



International Society
for Pharmacoepidemiology

ISPE COMMENTS ON THE FRAMEWORK FOR FDA'S REAL WORLD EVIDENCE PROGRAM

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, 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 resources such as the US Sentinel System, helping set the data standards for post-market medical product safety assessment.

RWE conceptually overlaps with Observational Research, an area in which ISPE members have been pioneers in developing and implementing improved study designs and methodologies for causal inference, such as pragmatic trials and hybrid studies (interventional/non-interventional), cluster randomized trials, self-controlled designs, use of negative controls or combined primary-secondary data source designs, and the New User Active Comparator design for retrospective studies. While randomized controlled clinical trials (RCTs) have and will continue to serve as the basis for evidentiary standards for most products, RWE and observational studies, with proper study design and methods for bias control for specific scientific questions, can result in balanced comparator groups and provide valid, reliable evidence to allow regulatory decision-making. Observational studies often address some of the known limitations of clinical trials (lack of generalizability of the narrow population studied, not reflective of routine medical practice), by studying broader populations in a routine clinical practice setting. In other circumstances, pragmatic randomized trials (pRCTs) harness advantages of both interventional and non-interventional designs (e.g., better baseline comparability of groups, but with broader population coverage) to address some of the limitations of RCTs, and masking (i.e., blinding) of key aspects of the pRCT can contribute to even greater validity without sacrificing their pragmatic nature [1].

The Framework for the Food and Drug Administration's (FDA's) RWE Program was reviewed by an ISPE RWE task force, convened to ensure clear communication around the importance of RWE to its members and external groups. This task force included representatives from academia, service providers, and the pharmaceutical industry in North America and Europe. We are pleased to have this opportunity to provide the FDA with recommendations to help improve the development of the planned RWE guidance(s) to be issued in 2021. No representatives from the FDA and EMA were involved in the development of these comments. These comments have been reviewed and endorsed by the ISPE Executive Committee on July 11, 2019.



ISPE commends the FDA for their response to the 21st Cures Act of 2016, in providing a draft framework on the fit-for-purpose use of RWE for regulatory purposes. While FDA has relied on RWE for safety evaluations for many years, it is encouraging to see the efforts of FDA to consider RWE for effectiveness to support new indications, labeling changes, and potentially for approval of new products.

Specifically, ISPE would like to suggest the following enhancements to the 2018 draft RWE Framework and forthcoming 2021 guidance(s) as areas for further clarity and definition to optimize their utility and facilitate consistency for Sponsors and FDA:

1. **What constitutes “evidentiary standards”** for submissions that include RWE? It would be helpful for this to be outlined by regulatory purpose, type of research question and study design.
2. More specific guidance is needed on **how to assess whether RWD is ‘fit for purpose’** and what constitutes ‘adequate’ RWD, balancing potential biases with effect size and better generalizability and representativeness. We recognize that this evaluation is unique to each scientific and/or regulatory question (i.e., a one size fits all database or study design cannot exist) [2]. A structured framework is therefore needed to guide both FDA reviewers and applicants on what may be ‘fit for purpose’ in the context of the specific research question and regulatory decision. In other words, what are the criteria that FDA intends to use to determine whether RWD is adequate for specific regulatory purposes, e.g., new indication, label change? Specifically, we recommend clarity on the following elements:
 - **Validation of Outcomes, Exposures, and Key Covariates.** To address issues around missing data, partial data, and data quality generally, are there preferred validation methods, particularly when using electronic medical records (EMR) and/or claims data? If it depends on the scientific regulatory question, guidance should be given on how to determine the adequate validation level.

Some common examples of issues in RWD studies include: lack of specific dosage information or lab/imaging test results; level of missing data and out of network health care encounters; the probability that the lack of specific data will impact the results in a meaningful way; and how and when validation of coding algorithms (“computable phenotypes”) is needed.

We recommend the FDA provide guidance on suggested minimum acceptable standards, including approaches to assess the potential impact of bias, such as e.g. sensitivity analyses, quantitative bias analyses [3,4,5], triangulation [2], replication/duplication, use of negative/positive exposure and outcome controls [6,7,8], or use of expert adjudicators, approaches that are commonly used by pharmacoepidemiologists to ensure interpretability and robustness of results [5,6,7,8]

- **Confounding Control.** For non-randomized studies, guidance would be helpful in terms of how the FDA will determine the adequacy of control for potential confounding, including confounding by indication or frailty, and unmeasured confounding, an area where ISPE and its membership have contributed a great deal to methodological advances. While this is a complex issue, guidance outlining the FDA view would be beneficial on key issues such as the use of propensity score approaches (matching, stratification and weighting) as well as other techniques for confounding control. Understanding FDA's position on specific detailed issues would also be helpful, such as: Is complete overlap of propensity score distributions (when used) sufficient after matching or other balancing techniques? Is it sufficient to have absolute standardized mean differences for all key individual confounders all <0.2 and predominantly <0.1 ? Is asymmetric trimming acceptable as it can help identify patients with contraindications or off-label use?
3. **Representativeness of the population of interest** is often brought up as justification for RWE and pRCTs. However, it is not necessarily desired in all cases [9] and having a strong scientifically valid study that allows interpretable treatment effect conclusions may outweigh the need to be 'representative'. In fact, classic RCTs are a prime example of this balance. We know that in many cases, RWD used to address research questions rely on either a commercially insured population, Medicare beneficiaries or specific health care systems treating patients who are seeking care. Interpretation must be made based on the population studied, as is the case with any RCT.
 4. We recommend in general a focus on effect estimates and their precision [10,11,12], not statistical significance and p-values. This is especially important in the context of rare diseases where adequate sample sizes will never be achievable. It is also particularly relevant to RWE in common diseases where non-randomized studies can have large sample sizes and very strong statistical significance for trivial effects sizes, in which case clinical significance and evaluation of potential bias should provide context.
 5. Guidance would also be beneficial around the use of medical records and related expectations for quality assurance (QA) and quality control (QC) of RWD sources, such as accurate linkage, and source verification. In many cases, when using EMR, RCT-style verification of RWD is impractical because the EMR is the actual data source. We recommend the FDA provide minimum acceptable standards for QA/QC and advocate approaches such as sensitivity analyses, linkage to other data sources or use of expert adjudicators, all of which are commonly used by pharmacoepidemiologists to ensure interpretability and robustness of results. In addition, more clarity is needed on the processes for FDA inspections and/or data set transfer to the FDA for reanalysis when the RWD sources are proprietary and external to both the Sponsor and FDA. Consideration should be given to the increasing use of "non-traditional" data sources such as actigraphy (e.g., "Fitbit™") or geocoding and how QA/QC processes would verify they are regulatory grade data.

6. Underpinning all of this is the need for meaningful transparency. This may include public registration of protocols. Protocols for RWE studies intended for regulatory purposes should include scientific regulatory objectives, pre-specified hypotheses, pre-specified analytic plans including pre-specified sensitivity analyses, and operational issues such as transformation of variables and handling of missing data [13,14,15,16,17]. In addition, the greater variability and heterogeneity in RWD should be recognized, and any unplanned deviations from the protocol along with the rationale should be adequately documented.
7. Additional guidance on the use of external control groups derived from RWD for single arm trials, is urgently needed. The guidance should address topics such as how to overcome difficulties in reliably selecting a comparable population, lack of standardized diagnostic criteria and equivalent outcome measures.

Obviously, the right expertise is needed to adequately plan, assess, and interpret RWD/RWE. ISPE would welcome the opportunity to work with the FDA to support RWE regulatory guidance document development and training of internal FDA staff on the use of RWE for decision-making, to help to ensure consistency across review divisions. We would encourage leveraging the existing internal expertise, particularly within OSE/OMP (many of whom are ISPE members), who have been using RWD for safety evaluation for decades, when expanding to effectiveness evaluations.

Finally, given the expertise of pharmacoepidemiologists and the ISPE membership in dealing with these issues for decades, we would welcome the opportunity to have direct dialogue with the FDA through public workshops, webinars, meetings or other forums to provide more extensive methodological input into the development of the forthcoming RWE 2021 guidance. Furthermore, ISPE would welcome the opportunity to partner with the FDA throughout the development and implementation of the guidance to help support FDA's various RWE initiatives and goals.

Respectfully submitted,

Drs. Rachel E. Sobel and Cynthia Girman, on behalf of the ISPE Real World Evidence Task Force

ISPE is a multi-national voluntary and professional society. It has adopted no official position on these matters. Only the Executive Committee of ISPE has endorsed these comments.

Authors

Rachel E. Sobel, DrPH, FISPE, Pfizer Epidemiology, NY, NY, USA

Cynthia Girman, DrPH, FISPE, CERobs Consulting, LLC, Chapel Hill, NC, USA

Vera Ehrenstein, MPH, DSc, Aarhus University, Denmark

Darren Toh, ScD, Harvard Medical School & Harvard Pilgrim Health Care Institute, Boston, MA, USA

Montse Soriano-Gabarró, MD, MSc, Bayer Epidemiology, Berlin, Germany

Fredrik Nyberg, MPH PhD MD, Astra Zeneca, Mölndal, Sweden

Alison Bourke, MSc, FISPE, President
Almut Winterstein, RPh, PhD, FISPE, President-Elect
Jesper Hallas, FISPE, Immediate Past President
Krista Huybrechts, FISPE, Vice President Finance
For the ISPE Executive Committee

¹ Christian, J. B., Brouwer, E. S., Girman, C. J., Bennett, D., Davis, K. J., & Dreyer, N. A. (2019). Masking in Pragmatic Trials: Who, What, and When to Blind. *Therapeutic Innovation & Regulatory Science*.
<https://doi.org/10.1177/2168479019843129>

² Girman CJ, Ritchey ME, Zhou W, Dreyer NA. Considerations in characterizing real-world data relevance and quality for regulatory purposes: A commentary. *Pharmacoepidemiol Drug Saf* 2018; 1-4.
<https://doi.org/10.1002/pds.4697>

³ Gatto NM, Reynolds RF, Campbell UB. A structured preapproval and postapproval comparative study design framework to generate valid and transparent real-world evidence for regulatory decisions. *Clin Pharmacol Ther* 2019; doi: 10/1002/cpt.1480

⁴ Franklin JM, Glynn RJ, Martin D, Schneeweiss S. Evaluating the use of non-randomized real-world data analyses for regulatory decision-making. *Clin Pharmacol Ther* 2019; 105(4); 867-87.

⁵ Funk MJ, Landi SN. Misclassification in Administrative Claims Data: Quantifying the Impact on Treatment Effect Estimates. *Curr Epidemiol Rep* 2014; 1(4): 175-185.

⁶ Sanderson E, Macdonald-Wallis C, Davey Smith G. Negative control exposure studies in the presence of measurement error: implications for attempted effect estimate calibration. *Int J Epidemiol*. 2018;47(2):587–596. doi:10.1093/ije/dyx213

⁷ Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies [published correction appears in *Epidemiology*. 2010 Jul;21(4):589]. *Epidemiology*. 2010;21(3):383–388. doi:10.1097/EDE.0b013e3181d61eeb

⁸ Eric Tchetgen Tchetgen, The Control Outcome Calibration Approach for Causal Inference With Unobserved Confounding, *American Journal of Epidemiology*, Volume 179, Issue 5, 1 March 2014, Pages 633–640, <https://doi.org/10.1093/aje/kwt303>

⁹ Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42(4):1012–1014. doi:10.1093/ije/dys223

¹⁰ Rothman, K and Greenland, S. Planning Study Size Based on Precision Rather Than Power. *Epidemiology*. 2018 Sep;29(5):599-603. doi: 10.1097/EDE.0000000000000876.

¹¹ ISPE guidance on p-values. <https://pharmacoepi.org/pub/?id=74e3d8d1-a054-7717-199a-9758f7f18dcf>



¹² Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature* 2019; 567: 305-307

¹³ Good Pharmacoepidemiology Practices. <https://www.ispor.org/heor-resources/good-practices-for-outcomes-research/article/good-practices-for-real-world-data-studies-of-treatment-and-or-comparative-effectiveness-recommendations-from-the-joint-ispor-ispe-special-task-force-on-real-world-evidence-in-health-care-decision-making>.

¹⁴ Berger ML, Sox H, Willke RJ, Brixner DL, Eichler H-G, Goettsch W et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the Joint ISPOR-ISPE Special Task Force on Real World Evidence in Health Care Decision-Marking. *PDS* 2017; 26(9): 1033-9. <https://www.ispor.org/heor-resources/good-practices-for-outcomes-research/article/good-research-practices-for-comparative-effectiveness-research-defining-reporting-and-interpreting-non-randomized-studies-of-treatment-effects-using-secondary-databases>.

¹⁵ Wang SV, Schneeweiss S, Berger ML. Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Value Health*. 2017;20(8):1009-1022. <https://www.ispor.org/heor-resources/good-practices-for-outcomes-research/article/reporting-to-improve-reproducibility-and-facilitate-validity-assessment-for-healthcare-database-studies-v1.0>.

¹⁶ ENCePP Guide on Methodological Standards in Pharmacoepidemiology. http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml

¹⁷ Dreyer NA, Schneeweiss S, McNeil B, Berger ML, Walker A, Ollendorf D, Gliklich RE: GRACE Principles: Recognizing high-quality observational studies of comparative effectiveness. *American Journal of Managed Care* 2010;16(6):467-471.