Duke Law Journal Online

VOLUME 72 NOVEMBER 2022

AI AND THE REGULATORY PARADIGM SHIFT AT THE FDA

CATHERINE M. SHARKEY†

KEVIN M.K. FODOUOP††

INTRODUCTION

Five years ago, Dr. Bakul Patel, the current Senior Director of Global Digital Health Strategy and Regulatory Affairs for Google Health, recruited "13 engineers—software developers, AI experts, cloud computing whizzes"—to prepare for "a future in which health care is increasingly mediated by machines."[1] At that time, artificial intelligence (AI) technologies were on their way to revolutionize drug development, medical diagnostics, and health care delivery—not only in the private sector,[2] but also at the federal Food and Drug

† Segal Family Professor of Regulatory Law and Policy, New York University School of Law. Professor Sharkey was one of the principal authors of *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, a 2020 Report commissioned by the Administrative Conference of the United States (ACUS), and is the author of a forthcoming 2023 ACUS Report on *Artificial Intelligence in Retrospective Review*.

†† J.D. candidate 2023, New York University School of Law. We thank the organizers and participants at *Duke Law Journal*'s 2022 Administrative Law Symposium, *Automating the Administrative State*, where Professor Sharkey commented on Professor Marks' draft article, *Automating FDA Regulation*. We also thank Nicholson Price and Rachel Sachs for further helpful comments.

1. Megan Molteni, *Medicine Is Going Digital. The FDA Is Racing to Catch Up*, WIRED (May 22, 2017), https://www.wired.com/2017/05/medicine-going-digital-fda-racing-catch [https://perma.cc/V2LE-UY4Z]. Dr. Batel joined Google in 2022, after working 13 years at the FDA. *See* Casey Ross, *Google Taps FDA's Former Digital Health Chief for Global Strategy Role*, STATNEWS (May 16, 2022), https://www.statnews.com/2022/05/16/bakul-patel-google-global-strategy-role [https://perma.cc/UR6C-VHME].

2. *See No Longer Science Fiction, AI and Robotics Are Transforming Healthcare*, PWC (Apr. 11, 2017), https://www.pwc.com/gx/en/industries/healthcare/publications/ai-robotics-new-health/transforming-healthcare.html [https://perma.cc/BFR8-YNQR] (describing advances in AI

Administration (FDA), which is in fact where Dr. Patel spearheaded this recruitment effort in 2017 as Director of the FDA's Digital Health Division.[3]

Tucked away in a chapter titled "Regulatory Analysis at the Food and Drug Administration" of a voluminous 2020 report commissioned by the Administrative Conference of the United States (ACUS), "Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies" (2020 *Government by Algorithm* Report), one of us showcased how the "FDA is in the vanguard among agencies in its experimentation with advanced AI techniques, including 'deep learning,'" and predicted that "[i]n the FDA's case, uptake of AI/ML tools may herald a broader shift away from premarket approval and toward postmarket surveillance efforts."[4]

Professor Mason Marks takes up both of these threads in "Automating FDA Regulation."[5] First, Marks describes rich case studies of FDA modeling and simulation to demonstrate that AI tools are changing the landscape of FDA regulatory decisionmaking.[6] Second, Marks insightfully probes how "the role of AI as a medical product regulator" has accelerated the shift of the FDA's focus away from premarket clearance of medical devices and drugs toward postmarket surveillance and review.[7]

We wholeheartedly agree with Marks that the FDA is at the forefront of an AI revolution in health safety. We likewise agree that the FDA is in the midst of a regulatory paradigm shift—one further propelled by the influx of AI technologies. Marks probes this shift with increasing alarm, warning of subpar safety and efficacy standards, eroding public trust in the FDA, and threats to the agency's transparency, accountability, objectivity, and legitimacy.[8] Whereas

---

for healthcare); Laura Lorenzetti, *Here's How IBM Watson Health Is Transforming the Health Care Industry*, FORTUNE (Apr. 5, 2016), https://fortune.com/longform/ibm-watson-health-business-strategy [https://perma.cc/ZWJ4-S7LF] (describing IBM Watson's attempts at applying AI to medical diagnostics).

3. Molteni, *supra* note 1; *Bakul Patel: Experience*, LINKEDIN, https://www.linkedin.com/in/bakul-patel/details/experience [https://perma.cc/FLG9-VVMM], (last visited Nov. 13, 2022).

4. Catherine M. Sharkey, *Regulatory Analysis at the Food and Drug Administration*, *in* GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 53, 54, 56 (Feb. 2020) [hereinafter Sharkey, *Regulatory Analysis at the FDA*].

5. Mason Marks, *Automating FDA Regulation*, 71 DUKE L.J. 1207 (2022).

6. *See id.* at 1219–45 (describing how the agency is "building an algorithmic FDA" through the use of "molecular models," "virtual humans and patient-specific models," and "simulated clinical trials").

7. *Id.* at 1207.

8. *Id.* at 1246–62, 1272–76.

Marks maintains a steadfast drumbeat of the perils of this shift, we draw attention to its promise. We highlight the transformative potential in terms of enhancing public health and safety. The FDA, by sustaining its investments in AI regulatory capabilities, could leverage AI to accelerate and broaden access to drugs and medical devices while preserving its "gold standard" of medical safety.

In Part I, we elaborate on the scope of the FDA's experimentation, which extends beyond the fascinating (albeit limited) case studies of FDA molecular modeling and clinical trial simulation presented by Marks. In particular, we highlight novel pilot projects in which the FDA used Natural Language Processing (NLP) to analyze data collected through its Adverse Event Reporting System for postmarket surveillance of drugs.[9]

In Part II, we describe the paradigm shift in resources and efforts at the FDA from stringent ex ante premarket approval to more dynamic and rigorous postmarket surveillance. Whereas Marks places exclusive emphasis on the potential perils from this shift, we fill out the picture by pointing to the potential promise of an AI-enabled postmarket surveillance regime.

In Part III, we explore the FDA's track record in building internal AI capacity and show how the agency's bold experimentation with the collection of structured "fit-for-purpose" data (as distinguished from unstructured text-based adverse event reports)[10] illustrates its transformation into an "information agency" of the twenty-first century.[11] Given that for most federal agencies the question is not whether the agency will eventually embrace AI technologies, but how

9. *See infra* notes 36–41 and accompanying text (surveying NLP pilots that FDA has built on top of its Adverse Reporting System); *Questions and Answers on FDA's Adverse Event Reporting System (FAERS)*, U.S. FOOD & DRUG ADMIN. [hereinafter *Q&A on FAERS*], https://www.fda.gov/drugs/surveillance/questions-and-answers-fdas-adverse-event-reporting-system-faers [https://perma.cc/EF69-9V3Q], (last updated June 4, 2018) (providing an introduction to the FDA Adverse Event Reporting System (FAERS) database, which includes adverse reports submitted for drugs and therapeutic biological products).

10. Unstructured data is data that is not organized in a pre-defined manner (often taking the form of unstructured text collected from free-form text inputs), as distinguished from structured data, which is data stored in a standardized format with a well-defined structure. *See* IBM Cloud Education, *Structured vs. Unstructured Data: What's the Difference*, IBM (June 29, 2021), https://www.ibm.com/cloud/blog/structured-vs-unstructured-data [https://perma.cc/3EHD-TRLU].

11. *See generally* Catherine M. Sharkey, *Direct-to-Consumer Genetic Testing: The FDA's Dual Role as Safety and Health Information Regulator*, 68 DEPAUL L. REV. 343 (2019) (arguing that the FDA has developed dual roles as "safety regulator" and "health information regulator").

and in which domains,[12] the FDA's experience provides a window into the future promise of AI in the administrative state.

I. THE FDA AT THE FOREFRONT OF AN AI REVOLUTION IN PUBLIC HEALTH AND SAFETY

The FDA is in the vanguard among federal agencies in its experimentation with advanced AI techniques, including "deep learning."[13] The FDA oversees products that represent over $3 trillion in annual consumption, or about 20 percent of household spending in the United States.[14] This vast regulatory scope means that even limited use of AI tools by the FDA has a substantial potential impact on human welfare. The AI revolution is affecting the FDA's regulatory operations in two ways. First, the FDA regulates medical devices that increasingly incorporate AI technologies. Second, the FDA leverages AI for its own internal uses. We explore each in turn.

A. Regulating AI Devices

AI medical devices are revolutionizing the practice of medicine. AI pattern recognition powers these devices to recognize certain types

12. *See* Ben Gansky, Michael Martin & Ganesh Sitaraman, *Artificial Intelligence Is Too Important To Leave to Google and Facebook Alone*, N.Y. TIMES (Nov. 10, 2019), https://www.nytimes.com/2019/11/10/opinion/artificial-intelligence-facebook-google.html [https://perma.cc/86AD-CHPE] (arguing that a "public option for artificial intelligence" would enable federal agencies to develop AI capacity, notably in the health care sector). The FDA is charting a path toward such a "public option for artificial intelligence." *Id.*

13. *See generally* DEP'T OF HEALTH & HUM. SERVS., ARTIFICIAL INTELLIGENCE USE CASES INVENTORY (2022), https://www.hhs.gov/sites/default/files/hhs-ai-use-cases-inventory.pdf [https://perma.cc/286K-U5HJ] (listing AI pilots at the FDA to search tobacco authorization applications, de-duplicate and extract clinical features from drug adverse event reports, and mine social media data to monitor opioid usage); Di Zhang, Jaejoon Song, Sai Dharmarajan, Tae Hyun Jung, Hana Lee, Yong Ma, Rongmei Zhang & Mark Levenson, *The Use of Machine Learning in Regulatory Drug Safety Evaluation*, STATS. BIOPHARM. RSCH. (2022) (describing the FDA's use of machine learning and "real-world data" to regulate drug safety); Zhaoyi Chen, Xiong Liu, William Hogan, Elizabeth Shenkman & Jiang Bian, *Applications of Artificial Intelligence in Drug Development Using Real-World Data*, 26 DRUG DISCOVERY TODAY 1256 (2021) (describing the FDA's efforts at promoting the use of "real-world data" and surveying resulting studies that have leveraged real-world data and AI, including deep learning); Pratik Shah, Francis Kendall, Sean Khozin, Ryan Goosen, Jianying Hu, Jason Laramie, Michael Ringel & Nicholas Schork, *Artificial Intelligence and Machine Learning in Clinical Development: A Translational Perspective*, 69 NPJ DIGIT. MED. 2, 4 (2019) (describing a partnership between the FDA and the MIT Media Lab to "engender AI and ML research for computational medicine and clinical development and [an] accompanying regulatory framework to improve health outcomes for patients").

14. *See Fiscal Year 2023: Justification of Estimates for Appropriation Committees*, U.S. FOOD & DRUG ADMIN. 2 (2022).

of injuries,[15] diagnose medical conditions,[16] increase the quality of medical imaging technology,[17] or predict future adverse medical events.[18] Such AI-powered devices increasingly outperform conventional medical devices in the hands of specialized doctors.[19]

The FDA has granted approval to a first generation of AI medical devices.[20] Utilizing the agency's de novo review process—an

15. *See, e.g.*, *FDA Permits Marketing of Clinical Decision Support Software for Alerting Providers of a Potential Stroke in Patients*, U.S. FOOD & DRUG ADMIN. (Feb. 13, 2018) [hereinafter *Viz.AI Approval*], https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm596575.htm [https://perma.cc/BN8G-GDZE] ("The Viz.AI Contact application is a computer-aided triage software that uses an artificial intelligence algorithm to analyze images for indicators associated with a stroke."); *FDA Permits Marketing of Artificial Intelligence Algorithm for Aiding Providers in Detecting Wrist Fractures*, U.S. FOOD & DRUG ADMIN. (May 24, 2018) [hereinafter *OsteoDetect Approval*], https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm608833.htm [https://perma.cc/E4N2-SG7N] ("The OsteoDetect software is a computer-aided detection and diagnostic software that uses an artificial intelligence algorithm to analyze two-dimensional X-ray images for signs of distal radius fracture, a common type of wrist fracture.").

16. *See, e.g.*, *FDA Permits Marketing of Artificial Intelligence-Based Device to Detect Certain Diabetes-Related Eye Problems*, U.S. FOOD & DRUG ADMIN. (Apr. 11, 2018) [hereinafter *IDx-DR Approval*], https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm604357.htm [https://perma.cc/8HAG-KHKC] ("IDx-DR[] is a software program that uses an artificial intelligence algorithm to analyze images of the eye taken with a retinal camera called the Topcon NW400.").

17. *See, e.g.*, *Aquilion Precision*, CANON, https://us.medical.canon/products/computed-tomography/aquilion-precision [https://perma.cc/UK97-UJ5B] (noting Canon's "Aquilion Precision" Ultra High Resolution CT Scanner provides more than twice the resolution of previous CT systems).

18. *See, e.g.*, Georgios Christopoulos, Jonathan Graff-Radford, Camden L. Lopez, Xiaoxi Yao, Zachi I. Attia, Alejandro A. Rabinstein, Ronald C. Petersen, David S. Knopman, Michelle M. Mielke, Walter Kremers, Prashanthi Vemuri, Konstantinos C. Siontis, Paul A. Friedman & Peter A. Noseworthy, *Artificial Intelligence–Electrocardiography to Predict Incident Atrial Fibrillation*, 13 CIRCULATION: ARRHYTHMIA AND ELECTROPHYSIOLOGY 1420 (2020); Nozomi Niimi, Yasuyuki Shiraishi, Mitsuaki Sawano, Nobuhiro Ikemura, Taku Inohara, Ikuko Ueda, Keiichi Fukuda & Shun Kohsaka, *Machine Learning Models for Prediction of Adverse Events After Percutaneous Coronary Intervention*, 12 NATURE: SCI. REPS. 1 (2022).

19. *See, e.g.*, Xiaoxuan Liu, Livia Faes, Aditya U. Kale, Siegfried K. Wagner, Dun Jack Fu, Alice Bruynseels, Thushika Mahendiran, Gabriella Moraes, Mohith Shamdas, Christoph Kern, Joseph R. Ledsam, Martin K. Schmid, Konstantinos Balaskas, Eric J. Topol, Lucas M. Bachmann, Pearse A. Keane & Alastair K. Denniston, *A Comparison of Deep Learning Performance Against Health-Care Professionals in Detecting Diseases from Medical Imaging: A Systematic Review and Meta-Analysis*, 1 LANCET DIGIT. HEALTH 271, 272, 291–93 (2019) (finding that deep learning algorithms using medical imaging provide equivalent diagnostic accuracy at increased diagnostic speed); Ravi Aggarwal, Viknesh Sounderajah, Guy Martin, Daniel S. W. Ting, Alan Karthikesalingam, Dominic King, Hutan Ashrafian & Ara Darzi, *Diagnostic Accuracy of Deep Learning in Medical Imaging: A Systematic Review and Meta-Analysis*, 65 NPJ DIGIT. MED. 1, 19–20 (2022) (finding that "[deep learning] currently has a high diagnostic accuracy"); *see also* Nan Wu et al., *Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening*, 39 IEEE TRANSACTIONS ON MED. IMAGING 1184, 1184–94 (2020).

20. As of September 2020, the FDA had approved sixty-four AI-based medical products. *See* Stan Benjamens, Pranavsingh Dhunnoo & Bertalan Meskó, *The State of Artificial Intelligence-*

alternative approval pathway for "novel devices of low to moderate risk"[21]—the FDA gave marketing clearance to Viz.AI (detects strokes), OsteoDetect (recognizes bone fractures), and IDx-DR (identifies diabetic retinopathy) after their respective manufacturers demonstrated certain performance criteria.[22] With regard to these first AI medical device approvals, the FDA sought to "creat[e] a regulatory framework for [clinical decision support] products that encourages developers to create, adapt, and expand the functionalities of their software to aid providers in diagnosing and treating diseases and conditions."[23]

The next generation of AI medical devices will present additional regulatory challenges. Thus far, the FDA has only approved "locked" devices, *i.e.*, devices that do not independently adapt to new data they observe, but rely, instead, on manufacturer updates. But, the FDA has recognized that "there's a great deal of promise beyond locked algorithms that's ripe for application in the health care space."[24] Enabling AI models to dynamically update through time promises to unlock better performance and more personalized health care

---

Based FDA-Approved Medical Devices and Algorithms: An Online Database, 3 NPJ DIGIT. MED. 1, 2 (2020).

21. This approval pathway provides authorization to "be marketed and used as predicates for future 510(k) submissions." *Evaluation of Automatic Class III Designation (De Novo) Summaries*, U.S. FOOD & DRUG ADMIN., http://www.fda.gov/about-fda/cdrh-transparency/evaluation-automatic-class-iii-designation-de-novo-summaries [https://perma.cc/PW3T-ZJ8X], (last updated Oct. 17, 2022).

The vast majority of medical devices enter the market via the FDA's 501(k) process, a streamlined "premarket notification" approval (PMN) pathway that imposes less stringent requirements than the FDA's Premarket Approval (PMA) process. It mirrors the streamlined Abbreviated New Drug Application (ANDA) regulatory approval pathway for generic drugs. *See* Sharkey, *Regulatory Analysis at the FDA*, *supra* note 4, at 54.

22. *See, e.g.*, *Viz.AI Approval*, *supra* note 15 ("The company submitted a retrospective study of 300 CT images that assessed the independent performance of the image analysis algorithm . . . against the performance of two trained neuro-radiologists for the detection of large vessel blockages in the brain."); *OsteoDetect Approval*, *supra* note 15 (explaining that OsteoDetect was approved based on "a retrospective study of 1,000 radiograph images that assessed the independent performance of the image analysis algorithm for detecting wrist fractures and the accuracy of the fracture localization of OsteoDetect against the performance of three board certified orthopedic hand surgeons"); *IDx-DR Approval*, *supra* note 16 (reporting that the FDA found that "IDx-DR was able to correctly identify the presence of more than mild diabetic retinopathy 87.4 percent of the time and . . . identify . . . patients who did not have more than mild diabetic retinopathy 89.5 percent of the time").

23. *Viz.AI Approval*, *supra* note 15.

24. Conor Hale, *FDA Lays Out Plans for a New Review Framework for AI and Machine Learning-based Devices*, FIERCE BIOTECH (Apr. 3, 2019), https://www.fiercebiotech.com/medtech/fda-lays-out-plans-for-a-new-review-framework-for-ai-and-machine-learning-based-devices [https://perma.cc/7XL9-869M].

delivery.[25] Anticipating approval of "dynamic AI" devices, the FDA, in January 2021, issued its "Software as a Medical Device Action Plan" (SaMD Plan) guidance, which outlines a "Predetermined Change Control Plan" process through which manufactures can get pre-approval for certain types of dynamic updates.[26] Critically, the FDA's SaMD Plan emphasizes the need for any dynamic AI medical device to be monitored "from its premarket development through postmarket performance."[27]

B. Internal AI Use

We turn now from the FDA's regulation of AI-powered medical devices for use in society to our main focus (as well as Marks'): the FDA's internal use of AI in its regulatory drug and medical device approval processes. Marks draws attention to the FDA's experimentation with uses of AI in molecular modeling (to preemptively identify potentially harmful drug substances), virtual humans and patient-specific models, and simulated clinical trials (to reduce the cost of *in vivo* clinical trials and accelerate drug

25. *See, e.g.*, Mihaela van der Schaar, Ahmed M. Alaa, Andres Floto, Alexander Gimson, Stefan Scholtes, Angela Wood, Eoin McKinney, Daniel Jarrett, Pietro Lio & Ari Ercole, *How Artificial Intelligence and Machine Learning Can Help Healthcare Systems Respond to COVID-19*, 110 MACH. LEARNING 1, 4 (2020) (explaining how machine learning models can "learn" through time how an individual's features can be "mapped into *personalized predictions of risk*"); Chris Giordano, Meghan Brennan, Basma Mohamed, Parisa Rashidi, François Modave & Patrick Tighe, *Accessing Artificial Intelligence for Clinical Decision-Making*, 4 FRONTIERS DIGIT. HEALTH 1, 4–5 (2021) (stating that clinical decisionmaking tools using "dynamic and personalized" AI models can improve patient outcomes); Fei Wang & Anita Preininger, *AI in Health: State of the Art, Challenges, and Future Directions*, 28 YEARBOOK MEDICAL INFORMATICS 16, 23 (2019) (explaining that combining state-of-the-art "dynamic AI" and "federated learning" techniques—which enables the training of personalized AI models without accessing sensitive patient data—can allow AI medical devices to continuously improve their performance based on troves of data collected by wearable or mobile devices).

Moreover, "locked" algorithms "can lead the AI/ML system to use a poor estimate of the true relationship between the inputs and outputs and thereby possibly cause harm to patients (for example, through misdiagnosis)." Boris Babic, Sara Gerke, Theodoros Evgeniou & I. Glenn Cohen, *Algorithms on Regulatory Lockdown in Medicine*, 366 SCI. 1202, 1202–03 (2019).

26. ARTIFICIAL INTELLIGENCE/MACHINE LEARNING (AI/ML)-BASED SOFTWARE AS A MEDICAL DEVICE (SAMD) ACTION PLAN, U.S. FOOD & DRUG ADMIN. 3 (Jan. 2021) [hereinafter SAMD PLAN], https://www.fda.gov/media/145022/download [https://perma.cc/4ZSY-AB2S] (describing the "Predetermined Change Control Plan" as including an SaMD Pre-Specifications (SPS) component describing "what" model aspects the manufacturer intends to change through learning and an Algorithm Change Protocol (ACP) component explaining "how" the algorithm will learn and change while remaining safe and effective).

27. *Id.* at 1.

development).[28] We have our quibbles with the conclusions Marks draws from his rich case studies. But, our greater concern is that by focusing solely on these specific use cases, Marks omits key ways in which the FDA leverages AI to more effectively operate in an increasingly complex and high-stakes regulatory environment, thus failing to acknowledge benefits the FDA is receiving through the use of AI in its regular activities.

To begin, we are skeptical of the main conclusion Marks draws from his case studies. Take his primary motivating example: the FDA's use of the Public Health Assessment via Structural Evaluation (PHASE) computational methodology in a potential regulatory decision to schedule and ban kratom.[29] Marks puts forth the PHASE/kratom saga to support his overarching thesis that the FDA is recklessly substituting algorithms for human judgment based on reliable evidence. Marks argues that PHASE is a "poor substitute for the eight-factor analysis"—the traditional means by which the FDA evaluates unknown substance risk, taking into account "complex historical, epidemiological, and psychological factors" in addition to purely chemical and physical properties.[30] While this is undoubtedly correct (especially given that PHASE is not designed for processing sociological and demographic data), PHASE may nonetheless serve as a "signaling" tool, incorporated into a hybrid human-machine review process that would likely be more accurate and reliable than the traditional eight-factor analysis.[31]

Moreover, as Marks details, in this specific case, the Department of Health and Human Services (HHS) overrode the FDA and asked

28. *See* Marks, *supra* note 5, at 1227–36, 1238–45 (outlining the current and potential roles of molecular models, virtual humans and patient-specific models, and simulated clinical trials in FDA regulation and policymaking).

29. *See id.* at 1227–36 (discussing how PHASE, a computational methodology adopted by the FDA, was used to assess risk to public health and its shortcomings). PHASE was an internally built model, later supplemented with Clarity, a third-party software, *i.e.*, "proprietary technology developed by a private drug company"—*i.e.*, one of the companies the FDA regulates. *Id.* at 1275. But, this type of conflict of interest is not endemic to all (or even most) forms of FDA experimentation with AI. *Id.*

30. *Id.* at 1231.

31. For example, PHASE can be useful in quickly identifying those substances that require further examination by the FDA because of their similarities to other harmful substances. In other words, the FDA could employ molecular modeling at the outset of unknown or understudied substance analysis to generate a framework through which to complete a more comprehensive, albeit streamlined analysis. The AI system would thereby guide the FDA in best utilizing its expertise and not to make a broader policy decision.

the Drug Enforcement Administration *not* to schedule the drug.[32] Could this be evidence that an existing internal oversight mechanism was deployed to recognize and correct for model deficiencies and lack of evidentiary support? Marks does not pause to consider this (and we thus do not have enough information to evaluate), nor does Marks appreciate the extent to which the FDA is subject to "high potential for judicial review and public scrutiny"[33] and thus subject to external constraints as well.[34]

And, even if one (like Marks) tends toward inherent skepticism regarding the FDA's AI use and administrative law constraints, an equally plausible interpretation of the PHASE/kratom incident is that then-Commissioner Gottlieb's preexisting hostility toward kratom led to a seemingly inevitable political outcome.[35] In other words, the PHASE/kratom saga fits an overarching story about the role of politics in regulatory policy and the influence of industry in agency decisionmaking and accountability—not having anything in particular to do with the FDA's use of AI.

---

32. *See* Marks, *supra* note 5, at 1233–36 (explaining that HHS instructed DEA not to schedule kratom because it did not find FDA's PHASE predictions conclusive and assessed that prohibiting the drug could prompt detrimental public health consequences from millions of users potentially switching to lethal opioids as alternative painkillers).

33. Nitisha Baronia, David Freeman Engstrom, Daniel E. Ho, Shawn Musgrave & Catherine M. Sharkey, *Building Internal Capacity*, *in* GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 74 (Feb. 2020) [hereinafter *Building Internal AI Capacity*].

34. Administrative law offers more agency-constraining tools than Marks lets on. Should the FDA actually use PHASE to decide not to approve a drug, a manufacturer could mount a challenge under the Administrative Procedure Act, arguing that FDA failed to consider methodological flaws of PHASE or that the FDA's conclusion is implausible given that it used PHASE to reach it. *See* 5 U.S.C. § 706(2)(A) (2018). Agency rules are arbitrary, capricious, and in violation of 5 U.S.C. § 706(2)(A) "if the agency failed to consider an important aspect of the problem, offered an implausible explanation that runs counter to the evidence before it, or relied on factors that Congress did not intend." Motor Vehicle Mfrs. Ass'n of U.S., Inc. v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 43 (1983). Marks does not consider the extent to which external constraints from administrative law doctrine can require the FDA to use reasoned decisionmaking, thus enhancing credibility and public trust in its decisions.

35. To take perhaps the most obvious example (not mentioned by Marks), consider the DEA's scheduling of marijuana as a Schedule I drug (defined by the DEA as those with "no currently accepted medical use and a high potential for abuse"). *See Drug Scheduling*, U.S. DRUG ENF'T ADMIN., https://www.dea.gov/drug-information/drug-scheduling [https://perma.cc/V9ZF-MEMN], (last visited Nov. 4, 2022). President Biden recently released a statement calling on the Secretary of HHS to review marijuana's scheduling classification. *See* President Joe Biden, Statement from President Biden on Marijuana Reform (Oct. 6, 2022), https:/www.whitehouse.gov/briefing-room/statements-releases/2022/10/06/statement-from-president-biden-on-marijuana-reform [https://perma.cc/NHT8-SU5K]. Marks' silence on the racial and societal implications of drug scheduling is surprising.

But, given that readers can evaluate for themselves the persuasiveness of Marks' case studies, more worrisome to us are Marks' sins of omission. We fill out the picture here by drawing attention to additional internal AI-use pilots conducted by the FDA that show great promise. The FDA has used AI-enabled models to mine data in order to assist with its postmarket surveillance targeting of drug safety issues and to uncover new relationships between drugs and medical conditions. The FDA's Adverse Event Reporting System (FAERS) database contains "adverse event reports, medication error reports and product quality complaints resulting in adverse events that were submitted to FDA."[36] These millions of reports and complaints are in freeform text-based format.

In one pilot, with the goal of using AI technology to identify postmarket safety concerns so as to prioritize safety review by FDA subject matter experts, the FDA, in collaboration with Stanford University data scientists, tested different NLP models to predict the probability that FAERS reports contained policy-relevant information about drug safety concerns.[37] The FDA considered the pilot, which identified six key features for priority review,[38] "the foundation" of an improved system that better allocates scarce agency resources in identifying postmarket safety concerns.[39] In another pilot, the FDA used similar NLP techniques to translate FAERS' unstructured data into structured data before attempting to model relationships between different drugs and hepatic failure, a medical condition affecting the liver.[40] For both pilots, the agency compared the performance of multiple NLP models (such as neural network, logistic regression, random forest, and support vector machine models) and optimized

36. *Q&A on FAERS*, *supra* note 9. The FDA received nearly 2.34 million FAERS reports in 2021. *See* U.S. FOOD & DRUG ADMIN., FDA ADVERSE EVENTS REPORTING SYSTEM (FAERS) PUBLIC DASHBOARD [hereinafter *FAERS Dashboard*], https://fis.fda.gov/sense/app/95239e26-e0be-42d9-a960-9a5f7f1c25ee/sheet/7a47a261-d58b-4203-a8aa-6d3021737452/state/analysis [https://perma.cc/2MJJ-9RNN].

37. The FDA-Stanford team trained their models on drug reports from the World Health Organization-Uppsala Monitoring Centre (WHO-UMC) manually labeled by FDA safety evaluators for drug causality assessment. *See* Sharkey, *Regulatory Analysis at the FDA*, *supra* note 4, at 55.

38. *See id.* at 56.

39. *Id.*

40. This second pilot proved less successful than the FDA-Stanford pilot to prioritize FAERS review, as it failed to confirm or disprove that the drugs involved in the FAERS reports had a causal relationship to hepatic failure and did not generate outputs accurate enough for deployment. *Id.*

model parameters until it reached satisfying predictive power, in some cases above 90 percent.[41]

Marks' dismissive characterization (in a footnote reference) that "the FDA concluded [the FAERS pilots] were unsuccessful" is thus misleading (at best).[42] Marks' sole reference to FAERS is in his scholarship review, in which he cites the 2020 *Government by Algorithm* Report.[43] But, the report uses far more positive qualified language—namely that "the FDA's FAERS efforts have been successful, to an extent."[44]

Moreover, the FDA fully recognizes that its experimentation with AI in this domain is an iterative, trial-by-error long-term approach. In 2020, the FDA engaged the public in a "precisionFDA challenge" to develop improved models for analyzing unstructured data from FAERS reports. precisionFDA challenges involve the public in solving technical challenges faced by the FDA in the domains of AI and bioinformatics. Public participation in such "community-sourced science" challenges helps bolster the transparency, accountability, objectivity, and legitimacy of the new AI methods used by FDA.[45] AI innovation is a particularly apt area for involving the public given the low startup costs of setting up computing-based challenges (as opposed to wet-lab scientific challenges). The "Gaining New Insights by Detecting Adverse Event Anomalies Using FDA Open Data" challenge, which started in early 2020, asked participants to develop AI/ML models to better detect "possible safety issues" from FDA's FAERS records.[46] Since then, the agency has expanded its use of NLP

41. The hepatic failure pilot achieved a true positive rate of 91 percent and false positive rate of 4.9 percent. *Id.* at 55. The review prioritization pilot did not achieve such high predictive performance (area under the curve of 0.66) but still provided actionable results to accurately prioritize human review. *Id.* at 55 & n.47.

42. *See* Marks, *supra* note 5, at 1213 n.30.

43. *Id.*

44. Sharkey, *Regulatory Analysis at the FDA*, *supra* note 4, at 56.

45. *Challenges*, PRECISIONFDA [hereinafter *precisionFDA Challenges*], https://precision.fda.gov/challenges [https://perma.cc/83E3-BV3U].

46. *See Gaining New Insights by Detecting Adverse Event Anomalies Using FDA Open Data*, precisionFDA [hereinafter *FAERS precisionFDA Challenge*], https://precision.fda.gov/challenges/9 [https://perma.cc/ZC8A-QKRL]. The challenge suggested that participants rely on NLP to extract relevant data features from the "text narrative portion of the adverse event reports," including basic information such as "symptoms, diagnosis, treatments, and dates." *Id.*

on unstructured text data to more data sources[47] and to cutting-edge "deep learning" language models.[48]

In addition to its investment in NLP technologies, the FDA has launched initiatives to restructure its data operations and reshape its technical data infrastructure, such as INFORMED (Information Exchanged and Data Transformation), which tasked entrepreneurs-in-residence, engineers, and data scientists with medical subject-matter expertise to strategize how the FDA should invest in big data analytics capabilities.[49]

II. A REGULATORY PARADIGM SHIFT AT THE FDA: FROM PREMARKET APPROVAL TO POSTMARKET SURVEILLANCE

The AI technological innovations for mining big data at the FDA will likely drive a paradigm shift at the agency—from heavy investment of resources and efforts focused on stringent ex ante premarket approval to more dynamic and rigorous postmarket surveillance. Marks, too, has recognized the significance of this paradigm shift. But, whereas Marks emphasizes the potential perils from this shift, we point to the promise not only in conventional terms of fostering innovation but also in more novel terms of enabling more effective institutional collaboration and regulatory enforcement in an AI-enabled

47. *See* A. Sorbello, R. Hasan, H. Francis, I. Chang, M. Ahadpour, M. Laponsky, J. Walsh & C. Trier, *A Novel Natural Language Processing and Machine Learning Classifier That Streamlines Extracting Drug-Adverse Event Data from Literature Reports*, U.S. FOOD & DRUG ADMIN., https://www.fda.gov/media/142029/download [https://perma.cc/83JQ-4KB9] (summarizing the FDA's analysis of the PubMed/MEDLINE database, a leading scientific literature database produced by the National Library of Medicine, to facilitate the identification of adverse drug events).

48. *See* Yiwen Shi, Ping Ren, Yi Zhang, Xiajing Gong, Meng Hu & Hualou Liang, *Information Extraction from FDA Drug Labeling to Enhance Product-Specific Guidance Assessment Using Natural Language Processing*, 6 FRONTIERS RSCH. METRICS & ANALYTICS 1, 1 (2021), https://www.frontiersin.org/articles/10.3389/frma.2021.670006/full [https://perma.cc/EXE9-2FSG] (using large language "transformer" model called BERT (Bidirectional Encoder Representations from Transformers) to extract information from regulatory text and help the development of product-specific guidances); Yue Wu, Zhichao Liu, Leihong Wu, Minjun Chen & Weida Tong, *BERT-Based Natural Language Processing of Drug Labeling Documents: A Case Study for Classifying Drug-Induced Liver Injury Risk*, 4 FRONTIERS A.I. 1, 10 (2021), https://www.frontiersin.org/articles/10.3389/frai.2021.729834/full [https://perma.cc/X92W-ZZBF] (using similar BERT models to classify risks of hepatic failures created by different drugs).

49. *See* Sean Khozin, Richard Padzur & Anand Shah, *INFORMED: An Incubator at the US FDA for Driving Innovation in Data Science and Agile Technology*, 17 NATURE REVS. DRUG DISCOVERY 529, 530 (2018) ("Our current objectives are twofold: first, to continue to expand and maintain organizational and technical infrastructure for data science and big data analytics; and second, to support systems thinking in oncology regulatory science research, [and develop] novel solutions for improving efficiency, reliability and productivity in related workflows.").

postmarket surveillance regime. Chief among these benefits are (1) AI increasing the speed at which FDA investigatory and corrective action can take place and (2) the public benefit of heightened product and manufacturing quality arising from firms being induced to have sound algorithmic monitoring systems in place.

A. Describing the Shift

1. *The FDA's traditional paradigm.* The FDA historically has operated as a strict ex ante regulator. The FDA's traditional regulatory framework imposes heightened ex ante premarket approval for both drugs and medical devices with relatively limited postmarket surveillance. The core of the premarket approval process for brand-name prescription drugs is the New Drug Application (NDA) process, which requires manufacturers to conduct three phases of premarket clinical trials to demonstrate the safety and efficacy of their drugs to the FDA's satisfaction.[50] A similarly stringent Premarket Approval (PMA) process applies to Class III (high risk) medical devices.[51] The FDA emerges as the most stringent ex ante safety regulator of any U.S. federal agency; moreover, its "gold standard" is higher than that of foreign medical product regulatory agencies.[52]

Under this traditional model, the FDA acts as a centralized federal safety gatekeeper for prescription medical devices and drugs.[53]

---

50. *See, e.g.*, *The FDA's Drug Review Process: Ensuring Drugs Are Safe and Effective*, U.S. FOOD & DRUG ADMIN (Nov. 24, 2017), https://www.fda.gov/Drugs/ResourcesForYou/Consumers/ucm143534.htm [https://perma.cc/NXL8-G3LN].

51. *See Premarket Approval (PMA)*, U.S. FOOD & DRUG ADMIN., https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/premarket-approval-pma [https://perma.cc/9M5U-NCJU] (stating that "PMA is the most stringent type of device marketing application required by FDA" because the "FDA has determined that general and special controls alone are insufficient to assure the safety and effectiveness of Class III devices"); Riegel v. Medtronic, Inc., 552 U.S. 312, 317–18 (2008) (noting that PMA is a "rigorous" process requiring a "multivolume application" and that "[t]he FDA spends an average of 1,200 hours reviewing each application").

52. *Merck and Vioxx: Putting Patient Safety First?: Hearings Before the Senate Committee On Finance*, 108th Cong., at 1 (2004) (statement of Sandra L. Kweder, Deputy Dir., FDA Off. of New Drugs) ("It is well recognized that FDA's drug review is a gold standard. Indeed, we believe that FDA maintains the highest worldwide standards for drug approval.").

53. This regulatory framework is largely federalized, with important swaths of federal preemption limiting the role of state law in promoting medical device and drug safety. *See generally* Catherine M. Sharkey, *Products Liability Preemption: An Institutional Approach*, 76 GEO. WASH. L. REV. 449, 464–66, 474 (2008) (describing the "pro-preemption" trend of U.S. Supreme Court cases); *Riegel*, 552 U.S. at 330 (holding that state tort laws seeking to add design or labeling requirements on top of the FDA's rigorous premarket approval scheme for medical devices were preempted); PLIVA, Inc. v. Mensing, 564 U.S. 604, 626 (2011) (holding that state law failure-to-warn claims against generic drug manufacturers were preempted).

Historically, the FDA has focused great attention on minimizing "Type I" errors—"false positives" or approval of drugs that turn out to have safety issues—at the expense of increasing corresponding "Type II" errors—"false negatives" or delaying or withholding drugs from the market that, in fact, would provide net safety benefits to patients.[54] Indeed, some have argued that, by doing so, the FDA has built up tremendous reputational capital.[55] It is precisely this institutional reputational capital that Marks fears the FDA has been eroding in recent times.[56]

Still, clinical trials—despite the FDA requiring three phases—nonetheless only provide information from a relatively limited population (even Phase III trials can be conducted with 1,000 participants) over a relatively brief period.[57] We should, therefore, be cognizant of their potential to mask serious safety risks and at least remain open to the possibility that AI-enabled simulated clinical trials might *improve* safety outcomes.

The stakes are particularly high for minority and marginalized communities due not only to their relatively small sample sizes but also selection bias in recruiting study participants for clinical trials.[58] Marks

54. *See, e.g.*, Andrew Flowers, *How the FDA Could Change The Way It Approves Drugs*, FIVETHIRTYEIGHT (Sept. 3, 2015, 12:34 PM), https://fivethirtyeight.com/features/how-the-fda-could-change-the-way-it-approves-drugs [https://perma.cc/4K3D-CZ57] (presenting research criticizing the FDA's focus on minimizing Type I error and arguing for higher false positive tolerance for severe diseases "to allow more drugs to hit the market even though some of them would be ineffective or harmful"); Jack Botting, *The History of Thalidomide*, 15 DRUG NEWS & PERSPS. 604, 604 (2002) (explaining that the FDA strengthened its drug premarket approval process following the thalidomide "disaster," where a sedative drug initially thought to be nontoxic caused an epidemic of deformities in children whose mothers had taken the drug while pregnant). For a corresponding analysis of the Type I versus Type II trade-off in the device approval realm, see Thomas J. Hwang, Elisaveta Sokolov, Jessica M. Franklin, & Aaron S. Kesselheim, *Comparison of Rates of Safety Issues and Reporting of Trial Outcomes for Medical Devices Approved in the European Union and United States: Cohort Study*, 353 BMJ 1, 1–7 (2016).

55. *See generally* DANIEL CARPENTER, REPUTATION AND POWER: ORGANIZATIONAL IMAGE AND PHARMACEUTICAL REGULATION AT THE FDA (2010) (summarizing a large-scale theoretical, historical, and statistical analysis of FDA pharmaceutical regulation (the FDA Project) and concluding that the FDA's organizational reputation has been the primary source of its power).

56. *See* Marks, *supra* note 5, at 1260 (stating that the FDA's reputation has been "under fire" and warning that substituting algorithmic models for clinical trials may further erode its reputation).

57. *See* Aaron S. Kesselheim, Michael D. Greene & Jerry Avorn, *Who is Now Responsible for Discovering and Warning About Adverse Effects of Generic Drugs?*, 310 JAMA 1023, 1023 (2013) (recognizing that "premarket testing does not reveal the full range of a drug's adverse effects").

58. *See* Theodore Eisenberg & Martin T. Wells, *Statins and Adverse Cardiovascular Events in Moderate-Risk Females: A Statistical and Legal Analysis with Implications for FDA Preemption*

is right to raise a cautionary flag—if not attended to, the use of AI could exacerbate biases against marginalized communities.[59] But, Marks understates the benefits of simulated trials, which can act as a complement (as opposed to replacement) to randomized controlled trials and allow trials to be run repeatedly or to reduce racial equity risks inherent in unrepresentative clinical trials.[60] Scientists conducting simulated trials can intentionally and proactively build representativeness into the virtual cohorts interacting with their models. Moreover, one might take a more sanguine view after examining how the FDA has been particularly attuned, as part of its Model-Informed Drug Development (MIDD) initiative, to leverage AI-based models "to bridge efficacy and safety for certain unstudied patient subpopulations or use scenarios"[61] with the goal of making clinical trial data not only more efficient to collect but also "more representative of diverse patient populations."[62] In one promising study in May 2022, a team of pharmaceutical researchers, taking up the MIDD challenge, confirmed the efficacy of an influenza drug on ethnic groups that were underrepresented in the drug's Phase III clinical studies.[63]

In addition to the high costs and potential unrepresentativeness of the FDA's existing pre-market approval regime, such a stringent ex

Claims, 5 J. EMPIRICAL LEGAL STUD. 507, 509–10 (2008) (explaining that the top-selling drug Lipitor was approved by the FDA although its clinical trials included few women and generated "inconclusive" drug efficacy results for women); Bassel Nazha, Manoj Mishra, Rebecca Pentz & Taofeek K. Owonikoko, *Enrollment of Racial Minorities in Clinical Trials: Old Problem Assumes New Urgency in the Age of Immunotherapy*, 39 AM. SOC'Y CLINICAL ONCOLOGY EDUC. BOOK 3, 4 (2019) (reporting that non-white participants represent only 20 percent of cancer clinical trial participants).

59. *See* Marks, *supra* note 5, at 1273–76.

60. As Marks himself seems to recognize—by referring to models predicting rituximab as a superior treatment to rheumatoid arthritis—simulated trials can also help drug manufacturers decide whether running a clinical trial is worth the risks and costs. *See id.* at 1243.

61. *See CDER Conversation: Model Informed Drug Development*, U.S. FOOD & DRUG ADMIN., https://www.fda.gov/drugs/news-events-human-drugs/cder-conversation-model-informed-drug-development https://perma.cc/KD44-VGCG], (last updated June 12, 2018).

62. *FDA's Technology Modernization Action Plan (TMAP)*, U.S. FOOD & DRUG ADMIN. 2 (Sept. 18, 2019) [hereinafter *TMAP*], https://www.fda.gov/media/130883/download [https://perma.cc/JFQ6-UC6W].

63. Sylvie Retout, Stefan De Buck, Sébastien Jolivet, Vincent Duval & Valérie Cosson, *A Pharmacokinetics–Time to Alleviation of Symptoms Model to Support Extrapolation of Baloxavir Marboxil Clinical Efficacy in Different Ethnic Groups with Influenza A or B*, 112 CLINICAL PHARMACOLOGY & THERAPEUTICS 372, 373, 380 (2022); Piet H. van der Graaf, *Diversity in Clinical Pharmacology Coming of Age*, 112 CLINICAL PHARMACOLOGY & THERAPEUTICS 191, 192 (2022) (stating that Retout and co-workers' success in using MIDD to support new drug applications in different ethnic patient groups demonstrates the potential of MIDD for drug development).

ante regulatory approach undeniably delays time to market for innovative drug and medical device products.[64] To date, the FDA has faced perhaps a Faustian dilemma with regard to tradeoffs between Type I and Type II errors, namely the impossibility of reducing one without increasing the other. Even its most adamant critics have acknowledged that, should the FDA lower its stringent premarket criteria, increased resources should be dedicated to postmarket surveillance.[65] While the FDA has been gradually expanding its postmarket surveillance regime over the last fifteen years, the AI revolution promises a more transformative shift.

2. *A shift, accelerated by AI, to postmarket surveillance.* The development of AI technologies has propelled the FDA's transformative shift toward increased postmarket surveillance. For the past several years, the FDA has relaxed the stringency of its premarket testing regime, in effect shifting resources from premarket to postmarket scrutiny.[66] In 2021, the FDA issued 74 percent of its new drug approvals under an "expedited program" loosening premarket requirements to speed drug commercialization, and 28 percent of approved drugs were approved through the Accelerated Approval program, one type of "expedited program" that enables earlier drug approval by relying on postmarket trials to confirm clinical efficacy.[67]

Back in 2007, responding to pleas by the FDA and legal scholars urging the need for postmarket surveillance of drugs, Congress enacted the Food and Drug Administration Amendments Act (FDAAA).[68]

---

64. Note the difficulty in measuring these Type II errors, which are not as salient as Type I errors. For some attempts to measure Type II errors in drug approvals, see generally Thomas J. Philipson & Eric Sun, *Cost of Caution: The Impact on Patients of Delayed Drug Approvals*, PROJECT FDA REPORT (June 2010).

65. *See* Richard A. Epstein, *Regulatory Paternalism in the Market for Drugs: Lessons from Vioxx and Celebrex*, 5 YALE J. HEALTH POL'Y, L. & ETHICS 741, 747–48 (2005) (arguing for a shift from premarket drug scrutiny by FDA to postmarket surveillance).

66. *See, e.g.*, Nathan Cortez, *Digital Health & Regulatory Experimentation at the FDA*, 18 YALE J. HEALTH POL'Y, LAW & ETHICS 4, 14 (2019) (describing the FDA's "shifting its focus from pre-market to post-market evidence gathering" as a significant experiment in medical product regulation); *see also* W. Nicholson Price III, *Regulating Black-Box Medicine*, 116 MICH. L. REV. 421, 458 (2017) (advocating an approach for "black-box medicine" that 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").

67. CTR. FOR DRUG EVALUATION AND RSCH., U.S. FOOD & DRUG ADMIN., *Advancing Health Through Innovation: New Drug Therapy Approvals 2021*, at 18 (Jan. 2022) [hereinafter *Advancing Health Through Innovation*].

68. *See generally* U.S. GOV'T ACCOUNTABILITY OFF., REPORT TO REQUESTERS: DRUG SAFETY, IMPROVEMENTS NEEDED IN FDA'S POSTMARKET DECISION-MAKING AND

The FDAAA granted the FDA authority to monitor safety risks from already approved drugs and require drug manufacturers to perform postmarket safety studies.[69] Concerns have been raised since then (including by Marks[70]) about the FDA's track record of postmarket surveillance.[71] We recognize that it will likely take increased financial resources along with institutional commitment for the FDA to up its postmarket surveillance game.

But, notwithstanding the FDA's constrained resources,[72] we see promise in the FDA's ability to harness AI-enabled tools (with the FAERS pilots being an early example) to improve the agency's ability to evaluate postmarket data at scale.[73] Since the late 2000s, the FDA has amassed large amounts of data, almost exclusively from manufacturers, into adverse event reports databases (including FAERS) to inform its postmarket surveillance efforts.[74] The FDA has

---

OVERSIGHT PROCESS (2006); *see also* Rebecca S. Eisenberg & W. Nicholson Price II, *Promoting Healthcare Innovation on the Demand Side*, 4 J.L. & BIOSCIENCES 3, 41–44 (2017) (describing the development of the FDA's postmarket surveillance authority).

69. *See* Catherine M. Sharkey, *The Fraud Caveat to Agency Preemption*, 102 NW. U. L. REV. 841, 863–64 (2008) ("Congress, after conducting numerous hearings . . . , has taken recent action to buttress the FDA's drug approval and oversight functions. The FDA Amendments Act . . . empowers the FDA with additional authority during the postapproval period to monitor drug side effects and to impose larger fines on companies that do not conduct postmarketing studies.").

70. *See* Marks, *supra* note 5, at 1216–17.

71. *See, e.g.*, Daniel Carpenter, *Reputation, Gatekeeping and the Politics of Post-Marketing Drug Regulation*, 8 ETHICS J. AM. MED. ASS'N 403, 404 (2006) (arguing that the FDA has struggled to ensure compliance postmarket because the agency's power over manufacturers decreases once a drug or device is approved); Sheila Kaplan, *FDA Faulted for Failure to Track Safety Issues with Drugs Already on Market*, STATNEWS (Jan. 14, 2016), https://www.statnews.com/2016/01/14/fda-postmarket-study-report [https://perma.cc/6XR3-FNGS] (noting that safety concerns arise after a drug goes on the market); Kesselheim et al., *supra* note 57, at 1023 (asserting that "post-market surveillance by the [FDA] is insufficient"); *see also* Prashant V. Rajan, Daniel B. Kramer & Aaron S. Kesselheim, *Medical Device Postapproval Safety Monitoring Where Does the United States Stand?*, 8 CIRCULATION: CARDIOVASCULAR QUALITY OUTCOMES 1, 3 (2015) (finding that "many of the [adverse event] reports" on which the FDA bases its postmarket surveillance "have flaws").

72. As the U.S. Supreme Court remarked: "The FDA has limited resources to monitor the 11,000 drugs on the market, and manufacturers have superior access to information about their drugs, especially in the postmarketing phase as new risks emerge." Wyeth v. Levine, 555 U.S. 555, 578–79 (2009).

73. *See supra* Part I.B.

74. The FDA started accepting electronic FAERS report submission in 2000 (though some of the data in FAERS date back to 1968), and the volume of received reports started increasing exponentially since around 2009. *See FAERS Dashboard*, *supra* note 36 (showing 107 reports for 1968, followed by a slow increase to tens of thousands in the 70s and 80s and hundreds of thousands in the 90s and 2000s, before an exponential increase from 490,412 in 2009 to 2.34 million in 2021). Manufacturers submit 95 percent of FAERS reports to fulfill FDA reporting requirements, whereas patients, caregivers, and healthcare professionals voluntarily submit the

used its postmarket analyses of adverse event reports data to update regulatory rulemaking and guidance.[75] On rare occasions, it even has relied on postmarket insights to reevaluate premarket approval decisions.[76]

Since incorporating AI into its postmarket surveillance regime, the FDA has announced notable progress on processing its backlog of postmarket surveillance analyses.[77] After focusing exclusively on premarket testing for most of its history, the FDA has made postmarket surveillance an increasing priority over the past fifteen years and now recognizes such ongoing review as a "critical part of the FDA's responsibilities."[78] While using NLP models to analyze FAERS data has been the FDA's principal foray into postmarket surveillance to date, this postmarket shift can expand to more data sources and leverage additional AI technologies.[79]

---

remaining 5 percent. *See* Anne Tobenkin, *An Introduction to Drug Safety Surveillance and FDA Adverse Event Reporting System*, U.S. FOOD & DRUG ADMIN. 24 (Apr. 10, 2018), http://www.learning.proclinical.com/wp-content/uploads/2019/04/PV-DDI-Webinar-FINAL.pdf [https://perma.cc/H88P-W53R].

75. *See* CTR. FOR DEVICES & RADIOLOGY HEALTH, U.S. FOOD & DRUG ADMIN., INFUSION PUMP IMPROVEMENT PROGRAM (Apr. 2010), https://www.fda.gov/medical-devices/infusion-pumps/white-paper-infusion-pump-improvement-initiative [https://perma.cc/33QR-QVX5] (describing how the FDA used its adverse event reports analyses to refine its regulation of infusion pumps).

76. *See* Sharkey, *Regulatory Analysis at the FDA*, *supra* note 4, at 53.

77. *See* SCOTT GOTTLIEB, U.S. FOOD & DRUG ADMIN., STATEMENT BY FDA COMM'R SCOTT GOTTLIEB, M.D., ON THE FDA'S EFFORTS TO HOLD INDUSTRY ACCOUNTABLE FOR FULFILLING CRITICAL POST-MARKETING STUDIES OF THE BENEFITS, SAFETY OF NEW DRUGS (Nov. 16, 2018), https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-fdas-efforts-hold-industry-accountable-fulfilling [https://perma.cc/NW8W-3MAR] (stating that, as of 2018, 76 and 81 percent of the FDA's "post-marketing requirements" and "post-marketing commitments," two types of post-approval studies, were respectively progressing on schedule).

78. *See* U.S. Food & Drug Admin., *The Public's Stake in Adverse Event Reporting*, https://www.fda.gov/drugs/questions-and-answers-fdas-adverse-event-reporting-system-faers/publics-stake-adverse-event-reporting [https://perma.cc/HY47-HM68] (presenting postmarket drug and device safety monitoring as a "critical part of FDA's responsibilities"). Especially with regard to medical devices, the FDA has explored ways to lower premarket barriers while ratcheting up postmarket scrutiny. *See generally*, U.S. FOOD & DRUG ADMIN., DEVELOPING A SOFTWARE PRECERTIFICATION PROGRAM: A WORKING MODEL (2019), https://www.fda.gov/downloads/MedicalDevices/DigitalHealth/DigitalHealthPreCertProgram/UCM629276.pdf [https://perma.cc/A9Y6-AVNA].

79. The FDA has created additional adverse event reporting databases, including the Sentinel Database (for drugs, vaccines, biologics, and medical devices), the Manufacture and User Facility Device Experience (MAUDE) database (for medical devices), and the Tobacco Product Problem database (for tobacco products). *See* U.S. FOOD & DRUG ADMIN., *FDA's Sentinel Initiative*, https://www.fda.gov/safety/fdas-sentinel-initiative [https://perma.cc/R4UE-AGM6]; U.S. FOOD & DRUG ADMIN., *Manufacturer and User Facility Device Experience*, OPENFDA, https://open.fda.gov/data/maude [https://perma.cc/6NSW-MQFX]; U.S. FOOD & DRUG ADMIN.,

B. Evaluating the Shift

As an initial matter, as the FDA confronts regulating a new generation of AI medical devices that incorporate dynamic AI, this shift to postmarket surveillance becomes a practical necessity. The FDA's regulatory guidance regarding novel medical AI devices emphasizes the need for postmarket monitoring of AI models.[80] Postmarket monitoring is accordingly one of the key principles for the FDA's "Good Machine Learning Practice for Medical Device Development" (co-published with Canadian and U.K. health safety regulators in October 2021).[81]

Moreover, the FDA's increase in postmarket surveillance opens up the possibility for more streamlined, less stringent premarket safety review—holding overall safety constant (or even improving it overall). For example, in the realm of regulating AI powered software-based medical devices, the FDA is experimenting with developing a more streamlined premarket review, coupled with ongoing heightened postmarket surveillance.[82] It is in this vein that we must evaluate Marks' concern that the FDA's relaxation of ex ante requirements for premarket clinical trial studies to allow AI-driven simulation models might lead to "erroneous conclusions" (*i.e.*, Type I errors) in drug approval.[83] We, too, might readily agree if viewed in isolation, but

---

Tobacco Product Problem Reports, https://www.fda.gov/tobacco-products/tobacco-science-research/tobacco-product-problem-reports [https://perma.cc/4XRP-29RY].

80. *See supra* notes 26–27 and accompanying text; *TMAP*, *supra* note 62, at 3 (mentioning the need to manage "workload both in the premarket and post-market space").

81. *Good Machine Learning Practice for Medical Device Development: Guiding Principles*, U.S. FOOD & DRUG ADMIN., https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles [https://perma.cc/5VLG-XQM2] (listing as a guiding principle that "Deployed Models Are Monitored for Performance and Re-training Risks Are Managed").

82. *See Informing the FDA's Digital Health Pre-Cert Program*, PRESIDENTIAL INNOVATION FELLOWS: PROJECTS, https://presidentialinnovationfellows.gov/projects/fda-precert [https://perma.cc/DML5-EWSP] (describing project by FDA to "streamline[]" the premarket review of software-based medical devices developed by trusted manufacturers who "are committed to [postmarket] monitoring [of the] real-world performance of their products"). The FDA's use of AI-powered postmarket surveillance thus will equip it to match the iterative improvement framework implemented by dynamically updating AI/ML-based software. *See* SAMD PLAN, *supra* note 26, at 1 ("This framework would enable FDA to provide a reasonable assurance of safety and effectiveness while embracing the iterative improvement power of artificial intelligence and machine learning-based software as a medical device.").

83. Marks, *supra* note 5, at 1223 ("[A]dopting [computer] models [for drug approval] hastily or haphazardly can produce erroneous conclusions. . . . [M]any existing and proposed algorithmic models have not been rigorously evaluated, and their credibility is unknown. Others have known deficiencies that negatively affect their credibility.").

surely such risks are mitigated in a context in which the FDA also simultaneously boosts its postmarket surveillance regime.

Our point here is, at least, to consider the extent to which the reduction of premarket regulatory requirements can thereby lower Type II errors—fostering innovation and minimizing delay to market—without sacrificing overall safety (or Type I errors) by increasing postmarket surveillance. The FDA is experimenting with the collection and evaluation of "real-world evidence" in the postmarket surveillance period to uncover risk evidence that inevitably may be missed from clinical trials or other premarket testing procedures.[84] The FDA has launched multiple projects to collect real-world data, such as the National Evaluation System for health Technology (NEST), which was designed to "help improve the quality of real-world evidence that FDA can use to detect emerging safety signals quickly and take appropriate actions,"[85] and the MyStudiesApp, which was built to "foster the collection of real world evidence via patients' mobile devices" with the goal of helping manufacturers in their design of new health care solutions while complying with "the FDA's regulations and guidance for data authenticity, integrity and confidentiality."[86] By leveraging such "real world evidence," the FDA "may be able to provide patients and providers with important answers much sooner by potentially identifying a broader range of safety signals more quickly."[87]

III. REALIZING THE PROMISE OF AI AT THE FDA

We recognize the tentative nature of our rebuttal to Marks' doomsday predictions. In our minds, the key to the FDA's realizing the

84. *See TMAP*, *supra* note 62, at 2 ("FDA is building the scientific and policy infrastructure to support increasing use of real-world evidence to support regulatory decisions. The 21st Century Cures Act, enacted in 2016, highlighted the importance of real-world evidence in the context of drug development."); *supra* notes 57–58 and accompanying text (emphasizing that clinical trials inevitably mask or fail to uncover drug risks).

85. *Medical Device Safety Action Plan: Protecting Patients, Promotion Public Health*, U.S. FOOD & DRUG ADMIN. 10 (2018), https://www.fda.gov/media/112497/download [https://perma.cc/UP5X-5UK7].

86. *FDA Launches New Digital Tool to Help Capture Real World Data from Patients to Help Inform Regulatory Decision-Making*, U.S. Food & Drug. Admin. (Nov. 6, 2018), https://www.fda.gov/news-events/fda-brief/fda-brief-fda-launches-new-digital-tool-help-capture-real-world-data-patients-help-inform-regulatory [https://perma.cc/4LPF-2GHT].

87. Scott Gottlieb, *Remarks Before the Bipartisan Policy Center, Breaking Down Barriers Between Clinical Trials and Clinical Care: Incorporating Real World Evidence into Regulatory Decision Making*, U.S. FOOD & DRUG. ADMIN. (Jan. 28, 2019), https://www.fda.gov/NewsEvents/Speeches/ucm629942.htm [https://perma.cc/3YC4-A62M].

promise of AI (and mitigating its perils) is the agency's success at building internal AI capacity.[88] Armed with the strong AI-embedded expertise that it has built over the past five years, the FDA has the opportunity to become an AI-enabled "information agency" if it restructures its data operations and adopts a technical data infrastructure relying on "fit-for-purpose" data.

A. Building Internal AI Capacity at the FDA

Whereas federal agencies often lack the technical capability necessary to regulate novel AI products or build AI tools in-house,[89] the FDA emerges as an outlier agency that has invested significant resources to develop internal AI capacity. More specifically (tracking the guidelines articulated in the 2020 *Government by Algorithm* Report), the FDA has: (1) "invest[ed] in [its] technical and data infrastructure," (2) "cultivate[d] in-house human capital to produce AI tools that are not only usable at the technical level but also compliant at the legal and policy levels,"[90] and (3) "invest[ed] in comprehensive and flexible AI strategies that allow [the] agenc[y] to learn strategically from failures and evolve."[91]

First and foremost, the FDA has invested in building human capital expertise in AI. As a protector of public health, the FDA must be prepared to respond to the health emergencies of the future using tools and processes that meet the sophistication of the industry that the FDA regulates. Starting in 2017, the FDA has recruited dozens of technical hires, including engineers, AI experts, and "cloud computing whizzes," to help it adjust to the AI revolution.[92] That same year, the

88. *See Building Internal AI Capacity*, *supra* note 33, at 71–74.

89. *See* Deirdre K. Mulligan & Kenneth A. Bamberger, *Procurement as Policy: Administrative Process for Machine Learning*, 34 BERKELEY TECH L.J. 781 (2019) (showing that agencies "most often lack the technical expertise to design or assess algorithmic systems on their own").

90. Success stories at other agencies, such as the Social Security Administration (SSA), the Internal Revenue Service (IRS), and the Securities and Exchange Commission (SEC), illustrate the need for agencies to have their staff with technical expertise collaborate with their staff with regulatory subject matter expertise in order to implement AI tools effectively for regulatory purposes. *See Building Internal AI Capacity*, *supra* note 33, at 71–73 & n.29–30 (showing that the SSA's success in building tools identifying potential errors in draft disability determinations hinged on its ability to hire lawyers with both regulatory and technical skills, that the IRS built in-house technical expertise to automate dynamic regulatory tasks, and that the SEC relied on internal expertise to iteratively update its AI enforcement tools and prevent regulated entities from gaming its violation detection models).

91. *Id.* at 71.

92. *See supra* note 1 and accompanying text.

agency created an Entrepreneurs in Residence program.[93] In 2021, HHS—the FDA's mother agency—appointed its first Chief AI Officer.[94] Among other initiatives, the Office of the Chief AI Officer released an AI strategy and created a framework for developing "trustworthy AI" within government.[95]

Leveraging this human capital with AI expertise, the FDA has developed a flexible and iterative approach to regulating AI-based medical products by experimenting with regulatory "sandboxes" and building partnerships to complement its own internal AI capacity.[96] Regulatory sandboxes present great advantages for regulating nascent and quickly evolving technologies such as AI. They enable regulators to "fail cheaply" and relatively safely,[97] de-risk projects early, define metrics to measure success, and iterate on their technical infrastructure for regulatory analytics. The FDA has partnered with entrepreneurs and private organizations to run its INFORMED program as a regulatory sandbox focusing on AI-driven oncology innovation. INFORMED created "a unique sandbox for networking, ideation and

93. *See Digital Health Innovation Action Plan*, U.S. FOOD & DRUG. ADMIN. 7 (2017), https://www.fda.gov/media/106331/download [https://perma.cc/2QRR-2YMY].

94. *See About the HHS Office of the Chief Artificial Intelligence Officer (OCAIO)*, U.S. DEP'T HEALTH HUMAN SERVS., https://www.hhs.gov/about/agencies/asa/ocio/ai/ocaio/index.html [https://perma.cc/LU9C-98S7], (last updated Mar. 4, 2022).

95. *Artificial Intelligence (AI) Strategy*, U.S. DEP'T HEALTH HUMAN SERVS. 3 (Jan. 2021), https://www.hhs.gov/sites/default/files/hhs-ai-strategy.pdf [https://perma.cc/R644-HFXV] (presenting HHS's AI strategy); *Trustworthy AI (TAI) Playbook*, U.S. DEP'T HEALTH HUMAN SERVS. 7 (Sept. 2021), https://www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook.pdf [https://perma.cc/R4QA-BV8S] (presenting HHS's "trustworthy AI" framework). The FDA's growing internal AI capacity could strengthen the agency's "innovation role" for AI products, with HHS's Office of the Chief AI Officer essentially serving as an innovation internal coordinator. *Cf.* Rachel E. Sachs, W. Nicholson Price II & Patricia J. Zettler, *Rethinking Innovation at FDA* (forthcoming) (manuscript at 8) (on file with the *Duke Law Journal*) (analyzing "the ways that FDA makes decisions and judgments that shape what products, or new uses of products, are developed (or are believed to be developed) in the future").

96. In its 2019 Technology Modernization Action Plan (TMAP), the FDA emphasized its strategy of adopting a flexible approach combining internal capacity building, collaboration with industry and other government agencies, or purchasing off-the-self solutions:

> "For some projects, FDA will perform the role of a traditional technology developer: seeking and taking a leading role in the technological modernization of our regulatory review system as well as the underlying infrastructure that supports it. . . . Other solutions will be catalyzed by FDA but otherwise built within the larger biomedicine ecosystem, including through collaboration with other government agencies. FDA will also continue to review the overall technology marketplace for new fit-for-purpose off-the-shelf solutions that can be efficiently adopted into the FDA environment."

See TMAP, *supra* note 62, at 6.

97. The FDA, however, must preserve a low risk tolerance even in such regulatory sandbox initiatives, given the potential public health consequences from any mistake in drug or medical device safety regulation.

sharing of technical and organizational resources, empowering project teams with the tools needed to succeed in developing novel data science solutions."[98] Perhaps even Marks would support such AI "sandboxes," which allow the FDA to take the necessary steps to develop stronger independent AI tools without endangering public safety.

With regard to partnerships, Marks raises a flag of caution with regard to the FDA's collaboration with industry partners.[99] Here, we agree that the FDA should take measures to ensure that it does not become reliant on "black box" proprietary AI technologies built by private companies, which could unduly favor these companies' regulatory objectives. Agencies that rely on third-party developers to build their AI tools take on a risk of allowing the regulated industry to gain access to that same developer, thus potentially compromising their tool. For this reason, the 2020 *Government by Algorithm* Report strongly encourages agencies to develop their own internal expertise on AI development and maintenance.[100]

The FDA has, moreover, built partnerships with other agencies and engaged the public to augment its internal technical capacity. In 2016, the FDA started a partnership with the National Institute of Standards and Technology (NIST), which has played an active role in creating standards to evaluate the risk and "trustworthiness" levels of AI tools.[101] While Marks criticizes the FDA's inattention to standards-setting in the realm of AI, he missed an opportunity to explore the possibility that the FDA might draw from NIST's AI Risk Management Framework (which gets nary a mention by Marks).[102]

---

98. Khozin et al., *supra* note 49, at 530.

99. *See* Marks, *supra* note 5, at 1245 (warning against the conflict-of-interest dangers inherent in the FDA's reliance on clinical trial simulation technologies developed by industry stakeholders to evaluate the safety and efficacy of their own products).

100. *See Building Internal AI Capacity*, *supra* note 33, at 71–74.

101. *See Memorandum of Understanding Between the National Institute of Standards and Technology, U.S. Department of Commerce and the Food and Drug Administration*, U.S. DEP'T HEALTH AND HUMAN SERVS., https://www.fda.gov/about-fda/domestic-mous/mou-225-21-006 [https://perma.cc/S2T6-Y2YG] (stating that the FDA and NIST will "collaborate in interdisciplinary research in . . . application of synthetic intelligence (e.g. neural networks, AI/ML), and adaptive process control strategies"); *AI Risk Management Framework*, NAT'L INST. OF STANDARDS AND TECH., https://www.nist.gov/itl/ai-risk-management-framework [https://perma.cc/4H9V-RW2Q] (describing NIST's work to "develop[] a framework to better manage risks . . . associated with artificial intelligence").

102. Instead, Marks looks to whether an industry framework developed by ASME's Verification & Validation 40 Committee could be adapted for FDA's internal use, specifically to the agency's internal assessment of model credibility. *See* Marks, *supra* note 5, at 1264–66. Given the V&V 40 model's shortcomings (as rehearsed by Marks), it is all the more surprising that he

Since 2014, the FDA has solicited input from outside AI experts to help solve technical challenges by running open challenges through its precisionFDA program, which provides "a secure, cloud-based platform where participants can access and share datasets, analysis pipelines, and bioinformatics tools, in order to benchmark their approaches and advance regulatory science."[103] The FDA's sustained efforts in strengthening its internal capacity and engaging partners to augment its expertise give us confidence in the agency's ability to avoid the dark path that Marks predicts.

The FDA faces perhaps its biggest future challenges to internal capacity building with regard to buttressing the agency's technical data infrastructure. Although the FDA has launched initiatives to increase the range of data available to its AI efforts—with a focus on "real-world" data[104]—the agency faces hurdles in turning "big" data into "smart" data. Traditionally, the FDA has used a limited variety of data submitted by manufacturers in randomized clinical trials. These data involved small sample sizes, were collected intermittently, and were stored in highly structured formats. In contrast, the FDA's AI-enabled postmarket surveillance plans will require the agency to collect data from a variety of sources—including real-world data collected directly from patients, in high volume, and on an ongoing basis.

B. AI and the Rise of the FDA as an "Information Agency"

The FDA has reached a significant fork in the road.[105] With FAERS and other postmarket databases, the FDA collects "big," high volume data on an ongoing basis.[106] Will the FDA continue along its traditional path of collecting continuous data streams in unstructured formats, or will it instead chart an alternative path? In our view, the way forward for the FDA to realize its full potential as an "information agency" of the twenty-first century is to chart a new path, leveraging

---

did not consider the NIST framework, which has the added benefit of being developed independently (and external to the FDA and the industry it regulates).

103. *About precisionFDA*, PRECISIONFDA, https://precision.fda.gov/about [https://perma.cc/R88E-V7FL]. The FDA has run a total of eighteen challenges under their precisionFDA program. *precisionFDA Challenges*, *supra* note 45.

104. *See supra* notes 84–87 and accompanying text (highlighting the FDA leadership's focus on collecting and analyzing real-world data).

105. The 2020 *Government by Algorithm* Report uncovered an internal split in potential approaches for the future of regulatory AI at the FDA. *See* Sharkey, *Regulatory Analysis at the FDA*, *supra* note 4, at 56 ("The FDA may be at a crossroads with respect to whether it continues to use NLP to handle unstructured data, or whether it instead restructures its data collection.").

106. In 2021, 2.34 million reports were submitted to FAERS. *See FAERS Dashboard*, *supra* note 36.

AI technologies using "fit for purpose" data, *i.e.*, data whose content and format are optimized for regulatory use. In order to leverage not only "big" but also "smart" data, the FDA should restructure its data collection protocols and collect structured "fit-for-purpose" data in the first instance rather than building out NLP-based tools to extract structured data from unstructured text, such as existing adverse event reports.[107]

The FDA's current data and AI approach to postmarket surveillance regulatory analytics—relying solely on NLP models applied to unstructured text data from databases such as FAERS—raises concerns regarding causality analyses and data interoperability. Although the FAERS pilots reached modest positive results from prioritizing human review of postmarket adverse event reports and uncovering new relationships between drugs and medical conditions,[108] they have proved less successful when attempting to make causal inferences based on unrepresentative data.[109] AI-based predictive analytics capabilities cannot substitute for conventional principles of causal inference, and Marks is thus right to warn that careless reliance on computer models could lead to flawed causation analysis.[110] The FDA's current reliance on unstructured data exacerbates these causal inference challenges. Whereas precision is critical in evaluating causal relationships between drugs and adverse health conditions, the first

107. While the FDA has focused on NLP techniques, it also has considered adopting a data infrastructure relying on "fit-for-purpose" data. *See Enabling More Efficient and Seamless Regulatory Review Processes*, PRESIDENTIAL INNOVATION FELLOWS, https://presidentialinnovationfellows.gov/projects/fda-cio [https://perma.cc/HG4Z-J72Z] (listing "shor[ing] up [FDA's] critical data assets to quality and 'fit for purpose' data" as one of the five challenges tackled by the project); Jacqueline Corrigan-Curay, *The FDA Real-World Evidence (RWE) Framework and Considerations for Use in Regulatory Decision-Making*, U.S. FOOD & DRUG ADMIN. 11 (May 12, 2021) https://www.fda.gov/media/148543/download [https://perma.cc/8HT9-48AS] (mentioning a project to "[d]evelop[] a Reusable Framework for transforming raw data in fit-for-purpose data"). NLP could still be used on top of fit-for-purpose data to perform predictive or causality analyses and would operate on "clean," standardized, and directly actionable data, instead of "messy," freeform unstructured text data requiring pre-processing.

108. *See supra* notes 37–41 and accompanying text.

109. FAERS submissions do not require demonstrating causation. *See* Tobenkin, *supra* note 74, at 25 (FAERS reports can be submitted "even if causality is uncertain"); *Q&A on FAERS*, *supra* note 9 ("FDA does not require that a causal relationship between a product and event be proven, and reports do not always contain enough detail to properly evaluate an event."). FAERS reports also suffer from quality (duplicate data), completeness (missing data), and reliability (unverified data) issues. *See* Tobenkin, *supra* note 74, at 26 ("Quality of the reports is variable and often incomplete; Duplicate reporting of the same case occurs."); *Q&A on FAERS*, *supra* note 9 (listing same issues).

110. *See* Marks, *supra* note 5, at 1243–44 ("[I]nstead of being surrogates for direct observation of clinical effects, models and simulations are surrogates for evidence of causation. Consequently, they may be even less reliable than traditional surrogates.").

step of translating unstructured data into actionable data labels (such as drug name or medical condition) leaves much room for error. Not surprisingly, then, the hepatic failure FAERS pilot fell short in terms of uncovering causal relationships between drugs and hepatic failures based on the available unstructured data.[111] Relying on unstructured data also poses significant challenges to data interoperability, as it is difficult to "join" different data sources that contain unstructured data.[112]

Training and using its AI models, instead, on fit-for-purpose data would alleviate—if not fully resolve—these causal inference and interoperability challenges. Structured data labels would make it easier to verify and enforce data quality, completeness, and reliability.[113] The FDA has taken some preliminary steps in this direction. Its Technology Modernization Action Plan emphasizes interoperability across multiple data sources and between the FDA and external stakeholders.[114]

The FDA has the means to transform most of the data it collects to structured data. Ninety-five percent of FAERS reports come from manufacturers on which the FDA could impose structured data submission requirements.[115] And, its Real World Data Enterprise Proposal involved a $100 million budget to expand the volume and variety of real-world data collected by the FDA to assist postmarket monitoring.[116] Despite a more significant investment required in the short term, the "fit-for-purpose" data approach provides superior prospects for powerful regulatory analytics use cases in the long term. Moving further along this path, moreover, would strengthen the FDA's regulatory role as an "information agency" of the twenty-first century.

---

111. *See supra* note 40 and accompanying text.

112. "Join" is a database operation performed to establish a connection between two or more database tables based on matching columns, thereby creating a relationship between the tables.

113. Based on structured data, the FDA could run straightforward analyses to identify duplicates and missing data and run quality assurance protocols to check for unexpected values.

114. *See TMAP*, *supra* note 62, at 4 ("FDA's action plan [includes] communication and collaboration between FDA and stakeholders, including the technology industry and other government agencies, to drive technological progress that is *interoperable across the system* and delivers value to consumers and patients." (emphasis added)); *id.* at 7 (mentioning the benefits of building "[c]lear technical interfaces for external stakeholders"); *id.* at 8 ("As FDA builds out increasingly advanced technologies, FDA will work with external partners to build appropriate application programming interfaces (APIs) and other tools to allow for the efficient submission of high-quality data to FDA.").

115. *See* Tobenkin, *supra* note 74, at 24.

116. Scott Gottlieb, *FDA Budget Matters: A Cross-Cutting Data Enterprise for Real World Evidence*, U.S. FOOD & DRUG. ADMIN. (June 10, 2018), https://www.fda.gov/news-events/fda-voices/fda-budget-matters-cross-cutting-data-enterprise-real-world-evidence [https://perma.cc/62RC-L5X5].

CONCLUSION

AI holds huge promise for the administrative state—not just for private industry. The question facing federal agencies is not whether, but how and in which domains, to adopt AI for their regulatory purposes. The FDA's early engagement with AI could serve as a model for other agencies. The agency has effectively built internal AI capacity; proactively outlined a regulatory framework tailored to novel AI medical devices featuring "dynamic AI"; and seen some moderate success with its use of NLP to analyze postmarket FAERS adverse safety event data.

The FDA's accelerating transition to a robust postmarket surveillance regime showcases the ability of AI technologies to accelerate regulatory paradigm shifts. The FDA is poised to further refine its data-driven regulatory approach to embrace "fit-for-purpose" data, thereby securing its transformation into an "information agency" of the twenty-first century and providing a window into the future promise of AI in the administrative state.