



2023

KOREA

INTERCHANGE

SEOUL | 11-14 DECEMBER



## Considerations in converting study data and lesson learned from our experiences

Presented by Mia Kim, Clinical Data Team Manager  
Clinical Data Team, DATA TEAMz Inc.



# Meet the Speaker

Mia Kim

**Title:** Clinical Data Team Manager

**Organization:** DATA TEAMz Inc.

Mia Kim, who has been working in the field of clinical data management for 13 years, majored in nursing at Yonsei University and started working in clinical data management at the local CRO in Korea. Inspired and influenced by her team leader and colleagues she worked with at the time, she became interested in data management.

She has also experience in clinical data management work at Asan Medical Center and served as head of the data quality innovation team and head of the data management team at a full service local CRO. Currently, she is working as a clinical data team manager at DATA TEAMz Inc., a company specializing in clinical data management, to bring new perspectives and technologies to data management work to make data science a reality. She is interested in energizing people and increasing job satisfaction for data managers and loves thinking about the future of data management as it evolves in response to a changing environment.



# 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.*



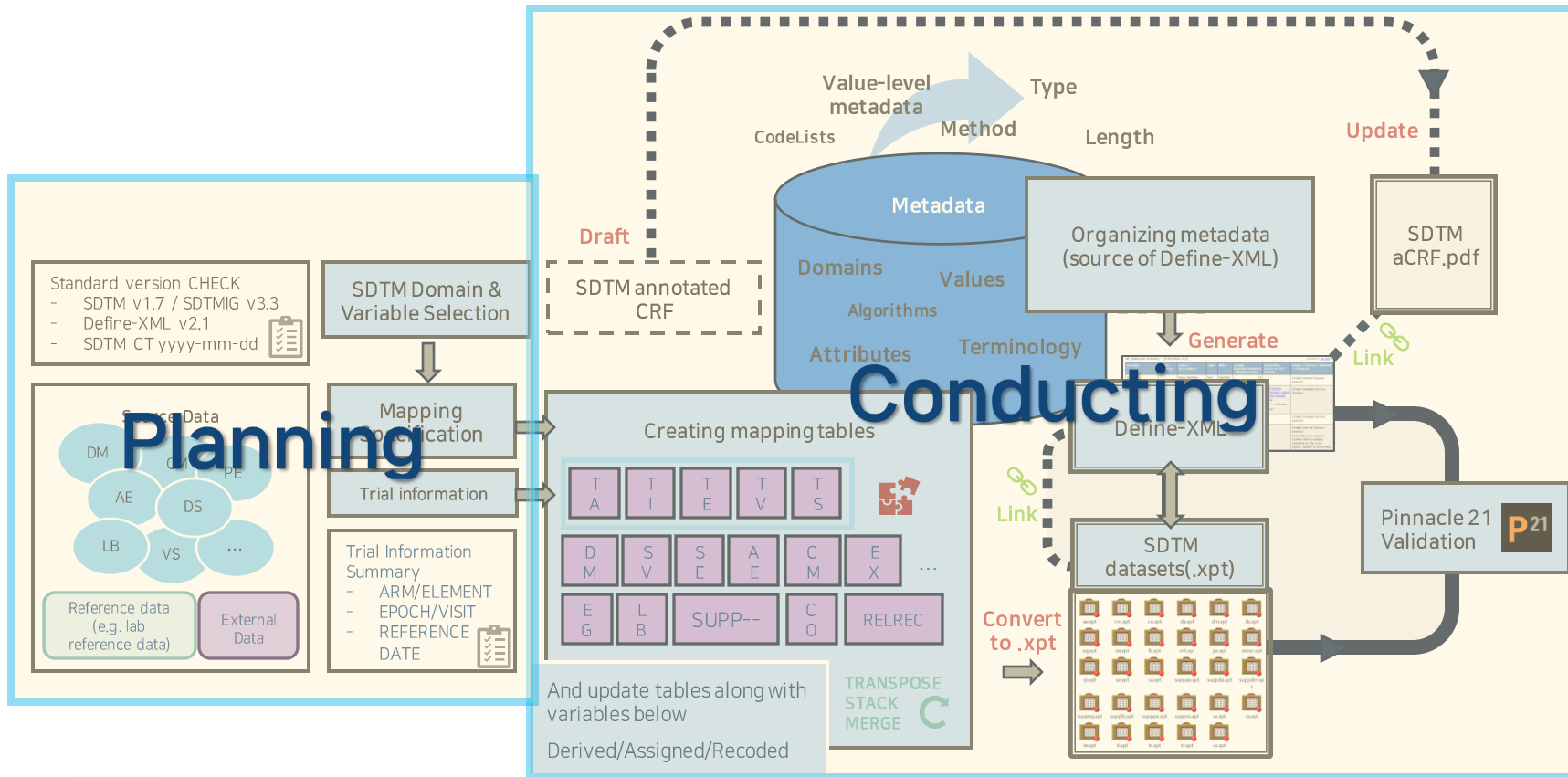
## Agenda

1. Considerations in converting study data and the cases you may encounter when you are in early stage of work
2. Lesson learned from our experiences



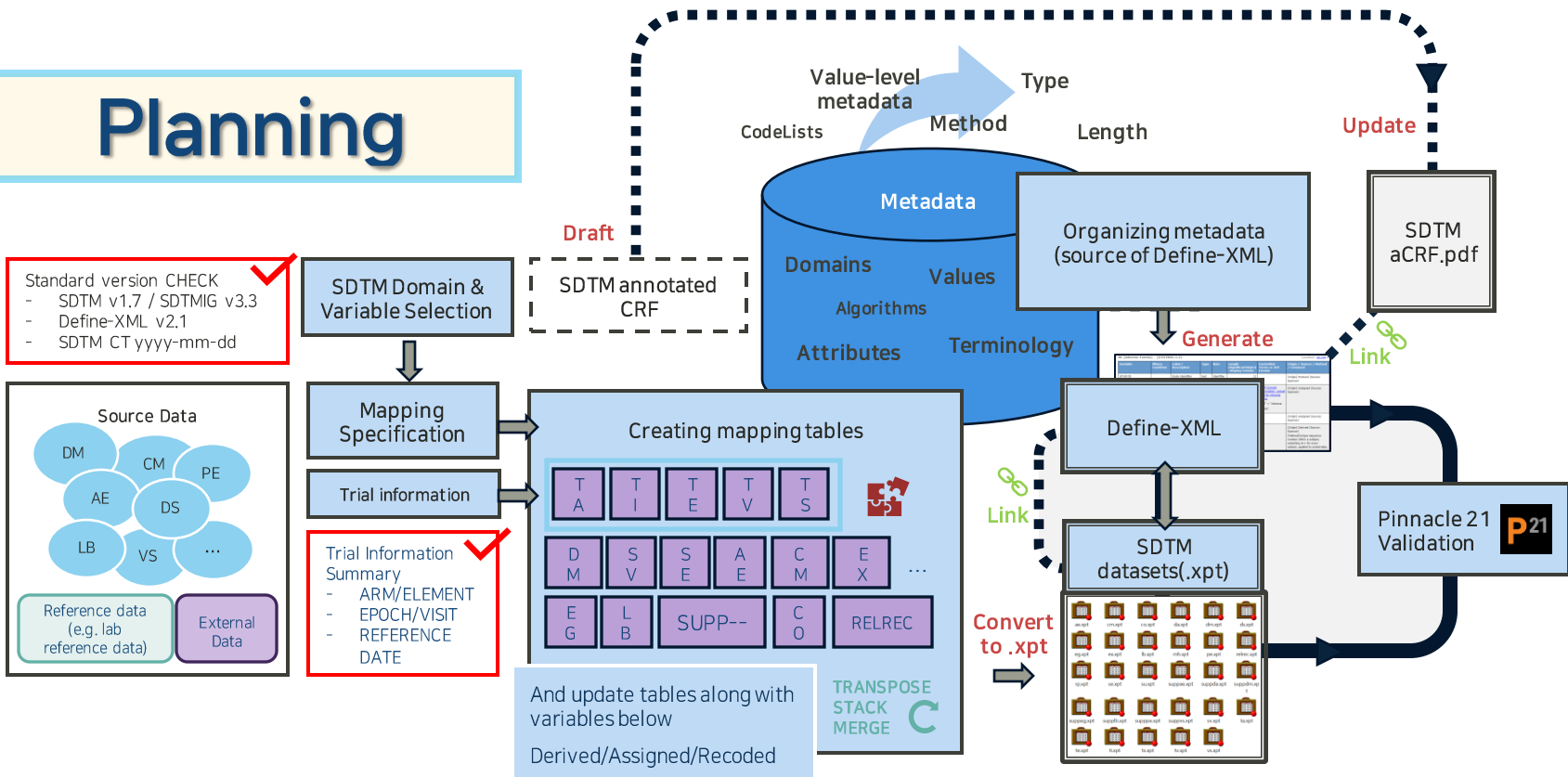
## Considerations in converting study data and the cases you may encounter when you are in early stage of work

# Data Conversion Process



# Data Conversion Process

## Planning



# Data Conversion Process

Standard version CHECK

- SDTM v1.7 / SDTMIG v3.3
- Define-XML v2.1
- SDTM CT yyyy-mm-dd

Trial Information Summary

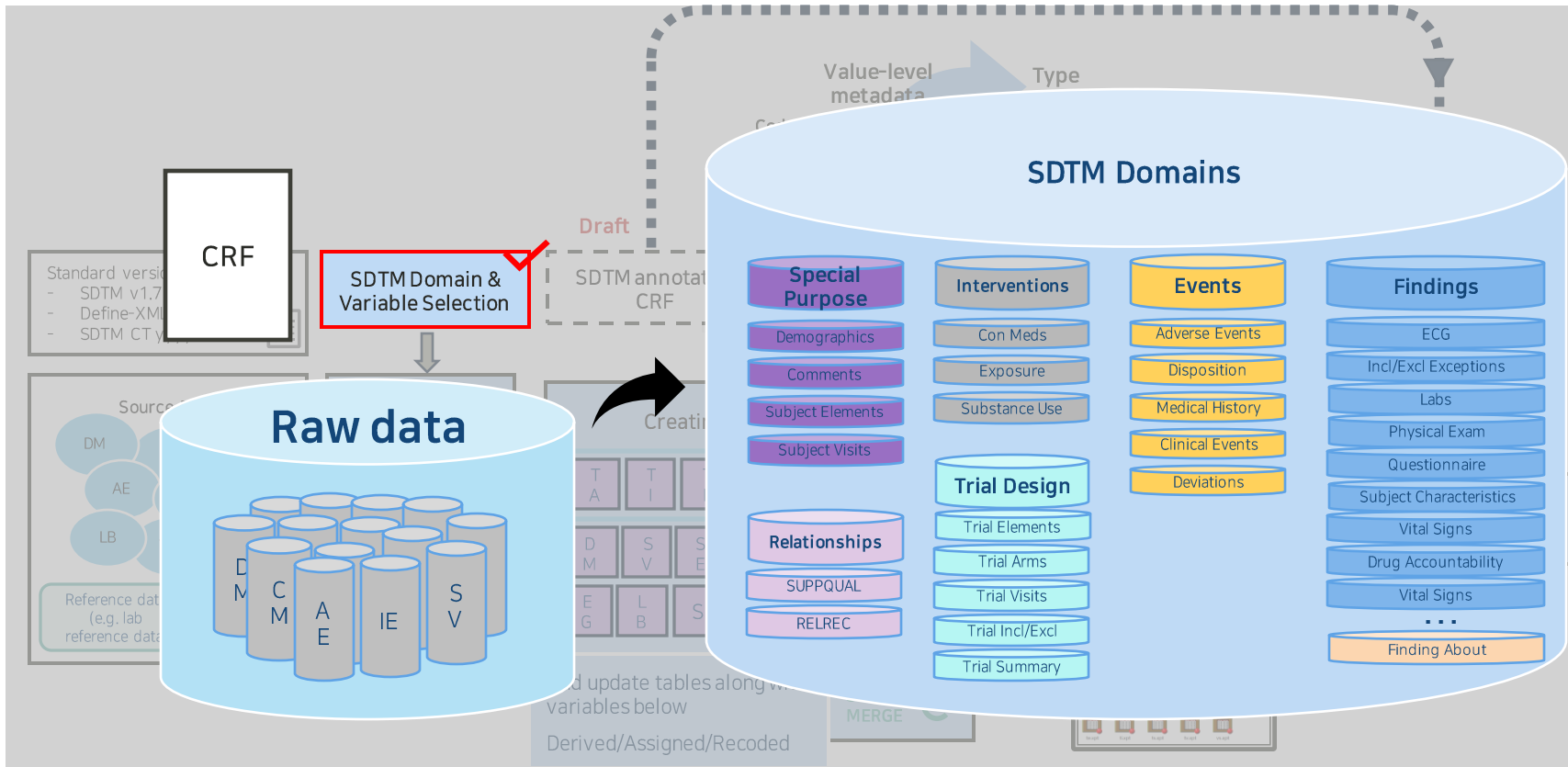
- ARM/ELEMENT
- EPOCH/VISIT
- REFERENCE DATE

Version					
SDTMIG	V3.3				
SDTM CT	2022-12-16				
Define-XML	V2.1				
MedDRA	V24.0				
WHO Drug Global	01-Mar-21				
STUDYID (Protocol Ver.)		TEAM288			
USUBJID		STUDYID  "  SUBJID			
INDICATION		hypertension			
PHASE					
No. of SITE		20			
No. of Subject		Remark			
Enroll	400				
Random	300	✓/Randomized subject only			
Exposure to IP	300				
Follow-up	none				
STRATIFICATION					
A					
B					
TRIAL INFORMATION			ARM		
TAETORD	ETCD	ELEMENT	ARMCD	ARM	
	1	SCRN	1	CT	
	2		2	CE	
			3	CE	
			4	CE	
			5	AT	
			6	AT	
			7	BT	
			8	BT	
ELEMENT			ELEMENT		
ETCD	ELEMENT	EPOCH			
		SCREENING			
		TREATMENT			
		TREATMENT			
		TREATMENT			
		TREATMENT			
		TREATMENT			
		TREATMENT			
		TREATMENT			
VISIT			VISIT		
RFXENDTC	MAX(DVOTC)	Equal to the latest value of EXENDTC (or the latest value of EXSTDTC if EXENDTC was not collected or is missing)	N		
RFICDTC	ICDTC	Date/time of informed consent	N		
RFPENDTC	DSDTC	Date/time when subject ended participation or follow-up in a trial	Y		Last date of all records for subject
					Remark
					Date/time when subject was first exposed to study treatment #if the subject has fall to SCREENING then NULL
					Date/time of last exposure to study treatment #if the subject has fall to SCREENING then NULL

- It is important to share a summary of the version and study information to be used within the team during the initial planning stage.



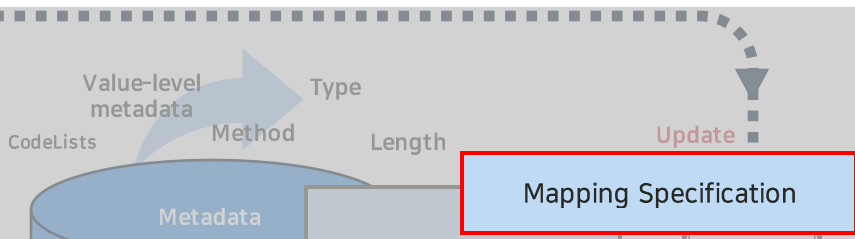
# Data Conversion Process



# Data Conversion Process

SDTM Domain & Variable Selection

Mapping Specification

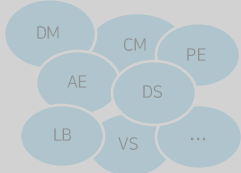


Standard version CHECK

- SDTM v1.7 / SDTMIG v3.3
- Define-XML v2.1
- SDTM CT yyyy-mm-dd



Source Data



Reference data  
(e.g. lab reference data)

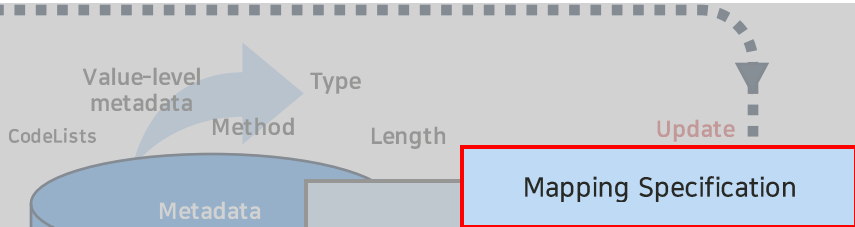
External Data

Version	Variable Order	Class	Dataset Name	Variable Name	Variable Label	Type	CDISC CT Codelist Code(s)/Described Value Domain(s)	CDISC Submission Value	Codelist Extensible (Yes/No)	Core	USED (Y/N)	Origin Type	Condition
SDTMIG v3.3	1	Findings	VS	STUDVID	Study Identifier	Char				Req	Y		
SDTMIG v3.3	2	Findings	VS	DOMAIN	Domain Abbreviation	Char				Req	Y		"VS"
SDTMIG v3.3	3	Findings	VS	USUBJID	Unique Subject Identifier	Char				Req	Y		
SDTMIG v3.3	4	Findings	VS	VSSEQ	Sequence Number	Num				Req	Y		
SDTMIG v3.3	5	Findings	VS	VSGRPID	Group ID	Char				Perm	N		
SDTMIG v3.3	6	Findings	VS	VSSPID	Sponsor-Defined Identifier	Char				Perm	N		
SDTMIG v3.3	7	Findings	VS	VSTESTCD	Vital Signs Test Short Name	Char	C66741	VSTESTCD	Yes	Req	Y		
SDTMIG v3.3	8	Findings	VS	VSTEST	Vital Signs Test Name	Char	C67153	VSTEST	Yes	Req	Y		
SDTMIG v3.3	9	Findings	VS	VSCAT	Category for Vital Signs	Char				Perm	N		
SDTMIG v3.3	10	Findings	VS	VSSCAT	Subcategory for Vital Signs	Char				Perm	N		
SDTMIG v3.3	11	Findings	VS	VSPPOS	Vital Signs Position of Subject	Char	C71148	POSITION	Yes	Perm	Y		
SDTMIG v3.3	12	Findings	VS	VSORRES	Result or Finding in Original Units	Char				Exp	Y		
SDTMIG v3.3	13	Findings	VS	VSORRESU	Original Units	Char	C66770	VSRESU	Yes	Exp	Y		
SDTMIG v3.3	14	Findings	VS	VSTRES	Character Result/Finding in Std Format	Char				Exp	Y		
SDTMIG v3.3	15	Findings	VS	VSTRESN	Numeric Result/Finding in Standard Units	Num				Exp	Y		
SDTMIG v3.3	16	Findings	VS	VSTRESU	Standard Units	Char	C66770	VSRESU	Yes	Exp	Y		
SDTMIG v3.3	17	Findings	VS	VSTAT	Completion Status	Char	C66789	ND	No	Perm	Y		
SDTMIG v3.3	18	Findings	VS	VSREASND	Reason Not Performed	Char				Perm	N		
SDTMIG v3.3	19	Findings	VS	VSLOC	Location of Vital Signs Measurement	Char	C74456	LOC	Yes	Perm	Y		
SDTMIG v3.3	20	Findings	VS	VSLAT	Laterality	Char	C99073	LAT	Yes	Perm	Y		
SDTMIG v3.3	21	Findings	VS	VSLOBXFL	Last Observation Before Exposure Flag	Char	C66742	NY	No	Exp	Y		
SDTMIG v3.3	22	Findings	VS	VSBLFL	Baseline Flag	Char	C66742	NY	No	Perm	N		
SDTMIG v3.3	23	Findings	VS	VSDRVFL	Derived Flag	Char	C66742	NY	No	Perm	Y		
SDTMIG v3.3	24	Findings	VS	VISITNUM	Visit Number	Num				Exp	Y		
SDTMIG v3.3	25	Findings	VS	VISIT	Visit Name	Char				Perm	Y		
SDTMIG v3.3	26	Findings	VS	VISITDY	Planned Study Day of Visit	Num				Perm	Y		
SDTMIG v3.3	27	Findings	VS	TAETORD	Planned Order of Element within Arm	Num				Perm	Y		
SDTMIG v3.3	28	Findings	VS	EPOCH	Epoch	Char	C99079	EPOCH	Yes	Perm	Y		
SDTMIG v3.3	29	Findings	VS	VSDTC	Date/Time of Measurements	Char	ISO 8601			Exp	Y		

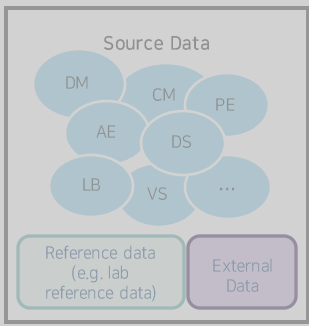
Derived/Assigned/Recorded

# Data Conversion Process

SDTM Domain & Variable Selection



- Standard version CHECK
- SDTM v1.7 / SDTMIG v3.3
  - Define-XML v2.1
  - SDTM CT yyyy-mm-dd



Version	Variable Order	Class	Dataset Name	Variable Name	Variable Label	Type	CDISC CT Codelist Code(s)/Described Value Domain(s)	CDISC Submission Value	Codelist Extensible (Yes/No)	Core	USED (Y/N)	Origin Type	Condition
SDTMIG v3.3	1	Findings	VS	STUDID	Study Identifier	Char				Req	Y		
SDTMIG v3.3	2	Findings	VS	DOMAIN	Domain Abbreviation	Char				Req	Y		
SDTMIG v3.3	3	Findings	VS	USUBJID	Unique Subject Identifier	Char				Req	Y		
SDTMIG v3.3	4	Findings	VS	VSSEQ	Sequence Number	Num				Req	Y		
SDTMIG v3.3	5	Findings	VS	VSGRPID	Group ID	Char				Perm	N		
SDTMIG v3.3	6	Findings	VS	VSSPID	Sponsor-Defined Identifier	Char				Perm	N		
SDTMIG v3.3	7	Findings	VS	VSTESTCD	Vital Signs Test Short Name	Char	C66741	VSTESTCD	Yes	Req	Y		
SDTMIG v3.3	8	Findings	VS	VSTEST	Vital Signs Test Name	Char	C67153	VSTEST	Yes	Req	Y		
SDTMIG v3.3	9	Findings	VS	VSCAT	Categ	Char				Perm	N		
SDTMIG v3.3	10	Findings	VS	VSSCAT	Subc	Char				Perm	N		
SDTMIG v3.3	11	Findings	VS	VSPPOS	Vital S	Char				Perm	N		
SDTMIG v3.3	12	Findings	VS	VSORRES	Resul	Char				Perm	N		
SDTMIG v3.3	13	Findings	VS	VSORRESU	Origin	Char				Exp	Y		
SDTMIG v3.3	14	Findings	VS	VSTSTRES	Chara	Char				Exp	Y		
SDTMIG v3.3	15	Findings	VS	VSTSTRESN	Num	Char				Exp	Y		
SDTMIG v3.3	16	Findings	VS	VSTSTRESU	Stand	Char				Exp	Y		
SDTMIG v3.3	17	Findings	VS	VSTSTAT	Comp	Char				Perm	N		
SDTMIG v3.3	18	Findings	VS	VSTREASND	Reaso	Char				Perm	N		
SDTMIG v3.3	19	Findings	VS	VSLLOC	Locat	Char				Perm	N		
SDTMIG v3.3	20	Findings	VS	VSLAT	Latera	Char				Perm	N		
SDTMIG v3.3	21	Findings	VS	VSLBFL	Last O	Char				Perm	N		
SDTMIG v3.3	22	Findings	VS	VSLBFL	Basell	Char				Perm	N		
SDTMIG v3.3	23	Findings	VS	VSDRVFL	Deriv	Char				Perm	N		
SDTMIG v3.3	24	Findings	VS	VISITNUM	Visit N	Char				Perm	N		
SDTMIG v3.3	25	Findings	VS	VISIT	Visit	Char				Perm	N		
SDTMIG v3.3	26	Findings	VS	VISITDY	Plann	Char				Perm	N		
SDTMIG v3.3	27	Findings	VS	TAETORD	Plann	Char				Perm	N		
SDTMIG v3.3	28	Findings	VS	EPOCH	Epoch	Char				Perm	Y		
SDTMIG v3.3	29	Findings	VS	VSDTC	Date/	Char				Exp	Y		

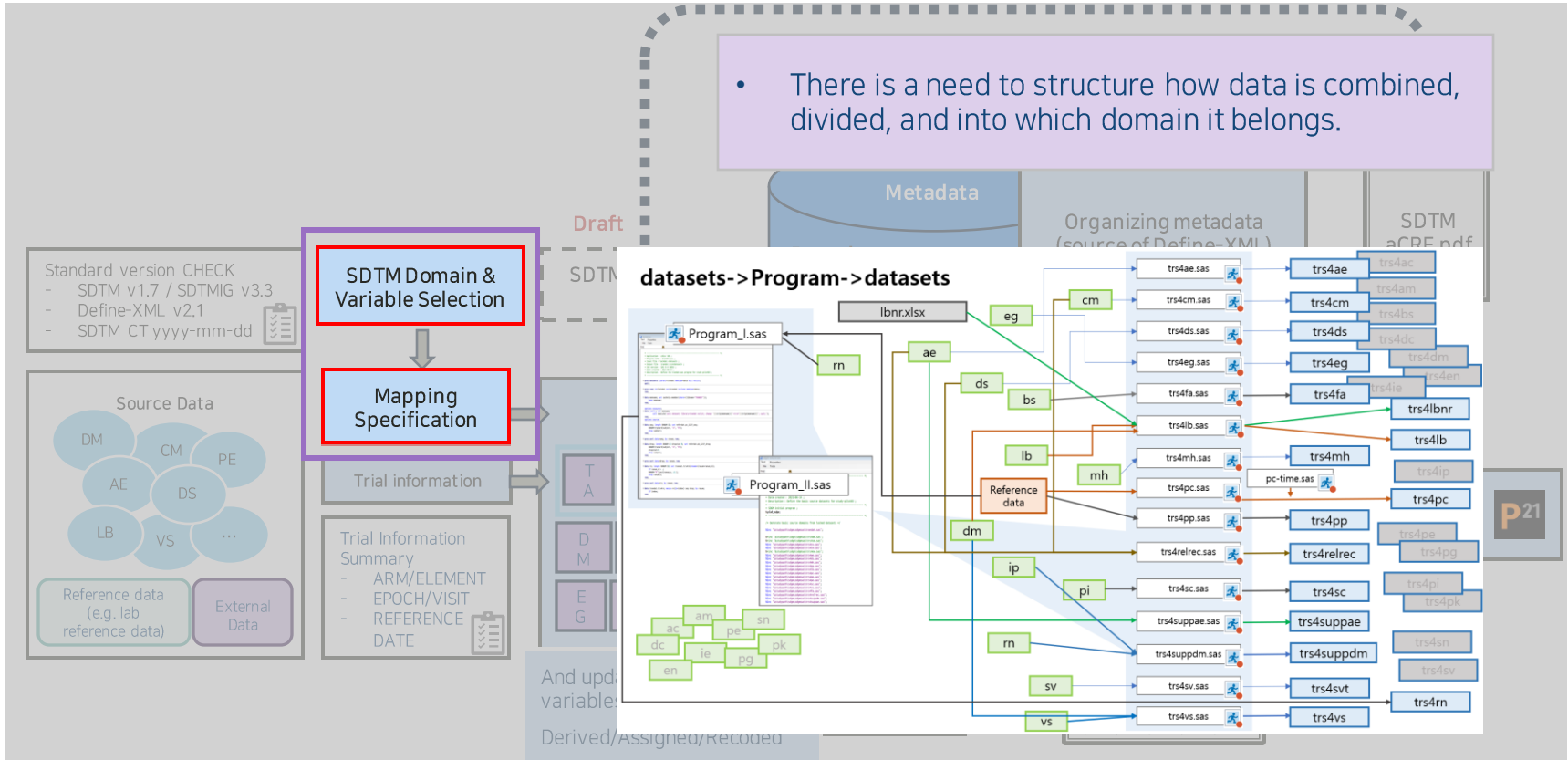
SDTM annotated CRF

- Collected
- Derived
- Assigned
- Protocol
- ..

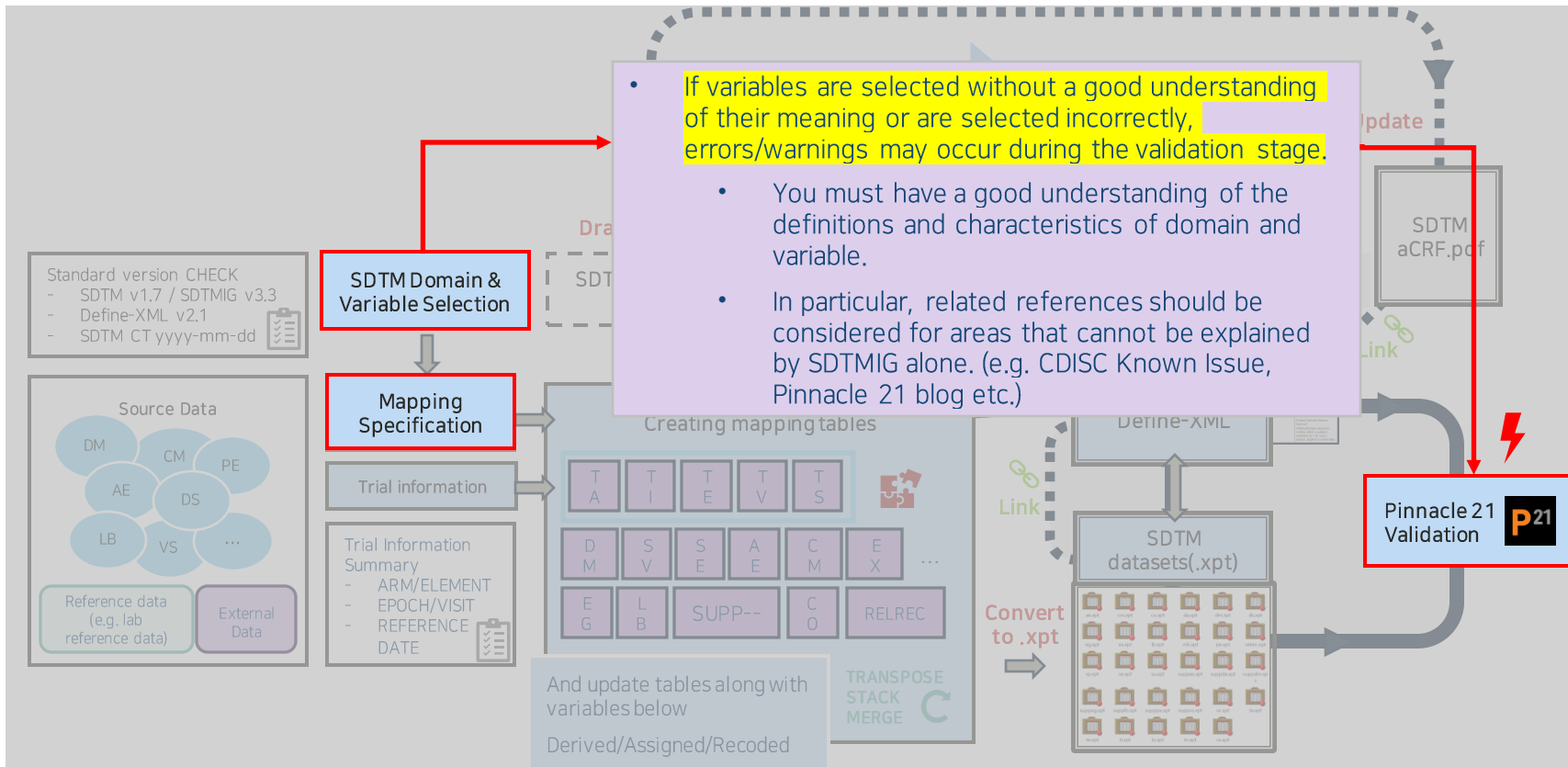
Describe how they should be programmed

# Data Conversion Process

- There is a need to structure how data is combined, divided, and into which domain it belongs.



# Data Conversion Process



- If variables are selected without a good understanding of their meaning or are selected incorrectly, errors/warnings may occur during the validation stage.
  - You must have a good understanding of the definitions and characteristics of domain and variable.
  - In particular, related references should be considered for areas that cannot be explained by SDTMIG alone. (e.g. CDISC Known Issue, Pinnacle 21 blog etc.)

# Case I.

## Unexpected value in IDVAR/IDVARVAL

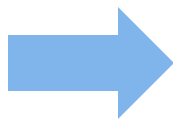
### Core

- Required
- Expected
- Permissible

When creating the SUPPDM dataset, the IDVAR and IDVARVAL variables were included because their core were Expected, but an error occurred in the validation report for this part.

### SDTMIG v3.3

Class	Dataset Name	Variable Name	Variable Label	Type	Core
Relationship	SUPPQUAL	IDVAR	Identifying Variable	Char	Exp
Relationship	SUPPQUAL	IDVARVAL	Identifying Variable Value	Char	Exp



Pinnacle 21 Validator Report		
Issue Summary		
Source	Pinnacle 21 ID	Message
SUPPDM	<a href="#">SD1051</a>	Unexpected value for IDVAR/IDVARVAL

## Case I.

# Unexpected value in IDVAR/IDVARVAL

### 8.4.1 Supplemental Qualifiers – SUPP-- Datasets [SDTMIG v3.3]

A record in a SUPP-- dataset relates back to its parent record(s) via the key identified by the STUDYID, RDOMAIN, USUBJID, and IDVAR/IDVARVAL variables. An exception is SUPP-- dataset records that are related to Demographics (DM) records, where both IDVAR and IDVARVAL will be null because the key variables STUDYID, RDOMAIN, and USUBJID are sufficient to identify the unique parent record in DM (DM has one record per USUBJID).

Exceptionally, in the SUPP-dataset for DM, STUDYID, RDOMAIN, and USUBJID are key variables, so IDVAR and IDVARVAL are null.

Solution: When creating a SUPP- dataset relating information to the DM domain, keep IDVAR and IDVARVAL blank!

## Case II.

# Missing DSDTC, when DSDY is presented

- DSDTC stands for collection date for disposition event. However, since this information is not confirmed in the original data, we decided not to include the DSDTC variable in the DS dataset.
- DSDY is a variable calculated as an integer day based on DSDTC. Since DSDTC is not collected, DSDY is all Null, but since the core is Expected, it must be included in the DS dataset.

SDTMIG v3.3					
Class	Dataset Name	Variable Name	Variable Label	Type	Core
Events	DS	DSDTC	Date/Time of Collection	Char	Perm
Events	DS	DSSTDC	Start Date/Time of Disposition Event	Char	Exp
	DS	DSDY	Study Day of Collection	Num	Exp



Pinnacle 21 Validator Report		
Issue Summary		
Source	Pinnacle 21 ID	Message
DS	<u>SD2270</u>	Missing DSDTC variable, when DSDY variable is present



## Case II.

# Missing DSDTC, when DSDY is presented

 Known Issues

### DSDY should be "Permissible", not "Expected"

<b>Short Name</b>	DSDY should be permissible, not expected
<b>Affected Standard</b>	SDTMIG v3.3
<b>Description of Error</b>	In the DS domain specification, the core value of DSDTC is "Perm" but the core value of the corresponding study day variable, DSDY, is "Exp". <u>The study day variable corresponding to a date/time variable that is "Perm" cannot be "Exp".</u> Note that DSDY was not included in the DS domain specification
<b>Efforts to Correct Error</b>	The core value for DSDY will be changed to be "Perm" in the ne

DSDY is a study day variable that corresponds to a date/time variable whose core is Permissible, so its core cannot be Expected.

#### Coping Strategy

Explain any validation errors or warnings in the Clinical Study Data Reviewers Guide (cSDRG).

Published Attribute Value


Revised Attribute Value

DS.DSDY

Solution: Describe this in cSDRG (Clinical Study Data Reviewers Guide)

## Case III.

# Treatment Emergent Adverse Event(TEAE) flag?



Pinnacle 21 Validator Report		
Issue Summary		
Source	Pinnacle 21 ID	Message
AE		
	<a href="#">SD1097</a>	No Treatment Emergent info for Adverse Event

- Among the variables in the AE domain, there is no variable that can display information about Treatment Emergent.
- Additionally, there is no guideline in SDTMIG that states that it is mandatory to include whether the AE is a treatment emergent.

## Case III.

# Treatment Emergent Adverse Event(TEAE) flag?


Hello,

There is an error in AE domain: No Treatment Emergent info for Adverse Event.

The description of rule 1097 states, it should be included in SUPPAE. But there is no imputation done at sdtm level, so how can one create emergent flag?

Inquiries and answers on the forum board of the Pinnacle 21 website

[No Treatment Emergent info for Adverse Event](#)

 Jozef on March 24, 2017

There SHOULD not be any imputed data in SDTM, but the sad reality is different ...

This is a requirement added by the FDA on top of what CDISC has as requirements.

According to the CDISC principles, "treatment emergent flags" should only appear in ADaM, but the FDA also wants it in SDTM.

Ref. <https://www.pinnacle21.com/forum/no-treatment-emergent-info-adverse-event>

## Case III.

# Treatment Emergent Adverse Event(TEAE) flag?

### FDA Validator rules v1.6

FDA Validator Rule ID	Publisher	Publisher ID	FDA Validator Rule Message
SD1321	FDA	FDAB001	No Treatment Emergent info for Adverse Event (missing SUPPAE)

### FDA Validator Rule Description

A treatment-emergent flag should be included in SUPPAE according to the FDA business rule that states, 'A treatment-emergent flag should be submitted.'



Although it is not included in the SDTM principles, a treatment-emergent flag must be included in the SUPPAE as specified in FDA business rules.

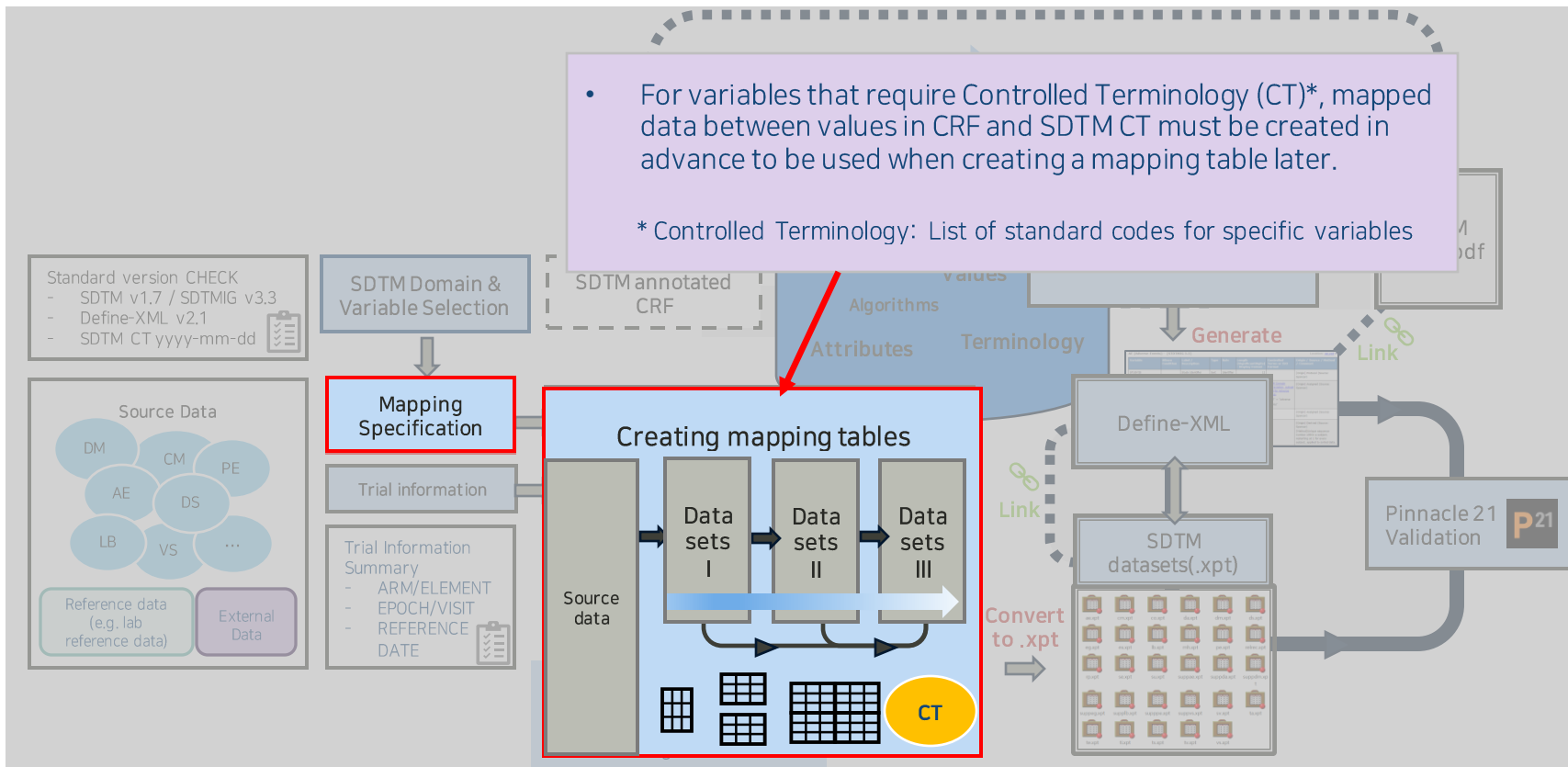
*Version 1.5 finalized June 2019*

FDA Business Rule ID	FDA Business Rule
FDAB001	A treatment-emergent flag should be submitted.

# Data Conversion Process

- For variables that require Controlled Terminology (CT)\*, mapped data between values in CRF and SDTM CT must be created in advance to be used when creating a mapping table later.

\* Controlled Terminology: List of standard codes for specific variables

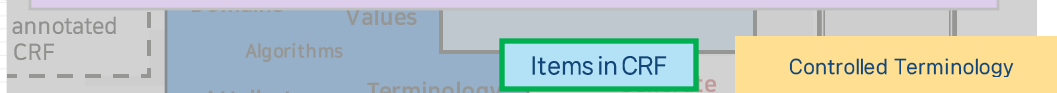


# Data Conversion Process

- For variables that require Controlled Terminology (CT)\*, mapped data between values in CRF and SDTM CT must be created in advance to be used when creating a mapping table later.

\* Controlled Terminology: The set of codelists and valid value used with data items within CDISC-defined datasets.

Code	Submission Value	Laboratory Test Code	Submission Value
C85047	Yes	Laboratory Test Code	LBTESTCD
C100429	C65047	Laboratory Test Code	
C181404	C65047	Laboratory Test Code	
C80167	C65047	Laboratory Test Code	
C186022	C65047	Laboratory Test Code	
C100462	C65047	Laboratory Test Code	
C100461	C65047	Laboratory Test Code	
C80168	C65047	Laboratory Test Code	
C172524	C65047	C67154	Yes
C154761	C65047	C179752	C67154
C154759	C65047	C179754	C67154
C100430	C65047	C179753	C67154
C189527	C65047	C132370	C67154
C184426	C65047	C124334	C67154
C111124	C65047	C154732	C67154
C150835	C65047	C163497	C67154
C150934	C65047	C103344	C67154
C125939	C65047	C186042	C67154
C135397	C65047	C186045	C67154
C184527	C65047	C186043	C67154
C74699	C65047	C186063	C67154
C74633	C65047	C186064	C67154
C80169	C65047	C186065	C67154
C135388	C65047	C186066	C67154
C82247	C65047	C186064	C67154
C147288	C65047	C186069	C67154
C74638	C65047	C186073	C67154
C96580	C65047	C186074	C67154
C96559	C65047	C142293	C67154
C80163	C65047	C186065	C67154
C147289	C65047	C186070	C67154
C189522	C65047	C147370	C67154
C103248	C65047	C186075	C67154
C186076	C67154	C186076	C67154
C186067	C67154	C186067	C67154
C186068	C67154	C186068	C67154



Domain	SEQ	TEST	TESTCD	TEST
LB	1	WBC	WBC	Leukocytes
LB	2	RBC	RBC	Erythrocytes
LB	3	Hemoglobin	HGB	Hemoglobin
LB	4	Hematocrit	HCT	Hematocrit
LB	5	Platelet	PLAT	Platelets
LB	6	Neutrophil	NEUT	Neutrophils
LB	7	Lymphocyte	LYM	Lymphocytes
LB	8	Monocyte	MONO	Monocytes
LB	9	Eosinophil	EOS	Eosinophils
LB	10	Basophil	BASO	Basophils
LB	11	AST	AST	Aspartate Aminotransferase
LB	12	ALT	ALT	Alanine Aminotransferase
LB	13	ALP	ALP	Alkaline Phosphatase
LB	14	γ-GTP	GGT	Gamma Glutamyl Transferase

Class	Dataset Name	Variable Name	Variable Label	Type	CDISC CT Codelist Code(s)/Described Value Domain(s)	CDISC Submission Value	Codelist Extensible (Yes/No)
Findings	LB	LBTESTCD	Lab Test or Examination Short Name.	Char	C65047	LBTESTCD	Yes
Findings	LB	LBTEST	Lab Test or Examination Name	Char	C67154	LBTEST	Yes

And update tables along with variables below  
Derived/Assigned/Recorded



# The importance of accurate CT mapping

There may be simple mappings such as gender (Male - M, Female - F),

but In the case of temperature measurement area,

## TYMPANIC MEMBRANE

Among the dosage forms, tablets vary as follows.

- TABLET
- TABLET, FILM COATED
- TABLET, CHEWABLE
- TABLET, EXTENDED RELEASE
- TABLET, MULTILAYER
- TABLET, ORALLY DISINTEGRATING ...

Since the code classification criteria can be very detailed or broad depending on the nature of the item, finding the correct terminology and reflecting it in the data to prevent data from being distorted or lost during the mapping process can lead to improved quality of results.



# CT mapping takes more time than you think

CT search

Code list Code or Name: RACE 검색

Extensible	Code	CDISC Submission Value	NCI Preferred Term
No	C74457	RACE	CDISC SDTM Race Terminology
	C41259	AMERICAN INDIAN OR ALASKA NATIVE	American Indian or Alaska Native
	C41260	ASIAN	Asian
	C16352	BLACK OR AFRICAN AMERICAN	Black or African American
	C41219	NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	Native Hawaiian or Other Pacific Islander
	C43234	NOT REPORTED	Not Reported
	C17938	UNKNOWN	Unknown
	C41261	UNKNOWN	White
	C128689	WHITE	CDISC SDTM Collected Race Terminology
	C154875	RACEC	Aboriginal Australian
	C128937	ABORIGINAL AUSTRALIAN	African American
	C41226	AFRICAN AMERICAN	African Caribbean
	C42331	AFRICAN CARIBBEAN	African
	C18207	AFRICAN	Alaska Native
	C41259	AFRICAN	American Indian or Alaska Native
	C43877	ALASKA NATIVE	American Indian
	C43876	AMERICAN INDIAN OR ALASKA NATIVE	Arab
	C16310	VE	Asian American
	C41262	AMERICAN INDIAN	Asian Indian
	C41260	ARAB	

내보내기

Term	Standard and Date
SDTM CT 2022-12-18	
SDTM CT 2022-12-18	

SDTM Terminology Website


CDISC SDTM Controlled Terminology

CDISC SDTM Controlled Terminology, 2023-06-30

OID	Name	Day/Type Extension	NCI Code	CDISC Synonym	CDISC Definition	Preferred Term
CLC141857	TENM91TC	10-Meter Walk/Run Functional Test Test Code	C141857	10-Meter Walk/Run Functional Test Test Code	10-Meter Walk/Run test code	CDISC Functional Test 10-Meter Walk/Run Test Code Terminology
TENM9102	TENM91-10m Walk/Run Performed	10-Meter Walk/Run - Was the 10-meter walk/run performed?	C141706	TENM91-10m Walk/Run Performed	10-Meter Walk/Run - Was Walk/Run Performed?	10-Meter Walk/Run - Was Walk/Run Performed
TENM9103	TENM91-10m Walk/Run 10 Meters	10-Meter Walk/Run - If yes, did subject walk or run 10 meters?	C141700	TENM91-10m Walk/Run 10 Meters	10-Meter Walk/Run - Time to Walk/Run 10 Meters	10-Meter Walk/Run - Time to Walk/Run 10 Meters
TENM9104	TENM91-10m Walk/Run Offenses	10-Meter Walk/Run - If yes, did subject wear offenses?	C141701	TENM91-10m Walk/Run Offenses	10-Meter Walk/Run - Wear Offenses	10-Meter Walk/Run - Wear Offenses
CLC141856	TENM91TN	10-Meter Walk/Run Functional Test Test Name	C141856	10-Meter Walk/Run Functional Test Test Name	10-Meter Walk/Run test name	CDISC Functional Test 10-Meter Walk/Run Test Name Terminology
TENM9101	TENM91-10m Walk/Run	10-Meter Walk/Run - Test grade	C141705	TENM91-10m Walk/Run	10-Meter Walk/Run - Test Grade	10-Meter Walk/Run - Test Grade
TENM9105	TENM91-10m Walk/Run 10 Meters	10-Meter Walk/Run - If yes, did subject walk or run 10 meters?	C141700	TENM91-10m Walk/Run 10 Meters	10-Meter Walk/Run - Time to Walk/Run 10 Meters	10-Meter Walk/Run - Time to Walk/Run 10 Meters
TENM9106	TENM91-10m Walk/Run Performed	10-Meter Walk/Run - Was the 10-meter walk/run performed?	C141706	TENM91-10m Walk/Run Performed	10-Meter Walk/Run - Was Walk/Run Performed?	10-Meter Walk/Run - Was Walk/Run Performed
TENM9107	TENM91-10m Walk/Run Offenses	10-Meter Walk/Run - If yes, did subject wear offenses?	C141701	TENM91-10m Walk/Run Offenses	10-Meter Walk/Run - Wear Offenses	10-Meter Walk/Run - Wear Offenses
CLC141983	AST91TC	4-Star Assend Functional Test Test Code	C141983	4-Star Assend Functional Test Test Code	4-Star Assend test code	CDISC Functional Test 4-Star Assend Test Code Terminology
AST9101	AST91-4star Assend	4-Star Assend - Was the 4-star assend performed?	C141985	AST91-4star Assend	4-Star Assend - Was 4-Star Assend Performed?	4-Star Assend - Was 4-Star Assend Performed
AST9102	AST91-4star Assend	4-Star Assend - Time to Do 4-Star Assend	C141986	AST91-4star Assend	4-Star Assend - Time to Do 4-Star Assend	4-Star Assend - Time to Do 4-Star Assend
AST9103	AST91-4star Assend	4-Star Assend - If yes, did subject wear offenses?	C141987	AST91-4star Assend	4-Star Assend - Wear Offenses	4-Star Assend - Wear Offenses
AST9104	AST91-4star Assend	4-Star Assend - Test grade	C141988	AST91-4star Assend	4-Star Assend - Test Grade	4-Star Assend - Test Grade
CLC141982	AST91TN	4-Star Assend Functional Test Test Name	C141982	4-Star Assend Functional Test Test Name	4-Star Assend test name	CDISC Functional Test 4-Star Assend Test Name Terminology

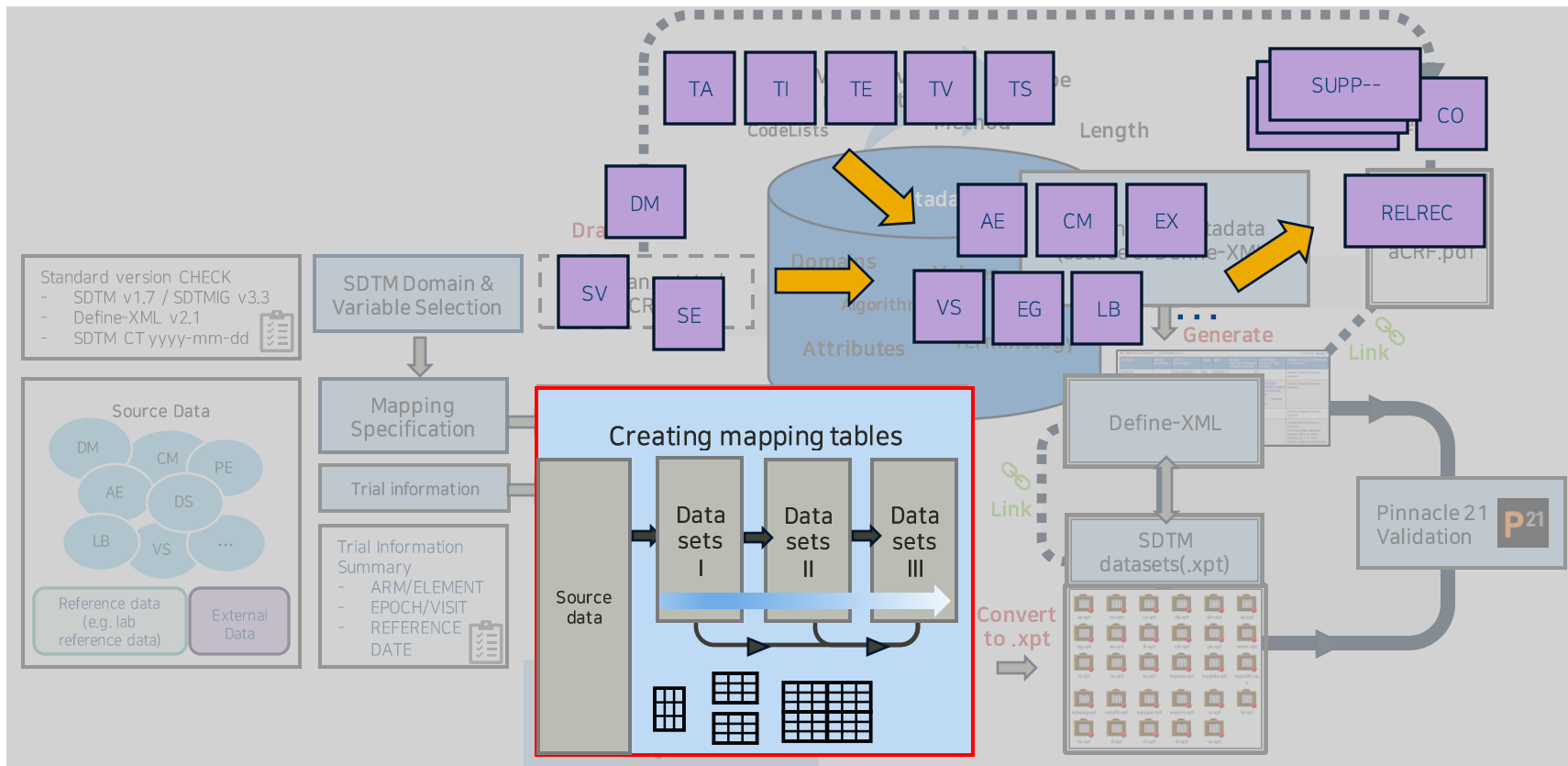
• Maintain and Utilize CT-mapped library

CRF ~ DB specification(Data dictionary)-->> Include CT!!





# Data Conversion Process



# Consider connectivity between tables

## EC domain

Domain Abbreviation	Unique Subject Identifier	Sponsor Device Identifier	Sequence Number	Name of Treatment	Pre-Specified	Occurrence
EC	CDISC009	DEV0009	122	PLACEBO	Y	N
EC	CDISC009	DEV0009	123	PLACEBO	Y	N
EC	CDISC009	DEV0009	124	PLACEBO	Y	N
EC	CDISC009	DEV0009	125	PLACEBO	Y	N
EC	CDISC009	DEV0009	126	PLACEBO	Y	N
EC	CDISC009	DEV0009	127	PLACEBO	Y	N
EC	CDISC009	DEV0009	128	PLACEBO	Y	N
EC	CDISC009	DEV0009	129	PLACEBO	Y	Y
EC	CDISC009	DEV0009	130	PLACEBO	Y	Y
EC	CDISC009	DEV0009	131	PLACEBO	Y	Y
EC	CDISC009	DEV0009	132	PLACEBO	Y	Y
EC	CDISC009	DEV0009	133	PLACEBO	Y	Y
EC	CDISC009	DEV0009	134	PLACEBO	Y	Y

## SUPPEC domain

Related Domain Abbreviation	Unique Subject Identifier	Identifying Variable	Identifying Variable Value	Qualifier Variable Name	Qualifier Variable Label	Data Value
EC	CDISC009	ECSEQ	122	ECREAS	Reason for Occur Value	INVESTIGATOR DECISION
EC	CDISC009	ECSEQ	123	ECREAS	Reason for Occur Value	INVESTIGATOR DECISION
EC	CDISC009	ECSEQ	124	ECREAS	Reason for Occur Value	INVESTIGATOR DECISION
EC	CDISC009	ECSEQ	125	ECREAS	Reason for Occur Value	INVESTIGATOR DECISION
EC	CDISC009	ECSEQ	126	ECREAS	Reason for Occur Value	INVESTIGATOR DECISION
EC	CDISC009	ECSEQ	127	ECREAS	Reason for Occur Value	INVESTIGATOR DECISION
EC	CDISC009	ECSEQ	128	ECREAS	Reason for Occur Value	INVESTIGATOR DECISION

## EC page (aCRF)

### EC (Exposure as Collected)

**EXPOSURE**

Did the subject take Zanomaline?  Yes  No ECOCUR

If No, Reason Not Administered  ECREASOC in SUPPEC

Date  ECSTDC ECENDTC

Lot Number  ECLOT

Dose ECDOSE 5 mL ECDOSU

- The order of creating tables
- key variable

# Consider how the table structure changes

TEMP	PULSE	SYSBP	DIABP
97.4	51	137	71



VSTESTCD	VSORRES
TEMP	97.4
PULSE	51
SYSBP	137
DIABP	71

**VITAL SIGNS** [NOT SUBMITTED]

Vital Signs Collected?  Yes  No VSSTAT = NOT DONE when VSTESTCD = VSALL

Visit: [dropdown] VISIT

Date: [dropdown] [dropdown] [dropdown] VSDTC

**VSTESTCD = WEIGHT:** Weight [VSORRES] pounds [VSORRESU]

**VSTESTCD = HEIGHT:** Height [VSORRES] inches [VSORRESU]

**VSTESTCD = TEMP:** Temperature [VSORRES] F [VSORRESU]

**VSPOS:** Pulse and Blood Pressure (Supine)

**VSTESTCD = PULSE:** Pulse [VSORRES] bpm [VSORRESU]

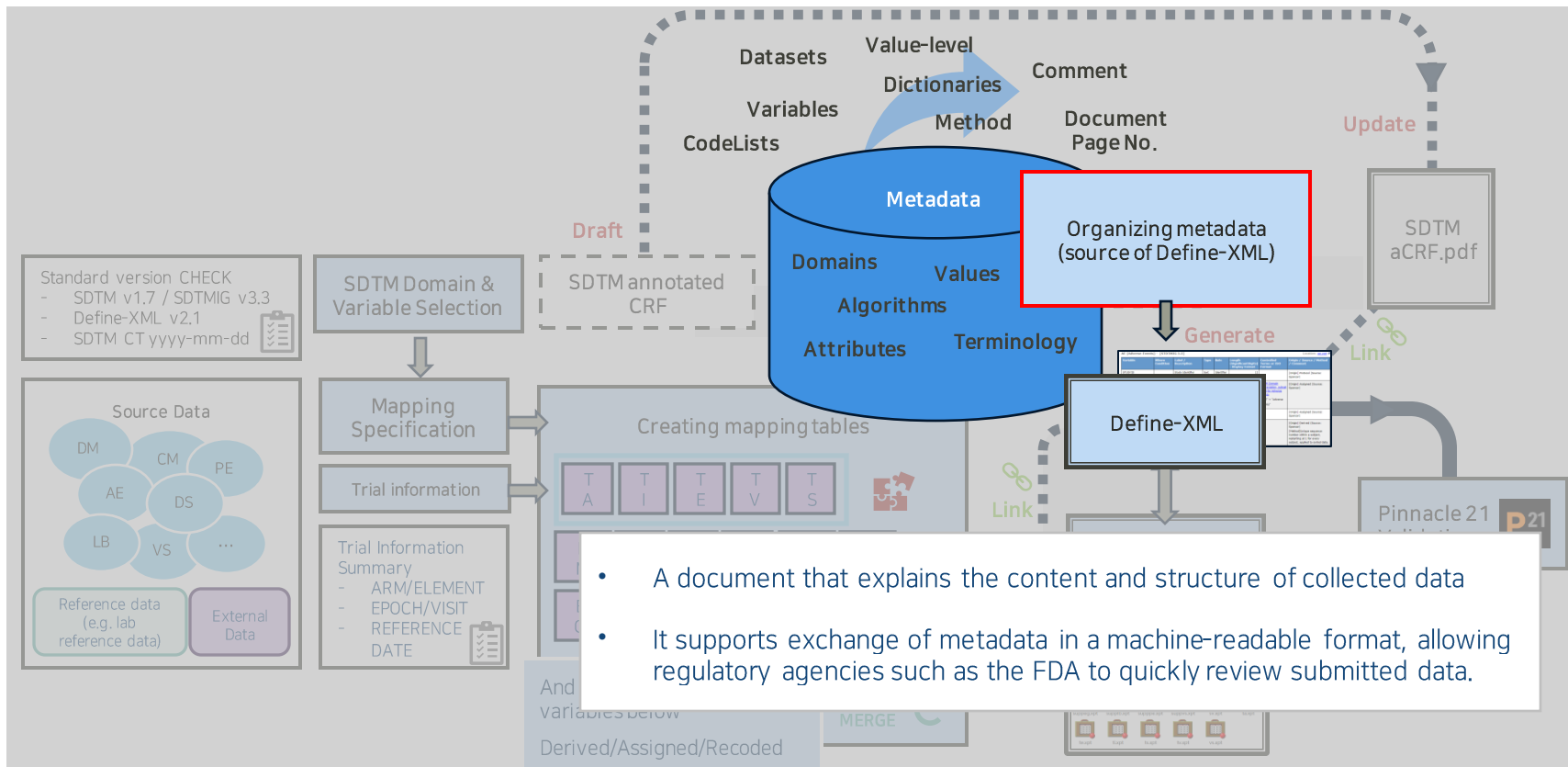
**VSTESTCD = SYSBP:** Systolic [VSORRES] mmHg [VSORRESU]

**VSTESTCD = DIABP:** Diastolic [VSORRES] mmHg [VSORRESU]

Unique Subject Identifier	Sequence Number	Vital Signs Test Short Name	Vital Signs Test Name	Vital Signs Position of Subject	Result or Finding in Original Units	Original Units
CDISC00	1	DIABP	Diastolic Blood Pressure	STANDIN	71	mmHg
CDISC00	15	HEIGHT	Height		71.5	in
CDISC00	16	PULSE	Pulse Rate	STANDIN	51	beats/min
CDISC00	30	SYSBP	Systolic Blood Pressure	STANDIN	137	mmHg
CDISC00	44	TEMP	Temperature		97.4	F
CDISC00	52	WEIGHT	Weight		173.5	LB
CDISC00	2	DIABP	Diastolic Blood Pressure	STANDIN	71	mmHg
CDISC00	17	PULSE	Pulse Rate	STANDIN	50	beats/min
CDISC00	31	SYSBP	Systolic Blood Pressure	STANDIN	137	mmHg
CDISC00	45	TEMP	Temperature		97.2	F
CDISC00	3	DIABP	Diastolic Blood Pressure	STANDIN	83	mmHg
CDISC00	18	PULSE	Pulse Rate	STANDIN	50	beats/min
CDISC00	32	SYSBP	Systolic Blood Pressure	STANDIN	130	mmHg
CDISC00	46	TEMP	Temperature		98.0	F

**Stack!**

# Data Conversion Process



# Define-XML

Machine-readable format

```

</ItemDef>
<ItemDef OID="IT.EX.DOMAIN" Name="DOMAIN" DataType="text" Length="2" SASFieldName="DOMAIN">
  <Description>
    <TranslatedText xml:lang="en">Domain Abbreviation</TranslatedText>
  </Description>
  <CodeListRef CodeListOID="CL.EX.DOMAIN"/>
  <def:Origin Type="Assigned" Source="Sponsor"/>
</ItemDef>
<ItemDef OID="IT.EX.EXDOSE" Name="EXDOSE" DataType="integer" Length="2">
  <Description>
    <TranslatedText xml:lang="en">Dose per Administration</TranslatedText>
  </Description>
  <def:Origin Type="Derived" Source="Sponsor"/>
</ItemDef>
<ItemDef OID="IT.EX.EXDOSFRM" Name="EXDOSFRM" DataType="text" Length="7">
  <Description>
    <TranslatedText xml:lang="en">Dose Form</TranslatedText>
  </Description>
  <CodeListRef CodeListOID="CL.FRM"/>
  <def:Origin Type="Predecessor" Source="Sponsor">
    <Description>
      <TranslatedText xml:lang="en">EC.ECDOSFRM</TranslatedText>
    </Description>
    <def:DocumentRef leafID="LF.acrf">
      <def:PDFPageRef PageRefs="20" Type="PhysicalRef"/>
    </def:DocumentRef>
  </def:Origin>
</ItemDef>
<ItemDef OID="IT.EX.EXDOSU" Name="EXDOSU" DataType="text" Length="2" SASFieldName="EXDOSU">
  <Description>
    <TranslatedText xml:lang="en">Dose Units</TranslatedText>
  </Description>
  <def:Origin Type="Derived" Source="Sponsor"/>
</ItemDef>
<ItemDef OID="IT.EX.EXENDTC" Name="EXENDTC" DataType="date" SASFieldName="EXENDTC">
  <Description>
    <TranslatedText xml:lang="en">End Date/Time of Treatment</TranslatedText>
  </Description>
  <def:Origin Type="Predecessor" Source="Sponsor">
    <Description>
      <TranslatedText xml:lang="en">EC.ECSTDTTC</TranslatedText>
    </Description>
    <def:DocumentRef leafID="LF.ecsttc">
      <def:PDFPageRef PageRefs="20" Type="PhysicalRef"/>
    </def:DocumentRef>
  </def:Origin>
</ItemDef>
  
```

CRISC01\_1

- ▶ Annotated CRF #
- ▶ Supplemental Documents Standards
- ▶ Datasets
- ▶ Controlled Terminology
- ▶ Methods

Expand all VLM

Collapse all VLM

Style sheet applied

EX (Exposure) - [SDTMIG 3.1.2]

Location: [ex.xml](#) #

Variable	Label / Description	Type	Length or Display Format	Controlled Terms or ISO Format	Origin / Source / Method / Comment
STUDYID	Study Identifier	text	7		Protocol (Source: Sponsor)
DOMAIN	Domain Abbreviation	text	2	<a href="#">Domain Abbreviation (EX)</a> • "EX" = "Exposure"	Assigned (Source: Sponsor)
USUBID	Unique Subject Identifier	text	14		Derived (Source: Sponsor) Concatenation of STUDYID and SUBID
EXSEQ	Sequence Number	integer	1		Derived (Source: Sponsor) Sequential number identifying records within each USUBID in the domain.
EXTRT	Name of Actual Treatment	text	20	<a href="#">Treatment</a> • "Miracle Drug" • "Placebo"	Derived (Source: Sponsor) Derived from ARM, ARMCD
EXDOSE	Dose per Administration	integer	2		Derived (Source: Sponsor) If ARMCD=WONDER10 then EXDOSE = Number of Tablets per Day (QVAL where QNAM=SMNO) * 10. If ARMCD=WONDER20 then EXDOSE = Number of Tablets per Day (QVAL where QNAM=SMNO) * 20. If ARMCD=PLACEBO then EXDOSE = 0.
EXDOSU	Dose Units	text	2		Derived (Source: Sponsor) Derived from ARM, ARMCD - equal to mg
EXDOSFRM	Dose Form	text	7	<a href="#">Pharmaceutical Dosage Form</a> • "TABLET" = "tab"	Predecessor (Source: Sponsor) EC.ECDOSFRM Annotated CRF [20] #
EXSTDTTC	Start Date/Time of Treatment	date		ISO 8601	Predecessor (Source: Sponsor) EC.ECSTDTTC Annotated CRF [20] #



# Data Conversion Process

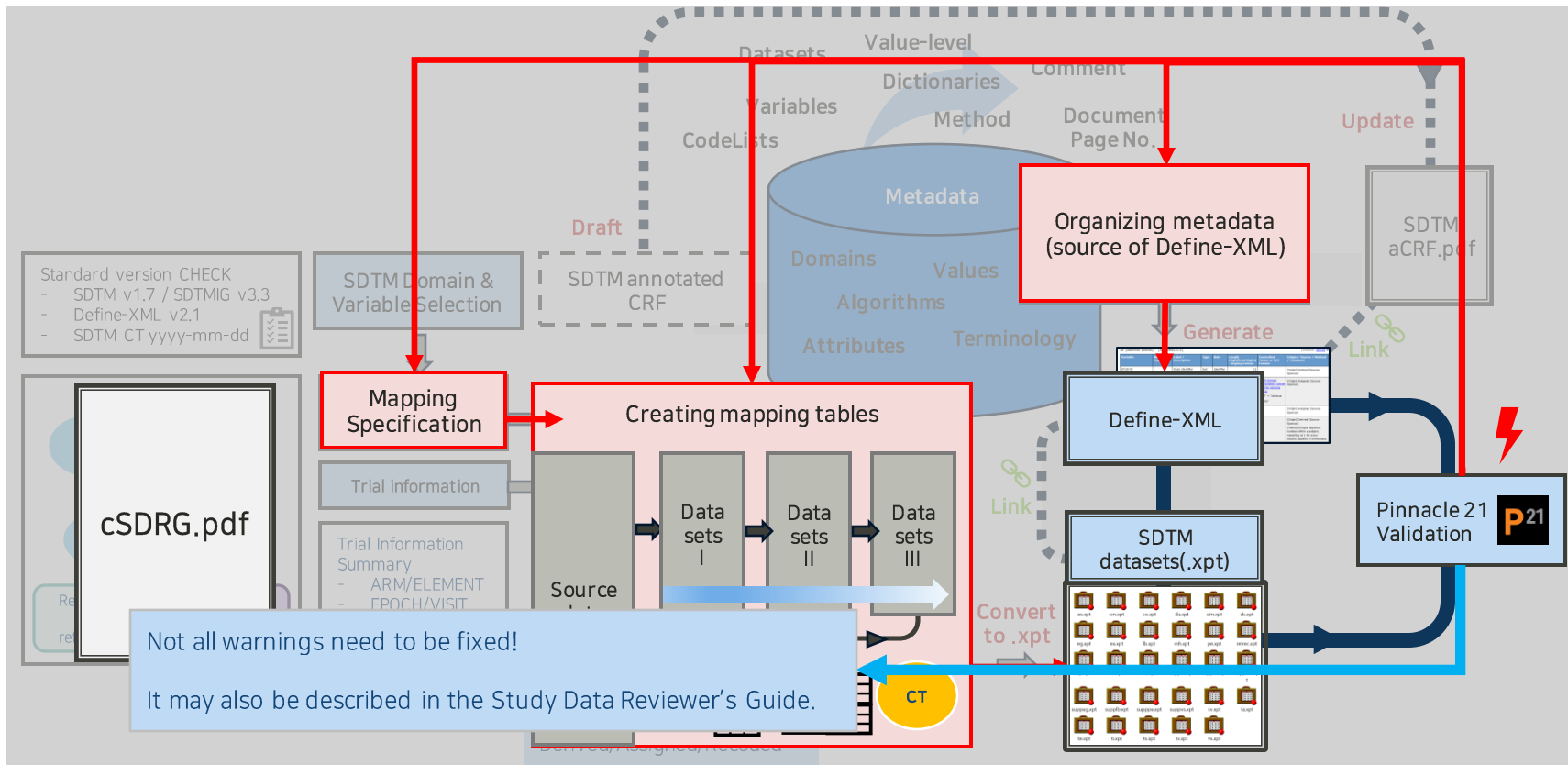
Pinnacle 21 Validator Report					
Processed Sources					
Domain	Label	Class	Source	Records	Rejects
GLOBAL	Global Metadata	--	--	0	0
AE	Adverse Events	EVENTS	ae.xpt	115	0
CM	Concomitant/Prior Medications	INTERVENTIONS	cm.xpt	1030	0
CO	Comments	SPECIAL PURPOSE	co.xpt	788	0
DA	Drug Accountability	FINDINGS	da.xpt	16108	0
DM	Demographics	SPECIAL PURPOSE	dm.xpt	413	0
DS	Disposition	EVENTS	ds.xpt	1652	0
EG	ECG Test Results	FINDINGS	eg.xpt	828	0
EX	Exposure	INTERVENTIONS	ex.xpt	1652	0
LB	Laboratory Test Results	FINDINGS	lb.xpt	37204	0
MH	Medical History	EVENTS	mh.xpt	1019	0
PE	Physical Examination	FINDINGS	pe.xpt	1466	0
RELREC	Related Records	RELATIONSHIP	relrec.xpt	442	0
RP	Reproductive System Findings	FINDINGS	rp.xpt	554	0
SE	Subject Elements	SPECIAL PURPOSE	se.xpt	826	0
SU	Substance Use	INTERVENTIONS	su.xpt	826	0
SUPPAE	Supplemental Qualifiers for AE	--	suppae.xpt	115	0
SUPPDA	Supplemental Qualifiers for DA	--	suppda.xpt	3220	0
SUPPDM	Supplemental Qualifiers for DM	--	suppdm.xpt	826	0
SUPPEG	Supplemental Qualifiers for EG	--	suppeg.xpt	308	0
SUPPLB	Supplemental Qualifiers for LB	--	supplb.xpt	3900	0
SUPPPE	Supplemental Qualifiers for PE	--	supppe.xpt	34	0
SUPPVS	Supplemental Qualifiers for VS	--	suppvs.xpt	413	0

Value-level Dictionaries  
Method  
Comment Document Page No.  
Organizing metadata (source of Define-XML)  
Values  
Terminology  
Generate  
Update  
SDTM aCRF.pdf

Pinnacle 21 Validator Report					
Issue Summary					
Source	Pinnacle 21 ID	Message	Severity	Found	
DS	<a href="#">SD2270</a>	Missing DSDTC variable, when DSDY variable is present		1	
EG	<a href="#">SD1023</a>	VISIT/VISITNUM values do not match TV domain data		2	
EX	<a href="#">SD1078</a>	Permissible variable with missing value for all records		1	
PE	<a href="#">SD0065</a>	USUBJID/VISIT/VISITNUM values do not match SV domain data		2	
PE	<a href="#">SD1023</a>	VISIT/VISITNUM values do not match TV domain data		2	
SU	<a href="#">SD1076</a>	Model permissible variable added into standard domain		1	
SUPPAE	<a href="#">CT2002</a>	QEVAl value not found in 'Evaluator' extensible codelist		115	
SUPPAE	<a href="#">SD0077</a>	Invalid referenced record		1	

# Data Conversion Process







## Lesson learned from our experiences

# 4-important things



## Understanding the Definitions correctly

And huge Resources..



## Starting with a pilot project

Experience SDTM project  
Face challenges as many as you can



## Utilizing technologies


Use software for repetitive and time-consuming tasks



## Library and Knowledge repository

Maintain inhouse library  
Update and Share Knowledge  
Learn from others' experiences

# 4-important things



**Understanding the Definitions correctly**

And huge Resources..

- CDISC
  - SDTM
  - SDTMIG, SDTMIG-MD, AP
  - SDTM-MSG
  - Conformance Rules
  - Define-XML specification..
- Regulatory agencies
  - Study Data Technical Conformance Guide
  - Study Data Standard Catalog
  - Business Rules, Validator Rules..
- Others
  - PHUSE WGs documents
    - cSDRG, SDSP packages..
  - PharmaSUG papers
  - Pinnacle 21 resources
    - Validation rules per engines, Blog, Community forum

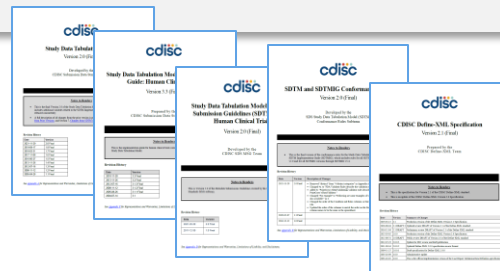
The requirements of the CDISC standard and the requirements of the regulatory agency exist separately and must be reflected when composing submission data to ensure that both are satisfied.

**FDA Guidance Documentations**

- Study Data Standards Resources
- Technical Rejection Criteria for Study Data - Incorporated into Study Data Technical Conformance Guide
- Study Data Technical Conformance Guide v4.9 (October 2022)
- FDA Data Standards Catalog v9.0
- Business Rules v1.5 (May 2019)
- Validator Rules v1.6 (December 2019)
- PHUSE Clinical Study Data Review
- PHUSE Analysis Data Reviewer's Guide
- Bioresarch Monitoring Technical
- PHUSE Bioresarch Monitoring Data

**FDA Presentations**

- FDA View Update on Study Data Submission
- Developing SEND for CBER - Collaborative Workstream Initiative
- Impact of SEND Data on FDA Review of Nonclinical Studies
- Using CDISC SEND Standardized Data in FDA Toxicology Review: The KickStart Service
- Study Data Topics at FDA/CDER
- What to look for at FDA: Pharma Industry Submissions in the PDUFA Era



# 4-important things



Understanding the Definitions correctly

And huge Resources..



**Starting with a pilot project**

Experience SDTM project  
Face challenges as many as you can



Utilizing technologies

Use software for repetitive and time-consuming tasks



Library and Knowledge repository

Maintain inhouse library  
Update and Share Knowledge  
Learn from others' experiences

- If you practice with actual data, you will become more familiar with the theory and it will be easier to understand how to use it.
- In the early stages of work, you can conduct practice by creating dummy data as a virtual study model.
  - Regular group meeting -> Sharing & discussion of results created by each person
  - Share questions or data processing skills during the table creation process.

# 4-important things



Understanding the Definitions correctly

And huge Resources..



Starting with a pilot project

Experience SDTM project  
Face challenges as many as you can



Utilizing technologies

Use software for repetitive and time-consuming tasks



Library and Knowledge repository

Maintain inhouse library  
Update and Share Knowledge  
Learn from others' experiences

- Proper use of technology and software improves efficiency compared to doing it manually. This can simplify repetitive, complex tasks and save time.
- Especially if you are new to SDTM work, applying standardization-related software to your procedures can help you quickly establish SDTM work processes.
- Sponsors can use vendors that provide SDTM conversion services.

# 4-important things



Understanding the Definitions correctly

And huge Resources..



Starting with a pilot project

Experience SDTM project  
Face challenges as many as you can



Utilizing technologies

Use software for repetitive and time-consuming tasks



Library and Knowledge repository

Maintain inhouse library  
Update and Share Knowledge  
Learn from others' experiences

- Utilizing the library enables efficient work progress while maintaining the quality of results.
- It is helpful to accumulate and manage various questions raised and new information learned while performing work.
- Accumulate and manage a Q&A log containing various inquiries regarding the entire SDTM tasks.
- For resolved content, specify an answer to the questions with references, or establish a provisional internal guidelines for parts for which no reference was found.
- It is also helpful to explore relevant experiences and share them within the team.



## On a final note..

- When looking for various examples or solutions to the numerous questions encountered while performing SDTM dataset conversion work, it was difficult to find shared data in Korea.
- In comparison, many questions about standardization work were being actively discussed on websites in other countries.
- I hope that standardization work will be activated in Korea as well, that many experienced people and experts in this field will be trained, and that an environment will be created where various practical cases can be shared and discussed actively.

**Thank you !**

[makim@datateamz.co.kr](mailto:makim@datateamz.co.kr)

**cdisc**