Received: 12 November 2021 | Revised: 16 February 2022 | Accepted: 22 April 2022

DOI: 10.1002/pds.5444

COMMENTARY

WILEY

Real-world data: Assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products

Cynthia J. Girman[1] | Mary E. Ritchey[1,2,3] | Vincent Lo Re III[4]

[1]Real World Evidence and Patient Outcomes, CERobs Consulting, LLC, Wrightsville Beach, North Carolina, USA

[2]Med Tech Epi, LLC, Philadelphia, Pennsylvania, USA

[3]Center for Pharmacoepidemiology & Treatment Science, Rutgers University, New Brunswick, New Jersey, USA

[4]Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence
Cynthia J. Girman, CERobs Consulting, LLC, 2612 Lumina Ave North, Wrightsville Beach, NC 28480, USA.
Email: cindy.girman@cerobs.com

1 | INTRODUCTION

On September 29, 2021, the U.S. Food and Drug Administration (FDA) released a draft guidance discussing "considerations when proposing to use electronic health records (EHR) or medical claims data in clinical studies to support regulatory decisions on effectiveness and safety."[1] The guidance is a result of years of FDA consideration of use of real-world data (RWD), including many stakeholder discussions. It addresses the adequacy and relevance of RWD to address specific research questions and validation and data quality/provenance of study question elements.[1] In addition, the guidance considers whether follow-up in RWD sources is sufficient to ascertain outcomes and if missing data limit interpretation of results.[1–4] Notably, FDA recommendations do not appear to be limited to new product indications or labeling changes, nor to non-interventional studies. Study design and analytic techniques, including handling of confounding will be addressed in forthcoming FDA documents. While this draft guidance moves the pharmaceutical industry closer to the use of RWD for regulatory decisions, it appears to incorporate inefficiencies common in randomized clinical trials (RCTs) into studies using RWD and attempts to be broadly applicable to studies in EHR and claims- and to RCTs and observational studies- without considering important differences between them. Here, we evaluate and provide commentary on the draft guidance with respect to sponsors aiming to use RWD sources to support regulatory decisions and labeling. We focus on key components of the guidance: adequacy and relevance of the data source, missing data, data linkage, and the ascertainment and validation of study variables along with quality assurance (QA). For illustrative purposes, we contrast the recommendations in this guidance for RWD studies to existing requirements for RCTs.

2 | ADEQUACY AND RELEVANCE OF THE DATA SOURCE

While all scientific research should pre-specify objectives, study design, and statistical analysis plans in a (preferably registered) protocol,[5] we agree with FDA that studies using claims or EHR should also detail the conceptual and operational definitions of all elements of the research question, possibly using the PICOTS framework as previously suggested.[3,4] As FDA stated, evaluation of adequacy of a data source must be in the context of the research question and regulatory decision (e.g., approval, labeling change, new indication) to be made.[2,3,6,7] The FDA could acknowledge that no RWD source is entirely complete with all variables that might be of interest, but RWD can be useful despite some missing data or variables. Lack of less important covariates may not matter and quantitative bias analyses (QBA) and analyses assessing sensitivity of findings to varying operational definitions and assumptions regarding missing data should be encouraged and pre-specified.[8–10] For example, acceptance of RWD research findings might depend on whether sensitivity analyses and QBA suggest minimal change when accounting for bias, missing

This Commentary is endorsed by the International Society for Pharmacoepidemiology (ISPE).

data or varying operational definitions, particularly if the effect size is moderately large.

3 | MISSING DATA

In Section V.C.2 and V.C.3 of the guidance, the FDA appropriately expresses concerns about missing data for prescriptions (e.g., samples, out-of-pocket purchase) and outcomes (e.g., out-of-network care), noting that linkage (discussed here in Section 4) to data sources may enhance capture of these elements.[1] The guidance should have noted that additional sensitivity analyses, akin to those in RCTs for missing exposure or outcome data, should be encouraged to understand whether results would have changed if such missing data were available.

Section IV.A. of the draft FDA guidance appropriately highlights critical information not in EHR or claims data that pharmacoepidemiologists routinely consider when assessing data sources for specific research questions, such as relevant formulary restrictions, tiering or stepped therapy, and prior authorizations that limit medication use.[1] Variation in medical practices, diagnostic criteria, and treatment patterns across health care systems are also acknowledged as important considerations. Because U.S. residents often switch medical insurance with change in employer, continuity of care is critical to allow sufficient longitudinal follow-up (person-time) and capture of study outcomes. Such longitudinal follow-up might not be achievable without linkage to other data sources.

4 | LINKAGE

To supplement EHR or claims data when key data elements are missing or poorly measured, the FDA guidance recommends linkage to other data sources (see FDA guidance Section IV.B.2). While important to consider, linkage may not always be an option due to lack of linking variables, privacy issues, legal considerations for sharing data, and/or data integrity. Even with linkage, the resulting cohort can be limited due to loss of adequate sample size and potential loss of generalizability. The FDA guidance acknowledges that information for some data elements may be contained in unstructured fields within EHR data (e.g., clinical notes), and recommends thoroughly describing the process for extracting and verifying such data, whether manually or using automated technology. However, this process is not detailed in the guidance.

5 | CHARACTERIZATION AND VALIDATION OF EXPOSURES, OUTCOMES, AND IMPORTANT COVARIATES AND QA

In Section V.C.1, the FDA guidance indicates that the medication exposure definition should include dose, formulation, strength, route of administration, timing, frequency, duration studied (if relevant), and

TABLE 1 FDA expectations when using real-world data to evaluate medical product effectiveness or safety compared to using randomized clinical trials for product approval.

FDA recommendation/ requirement	Recommendation in FDA guidance for effectiveness/ safety studies in EHR/claims[a]	Requirements for RCTs for product approval[b,c]
Pre-specified protocol and SAP	Yes	Yes
Pre-specified sensitivity and subgroup analyses	Yes	Yes
Definitions of outcomes	Yes	Yes
Verification of outcomes	"Most rigorous approach"	Dependent on study outcome, may or may not be required
Validation of outcomes	Yes	For outcomes trials; adjudication may be used
Validation of variables to define study population	Yes	No
Validation of treatment definitions	Yes	Data collection and pill counts; crossover assessed
Validation of covariates	Yes	No
QA/QC at time of data collection	Yes, but may be impractical for Sponsors (and Data Providers) to implement	Yes
QA/QC at data checking/ cleaning	Yes, but impractical for Sponsors to implement; documentation from data provider may not be obtainable	Yes, procedures documented
QA/QC at transformation to analytic file	Yes	Yes
Traceability/ Auditable	Yes, but detailed documentation from data provider may not be obtainable	Yes

Abbreviations: FDA, US Food and Drug Administration; QA/QC, quality assurance/quality control; RCT, randomized controlled trial; SAP, statistical analysis plan.

[a] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory.

[b] https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials.

[c] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trials-guidance-documents.

potentially manufacturer, some of which may not be available in EHR or claims databases. FDA should encourage consideration of just how crucial each of these exposure attributes is to the *specific* research question.

FDA pointed out that algorithms to identify exposures, outcomes, and important covariates should be operationally defined by data such as medical diagnoses, procedural codes, or dispensed medications. As with clinical assessments in RCTs, algorithms may not always be accurate, which can lead to misclassification. However, misclassification may or may not lead to biased estimates of treatment effect, and FDA should encourage proactive sensitivity analyses and QBA to inform the potential impact of such biases,[8–10] particularly when effect size is modest.

The guidance outlines the need to have documentation on all QA procedures during data accrual, curation, and transformation to the final analytic dataset (see FDA guidance Section VI.B). While data curation and provenance are important, it is extremely challenging for sponsors to obtain detailed documentation and verify data against original source records when using commercially licensed databases. Guidance for industry should focus on QA that sponsors can readily influence and document from the time of the receipt of data, and QA procedures should be driven by how potential misclassification, errors and bias might affect the conclusions of the study. Stringent QA, auditing and data provision requirements for all covariates, regardless of their association with outcomes, significantly increase inefficiencies, in contrast to the mandate to FDA to assess ways to use RWD to increase efficiencies of clinical research. Instead, FDA should recommend risk-based monitoring[11] and work with data providers to encourage transparency of QA procedures and to replicate analyses.

In general, the FDA draft guidance describes QA processes for data accrual and curation that are similar to or even beyond the requirements for primary data collection in phase 3 RCTs supporting product approval (see Table 1). For example, throughout the guidance, FDA refers to complete (100%) verification of all study variables (PICO and covariates) as the "most rigorous approach" to using RWD, acknowledging that it may not be feasible. The use of the term 'verification' may be conflated with 'validation.' Data *verification* typically involves confirming the correct data point for each patient (or subset), which is impossible for *all* variables in any RWD study. Such an approach is not expected for RCTs nor is it performed by the FDA in the Sentinel System to evaluate safety signals. In contrast, *validation* assesses the accuracy of an operational definition compared to a reference standard to yield test characteristics such as positive and negative predictive value, sensitivity, and specificity, contributing to understanding of potential misclassification. If the reference standard is the medical record, the guidance appropriately recommends blinding data abstractors and adjudicators to medication exposure to reduce bias and using standardized and reproducible abstraction and adjudication processes to minimize intra- and inter-rater error. However, the guidance should clarify the appropriate validation approaches to consider when EHR data are used for the study and the same EHR *is* the medical record reference standard. Of note is that these same EHR data are often used as the source of data collected in RCTs.

The guidance fails to distinguish between or prioritize critical variables in recommending validation. In RCTs, data are typically collected at each clinical site through direct assessment of a patient or from medical records. While procedures are available to ensure the accuracy of these data, risk-based quality assurance procedures drive focus on more critical variables,[11] and 100% verification is rare. Achieving 100% accuracy (no misclassification) takes extraordinary effort and expense and may never be achievable. Selecting only patients whose outcome can be 'validated' could result in a biased sample of patients with more complete records, decreasing generalizability. Validating each covariate in a study is inefficient, costly, and will yield little that might change the overall interpretation of the study. Instead, sponsors should justify the adequacy of critical study elements (exposure, primary outcomes, and key confounders) and perform analyses to assess the sensitivity of results to varying degrees of potential misclassification.

In Section IV.D, the FDA guidance suggests that the performance of algorithms or operational definitions for key variables should be demonstrated using sufficiently large samples, appropriate sampling techniques, and reasonable reference standards, but gives little guidance on how to define a 'sufficiently large' sample or the appropriate sampling methods. Appropriate reference standards are also not recommended.

FDA states that studies to validate health outcomes of interest can be performed, or a prior study or publication can be used if conducted in a "similar population or data source." The accuracy of case-identifying algorithms can vary across populations, healthcare settings, coding systems, or calendar time periods, and each of these factors can affect an algorithm's performance and transportability to other health databases. Thus, when considering applying algorithms in a different setting or database, researchers should consider whether the outcome prevalence differs, and whether restrictions on formulary access to medications are comparable between the validation population and the study population. The level of documentation on these points could be clarified.

FDA states that validation may be performed in cases when only false-positives are of concern, and in cases and non-cases when false-negatives are also of concern. However, validating non-cases in very large studies of rare outcomes may be impractical; a random sample of non-cases may be more practical.[12]

6 | CONCLUSION

The draft guidance outlining specific issues of concern to FDA is a welcome leap forward to carve a path for use of RWD to support regulatory decision-making. Throughout the draft guidance, FDA recommends that Sponsors discuss specific issues with the relevant review division. However, the process and timing for such discussions need explicit definition.

While providing a potential path forward, the path is arduous and unlikely to address the current inefficiencies of RCTs. The guidance recommends practices that would significantly limit the ability to use

RWD to generate reliable results of effectiveness and safety for regulatory purposes and imposes numerous processes for RCTs on observational studies. The guidance could be enhanced by implementing a risk-based approach to prioritize operational definition validation and QA as well as advising on the types of sensitivity analyses and QBA that would give FDA more confidence in the findings.

The pharmaceutical industry and regulators should advocate for more transparency in metrics and data curation procedures from data providers, and for routine agreement to provide patient-level data to FDA to support submissions from around the world. Resources available to assist in the identification of decision-grade, fit for purpose data and methods for RWE studies should be referenced.[2,4,7,13–15] Similarly, recommendations for justifying adequacy of unstructured data in RWE studies would be helpful. These additions would better address the congressional charge to FDA to identify ways that RWD can be used to *more efficiently and more rapidly generate evidence on product effectiveness and safety*.

ACKNOWLEDGMENTS

The authors would like to thank Molly Aldridge, MPH for her assistance with careful editing of the Commentary before and after response to reviewer comments, and the reviewers for their thoughtful comments to improve the Commentary. Authors acknowledge additional comments from review by the RWE & Regulatory Decisions workgroup of the ISPE RWE Task Force and ISPE members, particularly Drs. Kim Brodovicz, Cathy Anne Pinto, Magdalene Assimon, Helga Gardarsdottir and Rachel Sobel.

ORCID

Cynthia J. Girman https://orcid.org/0000-0002-5714-174X
Mary E. Ritchey https://orcid.org/0000-0003-0304-9304
Vincent Lo Re III https://orcid.org/0000-0001-7955-0600

REFERENCES

1. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) and Oncology Center for Excellence (OCE). Real-World Data: Assessing electronic health records and medical claims to support regulatory decision-making for drug and biological products: Guidance for Industry, Draft Guidance. Sept 2021. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory.
2. Daniel G, Silcox C, Bryan J, McClellan M, Romine M, Frank K. Characterizing RWD quality and relevancy for regulatory purposes. Duke-Margolis Center for Health Policy. Retrieved October 01, 2018. https://healthpolicy.duke.edu/sites/default/files/atoms/files/characterizing_rwd.pdf.
3. 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*. 2019;28(4):439-442. doi:10.1002/pdf.4697
4. Ritchey ME, Girman CJ. Evaluating the feasibility of electronic health records and claims data sources for specific research purposes. *Ther Innov Reg Sci*. 2020;54:1296-1302. doi:10.1007/s43441-020-00139-x
5. Wang SV, Patterson OV, Gagne JJ, et al. Transparent reporting on research using unstructured electronic health record data to generate 'real world' evidence of comparative effectiveness and safety. *Drug Saf*. 2019;42(11):1297-1309.
6. https://www.sentinelinitiative.org. Accessed Oct 3, 2021.
7. Girman CJ, Ritchey ME, McNeill AM, Andersson Sundell K, Meyer RJ. Demonstrating real-world evidence is fit for purpose to support labeling: parallels in pursuing labeling claims for patient reported outcomes. *Ther Innov Reg Sci*. 2021;55:561-567. doi:10.1007/s43441-020-00252-x
8. Girman CJ, Faries D, Ryan P, et al. Pre-study feasibility and identifying sensitivity analyses for protocol pre-specification in comparative effectiveness research. *J Comp Effect Res*. 2014;3(3):259-270.
9. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. Springer; 2009.
10. Jonsson-Funk M, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effects. *Curr Epidemiol Rep*. 2014;1:175-185.
11. Hurley C, Shiely F, Power J, et al. Risk based monitoring (RBM) for clinical trials: a systematic review. *Contemp Clin Trials*. 2016;51:15-27.
12. Thurin NH, Boxco-Levy P, Blin P, et al. Intra-database validation of case-identifying algorithms using reconstituted electronic health records from healthcare claims data. *BMC Med Res Methodol*. 2021;21:95. doi:10.1186/s12874-021-01285-y
13. 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 Pharm Ther*. 2019;106:103-115. doi:10.1002/cpt.1480
14. Bykov K, Patorno E, D'Andrea E, et al. Prevalence of avoidable and bias-inflicting methodological pitfalls in real-world studies of medication safety and effectiveness. *Clin Pharm Ther*. 2021;97:125-127. doi:10.1002/cpt.23
15. Hall GC, Sauer B, Bourke A, Brown JS, Reynolds MW, Casale RL. Guidelines for good database selection and use in pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf*. 2012;21(1):1-10.

How to cite this article: Girman CJ, Ritchey ME, Lo Re V III. Real-world data: Assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products. *Pharmacoepidemiol Drug Saf*. 2022;31(7):717-720. doi:10.1002/pds.5444