

Real-World Evidence to Guide the Approval and Use of New Treatments

Steven Galson, Amgen, and Gregory Simon, Group Health Research Institute

October 18, 2016

Current State

The Focus of Traditional Evidence Generation on Narrow Questions Regarding Efficacy and Safety

Research regarding new treatments (drugs, biological products, and high-risk devices) often begins with a broad assessment of disease epidemiology, disease burden, and shortcomings of existing treatments. That research may draw from diverse data sources, including real-world data generated by health system operations (see Figure 1).

The clinical research phase of treatment development typically follows a well-established pathway from initial evaluation of safety to preliminary evaluation of therapeutic efficacy to pivotal trials intended to support regulatory approval for marketing. Those pivotal trials focus on key questions of efficacy (typically in comparison to placebo or some analogous control condition) and safety (especially serious or previously unrecognized adverse effects). This focus is consistent with the responsibility of the U.S. Food and Drug Administration (FDA) for assuring the safety and efficacy of drugs, biological products, and medical devices at the time of approval.

Regulatory approval is sometimes followed by systematic postmarketing evaluation to address a wider range of practical or real-world questions. This more pragmatic research again draws from more varied sources of data drawn from diverse clinical settings.

A Lack of Information for Stakeholders (Patients, Providers, and Health Systems) to Guide Real-World Decisions

Evidence generated by traditional clinical research often fails to address key questions of patients, physicians, and health systems regarding the appropriate role of new treatments. Those unaddressed concerns include the following:

Effectiveness

Traditional efficacy trials typically aim to evaluate a single treatment rigorously. In contrast, patients, providers, and health systems choose among alternative treatments on the basis of net benefit in real-world practice. Real-world effectiveness may differ substantially from efficacy detected in the traditional clinical trial setting. Factors contributing to that efficacy–effectiveness gap include variation in practice settings, provider decision making, patient adherence, co-occurring conditions, and concomitant treatments. In a clinical trial designed to assess efficacy, these factors would be considered sources of noise or error, and trial design would attempt to minimize variation. In everyday clinical practice, these sources of variation are directly relevant to practical decisions by patients and providers and are central to the information stakeholders need to inform practical decisions. Many treatments would be expected to show some slippage or loss of benefit between efficacy trials and real-world practice. But we cannot presume that slippage is consistent across treatments, patient populations, or practice settings. Consequently, findings from efficacy trials regarding differences (or lack of differences) in efficacy do not necessarily translate to the same differences in real-world performance.

Tolerability

While patients and physicians are certainly concerned about less common and more serious adverse effects of new treatments, they are equally concerned about more common adverse effects—such as nausea, tremor, fatigue, weight gain, or interference with sexual function. Even when traditional efficacy trials evaluate these effects, the resulting evidence is rarely adequate to guide patients' and physicians' decisions regarding alternative treatments—especially if those treatments are to be continued for months or years.

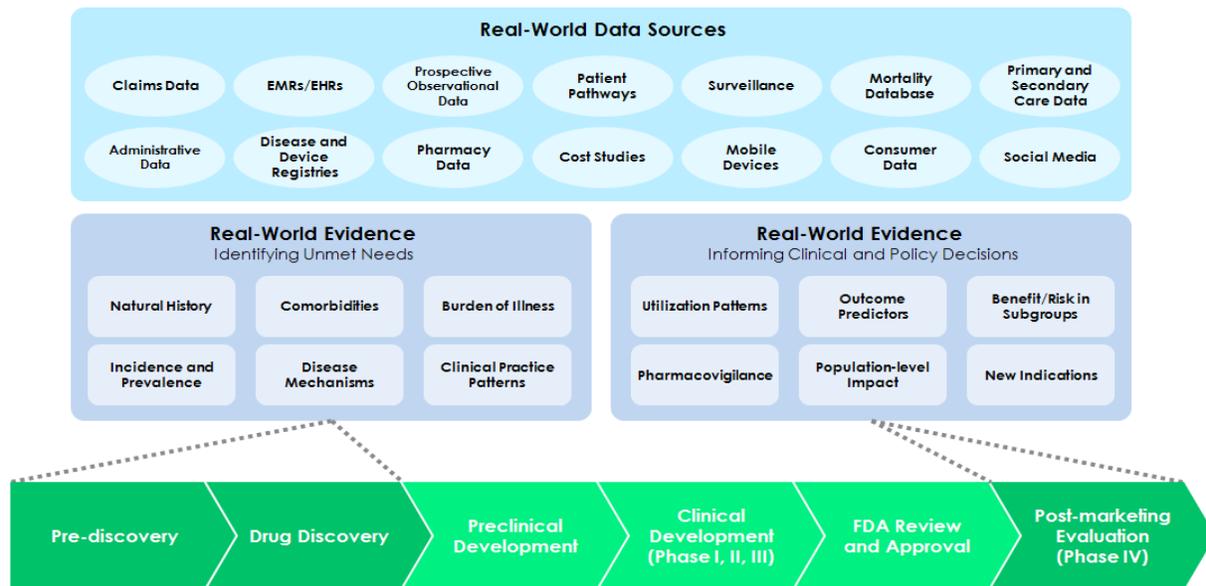


Figure 1 | Real-world evidence is derived from curating, standardizing, and analyzing real-world data to obtain high-quality and reliable information from diverse and complex sources. Real-world evidence could inform all phases of treatment discovery and development, although thus far has been more commonly used to inform the questions and data in early development decisions and in postmarketing safety surveillance or comparative effectiveness studies. In contrast, clinical development and review has tended to more use idealized and tightly controlled data sources for efficacy trials.

Value

Health care payers base coverage and reimbursement decisions regarding new treatments on the balance of net cost and benefit, although payers may not consider long-term factors, such as the benefits of prevention, over a typical lifespan. Given the increasing prevalence of high insurance deductibles and coinsurance arrangements, patients and families must also consider the value of alternative treatments. True net cost of a new treatment depends not only on its price but also on the net impact on overall cost of care. Traditional clinical trials, in which treatment protocols are highly controlled, usually offer little information on how new treatments affect overall or downstream use of health services.

Heterogeneity of Treatment Effects

An individual patient and his or her physician are naturally most interested in person-specific effects (“What are the expected benefits and harms of this treatment for someone like me?”). Traditional clinical trials focus instead on assessment of average effects. Particularly

in the early phases, heterogeneity of effects is more often a source of error to be minimized rather than an important signal to be detected.

Current Incentives Do Not Promote Necessary Innovation

Our current evidence-generating process has evolved to fit our traditional regulatory and business environment. Business imperatives of new treatment development drive research toward a relatively narrow focus: producing the data essential for regulatory approval. Traditional clinical trials are optimized to efficiently address key questions in the regulatory process: Is a new treatment superior to a placebo or other appropriate control treatment with respect to a specific clinical outcome? Is there evidence of a specific danger or harm—especially a harm not previously recognized?

Bringing a new treatment to market involves significant time and expense. For developers of new treatments, broadening research to address real-world questions may introduce additional uncertainty or delay and, in addition, require data not readily

available to industry. Expanding the evidence-generating process to address real-world effectiveness, value, tolerability, and heterogeneity of treatment effects would almost always require more flexible treatment protocols and more heterogeneous clinical populations. The “noise” introduced by that flexibility and heterogeneity could certainly interfere with the detection of primary “signals” regarding efficacy and safety.

Ideally, developers of new treatments would be rewarded for generating evidence more relevant to real-world decisions. Those rewards might include approval for labeling regarding improved effectiveness, tolerability, or value. Some European regulators may consider evidence regarding cost-effectiveness or value in regulatory or pricing decisions. In the United States, research to support those claims has been impeded by uncertainty regarding the types of evidence that will be acceptable to support approval of novel therapies and new indications (Bipartisan Policy Center, 2016).

Enabling Developments

The Increasing Use of Electronic Health Records and Development of Linked Data Resources

Data generated from research and practice have historically been siloed. However, as the concept of a learning health system continues to take hold, such distinctions are increasingly being reexamined. The nation’s electronic health information infrastructure has continued to mature over the past decades, and, as a result, a wealth of clinical data—residing in electronic health records, patient registries, and administrative claims databases—now provides an opportunity to generate evidence on the effectiveness of medical products directly from clinical experience, complementing the data generated through traditional randomized controlled trials. Electronic records systems can certainly facilitate the traditional clinical trials, expanding the scale and lowering the cost of participant recruitment and recording of clinical outcomes. In addition, the “data exhaust” of ordinary (nonresearch) health care has the potential to inform regulatory and reimbursement decisions. However, ensuring that the data are fit for the purpose of research will require thoughtful consideration. Most real-world data, including those from electronic health records and claims databases, are not currently generated for the purpose of research.

Leveraging the full potential of these data sources will depend on data integration capabilities. Analytic processes enabling patient-level linking of disparate clinical data sources are helping to address data quality issues (e.g., filling in missing data and vetting data by searching for inconsistencies across data sources). Networked systems (e.g., the Observational Health Data Sciences and Informatics [OHDSI] Collaboration) are enabling data aggregation on a scale not previously possible and are yielding information on patient population characteristics and health care utilization that may significantly impact future trial design and improve the generalizability of results (Hripcsak et al., 2016).

A key challenge of leveraging clinical data to support the evaluation of medical products resides in the fact that providers and health systems are not systematically assessing the impact of treatment. The routine collection of standard outcome measures is an important but tractable barrier.

The Increasing Use of Standard Patient-Reported Outcome Measures

In addition to traditional clinical data, patient experiences and perspectives are increasingly being incorporated as essential aspects of the medical product evaluation process. Patient-reported outcomes may include symptoms, quality of life, and functional status. Patient-centered outcomes research is a primary focus of the Patient-Centered Outcomes Research Institute (PCORI), which has initiated the development of minimum standards for patient-reported outcomes data. PCORI has also established PCORnet, a centrally coordinated network of research networks, to improve the speed and efficiency of clinical research by leveraging existing large data sources, including patient-powered research networks.

The Use of E-Health and M-Health Tools to Collect Real-World Data

The increasing popularity of wearable and other mobile devices that collect health-related data from individuals has opened new avenues for consumer and patient engagement and the collection of real-world data. Of course, such data only have value if they are meaningful. Mobile devices such as medical wearables that can passively and accurately collect data on

primary clinical endpoints can help to demonstrate the benefit of a particular intervention—a capability that is of great interest to not only patients and providers but also health systems and manufacturers of medical products. Online patient communities and consumer search behavior, too, may be sources of data on effectiveness, safety, use, and compliance. However, in addition to technical barriers related to data capture and integration, a number of legal and regulatory challenges must be addressed prior to routine integration of these real-world data sources with more traditional structured clinical datasets.

Promising Practices

Population-Based Surveillance Using Health System Data

The linkage of health system records has enabled large-scale observational research including representative samples of patients treated in under real-world conditions. Research to date has focused primarily on outcomes detectable through traditional insurance claims data, including billing diagnoses and procedures. The increasing availability of rich clinical data from electronic health records should enable research regarding a broader range of effectiveness outcomes, including laboratory results, vital signs, and standardized patient-reported outcomes.

Example: FDA Sentinel Initiative

The FDA Sentinel initiative was created with the goal of establishing an integrated, national, electronic system that monitors the safety of medical products, including small molecule drugs and biologics. The initiative was taken in response to recommendations of the Institute of Medicine (IOM) (IOM, 2006) and the 2007 FDA Amendments Act that mandated FDA to have a system in place for active postmarket risk identification and analysis (Food and Drug Administration Amendments Act of 2007). The Sentinel initiative for medical product safety surveillance was launched in 2008. In 2009, the Mini-Sentinel program was established as a pilot effort to test the core function of the future system—mainly the analysis of health care information obtained from multiple and varied data sources and the utilization of the data to inform FDA decision making (Sentinel Program Interim Assessment—FY15, 2015). A main goal for the program was to enable real-time queries while maintaining the privacy of patients and to build a system that would rely on existing infrastructure and

require minimal data transfer from data sources that would remain under the control and maintenance of their owners (Kuehn, 2016). Beginning in late 2014, FDA has been transitioning from the Mini-Sentinel system to the full Sentinel initiative. Along with the Harvard Pilgrim Health Care Institute, which was chosen as a data analytics partner to the program, the initiative has established relationships with 19 data partners that cover a combined 178 million lives.

Example: Observational Health Data Sciences and Informatics Program

The OHDSI program was created with the aim of facilitating better health decisions by using collaboratively generated evidence (1). OHDSI builds on the 5-year Observational Medical Outcomes Partnership (OMOP), a public–private partnership that focused on the use of observational datasets for investigating medical products. The program was relocated to the Reagan-Udall Foundation for the FDA and is the basis of the Innovation in Medical Evidence Development and Surveillance (IMEDS) program, in which the OHDSI information model was developed (Hripcsak et al., 2015). OHDSI now operates as an international collaborative that utilizes open-source data analytic tools applied to a network of databases contributed to by 90 participants for population- and patient-level analyses. In a recent demonstration of the potential for using the OHDSI system in large-scale, international observational research, the disease treatment pathways for type 2 diabetes mellitus, hypertension, and depression were investigated using data aggregated from 11 sources providing electronic health records of 250 million patients across four countries (Hripcsak et al., 2016).

Research Embedded in New Product Registries

Systematic registries of patients exposed to new treatments are more common for the postmarketing evaluation of medical devices than for drugs or biological products. These registries can support both observational research and randomized trials to evaluate the effectiveness and tolerability of new products.

Example: Transcatheter Valve Therapy Registry

The Society for Thoracic Surgeons and the American College of Cardiology collaborated with FDA and the Centers for Medicare & Medicaid Services to create the Transcatheter Valve Therapy (TVT) Registry (2) of

patients undergoing valve repair or replacement surgery. The registry was launched in 2011 to track patient safety and outcomes from transcatheter aortic valve replacement in real-world settings, enrolling nearly all patients receiving a device. One important reason for establishing the registry was to support gaps both in premarket trials of medical devices, which are typically held only in specialized centers and on a carefully selected patient samples, and in the postregulatory approval period. Postregulatory approval is aimed at optimizing outcomes, patient selection criteria, device safety monitoring, and possible expansion of device indications. While traditionally different stakeholder groups collected data to support such efforts in disparate ways, a collaboration among professional societies, regulators, payers, the medical device industry, clinicians, and patient groups enabled the selection and harmonization of the data elements comprising the TVT Registry as well as patient selection criteria (Carroll et al., 2013; Carroll et al., 2015).

In 2011 the TVT Registry began a partnership with the Duke Clinical Research Institute for registry data analytics, and in 2013 the first results from the TVT Registry were published, reporting on outcomes from 7,710 patients. By 2015, more than 319 medical centers participated in the TVT Registry involving more than 18,500 cases of transcatheter aortic valve replacement (Rumsfeld et al., 2015). As an example of the TVT Registry's role as an infrastructure for conducting postapproval studies, in 2013 FDA approved expanded labeling for a transcatheter heart valve (Edwards SAPIEN), allowing its use among a larger segment of patients with aortic stenosis; the decision was based partially on data provided from the TVT Registry (Society of Thoracic Surgeons, 2013).

Randomization Embedded in Real-World Practice

Despite common misperceptions, real-world evidence need not be generated solely through retrospective analysis of existing data. Increasingly, clinical trials are being conducted in real-world settings to improve the generalizability of results and to reduce inefficiencies related to separate research infrastructures. These pragmatic clinical trials use an existing clinical infrastructure to prospectively test interventions in everyday situations and enable randomization at the point of care (FOCR, 2016).

Example: Salford Lung Study

Initiated in 2012, the Salford Lung Study was a 12-month, open-label, phase III pragmatic randomized controlled trial (pRCT) sponsored by GlaxoSmithKline and conducted in the Salford borough of the greater Manchester area in the United Kingdom (3). The study represents the first time a pRCT was conducted before the registration of the treatment being investigated (New et al., 2014). The study included a series of trials that evaluated a new once-daily-administered dry-powder inhaler containing both the corticosteroid fluticasone furate and the long-acting β_2 agonist vilanterol (FF/VI). Previous studies demonstrated that a combined administration was more effective in treating chronic obstructive pulmonary disease (COPD) than each of the components administered separately (Bakerly et al., 2015). The trials were specifically designed to compare the real-world effectiveness of FF/VI to existing treatments for asthma and COPD in a large segment of the population of patients during routine clinical care. The primary outcomes measured in the trials were the rate of moderate and severe exacerbations and measured improvement in asthma control for COPD and asthma, respectively.

Patients were randomized and received the usual care for the duration of their study, including dispensing of trial medication at local pharmacies. The local technological and clinical infrastructure in Salford facilitated the implementation of this pRCT. While more than 60 primary care clinics were involved in the trial, patients in the study area are served by one regional hospital, and both primary and secondary health providers share one integrated electronic health record system. The Salford Integrated Record, originally created in 2001, is updated in real time and allowed the necessary safety monitoring required from a phase III trial. Additional data feeds into the system were created to capture information on mortality and access to health care services outside the region. Furthermore, all community pharmacies in the Salford area also participated in the study and provided information on medication adherence and prescriptions delivered. Following the Salford Lung Study pRCT, the European Commission granted marketing authorization to FF/VI treatment for asthma and COPD in November 2013 (Woodcock et al., 2015).

Desired Future State

Evidence Generation Driven by the Needs of Real-World Stakeholders

Real-world evidence regarding new treatments should address the practical questions of various stakeholders:

- For patients and physicians: When is this new treatment preferred over existing alternatives? Does effectiveness or tolerability vary in a predictable way among different groups of patients or different health care settings?
- For payers: How does the value of this new treatment compare to existing alternatives? What coverage or reimbursement policies will maximize overall value to taxpayers or insurance plan members?
- For industry: Where and when is this treatment likely to deliver the greatest value to our customers?
- For regulators: How can labeling be more informative for patients and clinicians?

Incentives Realigned to Promote Relevant and Efficient Research

If developers of new treatments are expected to broaden clinical trials to address real-world questions, they must be appropriately rewarded for bearing the additional expense and not burdened by additional delay in the approval process. Creating appropriate incentives for real-world evidence generation will require the following:

- Guidance regarding appropriate data sources, research methods, and analytic methods to support labeling regarding effectiveness, tolerability, value, and heterogeneity of treatment effects; and
- Consensus among payers regarding the role of real-world evidence and more specific labeling in decisions regarding coverage and reimbursement.

Guidance on the use of real-world data from regulatory agencies—such as that recently released for medical devices (FDA, 2016)—can help to define a new paradigm for evidence generation that improves the impact of research efforts.

More Flexible Boundaries between Premarket and Postmarket Research

A new model for medical product development may see the blurring of current demarcations between premarket and postmarket evaluation. Continued assessment for effectiveness, not just safety, in the postmarket setting using real-world data will enable continuous reevaluation of risk–benefit profiles and generate labeling changes and new indications. Premarket, real-world evidence may augment evidence from traditional RCTs, and increased use of pragmatic trials may improve external generalizability of results. Premarket and postmarket evaluations can form a feedback loop, enabling a rapid learning cycle.

Improved Public Health through More Widely Shared Real-World Data

Data from industry-sponsored clinical trials are increasingly available for secondary analyses by a wide range of users. Ironically, data from not-for-profit health care systems are typically not as freely shared. Many real-world questions (especially those involving cost and heterogeneity of treatment effects) require large samples only achievable with pooling of data across institutions. Data sharing and more open data access will certainly require appropriate protections for patients' privacy and health systems' proprietary information. Open access will also require researchers to set aside proprietary interests in order to facilitate more efficient learning.

True Integration between Research and Practice

Ultimately, the evidence necessary to guide policy decisions regarding new treatments overlaps substantially with the evidence necessary to guide everyday clinical decisions for individual patients. If we are to use data generated from everyday practice to create generalizable evidence, then we would need systematically recorded data regarding patients' risk factors or prognostic characteristics, providers' rationale for treatment choices, patients' actual treatment exposures, and a range of clinically relevant outcomes. More systematic collection and recording of those data would not only facilitate practice-based research but also significantly improve the quality and effectiveness of everyday health care.

Realizing this promise will necessitate an evolution to a true learning health system approach, as described by the IOM (IOM, 2013), where the traditional boundaries between clinical research and practice are blurred and where knowledge is generated as a by-product of each care experience.

Notes

1. For more information on the OHDSI program, see <http://www.ohdsi.org> (accessed September 23, 2016).
2. See <https://www.ncdr.com/WebNCDR/tvt/home> (accessed September 23, 2016).
3. For more information, see <https://clinicaltrials.gov/ct2/show/NCT01551758> (accessed September 23, 2016).

References

- Bakerly, N. D., A. Woodcock, J. P. New, J. M. Gibson, W. Wu, D. Leather, and J. Vestbo. 2015. The Salford Lung Study protocol: A pragmatic, randomised phase III real-world effectiveness trial in chronic obstructive pulmonary disease. *Respiratory Research* 16:101.
- Bipartisan Policy Center. 2016. *Using real-world evidence to accelerate safe and effective cures*. Washington, DC: Bipartisan Policy Center.
- Carroll, J. D., F. H. Edwards, D. Marinac-Dabic, R. G. Brindis, F. L. Grover, E. D. Peterson, E. M. Tuzcu, D. M. Shahian, J. S. Rumsfeld, C. M. Shewan, K. Hewitt, D. R. Holmes, Jr., and M. J. Mack. 2013. The STS-ACC transcatheter valve therapy national registry: A new partnership and infrastructure for the introduction and surveillance of medical devices and therapies. *Journal of the American College of Cardiology* 62(11):1026–1034.
- Carroll, J. D., J. Shuren, T. S. Jensen, J. Hernandez, D. Holmes, D. Marinac-Dabic, F. H. Edwards, B. D. Zuckerman, L. L. Wood, R. E. Kuntz, and M. J. Mack. 2015. Transcatheter valve therapy registry is a model for medical device innovation and surveillance. *Health Affairs (Millwood)* 34(2):328–334.
- FDA (Food and Drug Administration). 2016. *Use of real-world evidence to support regulatory decision-making for medical devices: Draft guidance for industry and Food and Drug Administration staff*. <http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf> (accessed October 12, 2016).
- FOCR (Friends of Cancer Research). 2016. *A blueprint for breakthrough: Exploring the utility of real world evidence (RWE)*. Paper presented at the 5th Annual Alexandria Summit and Friends of Cancer Research Blueprint for Breakthrough Forum, June 16, Washington, DC.
- Food and Drug Administration Amendments Act of 2007*, Public Law 110-85, 110th Cong., 1st sess. (September 27, 2007), §905, <https://www.gpo.gov/fdsys/pkg/PLAW-110publ85/pdf/PLAW-110publ85.pdf> (accessed September 23, 2016).
- Hripcsak, G., J. D. Duke, N. H. Shah, C. G. Reich, V. Huser, M. J. Schuemie, M. A. Suchard, R. W. Park, I. C. Wong, P. R. Rijnbeek, J. van der Lei, N. Pratt, G. N. Noren, Y. C. Li, P. E. Stang, D. Madigan, and P. B. Ryan. 2015. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for observational researchers. *Studies in Health Technology and Informatics* 216:574–578.
- Hripcsak, G., P. B. Ryan, J. D. Duke, N. H. Shah, R. W. Park, V. Huser, M. A. Suchard, M. J. Schuemie, F. J. DeFalco, A. Perotte, J. M. Banda, C. G. Reich, L. M. Schilling, M. E. Matheny, D. Meeker, N. Pratt, and D. Madigan. 2016. Characterizing treatment pathways at scale using the OHDSI network. *Proceedings of the National Academy of Sciences* 113(27):7329–7336.
- IOM (Institute of Medicine). 2006. *The future of drug safety: Promoting and protecting the health of the public*, edited by A. Baciuc, K. Stratton, and S. P. Burke. Washington, DC: The National Academies Press.
- IOM (Institute of Medicine). 2013. *Best care at lower cost: The path to continuously learning health care in America*, edited by M. Smith, R. Saunders, L. Stuckhardt, and J. M. McGinnis. Washington, DC: The National Academies Press.
- Kuehn, B. M. 2016. FDA's foray into big data still maturing. *Journal of the American Medical Association* 315(18):1934–1936.

- New, J. P., N. D. Bakerly, D. Leather, and A. Woodcock. 2014. Obtaining real-world evidence: The Salford Lung Study. *Thorax* 69(12):1152–1154.
- Rumsfeld, J. S., D. R. Holmes, Jr., W. G. Stough, F. H. Edwards, L. B. Jacques, and M. J. Mack. 2015. Insights from the early experience of the Society of Thoracic Surgeons/American College of Cardiology Transcatheter Valve Therapy Registry. *Journal of the American College of Cardiology Cardiovascular Interventions* 8(3):377–381.
- Sentinel Program Interim Assessment—FY15. 2015. <http://www.fda.gov/downloads/ForIndustry/UserFees/PrescriptionDrugUserFee/UCM464043.pdf> (accessed October 12, 2016).
- Society of Thoracic Surgeons. 2013. *FDA approval of device labeling highlights growing use of clinical registry data for post-market device surveillance: STS/ACC TVT Registry™ data cited as part of FDA decision process*. Press release, https://www.sts.org/sites/default/files/documents/pdf/pressreleases/STS_ACC%20TVT%20Registry_final.pdf (accessed September 23, 2016).
- Woodcock, A., N. D. Bakerly, J. P. New, J. M. Gibson, W. Wu, J. Vestbo, and D. Leather. 2015. The Salford Lung Study protocol: A pragmatic, randomised phase III real-world effectiveness trial in asthma. *BMC Pulmonary Medicine* 15:160.

Suggested Citation

Galson, S., and G. Simon. 2016. *Real-world evidence to guide approval and use of new treatments*. Discussion Paper, National Academy of Medicine, Washington, DC: <https://nam.edu/wp-content/uploads/2016/10/Real-World-Evidence-to-Guide-the-Approval-and-Use-of-New-Treatments.pdf>.

Author Information

Steven Galson is Senior Vice President, Global Regulatory Affairs and Safety, Amgen. **Gregory Simon** is Senior Investigator and Psychiatrist at Group Health Research Institute.

Disclaimer

The views expressed in this Perspective are those of the authors and not necessarily of the authors' organizations, the National Academy of Medicine (NAM), or the National Academies of Sciences, Engineering, and Medicine (The Academies). The Perspective is intended to help inform and stimulate discussion. It has not been subjected to the review procedures of, nor is it a report of, the NAM or the Academies. Copyright by the National Academy of Sciences. All rights reserved.