ACRO Innovate Advocate Collaborate

June 20, 2025

Grace R. Graham, Deputy Commissioner for Policy, Legislation, and International Affairs
Ethan Chen, Center for Drug Evaluation and Research
Hussein Ezzeldin, Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave
Silver Spring, MD 20993

RE: ACRO responses to RFI:
Exploration of Health Level Seven Fast Healthcare Interoperability Resources for Use in Study Data Created from Real-World Data Sources for Submission to the Food and Drug Administration
[Docket No. FDA–2025–N–0287]

Dear Ms. Graham, Mr. Chen, and Mr. Ezzeldin,

ACRO, founded in 2002, is a non-profit trade association representing the world's leading clinical research and technology organizations, which provide specialized services that are integral to the development of drugs, biologics and medical devices that enable patients to live longer, healthier, and more productive lives. ACRO members provide a wide range of specialized services across the entire spectrum of development—from pre-clinical, proof of concept, and first in human studies through post-approval, pharmacovigilance, and health data research. ACRO member companies employ nearly 400,000 people worldwide and conduct research in every global region.

We thank the Agency for this opportunity to provide input on the exploration of HL7 FHIR for use in study data created from RWD sources. ACRO's responses to the five questions are included below.

Question 1:

What challenges do you see for the pharmaceutical industry regarding the current state of submitting clinical study data collected from RWD sources to FDA?

In its final guidance *Data Standards for Drug and Biological Product Submissions Containing Real-World Data* (December 2023), FDA explains that for submission purposes RWD data should be treated, and formatted, like any other clinical study data. Currently, there are numerous challenges which prevent RWD sources from being utilized more broadly in clinical trials. Barriers include operational, technical, procedural and administrative complexities.

The primary challenge is that current FDA submission standards are oriented towards traditional clinical trials, making direct application to RWD difficult. Developing robust methodologies for mapping CDISC standards to RWD sources is crucial and some work has been done in this area including the HL7 to CDISC FHIR IG. https://www.cdisc.org/standards/real-world-data/fhir-cdisc-joint-mapping-implementation-guide-v1-0. The issue remains that there is still inconsistency in how these mappings could be potentially applied. There is enough flexibility in the implementation guide that could raise different interpretations across reviews and review teams.

Today, RWD sources tend to be varied, both in format and in fidelity. There is a lack of homogenized, well tested, and well accepted standards for the codification and communication of data, which makes direct data submission from an RWD source complex. The FDA recognizes this in its final guidance *Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products* (July 2024).

RWD sources were not initially designed for research or regulatory purposes; so, the onus falls to the sponsors to demonstrate that the resources are relevant and reliable enough to adequately address the study question and sufficiently characterize study populations, exposures, outcomes and covariates. Traditional electronic data collection mechanisms have relied on approaches such as source data verification (SDV) to establish the reliability of the data collection processes. Much of the RWD can be considered as electronic source data which has been defined as not requiring SDV. However, given that RWD sources were not necessarily defined with research intent, there remains an ambiguity as to the level of SDV required for a patient record. The SDV plan becomes more complex when diverse capabilities in multi-site investigations are taken into account.

In order to prepare the data for submission and verify its relevance, some data mapping is necessary to meet the structural and semantic requirements. An example of this is how health concerns are recorded at source. Healthcare systems will use vocabulary systems defined for health records or billing purposes (such as SNOMED-CT, ICD-10, CPT), whereas submissions to the FDA require the use of MedDRA. A similar dilemma exists for medications and other common codable concepts (e.g. sex, race, ethnicity). It is important to identify the provenance from the source data to the submission format. This mapping is complex to maintain and can be challenging to explain.

RWD collected during routine care requires adherence to privacy regulations, such as HIPAA, while also addressing secondary research use. Regulated research requires informed consent from participants and includes consent for direct access to source records by sponsors and regulators. Other complexities include:

- Flexible visit schedules with wide windows.
- Sites are often selected based on patient populations and they may not be familiar with the typical rigor that is assumed at clinical sites. For example, timeline for data entry, data cleaning, and query resolution.
- EMR/EHR data is not always ‘clinical research grade.’ It is collected, managed, and used primarily for patient care rather than clinical or reportable safety endpoints.

We agree with our colleagues at CDISC that the regulators’ ability to receive documents in machine readable format remains a challenge.

Question 2:

What opportunities and/or challenges do you see for the pharmaceutical industry on reaching a future state of clinical study data submissions collected from RWD sources using HL7 FHIR (e.g., business processes, technical considerations)?

While HL7 FHIR moves us towards greater adoption of RWD sources, more work is still required. Further, FHIR is an exchange standard, not a content standard.

Healthcare data is complex, variant, and often incomplete. FHIR creates a common language, structure and protocol for healthcare data exchange. It is not a data model designed for analysis or insight. FHIR does not account for the variances in documentation norms, a patient's propensity to visit multiple institutions, or general gaps in documentation.

Levels of support for different versions of FHIR are still widely varied across vendors and systems. There are many vendors who provide services for normalization of the underlying standards/data into common representations from healthcare systems; however, this introduces another challenge for provenance.

We believe that we are still early in a ubiquitous, well-aligned FHIR-facilitated exchange. The FHIR at Scale Taskforce (HL7 FAST) accelerator aims to bring scalable implementations of FHIR. They recognize issues such as a lack of blanket trust agreements, standard interpretations and implementations of FHIR, and endpoint directories as continued barriers.

Additionally, RWD data sources contain a wide breadth of information, both trial and non-trial related. There is no standard to identify data that is relevant to a clinical trial, and these data must be manually collected and organized.

Many of the challenges could be accounted for by using FHIR extensions and profiles and developing an implementation guide. This process can be lengthy and would require considerable engagement across many stakeholders but could represent a common semantic for enablement.

As an example, for the challenge of identification of data pertaining to a clinical trial, solutions have been proposed via FHIR implementation guides including HL7 Vulcan Schedule of Activity IG, and Evidence Based Medicine on FHIR. Consideration of these would be worthwhile within the context of using FHIR as the representation format for RWD. There is merit in looking beyond the FHIR model for data representation into looking at the FHIR Workflow model as a mechanism for deeper insight into data collection definition, planning and performing processes. Paradoxically, the FDA must keep in mind the burden that additional IGs may create on EHR vendors and health systems, when designing their applications.

The industry has already recognized the need for an interoperability bridge between FHIR and submission standards. CDISC has previously published a mapping specification between FHIR resources/attributes and SDTM Domains/Domain Variables and is currently (June 2025) working on an updated specification. This does not solve the interoperability problem but provides an important path for parties intending to use RWD in FHIR format for submission.

One significant challenge is that a single datapoint can now be leveraged in far-ranging sets of use cases and defining a source of 'truth' is not as clear as we typically desire it to be for a clinical study. For example, there is a registry being conducted for patients on GLP1 therapies for Diabetes and another registry for patients on GLP1 for weight loss. The current data flow is that data from the EMR would be transferred to the various Electronic Data Capture (EDC) systems for those studies, at which point the different teams/companies/sponsors would review, clean and query the data based on the research being done. Questions that would need to be answered include: If both teams come back with conflicting data queries or clarifications, what happens to the data in the EMR? If a change is made in the EMR and in the EDC, which then takes precedence? How does this impact the burden on sites? How do we align on data standards and cleanliness between the expectations of EMR data and clinical study data?

As an important first step to improve more consistent use of FHIR and overall more robust submissions, it will be important to include HL7 FHIR in the FDA Data Standards Catalog. FDA could signal an important shift to regulated industry by moving in this direction and better aligning and coordinating the use of FHIR with other government partners and stakeholders including but not limited to ASTP, CMS, and CDC.

We agree with our colleagues at CDISC that data cleaning of EHRs remains a challenge. Specifically, what is an adequate level of data cleaning, imputation, categorization, derivations, and other pre-processing (all while maintaining data lineage and traceability)? In other words, when does raw RWD become analysis-ready RWE?

Question 3:

What are your suggestions on how, from a data standards perspective, FDA might reach a future state of clinical study data submissions collected from RWD sources that aligns with ASTP/ONC health IT goals for HL7 FHIR-based exchange?

Historically, the data standards with the greatest impact have been those 'required' by the agency. Ensuring that the data collection and exchange components can easily align with submission requirements, like SDTM, or safety reporting like E2BR3 will drive greater adoption, accuracy and efficiency. In addition, we believe that first and foremost, aligning patient trial identifiers and study participation data should be considered.

USCDI may be expanded to include trial related data, and the patient's participation in-trial. Classes may be extended to include trial relevancy tags. Conventional data management processes are oriented around few primary sources for patient data, with identifiers strictly controlled; a CTMS system masters the patient identifier and that identifier is shared with other data providers (e.g., Central Laboratories, IXRS vendors). Patient data in healthcare systems can be distributed across multiple source healthcare systems such as multiple site EHR and independent third-party data source systems. Countries with top-down healthcare systems may allocate a single identifier, but that doesn't apply everywhere; careful consideration would need to be applied to how data may be joined. FHIR can facilitate the integration process, but it will not solve the single individual/multiple designations dilemma.

Schedule of activities alignment with a patient's Care Plan may also help demarcate trial-related activities. One existing project within the HL7 Vulcan Accelerator, Schedule of Activities -- HL7 Vulcan, has this as a use

case; the FHIR Workflow pattern can facilitate representation of intent to performance and outcomes of clinical procedures based on a Research Plan.

From a data standards perspective, the key to enabling the use of RWD is controlling the total cost of adoption. Ensuring a datapoint has as few transformations as possible lowers the cost of managing and auditing the data in sponsor systems.

As an example, expansion of controlled vocabularies acceptable for submission to include internationally accepted ontologies would be required. This would require extensions to the existing submission data models and process, as a move from terms defined in implicit coding systems (e.g. MedDRA, WHODrug) to terms defined in explicit coding systems (e.g. LOINC, SNOMED-CT, ICD-10) will need source and destination systems to be able to recognize and reconcile the intent/meaning of the terms.

Reconciling the representation formats for the different data management systems will need careful consideration; FHIR (as an exchange format) makes extensive use of relationships leading to nested and distributed representations (i.e. using URLs to other resources). The submission data formats have very different approaches for representation of relations; mediating the transformation may be an expensive process.

Aligning FDA's requirements with existing initiatives supported by ONC, such as USCDI and TEFCA, could strengthen support for not only medical product research and development, but also fortify the envisioned health technology ecosystem as discussed in a recent RFI by CMS/ASTP.

Question 4:

Does USCDI version 3 provide enough information for collecting RWD for research purposes? Is there information that USCDI version 3 does not sufficiently address?

USCDI 3 does a great job of capturing and articulating the patient's clinical summary, from a clinical lens, but does not necessarily capture trial specific activities.
Some examples of missing data include:

- Multiple Patient identifiers in a clinical trial
- Patient's participation in a study (start/end date, reason for discontinuation, etc.)
- Association of encounter and specific clinical data with trial activity
- Research Note type
- Consent for patient participation in research
- Higher levels of detail for some of the data elements – for example:
  - Laboratory data elements (e.g. units of measurement, reference ranges) that are core to interpretation of the results
  - Onset date
  - Specimen details
- Lack of research-specific data domains, such as adverse events

Subsequent versions of USCDI address many of these missing items through additional data elements and attributes.

ACRO Innovate Advocate Collaborate

Question 5:

Under TEFCA, a variety of "Exchange Purposes" are authorized. If "Research" was added as an "Exchange Purpose," what role could TEFCA play with using RWD for clinical research? How could TEFCA support more efficient collection and exchange of RWD for clinical research purposes? What challenges might exist with this approach?

There are many advantages for a Research Exchange Purpose (XP). TEFCA provides baseline governance, legal and technical requirements for secure information sharing. Adding a Research XP can leverage the existing infrastructure to provide an avenue for continuous data flow. Moreover, it would give a better representation of what is happening in a patient's day-to-day life during clinical trials without solely relying on that patient to provide all the data.

Formalizing on an agreed protocol and exchange standard, such as FHIR, shifts much of the cost from existing arrangements. Nothing like this has existed prior, with some EHR vendors providing 'capabilities' that were bespoke and required investment to set up, run and maintain long after a study has completed. Using the Research XP could make the decision around the use of RWD one of 'why not' rather than 'why.'

There would be challenges with this proposal:

- We can see issues with USCDI data elements and the required granularity to suit clinical research purposes. There would need to be consideration of how the USCDI might scale. In some cases, we would need to have an extension mechanism because many of the data elements required for research are highly specific and emergent.
- Research XP would not allay concerns around completeness, accuracy, consistency, bias or missing data.
- For multi-country studies, the Research XP will introduce complexities for US/non-US sites.
- Which platform becomes the source of truth?
- How do we manage conflicting data?
- Which set of standards do we follow?

The true value of data interoperability, particularly through HL7 FHIR, is linked to the actionability of the exchanged data. Mechanisms for ensuring data reliability and provenance should be integral to the data exchanges through TEFCA and submission process with FDA. Harmonized understanding of reliability will be important to continue scaling the use of FHIR across use cases. We strongly support a TEFCA Research XP. It could open the door to subject screening, recruitment and enrollment activities. TEFCA will continue to allow a variety of formats (recognizing ASTP's FHIR roadmap). However, there would also need to be an evolution and adoption of Bulk FHIR use cases to make the Research XP truly effective.

ACRO thanks the Agency for the opportunity to respond to this RFI. Please do not hesitate to contact ACRO (knoonan@acrohealth.org) if we can answer questions or provide additional information.

Respectfully submitted,

Karen Noonan

Karen Noonan, Senior Vice President, Global Regulatory Policy