

November 29, 2021

Lauren K. Roth  
Acting Principal Associate Commissioner for Policy  
Food and Drug Administration  
Dockets Management Staff  
5630 Fishers Lane, Rm. 1061,  
Rockville, MD 20852

**RE: ACRO comment submission on:**  
***Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products; Draft Guidance for Industry***  
[Docket No. FDA-2020-D-2307-0002]

Dear Ms. Roth,

The Association of Clinical Research Organizations (ACRO) represents the world's leading clinical research and technology organizations. Our member companies provide a wide range of specialized services across the entire spectrum of development for new drugs, biologics and medical devices, from pre-clinical, proof of concept and first-in-man studies through post-approval and pharmacovigilance research. ACRO member companies manage or otherwise support a majority of all FDA-regulated clinical investigations worldwide. The member companies of ACRO advance clinical outsourcing to improve the quality, efficiency, and safety of biomedical research.

ACRO thanks the Agency for releasing this draft guidance to advance the use of real-world data in regulatory decision-making. ACRO is pleased to provide the following feedback.

## **I. General comments:**

ACRO appreciates FDA's efforts to outline specific components regarding electronic health records (EHR) and medical claims data as "*part of its RWE program and to satisfy, in part, the mandate*". Although the draft guidance provides detail on EHRs and medical claims data, a summary of how all subsequent guidance documents will interconnect would be beneficial for stakeholders to understand the overall context of the FDA's RWE program. It would be particularly helpful if the FDA would summarize what they consider to be the minimum requirements for various RWD sources to support alignment on what is most important.

In a number of sections in the document, the FDA recommends that sponsors "*seek FDA input before conducting the study or request comments or a meeting to discuss.*" We suggest that the FDA further augment the draft guidance by providing specific recommendations on the process for engagement and examples of how and when a meeting or correspondence should take place. In cases where additional context can be addressed through offline methods, e.g., white papers, publications, pilot projects, or collaborative efforts, it would be helpful if the Agency could explore and consider the possibility of a central repository on the FDA website for such documents, if possible.

We understand that *“the contents of this document do not have the force and effect of law and are not meant to bind the public in any way.”* Nonetheless, we urge the FDA to reconsider some of the feasibility and operational challenges presented by the recommendations throughout the document. One example is with regard to *“technological advances in the field of artificial intelligence”*, where the FDA outlines that *“the protocol should specify the assumptions and parameters of the computer algorithm used.”* This may not be feasible as algorithms are often proprietary intellectual property. We recommend that the FDA reconsider how to approach providing adequate detail while also protecting proprietary information. Another example is the recommendation to provide feasibility and operational recommendations broadly across all studies, which in practice, is quite difficult as any such recommendations must be tailored to the type of data and data source.

EHR and medical claims are grouped together in this draft guidance *“for evaluating the relevance and reliability for both EHRs and medical claims data”*. This may be appropriate in some instances and may differ based on each case submitted; however, it would be beneficial for the FDA to further clarify how sponsors should assess the relevance, strengths, and limitations of each data source.

In a conventional clinical trial conducted under Investigational New Drug (IND) authority, 21CFR312.57 requires that *“A sponsor shall retain the records and reports required by this part for 2 years after a marketing application is approved for the drug; or, if an application is not approved for the drug, until 2 years after shipment and delivery of the drug for investigational use is discontinued and FDA has been so notified.”* The draft guidance does not address retention periods for the data, records and reports generated in connection with a clinical study based on real world evidence that is used for equivalent regulatory purposes. ACRO asks for the final guidance to clarify that the retention period for RWE data is no longer than that required for records and reports of conventional clinical trials conducted under IND authority. Consequently, we recommend that FDA include expectations for retention consistent with those in 21 CFR § 312.57.

This guidance refers to potential uses of RWE to help support approval of a new indication for approved products or to help support post-approval study requirements. ACRO asks the Agency to consider including language in the final guidance clarifying that, in addition, the Agency also recommends the use of this guidance for assessing EHR and Medical Claims data to support regulatory decision-making for *products yet to be approved*.

## II. Line-Specific Comments:

### Lines 42-44

The current text in the draft guidance states:

*“Examples of RWD include data derived from EHRs, medical claims data, data from product and disease registries, patient-generated data including from in-home use, and data gathered from other sources that can inform on health status, such as digital health technologies. This guidance focuses on health-related data recorded by providers that can be extracted from two sources: EHRs and medical claims data.”*

ACRO would welcome future guidance on additional sources – beyond EHRs and medical claims data. It would be helpful if the final guidance could indicate if additional guidance on use of other sources of RWD is planned.

## **Lines 163-174**

The current text in the draft guideline states:

*“There are differences in the practice of medicine around the world and between health care systems that may affect the relevance of the data source to the study question.....It is also important to identify whether the data sources cover all populations relevant to the study if those sources are to be used to examine the study hypothesis.”*

Differences in medical terminology and/or coding systems used in different health care systems should also be addressed, and ACRO asks the Agency to consider including this in the final guidance as follows (new language underlined)—

*There are differences in the practice of medicine around the world and between health care systems that may affect the relevance of the data source to the study question.....It is also important to identify whether the data sources cover all populations relevant to the study if those sources are to be used to examine the study hypothesis. Differences in terminology and coding systems used in different health care systems should also be addressed.*

## **Lines 237 – 239**

The current text in the draft guideline states:

*“Obtaining comprehensive drug coverage and medical care data on patients with certain types of privacy concerns (e.g., sexually transmitted infection, substance abuse, mental health conditions) can be challenging and failure to do so can result in incomplete or erroneous information.”*

ACRO asks the Agency to consider expanding this list to include language related to vulnerable participants such as children; adolescents; and caregivers for dementia and other conditions. EHR data may be more challenging in these population groups.

ACRO asks the Agency to consider the following text for the final guidance (new language underlined):

*Obtaining comprehensive drug coverage and medical care data on patients with certain types of privacy concerns (e.g., sexually transmitted infection, substance abuse, mental health conditions), as well as from vulnerable participants such as children, adolescents, and caregivers, can be challenging and failure to do so can result in incomplete or erroneous information.*

## **Line 259-266**

The current text in the draft guideline states:

*“Probabilistic and deterministic approaches to data linkage may result in different linkage quality, albeit both approaches can have value depending on the scenario. The deterministic approach for data linkage uses records that have an exact match to a unique or set of common identifiers, and the match status can be*

*determined using a single or multiple step process. The probabilistic approach for data linkage uses less restrictive steps in which the identifiers compared consist of fewer variables or part of them (Carreras et al., 2018). When a probabilistic approach is used, the analysis plan should include testing the impact of the degree of match and robustness of findings”*

ACRO asks for the FDA to reconsider the statement that when using a probabilistic approach, analysis plans should include testing the impact of degree of match and robustness of findings. The more stringent approach to probabilistic approach may discourage use and may have the perception that the FDA factors deterministic linkages. Furthermore, sensitivity analyses to test the robustness of the findings are useful regardless of the method of matching.

### **Line 271**

The current text in the draft guideline states:

*“ . . . integrated with acceptable quality, given the potential for heterogeneity in population . . . ”*

In order to clarify the meaning of “acceptable quality,” ACRO asks FDA to consider the following text for the final guidance (new language underlined):

*integrated to a level of quality such that there is no significant impact on the accuracy of the data, given the potential for heterogeneity in population*

### **Line 360**

The current text in the draft guideline states:

*“A computable phenotype definition should include metadata”*

It would be helpful to include information on expectations for the metadata required in this context, and ACRO asks the Agency to consider the following parenthetical clarification for the final guidance (new language underlined).

*A computable phenotype definition should include metadata (at a minimum, version, author, concept, identifier, data type, definition, and preferred label for a particular data element)*

### **Lines 414-416**

The current text in the draft guideline states:

*“An example of a potential proxy variable includes low-income subsidy under the Medicare Part D prescription drug program as a proxy for a patient’s socioeconomic status.”*

The guidance should make clear whether use of a proxy variable may be justified by rationale alone (as in the example given) or if validation is required. ACRO asks the Agency to consider the following language for the final guidance—

*An example of a potential proxy variable includes low-income subsidy under the Medicare Part D prescription drug program as a proxy for a patient's socioeconomic status. The protocol should explain the rationale for use of any proxy variable chosen.*

## **Lines 463-466**

The current text in the draft guideline states:

*"Although complete verification of a study variable is considered the most rigorous approach, there are scenarios where verifying a variable for every subject might not be feasible (e.g., a very large study population, lack of reference standard data for all study subjects) and assessing the performance of the variable's operational definition might suffice"*

We agree that complete verification is not always operationally feasible, especially, for example, when there is no source data beyond the RWD (e.g., EHR are often the source data). Sponsors would benefit from examples and use cases of incomplete verification and the level of acceptance depending on the different scenarios and varying methods of verifications.

## **Lines 711-718**

The current text in the draft guideline states:

*"Selecting an appropriate comparator..... of the exposed and comparator populations."*

It would be helpful for the guidance to make clear that, as the study is concerned with real world use, it may be appropriate to select a comparator that is commonly used off-label rather than a less commonly used comparator that is authorized for use in the specific indication. ACRO recommends the following revision of the text for the final guidance (new language underlined).

*Selecting an appropriate comparator..... of the exposed and comparator populations. There may be situations where it is appropriate to select a comparator that is commonly used off-label rather than a less commonly used comparator that is authorized for use in the specific indication. This should be justified in the protocol.*

ACRO thanks the Agency for this opportunity to provide feedback on the draft guidance on Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. Please do not hesitate to contact ACRO if we can answer any questions or provide additional details ([knoonan@acrohealth.org](mailto:knoonan@acrohealth.org)).

Respectfully submitted,



Karen A. Noonan  
Senior Vice President, Global Regulatory Policy