

SOUNDING BOARD

Real-World Evidence — What Is It and What Can It Tell Us?

Rachel E. Sherman, M.D., M.P.H., Steven A. Anderson, Ph.D., M.P.P.,
 Gerald J. Dal Pan, M.D., M.H.S., Gerry W. Gray, Ph.D., Thomas Gross, M.D., M.P.H.,
 Nina L. Hunter, Ph.D., Lisa LaVange, Ph.D., Danica Marinac-Dabic, M.D., Ph.D.,
 Peter W. Marks, M.D., Ph.D., Melissa A. Robb, B.S.N., M.S., Jeffrey Shuren, M.D., J.D.,
 Robert Temple, M.D., Janet Woodcock, M.D., Lilly Q. Yue, Ph.D., and Robert M. Califf, M.D.

The term “real-world evidence” is widely used by those who develop medical products or who study, deliver, or pay for health care, but its specific meaning is elusive. We believe it refers to information on health care that is derived from multiple sources outside typical clinical research settings, including electronic health records (EHRs), claims and billing data, product and disease registries, and data gathered through personal devices and health applications.^{1,2} Key to understanding the usefulness of real-world evidence is an appreciation of its potential for complementing the knowledge gained from traditional clinical trials, whose well-known limitations make it difficult to generalize findings to larger, more inclusive populations of patients, providers, and health care delivery systems or settings that reflect actual use in practice.³

Real-world evidence can inform therapeutic development, outcomes research, patient care, research on health care systems, quality improvement, safety surveillance, and well-controlled effectiveness studies. Real-world evidence can also provide information on how factors such as clinical setting and provider and health-system characteristics influence treatment effects and outcomes. Importantly, the use of such evidence has the potential to allow researchers to answer these questions efficiently, saving time and money while yielding answers relevant to broader populations of patients than would be possible in a specialized research environment.^{4,5}

As defined above, real-world evidence can be viewed as a means of incorporating diverse types of evidence into information on health care. However, the confluence of large data sets of uncertain quality and provenance, the facile analytic tools that can be used by nonexperts, and a

shortage of researchers with adequate methodologic savvy could result in poorly conceived study and analytic designs that generate incorrect or unreliable conclusions. Accordingly, if we are to realize the full promise of such evidence, we must be clear about what it is and how it can be used most effectively, and we must have appropriate expectations about what it can tell us. It is important to distinguish two key dimensions of real-world evidence. The first is the setting in which evidence is generated, which includes the population defined by the data source as well as the specific methods used to collect and curate the data on that population. The second is the methodologic approach used to conduct the surveillance or research.

 RESEARCH SETTINGS — TRADITIONAL TRIALS VS. REAL WORLD

“Traditional” clinical trials are often conducted with specific populations and in specialized environments that differ from the realities of clinical or home settings. These trials may take measures designed to control variability and to ensure the quality of the data they generate, such as the development of long lists of eligibility criteria, the use of detailed case-report forms that exist separately from ordinary medical records, and the use of intensive monitoring and specialized research personnel to ensure adherence to a well-characterized protocol that defines study procedures and ensures precision in data collection.

The clinical trial unquestionably remains a powerful tool for developing scientific evidence about the safety and efficacy of a medical product while informing our understanding of the

biologic mechanisms involved in its therapeutic action. These trials are often needed because they are designed to provide an essential element of the premarket evaluation of a medical product — namely, robust evidence that a treatment may “work.” However, the internal validity attained in these trials is often achieved at the expense of uncertainty about generalizability, especially since the populations enrolled in such studies may differ in significant ways from those seen in practice. In addition, there may be few data on interactions with concomitant illnesses and treatment, and adherence to therapies may be supported by intensive efforts that are infeasible in practice. Moreover, the expense of conducting large traditional trials has been growing steadily for years,⁶ and recent estimates suggest that the cost trajectory may be steepening,⁷ without any indication of a commensurate increase in the quantity of evidence produced to support decisions about health care.

Given these trends, many trialists, clinical researchers, and medical-product developers have become increasingly interested in expanding and integrating clinical research into more diverse, real-world settings by capitalizing on the exponential growth in access to data from EHRs, claims databases, electronic devices and software applications (or apps), registries embedded in clinical practice, and social media. These sources can provide new insight into states of health and illness. For instance, EHRs, registries, and claims databases contain rich data that are already being gathered in real-world settings at the point of care, personal devices and apps allow continuous monitoring and data capture⁸ and facilitate shared decision making,⁹ and data from social media can be used for epidemiologic purposes.¹⁰ But these data sources also raise concerns. EHR and claims data are not collected or organized with the goal of supporting research, nor have they typically been optimized for such purposes, and the accuracy and reliability of data gathered by many personal devices and health-related apps are unknown.^{11,12} Furthermore, the use of any of these sources, including social media, raises important questions about the quality of the data they provide and about privacy.

The technological and methodologic challenges presented by these new data sources are the focus of active efforts by researchers. For example, multiple stakeholders, including the Food

and Drug Administration (FDA), are working on ways to harmonize data collected from EHRs, claims data, and registries to create a unified system for monitoring the safety and effectiveness of medical devices.^{13,14} Others, such as the National Institutes of Health (NIH) Collaboratory (an NIH Common Fund initiative devoted to building infrastructure, operational knowledge, and capacity for pragmatic research in the context of health care systems),¹⁵ are developing and implementing methods for incorporating data from EHRs and other sources into research. Such efforts include the development of large-scale distributed research networks¹⁶ and “computable phenotypes” (i.e., conditions or patient characteristics that can be derived from EHRs and claims data without requiring external review or interpretation¹⁷) that allow researchers to identify cohorts of interest across multiple data sources.¹⁸

RESEARCH METHODS, TREATMENT ALLOCATION, AND THE DEFINITION OF REAL-WORLD EVIDENCE

We believe that real-world evidence can be used across a wide spectrum of research, ranging from observational studies to studies that incorporate planned interventions, whether with or without randomization at the point of care. At the same time, however, it is incorrect to contrast the term “real-world evidence” with the use of randomization in a manner that implies that they are disparate or even incompatible concepts.

As we adapt the tools and methods of traditional trials to real-world settings, we must consider the components of such trials that are critical to obtaining valid results and minimizing bias.¹⁹ Although real-world evidence can be used in multiple research scenarios, the selection of appropriate analytic approaches will be determined by key dimensions of the study design, including the use of prospectively planned interventions and randomization.²⁰ Planned interventions, whether randomized or not, can be used in both the tertiary care and academic environments, where much clinical research is typically performed in association with intensive support and expensive resources. These interventions can also be used in “real-world” settings with less labor-intensive clinical research support and possibly a lesser degree of familiarity with clini-

cal research. For this reason, discussions of real-world evidence must be informed by a clear understanding of the methods used, so that the best methods that have been developed and validated can be combined with the most appropriate research settings.

In traditional trials, randomization has long been an essential tool for minimizing bias by balancing underlying risk between treatment groups, but it can be just as useful and important in real-world studies. In fact, one of the first major randomized, controlled trials (RCTs) conducted in a real-world setting was the Salk field trial of the polio vaccine, which combined a large component comprising 750,000 children who were randomly assigned to receive vaccine or placebo (control group) with an even larger non-randomized “observed control” group of 1 million children, all of whom received the vaccine.²¹ A contemporary version of a large, simple trial performed in a real-world setting is the “Aspirin Study,” also known as ADAPTABLE (Aspirin Dosing: A Patient-centric Trial Assessing Benefits and Long-Term Effectiveness), which is being conducted by the National Patient-Centered Clinical Research Network. In this trial, 20,000 participants are being randomly assigned to one of two commonly used doses of aspirin in order to ascertain which of these two dose regimens is better for the secondary prevention of cardiovascular disease.²² There is extensive literature on pragmatic RCTs designed to inform decision making at the individual and the population level.²³ Many of the NIH Collaboratory’s demonstration projects involved innovative pilot approaches to performing pragmatic research within health systems.²⁴ Cluster randomization, which is particularly useful for evaluating interventions at the level of health systems, practices, or hospitals, was used for most of these projects.²⁵

In addition to its application in interventional studies, real-world evidence is also valuable in observational settings, where it is used to generate hypotheses for prospective trials,²⁶ assess the generalizability of findings from interventional trials (including RCTs),²⁷ conduct safety surveillance of medical products,²⁸ examine changes in patterns of therapeutic use, and measure and implement quality in health care delivery.¹ However, much of the current excitement about real-world evidence stems from the hope that access to sources of emerging data of adequate quality

will, when paired with the development of more robust methods, allow greater use of observational treatment comparisons in drawing causal inferences about the treatment effects of medical products.

Although observational studies are an essential tool for clinical epidemiologic investigations, quality improvement, and safety surveillance, their findings require judicious evaluation when used to assess treatment effects.²⁹ These limitations are particularly problematic when an observational study is used to evaluate the effectiveness of a medical product and the expected or observed effect is relatively small. When this is the case, it can be difficult to be confident that the effect is not due largely or wholly to confounding factors. This problem, compounded by the fact that observational studies often leverage existing rather than prospectively collected data (e.g., as part of a disease or product registry with well-established quality standards), can add to the uncertainty regarding findings and limit the usefulness of such data.

Awareness of the limitations of source data and analytic approaches¹ is fueling concern that when the term “real-world evidence” is used in such contexts, the allure of analyzing existing data may lead to flawed conclusions. This concern is especially salient in light of the growing proliferation of precision molecular medicine and treatments for rare diseases, many of which are anticipated to undergo review in accelerated approval programs. In such circumstances, real-world evidence will become an increasingly critical element in expediting the availability of data needed to confirm clinical benefit and value, because products will necessarily receive initial approval in an atmosphere in which there is greater uncertainty with regard to clinical outcomes. Although access to real-world data adds important dimensions to the assessment of therapies and important progress is being made in the methodologic arena,³⁰ these factors do not yet suffice to fully overcome the fundamental issues of confounding, data quality, and bias,³¹ unless other, specific countervailing features of the evaluation are relevant.

For example, prospective registries or single-group trials with planned external controls and high-quality data collection have been accepted for regulatory purposes in the evaluation of medical devices (e.g., a ventricular-assist system

that used propensity-score–matched controls from the Interagency Registry for Mechanically Assisted Circulatory Support³²). However, because medical devices are typically developed in an iterative fashion, building on earlier designs and incorporating refinements throughout the product life cycle, substantial knowledge of the effect of confounding factors is often available a priori. This availability in turn facilitates the evaluation of observed treatment effects, as exemplified by the use of data from the Transcatheter Valve Therapy Registry for postmarketing regulatory purposes, including labeling revisions.³³

Thus, although we are optimistic about long-term prospects for the evolution of mature, robust methodologic approaches to the incorporation of real-world evidence into therapeutic development and evaluation given the intensive efforts now under way, caution is still needed, and expectations of “quick wins” resulting from the use of such evidence should be tempered accordingly. Specifically, other analytic methodologies with varying levels of evidentiary requirements, such as historical controls or study designs with an open-label phase in which all patients receive the investigational product, fall within the spectrum of potentially useful approaches that will require careful consideration before they can be appropriately applied to answer important questions about the effects of treatment with medical products in real-world settings, including issues involving latent or rare outcomes and treatments for rare diseases.

To this end, the FDA is committed to robust policy development under the proposed reauthorization of the Prescription Drug User Fee Act VI (user-fee program) for drugs and biologic products.³⁴ This commitment includes convening public workshops involving participants on all bands of the research spectrum — from patients to providers to sponsors — to gather input on the use of real-world evidence in regulatory decision making. With this information, the agency will initiate activities to address key concerns and publish draft guidance on how such evidence can be used to assess safety and effectiveness in both premarketing and postmarketing regulatory requirements. Complementary efforts are included in Medical Device User Fee Amendments IV for devices.³⁵

CONCLUSIONS

We believe that when the term “real-world evidence” is used, the primary attribute that distinguishes it from other kinds of evidence is related to the context in which the evidence is gathered — in other words, in clinical care and home or community settings as opposed to research-intensive or academic environments. Most important, the distinction should not be based on the presence or absence of a planned intervention or the use of randomization. Real-world research and the concepts of a planned intervention and randomization are entirely compatible. Indeed, one of the most important advances in clinical trial methodology may be the broadening of the application of randomization outside more typical venues for clinical trials, such as academic research centers. But in order to gain collective confidence in the appropriate uses of this array of methods across disparate settings, we must first be clear about our terminology and its application.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Jonathan McCall, M.S. (Duke Clinical Research Institute, Durham, NC), for editorial assistance with an earlier version of the manuscript.

From the Office of the Commissioner (R.E.S., N.L.H., R.M.C.); the Office of Biostatistics and Epidemiology (S.A.A.) and the Office of the Center Director (P.W.M.), Center for Biologics Evaluation and Research; the Office of Surveillance and Epidemiology (G.J.D.P.), the Office of Biostatistics, the Office of Translational Sciences (L.L.), the Office of Medical Policy (M.A.R.), the Office of the Center Director (R.T., J.W.), Center for Drug Evaluation and Research, and the Division of Biostatistics (G.W.G., L.Q.Y.); and the Division of Epidemiology (D.M.-D.), Office of Surveillance and Biometrics (T.G.); and the Office of the Center Director (J.S.), Center for Devices and Radiological Health — all at the Food and Drug Administration, Silver Spring, MD. Address reprint requests to Ms. Robb at the Office of Medical Policy, Center for Drug Evaluation and Research, Food and Drug Administration, 10903 New Hampshire Ave., White Oak, Bldg. 51, Rm. 6346, Silver Spring, MD 20993, or at melissa.robb@fda.hhs.gov.

1. Network for Excellence in Health Innovation. Real world evidence: a new era for health care innovation. September 2015 (http://www.nehi.net/writable/publication_files/file/rwe_issue_brief_final.pdf).
2. Food and Drug Administration. Use of real-world evidence to support regulatory decision-making for medical devices: draft guidance for industry and Food and Drug Administration staff. July 27, 2016 (<http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM513027.pdf>).
3. Booth CM, Tannock IF. Randomised controlled trials and

- population-based observational research: partners in the evolution of medical evidence. *Br J Cancer* 2014;110:551-5.
4. Cziraky M, Pollock M. Real-world evidence studies. *Applied Clinical Trials*. October 12, 2015 (<http://www.appliedclinicaltrials.com/real-world-evidence-studies?pageID=1>).
 5. Cattell J, Chilukuri S, Levy M. How big data can revolutionize pharmaceutical R&D. McKinsey & Company, April 2013 (http://www.mckinsey.com/insights/health_systems_and_services/how_big_data_can_revolutionize_pharmaceutical_r_and_d).
 6. Kramer JM, Schulman KA. Appendix F — discussion paper: transforming the economics of clinical trials. In: Institute of Medicine. *Envisioning a transformed clinical trials enterprise in the United States: establishing an agenda for 2020: workshop summary*. Washington, DC: National Academies Press, 2012 (<http://www.ncbi.nlm.nih.gov/books/NBK114653/>).
 7. Berndt ER, Cockburn IM. Price indexes for clinical trial research: a feasibility study. Bureau of Labor Statistics Monthly Labor Review. June 2014 (<http://www.bls.gov/opub/mlr/2014/article/price-indexes-for-clinical-trial-research-a-feasibility-study-1.htm>).
 8. Zeevi D, Korem T, Zmora N, et al. Personalized nutrition by prediction of glycemic responses. *Cell* 2015;163:1079-94.
 9. Mora S, Ames JM, Manson JE. Low-dose aspirin in the primary prevention of cardiovascular disease: Shared decision making in clinical practice. *JAMA* 2016;316:709-10.
 10. Salathé M, Bengtsson L, Bodnar TJ, et al. Digital epidemiology. *PLoS Comput Biol* 2012;8(7):e1002616.
 11. Kuehn BM. Is there an app to solve app overload? *JAMA* 2015;313:1405-7.
 12. Plante TB, Urrea B, MacFarlane ZT, et al. Validation of the Instant Blood Pressure smartphone app. *JAMA Intern Med* 2016; 176:700-2.
 13. Food and Drug Administration. National Evaluation System for Health Technology (NEST). (<http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/ucm301912.htm>).
 14. Shuren J, Califf RM. Need for a national evaluation system for health technology. *JAMA* 2016;316:1153-4.
 15. National Institutes of Health Health Care Systems Research Collaboratory. About us (<https://www.nihcollaboratory.org/about-us/Pages/default.aspx>).
 16. National Institutes of Health Health Care Systems Research Collaboratory. Electronic health records (<https://www.nihcollaboratory.org/cores/Pages/ehr.aspx>).
 17. Richesson R, Smerek M. Electronic health records-based phenotyping. In: *Rethinking clinical trials: a living textbook of pragmatic clinical trials*. June 27, 2014 (<https://sites.duke.edu/rethinkingclinicaltrials/ehr-phenotyping/>).
 18. National Institutes of Health Health Care Systems Research Collaboratory. Phenotypes, data standards, and data quality (<https://www.nihcollaboratory.org/cores/Pages/phenotypes.aspx>).
 19. Peto R, Collins R, Gray R. Large-scale randomized evidence: large, simple trials and overviews of trials. *J Clin Epidemiol* 1995; 48:23-40.
 20. ClinicalTrials.gov. What is a clinical study? (<https://clinicaltrials.gov/ct2/about-studies/learn#WhatIs>).
 21. Meldrum M. "A calculated risk": the Salk polio vaccine field trials of 1954. *BMJ* 1998;317:1233-6.
 22. National Patient-Centered Clinical Research Network. ADAPTABLE, the aspirin study — a patient-centered trial (<http://www.pcornet.org/aspirin/>).
 23. Califf RM, Sugarman J. Exploring the ethical and regulatory issues in pragmatic clinical trials. *Clin Trials* 2015;12:436-41.
 24. National Institutes of Health Health Care Systems Research Collaboratory. Demonstration projects (<https://www.nihcollaboratory.org/demonstration-projects/Pages/default.aspx>).
 25. Anderson ML, Califf RM, Sugarman J. Ethical and regulatory issues of pragmatic cluster randomized trials in contemporary health systems. *Clin Trials* 2015;12:276-86.
 26. Gradman AH, Vekeman F, Eldar-Lissai A, Trahey A, Ong SH, Duh MS. Is addition of vasodilators to loop diuretics of value in the care of hospitalized acute heart failure patients? Real-world evidence from a retrospective analysis of a large United States hospital database. *J Card Fail* 2014;20:853-63.
 27. Tavis DR, Dey S, Albrecht-Gallauresi B, et al. Risk of local adverse events following cardiac catheterization by hemostasis device use — phase II. *J Invasive Cardiol* 2005;17:644-50.
 28. Food and Drug Administration. Pradaxa (dabigatran etexilate mesylate): drug safety communication — safety review of post-market reports of serious bleeding events (<http://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm282820.htm>).
 29. Ioannidis JP, Tarone R, McLaughlin JK. The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology* 2011; 22:450-6.
 30. Califf RM. Pragmatic clinical trials: Emerging challenges and new roles for statisticians. *Clin Trials* 2016;13:471-7.
 31. Pearl J. *Causality: models, reasoning, and inference*. 2nd ed. New York: Cambridge University Press, 2009:350.
 32. Food and Drug Administration. Summary of safety and effectiveness data: HeartWare ventricular assist device. Section X. Summary of primary clinical study (http://www.accessdata.fda.gov/cdrh_docs/pdf10/P100047b.pdf).
 33. Food and Drug Administration. FDA news release: FDA approval expands access to artificial heart valve for inoperable patients. September 23, 2013 (<http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm369510.htm>).
 34. Food and Drug Administration. PDUFA reauthorization performance goals and procedures fiscal years 2018 through 2022 (<http://www.fda.gov/downloads/ForIndustry/UserFees/PrescriptionDrugUserFee/UCM511438.pdf>).
 35. Food and Drug Administration. FDA — industry MDUFA IV reauthorization meeting. August 15, 2016 (<http://www.fda.gov/downloads/ForIndustry/UserFees/MedicalDeviceUserFee/UCM518203.pdf>).

DOI: 10.1056/NEJMs1609216

Copyright © 2016 Massachusetts Medical Society.

RECEIVE THE JOURNAL'S TABLE OF CONTENTS EACH WEEK BY E-MAIL

To receive the table of contents of the *Journal* by e-mail every Wednesday evening, sign up at NEJM.org.