An In-Depth Analysis of the Updates and Challenges in SDTM IG 3.3 and 3.4

Presented by Fan Yang, Senior Statistical Programmer
Novartis

Meet the Speaker

Fan Yang

Title: Senior Statistical Programmer

Organization: Novartis

Fan is currently working as a senior statistical programmer at Novartis in Shanghai. She received her master's degree in Statistics from the George Washington University in 2020. Over the past four years, she has participated in various Therapeutic Area (TA) studies, acquiring and applying expertise in the SDTM IG for the creation of SDTM datasets. Presently, she is leading a Phase III in-house study, utilizing SDTM IG v3.3 for the development of SDTM datasets.

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.*

- *The author(s) have no real or apparent conflicts of interest to report.*

Agenda

Introduction

A History of SDTM & SDTMIG

SDTM Version	Year	SDTMIG Version	Number of Domains*
1.0	2004	3.1	23
1.1	2005	3.1.1	30
1.2	2008	3.1.2	32
1.3	2012	3.1.3	32
1.4	2013	3.2	46
1.5	2016	N/A**	
1.6	2017	N/A***	
1.7	2018	3.3	61
2.0	2021	3.4	63

*   Includes all special-purpose, general-observation-class domains, and study references
**   Created for the SENDIG v3.1
*** Created for the SENDIG-DART v1.1

FDA and PMDA Data Standards

HA	Version	Date (mandatory)	Source
FDA	SDTMIG 3.3	15Mar2023	FDA Data Stds Catalog_8.01.2022 (v8.2)
PMDA	SDTMIG 3.3	01Apr2023	PMDA Data Stds Catalog (28.02.2023)
FDA	SDTMIG 3.4	15Mar2025	FDA Data Stds Catalog_04.12.2024 (v10.3)

For FDA Start date is assumed as **First Patient First Visit** on or after the specified date.

For PMDA Start date is assumed as **submission date**, Study submission on or after the specified date can be done on SDTM IG v3.3 and SDTM IG v3.2 as well until end date is provided.
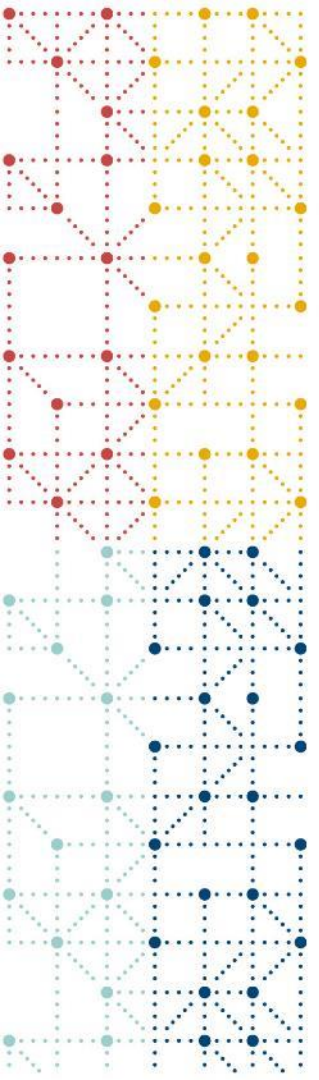
cdisc

SDTM IG v3.2 *vs* SDTM IG v3.3 & v3.4 Updates

Major Updates

❑ Newly added section & idea

❑ Newly added/decommissioned/reorganized SDTM domains

❑ New variables added in existing domains & details added clarifying derivation of existing variables

❑ Attribute updates

cdisc

SDTM IG v3.2 *vs* SDTM IG v3.3 & v3.4 Updates

Newly Added & Decommissioned & Reorganized Domains

Newly Added SDTM Domains

SDTM IG Version	Domain Type	Domain Name and Code
v3.3	Special Purpose	Subject Milestones (SM)
v3.3	Interventions	Meal Data (ML)
		Procedure Agents (AG)
v3.3	Findings	Cardiovascular Findings (CV)
		Musculoskeletal Findings (MK)
		Nervous System Findings (NV)
		Ophthalmic Examinations (OE)
		Respiratory System Findings (RE)
		Urinary System Findings (UR)
		Functional Tests (FT)
v3.3	Trial Design	Trial Milestones (TM)
v3.3	Relationships	RELSUB
v3.3	Study References	Device Identifiers (DI)
		Non-host Organism Identifiers (OI)
		Pharmacogenomic/Genetic Biomarker Identifiers (PB)
v3.4	Findings	Biospecimen Events (BE)
		Biospecimen Findings (BS)
		Cell Phenotype Findings (CP)
		Genomics Findings (GF)

Newly Added SDTM Domains
Example: Ophthalmic Examinations (OE)

Description:

- Contains tests that measure a person's ocular health and visual status
- Detect abnormalities in the components of the visual system to determine how well the person can see

Assumptions:

- Using variable FOCID to identify treatment sites 👀

(OEFOCUS) CDISC codelist:
OD: Oculus Dexter (Right Eye)
OS: Oculus Sinister (Left Eye)
OU: Oculus Uterque (Both Eyes)

Row	STUDYID	DOMAIN	USUBJID	FOCID	OESEQ	OETESTCD	OETEST	OEORRES	OESTRESC	OELOC	OELAT	VISITNUM	VISIT	OEDTC
1	XXX	OE	XXX-450110	OS	1	INTP	Interpretation	NORMAL	NORMAL	LENS	LEFT	1	SCREENING	2020-03-20
2	XXX	OE	XXX-450110	OD	2	INTP	Interpretation	ABNORMAL	ABNORMAL	LENS	RIGHT	2	SCREENING	2020-03-20

Row	STUDYID	DOMAIN	USUBJID	FOCID	PRSEQ	PRLINKID	PRTRT	PRPRESP	PROCCUR	PRLOC	PRLAT	VISITNUM	VISIT	PRSTDTC
1	XYZ	PR	XYZ-100001	OS	1	1	OCT	Y	Y	EYE	LEFT	2	VISIT 1	2012-05-25T08:30:00
2	XYZ	PR	XYZ-100001	OD	2	2	OCT	Y	Y	EYE	RIGHT	2	VISIT 1	2012-05-25T08:30:00
3	XYZ	PR	XYZ-100001	OU	3	3	OCT	Y	N			3	VISIT 2	

Reorganized Domains
Questionnaires, Ratings, and Scales (QRS) Domains

SDTM IG v3.2

QS

SDTM IG v3.3

QS   RS   FT
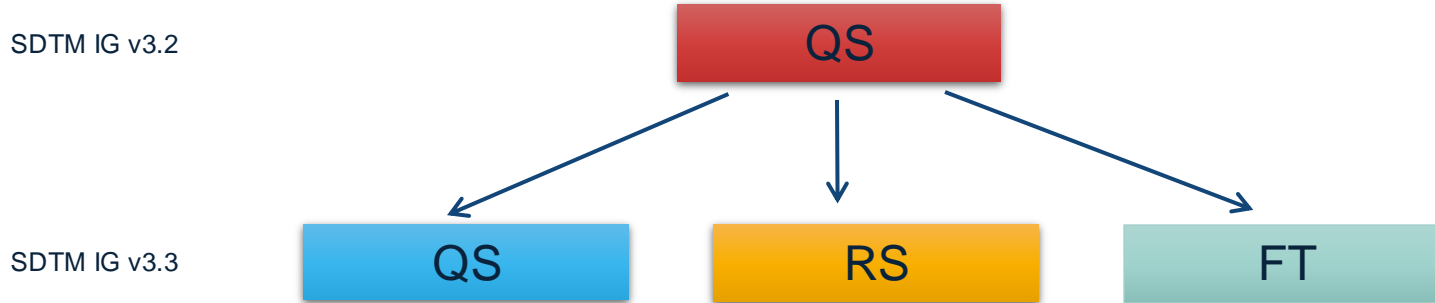
QS – True questionnaire data (questions with responses)

FT – Functional Tests (the subject is asked to perform a task and is evaluated)

RS – Disease Response and Clinical Classification (the assessment of disease response to therapy,

or clinical classification based on published criteria)

cdisc

Questionnaires, Ratings, and Scales (QRS) Domains

qs.xpt

Row	STUDYID	DOMAIN	USUBJID	QSSEQ	QSTESTCD	QSTEST	QSCAT	QSORRES	QSSTRESC	QSSTRESN	QSLOBXFL	VISITNUM	QSDTC
1	STUDYX	QS	P0001	1	FPQ01001	FPQ01-I Like Apples	FRUIT PREFERENCE QUESTIONNAIRE	Strongly Agree	4	4	Y	1	2012-11-16
2	STUDYX	QS	P0001	2	FPQ01002	FPQ01-I Like Oranges	FRUIT PREFERENCE QUESTIONNAIRE	Disagree	1	1	Y	1	2012-11-16
3	STUDYX	QS	P0001	3	FPQ01003	FPQ01-I Like Bananas	FRUIT PREFERENCE QUESTIONNAIRE	Agree	3	3	Y	1	2012-11-16

ft.xpt

Row	STUDYID	DOMAIN	USUBJID	FTSEQ	FTTESTCD	FTTEST	FTCAT	FTORRES	FTORRESU	FTSTRESC	FTSTRESN	FTSTRESU	FTLOBXFL	VISITNUM	FTDTC
1	STUDYX	FT	P0001	1	FYD01001	FYD01-Time	FORTY YARD DASH	5.2	sec	5.2	5.2	sec	Y	1	2012-11-16
2	STUDYX	FT	P0001	2	FYD01001	FYD01-Time	FORTY YARD DASH	5	sec	5	5	sec		2	2012-11-23
3	STUDYX	FT	P0001	3	FYD01001	FYD01-Time	FORTY YARD DASH	4.9	sec	4.9	4.9	sec		3	2012-11-30

rs.xpt

Row	STUDYID	DOMAIN	USUBJID	RSSEQ	RSLNKGRP	RSTESTCD	RSTEST	RSCAT	RSORRES	RSSTRESC	RSEVAL	VISITNUM	VISIT	RSDTC	RSDY
1	ABC	RS	44444	1		TRGRESP	Target Response	RECIST 1.1	PR	PR	INVESTIGATOR	40	WEEK 6	2010-02-18	46
2	ABC	RS	44444	2		NTRGRESP	Non-target Response	RECIST 1.1	SD	SD	INVESTIGATOR	40	WEEK 6	2010-02-18	46
3	ABC	RS	44444	3	A2	OVRLRESP	Overall Response	RECIST 1.1	PR	PR	INVESTIGATOR	40	WEEK 6	2010-02-18	46
4	ABC	RS	44444	4		TRGRESP	Target Response	RECIST 1.1	NE	NE	INVESTIGATOR	60	WEEK 12	2010-04-02	88
5	ABC	RS	44444	5		NTRGRESP	Non-target Response	RECIST 1.1	NE	NE	INVESTIGATOR	60	WEEK 12	2010-04-02	88
6	ABC	RS	44444	6		SYMPTDTR	Symptomatic Deterioration	PROTOCOL DEFINED RESPONSE CRITERIA	Pleural Effusion	PD	INVESTIGATOR	60	WEEK 12	2010-04-02	88
7	ABC	RS	44444	7	A3	OVRLRESP	Overall Response	PROTOCOL DEFINED RESPONSE CRITERIA	PD	PD	INVESTIGATOR	60	WEEK 12	2010-04-02	88

rs.xpt

Row	STUDYID	DOMAIN	USUBJID	RSSEQ	RSTESTCD	RSTEST	RSCAT	RSORRES	RSSTRESC	RSSTRESN	RSLOBXFL	RSEVAL	VISITNUM	RSDTC
1	STUDYX	RS	P0001	1	SSS01001	SSS01-Snoring Volume	SMITH SNORING SCALE	loud	3	3	Y	SPOUSE	1	2012-11-16
2	STUDYX	RS	P0001	2	SSS01002	SSS01-Snoring Extent	SMITH SNORING SCALE	25-50% of sleep time	2	2	Y	SPOUSE	1	2012-11-16
3	STUDYX	RS	P0001	3	SSS01003	SSS01-Snoring Pattern	SMITH SNORING SCALE	very regular	1	1	Y	SPOUSE	1	2012-11-16
4	STUDYX	RS	P0001	4	SSS01004	SSS01-Total Score	SMITH SNORING SCALE	6	6	6	Y	SPOUSE	1	2012-11-16

Decommissioned Domains
Example: Morphology (MO)

SDTM IG v3.2: SDS team found it more difficult to display morphological and physiological findings separately than to add some new values. This led to the decision to expand the body system domain to include all morphological and physiological findings, and removal of the MO domain in a future release.

SDTM IG v3.3: Added statement that this domain will be deprecated in a future version of the SDTMIG. Tests that were represented in the MO domain will be moved to morphology/physiology domains.

SDTM IG v3.4: Decommissioned MO domain

SDTM IG v3.2 *vs* SDTM IG v3.3 & v3.4 Updates

New variables added in existing domains & details added
clarifying derivation of existing variables

The Changes of Subject Visits (SV) Domain

NOV2021
New SV domain released in SDTM IG v3.4

JUL2020
Decided to add information in SV

APR2020
COVID-19 may cause missed visit;
Visit Events (VE) domain: interim solution
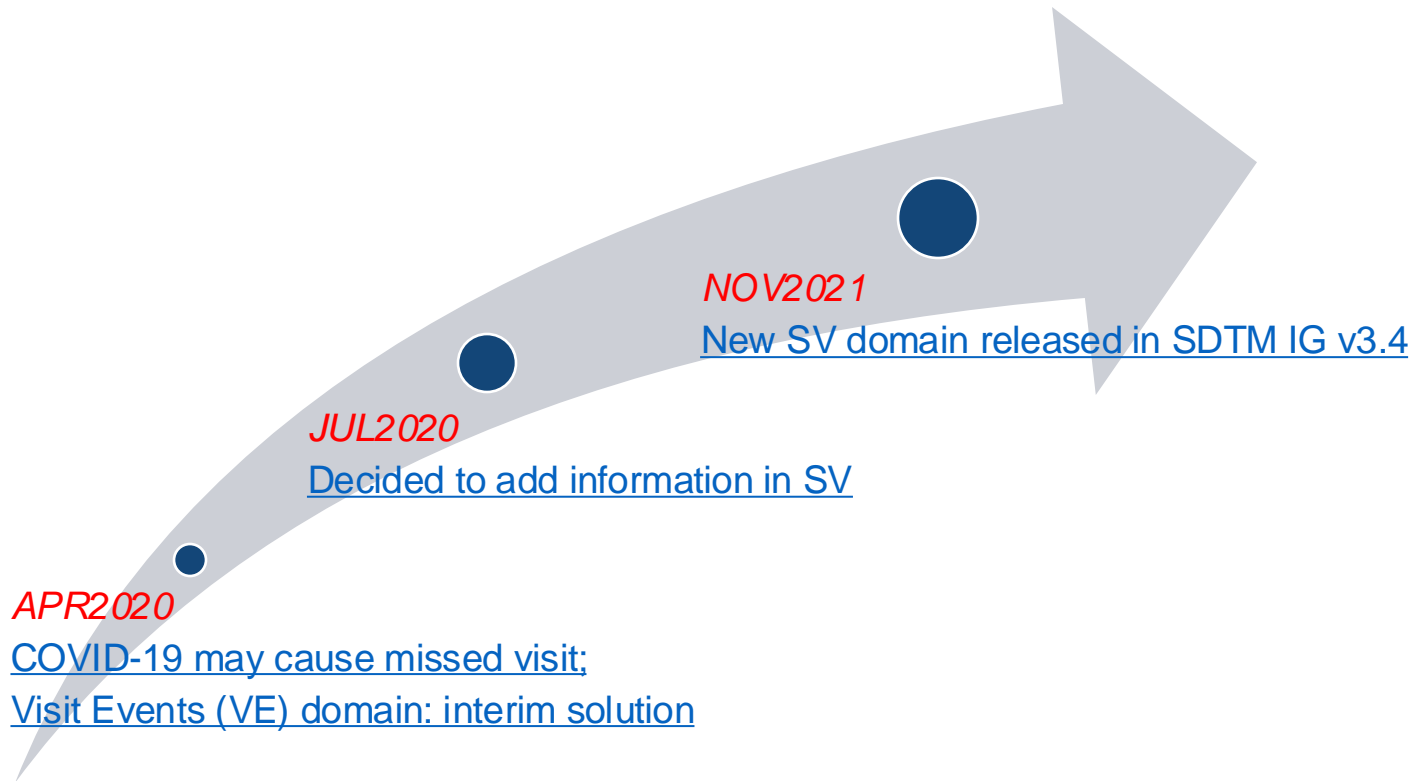
The Changes of Subject Visits (SV) Domain

➡️ Missed visits rationale

Guidance for Ongoing Studies Disrupted by COVID-19 Pandemic standard released in APR2020:

6 Missed Visits

The COVID-19 pandemic may cause missed visits, and in some cases may result in remote visits rather than in-person visits. Although the Protocol Deviations (DV) domain could be used to represent missed visits, the following example uses a custom events domain to document all visits. This approach accommodates amending protocols to alter visit schedules and/or allow remote visits, and permits information about visits before and after any protocol amendments to be represented in a single domain.

Known Issue: The Subject Visits (SV) domain is a standard domain that includes data about visits. However, the SV domain includes only visits that occurred and there is no way within the SDTM model to add variables to the SV domain as either standard variables or supplemental qualifiers. The custom Visit Events (VE) domain shown in the following example is an interim solution. In the future, it is possible that the solution for recording data about visits that did and did not occur may mean using a modified SV domain or using a new domain (e.g., the VE domain).

The Changes of Subject Visits (SV) Domain

➡ To counter this issue, Missed Visits (VE) was proposed by CDISC Team

Missed Visit (VE)

ve.xpt

Row	STUDYID	DOMAIN	USUBJID	VESEQ	VETERM	VEDECOD	VEPRESP	VEOCCUR	VISITNUM	VISIT	VISTDY	VEDTC	VESTDTC	VEDY	VESTDY	VEREASOC	VEEPCHGI	VECNTMOD
1	CVD-3	VE	301	1	ONSITE VISIT	PLANNED VISIT	Y	Y	1	BASELINE	1	2020-03-04	2020-03-04	1	1			
2	CVD-3	VE	301	2	ONSITE VISIT	PLANNED VISIT	Y	Y	2	WEEK 1	8	2020-03-11	2020-03-11	8	8			
3	CVD-3	VE	301	3	REPEAT VISIT FOR ABNORMAL LAB	UNSCHEDULED VISIT			2.5			2020-03-13	2020-03-13	10	10			
4	CVD-3	VE	301	4	ONSITE VISIT	PLANNED VISIT	Y	Y	3	WEEK 2	15	2020-03-19	2020-03-19	16	16			
5	CVD-3	VE	301	5	ONSITE VISIT	PLANNED VISIT	Y	N	4	WEEK 3	22	2020-03-25		22		Subject lacked transportation		
6	CVD-3	VE	301	6	ONSITE VISIT	PLANNED VISIT	Y	N	5	WEEK 4	29	2020-04-01		29		Subject refused due to fear of epidemic	Y	
7	CVD-3	VE	301	7	ONSITE VISIT	PLANNED VISIT	Y	N	6	WEEK 5	26	2020-04-08		36		Hospital restricted access to clinic	Y	
8	CVD-3	VE	301	8	VIRTUAL VISIT	PLANNED VISIT	Y	Y	7	WEEK 6	43	2020-04-15	2020-04-15	43	43		Y	REMOTE AUDIO
9	CVD-3	VE	301	9	VIRTUAL VISIT	PLANNED VISIT	Y	Y	8	WEEK 7	50	2020-04-22	2020-04-22	50	50		Y	REMOTE AUDIO VIDEO
10	CVD-3	VE	301	10	HOSPITAL RESTRICTED ACCESS TO RADIOLOGY DUE TO COVID-19	INCOMPLETE PLANNED VISIT	Y	Y	9	WEEK 8	57	2020-04-30	2020-04-30	59	59		Y	

VE NSV Metadata

Variable	Label	Type	Codelist	Role	Origin	Comment
VEREASOC	Reason for Occur Value	text		Non-Standard Record Qualifier	CRF	
VEEPCHGI	Epi/Pandemic Related Change Indicator	text	NY	Non-Standard Record Qualifier	CRF	
VECNTMOD	Contact Mode	text	CNTMODE	Non-Standard Record Qualifier	CRF	CNTMODE is a sponsor-defined codelist.

The Changes of Subject Visits (SV) Domain

➡️ Later, it was decided to add that information in SV as per study data technical conformance guide released by FDA on JUL2020.

FDA Guidance on Missed Visits

3. **CDISC Document: Guidance for Ongoing Studies Disrupted by COVID-19 Pandemic**

It is the current preference of the Agency that for all clinical studies, not limited to those impacted by COVID-19, subject visit data for scheduled (whether or not they occurred), and unscheduled visits be submitted in one single dataset structured as the current CDISC Subject Visits (SV) domain. It is also Agency preference that three non-standard variables (NSVs) for missed visits, --REASOC (Reason for Occur Value), --EPCHGI (Epi/Pandemic Related Change Indicator), and --CNTMOD (Contact Mode), outlined in the CDISC document "Guidance for Ongoing Studies Disrupted by COVID-19 Pandemic" be included within the SV domain and not within the supplemental SUPPSV domain or in other SDTM datasets. Submitting subject visits information in one single structured dataset allows both the human and technology consumer of this information to operate efficiently and with confidence that all visit data are considered during regulatory review.

Subject Visits (SV) Domain in SDTM IG v3.4

Now contains visits that did not occur as well did occur

SV – Specification

sv.xpt, Subject Visits — Special Purpose. One record per actual or planned visit per subject, Tabulation.

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	SV	Identifier	Two-character abbreviation for the domain most relevant to the observation. The domain abbreviation is also used as a prefix for variables to ensure uniqueness when datasets are merged.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
VISITNUM	Visit Number	Num		Topic	Clinical encounter number. Numeric version of VISIT, used for sorting.	Req
VISIT	Visit Name	Char		Synonym Qualifier	Protocol-defined description of a clinical encounter.	Perm
SVPRESP	Pre-specified	Char	(NY)	Variable Qualifier	Used to indicate whether the visit was planned (i.e., visits specified in the TV domain). Value is "Y" for planned visits, null for unplanned visits.	Exp
SVOCCUR	Occurrence	Char	(NY)	Record Qualifier	Used to record whether a planned visit occurred. The value is null for unplanned visits.	Exp
SVREASOC	Reason for Occur Value	Char		Record Qualifier	The reason for the value in SVOCCUR. If SVOCCUR="N", SVREASOC is the reason the visit did not occur.	Perm
SVCNTMOD	Contact Mode	Char	(CNTMODE)	Record Qualifier	The way in which the visit was conducted. Examples: "IN PERSON", "TELEPHONE CALL", "IVRS".	Perm
SVEPCHGI	Epi/Pandemic Related Change Indicator	Char	(NY)	Record Qualifier	Indicates whether the visit was changed due to an epidemic or pandemic.	Perm
VISITDY	Planned Study Day of Visit	Num		Timing	Planned study day of VISIT. Should be an integer.	Perm
SVSTDTC	Start Date/Time of Observation	Char	ISO 8601 datetime or interval	Timing	Start date/time of an observation represented in ISO 8601 character format.	Exp
SVENDTC	End Date/Time of Observation	Char	ISO 8601 datetime or interval	Timing	End date/time of the observation represented in ISO 8601 character format.	Exp
SVSTDY	Study Day of Start of Observation	Num		Timing	Actual study day of start of observation expressed in integer days relative to the sponsor-defined RFSTDTC in Demographics.	Perm
SVENDY	Study Day of End of Observation	Num		Timing	Actual study day of end of observation expressed in integer days relative to the sponsor-defined RFSTDTC in Demographics.	Perm
SVUPDES	Description of Unplanned Visit	Char		Record Qualifier	Description of what happened to the subject during an unplanned visit. Only populated for unplanned visits.	Perm

Updates to the DM Domain – from SDTM IG v3.3

ARMNRS and **ACTARMUD** have been added to Demography (DM) to provide more information on subjects who did not receive the expected study treatments.

Comparison of study treatment group derivation in SDTM IG v3.2 and SDTM IG v3.3

SDTM IG v3.2			
ARM	**ARMCD**	**ACTARM**	**ACTARMCD**
SCREEN FAILURE	SF	SCREEN FAILURE	SF
NOT ASSIGNED	NS	NOT ASSIGNED	NS
Drug A	ARMCD	NOT TREATED	NT

SDTM IG v3.3					
ARM	**ARMCD**	**ACTARM**	**ACTARMCD**	**ARMNRS**	**ACTARMUD**
				SCREEN FAILURE	
				NOT ASSIGNED	
Drug A	A			NOT TREATED	
Drug A	A			UNPLANNED TREATMENT	Drug B dispensed for part of Drug A element

cdisc

New Baseline Variable Added into Findings Domains
– from SDTM IG v3.3

Example from IG shows a set of similar flag variables and their usage across SDTM and ADaM

Variable	Structure Where It Is Defined	Requirement in That Structure	Definition	Intended Use
-- LOBXFL	SDTM Findings	Expected or Permissible	Last non-missing value prior to RFXSTDTC (Operationally derived)	Consistent pre-treatment reference value baseline for use across all studies and sponsors.
--BLFL	SDTM Findings	Permissible *(formerly expected in some domains)*	A baseline defined by the sponsor (Could be derived in the same manner as --LOBXFL or ABLFL, but is not required to be)	Any sponsor-defined baseline use

Multiple ECG Measures – from SDTM IG v3.3

Example of ECG Test Results (EG)

ROW	USUBJID	EGSEQ	EGTEST	EGORRES	EGORRESU	EGREPNUM	EGDTC
1	001	1	PR Interval, Aggregate	162	msec	1	2019-02-26T07:06
2	001	2	PR Interval, Aggregate	165	msec	2	2019-02-26T07:08
3	001	3	PR Interval, Aggregate	167	msec	3	2019-02-26T07:10

*Note: Some variables (e.g. STUDYID, DOMAIN, EGTESTCD, EPOCH, VISIT and EGDY) are dropped to aid readability.

Sometimes the protocol specifies that the ECG will be measured multiple times for a subject at the same visit/time point, for example the ECG may be measured in triplicate. To properly map this situation, EGREPNUM (Repetition Number) has been added to the EG domain

cdisc

General Assumptions – from SDTM IG v3.3

Guidance on Permissible Variables:

SDTM IG v3.2:

As long as no data was collected for Permissible variables, a sponsor is free to drop them and the corresponding descriptions from the Define-XML.

SDTM IG v3.3:

❑ If a study includes a data item that would be represented in a Permissible variable, then that variable must be included in the SDTM dataset, even if null. Indicate no data were available for that variable in the Define-XML document.

❑ If a study did not include a data item that would be represented in a Permissible variable, then that variable should not be included in the SDTM dataset and should not be declared in the Define-XML document.

SDTM IG v3.2 *vs* SDTM IG v3.3 & v3.4 Updates

Attribute Updates

Attribute Updates

SDTM IG v3.2 *vs* SDTM IG v3.3

|

SDTMIG v3.3 / SDTM v1.7

Summary of changes to the
standard since SDTMIG v3.2

PINNACLE 21

SDTM IG v3.3 *vs* SDTM IG v3.4

Exploring Changes in SDTMIG 3.4 & ADaMIG 1.3 | Pinnacle 21
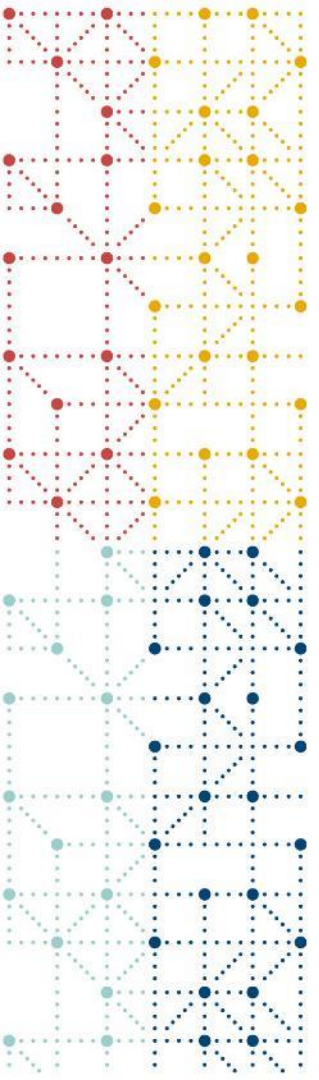
SDTM IG v3.2 *vs* SDTM IG v3.3 & v3.4 Updates

New Section & New Idea

A new idea: Disease Milestones
– from SDTM IG 3.3

❑ Disease milestones are observations or activities which are expected to occur in the course of the disease under observation, and these may occur before the study or during the study.

❑ Subject Disease Milestones (SM) and Trial Disease Milestones (TM) and, in most cases, these two domains should be presented at the same time.

❑ The concept of Disease Milestones has applicability to any type of event-driven data collection. This is in contrast to collection driven by a schedule (e.g., visits, daily diaries).
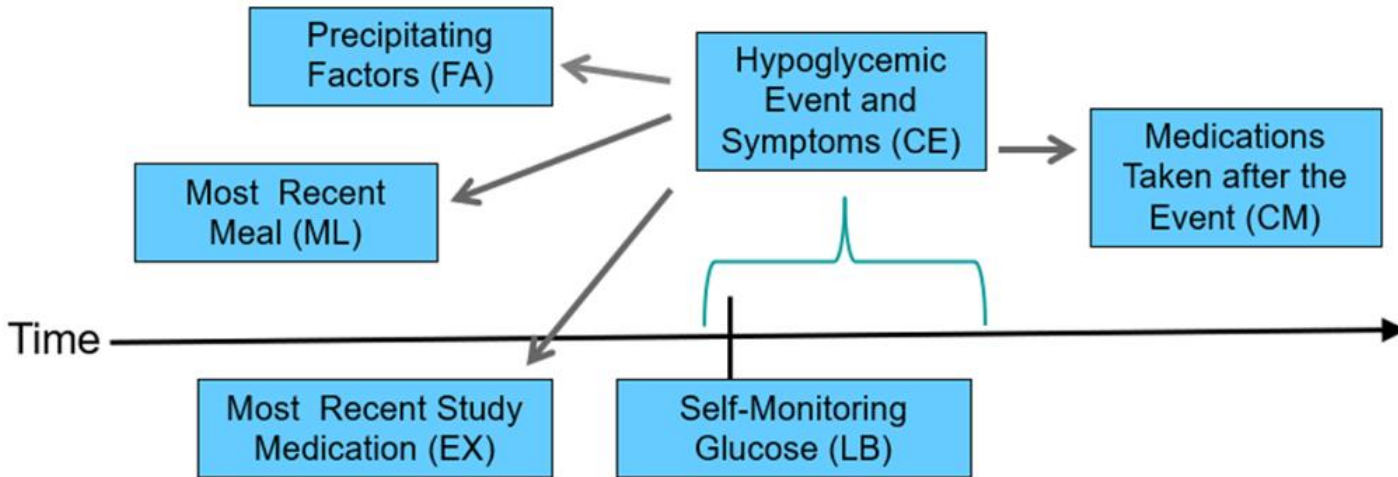
tm.xpt

STUDYID	DOMAIN	MIDSTYPE	TMDEF	TMRPT
ABC	TM	HYPOGLYCEMIC EVENT	Hypoglycemic Event, the occurrence of a blood glucose concentration below the specified (by study) level of hypoglycemia	Y

A new idea: Disease Milestones
– from SDTM IG 3.3

Data Collection for a Hypoglycemic Event

View Recording (gotowebinar.com)

A new section: Section 9 Study References – from SDTM IG 3.3

This section provides structures for representing study-specific terminology used in subject data and describes 3 domains:

1. Section 9.1: Device Identifiers (DI)
2. Section 9.2: Non-host Organism Identifiers (OI)
3. Section 9.3: Pharmacogenomic/Genetic Biomarker Identifiers (PB)

New Class

Summary

Take-Home Knowledge

✓ Mandatory date released from FDA/PMDA for application of SDTM IG v3.3 and v3.4

✓ New domains added for clearer categorization of collected data

✓ New variables added in existing domains to refine the dataset structure and present data more adequately

✓ Added clarifying details of existing variables to provide more standardization and clarity for variables derivation

✓ Introduced new concept "Disease Milestones" to clearly represent event-driven observations of interest

✓ SDTM version and SDTM IG version updated with the development of clinical trials and Pharmaceutical industry

cdisc

Reference

Reference

StudyDataTechnicalConformanceGuide_v4.5.1_July_2020_FINAL.pdf

Adobe Acrobat
Document

Guidance_for_Ongoing_Studies_Disrupted_by_COVID-19.pdf
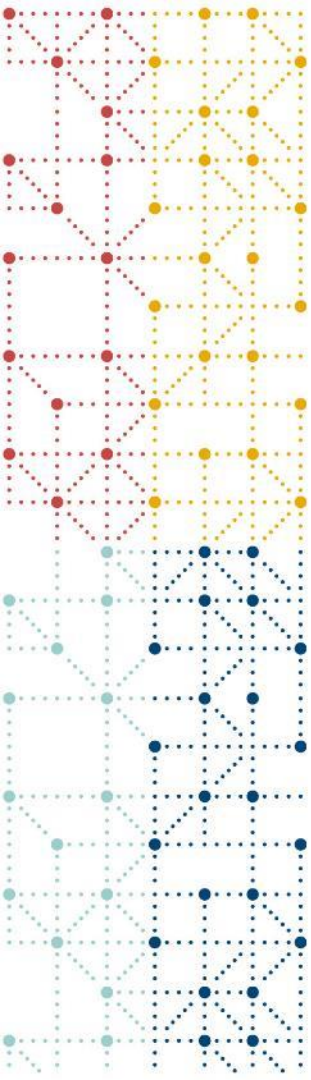
Adobe Acrobat
Document

SDTM IG v3.3

Adobe Acrobat
Document

SDTM IG v3.4

Adobe Acrobat
Document

Thank You!

Contact:

Email: fan-8.yang@novartis.com

cdisc