



2023

CHINA

INTERCHANGE

BEIJING | 25-26 AUGUST



## Metadata Repository (MDR): E2E Data Transparency and Traceability

Cong, Catherine. Principle Clinical Data Manager  
Zhao, Kai. Study Designer,  
Global Data Management and Standards,  
MSD R&D (China) Co., Ltd

# Meet the Speakers

Cong, Catherine

**Title:** Principle Clinical Data Manager , Associated Director

**Organization:** Global Data Management and Standards

Working and contributed in multiple TAs including Oncology, Diabetes, Vaccines and ID. With her extraordinary effort, Catherine cumulated solid End-to-End data management experiences from various InForm studies starting from Set-up till DBL.

Certified with CCDM.



Zhao, Kai

**Title:** Study Designer, Associated Director

**Organization:** Global Data Management and Standards

Working in Infectious Disease (HIV, COVID-19 etc) and Oncology therapeutic area.

SDTM Tabulate Certification.



# 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 MSD or position of CDISC.*
- *The authors have no real or apparent conflicts of interest to report.*



## Agenda

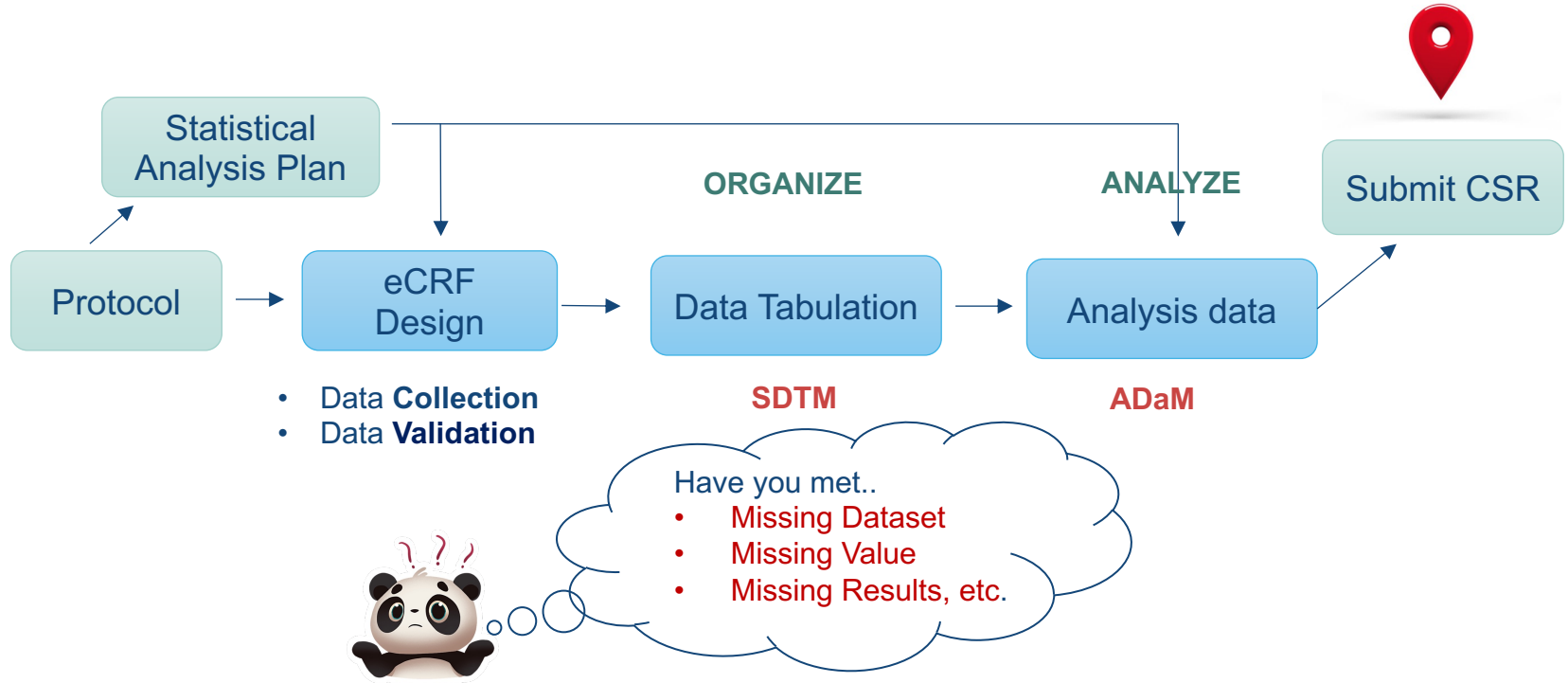
1. Start With the End in Mind
2. MDR: E2E Data Transparency and Traceability
3. Quality by Design by Using MDR
4. Deficiencies and Look Forward



# Module 1

Start With the End in Mind

# Simplified Clinical Data Flow



# Start With the End in Mind



**PROTOCOL**



**SAP**



**AGENCY  
REQUIREMENT**

**TRIAL DESIGN**

- ✓ Critical Data **Collection** & Validation
- ✓ Data Organizing & **Formatting**

**Guideline for Industry**

Structure and Content of Clinical  
Study Reports

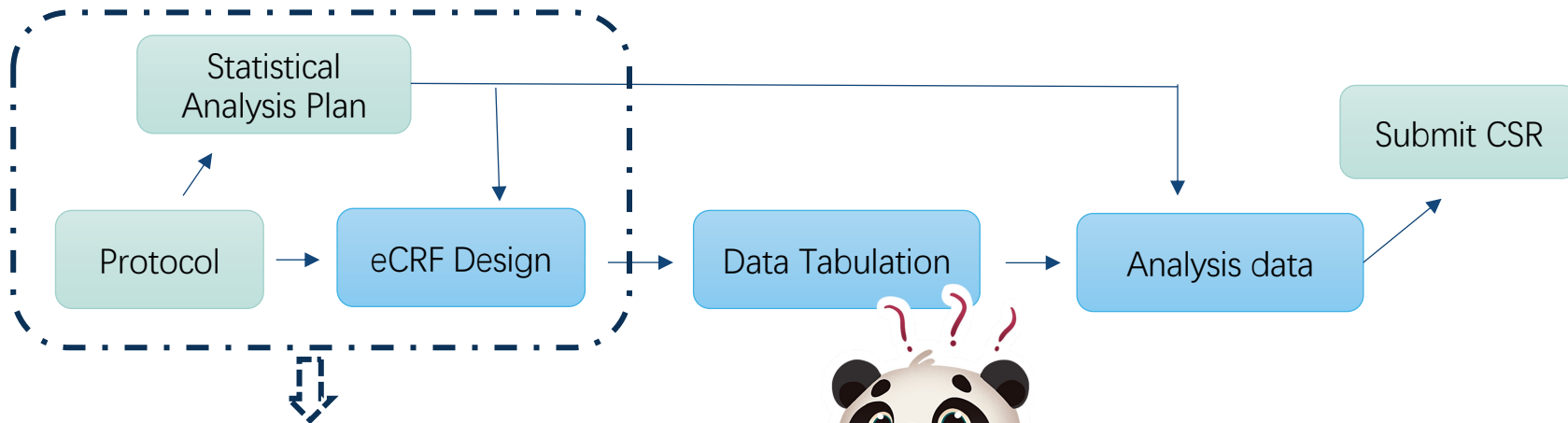


AGU 1996  
ICH E3



**Clinical  
Study  
Report**

# Start With the End in Mind

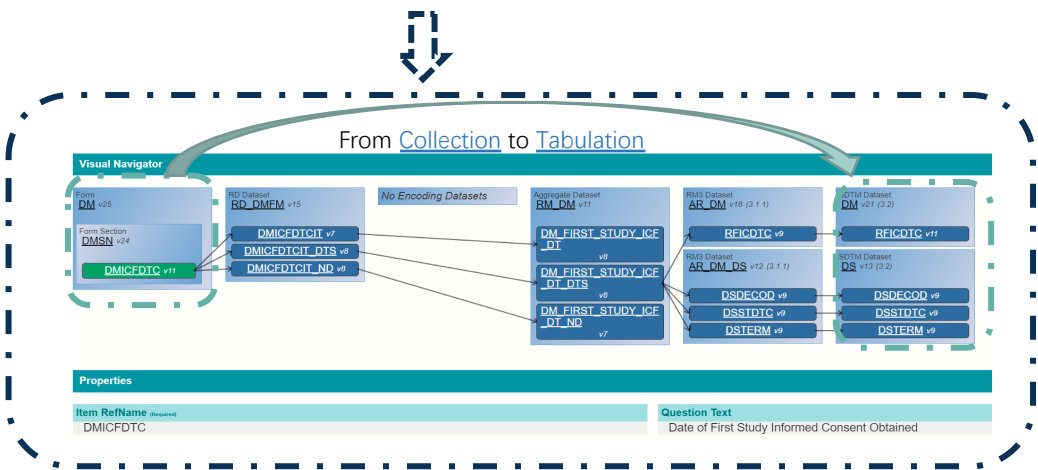
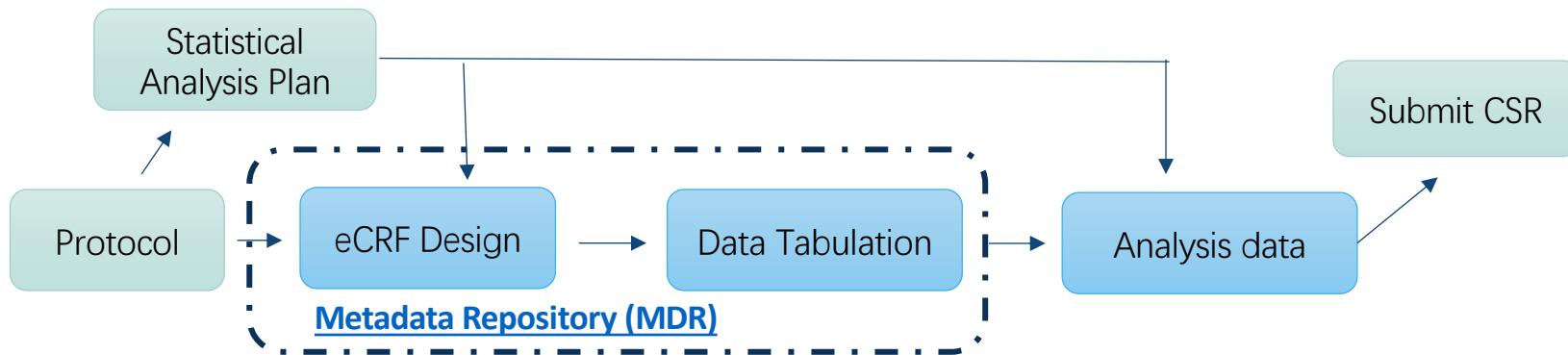


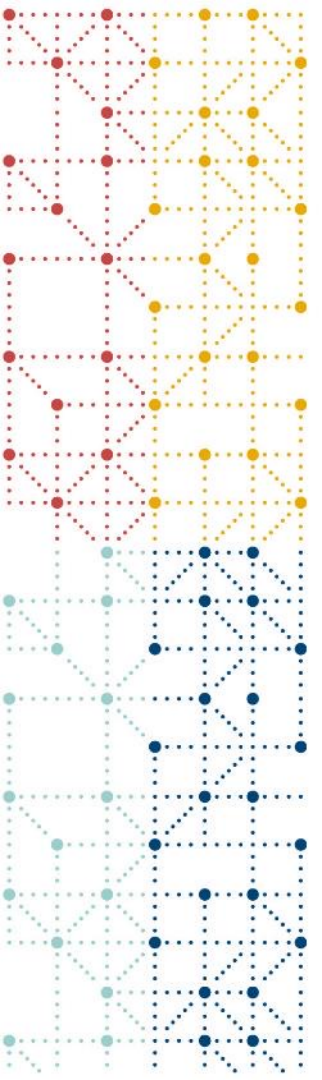
DEMOGRAPHICS (DM) [DMFM]	
Demographics [DMSN]	
1. <input checked="" type="checkbox"/> Date of Birth <i>[read-only]</i> [Date of Birth]	[DMBRTHDTCIT] (MM/YYYY) Req/Unk <input checked="" type="checkbox"/> / Req <input checked="" type="checkbox"/> (1900-2015)
2. <input checked="" type="checkbox"/> Calculated Age <i>[read-only]</i> [Calculated Age]	[DMAGEIT] NS
3. <input checked="" type="checkbox"/> Unit for Calculated Age <i>[read-only]</i> [Unit for Calculated Age]	[DMAGEUIT] [A:HOURS] <input type="radio"/> Hours [A:DAYS] <input type="radio"/> Days [A:WEEKS] <input type="radio"/> Weeks [A:MONTHS] <input type="radio"/> Months [A:YEARS] <input type="radio"/> Years
4. <input checked="" type="checkbox"/> Collected Age <i>[read-only]</i> [Collected Age]	[DMCAGEIT] NS
5. <input checked="" type="checkbox"/> Unit for Collected Age <i>[read-only]</i> [Unit for Collected Age]	[DMCAGEUIT] [A:HOURS] <input type="radio"/> Hours [A:DAYS] <input type="radio"/> Days [A:WEEKS] <input type="radio"/> Weeks [A:MONTHS] <input type="radio"/> Months [A:YEARS] <input type="radio"/> Years
6. <input checked="" type="checkbox"/> Sex Reported at Birth <i>[read-only]</i> [Sex Reported at Birth]	[DMSEXIT] [A:F] <input type="radio"/> Female [A:M] <input type="radio"/> Male [A:UNDIFFERENTIATED] <input type="radio"/> Undifferentiated [A:U] <input type="radio"/> Unknown
7.* <input checked="" type="checkbox"/> Gender Identity [Gender Identity]	[DMORRESGLISTIT] [A:LIST] <input type="radio"/> List [DMORRESGIPD] List [A:OTHER] <input type="radio"/> [DMORRESGIOTHTB] <input type="checkbox"/> Other, Specify: A200*





# Start With the End in Mind








## Module 2

MDR: E2E Data Transparency and Traceability

# Centralized and Reusable Standards

Home

-  Metadata Maintenance
-  Study Maintenance >
-  SDTM Maintenance >

## Version Summary

**Form Mnemonic** Search

---

**Status**

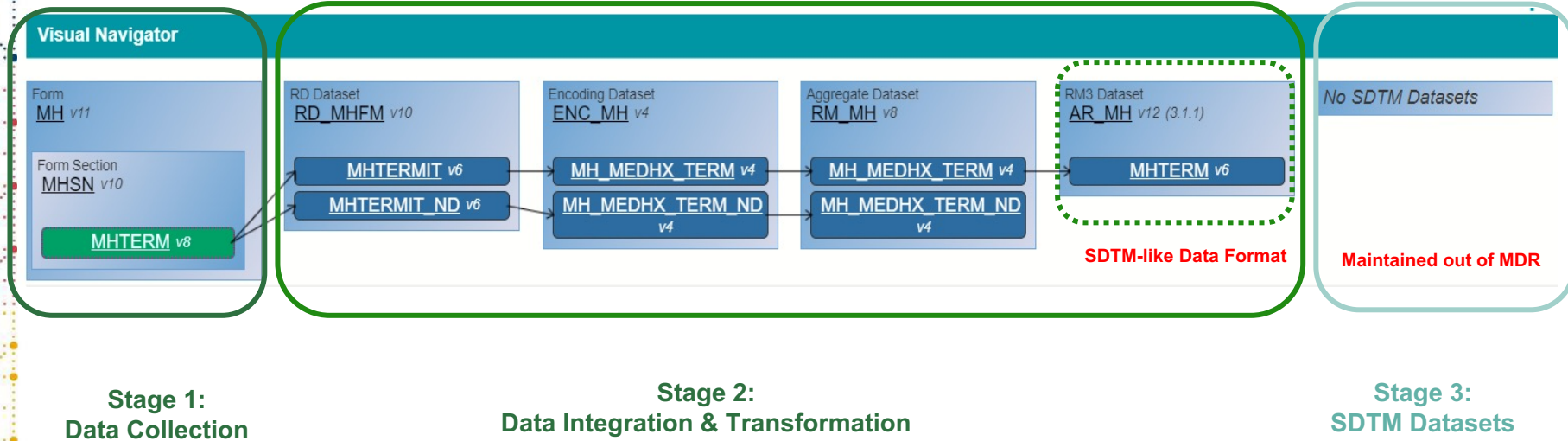
Show Preferred Version Only  
 Show Active Revisions Only

Show RM-Aggregate Only  
 Show RM-Original Only

## Data Definition Maintenance

Sponsor	SDTM Version	No. of Domains	Prod Revision	Draft Revision
Sponsor ↓				
Merck	<a href="#">3.1.1</a>	114	103	<a href="#">103.1</a>
Merck	<a href="#">3.1.3</a>	105	11	<a href="#">11.1</a>
Merck	<a href="#">3.2</a>	103	161	<a href="#">161.1</a>
Merck	<a href="#">3.3</a>	119	173	<a href="#">173.1</a>

# Data Lineage Shown via MDR



# From Collection to Tabulation

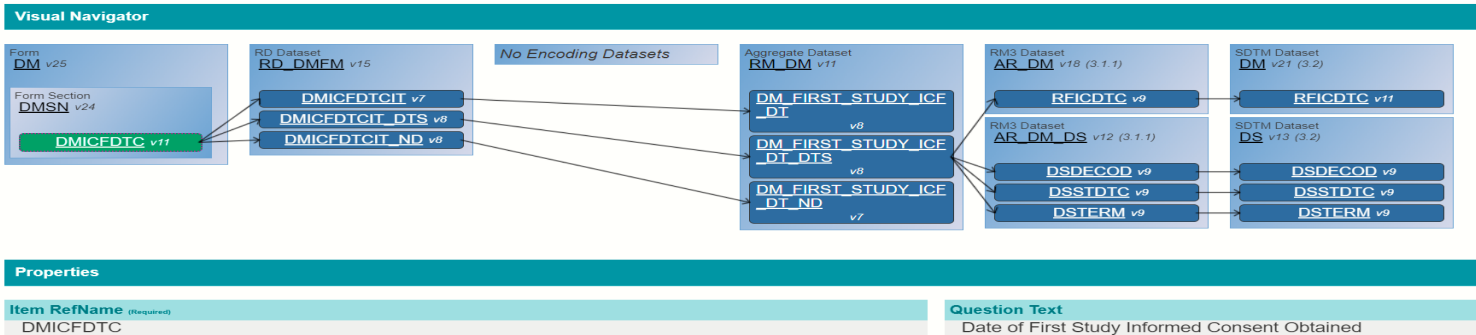
The screenshot displays the 'MH - Form' interface. At the top, there are navigation icons and buttons for 'GENERATE DRAFT' and 'FOCUS NAVIGATOR'. Below this is a breadcrumb trail: 'Source Scope: 8.1 > Scope: 9 > 20200430 > Production > RM-Aggregate > Wed Jul 01 2020 > Scope Type: Global'. A menu bar includes 'Metadata', 'Comments(0)', 'Change Summary', 'Spec Approval', 'Quality Check', 'CD Adherence', and 'AGGR Report'. The 'Visual Navigator' section shows a grid of data elements: 'Form MH v5', 'Form Section MHSN v4', 'RO Dataset ED\_MHFM v0', 'Encoding Dataset ENC\_MH v1', 'Aggregate Dataset RM\_MH v1', 'RM3 Dataset AR\_MH v5 (2.1.1)', 'SDTM Dataset MH v0 (2.1.1)', 'SDTM Dataset MH v6 (2.2)', and 'SDTM Dataset NOT MAPPED v4 (2.2)'. An 'Expand All' button is highlighted in red. Below the visual navigator is a 'Properties' section with a '+' icon, and a 'Form Section(s), SDTM Domain(s), RM3 Domain(s), Attachments, Form Details' section with a '-' icon. The bottom part of the interface shows a table with columns: 'Form Section(s)', 'SDTM Domain(s)', 'RM3 Domain(s)', 'Attachment(s)', and 'Form Details'. The table contains one row with 'MHSN' under 'Form Section(s)' and 'Medical History - Provide Any Background or ...' under 'Form Details'.

- E2E Data Transparency & Traceability
- Visualization for Stakeholders

# From Spreadsheet to Electronic Platform

RM3 Target Version: 3.1.1												
Item #	Itemref Name	Question	Response(s)	Database Value	Field Type (CONTROL ID, AIA, TYPE)	Field Format	Display Override	Derived in DMV	eCRF Design Guidelines	Mandatory/Optional	Required? (Yellow/Green)	SDV's?
		Demographics			Section Header				This item is mandatory for non-AT studies. This is a web-based form. Form is used once per study. This item should be placed in the first visit of the TBE. Form validation for TBE: ESD studies, all items must be set as SDV-Critical, not applicable to Late Stage studies.	Mandatory		
1	DMBIRTHDT	Date of Birth	REGURK: REQ	Date	Date	XXXX	Read-Only		This item is mapped to ICF item from the Screening (SCRE) form; however, the value on the DMV eCRF can be updated by the user.	Mandatory	Required	TRUE
2	DAGE	Calculated Age		Text Box	XXXX		Read-Only		This item is used to support Developmental Milestones (AGES, 5% SDSC, etc). This item is not reported in SDTM. The RPTAGE is derived in RM3 using the study DSSTDC; the value is inferred from the value in DSCT. Auto-populated in Inform by a derivation rule. Ensure the edit check validation is selected based on the calculated and to be used in the study. Notes: If any portion of the DOW is LBR, then the Calculated age may not be precise. The derivation logic is as follows: 1. Day is LBR or Blank, then use "0" as the day. 2. Month is LBR and Day is LBR or Blank, then use "June" as the Month and "00" as the Day. 3. The age is derived based on the first visit date. This question is used in conjunction with LBR for Calculated Age item.	Mandatory	Not Required	FALSE
3	DAGEU	Unit for Calculated Age	Hours	HOURS	Radio Button		Read-Only		This question is used in conjunction with the Calculated Age item. This item uses the unit (SDSC UNIT) dictionary in v10C. Study team to determine which unit to be populated for all subjects and ensure the edit check validation is updated accordingly. Remove any units that are not applicable for the study. This item is not reported in SDTM. The DSCTAGE Unit is derived by a DMV derivation using the study DSSTDC. Auto-populated in Inform by a derivation rule (see edit check spec in RM3). The edit check should be selected in the study. The standard calculation for age in years, however, the edit check derivation can be modified to be calculated in a different unit.	Mandatory	Not Required	FALSE
				Days	DAYS	Radio Button						
				Weeks	WEEKS	Radio Button						
				Months	MONTHS	Radio Button						
				Years	YEARS	Radio Button						
4	DMAGE	Collected Age		Text Box	XXXX				Collected age is optional. This is intended for sites to manually enter the age into Inform at the time of first DCOV. Note: This value must be different from the calculated age. Study team to determine if this item should be included. Some examples of when this might be needed: 1. When parent DCOB is collected, either by range or based on country-specific regulations that prevent a DCOB from being collected, and a precise age at time of enrollment is needed. 2. If age is being used for stratification and precision is required. Example: parent is 5 yrs and was born 23Mar and entered the study 01Jun. The collected age is 5 yrs but the calculated age would be 6 yrs.	Optional	Required	TRUE

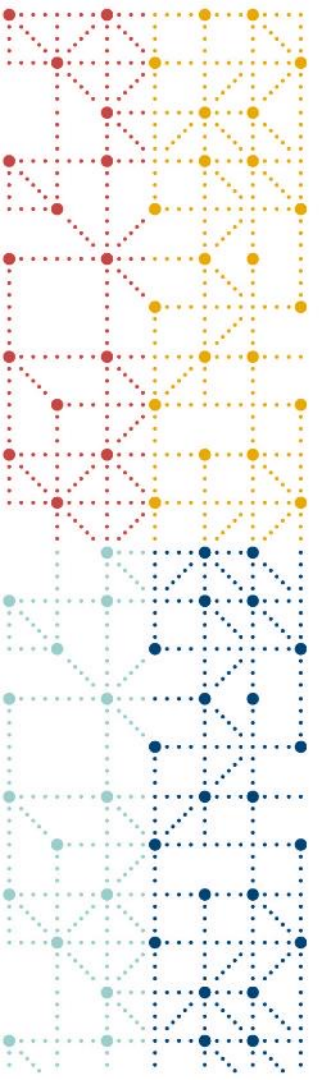
- Silos to Network
- Human-readable to Machine-readable



# MDR Application in Clinical Data Management

The screenshot displays the Metadata Management Repository (MDR) application interface. The top navigation bar is teal and contains the text "Metadata Management Repository" and a "Logout" button. A left sidebar menu includes "Home", "Metadata Maintenance", "Study Maintenance", and "SDTM Maintenance". The main content area is titled "Version Summary" and features a search box labeled "Form Mnemonic" and "Search". Below the search box are three input fields for "Form Type", "Status", and "Intended Use". A filter section contains several checkboxes: "Show Preferred Version Only", "Show RM-Aggregate Only" (checked), "Show Study Scope Only", "Include Abandoned Scopes", "Show Active Revisions Only", "Show RM-Original Only", "Show Global Scope Only" (checked), and "Include Retired Scopes". Below the filters, a message states: "Please enter filter criteria (ex. Form Mnemonic) and press enter for Version Summary results." followed by an "EXPORT" button. At the bottom, a table header is visible with columns: "Form", "Status", "Scope", "Scope Type", "RM Type", "CD Version", "Intended Use", "SDTM Versions", and "Req #".

- ✓ E2E Data Flow Transparency and Traceability
- ✓ Efficiency in Data Collection, Use and Sharing
- ✓ Data Management Capabilities



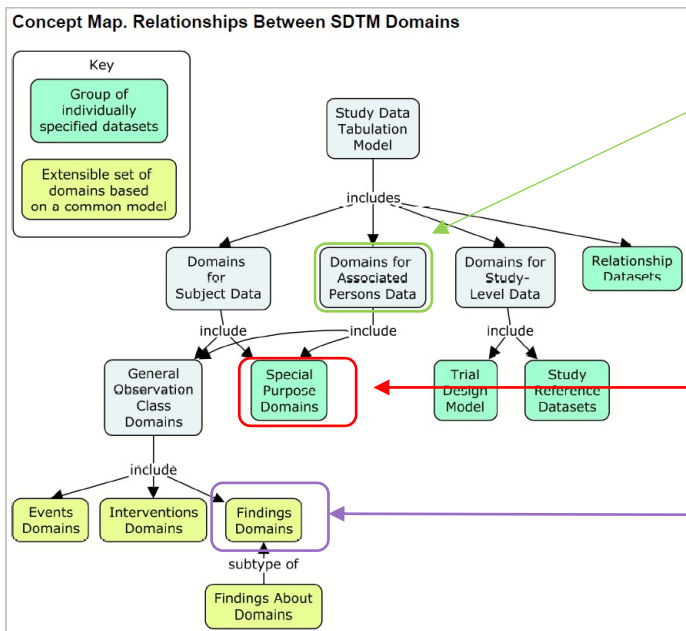
## Module 3

Quality by Design by using MDR



# What data will be collected and where will they go?

Fig.1 Concept map from SDTM 2.0



When a subject is pregnant or becomes pregnant, it is important and may be necessary to collect both prenatal and postnatal data on the infant.

Fig.2 Subject pregnancy-fetus/infant in MDR

Visual Navigator	
Form PRG02 v23	RD Dataset RD_PRG02FM v19
Form Section PRG022SN v1	No Encoding Datasets
Form Section PRG02SN v23	Aggregate Dataset RM_PRG02 v16
	RD Dataset AR_PRG02_APDM v1 (3,3)
	RD Dataset AR_PRG02_APRP v1 (3,3)
	RD Dataset AR_PRG02_APRS v1 (3,3)
	RD Dataset AR_PRG02_APSC v1 (3,3)
	RD Dataset AR_PRG02_APVS v1 (3,3)
	RD Dataset AR_PRG02_RP v1 (3,3)

Properties

Form Mnemonic (Required): PRG02

Form Title (Required): SUBJECT PREGNANCY - FETUS/INFANT

# E2E Data Visualization for an eCRF

Fig.1 DM eCRF

STND-009: DEMOGRAPHICS (DM) [DMFM]	
<b>Demographics [DMX]</b>	
1. Date of Birth (read-only) Date of Birth	[DMBIRTHDTCT] (MM/YYYY) Req/Unk <input checked="" type="checkbox"/> / Req <input checked="" type="checkbox"/> (1900-2025)
2. Calculated Age (read-only) Calculated Age	[DMCAGEIT] NS
3. Unit for Calculated Age (read-only) Unit for Calculated Age	[DMCAGEUT] (A.HOURS) <input type="radio"/> Hours (A.DAYS) <input type="radio"/> Days (A.WEEKS) <input type="radio"/> Weeks (A.MONTHS) <input type="radio"/> Months (A.YEARS) <input type="radio"/> Years
4. Collected Age (read-only) Collected Age	[DMCAGEIT] NS
5. Unit for Collected Age (read-only) Unit for Collected Age	[DMCAGEUT] (A.HOURS) <input type="radio"/> Hours (A.DAYS) <input type="radio"/> Days (A.WEEKS) <input type="radio"/> Weeks (A.MONTHS) <input type="radio"/> Months (A.YEARS) <input type="radio"/> Years
6. Sex Reported at Birth (read-only) Sex Reported at Birth	[DMSEXIT] (A.F) <input type="radio"/> Female (A.M) <input type="radio"/> Male (A.UNDERIDENTIFIED) <input type="radio"/> Underidentified (A.U) <input type="radio"/> Unknown
7. Gender Identity (read-only) Gender Identity	[DMGRESLISTIT] (A.LIST) <input type="radio"/> [DMGRESLIST] List (A.OTHER) <input type="radio"/> [DMGRESIDCTM] Other, Specify A200*
8. Ethnicity (read-only) Ethnicity	[DMETHNICIT] (A.HISPANIC OR LATINO) <input type="radio"/> Hispanic or Latino (A.NOT HISPANIC OR LATINO) <input type="radio"/> Not Hispanic or Latino (A.NOT REPORTED) <input type="radio"/> Not Reported (A.UNKNOWN) <input type="radio"/> Unknown
9. Race (read-only) Race	[DMRACEIT] (A.AMERICAN INDIAN OR ALASKA NATIVE) <input type="radio"/> American Indian or Alaska Native (A.ASIAN) <input type="radio"/> [DMRACEASIT] <input type="checkbox"/> Asian [DMRACEASL_VL_PEARLTRACE] <input checked="" type="checkbox"/> (A.BLACK OR AFRICAN AMERICAN) <input type="radio"/> Black or African American (A.NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER) <input type="radio"/> Native Hawaiian or Other Pacific Islander (A.WHITE) <input type="radio"/> White (A.OTHER) <input type="radio"/> [DMRACEMULTI] Multi-Race (Check at Least Two): (A.AMERICAN INDIAN OR ALASKA NATIVE) <input type="checkbox"/> American Indian or Alaska Native (A.ASIAN) <input type="checkbox"/> Asian (A.BLACK OR AFRICAN AMERICAN) <input type="checkbox"/> Black or African American (A.NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER) <input type="checkbox"/> Native Hawaiian or Other Pacific Islander (A.WHITE) <input type="checkbox"/> White
10. Date of First Study Informed Consent Obtained (read-only) Date of First Study Informed Consent Obtained	[DMFCOBTCT] (MM/DD/YYYY) Req <input checked="" type="checkbox"/> / Req <input checked="" type="checkbox"/> / Req <input checked="" type="checkbox"/> (2015-2025)

Fig.2 eCRF E2E Data Visualization in MDR

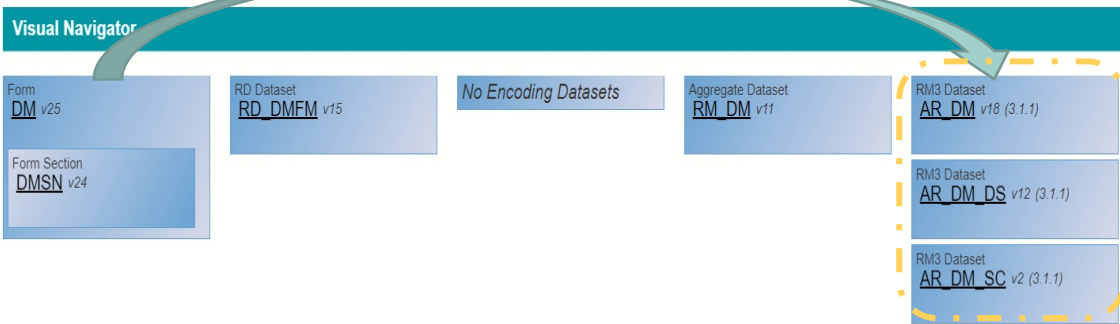
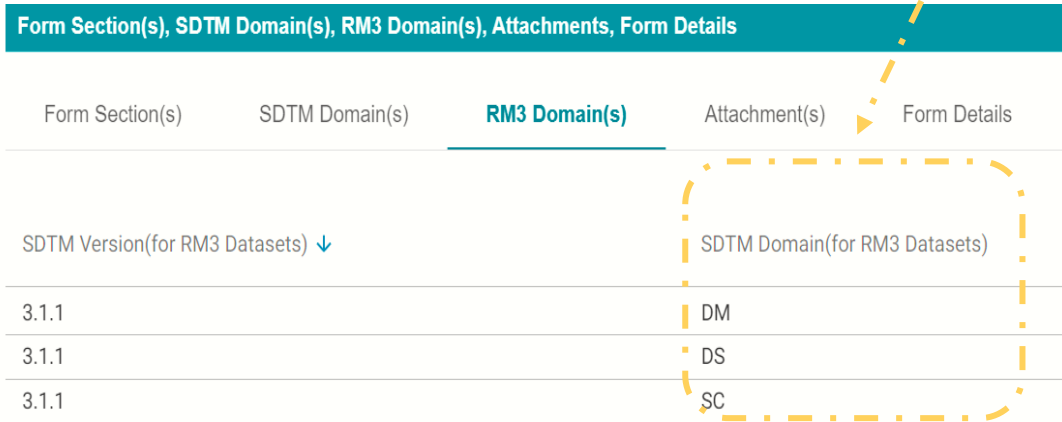


Fig. 3 SDTMIG version for domains



# E2E Data Visualization for an eCRF

Fig.1 eCRF E2E Data Visualization for an item in MDR

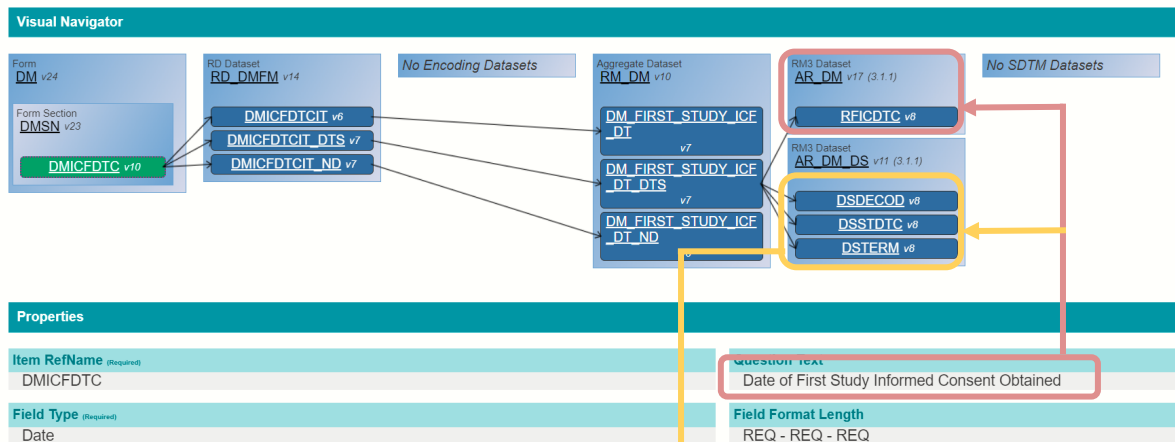


Fig.2 DSTERM, DSDECOD, DSCAT

RM3/SDTM SAS Name(s), Topic/Decode, RD Item(s)				
RM3/SDTM SAS Name(s)	Topic / Decode	RD Item(s)		
<input type="checkbox"/> CT Date ↓ 🔍	Topic# ↑ 🔍		Topic ↑ 🔍	Decode ↑ 🔍
<input type="checkbox"/> 2018-06-29			INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED
			CAT ↑ 🔍	SCAT ↑ 🔍
			PROTOCOL MILESTONE	

# E2E Data Visualization for an eCRF

Fig.1 DM eCRF

**SD-001: DEMOGRAPHICS (DM) [DMFM]**

**Demographics (DM)**

1. Date of Birth (Required)  
(Date of Birth)  
[MM/DD/YYYY] [MM/YY]

2. Calculated Age (Required)  
(Calculated Age)  
[Age]

3. Sex for Calculated Age (Required)  
(Sex for Calculated Age)  
[M] [F]

4. Collected Age (Not used)  
(Collected Age)  
[Age]

5. Sex for Collected Age (Not used)  
(Sex for Collected Age)  
[M] [F]

6. Sex Reported at Birth (Not used)  
(Sex Reported at Birth)  
[M] [F]

7. Sex Reported at Death (Not used)  
(Sex Reported at Death)  
[M] [F]

8. Gender Identity  
(Gender Identity)  
[Male] [Female] [Other]

9. Ethnicity  
(Ethnicity)  
[White] [Black or African American] [Hispanic or Latino] [Not Hispanic or Latino] [Other Hispanic or Latino] [Other Black or African American] [Other] [Other]

10. Race  
(Race)  
[American Indian or Alaska Native] [Asian] [Black or African American] [Hispanic or Latino] [Other] [Other]

11. Date of First Study Informed Consent Observed (Not used)  
(Date of First Study Informed Consent Observed)  
[MM/DD/YYYY] [MM/YY]

Fig.2 eCRF E2E Data Visualization for an item in MDR

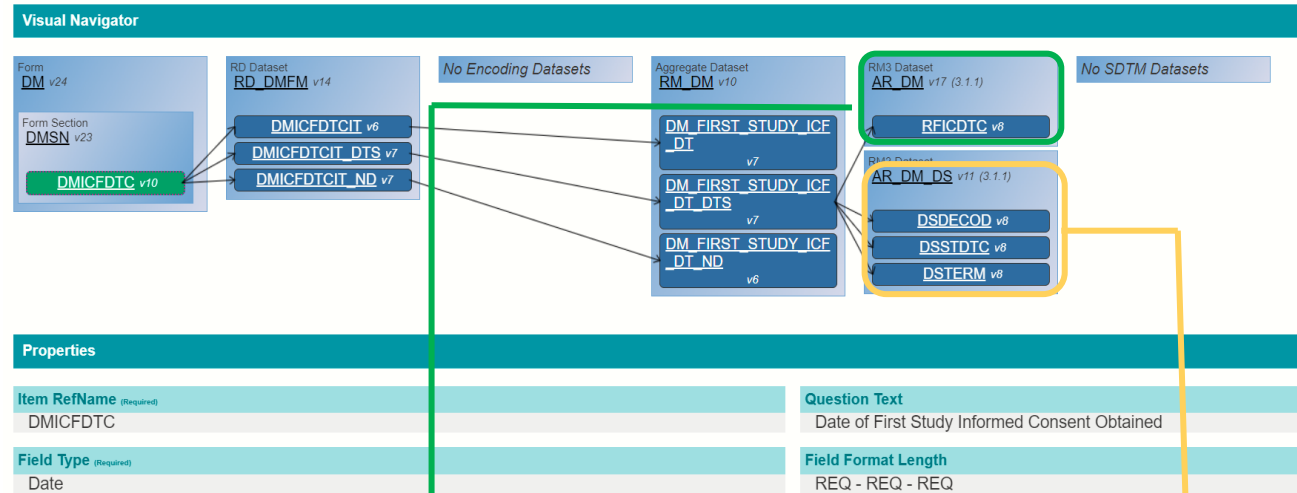


Fig.3 SDTM datasets

dm.xpt

Row	STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC	RFXSTDTC	RFXENDTC	RFICDTC	RFPENDTC	SITEID	INNVAM	BRTHDTC	AGE	AGEU
1	ABC123	DM	123101	01001	2006-01-12	2006-03-10	2006-01-12	2006-03-10	2006-01-03	2006-04-01	01	JOHNSON, M	1948-12-13	57	YEARS

ds.xpt

Row	STUDYID	DOMAIN	USUBJID	DSSEQ	DSTERM	DSDECOD	DSCAT	DSSCAT	EPOCH	DSDTC	DSSTDTC
1	ABC123	DS	123101	1	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE		SCREENING	2006-01-03	2006-01-03

# Data Conformance

## FDA-STUDY DATA TECHNICAL CONFORMANCE GUIDE

### 8.2.2 Support on Data Validation Rules

Sponsors should evaluate their study data before submission against the conformance rules published by an SDO, the eCTD Technical Rejection Criteria for Study Data (See Appendix F), and the FDA Business Rules. Sponsors may also wish to use the FDA Validator Rules to understand what is available to the FDA reviewer. Sponsors should either correct any discrepancies between study data and the standard or the business rules or explain meaningful discrepancies in the relevant Reviewer Guide (RG). Additional information about conformance to the standard, FDA Business Rules, or FDA Validator Rules that could facilitate review of the submitted data, or establish consistency and traceability between the study data and the Study Report, should also be provided in the relevant RG.

[Study Data Technical Conformance Guide – Technical Specifications Document | FDA](#)

## PMDA-Technical Conformance Guide on Electronic Study Data Submissions

### 3.6.1 Validation of data conforming to the CDISC standards

The PMDA will perform validation of data that conform to the CDISC standards using Pinnacle 21 Enterprise.

The rules used for validation have been classified by the level of importance taking into consideration the characteristics of each rule and based on conformity to the standards, ease of use of the data in review, quality of the clinical study data which the PMDA should know beforehand, and future uses of the clinical study data by the PMDA. The levels of importance are shown below.

- (a) Rules which, if violated, will cause the review to be suspended until corrections have been made  
Very basic rules such as the presence/absence of necessary datasets for each clinical study
- (b) Rules which, if violated, will require an explanation  
In many cases, these rules are clearly stated in each standard and implementation guide, and if violated, the applicant should explain in the reviewer's guide about the reason for the violation and the reason why it is not possible to correct it.
- (c) Rules which, even when violated, will not necessarily require any explanation

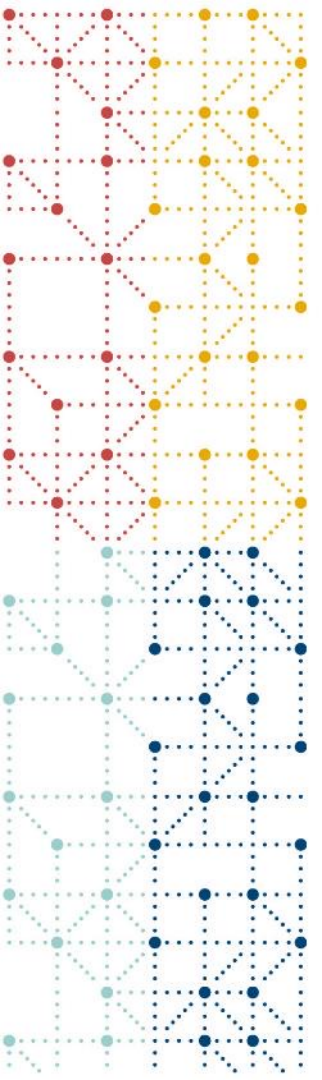
The reason for the violation possibly is requested separately for the above (c) from the perspective of the quality of the clinical study data.

# Data Conformance

Rule ID	Publisher ID	Message	Description	Domains	FDA	PMDA 2010.2	PMDA 2211.0	NMPA	FDA Severity	PMDA Severity	3.1.2	3.1.3	3.2	3.3	Notes
SD1035	FDAB014	Missing value for DSCAT	Category for Disposition Event (DSCAT) should be populated.	DS	X	X	X			Warning	X	X	X	X	
SD1315	CG0066	DSDECOD is not the same as DSTERM, when DSCAT=PROTOCOL MILESTONE	Reported Term for the Disposition Event (DSTERM) and Standardized Disposition Term (DSDECOD) should contain the same value, when Category for Disposition Event (DSCAT) is 'PROTOCOL MILESTONE'.	DS	X	X	X	X		Warning	X	X	X		

## RM3/SDTM SAS Name(s), Topic/Decode, RD Item(s)

RM3/SDTM SAS Name(s)	Topic / Decode	RD Item(s)
<input type="checkbox"/> CT Date ↓ 🔍	Topic# ↑ 🔍	
<input type="checkbox"/> 2018-06-29	<div style="display: flex; justify-content: space-between;"> <span>Topic ↑ 🔍</span> <span>Decode ↑ 🔍</span> </div>	<div style="display: flex; justify-content: space-between;"> <span>CAT ↑ 🔍</span> <span>SCAT ↑ 🔍</span> </div>
	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE



# Module 4

Deficiencies and Look Forward

# Deficiencies

- SDTM dataset is maintained out of MDR
- Controlled Terminology is maintained out of MDR

**Visual Navigator**

**Properties**

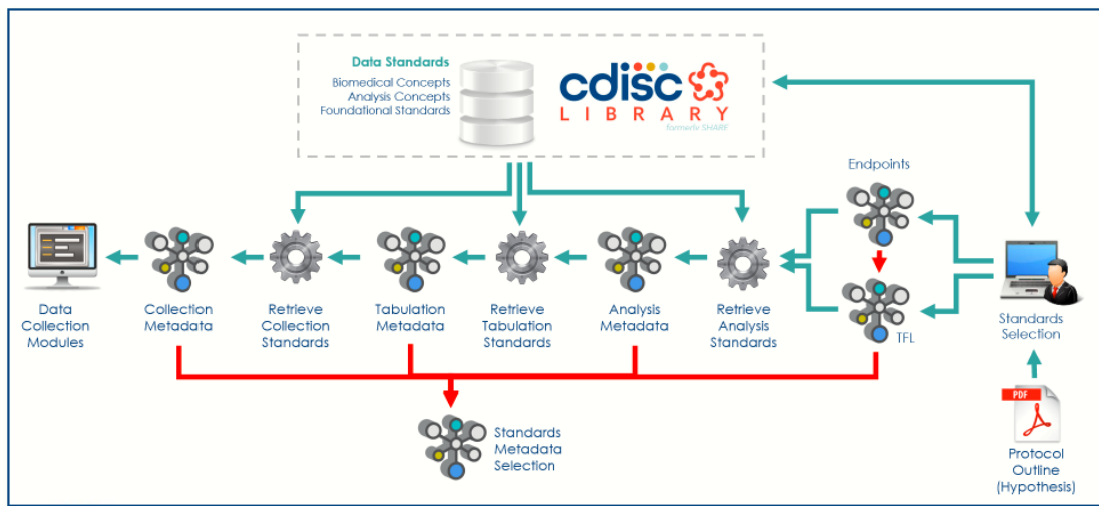
<b>Item RefName</b> <small>(Required)</small>	<b>Question Text</b>
DMICFDTC	Date of First Study Informed Consent Obtained
<b>Field Type</b> <small>(Required)</small>	<b>Field Format Length</b>
Date	REQ - REQ - REQ

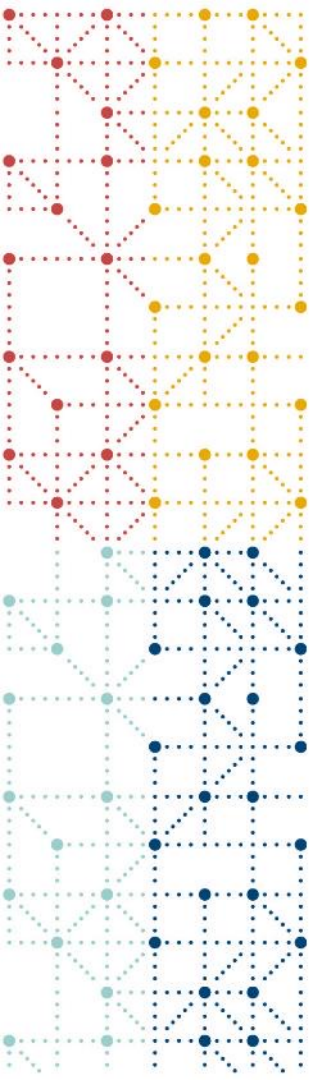
RM3/SDTM SAS Name(s)	Topic / Decode	RD Item(s)
<input type="checkbox"/> CT Date ↓ Q	Topic# ↑ Q	Topic ↑ Q
<input type="checkbox"/> 2018-06-29	INFORMED CONSENT OBTAI...	INFORMED CONSENT OBTAI... PROTOCOL MILESTONE



# Look Forward

## CDISC 360 Piloted development of linked biomedical concept metadata to enable end to end automation





**Thank You!**

**cdisc**