2017

Regulating Black-Box Medicine

W. Nicholson Price II

University of Michigan Law School, wnp@umich.edu

REGULATING BLACK-BOX MEDICINE

*W. Nicholson Price II**

Data drive modern medicine. And our tools to analyze those data are growing ever more powerful. As health data are collected in greater and greater amounts, sophisticated algorithms based on those data can drive medical innovation, improve the process of care, and increase efficiency. Those algorithms, however, vary widely in quality. Some are accurate and powerful, while others may be riddled with errors or based on faulty science. When an opaque algorithm recommends an insulin dose to a diabetic patient, how do we know that dose is correct? Patients, providers, and insurers face substantial difficulties in identifying high-quality algorithms; they lack both expertise and proprietary information. How should we ensure that medical algorithms are safe and effective?

Medical algorithms need regulatory oversight, but that oversight must be appropriately tailored. Unfortunately, the Food and Drug Administration (FDA) has suggested that it will regulate algorithms under its traditional framework, a relatively rigid system that is likely to stifle innovation and to block the development of more flexible, current algorithms.

This Article draws upon ideas from the new governance movement to suggest a different path. FDA should pursue a more adaptive regulatory approach with requirements that developers disclose information underlying their algorithms. Disclosure would allow FDA oversight to be supplemented with evaluation by providers, hospitals, and insurers. This collaborative approach would supplement the agency's review with ongoing real-world feedback from sophisticated market actors. Medical algorithms have tremendous potential, but ensuring that such potential is developed in high-quality ways demands a careful balancing between public and private oversight, and a role for FDA that mediates—but does not dominate—the rapidly developing industry.

TABLE OF CONTENTS

INTRODUCTION

Trauma patients often die from catastrophic hemorrhages.[1] If doctors and nurses are nearby to intervene promptly, death can frequently be prevented. But there are many patients, and it takes an observing expert to know when to intervene. How can we reduce these deaths? An emerging solution relies on computation—sophisticated algorithms that can find patterns in continuously monitored vital signs and call providers before it's too late.[2]

---

1. Nehemiah T. Liu et al., *Development and Validation of a Machine Learning Algorithm and Hybrid System to Predict the Need for Life-Saving Interventions in Trauma Patients*, 52 MED. & BIOLOGICAL ENGINEERING & COMPUTING 193, 193 (2014).

2. *Id.* at 193–94.

Medicine increasingly turns to algorithms to solve complex health problems. What pattern among a set of thousands of genes predicts which lung tumors will respond to treatment? How should scarce resources like inpatient beds be allocated to optimize patient care among patients with different illnesses and prognoses? What facial features can identify genetic disorders and suggest early interventions? More generally, how can algorithms be used to save lives and reduce suffering by improving medical practice? Sophisticated techniques are being developed to examine vast troves of health data, including genetic sequences, metabolic screens, and electronic health records, in search of answers to these and other questions. Frequently, these algorithms are opaque even to their developers, who may know that something reproducibly works, but not how or why.

This Article builds on my previous work introducing and analyzing this form of "black-box medicine" by canvassing current implementations and then asking: How should algorithmic medicine—that is, the use of such algorithms to guide care—be regulated? If providers rely on a trauma-monitoring algorithm that misses signs a provider could have observed, trauma patients might die from hemorrhage before anyone notices. Algorithms that predict the wrong drug to treat a cancer could waste months and hundreds of thousands of dollars on ineffective treatments. And if algorithms allocate hospital beds ineffectively, patients that need the resources most may not get them.

Patients and providers must trust that algorithms are safe and effective to rely on them, but they lack the experience or knowledge to evaluate algorithms at the point of care, creating a need for systemic regulation. Regulation can help but must walk a fine line: demonstrating safety and efficacy without destroying the flexibility and ongoing innovation that drive algorithmic medicine's development.

FDA is moving to regulate medical algorithms but looks to be moving in the wrong direction. In 2014, FDA proposed that complex laboratory-developed diagnostic tests should be subject to the same preapproval regime currently used to evaluate commercially sold diagnostic kits.[3] Under this regime, diagnostic tests, including algorithmic medicine, would be categorized by risk and then subjected to according regulatory scrutiny; complex algorithms would typically face the heaviest scrutiny and would require clinical testing and preapproval.[4] This regulatory model is poorly suited to drive the accurate, rapid, and safe development of algorithmic medicine. Algorithms can be developed quickly and tailored to the particular needs of health systems and patient groups. Furthermore, at least some algorithms evolve as they incorporate new data and learn to more accurately predict

---

3. U.S. Food & Drug Admin., Anticipated Details of the Draft Guidance for Industry, Food and Drug Administration Staff, and Clinical Laboratories: Framework for Regulatory Oversight of Laboratory Developed Tests (LDTs) (2014) [hereinafter FDA, LDT Draft Guidance], http://www.fda.gov/downloads/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/UCM407409.pdf [https://perma.cc/3CLA-LHRM].

4. *See infra* Section III.D.2.

relevant outcomes. Clinical-trial-based preapproval regimes—typically costly, slow, and designed for unchanging products—circumscribe these strengths. Rigidly imposing such a regime on algorithms would substantially slow their development and adoption.

This is not to argue that FDA should have no role in regulating algorithmic medicine. FDA has traditionally taken a command-and-control approach, exercising centralized authority to impose requirements for industry before allowing market access, and some elements of that approach are justified here.[5] In high-risk situations, moderate FDA preapproval requirements may be appropriate. In lower-risk situations, however, more modest registration requirements should suffice. In either case, robust postmarket surveillance will help ensure safety and quality as algorithms are developed and deployed in clinical practice. In general, more light-touch, iterative, adaptive regulation is likely to promote growth and innovation in the field.[6]

But command-and-control is not enough; FDA should consider adopting a collaborative governance approach in this area. Other evaluators, such as insurers, hospital systems, and providers, can and should be involved in helping to evaluate algorithms continually as they are implemented and used in clinical practice.[7] For these evaluators to help provide parallel oversight, they need information, and here FDA can play a central role by mediating the distribution of that information. FDA could require developers to disclose accurate information about their algorithms and could then mediate sharing of that information to insurers, hospitals, or providers. Algorithmic medicine demands a more flexible web of regulatory oversight to bring safe and effective algorithms to patients, providers, and the health system as a whole. In considering solutions, the new governance literature emerges as relevant, offering insights on how collaboration among multiple health care market participants may be useful for optimal oversight of complex medical algorithms.

This Article proceeds in five Parts. Part I describes medical algorithms, considers examples of current implementations, and lays out the overlapping subcategories of mobile-health algorithms and black-box algorithms. Part II asks why medical algorithms need to be regulated at all and, in the process, considers challenges faced by health-care market actors in evaluating medical algorithms in the absence of centralized regulatory authority. Part III begins by analyzing the relatively complex bases of FDA's regulatory

---

    5.   For a discussion on the background of FDA's command-and-control approach to regulation, see Eric R. Claeys, *The Food and Drug Administration and the Command-and-Control Model of Regulation*, 49 St. Louis U. L.J. 105 (2004).

    6.   *See, e.g.*, Robin Kundis Craig & J.B. Ruhl, *Designing Administrative Law for Adaptive Management*, 67 Vand. L. Rev. 1 (2014) (describing adaptive regulation and administrative law implications); *see also infra* Section V.C (describing the theory of "new governance").

    7.   *See generally* Jody Freeman, *Collaborative Governance in the Administrative State*, 45 UCLA L. Rev. 1 (1997) (describing collaborative public/private governance).

authority over software. It then describes FDA's existing and current approaches to regulation of medical algorithms, first canvassing FDA's historical approaches to software and in vitro diagnostic device regulations, and then considering FDA's current approaches to mobile-health software and laboratory-developed diagnostic tests. Part IV critiques aspects of current approaches, noting the dangers of too-rigid as well as too-permissive regulation. Part V presents suggestions for reform. It draws on the new governance literature to suggest possible approaches involving information forcing, collaborative governance, and iterative flexibility, but also suggests that FDA should retain a good deal of centralized command-and-control authority. A few brief thoughts conclude.

I.   What Are Medical Algorithms?

Medical algorithms are, somewhat tautologically, algorithms used in medicine. But since many decisionmaking processes can be described as algorithms,[8] what I mean here is more specific: computer-based algorithms that help make medical decisions or analyze medical information. Examples include computer-aided diagnostics, such as a classifier that diagnoses melanoma from pictures of skin lesions;[9] a program that evaluates a magnetic resonance image (MRI) for the presence of a tumor;[10] predictive analytics programs that attempt to identify high-risk patients based on a host of factors before that risk actually materializes;[11] diagnostic tests aimed at personalized medicine that calculate a drug dosage based on a patient's weight, sex, and genetic sequence;[12] or smartphone apps that recommend diet choices based on a patient's exercise patterns combined with baseline medical information.[13] In all these examples the central feature of the technology is an

---

8.   Simpler algorithms, sometimes paper-based, have a long history in medicine. More recently, efforts have been made to group, classify, and make available the spectrum of such simpler algorithms. The Medical Algorithms Project collected thousands of medical algorithms in the form of Microsoft Excel spreadsheets, *see* M. Sriram Iyengar & John R. Svirbely, The Medical Algorithms Project (2009), https://arxiv.org/pdf/0908.0932.pdf [https://perma.cc/3BWX-45G2], and has grown into Medal, a company with the same collection of algorithms with mobile apps and over 130,000 users, *see also* Medal, http://www.medicalalgorithms.com [https://perma.cc/7KYX-KCT4]. As older algorithms become more integrated with modern technology, distinguishing between simpler, conceptually paper-based algorithms and modern algorithms that require a computer becomes both more difficult and less important.

9.   *See* Andre Esteva et al., *Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks*, 542 Nature 115 (2017).

10.   *See* Shijun Wang & Ronald M. Summers, *Machine Learning and Radiology*, 16 Med. Image Analysis 933 (2012).

11.   *See* Alexander T. Janke et al., *Exploring the Potential of Predictive Analytics and Big Data in Emergency Care*, 67 Annals Emergency Med. 227 (2016).

12.   *See* Isaac S. Chan & Geoffrey S. Ginsburg, *Personalized Medicine: Progress and Promise*, 12 Ann. Rev. Genomics & Hum. Genetics 217 (2011) (describing dosing predictions for warfarin and other drugs).

13.   See, for example, MySymptoms, http://skygazerlabs.com/ [https://perma.cc/E35E-ZMR5], an app that tracks diet, exercise, symptoms, and other factors, and analyzes those data for patterns.

algorithm that analyzes information to tell patients or providers something about a medical situation or to recommend a course of medical action or inaction.

There are two types of algorithms involved in the process of using relationships in medical data to drive treatment. We might term the first a research algorithm—it is the process by which data are analyzed and relationships are discovered. The second we might call a prediction algorithm—it is the process by which relationships are applied to new data to generate predictions, recommendations, and the like. To take a very simple example, suppose we wish to examine the relationship between blood pressure and heart failure, and we have a dataset of many patients' blood pressure at the start of a year and whether they suffered heart failure during that year. The research algorithm might be simply, "Order all patients by blood pressure, separate them into ten groups by ascending average blood pressure, and calculate what fraction of each group experienced heart failure." We'd then have a table relating blood pressure ranges to likelihood of heart failure. The prediction algorithm would incorporate this data to make a prediction based on new data: take a patient's blood pressure, look up the relevant likelihood of heart failure in the table, and that's the prediction for the patient. In this Article, I am primarily examining the second type of algorithm, though in some of the techniques discussed here the two types cannot be readily separated.

An additional complication relates to changes in algorithms. In some cases, the algorithm may be static—that is, it is developed once and then applied over and over again to new data about new patients. In other cases, the algorithm itself evolves over time to improve its performance, incorporating new data as they are acquired.[14] Helping the algorithm improve over time is probably a good thing for the algorithm but makes it harder to regulate because the algorithm being regulated is a moving target. This Part goes into more depth on medical algorithms, giving current examples in Section I.A, then briefly describing two somewhat overlapping subphenomena: mobile-health applications in Section I.B, and inherently black-box algorithms, the principal focus of this Article, in Section I.C.

A. *Current Implementations*

Algorithmic medicine is a rapidly developing field. This Section describes a few complex algorithms in use or development today, including real-time patient-status analysis for hospital systems, monitoring devices that constantly track and analyze individuals' vital signs, and algorithms that predict or early diagnose cancer based on complicated sets of measurable biological characteristics.[15]

---

14. *See infra* Section IV.A.2.

15. This Article does not address the economics or business models of these development models. In general, they could create value by improving the quality or efficiency of care, or by identifying untapped opportunities for increased billing, as described throughout this Section.

Some complex medical algorithms are designed to aggregate information about individual patients with contemporaneous information from other sources in the health system. The company Lumiata aims to use "170 million data points to 'transform' data from insurance claims, electronic medical records, medical sensors and other sources into information that can be used to predict the best ways to treat individual patients and conditions."[16] The company's algorithms integrate data from many sources with information directly measured in or provided by the patient—including, potentially, through the use of wearable patient monitors[17]—to provide individual suggested diagnoses, predictions, and treatment recommendations.[18] Lumiata's system also suggests system-wide analytics, helping hospitals allocate resources to the patients who need them most—and also helping hospitals spot untapped billing opportunities.[19]

Daily measurements of individuals are also fodder that can be used to develop medical algorithms and their associated devices. Apple's iPhone collects data based on how the user moves throughout the day, and add-on devices can measure heart rates or rhythms; Apple offers a ResearchKit that enables these data to be connected to medical algorithms run by providers or by the individuals themselves.[20] In a more advanced tracker under development by the company Quanttus, a wristband blood-pressure monitor tracks blood pressure through the day and while and individual sleeps, up to 10,000 data points per hour.[21] The company's current model is based on

---

For a description of difficulties in appropriating that increased value, see W. Nicholson Price II, *Big Data, Patents, and the Future of Medicine*, 37 Cardozo L. Rev. 1401 (2016).

16.   Chris Rauber, *Lumiata Nabs $6 Million for Personalized Medical Care Software*, S.F. Bus. Times (Sept. 11, 2014, 7:04 AM), http://www.bizjournals.com/sanfrancisco/blog/2014/09/lumiata-6-million-funding-personalized-health-data.html [https://perma.cc/3PEG-FQZV?type=image]. The company's technology is based on a machine-learning based "Risk Matrix" using "175 million patient-record years" to predict risk of 20 conditions based on claims data, electronic health records, and other forms of medical data. *See* Press Release, Lumiata, Lumiata Launches Its AI-Powered Risk Matrix that Improves Care Coordination and Risk Management (Sept. 13, 2016), http://www.lumiata.com/lumiata-launches-its-ai-powered-risk-matrix-that-improves-care-coordination-and-risk-management-2/ [https://perma.cc/7K4T-AM2V].

17.   Scott Amyx, *Wearing Your Intelligence: How to Apply Artificial Intelligence in Wearables and IoT*, Wired (Dec. 4, 2014), http://www.wired.com/insights/2014/12/wearing-your-intelligence/ [https://perma.cc/2BBR-RDXQ].

18.   Press Release, Lumiata, *supra* note 16.

19.   *See Product*, Lumiata, http://www.lumiata.com/product/ [https://perma.cc/J5V9-S85B] (noting that Lumiata can "[i]mprove population health care management" and "realize earned reimbursements").

20.   *See* Chad Rudnick, *7 Medical Algorithms to Use with Apple ResearchKit Apps*, Medal, http://blog.medicalalgorithms.com/7-medical-algorithms-to-use-with-apple-researchkit-apps/ [https://perma.cc/TG4G-VSV4] (describing medical algorithms that can connect with the iPhone via ResearchKit).

21.   Vinod Khosla, *The Reinvention of Medicine: Dr. Algorithm V0-7 and Beyond*, Tech-Crunch (Sept. 22, 2014), http://techcrunch.com/2014/09/22/the-reinvention-of-medicine-dr-algorithm-version-0-7-and-beyond/ [https://perma.cc/Y2M9-DTMX].

tracking but it anticipates connecting the data to machine-learning algorithms at central servers to offer more complex predictive information to patients.[22]

Medical algorithms can also take advantage of the rapidly advancing ability to measure large numbers of biological markers in the human body. Applied Proteomics is leveraging the power of proteomics—measuring the proteins expressed by an individual—to develop new cancer tests.[23] The company intends to use its proteomics platform to predict and diagnose other disease states,[24] and can measure around 300,000 different markers from a single blood test or even a dried blood spot.[25] Once it determines the level of large numbers of proteins, it uses a complex algorithm to determine whether the patient has early-stage colorectal or pancreatic cancer.[26]

B. *Mobile Health*

Mobile health represents a tremendous area of growth and energy for medical algorithms.[27] Mobile health is worth describing separately for multiple reasons, but chiefly because FDA has treated it separately.[28] As Nathan Cortez has defined it, "'Mobile health,' or 'mHealth,' is the use of mobile communications devices like smartphones and tablet computers for health or medical purposes, usually for diagnosis, treatment, or simply well-being and maintenance."[29] Mobile health sometimes uses specialized devices, but more typically involves apps downloaded onto any number of mobile communications platforms.[30] Many types of mobile-health applications exist, with uses across the medical spectrum.[31] Patient-focused apps range from

---

22.   Rachel Metz, *This Fitness Wristband Wants to Play Doctor*, MIT Tech. Rev. (Feb. 19, 2014), http://www.technologyreview.com/news/524376/this-fitness-wristband-wants-to-play-doctor/ [https://perma.cc/C659-4PBX].

23.   *E.g.*, Jeffrey J. Jones et al., *A Plasma-Based Protein Marker Panel for Colorectal Cancer Detection Identified by Multiplex Targeted Mass Spectrometry*, 15 Clinical Colorectal Cancer 186 (2016).

24.   *Pipeline*, Applied Proteomics Inc., http://www.appliedproteomics.com/pipeline/ [https://perma.cc/6H23-VSNE].

25.   *See* Khosla, *supra* note 21.

26.   Press Release, Applied Proteomics Inc., Applied Proteomics Inc. Collaborates with German Cancer Research Center to Discover and Develop New Cancer Diagnostic Tests (Oct. 9, 2013), http://appliedproteomics.com/wp-content/uploads/2017/01/2013_10-09_api_GCRC_pr_1009.pdf [https://perma.cc/PM4F-7CYQ].

27.   For a detailed account of mobile health, see Nathan Cortez, *The Mobile Health Revolution?*, 47 U.C. Davis L. Rev. 1173 (2014) [hereinafter Cortez, *The Mobile Health Revolution?*]. For a brief overview of FDA regulation of mobile health technologies, see Nathan G. Cortez et al., *FDA Regulation of Mobile Health Technologies*, 371 New Eng. J. Med. 372 (2014) [hereinafter Cortez, *FDA Regulation of Mobile Health*].

28.   *See infra* Section III.D.1.

29.   Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1176.

30.   *Id.*

31.   *See id.* at 1181–90 (introducing an extensive typology of mHealth apps and classifying several types, with examples).

simple tracking apps like step monitors to complex apps for controlling
FDA-regulated medical devices such as insulin pumps. Provider-focused
apps let smartphones remotely view hospital monitors or provide diagnostic
assistance. Apps aimed at both markets may allow remote viewing of CT
scans and MRIs or may attempt to discern problems in recorded body
sounds like heartbeats or lung sounds. In 2015, over 165,000 health-related
apps were offered to consumers in the Apple and Android app stores.[32] Mo-
bile-health applications are especially significant in the field of algorithmic
medicine because they are rapidly proliferating, and because they have al-
ready been the subject of substantial recent regulatory attention.[33]

    This Article is principally concerned with those mHealth apps—
whether patient or provider focused—that use algorithms to monitor, sug-
gest, diagnose, or otherwise process medical information. For instance, de-
veloping technologies include an app that analyzes facial features of infants
to identify genetic disorders as early as possible;[34] an app that identifies au-
tism in young children through eye tracking;[35] and an app that tries to pre-
dict migraine attacks by using machine learning to analyze patterns of
triggers, symptoms, and physiological data and create individualized mod-
els.[36] All of these revolve around embedded algorithms—in these cases,
black-box algorithms as described in the next Section. Other apps may sim-
ply provide remote access to patient data, or may record patient informa-
tion; since these apps do not involve embedded algorithms in the same way,
I do not focus on them here.

C. *Black-Box Medicine*

    Black-box medicine is "the use of opaque computational models to
make decisions related to health care"[37] and is the focus of the remainder of
this Article. It is the subset of algorithmic medicine where the algorithms are

---

32. *Things Are Looking App*, Economist (Mar. 10, 2016), https://www.economist.com/
news/business/21694523-mobile-health-apps-are-becoming-more-capable-and-potentially-
rather-useful-things-are-looking [https://perma.cc/V2E5-9P4M].

33. *See infra* Section III.D.1.

34. Face2Gene, https://suite.face2gene.com [https://perma.cc/Y6TL-ZHBJ]; *see also*
Megan Molteni, *Thanks to AI, Computers Can Now See Your Health Problems*, Wired (Jan. 9,
2017, 1:00 PM), https://www.wired.com/2017/01/computers-can-tell-glance-youve-got-genetic
-disorders/ [https://perma.cc/H9QX-RWWW].

35. *Autism*, RightEye, https://www.righteye.com/tests-therapies/autism [https://perma.
cc/RYU6-MBYB].

36. Second Opinion Health, the developer of the Migraine Alert app, is currently recruit-
ing participants for a clinical trial to test its app in collaboration with the Mayo Clinic and
Allergan. *Individualized Prediction of Migraine Attacks Using a Mobile Phone App and Fitbit
(Migraine Alert)*, ClinicalTrials.gov, https://clinicaltrials.gov/ct2/show/NCT02910921
[https://perma.cc/7PBC-CFUT].

37. W. Nicholson Price II, *Black-Box Medicine*, 28 Harv. J.L. & Tech. 419, 421, 429–34
(2015) (introducing the concept of black-box medicine and exploring some of its policy
implications).

unavoidably opaque, whether those algorithms are used in an mHealth context or in other systems. Typically, such algorithms are derived from large datasets of health information using sophisticated machine-learning techniques and reflect complex underlying biological relationships. Algorithms can be opaque for multiple reasons. Sometimes, algorithms are nontransparent because, while they may rely on explicit rules, those rules are too complex for us to explicitly understand—for example, patients whose measurements place them in a particular region of n-dimensional (where n is large) characteristic-space are at a higher risk of stroke. In particular, these rules may be impossible to explain or to understand by following the process of scientific/medical discovery: mechanistic lab experiments followed by confirmatory clinical trials.[38] Other times, the relationships used in a black-box algorithm are literally unknowable because of the machine-learning techniques employed—that is, no one, not even those who programmed the machine-learning process, knows exactly what factors go into the ultimate decisions.[39] A key distinguishing feature of black-box algorithms, as the term is used here, is that it refers to algorithms that are inherently black box (i.e., their developers *cannot* share the details of how the algorithm works in practice)—rather than to algorithms that are deliberately black box (i.e., their developers *will not* share the details of how the algorithm works).[40] Black-box algorithms are especially likely to evolve over time as they incorporate new data into an integrated process of learning-and-applying.[41]

To understand black-box machine learning, take an example familiar from daily use in nonmedical contexts: image-recognition technology. Humans are very good at recognizing images, but it is hard to explain exactly what features of an image allow someone to recognize the subject. It is easy to look at a picture of a duck and say, "That's a duck." It is easy because we have seen many pictures of ducks—indeed, likely, many ducks—and know what ducks look like. But it is hard to state what a picture of a duck looks like with enough precision to tell someone who has never seen one to accurately and consistently identify one.[42] And it is very hard to tell a computer, hyperliteral and without any relevant experience, how to perform that task. But computers can learn how to do this; for everyday examples, we

---

38. *Id.* at 434; *see* Jenna Burrell, *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms*, Big Data & Soc'y, Jan.–June 2016, at 1, 5. Burrell notes that opacity in machine-learning algorithms can be either deliberate (i.e., secrecy), the result of technical illiteracy, or the result of the inherent incomprehensibility of machine-learning algorithms, Burrell, *supra*, at 1, and describes the third in some detail, *id.* at 5–9. I focus here on the third definition: the inherent incomprehensibility of machine-learning algorithms and their products.

39. Burrell, *supra* note 38, at 4–5.

40. For an excellent description of challenges with deliberately black-box algorithms in other contexts, see Frank Pasquale, The Black Box Society: The Secret Algorithms that Control Money and Information (2015).

41. *See infra* Section IV.A.2.

42. The curious reader may find the exercise enlightening.

need look no further than Google Image Search, which algorithmically identifies the subjects of images. Image-recognition algorithms are black-box machine-learning algorithms—a classifier is presented with a set of known images ("Here are 10,000 pictures of ducks"), develops complex internal rules based on nonlinear processes, tests those rules on a test set ("Which of these are ducks?"), adjusts the internal rules based on the success of the test, and repeats the process until it can accurately and consistently classify the images.[43] But at the end of the process, the computer can no more tell us how to identify a picture of a duck than we could tell it.

Computer algorithms can try to classify medical images in the same way.[44] Billions of medical images are taken each year, including MRIs, CAT scans, and x-rays. These images are interpreted by trained practitioners, typically radiologists, but their interpretations are expensive, time consuming, and have a relatively high error rate.[45] Algorithms that can classify radiological images are accordingly an area of significant research, though less substantially integrated into current clinical practice.[46]

To clarify, not all complex medical algorithms are black box. Any of the types of medical algorithms discussed above can be black-box algorithms, depending on how they are developed. Classifiers may be based on explicit characteristics (e.g., "if a yellow circle is surrounded by blue, it is the sun"; "in a brain MRI, an intense contiguous mass in the middle of the brain spanning the two hemispheres can be classified as the corpus callosum").[47] But these types of explicit rules work best for well-understood, simple characteristics and deal much less well with the complex relationships at the heart of much biology and medicine. Typically, black-box algorithms are most valuable for situations where the underlying scientific/medical relationships are especially complex, such that identifying and making these relationships explicit is overly difficult, very time consuming, or, in many circumstances, impossible with current tools—such as the relationships among thousands of genes or networks of environmental factors.

---

43.   *See* Gang Wang et al., *Learning Image Similarity from Flickr Groups Using Fast Kernel Machines*, 34 IEEE Transactions on Pattern Analysis & Machine Intelligence 2177 (2012) (describing a support vector machine method of classifying images from the image-sharing website Flickr by observing images in user-set groups (including sunsets, boats, penguins, and—indeed—ducks), learning, and predicting the likelihood that a future image will belong to that group).

44.   *See* Wang & Summers, *supra* note 10.

45.   Leonard Berlin, *Radiologic Errors, Past, Present and Future*, 1 Diagnosis 79 (2014) (noting persistent error rates of around 30% in radiological image interpretation, with roughly 70% perceptual errors (i.e., not seeing an abnormality) and 30% cognitive errors (i.e., failing to assign the correct significance to a perceived abnormality)).

46.   *See* Wang & Summers, *supra* note 10, at 946–47. Unsurprisingly, many radiologists are reluctant to embrace a technology that would automate one of their primary tasks. *See* Megan Molteni, *If You Look at X-Rays or Moles for a Living, AI Is Coming for Your Job*, Wired (Jan. 25, 2017, 1:00 PM), https://www.wired.com/2017/01/look-x-rays-moles-living-ai-coming-job/ [https://perma.cc/UN3F-X73J].

47.   *See* Alain Pitiot et al., *Expert Knowledge-Guided Segmentation System for Brain MRI*, 23 NeuroImage S85, S90–91 (2004).

II.   The Need for Regulation

Why does algorithmic medicine need to be regulated at all? The default for new technologies—at least in the free-market-based United States—is the absence of regulatory authority; if a producer wants to make something available and consumers wish to purchase it, that is the end of the story, barring some reason to deviate from the baseline.[48] Regulation in the form of premarket approval is especially rare. Drugs and medical devices are key examples of such premarket approval, for well-explored reasons relating to the status of health-care products as credence goods, whose efficacy must generally be taken on faith—or, more accurately, on the word of those with more knowledge.[49] This Section retells this familiar story in the context of algorithmic medicine.

As a prefatory note, I recognize that this discussion assumes a relatively neutral status quo ante—that is, that current medical practice is a neutral backdrop against which to consider the addition of medical algorithms. But, of course, current medical practice is rife with its own challenges, inaccuracies, and opacity.[50] Avoiding the implementation of algorithms because we fear the problems that might arise means leaving in place a system of medical errors that we know already exist, and foregoing the potential benefits of innovative treatment options that can save lives. Nevertheless, our medical-technology system tends to evaluate the adoption of new technologies against the backdrop of current practice. As a result, while I make the case for regulation here, that regulation should not be pursued as if the status quo were perfect. Later Parts attempt to address this balancing act.[51] For now, however, I present the case for regulation.

The quality and choice of medical algorithms can have substantial effects on the welfare of patients. Using a poor-quality algorithm[52] to direct

---

48.   In Europe, by contrast, the precautionary principle is applied in many circumstances, generally counseling against the adoption of a new technology that may have risks unless and until those risks are proven to be minimal or nonexistent. *See* Cass R. Sunstein, Laws of Fear: Beyond the Precautionary Principle 13 (2005) ("[I]t has become standard to say that with respect to risks, Europe and the United States can be distinguished along a single axis: Europe accepts the Precautionary Principle, and the United States does not.").

49.   *See* Stephen Breyer, Regulation and Its Reform 26–29 (1982) (noting information failures as a justification and invoking the examples of drugs and doctors); *see also* Uwe Dulleck & Rudolf Kerschbamer, *On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods*, 44 J. Econ. Lit. 5, 5–8 (2006).

50.   *See* Inst. of Med. of the Nat'l Acads., To Err is Human: Building a Safer Health System (Linda T. Kohn et al. eds., 2000).

51.   *See infra* Parts IV, V.

52.   How exactly to define the quality of an algorithm is itself a substantial challenge. I rely here on an operationalized definition—that is, a high-quality algorithm gives safe and effective recommendations or accurate, useful predictions, and a low-quality algorithm does not. But how to practically measure that quality may be hard, and the quality of the prediction algorithm may depend, of course, on the quality of both the incoming data and research algorithm. *See, e.g.*, Sharona Hoffman & Andy Podgurski, *Big Bad Data: Law, Public Health, and Biomedical Databases*, 41 J.L. Med. & Ethics 56, 57 (2013) (noting quality problems in medical big data). These questions are outside the scope of this paper.

care can result in a decision not to seek care when care is actually necessary, using a useless drug instead of one that might make a difference,[53] or even receiving fatal doses of radiation therapy.[54] On the other hand, high-quality algorithms can improve the quality of health care by reducing errors and generating better outcomes[55] and reduce costs by better targeting care and avoiding unnecessary treatment.[56]

But it is hard to know which algorithms are high quality. Ideally, purchasers can adequately and efficiently evaluate product quality in the marketplace, and they can choose to purchase only high-quality goods (or to accept a corresponding discount for lower quality goods). And ideally, producers guarantee the quality of their own products and send credible signals of that quality—but producers often face problematic incentives that limit our trust in their own evaluations.[57] Complex medical algorithms are generally hard for others to evaluate: their inner workings are (by definition) either complex or actually opaque, and in many circumstances, information about how they are developed and validated is kept secret.[58]

These barriers present challenges for those who might evaluate medical algorithms and suggest the need for regulation. Patients provide the most obvious parallel to consumers in ordinary technology purchases. Medical technology is, after all, designed to help patients. Patients, however, are not typically called upon to be informed consumers of medical technology in general, and especially not medical algorithms. They lack the training to evaluate information, even if it were available, and having each patient evaluate algorithms is much less efficient than some more centralized evaluator or evaluators. For patients, medical algorithms, like many other medical technologies, are "credence goods" that they cannot evaluate on their own.[59]

---

53. For example, some breast cancer patients that overexpress the *HER2/neu* receptor may be successfully treated with Herceptin (trastuzumab), while for other patients the side effects may outweigh the benefits. *See* Walter P. Carney, *HER2 Status Is an Important Biomarker in Guiding Personalized HER2 Therapy*, 2 Personalized Med. 317, 317–18 (2005); Melinda L. Telli et al., *Trastuzumab-Related Cardiotoxicity: Calling into Question the Concept of Reversibility*, 25 J. Clinical Oncology 3525, 3531 (2007).

54. *See* Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1218–19 (citing Nancy G. Leveson, Safeware: System Safety and Computers app. A (1995)); *infra* Section III.C.1 (discussing Therac-25). Such errors may be exacerbated by the tendency of decisionmakers to accept the recommendations of automated systems, *see* Linda J. Skitka et al., *Does Automation Bias Decision-Making?*, 51 Int'l J. Hum.-Computer Stud. 991 (1999), though of course such a bias is more understandable in the context of opaque algorithms.

55. *See* Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1192–95.

56. *See* I. Glenn Cohen et al., *The Legal and Ethical Concerns that Arise from Using Complex Predictive Analytics in Health Care*, 33 Health Aff. 1139, 1139 (2014); Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1195–97.

57. *See* George A. Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, 84 Q.J. Econ. 488, 488 (1970).

58. *See infra* Section II.B.

59. *See* Dulleck & Kerschbamer, *supra* note 49.

Notably, despite this challenge, many mHealth apps, including some relying on black-box algorithms, are marketed directly to consumers or patients;[60] this bolsters the need for rigorous evaluation. For instance, apps used to calculate insulin dosages can be used directly by patients with diabetes yet have frequently suffered from calculation errors.[61] Medical algorithms and mHealth apps that are deliberately consumer facing may need additional controls or demonstrations that they are suitable for unmediated use, beyond the regulatory mechanisms suggested throughout the remainder of this Article.[62]

Where patients are unlikely to be able to participate effectively in evaluating the quality of complex medical algorithms, other actors in the health system—such as providers, hospitals, or insurers—might. This Part discusses the general hurdles encountered in evaluating these technologies. Later I discuss the role these entities *could* play, if these challenges were overcome, in engaging in collaborative regulation of algorithms.[63]

A.　Inherent Complexity and Opacity

The biggest challenge in evaluating complex medical algorithms is definitional: they are complicated and often fully opaque.[64] Put simply, we use the tools of scientific understanding and clinical trials—including, ideally, comparative effectiveness trials—to choose whether, when, and which medical technologies to apply. But these methods just don't work well to evaluate quality for many complex medical algorithms.

60.　*See supra* notes 34–36 and accompanying text.

61.　*See infra* note 187 and accompanying text.

62.　Although this Article will not attempt to flesh out regulatory additions for consumer-facing black-box algorithms, two relevant contexts might provide some guidance: direct-to-consumer genetic testing and the process of making formerly prescription drugs available over the counter. For the former, see Kayte Spector-Bagdady & Elizabeth Pike, *Consuming Genomics: Regulating Direct-to-Consumer Genetic and Genomic Information*, 92 Neb. L. Rev. 677, 697–742 (2014). For the latter, see Holly M. Spencer, Comment, *The Rx-to-OTC Switch of Claritin, Allegra, and Zyrtec: An Unprecedented FDA Response to Petitioners and the Protection of Public Health*, 51 Am. U. L. Rev. 999, 1011–18 (2002).

63.　*See infra* Section V.C. This Article does not address the potential for product quality incentives driven by the tort system, which faces complicated issues in the context of complex medical algorithms. The combination of the learned intermediary doctrine (shielding medical product manufacturers from liability when well-informed providers direct treatment) and federal preemption of tort law suits against FDA-approved products makes the tort system a precarious source of direct incentives for product quality for medical software. *See, e.g.*, Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1218–19; Michael Greenberg & M. Susan Ridgely, *Clinical Decision Support and Malpractice Risk*, 306 JAMA 90, 90 (2011); Arnold J. Rosoff, *On Being a Physician in the Electronic Age: Peering into the Mists at Point-&-Click Medicine*, 46 St. Louis U. L.J. 111 (2002). In addition, many makers of medical information technology routinely use "hold-harmless" clauses in contracts with physicians to further limit liability. *See* Ross Koppel & David Kreda, *Health Care Information Technology Vendors' "Hold Harmless" Clause: Implications for Patients and Clinicians*, 301 JAMA 1276 (2009). These complex issues are beyond the scope of this Article, and remain the subject of future work.

64.　*See supra* Section I.C.

As described above, machine-learning methods often leave the mechanisms in the resulting algorithms fully opaque; even when they are not, they are likely so complex as to defy understanding.[65] Drugs and traditional medical devices are different; typically, we understand a significant amount about how they work.[66] This understanding can be used to evaluate the product and choose between products. But when medical algorithms are opaque or too complex to understand, this method of evaluation is substantially limited.[67] And while some algorithms may involve clinical trials for final validation, many will not for various reasons I have detailed previously.[68]

B.   *Lack of Procedural Information*

A key, related challenge is that the information useful to evaluate medical algorithms is often kept secret.[69] Although, as described above, many complex medical algorithms have inherently incomprehensible internal workings, other information can be used to evaluate them externally. Such information could include how the algorithms were developed, how they were validated, and, perhaps most usefully, the data on which they were trained.

In the context of drugs and many devices, information about how the drugs were developed and validated is developed and distributed as part of FDA's regulatory process.[70] Indeed, creating such information is a key

---

65.   *See supra* Section I.C.

66.   *See* Price, *supra* note 37, at 440–41. We certainly do not understand how all medical interventions work. *See id.* at 440 n.92.

67.   *See id.* at 441.

68.   *See* Roger A. Ford & W. Nicholson Price II, *Privacy and Accountability in Black-Box Medicine*, 23 Mɪᴄʜ. Tᴇʟᴇᴄᴏᴍᴍ. & Tᴇᴄʜ. L. Rᴇᴠ. 1, 16–20 (2016) (describing why clinical trials do not work for many forms of black-box medicine, including the potential requirement of hyper-specific sample populations and the conflict with black-box medicine's goals of being fast, cheap, and flexible).

69.   This lack of information has implications beyond regulation and evaluation. One related question, beyond the scope of this Article, is whether the standard for informed consent would need to change when providers use inevitably or deliberately opaque algorithms to shape care.

70.   *See, e.g.*, U.S. Fᴏᴏᴅ & Dʀᴜɢ Aᴅᴍɪɴ., Gᴜɪᴅᴀɴᴄᴇ ꜰᴏʀ Iɴᴅᴜsᴛʀʏ: Cʟɪɴɪᴄᴀʟ Sᴛᴜᴅɪᴇs Sᴇᴄᴛɪᴏɴ ᴏꜰ Lᴀʙᴇʟɪɴɢ ꜰᴏʀ Hᴜᴍᴀɴ Pʀᴇsᴄʀɪᴘᴛɪᴏɴ Dʀᴜɢ ᴀɴᴅ Bɪᴏʟᴏɢɪᴄᴀʟ Pʀᴏᴅᴜᴄᴛs—Cᴏɴᴛᴇɴᴛ ᴀɴᴅ Fᴏʀᴍᴀᴛ 2 (2006) [hereinafter FDA, Lᴀʙᴇʟ Gᴜɪᴅᴀɴᴄᴇ], http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm127534.pdf [https://perma.cc/N8Y5-3BSV] (noting that a label "must discuss those clinical studies that facilitate an understanding of how to use the drug safely and effectively. This is usually accomplished by providing concise, accurate summaries of information from studies concerning a drug's effectiveness (and sometimes safety) that practitioners consider important to clinical decision making"). The guidance also notes that clinical studies that show different effects in different subpopulations or different clinical settings should also be summarized. *Id.* There are exceptions; many medical devices are not independently shown to be safe and effective, but instead are shown to be equivalent to products already on the market. There are substantial criticisms of this approval mechanism, but at least the aim is to allow only safe and effective devices on the market. *See, e.g.*, Diana M.

justification for FDA's preapproval regime.[71] While providers may not typically read the clinical-trial reports of all the drugs they prescribe, the information is available should they choose to use it, and some do. And that information is certainly available to others who evaluate medical products in different contexts, including hospitals and insurers. Perhaps most importantly, this information is aggregated and analyzed by those with the relevant expertise, in the form of practice guidelines or other recommendation documents.[72]

This type of procedural information is less available for medical algorithms. First, information about how medical algorithms are developed looks substantially different from the clinical-trial data to which providers, insurers, and hospitals are accustomed. But developers of medical algorithms—especially those sold in the context of health software systems—keep information about those algorithms secret as a way to preserve competitive advantage.[73] This information can include algorithms' development, implementation, flaws, and weaknesses.[74] Even when providers gain independent experience about how algorithms perform in their own health-care settings, they may be kept from sharing that experience under nondisclosure agreements imposed by algorithm developers.[75] The lack of procedural information makes it substantially harder for players in the health-care system to evaluate medical algorithms.[76] Overall, black-box algorithms appear to be an appropriate subject of regulation. They have the potential to substantially shape medical practice for individual patients but face real difficulties in

---

Zuckerman et al., *Medical Device Recalls and FDA Approval Process*, 171 Archives Internal Med. 1006 (2011).

71.   *See* Rebecca S. Eisenberg, *The Role of the FDA in Innovation Policy*, 13 Mich. Telecomm. & Tech. L. Rev. 345, 366–72 (2007).

72.   *See, e.g.*, Nathan R. Every et al., *Critical Pathways: A Review*, 101 Circulation 461 (2000); Jeremy M. Grimshaw & Ian T. Russel, *Effect of Clinical Guidelines on Medical Practice: A Systematic Review of Rigorous Evaluations*, 342 Lancet 1317 (1993).

73.   Koppel & Kreda, *supra* note 63, at 1277; *see also* Price, *supra* note 15, at 1432–36 (describing the use of trade secrecy as a market exclusivity mechanism for black-box algorithms).

74.   Koppel & Kreda, *supra* note 63, at 1277.

75.   *Id.* A countervailing factor might be the reluctance of providers to adopt new technology without the ability to see how their peers are using it, shaped at least partially by malpractice risks from going outside the standard of care. As noted above at note 63, this Article does not address tort considerations in detail. For further consideration of the potential malpractice risks that might be faced by providers and health systems adopting, using, or failing to adopt black-box algorithms, see W. Nicholson Price II, *Medical Malpractice and Black-Box Medicine*, *in* Big Data, Health Law, and Bioethics (I. Glenn Cohen et al. eds., forthcoming 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2910417 [https://perma.cc/FU9S-PMWW].

76.   This lack of information has implications beyond regulation and evaluation. Although beyond the scope of this Article, deliberately or inevitably opaque algorithms that direct medical care may have implications for informed consent, either—most directly—because providers must disclose that they are using a complex algorithm they do not understand or—more controversially—because use of undisclosed algorithms actually requires a different conception of consent before adoption.

evaluation by actors in the health market. The next Part therefore turns to the role FDA plays in regulating related areas.

III. Existing Approaches

FDA is the primary regulator of medical technology, principally through its premarket review of drugs and medical devices.[77] And indeed, FDA is the regulator that has done the most to regulate algorithmic medicine so far. This Part briefly discusses the history of FDA regulation of algorithmic medicine. First, it describes FDA's general authority over medical devices under the Medical Device Amendments of 1976, and its baseline regulatory process for approving medical devices. Second, it addresses the extent of FDA's authority to regulate standalone medical algorithms. Third, it describes FDA's historical approach over the past few decades. Fourth and finally, it recounts recent changes in the agency's approach to medical technologies involving algorithms in two contexts: mobile health and laboratory-developed diagnostic tests.

A. *FDA Regulation of Medical Devices*

FDA has broad regulatory authority over "medical devices," which are defined as any "instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article which is 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."[78]

To determine whether products are medical devices, FDA must determine their intended use by looking to the objective intent of the manufacturer.[79] Objective intent, as defined by FDA, can be determined by examining any number of indicia including the product itself, manufacturer

---

77. Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1200–05. FDA is not the only regulator which could potentially regulate medical algorithms. *See id.* at 1211–17 (discussing other potential regulators of mHealth algorithms); Andrew Tutt, *An FDA for Algorithms*, 69 Admin. L. Rev. 83 (2017) (arguing for a new agency focused on regulating algorithms). This Article focuses on FDA's potential role as the most likely regulator of medical algorithms.

78. 21 U.S.C. § 321(h) (2012). The full statutory definition of a medical device is

> an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including any component, part, or accessory, which is—
>
> (1) recognized in the official National Formulary, or the United States Pharmacopeia, 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 is not dependent upon being metabolized for the achievement of its primary intended purposes.

Id.

79. Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1200–01.

claims, marketing materials, distribution patterns, and consumer use of the product.[80]

FDA regulates medical devices under a risk-based classification regime.[81] Under this regime, all devices are subject to registration, listing, and adverse-event-reporting requirements. Low-risk devices (Class I) are subject only to these requirements.[82] High-risk devices (Class III) must be approved through a full premarket-approval pathway (PMA),[83] which typically involves clinical trials and the presentation of extensive evidence of safety and efficacy to FDA.[84] Moderate-risk devices (Class II), unsurprisingly, fit in the middle; they can be cleared via a PMA or by demonstrating that they are equivalent to an already-approved device through a process known as a 510(k), named after its authorizing statutory provision.[85] The 510(k) notification process involves widely varying levels of regulatory scrutiny, ranging from a quick review of a short filing to searching scrutiny of applications in the thousands of pages.[86] The vast majority of devices are cleared via this process.[87]

---

80.   21 C.F.R. § 801.4 (2016); Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1201. FDA has testified that it will not determine intended use based on actual use by doctors or patients. Letter from Michele Mital, Acting Assoc. Comm'r for Legislation, to Tim Murphy, Chairman, Subcomm. on Oversight & Investigations, Comm. on Energy & Commerce, House of Representatives 2 (Mar. 20, 2013), http:// www.genomicslawreport.com/wp-content/uploads/2013/03/HousemHealthLetter.pdf [https://perma.cc/VV8Z-ATUD].

81.   21 C.F.R. § 860.

82.   *Id.* § 860.3.

83.   *Id.* § 860.3(c)(3).

84.   *See id.*

85.   *Id.* § 860.3(c)(2); *see also* Federal Food, Drug, and Cosmetic Act § 510(k), 21 U.S.C. § 360(k) (2012). The 510(k) pathway can also occasionally be used to approve a medical device through a modified process where the device is recognized as different enough from a precedent device to warrant a new categorization (under the categorization scheme of 21 C.F.R. pts. 862–892), but similar enough to allow 510(k) approval rather than a full PMA. *See* Nathan Cortez, *Analog Agency in a Digital World*, *in* FDA in the Twenty-First Century: The Challenges of Regulating Drugs and New Technologies 438, 446 (Holly Fernandez Lynch & I. Glenn Cohen eds., 2015) (describing the approval of the ingestible PillCam for intestinal videos).

86.   *See* Inst. of Med. of the Nat'l Acads., Public Health Effectiveness of the FDA 510(k) Clearance Process: Balancing Patient Safety and Innovation: Workshop Report 56 (Theresa Wizemann ed., 2010) ("It is common for a 510(k) submission to contain hundreds or thousands of pages of documentation . . . ."); *Content of a 510(k)*, U.S. Food & Drug Admin., http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/PremarketNotification510k/ucm142651.htm [https://perma.cc/2MVT-W8YD] (noting the average 510(k) is thirty-five pages long);.

87.   Inst. of Med. of the Nat'l Acads., Medical Devices and the Public's Health: FDA 510(k) Clearance Process at 35 Years 15, 85 (2011), http://www.nap.edu/catalog/13150/medical-devices-and-the-publics-health-the-fda-510k-clearance [https://perma.cc/Y94J-A8UE].

### B.	*Regulatory Authority over Algorithms*

FDA likely has the authority to regulate most medical algorithms. To be a regulable medical device, a product must meet the statutory definition above, including that it be "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."[88] Many medical algorithms are explicitly intended by their developer for use in the diagnosis, cure, treatment, or mitigation of disease or other conditions in humans and thus satisfy the intent requirement of the definition.[89]

One might ask whether there is also a meaningful statutory requirement that the medical device fall within the statutory categories of "instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article."[90] Algorithms and software do not fit obviously into those categories, which all seem to reference physical constructs.[91] But the listed categories are quite broad, and "contrivance" at the very least seems ambiguous enough that FDA's interpretation of whether such products can be "devices" would be entitled to *Chevron* deference.[92]

That question may be moot. The 21st Century Cures Act, enacted in December 2016, suggests that Congress intends certain medical software to fall within the definition of a medical device. A section of the law entitled "Clarifying Medical Software Regulation" lists a set of software functions that are *excluded* from the definition of a medical device.[93] This list includes software for administrative support, wellness maintenance, electronic health

---

88.	21 U.S.C. § 321(h)(2) (2012).

89.	*See supra* Part I.

90.	21 U.S.C. § 321(h).

91.	A group of international regulators, including FDA representatives, have referred to at least some types of software as medical devices. Int'l Med. Device Regulators Forum, "Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations 5, 13–15 (2014), http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf [https://perma.cc/7C5E-XEQD] (describing software as a medical device and describing a potential scheme for classifying such software by risk).

92.	21 U.S.C. § 321(h); *see* Chevron U.S.A. Inc. v. Nat. Res. Def. Council, Inc., 467 U.S. 837, 843 (1984). A counterargument would be that all of the categories appear to contemplate physical items, and that by the interpretive canon of *ejusdem generis*, authority over purely nonphysical software or algorithms should not be assumed. *See, e.g.*, Gooch v. United States, 297 U.S. 124, 128 (1936) ("The rule of ejusdem generis . . . limits general terms which follow specific ones to matters similar to those specified.").

93.	21st Century Cures Act, Pub. L. No. 114-255, § 3060 (2016) (amending 21 U.S.C. § 360j (2012)). For a useful explanation of the statutory language from the perspective of industry stakeholders, see Clinical Decision Support Coal., Voluntary Guidelines for the Design of Clinical Decision Support Software to Assure the Central Role of Healthcare Professionals in Clinical Decision-Making 20–26 (2017) [hereinafter CDS Guidelines], http://cdscoalition.org/wp-content/uploads/2017/04/CDS-3060-Guidelines-032717-with-memo.pdf [https://perma.cc/K3CN-BZN5] (draft guidelines prepared by the CDS Coalition).

records, and transferring and formatting medical information.[94] This language suggests, though it does not make explicit, that other software *does* fall within the definition of a medical device, because there would otherwise be no need for the exclusion.[95] In particular, the Act does *not* exclude from the definition of a medical-device software functions that are "intended to interpret or analyze clinical laboratory test or other device data, results, and findings."[96] In a confusing circumlocution, the Act also separately excludes software that analyzes medical information and supports or provides recommendations so long as the software enables providers "to independently review the basis for such recommendations so that it is not the intent that such [providers] rely primarily on any of such recommendations to make a clinical diagnosis or treatment decision regarding an individual patient."[97] Since black-box algorithms by definition cannot enable such independent review of the basis for their recommendations, they would appear *not* to be excluded from the definition of medical devices and thus are likely within FDA's regulatory authority.[98]

In many instances, the ambiguity over whether a certain algorithm is itself within FDA's regulatory authority may not matter. Whether or not FDA has authority to regulate algorithms as devices themselves under current law, it can often instead regulate them as accessories to associated medical devices, such as monitors, special-purpose medical computers, or wearable medical devices, which certainly fall within the definition of a device.[99]

FDA has since before the 21st Century Cures Act stated that at least some types of medical software are medical devices, most recently in a 2015

---

94. § 3060(a) (amending 21 U.S.C. § 360j to add subsections (o)(1)(A)–(E)). This type of software, which does not rely on algorithms, is outside the scope of this Article.

95. The fact that Congress excluded some software from the definition of algorithms is not proof that Congress believed others to be included under the language of the FDCA, or that such a belief would be accurate. *Cf.* Bilski v. Kappos, 561 U.S. 593, 644–48 (2010) (Stevens, J., concurring in the judgment) (noting that Congress' creation of a defense against assertion of business methods patents in 35 U.S.C. § 273 did not indicate that business methods should be patentable subject matter under 35 U.S.C. § 101 and that Congress may instead have intended merely to limit fallout from a recent judicial interpretation of § 101, and describing § 273 as "a red herring").

96. § 3060(a).

97. *Id.*

98. *But see* CDS GUIDELINES, *supra* note 93. The draft guidelines suggest design principles to assure both that "healthcare professional users remain fully in control of their own decision-making when using such software; and [that] *such software is not regulated by FDA.*" *Id.* at 2 (emphasis added). The guidelines suggest that machine-learning-based software must provide a physician with information about validation, databases, and limitations of machine-learning methods, and that this is sufficient to satisfy the "independent review" provisions of the statute. *Id.* at 7, 10–11.

99. *See* 21 U.S.C. § 321(h) (2012); U.S. FOOD & DRUG ADMIN., MEDICAL DEVICE ACCESSORIES—DESCRIBING ACCESSORIES AND CLASSIFICATION PATHWAY FOR NEW ACCESSORY TYPES: GUIDANCE FOR INDUSTRY AND FOOD AND DRUG ADMINISTRATION STAFF 4–5 (2017), https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm429672.pdf [https://perma.cc/M2V6-E33K].

guidance document describing proposed regulation of mobile medical applications.[100] It has applied this authority to software products used to analyze patient data and generate patient-specific information. In one relatively recent example, FDA sent a warning letter in 2010 to Knome,[101] a direct-to-consumer genetic testing firm (since acquired by Tute Genomics).[102] FDA stated that Knome's product KnomeCOMPLETE, "a software program that analyzes genetic test results that are generated by an external laboratory in order to generate a patient specific test report . . . [was] a diagnostic device," referencing only the intended use of the software.[103] To my knowledge, no court has addressed the validity of such determinations.

Nonetheless, FDA's authority to regulate the full range of complex medical algorithms remains somewhat contestable. FDA has long taken the position—and others have long understood—that it does not regulate the practice of medicine.[104] To the extent that using complex algorithms to make predictions and direct patient care begins to look more like the practice of

---

100. *See* U.S. Food & Drug Admin., Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff 6 (2015) [hereinafter FDA, Mobile Medical App Guidance], http://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf [https://perma.cc/E6TP-ENQN] ("Although the FDA has not issued an overarching software policy, the Agency has formally classified certain types of software applications that meet the definition of a device. . . . The FDA has previously clarified that when stand-alone software is used to analyze medical device data, it has traditionally been regulated as an accessory to a medical device or as medical device software."). FDA is a member of the International Medical Device Regulators, which works toward harmonizing medical device regulations around the world. *About IMDRF*, Int'l Med. Device Regulators F., http://www.imdrf.org/about/about.asp [https://perma.cc/QVS5-ETHZ]. The International Medical Device Regulators Forum has convened a working group to study Software as a Medical Device. *See* Int'l Med. Device Regulators Forum, Software as a Medical Device (SaMD): Key Definitions (2013), http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf [https://perma.cc/E8HQ-73DP].

101. Letter from Alberto Gutierrez, Ph.D., Ctr. for Device & Radiological Health, to Jorge Conde, Co-Founder & Chief Exec. Officer, Knome, Inc. (June 10, 2010), https://www.fda.gov/downloads/medicaldevices/resourcesforyou/industry/ucm215239.pdf [https://perma.cc/3GD4-99TG].

102. Reid Robison, *Tute Genomics Acquires Knome*, Tute Genomics (Nov. 2, 2015), https://www.tutegenomics.com/news/tute-genomics-acquires-knome/ [https://perma.cc/N74W-MP5T].

103. Letter from Alberto Guttierez to Jorge Conde, *supra* note 101, at 1; *see also* Spector-Bagdady & Pike, *supra* note 62, at 732–35 (describing FDA regulation of direct-to-consumer services offering software-based interpretation of genetic sequence information).

104. *See, e.g.*, Legal Status of Approved Labeling for Prescription Drugs, 37 Fed. Reg. 16,503 (proposed Aug. 15, 1972) ("Congress did not intend the Food and Drug Administration to interfere with medical practice and . . . the bill did not purport to regulate the practice of medicine as between the physician and the patient."); Patricia J. Zettler, *Toward Coherent Federal Oversight of Medicine*, 52 San Diego L. Rev. 427, 430–32 (2015) (noting "the conventional wisdom among courts, lawmakers, and administrative agencies is that states regulate medical *practice*, while the federal government regulates medical *products*," though later arguing that the federal government does substantially regulate the practice of medicine and an arguing for an increased federal role in such regulation); Patricia J. Zettler, *Pharmaceutical Federalism*, 92 Ind. L.J. 845 (2017) (noting the blurred distinction between practice and products in the contexts of regenerative medicine and genetic testing).

medicine, FDA regulation of those algorithms might run afoul of this distinction.[105] At least for the time being, however, FDA clearly believes it has the authority to regulate standalone software products, and recent legislation supports this view. Overall, FDA is likely to be able to exercise considerable regulatory control over complex medical algorithms.

C. *FDA's Historical Approach*

FDA regulation of medical devices is often difficult to divide into neat categories. This is especially true for regulation of medical algorithms, which may stand fully alone or be incorporated into many different types of devices. This Section discusses FDA's historical regulation in two parts. Section III.C.1 looks at FDA's general take on software and embedded algorithms, including both how the agency has thought about the topic and how FDA has implemented broader policy choices in practice.

"Software" and "algorithm" are flexible terms; here, I use software to mean the programmatic instructions that run on a computer or other device, and algorithm to mean the underlying method that some software uses to address a medical problem or suggest a medical decision. A software program that simply records patient records, for instance, would not include any algorithm as I use the term here.

Section III.C.2 considers a specific set of algorithm-related technology: diagnostic tests. Diagnostic tests are addressed separately both because a large portion of modern medical algorithms are related to diagnostics and because the agency's approach in this area illuminates the approach it may take in other, related technological implementations of algorithms.[106]

---

105. The question is complicated by the incomprehensibility of black-box algorithms to providers. A full analysis of this question is outside the scope of this Article. An additional possible complication, also outside the scope of this Article, is whether standalone medical algorithms might be considered commercial speech protected by the First Amendment. *See, e.g.*, Barbara J. Evans, *The First Amendment Right to Speak About the Human Genome*, 16 U. Pa. J. Const. L. 549, 590–630 (2014) (arguing that purely informational interpretations of genomic results are speech protected by the First Amendment); Spector-Bagdady & Pike, *supra* note 62, at 735–42; Tim Wu, *Machine Speech*, 161 U. Pa. L. Rev. 1495 (2013) (considering when the output of computer algorithm can be afforded First Amendment protection). First Amendment arguments have sharply limited FDA's ability to regulate promotion of off-label drug uses. *See* Christopher Robertson, *When Truth Cannot Be Presumed: The Regulation of Drug Promotion Under an Expanding First Amendment*, 94 B.U. L. Rev. 545, 550–55 (2014).

106. To some extent there is an artificial distinction between software/algorithms and the diagnostic tests that may use those algorithms. I maintain the distinction here because, to some extent, it reflects agency action and policy on these issues.

1.   Algorithms and Software

FDA has a long history of regulating medical software.[107] Computerized devices and software have been on the agency's radar since the 1970s, including a 1981 Task Force on Computers and Software as Medical Devices.[108] Early deaths from an algorithmically driven radiation therapy machine, the Therac-25,[109] prompted additional FDA attention in the mid-1980s.[110] But while FDA has considered software for close to four decades, its oversight has been relatively ad hoc. FDA has passed few regulations specifically directed to software or algorithms.[111] Instead, the agency has issued dozens of nonbinding guidance documents[112] and, more consistently, has regulated through case-by-case adjudication of individual instances.[113] But the lack of regulations and formal policy has meant that many implementations of medical algorithms pass outside the ambit of this form of oversight. When algorithms directed at medicine, such as web-based diagnostics or personal health apps, are implemented and opened to the public without someone seeking FDA approval or being otherwise actively noticed by FDA, the agency can provide little to no oversight.[114]

For those devices that do go through FDA's approval, most fall into Class I or Class II, and thus few go through full Class III PMA review.[115]

---

107.   For a brief but thorough history of FDA regulation of medical software, see Cortez, *supra* note 85.

108.   *Id.* at 444.

109.   The Therac-25 relied on software to the extent that it removed hardware failsafes, creating doses so high as to "sometimes burn[ ] literal holes in patients' bodies." Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1218–19 (citing Nancy G. Leveson, *Medical Devices: The Therac-25*, *in* Safeware: System Safety and Computers app. A (1995)).

110.   *Id.* at 1219–21.

111.   *See* Cortez, *supra* note 85, at 445 (noting that the only FDA regulations specifically mentioning software are in device classifications (21 C.F.R. pts. 862–892), rules for radiology products (21 C.F.R. pts. 1000–1050), and the Quality Systems Regulation requirements that software's specifications be documented and that it be subject to validation and manufacturing controls (21 C.F.R. pt. 820)).

112.   *See id.* at 446–47 (describing the author's review of fifteen original and eleven updated guidance documents on varied software regulatory topics and "perhaps dozens more" specifying "special controls" for Class II devices, and noting that guidance often incorporates guidance from standard-setting bodies).

113.   *See id.* at 445–46.

114.   To a certain extent, this is tautologically true of all FDA-regulated technologies; physical medical devices and drugs also must be brought to FDA's attention or noticed by the agency to be regulated. These more traditional forms of medical technology, however, typically follow more well-trodden paths and are thus subject more typically to FDA oversight. Perhaps more importantly, FDA has issued volumes of regulations clarifying what drugs and medical devices are, and when they are properly subject to regulation—information that is relatively absent in the field of medical software and algorithms.

115.   U.S. Food & Drug Admin., Performance Report to Congress for the Medical Device User Fee Amendments 11 (2016), https://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/UserFeeReports/PerformanceReports/UCM537819.pdf [https://perma.cc/43DT-VDV2].

Devices based on medical algorithms, including those at the heart of software-based diagnostics, have typically been approved through the 510(k) pathway, and FDA has often determined that software applications are equivalent to nonsoftware precedents, even though they perform tasks in a markedly different manner.[116]

2.   Diagnostic Tests

FDA has been more heavily involved in the regulation of diagnostic tests. FDA exercises regulatory authority over diagnostic tests as medical devices. But beginning in 1976 with the enactment of the Medical Device Amendments,[117] FDA divided these tests sharply into two different categories: diagnostic kits and laboratory-developed tests. The first were regulated under FDA's full medical-device regime. FDA exercised its discretion not to enforce regulatory requirements for the second category, allowing a proliferation of different tests until quite recently.

Diagnostic kits. FDA regulates in vitro diagnostics (IVDs) under the device approval classification and pathway described above.[118] FDA defines IVDs as products that are intended for use in the diagnosis of disease or other conditions and involve the collection, preparation, and examination of specimens taken from the human body.[119] Since this definition potentially covers a broad range of products, FDA can approve or clear IVDs as Class I, II, or III devices.[120]

Hospitals and providers use in vitro diagnostics in procedures ranging from cancer screens to drug tests. Companion diagnostics, such as assays for the *HER2/neu* test that identifies which breast cancers are likely to be responsive to the drug Herceptin, may likewise be considered IVD products.[121] IVDs also include kits marketed directly to consumers, including home pregnancy tests, blood-sugar testing kits, and many others. Most IVDs are classified in Class I or II as low to moderate risk, and therefore most go through the 510(k) framework.[122]

---

116.   *See* Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1219.

117.   Medical Device Amendments of 1976, Pub. L. No. 94-295, 90 Stat. 539 (codified as amended in scattered sections of 21 U.S.C.).

118.   *See supra* Section III.A.

119.   21 C.F.R. § 809.3 (2016).

120.   *Overview of IVD Regulation*, U.S. Food & Drug Admin. (Mar. 19, 2015), http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/IVDRegulatoryAssistance/ucm123682.htm [http://perma.cc/5S4R-MP58]; *see* 21 C.F.R. pts. 862–866 (2016).

121.   *See List of Cleared or Approved Companion Diagnostic Devices (In Vitro and Imaging Tools)*, U.S. Food & Drug Admin., http://www.fda.gov/MedicalDevices/ProductsandMedical-Procedures/InVitroDiagnostics/ucm301431.htm [https://perma.cc/5EHP-TPK7]; *supra* note 53.

122.   Jeffrey Gibbs, *Regulatory Pathway for Clearance or Approval of IVDs*, *in* In Vitro Diagnostics: The Complete Regulatory Guide 43, 44 (Scott D. Danzis & Ellen J. Flannery eds., 2010).

Laboratory-developed tests. By contrast, as soon as FDA acquired the authority to regulate medical devices, it chose not to enforce its requirements for laboratory-developed tests, announcing a policy of "enforcement discretion."[123] It defined laboratory-developed tests as IVDs that are "designed, manufactured, and used within a single laboratory."[124] FDA's rationale for exercising enforcement discretion was that laboratory-developed tests were relatively simple and used clinical reagents that were themselves FDA approved. Other diagnostic kits are manufactured and sold as products to providers or directly to patients. Laboratory-developed tests (LDTs), on the contrary, are marketed as services provided to the health-care provider, which appears more like the traditionally FDA-unregulated practice of medicine.[125] This Section describes regulation of LDTs in some detail for two reasons: first, because FDA's historical approach demonstrates the parameters of their hands-off enforcement discretion, and second, because if black-box algorithms do fit within the definition of medical devices, most are likely LDTs because they will be developed and deployed within a single laboratory.

As described above, getting products approved by FDA's Center for Devices and Radiological Health can be easy and quick, or it can be an expensive, difficult, and time-consuming process. But it is substantially easier to just avoid FDA approval process entirely.[126] Unsurprisingly, then, the ability of firms to avoid FDA regulatory requirements by using LDTs rather than selling diagnostic kits has been highly attractive. The vast majority of diagnostic tests developed in the last four decades have been LDTs.[127] In the process, firms have pushed the limits of what counts as an LDT, while staying with FDA's definition of a device "designed, manufactured, and used within a single laboratory."[128]

The category of LDTs has thus expanded dramatically. Rather than only being simple tests actually localized to a single facility, LDTs have grown in both complexity and scope. Hewing to the letter of FDA's qualifications for

---

123.   FDA, LDT Draft Guidance, *supra* note 3, at 5.

124.   *Id.* at 4.

125.   *See* Ronald L. Weiss, *The Long and Winding Regulatory Road for Laboratory-Developed Tests*, 138 Am. J. Clinical Pathology 20, 22–24 (2012) (discussing the perception among those offering LDTs that they are providing medical services rather than operating a medical device); Juliana Han, Note, *The Optimal Scope of FDA Regulation of Genetic Tests: Meeting Challenges and Keeping Promises*, 20 Harv. J.L. & Tech. 423, 427 (2007); *supra* note 104 and accompanying text (discussing the understanding that FDA does not regulate the practice of medicine). One could also argue that when a single laboratory performs a test, no device is being introduced into commerce. Peter M. Kazon, *Regulatory Issues Facing Genetic Testing*, J. Health & Life Sci. L., Jan. 2010, 111, 117; *see* 21 U.S.C. § 331(a) (2012) (prohibiting the introduction of adulterated or misbranded devices into interstate commerce).

126.   In addition to avoiding the approval process, staying outside FDA's regulatory ambit allows firms to avoid even the general controls and notice requirements applied to all devices, whether Class I, II, or III. *See* 21 U.S.C. § 360c(a)(1) (2012).

127.   Rachel E. Sachs, *Innovation Law and Policy: Preserving the Future of Personalized Medicine*, 49 U.C. Davis L. Rev. 1881 (2016).

128.   FDA, LDT Draft Guidance, *supra* note 3, at 4–7.

enforcement discretion, complex LDTs are designed, manufactured, and actually performed in a single centralized facility—but that facility exists not as a part of the hospital or medical school ordering the test, but rather as a standalone commercial entity.

The best example of this phenomenon, including how it can be used for technology centered on algorithms, may be Myriad Genetics. For many years, Myriad was the sole United States source of genetic testing for the *BRCA1/BRCA2* genes, which predict breast- and ovarian-cancer predisposition.[129] Then and now, Myriad performs "BRACAnalysis" testing for hospitals and providers' offices across the country. To help patients and providers decide on genetic testing, Myriad offers a simple algorithm to predict the likelihood of *BRCA1/BRCA2* mutations based on family-history factors.[130] Providers would send a patient's blood- or cheek-cell sample to Myriad's laboratory in Utah; Myriad would perform its genetic analysis, looking for markers using its proprietary predictive algorithms to compare those markers to its extensive database of mutations and disease predictors;[131] and then Myriad would return results to the providers to be shared with their patients.[132] Overall, BRACAnalysis was a multimillion-dollar annual business for Myriad, using complex laboratory techniques and sophisticated algorithms to test thousands of people nationwide per year.[133] Nonetheless, because of the centralization of Myriad's laboratory, BRACAnalysis has been considered an LDT and therefore exempted from FDA regulatory scrutiny.[134]

---

129.   Myriad was the only source of such testing because it held patents on the genes themselves and on using those genes to perform genetic predisposition testing. Those patents have since been struck down as covering unpatentable subject matter. Ass'n for Molecular Pathology v. Myriad Genetics, Inc., 133 S. Ct. 2107, 2111 (2013) (holding genomic DNA unpatentable); Ass'n for Molecular Pathology v. U.S. Patent & Trademark Office, 689 F.3d 1303, 1335 (Fed. Cir. 2012) (holding Myriad's method of genetic testing for breast cancer predisposition unpatentable), *aff'd in part, rev'd in part sub nom. Myriad Genetics, Inc.*, 133 S. Ct. 2107.

130.   *BRCA Risk Calculator*, Myriad, http://www.myriadpro.com/brca-risk-calculator/calc.html [https://perma.cc/KGF6-H4HK] (using data from Thomas S. Frank et al., *Clinical Characteristics of Individuals with Germline Mutations in BRCA1 and BRCA2: Analysis of 10,000 Individuals*, 20 J. Clinical Oncology 1480 (2002)).

131.   By virtue of being the sole test provider, Myriad also developed a far more extensive database of mutations than any other, an advantage it has kept after its patents were struck down. *See* Dan L. Burk, *Patents as Data Aggregators in Personalized Medicine*, 21 B.U. J. Sci. & Tech. L. 233, 239–40 (2015) (discussing the role of patents in allowing Myriad to assemble its dataset); *cf.* Barbara J. Evans, *Economic Regulation of Next-Generation Sequencing*, 42 J.L. Med. & Ethics 51 (2014) (discussing data aggregation and siloing).

132.   *BRACAnalysis® Overview*, MyriadPro, https://new.myriadpro.com/products/bra-canalysis-overview/#1479841758876-99a68d39-15f7 [https://perma.cc/6KMJ-V3CE] (explaining that results are sent to patient's health-care provider).

133.   J.J. Colao, *How a Breast Cancer Pioneer Finally Turned a Profit*, Forbes (Oct. 17, 2012, 11:50 AM), https://www.forbes.com/sites/jjcolao/2012/10/17/how-a-breast-cancer-pioneer-finally-turned-a-profit/#6b54bc1d254b [https://perma.cc/AJ4Y-SVUY].

134.   Myriad has also developed a companion diagnostic that uses genetic analysis of *BRCA1/BRCA2* to determine whether patients are eligible for treatment with Lynparza, an ovarian cancer treatment. Like Myriad's flagship BRACAnalysis predisposition product, this product, BRACAnalysis CDx, is an LDT that can only be performed at Myriad's laboratory in Salt Lake City. Note, *Diagnostic Methods Patents and Harms to Follow-On Innovation*, 126

FDA's exercise of regulatory discretion for LDTs matters for algorithmic medicine in two significant ways. First, to the extent that some algorithmic medicine is indeed centralized because the analyses are only performed at some central facility, such algorithms would have fit into the enforcement discretion exception neatly—at least, until FDA recently began discussing ending its policy of enforcement discretion for LDTs, as described below.[135] Second, FDA's exercise of enforcement discretion over LDTs provides an example of how FDA can pragmatically exempt a large class of medical devices from preapproval requirements, and demonstrates the rapid innovation—but also the quality problems—that can result from such a policy.

D.   *Current Approaches*

FDA's historical approach to regulating medical algorithms and diagnostic tests has changed recently in response to two technological developments. On the pure algorithm side, the past several years have seen an explosion of mobile-health apps and technologies.[136] On the diagnostic-test side, the field of laboratory-developed tests has grown as just described from relatively simple on-site tests to highly complex tests that frequently rely on shipping samples and returning data from far-flung sites to one centralized laboratory, and often involve complex informatics and algorithms. These separate developments have prompted distinctly different reactions from FDA. On mobile health, it has settled on an approach that exempts most players in the chain of products comprising mobile-health devices. On laboratory-developed tests, FDA is moving forward with a framework to bring those tests within its preexisting regulatory preapproval regime. This Section describes these two developments and addresses what they may mean for algorithms more generally.

1.   Algorithms and Mobile Health

Mobile health is currently the subject of a flurry of potential government oversight. Congress[137] and several federal agencies, including FDA and others,[138] have considered regulating mobile-health technologies.[139] Although the popular press has focused on the disconnect in attitude between

---

Harv. L. Rev. 1370, 1380 (2013); *BRACAnalysis CDx*, Myriad Genetics, https://www.myri ad.com/products-services/companion-diagnostics/bracanalysis-cdx/ [http://perma.cc/HJ63-2DJY].

135.   *See infra* Section III.D.2.

136.   *See* Cortez, *FDA Regulation of Mobile Health*, *supra* note 27, at 372; Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1177–78.

137.   Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1208–09, 1216–17 (describing congressional hearings and general legislative inaction on mHealth).

138.   *Id.* at 1211–16 (describing non-FDA action on regulating mHealth).

139.   *Id.* at 1179, 1208–09. Other interested agencies include the Federal Trade Commission, the Federal Communications Commission, the Department of Defense, the Department of Commerce, and multiple Department of Health and Human Services subagencies. *Id.* at 1179.

mobile-health entrepreneurs and agencies that may regulate them,[140] others have noted that regulating agencies are actually relatively friendly to mobile-health technologies.[141]

FDA issued draft guidance in 2011,[142] a final guidance document in September 2013,[143] and—after some pushback[144]—a new final guidance document in February 2015 addressing mHealth applications.[145] It specified that it would not exercise jurisdiction over those creating devices or networks on which mHealth apps would run, but only those developing the apps themselves.[146] The agency also stated that it would exercise jurisdiction only over "mobile medical app[s]," defined as those intended to cure, mitigate, diagnose, prevent, or treat diseases or other conditions, or to affect the body's structure or function[147]—a category that includes many mHealth apps but

---

140.   *See, e.g.*, Associated Press, *Silicon Valley Struggles to Speak FDA's Language*, Daily Mail (Sept. 11, 2014), http://www.dailymail.co.uk/wires/ap/article-2752144/Silicon-Valley-struggles-speak-FDAs-language.html [https://perma.cc/2YU6-N3C6]. A poster child for the clash between Silicon Valley culture and FDA's regulatory oversight role is the company 23andMe, which developed and sold a consumer-oriented genetic testing kit. After a long-running dispute between the FDA and 23andMe, FDA sent 23andMe a Warning Letter, prompting it to cease offering genetic testing results. *See* Letter from Alberto Gutierrez, Dir., Office of In Vitro Diagnostics & Radiological Health, Ctr. for Devices & Radiological Health, Dep't of Health and Human Servs., to Ann Wojicki, Chief Exec. Officer, 23andMe, Inc. (Nov. 22, 2013), http://www.fda.gov/ICECI/EnforcementActions/WarningLetters/2013/ucm376296.htm [https://perma.cc/A2G2-XJS5]; John Timmer, *23andMe Bows to FDA, Drops All Medical Information from New Tests*, Ars Technica (Dec. 6, 2013, 12:00 PM), http://arstechnica.com/science/2013/12/23andme-bows-to-fda-drops-all-medical-information-from-new-tests/ [https://perma.cc/XFR5-45RX]. In a postscript illustrating difficulties in regulating new health technologies, 23andMe now offers customers their raw sequence data and directs them to a crowdsourced website where they can interpret their own results—presumably not FDA's ideal result. After 23andMe went through FDA regulatory process, FDA approved its testing services for Bloom Syndrome in 2015. Cyrus Farivar, *FDA Allows 23AndMe to Use Its Genetic Kits to Test for Bloom Syndrome*, Ars Technica (Feb. 20, 2015, 2:09 PM), http://arstechnica.com/science/2015/02/fda-allows-23andme-to-use-its-genetic-kits-to-test-for-bloom-syndrome/ [https://perma.cc/5EXF-5N84].

141.   *See* Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1179, 1206 (describing agency responses as generally "optimistic and aspirational," but also describing the app industry as "relatively naïve to *any* kind of regulation").

142.   U.S. Food & Drug Admin., Draft Guidance for Industry and Food and Drug Administration Staff: Mobile Medical Applications (2011), ftp://ftp.cdc.gov/pub/CLIAC_meeting_presentations/pdf/Addenda/cliac0811/T_addendum_FDA_Draft_Guidance_for_Mobile_Medical_Devices.pdf [https://perma.cc/P86D-67AC].

143.   Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff, 78 Fed. Reg. 59,038 (Sept. 25, 2013).

144.   *E.g.*, Emily Wasserman, *Lawmakers Question FDA Oversight of Mobile Medical Apps as Market Booms*, FierceBiotech (June 17, 2014, 3:58 PM), http://www.fiercebiotech.com/medical-devices/lawmakers-question-fda-oversight-mobile-medical-apps-as-market-booms [https://perma.cc/39L9-544D].

145.   FDA, Mobile Medical App Guidance, *supra* note 100.

146.   *Id.* at 9–11.

147.   *Id.* at 7–8.

has tremendously unclear boundaries.[148] Apps that would otherwise not fit with FDA's definition—apps designed to provide reference materials, to track general well-being or health, or to perform office functions, for instance—fall under FDA jurisdiction if they "use patient-specific data to generate customized diagnoses or treatment recommendations."[149] Both FDA and industry recognize that FDA's traditional relatively rigid regulatory approaches may interact poorly with the fluid and fast-moving structure of the mobile-health industry.[150] Nevertheless, as the guidance currently stands, a large swath of mHealth apps appear subject to FDA's jurisdiction as medical devices subject to the requirements described above.[151]

2. Laboratory-Developed Tests

In the field of laboratory-developed tests, FDA has also moved to bring more technologies, including many centered on algorithmic medicine, within its stricter regulatory ambit. In July 2014, FDA announced that it intended to change its approach to regulating laboratory-developed tests substantially.[152] In its notice of proposed rulemaking, FDA stated that it intended to adopt a risk-based framework to laboratory-developed tests.[153]

FDA specifically described potential risk factors in modern LDTs, including that many are "used to direct critical treatment decisions (e.g., prediction of drug response)[, or are] highly complex (e.g., automated interpretation, multi-signal devices, use of non-transparent algorithms and/or complex software to generate device results)."[154] These characteristics are central to the concept of complex medical algorithms in general and black-box medicine in particular, and thus significantly increased FDA regulation of algorithmic medicine appears likely under this approach.

Under the risk-based classification regime, as described above,[155] low-risk devices (Class I) are subject to registration, listing, and adverse-event-reporting requirements. Moderate-risk devices (Class II) are subject to the same requirements and additional specific controls; they must also either demonstrate equivalence to an already-approved device (through the 510(k)

---

148. Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1203–06 (describing the unclear boundaries between categories of definitely, definitely-not, and maybe regulated device under FDA guidance).

149. FDA, Mobile App Guidance, *supra* note 100, at 15–25; Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1204. Notably, these general categories of software are excluded from the definition of medical devices under the 21st Century Cures Act, Pub. L. No. 114-255, § 3060, 130 Stat. 1033, 1130 (2016) (to be codified at 21 U.S.C. § 360j(o)(1)–(2)).

150. Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1206–08.

151. Even for apps that FDA has said are only potentially regulable, the agency has encouraged developers to abide by its quality requirements. FDA, Mobile Medical App Guidance, *supra* note 100, at 13.

152. FDA, LDT Draft Guidance, *supra* note 3.

153. *Id.* at 1.

154. *Id.* at 8.

155. *See supra* Section III.A.

pathway) or undergo a more rigorous de novo approval. High-risk devices (Class III) must undergo a full PMA process. LDTs that address rare diseases or unmet needs, or which resemble traditional LDTs, will need only to register, be listed, and report adverse events.[156]

The eventual classification of complex medical algorithms will await FDA action.[157] But complex algorithms and black-box-medicine implementations are quite likely to be considered higher-risk devices. In general, devices with new intended uses are classified as Class III as a matter of law,[158] though FDA may down classify those devices if appropriate, including if they serve an unmet need.[159] More specifically, FDA has identified as a generally higher-risk class those devices "that act like companion diagnostics," including

> those devices that claim to enhance the use of a specific therapeutic product, though selection of therapy, patient population, or dose, but which are not included in the therapeutic product labeling (e.g., devices that claim to predict who will respond to a therapy approved for use in a larger population).[160]

This group is a precise definition of most of black-box medicine and much of the scope of complex medical algorithms in general; implementations will thus be most likely classified as higher-risk devices, either Class II or Class III, and subject in turn to much higher regulatory barriers to approval. Overall, FDA's intent to more tightly regulate LDTs has been estimated to increase regulatory costs by one to two orders of magnitude,[161] and it remains an open empirical question whether such increased regulatory costs would be worth incurring.[162]

In late 2016, FDA announced that it had decided to postpone finalizing the draft guidance after substantial pushback from industry.[163] Final action

---

156.   FDA, LDT Draft Guidance, *supra* note 3, at 11, 19–22. FDA stated that it intended to issue classification guidance within 24 months after finalization of the LDT guidance, *id.* at 24, which has been postponed, *see infra* notes 163–164 and accompanying text.

157.   In determining the risk classification of an LDT, FDA will consider, among other factors, the risk level of the disease/patient population, use in screening versus diagnosis, what clinical decision will be based on the test, the availability of other information in making that decision, alternatives for diagnosis and treatment, the cost of error, and the existence of adverse events. FDA, LDT Draft Guidance, *supra* note 3, at 10–11.

158.   Federal Food, Drug, and Cosmetic Act § 513(f)(1), 21 U.S.C. 360c(f)(1) (2012).

159.   FDA, LDT Draft Guidance, *supra* note 3, at 25.

160.   *Id.* at 25–26. Opacity and risk are not necessarily linked; one could imagine a very low-risk opaque algorithm (for instance, predicting male-pattern baldness based on a large number of genetic factors).

161.   Sachs, *supra* note 127, at 1895.

162.   *See infra* notes 189–194 and accompanying text (describing LDT quality failures).

163.   *See* Sheila Kaplan, *FDA Puts off Closing Lab-Test 'Loophole,' Leaving Decision to Congress and Trump*, Stat (Nov. 18, 2016), https://www.statnews.com/2016/11/18/fda-lab-test-loophole/ [https://perma.cc/E4N8-WHYN].

will await input from Congress and from the new administration.[164] In January 2017, FDA did approve one machine-learning-based application for clinical use under the 510(k) clearance pathway, but the application does not diagnose disease or recommend treatment—it uses image analysis to calculate the volume of heart ventricles.[165] FDA did also distribute a discussion paper describing its current thinking with respect to LDTs in the same month.[166] Although the discussion paper notes potential slowdowns in implementation of full LDT oversight and suggests the possibility of grandfathering existing LDTs, it does not suggest that the agency's thinking has changed with respect to opaque algorithms that make diagnoses or predictions to direct clinical care.[167]

IV.   Challenges in Existing Approaches

FDA's existing piecemeal approach raises substantial challenges in dealing with the increasing pace of development and variety of medical algorithms. These challenges come from two directions, reflecting both common debates about regulation and deregulation and the two recent approaches FDA has taken. On the one hand is the problem of overregulation, exemplified by the adoption of FDA's full risk-based framework for LDTs, including its likely classification of black-box algorithms as Class III devices demanding full premarket approval. On the other is a worry about underregulation, as in FDA's decision to leave some fraction of mHealth applications largely unregulated (mirroring its earlier treatment of LDTs).

A.   *Overregulation*

FDA has indicated in its approach to LDTs that it may take a stricter regulatory approach, which I argue here could be overregulatory. Under this approach, algorithmic medicine would be treated roughly according to the risk-based framework applied to all medical devices; more complex algorithms would be more likely to receive more substantial regulatory scrutiny, including—in many cases—full premarket approval requirements as Class

---

164.   *Id.* For a description of potential benefits of agency decisions to postpone difficult enforcement decisions, see Sharon B. Jacobs, *The Administrative State's Passive Virtues*, 66 Admin. L. Rev. 565, 575–83 (2014).

165.   Bernard Marr, *First FDA Approval for Clinical Cloud-Based Deep Learning in Healthcare*, Forbes: Tech (Jan. 20, 2017, 2:11 AM), https://www.forbes.com/sites/bernardmarr/2017/01/20/first-fda-approval-for-clinical-cloud-based-deep-learning-in-healthcare/#407883d5161c [https://perma.cc/Q3D9-884P].

166.   U.S. Food & Drug Admin., Discussion Paper on Laboratory Developed Tests (LDTs) (2017) [hereinafter FDA, Discussion Paper on LDTs], http://www.fda.gov/downloads/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/LaboratoryDevelopedTests/UCM536965.pdf [https://perma.cc/NF8U-Z3PP].

167.   Note also that such algorithms are not excluded from the definition of medical devices under the 21st Century Cures Act, passed roughly a month after FDA's decision to defer finalizing the draft guidance. Pub. L. No. 114-255, § 3060 (2016) (to be codified at 21 U.S.C. § 260j(o)(1)(E)).

III devices. This approach faces two main challenges. The first arises from practical limits on FDA's ability to regulate algorithms. The second arises from concerns that excessive FDA regulation may stifle innovation in developing new medical algorithms, especially when such algorithms may constantly evolve.

1.   Limits on FDA's Regulatory Ability

FDA's attempts to regulate are likely to face substantial practical difficulties. From a political economy perspective, FDA's moves to more tightly regulate mHealth applications, in particular, have met with substantial pushback from entrepreneurs, commentators, and Congress, and corresponding walk backs of proposed policy changes.[168] Similarly, intense opposition to FDA's proposal to more strictly regulate LDTs led the agency to postpone final action, though without repudiating its view that opaque diagnostic algorithms should be considered high-risk Class III devices.[169] In addition, stricter premarket regulation demands expertise in algorithm development, validation, and operation—expertise that FDA largely lacks.[170] Especially for more complicated and black-box algorithms, implementing strict regulatory control efficiently and correctly is likely to be a substantial challenge for the agency. Nevertheless, it is not obvious that any other government agency would do a better job—nor that any other agency has the statutory mandate to do so.[171] Finally, there remain some doubts about the limits of FDA's statutory authority to regulate all black-box algorithms, though as described above, these are likely of only mild concern.[172]

2.   Limits on Innovation

The second major challenge to command-and-control premarket regulation of algorithmic medicine comes in the cost to innovation. FDA's current approach is likely to create substantial difficulties enabling innovation in the field of algorithmic medicine. If the status quo were ideal, this might

---

168.   *See supra* Section III.D.1 (noting FDA's change of position on regulating mobile health technologies). For more on the political economy of FDA regulations, see Daniel Carpenter, Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA 33–70 (2010).

169.   FDA, Discussion Paper on LDTs, *supra* note 166.

170.   *See* Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1206 (describing FDA's awareness that "it lacks technical expertise on mobile technologies"); *supra* note 27 and accompanying text.

171.   The Centers for Medicare and Medicaid Services regulate clinical laboratories analyzing specimens under the Clinical Laboratory Improvement Amendments, for instance, but this covers only quality controls at facilities "for the biological, microbiological, serological, chemical, immuno-hematological, hematological, biophysical, cytological, pathological, or other examination of materials derived from the human body for the purpose of providing information for the diagnosis, prevention, or treatment of any disease or impairment of, or the assessment of the health of, human beings." 42 U.S.C. § 263a(a) (2012).

172.   *See supra* Section III.B.

be fine—but the medical system is rife with uncertainty and errors, and black-box algorithms have the promise to bring major improvements to health and life.[173]

As described above, many more complex types of medical algorithms—especially black-box algorithms—are very likely to be classified as Class III devices requiring a full premarket approval process.[174] But this process is the most involved, most expensive, and most time-consuming form of regulation in FDA's arsenal of regulatory oversight. This has two impacts on this type of algorithmic medicine. First, it significantly increases the hurdles of getting particular algorithms developed and to the market. Second, it substantially limits the flexibility of algorithms to change with additional data.

The first concern is one familiar to many areas of regulation—when hurdles to bringing a product to market are increased, overall market incentives for firms to develop that product are decreased.[175] This pattern is typical for regulated industries but should nonetheless be noted. In other FDA-regulated contexts, intellectual-property protection, especially patents, may provide counterbalancing incentives to develop new technology. But intellectual-property incentives for algorithmic medicine are substantially weaker than those available for other FDA-regulated technologies.[176] The weight of FDA regulations on innovation is significant, as evidenced by the much more rapid development of laboratory-developed tests than diagnostic kits, despite the potentially broader market for the latter.[177] And to the extent that black-box algorithms may be developed cheaply, imposing higher regulatory burdens could negate that advantage.

The second concern is more distinct to algorithmic medicine. Ideally, medical algorithms—especially those disassociated from particular physical products—should be able to be flexibly updated as more data are collected. Some monitoring systems operate in real time or near-real time, taking into account how individual patients are doing across a medical system, such as a large hospital, and making recommendations about treatment choices accordingly.[178] One version of such a system would envision a static algorithm—given data *X*, *Y*, and *Z* about how patients are doing, and a shortage of hospital beds, preferentially discharge patients *A* and *B* over patient *C*; over months and years, that conditionality remains constant. More

---

173.   *See supra* note 51 and accompanying text.

174.   *See infra* notes 179–191 and accompanying text.

175.   This barrier to market entry has different effects on different market actors. Incumbent firms may be protected from new market entrants by substantial barriers to entry. *See* Eisenberg, *supra* note 71, at 356–57 (describing incumbent-protecting barriers to market entry created by hurdles to FDA approval).

176.   An in-depth consideration of the lack of strong intellectual property incentives for medical algorithms, especially for black-box algorithms, is outside the scope of this Article but is discussed in detail in Price, *supra* note 15.

177.   *See supra* Section III.C.2.

178.   *See* Cohen et al., *supra* note 56 (describing the use of predictive analytics to prioritize allocation of scarce hospital resources).

sophisticated versions, however, can be adapted over time: perhaps condition *Y* was an artifact of the original patient population and is actually irrelevant or actively unhelpful—the recommendation algorithm should then be modified to remove *Y*. If that algorithm were approved under the standard FDA device preapproval pathway, such a modification might require clinical trials, extensive regulatory filings, and potentially an entirely new preapproval process. This would substantially slow the development of adaptive algorithms that take account of new data.

In the broader context of medicine, one near-term goal is for a "learning health system."[179] In such a system, treatment and research are increasingly integrated rather than being strictly segregated.[180] As medicine is practiced, providers collect evidence about what treatments work, when, and how; that evidence can then be used to continuously update how providers practice medicine going forward.[181] Medical algorithms offer that capacity in a technological context; indeed, one way to actually develop a responsive learning health-care system is through the deployment of such algorithms, and vice versa.[182] If FDA uses the regulatory hook that these algorithms are medical devices, either themselves or when implemented on computers used in health care, the preapproval process stands to substantially limit the type of ongoing flexible innovation that algorithms enable for a learning health-care system.

The contours of market access and command-and-control regulation implicate a familiar debate about the speed with which FDA approves new technologies. Delay helps ensure safety, but it also limits care advances and patient access in the meantime.[183] The question here is whether, given the cost and speed advantages of black-box-algorithm development, the balance is better struck on the side of earlier access.

---

179.   *See, e.g.*, Harlan M. Krumholz, *Big Data and New Knowledge in Medicine: The Thinking, Training, and Tools Needed for a Learning Health System*, 33 Health Aff. 1163 (2014).

180.   *Id.* at 1164.

181.   *Id.* at 1167.

182.   *Id.* at 1168.

183.   This debate is long-standing. *See* Stephen Barlas, *Critics Assail FDA Medical Device Approval Process*, 36 J. Clinical Pharmacy & Therapeutics 395 (2011) (noting delays in the device approval process); Daniel P. Carpenter, *Groups, the Media, Agency Waiting Costs, and FDA Drug Approval*, 46 Am. J. Pol. Sci. 490 (2002) (providing an account of varying agency delays); Jonathan J. Darrow et al., *New FDA Breakthrough-Drug Category—Implications for Patients*, 370 New Eng. J. Med. 1252 (2014) (noting implications of FDA's faster approval pathway); Cassie Frank et al., *Era of Faster FDA Drug Approval Has Also Seen Increased Black-Box Warnings and Market Withdrawals*, 33 Health Aff. 1453 (2014) (noting safety costs of faster drug approval); J.D. Kleinke & Scott Gottlieb, *Is the FDA Approving Drugs Too Fast?*, 317 Brit. Med. J. 899 (1998) (noting debate about FDA speed versus safety); Samantha A. Roberts et al., *Despite Criticism of FDA Review Process, New Cancer Drugs Reach Patients Sooner in the United States than in Europe*, 30 Health Aff. 1375 (2011) (noting that FDA is faster than other parallel regulators).

B.   *Underregulation*

Coming from the other direction, there are also substantial problems with the possibility of FDA underregulation of algorithmic medicine. Part II described the need for regulation of algorithmic medicine; it has great potential to help but also to harm, and is difficult for nonregulators to evaluate, suggesting some role for regulation.

Nathan Cortez has described this problem in detail in the context of mHealth. Cortez argues that FDA's relative informality in regulating software has created substantial problems for the development of mHealth applications.[184] When regulators do not exercise strict oversight, firms have more room to produce poor-quality products. But the availability of those poor-quality products—before the market weeds them out—can result in injuries to customers and decreased consumer trust, slowing the process of high-quality technology development.[185]

Because black-box medicine is so new, we do not yet have examples of specific regulatory failures. But we do have potent examples of what can happen when diagnostic tests are poorly developed, whether in the mHealth context or outside it, especially with insufficient regulatory oversight.[186] In a recent study, for instance, two-thirds of insulin-calculation apps had a risk of outputting incorrect insulin dosages.[187] A 2017 review of advances in diabetes mHealth apps noted the "obvious oversupply of mHealth apps," described the difficulty for doctors and patients of choosing "a 'good' app for diabetes management," and lamented that most apps did not "meet the minimum requirements of an effective app" and could harm users.[188]

Unfortunately, many LDTs marketed under FDA's enforcement discretion have faced similar quality problems.[189] FDA collected several such case studies in 2015. For example, a test for Lyme disease incorrectly identified a substantial fraction of uninfected patients as having the disease.[190] Two couples were awarded a total of $30 million in damages for undergoing

---

184.   *See* Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1226–29.

185.   *Id.*; *see also* Cortez, *supra* note 85, at 453 ("FDA regulation can even be 'market-constituting,' in that it sustains consumer confidence that would otherwise erode if flooded with substandard products . . . ." (citing Daniel Carpenter, *Confidence Games: How Does Regulation Constitute Markets*, *in* Government and Markets: Toward a New Theory of Regulation 164 (Edward J. Balleisen & David A. Moss eds., 2009))).

186.   *See* U.S. Food & Drug Admin., The Public Health Evidence for FDA Oversight of Laboratory Developed Tests: 20 Case Studies (2015) [hereinafter FDA, The Public Health Evidence for FDA Oversight of LDTs], http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/UCM472777.pdf [https://perma.cc/6VRC-B3YK] (providing examples of failures from unregulated laboratory-developed tests).

187.   Kit Huckvale et al., *Smartphone Apps for Calculating Insulin Dose: A Systematic Assessment*, 13 BMC Med. 106 (2015).

188.   Farhad Fatehi et al., *Mobile Health (mHealth) for Diabetes Care: Opportunities and Challenges*, 19 Diabetes Tech. & Therapeutics 1 (2017).

189.   FDA, The Public Health Evidence for FDA Oversight of LDTs, *supra* note 186.

190.   *Id.* at 8–9.

"months of unnecessary treatment" based on such false positives.[191] Other problematic tests included tests for ovarian cancer with high false-positive rates, leading to the possibility of women needlessly choosing to have their ovaries removed;[192] a next-generation test for *HER2/neu* that missed some positive cases, with an estimated $775,278 cost of treatment foregone per false negative;[193] and a test to predict response to cholesterol-lowering statin drugs that relied on a genetic variation with no relationship to statin response—a lack of response only verified after over 150,000 tests had been performed on patients.[194]

The most dramatic recent story of quality problems for diagnostic tests is undoubtedly that of Theranos. Theranos promised to revolutionize diagnostic testing by taking tiny finger-stick blood samples, rather than large intravenous draws, and analyzing them for many possible problems.[195] Theranos's methods, including how precisely it performed its chemical reactions and the algorithms it used to analyze them, were kept completely secret.[196] Theranos did not publish quality controls for its tests, and since it used centralized testing procedures, for a long time it faced no FDA regulation as an LDT.[197] The entire enterprise came down when a journalist revealed widespread problems with the company's methods,[198] followed by a pair of devastating reports from FDA and CMS inspectors who found that Theranos's tests were inaccurate and unreliable.[199] Theranos's CEO, Elizabeth

---

191.   *Id.* at 9.

192.   *Id.* at 9–12.

193.   *Id.* at 14–15.

194.   *Id.* at 19–20.

195.   John Carreyrou, *Hot Startup Theranos Has Struggled With Its Blood-Test Technology*, Wall Street J. (Oct. 16, 2015, 3:20 PM), https://www.wsj.com/articles/theranos-has-struggled-with-blood-tests-1444881901 (on file with the *Michigan Law Review*).

196.   *See* John P.A. Ioannidis, *Stealth Research: Is Biomedical Innovation Happening Outside the Peer-Reviewed Literature?*, 313 JAMA 663 (2015).

197.   Theranos stated that it submitted all of its LDTs to FDA for approval, though it offered these tests before receiving any approval. *See* Letter from Elizabeth Holmes, Chief Exec. Officer, Theranos, to Margaret Hamburg, Comm'r of Food & Drugs, FDA (Mar. 1, 2015) (on file with author). Eventually, Theranos was granted FDA approval for its herpes simplex test under the 510(k) pathway in 2015. Letter from Sally A. Hojvat, Dir., Div. of Microbiology Devices, Dep't of Health & Human Servs., to Brad Arington, Assoc. Dir., Regulatory, Theranos (July 2, 2015), https://s3.amazonaws.com/s3.documentcloud.org/documents/2157669/k143236-letter-se-final-corrected-2015-07-02.pdf [https://perma.cc/7FQE-P2GP].

198.   Carreyrou, *supra* note 195.

199.   Ctrs. for Medicare & Medicaid Servs., Statement of Deficiencies for Theranos, Inc. (2015), https://assets.documentcloud.org/documents/2781859/Theranos-2567-With-Redactions-03-31-2016.pdf [https://perma.cc/3FDH-XDV8]; U.S. Food & Drug Admin., Inspection Report for Theranos, Inc. (2015) [hereinafter FDA, Inspection Report for Theranos, Inc.], http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-afda-orgs/documents/document/ucm469395.pdf [https://perma.cc/U5BL-FDGQ]. FDA noted that Theranos's "[n]anotainer" for capturing tiny blood samples as a Class II "uncleared medical device" which was shipped in interstate commerce. FDA, Inspection Report for Theranos, Inc., *supra*, at 1.

Holmes, has been barred from clinical diagnostic testing for two years.[200] Other companies relying on complex diagnostics similarly rely on the laboratory-developed-test exception to avoid oversight,[201] and FDA has collected examples of other problems arising from this practice.[202]

Underregulation may be especially likely when algorithm developers are the only ones who have or can provide information about how the algorithms were developed or trained.[203] Developers may have incentives to downplay potential problems in the quest for FDA approval.[204] In addition, whatever biases or mistakes were present in initial algorithm development may color developers' ability to verify their own algorithms and to identify problems.[205] Finding problems may depend on the ability of FDA—with its potential lack of expertise—or third parties to independently evaluate algorithm quality.[206]

V.   Reform

FDA faces substantial challenges in regulating algorithmic medicine, but it can and should still play an important role in that regulation. As described above, practical hurdles and potentially dampening effects on innovation limit the benefits of an approach that relies on strict classification into high-risk pre-approval categories.

So how should FDA regulate algorithmic medicine? To suggest answers, I turn to the growing literature on new governance, which explores and theorizes new techniques of regulation. Precise definitions of the new governance movement have proven somewhat contentious,[207] but broadly speaking, it addresses some moves "away from the familiar model of command-style,

---

200.   Andrew Pollack, *Elizabeth Holmes of Theranos Is Barred from Running Lab for 2 Years*, N.Y. Times (July 8, 2016), https://www.nytimes.com/2016/07/09/business/theranos-elizabeth-holmes-ban.html (on file with the *Michigan Law Review*). Theranos is appealing the decision. Sarah A. O'Brien, *Theranos Fights Back Against Sanctions*, CNN: Tech (Aug. 26, 2016, 9:03 AM), http://money.cnn.com/2016/08/26/technology/theranos-appealing-cms-sanctions/ [https://perma.cc/C5AB-JFSQ].

201.   *See* Arielle Duhaime-Ross, *Theranos Isn't the Only Diagnostics Company Exploiting Regulatory Loopholes*, Verge (Nov. 11, 2015, 8:28 AM), http://www.theverge.com/science/2015/11/11/9706356/fda-theranos-health-diagnostics-cancer-tests-regulation-loophole-ldt [https://perma.cc/XTJ2-KZXG] (describing three companies performing complex cancer genetics tests designed to direct cancer treatment).

202.   FDA, The Public Health Evidence for FDA Oversight of LDTs, *supra* note 186.

203.   *See* Ford & Price, *supra* note 68, at 12.

204.   *See id.* at 15 n.52.

205.   *Id.* at 14–15.

206.   *Id.* at 19.

207.   *Compare* Orly Lobel, *The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought*, 89 Minn. L. Rev. 342, 346–47 (2004) (noting that New Governance brings together into one field insights from scholarship on "'reflexive law,' 'soft law,' 'collaborative governance,' 'democratic experimentalism,' 'responsive regulation,' 'outsourcing regulation,' 'reconstitutive law,' 'post-regulatory law,' 'revitalizing regulation,' 'regulatory pluralism,' 'decentering regulation,' 'meta-regulation,' 'contractarian law,'

fixed-rule regulation by administrative fiat, and toward a new model of col-
laborative, multi-party, multi-level, adaptive, problem-solving New Govern-
ance."[208] Orly Lobel provides a (similarly contestable) set of organizing
principles that can also be thought of as tools for new governance: (1) par-
ticipation and partnership; (2) collaboration; (3) diversity and competition;
(4) decentralization and subsidiarity; (5) integration of policy domains; (6)
flexibility and noncoerciveness (or softness-in-law); (7) fallibility, adaptabil-
ity, and dynamic learning; and (8) law as competence and orchestration.[209]

   Though new governance offers a useful set of tools that could help im-
prove potential FDA regulation of black-box medicine, I am not arguing the
full set should be adopted uncritically. Indeed, new governance has been
subject to substantial criticism, including concerns about whether it is ap-
propriate for all contexts.[210] It has also been described as lacking a mecha-
nism to resolve normative conflicts;[211] in this context, new governance likely
cannot tell us where the appropriate balance is between risk and safety, in-
novation and caution. I do not aim to address these larger concerns with the
new governance project; instead, I seek to borrow a few tools that seem
particularly suited to the context of FDA regulation of black-box medicine.

   Broadly speaking, I argue that FDA should still exercise centralized com-
mand-and-control regulatory oversight. But a better approach for black-box
medicine would combine more moderate up-front regulation—graded by
risk but with lower barriers than the full premarket approval pathway—with
robust postmarket surveillance to monitor the performance of algorithms in
real-world settings. Crucially, this latter role should not fall to FDA alone.
Instead, other players in the health care ecosystem should collaboratively
help to monitor algorithmic safety and efficacy. For providers, hospital sys-
tems, and insurers to play a meaningful role in the postmarket oversight of
medical algorithms, they need access to meaningful information about those
algorithms. There, FDA should supplement its command-and-control regu-
lation by facilitating the sharing of information to those other players to

---

'communicative governance,' 'negotiated governance,' 'destabilization rights,' 'cooperative im-
plementation,' 'interactive compliance,' 'public laboratories,' 'deepened democracy and em-
powered participatory governance,' 'pragmatic lawyering,' 'nonrival partnership,' and 'a
daring legal system'" (footnotes omitted)), *with* Bradley C. Karkkainen, Reply, *"New Govern-
ance" in Legal Thought and in the World: Some Splitting as Antidote to Overzealous Lumping*, 89
Minn. L. Rev. 471, 479–81 (2004) (criticizing Lobel for largely failing to "acknowledge[ ] any
debate, difference, or diverging tendency among New Governance scholars").

   208.   Karkkainen, *supra* note 207, at 473.

   209.   Lobel, *supra* note 207, at 342, 371–404.

   210.   *See, e.g.,* Grainne de Burca, *New Governance and Experimentalism: An Introduction*,
2010 Wis. L. Rev. 227, 235–38 (canvassing some concerns about new governance); Martha
Minow, *Public and Private Partnerships: Accounting for the New Religion*, 116 Harv. L. Rev.
1229 (2003) (describing concerns with public and private partnerships).

   211.   *See, e.g.*, Samuel R. Bagenstos, *The Structural Turn and the Limits of Antidiscrimina-
tion Law*, 94 Calif. L. Rev. 1, 34–40 (2006) (criticizing a pragmatist-experimentalist approach
in the employment discrimination context); Lobel, *supra* note 207, at 455 (noting the illusion
that "comprehensive and widespread information on an issue will eventually lead people to
converge normatively on the same positions").

enable robust, flexible oversight. This Section describes each of these elements and the extent to which they draw on new governance tools; it does not address in detail the extent to which these reforms could be implemented under FDA's current statutory authority.[212]

A. *Command-and-Control Regulation*

FDA cannot and should not abandon its command-and-control role in directly regulating premarket access for at least some forms of algorithmic medicine. Although command-and-control regulation has been seriously criticized,[213] the challenges of evaluating health-related credence goods suggest at least some role for FDA-directed mandatory regulation.[214] Shaping preapproval requirements to risk level also fits the contours of new medical technology and comports with FDA's traditional gatekeeping role.[215] Particularly high-risk algorithms—those that directly influence critical treatment choices, such as whether to perform surgery or to prescribe chemotherapy—should probably face more stringent preapproval review than less risky algorithms. FDA works in a risk framework; its proposed laboratory-developed test framework applies a risk-based framework,[216] and it is exercising something like a hands-off approach to mobile-health algorithms except those that are directly involved in directing or replacing other FDA-approved medical devices.[217]

But in following this framework, FDA is still relying too much on categorical rules rather than a more flexible, adaptive approach. One strain of new governance offers the idea of such strategies, variously described as "incremental," or "experimentalist." The experimentalist subset of new governance scholars, including Charles Sabel and Michael Dorf, emphasizes the idea of devolving authority to more local levels, measuring the results, and adopting the most successful.[218] Here, I propose that a more modest—and

---

212. Whether FDA can regulate algorithms standing alone depends on their status as regulable medical devices, as addressed in Section III.B. The mechanisms by which FDA regulates algorithms similarly depends on current law (which establishes the classes of medical devices, for instance). FDA does have a great deal of flexibility, but this Article does not address which potential reforms arguably fit within that flexibility and which would require new statutory authority. Those questions remain for future work.

213. *See, e.g.*, Bruce A. Ackerman & Richard B. Stewart, Comment, *Reforming Environmental Law*, 37 Stan. L. Rev. 1333, 1335–37 (1985) (critiquing command-and-control regulation as expensive, rigid, and inefficient).

214. *See supra* note 49 and accompanying text.

215. *See generally* Carpenter, *supra* note 168, at 1–32 (describing FDA's role as gatekeeper).

216. FDA, LDT Draft Guidance, *supra* note 3, at 3.

217. *See* FDA, Mobile Medical App Guidance, *supra* note 100, at 4.

218. *See* Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 Colum. L. Rev. 267, 314–23 (1998); Lobel, *supra* note 207, at 379–81.

older—incremental approach may be more appropriate.[219] FDA can begin with a tentative policy on regulating black-box algorithms, then proceed iteratively, adapting that policy to new information as the technology develops.[220] Such a strategy especially suits the rapidly developing technology of black-box medical algorithms.[221]

For now, it looks like FDA is moving in the opposite direction, leaning toward a categorical approach.[222] FDA's proposed rules on laboratory-developed tests classified any tests used for diagnostics or using nontransparent algorithms as being categorically high-risk and therefore subject to the full PMA process.[223] But opacity can vary independently of risk. Tuning preapproval requirements to opacity—that is, applying more stringent requirements to nontransparent algorithms, whatever their underlying risk—means that more stringent rules will apply to those algorithms developed as black boxes precisely because that approach allows more flexible, quicker, cheaper development.

In the mobile-health context, FDA may err too much on the side of categorical permissiveness. Mobile-health applications can do substantial mischief without interfacing directly with medical devices or taking their place by, for instance, erroneously suggesting that patients avoid taking a particular medication, cancel a provider's appointment as unnecessary, or exercise strenuously in the face of cardiac warning signs.[224]

Striking this balance may seem a Goldilocksian task, but this is precisely the point of an incremental, adaptive response that responds to new information as the technology develops.[225] But even an incremental approach assumes some initial decision rule: What is the initial approach which will be modified once new information is received? In the case of black-box medicine, a baseline that opaque algorithms are Class III devices requiring

---

219. *See* Colin S. Diver, *Policymaking Paradigms in Administrative Law*, 95 Harv. L. Rev. 393, 399–409 (1981); Charles E. Lindblom, *The Science of "Muddling Through"*, 19 Pub. Admin. Rev. 79 (1959).

220. Diver, *supra* note 219, at 399–409. For an account of adaptive regulation, see, for example, Craig & Ruhl, *supra* note 6, which discusses how administrative law can take account of an iterative, adaptive management strategy.

221. *See* Sharon B. Jacobs, *The Energy Prosumer*, 43 Ecology L.Q. 519, 572 (2017) (noting that incremental or experimentalist approaches are appropriate for "complex and rapidly developing technology"). *But see* Cortez, *The Mobile Health Revolution?*, *supra* note 27, at 1222–23 (arguing that FDA should take a more stringent, brighter-line regulatory tack to discipline industry and to allow it to grow within a framework of more predictable rules).

222. It is hard to state definitively what approach FDA might or might not adopt, or might even be adopting already; I base these analyses on the information available through guidance and agency action.

223. *See supra* Section III.D.2.

224. *See, e.g.*, Torie Bosch, *mHealth Care Crisis: Should the FDA Regulate Smartphone Medical Apps?*, Slate (Aug. 2, 2011, 12:29 PM), http://www.slate.com/articles/technology/future_tense/2011/08/mhealth_care_crisis.html [https://perma.cc/F32Z-Z7V7].

225. *See* Lobel, *supra* note 207, at 388–95 (describing "soft law" concepts that can be applied alongside "hard law" regulations and rules, including an explicit goal of regulatory flexibility).

premarket approval would hamper the development of black-box medicine quite a bit. Given the promise of black-box medicine, and the realities of existing flaws in the medical system, a baseline of lighter-touch regulation seems a better approach.

One possible alternate approach to premarket access could rely on regulating algorithmic medicine by developer, rather than by individual algorithm. This would be similar to CMS's Clinical Laboratory Improvement Amendments framework, which requires laboratory certification based on the complexity of testing performed at facilities.[226] Certification requirements (including personnel qualifications and quality control) vary based on test complexity.[227] Adopting a similar system for black-box medicine might encourage developers to create and test a wide variety of algorithms under appropriate regulatory surveillance. FDA may be open to such an approach; when the agency approved 23andMe's first direct-to-consumer genetic test in spring 2017, it approved an initial slate of ten tests and stated that it "intends to exempt additional 23andMe [genetic-health-risk] tests from FDA's premarket review, and [that such] tests from other makers may be exempt after submitting their first premarket notification."[228] FDA noted that such an approach "would allow other, similar tests to enter the market as quickly as possible and in the least burdensome way, after a one-time FDA review."[229] FDA has expressed some interest in expanding this approach to some medical software, describing in mid-2017 a possible program where trusted software developers could face lighter premarket security.[230]

An incremental or experimentalist approach is not without its risks. The risks of low-quality algorithms have been discussed above.[231] Procedurally, incrementalism risks the possibility that the status quo becomes entrenched and adaptation fails to occur.[232] To help avoid these risks, both in terms of patient safety and in terms of procedural stagnation, a lighter touch on

---

226. Ctrs. for Medicare & Medicaid Servs., Clinical Laboratory Improvement Amendments (CLIA) and Medicare Laboratory Services 2–3 (2015), https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/CLIABrochure-TextOnly.pdf [https://perma.cc/MWH9-R73R].

227. *Id.*

228. Press Release, U.S. Food & Drug Admin., FDA Allows Marketing of First Direct-to-Consumer Tests that Provide Genetic Risk Information for Certain Conditions (April 6, 2017), https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm55185.htm [https://perma.cc/SJ7C-N28B].

229. *Id.*

230. U.S. Food & Drug Admin., Digital Health Innovation Action Plan 5–6 (2017), http://www.fda.gov/downloads/MedicalDevices/DigitalHealth/UCM568735.pdf [https://perma.cc/E5ZP-K98J].

231. *See supra* Section IV.B.

232. *See, e.g.*, Cristie Ford, *New Governance in the Teeth of Human Frailty: Lessons from Financial Regulation*, 2010 Wis. L. Rev. 441, 470–71 ("[I]t would be unwise to underestimate the amount of energy and focus required to push incremental change (or even to identify its direction, given the background noise) in the direction of prior commitments and empirically demonstrable improvement.").

premarket access—however implemented—should be complemented by continuing oversight once products are in use.

B.   *Postmarket Surveillance*

Relying solely on premarket regulation is likely insufficient to strike the right balance between innovation and safety for medical algorithms. A substantial strength of medical algorithms is that they can be developed relatively quickly, and at relatively low cost, and that they can change over time—this last especially true in the case of machine-learning-based blackbox medicine. Premarket approval limits these strengths, but postmarket surveillance complements them. When algorithms are based on existing data and observational studies, observing how those algorithms perform in the real world is critical to ensuring their safety and efficacy. New governance offers tools in this context, emphasizing the importance of flexibility and an iterative approach.[233] As Cristie Ford notes, "Learning by doing is the method, but it needs to be accompanied by actual mechanisms that make it possible for regulation to move."[234]

There is a useful synergy here. Medical algorithms based on observational data can themselves be improved by observing their performance and then incorporating more data. For essentially static algorithms, this can be a step-by-step process where, once enough data have been collected, the algorithm may be reworked and a new version distributed.[235] For dynamic black-box algorithms that flexibly update when given new data, constantly feeding such data back into the system should help improve performance over time.[236]

The idea of continuous feedback for medical algorithms meshes tightly with the idea of a learning health system.[237] In a learning health system, the distinction between clinical research and clinical care is blurred as data about care are folded back into the task of improving care going forward. In the context of medical algorithms, that ongoing information can be used both to improve the algorithms and to evaluate their performance, noting problems or flaws as they occur.

---

233.   *Id.* at 446 ("Flexibility is a key characteristic of new governance methods.").

234.   *Id.*

235.   The analogy to alpha and beta testing of software is clear; though patients may not appreciate being seen as alpha or beta testers, they already have this status de facto for many medical technologies. *See* Jonathan J. Darrow, *Crowdsourcing Clinical Trials*, 98 Minn. L. Rev. 805, 809–11 (2014) ("Because risks remain and data continues to be collected, new drugs are undergoing de facto 'human testing' *after* receiving the FDA's seal of approval.").

236.   This idea has a parallel for physical medical devices that are incrementally changed over the product's lifespan, making premarket review of each iteration impractical. *See* Frederic S. Resnic & Sharon-Lise T. Normand, *Postmarketing Surveillance of Medical Devices—Filling in the Gaps*, 366 New Eng. J. Med. 875 (2012) (proposing comprehensive review of postmarket surveillance systems to better account for the realities of device development and evolution).

237.   *See supra* note 179 and accompanying text.

FDA has already shown a willingness to engage in this type of pre- and postmarket blending in some contexts.[238] During premarket approval review of medical devices, the agency considers what clinical data are necessary to establish safety and effectiveness before approval, and what data can instead be collected postmarket.[239] FDA also weighs the probable benefits and risks of the device.[240] Recognizing the challenges devices face in achieving timely premarket approval, the agency created an Expedited Access Pathway (EAP) program in 2015.[241] Under EAP, qualified devices receive extra support and priority review from FDA.[242] To be considered for EAP, devices must "address unmet medical needs for life threatening or irreversibly debilitating diseases or conditions."[243] When exercising premarket approval over EAP devices, FDA is willing to accept greater upfront uncertainty regarding risks and benefits, given the importance of patient access to these critical medical devices.[244] Instead, FDA may use more extensive postmarket controls for EAP devices (including data collection).[245]

FDA's embrace of some premarket uncertainty when evaluating EAP devices provides a useful model for developing a similar program for medical algorithms. A limited number of algorithms might qualify under current EAP criteria, but FDA could include a far greater number of algorithms and allow for safe innovation by addressing complex medical algorithms specifically. The 21st Century Cures Act has laid some useful groundwork by expanding FDA's ability to consider more types of information in postmarket

---

238. U.S. Food & Drug Admin., Balancing Premarket and Postmarket Data Collection for Devices Subject to Premarket Approval: Guidance for Industry and Food and Drug Administration Staff (2015) [hereinafter FDA, Balancing Premarket and Postmarket Data Collection for Devices Subject to Premarket Approval], http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm393994.pdf [https://perma.cc/7JTF-SYWP]; *see also* 21 U.S.C. § 360c(a)(2)(C), 360c(a)(3)(D)(ii) (2012) (enabling such balancing).

239. FDA, Balancing Premarket and Postmarket Data Collection for Devices Subject to Premarket Approval, *supra* note 238, at 6–7.

240. U.S. Food & Drug Admin., Factors to Consider Regarding Benefit-Risk in Medical Device Product Availability, Compliance, and Enforcement Decisions: Guidance for Industry and Food and Drug Administration Staff (2016), http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm506679.pdf [https://perma.cc/7VBL-MVMA].

241. U.S. Food & Drug Admin., Expedited Access for Premarket Approval and *De Novo* Medical Devices Intended for Unmet Medical Need for Life Threatening or Irreversibly Debilitating Diseases or Conditions 3–4 (2015) [hereinafter FDA, Expedited Access for Premarket Approval], http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm393978.pdf [https://perma.cc/L6X6-37KB].

242. *Id.* at 18–19.

243. *Id.* at 4, 9.

244. *Id.* at 8–9; *cf.* Darrow et al., *supra* note 183 (describing FDA's accelerated drug approval program based on earlier-observable surrogate endpoints).

245. FDA, Expedited Access for Premarket Approval, *supra* note 241, at 8–9.

surveillance activities, including "real world evidence."[246] The Act also requires that FDA "consider the role of postmarket information in determining the least burdensome means of demonstrating a reasonable assurance of device safety and effectiveness."[247] Taken together, these provisions suggest support for the possibility of continuous monitoring in a learning health system, comporting with the new governance ideas of flexibility and iterative learning.[248] Along these lines, FDA is currently developing the National Evaluation System for Health Technology (NEST), intended to "generate evidence across the total product lifecycle of medical devices by strategically and systematically leveraging real-world evidence, and applying advanced analytics to data tailored to the unique data needs and innovation cycles of medical devices."[249]

There is a challenge with relying on postmarket surveillance; although FDA has long had postmarket-surveillance programs, they haven't tended to work especially well.[250] Compliance is low, especially because it is often voluntary and FDA lacks adequate enforcement resources.[251] This problem tracks the general concern described above that incremental approaches can be sticky at the status quo—though likely not as sticky as traditional command-and-control regulation.[252] So what would make postmarket surveillance work better for algorithms? Three things might help. First, Congress could give FDA additional authority and resources to make and enforce surveillance requirements.[253] Second, surveillance and monitoring should become easier as health-information systems become more integrated and data are easier to flag and share—and FDA can itself help this integration. FDA's Sentinel Initiative is one step in this direction, aiming to capture safety data from over 100 million Americans;[254] broader health-information integration

---

246.   *See* 21st Century Cures Act, Pub. L. No. 114-255, § 3022 (requiring FDA to "establish a program to evaluate the potential use of real world evidence" to fulfill post-approval surveillance requirements). Commentators have criticized this provision. *See, e.g.*, Jerry Avorn & Aaron S. Kesselheim, *The 21st Century Cures Act—Will It Take Us Back in Time?*, 372 NEW ENG. J. MED. 2473 (2015).

247.   21st Century Cures Act, Pub. L. No. 114-255, § 3058(b) (amending 21 U.S.C. § 360(e)(c)(5)(C) (2012)).

248.   *See, e.g.*, Lobel, *supra* note 207, at 388–400.

249.   *National Evaluation System for Health Technology*, U.S. FOOD & DRUG ADMIN., http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDR-HReports/ucm301912.htm [https://perma.cc/V9ND-XZP9].

250.   *See* Phil B. Fontanarosa et al., *Postmarketing Surveillance—Lack of Vigilance, Lack of Trust*, 292 JAMA 2647 (2004) (noting problems with postmarket surveillance system).

251.   *Id.*

252.   *See* Ford, *supra* note 232, at 470–71 (noting the difficulty of incremental change in a new governance model); Jason M. Solomon, *Law and Governance in the 21st Century Regulatory State*, 86 TEX. L. REV. 819, 822 (2008) (book review) ("[In] the command-and-control model the mechanisms for monitoring and adjusting the rules in light of experience are severely lacking.").

253.   *See* Resnic & Normand, *supra* note 236.

254.   Melissa Robb, *FDA's Mini-Sentinel Exceeds 100 Million Lives (and Counting) . . . a Major Milestone in Developing a Nationwide Rapid-Response Electronic Medical Product Safety*

should help with such surveillance, among other benefits.[255] Third, postmarket surveillance will be more effective if it involves not only FDA, but also other sophisticated health-market actors, a topic discussed in the next Section.

C.   *Information Forcing and Collaborative Governance*

To best regulate complex medical algorithms, FDA should turn to collaborative forms of governance involving other health-system actors. Collaborative governance requires information which is often now unavailable; to enable such collaborative governance, FDA could play an information-forcing role. Both of these concepts borrow from new governance.

Jody Freeman and other new governance scholars suggest a role for collaboration by different stakeholders, including industry.[256] Under this approach, regulators and other stakeholders are engaged in "continuous interaction and sharing of responsibility."[257] This approach recognizes that regulators lack perfect knowledge, and that others may have useful insights to contribute.[258] Although I do not suggest that FDA abdicate its command-and-control role, the agency could also play a coordinating function to enable such collaborative regulation by private actors, whether through input to FDA actions or through their own forms of parallel regulation.[259]

Which actors could be involved in such collaboration? Independent verification of algorithms by third-party expert developers, relying on information used to develop the algorithms, could help buttress FDA's own regulation; Roger Ford and I have discussed this approach extensively elsewhere.[260] But FDA can also enable real-world evaluation undertaken by health-system actors, including providers, hospitals, and insurers. The involvement of different health-system actors carries at least two benefits.

Surveillance Program, FDA Voice (June 29, 2012), https://blogs.fda.gov/fdavoice/index.php/2012/06/fdas-mini-sentinel-exceeds-100-million-lives-and-counting-a-major-milestone-in-de veloping-a-nationwide-rapid-response-electronic-medical-product-safety-surveillance-pro-gram/ [https://perma.cc/4VXQ-PXA7].

255.   *See* Rebecca S. Eisenberg & W. Nicholson Price II, *Promoting Healthcare Innovation on the Demand Side*, 4 J.L. & Bioscience 3, 14–23 (2017) (describing potential innovation benefits of health information integration).

256.   *See, e.g.*, Freeman, *supra* note 7 (proposing a model of collaborative governance); Jody Freeman, *The Private Role in Public Governance*, 75 N.Y.U. L. Rev. 543 (2000) (describing regulation by private parties on behalf of agencies); Lobel, *supra* note 207, at 376–79 (discussing collaboration as a principle of new governance); Anne Joseph O'Connell, *Bureaucracy at the Boundary*, 162 U. Pa. L. Rev. 841 (2014) (describing bureaucracy at the boundary between public and private).

257.   Lobel, *supra* note 207, at 377.

258.   *See, e.g.*, Ford, *supra* note 232, at 446–47 (recognizing that "humility about knowability" and "epistemological uncertainty" are linked to new governance).

259.   *See id.* at 445–46.

260.   *See* Ford & Price, *supra* note 68, at 30–31 (describing the benefits of independent third-party algorithm validation and noting the challenge of balancing third-party access to underlying data with patient privacy concerns).

First, different actors have different areas of expertise and may interact with algorithms in different contexts.[261] Second, to the extent that collaborative regulation invokes fears of agency capture, the involvement of health-system actors with distinctly different incentives may help allay those concerns.[262]

Insurer reimbursement decisions, provider adoption, and hospital choice can all help regulate forms of algorithmic medicine. But those forms of parallel regulation require accurate information to function well. FDA could enable each of these actors by requiring that those developing algorithmic medicine disclose substantial information about their algorithms rather than keeping them totally secret.[263] In the environmental context, Bradley Karkkainen has argued that government can play an information-forcing role through penalty default rules.[264] Here, I suggest instead directly requiring information disclosure both to the regulator and to other parties.

In this framework, FDA could require as a condition of marketing that developers disclose how the algorithm was developed, the data used for that development, and—to the extent known—how the algorithm works. FDA could also mediate information sharing of postmarket-surveillance information. Such mandatory sharing would likely require new statutory authority.[265]

Ideally, the agency would not only help enable those private actors to perform a private regulatory function in parallel but also learn from them in an ongoing collaborative enterprise. Indeed, FDA has suggested that it may be open to such an approach with respect to future regulation of laboratory-

---

261.   *See* Lobel, *supra* note 207, at 379–81 (describing "diversity and competition" as key elements of new governance).

262.   *See* Eisenberg & Price, *supra* note 255, at 17–18 (discussing the different incentives of insurers and other health-system actors).

263.   Mandatory disclosure regimes have been suggested in many contexts and have come under criticism for failing to actually meaningfully inform consumers. *See, e.g.*, OMRI BEN-SHAHAR & CARL E. SCHNEIDER, MORE THAN YOU WANTED TO KNOW: THE FAILURE OF MAN-DATED DISCLOSURE 33–54 (2014); *see also* Ryan Bubb, *TMI? Why the Optimal Architecture of Disclosure Remains TBD*, 113 MICH. L. REV. 1021 (2015) (reviewing BEN-SHAHAR & SCHNEI-DER, *supra*) (arguing that some forms of disclosure may still be effective). This proposal differs from standard disclosure regimes in that first, it would be mediated by a strong regulator; and second, the intended users of the information are not the final consumers—patients—but rather, other sophisticated intermediaries such as hospitals and insurers.

264.   *See* Bradley C. Karkkainen, *Information-Forcing Environmental Regulation*, 33 FLA. ST. U. L. REV. 861, 868–71 (2006) (extending Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 YALE L.J. 87, 91–93 (1989), to the regulatory context).

265.   This lack of statutory authority is not definite; one could argue creatively that FDA's authority goes that far—and the agency has been creative in its assertions of authority in the past. *See, e.g.*, FDA v. Brown & Williamson Tobacco Corp., 529 U.S. 120 (2000) (refusing to accept FDA's attempt to regulate tobacco products despite years of disavowing such authority). A rigorous analysis of this issue must await future work. Some precedent for mandatory disclosure does exist. *See* W. Nicholson Price II & Arti K. Rai, *Manufacturing Barriers to Biologics Competition and Innovation*, 101 IOWA L. REV. 1023, 1053 (2016) (discussing mandatory disclosure of manufacturing methods for biologics as part of the Biologics Price Competition and Innovation Act).

developed tests, noting in a 2017 discussion paper that the agency is open to the possibility of laboratory accreditation by third parties and promoting the idea of "clinical collaboratives" to, among other things, "crowdsource evidence to demonstrate clinical validity for specific types of tests."[266] FDA has also expressed some interest in exploring potential roles for third-party certification of medical software, thought it has not considered black-box algorithms explicitly.[267] The next three subsections address the possibilities of FDA engaging in collaborative regulation with providers, hospitals, and insurers.

1.   Providers

FDA-mandated information disclosure would increase the ability of providers to exercise control over the quality of algorithms. Providers are the traditional gatekeepers of medical technology. Providers order diagnostic tests, prescribe drugs, and administer treatments or send patients to specialists. At least ideally, doctors, nurses, and others who provide care possess the knowledge and training to evaluate and make choices between medical products (though of course sometimes they may not possess or exercise this knowledge). FDA regulation of drug-related information thus focuses on providers; FDA states explicitly that drug labels should provide clinical information "that practitioners consider important to clinical decision making."[268] The system of medical tort law also recognizes providers' key role in choosing technologies through the doctrine of learned intermediaries.[269]

---

266.   *See* FDA, Discussion Paper on LDTs, *supra* note 166, at 6–7.

267.   Scott Gottlieb, *Fostering Medical Innovation: A Plan for Digital Health Devices*, FDA Voice (June 15, 2017), https://blogs.fda.gov/fdavoice/index.php/2017/06/fostering-medical-innovation-a-plan-for-digital-health-devices/ [https://www.perma.cc/SE2D-CDY8].

268.   FDA, Label Guidance, *supra* note 70, at 2. Off-label drugs do not have the same information, of course. In off-label use, drugs are prescribed or administered in ways that are not included in the drug's label—that is, that have not been approved by FDA for use in that particular way. *See generally Understanding Unapproved Use of Approved Drugs "Off Label"*, U.S. Food & Drug Admin. (June 6, 2016), https://www.fda.gov/ForPatients/Other/Offlabel/default.htm [https://perma.cc/D6Z4-QS5M]. Some estimates are that off-label use may comprise 21% of all drug use. Randall S. Stafford, *Regulating Off-Label Drug Use: Rethinking the Role of the FDA*, 358 New Eng. J. Med. 1427 (2008) (citing David C. Radley et al., *Off-Label Prescribing Among Office-Based Physicians*, 166 Archives Internal Med. 1021 (2006)). Chemotherapy is replete with off-label treatments, and many drugs are prescribed off-label for children because manufacturers often do not conduct pediatric trials. *Id.*

269.   Under this doctrine, the treating provider is treated as the end user of a medical product because patients rely on the provider, not the manufacturer, and the provider can be "reasonably rel[ied] upon to pass on the warnings." *See* 2 Dan B. Dobbs et al., The Law of Torts § 466 (2d ed. 2011). Accordingly, manufacturers owe no duty directly to the patient, but have a duty adequately to warn the provider of potential dangers. *Id.* The learned intermediary doctrine is in substantial tension with the idea of complex medical algorithms. If providers cannot understand or independently reach the conclusions of an algorithm, which may fall outside the traditional standard of care, does the provider truly fill the role of a learned intermediary? A full account of this tension must await future work.

Currently, providers lack information about algorithms' mechanisms or development. Given that information—in a usable form—they could potentially help evaluate algorithm quality, much as they can do now when choosing between drugs or other treatment options.

While individual providers are traditional evaluators of medical technology, they should not be the principal evaluators of black-box algorithms. Theoretically, individual provider decisions could mediate the quality of algorithms used and also generate ex ante market pressure for high-quality development in the first place. But this possibility relies on both the availability of information and the ability of providers to evaluate that information. Providers are likely poorly equipped, at least for now, to directly evaluate the quality of medical algorithms. Specialized medical associations could, help alleviate this concern through practice guidelines, with adequate information. For instance, the American College of Radiology, which publishes guidelines on radiology practices,[270] could issue guidelines on the selection of high-quality algorithms to identify tumors in radiological images.[271] Such associations may be better at evaluating products ex ante rather than continuing to monitor product quality after introduction, but could potentially play either role. This collaborative governance process could help information flow in both directions; not only would FDA mediate information disclosure to providers, but provider and provider-group experiences and evaluations could also inform future decisions by FDA.

2.   Hospital Systems

Hospitals (or larger health systems) provide another opportunity for collaborative regulation. Hospitals are key gatekeepers for large medical systems, including information-technology systems. Hospitals may exert relatively little influence over the choice of particular drugs and medical devices, because those choices are typically left to providers to make directly from the set of available technologies. A substantial subset of algorithms, however, may be embedded in or associated with larger-scale hospital equipment, or associated with hospital information-technology and health-records systems.

In those cases, hospitals may be the principal decisionmakers in the choice of adapting certain types of medical algorithms. Presumably, they could take on the role of evaluating them before adoption. While hospitals are often large and sophisticated actors, their ability to judge between different medical algorithms and to evaluate quality is currently limited by the fact that much information is kept proprietary.[272]

---

270.   *See, e.g.*, Am. Coll. of Radiology, Best Practices Guidelines on Imaging Clinical Decision Support Systems (2012), https://www.acr.org/~/media/ACR/Documents/PDF/Economics/Managed-Care/Best-Practices-Guidelines-for-Imaging-Clinical-Decision-Support-Systems1.pdf [https://perma.cc/58BB-YDDJ].

271.   Of course, such evaluations carry the possibility of conflicting interests; the group of professional radiologists might not welcome the task of evaluating algorithms that could automate part of their job. *See, e.g.*, Molteni, *supra* note 46.

272.   *See supra* Section II.B.

Information sharing mediated by FDA could alleviate this concern and allow hospitals and health systems to more effectively evaluate algorithms, both before adoption and once they are integrated into practice. As large, sophisticated entities that implement and run information-technology systems, hospitals may be ideally placed to help evaluate algorithms before they are put into place.[273]

3.   Insurers

Finally, FDA's information-forcing role could also help enable improved collaborative decisionmaking on the part of insurers (defined broadly) for reimbursement purposes.[274] Insurers can serve as technological gatekeepers by deciding what technology will be reimbursed and for how much, within limits. Often, insurers only cover new technologies if their efficacy is demonstrated.[275] Without a reimbursement code, providers are much less likely to adopt a new technology.[276] Insurers' choices therefore help determine which technologies are adopted and which are not, consequently allowing them the opportunity to evaluate new technology.

To be effective, insurers' gatekeeping must account for ongoing complexities in the reimbursement of algorithms and diagnostic tests—that is, how and whether medical algorithms are reimbursed at all, and how they are paid for more generally.[277] Nonetheless, eliding those complexities, insurers

---

273.   *See* Price, *supra* note 75, at 13 (drawing a parallel between hospitals' duty to select qualified physicians and a potential duty to implement high-quality black-box medical algorithms).

274.   I use the term "insurer" for convenience, but this role could be filled by any large final payer for medical care, whether technically an insurer or not. The broader questions of how insurer reimbursement will work for black-box medicine are more complicated. For an initial take, see Price, *supra* note 37, at 462–65.

275.   Medicare, for instance, typically does not reimburse experimental medical devices unless FDA approves the device. 42 C.F.R. § 411.15(o)(1) (2016). Since November 1995, Medicare has expanded reimbursement to include some investigational devices before they are approved, so long as they are likely to be classified as lower-risk Class I or Class II devices by FDA, *see infra* Section III.A, and so long as they do not have underlying concerns regarding their safety or effectiveness, 42 C.F.R. §§ 405.201(b), 411.15(o)(1); *see also* Ctrs. for Medicare & Medicaid Servs., Pub. No. 100-02, Medicare Benefit Policy Manual, ch. 14, § 20 (2014), https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/bp102c14.pdf [https://perma.cc/43QR-497Q] (providing the criteria for approval of Medicare coverage of Approved Investigational Device Exemption studies).

276.   *See* Joshua P. Cohen & Abigail E. Felix, *Personalized Medicine's Bottleneck: Diagnostic Test Evidence and Reimbursement*, 4 J. Personalized Med. 163, 171–72 (2014).

277.   This large issue is outside the scope of this Article, but must be addressed soon, since the adoption of algorithms depends in large part on how they are paid for and by whom. *See* Price, *supra* note 37, at 462–65. For some descriptions of the complexity in reimbursing diagnostic tests and personalized medicine, see, for example, Cohen & Felix, *supra* note 276; Eric Faulkner et al., *Challenges in the Development and Reimbursement of Personalized Medicine—Payer and Manufacturer Perspectives and Implications for Health Economics and Outcomes Research: A Report of the ISPOR Personalized Medicine Special Interest Group*, 15 Value Health 1162 (2012); Lisa M. Meckley & Peter J. Neumann, *Personalized Medicine: Factors Influencing Reimbursement*, 94 Health Pol'y 91 (2010); Note, *supra* note 134.

have an opportunity to ensure that the products available to providers and patients are high quality.

Although insurers can make independent evaluations, reimbursement decisions actually often piggyback on the decisions of regulators like FDA.[278] This may be a matter of convenience: if regulators already make decisions for many types of medical technology, insurers can rely on those decisions. But insurers could instead serve a more powerful, independent evaluative role, given their incentives to reduce costs and ensure efficacious treatment.[279] This role might be especially important for black-box medical algorithms that could reduce costs, either by themselves (by replacing more expensive diagnostic methods) or by recommending more effective treatment and thus reducing the costs of superfluous treatment.

This evaluation, however, requires information that payers don't yet have.[280] FDA could act to reduce this information gap by requiring that developers of algorithms make their information available to payers, who could then more easily evaluate those algorithms. FDA disclosure requirements could also obviate the need for the industry-standard practice of nondisclosure agreements surrounding the operation of medical algorithmic technology.[281] The relaxation of nondisclosure agreements about algorithmic operation would, in turn, create the potential for information sharing by both providers and insurers about how well algorithms actually perform in real-world health settings. Insurers could pool knowledge that certain algorithms are low quality and not worth reimbursement or use, helping control quality in an ongoing manner.[282]

---

278. See Michael D. Graf et al., *Genetic Testing Insurance Coverage Trends: A Review of Publicly Available Policies from the Largest U.S. Payers*, 10 Personalized Med. 235 (2013); Rachel Sachs, *Your Weekly Reminder that FDA Approval and Insurance Coverage Are Often Linked*, Bill Health (Nov. 30, 2016), http://blogs.harvard.edu/billofhealth/2016/11/30/your-weekly-reminder-that-fda-approval-and-insurance-coverage-are-often-linked/ [https://perma.cc/T8HJ-ZZJ7].

279. See Eisenberg & Price, *supra* note 255, at 14–15 (discussing insurers' incentives to innovate toward more efficient care).

280. See *id.* at 11–12.

281. See *supra* note 75 and accompanying text.

282. This possibility raises the question of why insurers would be interested in sharing or pooling such information, given the reality of competition. While a valid concern, the answer may lie in informal disclosures mediated by health care providers. The providers who actually observe the operation of complex medical algorithms in practice are typically reimbursed by multiple insurers, and participate in health care settings with other providers and links to other insurers. Thus, these providers create natural information-sharing nodes between those entities. Nondisclosure agreements today may prevent providers from sharing information with each other—especially across health care systems—but if those agreements are removed, informal information sharing becomes much more likely. In the presence of informal information sharing, then, between providers and insurers, participating more formally to systematize such knowledge might be more palatable, especially if the result is the elimination of payments for low-quality algorithms and an emphasis on higher-quality algorithms.

D.   *Setting the Right Balance*

The challenge across the board with information forcing by FDA as a mechanism of enabling collaborative governance is, of course, that it limits the ability of algorithm developers to protect the value of their algorithms by keeping them secret. This implicates concerns further from the new governance movement, instead touching on how challenges from such regulation may interact with incentive mechanisms or political economy concerns. In particular, as I have discussed elsewhere, secrecy is a potentially important mechanism to provide ex post appropriability (that is, allowing innovators to capture the social value of an innovation) and therefore to provide ex ante incentives for the development of new medical algorithms.[283]

But secrecy is a problem for flexible multifaceted regulation of new technology. FDA can only do so much, and its traditional tools of premarket approval may well hinder innovation more than the lack of appropriability through keeping methods secret, though this is a contestable empirical question. Silicon Valley itself demonstrates this dynamic, citing as a major holdup not the inability to keep technology proprietary for long—for software can be reverse engineered and recoded in many circumstances, and proprietary appropriation incentives can be replaced by first-mover advantage and network effects—but rather the long and slow process of slogging through regulatory hurdles to reach market in the first place.[284] A disclosure mechanism that would decrease those regulatory delays might well be regarded as a good trade by many innovators, especially if the disclosure were shaped to preserve incentives where possible.[285]

How might disclosure best be shaped to preserve incentives? The simplest feature might be a delay in disclosure, which enables fast-moving firms to obtain first-mover advantages and to build network effects before potential competitors can see the disclosed technology—though this also delays the possibility of collaborative oversight by other health market actors.[286] Other more stringent restrictions could be developed as contractual restrictions on access to data through an FDA intermediary; for instance, data could be examined, but not used as the basis for further algorithms.[287] This arrangement already exists in the context of drug clinical-trial data, where some drug research firms permit access to data from their clinical trials on

---

283.   *See* Price, *supra* note 15, at 1432–36.

284.   *See* Associated Press, *supra* note 140.

285.   *See* W. Nicholson Price II, *Regulating Secrecy*, 91 Wash. L. Rev. 1769, 1808–12 (2016) (describing this possible tradeoff).

286.   In a parallel situation, patent applications are disclosed 18 months after filing; this gives competitors access to some technical information that they could use to invent-around the patent, but preserves a substantial lead time for the first inventor. 35 U.S.C. § 122(b) (2012).

287.   This arrangement would, by design, decrease the ability of other firms to generate precisely those later algorithms, which we might otherwise find desirable. In other work, I have suggested an infrastructure model of health big data to facilitate just such cumulative innovation. Price, *supra* note 15, at 1439–44.

the condition that such data not be used as the foundation of a regulatory filing for a competing product.[288] This does not remove all potential competitive harm, of course, but it helps limit the most obvious type of imitation.[289] Another answer, put briefly, recognizes that disclosure decreases some innovation incentives, and that offsetting incentives must be put into place to counteract this decrease, whether in the form of other appropriability mechanisms (patents or regulatory exclusivity) or direct incentives (grants, prizes, or tax incentives).[290]

An additional direction to enable innovation incentives might address decreasing the costs of innovation rather than trying to increase the rewards through appropriability or direct monetary subsidies. In particular, helping reduce the costs of data assembly and quality could enable the far easier development of medical algorithms.[291] Broad databases of health data, whether developed directly by public entities or by public-private partnerships, might serve as a form of innovation infrastructure allowing for easier development of medical algorithms while reducing the need for costly initial investments.[292]

Complicating matters, since trade secrets have traditionally been protected as state property interests, retroactive disclosure of trade secrecy could arguably be a government taking requiring compensation.[293] This problem would likely be avoidable in prospective regulation, since developers would lack the "reasonable investment-backed expectation" that this type of information could be kept as a trade secret.[294] Since complex medical algorithms are still a young science, moving to require such prospective disclosure relatively early might still govern the majority of algorithms.

---

288. *See, e.g.*, Inst. of Med. of the Nat'l Acads., Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk 147–48 (2015).

289. For instance, information about what does *not* work can be tremendously valuable, and is difficult to appropriate in any fashion other than secrecy. Amy Kapczynski & Talha Syed, *The Continuum of Excludability and the Limits of Patents*, 122 Yale L.J. 1900, 1923–29 (2013). At least one counterargument to firms' desire for trade secrecy, then could be that it is a waste of social resources to replicate paths that others already know will fail, and that avoiding such waste might be more important than maintaining some types of corporate competitive advantage.

290. *See* Price, *supra* note 15, at 1418–19 (discussing incentive challenges for black-box medicine).

291. *See id.* at 1439–44.

292. *Id.*

293. *See* Ruckelshaus v. Monsanto Co., 467 U.S. 986 (1984) (holding that EPA-required disclosure of pesticide regulatory filings were a taking when the filing company had reasonable expectations based in statute that the filings would not be disclosed); *see also* Price, *supra* note 285, at 1808–09 (describing takings barriers to mandatory regulatory disclosure).

294. *Monsanto*, 467 U.S. at 1005 (quoting Webb's Fabulous Pharmacies, Inc. v. Beckwith, 449 U.S. 155, 161 (1980)); *see also id.* at 1005–07 ("[A]s long as [someone submitting data] is aware of the conditions under which the data are submitted, and the conditions are rationally related to a legitimate Government interest, a voluntary submission of data by an applicant in exchange for the economic advantages of a registration can hardly be called a taking.").

Setting the appropriate balance between regulatory oversight—particularly oversight based on disclosure to allow participation by nongovernment actors—and innovation incentives is a challenging task. Requiring disclosure to enable third-party sources of oversight reduces some incentives, to be certain, but so does the imposition of slow, costly regulatory barriers to market entry. And the absence of any regulation itself hinders innovation by removing markers of quality and fostering consumer distrust in a burgeoning industry that depends very much on trust.[295]

The tools of new governance do not solve these problems, but they suggest paths for fruitful movement forward. FDA could benefit by following a flexible, incremental model relying on lighter up-front preapproval hurdles and more substantial premarket surveillance. Such an approach should ideally involve collaborative efforts with other health-system stakeholders.

Conclusion

Algorithms are rapidly changing the way medicine is practiced today and have the potential to even more quickly change practice in the future. They promise increases in efficiency, quality of care, and the precision with which interventions are targeted to individual patients. But black-box algorithms also bring substantial risks, where low-quality algorithms may recommend dangerous interventions without the long-available checks of explicit understanding or clinical trials. As the field explodes with possibilities, regulation is needed to allow providers, patients, and insurers to choose wisely from a wealth of options of varying quality.

Figuring out how best to regulate opaque, changing algorithms will not be easy. Such regulation will need to involve flexibility, collaboration among the agency and other stakeholders, and a willingness to reevaluate regulatory strategies in the face of changing technology. FDA standing alone as regulator runs the risk of stifling innovation, which bears its own cost in missed opportunity—and might not even ensure safety for all that. Black-box algorithms create a new context for the familiar debate about speed of access against the desire to ensure safety and quality. But the technology is fundamentally different, not least because speed and flexibility of development are a large part of what black-box medicine brings to the table. FDA's role should reflect this change by acknowledging that premarket controls can only do so much to ensure quality in opaque algorithmic systems. FDA's own oversight of real-world quality markers can help regulate quality after market entry. But FDA should also welcome help by acting to remove the veil of secrecy on commercial medical algorithms and enabling collaborative regulation by providers, hospitals, and health care payers. Such a regulatory role can help induce, and can later rely on, integrated health information systems to monitor real-time performance and quality.[296]

---

295.   Cortez, *FDA Regulation of Mobile Health*, *supra* note 27, at 374.

296.   FDA's Sentinel system was built off a mandate to better surveil drug safety after approval, and represents a substantial step for integrating health data. *See* Eisenberg & Price, *supra* note 255, at 40–42.

More broadly, black-box algorithms are not limited to medicine. Machine learning and artificial intelligence are buzzwords of modern technological development, representing the next step in using big data across whatever areas are available. And while medicine may be especially salient, machine-learning algorithms create risks and benefits that will need to be addressed, measured, and regulated in many contexts.[297] How should the NHTSA regulate machine-learning algorithms involved in self-driving cars?[298] How should police use machine-learning algorithms and big data to help prevent crime, and how can we ensure that those algorithms are high-quality and free from bias?[299] How can the SEC limit the risk that machine-learning trading algorithms will crash the stock market,[300] or the FTC the risk that opaque pricing algorithms will discriminate unacceptably against different buyers?[301] Perhaps closer to home, who can regulate the quality of machine-learning systems involved in legal practice?[302]

More generally, how should agencies—or private parties—ensure the quality of algorithms that are opaque and mutable? These are not easy questions, and they do not have easy answers. I have suggested here that, at least in the context of FDA and black-box medical algorithms, an answer may be a combination of flexibility, the establishment of mechanisms for ongoing, real-world performance monitoring, and the enablement of collaborative governance mechanisms with other sophisticated stakeholders. This prescription is not a panacea. But it is a start.

---

297.   *See generally* Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 Harv. J.L. & Tech. 353, 362–72 (2016).

298.   *See, e.g.*, Harry Surden & Mary-Anne Williams, *Technological Opacity, Predictability, and Self-Driving Cars*, 38 Cardozo L. Rev. 121 (2016).

299.   *See, e.g.*, Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 Ga. L. Rev. (forthcoming 2017), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2819182 [https://perma.cc/6YSR-VEDW].

300.   *See, e.g.*, Gregory Scopino, *Preparing Financial Regulation for the Second Machine Age: The Need for Oversight of Digital Intermediaries in the Futures Markets*, 2015 Colum. Bus. L. Rev. 439.

301.   *See, e.g.*, Salil K. Mehra, *Antitrust and the Robo-Seller: Competition in the Time of Algorithms*, 100 Minn. L. Rev. 1323 (2016).

302.   *See, e.g.*, Harry Surden, *Machine Learning and Law*, 89 Wash. L. Rev. 87 (2014).