices.

When such safety issues are found to be systematic, the manufacturer may issue a recall, under the FDA's oversight, which would remove the device from the market (either indefinitely or until remedial action can be taken) due to a problem with a medical device that violates FDA law. The types of "correction or removal" actions that may constitute recalls include inspecting a device for problems, repairing a device, notifying patients of a problem with their device, re-labeling a device, adjusting settings on a device, monitoring patients for health issues, etc.[29] The FDA's recall database provides detailed information on the date and severity of a recall. The FDA clearly defines three classes of medical device recalls:

- **Class I:** A situation where there is a reasonable chance that a product will cause serious health problems or death.
- **Class II:** A situation where a product may cause a temporary or reversible health problem or where there is a slight chance that it will cause serious health problems or death.
- **Class III:** A situation where a product is not likely to cause any health problem or injury.

The FDA's recall database was manually downloaded and scraped, such that all medical device recalls are linked to their associated product(s) via its 510(k) number. This allows recalls and specific products to be directly linked.[30] Similarly, the MAUDE database includes a flag for the 510(k) number of the device associated with each adverse event report, allowing each adverse event to be linked to its respective product.

The next steps use text analysis to identify devices with a software component as well as those that incorporate AI (i.e., "AI devices" as defined above). Both exercises rely on the availability of machine readable, publicly available summary documents (as described in section 2). Among the 30,779 top-8 specialty devices cleared during our period of analysis, 30,294 (or 98.4%) had such documents available and these form the basis of the text analysis used to flag those of interest. We apply the algorithm described in Stern & Foroughi 2020 to identify all software devices (both SiMD and SaMD) and then perform a further keyword search to identify AI devices as a subset of those. The keywords specifically selected for this exercise are "artificial intelligence," "deep learning,"

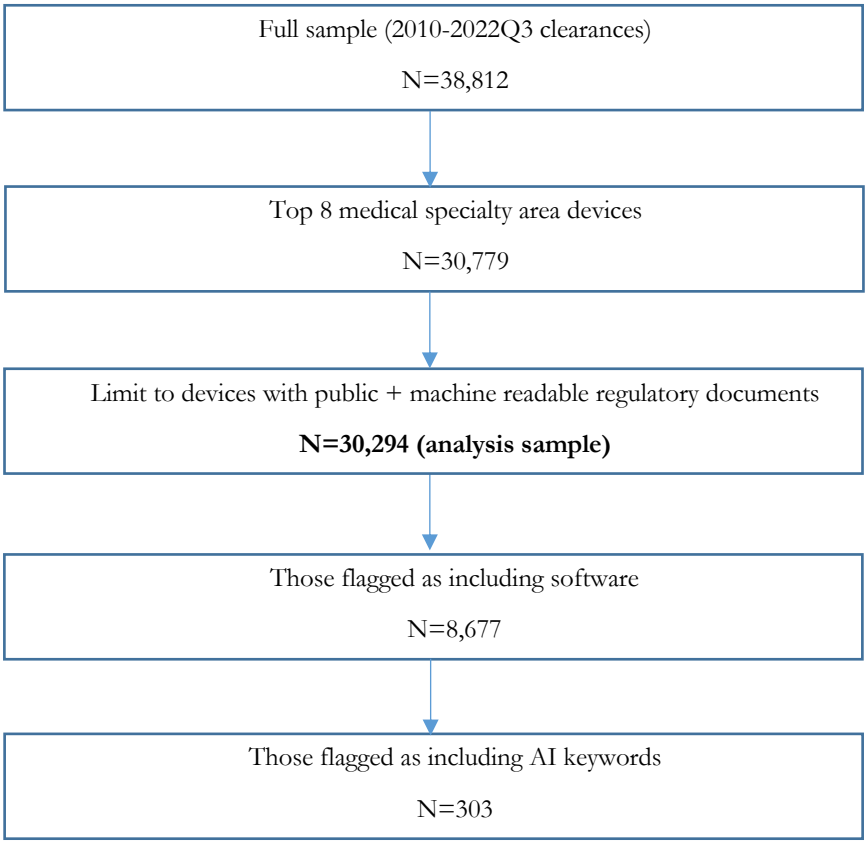
[28] https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities/about-manufacturer-and-user-facility-device-experience-maude

[29] https://www.fda.gov/medical-devices/medical-device-recalls/what-medical-device-recall

[30] The databases are merged in a "many-to-one" fashion, since an individual recall event can potentially impact more than one medical device. For example, a recall due to a safety issue with a material that is used in multiple devices would impact all devices that contain that material. Similarly, a recall impacting a piece of medical device software would impact all devices that run that software.

"machine learning," and "neural network." These terms were chosen for their direct and relationship with the description of AI algorithms and their likely lack of ambiguity when used, as such.[31] Manual inspection of a random sample of device summaries confirmed a 0% rate of false positives based on this method. Figure 1 presents a flow chart of how the analysis sample was constructed and Table 2 presents a breakdown of the analysis sample by medical specialty.

Figure 1: Analysis sample construction



[31] An initial version of the keyword list included the word "algorithm," however manual review suggested that it was being used in several cases where the device was collecting data that could be fed into an analysis program or used in a decision algorithm, where the product or method in question was not an AI tool. As such, the more conservative version of our keyword-based identification of AI devices does not incorporate the word "algorithm," however we continue to count its use among otherwise-identified AI devices, noting that there are virtually no "false positive" uses of the word conditional on it being used in the context of a product with other AI-related keywords in its description. For an excellent summary of deep learning techniques such as neural networks, see Hinton (2022).

Table 2: Analysis Sample Devices by Specialty

Specialty (FDA Advisory Committee Panels)	Analysis sample	All software	All AI software
Clinical Chemistry and Clinical Toxicology	1,408 (4.7%)	375 (4.3%)	1 (0.3%)
Cardiovascular	4,355 (14.4%)	1,515 (17.5%)	14 (4.6%)
Dental	2,928 (9.7%)	466 (5.4%)	0 (0.0%)
Gastroenterology-Urology	1,979 (6.5%)	428 (4.9%)	4 (1.3%)
General Hospital and Personal Use	3,796 (12.5%)	462 (5.3%)	0 (0.0%)
Orthopedic and Rehabilitation	6,893 (22.8%)	471 (5.4%)	0 (0.0%)
Radiology	4,752 (15.7%)	3,729 (43.0%)	283 (93.4%)
General and Plastic Surgery	4,183 (13.8%)	1,231 (14.2%)	1 (0.33%)
Total	**30,294 (100.0%)**	**8,677 (100.0%)**	**303 (100.0%)**

A few interesting findings emerge from reviewing summary statistics from the analysis sample (Table 3). First, it is notable that nearly 29% of the analysis sample devices included software. This is consistent with Stern and Foroughi (2020), who document significant digitization of the medical device industry with the highest rates of SiMD and SaMD seen in radiology devices, followed by cardiology devices. Having a digitized device that includes a software component is, of course, a necessary, but not sufficient condition for the incorporation of AI.

Table 3: Summary Statistics

	Analysis sample	All software	All AI software
Keyword-based flags (all binary)			
software	8,677 (28.6%)	8,677 (100.0%)	303 (100%)
algorithm	1,884 (6.2%)	1,724 (19.9%)	256 (85%)
artificial intelligence	118 (0.4%)	118 (1.4%)	118 (39%)
deep learning	132 (0.4%)	132 (1.5%)	132 (44%)
machine learning	122 (0.4%)	122 (1.4%)	122 (40%)
neural network	90 (0.3%)	90 (1.0%)	90 (30)
Device features (binary)			
Breakthrough Device Program	9 (0.03%)	4 (0.05%)	2 (0.7%)
De Novo	128 (0.4%)	83 (1.0%)	5 (1.7%)
Firm information			
US-based application	66.3%	57.7%	46.5%
Publicly listed, %	30.7%	39.1%	39.3%
Revenue in millions if public, mean	23,955	35,467	51,924
Employees if public, mean	75,013	118,476	163,436
R&D Spend if public, mean	1,633	2,083	3,219
Total number of devices	**30,294**	**8,677**	**303**

While the sample sizes are small, Table 3 also indicates that a) software devices and b) AI devices appear more likely to have received the Breakthrough Device Designation: while just 0.03% of all analysis sample devices received this designation, the rate rises to 0.05% among software devices and 0.7% among AI devices. The use of the De Novo pathway is also nearly two times more common among AI devices, a fact that is consistent with these devices being more likely to be novel and less likely to have a clear "predicate" product, although these summary statistics too are based on relatively small totals. Perhaps most interestingly, we note a significant difference in the likelihood that the manufacturer of a software or AI device is a publicly listed firm: the share of all AI devices brought to market by public firms is 39.3%, vs. 39.1% for software devices and 30.7% for all sample devices. That is, over 60% of AI and software devices (vs. roughly 70% of comparator devices from the same specialties) are being brought to market by privately held firms. This indicates that digital devices are being developed by a differentiated group of innovator firms, which may in turn have different needs and backgrounds.

Figure 2: Growth in Software and AI Devices

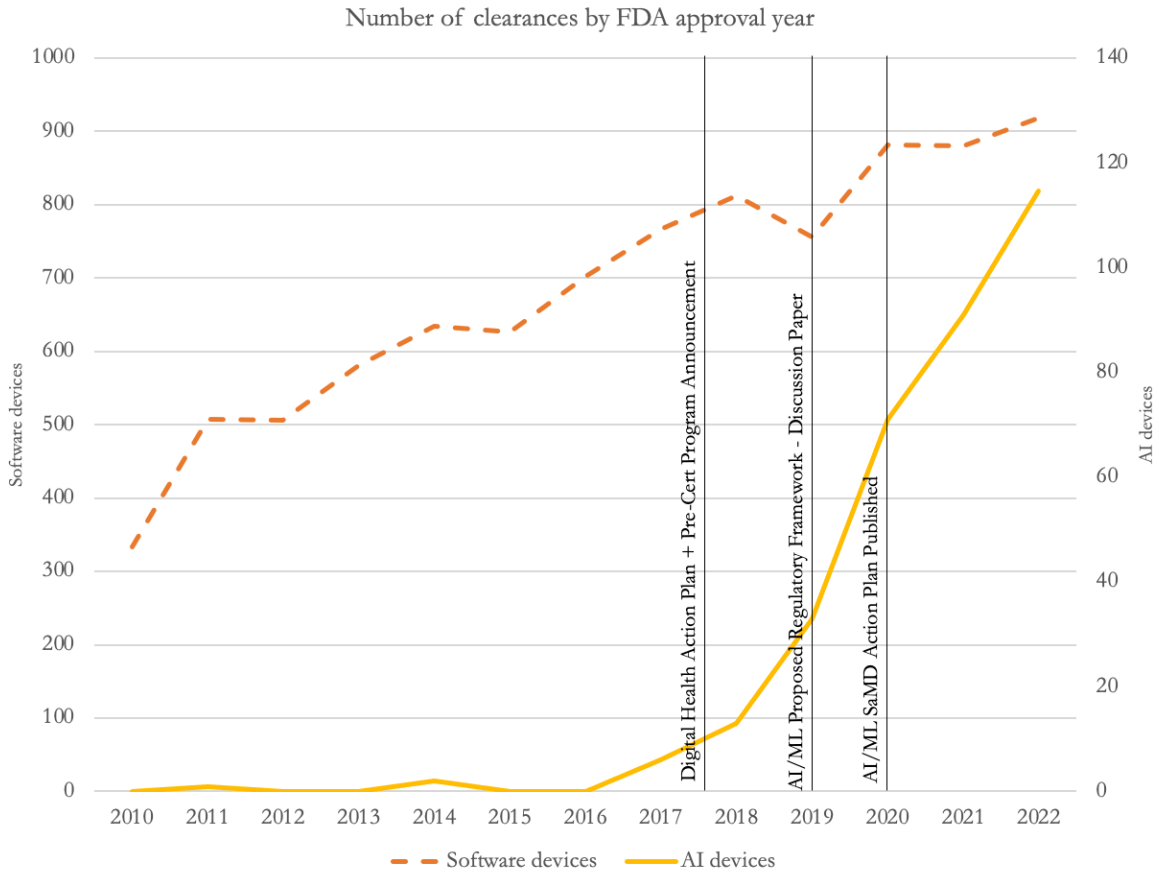


Figure 2 shows the growth in both software devices as well as AI devices over the sample period. Consistent with past work (Foroughi and Stern 2020), we observe significant, continued growth in software devices (left axis), with a more than doubling in the number of new devices cleared per year over the period of observation. AI device growth (right axis) shows even more dramatic growth (albeit off a much lower base), with annual clearances of no more than just a few devices per year through 2016 and then rising dramatically to 91 in just the first three quarters of 2021. (NB: in Figure 2, the totals for the datapoint used for 2022 were inflated but a factor of 4/3 to generate an annualized number).

Table 3: AI Devices by Country of Application

Country Code	Country	Count	% of AI Sample
US	United States	141	46.53
IL	Israel	33	10.89
JP	Japan	26	8.58
FR	France	15	4.95
CN	China	13	4.29
KR	South Korea	12	3.96
NL	Netherlands	10	3.3
SE	Sweden	7	2.31
CA	Canada	6	1.98
DE	Germany	6	1.98
GB	United Kingdom	6	1.98
AU	Australia	5	1.65
TW	Taiwan	5	1.65
FI	Finland	3	0.99
IN	India	3	0.99
SG	Singapore	3	0.99
BE	Belgium	2	0.66
IE	Ireland	2	0.66
AT	Austria	1	0.33
BG	Bulgaria	1	0.33
DK	Denmark	1	0.33
PT	Portugal	1	0.33
VN	Vietnam	1	0.33
Total		303	100

Because the 510(k) database includes a field for the applicant firm's addresses, we can also report on the address from which each application was submitted (NB: this is the address from which the firm submitted the application, not necessarily the same as the address of the headquarter of the manufacturer firm. However, in many cases, this may be a more accurate representation of where R&D took place. For example, a German medical device firm that has its software development team/division in the United States might submit an application from the address of that U.S. division). Table 3 reveals that AI devices are more likely to come from international applicants than either comparator group in the analysis sample. Overall, 66.3% of all sample devices submitted regulatory applications from a U.S. address, relative to 57.7% of software devices and just 46.5% of AI devices. A natural question to ask is: which other countries are developing AI devices? Table 3 presents the distribution of the nationalities of AI device submissions. Notably, while the United States is the most represented country in the sample, U.S. applications represent less than half (46.5%) of the AI products identified, highlighting the importance of other countries such as Israel, Japan, France, and China in the AI device development ecosystem.

A final set of descriptive statistics explores the safety profiles of our sample devices. As described above, we consider both adverse events and recalls as relevant safety outcomes. We consider the first two years after a product is cleared for marketing in order to capture the most relevant period of time after a new product's launch and to ensure that adverse event and recall outcomes are comparable across older vs. newer devices.

Table 4: Device Safety Outcomes

Panel I: Safety Outcomes (All years)	**Analysis sample**	**All software**	**All AI software**
Adverse events (AE)			
Any AE in the first two years	4,212 (13.9%)	1,047 (12.1%)	9 (3.0%)
Any mandatory AE report in the first two years	4,206 (13.9%)	1,044 (12.0%)	9 (3.0%)
Average # of AEs in the first two years (count)	15.861	46.250	0.046
Average # of Mandatory AEs in the first two years (count)	15.855	46.244	0.046
Recalls			
Any recall in the first two years	1,158 (3.8%)	523 (6.0%)	5 (1.7%)
…categorized as Class I	39 (0.1%)	14 (0.2%)	0 (0.0%)
…categorized as Class II	1,103 (3.6%)	509 (5.9%)	5 (1.7%)
…categorized as Class III	29 (0.1%)	4 (0.1%)	0 (0.0%)
Average # Recalls in the first two years (count)	0.052	0.092	0.046
…categorized as Class I	0.001	0.002	0.000
…categorized as Class II	0.050	0.090	0.046
…categorized as Class III	0.001	0.000	0.000
Total number of devices	**30,294 (100.0%)**	**8,677 (100.0%)**	**303 (100.0%)**

Panel II: Safety Outcomes (2010-2020 clearances only)	Analysis sample	All software	All AI software
Adverse events (AE)			
Any AE in the first two years	4,121 (15.9%)	1,011 (14.2%)	9 (7.1%)
Any mandatory AE report in the first two years	4,115 (15.9%)	1,008 (14.2%)	9 (7.1%)
Average # of AEs in the first two years (count)	18.454	56.388	0.111
Average # of Mandatory AEs in the first two years (count)	18.447	56.381	0.111
Recalls			
Any recall in the first two years	1,154 (4.4%)	519 (7.3%)	5 (4.0%)
…categorized as Class I	39 (0.1%)	14 (0.2%)	0 (0.0%)
…categorized as Class II	1,099 (4.2%)	505 (7.1%)	5 (4.0%)
…categorized as Class III	29 (0.1%)	4 (0.1%)	0 (0.0%)
Average # Recalls in the first two years (count)	0.061	0.112	0.111
…categorized as Class I	0.002	0.002	0.000
…categorized as Class II	0.058	0.109	0.111
…categorized as Class III	0.001	0.001	0.000
Total number of devices	**25,962 (100.0%)**	**7,108 (100.0%)**	**126 (100.0%)**

Panel I of Table 4 presents outcomes for the entire dataset (which is truncated due to the inability to observe two years of follow-up data for those devices approved during or after Q3 of 2020). Panel II includes only safety outcomes for devices cleared through 2020 (but includes outcome data through Q3 of 2022) for a subsample with a balanced follow-up period. The obvious tradeoff in building the balanced subsample is the quantity of data: since over half of AI devices in the sample were cleared in 2021 and 2022, the total number of AI devices considered (right-most column) drops significantly from Panel I to Panel II.

The safety outcome data reveal an interesting story: adverse events appear similarly likely to be recorded when comparing software devices to the full sample of devices, however the likelihood of an AI device having any adverse event reports drops dramatically in both samples. When considering *counts* of adverse events, both samples suggest a roughly 3x increase in adverse events among software devices relative to the overall sample, but a more than 100x decrease in adverse event counts per device among AI devices, again hinting at their relative safety.

Similarly, data on recalls (a thankfully much rarer outcome) suggest that software devices are more likely to experience recalls (and when they do, they experience more of them) relative to the full sample, however AI devices are either modestly or significantly less likely to experience recalls. While more work is needed to understand the drivers of adverse event reports and medical device recalls, the findings here paint a preliminarily sanguine picture of AI device safety: on average, AI devices appear to be at least as good or better than comparable

samples of devices in terms of the likelihoods with which negative safety outcomes are observed. These findings are consistent with Everhart and Stern (2022), who find somewhat higher rates of adverse events and recalls among software devices in many settings, however these results suggest an added layer of nuance in that the subset of AI-driven devices identified here are actually less likely to experience adverse safety-related outcomes.

4. INNOVATION INCENTIVES

How should innovation policy researchers think about the intersection of medical AI and medical product regulation and what can be learned from early data on regulated AI in the United States? A clear starting point for answering the first questions is the FDA's current position on the regulation of AI/ML products, which was articulated in early 2021 in the "Artificial Intelligence and Machine Learning (AI/ML) Software as a Medical Device Action Plan" published by the CDRH's Digital Health Center of Excellence (FDA 2021) and the "Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback" published previously (FDA 2019). Ideas in both documents rely on the risk categorization principles outlined by the IMDRF[32] and are based on the "total lifecycle approach" envisioned by the software Pre-Cert pilot program (see Section 2 for more detail).

A key part of the proposed framework for the future regulation of AI/ML devices is the "algorithm change protocol," which would rely on real-world performance data and ongoing monitoring to allow regulators to flexibly balance the FDA's dual mandate to protect public health while still providing timely access to new products. More generally, the Action Plan outlines five key actions that it sees as the focus of FDA's work going forward:

1. *Tailored Regulatory Framework for AI/ML-based SaMD*
2. *FDA to encourage harmonization of Good Machine Learning Practice (GMLP) development (an analog to GMP in physical devices)*
3. *Patient-Centered Approach Incorporating Transparency to Users*
4. *Regulatory Science Methods Related to Algorithm Bias and Robustness*
5. *Real-World Performance*

The fact that the Action Plan exists and that its first articulated goal is the establishment of a tailored regulatory framework for AI/ML-based SaMD send a clear signal that U.S. regulators are aware of the special challenges (and opportunities) inherent in the regulation of AI-based SaMD. And yet great uncertainty remains as to how precisely AI device regulation will be implemented in practice—and, as is always the case—the devil will be in the details. As just one example, understanding whether casual inference will play a role in the FDA's regulation of

[32] https://www.fda.gov/medical-devices/software-medical-device-samd/global-approach-software-medical-device

certain types of products will meaningfully shape the types of products brought to market—both within the regulated setting and beyond (Stern and Price 2019).

A laundry list of other questions that flow from the Action Plan's component sections remain unanswered: For example, what will be the scope and specificity of the tailored regulatory framework that emerges? CDRH plans to encourage the harmonization[33] of GMLPs, which are expected to include best practices for data management, feature extraction, training, interpretability, evaluation, and documentation that FDA acknowledges as being "akin to good software engineering practices or quality system practices" (FDA 2021)—how detailed and how burdensome (two related but ultimately distinct features) will these GMLPs be? How closely aligned will they be with the state-of-the-art in algorithm development? The answers to both questions—in particular, the latter—will meaningfully shape the barriers to entry for developers to enter the regulated SaMD space.

In addition to active participation in the IMDRF's AI Medical Device working group, the FDA maintains relationships with the Institute of Electrical and Electronics Engineers' (IEEE) AI medical device working group as well as the International Organization for Standardization/Joint Technical Committee's Sub-Committee on AI and the British Standards Institution's initiative on AI in medical technology (FDA 2021).[34] Such relationships are deeply important: (internationally recognized) standard setting organizations are known to be vital for identifying the most promising technologies and influencing the trajectory of technology adoption in other contexts (Rysman and Simcoe 2008) and regulatory clarity is known to accelerate time-to-market for other types of medical devices (Stern 2017).

With respect to patient-centricity in incorporating transparency for users, regulatory science methods to related to algorithm bias and robustness, and the creative, but still-rigorous application of real-world performance data, countless questions of a similar nature remain to be articulated by stakeholders and answered by regulators. A challenge with developing AI is often a lack of transparency as to how AI/ML algorithms such as neural networks performs their tasks; at present, it's simply not possible (Hinton 2022) . Nonetheless, there needs to be a mechanism to build trust in the outcomes via specific validation tests, which is where "real-world performance" studies can play an important role. On all of these topics, the broader digital medicine community can also collaborate to move medical AI forward; for example, in articulating best practices and priorities for the generation and use of real-world evidence in the assessment of digital health products (Stern et al. 2022). This work will necessarily touch on patient-centricity, transparency, and various aspects of regulatory science. For example, the

[33] FDA is also moving to harmonize regulatory guidelines with international standards for quality systems such as ISO 13485, which covers quality management systems for medical devices and requirements for regulatory purposes (https://www.iso.org/standard/59752.html).
[34] In 2021, the FDA officially joined the Xavier AI World Consortium Collaborative Community, the Pathology Innovation Collaborative Community, and also participates in the Collaborative Community on Ophthalmic Imaging (FDA 2021)

articulation of best practices for real-world evidence generation will lead to greater consensus on questions of how to handle and understand the implications of missing data, characterizing the generalizability and transportability of findings to broad populations, and understanding and standardizing hypothesis testing around whether digital health products are complements or substitutes to existing standards of care, to name just a few aspects. Finally, other questions that are not covered in the scope of the Action Plan also surface quickly: what about the large number of AI devices (as defined here) that are SiMD, rather than SaMD products? The scope of the Action Plan is explicitly limited to stand-alone software, with the corollary that a good deal of uncertainty may remain for AI SiMD products.

Against this backdrop, innovation policy questions associated with the concept of *regulatory uncertainty* become all the more trenchant: in the U.S. medical device setting, regulatory uncertainty—e.g., as experienced when a first-of-its-kind medical device goes through a regulatory approval process for the first time—has been associated with first-mover disadvantages and lower rates of novel device commercialization among small firms (Stern 2017). There is every reason to believe that regulatory uncertainty will negatively impact firms' willingness to engage in innovation in the regulated space broadly and to embark on more novel R&D projects (specifically, those without regulatory precedent) in particular. But this also indicates the great potential that proactive approaches could have in this setting: the provision of regulatory clarity through formal guidance documents has been shown to speed regulatory approval of new high-risk devices (Stern 2017), suggesting the vital role that regulatory innovation and policy clarification can play in shaping innovation incentives.

The value of regulatory innovation and regulatory clarity may be particularly important in the context of AI devices because such a large share of innovations to-date have emerged from smaller firms and those from other countries. Large (e.g., publicly listed) and domestic firms are far more likely to possess U.S. regulatory expertise in the first place and these firms also have, on average, more experience in the (U.S.) regulated product space than their smaller and/or internationally based peers. Yet evidence on the early years of AI device rollout in the United States suggests that privately held, international manufacturers are disproportionately likely to be developing AI devices relative to what is otherwise seen in the regulated space.

Taking an even larger step back, it is worth considering the margin between regulated and unregulated AI tools – that is, asking the "extensive margin question": what role does regulation *itself* play in innovation incentives for new AI devices. As is often the case, the answer is "it depends": in some situations, pursuing a non-regulated product development approach may be incentivized – e.g., to get a product to market more quickly. In other cases, however, having a *regulated* product on the market may itself be part of the commercialization and/or reimbursement strategy for a digital tool.

The gray area in which many devices reside and in which many manufacturers operate is substantial and there may be strategic reasons to pursue FDA regulation or to avoid it. One reason to pursue a strategy of formal regulation is that reimbursement by payers may be more straightforward for regulated medical products, increasing manufacturers' appetite for going through a clearance process, since it may be a significant component of their reimbursement strategy. In the extreme, this is already being seen in the case of "prescription digital therapeutics" (sometimes called PDTs), which are software-based (SaMD) therapeutic devices that are intended only for prescription by clinicians. Notably, the label "PDT" is one that was created by the SaMD industry – unlike drugs, the word "prescription" does not (necessarily) indicate anything about the product's risk level; rather the term is often used as part of a market access strategy for the manufacturer, which in turn is making a bet that products prescribed by physicians will be more likely to qualify for insurance coverage than direct-to-consumer apps/tools. Indeed, anecdotes suggest that other types of digital health companies are pursuing an FDA regulatory pathway in order to gain legitimacy in the eyes of clinician recommenders and health insurance companies. Such behavior has precedent: Eisenberg (2019) describes how a situation of "opting into device regulation" has been observed among developers of next generation sequencing (NGS) diagnostic tests for tumor DNA, where "understanding the rules and practices that govern health insurance coverage and the important role of FDA in assessment of new technologies" is key to understanding such a decision by a manufacturer.

On the other hand, manufacturers may deliberately select language so as to *avoid* making medical claims that would require premarket review in order to get products to market faster or at a lower cost, since FDA submissions are notoriously costly to file and take several months to years. Relatedly, many SaMD products will meet the definition of a medical device, but qualify for enforcement discretion (e.g., because they pose only a very low risk to patients), such that they are, *de facto*, not regulated by the FDA. A final set of innovations are likely to emerge outside of the regulated space, but may have spillover effects on the types of other AI products that can be commercialized. For example, Price, Sachs, and Eisenberg (2021) point out that AI tools such as those that might be used for quality improvement initiatives or for optimizing use of scarce hospital facilities are likely to be developed by health systems and insurers (rather than machine learning experts and/or medical device companies), since these parties already control the relevant training datasets and have the greatest operational interest in such products' development. Importantly, the authors note that "data possession and control play a larger role in determining capacity to innovate in this [more operationally-oriented] space" (Price, Sachs, and Eisenberg 2021) In certain cases, of course, there will not be any gray area: very high (or very low) risk products will (or will not) unambiguously qualify as regulated medical devices. and these realities, in turn, could impact the availability of datasets for AI development in other settings.

Notably, FDA does not actively regulate two significant types of AI systems that are likely to play a growing role in care delivery as well as in AI R&D (Stern and Price 2019). First, the FDA does not regulate certain types of clinical decision support software (CDS), software that "helps providers make care decisions by, for instance, providing suggested drug dosages or alerts about drug interactions." When providers have an opportunity to review the rationale behind the recommendation made by an AI, it is likely to be exempt from what FDA considers a device (21 U.S. Code § 360j). However recent regulatory guidance has served to significantly narrow the set of applications that qualify as CDS[35] and as a corollary, many more applications are likely to be considered devices. Further, some AI tools will be considered "laboratory-developed tests" which are those that are developed and used within a single health care facility such as hospital. For such tests, the FDA also holds back from exercising its regulatory authority. In both cases, we should expect to see notable growth in AI, although use cases with cost and/or comparative effectiveness have yet to be established.

5. CONCLUSION

A survey of the regulated medical device landscape suggests a dramatic uptick in the commercialization of AI products over recent years. At the same time, regulators have begun a germane and important discussion of how such devices could be regulated constructively in the future—ultimately under the umbrella of the FDA's dual mandate of ensuring public health, while facilitating patient access to important new medical technologies. There are several margins along which innovation incentives are likely to play a role. To the extent that many applications of AI for diagnosis and treatment of medical conditions are likely to meet the formal definition of a medical device, medical product regulation will be an ongoing presence in this space—and one that will shape incentives for software developers, established medical device companies, venture capital investors, and user-innovators.

Whether or not an AI tool qualifies as a medical device and the extent to which the answer occupies the gray area of ambiguity on this question will impact whether products are developed, what R&D decisions are made at the margin, and how new AI tools are paid for in the health care system once they are commercialized. Among regulated devices, meaningful differences in the burden of clinical evidence and paperwork required for commercialization already exist, depending on whether a device is classified as being of low, moderate, or high risk. These differences will be even more salient for AI developers, who are likely to have less medical device regulatory experience than those engaging in R&D activities in traditional medical technology firms. And of course, emergent guidelines from the FDA and best practices from the broader digital medicine community (including clinical

[35] https://www.fda.gov/media/109618/download

researchers, standard setting organizations, and companies) will impact the amount of uncertainty associated with new product development in this space and therefore the incentives for firms—both small and large—to innovate.

Such guidelines and best practices will of course emerge in the context of a larger societal conversation about AI in health care and must *necessarily* be considered against the backdrop of key issues such as those highlighted by the U.S. National Academy of Medicine (Matheny et al. 2019). These include the imperative of promoting population-representative data with accessibility, standardization, and quality; prioritizing ethical, equitable, and inclusive health care AI, while addressing explicit and implicit bias; contextualizing the dialogue of transparency and trust according to differential needs; cultivating a near-term focus on augmented intelligence vs AI autonomous agents; developing and deploying appropriate training and educational programs; leveraging frameworks and best practices for learning health care systems, human factors, and implementation science; and balancing innovation with safety via regulation and legislation to promote trust (Matheny et al. 2020)—a final priority that brings the discussion back to the role of regulatory and legal aspects of medical AI.

The emergence of regulated AI devices holds great promise for delivering higher-quality care to patients and addressing un- or under-served populations. It also can improve workflows, efficiency, and confidence in diagnosis and treatment decisions for clinicians, with secondary benefits to health care payers and clear financial benefits to product companies. Yet ongoing regulatory clarity and policy innovation will be necessary for regulation to keep pace with AI innovation in health care. Further research on innovation policy in medical AI should focus on understanding which aspects of regulatory clarity and regulatory policy are most likely to induce and facilitate the commercialization of welfare-enhancing innovations.

REFERENCES

Alon, Noy, Ariel Dora Stern, and John Torous. "Assessing the Food and drug administration's risk-based framework for software precertification with top health apps in the United States: quality improvement study." *JMIR mHealth and uHealth* 8, no. 10 (2020): e20482.

Benjamens, Stan, Pranavsingh Dhunnoo, and Bertalan Meskó. "The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database." *NPJ digital medicine* 3, no. 1 (2020): 1-8.

Brönneke, Jan Benedikt, Jennifer Müller, Konstantinos Mouratis, Julia Hagen, and Ariel Dora Stern. "Regulatory, legal, and market aspects of smart wearables for cardiac monitoring." *Sensors* 21, no. 14 (2021): 4937.

Eisenberg, Rebecca S. "Opting into Device Regulation in the Face of Uncertain Patentability." Marq. Intell. Prop. L. Rev. 23 (2019): 1.

Everhart, Alexander O. and Ariel D. Stern. "Post-Market Surveillance of Software Medical Devices: Evidence from Regulatory Data." (Forthcoming in Diagnosing in the Home: The Ethical, Legal, and Regulatory Challenges and Opportunities of Digital Home Health, Cambridge University Press)

Felber, Sarith and Dawid Maciorowski. "Technology Primers: Smart Wearables and Health." Working paper, Belfer Center for Science and International Affairs, Harvard Kennedy School, Spring 2023.

Foroughi, Cirrus, and Ariel Dora Stern. "Who Drives Digital Innovation?: Evidence from the US Medical Device Industry." Harvard Business School Working Paper, 2019.

Hinton, Geoffrey. "Deep learning—a technology with the potential to transform health care." *JAMA* 320, no. 11 (2018): 1101-1102.

International Medical Device Regulators Forum (IMDRF). "Software as a Medical Device (SaMD): Key Definitions." IMDRF SaMD Working Group, (2013).

Matheny, Michael, S. Thadaney Israni, Mahnoor Ahmed, and Danielle Whicher. "Artificial intelligence in health care: The hope, the hype, the promise, the peril." Washington, DC: National Academy of Medicine (2019).

Matheny, Michael E., Danielle Whicher, and Sonoo Thadaney Israni. "Artificial intelligence in health care: a report from the National Academy of Medicine." *JAMA* 323, no. 6 (2020): 509-510.

Price, W., I. I. Nicholson, Rachel E. Sachs, and Rebecca S. Eisenberg. "New Innovation Models in Medical AI." Wash. UL Rev. 99 (2021): 1121.

Rysman, Marc, and Timothy Simcoe. "Patents and the performance of voluntary standard-setting organizations." *Management science* 54, no. 11 (2008): 1920-1934.

Sanders, Samantha F., Mats Terwiesch, William J. Gordon, and Ariel D. Stern. "How artificial intelligence is changing health care delivery." *NEJM Catalyst* 5, no. 5 (2019).

Stern, Ariel Dora. "Innovation under regulatory uncertainty: Evidence from medical technology." Journal of public economics 145 (2017): 181-200.

Stern, Ariel Dora, and W. Nicholson Price. "Regulatory oversight, causal inference, and safe and effective health care machine learning." *Biostatistics* 21, no. 2 (2020): 363-367.

Stern, Ariel D., Jan Brönneke, Jörg F. Debatin, Julia Hagen, Henrik Matthies, Smit Patel, Ieuan Clay et al. "Advancing digital health applications: priorities for innovation in real-world evidence generation." *The Lancet Digital Health* 4, no. 3 (2022): e200-e206.

Torous, John, Ariel D. Stern, and Florence T. Bourgeois. "Regulatory considerations to keep pace with innovation in digital health products." *npj Digital Medicine* 5, no. 1 (2022): 1-4.

U.S. Food and Drug Administration. "Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback" Digital Health Center of Excellence, CDRH (2019).

U.S. Food and Drug Administration. "Artificial Intelligence and Machine Learning (AI/ML) Software as a Medical Device Action Plan." Digital Health Center of Excellence, CDRH (2021).