Published in final edited form as:

JAMA. 2017 July 25; 318(4): 332–334. doi:10.1001/jama.2017.8976.

Practical Improvements for Medical Device Evaluation

Daniel B. Kramer, MD, MPH and Robert W. Yeh, MD, MSc

Richard A. and Susan F. Smith Center for Outcomes Research in Cardiology, Beth Israel
Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts

Passage of the Medical Device Amendments Act in 1976 confirmed the US Food and Drug Administration's (FDA's) primary responsibility for evaluating the safety and effectiveness of medical devices in the United States.¹ Although there have been modest legislative updates in the ensuing decades, the broad structure of the FDA's risk-based framework for premarket evaluation has remained largely unchanged.² High-risk ("class III") devices (such as pacemakers and insulin pumps) are subject to stringent premarket requirements including demonstration of clinical effectiveness, typically done through the Premarket Approval pathway. Medium-risk devices (such as glucose monitors and CT scanners) generally earn marketing clearance through the "510k" pathway, which assesses whether a given device raises new safety or effectiveness concerns compared with a currently-marketed device to which the technology being evaluated is deemed "substantially equivalent."³

Even with diligent premarket review, introducing new therapeutics into clinical practice will involve important uncertainties. While the FDA has several tools for leveraging postmarket requirements in concert with the perceived risks of devices, this approach has been criticized as inefficient, lightly enforced, and ineffective. In addition, postmarket studies assume a nuanced understanding of the likely safety questions and effectiveness questions that need further evaluation and will be less effective in identifying unforeseen or rare events. The FDA's system for passive adverse event collection potentially serves as a mechanism for early detection of safety signals. But this approach is known to be limited by underreporting, incomplete event information, and dissociation from key data such as the true denominator for device use.

Recognizing these limitations, FDA and other interested parties have advocated for a strengthened National Evaluation System for Health Technology (NEST), envisioned as leveraging electronic health records, billing claims, professional society registries, and the wide range of expertise needed to generate informative clinical knowledge from diverse data

Corresponding Author: Daniel B. Kramer, MD, MPH, Beth Israel Deaconess Medical Center, Harvard Medical School, 375 Longwood Ave, Boston, MA 02215 (dkramer@bidmc.harvard.edu).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Kramer reports support from a Paul B. Beeson Career Development Award in Aging Research (NIH-NIA K23AG045963) and the Greenwall Faculty Scholars Program, and reports serving as a consultant to the Circulatory Systems Advisory Panel of the US Food and Drug Administration (FDA) and the Baim Clinical Research Institute for clinical trials of medical devices (unrelated to the current topic). Dr Yeh reports serving as principal investigator of several NHLBI-funded trials and serves as a consultant for Abbott Vascular and Boston Scientific. He reports receiving research funding from Abiomed and Boston Scientific, and salary from the Baim Clinical Research Institute. He has been a member of the methodology center at Harvard Medical School for the FDA's MDEpinet Initiative.

sources.⁴ Two Viewpoints in this issue of *JAMA* describe recent clinical examples highlighting opportunities and ongoing challenges for this proposed system.

In one Viewpoint, Redberg and coauthors⁵ describe the regulatory and clinical experience of power morcellators, a device type originally cleared via the 510k pathway in 1991 as a tool to support laparoscopic procedures for tissue removal that otherwise might require open surgery. Power morcellators became commonly used for gynecologic surgery, with hundreds of thousands of cases performed with this device prior to safety concerns arising in 2013 regarding the potential for malignant cells to be inadvertently disseminated in the abdomen. The device remains cleared for marketing with a black box warning, but not without controversy.⁶ As Redberg et al note, ongoing debate about the proper clinical role for power morcellators arises in part from the slow recognition of a legitimate safety concern not appreciated at the time of initial device review. The comprehensive system of device surveillance envisioned by NEST—capturing the true numerator, denominator, risk factors, and key outcomes—theoretically could provide more precise estimates of the risks and benefits of new devices to support clinical decisions than those that have been generated by current approaches. However, Redberg et al critique the slow pace with which such proposals to improve device surveillance are being implemented, including the adoption of unique device identifiers (UDIs) and the development of automated systems of safety surveillance.

In another Viewpoint in this issue of *JAMA*, Ibrahim and Dimick⁷ similarly raise the question of whether coordinated use of data from administrative claims might have led to the more timely identification of device malfunction and regulatory action involving gastric bands used for laparoscopic weight loss surgery. These devices were approved through the premarket approval process in 2001 and had rapid uptake, but they also had notable decline in usage after safety and effectiveness concerns emerged. Unlike most new devices, however, the gastric band acquired its own *ICD-9* code, allowing the authors to describe trends in initial procedures, revisions, and associated costs using publicly available Medicare claims data, precisely as the UDI initiative is intended to do for all medical devices. As the authors point out, observations from these claims data, which suggested relatively high proportions of revision procedures, contrasted sharply with the published clinical trial evidence—yet only the latter apparently informed subsequent FDA decisions regarding the device, which remains on the market.

These case examples describe similar regulatory and clinical trajectories: initial marketing authority based on very limited premarket evaluation; brisk uptake; appreciation of new concerns about safety, effectiveness, or both; and subsequent swift drop-off in utilization. Both cast new light on what the FDA and other observers have long acknowledged: passive, largely voluntary adverse event collection is deeply inadequate, yet remains the only system currently applicable to most medical devices. However, while these case studies expand the literature of postmarket failures, near-term solutions remain elusive.

These Viewpoints advocate the use of claims data to support postmarket assessment of several key questions: What are the best estimates of device safety and effectiveness (and for which specific end points?) over time, and how precise are those estimates? How does

clinical practice change in response to new marketing approval and emergence of new clinical data (either supportive or worrisome)? Large administrative claims, linked to specific devices via the UDI system, are a tantalizing target because of several theoretical strengths. These include the size and completeness of study populations, relative ease of collecting data with geographic, racial, and ethnic diversity (compared with a clinical trial, for example), and longitudinal measurement of multiple covariates and clinical end points. But this postmarket strategy poses several challenges for stakeholders seeking to distill a clear and actionable message from claims data.

Importantly, broad uptake of the UDI has been slow, and the vast majority of medical devices cannot be evaluated using a unique billing claim, as Ibrahim and Dimick were able to in their analysis of the gastric band. In addition, availability of these databases typically lags for 2 or more years and are commonly limited, as in the case of Medicare data, to specific payer or age groups. Moreover, although some end points such as all-cause mortality may be accurately identified by these data, other important safety concerns may be inaccurately recorded in administrative claims.

Even with a well-defined clinical question with clear end points, different observational research techniques may arrive at different conclusions. For example, 2 recent analyses of the clinical benefits for vascular closure devices—both using large national registries but different methodology—arrived at markedly different conclusions.^{8,9} Expanding these types of analyses for regulatory oversight to more complex physician-device-patient interactions will require even more attention to methods and underlying assumptions. Novel analytic methods, including the automated surveillance of devices using more detailed clinical registry data, have shown promise in leveraging large data sets for evaluating device performance.¹⁰ However, these approaches introduce their own assumptions and limitations. In particular, it is not clear how automated surveillance could be calibrated accurately to detect not just end points known to be of interest at the time of approval (eg, thrombosis of recently placed coronary stents) but unforeseen complications, such as those encountered with power morcellators.

Redberg et al correctly suggest that clinicians need to demand higher-quality data prior to adopting new devices, particularly those with established alternatives for which outcomes are well known. However, improvements in the design and conduct of postmarket studies can improve the value of information gained in the near term. For most devices, enrollment in postmarket studies, such as those embedded in national registries, will remain the most meaningful method of surveillance in the near future, and patients need to be more vigorously engaged in enrollment and follow-up. Registry-based postmarket studies provide more granular information than claims, are less expensive than de novo clinical trials, and may better capture representative patient populations.¹¹ As has been done with other new devices such as transcatheter aortic valves, payers such as Medicare can help reinforce enrollment by making access to a device with lingering safety or effectiveness questions conditional on enrollment in an approved postmarket surveillance study.

Additional pressure brought to bear on manufacturers might also improve the dismal rates of completion and public availability of postmarket studies.^{12,13} While the FDA can

theoretically withdraw marketing authority for devices for which sponsors have failed to fulfill postmarket commitments, there have been no recent cases in which this has occurred. To motivate postmarket study completion, select devices might be granted approval only for a set period, at which time failure to meet predetermined benchmarks for postmarket study completion would automatically rescind FDA approval.¹⁴ More strategic engagement of professional societies and patient advocacy groups may also help make the clinical community more aware of the critical importance of timely and rigorous postmarket studies.

Taken together, the Viewpoints discussing power morcellators and gastric bands in this issue of *JAMA* reinforce the shortcomings of current postmarket medical device regulation in the United States, the adverse effects on the health of patients, and the need for robust improvements including those suggested by the NEST proposal. Even though these strategies may yet provide more comprehensive, accurate, and precise assessments of device-related outcomes, challenges persist for integrating diverse data sources and a heterogeneous group of stakeholders. Almost inevitably, a medical device currently in use will prove to be the next unfortunate example arising from a system that might have been better, but for lack of urgency and regulatory will in the service of public health.

References

1. Medical Device Amendments of 1976 to the Federal Food, Drug, and Cosmetic Act. Pub L No. 94–295, 90 Stat 539 (1976).
2. Kramer DB, Kesselheim AS. User fees and beyond—the FDA Safety and Innovation Act of 2012. *N Engl J Med*. 2012; 367(14):1277–1279. [PubMed: 23034017]
3. Kramer DB, Xu S, Kesselheim AS. Regulation of medical devices in the United States and European Union. *N Engl J Med*. 2012; 366(9):848–855. [PubMed: 22332952]
4. Shuren J, Califf RM. Need for a national evaluation system for health technology. *JAMA*. 2016; 316(11):1153–1154. [PubMed: 27398696]
5. Redberg RF, Jacoby AF, Sharfstein JM. Power morcellators, postmarketing surveillance, and the US Food and Drug Administration. *JAMA*.
6. Rosenbaum L. N-of-1 policymaking: tragedy, trade-offs, and the demise of morcellation. *N Engl J Med*. 2016; 374(10):986–990. [PubMed: 26962735]
7. Ibrahim AM, Dimick JB. Monitoring medical devices: missed warning signs within existing data. *JAMA*.
8. Wimmer NJ, Secemsky EA, Mauri L, et al. Effectiveness of arterial closure devices for preventing complications with percutaneous coronary intervention: an instrumental variable analysis. *Circ Cardiovasc Interv*. 2016; 9(4):e003464. [PubMed: 27059685]
9. Farooq V, Goedhart D, Ludman P, de Belder MA, Harcombe A, El-Omar M. British Cardiovascular Intervention Society and the National Institute for Cardiovascular Outcomes Research. Relationship between femoral vascular closure devices and short-term mortality from 271 845 percutaneous coronary intervention procedures performed in the United Kingdom between 2006 and 2011: a propensity score-corrected analysis from the British Cardiovascular Intervention Society. *Circ Cardiovasc Interv*. 2016; 9(6):e003560. [PubMed: 27225421]
10. Resnic FS, Majithia A, Marinac-Dabic D, et al. Registry-based prospective, active surveillance of medical-device safety. *N Engl J Med*. 2017; 376(6):526–535. [PubMed: 28121489]
11. Yeh RW, Kennedy K, Spertus JA, et al. Do postmarketing surveillance studies represent real-world populations? A comparison of patient characteristics and outcomes after carotid artery stenting. *Circulation*. 2011; 123(13):1384–1390. [PubMed: 21422383]
12. Rathi VK, Krumholz HM, Masoudi FA, Ross JS. 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. *JAMA*. 2015; 314(6):604–612. [PubMed: 26262798]

13. Reynolds IS, Rising JP, Coukell AJ, Paulson KH, Redberg RF. Assessing the safety and effectiveness of devices after US Food and Drug Administration approval: FDA-mandated postapproval studies. *JAMA Intern Med.* 2014; 174(11):1773–1779. [PubMed: 25265209]
14. Kramer DB, Tan YT, Sato C, Kesselheim AS. Postmarket surveillance of medical devices: a comparison of strategies in the US, EU, Japan, and China. *PLoS Med.* 2013; 10(9):e1001519. [PubMed: 24223035]