THE REGULATION OF MEDICAL AI:

POLICY APPROACHES, DATA, AND INNOVATION INCENTIVES

Ariel Dora Stern†

September 2022

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 applications of AI tools already used 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 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 presents background on medical device regulation (especially as it relates to software products), quantitatively describes the emergence of AI among FDA-regulated devices, and discusses the implications of regulation for innovation incentives in medical AI.

* The author is grateful to Melissa Ouellet for excellent research and programming support.

† Harvard Business School & Harvard-MIT Center for Regulatory Science (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, with the overwhelming majority of applications falling into one of three broad categories: (1) administrative work, (2) diagnosis, and (3) treatment. Table 1 provides examples of the types of technologies, tools, and products that would fit into each of these categories.

Table 1: Applications of AI to health care delivery with 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: 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, administrative support tools will 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. 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. Building on descriptive statistics from the US 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 the forward-looking research agenda.

2. Background and Policy Approaches

Medical device regulation in the United States

[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).

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 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 510(k) applications.
- Finally, devices of the highest risk (Class III) are devices that are implantable and/or life sustaining and, as such, require significant evidence of safety and effectiveness to be approved for marketing.

[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)

> 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 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 Q2, 2022 a total of 54 devices with the Breakthrough Device Designation have 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 regulatory 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, a number of public documents are published online at the time that a device receives marketing approval. These include a device "Summary" for 510(k)-track devices, which "includes a description of the device such as might 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 a good source of information on a device's key technological characteristics. The PMA process

[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

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 device functionality at scale.

Regulation of Software-Driven Medical Devices

Because US medical device regulations are grounded in legislation from 1976, provisions for thinking about the regulation of software-driven products were not codified at the time of the MDA. 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 device 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 directly with consumer technology settings, where software programs are regularly, if not constantly, being improved upon and modified by developers. Nevertheless, a strict interpretation of 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 earlier for the addressing of safety and security issues associated with medical device software, it was not until 2017 that 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 driven by software that is inextricable from the device's

[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
[12] https://www.fda.gov/media/99785/download

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 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, standalone 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] 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 is the FDA's Digital Health Software Pre-Certification Program (Pre-Cert), which at the time of writing is still in its pilot phase. The pilot phase of the program is intended to "help inform the development of a future regulatory model that will provide more streamlined and efficient regulatory oversight of software-based medical devices" and includes 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.[14] While the program holds the promise of being more dynamic—an approach that surely makes sense for medical software—it is still in its nascency and observers have identified a number of concerns with the program, ranging from disagreement with the approach of not requiring individual product review to more practical concerns around implementation. For example, researchers have 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).

Medical Device Regulation in the European Union

[Optional additional section to be inserted here: background about EU Medical Device Regulation, including new MDR, risk-based classification, and AI in the MDR. This section could also potentially discuss the DiGA Fast-Track in Germany as an example of how regulators are approaching a more dynamic approach to SaMD regulation in Germany. The DiGA example is especially interesting at present because other European countries

[13] https://www.fda.gov/media/80958/download
[14] https://www.fda.gov/media/106331/download

are currently considering rolling-out a version of the Fast-Track domestically and an EU-level working group has already been convened to discuss the harmonization of app regulation across EU member countries.]

3. DATA

The empirical section of this chapter explores regulated, software-based, AI-supported & AI-driven medical devices (henceforward "AI devices")[15] 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 and only studies to survey FDA-regulated AI devices to date is Benjamens et al. 2020, which, in its own words published "the first comprehensive and open access database of strictly AI/ML-based medical technologies that have been approved by the FDA." Along with the initial publication, the authors commit that their database, hosted by The Medical Futurist Institute (TMF), [16] would be "constantly updated" (Benjamens et al. 2020). As of June 2022, the TMF database included 79 devices, updated from the 64 devices that had been included at the time of its initial publication. [17]

While the data used in this chapter were collected independently, the TMF database provides some useful information for informing how to identify AI devices and which medical specialties are likely to be most relevant for AI applications at present. For example, just two of the now 79 devices in the database were regulated through the PMA process for devices of the highest risk, with the remainder (77) having been 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 77 devices in the TMF database, 80.5% 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) devies regulated via the 510(k) pathway and b) the regulatory medical specialties for 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. The data assembly strategy and empirical analyses that follow are based on this data-informed approach.

[15] 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).

[16] The database is published online at https://medicalfuturist.com/fda-approved-ai-based-algorithms. At the time of writing, the most recent entry represented was a device approved in June, 2021.

[17] List pulled on June 7, 2022 from https://medicalfuturist.com/fda-approved-ai-based-algorithms/

As the basis for our analysis, we download the full 510(k) database for the years 2010 through 2021,[18] for a total of 36,452 unique device clearances over 12 full calendar years. We merge onto this database two additional sources of information on device outcomes: (1) the FDA's recall database[19] and (2) the FDA's medical device adverse event reporting database (the Manufacturer and User Facility Device Experience or MAUDE database).[20] 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 new devices clearances to be directly linked.[21] 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. We further flag devices that were part of the "Breakthrough Devices Program" (see section 2 for more detail).

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 18 FDA Medical Device Advisory Committees[22]) associated with the device. Given what is known about the distribution of devices in the TMF database, we would expect cardiology and radiology devices to represent roughly 80% of all AI devices. Further, these two regulatory medical specialties are of outsize importance for studying medical device innovation more generally: despite making up just 2 of the FDA's 18 Medical Device Advisory Committee panels, they constitute roughly a quarter of all medical device clearances during our period of observation (n=9,003). As such, we focus on just these two medical specialty areas in the analysis that follows.

The next steps involve using 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 and publicly available summary documents (see section 2 for more detail). Among the 9,003 cardiology and radiology devices cleared during our period of analysis, 8,834 (or 98.1%) had such documents available and these documents form the basis of the text analysis used to flag devices of interest. We apply the algorithm described in Stern & Foroughi 2020 to identify software devices (both SiMD and SaMD, as described in section 2 above) and then perform a further keyword search to identify AI devices. The keywords specifically selected for this exercise are "artificial intelligence," "deep learning," "machine learning," and "neural network." These terms were

[18] Data pulled on February 26, 2002 from https://www.fda.gov/medical-devices/510k-clearances/downloadable-510k-files

[19] Manually downloaded and scraped recall database from FDA in May 2021 https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfres/res.cfm

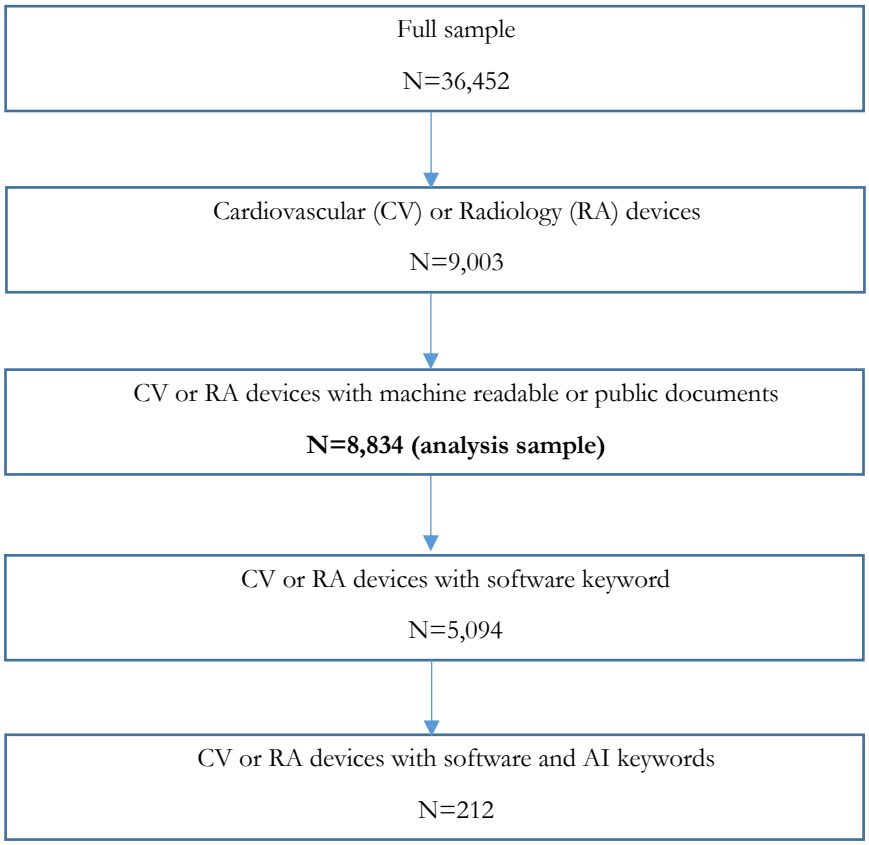
[20] Data pulled on March 2, 2022 from https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities/manufacturer-and-user-facility-device-experience-database-maude

[21] 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.

[22] https://www.fda.gov/advisory-committees/medical-devices/medical-devices-advisory-committee

chosen for their direct and relationship with the description of AI algorithms and their likely lack of ambiguity when used, as such.[23] Figure 1 presents a flow chart of how the analysis sample was constructed.

Figure 1: Analysis sample construction



A few interesting take-aways emerge by looking at the summary statistics from the sample of 8,834 cardiovascular and radiology devices with public summary documents (Table 2). First, it is notable that 57% of all cardiology and radiology devices cleared from 2010-2021 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.

While the sample sizes are quite small, Table 2 also indicates that a) software and b) AI devices appear more likely to have received the Breakthrough Device Designation: while just 0.01% of all cardiology and radiology

[23] 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.

devices received this designation, the rate rises to 0.3% among software devices and 0.9% among AI devices. The use of the De Novo pathway is also roughly 3 times more common among AI devices, a fact that resonates with these devices being more likely to be novel and not 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 an AI device is a publicly-listed firm: the share of all AI devices brought to market by public firms is just 26%, vs. 39% for software devices and 38% for all cardiovascular and radiology devices. That is, 74% of AI devices (vs. roughly 61-62% of comparator devices from the same specialties – both with and without conditioning on SiMD/SaMD profiles) are being brought to market by privately-held firms. This indicates that AI devices are being developed by a differentiated group of innovator firms, which may in turn have different needs and backgrounds.

Table 2: Summary Statistics

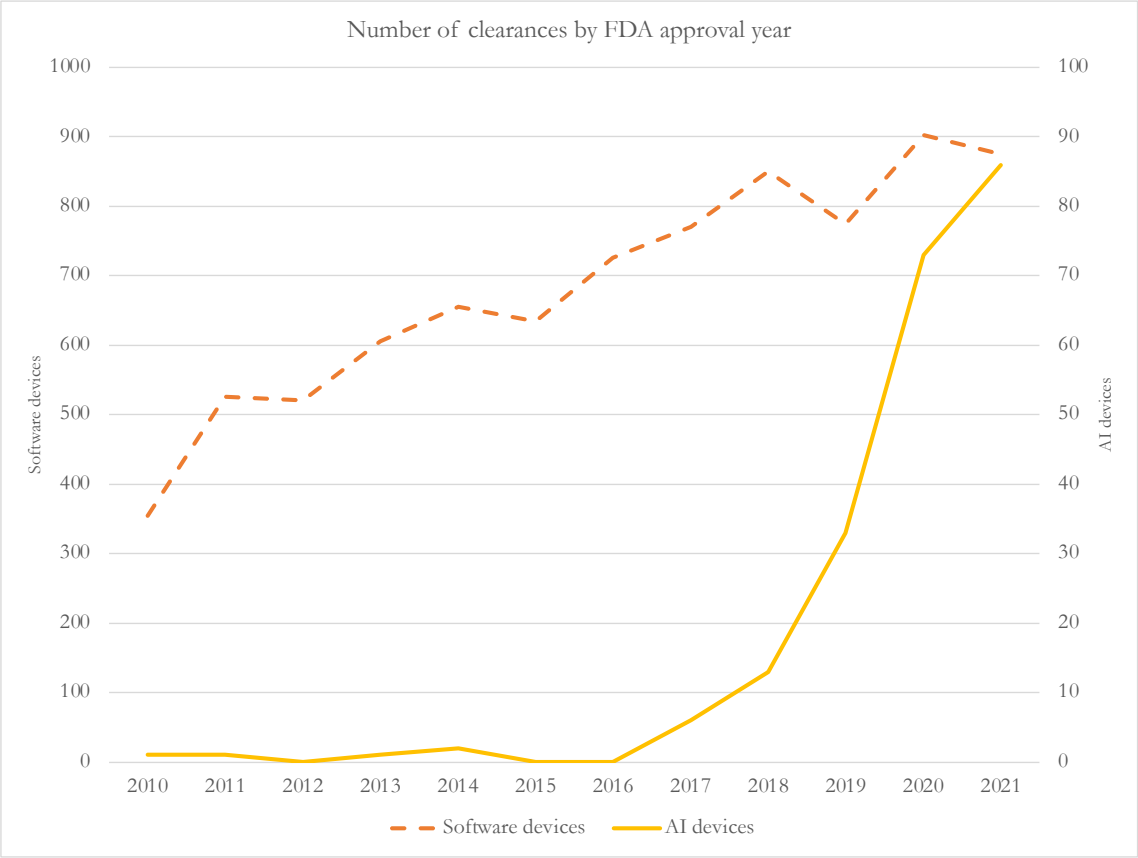
	All CV and RA in analysis sample	**All software CV and RA**	**All AI software CV and RA**
Keyword-based flags (all binary)			
software	5,094 (57%)	5,094 (100%)	212 (100%)
algorithm	1,456 (16%)	1,356 (27%)	176 (83%)
artificial intelligence	78 (0.9%)	78 (1.5%)	78 (37%)
deep learning	87 (1.0%)	87 (1.7%)	87 (41%)
machine learning	79 (0.9%)	79 (1.6%)	79 (37%)
neural network	64 (0.7%)	64 (1.3%)	64 (30%)
Device features (all binary)			
Included in TMF Database	62 (0.7%)	60 (1.2%)	29 (14%)
Breakthrough Device Program	3 (0.03%)	3 (0.1%)	2 (0.9%)
De Novo	32 (0.4%)	27 (0.5%)	3 (1.4%)
Firm information			
US-based application	64% (5,653/8,834)	55% (2,823/5,094)	46% (97/212)
Publicly listed*, %	38% (3,051/8,112)	39% (1,803/4,571)	26% (32/124)
Revenue in millions if public, mean	26,185	36,397	37,471
Employees if public, mean	83,392	116,328	171,319
R&D Spend if public, mean	438,684	680,070	63,480
Total number of devices	**8,834**	**5,094**	**212**

Notes: *Financial information is only available through 2020, i.e., not populated for devices cleared in 2021.

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 just 0-2 devices per year through 2016 and then rising dramatically to 86 by 2021.

Figure 2: Growth in Software and AI Devices



Because the 510(k) database includes applicant firm 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 US might submit an application from the address of that US division). Table 2 reveals that AI devices are more likely to come from international applicants than either comparator group in the analysis sample. 64% of all sample devices submitted regulatory applications from a US address, however that number drops to 55% for software devices and further to 46% for 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 US is the most represented country in the sample,

US applications represent less than half (46%) of the products identified, highlighting the importance of other countries such as China, Israel, and Japan in the AI device development ecosystem – even among devices that have gone through regulatory clearance in the United States.

Table 3: AI Devices by Country of Application

Coutry Code	Country	Count	% of AI Sample
AT	Austira	1	0.47
AU	Australia	4	1.89
BE	Belgium	1	0.47
CA	Canada	6	2.83
CN	China	12	5.66
DE	Germany	4	1.89
DK	Denmark	1	0.47
FI	Finland	1	0.47
FR	France	10	4.72
GB	United Kingdom	6	2.83
IL	Israel	23	10.9
IN	India	3	1.42
JP	Japan	19	8.96
KR	South Korea	8	3.77
NL	Netherlands	7	3.3
SE	Sweden	4	1.89
SG	Singapore	3	1.42
TW	Taiwan	2	0.94
US	United States	97	45.6
Total		212	100

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[24] 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 are facts that send a clear signal that US 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. Understanding whether casual inference will play a role in the FDA's regulation of 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 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 IA and the British Standards

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

Institution's initiative on AI in medical technology (FDA 2021).[25] Such relationships are deeply important: (internationally recognized) standards are known to be vital for incentivizing innovation in other contexts [Cites: Simcoe, etc.] 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. Here the broader digital medicine community can play an important role; 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 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 SiMDs.

Against the backdrop of so many open questions, innovation policy questions associated with the concept of *regulatory uncertainty* become all the more trenchant: in the US medical device setting, regulatory uncertainty—e.g., as emerges when a first-of-its-kind medical device goes through regulation 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 novel R&D projects (specifically, those without regulatory precedent) in particular.

But all of this also indicates room for proactive approaches: 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.

[25] 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)

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 are emerging from smaller firms and from other countries. Large (e.g., publicly listed) and domestic firms are far more likely to possess regulatory expertise and these firms also have more experience in the (US) 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 thinking about 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 true, the answer is "it depends": in some cases, 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. 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.

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. 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.

Further, 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 the 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). 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 preliminary survey of the medical device landscape suggests a dramatic uptick in the commercialization of AI medical devices over recent years. At the same time, regulators have begun a germane and proactive 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 it 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 be less experienced in dealing with medical device regulation than traditional medical technology firms. And of course, emergent guidelines from the FDA and best practices from the broader digital medicine community (including clinical researchers, standards organizations, and product firms) 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 do so.

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.

6. 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.

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

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

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).