PharmaSUG 2020 - Paper DS - 110

Demystifying SDTM OE, MI, and PR Domains

Lyma Faroz, Seattle Genetics, Inc., Bothell WA
Sruthi Kola, Seattle Genetics, Inc., Bothell WA

ABSTRACT

CDISC's SDTM IG is an extensive repository of domain metadata that helps organize clinical trial data into relevant and detailed classifications. With rapid advancements in new drug development, patients now have superior and expansive options for treatment. These innovations in medicine necessitate continual updates to the SDTM IG. However, study implementations may not always keep pace with these updates, thereby not fully utilizing valuable resources available through the IG. This paper highlights three such lesser-known SDTM domains which allow statistical programmers to more efficiently structure study data for downstream analysis and submission. We will also share sample CRFs as part of our case study on these domains:

Ophthalmic Examinations (OE)

Added in SDTM IG v3.3, this is part of the Findings class. It contains assessments that measure ocular health and visual status to detect abnormalities in the components of the visual system and determine how well the person can see.

Microscopic Findings (MI)

Also, part of the Findings class, it holds results from the microscopic examination of tissue samples performed on a specimen which is prepared with some type of stain. An example is biomarkers assessed by histopathological examination.

Procedures (PR)

This is part of the Interventions class. This domain stores details of a subject's therapeutic and diagnostic procedures such as disease screening (e.g., mammogram), diagnostic tests (e.g., biopsy), imaging techniques (e.g., CT scan), therapeutic procedures (e.g., radiation therapy), surgical procedures (e.g., diagnostic surgery).

INTRODUCTION

In the fast-paced world of data analysis where every second counts, it is sometimes challenging to keep pace with continuous changes in the SDTM IG and ensure updates that were suggested in newer versions of IG are incorporated in our study datasets. It is important to be aware of newly defined domains so that they can be used efficiently to report clinical trial data. This paper will discuss three such domains: MI (Microscopic Findings) and PR (Procedures) were newly added to SDTM-IG v3.2 and OE (Ophthalmic Examinations) is a more recent addition in SDTM-IG v3.3. Although the PR domain is already a well-established construct, especially in study teams that work in the oncology therapeutic area further awareness of its availability and optimal utilization will benefit many programmers.

BACKGROUND

Possibly because of lesser awareness of these domains it was challenging to find any references beyond the IG. This prompted us to share this paper where we look at these more closely, discuss their limitations, and describe the approach that we used to build these domains in a way that helped us analyze their data more efficiently with fewer workarounds.

Ophthalmic Examinations (OE)

Although the OE domain is of utmost importance to ophthalmic studies, huge amounts of ophthalmic data are also collected in studies in various other therapeutic areas. Hence, it is critical to structure the OE dataset as suggested by the SDTM IG for efficient analysis. Three kinds of eye-related information are often collected:

1. Any information that is collected by performing eye examinations must be stored in the PE (Physical Examination) domain.
2. Any eye-related adverse event should be placed in the AE (Adverse Events) domain and utilize the 'FOCID' identifier variable to store laterality information ('Left' or 'Right' eye) and link it to the OE domain.
3. If an ophthalmological evaluation is performed, it should be mapped to the OE domain, as it stores a subject's ocular health and visual status.

Microscopic Findings (MI)

This domain was originally a part of the IG for the Standard for Exchange of Non-clinical Data (SENDIG) and later was added to SDTM-IG v3.2 in section 6.3 of the Findings General Observation Class. Because information that can be incorporated in this domain may seemingly lend itself to be mapped to the LB domain as well, a careful evaluation should occur to determine where to map such ambiguous information. We will see what kind of data can be placed in this domain in the following section.

Procedures (PR)

Oncology studies collect myriad data to assess the safety profile of a study drug including medical history, adverse events, concomitant medications, prior systemic therapy, prior radiation therapy, prior surgical treatment, etc. Information from prior systemic therapy, prior radiation therapy, and prior surgical treatment CRFs sometimes gets mapped into the CM (concomitant medication) domain. This approach is understandable, as the SDTM IG does describe the CM domain as "CRF data that captures the concomitant and prior medication/**therapies** used by the subject." However, after the addition of the PR domain in SDTM-IG v3.2, its use should be encouraged, and an understanding of this domain should be shared amongst peers and colleagues.

There are three scenarios for which a procedures domain can be used.

1. To map various procedures/therapies directly in PR domain instead of presenting them in CM domain.
2. If a measurement is collected due to a specific procedure, then those results should be stored in the respective findings domain and a record for the procedure should be created in the PR domain. For example, if a biopsy is performed of a tissue sample and histopathological results are obtained, then the biopsy procedure can be mapped to the PR domain and the corresponding result can be mapped to the MI domain.
3. Sometimes information about a test method is collected in the --METHOD variable of a findings domain. This test method may also qualify as a procedure; examples are an MRI or CT scan. SDTM-IG v3.2 suggests capturing records in the PR domain if information such as start and end date or duration of the test is collected along with indicator variables such as PROCCUR (Occurrence), PRPRESP (Pre-specified), PRSTAT (Completion Status), and PRREASND (Reason Not Done). If only findings information is collected, then the decision lies with the sponsor whether to represent that procedure in the PR domain as it is optional to do so.

OPHTHALMIC EXAMINATIONS (OE): STRUCTURE AND IMPLEMENTATION

The OE domain is a Findings domain used for tests that measure a person's ocular health and visual status. It includes tests of visual acuity, color vision, ocular comfort (e.g., dryness, itching), intraocular pressure, etc. The purpose of the OE domain is to collect physiological ophthalmic examinations and the corresponding results only. When specialized ophthalmic examinations are performed in a trial, use the OE domain to map the data collected during these examinations. If eye examinations are performed as a part of general physical examinations, then it should be mapped to the PE domain rather than OE. Any data pertaining to morphological ophthalmic examinations should be mapped into the MO (Morphology) domain and not to OE.

Data Set Structure

OE has one record per ophthalmic finding per method per location, per time point per visit per subject. To uniquely identify each ophthalmic record, a new Identifier variable named FOCID (Focus of Study-specific Interest) was introduced as part of SDTM OE. It is used as a key identifier variable to store information such as right eye (oculus dexter: FOCID=OD), left eye (oculus sinister: FOCID=OS) and both eyes (oculus uterque: FOCID=OU).

Location variables (--LOC, --LAT, --DIR, --PORTOT) are permissible in the OE domain since this information may or may not be collected on a trial.

Although there is a potential duplication of information by having FOCID and --LOC/--LAT as permissible variables, populating these variables with any available data is still helpful in data aggregation and grouping.

SDTM Implementation

As a case study, this paper uses a sample CRF page that consists of assessments to be performed to record the results.

The assessments performed are:

- Visual Acuity
- Schirmer's Test
- Split Lamp
- Intraocular pressure
- Fundoscopy
- Ophthalmology diagnosis

Was an ophthalmology exam performed?	Yes No [SUPPOE]
Date of ophthalmology exam (DD/MMM/YYYY)	[OEDTC]
Visual Acuity	
Visual Acuity Right Eye (OD) [FOCID='OD'] [OELOC='EYE'] [OELAT='RIGHT'] [OETESTCD='VANAOD']	Normal Abnormal Not Done [OEORRES]
Visual Acuity Left Eye (OS) [FOCID='OS'] [OELOC='EYE'] [OELAT='LEFT'] [OETESTCD='VANAOS']	Normal Abnormal Not Done [OEORRES]
Vision changes since last evaluation [OETESTCD='VSCHG]	Yes No Not Applicable [OEORRES]
Schirmer's Test	

Schirmer's Test Right Eye (OD) [FOCID='OD'] [OELOC='EYE'] [OELAT='RIGHT'] [OETESTCD='SCHTOD']	Normal Abnormal Not Done [OEORRES]
Schirmer's Test Left Eye (OS) [FOCID='OS'] [OELOC='EYE'] [OELAT=LEFT'] [OETESTCD='SCHTOS']	Normal Abnormal Not Done [OEORRES]

Table 1. Sample CRF for Visual Acuity and Schirmer's Test

This example shows a general anterior segment examination performed on each eye at one visit, with the purpose of evaluating general abnormalities.

In Table 2 Below, assessments are performed for the left and right eye at each visit. FOCID of OS identifies the records pertaining to the left eye while FOCID of OD identifies the records for the right eye. The variables OELOC and OELAT also help in aggregating the records.

	STUDYID	DOMAIN	USUBJID	FOCID	OESEQ	OETESTCD	OETEST	OECAT	OEORRES	OESTRESC	OELOC	OELAT	VISITNUM	VISIT	OEDTC
2	xxxx	OE	XXX-XXX-001	OD	15	SCHTOD	Schirmer's Test Right Eye (OD)	SCHIRMERS	Normal	Normal	EYE	RIGHT	1	Baseline	2018-11-05
3	xxxx	OE	XXX-XXX-001	OS	33	SCHTOS	Schirmer's Test Left Eye (OS)	SCHIRMERS	Normal	Normal	EYE	LEFT	1	Baseline	2018-11-05
4	xxxx	OE	XXX-XXX-001	OD	20	VANAOD	Visual Acuity Right Eye (OD)	VISACU	Abnormal	Abnormal	EYE	RIGHT	1	Baseline	2018-11-05
5	xxxx	OE	XXX-XXX-001	OS	38	VANAOS	Visual Acuity Left Eye (OS)	VISACU	Normal	Normal	EYE	LEFT	1	Baseline	2018-11-05

Table 2. SDTM OE Data for a Subject with Assessments for Schirmer's Test and Visual Acuity (Left and Right)

All the assessments performed are collected as a test (OETEST/OETESTCD).

The grouping qualifier variable --CAT is used to indicate the function corresponding to the object of the test and --SCAT is used for further sub-classification, if available.

In our current sample, CRF OECAT would be populated with Visual Acuity, Schirmer's Test, Split Lamp, Intraocular pressure and Fundoscopy.

OESCAT may be populated with, for example, High Contrast or Low Contrast if such further qualifications are collected.

MICROSCOPIC FINDINGS (MI): STRUCTURE AND IMPLEMENTATION

This domain was originally a part of SENDIG and later was added to SDTM-IG v3.2 in Section 6.3 of the Findings General Observation Class. This domain model provides a record for each microscopic finding observed. It contains results from the microscopic examination of tissue samples performed on a specimen which is prepared with some type of stain. For example, any histologic or histopathological examination, such as the immunohistochemistry IHC method which treats a tissue with a stain that adheres to very specific substances, and its results should be stored in the MI domain. Special attention should be given to tests that could also be mapped to other domains such as LB. For example, examination of cells in fluid specimens should be mapped to LB if the fluid specimen is blood or urine.

Data Set Structure

The MI domain follows the data structure of one record per finding per specimen per subject. Important variables taken from SDTM-IG v3.2 are shown in Table 3.

MITESTCD	Microscopic Examination Short Name

MITEST	Microscopic Examination Name
MITSTDTL	Microscopic Examination Detail This variable is new in SDTM-IG v3.2 and implemented in the MI domain for the first time. It provides further details to the test performed, such as reaction score, intensity score or percentage of specific proteins (e.g., biomarkers of interest) present on the tissue sample, etc.
MIRESCAT	Result Category Categorize the result, such as if the finding is 'malignant' or 'benign'
MISPEC	Specimen Material Type Kind of specimen used for performing the test. Examples include 'tissue', 'bone marrow', etc.
MISPCCND	Specimen Condition Specimen created by performing any specific methods, example: 'AUTOLYZED'
MIMETHOD	Method of Test or Examination Technique used, or type of stain used for slides. For example: IHC (immunohistochemistry).

Table 3. List of Important variables in the MI Domain

SDTM Implementation

We'll present a case study of a biomarker of interest named Tissue Factor (TF), a transmembrane glycoprotein that is expressed in larger amounts during oncological transformation on the membranes of neoplastic cells and tumor- associated endothelial cells. In order to understand the relationship between this biomarker protein and cancer indication, a tumor tissue sample was collected. The IHC (immunohistochemistry) method was used to stain the tissue sample and results were reported for staining intensity of cytoplasm and membrane that ranged from 0 to 3+. The percentage of cells for each intensity level was reported and used to calculate an H-score ranging from 0 to 300. This information could have been easily been understood to map to the LB domain, as the dataset is received along with other lab parameters. However, while referring to SDTM-IG v3.2, it became apparent that this kind of data had an entire domain that could be used.

STUDYID	DOMAIN	USUBJID	MISEQ	MITESTCD	MITEST	MITSTDTL	MIORRES	MIORRESU	MISTRESC	MISTRESN	MISTRESU	MISPEC	MIMETHOD	VISIT
XYZ	MI	XYZ-1001	1	FACTIII	Factor III	Percentage of cells with 0 intensity of staining	10	%	10	10	%	Slide	IHC	Screening
XYZ	MI	XYZ-1001	2	FACTIII	Factor III	Percentage of cells with 1+ intensity of staining	5	%	5	5	%	Slide	IHC	Screening
XYZ	MI	XYZ-1001	3	FACTIII	Factor III	Percentage of cells with 2+ intensity of staining	80	%	80	80	%	Slide	IHC	Screening
XYZ	MI	XYZ-1001	4	FACTIII	Factor III	Percentage of cells with 3+ intensity of staining	5	%	5	5	%	Slide	IHC	Screening
XYZ	MI	XYZ-1001	5	FACTIII	Factor III	H-score of staining	180	%	180	180	%	Slide	IHC	Screening

Table 4. Presenting the Mapped SDTM MI Data per SDTM-IG v3.2

The intensity level was assessed in the cytoplasm. Since this information is important and cannot be placed in any available variable in the MI domain, it can be stored in the SUPPMI domain and represented in QNAM as CELLOC (Cellular Location).

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG	QEVAL
XYZ	MI	XYZ-1001	MISEQ	1	CELLOC	Cellular Location	CYTOPLASM	CRF Page	
XYZ	MI	XYZ-1001	MISEQ	2	CELLOC	Cellular Location	CYTOPLASM	CRF Page	
XYZ	MI	XYZ-1001	MISEQ	3	CELLOC	Cellular Location	CYTOPLASM	CRF Page	
XYZ	MI	XYZ-1001	MISEQ	4	CELLOC	Cellular Location	CYTOPLASM	CRF Page	
XYZ	MI	XYZ-1001	MISEQ	5	CELLOC	Cellular Location	CYTOPLASM	CRF Page	

Table 5. SUPPMI

PROCEDURES DOMAIN (PR): STRUCTURE AND IMPLEMENTATION

The Procedures domain belongs to the interventions class which stores collected data about a subject's therapeutic and diagnostic procedures. Any procedure-related data except measurements or results, whether therapeutic or diagnostic, can be stored in this domain. Careful consideration should be given to measurements or results collected from a procedure: these should be stored in the appropriate findings domain. SDTM-IG v3.2 gives some example procedures which are:

1. Disease screening (e.g., mammogram, pap smear)
2. Endoscopic examinations (e.g., arthroscopy, diagnostic colonoscopy, therapeutic colonoscopy, diagnostic laparoscopy, therapeutic laparoscopy)
3. Diagnostic tests (e.g., amniocentesis, biopsy, catherization, cutaneous oximetry, finger stick, fluorophotometry, imaging techniques (e.g., DXA scan, CT scan, MRI), phlebotomy, pulmonary function test, skin test, stress test, tympanometry)
4. Therapeutic procedures (e.g., ablation therapy, catheterization, cryotherapy, mechanical ventilation, phototherapy, radiation therapy/radiotherapy, thermotherapy)
5. Surgical procedures (e.g., curative surgery, diagnostic surgery, palliative surgery, therapeutic surgery, prophylactic surgery, resection, stenting, hysterectomy, tubal ligation, implantation)

Data Set Structure

The data structure of PR is one record per recorded procedure per occurrence per subject. A list of important variables taken from the SDTM-IG v3.2 are shown in Table 6.

PRLNKID	Link ID This variable can be used to link relationship between records or relationship between PR and another findings domain if the measurement for a procedure is represented in such other domains
PRTRT	Reported Name of Procedure
PRDECOD	Standardized Procedure Name
PRPRESP	Pre-specified
PROCCUR	Occurrence This variable is populated if pre-specified (PRPRESP) has a value of 'Y' When ADaM endpoints that look for the absence of a given procedure are designed, it is tempting to look for PROCCUR = N. However, PROCCUR is used only when there are pre-specified procedures, hence looking for PROCCUR=N often yields incorrect results as there will be no such records in SDTM for subjects who said on the CRF they didn't have the procedure. Thus, the correct strategy for such endpoints is to look for the absence of an SDTM record with that specific test, rather than the presence of a record with PROCCUR=Y.

Table 6. List of Important Variables in the PR Domain

SDTM Implementation

Example 1:
Information regarding prior surgical treatment, prior radiation therapy, and prior systemic therapy was collected for a patient before enrollment. Two example CRFs of surgical treatment and radiotherapy are as follows,

Prior Surgical Treatment [PRCAT]

Has the patient had surgical treatment for this cancer? [NOT SUBMITTED]	Yes No
Description of surgery	[PRTRT]

Site of surgery [PRSITE in SUPPPR]	Primary tumor Metastatic site for palliative/symptom management
Date of surgery (DD/MMM/YYYY)	[PRSTDTC] [PRENDTC]

Prior Radiation Therapy [PRCAT] [PRTRT = 'Radiotherapy']

Has the patient had any prior radiotherapy for this cancer? [NOT SUBMITTED]	Yes No
Site of radiation [PRSITE in SUPPPR]	Primary tumor Metastatic site for palliative/symptom management
Site/Location [PRLOC]	Mediastinum Lung Liver CNS-whole brain CNS-localized Bone Lymph node Other
If other, please specify [PRLOCOTH in SUPPPR]	
Date therapy started (DD/MMM/YYYY)	[PRSTDTC]
Date therapy stopped (DD/MMM/YYYY)	[PRENDTC]

Table 7. Sample CRF for Prior Surgical Treatment and Prior Radiation Therapy

STUDYID	DOMAIN	USUBJID	PRSEQ	PRTRT	PRCAT	PRLOC	PRSTDTC	PRENDTC
XYZ	PR	XYZ-10001	1	COLONOSCOPY	PRIOR SURGICAL TREATMENT		1998-12-24	1998-12-24
XYZ	PR	XYZ-10001	2	GASTROSCOPY	PRIOR SURGICAL TREATMENT		2018-10-08	2018-10-08
XYZ	PR	XYZ-10001	3	RADIOTHERAPY	PRIOR RADIATION THERAPY	LUNG	2019-01-14	2019-01-17

Table 8. SDTM PR Domain for a Subject With Prior Surgical Treatment and Radiation Therapy

Example 2:
For cases where procedure information is collected along with measurements/results, a PR record is created for the performed procedure and the corresponding measurement is represented in a findings domain. Table 9 below shows a biopsy procedure represented in the PR domain and Table 10 shows the corresponding histopathological measurements performed by immunohistochemistry in the MI domain. Records for both data sets are placed in the RELREC domain linked with an identifier variable such as PRLINKID from PR and MILINKID from MI.

STUDYID	DOMAIN	USUBJID	PRSEQ	PRLINKID	PRTRT	PRLOC	PRSTDTC	PRENDTC
XYZ	PR	XYZ-10002	1	XYZ_ABC_10002	BIOPSY	LUNG	2014-01-14	2014-01-14

Table 9. PR Dataset

STUDYID	DOMAIN	USUBJID	MISEQ	MILINKID	MITESTCD	MITEST	MITSTDTL	MIORRES	MIORRESU	MIMETHOD	VISIT
XYZ	MI	XYZ-10002	1	XYZ_ABC_10002	FACTIII	Factor III	Percentage of cells with 0 intensity of staining	10	%	IHC	Screening
XYZ	MI	XYZ-10002	2	XYZ_ABC_10002	FACTIII	Factor III	Percentage of cells with 1+ intensity of staining	5	%	IHC	Screening
XYZ	MI	XYZ-10002	3	XYZ_ABC_10002	FACTIII	Factor III	Percentage of cells with 2+ intensity of staining	80	%	IHC	Screening
XYZ	MI	XYZ-10002	4	XYZ_ABC_10002	FACTIII	Factor III	Percentage of cells with 3+ intensity of staining	5	%	IHC	Screening
XYZ	MI	XYZ-10002	5	XYZ_ABC_10002	FACTIII	Factor III	H-score of staining	180		IHC	Screening

Table 10. MI Dataset With IHC Results for a Tissue Sample Collected via a Biopsy Procedure

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
XYZ	PR	XYZ-10002	PRLINKID		ONE	1
XYZ	MI	XYZ-10002	MILINKID		MANY	1

Table 11. RELREC Dataset Showing the Link Between PR and MI

CONCLUSION

Utilizing SDTM-IG v3.2 and v3.3 to design these databases helps to create SDTM datasets efficiently and makes the review process easy by clear traceability. Awareness of these newly added domains in SDTM-IG v3.2 and v3.3 along with an understanding of when and how to use them especially for data that could, at first glance, deceivingly lend itself to be mapped into other more established domains will benefit higher-quality SDTM production and submission.

REFERENCES

CDISC Study Data Tabulation Model (SDTM) v1.4, Study Data Tabulation Model Implementation Guide (SDTMIG) v3.2 & v3.3
https://www.cdisc.org/standards/foundational/sdtm

Alyssa Wittle, Christine McNichol, Antonio Cardozo, 2016. "Moving up! – SDTM v 3.2 – What is new and how to use it" PharmaSUG 2016
https://www.pharmasug.org/2016-proceedings.html#DS

Fred Wood, Diane Wold, Rhonda Faile, Wayne Kubick, 2014. "Data Standards Development for Therapeutic Areas: A Focus on SDTM-Based Datasets"
https://www.pharmasug.org/2014-proceedings.html#DS

Marcelina Hungria. "Ophthalmic Examinations (OE) Domain Overview and Status"
https://www.phusewiki.org/docs/New%20Jersey%20SDE%20Presentations%202014/01%20-%20Presentation(Marcelina%20Hungria-DIcore%20group).pdf

ACKNOWLEDGMENTS

We would like to acknowledge our manager Shefalica Chand and John Shaik for inspiring us to contribute to PharmaSUG and for their constant support.

CONTACT INFORMATION

Your comments and questions are valued and encouraged.

Contact the authors at:

Lyma Faroz
Seattle Genetics, Inc.
21823 - 30th Drive S.E.
Bothell, WA 98021
lfaroz@seagen.com

Sruthi Kola
Seattle Genetics, Inc.
21823 - 30th Drive S.E.
Bothell, WA 98021
srukola@seagen.com