**ISPE Comments on "Key Considerations in Using Real-World Evidence to Support Drug Development (Draft for Public Review)" by Center for Drug Evaluation, National Medical Product Administration, China**

August 27, 2019

The International Society for Pharmacoepidemiology (ISPE) is a global, nonprofit, professional membership organization dedicated to advancing public health by providing an open exchange forum within pharmacoepidemiology, the science applying epidemiological approaches to study the use, effectiveness and safety of drugs, vaccines, devices and other medical interventions in human populations. With more than 1,500 members from academia, government, service providers and the pharmaceutical industry in over 50 countries, ISPE has contributed to the development of policy, education and advocacy for the field for 30 years. Typical ISPE members are researchers with background and training in pharmacoepidemiology and the use of real-world data (RWD) collected as part of routine clinical care for the generation and interpretation of real-world evidence (RWE) intended for medical, payer and regulatory decision-making. ISPE has strong focus on data science and causal inference methods, and many ISPE members have contributed to the development of global data resources, helping set the data standards for post-market medical product assessment.

The "Key Considerations in Using Real-World Evidence to Support Drug Development" is a well-conceived and clearly-written document describing how RWD and RWE will be utilized for drug regulation in China. Important items to be considered in assessment of data quality and relevant research methods are described, and they are consistent with similar documents published by the US Food and Drug Administration (FDA) and in international scientific journals. ISPE sincerely appreciates the opportunity to provide written feedback to the document. The comments below reference the Line Numbers in the English version of the document, and they are grouped into two broad categories – **data** and **methods**.

I.     **RWD, Data Development and Validation**

RWD vs. Pivotal Clinical Trial Data

Research data collected in pivotal randomized clinical trials (RCT) are evaluated according to well established quality standards. In clinical trials that are compliant with Good Clinical Practice (GCP), data are manually compared against the source document, which are usually medical records. However, RWD are not necessarily GCP data and the evaluation of RWD quality should follow a different approach. Clarification on how GCP terminology may have different implications for RWD would be very helpful. Better yet, stating that RWD quality is evaluated according to a unique set of guidelines that are essentially different from GCP would be most appropriate.

## Data quality and heterogeneity

The proposed method for evaluating data quality depends on assessments of relevance and reliability, and the concept of accuracy is essential to both. A crucial element is that the clinical outcomes of interest that are being studied should be accurate and clinically significant (Line 165), and it will be desirable, perhaps in a follow-up document, to further specify how those outcomes could be evaluated. As RWD are derived from disparate sources, data quality may vary for different types of data. It is important to recognize the heterogeneity of RWD and different quality evaluation approaches will be required for different types of data.

## Data accuracy

Assessing the accuracy of RWD is proving to be challenging for all regulatory agencies. In the US, the FDA has launched or participated in several activities intended to evaluate the accuracy of RWD, such as the recent project by Friends of Cancer Research.[1] In this study which compared 6 sources of RWD with published data from traditional RCT of immune checkpoint inhibitors used in advanced non-small cell lung cancer, the RWD conclusions generally were consistent with the RCT data. The issues that generated the most lively discussions about accuracy, however, came from subtle differences between the RWD and the RCT source data. As Dr. Amy Abernathy, the now US FDA Principal Deputy Commissioner, commented at a public meeting in July 2018[2], some observed deviations of RWD from what was observed in the RCT were not wrong, but just different. For example, the time to treatment discontinuation was longer in real world studies than in the protocol-driven behavior seen in RCT because in real world settings, physicians commonly treat beyond cancer progression in hopes that longer treatment will trigger remission. True and important differences like this make it difficult to assess the accuracy of RWD, particularly in the context of their comparison "against authoritative sources" (line 179) like RCT, which will reflect protocol driven behavior and assessments that differ from those used in real world settings.

## Accuracy of wearable and other devices

The importance of calibrated measurement devices, like sphygmomanometers, was noted in the document. It is worth highlighting, particularly in the context of readily accessible wearables and other recorders (such as those connected with the Internet of Things and smartphones), that many measures may be available that have not been calibrated, but which may still be broadly indicative of the variable

---

[1] Stewart M, Norden AD, Dreyer N, et al. An Exploratory Analysis of Real-World Endpoints for Assessing Outcomes Among Immunotherapy-Treated Patients with Advanced Non-Small Cell Lung Cancer. **JCO Clinical Care Informatics** 2019. Epub July 23, 2019. https://ascopubs.org/doi/full/10.1200/CCI.18.00155

[2] https://www.focr.org/events/future-use-real-world-evidence (last accessed on August 26, 2019)

of interest. For example, home blood pressure machines are generally inexpensive, and while not as accurate as a calibrated device, can give consistent estimates that can classify participants as generally normotensive, hypotensive or hypertensive, at a low cost and with the potential for broad utility. Similarly, the health app on the iPhone measures steps and mileage, but couples who walk together may find quite different counts of steps and distance; nonetheless, those measurements could be used to classify study subjects' broad levels of physical activity since the devices clearly distinguish little physical activity from intense physical activity, and are probably more accurate than self-reports.

Patient Reported Outcomes

Patient Reported Outcomes (PROs) play critical roles in evaluation of drugs in real world settings, and the data could conveniently be entered by patients, families or caregivers via mobile devices or desktop computers, resulting in the term ePRO. Chinese versions of the relevant PRO instruments are already used in clinical trials. Technical challenges with more widespread use of ePROs to support drug regulations are well documented and highlighting how they fit into the current RWE framework would be helpful.

Source document

"Source document" is a GCP concept that fits into the traditional RCTs. During the trial, protocol-driven data are collected and are then verified against information in medical records. In the RWD era, "source document" becomes a much broader entity. Some types of RWD that may serve as RCT end points are described above, and instead of trying to implement the "source document verification" process into RWD, it will be useful to highlight the fact that RWD sources are heterogeneous and each type of RWD would require its own type of verification. Moreover, it is important to focus any source data verification on key exposure, outcomes, and confounders, rather than try to trace every single data element that may be of interest.

System validation

This concept is well-known to those involved in data systems to support RCTs and it can be expanded to the evaluation of different types of RWD, such as health insurance claims. For example, there may be built-in fraud detection and audits incorporated into health insurance claims processing, and evaluating how such quality assurance system may affect data quality would provide insight into the potential of using health insurance claims to generate RWE. With such system validation in place, there may not be a need for full scale data validation for individual studies.

For health insurance claims and electronic medical records, it will be useful to document health system elements that may compromise follow-up of individuals, and how those may have changed over time.[3] For example, patients discharged from a hospital after an acute event may not come back to the same hospital for long term follow-up, or a family may move from one province to another and the health insurance records in two locations may not be linked.

Data standard

While there is a Clinical Data Interchange Standards Consortium coding standard for clinical trial data, RWD are derived from disparate sources. Some RWD come with standard coding schemes, while other data sources, such as data generated from wearable and other devices, are newly developed and coding standards may be evolving.

II.      Research Methods and applications

The document offers several scenarios where RWE may be used to support drug development and regulatory decisions, including use of parallel external controls (also known as "contemporary external comparators") for approval of rare disease treatments where randomization is not feasible or ethical. The use of RWE for label expansions and especially their use for evaluating drug combination labeling, pediatric use, and subgroups of special interest are all consistent with objectives of RWE use specified in US FDA documents.

China is in a unique position to evaluate traditional herbal medications among large numbers of patients in a systematic manner with real world studies and a large volume of high quality evidence about these regimens will be generated. The fact that most formulations are prepared locally at individual hospitals would be a challenge. Regardless, there are statistical methods to take into account the batch-to-batch variability of the formulations and the between-hospital variability. Within the same section in the document there is a clear and broadly useful decision tree on where pragmatic trials fit in the context of real world studies and the section also describes the utility of RWD for guiding clinical trial design and supporting more precise identification of target populations of interest.

The document also does an excellent job of describing the basics of real world research design, including design choices for pragmatic trials (e.g., individual randomization or allocation per pre-defined criteria

---

[3] Bourke A, Bate A, Sauer BC, et al. Evidence generation from healthcare databases: recommendations for managing change. **Pharmacoepidemiol Drug Saf** 2016;25:749-54

pharmacoepi.org

[line 383]), with a statement that blinding is not used in most pragmatic trials. Further, the description of external comparators states that "the use of parallel [contemporary] external controls is generally superior to historical controls" (line 411). The robust and sensible endorsement of the desirability of recent RWD is an important concept that has not been described nearly as clearly by regulatory agencies in other countries.

## Difference between RWE findings and RCT findings

The fact that RWD may be different from RCT data was described above. Similarly, research findings from RCT may be different from those of real world studies. Although RCT findings are widely regarded as the gold standard, an observed difference between the two on the same clinical topic may not mean that the RWE is invalid, as the two sources of data may be answering different questions. If there are differences between findings from the two approaches, it will be most useful to evaluate the reasons behind the differences; whether it is due to differences in study population, how treatment regimens are used in typical care settings, operational definitions of study parameters, or it is a result of intractable confounding in observational studies. Ongoing research, such as the US FDA-funded RCT DUPLICATE program (rctduplicate.org), will provide more insight in the near future.

## Sensitivity analysis

Sensitivity Analysis, also known as quantitative bias analysis, has been recognized as an important tool in evaluating RWE. The traditional RCT paradigm is to have one set of primary analysis to guard against Type I error in statistical inference. In RCT, sensitivity analyses are sometimes carried out in cases of loss-to-follow-up or missing data as supplementary analyses. With RWD, sensitivity analysis could be applied more extensively, ranging from operational definitions for exposures and outcomes of interest to potential confounders, in order to identify key factors that may affect the main findings. Judicious use of sensitivity analysis can further demonstrate robustness of RWE.

## Protocol development and transparency

The document provides some very reasonable, high level guidance for evaluating RWE, including the importance of developing protocols and analysis plans in advance of study conduct. However, the document would benefit from highlighting the concept of transparency. It is worth noting that it is difficult to evaluate whether a protocol and statistical analysis plan were created before the study was conducted unless such documents are made available through a public website or publicly-accessible registry, like the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. An

article prepared by a joint task force from two professional societies – International Society for Pharmacoeconomics and Outcomes Research and ISPE has addressed this issue.[4]

Overall, this document is clear, well-constructed and provides pragmatic guidance including its request for consultation in advance of conducting real world studies to support regulatory decisions. This clarity should make it easier for pharmaceutical companies and others to bring well-thought out designs to the Center for Drug Evaluation (CDE) for review before execution, since the broad specifications are well laid out here.

In the future ISPE will humbly appreciate the opportunity to provide further professional input to CDE on RWD- and RWE-related topics.

Respectfully prepared by:

K. Arnold Chan, MD, ScD, FISPE

Director, Health Data Research Center, National Taiwan University

Nancy A. Dreyer, PhD, FISPE

Chief Scientific Officer, IQVIA Real World Solutions

---

[4] Wang SV, Brown J, Gagne JJ, et al. Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies. **Value Health** 2017;20:1009–22

pharmacoepi.org