Submitting RWD: Where We Are and Where We Are Going

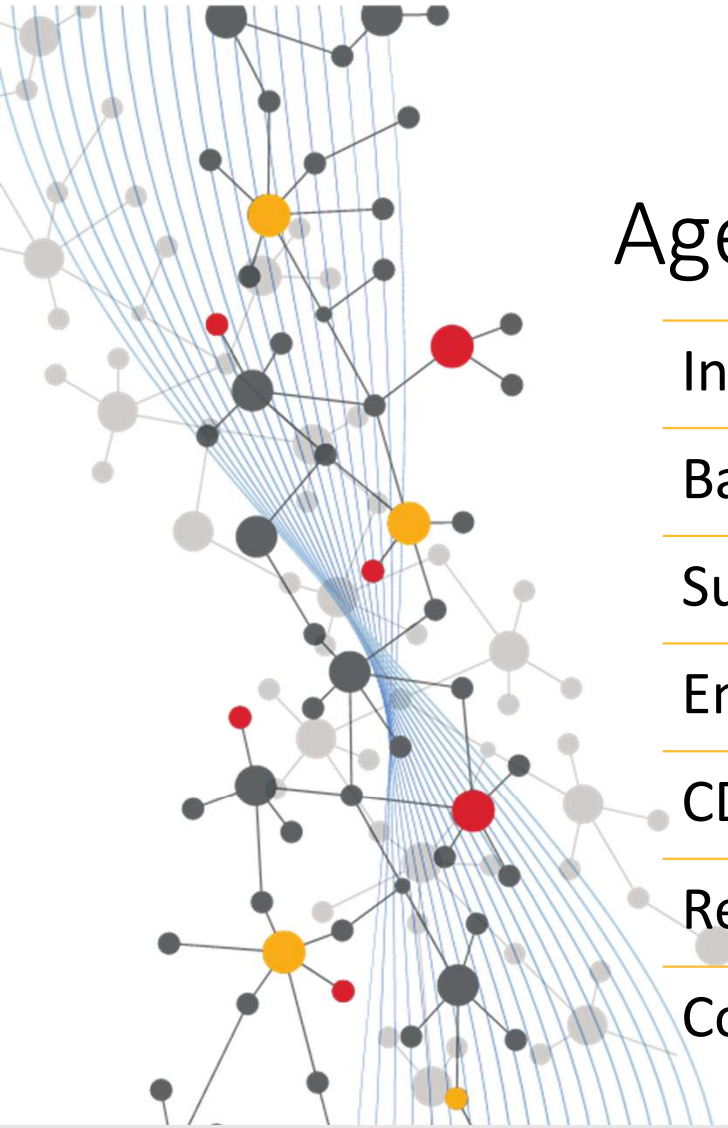
Jeff Abolafia | Director of Product Innovation, *Certara*

October 2, 2025

CERTARA

Disclaimer

The views and opinions presented here represent those of the speaker and should not be considered to represent advice or guidance on behalf of Certara



Agenda

- Introduction
- Background
- Submission Challenges
- Environmental Scan
- CDISC, FHIR, or OMOP?
- Recommendations
- Conclusions

Introduction

Introduction

- Can our current submission standard (CDISC) support diverse study designs?
- Is CDISC evolving to support diverse study designs?
- Can a single standard support all study designs
- Is there another standard that may be a better fit
- Is a hybrid approach (multiple standards) a better option?
- Presentation goals
  - Reframe the discussion around future submission standards
  - Current State: How do we fit RCT, RWD, and other non-interventional designs into CDISC standards?
  - Should data from all study designs be submitted using CDISC standards
  - Should we choose a different standard to support regulatory submissions
  - Should we consider a hybrid approach

Clinical trial vs observational study

Reference: “Framework for FDA’s Real-World Evidence Program”, December 2018. https://www.fda.gov/media/171667/download

“Interventional study (also referred to as a clinical trial) is a study in which participants, either healthy volunteers or volunteers with the condition or disease being studied, are assigned to one or more interventions, according to a study protocol to evaluate the effects of those interventions on subsequent health-related outcomes”

“Non-interventional study (also referred to as an observational study) is a type of study in which patients received the marketed drug of interest during routine medical practice and are not assigned to an intervention according to a protocol”

Non interventional study data sources

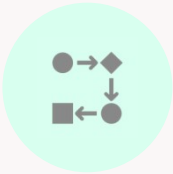
Electronic health records (EHRs) / medical records

Medical claims and billing data

Product, disease, and population-based registries

Cohort study

Case-control study



Other non-interventional designs

Why RWE? RWE Trends

- 95% of approved NDAs and BLAs in 2021 included an RWE study
- 75% in 2019
- 58% of approvals in 2021 used RWE to support safety and/or efficacy

Table 3

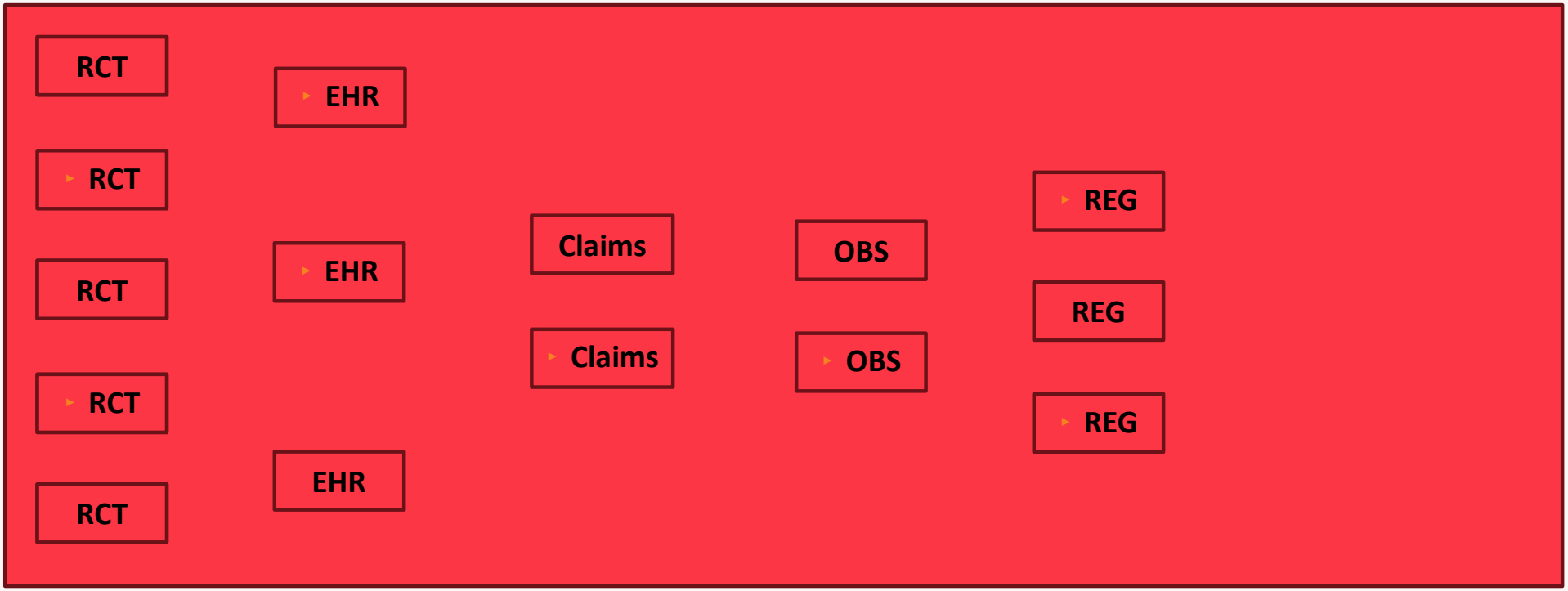
Observed intended use of RWE in included NDAs/BLAs, January 2019 to June 2021

Included NDAs and BLAs	2019	2020	2021 through June 30	Total
	n = 51 approvals	*n* = 59 approvals	*n* = 26 approvals	*n* = 136 approvals
Incorporated RWE for any purpose	38 (75%)	53 (90%)	25 (96%)	116 (85%)
Used RWE to provide therapeutic context	25 (49%)	36 (61%)	22 (85%)	83 (61%)
Used RWE to support safety and/or effectiveness	27 (53%)	46 (78%)	15 (58%)	88 (65%)
Safety only	17 (33%)	21 (36%)	5 (19%)	43 (32%)
Effectiveness only	7 (14%)	6 (10%)	2 (8%)	15 (11%)
Safety and effectiveness	3 (6%)	19 (32%)	8 (31%)	30 (22%)

Purpura CA, Garry EM, Honig N, Case A, Rassen JA. The Role of Real-World Evidence in FDA-Approved New Drug and Biologics License Applications. Clin Pharmacol Ther. 2022 Jan;111(1):135-144. doi: 10.1002/cpt.2474. Epub 2021 Nov 22. PMID: 34726771; PMCID: PMC9299054.

CERTARA

Clinical data Submission Trends



CERTARA

Background Information

Current Submission standards (1)

Reference: https://www.fda.gov/media/153341/download

- October 2021 FDA draft guidance (finalized in 2023)

  " Data Standards for Drug and Biological Product Submissions Containing Real-World: Data Guidance for Industry“

- Guidance outlined data standards required when submitting RWD (or data from other non-interventional studies) in support of a marketing application
- **Currently**
  - According to the guidance, RWD must be submitted using the standards documented in the FDA Data Standards Catalog
  - For now, that means **RWD must be submitted using CDISC standards**

Current Submission standards (2)

- CDISC required for data from all interventional & non-interventional studies
- Data from non-interventional studies must be converted to CDISC format
- Agency acknowledges
  - Current catalog of standards does not necessarily reflect data derived from real-world sources
  - It is considering updates (i.e., FRN to submit RWD in FHIR)
- Presents numerous challenges to sponsors
- CDISC designed for RCT data

Submission Challenges

Non-RCT data presents new challenges

Topic	RCT	NON-RCT
Data Collection	Collected under a protocol	Collected in real world settings
Data Monitoring	Data monitored and cleaned	No monitoring or cleaning
Data Entry	Data collected via CRF	Data entered in EHR
Visits/Encounters	Visits at protocol defined schedule	No defined length between encounters
Treatment Schedule	Pre-defined treatment	As-needed treatment
Data Source	Data designed and collected by sponsor	Curated data acquired by sponsor from vendor
Availability or Source Data	Owned by sponsor	Owned by HCPs / Aggregators
Documentation	Protocol / CRFs / DMP	Aggregators
Sites	Single site	Multiple healthcare systems
Data Uniformity	Uniform data entry across sites	Standards/formats differ by site
Terminologies/Vocabularies	Uniform across sites	May differ by site/type of RWD/region
Data Standards	CDISC	HL7 FHIR, OMOP

Challenges for submitting non-rct data

Challenges are presented in our previous five papers:

Abolafia, J, Ferko, S, & Holt, I. (2022). “Submission Standards for RWD: The Good, the Bad and the Ugly”. Paper presented at the PHUSE Annual Conference 2022, Belfast, United Kingdom.
https://phuse.s3.eu-central-1.amazonaws.com/Archive/2022/Connect/EU/Belfast/PRE_RE09.pdf

Ferko, S., Holt, I., & Abolafia, J., (2023). “Challenges and Considerations for Submitting Real World Data”. Paper presented at the PHUSE US Annual Conference 2023, Orlando, FL.
https://phuse.s3.eu-central-1.amazonaws.com/Archive/2023/Connect/US/Florida/PRE_RE05.pdf

Abolafia, J, Ferko, S, & Holt, I. (2023). “Submission Standards for Real World Data: Gaps, Limitations and Recommendations”. Paper presented at the PHUSE Annual Conference 2023, Birmingham, United Kingdom. https://phuse.s3.eu-central-1.amazonaws.com/Archive/2023/Connect/EU/Birmingham/PAP_RE03.pdf

Abolafia, J, Ferko, S, & Holt, I. (2024). “Considerations for the Submission of RWD using CDISC with Insights from HL7 FHIR and OMOP”. Paper presented at the PHUSE Annual Conference 2024, Strasbourg, France.
https://phuse.s3.eu-central-1.amazonaws.com/Archive/2024/Connect/EU/Strasbourg/PAP_RE03.pdf

Holt, I., Ferko, S., & Abolafia, J. (2025). “Considerations for the Submission of RWD using CDISC with Insights from HL7 FHIR and OMOP”. Paper presented at PHUSE US Annual Conference 2025, Orlando, FL.
https://www.lexjansen.com/phuse-us/2025/re/PAP_RE02.pdf

Environmental scan

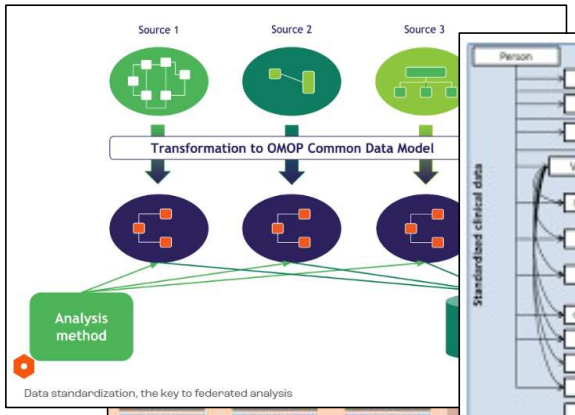
Existing Data Models and Standards for RWE and Clinical Trials

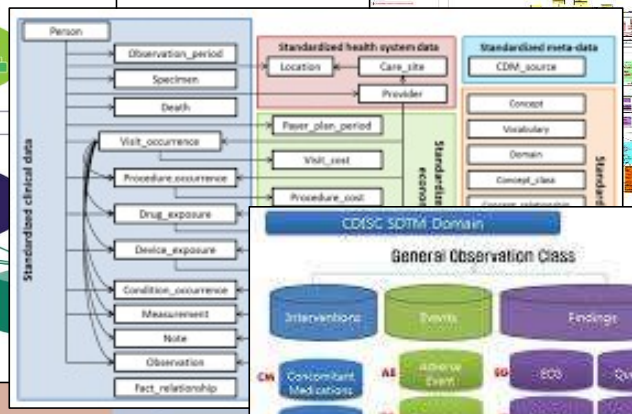
SDOs

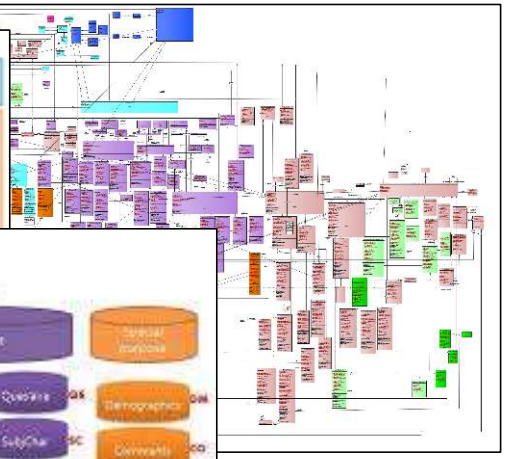
Healthtech / Industry

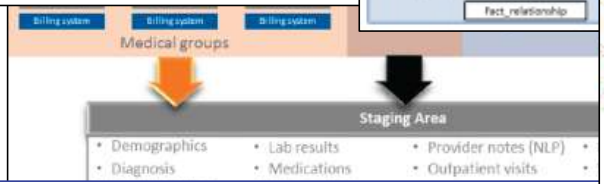
Government

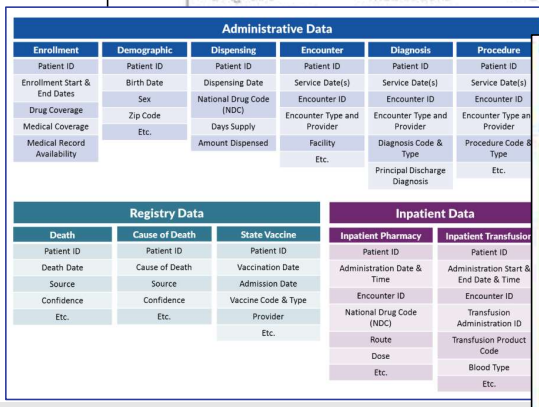
Consortium

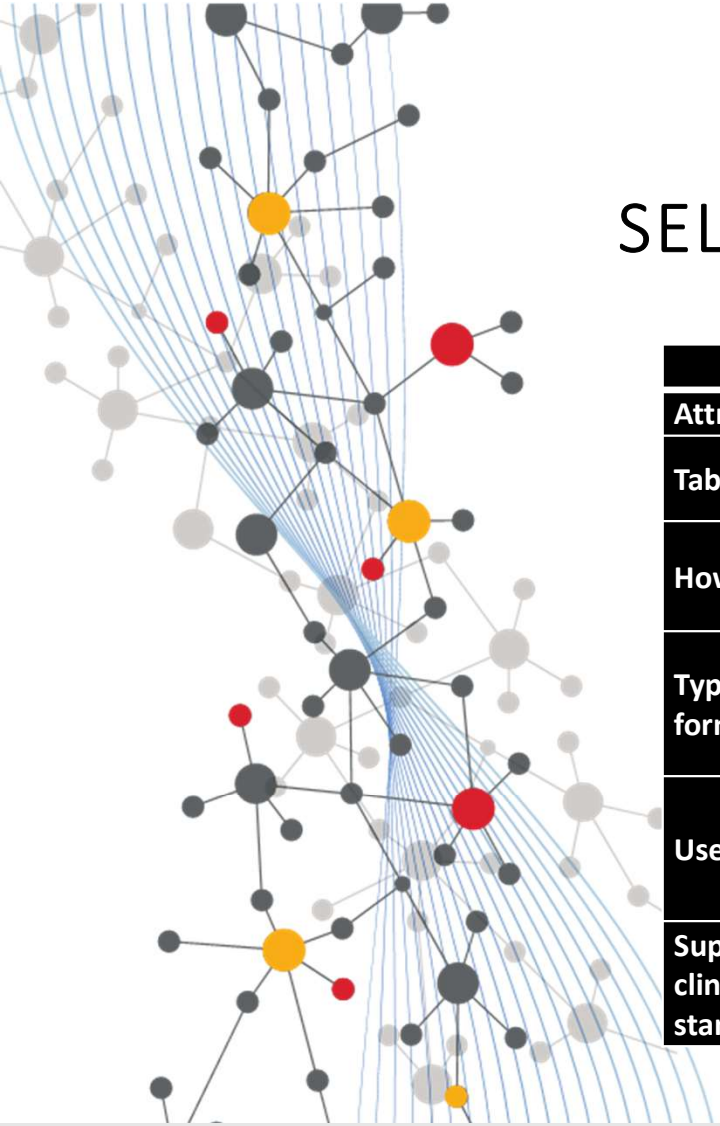








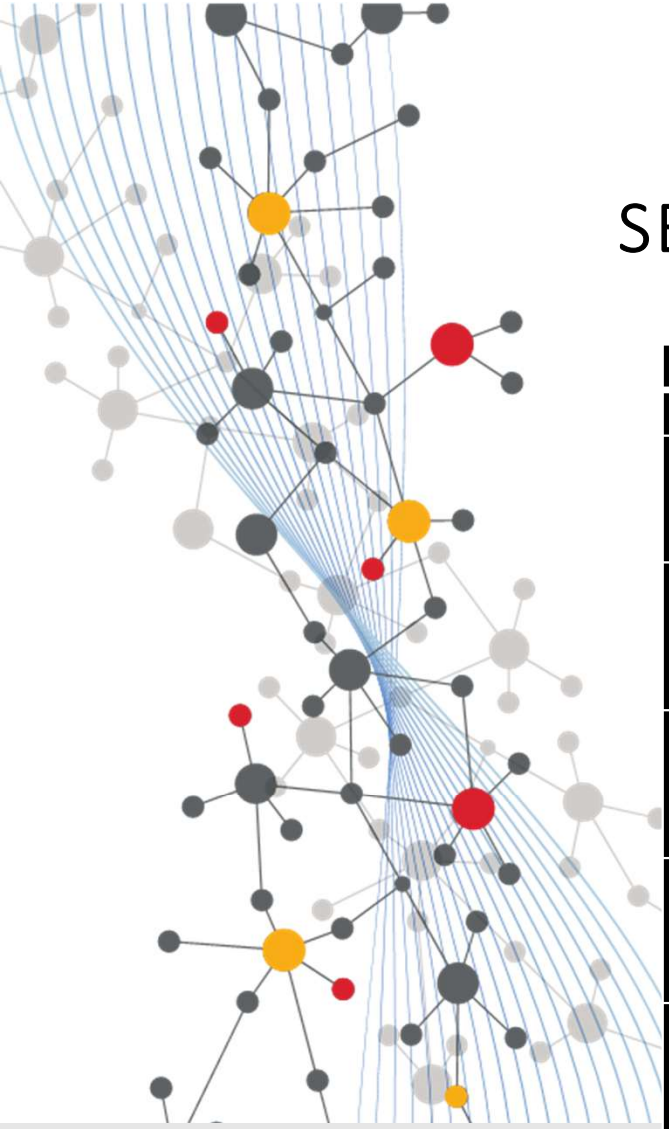




SELECT CLINICAL DATA STANDARDS (1)

	CDISC	HL7	OMOP
Attribute			
Tabulation Standard	Yes (SDTM)	Yes (FHIR)	Yes
How organized	Topic Related Domains in SDTM	Resources	Relational database tables for clinical data, vocabulary, and other healthcare concepts
Typical exchange format	SAS V5 Transport / Metadata ODM XML	JSON	Each organization or user builds their own CDM, tools are provided for assistance with ETL to transform data
Use case(s)	Submission of RCT data to regulatory authorities including FDA and PMDA	Electronic health records, claims data	Enable efficient analyses of observational data
Supported FDA clinical data standard	Yes	No	No

CERTARA



SELECT CLINICAL DATA STANDARDS (2)

	CDISC	HL7	OMOP
Attribute			
Supports Non-rectangular structures	No	Yes	Not directly, but data could be extracted in non-rectangular format defined by user
Record linkage	Implicit (RELREC domain in SDTM)	Explicit (Data elements are linked via the model)	Yes, identifiers are used to link the data as needed.
Non-standard data elements	Supplemental variables	Extensions	No
Analysis standard	Yes (ADaM)	No	No
Typical dictionaries	MedDRA, WHODrug, CDISC Controlled Terminology	SNOMED, LOINC, ICD, RxNorm	Over 100 supported, see https://github.com/OHDSI/Vocabulary-v5.0/wiki/Standardized-Vocabularies

Exchange standards

SAS V5 Transport

Openly documented specification developed by SAS in late 1980's

Required format for submitting study data since 1999

Imposes several restrictions on submission datasets

XML (Extensible Markup Language)

"Text" format, where you can define your own tags to meet specific needs

Current standard for the define file

CDISC published Dataset-XML standard in 2014 > can be used to submit data (but not accepted standard)

JSON ((JavaScript Object Notation)

Consists of human-readable text to store and transmit data objects comprised of attribute-value pairs

Used extensively when exchanging data using the FHIR model

De facto standard for data exchange using APIs

Dataset-JSON was adapted from the Dataset-XML specification, but instead uses JSON format for regulatory submission needs

Ongoing project in collaboration with FDA, PHUSE, CDISC to pilot the use of JSON for submitting study data

CDISC, FHIR or OMOP?

CDISC: Advantages

- Supported by the FDA as defined in the Data Standards Catalog
- Industry familiarity with the standards across pharma, SDOs, regulatory agencies, and other industry groups
- Contains concepts needed to represent clinical research data; currently the only 'submission ready' standard
- CDISC defines comprehensive end-to-end standards from data collection through analysis; both nonclinical & clinical
- Documentation by SDOs (e.g., IGs, TAUGs, QRS supplements, controlled terminology) and regulatory agencies (e.g., tech specs)
- Tools developed and in use by SDOs, pharmaceutical companies, & regulatory agencies that use CDISC
- Analysis standard (ADaM) exists and is harmonized with SDTM (unlike FHIR)
- Many sponsors have designed end-to-end processes around CDISC standards from collection through analysis

CDISC: Disadvantages

- ‘Siloed’ for clinical research; data elements relevant to RCTs
- Not optimized for RWD
- A lot of work to convert RWD into CDISC format
- CDISC SDTM is rigid; hard to add new elements
- Cumbersome and burdensome to link data (RELREC)
- Rectangular structure
- Limitations & constraints with SAS Version 5 transport format

FHIR: ADVANTAGES

FHIR is optimized for RWD (and, for EHR and claims data). Over time, there is a trend for more RWD to be submitted as part of a marketing application.

There is a trend toward collecting data for RCTs and other observational study designs in an EHR.

Most data in EHRs are already exchanged using FHIR, and in the future, it is likely that almost all EHR data will be represented in FHIR.

FHIR is a more "modern" standard, that leverages web standards such as JSON, HTTP, Atom, OAuth and others.

Data exchanged using FHIR is usually represented in JSON and is not limited by the constraints on SAS Version 5 transport format

Linking to related data w/i a marketing application and external data sources is much easier than in CDISC

FHIR is already being used at FDA for several submission-related activities.

OMOP: ADVANTAGES

OMOP has been optimized for RWD and its analysis and has a track record for use with epidemiological studies.

OMOP was designed to support interoperability between different terminologies and dictionaries.

OMOP CDM supports large-scale collaborative research efforts and the OHDSI community is actively involved in creating tools and providing user support.

OHDSI offers a wide range of open-source tools designed for use with one or more of the databases of the Common Data Model (CDM).

The OMOP CDM is being used by some major health systems including The University of California Health System.

Designed as a relational database schema facilitates easy manipulation of the data for aggregate or individual patient analysis.

Given the relational structure is similar to that of CDISC SDTM, creating a submission package with data similar to CDISC should be straightforward.

FHIR/OMOP: DISADVANTAGES

- Lacks many of the concepts necessary to represent clinical research data
- Many non-standard variables and domains must be created
- Many validation/business rules do not apply
- FDA /sponsors are less familiar with FHIR/OMOP than CDISC
- What would a FHIR/OMOP submission package would look like ?
- New tools needed to facilitate a FHIR/OMOP submission
- FDA would have to update documents/tech specs
- No model for analysis data > just a collection/exchange standard
- Many sponsors have designed end-to-end processes around CDISC standards

SINGLE OR HYBRID?

Can one of these standards adequately represent all designs?

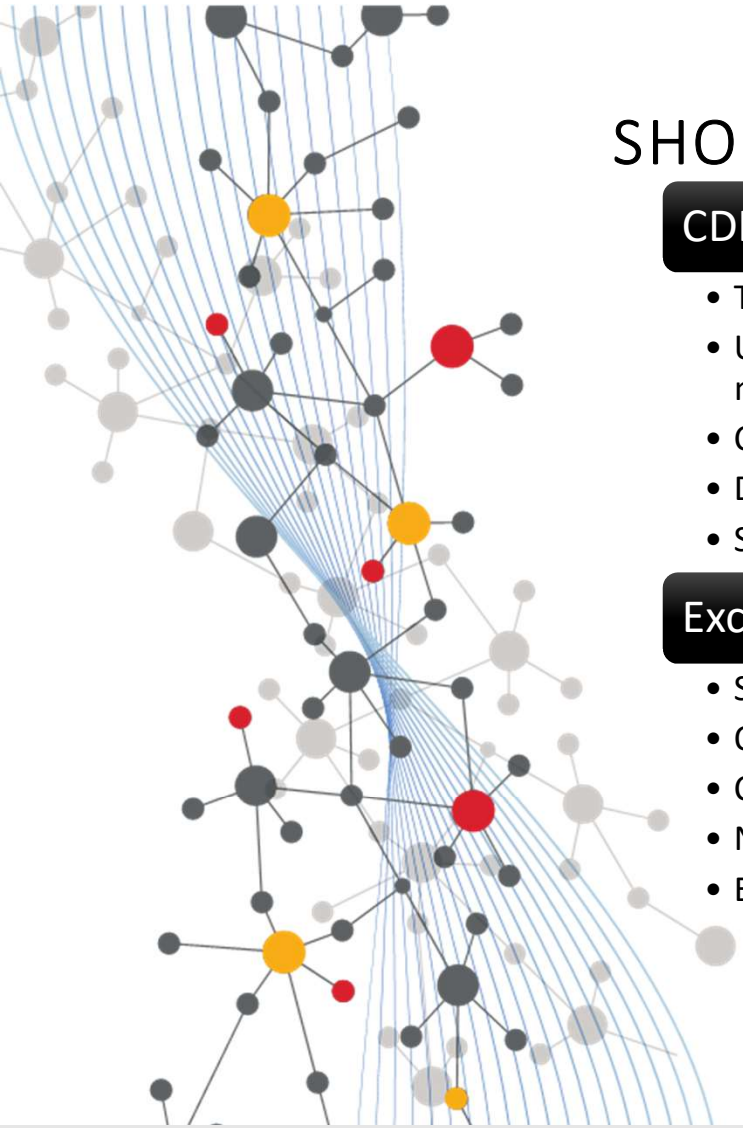
If not a single standard, would a hybrid approach be better?

Single standard option encompasses choosing either the CDISC SDTM, OMOP or HL7 FHIR standard to represent collected data for all study designs

Hybrid approach

- Collected data is submitted in the standard for which it is optimized
- RCT data are submitted using the SDTM standard
- EHR and claims data are submitted using the FHIR standard
- Data from registry studies and other observational study designs are submitted using OMOP standard
- CDISC ADaM is still the analysis standard (for now)
- May be challenging to harmonize collected data into a single analysis standard

Recommendations



SHORT TERM CONSIDERATIONS/SOLUTIONS

CDISC currently required

- Transform all collected data to STDM
- Use ADaM for analysis > required and FHIR/OMOP do not have an analysis model
- Create non-standard variables/domains
- Document violations of validation rules in reviewer's guides
- See our previous 5 papers (documented above) for short term solutions

Exchange standard >Transition to JSON as soon as possible

- So far, results of pilot look promising
- Overcomes many of the shortcomings of SAS V5 transport
- Can capture RWD more accurately
- Not a large impact on FDA/sponsors
- Extend JSON to accommodate non-rectangular data

LONG TERM CONSIDERATIONS/SOLUTIONS

Recommendation: Re-examine current submission standards (CDISC)

- Change the conversation
  - From > how to we fit all study designs into CDISC standards
  - To > What should the paradigm be for submitting study data, when considering the increasing diversity of study designs
- Can a single standard accommodate data from all study designs?
- Would a hybrid approach be better?

LONG TERM CONSIDERATIONS/SOLUTIONS

CDISC has developed core data elements for representing data from RCTs needed for regulatory submission. Can industry continue to develop core data elements and terminologies for non-interventional studies needed for regulatory submission and review?

Can CDISC standards be modified to accommodate data from all study designs? Can CDISC develop “profiles” for each type of study design. Profiles would contain core elements and validation rules for a given study design.

- If so, what changes are needed?
- What progress has CDISC made to date to accommodate non-RCT study designs?

Is OMOP or FHIR a more viable future submission standard?

- If so, what changes are needed?
- How have these standards evolved in the past several years?

LONG TERM CONSIDERATIONS/SOLUTIONS

- How will changes in technology affect future submission standards?
- Will we be able to eliminate the artificial packaging of clinical data and enable real time data steaming?
- Can we develop technologies and tools to eliminate the need for analysis datasets?
- Can we harmonize CDISC, FHIR, and OMOP, so any of these standards can easily be mapped to the other two standards?
- Over time, will more and more data be collected in EHRs?
- If so, does it make sense to transform this data from FHIR to another standard?

LONG TERM CONSIDERATIONS/SOLUTIONS

Can industry, regulatory authorities, and standards organizations work together more effectively on pilot projects related to the future of submission standards?

How can mapping, especially dictionary to dictionary, be resolved to ensure issues related to loss of granularity as well as the introduction of granularity that may result in signals (both of safety and efficacy) being diluted

Can our industry and regulatory authorities move beyond their comfort zone?

Right now, CDISC is familiar to both sponsors and FDA. Sponsors and FDA are much less familiar with other standards such as FHIR and OMOP. Though CDISC standards are currently required, this should not inhibit our industry from considering the best future state and evaluating other alternatives. Collectively, we should continue to monitor how each standard is evolving and how this affects its ability to meet future submission needs.

Conclusions

Submitting non-RCT presents a number of new challenges

There are significant gaps in the current standards for submitting non-RCT data

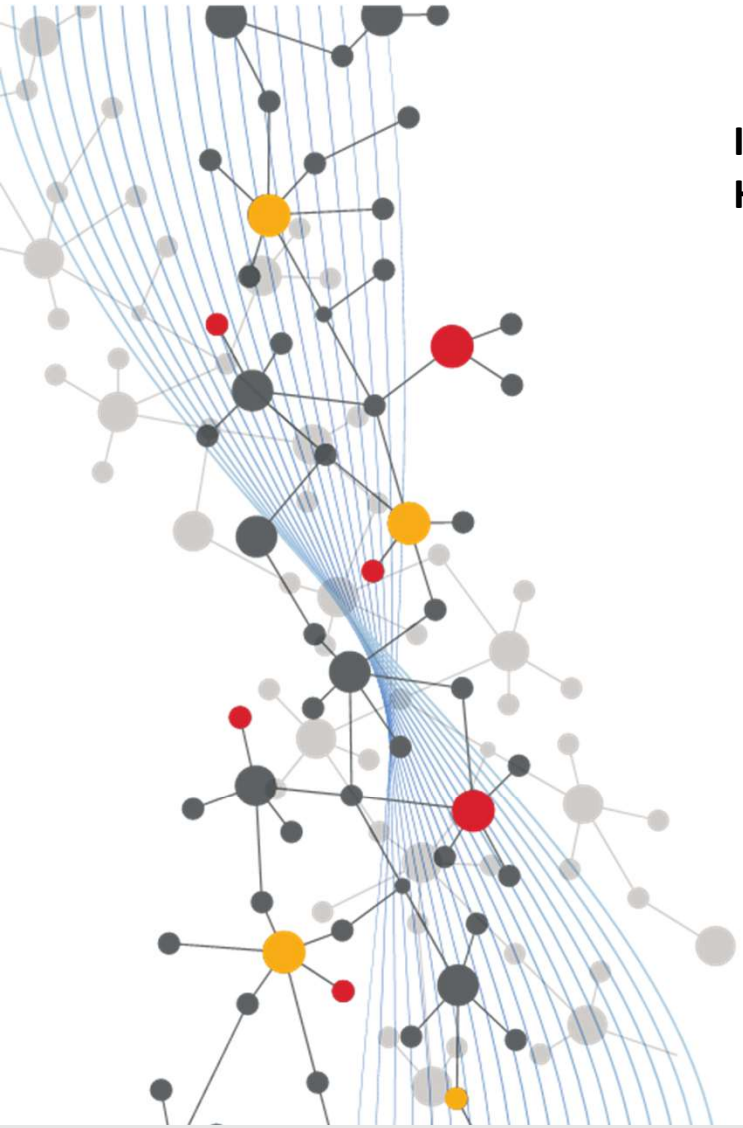
Future is not complex transformations

Are we asking the right questions to solve these challenges?

It's time to re-evaluate the use of CDISC as the single submission standard?

Need for pilot projects related to the future of submission standards across industry and regulatory agencies to help make these decisions

Let's collaborate!!



I would like to acknowledge Sarah Ferko and Ingeborg Holt for their input and work on this presentation

Thank you ;)

Keep In Touch!

Jeff Abolafia

Jeffrey.Abolafia@certara.com

CERTARA