Examining the Impact of Real-World Evidence on Medical Product Development

The Forum on Drug Discovery, Development, and Translation of the National Academies of Sciences, Engineering, and Medicine hosted a three-part workshop series in Washington, DC, Examining the Impact of Real-World Evidence (RWE) on Medical Product Development. The series, which was sponsored by the Food and Drug Administration (FDA), was designed to examine how RWE development and uptake could enhance medical product development and evaluation. Workshop participants discussed the current system of evidence generation and its limitations, shared lessons learned from successful initiatives that have incorporated RWE, and explored under what conditions RWE may be appropriate for informing medical product decision-making.

PROCEEDINGS OF A WORKSHOP SERIES

Examining the Impact of Real-World Evidence on Medical Product Development

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

- Workshop 1 (September 2017) focused on how to align incentives to support collection and use of RWE in health product review, payment, and delivery;
- Workshop 2 (March 2018) covered what types of real-world data (RWD) and RWE might be appropriate for specific purposes;
- Workshop 3 (July 2018) examined approaches for operationalizing the collection and use of RWE.

FDA Commissioner Scott Gottlieb spoke at workshop 1, acknowledging that while RWE may not replace data from traditional clinical trials in many cases, FDA is working to develop policies to achieve more appropriate adoption of RWE to support regulatory decision-making, including new indications for approved drugs. He emphasized the importance of expanding the use of RWE in ways that could make medical product development more efficient and cost effective.

STAKEHOLDER PERSPECTIVES ON RWE

Several workshop 1 participants, including representatives of payers, health care delivery systems, and patients, presented perspectives on incentives for using RWE. Michael Sherman, Harvard Pilgrim Health Care, highlighted that payers must find a balance between access and affordability while driving innovation. He suggested that in cases for which a product approval may be based on limited evidence, FDA could consider requiring manufacturers to enter into value-based agreements that tie reimbursement to performance and encouraged post-marketing collaborations between payers and pharmaceutical companies.

Michael Horberg, Kaiser Permanente (KP) Mid-Atlantic Permanente Medical Group, and Daniel Ford, Johns Hopkins Health System, described delivery system perspectives. They noted that delivery systems value medical practices that are supported by quality, relevant evidence that demonstrates value to patients and discussed typical evidence generation processes.

Sharon Terry, Genetic Alliance, explained that patient-generated data and community-led registries can be an important source of evidence generation because they focus on patient priorities and lived experiences. These data sources still require rigorous validation, she said, but they should be integrated into clinical decision-making.

LESSONS LEARNED FROM RWE INITIATIVES

In workshop 1, Martin Gibson and Marie Kane, Northwest EHealth, presented on the Salford Lung Studies, which are two late-phase randomized controlled trials (RCTs)—one for asthma and another for chronic obstructive pulmonary disease. These studies were the first to evaluate the effectiveness of a pre-license medication in a real-world setting. Gibson and Kane described broad stakeholder engagement as a reason for the success of the studies and the challenges of developing a suitable data platform.

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Richard Platt, Harvard Medical School, described Sentinel, an FDA monitoring system that uses electronic health data to support postmarketing medical product evaluation. He said the distributed system allows external data partners to retain private data prior to curation and can be used on its own or linked to other data sources, such as electronic health records (EHR) or patient-reported data.

Rachael Fleurence, National Evaluation System for Health Technology Coordinating Center, described the use of RWD and RWE for devices. She said that both are crucial for identifying problems with devices early in their use, and reliable RWE could shift device approval timelines and improve surveillance. Fleurence highlighted that registries are used widely for devices, and increased use of RWE could link existing registries to other data sources through Coordinated Registries Networks.

BARRIERS TO IMPLEMENTATION

Brian Bradbury and Elliot Levy, Amgen Inc., described barriers to RWD and RWE implementation, including a lack of: knowledge and awareness about RWE methods; capacity and expertise in relevant areas of research; and systems and processes to support RWE collection and use. Ford and John Doyle, IQVIA, identified RCT–RWE hybrid studies, such as pragmatic trials and cluster randomized designs, as possible approaches that combine advantages of both types of studies. Hui Cao, Novartis, suggested that evidence hierarchies that currently exist in medical product research could be revisited.

Marcus Wilson, HealthCore, described defragmentation as a process to integrate data sources from distinct stakeholders to provide a more complete understanding of a medical product. The process still requires data security and protection of patient privacy and business interests, he said. Anna McCollister-Slipp, Scripps Translational Science Institute, highlighted the lack of urgency around RWE adoption as problematic, as well as the hesitancy to include nontraditional stakeholders in research.

Addressing current evidence generation practices, Robert Califf, Verily Life Sciences, said the system should move past precision to focus on reliability. Potential steps to meet this goal could include the creation of a learning health care system, the use of quality by design, the use of automation, and operating from basic principles of scientific research, he said. Reflecting on the use of observational data networks, Patrick Ryan, Janssen, said analyses that incorporate the entire breadth of data on a particular set of medical products, including those that are not statistically significant, could be used to reflect a fuller understanding of those medical products.

Rory Collins, University of Oxford, focused on methods to improve RCTs, rather than replacing them with observational studies. He said RCTs are good at discovering moderate treatment effects and, while costly, innovative RCT designs that do not create data verification burdens could be useful. Janet Woodcock, FDA, ended the session by stating that the current evidence generation system needs improvement, and said opportunities to test product effectiveness using RWE could arise. She mentioned master protocols as a particular platform of interest.

PRACTICAL APPROACHES AND APPLICATIONS

While workshop 1 explored broad issues concerning barriers and incentives for the use of RWE, workshops 2 and 3 focused on specific questions stakeholders might consider before incorporating RWD and RWE into a study design. Questions raised at the first two workshops were incorporated into draft "decision aids" used to prompt further discussion during workshop 3 (to access the decision aids as well as additional details and resources, please see the Proceedings). The "decision aid" topics included: (1) when a particular real-world data element may be fit to assess study eligibility, treatment exposure, or outcomes; (2) some considerations for controlling or restricting treatment quality in real-world trials; (3) some considerations for obscuring intervention allocation in trials to generate RWE; and (4) potential ways to assess and minimize bias in observational comparisons.

WHEN CAN DECISION MAKERS RELY ON RWD?

At workshop 2, Adrian Hernandez, Duke University School of Medicine, presented on a suite of trials that compared novel oral anticoagulants (NOACs) to warfarin, all of which utilized RWD and consistently showed that NOACs were non-inferior to warfarin. He posed a question for consideration: What questions characterize the use of a RWD source and signal reliability before a study is performed? At workshop 3, Jeff Allen, Friends of Cancer Research, presented a pilot project that investigated the performance of real-world endpoints among patients with advanced non-small cell lung cancer treated with immune checkpoint inhibitors. The project demonstrated that several real-world endpoints correlate well with overall survival, and showed that overall survival rates assessed from EHR and claims data were consistent with rates observed in clinical trials.

Aylin Altan, OptumLabs, and Brande Yaist, Eli Lilly and Company, said the usefulness of an RWD source for a particular question depends on whether it has information about the correct population, exposures, and outcomes, and Platt, Yaist, and Robert Temple, FDA, pointed out that it may be acceptable for RWD to be of different quality for different purposes. Cao said accuracy of RWD varies predictably, depending on factors such as treatment administration methods or the outcomes being measured.

Hernandez and Gregory Simon, KP Washington Health Research Institute, said provider-collected RWD is affected by the experience of the provider and the incentives they face. Luca Foschini, Evidation Health, spoke about patient-generated health data, noting that—while it has the potential to answer difficult research questions, facilitate broader participation in health research, and incorporate new data sources—it is subject to different biases than data collected within the health care system.

Other workshop participants discussed issues with the analysis of RWD. Marc Berger, formerly of Pfizer, Grazyna Lieberman, Genentech, and Deven McGraw, Ciitizen, explained that data sharing and transparency in data curation and analysis could be improved to encourage broader use of

reliable RWE. Many speakers—including Altan, Berger, Foschini, Simon, and Yaist—pointed out that RWD can be affected by systemic and random bias, and are unique from other data sources because of their dynamic nature. Researchers can compensate for these, they said, but should remain mindful of potential biases when using RWD.

WHEN CAN DECISION MAKERS RELY ON REAL-WORLD TREATMENT?

At workshop 2, Ira Katz, Department of Veterans Affairs (VA), presented on a VA RCT that tested lithium as a treatment for suicide prevention. Katz described key questions that emerged through the study design process and emphasized the difficulty of making the trial generalizable to patients in real-world settings. At workshop 3, Larry Alphs, Newron Pharmaceuticals, presented on two real-world mental health trials (PRIDE and INTERCEPT) that grappled with issues around patient restriction to answer questions about safety and efficacy.

Horberg, Katz, Califf, and Alex London of Carnegie Mellon University discussed inclusion and exclusion criteria in real-world treatment settings. They argued for broadening these criteria in real-world trials to include patients with comorbidities or concomitant treatments to make the results more generalizable.

Alphs and Peter Stein, FDA, discussed a potential approach to choosing real-world trial restrictions, explaining that researchers could consider a specific set of categories that answer the research question while still honoring participant safety and autonomy. W. Benjamin Nowell, Global Healthy Living Foundation, also expressed concern about the role of patients in real-world research, emphasizing that research driven by patients is iterative and considers patients' needs, priorities, and experiences. The purpose of patient-centered research, he said, is to enable patients to make informed decisions about their own health care.

Alphs, Katz, and Simon described the role of researchers in real-world trials: Maintaining the trial protocol and caring for the well-being of patients, with patient safety coming first should the two conflict. Califf, Hernandez, and Stein discussed the importance, and ethical obligation, of setting a standard of care for the control arm of a study when designing real-world trials despite variance in standards across regions and treatment settings.

WHEN CAN DECISION MAKERS LEARN FROM REAL-WORLD TREATMENT ASSIGNMENT?

At workshop 3, Orly Vardeny, University of Minnesota and Minneapolis VA Center for Chronic Disease Outcomes Research, presented on the INVESTED trial, which explored the connection between influenza vaccine and cardiovascular events. The research team hypothesized that a stronger immune response from the high dose flu vaccine would translate into better cardiovascular outcomes, she said; they conducted a double-blinded RCT to prevent systematic biases inevitable in dispensing standard versus high dose vaccines.

Jonathan Watanabe, University of California, San Diego, and London said that blinding allows researchers to study the effects of an intervention without influence from patients or providers, but it may not always be appropriate or feasible. Cathy Critchlow of Amgen, Nancy Dreyer of IQVIA, and James Smith of FDA noted that the appropriateness of blinding is dependent on a study's context and uncertainties. These uncertainties, said London, can be classified along two axes: ensemble efficacy and utilization factors. The interaction of these two categories can indicate the appropriateness of blinding.

Dreyer, Rob Reynolds of Pfizer, and Smith explained that decisions on blinding can also be influenced by practical considerations, such as study cost, feasibility of masking treatment delivery, patient preferences, and data generalizability. Critchlow, Dreyer, John Graham of GlaxoSmithKline, and Smith said patient and provider bias can be difficult to predict, and it may not affect all outcomes, such as quantitative lab readings or all-cause mortality. However, it can affect subjective outcomes or have other effects such as in ascertainment or treatment bias, they said.

GAINING CONFIDENCE IN OBSERVATIONAL COMPARISIONS

At workshop 2, Sebastian Schneeweiss, Harvard Medical School, presented on the use of health care databases for regulatory decision-making. He explained that confidence in database studies is related to the type of effect being detected, and said such studies may be more appropriate when the outcomes and exposures are measurable in the data, when two active treatments are compared, and when the key confounding variables are measurable.

At workshop 3, Hector Izurieta, FDA, described a real-world study using Medicare Part D beneficiary data on the effectiveness and duration of effectiveness of the shingles vaccine, Zostavax. Izurieta explained how the investigators achieved balance between the treatment cohorts using propensity score matching and Mahalanobis metric matching, and conducted a secondary analysis to account for unmeasured confounders.

During discussion, David Madigan, Columbia University, noted that in disease areas for which RCTs are impractical, evidence from observational studies could be particularly valuable. Several participants discussed methods for observational data analysis. Madigan and Schneeweiss said transparent reporting of study methods can promote replicability and aid in assessing study validity. Speaking to a project currently under way, Jessica Franklin, Harvard Medical School, said replication of RCT results using observational databases can help establish criteria for conducting such studies more widely. Looking toward the future of observational studies, Javier Jiminez, Sanofi, and Mark van der Laan, University of California, Berkeley, said new methods such as predictive analytics and machine learning can potentially be used to predict outcomes for individual patients or identify associations.

Nicole Gormley and Heng Li, FDA, spoke from a regulatory perspective. Gormley described FDA's regulatory criteria for evaluating observational evidence: the data's relevance for a product's proposed indication; well-assessed outcomes; methods used to minimize bias; and rigorous statistical analysis.

REGULATORY PERSPECTIVES AND FUTURE OPPORTUNITIES

At workshop 3, Pall Jonsson, National Center for Health and Care Excellence (UK), described the Innovative Medicines Initiative GetReal project, explaining that health technology assessment relies on understanding the comparative effectiveness of new treatments. He noted that RWE can play a role in supplementing evidence from RCTs.

Komathi Stem, monARC Bionetworks, noted that using RWE can potentially engage patients more deeply in their care and in research, particularly with increases in usage of mobile technology and patients' ability to aggregate and store data about their own health. She explained that supporting a patient-centric shift in health research and care may require rethinking legislation, incentives, and partnerships. Levy described how new methods—such as adaptive designs, platform trials, or greater incorporation of RWE—have the potential to significantly reduce cost and time investments required for medical product development.

Concluding the workshop series, a panel of FDA leaders reacted to the workshop discussions. Jacqueline Corrigan-Curay, FDA Center for Drug Evaluation and Research (CDER), said CDER routinely uses RWE to support postmarketing safety evaluation and, to a limited extent, to evaluate effectiveness in certain rare diseases (including oncology). She emphasized that CDER's experience with Sentinel and other demonstration projects can inform policies going forward. Steve Anderson, FDA Center for Biologics Evaluation and Research (CBER), noted that CBER uses population-based data systems to conduct RWE safety and effectiveness studies, including the Biologics Effectiveness and Safety Sentinel Initiative to expand CBER's capabilities by providing data infrastructure, tools, and expertise.

Last, Jeffrey Shuren, FDA Center for Devices and Radiological Health (CDRH), said CDRH uses RWE in its product evaluations in pre- and postmarket decisions; it has started two programs combining registry data with other RWD to address regulatory needs. CDRH's 2017 RWE guidance, Shuren said, highlighted relevance and reliability as two critical considerations in evaluating RWE. All three FDA representatives said their Centers are interested in continuing to use RWE, but acknowledged that evidence used for regulatory purposes is necessarily different.

Mark McClellan, Duke-Margolis Center for Health Policy, touched on the idea of fit-for-purpose RWE in an environment with more readily available tools. He noted that clarity and specificity about when RWE is appropriate—and which data sources and methods are appropriate to address different types of questions—is the key to developing a framework for generating relevant evidence. Simon explained that, ultimately, delivering better health care to patients is the goal of using RWE.

Examining the impact of real-world evidence on medical product development: Proceedings of a workshop series can be purchased or downloaded from the National Academies Press, 500 Fifth Street, NW, Washington, DC 20001; (800) 624-6242; www.nap.edu.

For more information, visit www.nationalacademies.org/RWEseries.

Health and Medicine Division