Incorporating Evidence from Clinical Experience into Regulatory Decision-Making: A Pragmatic Approach to Randomization in the Clinical Setting

Washington Plaza Hotel • Washington, DC
January 15, 2016

Meeting Summary

Introduction

Throughout the development, regulatory review, and clinical use of medical products, large quantities of increasingly diverse data contribute to the scientific body of evidence around each drug, biologic, or device. The traditional cornerstone for each product is the randomized controlled trial (RCT), which is considered the gold standard for developing evidence to support the Food and Drug Administration’s (FDA) determination that a drug is safe and effective for its intended use. However, RCTs are increasingly time and resource intensive and often do not assess long-term outcomes, the drug’s effectiveness in different practice settings, or the benefits and risks to populations not adequately represented in the trial.

A range of opportunities to develop real-world evidence (RWE) from real-world data (RWD) sources are needed to complement and expand our knowledge on the performance of medical products in diverse patient populations and care settings. Routinely collected data from payer claims databases, electronic health records (EHRs), patient registries, and other sources enable the development of RWE that can help providers, payers, and patients make treatment decisions, enhance care quality, and support innovative payment and delivery models. These sources of data have the potential to improve the efficiency of the evidence development process overall and increase value for all stakeholders.

Potential regulatory applications of RWE have not been firmly established and are the focus of a number of emerging policy initiatives. While FDA can and has relied on RWE to support regulatory decision making for rare diseases or through post-market safety surveillance activities, there is a growing need for comprehensive examination of the ways in which RWE could further enhance FDA’s regulatory framework. Congressional proposals for tackling this issue have recently been included in legislative packages such as the 21st Century Cures Act, while FDA and its industry partners agreed to explore opportunities to leverage RWE as part of the sixth Prescription Drug User Fee Act commitments. Taken together, these developments signal a willingness to work collaboratively to improve evidence development efforts to facilitate patients’ access to the treatments they need.

The Duke-Robert J. Margolis, MD, Center for Health Policy is dedicated to helping tackle these issues through a number of data, methods, and policy development efforts aimed at meaningful progress for the use of RWE in the regulatory process. Under a cooperative agreement with the FDA, Duke-Margolis has formed an RWE working group and initiated a series of convening and white paper activities on priority RWE topics. This meeting summary represents stakeholder discussion from the first of these FDA-sponsored activities, a January 15, 2016, expert workshop on “Incorporating Real-World Evidence in Regulatory Decision-Making: A Pragmatic Approach to Randomization in the Clinical Setting.” The event was designed with two primary objectives: to initiate a conversation around clarifying key terms and
concepts in the RWE space, and to explore a single potential regulatory use case for RWE. The views expressed in this summary are those of the individual participants and do not necessarily reflect the official policies of the Department of Health and Human Services, nor does mention of specific projects or organizations imply endorsement by the U.S. Government or other organizations.

**Describing Real World Evidence**

The terminology surrounding RWE is often used inconsistently or interchangeably, and workshop discussion began with a session aimed at delineating important differences between definitions and the potential impact that inconsistent use can have on understanding the applications of RWE. Multiple accepted definitions for real-world data (RWD) and RWE exist (Table 1), and participants underscored the importance of clearly establishing differences between the two concepts. RWD refers to the underlying data elements that are collected, stored, and subsequently used to generate knowledge in the form of evidence. In clinical practice, RWD are routinely collected and stored in EHRs, claims, registries, or other repositories and include, but are not limited to, clinical data in structured and unstructured forms, administrative billing data, and patient-generated data. The baseline definition for RWE is, in turn, evidence developed through collection of RWD outside of randomized clinical trials.

While data is necessary, by itself it is not sufficient to generate RWE. Analytic methodologies and techniques must be applied to extract insight from RWD that is credible and fit-for-purpose. Workshop participants noted that both data and methods should be considered as part of an RWD/RWE definitional framework. Given the varied sources of RWD and how it is captured, participants thought a threshold for data quality may be helpful to further RWD applications (e.g., varying levels of RWD quality may impact whether or not it is considered fit-for-purpose for certain applications). Criteria for developing a threshold for RWD quality might include completeness, availability, reliability, and credibility of RWD sources.

The methods used for RWD collection and analysis, and ultimately the generation of RWE, can also impact how stakeholders define these terms. For instance, RWE is often common shorthand for evidence developed through purely observational methods. Participants noted that a much broader spectrum of study designs and methods contribute to the larger body of RWE. Studies that utilize randomization in the clinical setting or a large simple trial, for example, do not fit neatly into the traditional RCT and observational dichotomy, but still generate what many would consider “real-world” evidence.

While current working definitions for RWD and RWE can provide stakeholders with a common starting point, participants reiterated throughout the workshop that a more nuanced approach to terminology may be needed in order to clearly match real-world data collection activities with stakeholder use cases. Clearly mapping the varying levels of data quality, processes and tools for data capture, analysis methods, and downstream applications would illuminate the large number of ways in which RWD and RWE could contribute to a more efficient system of evidence development and stakeholder decision-making.
Table 1. Examples of Stakeholder Terms and Descriptions for RWD and RWE

<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMS Health(^1)</td>
<td>RWD, RWE</td>
<td>RWD is patient-level data not collected in conventional randomized controlled trials. RWE are insights generated from RWD using appropriate scientific and/or commercial analytics.</td>
</tr>
<tr>
<td>International Society for Pharmacoeconomics and Outcomes Research (ISPOR)(^2)</td>
<td>RWD, RWE</td>
<td>RWD reflects data used for decision making that are not collected in conventional randomized controlled trials. RWE connotes the organization of the information to inform a conclusion or judgment, and generated according to a research plan and interpreted accordingly.</td>
</tr>
<tr>
<td>Network for Excellence in Health Innovation (NEHI)(^3)</td>
<td>RWE</td>
<td>RWE is evidence from any and all sources of data that may contribute to more effective health care, including health care best tailored to the needs of individual patients.</td>
</tr>
<tr>
<td>21(^{st}) Century Cures Act(^4)</td>
<td>RWE</td>
<td>The term real world evidence means data regarding the usage, or the potential benefits or risks, of a drug derived from sources other than randomized clinical trials.</td>
</tr>
</tbody>
</table>

Exploring a Regulatory Use Case

In order to focus the remainder of workshop discussion on identifying concrete challenges and opportunities in the generation and use of RWE for regulatory purposes, the RWE working group chose a single regulatory application to serve as a test case: the approval of a new indication for an already-marketed drug or biologic. This use case was chosen because it represents a clear and feasible opportunity for making near-term progress. In addition, an indication expansion has the benefit of: 1) an established evidence profile with existing safety and efficacy endpoints, and 2) post-market evidence collection that would reflect a broader population taking the drug compared to the clinical trials used for the initial new drug approval.

The RWE working group was also interested in this use case as it provided a setting for further exploring how best to embed randomization within clinical practice settings. For this reason, example pragmatic clinical trial designs (PCTs) were presented as illustrative methods for generating the necessary RWE. PCTs are defined as trials “for which the hypothesis and study design are formulated based on information needed to make a decision”. While pragmatic trials fall within a spectrum of randomization designs, they do entail a different set of trial features compared to traditional RCTs. Traditional RCTs are highly controlled and conducted under “ideal settings” to improve the study’s internal validity whereas PCTs are often designed to be flexible and relax control over the intervention to improve the study’s external validity. Examples of flexibility include broad inclusion criteria, diverse settings, and variation in how the intervention or treatment is delivered. By retaining randomization, however, results of PCTs may also have higher internal validity relative to observational studies, since they reduce the selection bias and unmeasured baseline confounding inherent in observational designs.
Illustrative pragmatic designs presented at the workshop included the Salford Lung Study and the INFluenza Vaccine to Effectively Stop Cardio Thoracic Events and Decompensated Heart Failure (INVESTED). The Salford Lung Study, conducted through a multi-stakeholder partnership in Salford, UK, is the world’s first pragmatic trial of an investigational medication. Using a broad and inclusive population of patients, the study evaluated the effectiveness and safety of a new treatment for chronic obstructive pulmonary disease (COPD) across participating primary care sites. INVESTED is a multi-center trial based in the U.S. that will evaluate the effectiveness of a high dose influenza vaccine compared to a standard dose influenza vaccine in adult participants with a history of myocardial infarction or heart failure. These examples were used as a backdrop for exploring when and how to potentially use a PCT for the regulatory use case of interest.\(^1\)

When to Conduct a PCT for Indication Expansion

In discussing the suitability of PCTs for supporting an indication expansion, participants began to lay out certain criteria for when such an approach could be a viable option. Experts felt that the following factors lend themselves to contexts when a PCT might be a viable approach to developing evidence to support this use case (this is not intended to be an exhaustive list):

- Studies with a single intervention;
- The existence of clinical equipoise to facilitate ethical randomization, making informed consent more straightforward;
- Clinical agreement on a set of standardized data elements that could be easily verified and collected from RWD sources;
- Short term outcomes that are easily captured and measured within RWD sources; and
- A mechanism to engage patients and provide feedback to help sustain the study.

While these criteria could help determine when PCTs are viable approaches, the ultimate decision will likely be based on the contextual needs for evidence.

How to Generate Robust and Reliable RWE from a Pragmatic Trial

To identify and explore key PCT design and implementation factors relevant for the use case, workshop participants worked through four dimensions of pragmatism adapted from Tunis et al\(^8\): study design, patient identification and selection, data collection and quality, and outcomes or endpoints of interest. The following is a distillation of conversation around these dimensions, which also highlights the challenges that should be considered under each when pursing a PCT.

**Study Design Considerations**

An important design consideration involves defining and prioritizing clear study objectives to determine if a PCT is fit-for-purpose. Potential purposes might include filling a specific knowledge gap, assessing potential hypotheses to determine whether an RCT or other study design is needed, or generating rapid evidence at a lower cost to confirm study results. If the study’s objective is to produce rapid evidence for a narrowly defined subgroup of patients, the study design might be tailored towards use of analytic methods and tools that can analyze large, diverse data sets across different settings of care.

Numerous promising methods are emerging to generate rapid and efficient evidence generation. Sequential, multiple assignment, randomized trials (SMART) provide one such example.\(^9\) SMART designs allow investigators to refine or adapt the trial’s design – while it is being conducted – by assessing results at different stages of the trial. This allows researchers to monitor key study components such as

\(^1\) Additional study details are provided in Appendix A.
patient outcomes for appropriate treatment responses and modify the treatment if the patient is not responding with the desired effect. While propensity score matching and instrumental variable analysis are other innovative methods for pragmatic trials, more understanding is needed to determine their appropriate use to approve a new indication.

Patient Identification and Selection
As noted above, a compelling feature of pragmatic trials is the ability to enroll a diverse study population. Ideally, electronic health records could support broad enrollment using computable phenotypes based on key clinical markers found in patient records such as clinical conditions, outcomes, or prescribed medications.\textsuperscript{10} Utilizing health IT strategically, PCTs could be designed to streamline patient identification and selection, but a number of key technical and operational issues must be addressed to realize this aim.

Enrolling eligible populations into large, multi-site clinical trials is complex and challenging. An important consideration for both study design and patient enrollment involves a bioethical concept known as clinical equipoise. Clinical equipoise is the “genuine uncertainty within the expert medical community – not necessarily on the part of the individual investigator – about the preferred treatment.”\textsuperscript{11} This assumes more than one treatment is available for the condition, but that providers genuinely disagree whether one treatment provides a clear clinical benefit or advantage to the patient. Target populations may not be well represented in clinical trials when investigators have differing opinions on which patients should be enrolled for the study. It is therefore critical for investigators to have a shared understanding of study objectives and enrollment needs across participating study sites.

The level or unit at which randomization takes place is another area of complexity. With traditional RCTs the unit of randomization is typically the patient. “Cluster” randomization, a feature common to PCTs, allows for additional levels of randomization such as by practicing health care providers and their patient base, clinics and hospitals, or even by communities to quickly increase the trial’s size. However, a participant noted that one of the difficulties with cluster randomization is that sample sizes tend to be smaller than when the randomization is at the individual level. Many operational challenges remain with use of this technique, including how informed consent is obtained and the inability of EHRs to randomize patient populations, which can limit the efficiency of clustering.

The feasibility of using EHR data to facilitate the screening and selection of individuals for trial enrollment is another important consideration. Workshop participants noted some concerns with the timeliness and completeness of EHR data for this purpose. In particular, a participant commented that according to statistics collected by internally for their organization, potentially up to 35 percent of patient notes are not completed by providers during the time of visit. Others stated that almost 80 percent of clinical documentation in the EHR is captured as free text.\textsuperscript{12} These challenges could limit or hinder investigators from recruiting from the full pool of eligible study participants. There are opportunities to improve both the timeliness and completeness of patient data, including technological solutions such as natural language processing, which can extract data from unstructured text. However, this technology is still relatively nascent and more experience and development of best practices are still needed to successfully implement patient identification and selection through the EHR in pragmatic trials.

Data Collection and Quality
Developing RWE requires access to large, diverse sets of data, but the usefulness of these data is contingent on their overall level of quality. Data standards (e.g., ICD 10, SNOMED, LOINC) help ensure
data are captured systematically and with consistent meaning across patient records to support data aggregation, linkage, and analysis. This level of standardization can help improve care delivery, as well as facilitate clinical trial research. Despite the availability of data standards, data capture and overall quality remains highly variable for numerous reasons. The current incentives for data collection are focused on how certain data must be collected either for care delivery or for payment of rendered services. This can cause asymmetries with the capture of data elements needed for clinical research.

Some participants felt that claims data could serve as a comprehensive data source for study recruitment using patient demographics, patient histories of hospitalizations, or related events. Claims, however, might provide less complete data for vaccination records or other important clinical values (such as the lower ventricular ejection fraction, or LEVF) that are critical to some studies, such as INVESTED. While the EHR could serve as a better source for these data elements, they are less likely to be systematically captured across EHR systems, which can limit researchers’ ability to aggregate or link data for analysis. The INVESTED study discussed during the workshop is currently being conducted to demonstrate the advantages and weaknesses of claims and EHR data.

To address these data collection challenges, data needed for PCTs should be limited to a core set of data elements that are easily captured within existing clinical workflows. Defining this core list of elements as part of the initial study’s design could improve trial sustainability and overall data quality as a result of clear data collection expectations for clinical investigators. Ideally, identifying core data elements that are also needed for payment purposes would provide added incentives for systematic data capture.

Also of note are emerging data networking initiatives such as the National Patient-centered Clinical Research Network (PCORNet) and FDA’s Sentinel System, both of which standardize existing data collected across health systems through a common data model (CDM). The CDM is used by all collaborating partners in the networks and ensures data elements are maintained in a standard format. This enables the development of standardized analytical methods and tools to curate and validate data and tools for analytics. Implementing a CDM requires data governance processes and stakeholder cooperation to harmonize meaning across data elements.

**Outcomes and Endpoints of Interest**

Endpoint collection, measurement, and reporting are critical for the conduct and analysis of clinical trials as well as the translation of study results into clinical and health policy decision-making. Inherent difficulties in collecting outcomes, such as overall variability of outcomes across trials, limits both the interpretation and generalizability of results. Poorly described outcomes also hinder their consistent collection. While PCTs impose less control over data collection compared to traditional RCTs, the inherent complexity of the methods can further exacerbate these challenges.

Standardizing the measurement and interpretation of outcomes and endpoints through core outcome sets (COS) can improve their collection and reporting. Developing a COS for specific therapeutic areas of research requires stakeholder agreement. Standardized disease definitions across therapeutic areas will also be required as a precursor to standardized measurement of clinical trial outcomes. The stakeholder community appears to be moving towards such standardization, given FDA’s recent guidance to assess cardiovascular endpoints because of the wide variation in how these endpoints are interpreted across studies.

This level of standardization could be particularly useful for trials like INVESTED that use “composite endpoints” where the primary endpoint reflects a range of outcomes. Endpoints are weighted to account for an assumed larger impact on an overall composite leading to a disease score. Interpreting
these scores can be difficult however. Key assumptions used for weighting endpoints along with their ordered sequence can affect trial results potentially making them appear statistically stronger or weaker. Therefore, standardized COS could possibly support better interpretation and use of composite endpoints by reducing the number of assumptions specific to individual trials.

Next Steps
Throughout the workshop, participants identified areas where additional work is needed to advance the specific methods used to generate RWE as well as address challenges with the use of RWE in regulatory decision making. In particular, they noted the lack of clarity of the use of terms such as RWE and RWD. Clearer definitions should address the full range of potential RWE activities beyond the narrow category of retrospective studies conducted with purely observational designs, and consider other approaches that include randomization.

With clearer terminology, a key next step will be developing a high-level vision that establishes specific instances where RWE could complement and enhance more traditional types of evidence in regulatory decision-making. This vision could lay out the potential application and role of RWE across a range of regulatory decisions, taking into consideration the context of use. It would be useful to lay out examples of how different RWD elements, methods for data collection, and analytical techniques would inform a specific regulatory use such as an indication expansion study. Identifying key contextual factors such as specific therapeutic areas or outcomes could also help determine instances where RWE might be more practical, meaningful, and consistently applied. While the aim of these activities is to develop an ideal data infrastructure to generate RWE for regulatory decision-making processes, it would also need to identify the current gaps which hinder making that vision a reality.

The workshop also highlighted the need for best practices in generating RWE across a number of study designs or methods. While pragmatic trials potentially represent an important methodology for evidence development in regulatory settings, many questions remain on how these designs would be evaluated by the FDA. More direction from FDA on key aspects of pragmatic trial designs and how they fit within evidentiary hierarchies is needed. In addition to more regulatory clarity, the stakeholder community must also identify best practices and lessons learned from the growing implementation experience with pragmatic trials. Examples such as Salford and INVESTED and broader initiatives like the NIH Collaboratory, which is implementing 10 active pragmatic trials, will be instrumental in distilling important insights and opportunities to improve study methods and designs.  

Finally, a common theme across all workshop sessions was the need to develop a value case for using pragmatic trials to generate RWE. The present cost of many PCTs are likely similar to RCTs given the substantial infrastructure needed to run trials with large populations across diverse settings. The development of research and surveillance platforms leveraging a distributed data infrastructure, such as PCORnet and the Sentinel System, could help scale this type of research by executing multiple trials simultaneously. If these trials were designed to meet varying evidentiary needs of stakeholders, distributed platforms could provide a compelling business case for participation and help reduce fixed costs across partners interested in using this infrastructure. RWE must support critical evidentiary needs and demonstrate gains in efficiency through faster, cheaper, and more scalable data sources.
## Appendix A. Overviews of the Salford Lung Study and INVESTED Trial

<table>
<thead>
<tr>
<th>Project Title</th>
<th>Research Question</th>
<th>Example Pragmatic Features of Studies</th>
</tr>
</thead>
</table>
| **Salford Lung Study (SLS)**                                                 | 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. |
| **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 |
References

15. NIH Collaboratory. Retrieved November 17, 2016, from [https://www.nihcollaboratory.org/Pages/default.aspx](https://www.nihcollaboratory.org/Pages/default.aspx)