Duke-Margolis Center for Health Policy

Enhancing the Application of Real-World Evidence in Regulatory Decision-Making

March 3-4, 2016

Discussion Guide

Introduction

The nation's growing electronic health information infrastructure has enabled routine and increasingly robust collection of digital data at the point of patient care. These data in turn are facilitating the development of an expanding body of real world evidence (RWE) on the use of medical products – evidence that is used increasingly by providers, payers, patients and other stakeholders to improve treatment decisions, enhance care quality, and support innovative payment and delivery models. These applications of RWE are important and complementary to the existing body of clinical evidence related to a medical product at the time of its approval by the U.S. Food and Drug Administration (FDA). With this growing capacity to collect and access richer sources of data, there are unprecedented opportunities in the near future to advance understanding on how a variety of medical products perform in the real world and on the course of underlying diseases they are designed to treat.

There is great potential for RWE to improve clinical research and evidence development. Presently, evidence generation can be costly and often is not available or not relevant to clinical decision making during routine care delivery. Reviews of clinical practice guidelines, for example, point to the prevalence of evidence gaps in routine care delivery.[1] It is thought RWE could improve the existing evidence base by enhancing the efficiencies of studies and gaining a better understanding of the clinical, economic, and humanistic outcomes associated with medical products across a range of practice settings and diverse patient populations.

A number of different stakeholder groups have already made headway in leveraging RWE to generate knowledge and improve decision-making; however, its application in the regulatory setting has been relatively less-well-explored. While RWE has been used to support drug development and regulatory review in rare diseases and is routinely used in the assessment of safety (pre- and postmarket safety surveillance), there are a number of other potential applications of RWE within the existing regulatory framework. Indeed, enhanced use of RWE by FDA to support the agency's decision-making capabilities has been the animating goal of a number of recent proposals found in the U.S. House of Representatives' 21st Century Cures Act and the sixth renegotiation of 2017's Prescription Drug User Fee Act (PDUFA).

This public conference, convened by the Duke-Margolis Center for Health Policy under a cooperative agreement with FDA, is intended to support broader discussion around clarifying and expanding how and when RWE might be used to support regulatory science and decision-making. Throughout the conference, experts and participants will highlight promising regulatory use cases for applying RWE, the study designs and methods that could be used to support development of RWE fit for such use, and the broader infrastructure and sustainability challenges that impact evidence development and use across all health care stakeholder needs. The conference and related Duke-convened RWE activities will inform a number of downstream papers and other products focused on these important issues.

Defining Real-World Evidence

Data collected from sources outside of conventional randomized controlled trials – for example, from electronic systems used in health care delivery and to track patient experience with care – are commonly referred to as real-world data (RWD) while the evidence derived from such data elements are commonly referred to as real-world evidence (RWE).[2] While stakeholders seem to agree on these general terms, meaningful discussion about what is actually meant and included within the scope of these terms and the most promising opportunities for regulators to use RWE is needed. One approach to further the stakeholder community's common understanding of RWE is to identify and describe the multiple components and processes entailed with generating RWE from data and how it can fulfill evidentiary needs. Figure 1 provides a graphical depiction of the evidence spectrum, which ranges from traditional randomized-controlled clinical trials (RCTs) on one end and anecdotal case reports on the other, and illustrates the growing RWE landscape as part of this spectrum. Data collected in routine care settings is being leveraged by numerous stakeholder groups using a range of methodological approaches to generate RWE.
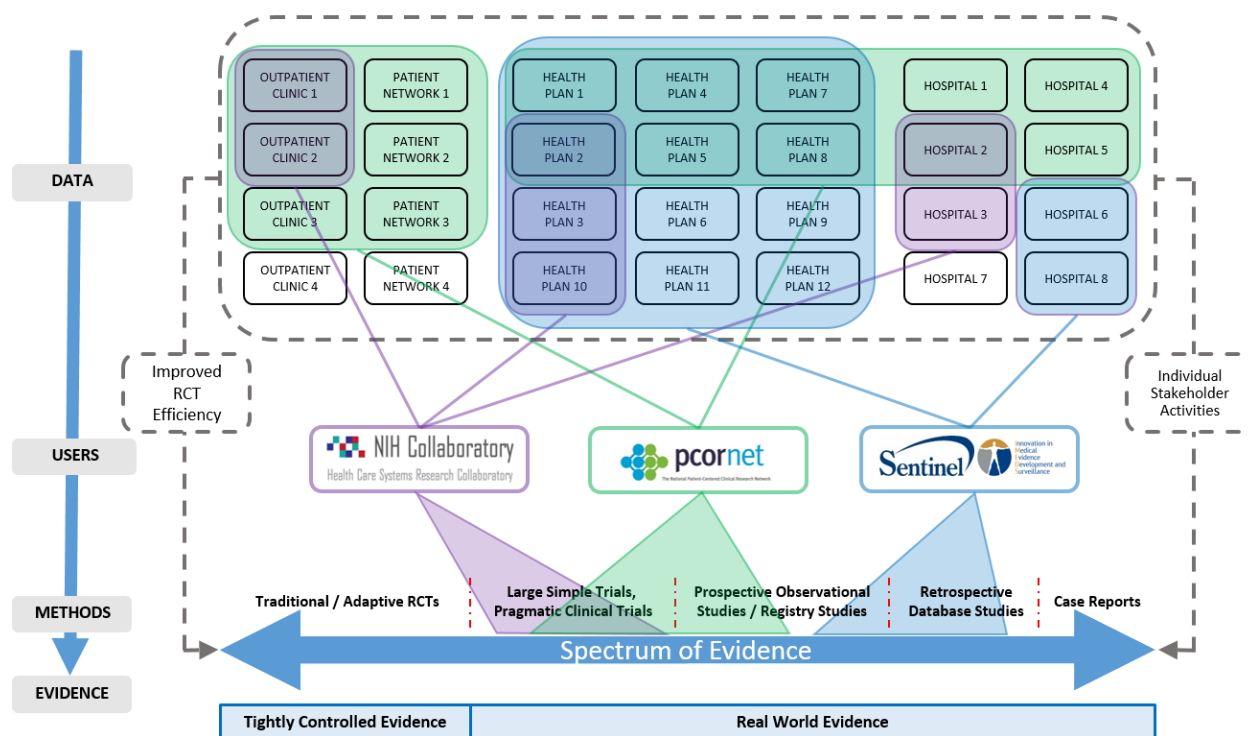

Figure 1: RWE and the Spectrum of Evidence

Real world data represents the first key component of developing RWE. There are many sources of RWD as represented in this figure. These sources include routinely collected electronic clinical, administrative, and economic data captured by health systems, individual clinics, payers, and patients through registries, electronic health records, diagnostic and laboratory results, billing claims, health plan enrollment files, and pharmacy dispensing data. Additionally, as the result of emerging technologies such as wearable devices and the widespread use of social media networks and patient communities, new streams of data

on health and patient experience are being produced and becoming more accessible for research purposes. With growing amounts of data increasingly available for a variety of secondary uses, innovations in technology such as distributed database architectures ensure that personal health information remains private, confidential and protected. As a result, there is a growing ecosystem of RWD, which is providing new opportunities to support a range of methodological approaches to generate evidence.

Many examples of individual organizations and researchers generating RWE from these data exist today, as well as multi-organizational collaborations or "systems" aimed at more efficiently developing such evidence from a broad array of settings and data sources. While payers, health systems and other data stakeholders are using data collected within their systems for a variety of purposes, in recent years there has been increasing development of collaborative multi-organizational data networks that enable more efficient generation of evidence from larger and more diverse patient populations and clinical settings through a common data infrastructure. Examples of developing data networks and their evidence aims include (but are not limited to):

- The FDA's **Sentinel System**, which actively tracks the performance of approved medical products (e.g. drugs) once they reach the market through querying data from diverse data sources, including electronic health record systems, health insurance claims databases, and electronic laboratory results—to evaluate possible medical product safety issues quickly and securely.
- The **Innovation in Medical Evidence Development and Surveillance (IMEDS) Evaluation Program** was designed around two key components: a methods component to facilitate methods research aimed at monitoring safety of marketed medical products; and an evaluation component to provide opportunities for non-FDA stakeholders to conduct active surveillance and other medical product safety evaluations using the same tools, methods, and data infrastructure as Sentinel.
- The Patient-Centered Outcomes Research Institute (PCORI) **National Patient-Centered Clinical Research Network** (PCORnet) has been designed to make it faster, easier, and less costly to conduct clinical research by harnessing the power of large amounts of health data with a particular emphasis on patient partnership.
- The **National Institutes of Health (NIH) Collaboratory** shares similar aims with PCORnet to conduct streamlined clinical research, but with a specific focus on developing and implementing demonstrations leveraging pragmatic study designs.
- **The National Medical Device Evaluation System,** FDA's developing medical device surveillance network, is being designed to utilize RWE comprised of electronic health information such as electronic health records (EHRs), registries, and health insurance claims in which device identifiers (such as the Unique Device Identifier) will be incorporated. The aims of the network are to enable systematic assessments of the benefits and risks of medical devices throughout the product life cycle, reduce the burden and cost of postmarket surveillance, as well as to facilitate the clearance and approval of novel devices and new uses for existing devices. The system would leverage and integrate with other national resources such as PCORnet, Sentinel, and national and international registries.
- A variety of **private sector-based data networks** are being developed by firms such as Optum and HealthCore that integrate varied sources of data, including claims linked to clinical detail from providers to support analytics for improving care delivery.

These data networks employ a range of methodological approaches to generate RWE to support, complement, and improve traditional evidence generation. To improve upon or support well-controlled

RCTs and observational studies, stakeholders are developing a range of methodologies that can be more flexibly designed while still incorporating elements of randomization. Examples include large simple trials, which are characterized as having large sample sizes and relatively simple approaches to randomization and data collection, to study designs that involve flexible eligibility criteria, cluster randomization (e.g., at the health care provider or plan level), and patient differences in use of the intervention(s).[3,4] These designs are generally considered to have "pragmatic" aims because they incorporate randomization in routine clinical care settings and rely on data collected during patient care. Such approaches can be considered a hybrid between traditional randomized trials and observational studies, producing results that can be generalized to an extended population and clinical setting while minimizing biases.[5]

There are many varieties of these designs, which have been extensively detailed in the literature that may be more or less well suited to a particular question or evidence need. The starting point for determining a suitable design depends on the pragmatic intention of the study. The PRECIS Wheel is a tool to help trialists make design decisions consistent with the intended purpose of their trial and specifies nine key domains of pragmatic considerations, which are provided in Appendix A.[6] While each method has its own advantages and limitations, there is increasing interest to understand how clinical research could incorporate pragmatic features and how they could be used in randomized trials more widely. Appendix B provides a few examples of trials that are being designed or implemented with such features.

Opportunities to Incorporate RWE into Regulatory Decision Making

Regulators at FDA traditionally assess clinical benefit and risk of medical products using an evidence hierarchy that consists of interventional clinical trials and non-interventional study designs. Through this approach, the FDA often relies on clinical trials to provide the basis for approval decisions while using non-interventional studies to assess larger and more longitudinal data sets – often derived from real-world settings – to review the product's postmarket safety and effectiveness. This approach acknowledges that real-world use of medical products introduces a range of complicating factors, such as co-morbidities or patient adherence, which might impact the intervention's effect in the indicated population compared to the outcomes observed in a clinical trial. This is especially true for disease areas in which a high percentage of patient use occurs in off-label settings that were not well-studied in premarket trials.[7]

As data on outcomes associated with different patient subgroups, uses outside the indicated populations, or in different clinical settings continue to accumulate, identifying how such data could provide evidence to support additional regulatory decisions is needed. Doing so would provide a clearer path to continuous learning about products once on the market and would help ensure that such evidence is used not only by FDA to support their growing evidentiary needs, but also by patients, providers, and payers to support treatment decisions.

There are a range of opportunities for exploring the potential application of RWE within regulatory science, starting with the use of RWD elements to improve the efficiency of RCTs. The use of RWD from EHRs, registries, and claims data can support more efficient and targeted patient recruitment, centralized consent management, or reduced burden of data collection and reporting could have a direct impact on lowering the significant cost and time associated with clinical development. While not necessarily directly supporting regulatory decision-making, such practices will likely need to be shaped with input or guidance from FDA on acceptable applications as any resultant trial data will need to be fit for regulatory submission.

More squarely within FDA's regulatory framework are opportunities to harness RWE to support decisions related to the approval of new indications or label revisions based on data from clinical experience in broader patient populations, as well as completion of postmarket commitments or Phase IV confirmatory trials. Of particular interest for this public conference is the use of RWE in approving a new indication for an already marketed product: through two sessions, experts will begin to outline criteria for when sponsors and FDA might consider pursuing such an RWE development strategy and the potential methods and designs that might be most appropriate for generating the desired data. A conference session will also explore how RWE generated from more purely observational studies and methods could potentially be used to support a range of FDA decisions.

Realizing these potential uses of RWE in regulatory decision making will require continued work on methods and data standards, as well as a potential framework for matching certain decisions with the methods or designs most well-suited for answering the regulatory question of interest. For example, randomization designs employed in clinical settings might be most appropriate for teasing out subpopulation effects within a targeted population. Observational designs may be appropriate in certain cases where randomization may be unethical or infeasible, when effect sizes of the intervention are very large, or for hypothesis generation. A wide array of stakeholders will need to be engaged in discussion around these topics in order to make meaningful progress toward a regulatory framework to optimally develop and make full use of such practices and the RWE they can generate.

Moving towards a 21st Century evidence development system

Essential to realizing the potential of RWE is the need to develop a sustainable infrastructure to support not only regulatory decision making of medical products, but that could support the evidentiary needs of a variety of stakeholders including patients and providers making clinical care decisions and health plans and purchasers of health services. This is largely in line with the continued development of what many refer to as "the learning health care system."[8] To garner broad support for such a system, a viable business case is needed to mobilize partners to develop the necessary infrastructure and sustain the system's maturation and growth. This will necessitate an environment in which RWE development is not just a potentially interesting option for exploring health care questions, but is truly a more cost-effective, efficient, and thus widely-used mechanism for pursuing or supporting a wide range of important evidence gaps. Key questions that will need to be addressed include:

- How can improvement to the connectivity and scalability of research and data collection methods help to achieve these ends?
- Are there specific policy levers or incentives that could bolster infrastructure support and improvements?
- What additional incentives are needed to encourage generation of more robust and reliable RWE in the post-market?

To ensure that further steps involving RWE maximize the development of meaningful evidence, it is critical to engage providers and patients, each playing a critical role in data collection and use of the evidence. This will require providers to be actively engaged in all stages of system development and evidence generation and patients are not only informed of the system's activities, but serve as key data partners. The importance of patient-centeredness cannot simply be at face-value, but woven into the core of the system's identity and value.

Appendix A: Precis-2 Domains

1. Eligibility—To what extent are the participants in the trial similar to those who would receive this intervention if it was part of usual care?

2. Recruitment—How much extra effort is made to recruit participants over and above what would be used in the usual care setting to engage with patients?

3. Setting—How different are the settings of the trial from the usual care setting?

4. Organization—How different are the resources, provider expertise, and the organization of care delivery in the intervention arm of the trial from those available in usual care?

5. Flexibility (delivery)—How different is the flexibility in how the intervention is delivered and the flexibility anticipated in usual care?

6. Flexibility (adherence)—How different is the flexibility in how participants are monitored and encouraged to adhere to the intervention from the flexibility anticipated in usual care?

7. Follow-up—How different is the intensity of measurement and follow-up of participants in the trial from the typical follow-up in usual care?

8. Primary outcome—To what extent is the trial's primary outcome directly relevant to participants?

9. Primary analysis—To what extent are all data included in the analysis of the primary outcome?

Adapted from: Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K. E., & Zwarenstein, M. (2015). The PRECIS-2 tool: designing trials that are fit for purpose. BMJ, 350. doi:10.1136/bmj.h2147.

Appendix B: Pragmatic features of trials being planned or implemented

Project Title	Research Question	Example Pragmatic Features of Studies
Salford Lung Study (SLS)[9]	Is a once-daily combination of ICS fluticasone furoate (FF) and novel LABA vilanterol (VI) (Relvar®) in a dry-powder inhaler (Ellipta®) safer and more effective for the treatment of chronic obstructive pulmonary disorder (COPD) compared to existing twice-daily ICS/LABA combinations, the current standard of care?	Study Size and Setting: • Approx. 3,000 people with (COPD) living in Salford and the surrounding area have been enrolled in the study. • Study conducted in primary and secondary care settings in Salford, United Kingdom Study Design: • Minimal exclusion criteria: Patients with chronic obstructive pulmonary disease (COPD), ≥40 years old, with exacerbation in the previous 3 years are randomized 1:1 to once-daily FF 100 μg/VI 25 μg in a novel dry-powder inhaler (Ellipta®) versus continuing their existing therapy. • GPs may make treatment adjustments according to their clinical opinion. • The primary endpoint is mean annual rate of COPD exacerbations; an electronic medical record allows real-time collection and monitoring of endpoint and safety data.
Time to Reduce Mortality in End-Stage Renal Disease (TiME)[10]	Does systematically implementing a hemodialysis session duration of at least 4.25 hours improve survival, reduce hospitalizations and improve quality of life for patients with end-stage kidney disease?	Study Size and Setting: • Cluster-randomized at facility level to either: Intervention arm of recommended dialysis session durations of at least 4.25 hours for all patients initiating hemodialysis treatment regardless of body size or dialysis solute clearance measurements, or control arm (usual care) of no trial-driven approach to session duration. Study Design: • Approx. 6,500 patients with end stage renal disease treated by thrice weekly maintenance hemodialysis at two large dialysis provider organizations.
A Pragmatic Trial of Lumbar Image Reporting with Epidemiology (LIRE)[11]	Does adding epidemiologic benchmark data to spine imaging reports decrease subsequent back-related healthcare utilization?	Study Design: • Cluster randomized trial comparing typical imaging reports to those that include benchmarks prevalence data of findings in patients without back pain. Study Size and Setting: • Approx. 150,000 adults for whom a primary care provider has requested imaging of the lumbar spine • Study conducted at primary care clinics within the Kaiser Permanente-Northern California, Group Health Cooperative, Mayo Clinic Health System, and Henry Ford Health System
Influenza Vaccine to Effectively Stop CardioThoracic Events and Decompensated heart failure (INVESTED) Trial	Does a higher dose of influenza vaccine compared with standard does vaccine reduce cardiopulmonary events in a high-risk cardiovascular population?	Study Size and Setting: • Approx. 9,300 patients will enroll in the study across 180 sites in North America through four networks: University of Toronto-based Pan-Canadian Network, University of Wisconsin-Madison Network, Veterans Administration Consortium, and PCORnet. Study Design: • Large, "simple", adequately powered, double-blind comparative multicenter trial to assess whether high-dose influenza vaccine compared with standard dose vaccine will reduce cardiopulmonary events in a high-risk cardiovascular population • Endpoints: (Primary outcome) composite of time to first occurrence of death or cardiopulmonary hospitalization; (Secondary outcomes) composite death or cardiovascular hospitalization; composite of death or all-cause hospitalization; and all-cause death) • Patient reported data: participants self-report for hospitalizations via dedicated voicemail and website

[1] Tricoci, P., Allen, J. M., Kramer, J. M., Califf, R. M., & Smith, S. C. (2009). Scientific evidence underlying the ACC/AHA clinical practice guidelines.*Jama*, *301*(8), 831-841.

[2] Califf, R. & Sherman, R. (2015). What We Mean When We Talk About Data [Blog post]. Retrieved from http://blogs.fda.gov/fdavoice/index.php/2015/12/what-we-mean-when-we-talk-about-data/

[3] IOM (Institute of Medicine). 2013. *Large simple trials and knowledge generation in a learning health system: Workshop summary.*Washington, DC: The National Academies Press.

[4] Johnson, K. E., Neta, G., Dember, L. M., Coronado, G. D., Suls, J., Chambers, D. A., Glasgow, R. E. (2016). Use of PRECIS ratings in the National Institutes of Health (NIH) Health Care Systems Research Collaboratory. *Trials, 17*(1), 1-11. doi:10.1186/s13063-016-1158-y

[5] Tunis, S. R., Stryer, D. B., & Clancy, C. M. (2003). Practical clinical trials: Increasing the value of clinical research for decision making in clinical and health policy. *JAMA, 290*(12), 1624-1632. doi:10.1001/jama.290.12.1624

[6] Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K. E., & Zwarenstein, M. (2015). The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ, 350*. doi:10.1136/bmj.h2147.

[7] Riley JB, Basilius PA. *Physicians' liability for off-label prescriptions*. Hematology and Oncology News and Issues 2007;6(5):24-37.

[8] Friedman, C., Rubin, J., Brown, J., Buntin, M., Corn, M., Etheredge, L., . . . Van Houweling, D. (2015). Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. *Journal of the American Medical Informatics Association, 22*(1), 43-50. doi:10.1136/amiajnl-2014-002977

[9] Bakerly, N. D., Woodcock, A., New, J. P., Gibson, J. M., Wu, W., Leather, D., & Vestbo, J. (2015). The Salford Lung Study protocol: a pragmatic, randomised phase III real-world effectiveness trial in chronic obstructive pulmonary disease. *Respiratory Research, 16*(1), 1-5. doi:10.1186/s12931-015-0267-6

[10] Johnson, K. E., Neta, G., Dember, L. M., Coronado, G. D., Suls, J., Chambers, D. A., . . . Glasgow, R. E. (2016). Use of PRECIS ratings in the National Institutes of Health (NIH) Health Care Systems Research Collaboratory. *Trials, 17*(1), 1-11. doi:10.1186/s13063-016-1158-y

[11] Johnson, K. E., Neta, G., Dember, L. M., Coronado, G. D., Suls, J., Chambers, D. A., . . . Glasgow, R. E. (2016). Use of PRECIS ratings in the National Institutes of Health (NIH) Health Care Systems Research Collaboratory. *Trials, 17*(1), 1-11. doi:10.1186/s13063-016-1158-y