

Considerations for the Submission of RWD using CDISC, With Insights from HL7 FHIR and OMOP

Jeffrey Abolafia, Pinnacle21 by Certara
Sarah Ferko, IBM Consulting
Ingeborg Holt, Orizaba Solutions



Disclaimer

The views and opinions presented here represent those of the speakers and should not be considered to represent advice or guidance on behalf of IBM Consulting, Pinnacle21, or Orizaba Solutions.

Contents

- Introduction / Background
- Challenges Submitting RWD
- Environmental Scan
- Exposure Data
- RWD Gaps in SDTM and Recommendations
- Summary/Conclusions

Introduction

Introduction

More health-related data from a variety of study designs now available

Can provide more evidence about the benefits/risks of a medical project

Sharp increase in NDAs/BLAs that include non-RCT study designs

- EHR
- Registry and observational studies

CDISC standards developed over 20 years ago

- Developed to support marketing applications
- Developed and optimized to primarily support RCT study designs

Submissions of non-RCT studies has presented challenges to FDA/sponsors

HL7 FHIR, OMOP, and PCORNet designed to represent RWD and non-RCT designs

We can learn from FHIR, OMOP, and PCORNet to improve CDISC submissions

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
	<i>n</i> = 51 approvals	<i>n</i> = 59 approvals	<i>n</i> = 26 approvals	<i>n</i> = 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.

Current Submission standards (1)

- October 2021 FDA draft guidance (finalized in 2023)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**

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

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
- Presents numerous challenges to sponsors
- CDISC designed for RCT data

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 three 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-](https://phuse.s3.eu-central-1.amazonaws.com/Archive/2023/Connect/EU/Birmingham/PAP_RE03.pdf)

[1.amazonaws.com/Archive/2023/Connect/EU/Birmingham/PAP_RE03.pdf](https://phuse.s3.eu-central-1.amazonaws.com/Archive/2023/Connect/EU/Birmingham/PAP_RE03.pdf)

Existing Data Models and Standards for RWE and Clinical Trials



SDOs



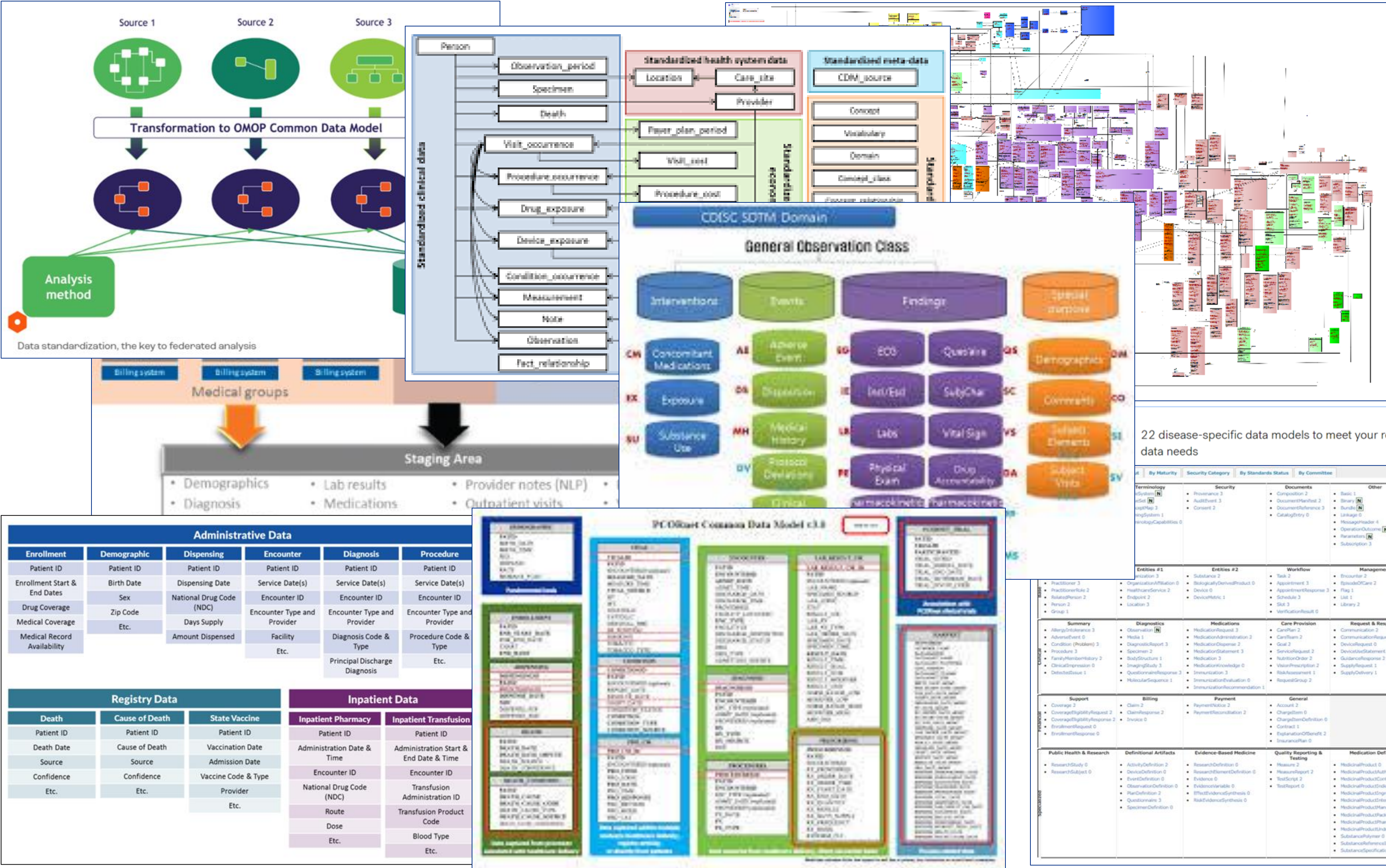
Healthtech /
Industry



Government



Consortium



Comparison of Data Models Across Select Attributes

Attribute	Data Models			
	CDISC	HL7 FHIR	OMOP	PCORnet
Tabulation Standard	Yes (SDTM)	Yes (FHIR)	Yes	Yes
How organized	Topic Related Domains in SDTM	Resources	Relational database tables for clinical data, vocabulary, and other healthcare concepts	Topic related domains each containing patient id allowing linking of information between domains; the CDM is intended to support multiple RDBMS
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	Participating organizations transform their own data into the PCORnet CDM and allow querying of the data (based on DSAs) but data is not typically exchanged.
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	Facilitate comparative effectiveness research (CER) and drive US clinical research to become more patient centric
Supported FDA clinical data standard	Yes	No	No	No

Comparison of Data Models Across Select Attributes (Cont'd)

Attribute	Data Models			
	CDISC	HL7 FHIR	OMOP	PCORnet
Supports Non-rectangular structures	No	Yes	Not directly, but data could be extracted in non-rectangular format defined by user	Organizations need to maintain the CDM for research purposes but can make their own additions w/o impacting others
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.	Yes, identifiers and pseudointifiers are used to link the data as needed.
Non-standard data elements	Supplemental variables	Extensions	No	The CDM has the concept of expanded value sets for fields providing standard terms for additional information to be stored.
Analysis standard	Yes (ADaM)	No	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	Explicit support for mapping and for ICD9/10, SNOMED CT, LOINC, CTP, HCPCS, Human Phenotype Ontology, RXNORM, NDC.

Exposure Across Data Standards: Gaps and Recommendations

Exposure Data

RCT	RWE
Well-characterized in protocol	No Protocol
Collected via Uniform CRF	Collected in EHR/Claims database
Standardized/uniform across sites	No uniformity/standardization across sites
Monitored and cleaned	Not Monitored or cleaned
Primary business case is research	Primary business case is health care

*How is exposure data represented
in CDISC SDTM, HL7 FHIR R4,
OMOP, and PCORnet?*

Exposure Domains in CDISC SDTM

6.1.3 Exposure Domains

Clinical trial study designs can range from open label (where subjects and investigators know which product each subject is receiving) to blinded (where the subject, investigator, or anyone assessing the outcome is unaware of the treatment assignment(s) to reduce potential for bias). To support standardization of various collection methods and details, as well as process differences between open-label and blinded studies, 2 SDTM domains based on the Interventions General Observation Class are available to represent details of subject exposure to protocol-specified study treatment(s): Exposure (EX) and Exposure as Collected (EC).

6.1.3.1 Exposure (EX)

EX – Description/Overview

An interventions domain that contains the details of a subject's exposure to protocol-specified study treatment. Study treatment may be any intervention that is prospectively defined as a test material within a study, and is typically but not always supplied to the subject.

6.1.3.2 Exposure as Collected (EC)

EC – Description/Overview

An interventions domain that contains information about protocol-specified study treatment administrations, as collected.

CDISC SDTM: What are we used to seeing in an RCT?

6.1.3.1 Exposure (EX)

EX – Description/Overview

An interventions domain that contains the details of a subject's exposure to protocol-specified study treatment. Study treatment may be any intervention that is prospectively defined as a test material within a study, and is typically but not always supplied to the subject.

EX – Specification

ex.xpt, Exposure — Interventions. One record per protocol-specified study treatment, constant-dosing interval, per subject, Tabulation.

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format ¹	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	EX	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
EXSEQ	Sequence Number	Num		Identifier	Sequence number given to ensure uniqueness of subject records within a domain. May be any valid number.	Req
EXGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain for a subject.	
EXREFID	Reference ID	Char		Identifier	Internal or external identifier (e.g., kit number, bottle label, vial identifier).	
EXSPID	Sponsor-Defined Identifier	Char		Identifier	Sponsor-defined reference number. May be preprinted on the CRF as an external identifier defined in the sponsor's operational database. Example: Line number on a label.	
EXLNKID	Link ID	Char		Identifier	Identifier used to link related records across domains.	
EXLNKGRP	Link Group ID	Char		Identifier	Identifier used to link related, grouped records across domains.	
EXTRT	Name of Treatment	Char	*	Topic	Name of the protocol-specified study treatment given during the dosing period.	
EXCAT	Category of Treatment	Char	*	Grouping Qualifier	Used to define a category of EXTRT values.	
EXSCAT	Subcategory of Treatment	Char	*	Grouping Qualifier	A further categorization of EXCAT values.	
EXDOSE	Dose	Num		Record Qualifier	Amount of EXTRT when numeric. Not populated when EXDOSTXT is populated.	
EXDOSTXT	Dose Description	Char		Record Qualifier	Amount of EXTRT when non-numeric. Dosing amounts or a range of dosing amounts. Example: "200-400". Not populated when EXDOSE is populated.	
EXDOSU	Dose Units	Char	(UNIT)	Variable Qualifier	Units for EXDOSE, EXDOSTOT, or EXDOSTXT representing protocol-specified units. Examples: "ng", "mg", "mg/kg", "mg/m2".	
EXDOSFRM	Dose Form	Char	(FRM)	Variable Qualifier	Dose form for EXTRT. Examples: "TABLET", "LOTION".	
EXDOSFRQ	Dosing Frequency per Interval	Char	(FREQ)	Record Qualifier	Usually expressed as the number of repeated administrations of EXDOSE within a dosing interval. Examples: "Q2H", "QD", "BID".	
EXDOSRGM	Intended Dose Regimen	Char		Record Qualifier	Text description of the intended schedule or regimen for the intervention. Examples: "QW", "TWO WEEKS OFF".	
EXROUTE	Route of Administration	Char	(ROUTE)	Variable Qualifier	Route of administration for the intervention. Examples: "ORAL", "INTRAVENOUS".	
EXLOT	Lot Number	Char		Record Qualifier	Lot number of the intervention product.	
EXLOC	Location of Dose Administration	Char	(LOC)	Record Qualifier	Specifies location of administration. Examples: "ARM", "LIP".	
EXLAT	Laterality	Char	(LAT)	Variable Qualifier	Qualifier for anatomical location further detailing laterality of the intervention. Examples: "LEFT", "RIGHT".	
EXDIR	Directionality	Char	(DIR)	Variable Qualifier	Qualifier for anatomical location further detailing directionality. Examples: "ANTERIOR", "LOWER", "PROXIMAL", "UPPER".	Perm
EXFAST	Fasting Status	Char	(NY)	Record Qualifier	Indicator used to identify fasting status. Examples: "Y", "N".	Perm
EXADJ	Reason for Dose Adjustment	Char	*	Record Qualifier	Describes reason or explanation of why a dose is adjusted.	Perm
TAETORD	Planned Order of Element within Arm	Num		Timing	Number that gives the planned order of the element within the arm.	Perm
EPOCH	Epoch	Char	(EPOCH)	Timing	Trial epoch of the exposure record. Examples: "RUN-IN", "TREATMENT".	Perm
EXSTDTC	Start Date/Time of Treatment	Char	ISO 8601 datetime or interval	Timing	The date/time when administration of the treatment indicated by EXTRT and EXDOSE began.	Exp
EXENDTC	End Date/Time of Treatment	Char	ISO 8601 datetime or interval	Timing	The date/time when administration of the treatment indicated by EXTRT and EXDOSE ended. For administrations considered given at a point in time (e.g., oral tablet, pre-filled syringe injection), where only an administration date/time is collected, EXSTDTC should be copied to EXENDTC as the standard representation.	Exp
EXSTDY	Study Day of Start of Treatment	Num		Timing	Study day of EXSTDTC relative to DM.RFSTDTC.	Perm
EXENDY	Study Day of End of Treatment	Num		Timing	Study day of EXENDTC relative to DM.RFSTDTC.	Perm
EXDUR	Duration of Treatment	Char	ISO 8601 duration	Timing	Collected duration of administration. Used only if collected on the CRF and not derived from start and end date/times.	Perm
EXTPT	Planned Time Point Name	Char		Timing	Text description of time when administration should occur. This may be represented as an elapsed time relative to a fixed reference point, such as time of last dose. See EXTPTNUM and EXTPTREF.	Perm
EXTPTNUM	Planned Time Point Number	Num		Timing	Numerical version of EXTPT to aid in sorting.	Perm
EXELTM	Planned Elapsed Time from Time Point Ref	Char	ISO 8601 duration	Timing	Planned elapsed time relative to the planned fixed reference (EXTPTREF). This variable is useful where there are repetitive measures. Not a clock time.	Perm
EXTPTREF	Time Point Reference	Char		Timing	Name of the fixed reference point referred to by EXELTM, EXTPTNUM, and EXTPT. Examples: PREVIOUS DOSE, PREVIOUS MEAL.	Perm
EXRFTDTC	Date/Time of Reference Time Point	Char	ISO 8601 datetime or interval	Timing	Date/time for a fixed reference time point defined by EXTPTREF.	Perm

Medication Data in HL7 FHIR R4



11.0.2.1 Medications

Name	Description
MedicationRequest	Represents an instruction for the administration of medication to a patient - both in the inpatient (hospital) and community setting. It can also include instructions for the dispensing, the reasons why the administration should occur and other data. It is called an 'Request' to be consistent with other FHIR resources and the workflow pattern, but a common alias for this resource is a 'Prescription' or an 'Order'. The Order itself represents the content of the instruction and is not, by itself, actionable. The workflow process around 'fulfilling' the order is part of the generic FHIR workflow (see below), with the MedicationRequest representing the contents.
MedicationDispense	The provision of a supply of a medication with the intention that it is subsequently consumed by a patient (usually in response to a prescription).
MedicationAdministration	A record of a patient actually consuming a medicine, or if it has otherwise been administered to them
MedicationStatement	This is a record indicating that a patient may be taking a medication now, has taken the medication in the past, or will be taking the medication in the future. The source for this information can be the patient, significant other (such as a family member or spouse), or a clinician. A common scenario where this information is captured is during the history taking process during a patient visit or stay. A medication statement is not a part of the prescribe->dispense->administer sequence, but is a report that such a sequence (or at least a part of it) did take place, resulting in a belief that the patient has received a particular medication. It may be used to construct a patients 'Current Medications' list.
Medication	The medication resource represents an actual medication that can be given to a patient, and referenced by the other medication resources. In many cases, this resource is not needed and the drug is indicated by a reference to the appropriate terminology and so can be represented using a codeable concept. In other cases, however, it may be desired to indicate more details than the simple drug (such as the packaging, whether it is a generic medication or the active and inactive ingredients) and so the Medication resource can be used for this.
MedicationKnowledge	The MedicationKnowledge resource is draft and is included for comment purposes. This resource represents information about a medication, for example, details about the medication including interactions, contraindications, cost, regulatory status, administration guidelines, etc.

Medication Data in PCORnetv6.1

DISPENSING Domain Description:

Prescriptions filled through a community, mail-order or hospital pharmacy. Outpatient dispensing may not be directly captured within healthcare systems.

PRESCRIBING Domain Description:

Provider orders for medication dispensing and/or administration. These orders may take place in any setting, including the inpatient or outpatient basis.

MED_ADMIN Domain Description:

Records of medications administered to patients by healthcare providers. These administrations may take place in any setting, including inpatient, outpatient or home health encounters.

Drug Exposure Data in OMOPv5.4

drug_exposure

Table Description

This table captures records about the exposure to a Drug ingested or otherwise introduced into the body. A Drug is a biochemical substance formulated in such a way that when administered to a Person it will exert a certain biochemical effect on the metabolism. Drugs include prescription and over-the-counter medicines, vaccines, and large-molecule biologic therapies. Radiological devices ingested or applied locally do not count as Drugs.

RWD Gaps in CDISC SDTM

...and how are these handled by other data models?

Gap: Need for details on record type

- Within the CDISC SDTM EX domain specification, no details exist to indicate the type of record for a medication (e.g., administration record, dispense record, prescription record, medication statement).
- How are these handled in other data models?
 - HL7 FHIR R4: Separate resources for MedicationRequest, MedicationDispense, MedicationAdministration, and MedicationStatement
 - PCORnet v6.1: Separate tables for DISPENSING, PRESCRIBING, and MED_ADMIN
 - OMOP v5.4: Separate tables not present for exposure, however the DRUG_EXPOSURE table contains the drug_type_concept_id concept to provide info about the provenance of a record
- **CDISC Recommendation:** Expand the EX domain specification to include a data element for Record Type or consider the use of multiple domains.

Gap: Additional source data variables needed when tabulating exposure data

- Within the CDISC SDTM EX domain, EXTRT (Name of Treatment) is the name of the protocol-specified study treatment given during the dosing period for the observation.
- When present in the RWD source data, additional treatment variables are needed to account for these, which may include the source medication name, code, and coding system:

Source Medication Name
Source Medication Code
Source Medication Coding System

- These additional treatment variables do not currently exist in the CDISC SDTM domain specification.

Example of additional source variables in FHIR

```
{
  "resourceType" : "MedicationRequest",
  "id" : "medrx0333",
  "contained" : [{
    "resourceType" : "Medication",
    "id" : "med0310",
    "code" : {
      "coding" : [{
        "system" : "http://snomed.info/sct",
        "code" : "430127000",
        "display" : "Oxycodone-containing product in oral dose form"
      }]
    }
  }]
},
```

How is this gap handled in other data models?

- HL7 FHIR R4: Source vs. standardized medication information is not represented.
- PCORnet v6.1 includes the following source-related concepts:

Table	Concept	Note
DISPENSING	RAW_NDC	Field for originating value, prior to mapping into the PCORnet CDM value set.
PRESCRIBING	RAW_RX_MED_NAME	Field for originating, full textual medication name from the source.
PRESCRIBING	RAW_RX_NDC	Field for originating value, prior to mapping into the PCORnet CDM value set
MED_ADMIN	RAW_MEDADMIN_MED_NAME	Field for originating, full textual medication name from the source.
MED_ADMIN	RAW_MEDADMIN_CODE	Field for originating value, prior to mapping into the PCORnet CDM value set.

How is this gap handled in other data models?

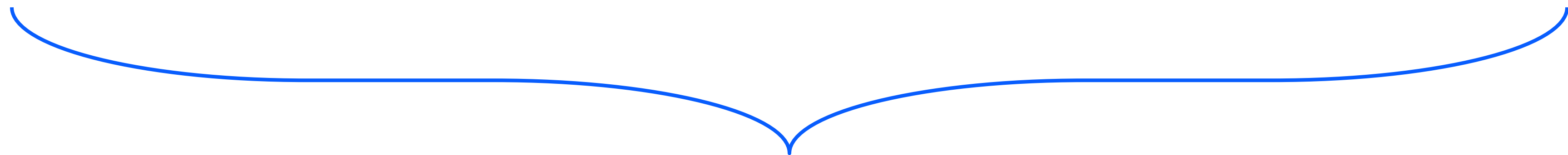
- OMOP v5.4:

Concept	Note
drug_source_value	This field houses the verbatim value from the source data representing the drug exposure that occurred. For example, this could be an NDC or Gemscrip code.
drug_source_concept_id	This is the concept representing the drug source value. For example, the concept_id value 750264 stands for the NDC code for “Acetaminophen 325 MG Oral Tablet”.

Example with additional source variables provided within tabulated dataset

Example of tabulated data:

Source Medication Name	Source Medication Code	Source Medication Coding System	Standardized Name of Treatment
Oxycodone-containing product in oral dose form	430127000	http://snomed.info/sct"	Oxycodone



Does not currently exist in SDTMIG

↑
EXTRT

CDISC Recommendation: Expand the EX domain specification to include additional source variables.

Gap: Additional dosing information needed for certain record types

- Within the CDISC SDTM EX domain and other related interventions domains, no details currently exist to indicate ‘day supply’ or refill information for prescribed medication, which may be useful in deriving end date when data is missing.
 - Possibly useful variables include:

Days Supply: The amount days the drug will be used if the intended dosage is followed.

Fill Quantity: The quantity dispensed to fill the prescription.

Total Number of Refills: The number of refills given in the originating prescription.

How is this gap handled in other data models?

- HL7 FHIR R4:
 - MedicationRequest:

⌄	dispenseRequest	0..1	BackboneElement	Medication supply authorization
⌄	initialFill	0..1	BackboneElement	First fill details
⌄	quantity	0..1	SimpleQuantity	First fill quantity
⌄	duration	0..1	Duration	First fill duration
⌄	dispenseInterval	0..1	Duration	Minimum period of time between dispenses
⌄	validityPeriod	0..1	Period	Time period supply is authorized for
⌄	numberOfRepeatsAllowed	0..1	unsignedInt	Number of refills authorized
⌄	quantity	0..1	SimpleQuantity	Amount of medication to supply per dispense
⌄	expectedSupplyDuration	0..1	Duration	Number of days supply per dispense
⌄	performer	0..1	Reference(Organization)	Intended dispenser

–MedicationDispense:

⌄	type	0..1	CodeableConcept	Trial fill, partial fill, emergency fill, etc. V3 Value SetActPharmacySupplyType (Example)
⌄	quantity	0..1	SimpleQuantity	Amount dispensed
⌄	daysSupply	0..1	SimpleQuantity	Amount of medication expressed as a timing amount

How is this gap handled in other data models?

- PCORnet v6.1:

Table	Concept	Note
DISPENSING	DISPENSE_SUP	Days supply. Number of days that the medication supports based on the number of doses as reported by the pharmacist. This amount is typically found on the dispensing record. Integer values are expected.
DISPENSING	DISPENSE_AMT	Number of units (pills, tablets, vials) dispensed. Net amount per NDC per dispensing. This amount is typically found on the dispensing record. Positive values are expected.
PRESCRIBING	RX_QUANTITY	Quantity ordered.
PRESCRIBING	RX_REFILLS	Number of refills ordered (not including the original prescription). If no refills are ordered, the value should be zero.
PRESCRIBING	RX_DAYS_SUPPLY	Number of days supply ordered, as specified by the prescription.

- OMOP v5.4 within the drug_exposure table includes ‘refills’, ‘quantity’, and ‘days_supply’.

Example with additional dosing variables provided within tabulated dataset

- Example of tabulated data for a prescription record

Standardized Medication Name	Start Date/Time of Medication	Days Supply	Fill Quantity	Total Number of Refills
Amoxicillin	2024-09-10	14	28	0

↑
CMDECOD

↑
CMSTDTC

Does not currently exist in SDTMIG

- **CDISC Recommendation:** Expand relevant interventions domain specifications (EC, EX, CM) to include additional dosing variables.

Gap: Need for additional contextual variables

- Within an RCT, the protocol specifies details such as how an assessment is conducted, who performs the assessment, and where the assessment is completed. Thus, this data is not captured during data collection.
- Since RWD is not initially collected under a protocol, submitting this contextual information can help increase a reviewer's confidence and understanding of the submitted RWD.
- Further, 'not all sites are created equal' in RWD (for example, consider a large medical center versus a corner drug store) and clinical staff have different skill sets and specializations. Thus, this information can be especially useful in the context of understanding RWD.
- **CDISC Recommendation:** Submit these supplemental “contextual” variables as additional qualifiers to the SDTM EX/EC domains, which may include who conducted the assessment and the name and location of the site where the assessment was done.

Gap: Need for additional contextual variables (Cont'd)

- FHIR, OMOP, and PCORnet take a similar approach to providing this information
 - Medication-related resources or tables contain data elements for Provider ID, Encounter ID, and Facility or Location ID.
 - These data elements link to either a Provider table, Encounter table, or a Facility/Location table (or resource) that contains extensive information about a given provider, encounter, or facility.
- For example, the Prescribing Table in PCORnet contains the variables RX_PROVIDERID and ENCOUNTERID.
 - The RX_PROVIDERID field stores a code for the provider who prescribed the medication and links to the Provider table, which contains extensive details about the provider.
 - Similarly, the ENCOUNTERID field holds a code for the encounter associated with the prescribing event and links to the Encounter table, which provides details about both the facility and encounter.
- OMOP and FHIR use the same methodology for linking a medication event with a provider, facility, and encounter.

Gap: Need for additional contextual variables (Cont'd)

- **CDISC Recommendation:** Implement the FHIR/PCORnet/OMOP strategy:
 - 1) Add Provider ID and Facility ID in the SDTM EX/EC domains (as well as other applicable domains) as supplemental qualifiers.
 - 2) Create a PROVIDER domain to store details about a given provider.
 - 3) Either expand the Subject Visits (SV) domain by adding facility related data elements or create a new domain which contains details about a given facility.
 - 4) Provider ID will link EX/EC to the PROVIDER domain.
 - 5) Facility ID will link EX/EC to the domain storing the facility information.

Additional traceability variables

- Additional site, subject, and record identifiers
 - Facilitate traceability to the source
 - Provide increased confidence in RWD to reviewers
 - Needed for RWD since records can come from a variety of sources
- **RWD source**
 - indicates where the RWD comes from
 - can have values like EHR, Registry, Claims, etc.
- **Source subject identifier**
 - subject identifier contained in the source data.
 - likely be encrypted or scrambled
- **Source record identifier**
 - record identifier from the source data (also likely encrypted)
 - used to link back to the original record in the source data
 - in many cases, this will also be the encounter identifier for a subject

Summary

- Gaps are present in the CDISC SDTM domain specifications for exposure and other interventions domains, including but not limited to:
 - Lack of record type information
 - Lack of source medication information
 - Lack of additional dosing variables
 - Lack of additional contextual variables
 - Lack of additional linking variables
- This data is captured in the real-world setting, and sponsors and regulators alike need to have a high level of confidence in the data they are evaluating.
 - This differs from RCTs, where the captured data is planned and known.
- To address these gaps, CDISC may consider expanding relevant domain specifications to better incorporate the use of RWD within regulatory review.

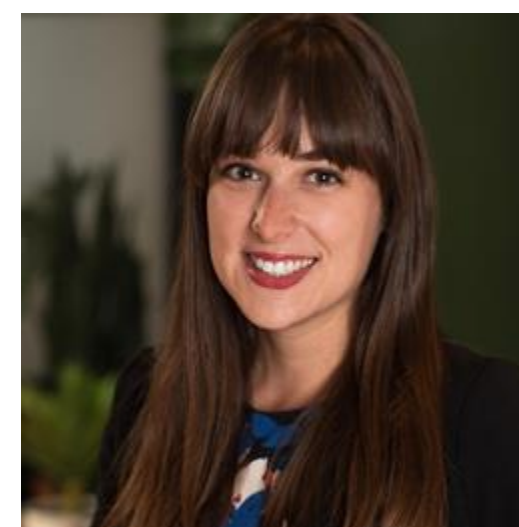
Conclusion

- Guidance for submitting RWD is evolving
- Standards for representing RWD is also evolving
- However, current submission standards have numerous gaps for representing RWD
- CDISC standards must be supplemented to represent RWD
- Learn about RWD CDMs like FHIR, OMOP, and PCORnet
- Incorporate solutions from CDMs designed for RWD
- Early communication with regulatory authorities is key to meet review's needs
- Reach out to SDOs like CDISC to improve RWD solutions
- Join the PHUSE RWE group to collaborate on submitting RWD

Thank you!



Jeff Abolafia
Pinnacle21
jeffrey.abolafia@certara.com



Sarah Ferko
IBM Consulting
smferko@us.ibm.com



Ingeborg Holt
Orizaba Solutions
iholt@orizabasolutions.com

Questions?