

July 2017

Trading Safety for Innovation and Access: An Empirical Evaluation of the FDA's Premarket Approval Process

George Horvath

Follow this and additional works at: <https://digitalcommons.law.byu.edu/lawreview>

 Part of the [Consumer Protection Law Commons](#), [Health Law and Policy Commons](#), and the [Medical Jurisprudence Commons](#)

Recommended Citation

George Horvath, *Trading Safety for Innovation and Access: An Empirical Evaluation of the FDA's Premarket Approval Process*, 2017 BYU L. Rev. 991 (2018).

Available at: <https://digitalcommons.law.byu.edu/lawreview/vol2017/iss5/4>

This Article is brought to you for free and open access by the Brigham Young University Law Review at BYU Law Digital Commons. It has been accepted for inclusion in BYU Law Review by an authorized editor of BYU Law Digital Commons. For more information, please contact hunterlawlibrary@byu.edu.

Trading Safety for Innovation and Access: An Empirical Evaluation of the FDA's Premarket Approval Process

*George Horvath**

Congress created the premarket approval process (PMA) to provide a rigorous safety evaluation of high-risk medical devices before they may be sold on the U.S. market. Evaluating a PMA application requires the Food and Drug Administration (FDA) to conduct a lengthy, complex, and costly assessment of the extensive data a manufacturer must submit. But other policy concerns, notably a fear of hampering innovation and a desire to assure timely access to new technologies, have led Congress to relax some of the rigorous data requirements the PMA process imposes on manufacturers. Congress mandates that the FDA employ the “least burdensome” approach to regulation that allows a reasonable assurance of safety. The FDA has interpreted this as permitting it, among other things, to approve high-risk devices based on small, short-duration clinical trials the designs of which fall short of the most rigorous scientific standards. Congress also created “PMA Supplement” pathways that allow manufacturers to modify their PMA-approved devices with only limited supporting data. And Congress included several provisions in the recently-enacted 21st Century Cures Act that further tip the balance away from ensuring device safety.

Scholars writing in the medical literature have raised concerns that the standards for PMA approval have become too relaxed, potentially compromising device safety. But most empirical studies have focused on the less rigorous 510(k) pathway, which is designed for low- and medium-risk devices. These studies provide limited evidence about how frequently PMA-approved devices fail. And no empirical work has examined whether these failures are related to the statutes and

* Darling Foundation Fellow in Public Law and Lecturer, University of California, Berkeley, School of Law. J.D. University of California, Berkeley. M.D. Temple University. B.A. Brown University. I would like to thank Kathy Abrams, Mark Gergen, Kevin Quinn, Bertrall Ross, Avani Meta Sood, Karen M. Tani, and John C. Yoo for their helpful comments, suggestions, and encouragement.

regulations through which Congress and the FDA have attempted to balance safety against innovation and access. This Article begins such an examination, presenting the results of a new empirical study of PMA-approved devices. The study finds that under a best-case scenario at least 4.6%–6% of PMA-approved devices will fail in such a way as to threaten death or serious and permanent harm. Complex cardiovascular devices and devices that have been frequently and rapidly modified through certain PMA supplements are most likely to fail.

Based on the concerns that have been raised and on the findings of this study, this Article suggests that Congress and the FDA should take steps to readjust the balance between safety on one hand and innovation and access on the other. The FDA should insist on scientifically rigorous, longer-duration clinical trials before approving PMA applications. Further, the FDA should limit the number of significant modifications that manufacturers of certain devices are permitted to make to a device through PMA supplements before a thorough safety assessment is required, and should limit how soon after one significant modification is approved that a second modification will be considered. Finally, Congress should amend the 21st Century Cures Act to avoid further tipping the balance between safety, innovation, and access away from the FDA's primary mission of ensuring medical device safety.

CONTENTS

I. INTRODUCTION.....	993
II. THE STATUTORY AND REGULATORY FRAMEWORK OF PREMARKET APPROVAL	999
A. The PMA Process: A Rigorous Framework to Ensure Medical Device Safety	1001
B. The Least Burdensome Approach: Relying on Unreliable Data	1006
C. PMA Supplements and Iterative Device Modification.....	1010
D. The 21st Century Cures Act: Potential to Worsen both Problems	1016

III. THE EMPIRICAL LITERATURE ON DEVICE SAFETY	
FOLLOWING PMA REVIEW	1018
A. FDA-Classified Class 1 Recalls: A Proxy for Failure	
of the PMA Process to Detect Problems	1018
B. Literature Review.....	1021
IV. AN EMPIRICAL STUDY OF PMA-APPROVED	
MEDICAL DEVICES	1025
A. Study Aims and Hypotheses	1025
B. Methodology	1028
C. Results	1032
1. Descriptive findings.....	1032
2. Hypothesis testing	1034
V. A DATA-DRIVEN ANALYSIS OF THE PMA FRAMEWORK....	1039
VI. LIMITATIONS OF THE EXISTING EMPIRICAL DATA ON	
THE SAFETY OF PMA APPROVED DEVICES:	
CONSIDERATIONS FOR FUTURE RESEARCH	1048
VII. CONCLUSION	1052

I. INTRODUCTION

The central question concerning the performance of the Food and Drug Administration (FDA) is whether its regulation of a broad range of products sold in the United States achieves a desirable balance between the often-competing policy objectives of assuring safety, efficacy, innovation, and access.¹ Congress, through the Medical Device Amendments of 1976 (MDA),² assigned to the FDA the primary—and in some cases the sole—responsibility of assuring that medical devices are safe and effective.³ Fulfilling this responsibility requires the FDA to conduct a lengthy, complex, and

1. See generally Robert M. Califf, *Benefit-Risk Assessments at the U.S. Food and Drug Administration: Finding the Balance*, 317 J. AM. MED. ASS'N 693 (2017).

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

3. See 121 CONG. REC. 1859 (1975) (statement of Sen. Kennedy).

costly assessment before a device is marketed, including scrutiny of the uses for which a device is intended; the design, components, manufacturing plans, and labeling; and non-clinical and clinical studies.⁴ But other policy objectives weigh against such a rigorous process.⁵ First, the costs imposed by the premarket review process must not be so high as to discourage manufacturers from engaging in the research and development necessary to create new, potentially life-saving technologies. Clinical studies of sufficient quality to establish device safety and efficacy are expensive and complex.⁶ And the regulatory process itself is expensive.⁷ Second, the delays imposed by the review process must not prevent patients from gaining access to newly-developed technologies in a timely fashion.⁸ Thus, rigorous premarket evaluation must be balanced against the dangers of overly-stringent regulation.⁹

Congress and the FDA have sought to balance these concerns through several mechanisms. The amount and types of data required

4. 21 U.S.C. § 360e(c)(1) (2012) (establishing requirements for premarket evaluation); Aaron V. Kaplan et al., *Medical Device Development: From Prototype to Regulatory Approval*, 109 CIRCULATION 3068, 3072 (2004) (“[T]he demonstration of safety and efficacy for a new medical device is a long, arduous, and expensive developmental path from early concept to introduction into clinical practice.”).

5. Richard A. Deyo, *Gaps, Tensions, and Conflicts in the FDA Approval Process: Implications for Clinical Practice*, 17 J. AM. BOARD FAM. PRAC. 142, 145–47 (2004).

6. See David Steinberg et al., *Building a Business Model in Digital Medicine*, 33 NATURE BIOTECHNOLOGY 910, 914 box2 (2015) (noting that the premarket studies required to bring a new device to market can cost nearly \$100 million).

7. Application fees for approval to market a medical device can be as high as \$310,764. *FY 2018 MDUFA User Fees*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/ForIndustry/UserFees/MedicalDeviceUserFee/ucm452519.htm> (last updated Oct. 3, 2017).

8. Califf, *supra* note 1, at 693.

9. 21 C.F.R. § 814.2 (2017) (stating that premarket review must be both “efficient” and “thorough”); see also *Riegel v. Medtronic, Inc.*, 552 U.S. 312, 326 (2008) (“[T]he solicitude for those injured by FDA-approved devices . . . was overcome in Congress’s estimation by solicitude for those who would suffer without new medical devices”); Califf, *supra* note 1, at 693; Diana M. Zuckerman et al., *Medical Device Recalls and the FDA Approval Process*, 171 ARCHIVES INTERNAL MED. 1006, 1006 (2011). Other concerns include the possibility that over-regulation may have far-reaching effects on the economy. See William H. Maisel, *Medical Device Regulation: An Introduction for the Practicing Physician*, 140 ANNALS INTERNAL MED. 296, 301 (2004); Sriram Thirumalai & Kingshuk K. Sinha, *Product Recalls in the Medical Device Industry: An Empirical Exploration of the Sources and Financial Consequences*, 57 MGMT. SCI. 376, 376 (2011) (citing 2006 data showing overall medical device industry revenue of \$90 billion); cf. Eva Stensvad & Ralph F. Hall, *Left to Their Own Devices: IOM’s Medical Device Committee’s Failure to Comply*, 13 MINN. J.L. SCI. & TECH. 75, 76 (2012) (citing 2011 data that the U.S. device market was \$105.8 billion).

for device approval varies based on risk.¹⁰ Manufacturers may market the lowest-risk (Class I) devices without informing the FDA beforehand.¹¹ Manufacturers seeking to market moderate-risk (some Class I and most Class II) devices must submit a “510(k) notification,” which requires evidence that the device is “substantially equivalent” to a device already on the market.¹² Manufacturers seeking to market the highest-risk (Class III) devices must submit a lengthy and detailed “Premarket Approval” (PMA) application.¹³ This tiered premarket evaluation scheme allows for faster, less expensive approvals for devices that pose only low to moderate risks or that are similar to devices already on the market.¹⁴

Even for the highest-risk, Class III, devices, Congress requires the FDA to choose the “least burdensome” regulatory approach that provides “a reasonable assurance of safety.”¹⁵ Following the least burdensome principle, the FDA frequently grants approvals based on clinical trials the design or execution of which falls short of the generally accepted standards of scientific rigor.¹⁶ Further, Congress

10. 21 U.S.C. § 360c(a) (2012) (establishing classification scheme for devices); *see also* Robert B. Leflar, *Public Accountability and Medical Device Regulation*, 2 HARV. J. L. & TECH. 1, 7–8 (1989).

11. *Class I/II Exemptions*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/ucm051549.htm> (last updated June 26, 2014). These devices, such as tongue depressors and bandages, are “exempted” from premarket review. Manufacturers must often comply with post-marketing requirements, including registering annually with the FDA, ensuring the device labeling comports with FDA regulations, and following determined Good Manufacturing Practices. *Id.*

12. 21 C.F.R. § 807.87(l) (2017).

13. 21 U.S.C. § 360c(a).

14. *See* Stensvad & Hall, *supra* note 9, at 82 (reporting that in 2012 the fee charged to a manufacturer for a 510(k) application was \$4049, compared with \$220,050 for a PMA application); Zuckerman et al., *supra* note 9, at 1009 (citing 2005 data showing that “the average cost for the FDA to review a 510(k) submission was estimated at \$18,200, while a PMA submission cost the agency \$870,000 to review”). For 2018, the user fees are \$10,566 for a 510(k) application and \$310,764 for a new PMA application. *FY 2018 MDUFA User Fees*, *supra* note 7.

15. 21 U.S.C. § 360c(a)(3)(D)(ii); 21 C.F.R. § 860.7; U.S. FOOD & DRUG ADMIN. THE LEAST BURDENSOME PROVISIONS OF THE FDA MODERNIZATION ACT OF 1997: CONCEPT AND PRINCIPLES; FINAL GUIDANCE FOR FDA AND INDUSTRY (2002) [hereinafter LEAST BURDENSOME PROVISIONS], <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm085999.pdf>.

16. *See* Jonas Zajac Hines et al., *Left to Their Own Devices: Breakdowns in United States Medical Device Premarket Review*, PLOS MED., July 2010, at 1, 6 (detailing how “the FDA appears to have permitted scientific approaches that fall short of rigorous” in the PMA

has created pathways—PMA Supplements—through which manufacturers can modify PMA-approved devices without submitting data establishing the safety of “the entire device.”¹⁷ Instead, they need only provide data that is “relevant to the proposed changes.”¹⁸ The FDA is committed to reviewing Supplemental PMA applications quickly, through the use of the most limited form of review available.¹⁹

The impact of these attempts to balance safety with innovation and access through the use of less rigorous review standards and mechanisms is unclear. Courts, legal scholars, and the medical literature have all examined aspects of this balance. Courts, fashioning an expansive preemption doctrine, have been reluctant to disturb the balance Congress and the FDA have struck.²⁰ Legal scholars, criticizing the expansive application of preemption doctrine, have focused on the possibility that limitations in the FDA’s premarket review, coupled with a lack of state oversight, compromises safety.²¹ The medical literature has focused on specific weaknesses in the PMA review process.²²

approval process). As one example, the FDA may approve a device on the basis of a single clinical trial, 21 C.F.R. § 814.20(b)(7), whereas for drugs at least two trials are generally required.

17. 1 CHARLES S. ZIMMERMAN, PHARMACEUTICAL AND MEDICAL DEVICE LITIGATION § 1:17 (2017).

18. 21 C.F.R. § 814.39(c)(1).

19. *Office of Device Evaluation*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHOffices/ucm115879.htm> (last updated Nov. 1, 2017).

20. *Riegel v. Medtronic, Inc.*, 552 U.S. 312, 326 (2008) (noting that Congress’s “solicitude for those injured by FDA-approved devices . . . was overcome in Congress’s estimation by solicitude for those who would suffer without new medical devices”); *Medtronic, Inc. v. Lohr*, 518 U.S. 470, 477 (1996) (“Despite its relatively innocuous phrasing, . . . the ‘premarket approval,’ or ‘PMA’ process, is a rigorous one.”); see also *A Delicate Balance: FDA and the Reform of the Medical Device Approval Process: Hearing Before the Spec. Comm. on Aging*, 112th Cong. 81, 88 (2011) [hereinafter *Hearings*] (statement of Ralph D. Hall, Distinguished Professor and Practitioner, University of Minnesota Law School).

21. See, e.g., Jean Macchiaroli Eggen, *Navigating Between Scylla and Charybdis: Preemption of Medical Device “Parallel Claims,”* 9 J. HEALTH & BIOMEDICAL L. 159, 211 (2013).

22. See *infra* Part III.

But little empirical work has examined how well the PMA pathway serves to ensure device safety.²³ And no empirical work has attempted to link the outcomes of PMA review with the statutory and regulatory framework that structures that review. Of the few empirical studies that have examined the PMA pathway, all but one used a methodology that underestimates the rate at which the overall safety assessment fails to detect problems that can lead to fatal or seriously-harmful failures. The one prior study that used a better-adapted methodology examined only a small number of devices for a short period of time. And no prior study has attempted to uncover the factors that may predict failures in PMA-approved devices.²⁴

This Article presents a novel empirical study of serious failures of PMA-approved devices. Under a set of best-case assumptions,²⁵ 4.6%–6% of PMA-approved, high-risk devices have failed in ways the FDA determined to have created the most serious threats to patients' health and lives.²⁶ These failures have involved widely-distributed devices: over 660,000 individual device units, including life-sustaining products such as artificial hearts, vascular stents, and anesthesia administration systems were potentially at risk of failure.²⁷ These study findings suggest that Congress and the FDA have not managed to strike an acceptable balance between safety on the one hand and innovation and access on the other. I use the study findings to suggest several regulatory and statutory changes in the PMA framework that would result in improved safety for these devices.

The empirical study presented here is particularly relevant because the mechanisms available to ensure safety once a device reaches the market—ex post regulatory methods—are limited. The FDA can require manufacturers of PMA-approved devices to conduct post-market studies and surveillance.²⁸ But the FDA's authority is circumscribed both by statutory provisions and by

23. Most studies of how well FDA premarket evaluation functions to ensure safety have focused on the 510(k) clearance pathway. *See infra*, Part III.

24. *See infra* Part IV.

25. *See infra* Part VI.

26. *See infra* Part IV.

27. *Id.*

28. 21 C.F.R. § 814.82 (2017); JUDITH A. JOHNSON, CONG. RESEARCH SERV., R42130, FDA REGULATION OF MEDICAL DEVICES 11 (2016) [hereinafter JOHNSON 2016].

resource constraints. Commentators in the medical literature have questioned how effective these studies—and the FDA’s responses to their findings—have been.²⁹ A recent U.S. Government Accountability Office (GAO) study found that delays were frequent.³⁰ Even where studies are progressing on target, the FDA’s limited resources and their allocation may limit the Agency’s ability to monitor study progress.

In addition, the other ex post regulatory system that might serve to expose device problems—litigation under state tort and product liability theories—is largely barred from application to PMA-approved devices. Most state-law claims, such as those based on design defect and failure to warn, are expressly preempted by 21 U.S.C. § 360k(a) and the Supreme Court’s 2008 decision in *Riegel v. Medtronic*.³¹ Other claims that might escape express preemption are likely barred under the Court’s implied preemption doctrine set out in *Buckman Co. v. Plaintiffs’ Legal Committee*.³² Thus, so long as we rely on FDA premarket evaluation to ensure that the highest-risk devices are safe, it is important to examine critically how successfully the Agency and Congress have balanced safety with innovation and access.

This Article begins in Part II with an examination of the ways in which Congress and the FDA have attempted to balance the competing priorities of assuring safety and avoiding the inhibition of

29. See Prashant V. Rajan et al., *Medical Device Postapproval Safety Monitoring: Where Does the United States Stand?*, 8 CIRCULATION 124, 128 (2015) (“[I]t is unclear whether [post-market] studies themselves tend to not get done properly or whether the knowledge of their progress is just inexact. Moreover, the exact nature of FDA action after completion of the PAS and 522 studies is also not routinely disseminated.”); Rita F. Redberg et al., *Power Morcellators, Postmarketing Surveillance, and the US Food and Drug Administration*, 318 J. AM. MED. ASS’N 325 (2017); Ian S. Reynolds et al., *Assessing the Safety and Effectiveness of Devices After US Food and Drug Administration Approval: FDA-Mandated Postapproval Studies*, 174 J. AM. MED. ASS’N INTERNAL MED. 1773, 1776–77 (2014) (finding that delays in initiating PAS were common and noting that the FDA has never sanctioned a manufacturer for failing to start or complete a study, “which may undermine its authority in ordering these studies”).

30. U.S. GOV’T ACCOUNTABILITY OFF., GAO-15-815, MEDICAL DEVICES: FDA ORDERED POSTMARKET STUDIES TO BETTER UNDERSTAND SAFETY ISSUES, AND MANY STUDIES ARE ONGOING 14–15 (2015); Rajan et al., *supra* note 29, at 127 (finding twenty percent of FDA-ordered, post-market studies to be delayed). Most delays were attributed to low patient enrollment. *Id.* at 125.

31. 21 U.S.C. § 360k(a) (2012); *Riegel v. Medtronic, Inc.*, 552 U.S. 312, 330 (2008).

32. *Buckman Co. v. Plaintiffs’ Legal Comm.*, 531 U.S. 341 (2001).

innovation and the restriction of access. Part III reviews the available empirical literature on the performance of PMA evaluation as it related to device safety. Part IV presents the methodology and results of my study of high-risk medical devices approved through the PMA process. Part V then interprets the study findings in light of the statutory and regulatory attempts at striking an appropriate balance. I make several suggestions for modifications to the FDA's regulatory requirements and to Congress's statutory framework that govern the PMA process. Finally, Part VI suggests that PMA-approved devices may fail even more frequently than the study data indicates, and outlines some directions for future empirical projects.

II. THE STATUTORY AND REGULATORY FRAMEWORK OF PREMARKET APPROVAL

The FDA bears the primary responsibility for regulating medical devices marketed in the United States.³³ In this role, the agency must balance competing concerns: ensuring that devices are safe and effective while not hampering innovation of and access to new technologies. These goals often conflict,³⁴ simultaneously pushing the FDA's premarket evaluation of devices toward being overly stringent and overly lax.

On one hand, medical device failures impose substantial costs on patients and the health care system. Hundreds of thousands of U.S. patients have been treated with devices that may be more prone to failure than was recognized at the time the FDA granted its approval.³⁵ Government data indicate that over 100,000 people

33. Marcia Boumil, *FDA Approval of Drugs and Devices: Preemption of State Laws for "Parallel" Tort Claims*, 18 J. HEALTH CARE L. & POL'Y 1, 15 (2015). Regulatory authority is shared with other federal agencies, including the Federal Trade Commission, *see* 15 U.S.C. § 52 (2012) (conferring authority over false advertising of drugs and devices), and with common-law tort systems, *see* *Medtronic, Inc. v. Lohr*, 518 U.S. 470 (1996) (federal courts in diversity jurisdiction applying state common law); *Levine v. Wyeth*, 944 A.2d 179 (Vt. 2006) (state court applying common law), *aff'd*, 555 U.S. 555 (2009). The complex issues surrounding the allocation of regulatory authority, especially as between the states and the federal government, are largely beyond the scope of this article.

34. JOHNSON 2016, *supra* note 28, at 1–2; Califf, *supra* note 1, at 1.

35. JOHNSON 2016, *supra* note 28, at 2 (noting that many of the more than 500,000 people in the United States who have received artificial hip joints have required a re-operation because of breakdowns of device components).

suffer a device-related injury annually.³⁶ Independent analysts estimate that in a recent year medical device failures caused over 4500 deaths.³⁷ And the failure of just one medical device, the Medtronic Sprint Fidelis defibrillator lead, may cost Medicare anywhere from \$287 million to almost \$1.2 billion.³⁸ These problems suggest the need for a stringent premarket evaluation of high-risk medical devices.

On the other hand, Congress—and the public—want access to cutting-edge technology as quickly as possible after it is developed. But if the costs imposed by the approval process are too great, innovators may direct their research and development effort elsewhere. Bringing a new, high-risk device “from concept to market” currently costs nearly \$100 million.³⁹ Requiring more or larger or longer clinical studies for approval would increase this cost.⁴⁰ An overly-stringent premarket evaluation process might delay newly-developed technologies reaching the market. Patients would suffer as a result. Further, over-regulation can have broad impacts on

36. JUDITH A. JOHNSON, CONG. RESEARCH SERV., R42130, FDA REGULATION OF MEDICAL DEVICES 2–3 (2012), <https://ia801306.us.archive.org/35/items/R42130FDARegulationofMedicalDevices-crs/R42130%20FDA%20Regulation%20of%20Medical%20Devices.pdf> [hereinafter JOHNSON 2012]; see also *Medical Device Reporting (MDR)*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/MedicalDevices/Safety/ReportaProblem/ucm2005291.htm> (last updated Nov. 7, 2016) (“Each year, the FDA receives several hundred thousand medical device reports of suspected device-associated deaths, serious injuries and malfunctions.”).

37. JOHNSON 2012, *supra* note 36, at 3 (citing data from 2009).

38. Amit K. Mehrotra et al., *Medtronic Sprint Fidelis Lead Recall: Determining the Initial 5-Year Management Cost to Medicare*, 8 HEART RHYTHM 1192, 1196 (2011), <http://www.ncbi.nlm.nih.gov/pubmed/21377552>. This study likely underestimates the total cost to Medicare because it considers only the cost of surgical removal of the leads and not the costs associated with increased monitoring, patient counseling, treatment of anxiety disorders related to the risk of failure and of suffering multiple shocks, and other care. However, similar estimates of the cost to Medicare have been derived by Fidelis plaintiffs’ lawyers. See H. DENNIS TOLLEY, EXAMINING THE SPRINT FIDELIS EFFECT ON MEDICARE COSTS 18 (2010) (estimating cost to Medicare system of Sprint Fidelis failure from \$375 million to \$1 billion).

39. Steinberg et al., *supra* note 6, at 914.

40. Manufacturers also object to a requirement for longer studies because the inevitable postponement in reaching the market can result in significant losses in the highly-competitive medical device marketplace. U.S. FOOD & DRUG ADMIN., UNDERSTANDING BARRIERS TO MEDICAL DEVICE QUALITY 28 (2011), <http://www.fda.gov/downloads/AboutFDA/CentersOffices/CDRH/CDRHReports/UCM277323.pdf> (reporting industry views that “the primary focus for R&D is on timelines and that R&D is not incentivized on embedding quality”).

the economy. The medical device industry employs well over half-a-million workers and generates over \$197 billion in revenues annually.⁴¹ The FDA estimates that the products it regulates account for nearly twenty-five percent of consumer spending.⁴² These considerations argue in favor of a more limited premarket evaluation.

Section II.A reviews the premarket evaluation process that Congress and the FDA have established to ensure the safety of high-risk medical devices. Sections II.B and II.C then examine two subsequent changes in the statutory and regulatory framework that might compromise device safety by permitting approvals based on less-than-rigorous data. Section II.D discusses a provision of the recently-enacted 21st Century Cures Act that may further compromise device safety.

A. The PMA Process: A Rigorous Framework to Ensure Medical Device Safety

Congress enacted the Medical Device Amendments of 1976 (MDA)⁴³ in order to establish a unitary, nationwide regulatory floor ensuring device safety.⁴⁴ The MDA vests all premarket device evaluation and approval authority in the FDA.⁴⁵ For high-risk, “Class III” devices, the MDA outlines a rigorous premarket evaluation process.⁴⁶ Manufacturers seeking to market a Class III device must submit a lengthy and detailed Premarket Approval (PMA)

41. JOHNSON 2016, *supra* note 28, at 2.

42. This includes all FDA-regulated products, including human and animal drugs, food, and cosmetics. *The Food and Drug Administration Celebrates 100 Years of Service to the Nation*, PHARMACEUTICAL ONLINE (Jan. 6, 2006), <https://www.pharmaceuticalonline.com/doc/the-food-and-drug-administration-celebrates-1-0001>.

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

44. *See id.* at 539 (describing purpose as being “to provide for the safety and effectiveness of medical device [sic]”); 121 CONG. REC. 1859 (1975) (statement of Sen. Kennedy) (describing that Congress’s purpose was to “give[] the Food and Drug Administration the necessary authority to require that medical devices be proven safe and effective before they reach the American consumer”).

45. *See* 21 U.S.C. § 360e (2012).

46. *Id.* § 360c(a)(1)(C) (defining Class III devices as those which are intended “for a use in supporting or sustaining human life or for a use which is of substantial importance in preventing impairment of human health, or” which carry “a potential unreasonable risk of illness or injury”).

application.⁴⁷ Although only one percent of devices have been approved through the PMA process, these devices are among the most critical, complex, and costly devices used in health care.⁴⁸ Applicants are required to submit clinical data demonstrating that the device is safe and effective.⁴⁹ This includes “full reports of all information . . . concerning investigations which have been made to show whether or not such device is safe and effective.”⁵⁰ Once a device is approved, the manufacturer is required to adhere to the design, manufacturing, and labeling specifications contained in its PMA application.⁵¹

The extensive information in a new PMA application, typically running to several thousand pages, is often reviewed by an independent panel of experts in the field into which the device is

47. *Id.* New devices for which no predicate device can be identified are by default classified as Class III. 21 U.S.C. § 360c(f); JOHNSON 2016, *supra* note 28, at 22. Manufacturers of these devices can use the “de novo” process to attempt to avoid the expense and delay of a PMA application. JOHNSON 2016, *supra* note 28, at 22. Under the original de novo process, the manufacturer submitted a 510(k) application and a petition for reclassification of the device as a Class I or Class II device. *Id.* The FDA Safety and Innovation Act of 2012 (FDASIA) created a simplified de novo process through which a manufacturer may petition for reclassification without submitting a 510(k) application. *Id.*

48. JOHNSON 2016, *supra* note 28, at 4–5.

49. 21 U.S.C. § 360e(d).

50. *Id.* § 360e(c)(1)(A). Manufacturers must also submit

(B) a full statement of the components, ingredients, and properties and of the principle or principles of operation, of such device;

(C) a full description of the methods used in, and the facilities and controls used for, the manufacture, processing, and, when relevant, packing and installation of, such device;

. . . .

(E) such samples of such device and of components thereof as the Secretary may reasonably require . . . ;

(F) specimens of the labeling proposed to be used for such device; [and]

. . . .

(H) such other information relevant to the subject matter of the application as the Secretary, with the concurrence of the appropriate panel under section 360c of this title, may require.

Id. § 360e(c)(1).

51. 21 U.S.C. § 360e(d)(5)(A)(i) (“A supplemental application shall be required for any change to a device subject to an approved application under this subsection that affects safety or effectiveness”); 21 C.F.R. § 814.80 (2017) (“A device may not be manufactured, packaged, stored, labeled, distributed, or advertised in a manner that is inconsistent with any conditions to approval specified in the PMA approval order for the device.”).

classified.⁵² The panel includes individuals with the specific “training and experience to evaluate the safety and effectiveness of the device[],” including those with “expertise in such fields as clinical and administrative medicine, engineering, biological and physical sciences, and other related professions.”⁵³

In determining the safety (and efficacy) of a device, Congress requires that the FDA weigh the probable health benefits against the probable risk of injury or illness resulting from the use of the device.⁵⁴ The FDA’s decision whether to approve a PMA application is to be made “on the basis of well-controlled investigations, including 1 or more clinical investigations.”⁵⁵ Through these requirements, Congress created a stringent process for the evaluation of the safety of high-risk devices.⁵⁶ But safety was not Congress’s

52. 21 U.S.C. §§ 360c(b)(2), 360e(c)(3).

53. *Id.* § 360c(b)(2). Panels must also include non-voting representatives of consumers and the device industry, and may include members nominated by “[s]cientific, trade, and consumer organizations.” *Id.*

54. *Id.* § 360c(a)(2)(C); Califf, *supra* note 1, at 1.

55. 21 U.S.C. § 360c(a)(3)(A). The provision that follows provides some latitude for the optional use of “valid scientific evidence” other than studies conducted in accord with § 360c(a)(3)(A). *See Id.* § 360c(a)(3)(B).

56. Congress has continued to tinker with the premarket evaluation process, at times conferring additional authority on the FDA that tips the balance in favor of safety. Most of these changes are not directly related to the focus of this Article. The Safe Medical Devices Act (SDMA), addressed several weaknesses of the MDA that had become apparent. Ralph F. Hall & Michelle Mercer, *Rethinking Lohr: Does “SE” Mean Safe and Effective, Substantially Equivalent, or Both?*, 13 MINN. J.L. SCI. & TECH. 737, 747–48 (2012). The Act permitted the FDA to impose a broader range of requirements on devices submitted through the 510(k) process, and to require safety and effectiveness data on devices with different technological characteristics from their predicate device. *Id.* at 748; JOHNSON 2016, *supra* note 28, at 45. The Act also authorized the FDA to require manufacturers of Class III devices that had been on the market when the MDA took effect to submit safety and effectiveness data, and to consider whether to reclassify those devices as Class I or II, or to retain the device as Class III (in which case the manufacturer would be required to submit a PMA application). Hall & Mercer, *supra*, at 750. Further, the Act required hospitals and other facilities that used medical devices to report device-related death, illness, or injury to the FDA. JOHNSON 2016, *supra* note 28, at 45. Under current regulations, manufacturers must report device-related deaths, injuries, and malfunctions to the FDA, while facilities must report only deaths to the FDA. 21 C.F.R. § 803.10 (2017). The FDA Amendments Act of 2007 (FDAAA) subjected clinical trials of medical devices to the reporting requirements related to the clinicaltrial.gov database. ERIN D. WILLIAMS & SUSAN THAUL, CONG. RESEARCH SERV., RL34465, FDA AMENDMENTS ACT OF 2007 (P.L. 110-85), at CRS-67 (2008), research.policyarchive.org/18795.pdf. Before this, manufacturers were not required to report negative studies—studies showing that a device was not effective or not safe. FDAAA required that “results must be posted for clinical trials that form the primary basis of an efficacy claim” for medical device approvals. *Id.* at CRS-68. The

only concern. The need to balance ensuring safety with avoiding a detrimental effect on innovation and access was recognized during the debate over the MDA.⁵⁷ Thus, Congress sought to create a system that provided for “the least regulation necessary to assure safety and effectiveness” of medical devices.⁵⁸

Congress utilized two “sliding scale” structures in its attempt to balance these concerns.⁵⁹ First, the MDA created a three-tiered, risk-based system of device classification.⁶⁰ Class I devices are those that do not pose an unreasonable risk and are not intended for sustaining life, or for which general controls provide reasonable assurance of safety and efficacy.⁶¹ These are simple devices such as examination gloves and bandages.⁶² For these “exempted” devices, no premarket approval or clearance is necessary.⁶³

Class II devices, such as X-ray machines and certain hearing aids,⁶⁴ are considered to present an intermediate level of risk. For these devices, “general controls by themselves are insufficient to provide reasonable assurance of the safety and effectiveness.”⁶⁵ Rather, compliance with “special controls” is deemed sufficient to

FDAAA also sought to reduce conflicts of interest regarding members of the FDA advisory committees that evaluate and recommend whether to approve a PMA application. *Id.* at CRS-60 to CRS-61. The Act also required the FDA to establish a unique device identifier system that would permit the tracking of individual devices. *Id.* at CRS-16.

57. See Joseph R. Radzius, *Medical Devices and Judicial Legislation*, 27 FOOD DRUG COSM. L.J. 639, 640–42 (1972) (emphasizing the need to design a system that balanced safety and innovation).

58. 122 CONG. REC. 13778 (1976) (remarks of Rep. Rogers, Chairman of the Health Subcommittee).

59. See Hall & Mercer, *supra* note 56, at 747; Leflar, *supra* note 10, at 7–8.

60. 21 U.S.C. § 360c(a) (classifying devices based on their intended use, the availability of general controls to ensure safety, the risk inherent in the use of the device, and the seriousness of the condition for which its use is intended); see also Leflar, *supra* note 10, at 7. “Controls” are simply regulations that may apply to all devices or only to certain device types. See Hall & Mercer, *supra* note 56, at 746.

61. 21 U.S.C. § 360c(a)(1)(A); Hall & Mercer, *supra* note 56, at 746.

62. 21 C.F.R. §§ 880.5075, 880.6250 (2017).

63. *Id.* Manufacturers must comply with post-marketing requirements, including registering annually with the FDA, ensuring the device labeling comports with FDA regulations, and following determined Good Manufacturing Practices. *Class I/II Exemptions*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/ucm051549.htm> (last updated June 26, 2014).

64. 21 C.F.R. §§ 892.1630, 874.3300.

65. 21 U.S.C. § 360c(a)(1)(B).

provide an assurance of safety.⁶⁶ Manufacturers seeking to market some Class I and most Class II devices must submit a 510(k) notification, which requires evidence that the device is “substantially equivalent” to a “predicate” device already on the market.⁶⁷ Class III devices, as already noted, must typically be submitted through the rigorous PMA process.⁶⁸

The second sliding-scale that the MDA created was that differing amounts and types of data would be required for approval of devices in each class. The difference between the burdens imposed on Class I and II devices by the 510(k) substantial equivalence process and on Class III devices by the PMA process is profound. Applicants seeking 510(k) clearance are required only to demonstrate that their device is “substantially equivalent” to an already-marketed device (a “predicate” device).⁶⁹ Most importantly, clinical trial data is not typically required.⁷⁰

By providing manufacturers with a far less demanding pathway to the market through the 510(k) process, the MDA balanced rigorous safety assessment with the concerns over innovation and access. Not surprisingly, manufacturers often prefer to submit their devices through the faster and less expensive 510(k) pathway.⁷¹ Unfortunately, the dividing line between devices required to undergo PMA evaluation and devices permitted to utilize the 510(k) clearance pathway is indistinct, affording manufacturers (and the FDA) latitude in deciding what pathway to use. As a result, the Court in *Medtronic, Inc. v. Lohr* noted that the vast majority of new Class III devices, many of which should have been evaluated through the PMA process, had been channeled through the 510(k) pathway.⁷² For many of these evaluations, the predicate device was a

66. Hall & Mercer, *supra* note 56, at 746. Class II devices must also comply with all relevant general controls. *Id.*

67. *Id.* at 739–40.

68. JOHNSON 2016, *supra* note 28, at 22.

69. *Id.* at 4.

70. For a discussion of a different viewpoint, see *infra*, notes 76–77 and accompanying text.

71. Stensvad & Hall, *supra* note 9, at 83–84.

72. *Medtronic, Inc. v. Lohr*, 518 U.S. 470, 479 (1996). This was true even when, by statute, review should have been through the PMA process. See, e.g., Hines et al., *supra* note 16, at 6 (recommending statutory changes aimed at eliminating the permissive approach of the 510(k) pathway to technological differences and “steering such devices toward the PMA

pre-amendment device (one that was already on the market when the MDA was enacted in 1976) for which a safety evaluation had never been completed. Even where the safety of the predicate device had been evaluated, a 510(k) applicant was not required to submit data establishing that the new device is safe and effective.⁷³

The legal and medical literatures have focused extensively on how and to what extent this porous border between the 510(k) and PMA pathways might compromise device safety in an effort to sustain innovation and access. But for devices evaluated through the PMA process, the MDA created a formally rigorous framework of premarket review. The next three sections examine ways in which Congress and the FDA have, both formally and in practice, tipped the balance away from the MDA's priority of assuring device safety.

B. The Least Burdensome Approach: Relying on Unreliable Data

The FDA Modernization Act of 1997 (FDAMA) required the FDA to employ the “least burdensome” approach to regulating devices submitted through the PMA and 510(k) processes.⁷⁴ This statutory provision, and the FDA's implementation of it, could have far-reaching, detrimental effects on the Agency's premarket safety evaluation. Applying the least burdensome principle, the FDA has refrained from exercising its authority to require clinical data when evaluating 510(k) devices and has accepted less-than-rigorous data when evaluating PMA devices.

The Safe Medical Device Act of 1990 (SMDA) had granted the FDA authority to “request additional information,” including clinical data, when conducting a 510(k) review.⁷⁵ The FDA asserts the authority to review “all available safety and effectiveness information available for the medical device.”⁷⁶ But the FDA's own

route”); Zuckerman et al., *supra* note 9, at 1009 (criticizing the FDA for “not fully implementing the law that requires high-risk medical devices to be approved through the PMA process and frequently us[ing] the 510(k) process instead”).

73. Riegel v. Medtronic, Inc., 522 U.S. 312, 317 (2008).

74. 21 U.S.C. §§ 360c(a)(3)(D)(ii), (i)(1)(D)(i) (2012); Hall & Mercer, *supra* note 56, at 750.

75. Hall & Mercer, *supra* note 56, at 748–49, 776. Hall and Mercer argue that these and other changes converted the 510(k) process into one which provides a robust assessment of device safety. *Id.* at 750.

76. *Id.* at 776.

guidance documents that describe its implementation of the least burdensome principle state that the agency will only rarely review the safety data that the manufacturer possesses: “ODE reviewers normally should accept” a manufacturer’s declaration that its device meets the relevant 510(k) requirements, and “normally should not require the submission of information demonstrating conformity with the standard.”⁷⁷

Although much of the literature examining the PMA and the 510(k) approval processes has assumed that the PMA process represents the gold standard for establishing device safety,⁷⁸ the FDA’s implementation of the least burdensome principle has also impacted the formally rigorous design of the PMA process.⁷⁹ In a guidance document originally released in 2002, the Agency outlined its view that the principle required it to approve high-risk devices based on studies that employed less-than-rigorous protocols. These included studies that did not directly compare treatment and non-treatment groups, and studies that were analyzed using alternative statistical methods.⁸⁰ Further, the guidance noted that the FDA would accept studies that utilized “surrogate endpoints,” that is, outcomes other than the clinical outcome the device was supposed to achieve.⁸¹

77. U.S. FOOD & DRUG ADMIN., USE OF STANDARDS IN SUBSTANTIAL EQUIVALENCE DETERMINATIONS 6 (2000), <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm073756.pdf>.

78. Benjamin N. Rome et al., *FDA Approval of Cardiac Implantable Electronic Devices via Original and Supplement Premarket Approval Pathways, 1979-2012*, 311 J. AM. MED. ASS’N 385, 390 (2014) (“The PMA process supporting FDA approval has long been considered the gold standard for rigorously establishing the safety and effectiveness of high-risk medical devices.”); see also Hines et al., *supra* note 16, at 6 (recommending statutory changes aimed at eliminating the permissive approach of the 510(k) pathway to technological differences and “steering such devices toward the PMA route”); Zuckerman et al., *supra* note 9 at 1009 (criticizing the FDA for “not fully implementing the law that requires high-risk medical devices to be approved through the PMA process and frequently us[ing] the 510(k) process instead”).

79. See LEAST BURDENSOME PROVISIONS, *supra* note 15 at 3–4.

80. *Id.*

81. *Id.* at 4. See also *FDA Facts: Biomarkers and Surrogate Endpoints*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/AboutFDA/Innovation/ucm512503.htm> (last updated July 21, 2017). For a comparison of the rigorous clinical trial requirements for new drug applications, see 21 C.F.R. § 314.126(b) (2017). This regulation defines “adequate and well-controlled” studies as those that have a clearly-stated, predefined protocol, control of all relevant variables, comparison of treatment to placebo, blinding of participants and

Work by academic medical researchers has demonstrated how widespread the FDA's reliance on less-than-rigorous data has become.⁸² One study found that nearly two-thirds of original PMA approvals for cardiovascular devices had been granted on the basis of a single study.⁸³ This and another study of cardiovascular devices concluded that premarket approval "by the FDA is often based on studies that lack adequate strength and may be prone to bias."⁸⁴ Few (27%) of the studies examined were randomized and fewer (14%) were blinded.⁸⁵ Many were poorly designed and did not adequately control for the relevant variables.⁸⁶ The majority of studies (88%) used "surrogate end points" to assess device safety and efficacy.⁸⁷ Thus, most device approvals were granted on the basis of studies that never examined the actual effects for which the device was to be marketed. And many of these studies were flawed in other ways—

investigators, and randomized assignment to treatment or placebo. *Id.* Other methodologies, such as the use of historical controls, are disfavored. *Id.*

82. See Rome et al., *supra* note 78.

83. Sanket S. Dhruva et al., *Strength of Study Evidence Examined by the FDA in Premarket Approval of Cardiovascular Devices*, 302 J. AM. MED. ASS'N 2679, 2684 (2009). By contrast, the new drug application (NDA) process, which the PMA process was designed to emulate, Hall & Mercer, *supra* note 56, at 747, typically requires at least two clinical studies, Hines et al., *supra* note 16, at 2.

84. Dhruva et al., *supra* note 83, at 2679; Daniel B. Kramer et al., *Premarket Clinical Evaluation of Novel Cardiovascular Devices: Quality Analysis of Premarket Clinical Studies Submitted to the Food and Drug Administration 2000–2007*, 17 AM. J. THERAPEUTICS 2, 4–6 (2010).

85. Dhruva et al., *supra* note 83, at 2680. "Blinding" prevents study subjects' and investigators' knowledge from influencing the outcome. "Single-blind" studies are those in which subjects do not know if they are receiving the treatment under study or a placebo. Without blinding, subjects who know they are in the active treatment group are more likely to report improvement than those in the placebo-treated group. In "double-blind" studies, neither the subjects nor the investigators know whether each subject's treatment is real or placebo. This prevents the investigators' knowledge, and potential desire for a particular outcome, from influencing the study outcome. Blinding is often considered impossible in studies involving devices. For example, the only way to blind patients and investigators in a study of an implanted device would be to perform a "sham" surgery in the control group.

86. Kramer et al., *supra* note 84, at 4.

87. Dhruva et al., *supra* note 83, at 2682. The FDA currently notes that "endpoints denoting clinical benefit are usually measured directly, but in some cases may be demonstrated by use of validated surrogate endpoints." U.S. FOOD & DRUG ADMIN., GUIDANCE FOR INDUSTRY AND FOOD AND DRUG ADMINISTRATION STAFF: FACTORS TO CONSIDER WHEN MAKING BENEFIT-RISK DETERMINATIONS IN MEDICAL DEVICE PREMARKET APPROVAL AND DE NOVO CLASSIFICATIONS 8 (2016), <http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm517504.pdf>.

they were of short duration, contained discrepancies in the numbers of patients enrolled and reported, and used “post hoc” data analyses.⁸⁸ In sum, “[p]oorly defined safety and effectiveness end points, poor patient accounting, and incomplete collection of important patient comorbidities make device safety and effectiveness assessments more challenging.”⁸⁹

Two other groups that examined larger sets of Class III devices (not limited to cardiovascular devices) found similar problems. Vinay Rathi and colleagues at the Yale Medical School studied twenty-eight devices for which the FDA granted an original PMA approval in 2010 and 2011, finding that many relied on a single pivotal study, failed to use a blinding protocol, and did not include a comparator group.⁹⁰ They also noted that all of the pivotal studies had been funded by the device industry.⁹¹ Another group discussed examples of device approvals based on poorly designed research protocols that were possible because “for devices, the regulations permit ‘reliance upon other valid scientific evidence . . . even in the absence of well-controlled investigations.’”⁹²

Thus, in spite of a formally demanding and thorough PMA review process, the FDA, implementing the least burdensome

88. Dhruva et al., *supra* note 83, at 2683. Problems that develop after a few years are unlikely to be detected by a study that follows subjects for six months. The problem with discrepancies between the numbers of subjects enrolled and the number reported is that the unreported subjects’ data could completely change the study results: A study that enrolled 100 patients, reported data on 90, and found no adverse events in those 90 might be interpreted to show that a device is safe. However, if the 10 patients whose data was not reported suffered significant harm from the device, an adverse event rate of 10%—likely enough to preclude approval—goes undetected. “Post hoc” data analysis refers to analyses that were not planned in advance. The risk here is that the study design may not have been sufficient to ensure that all relevant variables were controlled for and that investigators may “slice and dice” the data to get the desired outcome.

89. Kramer et al., *supra* note 84, at 2.

90. Vinay K. Rathi et al., *Characteristics of Clinical Studies Conducted Over the Total Product Life Cycle of High-Risk Therapeutic Medical Devices Receiving FDA Premarket Approval in 2010 and 2011*, 314 J. AM. MED. ASS’N 604, 607–08 (2015).

91. *Id.* at 608.

92. Hines et al., *supra* note 16, at 2. Hines presented the example of a “vagus nerve stimulator (VNS), a surgically implanted device for treatment-resistant depression.” *Id.* Only one rigorous, randomized controlled trial had been conducted, and it had failed to demonstrate an improvement. *Id.* The FDA granted PMA approval on the basis of a non-blinded follow up that compared VNS-treated patients with other patients who were enrolled at different times and who received other treatments as well. *Id.*

principle, has frequently approved Class III medical devices based on limited and low-quality data. Some of these compromises are inescapable in device evaluations. Conducting randomized, double-blinded studies is sometimes impossible for medical devices.⁹³ Especially for implanted devices, a double-blind study would require the surgical implantation of a sham device, an ethically unacceptable procedure. But as it has implemented the least burdensome principle, the FDA has granted PMA approvals based on less-than-rigorous data, potentially compromising the balance between device safety and innovation/access.⁹⁴

C. PMA Supplements and Iterative Device Modification

Medical device development is often iterative, occurring in small, evolutionary (as opposed to large, revolutionary) steps. A requirement that manufacturers submit a new PMA application, complete with clinical studies and thousands of pages of documentation, for a modified version of an approved device would maximize device safety but would also impede the iterative model of device development. To balance safety with innovation and access, FDAMA created a number of abbreviated “PMA Supplement” processes through which manufacturers may gain approval for relatively minor modifications to their PMA-approved devices without a full PMA evaluation.⁹⁵ The Medical Device User Fee and Modernization Act of 2002 (MDUFMA) added several additional categories of supplemental PMAs that allow for more significant modifications.⁹⁶

93. Dhruva et al., *supra* note 83, at 2683.

94. *Id.* at 2684.

95. Food and Drug Administration (FDA) Modernization Act of 1997, Pub. L. No. 105-115, 111 Stat. 2296 (codified in scattered sections of 21 U.S.C.); 21 C.F.R. § 814.39 (2002) (allowing manufacturers to make changes in device labeling, indications for use, and manufacturing site or processes, and “changes in the performance or design specifications, circuits, components, ingredients, principles of operation, or physical layout of the device,” among other reasons).

96. Medical Device User Fee and Modernization Act of 2002, Pub. L. No. 107-250, 116 Stat. 1588 (codified in several provisions of the Food, Drug, and Cosmetic Act); U.S. FOOD & DRUG ADMIN., GUIDANCE FOR INDUSTRY AND FDA STAFF: MODIFICATIONS TO DEVICES SUBJECT TO PREMARKET APPROVAL (PMA) - THE PMA SUPPLEMENT DECISION-MAKING PROCESS 2 n.2 (2008), <https://www.fda.gov/downloads/MedicalDevices>

The supplemental categories now available to manufacturers can be divided into two groups. The first are those indicated for significant modifications. Panel Track Supplements are indicated for modifications that result in “a significant change in design or performance of the device, or a new indication for use of the device, and for which substantial clinical data are necessary to provide a reasonable assurance of safety and effectiveness.”⁹⁷ Panel Track Supplements require the submission of “substantive clinical data.”⁹⁸ And 180-Day Supplements are indicated for modifications involving “a significant change in components, materials, design, specification, software, color additives, or labeling.”⁹⁹ Typically, “only new preclinical testing is needed to demonstrate reasonable assurance of safety and effectiveness of the modified device.”¹⁰⁰ The second group consists of supplemental processes indicated for minor design and labelling modifications.¹⁰¹ These include Real-Time Supplements and

/DeviceRegulationandGuidance/GuidanceDocuments/UCM270606.pdf [hereinafter FDA 2008 GUIDANCE].

97. FDA 2008 GUIDANCE, *supra* note 96, at 7–11 (citing as examples the request for approval of an additional indication—and a new patient population—for a ventricular assist device which had originally been approved only for patients awaiting a heart transplant, adding patients who were not transplant candidates; the request for approval of a new surgical usage of a laser originally used to reshape the outer surface of the cornea to correct poor vision, adding a surgical usage to cut the cornea and then reshape the cornea’s inner surface; and a request for approval for use an artificial heart valve originally approved for use in the aortic position, for use in the mitral position).

98. *Id.* at 7.

99. *Id.* at 11–14 (citing as examples changes in the way the material used in aortic stent was woven together, and changes in the power supply of a ventricular assist device from compressed air to electricity).

100. *Id.* at 11 (citing as examples changes in the way the material used in aortic stent was woven together, and changes in the power supply of a ventricular assist device from compressed air to electricity).

101. Real-Time Supplements are indicated for “minor change[s] to the design of the device, software, sterilization, or labeling.” *Id.* at 15–17 (citing as examples changes to the circuitry controlling a ventricular assist device’s battery usage, the sterilization procedure used for cardiac ablation catheters, and the method of bonding a balloon to the catheter in a transurethral microwave ablation system). Typically, these applications are supported only by bench testing or testing done by the methodologies of a single scientific field. *Id.* Manufacturing Site Change Notices are indicated for a change to the “facility or establishment to manufacture, process, or package the device.” *Id.* at 21. Changes-Being Effected are used when a manufacturer first learns of information about device safety that was not previously submitted to the FDA, and which prompts “labeling changes that add or strengthen a contraindication, warning, precaution, or information about an adverse reaction for which there is reasonable evidence of a causal association.” *Id.* at 17. Finally, 30-Day Notices (which

Manufacturing Site Change Notices. For minor changes to manufacturing processes, the 30-Day Notice process allows a manufacturer to inform the FDA of relatively trivial changes and begin marketing the device using the new process thirty days later.¹⁰²

Only limited data is necessary to support a PMA Supplement: the manufacturer must present only data related to the specific changes being made.¹⁰³ Clinical data about the safety of the modified device is often not required.¹⁰⁴ The underlying assumption is that the overall safety of a modified device is evaluated in steps, with an initial, complete assessment including full clinical trial data at the time of an original PMA application and a limited assessment at the time of each PMA Supplement application. This cumulative assessment process reduces the burden on a manufacturer: by allowing manufacturers to rely on previously submitted data concerning aspects of a device that have not been directly modified,¹⁰⁵ the process permits modifications without large new studies and duplicative paperwork.

But this process carries a risk: a change to one component of a complex device may have unpredictable effects on the functioning of the device as a whole, which limited testing may fail to disclose. Although a full consideration of the field of complexity is beyond the scope of this article, it is useful to flag two kinds of complexity that are relevant. “Connectivity complexity” refers to “the number of

are not considered a Supplemental PMA application) permit “a PMA applicant to submit written notification to the agency of a modification to the manufacturing procedure or method of manufacture affecting the safety and effectiveness of the device rather than submitting such change as a PMA supplement.” *Id.* at 19.

102. *Id.*

103. *Guidance for Industry - Supplements to Approved Applications for Class III Medical Devices: Use of Published Literature, Use of Previously Submitted Materials, and Priority Review*, U.S. FOOD & DRUG ADMIN. (May 20, 1998), <http://www.fda.gov/RegulatoryInformation/Guidances/ucm080183.htm> (“The abbreviated regulatory requirements for PMA supplemental applications are established under 21 C.F.R. § 814.39(c), which states that ‘all procedures and actions that apply to an application under Sec. 814.20 also apply to PMA supplements except that the information required in a supplement is limited to that needed to support the change.’”).

104. *Id.* (“Nonclinical data may be sufficient to demonstrate that the design/product modification creates the intended additional capacity, function, or performance of the device. The new provision clarifies, however, that FDA may require, when necessary, additional clinical data to evaluate the modification of the device to provide a reasonable assurance of safety and effectiveness.”).

105. *Id.*

relations/interconnections between the components of a given system.”¹⁰⁶ Some medical devices incorporate a large number of components that are closely related in space and through multiple connections. As connectivity complexity increases, a change to one component becomes more likely to change the function of the whole.¹⁰⁷ “User complexity” refers to the cognitive demands placed on a managing health care provider or a device user.¹⁰⁸ Some devices are fully-automated, thus placing minimal demands and allowing few opportunities for human error. Other devices require a great deal of sophistication and training.¹⁰⁹ A small change to a single component can have unpredictable effects on both the device function and the possibility for user error.

There are currently no limits on how many times a manufacturer may modify a PMA-approved device through the use of PMA supplements. This iterative process, with only a limited safety assessment performed for each modification, magnifies the risk that complex interactions between device components or between the device and its users may not be completely understood because the safety evaluation required for a PMA Supplement is limited.

Commentators have raised related concerns that the 510(k) pathway, which permits iterative changes to a device without an assessment of safety, has compromised device safety through “predicate creep.”¹¹⁰ The MDA permits 510(k) clearance for devices incorporating “significant change[s] in the materials, design, energy source, or other features of the device from those of the predicate device.”¹¹¹ Thus, a new device (“ D_i ”) may be approved despite having significant differences from its predicate device (“ D_0 ”).

106. J.W. Senders, *On the Complexity of Medical Devices and Systems*, 15 QUALITY & SAFETY HEALTH CARE i41, i42 (2006).

107. *Id.* at i42–i43.

108. *See id.* at i41.

109. *See CLIA Categorizations*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/IVDRegulatoryAssistance/ucm393229.htm> (last updated Apr. 16, 2014) (categorizing clinical laboratory testing devices based on, inter alia, the amount of knowledge and training required by users and the degree of automation).

110. *See INST. OF MED., MEDICAL DEVICES AND THE PUBLIC’S HEALTH: THE FDA 510(K) CLEARANCE PROCESS AT 35 YEARS* 89 (2011); Hines et al., *supra* note 16, at 4.

111. 21 U.S.C. §§ 360c(i)(1)(A)(ii), (B) (2012). The FDCA does require that the technological changes “not raise different questions of safety and effectiveness than the predicate device.” *Id.* § 360c(i)(A)(ii).

Predicate drift arises because of iterative 510(k) approvals: a newer device (“ D_2 ”) can be approved based on substantial equivalence to D_1 in spite of technological differences between D_1 and D_2 . Through dozens of iterations, device D_{25} or D_{50} may incorporate technology that is radically different from the original predicate device, D_0 . Even if D_0 had been subjected to a thorough safety evaluation, which is frequently not the case because many original predicates were pre-amendment devices, D_{25} or D_{50} have not.¹¹²

The academic medical literature has raised the concern that repeated device modification through the PMA Supplement pathways may similarly compromise device safety. Ben Rome and colleagues found that the seventy-seven cardiovascular devices that received original PMA approval between 1979 and 2012 had been modified through a total of 5829 supplements, with a median of 50 supplements per original device.¹¹³ They noted that “clinical data are rarely collected as part of PMA supplement applications prior to marketing.”¹¹⁴ They raised the concern that through iterative PMA supplements “minor design changes may accumulate over time and in some cases may add up to substantial changes from the device approved in the original PMA application.”¹¹⁵

An example of this analog to predicate creep is provided by the evolution of one complex device over a fifteen-year period. The FDA granted an original PMA approval to Medtronic for its Transvene defibrillator lead system in 1993.¹¹⁶ The original Transvene lead had a diameter of 3.3 to 4 mm, polyurethane insulation, and a co-axial design in which three conducting wires ran the length of the lead

112. See Hines et al., *supra* note 16, at 4 (providing the example of “a screening test for illicit drugs [that] ultimately allowed for the clearance of a malignancy diagnostic test,” after several 510(k) iterations, all without safety vetting); William H. Maisel, Premarket Notification: Analysis of FDA Recall Data 30 (July 28, 2010) (unpublished presentation), <http://iom.nationalacademies.org/-/media/Files/Activity%20Files/PublicHealth/510kProcess/2010-JUL-28/05%20Maisel.pdf> (concluding that a large number of predicates was correlated with an increased risk of recall).

113. Rome et al., *supra* note 78, at 387.

114. *Id.* at 389.

115. *Id.* at 390.

116. *Premarket Approval* (PMA), U.S. FOOD & DRUG ADMIN., <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P920015> (last updated Jan. 8, 2018) [hereinafter *Premarket Approval*]. A defibrillator lead conducts electrical signals from a patient’s heart to a defibrillator implanted under the skin over the chest, and conducts a large shock from the defibrillator to the heart if the heart develops a dangerous arrhythmia.

wound concentrically around a central hollow core.¹¹⁷ By late 2007, Medtronic had submitted thirty-eight supplemental PMA applications, of which the FDA had granted thirty-three.¹¹⁸ The lead, by then marketed as the Sprint Fidelis, contained as many as six conductors.¹¹⁹ Instead of a co-axial design, the lead utilized a multilumen design in which each conductor was tightly wound around its own core, each running parallel to the others.¹²⁰ The insulation incorporated both silicone and polyurethane.¹²¹ And the diameter of the lead was markedly smaller at 2.2 mm.¹²² Even if each of these individual changes—and others not discussed here—did not adversely affect the safety of the lead, the combination may have.¹²³ In October 2007, the FDA announced a Class 1 recall because the lead was exhibiting an unacceptably high failure rate.¹²⁴

Rome and colleagues raised an important question: When should manufacturers be required to submit device modifications—regardless of how seemingly minor—through an original PMA

117. Robert G. Hauser, *A Better Method for Preventing Adverse Clinical Events Caused by Implantable Cardioverter-Defibrillator Lead Fractures?*, 118 CIRCULATION 2117, 2117 (2008).

118. See *Premarket Approval*, *supra* note 116. This webpage lists the basic information for the original PMA approval and the Supplemental PMA applications granted by the FDA. As of October 25, 2007, the date of the Fidelis recall, Medtronic had submitted thirty-eight Supplemental PMA applications. See *id.* (follow link to S038 (application submitted Sept. 12, 2007, approved Feb. 6, 2008)). As of October 25, 2007, the FDA had approved thirty-five Supplemental PMA applications. See *id.* (follow link to S035 (application approved Nov. 7, 2006)). No FDA action is listed for Supplemental PMA applications number 18 and 19. *Id.*

119. See KENNETH A. ELLENBOGEN ET AL., CLINICAL CARDIAC PACING, DEFIBRILLATION, AND RESYNCHRONIZATION 174 (3d ed. 2007); Joshua D. Lovelock et al., *ICD Leads Prone to Failure: Weighing the Risks at the Time of Pulse Generator Exchange*, 4 J. INNOVATIONS CARDIAC RHYTHM MGMT. 1108, 1109 (2013); Charles D. Swerdlow & Kenneth A. Ellenbogen, *Implantable Cardioverter-Defibrillator Leads: Design, Diagnostics, and Management*, 128 CIRCULATION 2062 (2013).

120. See ELLENBOGEN ET AL., *supra* note 119, at 174.

121. Robert G. Hauser et al., *Early Failure of a Small-Diameter High-Voltage Implantable Cardioverter-Defibrillator Lead*, 4 HEART RHYTHM 892, 892–93 (2007).

122. See ELLENBOGEN ET AL., *supra* note 119, at 174.

123. Christopher M. Janson et al., *Implantable Cardioverter-Defibrillator Lead Failure in Children and Young Adults: A Matter of Lead Diameter or Lead Design?*, 63 J. AM. C. CARDIOL. 133, 136 (2014).

124. *Medical Device Recalls*, U.S. FOOD & DRUG ADMIN., <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfres/res.cfm> (input October 25, 2007 into the “Recall Date” fields, and narrow the “Recall Class” to “Class 1”) (last updated Jan. 5, 2018) [hereinafter *Medical Device Recalls*].

application?¹²⁵ Stated in another way, when should the FDA evaluate a modified device as though it were a new device? According to the FDA guidance, a new device is one in which the

design [is] so different from the original version that the pre-clinical . . . and clinical data . . . previously submitted on [the] original device are not applicable (i.e., not supportive) for the specific change in demonstrating a reasonable assurance of the safety and effectiveness of the modified device.¹²⁶

But determining the overall effect of any small change on a complex system is difficult. And determining the overall effect of repeated small changes is vastly more difficult. Thus, the availability of PMA Supplement pathways as means by which manufacturers can iteratively and extensively modify high-risk devices carries the potential to compromise device safety.

D. The 21st Century Cures Act: Potential to Worsen both Problems

The recently-enacted 21st Century Cures Act has the potential to exacerbate the data insufficiencies outlined in section II.B and section II.C of this Article. Section 3058 of the Act emphasizes the least burdensome approach, proscribing the FDA from demanding more than “the minimum required information that . . . provides a reasonable assurance of the safety and effectiveness of the device.”¹²⁷ Further, for “Breakthrough Devices,” the FDA is permitted to “facilitate, when scientifically appropriate, expedited and efficient development and review of the device through utilization of timely *postmarket* data collection,” thus minimizing premarket data requirements.¹²⁸ And for other Class III medical devices, this shift to post-market data is emphasized: “[T]he Secretary *shall* consider the role of postmarket information in determining the least burdensome

125. Rome et al., *supra* note 78, at 390.

126. FDA 2008 GUIDANCE, *supra* note 96, at 5 (footnotes omitted).

127. 21st Century Cures Act, Pub. L. No. 114-255, § 3058(a), 130 Stat. 1033, 1129 (2016) (codified at 21 U.S.C. § 360e(c)) (requiring training for every FDA employee who is involved in the premarket assessment of medical devices and including the medical device industry in periodic audits of how well the Agency is implementing this approach).

128. *Id.* § 3051 (codified at 21 U.S.C. § 360e-3) (emphasis added).

means of demonstrating a reasonable assurance of device safety and effectiveness.”¹²⁹

The problem with relying on post-market studies and data is that under this approach the entire population is enlisted as a study group. Newly-approved devices are often used in large numbers of patients in a very short time. The Sprint Fidelis lead was implanted into 268,000 patients within three years of approval.¹³⁰ It was only then—after hundreds of thousands of patients had received these leads—that the increased risk of failure was detected. A premarket study that required longer follow up would have detected this problem.¹³¹ Reliance on post-market data collection is not sufficient to prevent widespread harm—by the time data is collected and analyzed, the failures and harms will already have occurred. The emphasis on the least burdensome approach and the shifting of data production burdens to the post-market period may impact the evaluation of both new and supplemental PMA applications.

* * *

Consumers, manufacturers, and Congress constantly pressure the FDA to tip the scales in favor of permitting innovation and access, often at risk of compromising its mission of assuring device safety. At a minimum, these pressures suggest that we should closely evaluate how well the PMA process functions to ensure the safety of the most high-risk devices sold in the United States. The next Part examines the available empirical studies that might shed light on how well the balance has been struck.

129. *Id.* § 3058(b) (emphasis added).

130. Dhruva et al., *supra* note 83, at 2684.

131. The Sprint Fidelis lead failure rate currently stands at approximately 20%. *CRHF Product Performance: 6949 Sprint Fidelis*, MEDTRONIC, <http://wwwp.medtronic.com/productperformance/model/6949-sprint-fidelis.html> (last updated July 31, 2017). A study involving a few thousand subjects followed for three years, at which time the failure rate was 6.5% (which is significantly higher than other leads), could have found this problem, exposing far fewer individuals to harm.

III. THE EMPIRICAL LITERATURE ON DEVICE SAFETY FOLLOWING PMA REVIEW

Most empirical work on device safety has focused on devices cleared by the 510(k) pathway.¹³² Of the studies that did examine PMA-approved device failures, several did not report the number of devices that were at risk of failure, making it impossible to determine the proportion of approved devices that failed and, thus, how reliably the PMA pathway functions.¹³³ In this Part, I first describe the proxy that all studies of device safety have used to highlight the limitations of these studies. Then I review the few studies that have calculated the risk that the PMA process fails to ensure the safety of high-risk devices.¹³⁴ These studies provide no definitive answers, but they do form a starting point for the study I present in Part IV.

A. FDA-Classified Class I Recalls: A Proxy for Failure of the PMA Process to Detect Problems

An optimal assessment of how well FDA premarket evaluation serves to ensure device safety would directly identify device failures from an accurate and comprehensive source, determine how many of these failures were systematic (as opposed to sporadic), and calculate the ratio of devices that exhibited significant systematic failures to the number of devices at risk of failure. But in spite of mandatory reporting requirements, device failures most often go unreported.¹³⁵ Even where failures are reported, the available databases in which reports are collected are not easily searchable. Further, the poor quality of the reports in those databases often makes determining the

132. See Zhivko Bliznakov et al., *Analysis and Classification of Medical Device Recalls*, 14 INT'L FED'N FOR MED. & BIOLOGICAL ENGINEERING PROC. 3782 (2006); Marta L. Villarraga et al., *Medical Device Recalls from 2004 to 2006: A Focus on Class I Recalls*, 62 FOOD & DRUG L.J. 581 (2007).

133. See, e.g., Zuckerman et al., *supra* note 9.

134. I will not discuss here studies that did not distinguish recalls of PMA approved from 510(k) cleared devices, including those mentioned *supra*, note 132.

135. See James R. Ward & P. John Clarkson, *An Analysis of Medical Device-Related Errors: Prevalence and Possible Solutions*, 28 J. MED. ENGINEERING & TECH. 1, 5 (2009) (describing underreporting rates for device-related medical errors ranging from 10 to 80%); *Hearings*, *supra* note 20, at 11 (statement of Frederic S. Resnic) (citing GAO estimate that only 0.5% of medical device adverse events are reported to the FDA).

root cause and even the seriousness of the failure impossible.¹³⁶ And because manufacturers change the names under which they sell their devices after often-trivial modifications, examining individual reports may fail to detect patterns arising from systematic failures of a single technology that is sold as several apparently distinct devices. Rather than attempt to overcome these obstacles, researchers have relied on FDA-classified recalls as a proxy for device failures.

A “recall” is an action, typically initiated by a manufacturer, for the “removal or correction of a marketed product that the Food and Drug Administration considers to be in violation of the laws it administers.”¹³⁷ Where the FDA determines that there is a “reasonable probability” that a device failure “will cause serious adverse health consequences or death,”¹³⁸ it may classify the action as a “Class 1” recall.¹³⁹ For a situation in which a product “is unlikely to cause adverse health consequences,” the Agency may classify the action as a “Class 3” recall.¹⁴⁰ “Class 2” recalls form the middle ground, in which device failures “may cause temporary or medically reversible adverse health consequences or where the probability of serious adverse health consequences is remote.”¹⁴¹ Since November 1, 2002, all FDA-classified recalls (including firm-initiated and FDA-ordered) have been recorded in a publicly-accessible, easily

136. See *Hearings*, *supra* note 20, at 23 (testimony of William Maisel) (“[T]he adverse event reports we get are cryptic and don’t contain enough information.”).

137. 21 C.F.R. §§ 7.3, 810.2(k) (2017). Most recalls are initiated by the manufacturer. JOHNSON 2016, *supra* note 28, at 39. A manufacturer must report to the FDA the “[r]eason for the removal or correction and . . . [the manufacturer’s e]valuation of the risk associated.” 21 C.F.R. § 7.46(a). The FDA requires the manufacturer to cease distribution of the device and to notify health professionals and facilities of the order. The manufacturer may request a hearing to be held within ten days of the order. If, after the hearing, the FDA still finds cause, it must amend the original order to include the recall of the device. 21 U.S.C. § 360h(e) (2012). The FDA may also order a manufacturer to initiate a recall. 21 C.F.R. § 810.13; see also *Recalls, Corrections & Removals (Devices)*, U.S. FOOD & DRUG ADMIN., <https://www.fda.gov/medicaldevices/deviceregulationandguidance/postmarketrequirements/recalls corrections and removals/default.htm> (last updated Aug. 25, 2017).

138. *Industry Guidance: Recalls Background and Definitions*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/Safety/Recalls/IndustryGuidance/ucm129337.htm> (last updated Aug. 22, 2014) [hereinafter *Industry Guidance*].

139. 21 C.F.R. § 7.3(m)(1).

140. *Id.* § 7.3(m)(3).

141. *Id.* § 7.3(m)(2).

searchable database.¹⁴² Most investigators have elected to use only Class 1 recalls as a proxy for device failure.

This choice has distinct benefits. Using the FDA's classification removes one level of subjectivity from the analysis because investigators' decisions about which failures are significant are made by the FDA. Other than manufacturers, who might be reluctant to divulge information about device failures, the FDA is in the best position to accumulate and offer access to this data. And restricting the analysis to Class 1 recalls, which the FDA considers to be appropriate for the most serious device failures, minimizes the risk of overestimating the rate of serious device failures.

The choice also carries distinct drawbacks, which are largely mirror images of the benefits. The FDA's assignment of recalls is based on overlapping categories: consequences that are "serious" are not mutually exclusive from those that are "temporary" or "reversible." Thus, the FDA may classify a firm-initiated action as a Class 2 recall for a serious, even life-threatening, failure that also happens to be temporary or reversible.¹⁴³ Relying on the FDA's classification of recalls merely shifts the subjectivity to the FDA. Further, in attempting to define the risk that a premarket assessment failed to detect serious device flaws, some kinds of failures that are properly classified as Class 2 recalls are relevant. Temporary or reversible consequences may expose patients to threats to their lives, the need for corrective surgery, and many other problems.¹⁴⁴ And because failures are underreported,¹⁴⁵ manufacturers and the FDA almost certainly underestimate their true incidence. Ultimately, these drawbacks ensure that any study (including the one I present in Part IV) that relies solely on Class 1 recalls to identify serious device failures will underestimate the frequency of those failures, and hence will overestimate how well the PMA process functions to ensure device safety.¹⁴⁶ Thus, in interpreting the findings of these studies it

142. *Medical Device Recalls*, *supra* note 124.

143. *See infra* notes 241–45 and accompanying text.

144. *See infra* Part V.

145. *See supra* note 135 and accompanying text.

146. *See infra* Part V for a more complete discussion of the ramifications of the choice to use only Class 1 recalls in empirical studies of FDA premarket evaluation effectiveness. *But see Hearings*, *supra* note 20, at 87 (statement of Ralph F. Hall) (arguing that Class 2 recalls are

is important to recognize that they determine only a “best-case scenario” that defines the upper bound for how well FDA premarket assessments function to ensure device safety.

B. Literature Review

Four empirical studies have attempted to determine the frequency with which the FDA has classified firm-initiated or FDA-ordered actions as Class 1 recalls for PMA-approved devices.¹⁴⁷ Three of these studies not only focused on the 510(k) pathway but also examined the PMA pathway for comparison. These studies, by Dr. John Somberg,¹⁴⁸ Professor Ralph Hall,¹⁴⁹ and the medical device industry trade organization AdvaMed,¹⁵⁰ reported low rates of recall, ranging from 0.45 to 0.85%.¹⁵¹

While a failure rate under one percent would seem to indicate that the PMA process functions well to ensure device safety, these studies considered only Class 1 recalls, which, as discussed in section II.A, *supra*, allows only for a determination of the best-case scenario.

“for remote risks or low impact problems” and that even using Class 1 recalls includes many low-probability failures).

147. See Daniel B. Kramer et al., *How Does Medical Device Regulation Perform in the United States and the European Union?: A Systematic Review*, PUB. LIBR. SCI. MED., July 2012, at 1, for a review of the literature on this topic as of 2012. Other studies have reported numbers of failures without reporting the number of devices at risk. See, e.g., Zuckerman et al., *supra* note 9.

148. John C. Somberg et al., *Assessment of Cardiovascular and Noncardiovascular Medical Device Recalls*, 113 AM. J. CARDIOLOGY 1899 (2014).

149. *Hearings*, *supra* note 20, at 81 (statement of Ralph F. Hall). Professor Hall presented the results of a study of a study on FDA premarket device evaluation and subsequent Class 1 recalls to a congressional committee investigating the FDA’s 510(k) process. Hall reported that fewer than 0.5% of devices receiving PMA approval between 2005 and 2009 were recalled during the study period. *Id.*

150. BATTELLE MEMORIAL INST., 510(K) PREMARKET NOTIFICATION EVALUATION (2010), <http://thehill.com/images/stories/blogs/battelle.pdf> (study prepared for AdvaMed).

151. Somberg examined recalls issued over an eight-year period. Somberg’s data included 249 recalls, 246 of which were Class 1. Somberg et al., *supra* note 148, at 1900. Given the very small number of Class 2 recalls, I consider this study to have examined only Class 1 recalls. *Id.* Professor Hall presented the results of a study on FDA premarket device evaluation and subsequent Class 1 recalls to a congressional committee investigating the FDA’s 510(k) process. Hall reported that fewer than 0.5% of devices receiving PMA approval between 2005 and 2009 were recalled during the study period. *Hearings*, *supra* note 20, at 82–83 (statement of Ralph F. Hall) (concluding that over 99.5% of PMAs/PMA approved devices were not recalled). The AdvaMed study reported that the risk of Class 1 recall of a PMA-approved device was 0.85%. BATTELLE MEMORIAL INST., *supra* note 150, at 4 tbl.1.

Beyond this, these studies all employed a methodology that leads to an even greater underestimation of the true risk of recall. The risk of recall is determined by the following ratio:

Devices Recalled During the Observation Period

Devices at Risk of Being Recalled During the Observation Period

The Somberg, Hall, and AdvaMed studies all counted the number of devices at risk (the denominator) in a way that biased the findings toward a lower calculated value. These studies inflated the denominator by considering certain types of PMA Supplements as creating unique devices at risk for recall. This problem is best understood if one were to consider a hypothetical study in which all original and all PMA Supplement approvals, including supplements for trivial changes to the packaging or labeling of devices, were counted as having created unique devices at risk of failure. A device marketed in three different packages or bearing one of three labels with insignificant differences should be counted as one, not three, devices for the purpose of determining the risk that the premarket safety evaluation had failed. But in the hypothetical, the denominator of the risk equation would count such a device as three different devices at risk.

The Somberg and AdvaMed studies excluded trivial modifications, but counted each original PMA approval and each Panel-Track and 180-Day PMA Supplement approval as having created unique devices at risk of failure.¹⁵² Even this inclusion only of more significant modifications, however, leads to an artificially low calculated risk. Considering every modification approved by a Panel

152. Based on the number of devices in the Somberg et al. study, the authors appear to have excluded 30-day supplements, which as above are used for the smallest of changes to device labelling and are not considered by the FDA as being supplemental PMAs. *See generally* Somberg et al., *supra* note 148. It is not clear whether they included Real Time, Changes Being Effected, and 135-Day supplements, but at least some of these may have been included. Somberg almost certainly included 180-Day and Panel Track supplements in the devices at risk total. The AdvaMed study used original PMA approvals, 180-day supplements, and panel-track supplements to calculate the total number of PMA-approved devices at risk. Hall did not specify his method for determining the denominator of the risk calculation. His methodology has been criticized for relying on an indirect estimate of the number of devices at risk, based on historical averages of the number of device approvals. *Hearings, supra* note 20, at 69 (statement of Diana Zuckerman).

Track or 180-Day supplement as a new device is not consistent with the nature of the safety review conducted for PMA Supplement applications. As noted above, for PMA Supplement applications the FDA does not review the safety of the entire device. Rather, the FDA reviews only the changes the manufacturer is proposing to make to an existing device. The specific device that is marketed (assuming the application is approved) has not, by itself, been subjected to a comprehensive safety evaluation. Only the entire line of products extending back through every applicable supplement to the original PMA application has been subjected to a comprehensive safety evaluation. Treating a PMA Supplement approval as an approval to market a new, unique device with its own complete safety evaluation is thus inconsistent with the review process.

Nor is this methodology consistent with Congress's and the FDA's approach. By statute, Congress established that PMA Supplements are to be used when a manufacturer makes "any change to a device subject to an approved application under this subsection that affects safety or effectiveness."¹⁵³ The FDA does not consider a supplemental PMA application as appropriate for a new device: FDA regulations define a PMA Supplement as an "application to an approved PMA for approval of a *change or modification* in a class III medical device."¹⁵⁴ Both Congress and the FDA thus consider a device evaluated through a PMA Supplement pathway as an originally-approved device that has been modified rather than a new device.

A more suitable methodology for studying how well the FDA's premarket evaluation process serves to ensure device safety would consider each new PMA approval and all significant supplements as creating a single device that has been modified over time and that has been subjected to a single, albeit segmented, safety assessment. Dr. Vinay Rathi and colleagues used such a methodology to study a subset of PMA-approved devices.¹⁵⁵ Rathi counted only original PMA approvals as devices at risk in the denominator of the risk equation.¹⁵⁶ Based on a small sample size, they calculated that 3.6%

153. 21 U.S.C. § 360e(d)(6) (2012).

154. 21 C.F.R. § 814.3(g) (2017) (emphasis added); *see also* 21 C.F.R. § 814.39(a).

155. Vinay K. Rathi et al., *supra* note 90, at 607. Rathi excluded non-therapeutic, typically diagnostic, devices from the study. *Id.* at 605.

156. *Id.* at 607.

of devices that were PMA approved in 2010 and 2011 had a Class I recall prior to June 2015.¹⁵⁷ This suggests that the PMA pathway might fail to prevent flawed devices from reaching the market at a much higher rate than the studies by Somberg, Hall, and AdvaMed would indicate.

These studies allow for an even more limited assessment of the factors that might underlie an increased risk of failure. One such factor is device age: devices that have been in use for a long time might be expected to fail more often as components deteriorate and latent problems become manifest. Somberg reported, however, that Class I recalls tended to occur early in the lifetime of a device: the majority of Class I recalls (71%) in that study occurred within the first three years of device approval.¹⁵⁸ However, Somberg did not report how time on the market was calculated, limiting the interpretation of this finding.¹⁵⁹

Another factor that previous studies have examined is device type. Two studies considered whether the FDA review panel to which devices were assigned correlated with the risk of recall. In Somberg's study, Cardiovascular devices did not appear to have a higher risk of a Class I recall (0.45% overall risk for all PMA approved devices and 0.48% for devices approved by the Cardiovascular panel).¹⁶⁰ Hall's study, by contrast, suggested that Cardiovascular and General Hospital devices accounted for most recalls, although Hall did not separate devices approved through the PMA pathway from those cleared through the 510(k) pathway.¹⁶¹ No studies of which I am aware have sought to assess other factors that might contribute to an increased risk that a Class III device will exhibit a significant failure, including the number of times or rate at

157. *Id.* at 605 n.19, 607.

158. Somberg et al., *supra* note 148, at 1902. The investigators did not state whether their time on the market analysis used the interval from the original PMA approval to the recall, or the interval from the most recent supplemental PMA approval to the recall.

159. If the period was calculated as extending from the most recent supplemental approval to the recall, the frequent granting of supplemental PMA approvals is likely to have biased the findings in favor of finding a short interval. On the other hand, if the period was calculated from the original PMA approval, the finding strongly suggests that failures are due to flaws already present at the time the device was manufactured.

160. Somberg et al., *supra* note 148, at 1900–01.

161. *Hearings, supra* note 20, at 93 (statement of Ralph F. Hall).

which a manufacturer has modified its device through the PMA Supplement pathways.

Because the scant empirical literature on the PMA process does not provide clear answers to these and other questions that are essential to understanding the impact of Congress's and the FDA's attempts to balance safety against innovation and access, I performed an empirical study of PMA-approved, Class III devices. I present this study in Part IV.

IV. AN EMPIRICAL STUDY OF PMA-APPROVED MEDICAL DEVICES

A. Study Aims and Hypotheses

The foregoing review of the available empirical literature shows that most questions relating to how well Congress and the FDA have struck the balance between assuring safety and assuring innovation and access remain unanswered. Foremost among these is the rate at which PMA evaluation of high-risk, Class III devices fails to prevent flawed devices that present the threat of death or serious harm from reaching the market. As discussed in Part III, the methodology that most investigators have used biases their findings toward lower-than-actual values by considering each PMA Supplement as having created a new, distinct device. The only study that used an appropriate methodology examined only a very small number of devices for a relatively short time. Therefore, the primary aim of the present study was to determine the risk of Class I recall for as large a number of high-risk (Class III, PMA-approved) devices as could reliably be evaluated.

In addition to determining the failure rate of PMA-approved devices, I sought to evaluate several hypotheses related to the concerns raised in Parts II and III of this Article. As discussed in Part II, the PMA Supplement pathways created by the FDA Modernization Act and the Medical Device User Fee and Modernization Act might compromise safety in a manner analogous to the predicate drift problem identified in devices cleared through the 510(k) pathway. Because the PMA Supplement process allows manufacturers to modify devices without presenting full safety data and clinical data, devices that are marketed after multiple supplements might not have been adequately vetted from a safety perspective. These devices may be more likely to fail, even if the originally approved device was safe. No prior study has examined

whether the iterative process of altering devices through PMA supplements impacts device safety. Therefore, I sought to test the hypothesis that:

H1: An increased number of PMA Supplement approvals is associated with a higher likelihood of Class I recall.¹⁶²

Related to H1, I postulated that even if a manufacturer makes iterative modifications to a device, if those modifications are made at long intervals, the safety impact of each can be evaluated and subsequent PMA Supplements can incorporate additional changes that ensure safety. Conversely, where a manufacturer makes changes in short-order, it is unable to assess the safety ramifications of each. Therefore, I sought to investigate whether:

H2: Shorter intervals between PMA Supplements is associated with a higher risk of Class I recall.¹⁶³

Based on the existing literature on the PMA process, I also sought to determine whether device age impacted the risk of failure. Devices might fail due to mechanical or other time-dependent modes of deterioration—for example, artificial joints might fail after years of use due to ordinary wear and tear. Conversely, devices might fail because of poor design features or manufacturing processes, in which case the failures would occur because of flaws that are already present when the device leaves the manufacturer. These devices might fail shortly after they are marketed. The only study to have addressed this question reported that devices that failed tended to do

162. I pre-specified that comparisons to test hypothesis H1 would compare the risk of recall for devices in the upper quartile of supplements granted with the risk of all other devices. Because some supplement types are used solely for trivial or minor modifications, most investigators have not considered these supplements. *See supra* note 151 and accompanying text. Others have raised the concern that even trivial changes, when made in sufficient numbers, may compromise device safety. *See supra* notes 113–115 and accompanying text. I pre-specified two comparisons that address these issues. The first would compare the total number of supplements granted to an original PMA-approved device with the likelihood of recall, while the second would compare only supplements used for more significant changes (the total number of 180-Day and Panel-Track supplements, *see supra* Part II.C) with the likelihood of recall.

163. I pre-specified that comparisons to test hypothesis H2 would compare the risk of recall for devices with the mean interval between supplements in the lowest quartile (i.e., those with the shortest intervals) with all other devices.

so within a few years of approval.¹⁶⁴ However, that study may have been biased toward finding a shorter interval.¹⁶⁵ Because older devices might fail because of time-dependent component breakdown and because safety practices may have improved over time, I hypothesized that

H3: Longer time on the market correlates with the occurrence of Class I recalls.

The type of device and its indicated uses might predict the risk of failure—devices that are used to treat certain problems might be more prone to fail or more likely to have their failures recognized. Somberg's study reported that device type did not correlate with the risk of recall.¹⁶⁶ In contrast, Hall's study of 510(k)-cleared devices found that Cardiovascular and General Hospital devices had much higher failure rate than other devices.¹⁶⁷ I hypothesized that the panel assignment would predict a Class I recall:

H4: Assignment to the Cardiovascular and General Hospital Panels correlates with the occurrence of Class I recalls.

I also generated several hypotheses de novo. Because implanted devices carry risks that non-implanted devices do not, such as infection and inflammatory reactions, I hypothesized that:

H5: Classification as an Implanted Device correlates with the occurrence of Class I recalls.

I also hypothesized that because failures of devices that are used to sustain lives will manifest more overtly, failures of these devices will be more rapidly identified and will trigger more aggressive FDA action. Thus,

H6: Classification as a Life Sustaining Device correlates with the occurrence of Class I recalls.

Finally, I postulated that a device that is the subject of a Class I recall is more likely to have subsequent Class I recalls. Once a manufacturer has taken the necessary steps to complete a device

164. Somberg et al., *supra* note 148, at 1902.

165. See *supra* note 158 and accompanying text.

166. Somberg et al., *supra* note 148, at 1902.

167. *Hearings, supra* note 20, at 93 (statement of Ralph F. Hall).

recall, the manufacturer may continue to manufacture and sell the device. Thus, simply because a device has been recalled once does not mean that it will not be recalled again. A device recall might identify a technology that is so complex or a design or manufacturing process that is so inherently flawed as to continue to pose an increased risk in spite of a manufacturer's corrective actions. On the other hand, the manufacturer's corrective actions may eliminate the problem. None of the studies discussed here examined the impact of an initial Class 1 recall on the risk of a subsequent Class 1 recall.¹⁶⁸ Thus, I hypothesized that:

H7: Following a Class 1 recall, devices will have a higher risk of a second Class 1 recall compared with the risk that all devices have a first Class 1 recall.

B. Methodology

I used the FDA-maintained, publicly-accessible PMA database as the primary source of device-related data.¹⁶⁹ The PMA database can be searched by individual device or device type, or downloaded *in toto*. The PMA database contains information about each original PMA, along with every PMA Supplement granted by the FDA. The file includes the submission and approval dates, the FDA's assignment of product code,¹⁷⁰ and the review panel for each PMA and PMA Supplement. For supplements, the file also contains the type of supplement (e.g., Panel-Track, 180-Day, Real Time). This database, which has served as a source of information that all previous investigators have used, is considered to contain complete

168. Most studies fail to explicitly discuss how the investigators treated multiple Class 1 recalls issued for a single device. See *Hearings*, supra note 20 (statement of Ralph D. Hall); BATTELLE MEMORIAL INST., supra note 150; Somberg et al., supra note 148. Rathi et al. included only the highest-risk recall in their data, eliminating consideration of repeated device recalls. See Rathi et al., supra note 90, at 605.

169. *Premarket Approval*, supra note 116.

170. The FDA categorizes each PMA device using a three-letter Product Code at the time of approval. This code serves to group or subgroup similar devices. See *Medical Device Classification Product Codes: Guidance for Industry and Food and Drug Administration Staff*, U.S. FOOD & DRUG ADMIN. (Apr. 11, 2013), <http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm285325.pdf>.

and accurate information.¹⁷¹ I downloaded the entire PMA database, containing information on all original and supplemental PMAs granted through September 30, 2016.

I obtained additional information about each PMA and PMA Supplement from the FDA's Product Classification database.¹⁷² This database lists certain information associated with each product code, including the Agency's classification of devices as implanted or non-implanted, and as life-sustaining or non-life-sustaining. I combined the information in this database with that contained in the PMA database.

I obtained data on Class 1 recalls from two sources.¹⁷³ The first was the PMA database itself. When searched by individual PMA numbers, the original PMA web pages contain hyperlinked information on device recalls. The hyperlinks take a user to the FDA Recalls website,¹⁷⁴ which contains information about all recalls for the individual device dating back to November 1, 2002. For each Class 1 recall entry, I extracted the date at which the manufacturer initiated the correction action that led to the recall, the date on which the FDA announced the recall, the manufacturer's explanation for the recall, the FDA's determination of the root cause, and the number of device units affected. I supplemented this data by downloading all Class 1 recall data directly from the FDA Device Recalls website. The device recall website contains information about all device recalls since November 1, 2002.¹⁷⁵ I combined the recall

171. *Hearings, supra* note 20, at 86 (statement of Ralph F. Hall); BATTELLE MEMORIAL INST., *supra* note 150, at 12; Rathi et al., *supra* note 90, at 605; Somberg et al., *supra* note 148, at 1900.

172. *Product Classification*, U.S. FOOD & DRUG ADMIN., <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpd/classification.cfm> (last visited Jan. 8, 2018).

173. A word of explanation about my decision to rely on Class 1 recalls as a proxy for device failure is in order. Two reasons motivated this decision. First, using Class 1 recalls is consistent with nearly all prior studies of the safety of 510(k) cleared and PMA approved devices. By using Class 1 recalls, the findings of my study can be compared to the prior studies to ensure that my database is similar to those assembled by previous investigators. Second, although using Class 1 recalls will be underinclusive, using all Class 2 recalls would be drastically overinclusive because most Class 2 recalls will have been ordered for less serious problems.

174. *Medical Device Results, supra* note 124.

175. *See id.*

information from the Device Recalls database with the data from the PMA database to create a master study database.¹⁷⁶

Throughout the history of the MDA, Congress has required and the FDA has been engaged in a process of re-evaluating device classification.¹⁷⁷ Many devices that were first marketed after the enactment of the MDA were presumptively classified as Class III.¹⁷⁸ The FDA has reclassified some of these devices to Class I or Class II.¹⁷⁹ Because reclassification of a Class III device to Class I or II reflects the FDA's determination that these devices were not high risk, I excluded those devices from the data set. I maintained a separate data set for these devices, which was not used in the present study. I also excluded devices approved after November 1, 2015, to ensure that all devices had at least one year of observation. The study period was from November 1, 2002, to November 1, 2016.¹⁸⁰

The "Primary Data Set" used in this study consisted of all PMA-approved Class III devices originally approved on or after November 1, 2002. This date was selected because the data in the FDA's Device Recalls database begins then. Devices approved prior to November 1, 2002, might have had Class I recalls that would not be included in any database.¹⁸¹ Further, devices approved decades ago

176. Using the PMA database as the primary source of information about device recalls allowed me to circumvent one commonly cited difficulty confronted by several other investigators who instead used the Recall database as their primary source: The data contained in the FDA Recall database frequently does not contain the approval history for the recalled device. Because device names may change and several related technologies may be approved under different PMAs or even under a 510(k), an investigator starting from the Recalls database is forced to assign recalls to a given PMA. See BATTELLE MEMORIAL INST., *supra* note 150, at 3; Somberg et al., *supra* note 148, at 1900 (using data from FDA's medical device recalls website); Zuckerman et al., *supra* note 9, at 1007. These assignments often leave a residual uncertainty as to whether the correct approval type (510(k) or PMA) and device have been associated with the recall. See BATTELLE MEMORIAL INST., *supra* note 150, at 12–13. By using the PMA database as the primary source, I ensured that no 510(k) cleared devices were included in the study.

177. See JOHNSON 2016, *supra* note 28, at 6–7.

178. *Reclassification*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHTransparency/ucm378724.htm> (last updated Dec. 19, 2017).

179. *Id.*

180. The dates were selected to ensure at least one year of observation for all devices. Thus, supplements granted between November 1, 2015, and November 1, 2016, for devices granted an original PMA before November 1, 2015, were included in the data sets.

181. For example, the St. Jude Silzone valve (P810002) was recalled in 2000, but is not included in the FDS recalls database. See *In re St. Jude Silzone Heart Valves Prod. Liab. Litig.*,

might not have been on the market and at risk for recall during the study period. Including these devices in the denominator of the risk calculation would bias the results toward a lower-than-actual risk value. The Primary Data Set contained 389 Original PMA approved devices and 9217 PMA Supplements.

To examine the performance of FDA premarket evaluation under recent statutory and regulatory frameworks, I created a subset of the Primary Data Set, the “Restricted Data Set.” This consisted of devices approved at least one year after the significant statutory overhaul made by the FDA Amendments Act in 2007.¹⁸² The Restricted Data Set contained 199 original and 3666 PMA Supplements.

I also created an “Expanded Data Set,” consisting of all PMA-approved Class III devices in the FDA database, including devices approved before November 1, 2002. Again, devices that had been reclassified to Class I or Class II were excluded. Further, to avoid the inclusion of devices that had been approved long before but were no longer marketed by the beginning of the study period, I excluded devices originally approved prior to November 1, 2002, for which no supplemental approvals were granted during the study period. Thus, each device in the Expanded Dataset had at least one approval (an original or a supplemental PMA) during the study period. The Expanded Data Set contained 755 Original PMA approvals and 26,674 PMA Supplements.

For the Primary Data Set I calculated the risk of Class 1 recall as:

$$\frac{\text{Number of devices approved between Nov. 1, 2002,} \\ \text{And Nov. 1, 2015, that had a recall during the study period}}{\text{Number of devices approved between} \\ \text{Nov. 1, 2002, and Nov. 1, 2015}}$$

An analogous calculation was performed for the Restricted Data Set. For the Expanded Data Set, I calculated the risk as:

No. MDL 01-1396 JRTFLN, 2004 WL 45503, at *2 (D. Minn. Jan. 5, 2004) (“In response to the voluntary recall, the FDA informed St. Jude that its actions would be considered a ‘recall.’”).

182. See *supra* note 56.

Number of devices that had a recall during the study period
Number of devices with an original or supplemental
PMA approved during the study period

Devices that were subjected to more than one Class 1 recall during the study period were counted only once, with all data entries after the first Class 1 recall being censored.

Statistical analyses were performed using Stata. Unless otherwise noted, comparisons of continuous variables were performed using two-tailed t-tests with unequal variances, and comparisons of categorical variables were performed using Pearson's chi-squared tests. A *p* value of .05 was considered to be statistically significant. Multiple logistic regression analysis was performed to determine the relationship between recalls and the variables found to independently predict recall.

C. Results

1. Descriptive findings

Manufacturers gained approval for a wide range of devices through the PMA pathway. As of September 30, 2016, the entire FDA PMA Approval database contained 1407 devices that were granted an original PMA approval. In the Primary Data Set, 389 devices had received an original PMA. Seventeen different review panels evaluated the PMA applications for the devices in the Primary Data Set. The number of original applications approved by each panel varied from 2 (General Hospital, Physical Medicine) to 135 (Cardiovascular). Six panels reviewed 25 or more devices (Cardiovascular, Microbiology, Ophthalmic, Orthopedic, Pathology, General and Plastic Surgery). The Cardiovascular panel accounted for the plurality of devices reviewed.¹⁸³

Confirming the findings of prior studies,¹⁸⁴ manufacturers have made extensive use of the PMA Supplement pathways. The entire FDA PMA database contained 31,879 approved PMA

183. The cardiovascular panel reviewed 34.7% of devices in the primary data set.

184. *See supra* notes 113–15 and accompanying text.

Supplements.¹⁸⁵ The number of supplements granted by each panel varied widely, from 8 (Physical Medicine) to 5279 (Cardiovascular) in the Primary Data Set. In the Primary Data Set, the median number of supplements (including those used for minor modifications, such as 30-Day Notice and 135-Day supplements) was 12 per original PMA approval, ranging from 0 to 291. Considering only Panel Track and 180-Day Supplements, which are used for more significant modifications, the median number of supplements was 3 (range 0 to 49).

In the Primary Data Set, 4.6% of devices at risk were subjected to a Class 1 recall (18 out of 389 devices) during the study period. Of the 18 devices that were the subjects of Class 1 recalls, 16 had one recall, one device had 2, and one device had 6 recalls.¹⁸⁶ Limiting the analysis to devices approved after implementation of FDAAA, the devices in the Restricted Data Set had a rate of recall of 5.0% (10 out of 199).

The 24 Class 1 recalls of devices in the Primary Data Set involved a wide variety of products.¹⁸⁷ The most frequently recalled devices were stents (5 recalls) and artificial hearts/ventricular assist devices (3 recalls). Other recalled devices included heart valves (2), angioplasty catheters (2), intraocular lenses (1), invasive blood glucose monitors (1), infant cooling caps (1), anesthesia administration systems (1), embolization coils (1), and an injectable bulking agent (1).¹⁸⁸ These recalls covered over 660,000 individual device units.¹⁸⁹

Of the 24 total Class 1 recalls involving devices in the Primary Data Set, the FDA had not established the root cause for 3 recalls as of June 2017.¹⁹⁰ Of the 21 recalls for which the FDA had determined a cause, 7 (33.3%) were attributed to design defects, 3 (14.3%) were attributed to component design/selection, and 2 (9.5%) were attributed to process design.¹⁹¹ Thus, over half (57.1%)

185. For simplicity, I include 30-Day Notices in the PMA Supplement total, unless otherwise specified.

186. See App. 1 for details.

187. *Id.*

188. *Id.*

189. *Id.*

190. *Id.*

191. *Id.*

of the Class I recalls were attributed to design problems that might have been detected with a more robust premarket evaluation.¹⁹² These recalls involved nearly 400,000 individual units.¹⁹³ A smaller number of recalls were due to problems that likely would not have been detected by a more robust premarket evaluation, including one recall attributed to “use error” and two attributed to a “nonconforming material or component.”¹⁹⁴

By using the Expanded Data Set, I assessed the performance of PMA review for as large a set of marketed devices as possible. In the Expanded Data Set, 6.0% of devices were subjected to a Class I recall (45 out of 755 devices) during the study period. The 45 recalled devices had a total of 66 recalls. Eight devices had more than one Class I recall (range 2 to 11). Since nearly half of the devices in the Expanded Data Set were marketed before the FDA’s Recall Database began collecting data, it is likely that the actual risk of recall for all Class III devices on the market during the study period was higher than 6.0%.

2. Hypothesis testing

H1: An increased number of PMA Supplements is associated with a higher likelihood of Class I recall.

Considering all modifications, including minor changes made through 30-Day Notifications and 135-Day and Real-Time supplements, the number of supplements did not correlate with the likelihood of an initial Class I recall. In the Primary Data Set, devices that were recalled had an average of 32.4 ± 26.8 supplements, while those not recalled had an average of 23.3 ± 33.2 supplements ($p=.25$). Although devices in the upper quartile of number of total supplements were more likely to be recalled than other devices, 7.4% versus 3.7%, the difference was not statistically significant ($p=.15$).

However, when only the most significant modifications (those made through Panel-Track and 180-Day Supplements) are considered, the number of supplements does correlate with the likelihood of a first Class I recall. Devices that were recalled had

192. *Id.*

193. *Id.*

194. *Id.*

received a median of 5.5 supplements (range 0 to 15), compared with two supplements (range 0 to 49) for non-recalled devices.¹⁹⁵ Devices in the upper quartile of the number of Panel-Track and 180-Day Supplements (seven or more) were significantly more likely to be recalled than other devices: 9.2% versus 3.3% for devices with 6 or fewer supplements ($p=.027$). Thus, a large number of significant modifications made to a device through PMA Supplements appears to correlate with the risk of recall.

H2: More rapid supplementation is associated with a higher risk of Class 1 recall.

For each device, I calculated the “supplementation interval” for all supplements granted as

$$\frac{\text{Days from Original Approval to Nov. 1, 2016}}{\sum(\text{All Supplements Granted}) + 1}$$

and the supplementation interval for only 180-Day and Panel Track supplements as

$$\frac{\text{Days from Approval to Nov. 1, 2016}}{\sum(\text{180-Day and Panel Track Supplements}) + 1}$$

Including all supplements, the median supplementation interval for recalled devices was 75 days (range 18.5 to 988.4), and for non-recalled devices was 153.3 days (range 9.4 to 5051). Limiting the analysis to the most significant modifications (those made through 180-Day and Panel-Track Supplements), recalled devices had a median supplementation interval of 293.2 (range 103 to 1829) days, while non-recalled devices had a median interval of 596 (range 79.2 to 5073) days.¹⁹⁶ Devices whose supplementation interval for 180-Day and Panel-Track supplements was in the lowest quartile (335 days or less) were more likely to be recalled, 9.4% versus 3.1% ($p=.01$).¹⁹⁷

195. Because the distribution was strongly non-normal, t-testing is not appropriate.

196. Because the distribution was strongly non-normal, t-testing is not appropriate.

197. For administrative convenience, an interval based on a cut-off of one year might be more useful. Devices with a 180-Day and Panel Track supplementation interval of one year (365 days, $n=116$) or less had a recall risk of 8.7%, while devices with a longer interval ($n=273$)

If the shorter interval for recalled devices resulted from manufacturers' use of post-recall PMA Supplements to correct problems, the intervals should be shorter after a recall than before. To assess this, I examined the supplementation intervals for recalled devices before and after the first Class I recall. The intervals for 180-Day and Panel-Track supplements was not significantly different (266 ± 236 versus 247 ± 145 days, $p = .79$) in these two periods.¹⁹⁸ Thus, the supplementation intervals for these devices were essentially the same before and after their first (or only) Class I recall.

H3: Longer time on the market correlates with the occurrence of Class I recalls.

Dividing the devices into three subgroups based on their approval dates (before November 1, 2002; November 1, 2002, to September 27, 2008; and after September 27, 2008), the risk of recall was 7.4%, 4.2%, and 5.0%, respectively. The differences were not statistically significant ($p = .25$ by Pearson's chi-square test). Thus, the study does not suggest that an earlier date of original PMA approval correlates with the risk of Class I recall. However, the hypothesis as to devices approved prior to November 1, 2002, cannot be rejected, because the FDA database does not include recalls prior to this date. Thus, whether devices sold under earlier-granted PMAs are more likely to be recalled remains an open question.

In the Primary Data Set, recalled devices had been on the market 1198 ± 1070 days until the first recall. The shortest interval from approval to first Class I recall was 119 days, and the longest was over 10.5 years. Fifteen of the eighteen recalled devices were recalled less than five years after original PMA approval. Thus, in the Primary Data Set, the suggestion that device failures tend to occur within a

had a recall risk of 2.9%. By Pearson's chi-squared analysis, this difference is significant, with a p value of .013.

198. Of the eighteen recalled devices, sixteen remained on the market and were granted subsequent supplemental PMAs. The analysis was limited these devices. The pre-recall supplementation interval was calculated as "Time from Approval to Recall" divided by the "Total Number of 180-Day and Panel Track Supplements + 1." The post recall interval was calculated as "Time from Recall to Final Supplement" divided by the "Total Number of 180-Day and Panel Track Supplements + 1." Comparisons were performed using two-tailed t-tests with equal variance.

relatively short time after original PMA approval finds at least some support.

H4: Assignment to the Cardiovascular and General Hospital Panels correlates with the occurrence of Class 1 recalls.

The Primary Data Set included only two devices approved by the General Hospital panel. Because of this small number, I could not test the hypothesis that these devices have a higher risk of Class 1 recall.

When all devices are included in the analysis, the advisory panel that reviewed the original PMA application did not correlate with the risk of recall (chi-squared $p=.073$). However, considering only the six panels that reviewed twenty-five or more devices, Cardiovascular devices were more likely to be recalled (8.9% versus 0.6%) than devices approved by the other panels (Gastroenterology/Urology, Microbiology, Ophthalmic, Orthopedic, Pathology, General/Plastic Surgery; $p=.026$).¹⁹⁹

Certain attributes of Cardiovascular devices might account for the higher frequency of Class 1 recall. Cardiovascular devices are more often classified by the FDA as life-sustaining (76%, versus 37% for non-Cardiovascular devices, $p<.01$) and are more often implanted than other devices (54% versus 3%, $p<.01$). However, neither of these classifications correlated with the occurrence of recalls of Cardiovascular devices.²⁰⁰

Another possibility is that the FDA applies a lower threshold to label corrective actions taken by manufacturers of cardiovascular devices as Class 1 recalls. If so, one would expect that a larger fraction of Class 1 recalls of Cardiovascular devices would be for less serious problems. Using the information available in the FDA Recall Database and other FDA communications about the eighteen recalls in the Primary Data Set, and reports available in the medical literature at the time of the recall, this does not appear to be the case. The problem for which the FDA ordered a Class 1 recall had

199. The risk of recall of all non-cardiac devices, including those reviewed by panels that approved fewer than 25 original PMAs, was 2.6%.

200. Cardiovascular devices classified as implanted had a 9.8% risk of recall, while non-implanted Cardiovascular devices had a 6.0% risk ($p=.51$). Cardiovascular devices classified as life sustaining had a 6.8% risk of recall, while non-life sustaining Cardiovascular devices had an 11.3% risk ($p=.37$).

manifested in clinical practice (as opposed to manifesting only in bench testing) with nearly equal frequency: The Class 1 recalls for seven of the twelve recalled Cardiovascular devices and for three of the six recalled non-Cardiovascular devices were for a problem that had manifested in clinical practice ($p=.74$). Reports of serious injuries that occurred prior to the Class 1 recall could be located in 50% of cardiovascular device recalls and 33% of non-cardiovascular device recalls ($p=.74$). Thus, it does not appear that the FDA applies a lower threshold to classify a corrective action for Cardiovascular devices as a Class 1 recall.²⁰¹

H5: Classification as an Implanted Device correlates with the occurrence of Class 1 recalls.

H6: Devices classified as Life Sustaining will be more likely to have a Class 1 recall.

Neither of these classifications correlated with the risk of recall. Implanted devices had a recall risk of 5.6% (11 out of 195), while non-implanted devices had a risk of 3.6% (7 out of 193, $p=.63$). Life-sustaining devices had a risk of recall of 6.3% (5 out of 80), and non-life sustaining devices had a risk of 4.2% (13 out of 309, $p=.44$). Stratifying the devices by cardiovascular/non-cardiovascular type, neither implant nor life-sustaining classification correlated with the risk of recall.

H7: Following a Class 1 recall, devices will have a higher risk of a second Class 1 recall compared with the risk that all devices have a first Class 1 recall.

In the Primary Data Set, once a device had a Class 1 recall, the risk of a subsequent Class 1 recall was 11.1% (2 out of 18). Treating the eighteen devices that had an initial Class 1 recall as a separate population and second recalls as independent events, this frequency was not significantly different from the risk that any device had an initial Class 1 recall ($p=.21$). In the Expanded Data Set, eight devices with a Class 1 recall (out of forty-five) had a subsequent Class 1 recall,²⁰² yielding a risk of a subsequent recall of 17.8% (8 out of 45).

201. See *infra*, Part V, for a discussion of other possible reasons.

202. One device had a subsequent Class 1 recall (after an initial Class 1 recall) that occurred as a result of the manufacturer's attempt to correct the initially-identified problem. I included this recall since it was not merely an expansion of an earlier recall to other lots.

This was significantly higher than the 6.0% risk of an initial Class 1 recall in the full data set ($p=.048$). The study findings are thus equivocal as to whether hypothesis *H7* should be rejected.

* * *

To determine whether device type (Cardiovascular versus non-Cardiovascular), number of 180-Day and Panel-Track supplements, and supplementation interval were independent predictors of recall, I performed a multiple logistic regression analysis. I assigned a dummy variable of “1” for Cardiovascular devices, devices with that number of supplements in the upper quartile, and devices with the supplementation interval in the shortest quartile, and a dummy variable of “0” for non-Cardiovascular devices, devices with the number of 180-Day and Panel Track supplements not in the upper quartile, and devices with a supplementation interval not in the shortest quartile. Cardiovascular device type (coefficient 1.18, $p=.026$) remained as an independent predictor of recall.

Stratifying by device type, Cardiovascular devices had been granted a median of four (range 0 to 49) Panel-Track and 180-Day supplements. Devices with four or fewer supplements had a recall rate of 2.8%, while devices with five or more supplements had a rate of 15.6% ($p=.009$). Cardiovascular devices had a median supplementation interval of 477.6 days. Devices with intervals shorter than this had a recall rate of 17.7%, while devices with longer intervals had a recall rate of 3.0% ($p=.017$).

For non-Cardiovascular devices, the median number of 180-Day and Panel Track supplements was two (range 0 to 37). Although devices at or below the median were recalled less frequently than those above the median, 0.7% versus 4.3%, this difference was not statistically significant ($p=.061$). The recall rate for devices with supplementation intervals at or shorter than the median was not significantly different than for devices with longer intervals, 3.1% versus 1.6% ($p=.41$).

V. A DATA-DRIVEN ANALYSIS OF THE PMA FRAMEWORK

The PMA approval process is designed to provide a rigorous assessment of high-risk medical device safety. In spite of this, the PMA process must accommodate other policy goals. Chief among these is that premarket evaluation neither impose such high costs as

to impede the development of new, potentially life-saving technologies nor impose such high regulatory burdens as to delay the availability of these new technologies to millions of Americans.

In order to situate the findings of the empirical study presented here, it is necessary to recall the ways in which the rigors of premarket evaluation have been relaxed in an attempt to balance these competing objectives. Congress requires the FDA to adopt the least burdensome approach to regulating each Class III device.²⁰³ The FDA has interpreted this as permitting it to make a determination of safety and efficacy based on a single study, and on studies that used surrogate endpoints, limited durations, and other protocol elements that do not meet strict scientific standards.²⁰⁴ Further, manufacturers can modify devices through the PMA Supplement pathways, which impose even less rigorous data requirements. And by iteratively modifying devices through PMA supplements, an equivalent to predicate drift may occur, resulting in devices that bear little resemblance to their progenitors.

The resulting question is whether Congress and the FDA have struck a desirable balance between ensuring the safety of, and not hampering the innovation of and access to, new technologies. In brief, the concerns raised in Parts II and III, and the study presented in Part IV of this Article suggest that the answer is “No.” The first finding of the study is that the chance that a PMA-approved, Class III medical device will be recalled is higher than most prior studies have suggested.²⁰⁵ The FDA classified manufacturers’ actions as Class I recalls for 4.6% of devices approved after November 1, 2002. Considering all Class III devices, including those approved before

203. 21 U.S.C. § 360c(a)(3)(D)(ii) (2012).

204. 21 C.F.R. § 860.7 (2017); LEAST BURDENSOME PROVISIONS, *supra* note 15.

205. Most prior studies have reported an incidence of Class I recalls for high-risk, PMA approved devices of 0.45% to 0.85%. These studies utilized a methodology that considered every modification made through a PMA supplement to be a new, distinct device. *See supra* Section II.B. My study is consistent with these findings: re-analyzing my data using the total number of original PMA, 180-Day, and Panel Track approvals as the denominator for the risk calculation, the devices in my Primary Data Set had risk of recall of 0.8%. In the only prior study to use a similar methodology to the one employed here, Rathi and colleagues found a 3.6% chance of recall. However, Rathi’s sample size was small and the study duration was limited. *See supra* notes 155–57. Not surprisingly, since my study followed devices for a much longer period of time, I found a higher risk of recall. However, my findings are also consistent with those of Rathi: Including only recalls issued within five years of approval (close to the follow-up duration of Rathi’s study), the risk of recall is 3.9% in my data.

the FDA began collating data on recalls, my study finds that the risk of recall is 6.0%. This underestimates the true risk of recall, since the older devices may have been recalled prior to November 1, 2002, when the data in the FDA's publicly-accessible database on recalls begins. And as discussed in Part VI, *infra*, several factors ensure that the true risk of failure is even higher.

These recalls reflect device failures that the FDA judged to pose a "reasonable probability" of "serious adverse health consequences or death."²⁰⁶ In the Class III devices approved after November 1, 2002, these recalls affected more than 660,000 individually marketed units.²⁰⁷ Thus, the current study suggests that weaknesses in the premarket evaluation of medical devices pervade the PMA process as well as the 510(k) process, leading to serious failures that impact large numbers of people.

Professor Hall has implied that using Class I recalls as a proxy for device failures that compromise safety tends to overestimate the risks presented by device failures. In a statement to a Senate subcommittee investigating the performance of the 510(k) pathway, Hall argued that Class I recalls "represent all recalls with any meaningful risk to patients."²⁰⁸ Continuing, he noted that devices failing at rates "as low as 1/20,000 have been classified as Class [1] recalls."²⁰⁹ However, anecdotal accounts of the human costs of these failures are easy to find,²¹⁰ and the scholarly literature indicates that deaths and injuries related to medical devices are frequent. Artificial, metal-on-metal hip joints have been implanted in over 500,000 people in the United States, many of whom have required a re-operation to replace the prosthesis because of component breakdowns that caused severe pain and decreased mobility.²¹¹ Data from 2006 indicate that over 100,000 people suffered a device-

206. *Industry Guidance*, *supra* note 138.

207. *See* App. 1.

208. *Hearings*, *supra* note 20, at 87 (statement of Ralph F. Hall).

209. *Id.*

210. *See* Stengel v. Medtronic, Inc., 704 F.3d 1224, 1227 (9th Cir. 2013) (partial paralysis due to an infusion pump); Barry Meier, *Maker of Heart Device Kept Flaw from Doctors*, N.Y. TIMES (May 24, 2005), http://www.nytimes.com/2005/05/24/business/maker-of-heart-device-kept-flaw-from-doctors.html?_r=0 (sudden death due to malfunction of an implanted defibrillator).

211. *Hearings*, *supra* note 20, at 67–68 (statement of Diana Zuckerman); JOHNSON 2016, *supra* note 28, at 2.

related injury.²¹² And independent analysts have estimated that in 2009 alone, over 4,500 device-related deaths occurred.²¹³ Likewise, examples of the economic costs imposed by these failures are easy to find. The failure of a single medical device, the Medtronic Sprint Fidelis defibrillator lead, has the potential to cost Medicare between \$287 million and \$1.2 billion.²¹⁴

The study findings suggest that Congress's least burdensome approach directive and the FDA's implementation of that directive are compromising device safety. Unfortunately, several provisions of the recently-enacted 21st Century Cures Act have the potential to tip the balance further in an unfavorable direction. Section 3058 of the Act emphasizes the application of the least burdensome approach.²¹⁵ Manufacturers already may gain market approval on the basis of a single pivotal study that is not blinded, does not have a comparator group, uses surrogate endpoints, contains significant methodological flaws and inconsistencies, and is of short duration. The Cures Act will push the FDA even more strongly toward requiring less rigorous safety data.

Some of these compromises are unavoidable. Blinding may be impossible for implanted devices (or at least raise prohibitive ethical restraints).²¹⁶ But other protocol weaknesses could be addressed by the FDA, even within the constraints of the least burdensome approach: If the failure rate of PMA-approved devices is too high, then the information currently being required does not provide the statutorily required "reasonable assurance of . . . safety."²¹⁷ The FDA could require longer studies and study designs that examine actual, not surrogate, outcomes. And the FDA could reject studies that contain flaws such as discrepancies in subject numbers and the reliance on post-hoc analysis to establish safety and efficacy. These

212. JOHNSON 2012, *supra* note 36, at 2–3.

213. *Id.* at 3.

214. Mehrotra, *supra* note 38, at 1195.

215. 21st Century Cures Act, Pub. L. No. 114–255, § 3058(a), 130 Stat. 1033, 1128 (2016) (codified as amended at 21 U.S.C. § 360c). The Act defines "least burdensome" to mean the FDA could require only "the minimum required information that would support a determination by the Secretary that an application provides a reasonable assurance of the safety and effectiveness of the device." *Id.* § 3058(b), 130 Stat. at 1129.

216. Kramer et al., *supra* note 84, at 3.

217. 21st Century Cures Act § 3058(a), 130 Stat. at 1128–29.

changes should be possible at the administrative level, without congressional action. However, the Cures Act, by emphasizing the least burdensome approach, may limit the FDA's flexibility.

The Cures Act contains another provision that may tip the balance between safety and innovation/access unfavorably: for breakthrough technologies, the FDA is permitted to reduce the pre-market burden on manufacturers to produce clinical trial data "through utilization of timely postmarket data collection."²¹⁸ And for other Class III medical devices, this shift to reliance on post-market data is strongly encouraged: "[T]he Secretary shall consider the role of postmarket information in determining the least burdensome means of demonstrating a reasonable assurance of device safety and effectiveness."²¹⁹ These directives push the FDA toward granting PMA applications based on smaller and shorter-duration premarket studies.

Post-market studies and surveillance will function under section 3058 as early warning systems, detecting the first signs of a pattern of failures which will trigger a corrective action by the manufacturer and the FDA. However, the ability of such a system to minimize harms caused by failing devices is limited. New devices are now rapidly incorporated into clinical practice—within a few years, hundreds of thousands of patients may be exposed to a newly-approved device.²²⁰ Unfortunately, device failures may take longer than this to manifest and trigger recognition.²²¹ Thus hundreds of thousands of units of an artificial hip or a defibrillator lead that break down prematurely after only two or three years may be implanted into patients before a pattern of premature failure is recognized.

This problem arises from two sources. First, short-duration premarket studies of devices intended for long-term use—especially for implanted devices—cannot detect failures that manifest only after a period of time longer than the studies themselves. Second, health care providers rapidly adopt new technologies without fully understanding the limitations of the data that supported FDA approval.

218. *Id.* § 3051, 130 Stat. at 1121–23.

219. *Id.* § 3058(b), 130 Stat. at 1129.

220. *See supra* notes 130–31 and accompanying text.

221. *See supra* notes 130–31 and accompanying text.

These temporal patterns, in turn, suggest two possible solutions. The first would address the problem of short-duration studies. Congress could amend the Cures Act, eliminating the shift in data production to the post-market period required by section 3058. In fact, Congress could require that reliance on post-market data be a last, not a first, resort. This would incentivize the FDA to require the results of longer-term pre-market studies before declaring a device safe. Or Congress could permit the FDA to use its expertise to determine the premarket/post-market data requirement for each device, free from the push toward post-market data provided by the Cures Act.

The second possible solution would address the rate at which patients are exposed to newly-approved devices. The FDA has limited authority to impose restrictions on the sale and distribution of devices at the time of PMA approval.²²² This authority is triggered when “there cannot otherwise be reasonable assurance of [a device’s] safety and effectiveness.”²²³ An FDA determination that a shift in data production to the post-market period satisfied this provision could provide the agency with the authority to impose restrictions on how many device units may be sold until adequate long-term safety data has been collected.

This suggestion raises the same balancing issues that the present study examines: If the FDA restricts the number of devices allowed onto the market, the collection of long-term, “real world” safety data will be delayed. Further, “real world” experience with device efficacy would be reduced, potentially delaying the implementation of design and labeling improvements. Thus, the FDA would need to carefully consider, on a case-by-case basis, whether and to what extent it should impose restrictions on the number of newly-approved devices that may be sold.

The study presented in Part IV also provides support for other, targeted alterations of the PMA process. Based on the finding that devices approved by the Cardiovascular panel are significantly more likely than other devices to be subjected to a Class 1 recall, it would be prudent for the FDA to require a more robust set of safety data for these devices. This could include requiring more than one pivotal

222. See 21 U.S.C. § 360e(d)(1)(B)(ii) (2012).

223. *Id.* § 360j(e)(1)(B).

study, longer studies, and strict adherence to standard study protocols. This should be within the FDA's existing authority.

Looking more deeply, it is useful to consider why Cardiovascular devices have a higher risk of recall. One possibility is that these devices are more often used for life-sustaining functions, and thus that their failures are more rapidly and routinely detected. But the study findings refute this: devices that the FDA classifies as life sustaining were not significantly more likely to be recalled than other devices.²²⁴ Another possibility is that because many of these devices, including artificial valves, stents, leads, pacemakers, and defibrillators, are implanted, they carry higher risks of serious complications. Again, the study findings refute this: implanted devices were no more likely to be recalled than other devices.²²⁵

Cardiovascular devices have other characteristics that likely underlie the higher failure rate. These devices are among the most complex on the market, often featuring multicomponent composition, miniaturization, the incorporation of computer hardware and software, and a rapid rate at which technology is developing.²²⁶ These devices have very high “connectivity complexity,” which measures “the number of relations/interconnections between the components of a given system.”²²⁷ These devices also have high “user complexity,” placing substantial cognitive demands on health care providers and device users. These and other complexity factors might serve as triggers for heightened FDA scrutiny. The FDA currently uses a scoring algorithm to categorize clinical laboratory testing devices based on their complexity.²²⁸ As required by the Clinical Laboratory Improvement Amendments of 1988,²²⁹ device complexity is scored based on several factors specific to laboratory devices.²³⁰ The agency could

224. *See supra* Part III.

225. *See supra* Part III.

226. *See supra* notes 106–09 and accompanying text.

227. Senders, *supra* note 106, at i42 app. 1.

228. *Id.*

229. Clinical Laboratory Improvement Amendments of 1988, Pub. L. No. 100-578, § 2, 102 Stat. 2903, 2903 (codified at 42 U.S.C. § 263a (1997)).

230. *CLIA Categorizations*, U.S. FOOD & DRUG ADMIN., <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/IVDRegulatoryAssistance/ucm393229.htm> (last updated Apr. 16, 2014) (scoring complexity based on “the scientific and technical knowledge . . . required to perform the test,” the “[t]raining and experience” required, the ease with which

develop a broadly applicable algorithm that could determine when a device category (defined by product code) has become sufficiently complex as to trigger heightened premarket study data requirements. As devices bearing a particular product code incorporate computer technology and multiple components, undergo miniaturization, and place higher cognitive demands on providers and users, the complexity score of that product code would increase. Eventually, the aggregate score would trigger more rigorous premarket scrutiny for all devices bearing that product code.

Based on my findings that Cardiovascular devices with more 180-Day and Panel-Track supplements, and with shorter intervals between those supplements, are more likely to be subjects of Class I recalls, modifications to the PMA Supplement pathways should also be considered. These findings are consistent with the possibility that rapid, iterative changes to the most complex devices through the PMA Supplement pathways allows devices to reach the market without a thorough evaluation for safety. One way to mitigate this risk would be to impose a limit on how many 180-Day and Panel-Track supplements may be granted, after which a manufacturer would be required to either submit an original PMA application with complete information or a more robust set of clinical data in a supplemental application.²³¹ My study findings suggest that after a Cardiovascular device manufacturer has made four significant changes to a device, a fifth supplemental application should be converted to an original PMA or should be required to contain data demonstrating the overall safety of the device. Alternatively, the increased risk might be mitigated by requiring a specified period of clinical use following any significant modification before another Panel-Track or 180-Day supplement will be granted. This would

the reagents may be used, the degree of automation, the stability and reliability of the calibration process, the ease of troubleshooting, and the amount of interpretation and judgement that must be used).

231. Ben Rome and colleagues suggested that an expert panel review PMA-approved devices every five to seven years to determine whether additional clinical data should be required for supplemental PMA applications. Rome et al., *supra* note 78, at 390. However, in my study fifteen of eighteen recalled devices were recalled within five years of approval. One possible response would be to require expert review even earlier than five years. However, since most devices are not recalled, this would impose a large burden in order to identify a relatively small number of risks. Using the number of or interval between 180-Day and Panel-Track PMA supplements permits a more targeted intervention.

provide an opportunity for manufacturers, prescribers, and the FDA to recognize safety issues that the first modification created; this recognition might guide future changes to the device, allowing better assurance of safety. Alternatively, applications for 180-Day or Panel-Track supplements within a short time of an earlier supplement might trigger closer scrutiny by the FDA, with an increased quantum or quality of data required in this circumstance. These requirements should currently apply to Cardiovascular devices. When other devices become sufficiently complex, these requirements should apply to them as well.

It is important to recognize that placing new limitations on device modifications made through PMA Supplement pathways would not deprive patients of access to breakthrough technologies. Devices based on novel technologies that offer treatments for currently untreatable problems would be evaluated through a new, original PMA application. The limits on PMA Supplement approvals that I am suggesting would certainly delay access to modified devices but would have no direct impact on the availability of truly new devices.

Other findings provide reassurance about the PMA Supplement pathways. For non-Cardiovascular devices, neither the absolute number of 180-Day and Panel-Track supplements nor the rate at which the FDA has granted them appears to predict the occurrence of a Class I recall. For devices with low complexity, the supplemental PMA pathways appear to offer a relatively inexpensive and timely route for making incremental, evolutionary changes to medical devices. The limited safety evaluation required under these pathways appears sufficient to ensure device safety.

Some commentators have raised concerns that even minor changes, when made iteratively, might compromise device safety.²³² However, the total number of supplements (including those used for trivial alterations) and the rate at which all supplements were granted did not correlate with the occurrence of Class I recalls, even in Cardiovascular devices.²³³ This should provide reassurance that

232. *See supra* Section II.C.

233. Cardiovascular devices in the upper quartile of total number of supplements granted had a recall risk of 9.1%, compared with 8.8% for other devices (p=.96).

frequent use of 30-Day, Real-Time, and other pathways for minor device modifications does not compromise safety.

Finally, the present study suggests that following a Class I recall, the FDA should consider requiring additional safety assurance if a device is to remain on the market. The conditional phrasing of this suggestion is deliberate: The finding that a Class I recall predicts a higher risk of subsequent Class I recalls is more tenuous, with statistical significance only in some data sets.

VI. LIMITATIONS OF THE EXISTING EMPIRICAL DATA ON THE SAFETY OF PMA APPROVED DEVICES: CONSIDERATIONS FOR FUTURE RESEARCH

It is possible to read the empirical literature on the PMA process as providing evidence that the formally rigorous pathway functions well to ensure the safety of the highest-risk medical devices. Most prior studies have suggested a low risk of device failure. Even the study presented here indicates that roughly 95% of devices will not fail in such a way as to endanger patients' lives and health.

Unfortunately, all of these studies define only a best-case scenario. This is because all have used the FDA classification of a corrective action as a Class I recall as the proxy for costly, potentially life-threatening device-related problems that were not detected during the PMA process. Limiting the analysis to FDA data on Class I recalls underestimates the number of relevant failures. First, device failures must be reported in order to be entered into the FDA's databases. Estimates of the underreporting of device problems range from 10% to 99.5%.²³⁴ Although Congress has at times strengthened manufacturers' and hospitals' tracking and reporting obligations,²³⁵ several factors contribute to the persistent underreporting of device failures. Reporting depends in part on patients, physicians, and hospitals making efforts to alert the Agency to problems. Patients

234. See Ward & Clarkson, *supra* note 135, at 5 (describing underreporting rates for device-related medical errors ranging from 10 to 80%); *Hearings, supra* note 20, at 11 (statement of Frederic S. Resnic) (citing GAO estimate that only 0.5% of medical device adverse events are reported to the FDA).

235. See Safe Medical Devices Act of 1990, Pub. L. No. 101-629, § 2, 104 Stat. 5411, 5411 (codified as amended at 21 U.S.C. § 360i) (requiring hospitals and other facilities to report information to the manufacturer and the FDA that suggests a device caused or contributed to a patient's death, serious illness, or serious injury).

may never recognize a device flaw; in the most serious situations, the abrupt failure of a device upon which a patient's life depends can lead to sudden death with no obvious cause. In many patient populations—the elderly and those with comorbidities—deaths may be attributed to other causes, delaying or even preventing discovery of a device problem. Physicians may fail to realize the importance of reporting device problems.²³⁶ Congress has eased manufacturers' and hospitals' reporting requirements.²³⁷ And device manufacturers have strong incentives to avoid alerting the FDA to problems with their medical devices.²³⁸

In addition, even when device failures are reported to the FDA, restricting the scope of a study to Class 1 recalls assumes the FDA correctly classified the event. However, the low quality of many of the individual reports may prevent or delay the Agency's recognition of a pattern of device failures;²³⁹ low quality reports could also prevent the FDA from appropriately classifying a recall. Further, manufacturers may influence the FDA's classification of a recall through the strategic presentation of information to the Agency and the public. Downplaying the extent or severity of a device failure, suggesting that monitoring can detect clinical risk before it manifests, and emphasizing the adverse consequences of corrective action are among many strategies that could tilt the FDA toward classifying a recall as Class 2 rather than Class 1.²⁴⁰ Finally, FDA classification of medical device recalls is inherently subjective.²⁴¹ The

236. See *Hearings*, *supra* note 20, at 68 (statement of Diana Zuckerman).

237. See Food and Drug Modernization Act of 1997, Pub. L. No. 105-115, § 213, 111 Stat. 2296, 2346 (codified as amended at 21 U.S.C. § 360i) (removing certain reporting obligations for hospitals and device manufacturers).

238. See Thirumalai & Sinha, *supra* note 9, at 379 (discussing factors including the costs of replacing defective devices, the costs of the recall itself, costs related to inventory of the recalled device that cannot be sold, loss of market share, reputational harm). All of the factors that Thirumalai and Sinha discuss could incentivize manufacturers to refrain from informing the FDA of device-related problems in spite of their statutory duties to do so.

239. See, e.g., Kramer et al., *supra* note 147 (reporting that FDA had accumulated 679 reports over three years of Sprint Fidelis lead failures before it issued a recall).

240. For example, Guidant justified not informing the FDA, physicians, and patients about failures of certain implantable defibrillators because of the risks associated with surgical replacement of the devices. See Robert G. Hauser & Barry J. Maron, *Lessons from the Failure and Recall of an Implantable Cardioverter-Defibrillator*, 112 CIRCULATION 2040 (2005).

241. See *Hearings*, *supra* note 20, at 105 (statement of Ralph F. Hall) (advocating the adoption of “more objective criteria for the classification of recalls”).

border between these categories is defined by the likelihood and the severity of harm that may occur. Device failures that present a “reasonable probability” of harm should trigger a Class 1 recall, while failures that “may” cause harm should lead to a Class 2 recall.²⁴² Failures that result in “serious adverse health consequences or death” should trigger a Class 1 recall while those that “cause temporary or medically reversible adverse health consequences” should lead to a Class 2 recall.²⁴³ Both of these definitions obviously offer latitude that could result in corrections for serious device failures being classified as Class 2 recalls.

One example of a Class 2 recall demonstrates that these concerns are germane. The Class 2 recall for the Guidant Contak Renewal 3 RF implantable cardioverter defibrillator was for “a potential for malfunction of a high voltage wire, which could compromise effectiveness of shock therapy.”²⁴⁴ Such a failure could result in the death of a device recipient in the event the device failed to treat a potentially fatal heart rhythm disorder. In fact, this catastrophic outcome had already occurred to the recipient of a similar implanted cardioverter defibrillator, the Guidant Prizm 2 DR, the correction of which the FDA had previously classified as a Class 1 recall.²⁴⁵

One recent study suggested that the problem is systematic, finding that in a subset of devices cleared or approved through the 510(k) and PMA pathways, the FDA more often than not classified

242. 21 C.F.R. § 7.3(m) (2017).

243. *Id.*

244. See *Class 2 Device Recall Guidant CONTAK RENEWAL 3 RF & 4 RF*, U.S. FOOD & DRUG ADMIN., <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRES/res.cfm?id=48748> (last updated Jan. 8, 2018).

245. See *Class 1 Device Recall Ventak PRIZM 2 DR ICD*, U.S. FOOD & DRUG ADMIN., <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRES/res.cfm?id=39930> (last updated Jan. 8, 2018). In other cases, corrective actions for some lots of a device were assigned as a Class 1 recall while others were assigned as a Class 2 recall, even though the failure mode was the same. See *Medical Device Recalls*, U.S. FOOD & DRUG ADMIN., https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRES/res.cfm?start_search=1&pnumber=P950020 (last updated Jan. 8, 2018). The manufacturer initiated a corrective action on December 7, 2005, for thousands of lots of the device, which had exhibited an increased risk that a malfunctioning angioplasty catheter could separate, necessitating emergency open heart surgery. The FDA assigned most lots as a Class 1 recall, but assigned others a Class 2 recall. The information available on the Recalls webpage indicated that problem was the same in all lots.

potentially serious device problems as Class 2 instead of Class 1.²⁴⁶ This study investigated computer-related failures and recalls, suggesting that the devices were relatively complex.²⁴⁷ The investigators used a broad definition of “safety-critical” problems to include not only Class 1 recalls but also recalls for which the FDA’s recall database “indicated a patient safety issue such as injury or death” and the “potential for exposing patients or users to immediate physical safety hazards such as overdose, overexposure, electrical shock, burning, or fire.”²⁴⁸ Out of 197 recalls classified as safety-critical by the investigators, 155 (79%) were classified by the FDA as a Class 2 recall.²⁴⁹

This data does not directly address the study I present here. A majority of the 197 devices in the study were cleared through the 510(k) pathway. Further, the investigators employed a broader definition of a serious safety issue than does the FDA.²⁵⁰ But even if more stringent criteria were employed, their findings suggest that the FDA classifies corrective actions as Class 2 recalls for a significant portion of potentially fatal device problems. Thus, it is clear that using Class 1 recalls as a proxy for compromised device safety will underestimate the true risk, and will provide only an upper bound for effectiveness of FDA review to ensure device safety.

Beyond these considerations of underreporting and classification errors, there is a more fundamental problem with relying on Class 1 recalls as a proxy for failures of the premarket evaluation processes to detect design and manufacturing flaws. The regulatory standard for a Class 2 recall includes situations in which device failures “may cause temporary or medically reversible adverse health consequences.”²⁵¹ Yet even temporary or reversible consequences may subject patients to fear, anxiety, painful complications, and the need for surgery with its attendant risks. And these failures may be quite costly to correct. For example, the FDA has classified the corrective actions to address failures of some hip prostheses as Class 2 recalls, which conforms

246. Homa Alemzadeh et al., *Analysis of Safety-Critical Computer Failures in Medical Devices*, IEEE SECURITY & PRIVACY, July/August 2013, at 14, 14.

247. *See id.* at 16.

248. *Id.* at 15.

249. *Id.* at 21.

250. *Id.*

251. *Industry Guidance*, *supra* note 138; *see also* 21 C.F.R. § 7.3(m)(2) (2017).

with the definition in 21 C.F.R. § 7.3(m)(2) because these failures were unlikely to cause death and could be corrected by surgical replacement. However, as Diana Zuckerman has noted, surgery “costs an average of \$35,000 and results in a 3–5 day hospital stay, at least 6 weeks walking with crutches or a walker or cane, 4 weeks where the patient is not allowed to drive, and several weeks or months of rehab or physical therapy.”²⁵² Just one model of hip prosthesis that failed had been implanted into thousands of U.S. patients.²⁵³ A well-functioning premarket evaluation system should minimize device problems that lead to these physical, psychological, and financial consequences.²⁵⁴ Studies that rely on Class 1 recalls cannot assess these costs of device failures.

These problems with the use of Class 1 recalls as the sole proxy for relevant device failures lead to two suggestions. First, policy considerations must be based on the assumption that a risk of 4.6%–6%, which my study found, represents the best-case estimate of how well the PMA process ensures device safety. Second, assuming that we cannot define a single, “true” value for this risk, it is imperative that future projects attempt to define a worst-case scenario as well.

VII. CONCLUSION

Undoubtedly, the PMA pathway, through which many high-risk medical devices are approved for marketing in the United States, is rigorous, imposing significant costs and obligations on manufacturers. But the medical literature has raised significant concerns about the ability of the PMA process to ensure device safety. The study presented here validates many, but not all, of those concerns. The best-case scenario is that a significant percentage and large numbers of devices fail in ways that present real dangers to patients and that impose real costs on health care financing systems.

252. *Hearings, supra* note 20, at 65–67 (statement of Diana Zuckerman); JOHNSON 2016, *supra* note 28, at 2.

253. *Hearings, supra* note 20, at 2 (statement of Frederic Resnic); JOHNSON 2016, *supra* note 28, at 2.

254. See *Hearings, supra* note 20, at 87 (statement of Ralph F. Hall), for an opposing argument. Hall justified the use of Class 1 recalls as a proxy for regulatory failure, citing the FDA’s criteria for Class 2 recalls as involving problems that “might cause a temporary health problem, or pose only a slight threat of a serious nature,” and by noting that some Class 1 recalls were for problems carrying only a 1/20,000 risk of death. *Id.*

Two aspects of the statutory and regulatory framework that governs the PMA evaluation process appear to contribute to this problem. First, Congress's requirement that the FDA employ the "least burdensome" approach to regulation, even of the highest-risk devices, has led to device approvals based on less-than-rigorous data. Second, the PMA Supplement pathways that allow for repeated, significant device modifications without a demonstration of overall safety create a "predicate drift" problem in the most complex devices.

The changes to the PMA processes I have suggested here—requiring longer, larger, and more rigorous study data, limiting the number of and rate at which 180-Day and Panel-Track supplements are granted to complex devices, and averting the shift to a reliance on post-market data encouraged by the 21st Century Cures Act—would partially address these problems. But all of these changes would come at a cost: they would tip the balance between safety on one hand and the innovation of and access to new technology on the other. Manufacturers might not invest the resources need to develop new technologies, and patients' access to those technologies might be delayed.

Balancing these conflicting goals requires a fine-grained analysis. How we weigh safety, innovation, and access may—in fact should—differ depending on the context: a different balance should be struck for breakthrough technologies—those that offer effective treatments for serious conditions that currently have no effective treatments—than for the fourth or eighth market entrant providing the same treatment as the previous three or seven devices. And how closely we scrutinize devices needs to be determined on a more granular level: even within the group of high risk, Class III devices, differences in complexity and the rate of technological change call for different quanta of data to support a finding of a reasonable assurance of safety.

Any such analysis should be informed by reliable data. If we are truly to determine how well or poorly the PMA process is working, a more searching analysis of failures of PMA-approved high-risk devices is needed. Using Class I recalls as a proxy is sure to underestimate the true incidence of failures. Adding the relevant failures to which the FDA has responded with a Class 2 recall classification might be a first step. Examining other data sources, while labor-intensive, might also be useful. But as a matter of policy,

it would be advisable to know the best-case and the worst-case scenarios for how well the PMA evaluation process serves to ensure device safety before weakening the existing framework that governs that process.

Appendix 1

Class 1 Recalls Announced Between November 1, 2002, and November 1, 2016, for Products in Primary Data Set

Device PMA	Manufacturer and Device Name	Date of Firm-Initiated Action	FDA-Determined Cause	Manufacturer Explanation of Reason for Recall	Number of Units Affected
P020006	Boston Scientific Enteryx Injectable Bulking Agent	9/23/05	Other	“Incorrect implantation by cause serious health complications”	8703
P030011	Temporary Total Artificial Heart with Freedom Driver System	8/6/15	Component Design/ Selection	Component failure “may fail and cause the drive mechanism to stop pumping”	56
P030025	Boston Scientific Taxus Express Paclitaxel-Eluting Stent	7/1/04	Other	Balloon failed to deflate after stent deployment	209,373
P040020 *	Alcon AcrySof IQ Toric Intraocular Lens	4/15/15	Process Design	Toxic Anterior Segment Syndrome (Post-Operative Inflammation of eye)	45,391
		9/29/15	Process Design	Toxic Anterior Segment Syndrome (Post-Operative Inflammation of eye)	43,651
P040025	Natus Medical Olympic Cool Cap System	5/9/12	Non-conforming Material/ Component	Control screen freezes, displaying information although device stopped functioning	324

P050018	Angioscore PTCA Scoring Balloon Catheter	12/4/09	Process Design	Catheter shaft separates, allowing fragments to lodge in coronary arteries	3870
P050025	Boston Scientific Nexstent Carotid Stent	6/6/08	Component Design/ Selection	Catheter tip detaches	1946
P060040 (original, S004, S005)	Thoratec Heartmate II Left Ventricular Assist Device	2/23/12	Device Design	Disconnected bend reliefs	3852
P060040		3/4/14	Device Design	Death and injury during change of controller unit	13,938
P070014 S032	Bard Lifesent Solo Femoral Artery Stent	9/30/13	Device Design	Partial deployment	4112
P080009	Ethicon SEDASYS Computer- Assisted Personalized Sedation System	5/13/15	Device Design	Disinfection degrades plastic	14
P100009	Abbott Mitraclip Clip Delivery System	2/4/16	Use Error	Clip could not be detached from delivery system	7600
P100018	Micro Therapeutics Inc. Pipeline Embolization Device for intracranial aneurysms	4/1/14	Process Change Control	Coating might detach from delivery wire	32
P100022	Cook Zilver PTX Drug Eluting Peripheral Stent	4/18/13	Process Control	Fracture of delivery system leading to tip separation	1592

P100047	HeartWare Ventricular Assist System	12/6/13	Process Control	Driveline connector separation	4586
		1/16/15	Device Design	Controller sensitivity to ESD	4845
		4/29/15	Device Design	Damaged or bent connection pins within power supply ports	3782
		7/8/16	Under Investigation	Premature Battery Depletion	18,631
		8/17/16	Component Design/Selection	Foreign material causing electrical faults	344
		10/17/16	Under Investigation	Loose connector	8799
P120005	Dexcom Inc. G4 S002, S011, S018, S028, S031, S033	2/23/16	Device Design	Failure of audible warning of severe low or high blood glucose	270,111
P100006	Ovation Prime Abdominal Stent Graft	9/20/14	Under Investigation by Firm	Rapid emptying of polymer syringe during implant procedure	0
P130005	Cardiovascular Systems, Inc. Diamondback 360 Coronary Orbital Atherectomy System	3/4/15	Non-confirming Material/Component	Outer sheath may flake off resulting in embolism	30
P130021	Medtronic EnVeo R Aortic Valve Prosthesis Delivery System	7/15/15	Process Control	Particular Matter	6912

Totals		
18	Artificial Heart/VAD—3	662,494
Devices	Stent—5	
	Heart Valve—2	
	Intraocular Lens—1	
	Glucose Monitor—1	
	Infant Cooling Cap—1	
	Sedation System—1	
	Embolization Device—1	
	Injectable Bulking Agent—1	
	Angioplasty—2	

- * This device was considered to have had one recall because the second was merely an expansion of the first.