LPO Q&A: Which Domains and Variables Should be Used in a Study

NCI CDISC Harmonization Working Group Meeting

13 May 2020

NIH NATIONAL CANCER INSTITUTE

Topics

1. *YN vs PERF vs OCCUR*
2. *When to use Procedures Domain*
3. *General Q&A about SDTM requirements*

--PERF, --YN, and--OCCUR

What's the difference?

Comparing --YN, --PERF, and --OCCUR: **What do they have in common**?

- Answer: They all use a **"Yes" or "No"** response

Were any concomitant medications taken? NOT SUBMITTED CMYN	o Yes o No
What was the medication? CMTRT	_____
For what indication was the medication taken? CMINDC	_____
What was the individual dose of medication? CMDOSE	_____
Dose Unit CMDOSU	o mg o g
What was the frequency of the medi CMDOSFRQ	
What was the route of administratio CMROUTE	
What was the medication start date CMSTDTC CMSTDAT	
What was the medication end date? CMENDTC CMENDAT	

Does the subject have cirrhosis? MHOCCUR	o Yes o No
Start Date MHSTDTC MHSTDAT	__/___/____
Ongoing MHENRF / MHENRTPT MHENTPT MHONGO	o Yes o No
End Date MHENDTC MHENDAT	__/___/____

Was the ECG performed? EGPERF	o Yes NOT SUBMITTED o No EGSTAT = "NOT DONE" WHERE EGTESTCD = "EGALL"
What was the ECG reference identifier? EGREFID	_____
What was the ECG date? EGDTC EGDAT	__/___/____
What was the ECG time? EGDTC EGTIM	__:__
What was the method used to measure ECG? EGMETHOD	o 6-Lead Standard o 12-Lead Standard
What was the position of the subject during ECG measurement? EGPOS	o Supine o Standing

Comparing --YN, --PERF, and --OCCUR: **How are they different**?

--YN	--PERF	--OCCUR
Asks: "Is there **any** data to record"	Asks "Was the protocol-specified measurement, test or assessment done **(perf**ormed)?**"**	Asks "Does (has) the participant have (had) [**specific** medical condition or symptom]" or "Does (has) the participant use(d) / take(n) [**specific** medication / procedure / other specific Intervention]"
Typically used in Interventions (e.g., CM, SU, PR) and Events (e.g., MH, AE)	Typically used in Findings class data (because these are the *planned measurements, tests, observations required by protocol)*	**Always** used to ask about pre-specified (**specific**) Interventions or Events.
If the answer is 'No', no data would be expected in that form	If the answer is "No", no data would be expected in the form (except perhaps Reason Not Done)	If the answer is "No", maybe "Reason for Occur" could be collected
If the answer is "No", there is no SDTM record created	If "No" an SDTM Record can be created for the planned test to indicate the planned test was Not Done (and the reason not done, if collected)	If "No" an SDTM Record can be created to show they did or did not have that specific condition (or that they did or did not take that specific medication).

Examples: CMYN asks "are there **any** concomitant medications to record?"

Were any concomitant medications taken? NOT SUBMITTED CMYN	o Yes o No
What was the medication? CMTRT	_____
For what indication was the medication taken? CMINDC	_____
What was the individual dose of medication? CMDOSE	_____
Dose Unit CMDOSU	o mg o g o mL
What was the frequency of the medication? CMDOSFRQ	o PRN o BID o QD
What was the route of administration of the medication? CMROUTE	o ORAL o SUBCUTANEOUS o TOPICAL
What was the medication start date? CMSTD	/ /
What CMEND	

- If "Yes", an edit check would confirm that at least one complete CM record was entered in the form
- If "No", no data would be expected in the CRF
- In the SDTM CM domain, there would be **no CM records because nothing was collected**
  - Note in the example there are only records for two participants displayed. The gap in between the USUBJIDs could indicate that USUBJIDs ABC123-766 through ABC123-898 did not have **any** CMs to record.

cm.xpt

Row	STUDYID	DOMAIN	USUBJID	CMSEQ	CMGRPID	CMTRT	CMCAT	CMDOSFRM	CMROUTE	CMRSDISC
1	ABC123	CM	ABC123-765	1	1	PEGINTRON	HCV TREATMENT	INJECTION	SUBCUTANEOUS	COMPLETED SCHEDULED TREATMENT
2	ABC123	CM	ABC123-765	2	1	RIBAVIRIN	HCV TREATMENT	TABLET	ORAL	COMPLETED SCHEDULED TREATMENT
3	ABC123	CM	ABC123-765	3	1	BOCEPREVIR	HCV TREATMENT	TABLET	ORAL	COMPLETED SCHEDULED TREATMENT
4	ABC123	CM	ABC123-899	1	1	PEGINTRON	HCV TREATMENT	INJECTION	SUBCUTANEOUS	TOXICITY/INTOLERANCE
5	ABC123	CM	ABC123-899	2	1	RIBAVIRIN	HCV TREATMENT	TABLET	ORAL	TOXICITY/INTOLERANCE
6	ABC123	CM	ABC123-899	3	1	BOCEPREVIR	HCV TREATMENT	TABLET	ORAL	TOXICITY/INTOLERANCE

Examples: EGPERF asks “was the ECG performed?”

Was the ECG performed? EGPERF	o Yes NOT SUBMITTED o No EGSTAT = "NOT DONE" WHERE EGTESTCD = "EGALL"
What was the ECG reference identifier? EGREFID	_____
What was the ECG date? EGDTC EGDAT	__/___/_
What was the ECG time? EGDTC EGTIM	__:__
What was the method used to measure ECG? EGMETHOD	o 6-Lead St o 12-Lead S
What was the position of the subject during ECG measurement? EGPOS	o Supine o Standing

- If “Yes”, an edit check would confirm that the rest of the mandatory fields in the form are completed
- If “No”, no other data would be expected (except, perhaps, a “Reason Not Done”)
- In the SDTM EG domain, a record can show that the ECG was not performed (at that VISIT or Date, or Planned Timepoint)

eg.xpt

Row	EGTESTCD	EGTESET	EGORRES	EGORRESU	EGSTRESC	EGSTRESN	EGSTRESU	EGSTAT	VISITNUM	EGDTC
1	QRSAG	PR Interval, Aggregate	0.362	sec	0.362	0.362	sec		1	2015-03-07
2	QTAG	QT Interval, Aggregate	221	msec	0.221	0.221	sec		1	2015-03-07
3	QTCBAG	QTcB Interval, Aggregate	412	msec	0.412	0.412	sec		1	2015-03-07
4	SPRTARRY	Supraventricular Arrhythmias	ATRIAL FLUTTER		ATRIAL FLUTTER				1	2015-03-07
6	INTP	Interpretation	ABNORMAL		ABNORMAL				1	2015-03-07
5	PRAG	PR Interval, Aggregate						NOT DONE	2	2015-03-14
7	EGALL	ECG Test Results						NOT DONE	3	2015-03-21

Examples: MHOCCUR "do they have **[cirrhosis]**?"

Specific medical condition

Does the subject have cirrhosis? MHOCCUR	o Yes o No
Start Date MHSTDTC MHSTDAT	__/___/____
Ongoing MHENRF / MHENRTPT MHENTPT MHONGO	o Yes o No
End Date MHENDTC MHENDAT	__/___/____

- If "Yes", an edit check would confirm that the rest of the mandatory fields are completed for the cirrhosis record
- If "No", no other data would be expected (perhaps Reason for Occur - more typical in Interventions)
- In the SDTM MH domain, a record would be created to show either that this person does NOT have cirrhosis, or to show the complete record if they do have cirrhosis.

mh.xpt

Row	STUDYID	DOMAIN	USUBJID	MHSEQ	MHGRPID	MHTERM	MHPRESP	MHOCCUR	MHSTDTC	MHENDTC	MHENRTPT	MHENTPT
1	ABC123	MH	123101	1		CIRRHOSIS	Y	N				
2	ABC123	MH	123102	1		CIRRHOSIS	Y	N				
3	ABC123	MH	123103	1		CIRRHOSIS	Y	N				
4	ABC123	MH	123104	1	CHF	CIRRHOSIS	Y	Y	9/17/2004		ONGOING	VISIT 1
5	ABC123	MH	123105	1	CHF	CIRRHOSIS	Y	Y	9/19/2004		ONGOING	VISIT 1

Examples: MHOCCUR "do they have **[cirrhosis]**?"

Does the subject have cirrhosis? MHOCCUR	o Yes o No
Start Date MHSTDTC MHSTDAT	__/___/____
Ongoing MHENRF / MHENRTPT MHENTPT MHONGO	o Yes o No
End Date MHENDTC MHENDAT	__/___/____

- Alternative representation in SDTM is to use Findings About (FA) to show all the MHOCCUR responses, and then only include the "Y" records in the MH domain.
  - *This approach is used for AE because we are not allowed to show AEs that did NOT happen in the AE domain. You can use it for any Interventions or Events.*

fa.xpt

Row	STUDYID	DOMAIN	USUBJID	FASEQ	FATESTCD	FATEST	FAOBJ	FAORRES	FASTRESC	FASTAT	VISITNUM	VISIT
1	ABC123	FA	123101	1	OCCUR	Occurrence	CIRRHOSIS	N	N		1	VISIT 1
2	ABC123	FA	123102	1	OCCUR	Occurrence	CIRRHOSIS	N	N		1	VISIT 1
3	ABC123	FA	123103	1	OCCUR	Occurrence	CIRRHOSIS	N	N		1	VISIT 1
4	ABC123	FA	123104	1	OCCUR	Occurrence	CIRRHOSIS	Y	Y		1	VISIT 1
5	ABC123	FA	123105	1	OCCUR	Occurrence	CIRRHOSIS	Y	Y		1	VISIT 1

mh.xpt

Row	STUDYID	DOMAIN	USUBJID	MHSEQ	MHGRPID	MHTERM	MHPRESP	MHSTDTC	MHENDTC	MHENRTPT	MHENTPT
1	ABC123	MH	123104	1	CHF	CIRRHOSIS	Y	9/17/2004		ONGOING	VISIT 1
2	ABC123	MH	123105	1	CHF	CIRRHOSIS	Y	9/19/2004		ONGOING	VISIT 1

Questions?

When Should We Use the Procedures (PR) Domain

How do we know when to use Procedures (PR)?

- Use Procedures (PR) when information collected about the **procedure *is the record of interest:***
  - Usually (not always) additional questions would be asked about the procedure, for example
    - Start Date (Start Time)
    - End Date (End Time)
    - Duration
    - Indication
    - Dose (e.g., Radiation Therapy, Transfusions)
    - Dose Unit

Did the participant have:	
Radiation Therapy	Yes / No
Mastectomy	Yes / No
Other, Specify	

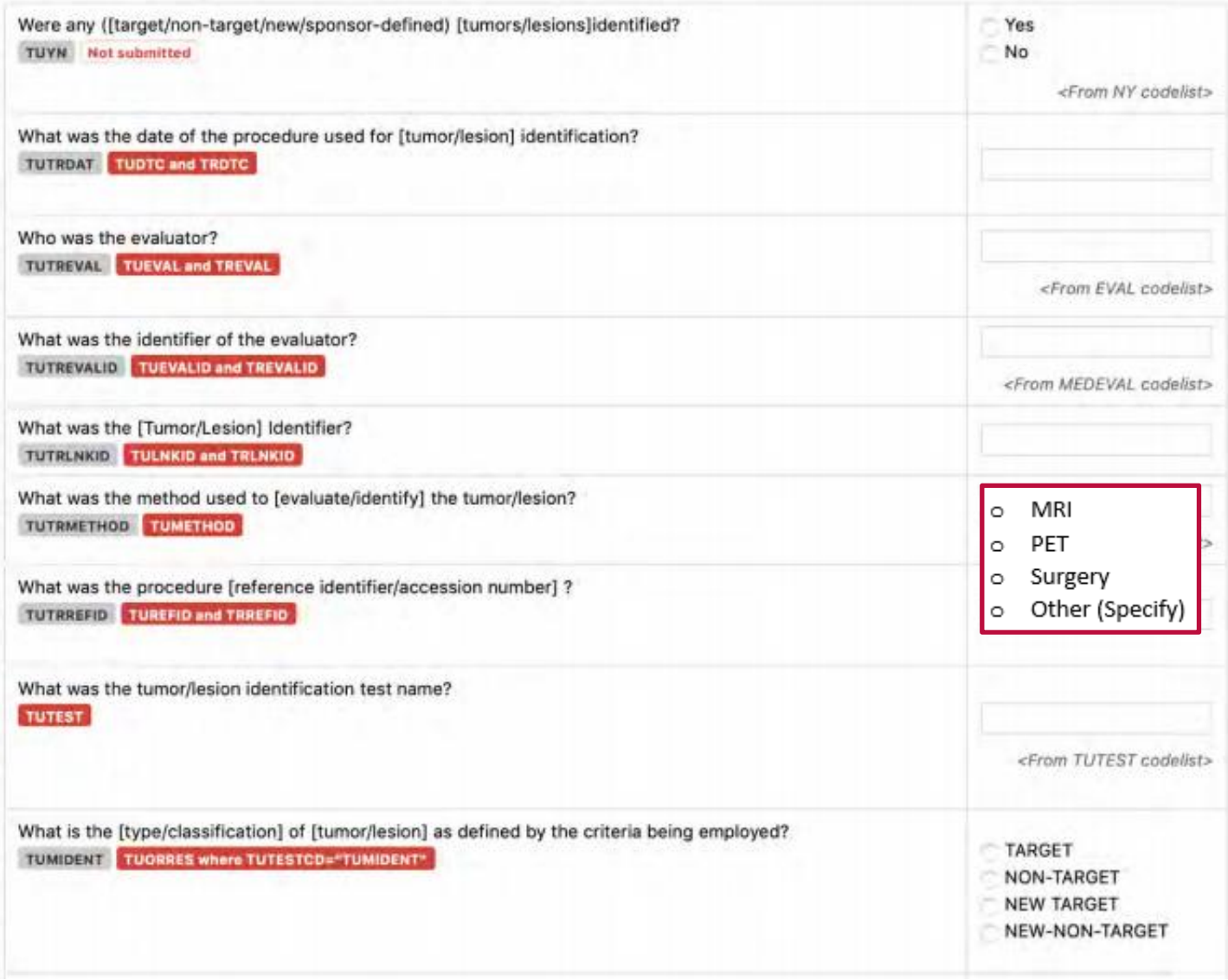
Sometimes we just want to know whether they had the procedure, but *that's all we collect.*

Requires PR

Did the participant have radiation therapy?	Yes / No
Date of most recent RT	__ __ - __ __ __ - __ __ __ __
Type of RT	
Amount of RT	
Duration of RT treatment	

How do we know when to use Procedures (PR)?

- Procedures (PR) is NOT required when the procedure is collected *only as the method for obtaining results*.
- Use the **relevant Findings Class domain** when the procedure is just a **method** of obtaining results



How do we know when to use Procedures (PR)?

- In <u>some</u> cases you might need both

Were any tumors identified	○ Yes ○ No	
Tumor Identification Date	▼ Day / ▼ Month / ####	
Tumor Type	○ Target ○ Non-Target	
Evaluation Method	○ MRI ○ PET ○ Surgery ○ Other (Specify)	(If surgery is selected, complete the surgery form)

Tumor Identification (TU)

Surgery Form	
Start Date	▼ Day / ▼ Month / ####
Start Time	__:__
End Date	▼ Day / ▼ Month / ####
End Time	__:__

Procedures (PR)

Questions?

Requirements, In General

Which Domains and Variables Should We Always Use?

Requirements Q&A: Which Domains and Variables to Use

- **Question**: "*The SDTMIG specifies required fields for all domains. I wanted to some clarifications on "required" fields. If you are not using specific domains in your study, then is it OK to leave those required fields out? For example, the SDTMIG list the required fields for the Skin Response (SR) domain, but most of our studies would not use this domain as we don't usually collect data that would fall under the SR domain. I would assume if you are not using specific domains in your data collection that you would not need to include required fields from these domains in your data set. Is this correct or would we need to include these required fields for domains we are not using, and these fields would be left blank?"*
- **RESPONSE:**
  - The data you *need to collect / have collected for your study* will determine *which domains* you use. If you are not collecting data that goes into the domain, you will not use ANY of the variables in that domain.
  - However, there are three domains that are "required" by SDTM rules
    - Demographics - to describe all of the study participants
    - Exposure (**if it is an interventional study with an IP**) - to describe each participant's exposure to your study treatment
    - Disposition - to describe how each person described in Demographics exited the study
  - One additional domain - Protocol Deviations (DV) - is also required by FDA (TCG)

Requirements Q&A: Which Domains and Variables to Use

- **Question:** Is there a spreadsheet or document that exists that lists all the SDTM variables that are required for FDA submission? I guess I'm wondering if some of the variables that are not listed as required in SDTM may actually be required for FDA submission and vice versa.
- **RESPONSE:**
  - The short answer is "no" (see next slide) - there is not a single spreadsheet or other document that lists all the SDTM variables that need to be in a submission package:
    - The basic, foundational requirements for the CDISC standards are published in the CDISC standards documentation.
    - FDA has additional requirements that they publish on their website
    - **Your exact requirements are based on the science and regulation that govern your study:**
      - **Data needed for an oncology study will be different from data needed for a vaccine study**
      - **There may be some differences between FDA and PMDA requirements**

Requirements Come From Multiple Sources

- **Foundational**, basic **requirements** for submission data come from the published Standards, e.g.:
  - CDISC Standards (e.g., SDTM, ADaM) come from CDISC - https://www.cdisc.org/
  - MedDRA coding dictionaries come from MedDRA Maintenance and Service Organization (MedDRA MSSO) - https://www.meddra.org/
- **FDA requirements** come from FDA and are published on their website (updated semi-annually):
  - https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources
  - The standards *that are required* are listed in FDA's Data Standards Catalog (DSC)
    - https://www.fda.gov/media/85137/download (latest version April 2020)
  - Implementation rules that are *really important to FDA (to support their review and oversight)* are published in a document called the FDA Study Data Technical Conformance Guide (TCG)
    - https://www.fda.gov/media/136460/download (latest version March 2020)
    - Some of the TCG rules are *in addition to the foundational rules from CDISC*
- *PMDA, Health Canada, EMEA, MHRA, etc. each may publish their own requirements for data*

Requirements Q&A: Which Domains and Variables to Use

- **Question:** Is there a spreadsheet or document that exists that lists all the SDTM variables that are required for FDA submission? I guess I'm wondering if some of the variables that are not listed as required in SDTM may actually be required for FDA submission and vice versa.
- **RESPONSE:**
  - The longer answer is that we are putting together a spreadsheet that may help a little bit
  - The spreadsheet will include
    - All of the variables that are ***allowed*** to be used in that domain (because, following the published **rules** for doing so, you can *bring variables in from the Model* if they are not already published in the domain)
    - Extra columns to show additional requirements from FDA (see examples following slides)

Example: SDTMIG "Allowed" Variables

Observation Class	Domain Prefix	Variable Name (minus domain prefix)	Variable Name	Variable Label	Ty	Controlled Terms or Form	Role	CDISC Notes (for domains) Description (for General Classes), or Assumptions (for Allowed Variables)	Core and Allowed Variables	Reference for Adding Allowed Variables (SDTMIG 1.4.1)	FDA Requirement	FDA Referenc
Interventions	AG	INDC	AGINDC	Indication	Char		Record Qualifier	Although allowed, AGINDC would not generally be used since substance administrations represented in AG are given as part of a testing procedure rather than with therapeutic intent.	Allowed	SDTMIG Section 6.1.1 AG Domain, Assumption 5a		
Interventions	AG	DOSTOT	AGDOSTOT	Total Daily Dose	Num				Allowed	SDTMIG Section 6.1.1 AG Domain, Assumption 5b		
Interventions	AG	DOSRGM	AGDOSRGM	Intended Dose Regimen	Char				Allowed	SDTMIG Section 6.1.1 AG Domain, Assumption 5b		
Interventions	AG	LOT	AGLOT	Lot Number	Char				Allowed	SDTMIG Section 6.1.1 AG Domain, Assumption 5		
Interventions	AG	LOC	AGLOC	Location of Dose Administration	Char	(LOC)			Allowed	SDTMIG Section 6.1.1 AG Domain, Assumption 5		
Interventions	AG	LAT	AGLAT	Laterality	Char	(LAT)			Allowed	SDTMIG Section 6.1.1 AG Domain, Assumption 5		
Interventions	AG	DIR	AGDIR	Directionality	Char	(DIR)			Allowed	SDTMIG Section 6.1.1 AG Domain, Assumption 5		
Interventions												

This spreadsheet will combine the "Allowed" variables from the Model (SDTM) into each SDTMIG domain

- In the proper order (i.e., the same order as published in the Model)
- With "Allowed" designated in the "Core and Allowed Variables" column
- With notes from general or domain Assumptions (if any) about the variables
- With any required SDTM CT codelists designated as Controlled Terminology for the allowed variable

Identifiers → Topic → Qualifiers → Timing

Example: Additional FDA Requirements (required "on top of" foundational rules)

Observation Class	Domain Prefix	Variable Name (minus domain prefix)	Variable Name	Variable Label	Ty	Controlled Terms or Form	Role	CDISC Notes (for domains) Description (for General Classes), or Assumptions (for Allowed Variables)	Core and Allowed Variables	Reference for Adding Allowed Variables (SDTMIG 1.4.1)	FDA Requirement	FDA Referenc
Interventions	CM	EPOCH	EPOCH	Epoch	Char	(EPOCH)	Timing	Epoch associated with the start date/time of the medication administration. Null for medications that started before study participation.	Perm		EPOCH should be included for clinical subject-level observations	FDA Study Data Technical Conformance Guide (March 2020) Section 4.1.4.1
Interventions	CM	STDY	CMSTDY	Study Day of Start of Medication	Num		Timing	Study day of start of medication relative to the sponsor-defined RFSTDTC.	Perm		Populate if CMSTDTC contains a date	Study Data Technical Conformance Guide (March 2020) Section 4.1.4.1 #3
Interventions	CM	ENDY	CMENDY	Study Day of End of Medication	Num		Timing	Study day of end of medication relative to the sponsor-defined RFSTDTC.	Perm		Populate if CMENDTC contains a date	Study Data Technical Conformance Guide (March 2020) Section 4.1.4.1 #3
Interventions	EX	EPOCH	EPOCH	Epoch	Char	(EPOCH)	Timing	Trial Epoch of the Exposure record. Examples: "RUN-IN", "TREATMENT".	Perm		EPOCH should be included for clinical subject-level observations	FDA Study Data Technical Conformance Guide (March 2020) Section 4.1.4.1
												Study Data

In most cases, the additional FDA Requirements are to tell us that SDTMIG has a Permissible (Perm) variable that FDA wants to effectively make Expected (Exp) - meaning that variable should always be in your dataset, and you should populate it if you have collected or can derive the data. If you cannot populate it, you should explain why in the Study Data Reviewer's Guide (SDRG).

Example: In a few cases, FDA Requirements may break SDTM Conformance

Observation Class	Domain Prefix	Variable Name (minus domain prefix)	Variable Name	Variable Label	Ty	Controlled Terms or Form	Role	CDISC Notes (for domains) Description (for General Classes), or Assumptions (for Allowed Variables)	Core and Allowed Variables	Reference for Adding Allowed Variables (SDTMIG 1.4.1)	FDA Requirement	FDA Referenc
Special-Purpose	DM	SUBJID	SUBJID	Subject Identifier for the Study	Char		Topic	Subject identifier, which must be unique within the study. Often the ID of the subject as recorded on a CRF.	Req		SDTMIG rules have SUBJID *only* in the DM domain. FDA requirement is in addition to this: If a single subject is screened and/or enrolled more than once in a study, then the subject's SUBJID should be different for each unique screening or enrollment. For a study with multiple screenings and/or multiple enrollments per subject, SUBJID should be included in other related domains besides DM even though it may cause validation errors. It is recommended to include a table linking each SUBJID for a single subject to that subject's USUBJID with any additional necessary explanation included in the relevant RG.	FDA Study Data Technical Conformance Guide (March 2020) Section 4.1.1.2

Example: FDA's preference for handling participants with multiple screenings and / or multiple enrollments will cause validation errors. Explain the errors in the SDRG.

Questions?

Send additional questions to the support email:

NCICDISCSupport.nih.gov