



2023

US

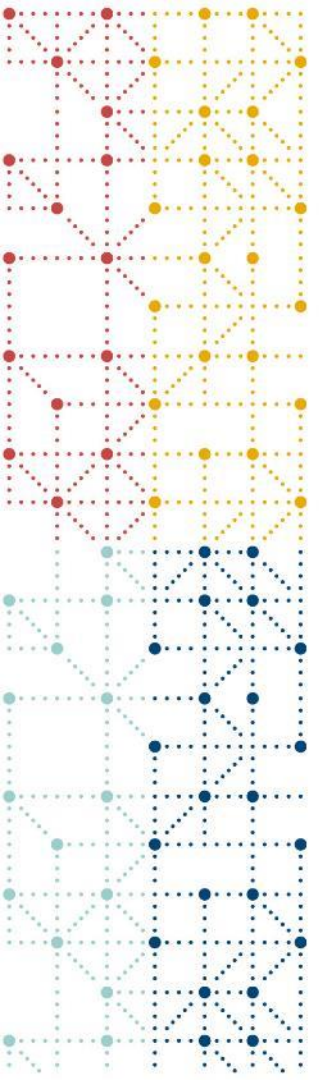
INTERCHANGE

FALLS CHURCH, VA | 18-19 OCTOBER



CDISC Biomedical Concepts and Dataset Specializations Pragmatic Implementation and Tangible Value

Jon Neville, Senior Director Standards Development, CDISC
Lex Jansen, Senior Director Data Science Development, CDISC
Linda Lander, Director Data Science, CDISC



Meet the Speakers

Jon Neville

Title: Senior Director, Standards Development

Organization: CDISC

Jon Neville has been working in CDISC standards development since 2009. He has been working at CDISC for 6 years, where he is currently Senior Director, Standards Development at CDISC



Lex Jansen

Title: Senior Director, Data Science Development

Organization: CDISC

Lex Jansen is an independent consultant, currently working as Senior Director, Data Science Development at CDISC. Before he was a Principal Solution Consultant and Principal Software Developer at SAS Institute.



Linda Lander

Title: Director, Data Science

Organization: CDISC

Linda Lander is an independent contractor, currently working as Director Data Science and Biomedical Concepts Product Owner at CDISC. Before she was Director Data Standards at GlaxoSmithKline



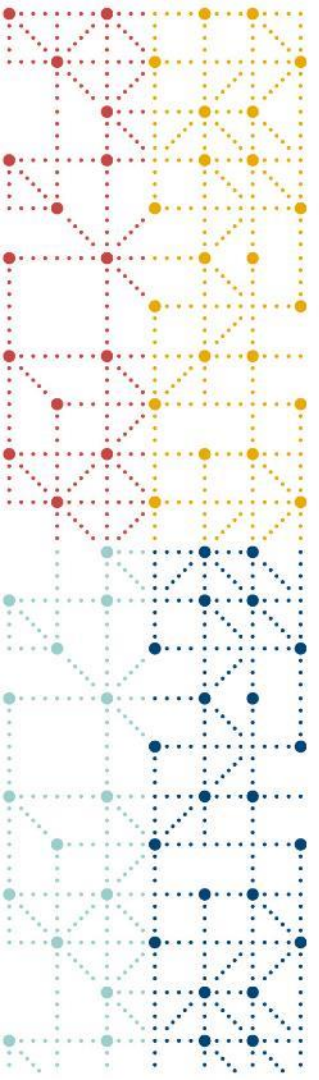
Disclaimer and Disclosures

- *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.*
- *The authors have no conflicts to disclose*



Agenda

1. CDISC Biomedical Concepts - Introduction and Background
2. SDTM Dataset Specializations as Define-XML Building Blocks
3. Retrieval of BCs and SDTM Dataset Specializations using CDISC Library APIs
4. Key Accomplishments, Current Status and What's Coming

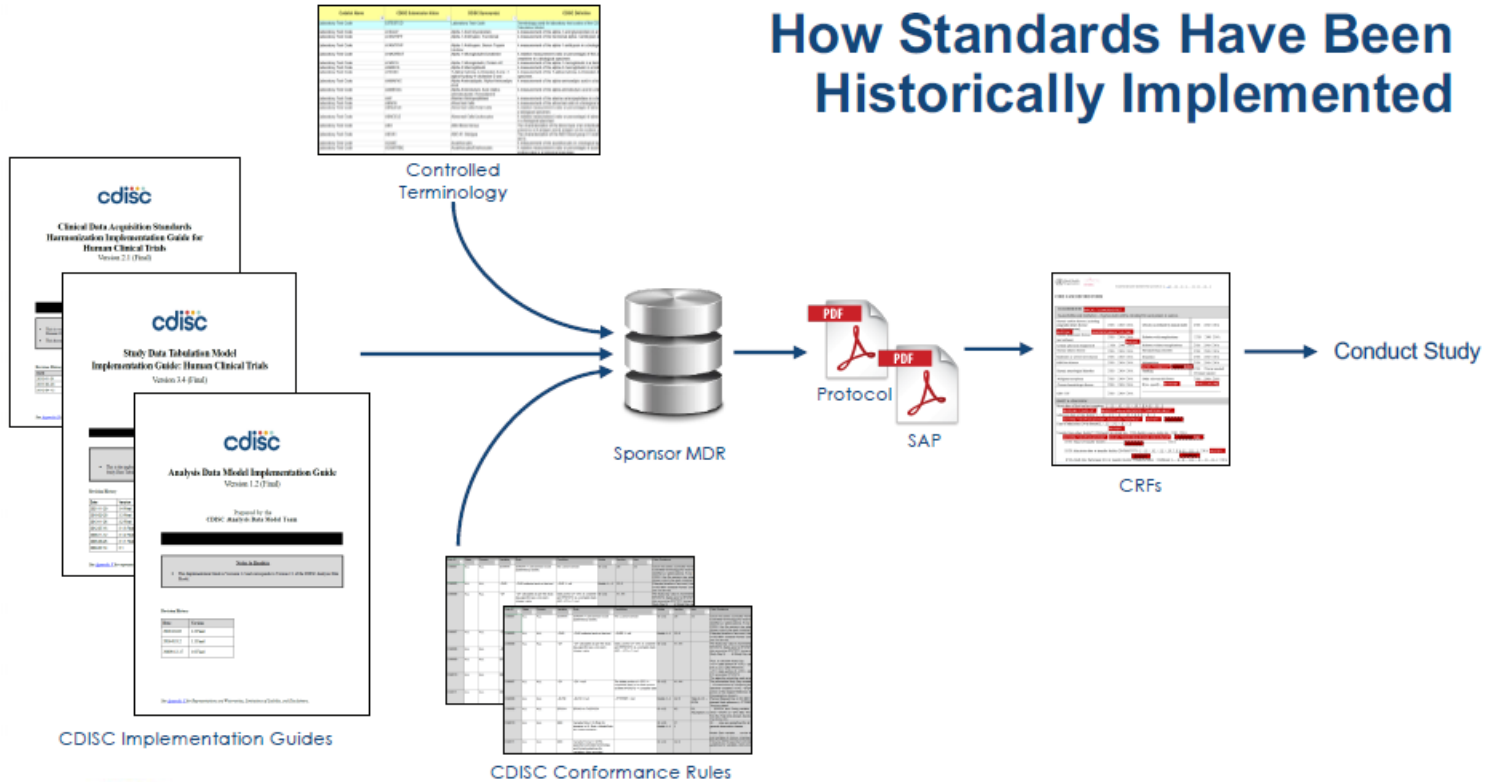


CDISC Biomedical Concepts and SDTM Dataset Specializations

Background and Introduction

CDISC Biomedical Concepts and SDTM Dataset Specializations

How Standards Have Been Historically Implemented





CDISC Biomedical Concepts and SDTM Dataset Specializations

Problem:

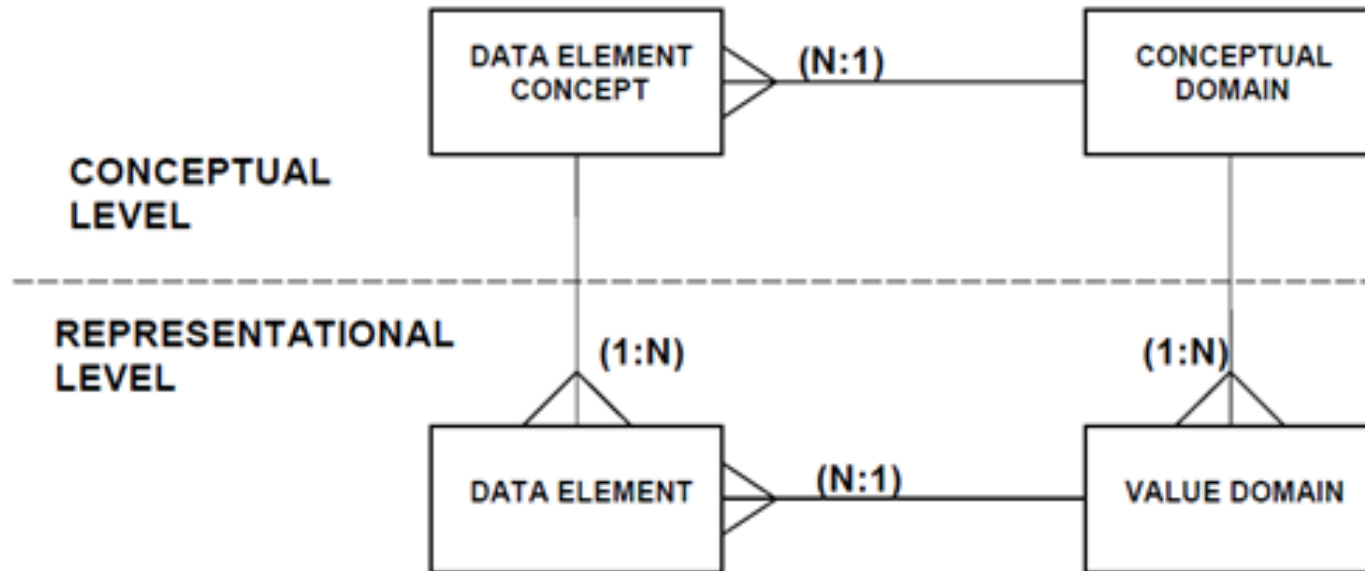
- Labor-intensive; requires extensive knowledge of standards documents
- Subject to interpretation (and therefore, misinterpretation)
- Can result in inconsistent implementation
- The intense effort required is a barrier to standards adoption

CDISC has evolved:

- CDISC Library has published data standards as groups of linked metadata
- Defined relationships between variables, associated terminology codelists, and linkages across standards
- CDISC 360 piloted the development of linked **Biomedical Concept** metadata to enable end to end automation

What Is a Biomedical Concept (BC)?

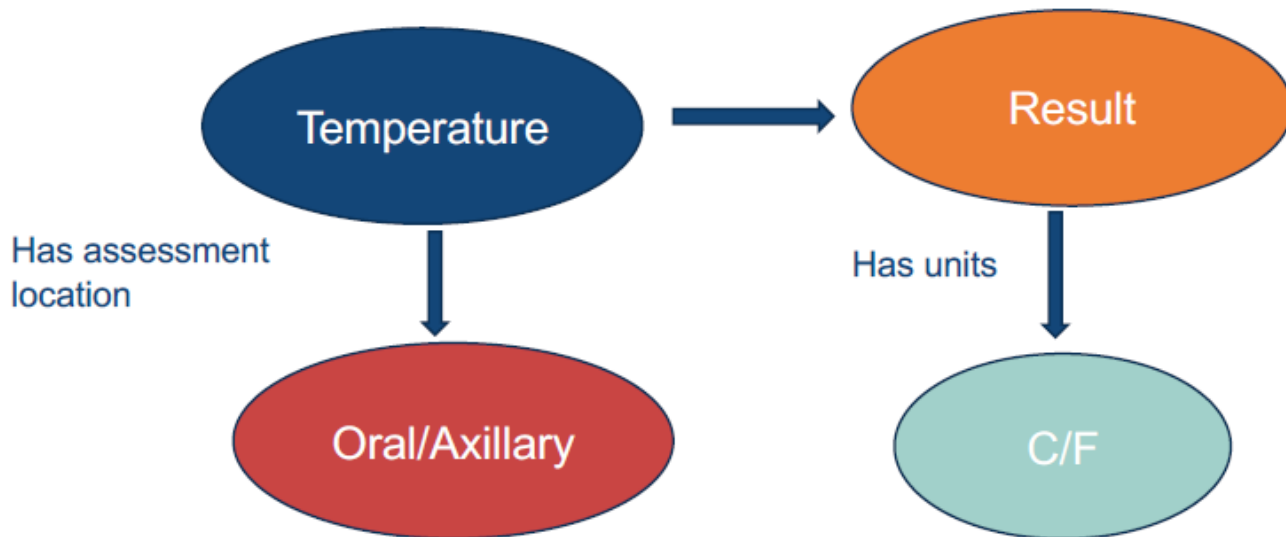
ISO 11179 Definition: A unit of knowledge created by a unique combination of characteristics



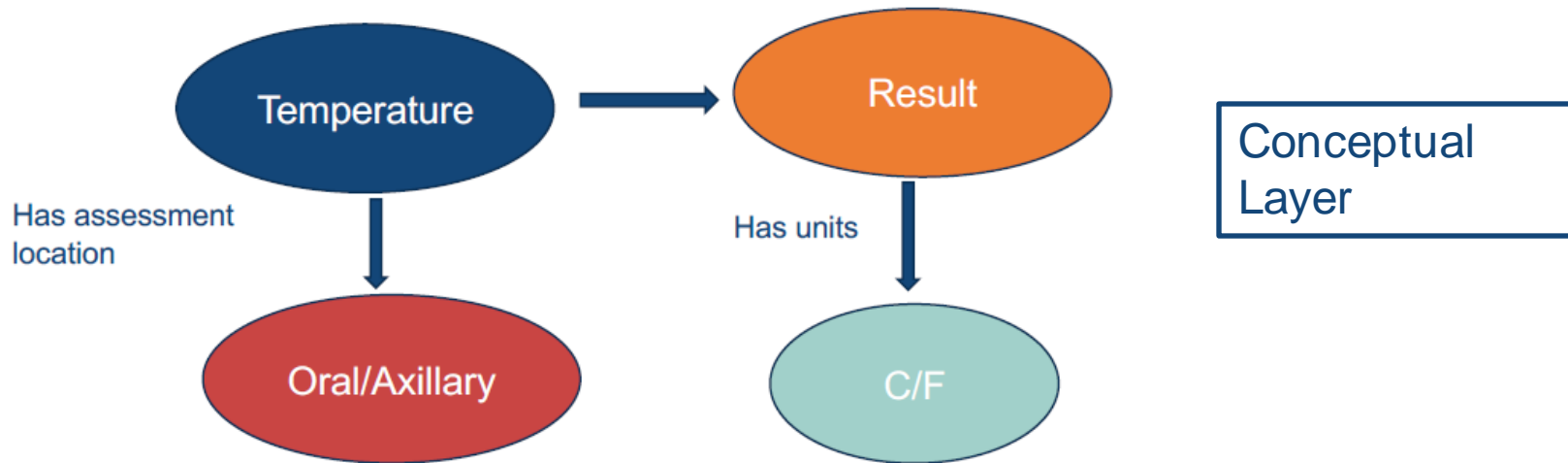
What Is a Biomedical Concept (BC)?

ISO 11179 Definition: A unit of knowledge created by a unique combination of characteristics

- Independent of study
- Independent of a representation in any standard, but can be tethered to a standard



What Is a Biomedical Concept (BC)?



VSTEST	VSTESTCD	VSORRES	VSUNIT	VSLOC
Temperature	TEMP	101.3	F	ORAL

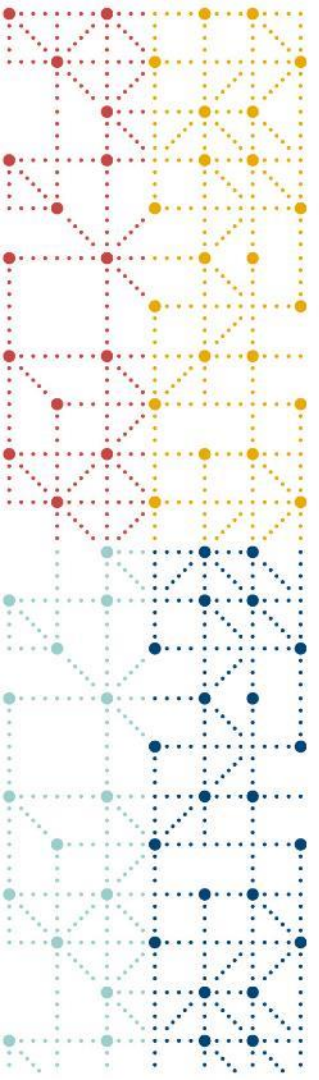
Implementation Layer

CDISC Biomedical Concepts and SDTM Dataset Specializations

Developing Biomedical Concepts allows accurate and **more consistent implementation** of the *conceptual content* being implemented

3 Key pieces of the **Pragmatic Implementation**:

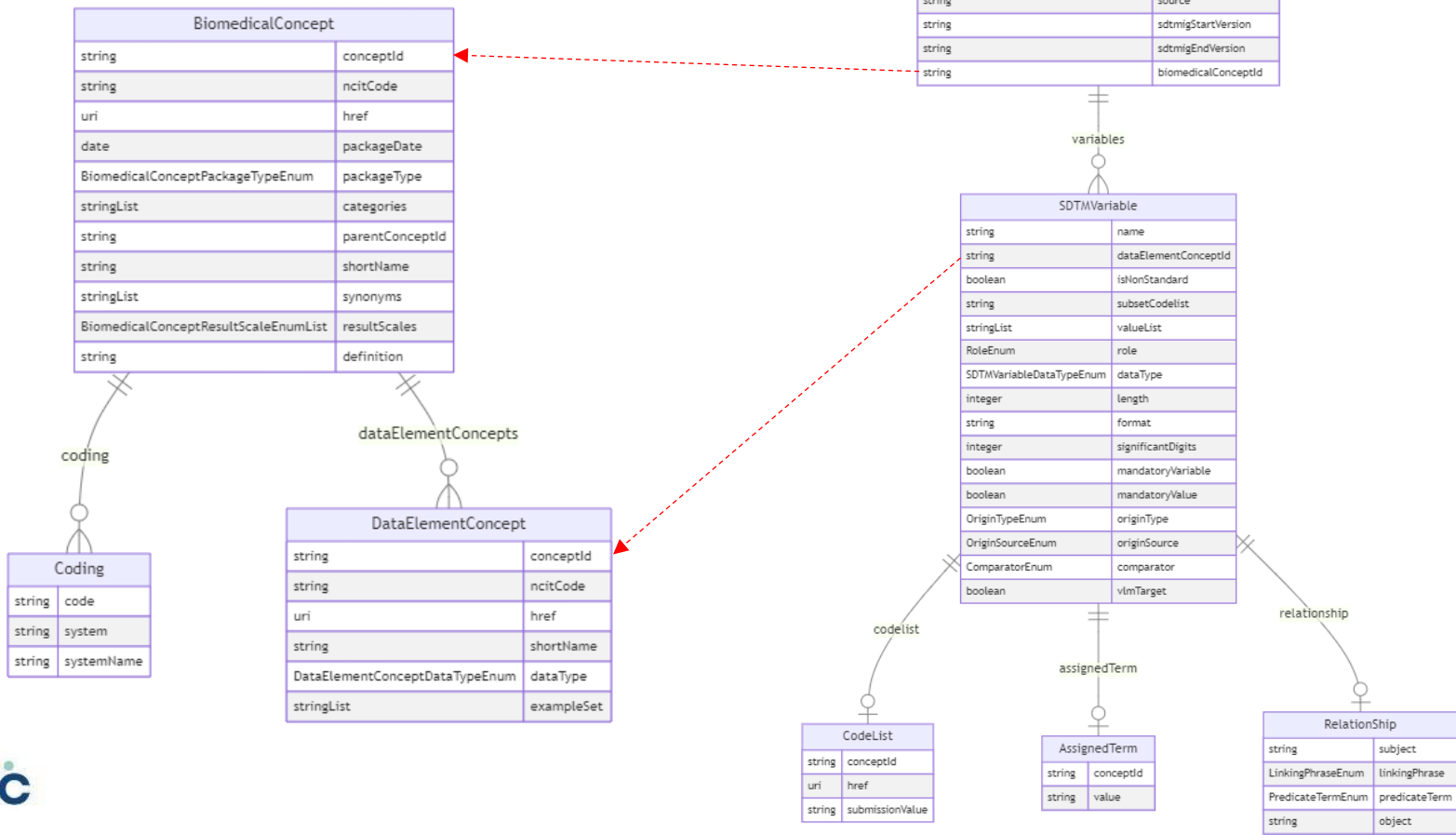
- Extend foundational standards
 - Add explicit relationships between variables
 - Additional operational metadata, e.g., data type, etc.
- Conceptual Layer – abstract BC's
 - Provides semantics - aligned with NCI terminology
 - Supports study design, Schedule of Activities (SOA)
- Implementation Layer - Dataset Specializations with VLM definitions
 - Supports programmers
 - Pre-configured building blocks for Define-XML
 - Tailored to BCs to link with unambiguous semantics & definitions
 - Dataset specializations as an extended dataset structure



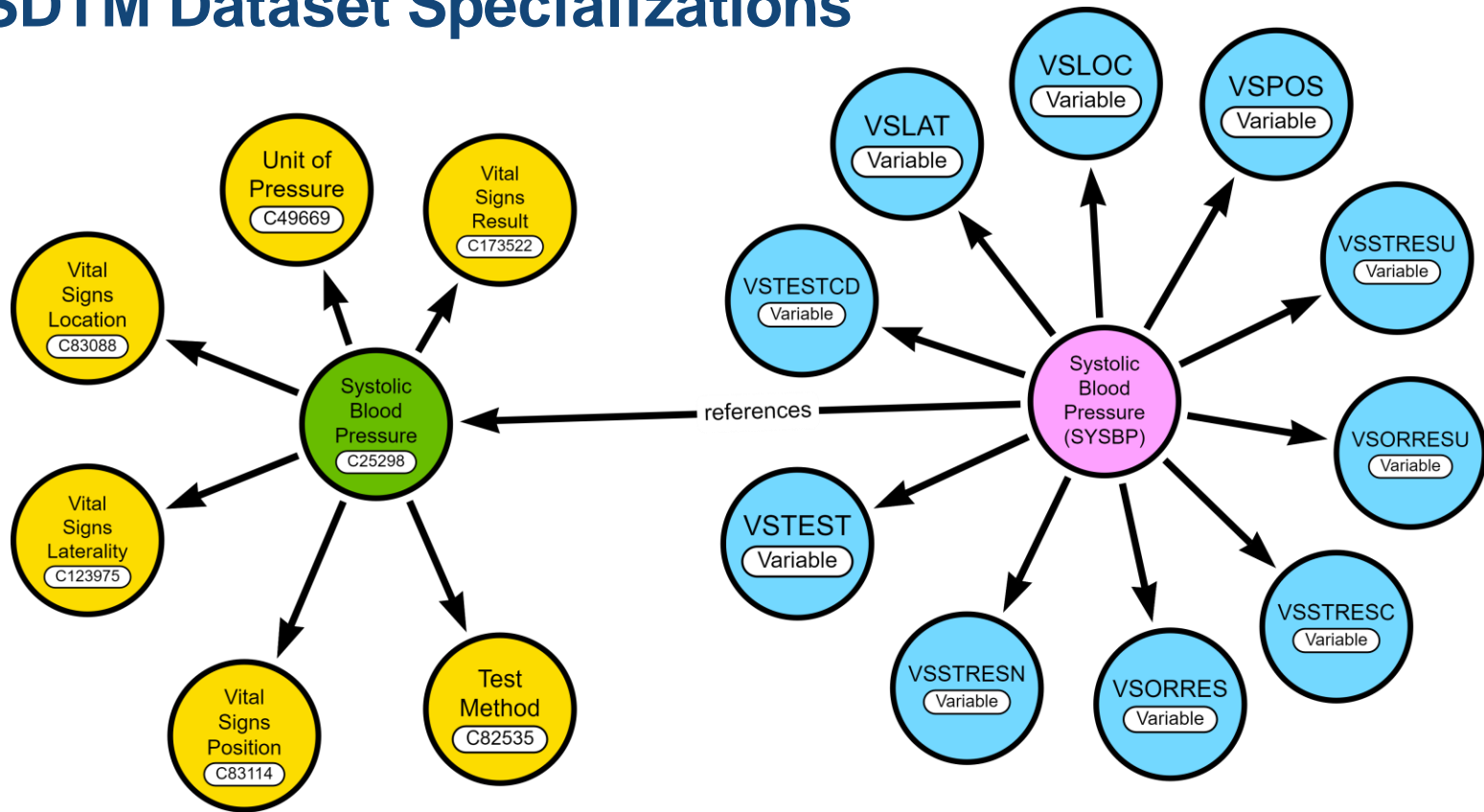
SDTM Dataset Specializations

Building Blocks for Define-XML

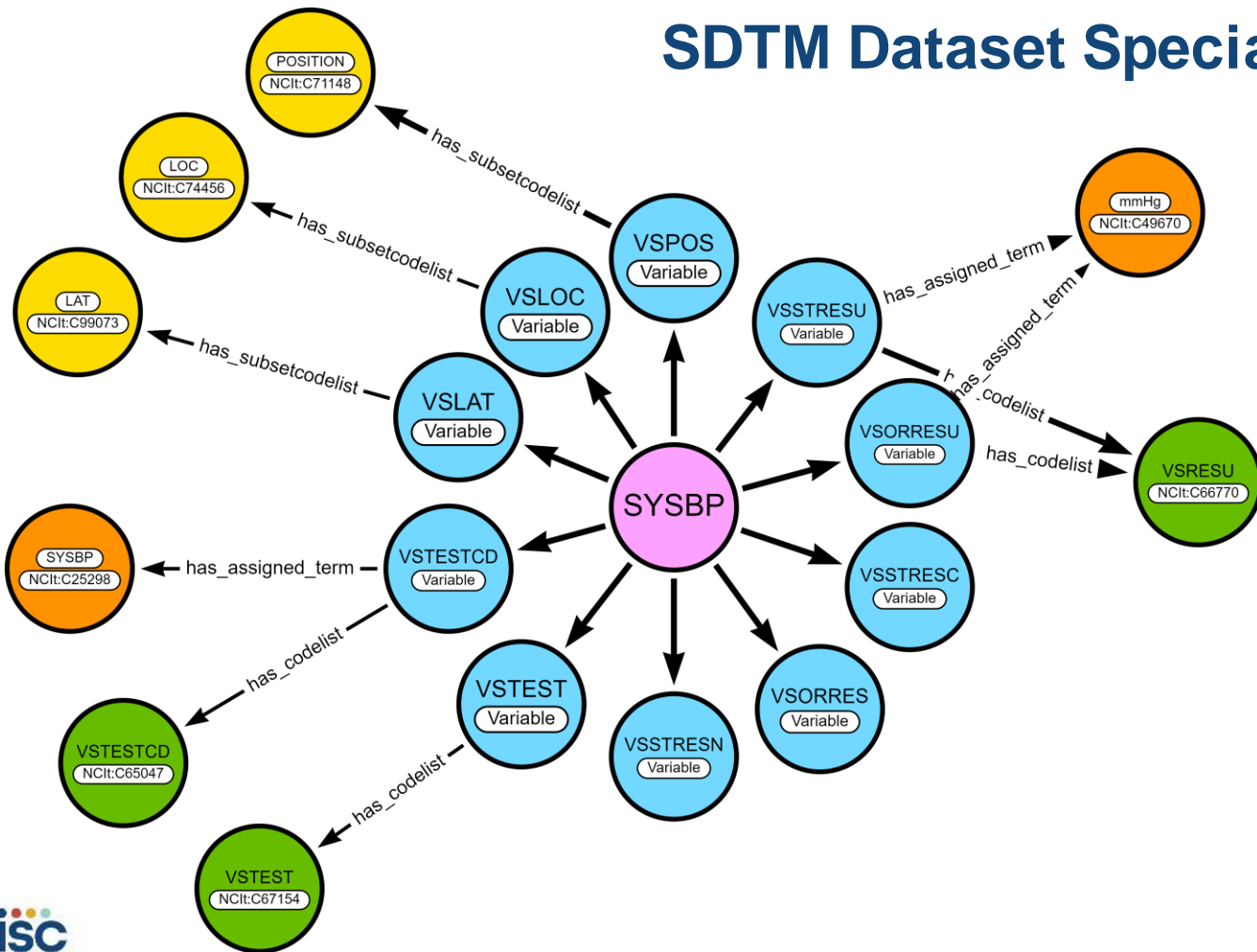
The logical Model



CDISC Biomedical Concepts and SDTM Dataset Specializations



SDTM Dataset Specializations



SDTM Dataset Specializations

Attribute	Description
datasetSpecializationId	Identifier for SDTM Value Level Metadata group
domain	Domain for the SDTM specialization group
shortName	SDTM group short name which provides a user friendly and intuitive name for the datasetSpecializationId
source	SDTM VLM Source which categorizes VLM groups by topic variable
sdtmigStartVersion	The earliest SDTMIG version applicable to the SDTM dataset specialization
sdtmigEndVersion	The last SDTMIG version that is applicable to the SDTM dataset specialization
biomedicalConceptId	Biomedical Concept identifier

SDTM Dataset Specializations

Attribute		Description
Name		Name of the variable included in the SDTM dataset specialization
dataElementConceptId		Biomedical Concept Data Element Concept identifier
isNonStandard		Flag that indicates if the variable is a non-standard variable
codelist	conceptId	C-code for a codelist in NCIt
	href	Link to NCIt for the codelist
	submissionValue	CDISC submission value for the codelist
subsetCodelist		Subset codelist short name
valueList		List of SDTM submission values used if subset codelist is not applicable
assignedTerm	conceptId	C-code for assigned term in NCIt
	value	Submission value for assigned term in NCIt if it exists, or an assigned value which will be the default value
role		SDTM variable role

SDTM Dataset Specializations

Attribute		Description
relationship	Subject	Subject in a variable relationship
	linkingPhrase	Variable relationship descriptive linking phrase
	predicateTerm	Short variable relationship linking phrase for programming
	object	Object in a variable relationship
datatype		Variable data type
length		Variable length
format		Variable display format
significantDigits		Variable significant digits
mandatoryVariable		Indicator that variable must be present within the SDTM group
mandatoryValue		Indicator that variable must be populated within the SDTM group
originType		Variable origin type (Assigned, Collected, Derived, Protocol, Predecessor)
originSource		Variable origin source (Investigator, Sponsor, Subject, Vendor)
comparator		Comparison operator for SDTM group variables included in VLM (EQ, IN)
vImTarget		Target variable for VLM (true/false)

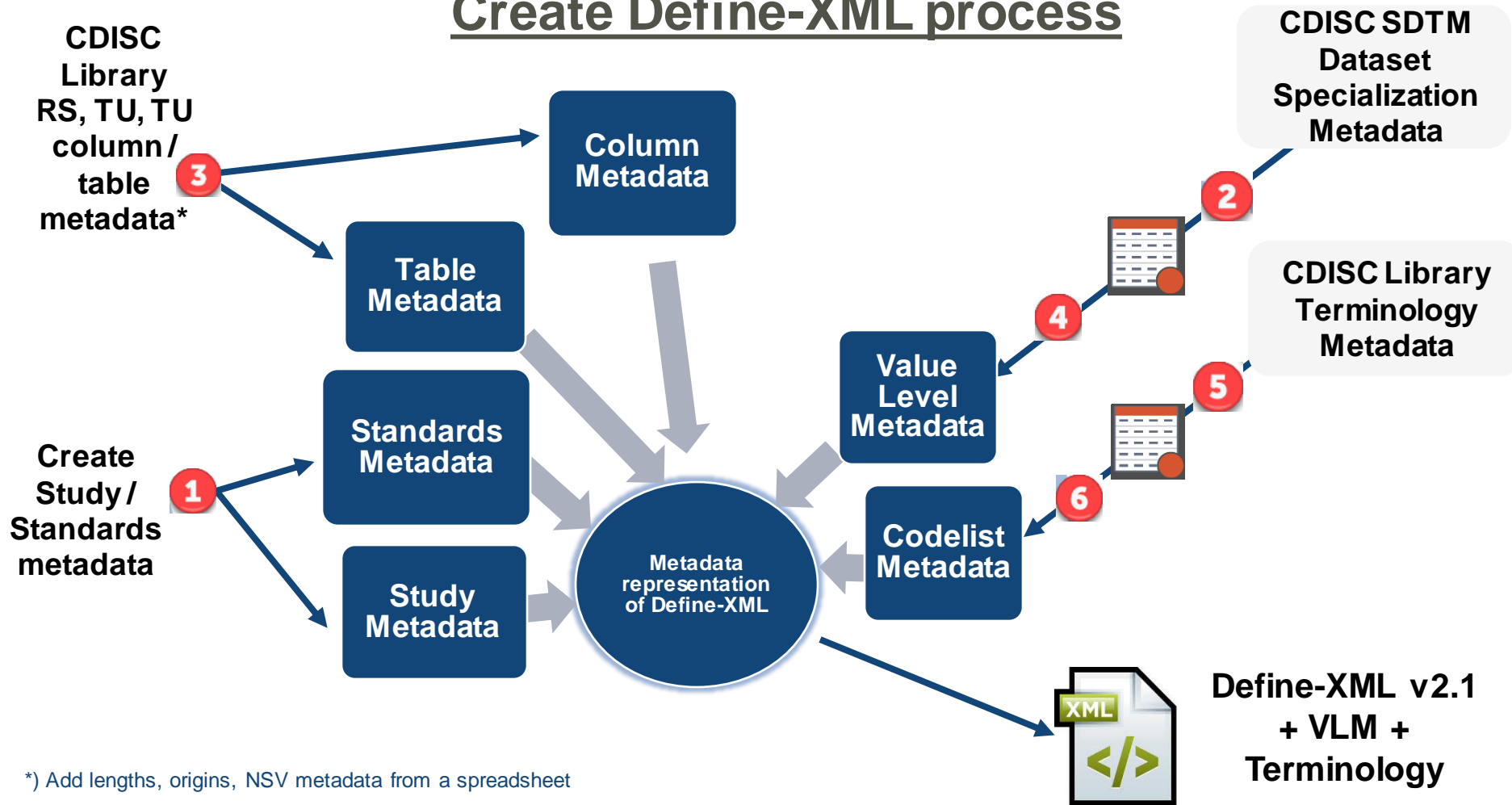
Define-XML v2.1 document with SDTM Dataset Specialization:

- Value Level Metadata and

- Controlled Terminology metadata for the RS, TR, and TU domains

- SDTM Dataset Specializations are considered pre-configured building blocks, from which end-users can select and configure to build Define-XML Value Level Metadata
- Exercise: present Oncology RECIST 1.1 SDTM Dataset Specializations as Value Level Metadata in Define-XML v2.1
- Oncology RECIST 1.1
 - 13 Biomedical Concepts
 - 13 SDTM Specializations (RS, TR, TU))
- REST API:
 - GET Biomedical Concepts:
`/mdr/bc/biomedicalconcepts?category=RECIST 1.1`
 - GET SDTM Specializations:
`/mdr/specializations/sdtm/datasetspecializations?domain=RS`

Create Define-XML process



*) Add lengths, origins, NSV metadata from a spreadsheet

Define-XML v2.1 document with SDTM Dataset Specialization:

- Value Level Metadata and
- Controlled Terminology metadata for the RS, TR, and TU domains

CDISC01

Standards

▼ Datasets

RS (Disease Response and Clin Class)

TR (Tumor/Lesion Results)

TU (Tumor/Lesion Identification)

► Controlled Terminology

Expand all VLM

Collapse all VLM

Date/Time of Define-XML document generation: 2023-10-11T14:27:04-04:00

Define-XML version: 2.1.6

Define-XML Context: Submission

Stylesheet version: 2019-02-11

Study Name	CDISC01
Study Description	CDISC Test Study
Protocol Name	CDISC01
Metadata Name	Study CDISC01_1, Data Definitions V-1
Metadata Description	Data Definitions for CDISC01-01 SDTM datasets

This Define-XML document is based on RS, TR and TU dataset and column metadata extracted from the CDISC Library. Value level metadata (VLM) and codelists were programmatically created by extracting metadata from CDISC SDTM Dataset Specializations and the CDISC Library.

Standards for Study CDISC01

Standard	Type	Status	Documentation
SDTMIG 3.3	IG	Final	
CDISC/NCI SDTM 2023-09-29	CT	Final	
CDISC/NCI DEFINE-XML 2023-06-30	CT	Final	

Datasets

Dataset	Description	Class	Structure	Purpose	Keys	Documentation	Location
RS [SDTMIG 3.3]	Disease Response and Clin Classification	FINDINGS	One record per response assessment or clinical classification assessment per time point per visit per subject per assessor per medical evaluator	Tabulation	STUDYID, RSDTC, USUBJID, RSTESTCD, RSNAM, RSEVAL, RSEVALID, RSRGPID, VISITNUM		rs.xpt 🔗
TR [SDTMIG 3.3]	Tumor/Lesion Results	FINDINGS	One record per tumor measurement/assessment per visit per subject per assessor	Tabulation	STUDYID, VISITNUM, TRDTC, USUBJID, TRTESTCD, TRMETHOD, TRNAM, TREVAL, TREVALID, TRLNKID		tr.xpt 🔗
TU [SDTMIG 3.3]	Tumor/Lesion Identification	FINDINGS	One record per identified tumor per subject per assessor	Tabulation	STUDYID, TUEVALID, TULNKID, VISITNUM, TUDTC, USUBJID, TUTESTCD, TULOC, TULAT, TUMETHOD, TUNAM, TUEVAL		tu.xpt 🔗

Define-XML v2.1 document with SDTM Dataset Specialization:

- Value Level Metadata and
- Controlled Terminology metadata for the RS, TR, and TU domains

CDISC01

Standards

▼ Datasets

RS (Disease Response and Clin Cla

TR (Tumor/Lesion Results)

TU (Tumor/Lesion Identification)

▼ Controlled Terminology

▼ CodeLists

Directionality

Epoch, subset

Evaluator, subset

Laterality

Anatomical Location

Medical Evaluator Identifier

Method, subset

Not Done

No Yes Response, subset

No Yes Response, subset for Non

No Yes Response, subset for Targ

No Yes Response, subset for Non

No Yes Response, subset for Targ

Category of Oncology Response /

Oncology Response Assessment /

Oncology Response Assessment /

Oncology Response Assessment /

RSSCAT		Subcategory	text	Grouping Qualifier	200		Collected (Source: Investigator)
RSORRES VLM		Result or Finding in Original Units	text	Result Qualifier	200		Collected (Source: Investigator)
	EPOCH = "TREATMENT" and RSCAT = "RECIST 1.1" and RSEVAL = "INVESTIGATOR" and RSTESTCD = "NEWLPROG" (New Lesion Progression)	New Lesion Progression	text	Qualifier		Oncology Response Assessment Result, subset for New Lesion Progression - Original (Res) <ul style="list-style-type: none"> "EQUIVOCAL" = "Equivocal" "UNEQUIVOCAL" = "Unequivocal" 	Collected (Source: Investigator)
	EPOCH = "TREATMENT" and RSCAT = "RECIST 1.1" and RSEVAL = "INVESTIGATOR" and RSTESTCD = "NTRGRES" (Non-target Response)	Non-Target Response	text	Qualifier		Oncology Response Assessment Result, subset for Non-Target Response - Original (Res) <ul style="list-style-type: none"> "CR" = "Complete Remission" "NE" = "Unevaluable" "NON-CR/NON-PD" = "Non Complete Response/Non Progressive Disease" "PD" = "Progressive Disease" 	Collected (Source: Investigator)
	EPOCH = "TREATMENT" and RSCAT = "RECIST 1.1" and RSEVAL = "INVESTIGATOR" and RSTESTCD = "OVLRESP" (Overall Response)	Overall Response	text	Qualifier		Oncology Response Assessment Result, subset for Overall Response - Original (Res) [7 Terms]	Collected (Source: Investigator)
	EPOCH = "TREATMENT" and RSCAT = "RECIST 1.1" and	Target Response	text	Qualifier		Oncology Response Assessment Result, subset for Target Response - Original (Res)	Collected (Source: Investigator)

Define-XML v2.1 document with SDTM Dataset Specialization:

- Value Level Metadata and
- Controlled Terminology metadata for the RS, TR, and TU domains

CDISC01

- Standards
- ▼ Datasets
 - RS (Disease Response and Clinical Response)
 - TR (Tumor/Lesion Results)
 - TU (Tumor/Lesion Identification)
- ▼ Controlled Terminology
 - ▼ CodeLists
 - Directionality
 - Epoch, subset
 - Evaluator, subset
 - Laterality
 - Anatomical Location
 - Medical Evaluator Identifier
 - Method, subset
 - Not Done
 - No Yes Response, subset
 - No Yes Response, subset for
 - No Yes Response, subset for
 - No Yes Response, subset for
 - No Yes Response, subset for
 - Category of Oncology Response
 - Oncology Response Assessment
 - Oncology Response Assessment
 - Oncology Response Assessment
 - Oncology Response Assessment
 - Oncology Response Assessment

TRSTRESC VLM		Character Result/Finding in Std Format	text	Result Qualifier	200	Tumor or Lesion Properties Test Result [22 Terms]	Derived (Source: Sponsor)
	<p>EPOCH IN ("SCREENING", "TREATMENT") and TREVAL IN ("ADJUDICATOR", "INDEPENDENT ASSESSOR", "INVESTIGATOR") and TRMETHOD IN ("CALIPER MEASUREMENT METHOD", "CT SCAN", "ENDOSCOPY", "LYMPHANGIOGRAPHY", "MAMMOGRAPHY", "MRI", "NUCLEAR RADIOLOGY", "PET SCAN", "PET/CT SCAN", "PET/MRI SCAN", "PHOTOGRAPHY", "SCINTIGRAPHY", "TOTAL BODY RADIOGRAPHY", "ULTRASOUND", "X-RAY") and TRTESTCD = "LNSTATE" (Lymph Node State)</p>	Lymph Node State	text	Qualifier		<p>Tumor or Lesion Properties Test Result, subset for Lymph Node State - Standardized (Char Res)</p> <ul style="list-style-type: none"> "NON-PATHOLOGICAL" "PATHOLOGICAL" 	Derived (Source: Sponsor)

Define-XML v2.1 document with SDTM Dataset Specialization:

- Value Level Metadata and

- Controlled Terminology metadata for the RS, TR, and TU domains

CDISC01

Standards

▼ Datasets

RS (Disease Response and Clin

TR (Tumor/Lesion Results)

TU (Tumor/Lesion Identifier

▼ Controlled Terminology

▼ CodeLists

Directionality

Epoch, subset

Evaluator, subset

Laterality

Anatomical Location

Medical Evaluator Identifier

Method, subset

Not Done

No Yes Response, subset

No Yes Response, subset for

No Yes Response, subset for

No Yes Response, subset for

No Yes Response, subset for

Category of Oncology Respor

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

Oncology Response Assessm

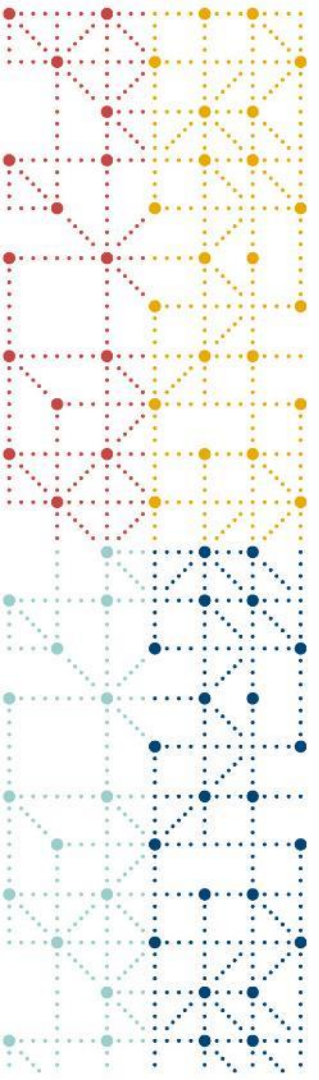
Oncology Response Assessm

TUSTRESC VLM		Tumor/Lesion ID Result Std. Format	text	Result Qualifier	200	Tumor or Lesion Identification Test Results [28 Terms]	Derived (Source: Sponsor)
	EPOCH = "SCREENING" and TUEVAL = "INVESTIGATOR" and TUTESTCD = "NTIND" (Non-Target Indicator)	Non-Target Indicator	text	Qualifier	24	No Yes Response, subset for Non-Target Indicator - Standardized (Char Res) • "N" = "No" • "U" = "Unknown" • "Y" = "Yes"	Derived (Source: Sponsor)
	EPOCH = "SCREENING" and TUEVAL = "INVESTIGATOR" and TUTESTCD = "TIND" (Target Indicator)	Target Indicator	text	Qualifier	24	No Yes Response, subset for Target Indicator - Standardized (Char Res) • "N" = "No" • "U" = "Unknown" • "Y" = "Yes"	Derived (Source: Sponsor)
	EPOCH = "TREATMENT" and TUEVAL IN ("ADJUDICATOR", "INDEPENDENT ASSESSOR", "INVESTIGATOR") and TUMETHOD IN ("CALIPER MEASUREMENT METHOD", "CT SCAN", "ENDOSCOPY", "LYMPHANGIOGRAPHY", "MAMMOGRAPHY")	Tumor Merged	text	Qualifier	24	Tumor or Lesion Identification Test Results, subset for Tumor Merged - Standardized (Char Res) • "TARGET"	Derived (Source: Sponsor)



Conclusion

- SDTM Dataset Specializations can be represented as Value Level Metadata definitions in Define-XML v2.1.
- These definitions contain detailed metadata, including Controlled Terminology subsets.
- The SDTM Dataset Specializations can be considered pre-configured building blocks, from which end-users can select and configure to build Define-XML Value Level Metadata
- SDTM dataset specializations are ready to be used as building blocks for Define-XML.
- This provides immediate benefits to SDTM programmers and opens the door to efficient programming and automation



Retrieval of Biomedical Concepts and SDTM Dataset Specializations

Using CDISC Library APIs

API Endpoints in CDISC Library

- Biomedical Concepts and SDTM Specialization are published in packages
- Packages have **new content** and **updates to existing content**
- Not cumulative!

2022-10-26

API request template for Biomedical Concepts

2023-02-13

`/mdr/bc/packages`

2023-03-31

`/mdr/bc/packages/{package}/biomedicalconcepts`

2023-07-06

`/mdr/bc/packages/{package}/biomedicalconcepts/{biomedicalconcept}`

2023-10-03

API request template for SDTM Specializations

`/mdr/specializations/sdtm/packages`

`/mdr/specializations/sdtm/packages/{package}/datasetspecializations`

`/mdr/specializations/sdtm/packages/{package}/datasetspecializations/{datasetspecialization}`

API Endpoints in CDISC Library

- Biomedical Concepts and SDTM Specialization can now also be requested through the API (**v2 only**) with all the **latest versions**

API request template for Biomedical Concepts	API v2 Only?	Return Latest Version Only?
<code>/mdr/bc/biomedicalconcepts</code>	✓	✓
<code>/mdr/bc/biomedicalconcepts/{biomedicalconcept}</code>	✓	✓
<code>/mdr/bc/categories</code>	✓	
<code>/mdr/bc/biomedicalconcepts?category={category}</code>	✓	✓

API Endpoints in CDISC Library

- Biomedical Concepts and SDTM Specialization can now also be requested through the API (**v2 only**) with all the **latest versions**

API request template for SDTM Specialization	API v2 Only?	Return Latest Version Only?
<code>/mdr/specializations/sdtm/datasetspecializations</code>	✓	✓
<code>/mdr/specializations/sdtm/datasetspecializations/{datasetspecialization}</code>	✓	✓
<code>/mdr/specializations/sdtm/domains</code>	✓	
<code>/mdr/specializations/sdtm/datasetspecializations?domain={domain}</code>	✓	✓

API request template for Specializations	API v2 Only?	Return Latest Version Only?
<code>/mdr/specializations/datasetspecializations?biomedicalconcept={biomedicalconcept}</code>	✓	✓


API Requests in SAS


```
%let ApiKey=<your_personal_api_key>;  
%let baseUrl=https://library.cdisc.org/api/cosmos/v2;
```


```
filename json_out temp;  
proc http  
  method = 'GET'  
  url="&baseUrl/mdr/specializations/sdtm/datasetsspecializations/SYSBP"  
  out=json_out;  
  headers  
    "api-key" = "&ApiKey"  
    "Accept" = "application/json";  
run;
```


```
filename json_map temp;  
libname json_out json map=json_map automap=create fileref=json_out;
```


```
proc copy in = json_out out = work;  
run;
```


 _links_parentbiomedicalconcept.sas7bdat


 _links_self.sas7bdat


 root.sas7bdat


 variables_assignedterm.sas7bdat


 variables_relationship.sas7bdat

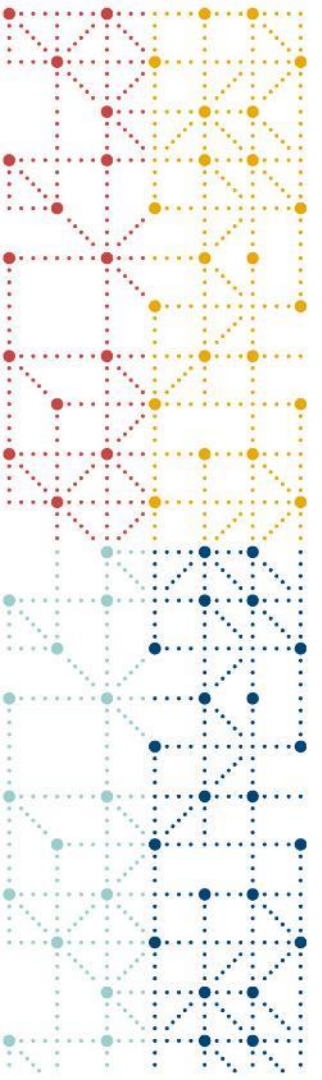
 _links_parentpackage.sas7bdat

 alldata.sas7bdat

 variables.sas7bdat

 variables_codelist.sas7bdat

 variables_valuelist.sas7bdat



Key Achievements and Next Steps

Current Status and Future Plans



CDISC Biomedical Concepts

2023 Key Achievements

- Curation of new Biomedical Concepts (BCs) and SDTM Dataset Specializations
 - 208 BCs and 171 SDTM Dataset Specializations currently available via CDISC Library APIs
 - Additional BCs and SDTM will be loaded end of Q42023
 - Streamlined curation process and documentation
- New API Endpoints to find and retrieve BCs/SDTM Dataset Specializations
- Established versioning and change management for BCs and APIs
 - Loaded multiple packages of BCs/SDTM Dataset Specializations
 - Retrieval of the latest version of BCs/SDTM across packages
 - Established lineage across versions
- Supported CDISC Oncology SDS Team to create, validate and release BCs and SDTM Dataset Specializations for RECIST 1.1 Disease Response Criteria
 - Retrievable via CDISC Library APIs
- Supported Digital Data Flow Phase 2-3 (DDF)
 - Created BCs/SDTM Dataset Specializations for example COVID and Alzheimer's protocols
 - Integration of CDISC BC model into USDM to support eDC automation

Supporting Digital Data Flow Phase 3 (DDF3)

- DDF 3 Pilot Study
 - Development of additional biomedical concepts to cover the CDISC Pilot Study (LZZT)
 - Allows for a full exemplar USDM protocol Design



**Protocol Attachment LZZT.1
Schedule of Events for Protocol H2Q-MC-LZZT(c)**

	WEEK	1	2	3	4	5	7	8
ACTIVITY	WEEK	-2	-1	0	2	4	6	8
Informed consent	X							
Patient number assigned	X							
Baseline vitals	X							
MMSSE 10-23	X							
Physical examination	X							
Medical History	X							
Habits	X							
Chest x-ray	X							
Site & packaging					X			
Patient randomized				X				
Visit signs/temperature	X	X	X	X	X		X	X
Ambulatory ECG placed		X						
Ambulatory ECG removed			X					
ECG	X			X	X		X	X
Physical TTS test	X							
CT Scan (if not within last year and patient passes all other systems)	X							
Concomitant Medications	X	X	X	X	X		X	X
Laboratory (Chem/Hemat)	X			X	X		X	X
Laboratory (Urinary)	X			X				
Plasma Specimen (Xanomeline)			X	X	X		X	
Hemoglobin A1c	X ^o							
Study drug record			X	X	X		X	X
Medications dispensed								
Medications returned								
TTS Acceptability Survey								
ADAS-Cog	P	X						X
CBIC ^o	P	X						X
DAD	P	X						X
NPI-S	P	X	X	X	X		X	X ^o
Adverse events	X	X	X	X	X		X	X

Abbreviations: CT = computed tomography, ECG = electrocardiogram
X = Performed at this visit.
X^o = Performed at this visit if patient is an insulin-dependent diabetic.
X^o = Performed at this visit and via telephone interview 2 weeks following this visit.
P = Practice only - It is recommended that a sampling of the CBIC^o, ADAS-Cog, DAD,
and NPI-S be administered at Visit 1. Data from this sampling would not be
considered as study data and would not be collected.

Supporting Digital Data Flow Phase 3 (DDF3)

Supporting Study Design and eDC automation

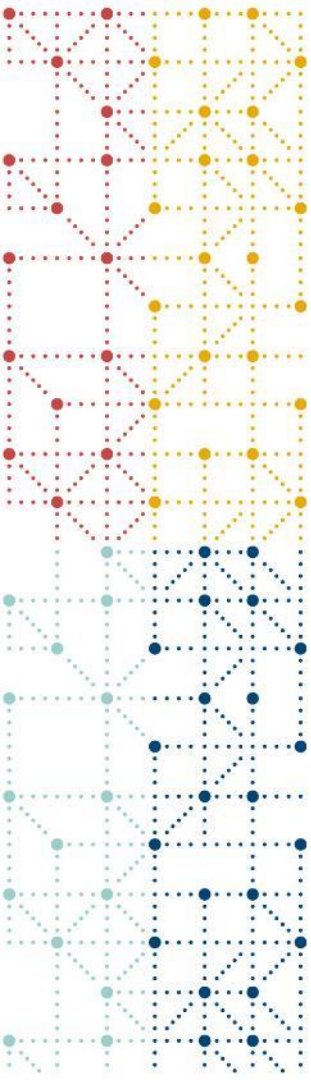
- Integration of the CDISC BC model into USDM to support EDC automation
- Building a complete set of BCs and SDTM Dataset Specializations for the Pilot Study (Alzheimer's Disease)
- Attaching CDISC BCs/SDTM Dataset Specializations to Activities which are in turn attached to timepoints in the SoA which provides:
 - Decision occurs at either protocol design or eDC stage
- Ends with a detailed study specification (next level after SoA)
- What's the big deal?
 - Precise definition of all data points
 - Increased standardization and cross-study consistency
 - Higher quality specifications



Looking to Future

Upcoming Activities and Future Plans

- CDASH dataset specializations
- FHIR dataset specializations
- Include metadata for transformational algorithms in CDISC Dataset Specializations
 - Participation and alignment with OAK project to begin building end to end data flow CDASH–SDTM, etc.
- Support new TAUGs
- Support CDISC Oncology SDS Team to create more BCs/SDTM Specializations
- Migration of BC content from CDISC community
 - Currently working with Novo-Nordisk to migrate/load a sampling of their BCs to CDISC Library
- Establish User friendly search, visualizations and exports
- Continue to improve and streamline the BC curation process
- Hoping to get more ‘BCs’ from the community



спасибо 谢谢
GRACIAS 谢谢
THANK YOU
ありがとうございました MERCI
DANKE धन्यवाद
شُكْرًا **OBRIGADO**