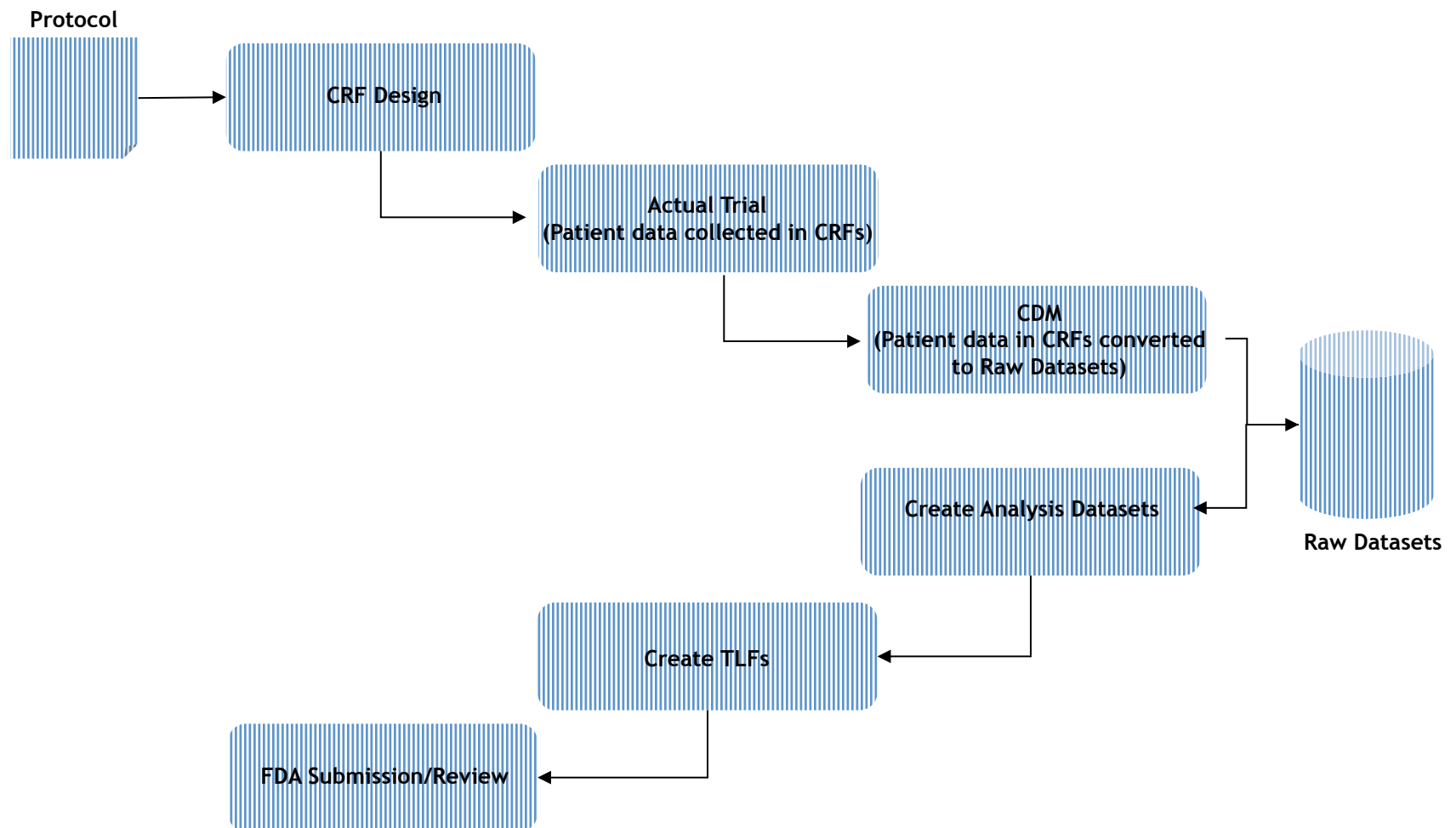


**A Brief Introduction to
CDISC – SDTM
and
Data Mapping**

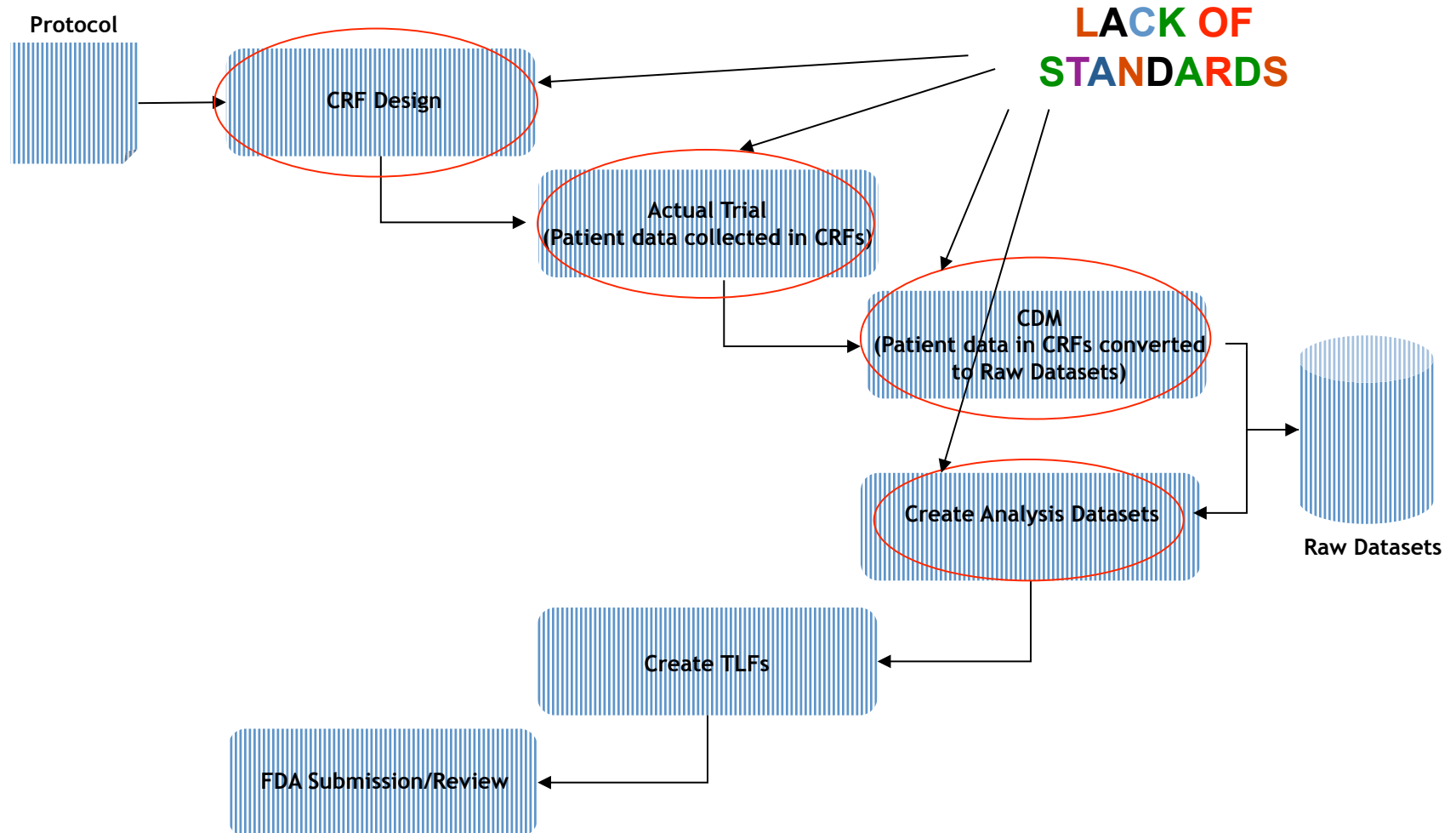
Agenda

- Flow of Clinical Trials Data
- The Problem
- Introducing CDISC
- Understanding SDTM
- Concepts of Data Mapping
- References

Flow of Clinical Trials Data



The Problem





Introducing CDISC

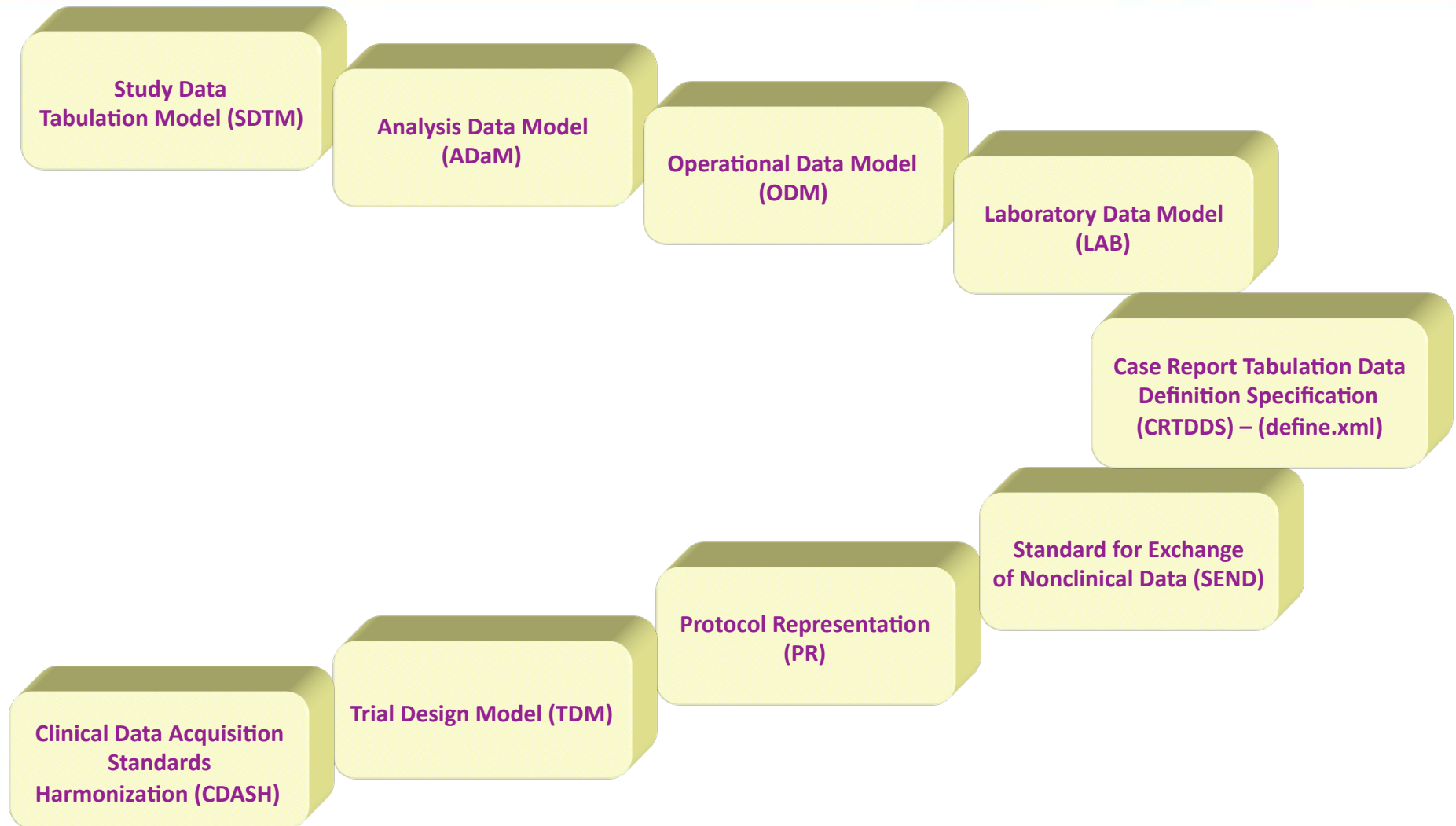
Introduction to CDISC

Non-profit organization
in 2000

Clinical Data Interchange
Standards Consortium
(CDISC)

1. Development of **industry standards** to support the
 - a) electronic acquisition
 - b) exchange
 - c) submission and
 - d) archivingof **clinical trials data**
2. Development of global, **platform independent** standards
5. Improve **data quality** and accelerate product development

CDISC Standards



Understanding SDTM

Study Data Tabulation Model (SDTM)

Operational Data Model (ODM)

Laboratory Data Model (LAB)

Case Report Tabulation Data Definition Specification (CRD) - (define.xml)

Clinical Data Acquisition Standards Harmonization (CDASH)

Trial Data

- Describes contents and structure of data collected during a clinical trial
- Purpose is to provide regulatory authority reviewers (FDA) a clear description of the structure, attributes and contents of each dataset and variables submitted as part of a product application

1. Study Data Tabulation Model V1.1
2. Study Data Tabulation Model Implementation Guide: Human Clinical Trials – V3.1.1

<http://www.cdisc.org/sdtm/>

Before and After - SDTM

Before

- ✓ Domains
- X Standard Domain Names
- X Standard Variables
- X Standard Variable Names

After

- ✓ Domains
- ✓ Standard Domain Names
- ✓ Standard Variables
- ✓ Standard Variable Names

Before and After - SDTM

Before

- Reviewers had to familiarize themselves with unique domain names, variables and variable names used in an application - **TIME CONSUMING**
- Good portion of review time spent “cleaning up the data”
- **Inefficient, error-prone**

After

- Standard Domain Names = Easy to Find Data
- Standard Variables/Variable Names = Immediate Familiarity with the Data
- Consistency
- Minimal learning curve
- **TIME EFFICIENT**

CDISC Submission Metadata Model

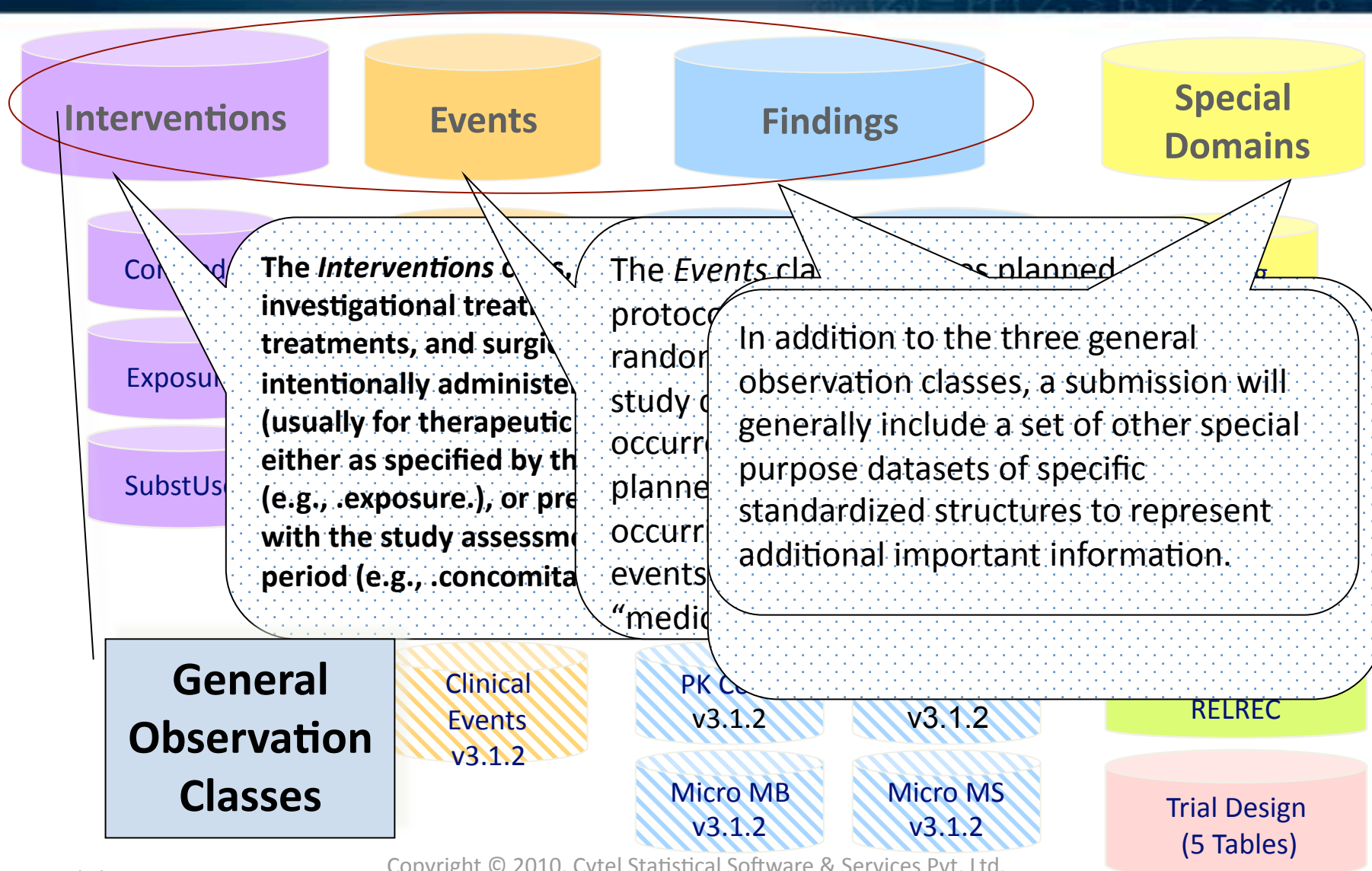
Seven distinct metadata attributes :

1. The **Variable Name**
2. A descriptive **Variable Label**, up to 40 characters
3. The **data Type**
4. The set of **controlled terminology** for the value or the presentation format of the variable (Controlled Terms or Format)

Before and After - SDTM

5. The **Origin** or source of each variable
6. The **Role** of the variable, which determines how the variable is used in the dataset.
7. **Comments** or other relevant information about the variable or its data.

Types/Classes of Domains in SDTM

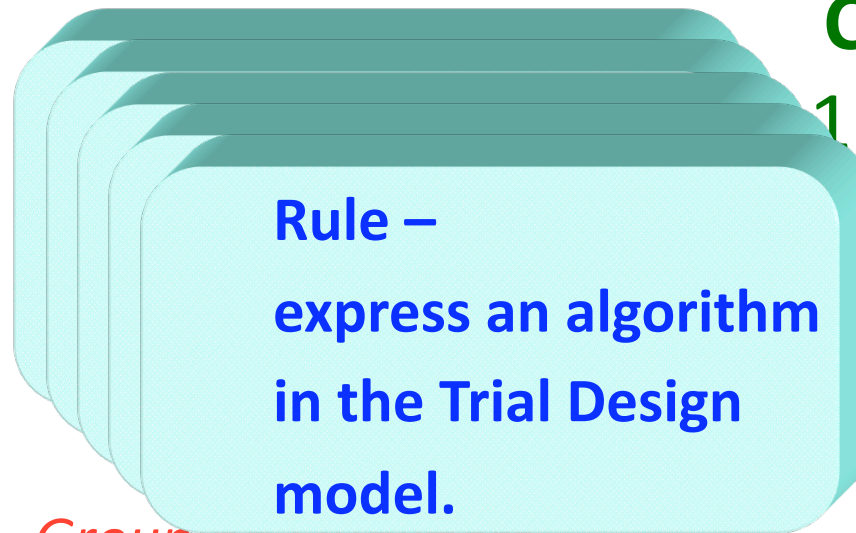


Categorization Of Variables in SDTM

Variables

Roles

1. Identifier
2. Topic
3. Timing
4. Rule
5. Qualifier -



Core/Being

1. Required
- Expected
- Permissible

Grouping Qualifiers
Result Qualifiers
Synonym Qualifiers
Record Qualifiers
Variable Qualifiers

Categorization Of Variables in SDTM

Variable Qualifiers –
are used to further modify or describe a specific variable within an observation and is only meaningful in the context of the variable they qualify.

5. E. g. --ORRESU, --ORNHI, and --ORNLO

Core/Being
. Required
. Expected
. Permissible

Record Qualifiers
Synonym Qualifiers
Record Qualifiers
Variable Qualifiers

Categorization Of Variables in SDTM

Variables

Roles

1. Identifier
2. Topic
3. Timing
4. Rule
5. Qualifier -

Permissible Variables –
– Variables may be present
in the domains –
can be included as needed

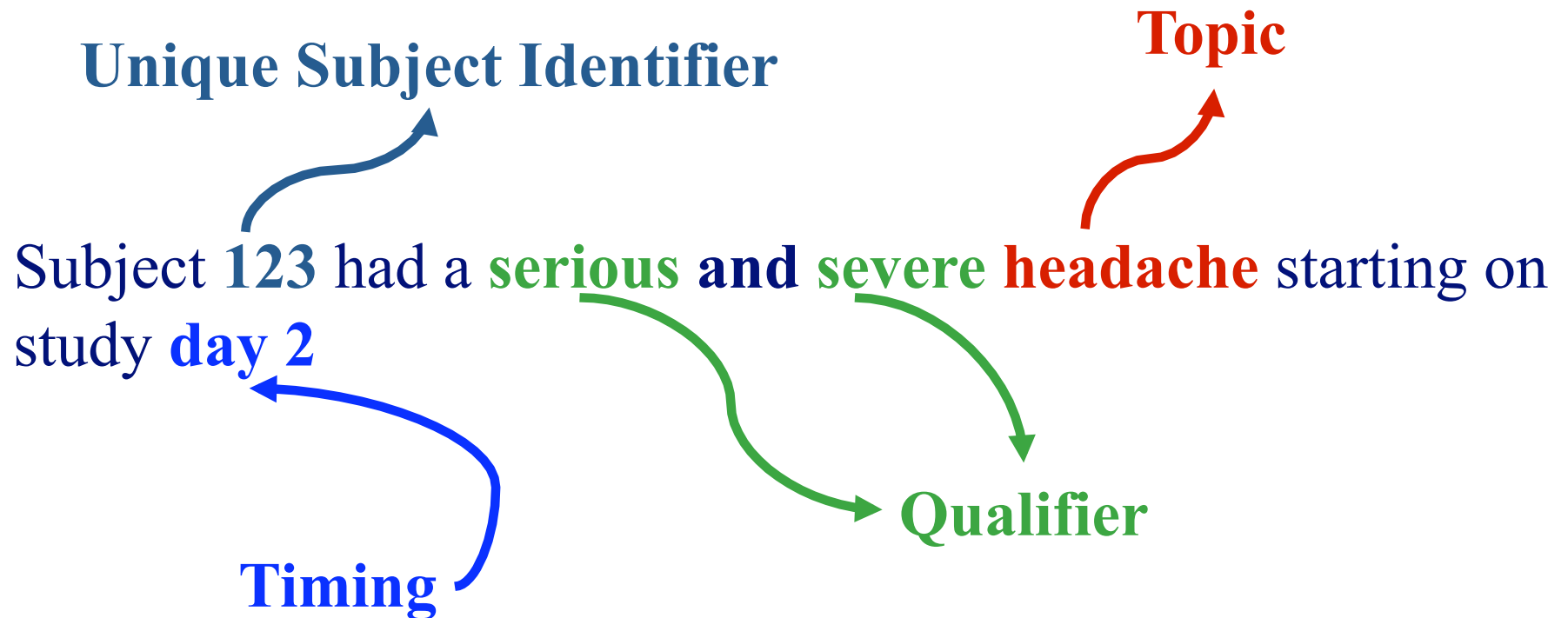
Core/Being

- . Required
- . Expected
- . Permissible

Grouping Qualifiers
Result Qualifiers
Synonym Qualifiers
Record Qualifiers
Variable Qualifiers

Categorization Of Variables in SDTM

An Example Observation



Dataset Structure

Subject Identifier

Topic

Qualifier

Timing

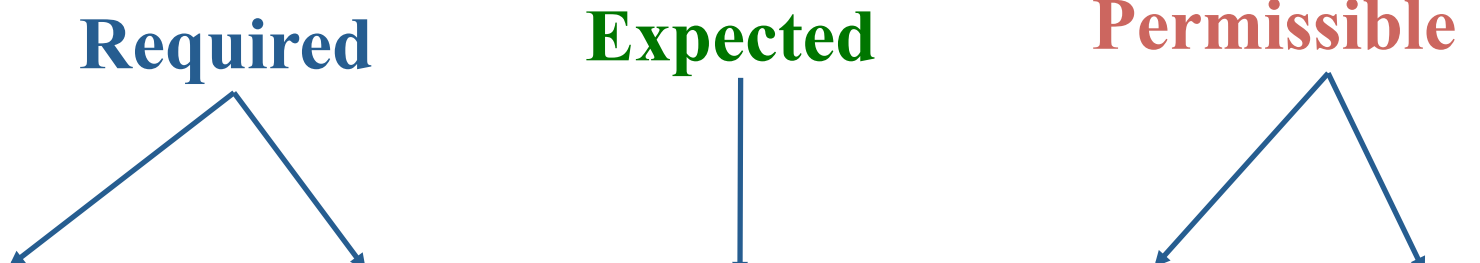
USUBJID	AETERM	AESER	AESEV	AESTDY
Unique Subject Identifier	Reported Term for the Adverse Event	Serious Event	Severity/ Intensity	Study Day of Start of Event
123	HEADACHE	YES	SEVERE	2

Dataset Structure

Required

Expected

Permissible



USUBJID	AETERM	AESER	AESEV	AESTDY
Unique Subject Identifier	Reported Term for the Adverse Event	Serious Event	Severity/ Intensity	Study Day of Start of Event
123	HEADACHE	YES	SEVERE	2

Example Events Data (MH)

USUBJID	MHSEQ	MHCAT	MHTERM	MHSTAT	MHONGO	VISITNUM
1111-0001	1	ENDOCRINE/METABOLIC MEDICAL HISTORY	HYPERTHYRO IDISM	PAST DISEASE / SYNDROME	N	-2
1111-0001	2	ENDOCRINE/METABOLIC MEDICAL HISTORY	HYPOTHYROI DISM	PAST DISEASE / SYNDROME	N	-2
1111-0001	3	RESPIRATORY MEDICAL HISTORY	ASTHMA	PRESENT DISEASE / SYNDROME	Y	-2

Identifiers

Topic

Timing

Qualifiers

Domains –

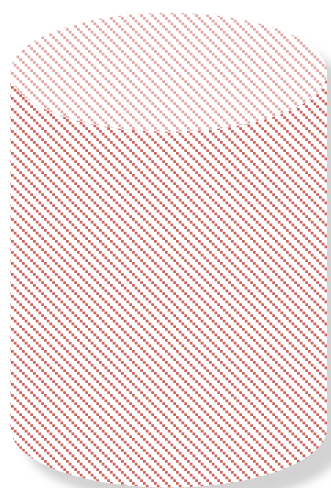
1. Domain: Collection of **observations** with **common topic**. Generally each domain is represented by a **dataset**.
2. Each domain has a **unique two-character** domain name (e.g., AE, CM, VS)
3. Variables in domain begin with the **domain prefix**: (e.g., VSTESTCD)

Characteristics of SDTM Variables

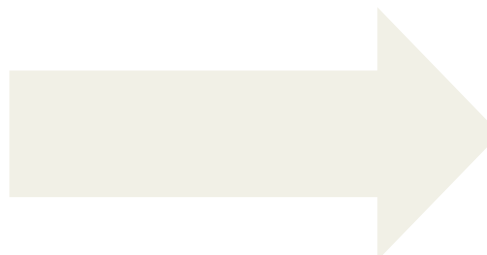
Variables –

1. Variable Name: limited to **8-chars**
2. Date/time format – **ISO 8601** is a text string
YYYY-MM-DDT hh:mm:ss (not a SAS format)
3. Controlled terminology –
 1. CT (**) published externally (ex: MedDRA or follow **CDISC-specific terminology**)
 2. CT (*) value from a **sponsor-defined code list**

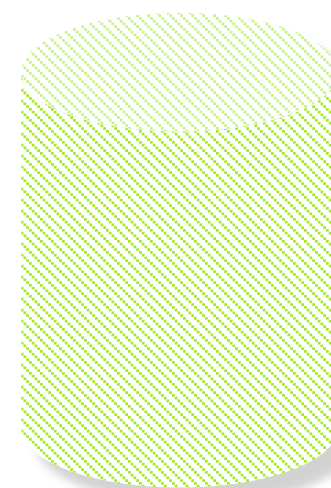
Concepts of Data Mapping



Source Datasets: Standard A
e.g.: Client's legacy standard datasets



Mapping Process



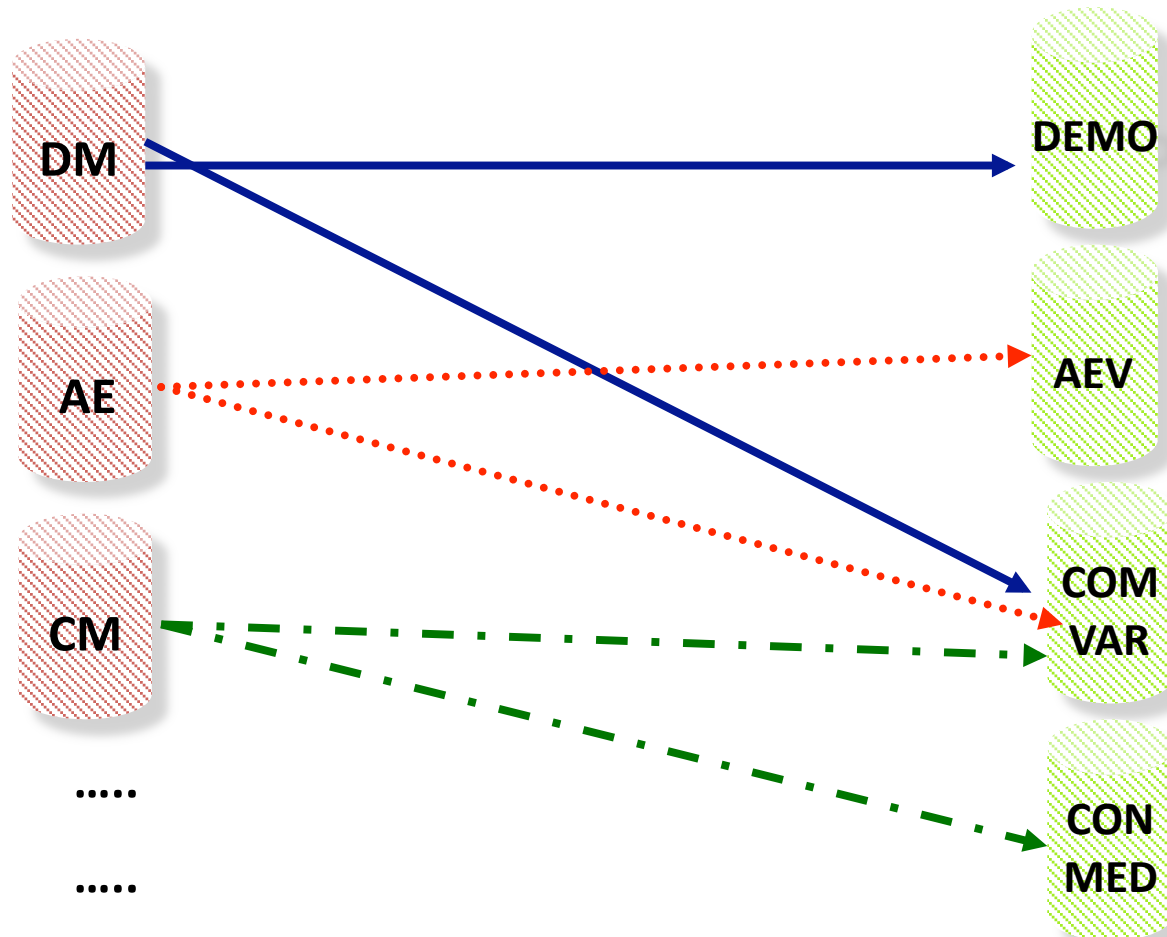
Target Datasets: Standard B
e.g.: CDISC SDTM
CDISC ADaM
Client's new standards

Data Mapping – Step 1

Mapping Domains

Target Domains

Source Domains



Data Mapping – Step 2

Mapping Variables

DM

DEMO

COM
VAR

Target Domain Variables	Source Domain Variables
STUDYID	COMVAR.STUDYID
DOMAIN	
USUBJID	COMVAR.USUBJID
SUBJID	DM.SUBJID
RFSTDTC	COMVAR._RFSTDTC
RFENDTC	COMVAR._RFENDTC
SITEID	DM.SITEID
INVID	
INVNAM	
BIRTHDTC	DM.BIRTHDTC
AGE	DM.AGE
AGEU	DM.AGEU
SEX	DM.SEX

Data Mapping – Step 3

Mapping Variables - Logic

DM

DEMO

**COM
VAR**

Mapping Logic

STUDYID	COMVAR.STUDYID	Direct
DOMAIN		Compute
USUBJID	COMVAR.USUBJID	Direct
SUBJID	DM.SUBJID	Direct
RFSTDTC	COMVAR._RFSTDT	Date Conversion
RFENDTC	COMVAR._RFENDTC	Direct
SITEID	DM.SITEID	Direct
INVID		Not mapping
INVNAM		Not mapping
BRTHDTC	DM.BRTHDTC	Direct
AGE	DM.AGE	Direct
AGEU	DM.AGEU	Direct
SEX	DM.SEX	Value Conversion

Types of Mapping

One to One Mapping

Source CRF Variables

Date of birth

SEX

Weight

SDTM Variables

BRTHDTC

SEX

VSORRES



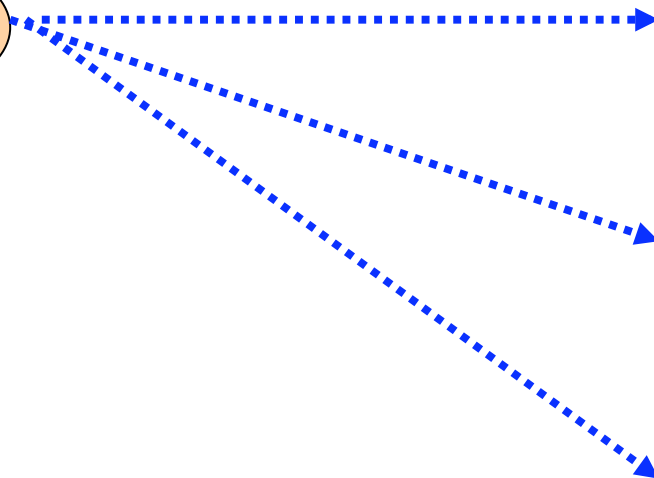
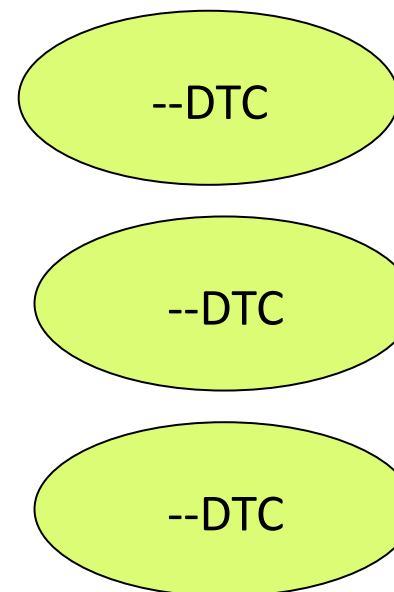
Types of Mapping

One to Many Mapping

Source CRF Variables



SDTM Variables



Types of Mapping

Many to One Mapping

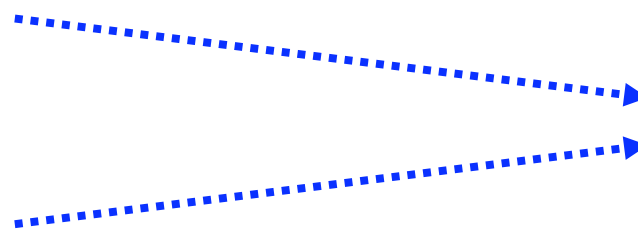
Source CRF Variables

FU Visit Date

Baseline Visit
Date

SDTM Variables

STUDY DAY



Automated Tool - MapGenie

Microsoft Excel - MapWizard_8feb_fill_test.xls										
File Edit View Insert Format Tools Data Window Help										
J4 AUTO										
Target Information					Source Metadata					
Seq. No.	Domain	Variable Name	Core	User Defined Format	Source Variable(ds.varname)	Mapping Type	Macro Calls/ SAS Expressions	User Comments	Mapped? (Y/N)	
		Details on			Show Source Variable List	Show Mapping Form				
1	DM	STUDYID	Req		COMVAR.STUDYID	AUTO	_UDYID=STUDYID;		Y	
2	DM	DOMAIN	Req			AUTO	main = "DM";		Y	
3	DM	USUBJID	Req		COMVAR.USUBJID	DATA TYPE	UBJID=USUBJID;		Y	
4	DM	SUBJID	Req		DM.SUBJID	VALUE CONVERSION	BJID=SUBJID;		Y	
5	DM	RFSTDT	Exp		COMVAR._RFSTDT	ISODATE CONVERSION	so8601_date(output_v		Y	
6	DM	RFENDTC	Exp		COMVAR._RFENDTC	CONCATENATE	=RFSTDT,cdm_var_typ		Y	
7	DM	RFENDTC	Exp		COMVAR._RFENDTC	EXTRACT	num,cdm_var_name=_		Y	
8	DM	RFENDTC	Exp		COMVAR._RFENDTC	COMPUTE	RFSTDT,cdm_informat=D		Y	
9	DM	RFENDTC	Exp		COMVAR._RFENDTC	AUTO	_RFENDTC=RFENDTC;		Y	
10	DM	SITEID	Req		DM.SITEID	AUTO	SITEID=SITEID;		Y	
11	DM	INVID	Perm							
12	DM	INVNAM	Perm							
13	DM	BRTHDTC	Perm		DM.BRTHDTC	AUTO	BRTHDTC=BRTHDTC;		Y	
14	DM	AGE	Exp		DM.AGE	AUTO	AGE=AGE;		Y	
15	DM	AGEU	Exp		DM.AGEU	AUTO	AGEU=AGEU;		Y	
16	DM	SEX	Req		DM.SEX	VALUE CONVERSION	%ValueConversionMap(o		Y	
17	DM	RACE	Exp		DM.RACE	AUTO	output_var=SEX,cdm_var_n		Y	
18	DM	ETHNIC	Perm		DM.RACEOTH	COMPUTE	ame=SEX,cdm_var_value		Y	
19	DM	ARMCD	Req		COMVAR._ARMCD	AUTO	s=M F,sdtm_var_values=		Y	
20	DM	ARM	Req		COMVAR._ARM	AUTO	M F);		Y	
21	DM	COUNTRY	Req		DM.COUNTRY	COMPUTE	RACE=RACE;		Y	
22	DM	DMDTC	Perm				if upcase(raceoth) =		Y	
23	DM	DMDY	Perm				"MOROCCAN" then ethnic		Y	
							= "HISP";			
							_ARMCD=ARMCD;			
							_ARM=ARM;			
							country = "RUS";			

References

1. Practical Methods for Creating CDISC SDTM Domain Data Sets from Existing Data - Robert W. Graebner, Quintiles, Inc., Overland Park, KS
2. SDTM--3.1.1ImplementationGuide.pdf
3. StudyDataTabulationModelv1.1.pdf

Questions?



Email Address

Anindita Bhattacharjee –
anindita.b@cytel.com