cdisc

2022
US
INTERCHANGE
26-27 OCTOBER | AUSTIN

The Continuing Evolution of SDTM Metadata

Diane Wold, PhD, Sr. Director, Standards Development, CDISC

Meet the Speaker

Diane Wold

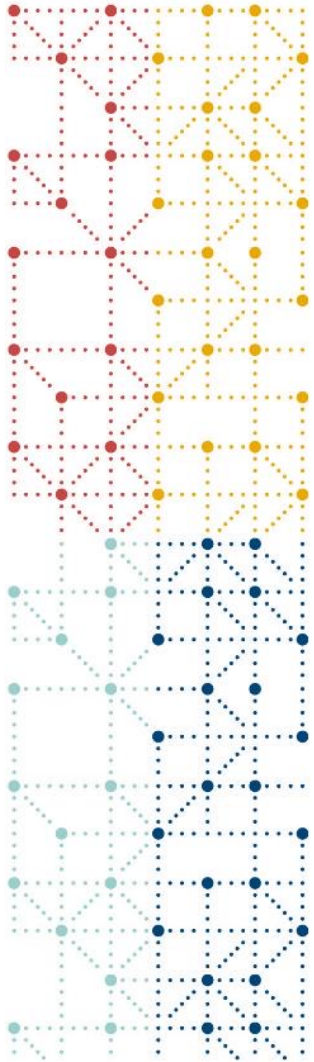**Title:** Sr. Director, Standards Development

Organization: CDISC

Diane Wold received her Ph.D. in Statistics from the University of North Carolina at Chapel Hill. She worked for Burroughs Wellcome/Glaxo Wellcome/Glaxo Smith Kline in a variety of roles for over 30 years. At the Glaxo Smith Kline merger, she joined the data standards group, and in 2002 she joined the CDISC SDS team. She was also involved in other CDISC teams, including the Protocol Representation Group and SHARE. In 2012 she became involved in the CFAST initiative to develop therapeutic area standards. In 2015 she joined CDISC as an employee. Among other activities, she is currently working on SDTM variable definitions, SDTM variable roles, and the CDISC Knowledge Base.
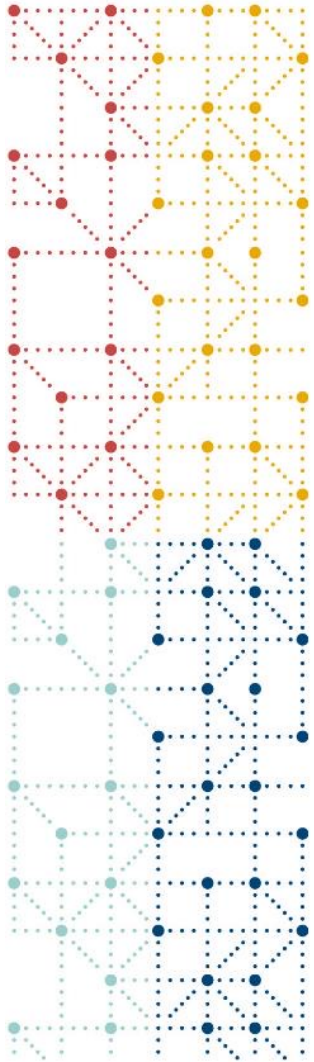
Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*

- *{Please disclose any financial relationship or conflict of interest relevant to this presentation here OR}*

- *The author(s) have no real or apparent conflicts of interest to report.*

Agenda

1. Introduction
2. Evolution of SDTM metadata in versions 1.x
3. Changes in SDTM v2.0
4. Changes proposed for SDTM with SDTMIG v4.0
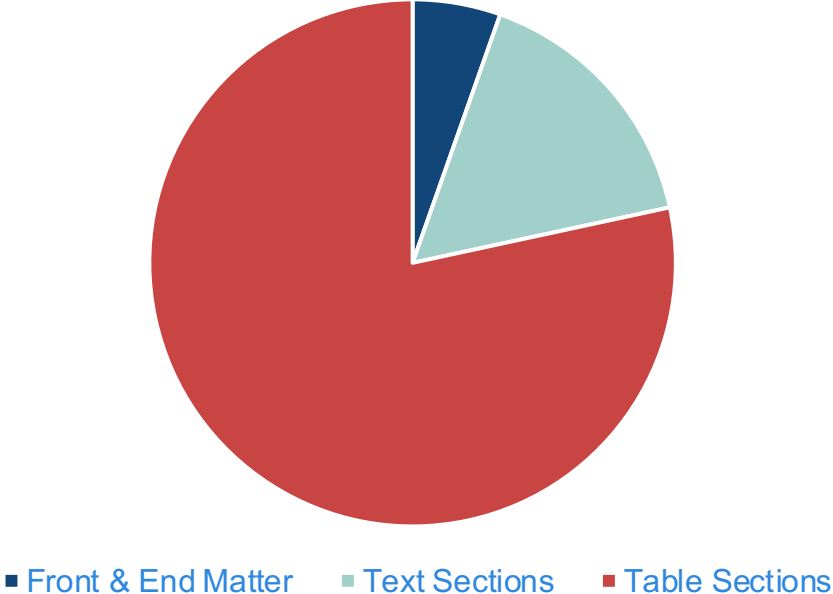5. Possible future changes

Introduction
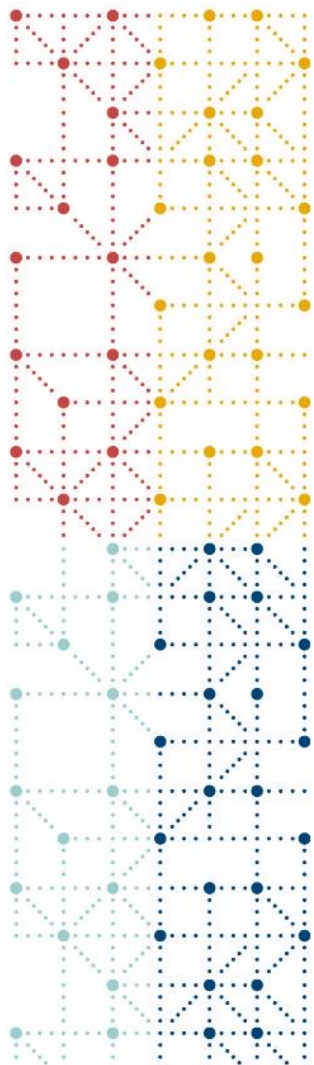
The SDTM is mostly variable metadata

Introduction
The SDTM is Mostly Variable Metadata

SDTM v2.0 Pages

■ Front & End Matter ■ Text Sections ■ Table Sections

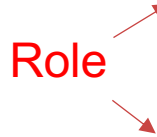# Evolution of Metadata in SDTM versions 1.x

SDTM 1.0
Four metadata items (columns)

2.2.3 The Findings Observation Class

Table 2.2.3: Findings — Topic and Qualifier Variables, One Record per Finding (--TESTCD)

Variable Name	Variable Label	Type	Description
	Topic Variable		
--TESTCD	Short Name of Measurement, Test or Examination	Char*	Short character value for --TEST used as a column name when converting a dataset from a vertical format to a horizontal format. The short value can be up to 8 characters. Examples: PLATELET, SYSBP, PR, EYEEXAM.
	Qualifier Variables		
--TEST	Name of Measurement, Test or Examination	Char*	Verbatim name, corresponding to the topic variable, of the test or examination used to obtain the measurement or finding. Examples: Platelet Count, Systolic Blood Pressure, PR Interval, Eye Examination.
--MODIFY	Modified Term	Char	If the value of --TEST is modified as part of a defined procedure, then the modified text is placed here.
--CAT	Category	Char*	Used to define a category of related records. Examples: HEMATOLOGY, URINALYSIS, CHEMISTRY, HAMILTON DEPRESSION SCALE, SF36.
--SCAT	Subcategory	Char*	Used to define a further categorization level for a group of related records. Example: DIFFERENTIAL.
--POS	Position of Subject During Observation	Char*	Position of the subject during a measurement or examination. Examples: SUPINE, STANDING, SITTING.

Role

SDTM v1.2

Table 2.2.6: Subject Demographics Domain Variables

Variable Name	Variable Label	Type	Description
Identifier Variables			
STUDYID	Study Identifier	Char	Unique identifier for a study.
DOMAIN	Domain Abbreviation	Char	Two-character abbreviation for the domain which must be DM.
USUBJID	Unique Subject Identifier	Char	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.
Topic Variable			
SUBJID	Sub...		
RFSTDTC	Su...		

Role

Table 2.2.7: Comments Domain Variables

Variable Name	Variable Label	Type	Role	Description
STUDYID	Study Identifier	Char	Identifier	Unique identifier for a study.
DOMAIN	Domain Abbreviation	Char	Identifier	Two-character abbreviation for the domain which must be CO.
RDOMAIN	Related Domain Abbreviation	Char	Record Qualifier	Domain Abbreviation of the parent record(s). Null for records collected on general comments or additional information CRF page.
USUBJID	Uni...	Char	Identif...	Identif...

SDTM v1.7

Format column

Table 2.2.6.1 Subject Demographics Domain Variables

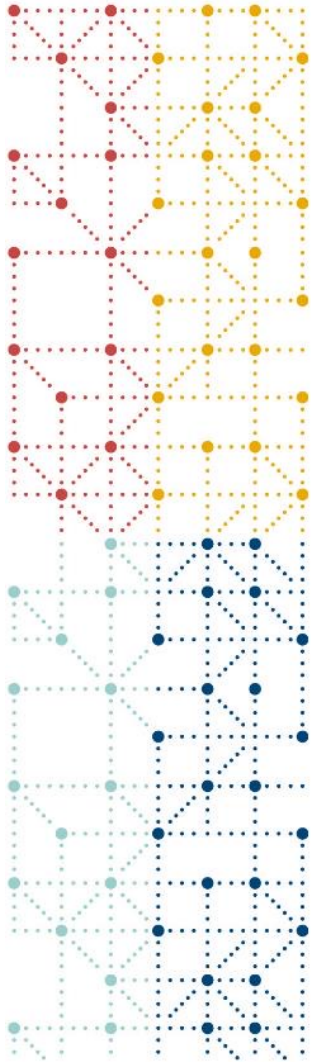| Variable Name	Variable Label	Type	Format	Role	Description
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.
DOMAIN	Domain Abbreviation	Char		Identifier	2-character abbreviation for the domain, which must be "DM".
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.
SUBJID	Subject Identifier for the Study	Char		Topic	Subject identifier, which must be unique within the study. Often the ID of the subject as recorded on a CRF.
RFSTDTC	Subject Reference Start Date/Time	Char	ISO 8601	Record Qualifier	Reference start date or start date and time for the subject. Usually equivalent to date/time when subject was first exposed to study treatment. Required for all randomized subjects; will be null for all subjects who did not meet the milestone the date requires, such as screen failures or unassigned subjects.
RFENDTC	Subject Reference End Date/Time	Char	ISO 8601	Record Qualifier	Reference end date/time for the subject. Usually equivalent to the date/time when subject was determined to have ended the trial, and often equivalent to date/time of last exposure to study treatment. Required for all randomized subjects; null for screen failures or unassigned subjects.
RFXSTDTC	Date/Time of First Study Treatment	Char	ISO 8601	Record Qualifier	First date/time of exposure to any protocol-specified treatment or therapy for the subject.
RFXENDTC	Date/Time of Last Study Treatment	Char	ISO 8601	Record Qualifier	Last date/time of exposure to any protocol-specified treatment or therapy for the subject.
RFICDTC	Date/Time of Informed Consent	Char	ISO 8601	Record Qualifier	Date/time of informed consent.
RFPENDTC	Date/Time of End of Participation	Char	ISO 8601	Record Qualifier	Date/time when subject ended participation or follow-up in a trial. Should correspond to the last known date of contact.
DTHDTC	Date/Time of Death	Char	ISO 8601	Record Qualifier	Date/time of death for any subject who died, in ISO 8601 format. Should represent the date/time that is captured in the clinical-trial database.
DTHFL	Subject Death Flag	Char		Record Qualifier	A value of "Y" indicates the subject died. Should be "Y" or null. Should be populated even when the death date is unknown.
SITEID	Study Site Identifier	Char		Record Qualifier	Unique identifier for a site within a study.
INVID	Investigator Identifier	Char		Record Qualifier	An identi... ition to the SITEID. Not needed if SITEID is equivalen...
INVNAM	Investigator Name	Char		Synonym Qualifier of INVID	Name of...
BRTHDTC	Date/Time of Birth	Char	ISO 8601	Record Qualifier	Date/time of birth of the subject.

Variable(s) Qualified

SDTM v1.7

2.2.12 Domain-Specific Variables for the General Observation Classes

The concept of domain-specific variables was first introduced in SDTM v1.5. These variables are for use only in a specific domain and will be identified in the appropriate implementation guide. The variable names include the specific domain prefix. Table 2.2.12.1 lists the proposed domain-specific variables.

Table 2.2.12.1 Domain-Specific Variables for General Observation Class Domains

Observation Class	Domain	Variable Name	Variable Label	Type	Role	Description	Position
Events	MH	MHEVDTYP	Medical History Event Date Type	Char	Variable Qualifier of MHSTDTC and/or MHENDTC	Specifies the aspect of the medical condition or event by which MHSTDTC and/or the MHENDTC is defined. Examples: "DIAGNOSIS", "SYMPTOMS", "RELAPSE", "INFECTION".	After MHDECOD
Interventions	EX	EXMETHOD	Method of Administration	Char	Record Qualifier	Method of administration of the treatment. Not to be used with human clinical trials.	After EXLOC
Findings	EG	EGBEATNO	ECG Beat Number	Num	Variable Qualifier of EGORRES	A sequence number that identifies the beat within an ECG.	After EGPOS
Findings	IC	ICIMPLBL	Implantation Site Label	Char	Record Qualifier	Label or identifier that describes the location or position of a fetal implantation site in the uterus (or uterine horn) when classifying implantations during a uterine examination in a reproductive toxicology study. Not to be used with human clinical trials.	After ICSTRESC
Findings	MS	MSAGENT	Agent Name	Char	Record Qualifier	The name of the drug or other material for which resistance is tested. The agent may be used for in vitro testing or may be used in tests for genetic markers or in direct phenotypic drug-sensitivity testing.	After MSTEST
Findings	MS	MSCONC	Agent Concentration	Num	Variable Qualifier of MSAGENT	The amount of drug or other material listed in MSAGENT per unit volume or weight. Used when the agent is part of the pre-specified test. Not to be used when the concentration is a result of a test such as minimal inhibitory concentration, IC50, or EC50.	After MSAGENT
Findings	MS	MSCONCU	Agent Concentration Units	Char	Variable Qualifier of MSCONC	Unit of measure for MSCONC.	After MSCONC

SDTM v2.0

Big Changes to Variable Metadata
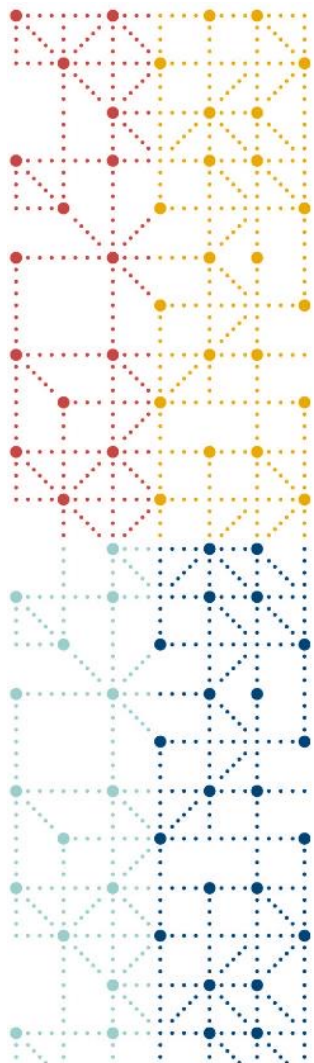
SDTM v2.0 Changes

- Domain-specific variables rolled into General Observation Class tables
- All tables have the same metadata items
- Role split
- C-code and Definition added
- Description replaced by Usage Restrictions, Notes, Examples

- Motivated in part by desire for more usable metadata in CDISC Library

cdisc

SDTM 2.0 Variable Metadata

3.1.1 The Interventions Observation Class

Interventions—Topic and Qualifier Variables—One Record per Constant-dosing Interval or Intervention Episode

#	Variable Name	Variable Label	Type	Format	Role	Variable(s) Qualified	Usage Restrictions	Variable C-code	Definition	Notes	Examples
1	--TRT	Name of Treatment	Char		Topic			C82542	The reported name of the drug, procedure, or therapy.	The topic for the intervention observation, usually the verbatim name of the treatment, drug, medicine, or therapy given during the dosing interval for the observation.	
2	--MODIFY	Modified Treatment Name	Char		Synonym Qualifier	--TRT		C170998	A value which represents an alteration to a collected value for coding purposes.	If the value for --TRT is modified for coding purposes, then the modified text is placed here.	
3	--DECOD	Standardized Treatment Name	Char		Synonym Qualifier	--TRT		C170991	Standardized or dictionary-derived text for the description of an event or intervention.	Equivalent to the generic drug name in WHODrug, or a term in SNOMED, ICD-9, or other published or sponsor-defined dictionaries.	
4	--MOOD	Mood	Char		Record Qualifier			C117051	The state that may be applied to a record to indicate its phase in a life cycle or business process (e.g., scheduled, performed).		"SCHEDULED", "PERFORMED"
5	--CAT	Category	Char		Grouping Qualifier			C25372	A grouping or classification of the topic of the finding, event, or intervention.		
6	--SCAT	Subcategory	Char		Grouping Qualifier				A further grouping or classification of the category for the topic of the finding, event, or intervention.	The category is in --CAT.	
7	--PRESP	Pre-Specified	Char		Variable Qualifier	--TRT	Not in nonclinical trials	C82510	An indication that the event or intervention was prospectively stated or detailed on the CRF.	Values should be "Y" or null.	

Proposed Changes

To be introduced with SDTMIG v4.0

Proposed Changes

- Remove Qualifier role sub-classes
    - Grouping Qualifier, Result Qualifier, Synonym Qualifier, Variable Qualifier, Record Qualifier
    - All will have a role of simply "Qualifier"
- Add Variable Groups
- Add Variable Relationships

Rationale for Proposed Changes

- Role is not submitted, so minimal impact
- Qualifier role subclasses were never well defined
- Qualifier role subclasses were assigned inconsistently
- Assigning qualifier role subclasses was sometimes controversial

- Conclusions:
  - Qualifier role subclasses were little used, poorly understood, and not worth trying to reform
  - A different classification of variables (Variable Groups) could be more useful for understanding how variables are used in the standard
  - The small amount of information in "Variable(s) Qualified" would be better represented by more explicit representation of relationships between variables.

Variable Groups

- Intended as informative

- Reflect concepts and wording already being used informally. e.g.
    - Visit variables
    - Relative timing variables
    - Coding variables
    - Method variables

- Allow grouping of similar groups not grouped in standard variable order
    - For example, several Findings qualifier describe a test's method, but are scattered in order

- May help users more easily pick out the variables needed in a particular case

- Will replace "Variable(s) Qualified" column in SDTM tables

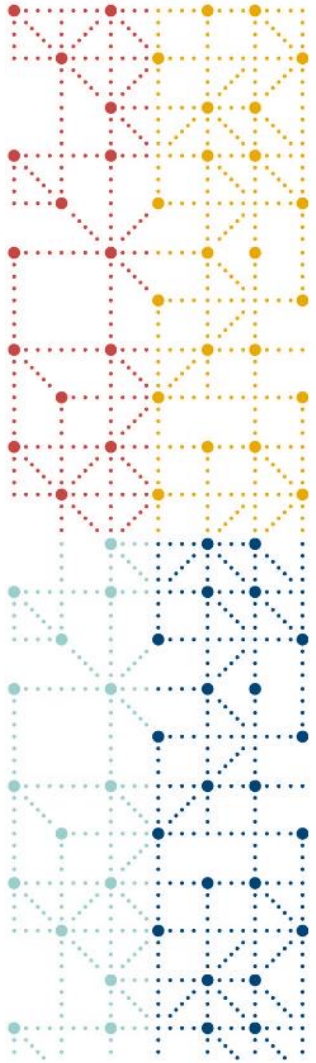cdisc

Variable Relationships

- "Variable(s) Qualified" provides a small amount of information about relationships between variables
  - Does not describe the nature of the relationship
  - Describes only a few of the relationships between variables

- Since a variable can be part of multiple relationships, variable relationships will be in separate tables, something like this:

Source Variable Name	Source Variable Label	Linking Phrase	Target Variable Name	Target Variable Label
--TEST	Name of Measurement, Test, or Exam	Decodes the value in	--TESTCD	Short Name of Measurement, Test, or Exam
--CAT	Category	Groups values in	--TESTCD	Short Name of Measurement, Test, or Exam
--POS	Position of Subject During Observation	Is the subject position during performance of the test in	--TESTCD	Short Name of Measurement, Test, or Exam
--ORRESU	Original Units	Is the unit for the value in	--ORRES	Result or Finding in Original Units

Variable Relationships Strategy

- Uses
  - Enable semantic reasoning
  - Used to construct Biomedical Concepts
  - Enhance understanding and correct use of variables

- Development
  - There are many more relationships than currently reflected in "Variable(s) Qualified"
  - Developing explicit representations of currently implicit relationships will take a lot of work
  - First phase will cover relationships currently reflected in "Variable(s) Qualified," those needed for initial Biomedical Concept development, and very low-hanging fruit
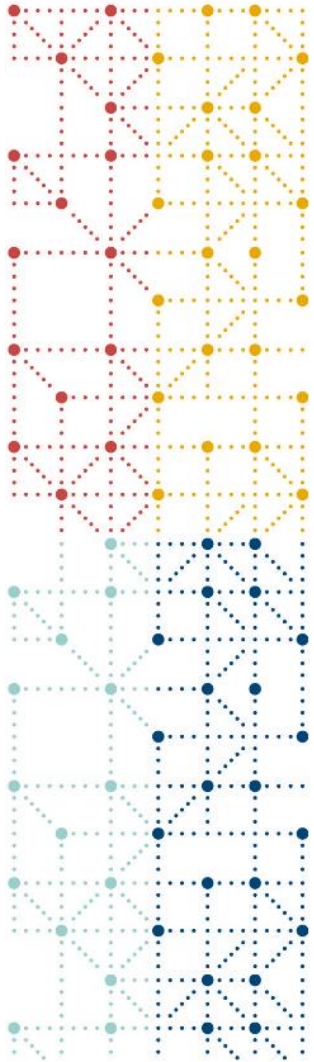
cdisc

Possible Future Changes

Possible Future Additions to SDTM Metadata

- Codelists
  - Some variables inherently use particular codelists (e.g., flag variables)
  - Codelists for some general observation class variables are domain-specific, but a default codelist is used for most domains (e.g, Anatomic Location for –LOC)
    - Similar to SDTM variable labels, which are a default that may be overridden at the domain level
  - Could help with constructing a custom domain

- Allowed Values
  - Value for many variables that use the "No Yes Response" codelist

cdisc

Thank You!

Diane Wold diane.wold@cdisc.org

cdisc