Examining the Impact of Real-World Evidence on Medical Product Development: I. Incentives: Proceedings of a Workshop in Brief

DETAILS

10 pages | 8.5 x 11 | PDF
ISBN 978-0-309-47205-0 | DOI 10.17226/25024

CONTRIBUTORS

Amanda Wagner Gee, Ben Kahn, and Carolyn Shore, Rapporteurs; Forum on Drug Discovery, Development, and Translation; Board on Health Sciences Policy; Health and Medicine Division; National Academies of Sciences, Engineering, and Medicine

Proceedings of a Workshop

IN BRIEF

February 2018

Examining the Impact of Real-World Evidence on Medical Product Development: I. Incentives

Proceedings of a Workshop—in Brief

On September 19–20, 2017, the National Academies of Sciences, Engineering, and Medicine held the first workshop of a three-part series titled Examining the Impact of Real-World Evidence on Medical Product Development. The workshops are convened under the auspices of the Forum on Drug Discovery, Development, and Translation and sponsored by the U.S. Food and Drug Administration (FDA). The workshops are intended to advance discussions and common knowledge among key stakeholders about complex issues relating to the generation and use of real-world evidence (RWE). The first workshop focused on how to align incentives to support the collection and use of RWE in health product review, payment, and delivery and how to address the barriers that impede the uptake and application of RWE.

Gregory Simon of the Kaiser Permanente (KP) Washington Health Research Institute told the workshop participants that establishing a common language will be key to understanding and changing the traditional paradigm of evidence generation, and he emphasized that real-world data (RWD) are distinct from RWE. He noted that RWE may come from many sources, including randomized controlled trials (RCTs). Andrew Bindman of the University of California, San Francisco, and Robert Califf of Duke University and Verily Life Sciences said that all stakeholders want access to scientifically derived evidence, but there may be different ways to obtain such evidence and the required degree of confidence in RWE may depend on the needs of the person making a decision. Simon characterized the core qualities of RWE as:

- It is generalizable to correctly predict an outcome for patients; the ability to assess the accuracy of the prediction after a patient was treated imbues an implicit accountability to the evidence-generating system.
- It is relevant to decision makers' specific information needs and stems directly from their priorities. This implies that it is "fit for purpose," meaning that the evidence is designed to answer the question regardless of its source.
- It is adaptable to embrace the heterogeneity in RWD.
- It is efficient in the sense that the evidence can be produced more quickly and less costly than through traditional methods; this efficiency is necessary because answering fit-for-purpose questions requires the generation of more evidence types.

ADVANCING PUBLIC HEALTH OPPORTUNITIES WITH REAL-WORLD EVIDENCE

Scott Gottlieb of FDA, the workshop's first keynote speaker, laid out FDA's current thinking about RWE. Gottlieb said that RWE is being more widely used for coverage and reimbursement decisions and its rigor is therefore increasing. "As the breadth and reliability of RWE increases, so do the opportunities for FDA to make use of information," he said. Clinical care choices are made based on many sources of information that have varying degrees of uncertainty. FDA could therefore support the development of and access to reliable evidence that meets standards for approval. Gottlieb

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

emphasized that FDA will uphold and promote the "gold standard" for evidence; however, the source of that evidence is not mandated.

Gottlieb said that pre- and postmarket evaluations should be thought of as parts of a continuum rather than as two separate and distinct processes; in particular, he said, the "need for a point of regulatory accountability" should not preclude the possibility of evaluating products over their life cycle of use. RWE offers a way to better inform the benefit–risk profiles for medical products and is already used routinely by FDA to evaluate safety and emerging risk.

To encourage regulated industry to take the risk on RWE, FDA is taking steps to offer more clarity by issuing final guidance for devices and it is developing policies that support the use of RWE in indication expansions, especially in cases of unmet medical need and rare diseases and in meeting postapproval requirements. Rachel Sherman of FDA added that the goal is to achieve "a better evidence base on which to make our medical product approval or clearance decisions."

SEEING THE DESTINATION

During the workshop's first session, speakers and discussants explored the knowledge base that may be necessary to make informed decisions about the use of medical products. Bindman introduced the speakers, noting that if one seeks to redesign the evidence-generation system, it is critical to hear from multiple perspectives.

A Payer Perspective

Michael Sherman of Harvard Pilgrim Health Care said that Harvard Pilgrim strives to find a balance between treatment access and affordability for patients while also encouraging treatment innovation. M. Sherman said that payers want to pay for new treatments proven to add value for patients, but these therapies must be selected carefully with cost in mind (e.g., payers should evaluate whether a given treatment would lead to premiums increases for employers or individuals). Ultimately, he said, payers should focus on treatments backed by good data and evidence, while also considering the patients' perspective when determining coverage. M. Sherman acknowledged that some conditions are too rare for treatments to be tested in RCTs, which creates an ethical tightrope for payers.

M. Sherman advocated for manufacturers to enter into value-based agreements for reimbursement that are tied to product success. He suggested that manufacturers could be required to submit data to third-party value analysts, such as the Institute for Clinical and Economic Review,[1] so that pricing could be assessed in a value-based way. Finally, M. Sherman said that payers could encourage FDA to collect postmarketing data for cases in which there are limited data on a given treatment—a practice the agency often already engages in and one that creates RWD that can be used to better study patient outcomes.

An Integrated Health Care System Perspective

Michael Horberg of KP described the perspective of a health care delivery system on RWE, focusing on KP's integrated care model. Horberg said that KP makes care-altering decisions after considering the answers to several key questions, which would also apply to RWD assessment and RWE generation:

- What are the efficacy and effectiveness of a new treatment?
- Who conducted the research?
- What is the population at risk?
- What does KP do currently to address a disease/condition?
- What is the added cost of the new treatment?
- If implemented, how would the change in practice be operationalized?

Horberg said that KP recognizes that a gap exists between clinical trial efficacy and real-world effectiveness of new treatments; it also recognizes, he said, the cultural tendencies for patients and providers to rely on anecdotal evidence in making care decisions. Horberg explained that in KP's integrated care model, decisions to change practice may be based on one of several factors: new knowledge that appeared in the medical literature, pharmacy or physician requests, patient demand, changes to state or federal statutes, and especially KP's own data or research findings. Changes in practice require the input of multiple stakeholder groups that think and act independently, with the hope that opinions will ultimately converge about treatment recommendations, benefits decisions, and formulary decisions.

[1] https://icer-review.org (accessed November 20, 2017).

An Academic Health System Perspective

Daniel Ford of the Johns Hopkins University School of Medicine presented an academic health care delivery system perspective. He described the Johns Hopkins Health System (JHHS) Corporation—which includes 2.5 million people in its electronic health records (EHRs) and 300,000 people enrolled in clinical protocol—as an evidence generator and evidence consumer. He said that the system incorporates and relies on both internal and external evidence. In generating evidence, Ford said, JHHS still relies heavily on traditional clinical research, which is often funded by the National Institutes of Health (NIH). However, recent efforts to conduct clinical research at community hospitals opened up new opportunities to gather clinical evidence. Ford discussed JHHS's data needs, noting that while the system's data warehouse is improving, it would be difficult to rely solely on its observational data to make judgments about clinical effectiveness. He mentioned several collaborative efforts, including the Chesapeake Regional Information System for our Patients,[2] a health information exchange that serves Maryland and Washington, DC, and said that such efforts provide an important data tool for Johns Hopkins both in research and clinical practice.

Ford discussed the roles and expectations of patients and providers in the generation of RWE. Patients often desire information about their treatment plans in order to inform their personal decision making, but this information is not always accessible in the current environment. Academic researchers, meanwhile, frequently express interest in researching drug effects on off-label indications and applying their findings to usage recommendations, but providers still rely on traditional RCT evidence rather than other potential sources of evidence. Ultimately, Ford said, altering the evidence-generation system will require changes on the part of providers.

A Patient Perspective

Sharon Terry of Genetic Alliance reminded participants that the term "patient" implies a power imbalance between the practitioner and the individual seeking treatment, saying that many health professionals would approach decisions differently if they put their personal health before business or industry preferences. Patients, she said, are concerned about their own health, and groups of them have created disease- or condition-centric communities to represent their interests and generate data to inform health care decisions. However, the data generated by these groups are often discounted when they should be systematically included with other, traditional sources of data, she said.

Terry described the importance of community-based registries to her work and to the evolution of RWD. She noted that community registries, such as those run through PCORnet[3] or Genetic Alliance's Platform for Engaging Everyone Responsibly (PEER),[4] are distinct from industry- or hospital-led registries in that they are led by community members, they respect and highlight the priorities and concerns of the individuals participating in the registry, and they focus on the education of members through consistent outreach over social media and other media. Terry admitted that one area in which community registries could improve would be to provide more rigorous validation and she suggested that developing a methodology for this could become a priority since it is critical to establish a RWE-generating system that takes into consideration the lived experience of those seeking treatment.

A Reaction from Data Generators

Following the panel, invited discussants Joanne Waldstreicher of Johnson & Johnson and Eleanor Perfetto of the National Health Council (NHC) provided additional observations in reaction to the presentations. Waldstreicher agreed with Terry's emphasis on creating evidence that maximizes value and minimizes risk for individuals seeking treatment. Waldstreicher encouraged the increased sharing of data between stakeholders as a way of developing an improved understanding of various medical products. The follow-up analysis of product data post-licensure is an important step in building a learning health care system, she said.

Perfetto referred workshop participants to an NHC white paper stemming from a July roundtable on RWE and patient perspectives.[5] She highlighted several takeaways from the meeting and paper. First, patients often do not distinguish between RWE and RCT-derived evidence; they care primarily about answering the questions "Will this work for me, and [will I experience] side effects?" Second, Perfetto said, patients believe that they own their own data and deserve to dictate who can use them and when, in an "opt-in" environment. Third, she said that patients do not view data from EHR and claims as authentic patient data because those data do not reflect patients' preferences or experiences in a meaningful

[2] https://www.crisphealth.org (accessed November 20, 2017).
[3] http://www.pcornet.org (accessed November 20, 2017).
[4] http://www.geneticalliance.org/programs/biotrust/peer (accessed November 20, 2017).
[5] http://www.nationalhealthcouncil.org/sites/default/files/Patient%20Perspectives%20on%20Real-World%20Evidence.pdf (accessed November 20, 2017).

way. Fourth, Perfetto said, patients need clearer definitions of RWD and RWE if they are to be credible stakeholders and make informed decisions. Last, she emphasized Terry's point that patient communities are a great source of RWD and one that should not be discounted in the generating of evidence from a diverse set of sources.

FIT-FOR-PURPOSE EVIDENCE

Mark McClellan of the Duke–Margolis Center for Health Policy gave a keynote address on fit-for-purpose evidence after the Session I panel and discussion, reiterating many of the points made by the Session I panelists. He noted that fit-for-purpose evidence was a topic that Duke–Margolis addressed at a separate September 2017 workshop[6] and in a September 2017 white paper.[7] Addressing the uncertainty concerning terminology, he offered a crucial distinction between RWD and RWE: RWD are related to patient health status and the delivery of care, are routinely collected through clinical records or similar sources, and should be credible and trustworthy. RWE is derived by culling RWD by applying sound, rigorous analytical methods to RWD. McClellan, referring to a framework developed in the Duke–Margolis white paper,[8] emphasized that successfully generating fit-for-purpose evidence relies on understanding the regulatory and clinical contexts surrounding the data.

Concluding his presentation, McClellan said that steps being taken to develop fit-for-purpose and patient-centered RWE not only may lead to new clinical and regulatory decision-making tools, but also could support the development of new payment models based on the value and quality of care. Ultimately, he said, the use of RWE is limited compared to the amount of RWD currently available, so further stakeholder investment in this area is needed.

LEARNING FROM SUCCESS

In Session II, Learning from Success, workshop participants discussed examples of the successful use of RWE in decision making. The presentations focused on how the methods and techniques used in the successful examples could apply to future applications.

Generalizing and Scaling the Salford Lung Studies

Martin Gibson and Marie Kane of Northwest eHealth discussed the Salford Lung Studies (NASEM, 2017; Vestbo et al., 2016; Woodcock et al., 2017). Gibson said that these studies demonstrated a way to bridge the traditional gap between pre- and postmarketing data collection.

Gibson and Kane emphasized that the keys to the success of these trials were coordinating and connecting patient care in the hospital with care in community health centers and planning the trials from the beginning to answer questions from regulators, payers, and researchers. This required working before the studies began to engage or train the health system, sponsor, regulators, providers, pharmacists, and information technology departments. Kane said that these studies required far more investment in data processing and error management than traditional clinical studies. This was due to variability in the data as they were collected as well as the scale and complexity of the data linkages required to determine patient outcomes. Kane said that it was important to focus on collecting the right data rather than on collecting more data. Kane and Gibson reported that Northwest eHealth has converted the platform developed for the Salford studies into individual, cloud-based "configurable and modular" applications to make it easier to adapt the infrastructure for future studies.

During the discussion period, participants debated how the Salford model could function as a "franchise" that could be exported to other health systems and other disease questions. Kane added that the "franchise" part of the studies would consist of the methods developed and the lessons learned from each variation. John Graham of GlaxoSmithKline (GSK) agreed that the studies had generated both a reusable infrastructure and lessons learned that could be applied in other disease areas; GSK is already applying a similar model in cardiovascular and renal disease studies in the United States. Gibson added that recruiting health systems to participate became easier with each study because of the positive experiences of the study participants and the investments made in building relationships as well as the desire of health systems to participate in exciting, new, relevant studies. He said that experience alleviated the fear of doing a trial among community hospitals and other systems where research is not the main focus.

[6] https://healthpolicy.duke.edu/events/public-workshop-framework-regulatory-use-real-world-evidence (accessed November 20, 2017).
[7] https://healthpolicy.duke.edu/sites/default/files/atoms/files/rwe_white_paper_2017.09.06.pdf (accessed November 20, 2017).
[8] *Considerations for generating RWE fit for regulatory purposes* (p. 10), https://healthpolicy.duke.edu/sites/default/files/atoms/files/rwe_white_paper_2017.09.06.pdf (accessed November 20, 2017).

Using Sentinel to Evaluate Effectiveness or Efficacy

Richard Platt of Harvard Medical School discussed Sentinel's design and its potential for use in assessing efficacy and effectiveness in addition to safety. Sentinel[9] is built on a highly curated, distributed data network with a common data model used across the network (NASEM, 2017). While at this time Sentinel is used primarily for safety surveillance, monitoring effectiveness is also part of its mandate. Platt shared six examples of how Sentinel has been, or is being, used to answer questions of efficacy and effectiveness in addition to questions concerning safety.

1. Sentinel data alone were used to assess the comparative safety of rivaroxaban versus warfarin.[10] Comparing the results of that study with the ROCKET-AF RCT demonstrated that the observational analyses in Sentinel can correlate well with RCT results (Patel et al., 2011). Platt said that this technique may offer a mechanism to explore populations that are not well represented in RCTs.
2. Sentinel linked with adjudicated medical records was used to assess intussusception with rotavirus vaccine (Yih et al., 2014). This required expert adjudication to verify exposure to the vaccine and presentation with symptoms and it resulted in a change to the vaccine label.
3. Sentinel in combination with registries was used to link infants with their mother's health records.[11] Linking Sentinel data with state registry data incrementally improved the results over what was possible with Sentinel data alone. Echoing experience from the Salford studies, Platt said that this illustrates the benefit of collecting the right data.
4. Sentinel-like data are being linked to EHRs in PCORnet's ADAPTABLE trial to assess low-dose versus high-dose aspirin for the prevention of coronary artery disease.[12]
5. Sentinel linked to patient-generated data is being tested in the development of a mobile app for patients to collect and report information, such as information about the daily lives of pregnant women, that will later be merged with the Sentinel databases.[13]
6. Sentinel is being tested as a platform for randomized trials through IMPACT-Afib (Pokorney, 2017). This trial will reach out directly to Sentinel members who have atrial fibrillation and a high risk of stroke but no evidence of oral anticoagulant use.

During the discussion period, workshop participants considered which aspects of Sentinel could inform RWD/RWE use. Platt and R. Sherman pointed out that everything done in Sentinel is in the public domain and is intended as a public resource. The distributed data model allows data partners to choose whether to contribute their data to each inquiry. However, Platt said that this model is expensive to maintain and is only successful if the data partners continue to find it useful.

Applying Lessons Learned from Device Registries to Other Treatment Types

Rachael Fleurence of the National Evaluation System for Health Technology (NEST) Coordinating Center spoke about using RWE in devices and how the lessons learned from such use could be applied to other treatment types.

- The processes of device approval and regulation are different from the corresponding processes for drugs in a number of ways, including that there are several different pathways to approval for devices; different devices may have different requirements for postmarket safety studies; it is difficult to track device implantation because the system of unique device identifiers only began in 2015 and these identifiers are not universally required in EHRs; adherence is not a concern for implanted devices; it can be difficult to disentangle the "learning curve effect" of providers iteratively improving their implantation techniques with practice from actual problems with the device; and the current surveillance system for problems with devices depends on voluntary reporting.

[9] https://www.sentinelinitiative.org (accessed December 4, 2017).
[10] https://www.sentinelinitiative.org/sites/default/files/Communications/Publications/Sentinel-ICPE-2017-Presentation-Rivaroxaban.pdf (accessed December 4, 2017).
[11] https://www.sentinelinitiative.org/sites/default/files/Sentinel-ICPE-2017-Presentation-PRISM-Mother-Infant-Cohort.pdf (accessed December 4, 2017).
[12] ClinicalTrials.gov: NCT02697916: https://clinicaltrials.gov/ct2/show/NCT02697916 (accessed December 4, 2017).
[13] https://www.sentinelinitiative.org/sites/default/files/Communications/Publications/Sentinel-ICPE-2017-Presentation-Mobile-App.pdf (accessed December 4, 2017).

- The device community has some experience already with RWD from the widespread use of registries. Device registries have historically been important in device regulation, in part due to mandates for registries for expensive, high-risk device implants.
- Registries have been considered by some in medical product development as a potential answer to a number of questions about how to use RWD more widely, Fleurence said. The benefits of registries include the provision of high-quality, fit-for-purpose data; easy linking with other data sources through coordinated registries networks; the ability to support both pre- and postmarketing observational studies, potentially at lower costs; and the potential to serve as a platform for automated safety surveillance
- However, Fleurence said, registries are not a blanket solution to be applied to all treatment types, or even to all devices, because of several significant drawbacks that registries have. They are expensive to develop and maintain; they are not practical in some kinds of treatments and patient populations; they vary in their data quality and methods; they can pose significant administrative challenges; and, as is the case with other data sources, those who operate registries need to grapple with safeguarding patient privacy and security.

Rather than further developing registries that may become increasingly burdensome as they become larger, Fleurence suggested that the key to RWE use in the future could be to focus on developing opportunities to work more directly with the systems generating data in the course of clinical care or at home. In device development and use, RWD could be used to generate robust postmarket data to support earlier premarket decisions, help researchers recognize and assess safety problems sooner, help medical professionals determine better ways of using a device, and help researchers design rigorous studies that will be able to reliably detect safety and efficacy outcomes. NEST was established, Fleurence said, to promote these applications and "serve as a catalyst to support timely and reliable development of high-quality RWE."

GETTING UNSTUCK: ALIGNING INCENTIVES

In Session III, Getting Unstuck: Aligning Incentives, workshop participants discussed incentives for maintaining the current data generation processes and potential barriers to the use of new methods of evidence generation. Anna McCollister-Slipp of the Scripps Translational Research Institute and VitalCrowd, Inc., framed the discussion by saying that the biggest barriers to the system as a whole are a lack of a sense of urgency to accept RWE, which she suggested is primarily due to bias favoring the traditional evidence generation system and limiting access to data after they have been generated. Like Terry and Perfetto, she observed that there are consequences to a heavy reliance on RCTs for data generation and it is no longer reasonable to exclude the consideration of patient-generated data sources.

A Contract Research Organization Perspective

John Doyle of QuintilesIMS told the workshop participants that contract research organizations (CROs) are interested in using RWD/RWE to improve the process and increase the efficiency of delivering trials as well as to design better studies. He said that CROs are interested in implementing RWE studies in a scalable and systematic way, and he added that when CROs start a new study, they consider the needs and requirements of regulators, policy makers, payers, patients, and others who must make decisions about a medical product. RWD have already been used to optimize recruitment, to shorten the time it takes to start a study, and to reduce costs through risk-based monitoring. Furthermore, Doyle said, using RWD/RWE to bridge the evidence discrepancies between clinical trial patients and real-world patients could answer questions about subpopulations for precision therapy treatments or demonstrate proof of value to payers and patients. Doyle suggested that FDA's recently released final guidance on RWE for devices could be a source of ideas for how other therapy modalities could incorporate RWE into study designs. He offered several examples of methods that blend the RWD and RCT approaches to evidence generation, such as running single-arm open trials with historical controls rather than concordant placebo study arms or pragmatic randomization designs.

A Product Developer Perspective

Elliott Levy of Amgen Inc. discussed aligning incentives from the perspective of a product developer. Companies are already taking advantage of big data and RWD internally, he said, by using methods similar to those described by Doyle to improve product development, patient experience and outcomes, and value to the health care system. Levy emphasized that RWE and clinical trial evidence are not in conflict, but rather are complementary. Clinical trials inform development of rigorous RWE generation and RWE improves the pragmatism and relevance of clinical trials. Levy identified the barriers within companies that impede change as (1) a lack of knowledge and awareness of RWE methods, because product teams

are usually led and staffed by scientists without training or trust in those approaches, (2) a lack of talent and capabilities in the relevant areas, because there are few individuals in organizations who have experience in observational research and these tend to be found on safety or health economics teams rather than in product development, and (3) systems and processes are generally not set up for RWE, because organizations tend to be optimized around generating RCT data and new approaches can become unnecessarily complicated once they are fit with existing company processes, such as procurement. Levy emphasized that these barriers are all addressable. He suggested that senior leadership promote and support RWE adoption, companies invest in training team members in RWD management and analytics capabilities, and the company leadership directly address challenges in organization processes.

An Academic Researcher Perspective

Ford discussed the misaligned incentives and barriers to RWE use from an academic researcher perspective. Ford said that younger investigators are often less willing to try new methods because they are more risk averse as they are establishing their careers. Furthermore, the career incentives for established investigators favor performing high-quality RCTs. RWD requires that researchers relinquish some control over a study, Ford said, because the analytical methods require more dependence on statistician colleagues and the nature of the data requires a willingness to accept less precise data as well as to discard extraneous data in EHRs. These practicalities of RWD use can be difficult to accept for researchers trained in RCT methods.

Ford suggested that pragmatic study designs might be a good way for traditional clinical trialists to gain experience with RWD. He acknowledged, however, that in addition to the initial time investments described by other presenters, it may be difficult to find partnering health systems because they are probably capable of accommodating only a few trials at a time and demand is increasing. In a different approach, Johns Hopkins is beginning to develop capabilities for physicians to query their own patients' EHR data, Ford said; the goal is to encourage greater interest in data collection.

A Big Data Perspective

Marcus Wilson of Anthem's HealthCore observed that most patients get care from a highly fragmented health system. Payer companies often do not have the relevant evidence available to determine how RCT populations relate to their own patient populations, nor is the evidence generated until after a product is marketed. Wilson emphasized that the gravitational pull back to the familiar is a major underestimated systemic obstacle to overcoming this fragmentations. In organizations with large data sources like HealthCore, as well as in the developer companies as described by Levy, this pull can influence and affect decisions at every step. McCollister-Slipp added that this particular barrier also affects funding decisions made by reviewers at funding agencies, so the bias often extends to what types of studies can be run.

Wilson argued that the solution is to defragment the patient view by sharing data responsibly and creating value by linking data from disparate parts of the health system as well as patient-provided information. He said that institutions that collect and share these data should adhere to core principles, including protecting patient privacy and security, using data only for those purposes for which they are fit, and actively creating a learning health system. This data work, and its associated cost, can be planned for prior to marketing approval so that better decisions can be made by all stakeholders earlier. Wilson argued that these steps would benefit both patients and the business interests of the data sources.

The Perception of Evidence Hierarchies

Workshop attendee Hui Cao of Novartis observed that the hierarchical rating of evidence by data source begins in medical school training and is further promoted in the peer-reviewed literature and asked workshop presenters for comments and possible solutions to this phenomenon. Wilson suggested a system that grades evidence by analytical method rather than by data source, and Waldstreicher commented that observational studies should incorporate higher standards for rigor, transparency, and reproducibility. Ford suggested publishing the costs of RCTs along with the resulting data to encourage consideration about whether the extra investment required is justified by the perceived improvement in data quality. Many individual workshop participants emphasized that the most important point was to find the right method, whether RCT or RWD techniques and data sources, for addressing each question. During his keynote address, Gottlieb said that the historical hierarchy of evidence is changing as the reliability of forms of evidence other than fully randomized, prospective, placebo-controlled trials increases with improvements in the methods used to evaluate them. He said that FDA could therefore support changes by releasing consensus definitions of terms and describing RWE and its applications for satisfying FDA requirements as part of a developing guidance document.

The Potential of Building Trust

Wilson discussed the reluctance of health care systems to share data with other health care systems as one barrier to defragmentation. Gibson, Horberg, and Wilson attributed this reluctance to a lack of trust between the health systems and they emphasized the importance of establishing relationships to facilitate data sharing. Deven McGraw of the Office for Civil Rights at the U.S. Department of Health and Human Services agreed that creating trust was paramount for data sharing to develop RWE. She said that a framework could be in place to facilitate the development of trust and responsible data sharing and to reduce the uncertainty around putting patients at risk or violating the Health Insurance Portability and Accountability Act of 1996 (HIPAA). She suggested considering some type of credit or reward for organizations that were already doing this well.

GETTING UNSTUCK: MYTHBUSTING

On the second day of the workshop, participants examined ideas and misconceptions about established evidence-generation practices. Califf opened the session with a keynote address about false precision and estimating the reliability of the evidence-generating process. He encouraged changing the goal of evidence generation from precision to reliability, which will require focusing on shedding practices and portions of the old system that increase cost without improving evidence quality and emphasizing rigorous science over standard operating procedures. Califf listed four key principles that could underlie any such evidence-generating system:

1. Build a reusable system embedded in clinical practice and learn from every encounter, but also ensure that lessons learned are actually spread to the point of care. Califf observed that several of the public–private partnerships and integrated health systems discussed on the first day of the meeting are developing this capability.
2. Use quality by design to eliminate errors that bias results and ignore those errors that do not affect the outcome so that effort and resources are spent efficiently.
3. Use automation for repetitive tasks, real-time analysis of comparison data embedded within health care, and infrastructure to share the results with practitioners to support a constantly learning system.
4. Operate from basic principles rather than merely establishing different standard operating procedures.

A Data Aggregator Perspective

Patrick Ryan of Janssen Research and Development discussed some of the methods being developed at Observational Health Data Sciences and Informatics (OHDSI) (NASEM, 2017).[14] These methods and the related OHDSI databases support three types of analytic use cases: (1) clinical characterization to describe outcomes in a specific population, (2) patient-level predictions to anticipate what will happen for an individual patient, and (3) estimates of population-level effects for safety surveillance, comparative effectiveness, or causal inference.

Ryan performed a live demonstration of an analysis of results from published, peer-reviewed literature to show that even after years of research, meta-analyses conclude that the answers to many clinical questions are unknown. Ryan said that the purpose of this demonstration was to show that "we can't necessarily trust the process that we are using to generate evidence as a community," regardless of the data source. OHDSI databases contain raw data from four different sources in order to minimize bias and allow for easy comparisons. This systematic standardized approach of analyzing multiple data sources simultaneously and asking specific, fit-for-purpose questions can generate more trustworthy answers and potentially answer patients' question of "What will work for me?"

A Medical Product Developer Perspective

Graham said that the goal for developers is to have the right answers to the right questions at the right time. He emphasized that RWE is now necessary to answer many questions, but it is not a replacement for traditional research. He advocated for focusing on a challenge-based, holistic thinking process about what will improve an individual patient's outcome rather than the traditional assessment of individual studies. RWE can help do this by informing which disease states to focus on in development, how to develop a particular treatment, or how to explain benefits and risks to patients in a meaningful way.

[14] https://ohdsi.org (accessed December 4, 2017).

An Academic Researcher Perspective

Rory Collins of the University of Oxford said that randomized trials are necessary to detect moderately beneficial or adverse effects of new treatments and establish causality, particularly when trial populations are widely diverse. He argued, however, that under current regulations the burden of conducting randomized trials is too cumbersome, in large part because of the widespread misapplication of the good clinical practice (GCP) guidelines issued by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Collins said that these guidelines incentivize a focus on standard operating procedures, rather than on innovation in the design and conduct of studies, and on verifying and adjudicating source data rather than on generating reliable results. He therefore advocated for a focus on developing evidence-based strategies to make conducting randomized trials easier. Other individual workshop participants noted that a purpose of GCP was to give providers guidelines for how to conduct research well. Those participants said that in the context of pragmatic or real-world trials, this becomes a mission to train providers to recognize whether to intervene with an individual patient's treatment course.

A Regulator Perspective

Janet Woodcock of FDA reiterated that current data generation methods are costly and time consuming and often cannot generate the evidence necessary to answer the relevant question. Woodcock said that FDA is committed to exploring the use of RWE in regulatory decisions and she cited the recently issued guidance for devices. She said that the incentive embedded within that guidance was for developers to invest in developing RWE methods for regulatory purposes. As drug development changes with the onset of precision medicine and there is a push for accelerated development programs, the need for innovation in study methods is becoming more apparent. Woodcock emphasized the promise of hybrid RCT–RWD approaches to studying investigational drugs and she pointed to the NIH collaboratories for examples of using innovative method applications. She also discussed innovative trial designs, such as master protocols or platform trials, as promising ways to incorporate RWD. To carry out innovative trial designs will require additional work in standardization, verification, and training as well as potentially different strategies for development or funding structures and academic rewards, but these designs offer tangible opportunities to more easily adopt new practices in the clinic; to answer multiple questions simultaneously, including comparative effectiveness; and to maintain the focus on patients.

FINAL REMARKS

Simon concluded by commenting that change is held back not by greed, but by fear. Traditional evidence-generating methods fail in familiar ways and so are perceived as more reliable than RWE. The benefits to the broader community may not translate to benefits for the data generator. Finally, Simon echoed the discussion about the difficulty of building trust across stakeholders. He listed potential next steps discussed by some workshop participants which could help address these challenges and encourage the wider use of RWE. First, Simon said that asking fit-for-purpose questions will be critical in determining context-specific value in the health care system. Second, he stressed the importance of using appropriate methods, including making choices about when randomization is needed to answer a particular question. Regarding observational studies, he said that elevating the rigor of trial designs through transparency and sharing of methods and data will become more important. Finally, Simon suggested accommodating diverse evidence needs across stakeholders, defining smaller studies based on simple questions and sound research design, and reconsidering the delineation between the preapproval and postapproval process.♦♦♦

REFERENCES

NASEM (National Academies of Sciences, Engineering, and Medicine). 2017. *Real-world evidence generation and evaluation of therapeutics: Proceedings of a workshop.* Washington, DC: The National Academies Press.

Patel, M. R., K. W. Mahaffey, J. Garg, G. Pan, D. E. Singer, W. Hacke, G. Breithardt, J. L. Halperin, G. J. Hankey, J. P. Piccini, R. C. Becker, C. C. Nessel, J. F. Paolini, S. D. Berkowitz, K. A. A. Fox, R. M. Califf, and the ROCKET AF Steering Committee. 2011. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *New England Journal of Medicine* 365(10):883–891.

Pokorney, S. 2017. *Sentinel IMPACT-AFib: Transforming pragmatic clinical trials using a nationwide distributed claims database.* Presented at the Clinical Effectiveness Research Innovation Collaborative Meeting, Washington, DC.

Vestbo, J., D. Leather, N. Diar Bakerly, J. New, J. M. Gibson, S. McCorkindale, S. Collier, J. Crawford, L. Frith, C. Harvey, H. Svedsater, and A. Woodcock. 2016. Effectiveness of fluticasone furoate–vilanterol for COPD in clinical practice. *New England Journal of Medicine* 375(13):1253–1260.

Woodcock, A., J. Vestbo, N. D. Bakerly, J. New, J. M. Gibson, S. McCorkindale, R. Jones, S. Collier, J. Lay-Flurrie, L. Frith, L. Jacques, J. L. Fletcher, C. Harvey, H. Svedsater, D. Leather, D. Adams-Strump, L. S. Addlestone, A. Afshar, J. Amin, R. Archer, M. Austin, A. Bakhat, J. Behardien, J. M. Borg-Costanzi, G. Breen, N. Browne, C. Brunt, K. H. Buch, P. Budden, J. Chandy, A. Chaudhry, L. Cheema, N. Chennupati, S. Coulson, L. Cribbin, D. Dillon, A. El-Kafrawy, E. Elliott, B. Farooq, N. A. Finegan, P. Fink, A. Fletcher, S. Frier, C. Gibbons, L. Gill, D. Herron, B. Hope, R. E. Howard, C. Hughes, S. Iles, P. Jackson, M. Jarvis, V. Joshi, N. Kanumilli, R. A. Khan, M. Khan, S. Kwok, N. Lord, C. Mafunga, C. I. Malcomson, D. K. McCarthy, H. S. Milligan, P. Patel, S. J. Patel, V. B. Raj, K. A. Richardson, R. Salim, R. B. Seaton, D. Shah, M. Sharma, H. Singh, N. Smith, N. N. Smyrniou, M. Stamp, P. Stratford-Smith, M. Sultan, R. S. Sumra, J. Tankel, U. I. N. Umeadi, C. Westwood, J. White, H. C. Wilkinson, R. G. Wilson, S. A. Wright, and A. T. Wright. 2017. Effectiveness of fluticasone furoate plus vilanterol on asthma control in clinical practice: An open-label, parallel group, randomised controlled trial. *The Lancet* 390(10109):2247–2255.

Yih, W. K., T. A. Lieu, M. Kulldorff, D. Martin, C. N. McMahill-Walraven, R. Platt, N. Selvam, M. Selvan, G. M. Lee, and M. Nguyen. 2014. Intussusception risk after rotavirus vaccination in U.S. infants. *New England Journal of Medicine* 370(6):503–512.

DISCLAIMER: This Proceedings of a Workshop—in Brief was prepared by **Amanda Wagner Gee, Ben Kahn,** and **Carolyn Shore** as a factual summary of what occurred at the workshop. The statements made are those of the rapporteurs or individual workshop participants and do not necessarily represent the views of all workshop participants; the planning committee; or the National Academies of Sciences, Engineering, and Medicine.

REVIEWERS: To ensure that it meets institutional standards for quality and objectivity, this Proceedings of a Workshop—in Brief was reviewed by **John Doyle,** QuintilesIMS; and **Rachael Fleurence,** National Evaluation System for Health Technology (NEST) Coordinating Center. **Lauren Shern,** National Academies of Sciences, Engineering, and Medicine, served as the review coordinator.

SPONSOR: This workshop was partially supported by the U.S. Food and Drug Administration.

For additional information regarding the workshop, visit http://nationalacademies.org/hmd/Activities/Research/DrugForum/2017-SEP-19.aspx.

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2018. *Examining the impact of real-world evidence on medical product development: I. incentives: Proceedings of a workshop—in brief.* Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/25024.

Health and Medicine Division

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people's lives worldwide.
www.national-academies.org