Enhancing Data Collection in Clinical Trials: Tips, Best Practices and Real-Life Applications

Maria Sekac, Associate Director Clinical Data Management, Merck & Co., Inc., USA, and
Amy Palmer, Associate Director Clinical Data Management, Merck & Co., Inc., USA

Meet the Speakers

Maria Sekac

Title: Associate Director Clinical Data Management

Organization: Merck & Co., Inc., USA

Maria is a Clinical Data Standards Lead at Merck, bringing over 25 years of experience in the pharmaceutical industry. With a strong foundation in clinical data management and industry standards, she has led initiatives across the clinical trial lifecycle—from protocol development to regulatory submissions—and has developed hundreds of CDASH and SDTM-compliant data collection specifications. Maria also serves on the Leadership Team of the CDISC Submission Data Standards team, contributing to the advancement of industry-wide data standardization efforts.

Amy Palmer

Title: Associate Director Clinical Data Management

Organization: Merck & Co., Inc., USA

Amy is a Clinical Data Standards Lead at Merck, bringing over 31 years of experience in the pharmaceutical industry. She joined Merck in January 2024 after 11 years at CDISC, where she contributed to advancing clinical data standards and industry collaboration. With a broad background in therapeutic area leadership, Amy supports oncology-focused clinical research while actively volunteering with the Events Adjudication and Non-Standard Variable CDISC team, as well as the CDISC QRS Controlled Terminology Team, helping to improve data quality and standardization across clinical trials.

Disclaimer and Disclosures

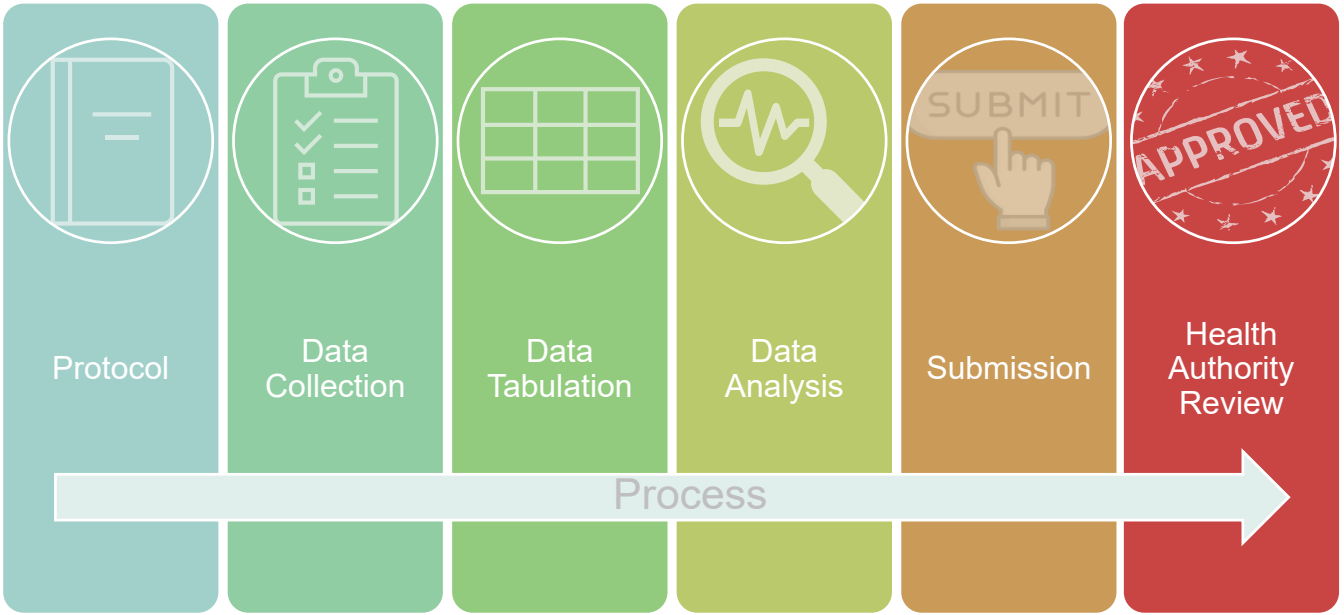
- *The views and opinions expressed in this presentation are those of the authors and do not necessarily reflect the official policy or position of CDISC.*

- *The authors have no real or apparent conflicts of interest to report.*

Agenda

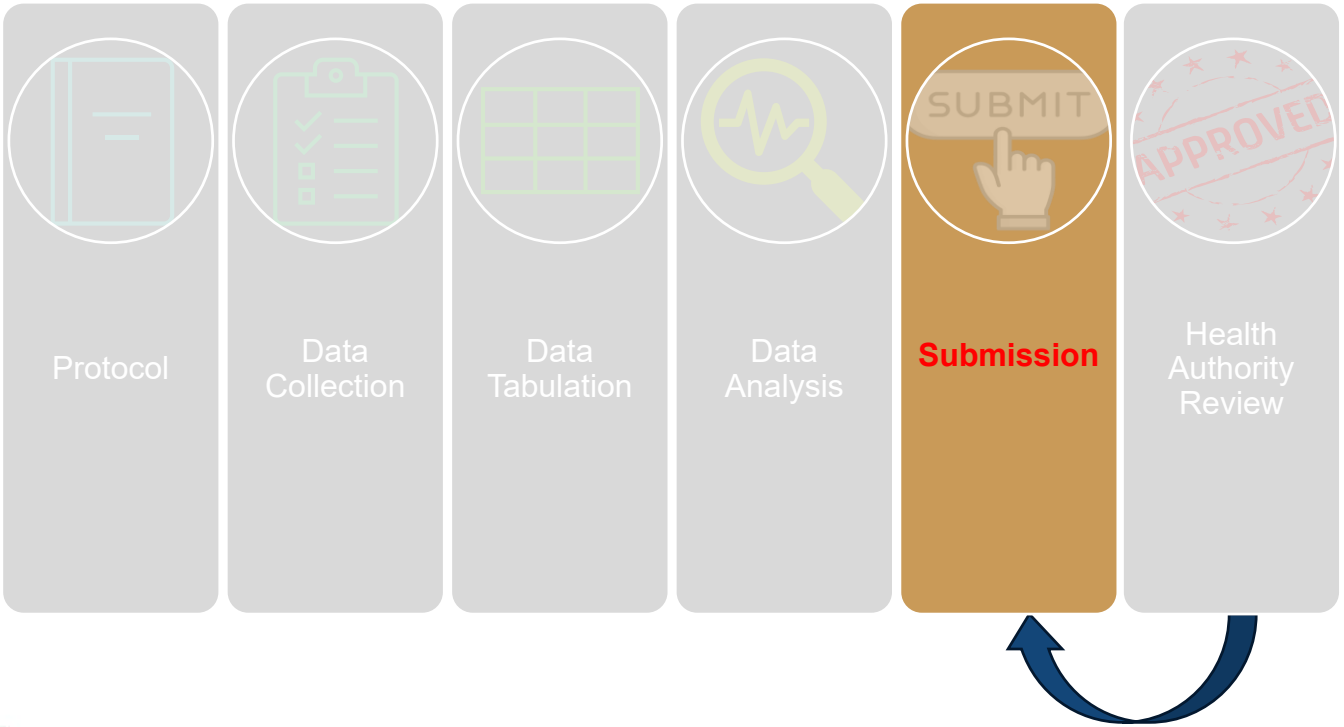
1. Clinical Trial Data Flow Process
2. Key Data Collection Principles
3. Resources and Tips to Support the Use of Standards in Data Collection

Clinical Trial Data Flow Process

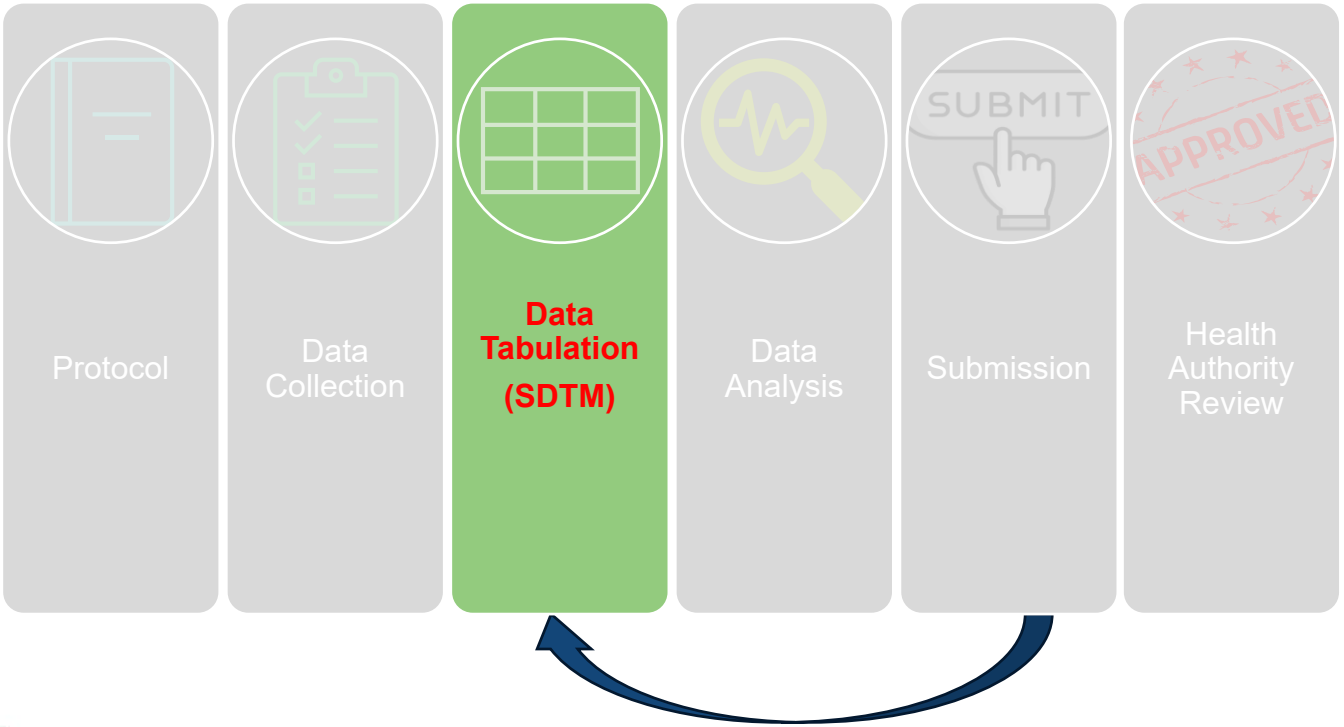


We Start With The End In Mind…

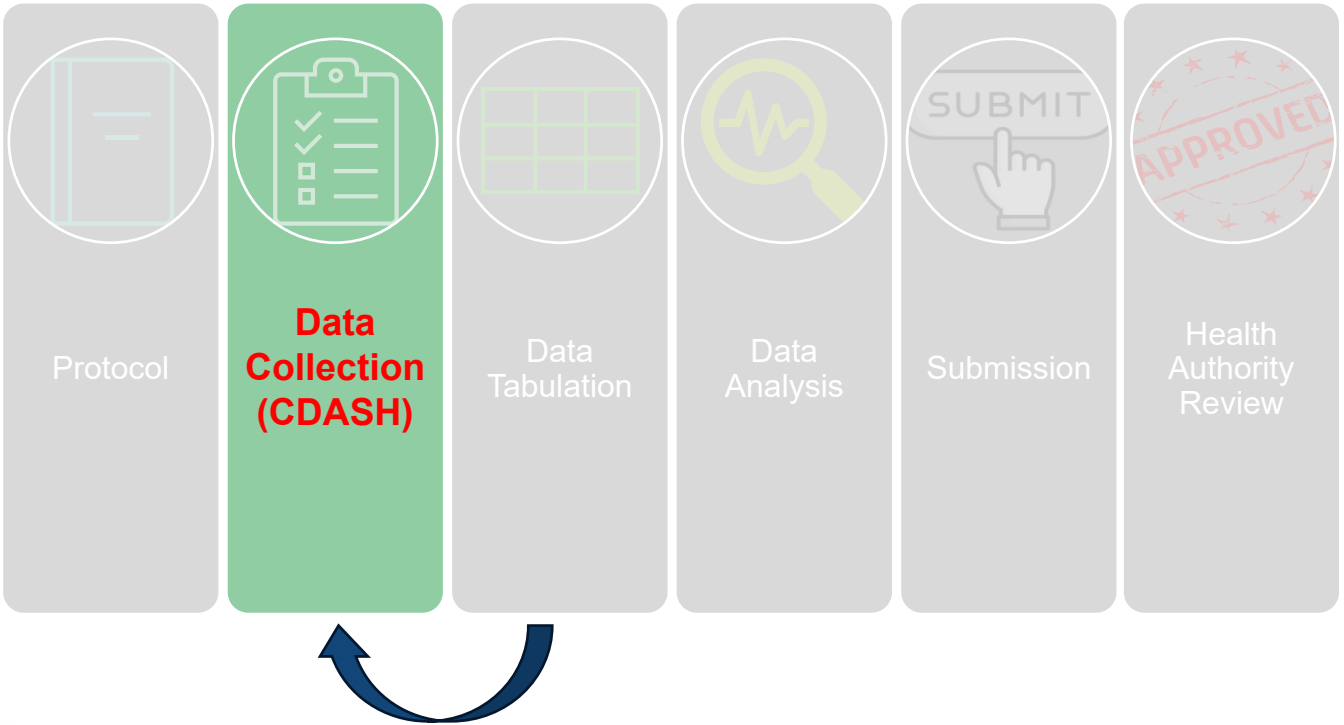
And Trace the Data Backwards

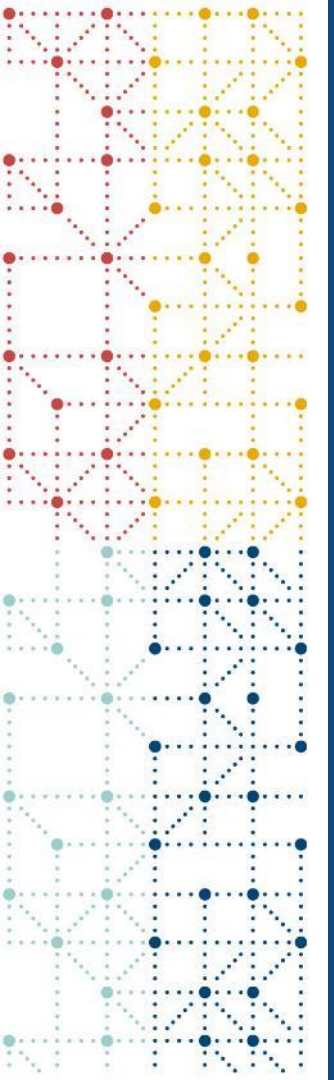


And Trace the Data Backwards



And Trace the Data Backwards





Key Data Collection Principles

Key Data Collection Principles

1. Encourage Lean Design (lean protocol and lean data collection)
   - Collect data that is fit for purpose (i.e., is relevant and provides value)
   - Avoid irrelevant or low-value data that consumes resources and time
2. Implement Quality by Design
   - Emphasizes quality from the start, including how the data will be cleaned
   - Ensure cross-functional agreement across the stakeholders (e.g., Clinical, Global Safety, Statistics, etc.)
   - Ensure harmonization and impact across Data Collection, CT, SDTM, ADaM
3. Implement and Utilize Standards

❖Note: Refer to the GCDMP and CDASHIG for additional data collection principles

Benefits of Standards

1. Improves Data Quality
2. Enables Consistency & Reliability
3. Reduces Time and Resources (and Cost) for Data Processing
4. Supports Reuse
5. Facilitates Ability to Compare Data Across Studies
6. Optimize Processes
7. Promotes Efficiency
   - Facilitates Automation
8. Enhances Traceability
9. Streamlines Regulatory Review and Approval

Resources and Tips to Support the Use of Standards in Data Collection

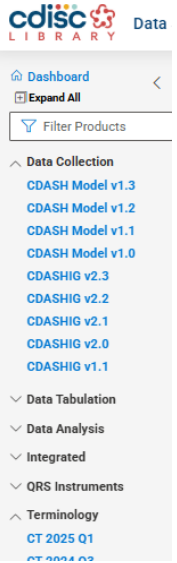
Resources to Support the Use of Standards in Data Collection

- ❑ CDISC Library Data Standards Browser
- ❑ CDISC SDTMIG
- ❑ CDISC SDTM Model
- ❑ CDISC SDTMIG Associated Persons
- ❑ CDISC SDTMIG Medical Devices
- ❑ CDISC SDTMIG Tobacco
- ❑ CDISC QRS Supplements
- ❑ CDISC QRS Naming Rules
- ❑ CDISC TAUGs
- ❑ CDISC NSVs & Fragments
- ❑ CDISC SDTM Controlled Terminology
- ❑ CDISC Codelist Rules
- ❑ CDISC CT Requests Denied
- ❑ CDISC Codetable Mapping Files
- ❑ CDISC CDASH Model
- ❑ CDISC CDASHIG
- ❑ CDISC Knowledge Base Articles (KBA)
- ❑ CDISC Errata
- ❑ Regulatory Guidance's (DSC, Tech Specs, Validation Rules)
- ❑ LOINC
- ❑ Web Browser Searching & AI
- ❑ GCDMP by SCDM
- ❑ CDISC Primer Videos

*Tip: Don’t panic! Have the awareness that these documents exist, refer to them when needed, and utilize automation where possible

CDISC Library Data Standards Browser

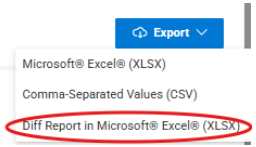
- Metadata repository of CDASH, SDTM, ADaM, QRS and CT
- Tips:
  - Utilize key words to identify appropriate variables
  - The variables include hyperlinks to other linked metadata (e.g., a CDASH variable hyperlinks to it's SDTM target and associated CT)



route

Name	Label	Definition	Question Text	Prompt	Data Type	Core	Completion Instructions	Implementation Notes	Mapping Instructions	Mapping Targets	Code List
CMROUTE	CM Route of Administration	The route of administration of the concomitant medication/treatment/therapy.	What was the route of administration of the (concomitant) [medication/treatment/therapy]?	Route	Char	R/C	Provide the route of administration for the (concomitant) [medication/treatment/therapy].	This additional information may be important to collect on the CRF when the sponsor wants to capture a medication's/treatment's route of administration, for purposes such as coding; also, the medication/treatment may have more than 1 route. Some companies may use route in coding medications/treatments, to be able to choose a precise preferred name and ATC code.	Maps directly to the SDTMIG variable listed in the SDTMIG Target column.	CMROUTE	C66729; C78420

- Compare any version with its previous version via diff report



SDTM Documents

- CDISC SDTMIG
  - https://www.cdisc.org/standards/foundational/sdtmig
  - Tip: Search for keywords for examples of data e.g., “bronchodilator”,
  - “Pharmaceutical Strength” is in the EC domain but not in the CM domain. Refer to the Model to confirm it can be used within the CM domain.
- CDISC SDTM Model
  - https://www.cdisc.org/standards/foundational/sdtm
  - Superset of all industry standard submission compliant SDTM variables

#	Variable Name	Variable Label	Type	Definition	Examples
34	--PSTRG	Pharmaceutical Strength	Num	The amount of active ingredient per unit of pharmaceutical dosage form.	"50", "300"

  - Review usage restrictions for variables (e.g., “Test Condition" for CP, IS and LB only)

#	Variable Name	Variable Label	Type	Format	Role	Variable(s) Qualified	Usage Restrictions	Variable C-code	Definition	Notes	Examples
8	--TSTCND	Test Condition	Char		Variable Qualifier	--TESTCD	CP, IS, and LB domains only			Identifies any planned condition imposed by the assay system on the specimen at the time the test is performed.	Stimulating or activating agents, assay temperature, incubation time "STIMULATED", "NON-STIMULATED", "25 C", "37 C".

Tips about SDTM

- Think of “SDTM” as “standard, structured formats”
  - How best to mockup the sample data to get it into a standardized, structured format
- Don’t need to be SDTM experts, mappers, or programmers to understand SDTM
- Learn these 5 basics:
  - What type of data (events, interventions or findings)
  - What are the standard domains to store the different types of data
  - What are the common standard item/variable names
  - Which of those common items expect to use controlled terminology
  - Which items are required, expected or permissible per domain

Tip: Create Sample Data

- Most effective tool to drive a conversation about **data**
- Data is the driver and asset for every clinical trial
- Understand how data should be compiled before designing a data collection instrument
- Provides stakeholders a visual and shared understanding of the compiled data, ability to identify gaps, prevents design flaws
- Supports quality by design

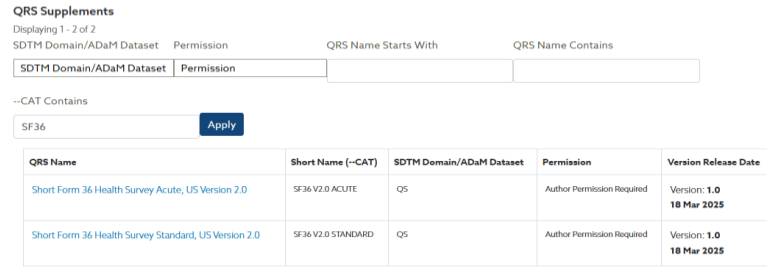
Example of Sample Data & Questions to Ask

Study	Subjid	Category	Subcategory	Test	Result	Unit	Evaluation Interval	Start Date	Start Time	End Date	End Time
ABC	100	Fitness Test	Endurance	Jumping Jacks	50		1 minute	1/1/2023	8:00 AM	1/1/2023	9:00 AM
ABC	100	Fitness Test	Endurance	Pushups	20		1 minute	1/1/2023	9:10 AM	1/1/2023	9:11 AM
ABC	100	Fitness Test	Endurance	Squats	30		1 minute	1/1/2023	9:15 AM	1/1/2023	9:16 AM
ABC	100	Fitness Test	Endurance	Distance Run	3	miles	30 minutes	1/1/2023	9:30 AM	1/1/2023	9:59 AM

- *Is the data meaningful and does it make sense?*
- *Does it tell a complete story?*
- *Are there any datapoints missing?*
- *Are there data quality issues?*
- *Could the data potentially result in site (or participant) data entry burden?*
- *Could the data result in any potential query burden?*
- *Will this data support the analysis?*
- *Will the data support submission compliance?*
- *Anticipate and identify potential questions from reviewers of data, auditors, and inspectors*

CDISC QRS Supplements

- CDISC QRS (Questionnaires, Ratings, Scales)
  - Supplemental documents have been created for QRS instruments separate from the SDTMIG
  - https://www.cdisc.org/standards/foundational/qrs
  - Data mapped to the QS, RS or FT domains
  - Tips:
    - Search for specific keywords across the fields



    - Also check SDTM CT as terminology is developed ahead of the supplements
- CDISC QRS Naming Rules
  - Reference when defining and mapping QRS instruments not published by CDISC

Non-Standard Variables (NSVs) & Fragments

- Also known as Supplemental Qualifiers
- Used when an industry standard variable does not exist
- Registry of NSVs used in TAUGs and other foundational standards
  - https://www.cdisc.org/standards/terminology/non-standard-variables
  - Recommend referring to this list to reuse existing NSVs
  - Curated and governed for promotion to a standard variable
- If not in registry, sponsors can create new NSVs as needed
  - CDISC has published naming fragments and principles for creating NSVs
- Tip: Ensure the variable does not already exist in the CDASH or SDTM Models and review Model restrictions

SDTM Controlled Terminology

- https://www.cdisc.org/standards/terminology/controlled-terminology
- Controlled Terminology (CT): set of codelists and valid values used with data items
- Industry standard Controlled Terminology exists for many industry standard variables:

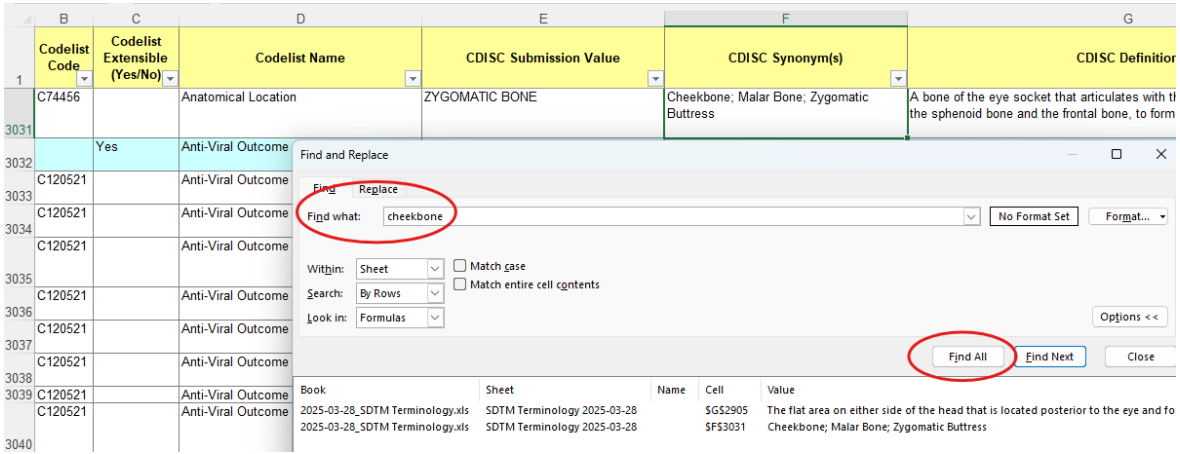
Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes
VSTESTCD	Vital Signs Test Short Name	Char	(VSTESTCD)	Topic	Short name of the measurement, test, or examination described in VSTEST. It can be used as a column n horizontal format. The value in VSTESTCD cannot be longer than 8 characters, nor can it start with a num contain characters other than letters, numbers, or underscores. Examples: "SYSBP", "DIABP", "BMI".
VSORRESU	Original Units	Char	(VSRESU)	Variable Qualifier	Original units in which the data were collected. The unit for VSORRES. Examples: "in", "LB", "beats/min".

- CT is required for submission compliance to FDA and PMDA
- Using CT from the start during data collection builds in traceability and transparency and reduces the challenge with converting or remapping the data to compliant terminology

- Tip: Refer to the latest version for current terminology (published 2x/year)
  - Work with Standards and/or Submission Team on impact of using latest version for backwards compatibility

Searching for Values in the SDTM CT Metadata

- Tips:
  - Export the latest SDTM CT from the CDISC Library
  - Search broadly across all properties of the CT
    - Only searching within the Submission Value column will limit the results, as the keywords might appear as a Synonym or part of the definition
    - e.g., if searching for a location of “Cheekbone”, notice it’s a synonym to “Zygomatic Bone”, so will not appear as a submission value

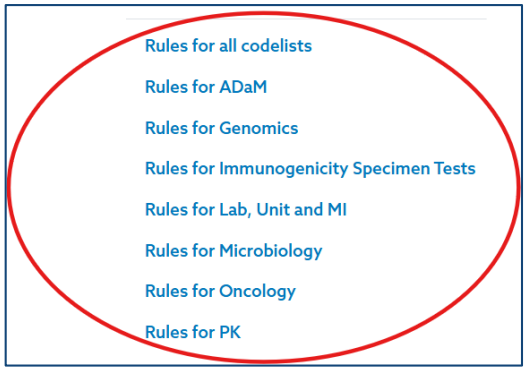


Rules for Creating New Controlled Terminology

- New terms can be created if the codelist associated to the SDTM target variable, is considered extensible

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]
FALOC	Location of the Finding About	Char	(LOC)

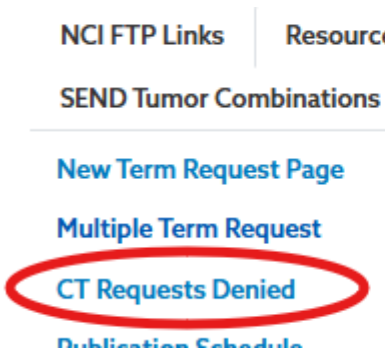
Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition
C74456		Yes	Anatomical Location	LOC	Anatomical Location	Terminology codelist used for anatomical location within CDISC.
C12401	C74456		Anatomical Location	EYE	Eyeball	The sensory organ of vision.



- Avoid creating new CT which is redundant to an existing CDISC Synonym

CT Requests Denied

- https://www.cdisc.org/standards/terminology/controlled-terminology
- Tip: search this file for keywords before creating new extensible terms



Codelist Name	Original Request	Original Requestor Notes	Request Type	Final Decision Reason for Request Denied
LOC	Please add abdomen	We require an anatomical location for reported abdominal pain. Until diagnostics are performed, the cause of the pain could be located either in the "Abdominal Cavity" (for which abdomen is currently listed as a synonym) or the "Abdominal Wall" (which is not now noted as a synonym). We feel that the more broad term is necessary to accurately represent the reported symptom.	New Term Request	Do not add. Already a synonym to another term.
LOC	THALAMUS GREY	The grey matter component of the thalamus.	New Term Request	Do not add. This would map to a LOC of Thalamus (C12459) - already published, and a SPECTYPE of Grey Matter (already requested).

Codetable Mapping Files

- Codetable Mapping Files show relationships of data across different codelists
- Supports requirements gathering, design, and data cleaning to ensure combinations of data are appropriate
- Tip: Take advantage of these useful tools when defining data collection requirements ☺

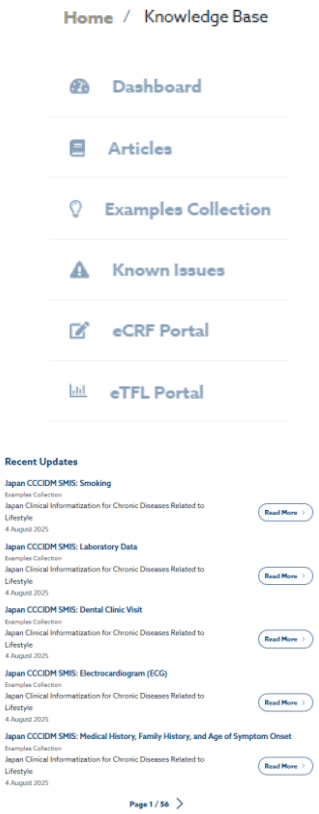
Vital Signs Test Code (VSTESTCD) (codelist code = C66741)	Vital Signs Test Name (VSTEST) (codelist code = C67153)	Units for Vital Signs Results (VSRESU) (codelist code = C66770)	Anatomical Location (LOC) (codelist code = C74456)
TEMP	Temperature	C	RECTUM
TEMP	Temperature	C	AXILLA
TEMP	Temperature	C	EAR
TEMP	Temperature	C	FOREHEAD
TEMP	Temperature	C	SUBLINGUAL REGION
TEMP	Temperature	F	RECTUM
TEMP	Temperature	F	AXILLA
TEMP	Temperature	F	EAR
TEMP	Temperature	F	FOREHEAD
TEMP	Temperature	F	SUBLINGUAL REGION

CDASH Model and IG

- CDISC CDASH Model
  - Superset of all industry standard data collection root variables per observation class
  - Facilitates mapping collected data to SDTM
- CDISC CDASHIG
  - Domain specific
  - Most commonly used variables in domains
  - Due to timing of releases, there are domains and variables in SDTMIG that are not yet in the CDASHIG
  - Tip: identify the SDTM domain/variables and then apply the CDASH conventions
  - ~89% of CDASH maps directly to SDTM variables, the 11% difference is due to differing data capture needs
  - Review sections:

CDISC Knowledge Base

- https://www.cdisc.org/kb
- Resource hub supporting CDISC standards implementers
- Includes:
  - Articles tailored to specific interests
  - Collection of topic-specific examples
  - Known Issues
  - eCRF Portal with example case report forms
- Ability to search and filter
- Tip: Add a calendar reminder to review “Recent Updates” which is sorted by date order



Knowledge Base Articles of Interest to Data Collection

1. https://www.cdisc.org/kb/articles/magic-happens-between-cdash-and-sdtm
2. https://www.cdisc.org/kb/articles/cdisc-published/why-would-standards-development-organization-make-non-standard
3. https://www.cdisc.org/kb/articles/cdisc-published/why-study-subject-instead-study-participant
4. https://www.cdisc.org/kb/articles/cdisc-published/why-are-certain-sdtm-variables-not-use-human-clinical-trials
5. https://www.cdisc.org/kb/articles/sex-and-gender

Key Regulatory Guidance's for Data Collection Awareness

- https://www.fda.gov/industry/fda-data-standards-advisory-board/study-data-standards-resources
  1. FDA Data Standards Catalog (DSC)
     - Which SDTM IG version is required to be followed for submissions
  2. FDA Study Data Technical Conformance Guide (TCG)
     - Explains the Study Data Reviewers Guide (SDRG)
     - SDTM Considerations
     - Supported TAUGs
     - List of Technical Specification Documents
     - Use of CT and "OTHER"
     - Study Data Validation and Traceability
  3. FDA Business Rules
  4. FDA Validator Rules
- PMDA has similar documents

Tip: Refer to these documents to drive discussions related to the impacts of data collection through submission

Study Data Reviewer's Guides

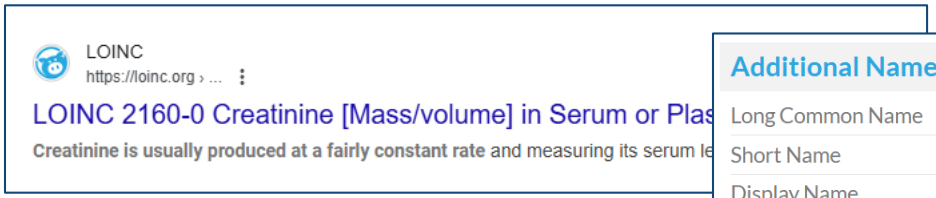
The preparation of the relevant Reviewer Guides (RG)[12] is recommended as an integral part of a standards-compliant study data submission. An RG should describe any special considerations or directions or conformance issues that may facilitate an FDA reviewer's use of the submitted data and may help the reviewer understand the relationships between the study report and the data.

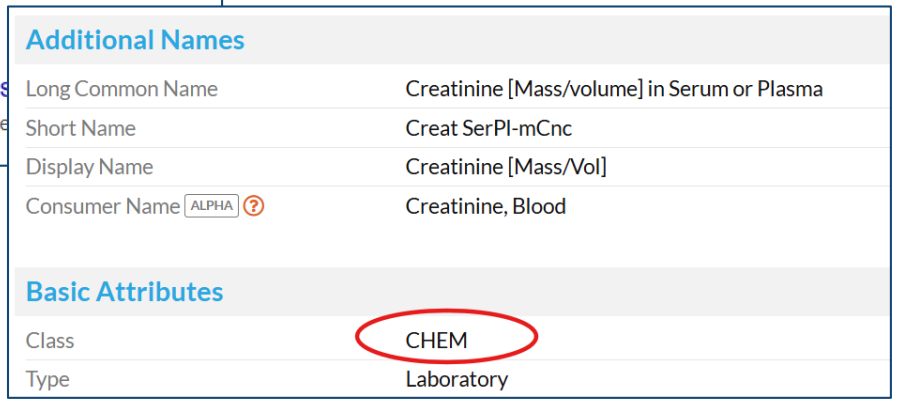
Examples of Regulatory Business Rules

1. Controlled terms should use the exact term (case, spelling, and punctuation) used by the terminology maintenance organizations (e.g., MedDRA, CDISC controlled terminology)
2. When a test is not done, the result variable should not be populated, and the reason not done should be provided in the appropriate variable
3. All death information should be populated for subjects that died during the study including any post treatment follow-up
4. Category for Disposition Event (DSCAT) should be populated
5. For interventional studies all treated subjects should have exposure data
6. Randomized subjects are expected to receive a study treatment
7. Trial participants should self-report race and ethnicity and they should not be assigned by the study team
8. When collecting ethnicity demographic data from clinical trial participants, the following two minimum choices should be offered: "HISPANIC OR LATINO" or "NOT HISPANIC OR LATINO"
9. When collecting racial demographic data from clinical trial participants, the following five minimum choices should be offered: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White
10. Tip: Utilize automation and tools such as CDISC CORE to execute the rules

Using LOINC to Determine a Specimen Category

- LOINC = Logical Observation Identifiers Names and Codes
- Tip: Use to determine a CAT for the Specimen Based Findings Test
- Which CAT does the test “Creatinine” go to?
  - Web browser search: “LOINC Creatinine”. The LOINC Class will help determine the CAT = CHEMISTRY



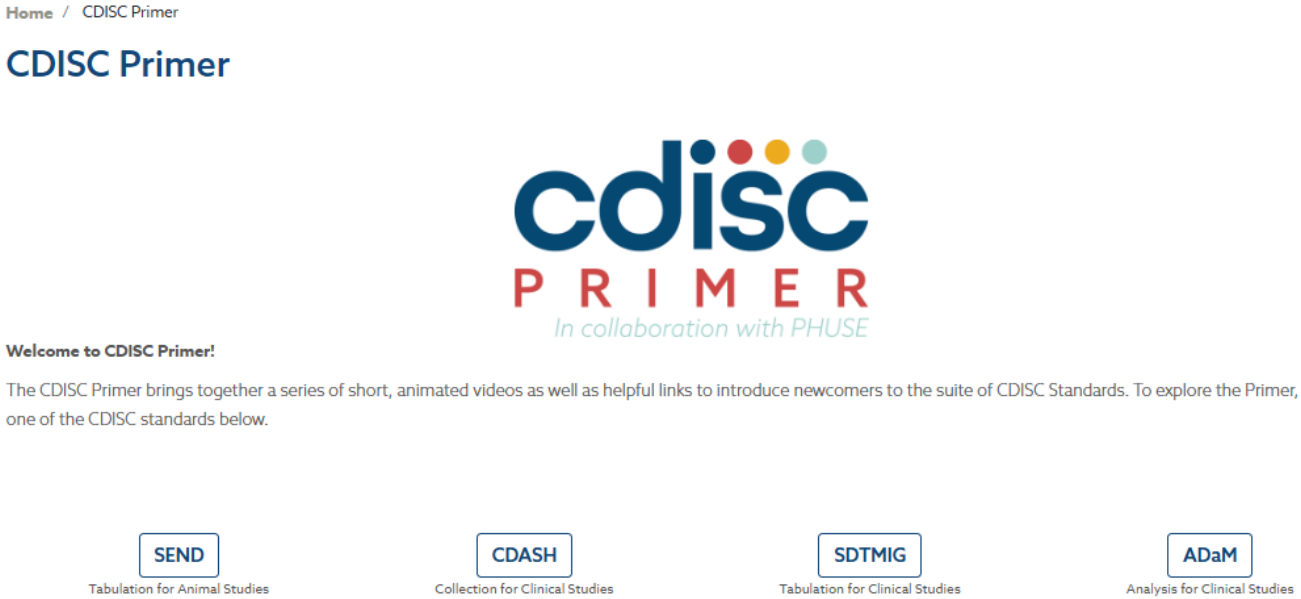


GCDMP by SCDM

- https://scdm.org/gcdmp-chapters/
- The GCDMP© provides comprehensive best practices covering essential data management topics:
  - “If the data points specified in the protocol are not accurately collected, a meaningful analysis of the study’s outcome will not be possible. Therefore, the design, development, and quality assurance processes of a *data collection instrument* must receive the utmost attention.”
  - “Design *data collection instruments* with safety and efficacy endpoints in mind.”
  - “Keep the *data collection instruments* questions, prompts, and instructions clear, concise and conformant to CDISC CDASH standards, where possible.”

CDISC Primer Videos

- https://www.cdisc.org/primer

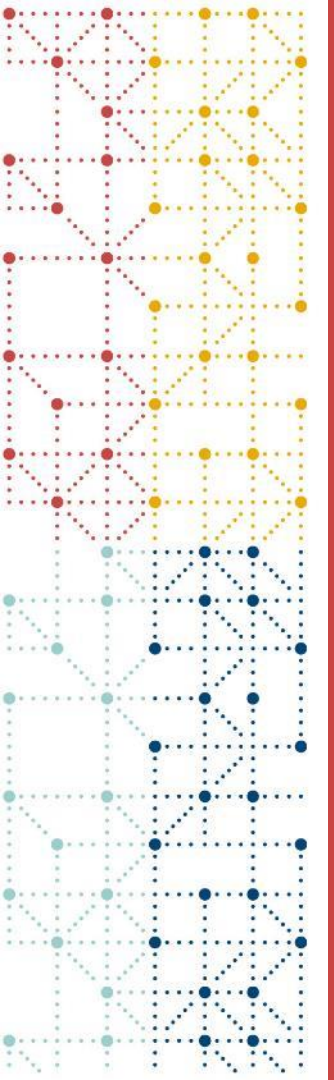


CDISC Primer Videos

- Introduction
- Versions
- Therapeutic Area User Guides
- Controlled Terminology
- Traceability
- Regulatory Requirements
- Standards in Development
- Knowledge Base
- Education

Introduction

CDASH: The Clinical Data Acquisition Standards Harmonization Implementation Guide (CDASHIG) establishes a standard way to collect data consistently across studies and sponsors so that data collection formats and structures provide clear traceability of submission data into the SDTM, delivering more transparency to regulators and others who conduct data review. The following videos introduce you to the CDASH model and the CDASHIG.



THANK YOU!

Backup Slides To Support Questions

CDASH and SDTM

3.1 How CDASH and SDTM Work Together

1. The SDTM and the SDTMIG provide a standard for the submission of data. CDASH is earlier in the data flow and defines a basic set of data collection fields that are expected to be present on the majority of CRFs. SDTM and CDASH are clearly related. The use of CDASH data collection fields and variables is intended to facilitate mapping to the SDTM structure. When submitted data can be collected as represented in an SDTM dataset, with no transformations or derivations, the SDTMIG variable names are presented in the CDASHIG metadata table and should be used to collect the data. In cases where the collected data must be transformed or reformatted prior to inclusion in an SDTM dataset, or where a corresponding SDTMIG variable does not exist, CDASH has created standardized data collection variable names.

- During the development of CDASH-conformant collection instruments (e.g., CRFs, eCOA screens), the SDTMIG domain to which the collected data is to be mapped must be determined. The choice of the SDTMIG domain to use does *not* depend upon the mode of transmission, the methodology used to generate the data, the medium used to store the data, the person who recorded the data, or the subject described by the data. The SDTMIG domain to be used affects what CDASH variable names, question texts, prompts, controlled terminology, and so on, to use. CDASH suggests a format to be presented to those entering the data, but it does not dictate any data structure in which to store the collected data (often referred to as a "data management operational database").

Difference Between CDASH and SDTM

- CDASH and SDTM are closely aligned but have different purposes
  - CDASH optimized for data capture but flexible for different types of data capture scenarios.
    - e.g. Normalized vs Denormalized; eCRF vs eDT
    - Utilizes different variable names when the data is not a direct mapping from source to target (e.g. DAT vs DTC)
  - SDTM optimized for organizing, aggregating and tabulating data
    - E.g. Merge Local Lab RBC data and Central Lab RBC data together
    - Designed to be machine readable (e.g. date collected as 07Jan2025 but converted to machine readable 2025-01-07 format)

DAT *(CDASH)*	->	DTC *(SDTM)*
07Jan2025, 2:30pm	->	2025-01-07T14:30
(Human readable)	->	(Machine readable; ISO format)

CDASH Denormalized Data Collection Design

Vital Signs (VS) *[Flat: Each Test is Listed as It's Own Item]*			
Item #	**Question Text**	**Collected Data**	**Metadata**
1	Height	XXX.XX	Name: VSORRES_HEIGHT
2	Height Unit	o in o cm	Name: VSORRESU_HEIGHT Codelist: C71620
3	Weight	XXX.XX	Name: VSORRES_WEIGHT
4	Weight Unit	o kg o LB	Name: VSORRESU_WEIGHT Codelist: C71620
5	Systolic Blood Pressure	XXX	Name: VSORRES_SYSBP
6	Diastolic Blood Pressure	XXX	Name: VSORRES_DIABP
7	Blood Pressure Unit	mmHg	Name: VSORRESU_BP

CDASH Collect → SDTM Organize (Tabulate)

Raw EDC Data (Denormalized; Horizontal; Unpredictable Structure as the Amount of Tests/Columns can change from study to study; sponsor to sponsor)

VSORRES_HEIGHT	VSORRESU_HEIGHT	VSORRES_WEIGHT	VSORRESU_WEIGHT	VSORRES_SYSBP	VSORRES_DIABP	VSORRESU_BP
65.50	In	135.25	LB	120	90	mmHg

A denormalized structure must be transposed to a normalized structure

SDTM Dataset
(Normalized; Vertical; Predictable, Reuseable, Same Structure Regardless of the Amount of Tests)

VSTEST	VSTESTCD	VSORRES	VSORRESU	VSSTRESN	VSSTRESU
Height	HEIGHT	65.50	in	166.37	cm
Weight	WEIGHT	135.25	LB	61.34	kg
Systolic Blood Pressure	SYSBP	120	mmHg	120	mmHg
Diastolic Blood Pressure	DIABP	90	mmHg	90	mmHg

Example of Sample Data with a Potential Quality Issue

STUDY	SUBJID	ECTRT	ECROUTE	ECLOC
ABC	100	DRUG X	SC	STOMACH
ABC	100	DRUG X	IV	
ABC	200	DRUG X	SC	LEG
ABC	200	DRUG X	IV	ARM
ABC	300	DRUG X	SC	ABDOMINAL WALL
ABC	300	DRUG X	IV	

- EC form where ECLOC is not required for IV administrations
- Think about the data merged in a dataset
- LOC is permissible in SDTM but there is an inconsistency in its collection:
  - (1) LOC is present for all SC routes
  - (2) LOC is inconsistently present for IV routes
- Could this raise questions from a reviewer, auditor, inspector?
- Not a submission compliance issue, but a data quality issue
- Should the site be burdened with a query to remove the ECLOC for “ARM” or burden the site with an entry for all IV records?
- No right or wrong answer. The sample data supports driving the conversation.

Currently Available Codetable Mapping Files

DD Codetable	DS Codetable	CV Codetable	EG Codetable	GF Codetable
GI Codetable	IG Codetable	IS Codetable	MK Codetable	Oncology Codetable
Race Ethnicity Codetable	RE Codetable	RP Codetable	SC Codetable	SR Codetable
SS Codetable	TS Codetable	UR Codetable	VS Codetable	

CDISC TAUGs

- Therapeutic Area User Guides
  - Extensions of CDISC Foundational Standards
  - https://www.cdisc.org/standards/therapeutic-areas

Acute Kidney Injury
Alzheimer's
Asthma
Breast Cancer
Cardiovascular
CDAD
Colorectal Cancer
COPD
COVID-19
Crohn's Disease
Diabetes
Diabetes Type 1 - Exercise and Nutrition
Diabetes Type 1 - Pediatrics and Devices
Diabetes Type 1 - Screening, Staging and Monitoring of Pre-clinical Type 1 Diabetes
Diabetic Kidney Disease
Duchenne Muscular Dystrophy
Dyslipidemia
Ebola
Heart Failure
Hepatitis C
HIV
Huntington's Disease
Influenza
Kidney Transplant
Lung Cancer
Major Depressive Disorder
Malaria
Multiple Sclerosis
Nutrition
Pain
Pancreatic Cancer
Parkinson's Disease
Pediatrics
Polycystic Kidney Disease
Post Traumatic Stress Disorder
Prostate Cancer
Psoriasis
QT Studies
Rare Diseases
Rheumatoid Arthritis
Schizophrenia
Traditional Chinese Medicine - Acupuncture
Traditional Chinese Medicine - Coronary Artery Disease-Angina
Traumatic Brain Injury
Tuberculosis
Vaccines
Virology

Other SDTM IG Documents

- CDISC SDTMIG Associated Persons (AP)
  - https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-ap-v1-0
  - For data other than the subject/participant (e.g., parents, child, etc.)
- CDISC SDTMIG Medical Devices (MD)
  - https://www.cdisc.org/standards/foundational/medical-devices
  - For data about medical devices
- CDISC SDTMIG Tobacco (TIG)
  - https://www.cdisc.org/standards/foundational/tobacco-implementation-guide
  - Used for tobacco product research and submissions

CDISC Errata

An erratum (plural: errata) is a correction to a mistake made in a published standard, which was identified after it was published on the CDISC website, and which would have been corrected in the current version had it been identified prior to publication.

SDTMIG v3.4

Section 1.5 - Known Issues

- The issue "Use of --LNKID and --LNKGRP" has text that reads, "…; a row where RELTYPE = MANY will have an IDVAR of --LNKID (and not --LNKGRP).", but should be, "… ; a row where RELTYPE = MANY will have an IDVAR of --LNKGRP (and not --LNKID)."

Section 3 - Submitting Data in Standard Format

- Section 3.2.1 - Dataset-Level Metadata
  - In the row of the table for the SV dataset, the value given for Keys is "STUDYID, USUBJID, SVTERM" but should be "STUDYID, USUBJID, VISITNUM".
  - UR metadata row, in the column "Structure", "per per" should be changed to "per".

Section 5.3 - SE Domain

- Example 1, dm.xpt, Variable BIRTHDTC should be BRTHDTC.
- Example 2, dm.xpt, Variable BIRTHDTC should be BRTHDTC.

Section 6.3.7.8 Urinary System Findings

- UR - Specifications: In the table title, "per per" should be changed to "per".

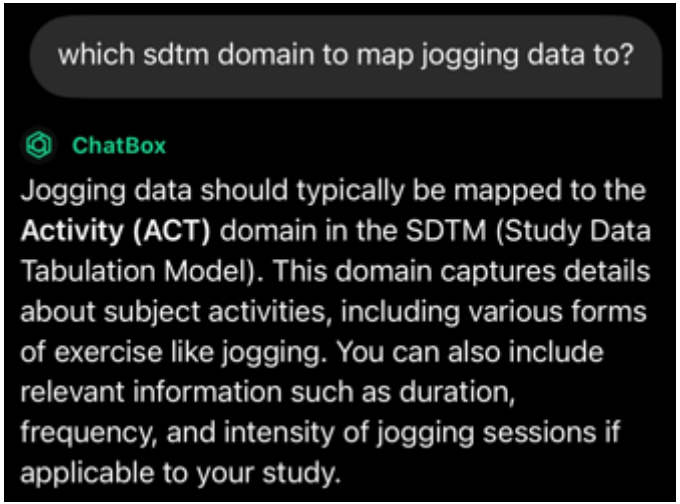
Section 8.4.3, Trial Summary

- --SUPP Examples, Example 1, suppae.xpt: The value for QEVAL in Row 2 should be changed from "SPONSOR" to null.

How Web Browsing Can Help

- Tip: Web browsing can assist in finding articles or papers written to support the type of datapoints of interest...Examples:
- Google searches:
- “How do I know if an instrument is considered an SDTM clinical classification?”
  - Google Search Result: PharmaSug paper “Clinical Classifications” which describes mapping to the RS domain
- “Which SDTM domain should I map adjudication data to?”
  - Google Search Result: Phuse paper “Best Practices for Submission of Event Adjudication Data”

Example of an Artificial Intelligence Application Search



- When asked “Which SDTM domain to map jogging data to?”
  - It responded with a domain that does not exist
  - Tip: Be cautious about utilizing AI to identify SDTM mapping for data concepts which haven’t yet been defined in SDTM