CDISC Biomedical Concepts

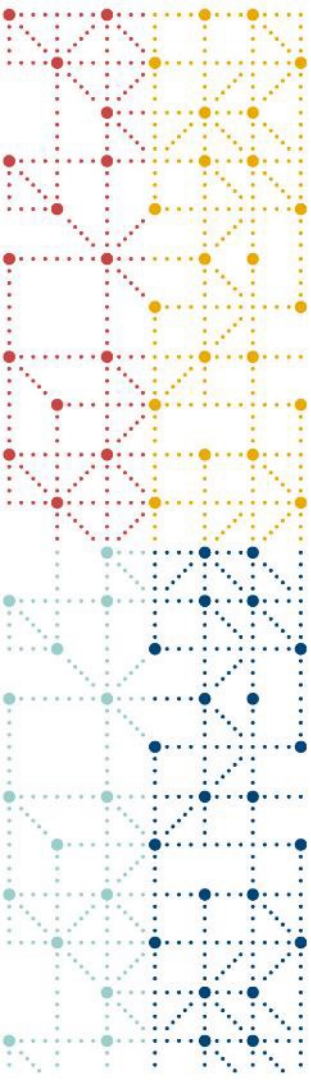
Public Review Webinar

Bess LeRoy, Jon Neville, Linda Lander

2024-05-16

Agenda

1. Introductions and Background
2. Overview of CDISC Biomedical Concepts
3. Demo of Export Files
4. Reviewer Hints and Tips
5. Guidance on JIRA Feedback



CDISC Biomedical Concepts

Introduction and Background

Background

Current Problem with Standards Adoption:

- Variation across studies
- Poor quality data from various sources, e.g., venders, CROs, etc.
- Lengthy cycle times to clean and standardize across studies in submissions
- Costly manual efforts
- Insufficient linkages across standards (end-to-end)
- Demand for more standards
- Data re-use challenging

CDISC has evolved:

- CDISC 360 piloted development of linked **Biomedical Concept** metadata to enable end to end automation
- CDISC Library has published data standards as groups of linked metadata

Biomedical Concepts - Benefits to the CDISC Community

Part of the overarching CDISC vision enabling **connected standards**

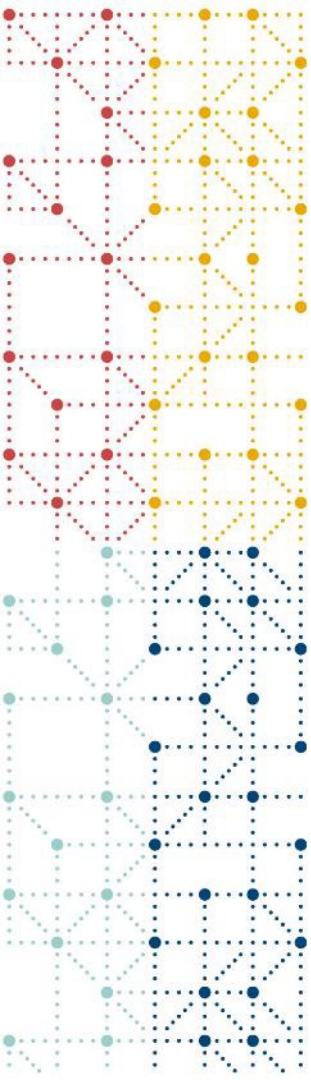
Facilitates **accurate** and **more consistent implementation** by reducing unnecessary variability

Facilitates **metadata-driven automation**

Increases **quality and efficiency** throughout end-to-end study delivery process

Enables **data reuse**

cdisc

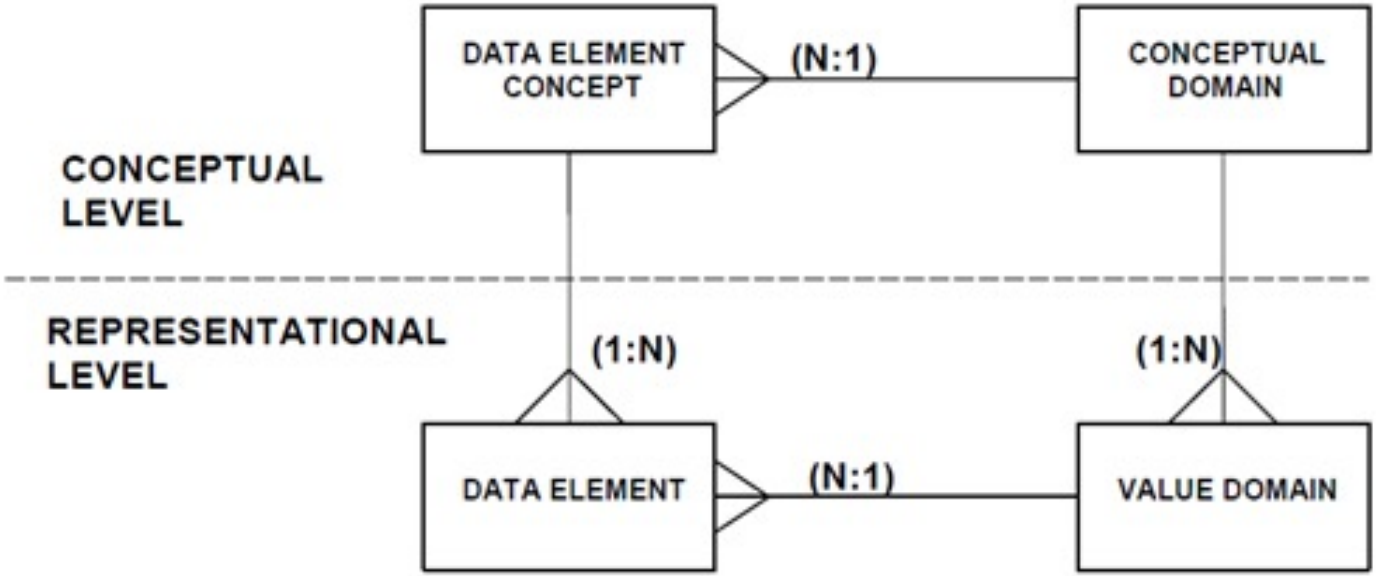


CDISC Biomedical Concepts

Overview

What Is a Biomedical Concept?

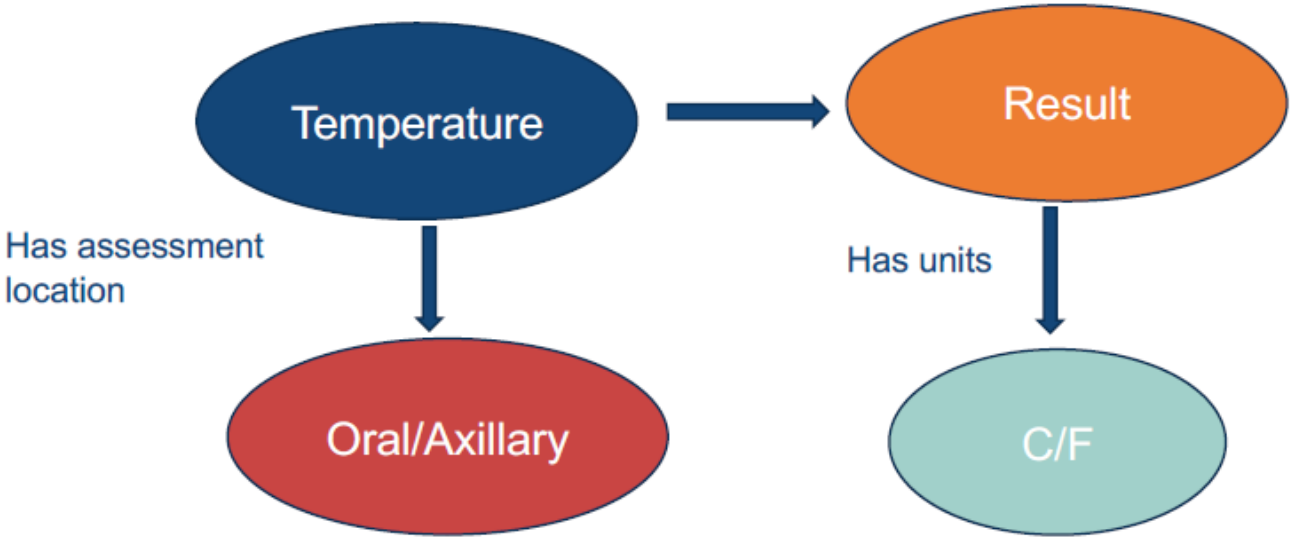
ISO 11179 Definition: A unit of knowledge created by a unique combination of characteristics



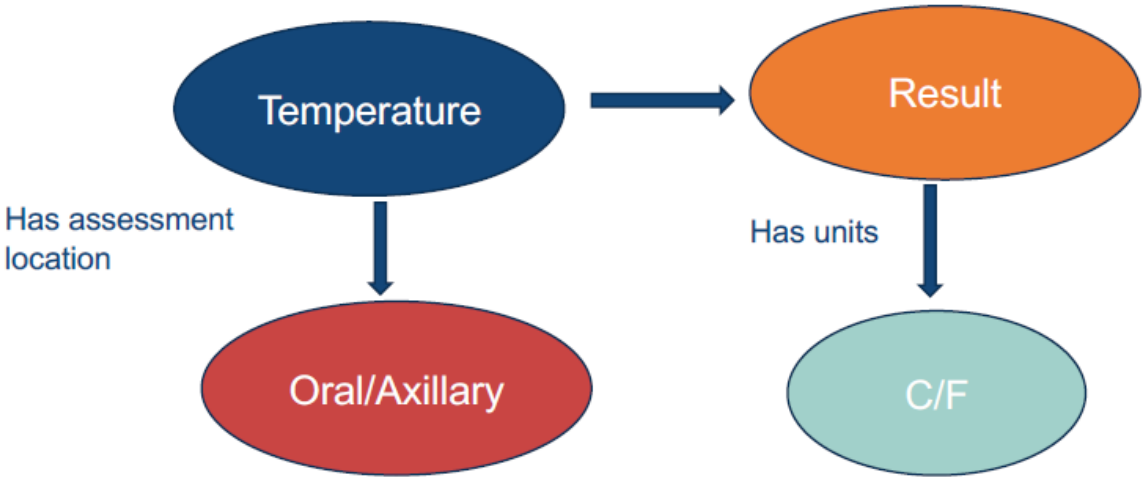
What Is a Biomedical Concept?

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?



Conceptual Layer

VSTEST	VSTESTCD	VSORRES	VSUNIT	VSLOC
Temperature	TEMP	101.3	F	ORAL

Implementation Layer

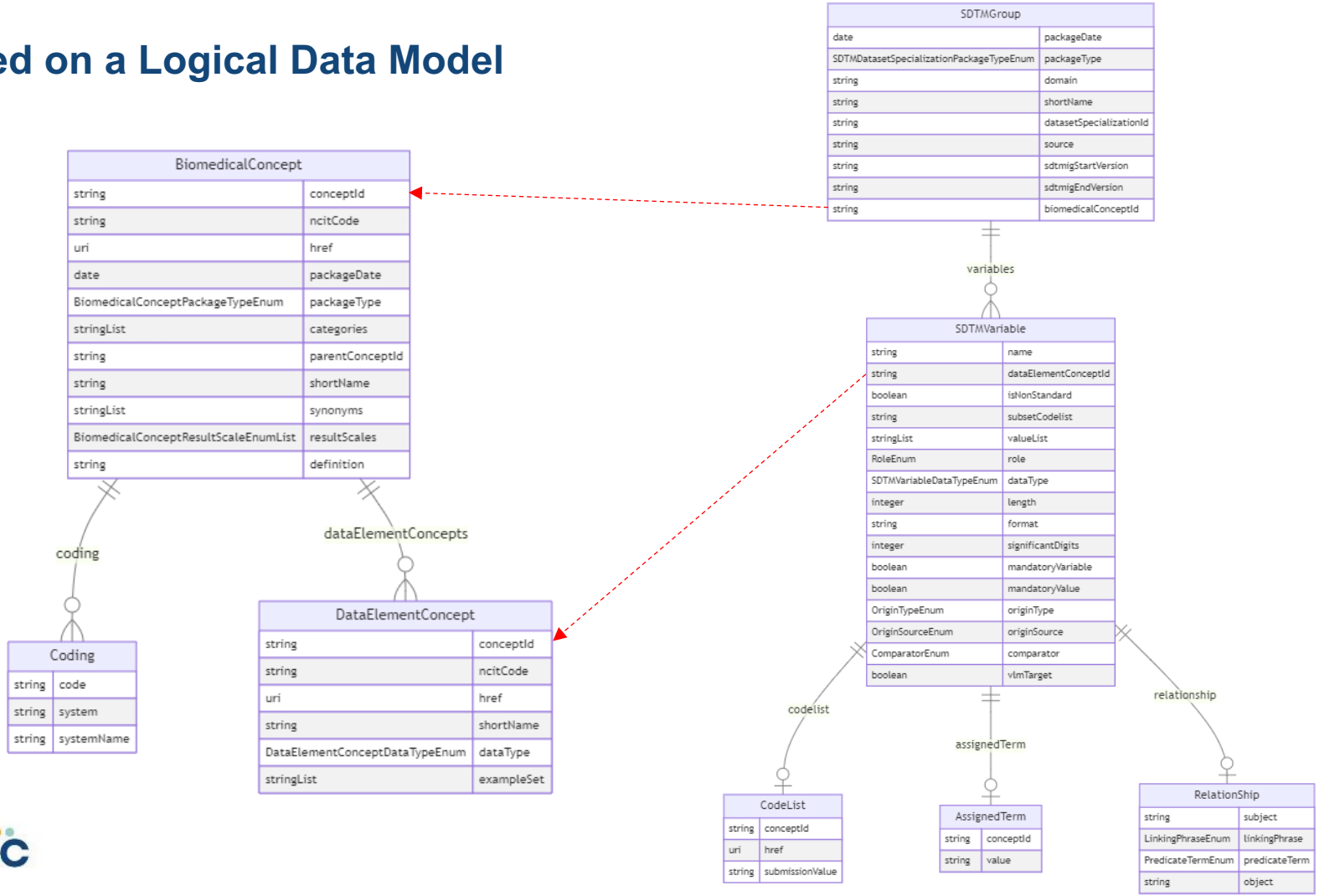
CDISC Biomedical Concepts & SDTM Dataset Specializations

Pragmatic Implementation:

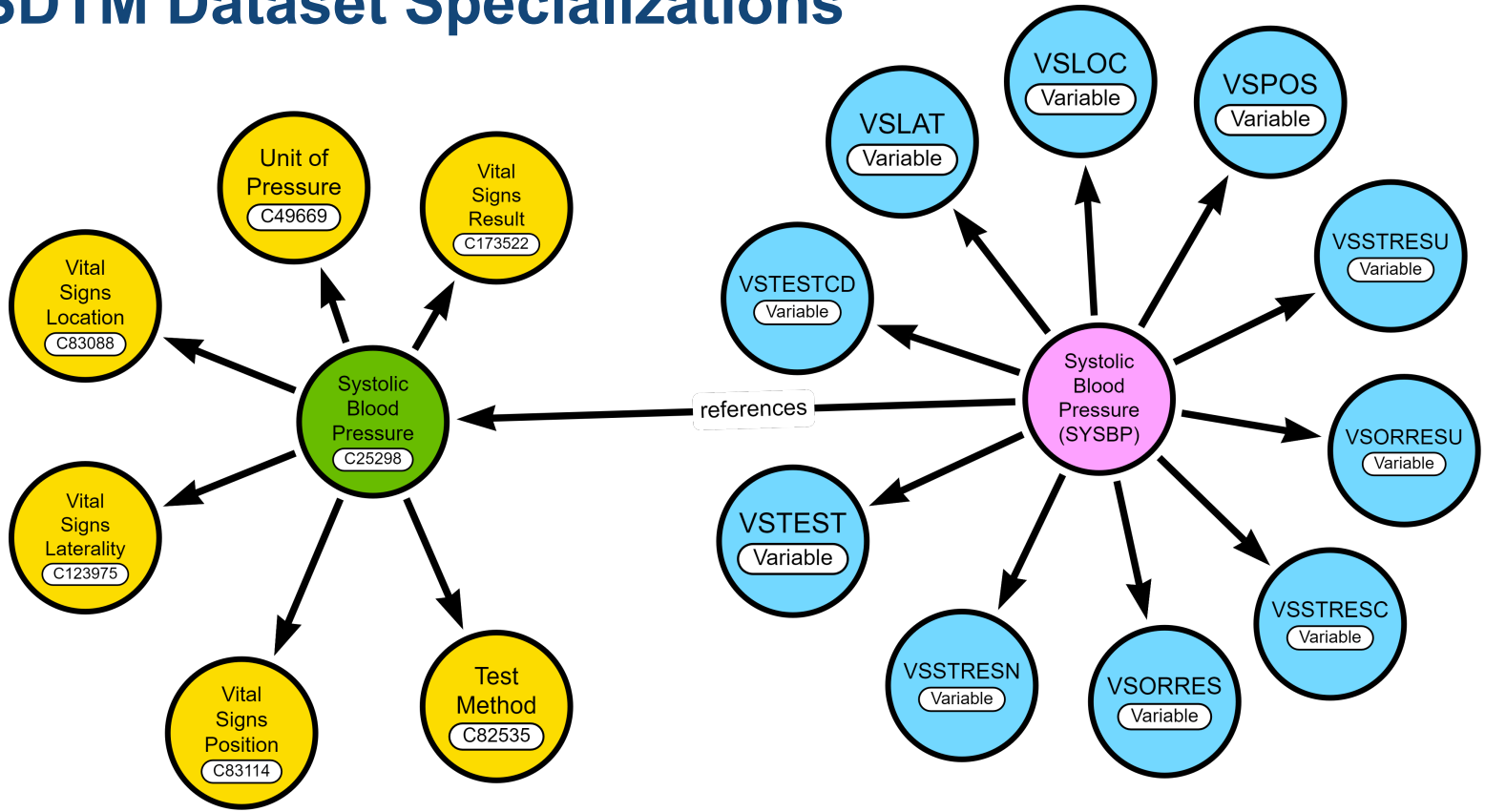
- **Conceptual Layer** – abstract BC's
  - Provides semantics - aligned with NCI terminology
  - Supports study design, Schedule of Activities (SOA)

- **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
  - Extend foundational standards
    - Add explicit relationships between variables
    - Additional operational metadata, e.g., data type, etc.

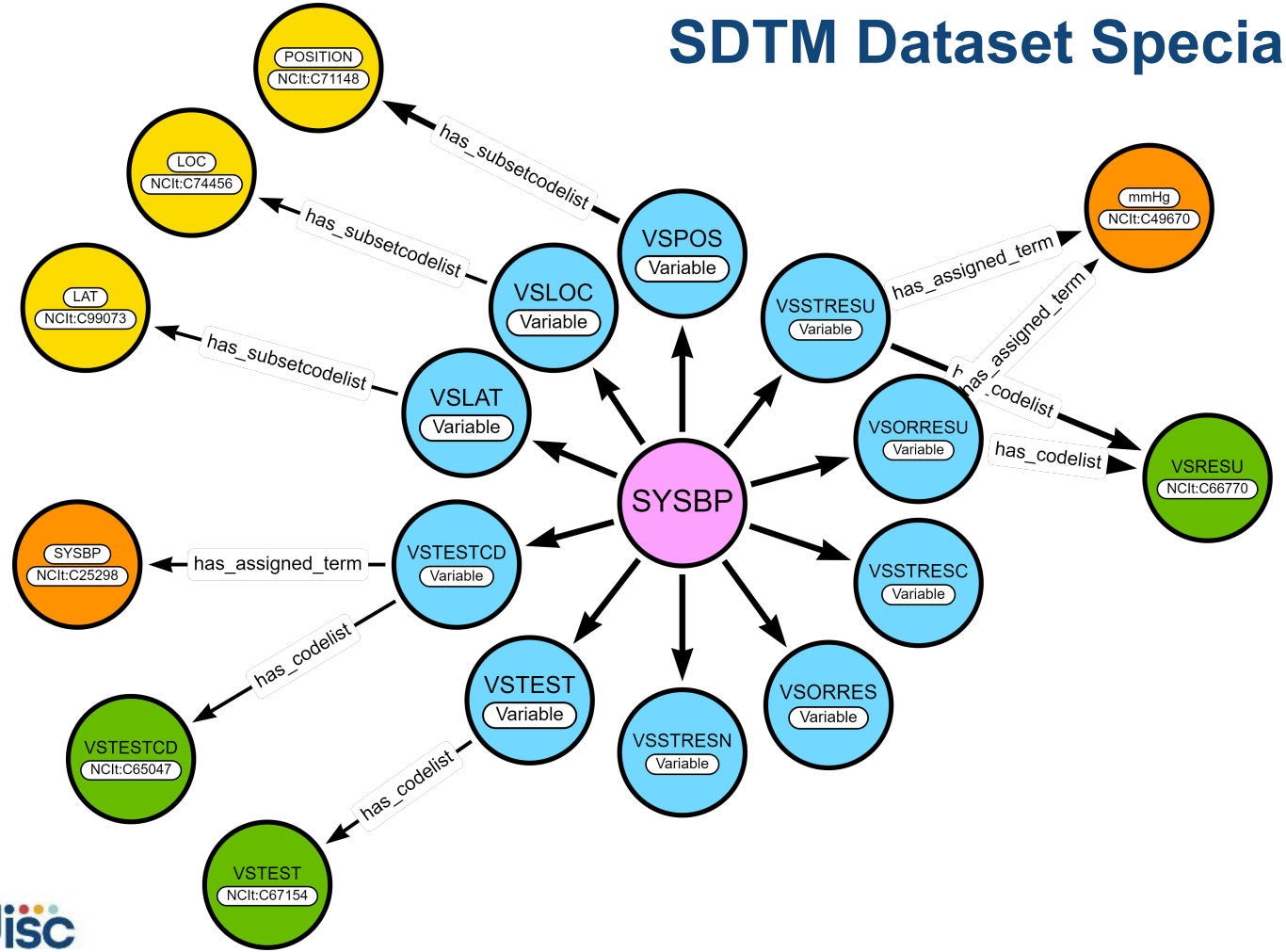
Based on a Logical Data Model

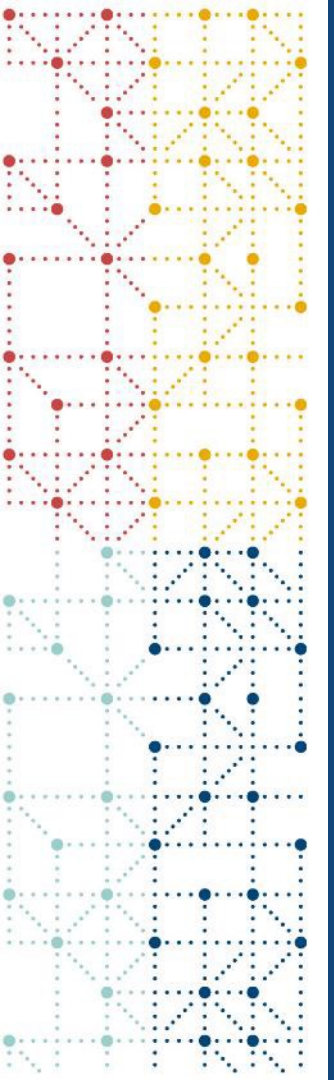


CDISC Biomedical Concepts and SDTM Dataset Specializations



SDTM Dataset Specializations





CDISC Library APIs and Export Files

Retrieval of BCs and SDTM Dataset Specializations

- **export**

  The export folder contains spreadsheets with the latest versions of all BCs and SDTM Dataset Specializations in the CDISC Library:

  - Latest CDISC Biomedical Concepts (Excel spreadsheet)
  - Latest SDTM Dataset Specializations (Excel spreadsheet)

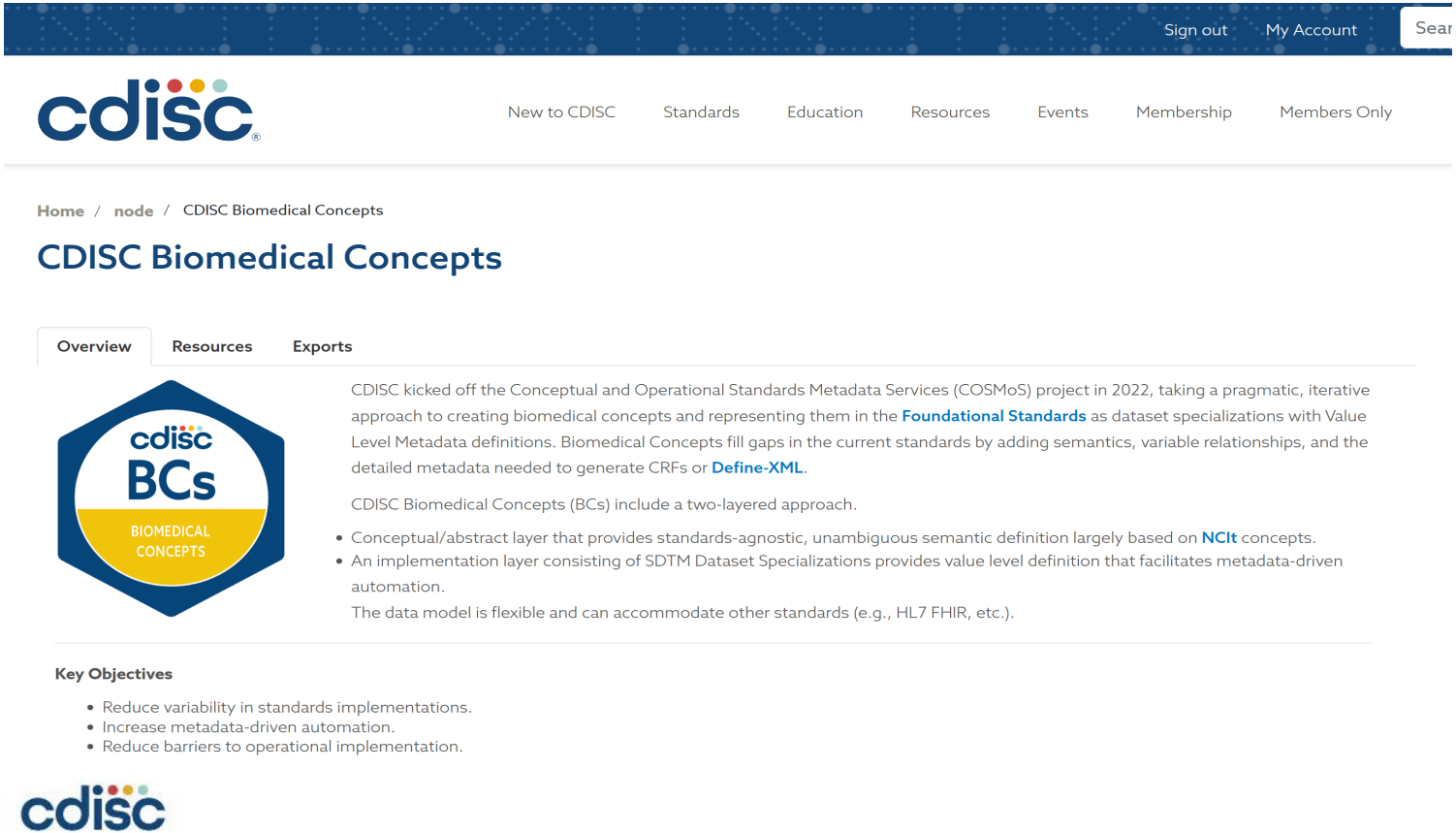
- **openapi**

  The openapi folder contains the OpenAPI definition of the COSMoS API (cosmos.yaml)

  More information about using the CDISC Library API can be found on the CDISC Library API Portal and the Biomedical Concepts and Dataset Specializations Release Notes on the CDISC Wiki. Use your cdiscID credentials to login to the CDISC Library API Portal.

CDISC Biomedical Concepts Website

CDISC Biomedical Concepts | CDISC



API Endpoints in CDISC Library

- Biomedical Concepts and SDTM Specialization can 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?
/mdr/bc/biomedicalconcepts	✓	✓
/mdr/bc/biomedicalconcepts/{biomedicalconcept}	✓	✓
/mdr/bc/categories	✓	
/mdr/bc/biomedicalconcepts?category={category}	✓	✓

API Endpoints in CDISC Library

- Biomedical Concepts and SDTM Specialization can 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?
/mdr/specializations/sdtm/datasetspecializations	✓	✓
/mdr/specializations/sdtm/datasetspecializations/{datasetspecialization}	✓	✓
/mdr/specializations/sdtm/domains	✓	
/mdr/specializations/sdtm/datasetspecializations?domain={domain}	✓	✓

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

Demo of Excel Export Files

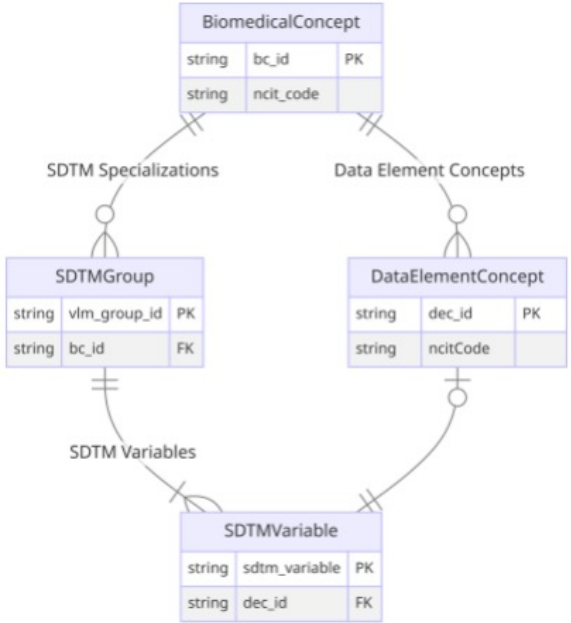
Biomedical Concepts

SDTM Dataset Specializations

Export Files: CDISC Biomedical Concepts

This spreadsheet contains the latest versions of CDISC Biomedical Concepts in the CDISC Library as of 2024-04-02.
There are currently 302 unique CDISC Biomedical Concepts in the CDISC Library.
The image on the right shows the relation between Biomedical Concepts and SDTM Dataset Specializations.
Only a few attributes are shown in the image.

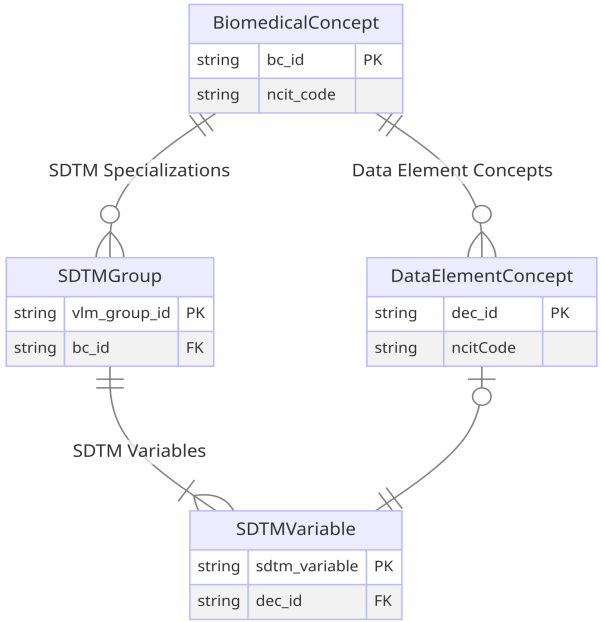
Group	Column	Description	Class
Biomedical Concept	package_date	Biomedical Concept package release date indicating when the BC package was published to production	BiomedicalConcept
	short_name	NCI Preferred Name for the concept; provisional name will be used if concept is not available in NCIt	
	bc_id	A unique identifier for a Biomedical Concept which will be assigned as the NCIt code if it exists or a placeholder identifier if the concept is not yet available in NCIt	
	ncit_code	NCIt C-code for the Biomedical Concept	
	parent_bc_id	C-code for the parent concept in the NCIt hiearchy; blank if concept is not available in NCIt	
	bc_categories	Biomedical Concept category for the faciliation of API search and extract	
	synonyms	Biomedical Concept synonym equivalent to BC short name for the facilitation of API search and extraction	
	result_scales	Scale of measurement for the Biomedical Concept result	
	definition	NCIt definition for the Biomedical Concept; provisional defintion if concept is not available in NCIt	
	system	Identifies the code system for the synonym concept. The URL of the code system should be used if it exists	Coding
	system_name	Human-readable name for the code system	
	code	Synonym concept for the Biomedical Concept as defined in a code system	
Data Element Concept (DEC)	dec_id	An identifier for a Data Element Concept (DEC) which will be assigned as the NCIt code if it exists or a placeholder identifier if the concept is not yet available in NCIt	DataElementConcept
	ncit_dec_code	NCI C-code for the BC Data Element Concept	
	dec_label	NCI Preferred Name for the concept; provisional name will be used if concept is not available in NCIt	
	data_type	Data Type for the Data Element Concept	
	example_set	Example values for the Data Element Concept	



Export Files: SDTM Dataset Specializations

This spreadsheet contains the latest versions of CDISC SDTM Dataset Specializations in the CDISC Library as of 2024-04-02.
There are currently 290 unique CDISC SDTM Dataset Specializations in the CDISC Library.
The image on the right shows the relation between Biomedical Concepts and SDTM Dataset Specializations.
Only a few attributes are shown in the image.

Group	Column	Description	Class
SDTM Group	package_date	Biomedical Concept package release date indicating when the BC package was published to production	SDTMGroup
	bc_id	Biomedical Concept identifier foreign key	
	sdtmig_start_version	The earliest SDTMIG version applicable to the BC dataset specialization	
	sdtmig_end_version	The last SDTMIG version that is applicable to the BC dataset specialization	
	domain	Domain for the SDTM specialization group	
	vlm_source	SDTM VLM Source which categorizes VLM groups by topic variable	
	vlm_group_id	Identifier for SDTM Value Level Metadata group	
	short_name	SDTM group short name which provides a user friendly and intuitive name for the vlm_group_id	
SDTM Variable	sdtm_variable	Variable included in the SDTM dataset specialization	SDTMVariable
	dec_id	Biomedical Concept Data Element Concept idenfifier foreign key	
	nsv_flag	Flag that indicates if the variable is a non-standard variable	
	codelist	C-code for a codelist in NCIt	CodeList
	codelist_submission_value	CDISC submission value for the codelist	
	subset_codelist	Subset codelist short name	SDTMVariable
	value_list	List of SDTM submission values used if subset codelist is not applicable	
	assigned_term	C-code for assigned term in NCIt or left blank when CDISC terminology does not apply	AssignedTerm
	assigned_value	Submission value for assigned term in NCIt if it exists, or an assigned value which will be the default value	



Retrieval of BCs and SDTM Dataset Specializations

bc_id	_sta	_enc	doma	vlm_source	vlm_group_i	short_name	sdtm_variabl	dec_id	codelis	codelist_submission_valu	value_lis	assigned_tern	assigned_value
C161483	3-2		TU	TU.TUTESTCD	NTIND	Non-Target Indicator	TUTESTCD		C96784	TUTESTCD		C161483	NTIND
C161483	3-2		TU	TU.TUTESTCD	NTIND	Non-Target Indicator	TUTEST		C96783	TUTEST		C161483	Non-Target Indicator
C161483	3-2		TU	TU.TUTESTCD	NTIND	Non-Target Indicator	TUORRES	C117221	C66742	NY	Y;N;U		
C161483	3-2		TU	TU.TUTESTCD	NTIND	Non-Target Indicator	TUSTRESC	C117222	C66742	NY	Y;N;U		
C161483	3-2		TU	TU.TUTESTCD	NTIND	Non-Target Indicator	TUEVAL	C51824	C78735	EVAL		C25936	INVESTIGATOR
C161483	3-2		TU	TU.TUTESTCD	NTIND	Non-Target Indicator	EPOCH	C71738	C99079	EPOCH		C48262	SCREENING
C178053	3-2		TU	TU.TUTESTCD	TIND	Target Indicator	TUTESTCD		C96784	TUTESTCD		C178053	TIND
C178053	3-2		TU	TU.TUTESTCD	TIND	Target Indicator	TUTEST		C96783	TUTEST		C178053	Target Indicator
C178053	3-2		TU	TU.TUTESTCD	TIND	Target Indicator	TUORRES	C117221	C66742	NY	Y;N;U		
C178053	3-2		TU	TU.TUTESTCD	TIND	Target Indicator	TUSTRESC	C117222	C66742	NY	Y;N;U		
C178053	3-2		TU	TU.TUTESTCD	TIND	Target Indicator	TUEVAL	C51824	C78735	EVAL		C25936	INVESTIGATOR
C178053	3-2		TU	TU.TUTESTCD	TIND	Target Indicator	EPOCH	C71738	C99079	EPOCH		C48262	SCREENING
C94523	3-2		TU	TU.TUTESTCD	TUMERGE	Tumor Merged	TUTESTCD		C96784	TUTESTCD		C94525	TUMERGE
C94523	3-2		TU	TU.TUTESTCD	TUMERGE	Tumor Merged	TUTEST		C96783	TUTEST		C94525	Tumor Merged
C94523	3-2		TU	TU.TUTESTCD	TUMERGE	Tumor Merged	TUORRES	C117221	C123650	TUIDRS		C94520	TARGET

Reviewer Hints and Tips

Reviewer Guidance

CDISC BC Model Includes 2 Separate Layers

- **SDTM Dataset Specializations** are 'value level instances' of a parent SDTM dataset, e.g., For Findings classification content, the specialization is centered around the –TESTCD.
- **Biomedical Concepts** provide the standards agnostic, semantic definition and are based on NCI Thesaurus. Keep in mind that there can be multiple SDTM Dataset Specializations associated with a single BC (refer BC Principles and Guidelines).

Note: SDTM Dataset Specializations include a link (bc_id) to the Biomedical Concepts

Reviewer Guidance

- Familiarize yourself with the documentation in *CDISC Biomedical Concepts (BC) Starter Package*.

- Keep *BC Curation Principles and Completion Guidelines* handy.
  - These describe the columns in the BC and SDTM Dataset Specialization export files and provide insights into how they should be populated.

- Total of 302 BCs and 290 SDTM Dataset Specializations in public review.
  - Recommended that you begin by selecting a specific SDTM domain within the SDTM Dataset Specialization export file.
  - After reviewing the SDTM content, then trace back to an associated BC.
  - The link between the SDTM Dataset Specializations and BCs is the 'bc_id'. This is a unique NCIt code or in rare cases, a placeholder.

CDISC Website

Home / CDISC Biomedical Concepts

CDISC Biomedical Concepts

Overview | Resources | Exports

CDISC kicked off the Conceptual and Operational Standards Metadata Services (COSMoS) project in 2022, taking a pragmatic, iterative approach to creating biomedical concepts and representing them in the **Foundational Standards** as dataset specializations with Value Level Metadata definitions. Biomedical Concepts fill gaps in the current standards by adding semantics, variable relationships, and the detailed metadata needed to generate CRFs or **Define-XML**.

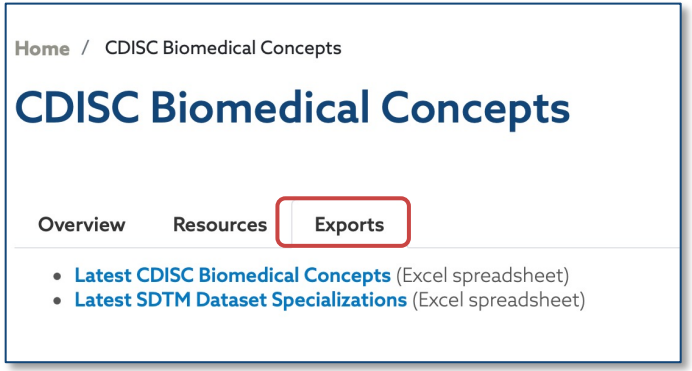
CDISC Biomedical Concepts (BCs) include a two-layered approach.

- Conceptual/abstract layer that provides standards-agnostic, unambiguous semantic definition largely based on **NCIt** concepts.
- An implementation layer consisting of SDTM Dataset Specializations provides value level definition that facilitates metadata-driven automation.

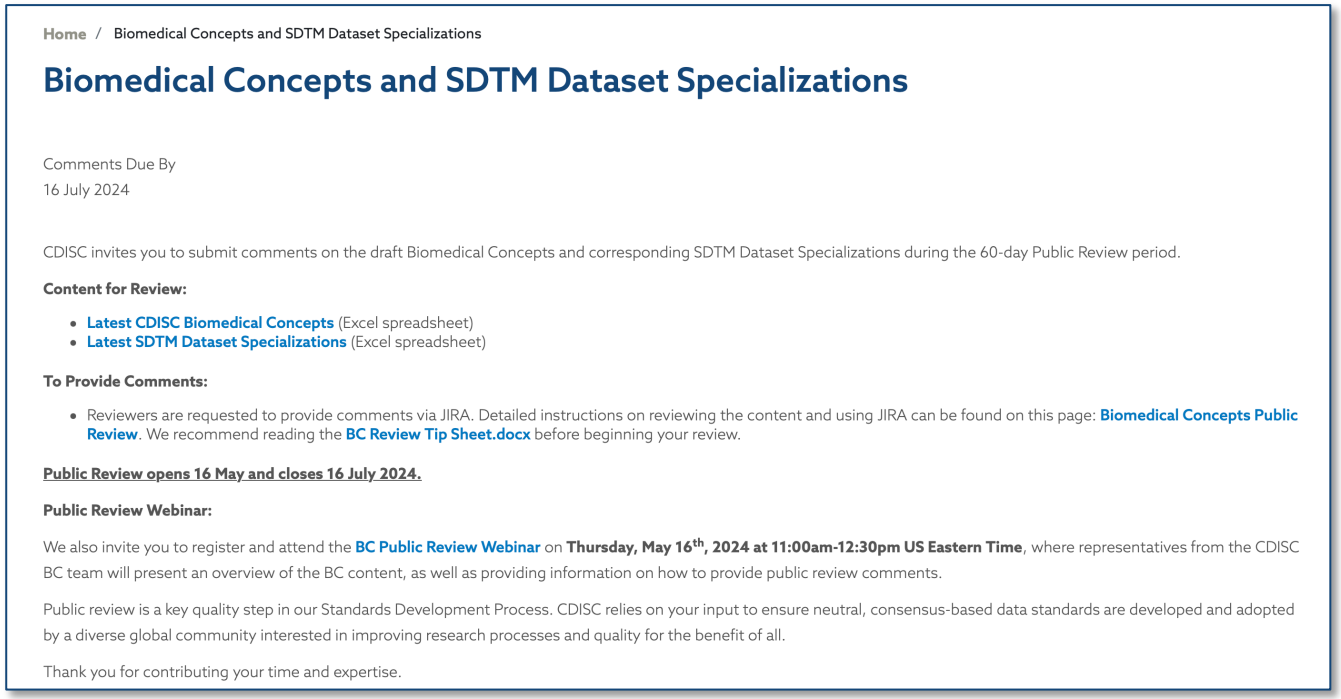
The data model is flexible and can accommodate other standards (e.g., HL7 FHIR, etc.).

https://www.cdisc.org/cdisc-biomedical-concepts

Resources and Exports



Public Reviews



Home / Biomedical Concepts and SDTM Dataset Specializations

Biomedical Concepts and SDTM Dataset Specializations

Comments Due By
16 July 2024

CDISC invites you to submit comments on the draft Biomedical Concepts and corresponding SDTM Dataset Specializations during the 60-day Public Review period.

Content for Review:

- **Latest CDISC Biomedical Concepts** (Excel spreadsheet)
- **Latest SDTM Dataset Specializations** (Excel spreadsheet)

To Provide Comments:

- Reviewers are requested to provide comments via JIRA. Detailed instructions on reviewing the content and using JIRA can be found on this page: **Biomedical Concepts Public Review**. We recommend reading the **BC Review Tip Sheet.docx** before beginning your review.

Public Review opens 16 May and closes 16 July 2024.

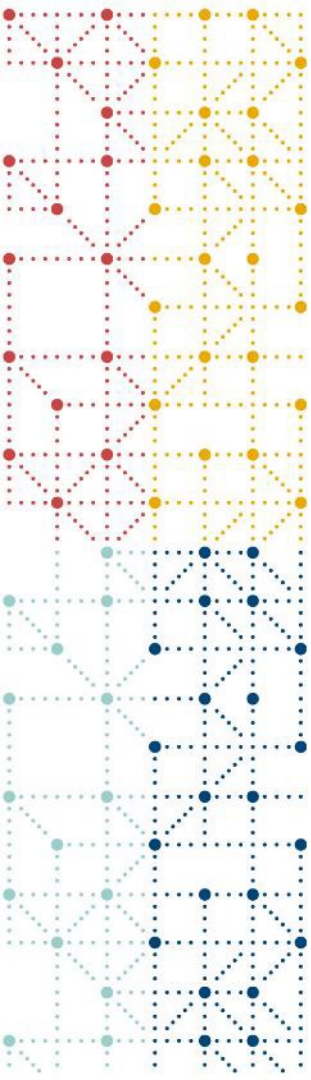
Public Review Webinar:

We also invite you to register and attend the **BC Public Review Webinar** on **Thursday, May 16th, 2024 at 11:00am-12:30pm US Eastern Time**, where representatives from the CDISC BC team will present an overview of the BC content, as well as providing information on how to provide public review comments.

Public review is a key quality step in our Standards Development Process. CDISC relies on your input to ensure neutral, consensus-based data standards are developed and adopted by a diverse global community interested in improving research processes and quality for the benefit of all.

Thank you for contributing your time and expertise.

https://www.cdisc.org/public-review/biomedical-concepts-and-sdtm-dataset-specializations



JIRA Feedback

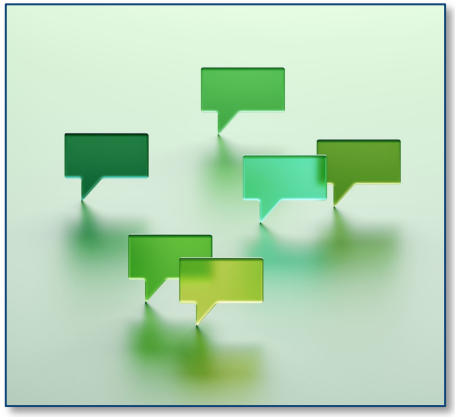
Reviewer Guidance

Public Review Timeline: 5/16/2024 through 7/16/2024

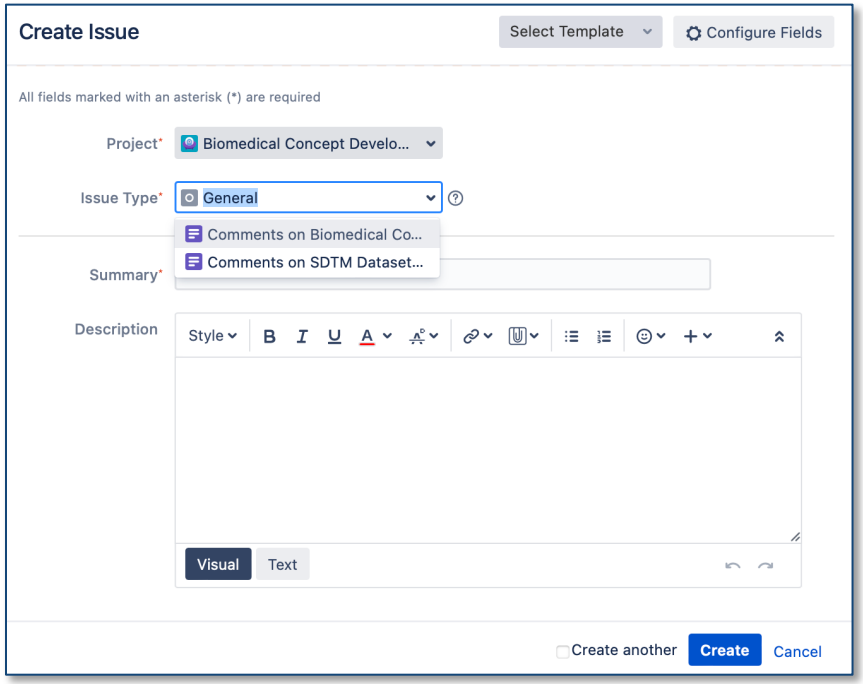
- Go to BCD JIRA website: Biomedical Concept Development - JIRA (cdisc.org) and click the 'Create' button located in the banner.

- There are 3 Issue Types:

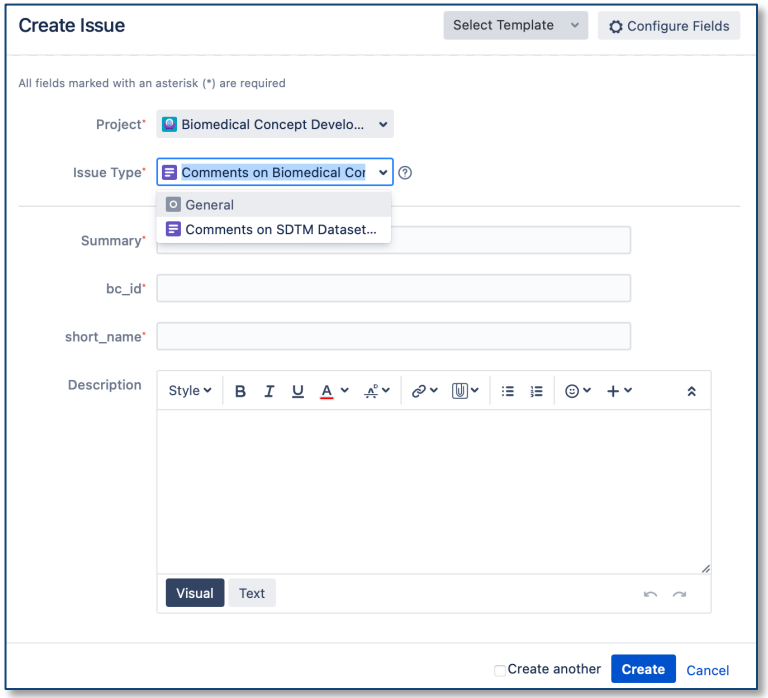
  1. General Comments: Feedback not related to a specific SDTM Dataset Specialization or Biomedical Concept. Include only a Summary and Description
  2. Comments on Biomedical Concepts, include the 'bc_id' and 'short_name'
  3. Comments on SDTM Dataset Specializations: Feedback on specific SDTM Dataset Specialization(s). In addition to a Summary and Description, requirements include 'Domain' and 'vlm_group_id'

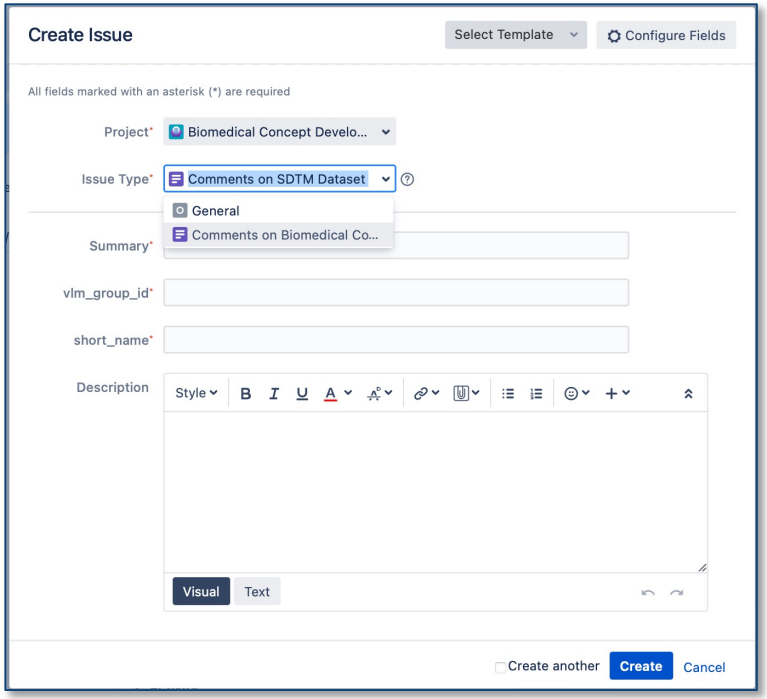


General Issues



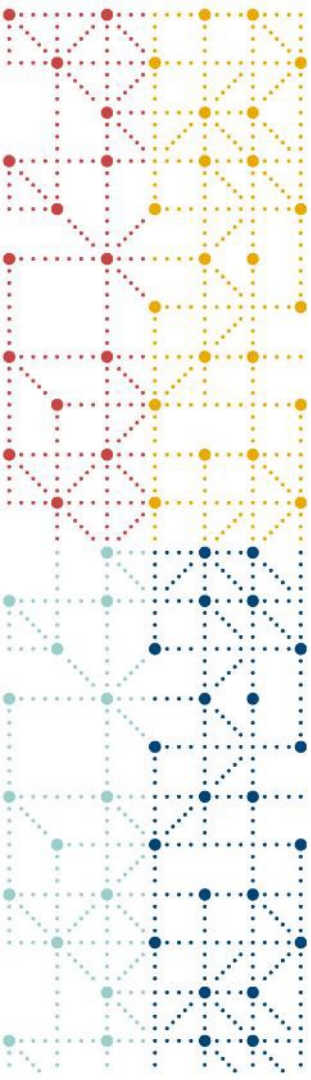
BC and SDTM Specializations





Resources:

- CDISC Biomedical Concepts | CDISC
- https://jira.cdisc.org/projects/BCD/summary
- COSMoS GitHub repository and Documentation
  https://cdisc-org.github.io/COSMoS



Backup Slides

Benefits to the CDISC Community

- **Regulators**
  - Reduces variability in standards implementation due to preconfigured value level metadata.
  - Promotes reuse and improves adherence to standards.
  - Fills gaps in the CDISC Foundational Standards, including explicit relationships and improved semantic definition.
- **EDC/CRF Designers**
  - Reduces time spent configuring the standards for data collection.
  - Promotes consistency in data collection and transformation.
  - Opens avenues for automation around EDC build.
- **Data Managers**
  - Reduces time spent interpreting the clinical protocol (e.g., BCs linked to a Schedule of Activities).
  - Promotes consistent protocol design and data collection.
  - Enables improved study specifications for vendors.
- **Statistical Programmers**
  - Facilitates end-to-end standards with improved data flow transparency.
  - Promotes smart/efficient programming code which help facilitate automation.
  - Provides building blocks for Define-XML through preconfigured value level data.
  - Promotes standardized reporting and reuse.
- **Academia**
  - Promotes reuse and improved adherence to standards.
  - Offers a more granular level of data standards which includes value level data.
  - Fills gaps in the CDISC Foundational Standards, including explicit relationships and improved semantic definition.