

SDTM/ADaM IG Nuances

Question: Adverse Event Analysis Dataset Sub-Classification

Please confirm instances that would define the OCCDS analysis dataset to be the Adverse Event sub-classification.

	Adverse Event sub-class under OCCDS
ADAE only contains the data from AE, SUPPAE and ADSL	Yes
ADAE only contains the data from AE, SUPPAE, FA and ADSL	Yes
ADAE only contains the data from AE, SUPPAE, CM and ADSL	No

PHUSE Team Response: 02 September 2025

The sub-class definition stated in the CDISC ADaM OCCDS IG v1.1 is:

“

3.1.2 SubClass ADVERSE EVENT

The intent of the ADVERSE EVENT SubClass is to have a consistent way to represent data needed for typical adverse event analyses. Examples from Sections 4-9 can be produced from a dataset that is of SubClass ADVERSE EVENT.

Datasets in the SubClass ADVERSE EVENT must have a Class of OCCURRENCE DATA STRUCTURE: All the principles described in Section 1.1, Purpose, must be met, and the structure is usually 1 record per each record in the corresponding SDTM domain. Additionally,

- The SDTM input dataset for the ADVERSE EVENT SubClass is always AE, with some additional information from SUPPAE, FA, and ADSL.
- Data in other event domains, such as Medical History (MH) or Clinical Events (CE), are **not** included in the ADVERSE EVENT SubClass.

When adverse event-related information is collected in the Findings domain, every record in SubClass ADVERSE EVENT will have an AESEQ, and records from FA will also have a unique identifier variable, such as FASEQ or FASPID, for traceability.

Not all OCCDS datasets that contain adverse event data will necessarily be of SubClass ADVERSE EVENT. In the example in Section 10, Example 7: Analysis of Adverse Events from Multiple Input Domains, the OCCDS dataset contains input

SDTM/ADaM IG Nuances

rows from CE in addition to AE. Although this is an OCCDS dataset, it is not of SubClass ADVERSE EVENT.

1.1 Purpose

The statistical analysis data structure presented in this document describes the general data structure and content typically found in occurrence analysis. *Occurrence analysis* is the counting of subjects with a given record or term, and often includes a structured hierarchy of dictionary coding categories. Examples of data that fit into this structure include those used for typical analysis of adverse events, concomitant medications and medical history. The structure is based on the Analysis Data Model (ADaM) v2.1 and the ADaM Implementation Guide (ADaMIG) v1.2, available at <https://www.cdisc.org/standards/foundational/adam>.

As presented in the ADaMIG, many analysis methods can be performed using the ADaM Basic Data Structure (BDS), including Parameter (PARAM) and Analysis Value (AVAL). However, data analyzed as described above do not fit well into the BDS and are more appropriately analyzed using a Study Data Tabulation Model (SDTM) structure with added analysis variables.

”

Per the description of adverse event sub-class in the OCCDS IG, the required variables needed to classify the sub-class to the adverse event should first be included in the analysis dataset, and, without them, the sub-class of adverse event would not be valid. The following can help determine the sub-classification:

1. If an ADAE only contains the data from AE, SUPPAE and ADSL, then this ADAE is an adverse event sub-class under OCCDS.
2. If an ADAE only contains the data from AE, SUPPAE, FA and ADSL, then this ADAE is an adverse event sub-class under OCCDS.
3. If an ADAE only contains the data from AE, SUPPAE, CM and ADSL, then this ADAE is not an adverse event sub-class under OCCDS. This assumes that records from the CM domain are present in the analysis dataset and have

SDTM/ADaM IG Nuances

generated additional records in the analysis dataset that cannot be traced back to the AE domain.

Question

Use of Different Versions of MedDRA SMQs and CMQs for a Single Study

For a study, when the latest SMQ version is 27_1 and the latest CMQ version is 27_0_5, what MedDRA version do we need to use – 27.0 or 27.1?

PHUSE Response: 19 November 2024

The Standardised MedDRA Queries (SMQs) are published by MedDRA to match each new version of the dictionary. The customised queries (often referred to as CQ variables in CDISC ADaM standards) are usually maintained by individual sponsors, and it is up to the sponsor to maintain or upversion them to match the MedDRA dictionary version the sponsor is using.

Question

Missing Doses in the SDTM EX Domain

The SDTMIG v3.4 EX assumption 6.a states: “EX contains medications received; **the inclusion of administrations not taken, not given, or missed is under evaluation.**”

Does anyone know if there’s been progress on this topic? Will the next version of the IG address the inclusion of missing doses in EX?

PHUSE Response: 22 October 2024

EX is designed to capture what was taken, while any missed dose should be included in the EC domain only if the CRF collects for missing dosing. In the event the patient CRF does not collect the missing information, any planned dosing regimen determination can be made in the ADaM dataset for exposure calculation purposes. No derivation is recommended at the SDTM level for missed doses when missed dosing information is not collected.

Question

Collapsed AE Dataset

In an AE dataset, such a scenario exists: multiple AEs are linked together through AEGRPID (group ID, or identifier of linked AE). A collapsed AE record based on these multiple AE records is created. The values of the variables of this collapsed AE record are taken from different records of these multiple AE records. My question is, do we use one ADAE dataset (original records + new collapsed records), or do

SDTM/ADaM IG Nuances

we use two datasets – one original ADAE and one new ADAECLPS (only include the collapsed AE)?

If we use two datasets, how do we express the traceability? [Click image to enlarge:](#)

ROW#	USUBID	AESEQ	AEGRPID	AERTFID	AEDCOD	AESV	AESR	ADACH	ADREL	AEOUT	RECOVERED/NOTRECOVERED	ACTIONR	EPOCH	AESTDTC	AEENDTC	ANALDTC
1	M20K21-115012-023001	01	022001007	023001017	Neutropenia	N		MULTIPLE	MULTIPLE		RECOVERED/NOTRECOVERED		1	TREATMENT	2/8/2024	2/8/2024
2	M20K21-115012-023001	01	022001007	023001040	Neutropenia	N		MULTIPLE	MULTIPLE		NOT RECOVERED/NOT RECOVERED		2	TREATMENT	2/8/2024	2/8/2024
3	M20K21-115012-023001		022001007		Neutropenia	N	per the logic	per the logic	per the logic			high	TREATMENT	2/8/2024	2/8/2024	

PHUSE Response: 22 October 2024

Per the CDISC SDTMIG v3.3, it is acceptable to collapse AE records in SDTM. SDTM IG page 137 provides the following: [Click image to enlarge:](#)

Domain and in Appendix C2, [Supplemental Qualifiers Name Codes](#).

- d. In studies using toxicity grade according to a standard toxicity scale such as Common Terminology Criteria for Adverse Events v3.0 (CTCAE), published by the NCI (National Cancer Institute) at https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcae3.pdf, AETOXGR should be used instead of AESEV. In most cases, either AESEV or AETOXGR is populated but not both. There may be cases when a sponsor may need to populate both variables. The sponsor is expected to provide the dictionary name and version used to map the terms utilizing the external codelist element in the Define-XML document.

- e. AE Structure

The structure of the AE domain is one record per adverse event per subject. It is the sponsor's responsibility to define an event. This definition may vary based on the sponsor's requirements for characterizing and reporting product safety and is usually described in the protocol. For example, the sponsor may submit one record that covers an adverse event from start to finish. Alternatively, if there is a need to evaluate AEs at a more granular level, a sponsor may submit a new record when severity, causality, or seriousness changes or worsens. By submitting these individual records, the sponsor indicates that each is considered to represent a different event. The submission dataset structure may differ from the structure at the time of collection. For example, a sponsor might collect data at each visit in order to meet operational needs, but submit records that summarize the event and contain the highest level of severity, causality, seriousness, etc. Examples of dataset structure:

1. One record per adverse event per subject for each unique event. Multiple adverse event records reported by the investigator are submitted as summary records "collapsed" to the highest level of severity, causality, seriousness, and the final outcome.
2. One record per adverse event per subject. Changes over time in severity, causality, or seriousness are submitted as separate events. Alternatively, these changes may be submitted in a separate dataset based on the Findings About Events and Interventions model (see Section 6.4, [Findings About Events or Interventions](#)).
3. Other approaches may also be reasonable as long as they meet the sponsor's safety evaluation requirements and each submitted record represents a unique event. The domain-level metadata (see Section 3.2, [Using the CDISC Domain Models in Regulatory Submissions — Dataset Metadata](#)) should clarify the structure of the dataset.

7. Use of EPOCH and TAETORD

When EPOCH is included in the Adverse Event domain, it should be the Epoch of the start of the adverse event. In other words, it should be based on AESTDTC, rather than AEENDTC. The computational method for EPOCH in the Define-XML document should describe any assumptions made to handle cases where an adverse event starts on the same day that a subject starts an Epoch, if AESTDTC and SESTDTC are not captured with enough precision to determine the epoch of the onset of the adverse event unambiguously. Similarly, if TAETORD is included in the Adverse Events domain, it should be the value for the start of the adverse event, and the computational method in the Define-XML document should describe any assumptions.

Any collapsing methodology for severity, causality, seriousness, action taken, and final outcome should be stated in the study data reviewer's guide (cSDRG). Sometimes there is no way to show traceability from multiple records.

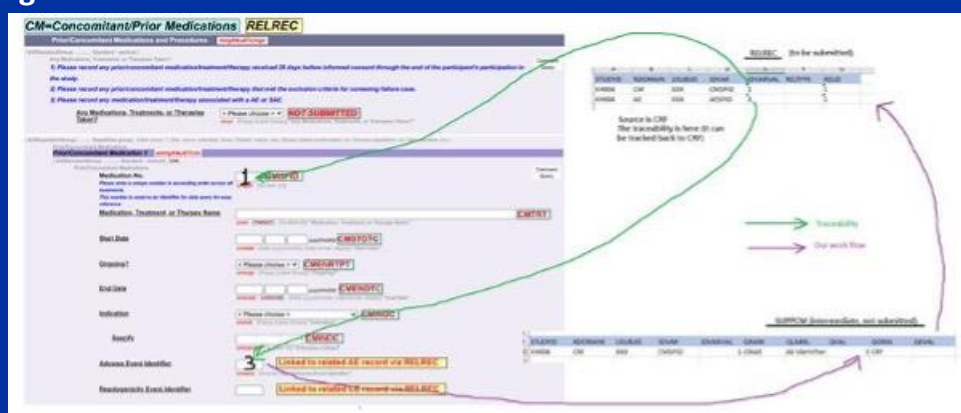
SDTM/ADaM IG Nuances

Therefore, it is not necessary to create the collapsed record in the ADaM dataset. One ADAE ADaM dataset that includes all records from the SDTM AE, including the original records and the collapsed records, is recommended.

Note that some regulatory agencies or divisions within regulatory agencies may have specific requirements for submitting unique adverse events. For example, CBER (Submitting Study Datasets for Vaccines to the Office of Vaccines Research and Review – Guidance for Industry) asks for the AE to contain the collapsed/summary record while the ‘day-to-day’ details go to the FAAE. Sponsors should ensure any such requirements relevant to their compound are taken into consideration before submission and should contact the regulatory agencies for clarification of such requirements.

Question: RELREC Implementation for Medications Prescribed for an Event

While programming CM/RELREC, is it feasible to use the CM.sasprogram to generate an ‘intermediate SUPPCM dataset’ containing linkage information (SUPPCM.CMAE/AE Identifier, SUPPCM.CMMH/MH Identifier) and subsequently use this ‘intermediate SUPPCM dataset’ to create the RELREC domain? **Click image to enlarge:**



Meanwhile, the ‘final/formal SUPPCM’ intended for submission won’t contain those linkage variables (SUPPCM.CMAE/AE Identifier, SUPPCM.CMMH/MH Identifier). The RELREC would still be based on CRF collect data in this case. Would this method pose any risks concerning CDISC compliance or result in an incorrect programming process according to FDA requirements?

SDTM/ADaM IG Nuances

I am exploring if we can decrease the number of qualifiers in the SUPPCM by implementing this method, even though the program CM.sas would create two SUPPCM datasets.

PHUSE Response: 08 October 2024

In general, the AE identifier would not be stored in the SUPPCM domain, particularly when the AE identifier can be stored as the LINKID in the CM domain. However, there is no harm in adding it to the SUPPCM. Ensure the RELREC is sourced from the parent domains (e.g. CM and AE).

Question

Is there a recommended standard for how sponsor organisations should be handling the mapping of inclusion/exclusion criteria into the SDTM IE domain? Should 'like' or 'similar' inclusion/exclusion criteria be mapped into a similar IETESTCD?

PHUSE Team Response: 25 April 2023

The SDTM TI and IE domains together reflect the inclusion and exclusion criteria data for any given study. The TI SDTM domain should reflect what is/was in the protocol at the time (assuming different versions are present due to amendments). The SDTM IG v3.4 mentions the following assumptions for the TI domain in section 7.4.1 with respect to protocol amendments:

1. If inclusion/exclusion criteria were amended during the trial, then each complete set of criteria must be included in the TI domain. TIVERS is used to distinguish between the versions.
2. Protocol version numbers should be used to identify criteria versions, though there may be more versions of the protocol than versions of the inclusion/exclusion criteria. For example, a protocol might have versions 1, 2, 3 and 4. Still, if the inclusion/exclusion criteria in version 1 were unchanged through versions 2 and 3, and only changed in version 4, then there would be two sets of inclusion/exclusion criteria in TI – one for version 1 and one for version 4.
3. Individual criteria do not have versions. If a criterion changes, it should be treated as a new criterion, with a new value for IETESTCD. If criteria have been numbered and values of IETESTCD are generally of the form INCL00n or EXCL00n.

SDTM/ADaM IG Nuances

New versions of a criterion have not been assigned new numbers; separate values of IETESTCD might be created by appending letters, e.g., INCL003A, INCL003B.

There are no additional expectations from the regulatory agencies. The FDA expects sponsors to manage updates to the inclusion/exclusion criteria.

Question

Historical Data Consideration in the SV Domain (under SDTMIG v3.4)

Assumption 13 under SDTM IG v3.4 for the SV domain states: “Therefore, dates before informed consent are not part of the determination of SVSTDTC.” However, some protocols permit historical results within a specified period (e.g., test results within 4 weeks before the informed consent date) as valid screening results. Protocols also allow for the collection of these pre-screening tests, primarily for verifying inclusion/exclusion criteria (e.g., a specific gene mutation test performed two or three years before informed consent for a study). If we don’t consider these dates as the determination of SVSTDTC, VISIT in that particular domain will have to be set to null. Is this reasonable?

PHUSE Team Response: 09 December 2022

Pre-study findings, such as tests performed at the time the disease was diagnosed, can be assigned to the initial screening visit. In this case, the content of the visit variable represents the visit when the test result was recorded in the CRF. The date of the test (or sample collection date) will be stored in the –DTC variable of the applicable domain (e.g. MIDTC).

In cases where historical data is stored as a finding, these historical test/sampling dates should not be taken into account when populating SVSTDTC for the particular visit.

In your case, you can set MI.VISIT to ‘Screening’, MIDTC=date of test, and SV.SVSTDTC will be the date of the first day of the screening visit and will not include MIDTC.

References:

SDTM/ADaM IG Nuances

CDISC guidelines: <https://www.cdisc.org/kb/articles/sdtm-timing-variables-pre-study-findings>

Assumptions 13 for the SV domain in SDTMIG v3.4:

“13. Algorithms for populating SVSTDTC and SVENDTC from the dates of assessments performed during a visit may be particularly challenging for screening visits, as baseline values collected at a screening visit are sometimes historical data from tests performed before the subject began screening for the trial. **Therefore, dates before informed consent are not part of the determination of SVSTDTC.**”

Question

How should the sex of transgender patients be collected and analysed in clinical trials? Should the sex at birth be collected only, or should the gender preference also be collected? Which laboratory normal ranges should be assigned to transgender patients' laboratory test results? How does hormone therapy affect data collection and/or analysis for transgender patients?

PHUSE Team Response: 30 June 2022

The CDISC CDASH team is currently working to publish either an updated guidance or white paper **planned for 2025** on recommendations for capturing the sex for transgender patients. In the draft version, the recommendation would be to collect a two-stage question (note that the controlled terminology and collection text are a draft stage and not finalised): 1. “Sex at Birth” (Male | Female | Don't know | Prefer not to answer) and 2. “Sexual Identity” (Male | Female | Intersex | Transgender | ... | Don't know | Prefer not to answer | Self-describe). In the interim, each sponsor should determine how the data should be collected. It is recommended to provide clarity on the definition of each question, perhaps within the CRF Completion Guidelines. For example, does Sex at Birth pertain to sex stated on the birth certificate, and how to complete the data entry if a patient does not have a birth certificate?

The following articles may be reviewed to determine how hormone therapy affects laboratory results and, in general, analysis for transgender subjects:

1. “Common Hormone Therapies Used to Care for Transgender Patients Influence Laboratory Results”, Humble, R. et al, 2018, *American Association for Clinical Chemistry*.

SDTM/ADaM IG Nuances

2. "Interpreting Laboratory Results in Transgender Patients on Hormone Therapy", Roberts, T. et al, 2014, *The American Journal of Medicine*.
3. "Impact of Hormone Therapy on Laboratory Values in Transgender Patients", SoRelle, J. et al, 2019, *Clinical Chemistry*.
4. "Approach to Interpreting Common Laboratory Pathology Tests in Transgender Individuals", Cheung, A. et al, 2021, *The Journal of Clinical Endocrinology & Metabolism*.
5. "Endocrine Treatment of Gender-Dysphoric/Gender-Incongruent Persons: An Endocrine Society* Clinical Practice Guideline", Hembree, W. et al, 2017, *The Journal of Clinical Endocrinology & Metabolism*.

Question

How do you proceed in providing the reason for the missing code? Do you collect the reason for the missing LOINC code, or do you just provide a predetermined reason?

The LOINC working group recommend providing a reason for the missing code in the cSDRG. (See the extracted text from the

Reference: <https://www.fda.gov/media/109376/download>)

For any lab test where a LONIC code is not submitted, the reason for its omission should be noted in the Clinical Study Data Reviewers Guide.

The Working Groups propose that a starter set of reasons be predetermined (perhaps as CDISC terms) for consistency of reporting, including:

The performing laboratory is unable to determine if an appropriate LONIC code exists

Performing laboratory indicates that no appropriate LONIC code currently exists.

The FDA TCG 4.6 recommends providing the LOINC code of the laboratory parameters for studies starting after March 2020, but nothing is mentioned in the case of a missing code.

SDTM/ADaM IG Nuances

PHUSE Team Response: 07 February 2022

If the laboratory hasn't sent the LOINC code, it is recommended to go back to the laboratory to obtain it. According to the team members' experience, the FDA accepts the submission if the laboratory hasn't provided the LOINC code, and it is missing. In the cSDRG, it notes the reason for its missing as "Lab did not provide the code", or as indicated by the LOINC working group's screenshot.

(Reference: <https://www.fda.gov/media/109376/download>)

One solution would be to request the LOINC code from the lab during the study initiation phase; however, it is expected that not all lab tests will have a corresponding LOINC code assigned.

Question

There are a couple of papers that offer guidance for maintaining 1:1 maps between AVAL and AVALC. Things like:

https://www.lexjansen.com/wuss/2017/79_Final_Paper_PDF.pdf

<https://www.pharmasug.org/proceedings/2012/DS/PharmaSUG-2012-DS16.pdf>

However, neither of these papers explains how to consistently create derived records, where AVALC is a rounded version of AVAL, which satisfies the 1-1 criteria. For example, suppose (within a single PARAMCD) I need to compute an average and then present it in a list to 1 decimal place. For example, let's say AVAL = 45.333333. So, for the listing, I want to show 45.3. I've computed an average for another subject, where AVAL = 45.26, which I also want to show as 45.3 in a listing. If AVALC = 45.3 for both records, then this is not a one-to-one mapping. I obviously can't round AVAL, because that would represent a loss of numerical precision in other calculations. One solution might be 'do not populate AVALC, do the rounding when producing the report'. However, this leaves a lot of work in the reporting program if many parameters are to be listed; the programmer would have to determine the rounding on a per-parameter basis. Ideally, the 'heavy lifting' should already have been done at the dataset level.

PHUSE Team Response: 08 July 2020

Rounding values of AVAL for listing purposes - where to do the rounding and how/where to store the rounded value.

Storing a rounded value in AVAL is not good practice, as it typically results in a loss of precision for calculations in the tables. Storing rounded values in AVALC contradicts the ADaM rule that requires a 1:1 mapping of AVAL to AVALC. Also, it is not the intent to store the character version of a numeric analysis value in AVALC.

SDTM/ADaM IG Nuances

AVALC should be populated only when the character value is used for analysis. See ADaM IG v1.1, section 3.3.4, 'PARAM, AVAL, AVALC' paragraph 3.

There is no ADaM guidance on variable naming for variables used solely for listing purposes.

Rounding the analysis result can be done in the listings program. Alternatively, suppose one wants to store the rounded value in the ADaM dataset. In that case, a custom variable can be added with an intuitive meaning, such as LISTVAL, to store the rounded value.

Question

The study treatment regimen will be A-B-C-D; therefore, the planned ARMCD can be ABCD. Most of the patients' actual ARM ACTARMCD will also be ABCD. However, a few patients may skip D or repeat the ABC part, which is either ABC or ABCABCD. Shall we put 'UNPLANN' in the actors or put the real ABC or ABCABCD in the ACTARMCD?

PHUSE Team Response: 09 January 2020

The planned treatment should be reflected in ARMCD/ARM, while the actual regimen received should be reflected in ACTARMCD/ACTARM. In general, TA should reflect the protocol-specified treatment regimens to be administered. If the protocol specified the skipping of a treatment regimen by design, then it is acceptable to find inconsistencies between ARMCD and ACTARMCD. However, these should be noted in the cSDRG and explained in further detail.

Question

In the SV domain, we search all the by-visit source data to obtain the minimum and maximum dates of each CRF visit. If, due to some reason, there is a 1-2 days overlap among two consecutive CRF visits in the SV domain, we can explain in SDRG or always make visits in SV without overlap, which means we assign the overlapped days to 1 CRF visit in SV rather than keep the days in both visits, as the source data shows?

PHUSE Team Response: 09 January 2020

Acceptable to have the overlap on the visits in the SV domain. There will be no P21 consequences due to this, and therefore, it is not necessary to explain further in the DRG. The explanation in the DERG would be left to the Sponsor's determination.

SDTM/ADaM IG Nuances

Question

For the Table like 'Summary of Common ($\geq$ X%) Adverse Events by Overall Frequency', should the flags for common AEs be created in the ADAE dataset?

PHUSE Team Response: 31 July 2019

5% and 2% custom flag variables can be added to ADAE. The derivation of the flag variable depends on the definition in SAP/table Janssen. If the derivation rule is complicated enough, include it in ADAE.

- Have an internal macro to derive the variable with parameter being the x of x%
- Internal macro is a reporting macro, not tied to the ADAE

Other companies are not included in the ADAE and handle it in the table-generating programs.

- Can also explain in the ADRG
- If this table falls into the category of the primary/secondary key safety and efficacy, you will need to submit the program.

Can also be included in the ARM.

Question

The FDA expressed a desire to keep ETCD/ELEMENT to facilitate reviewers in reviewing the data in the 2011 CDER Common Data Standards Issues Document. [LW1] [NN2] However, in all later FDA-published Study Data Technical Conformance Guides up to V4.1, published in 2018, only EPOCH is required.

Epoch, by itself, should have been informative enough. The FDA validator rules V1.2, published in December 2017, still mention that variables requested by the FDA in Policy documents should be included in the dataset, e.g., EPOCH and ELEMENT. Do you know if the FDA still require ELEMENT/ETCD in all domains? If so, I would suggest that the CDISC SDTM team include those two variables in the parent domain, rather than the SUPP domain.

PHUSE Team Response: 04 July 2018

ETCD/ELEMENT Variables:

The reference to the 2011 CDER Common Data Standards Issues document is no longer relevant and has been superseded by the FDA Study Data Technical Conformance Guide. Therefore, any such references must be in alignment with current FDA guidelines. The inclusion of ETCD/ELEMENT within other domains other than those identified within the SDTM/SDTMIG** is not recommended.

SDTM/ADaM IG Nuances

EPOCH Variables:

Section 2.2.5 of the SDTM* allows for the timing variable EPOCH within any of the three general observation class domains, except where explicitly stated otherwise in the SDTMIG. Therefore, EPOCH inclusion to facilitate the recommendations identified in section 4.1.4.1 of the FDA Study Data Technical Conformance Guide** is in alignment with CDISC SDTM/SDTMIG*.

Additional References:

CDISC SDTM V1.4/SDTMIG V3.2

FDA Study Data Technical Conformance Guide V4.1

Question

How should OTHER be represented for variables bound by non-extensible codelists?

PHUSE Team Response: 07 June 2017

Existing SDTMIGs (e.g., v3.1.2, v3.1.3, v3.2) do not explicitly define how "OTHER" should be implemented universally for all non-extensible codelists.

Question

How should MULTIPLE be used for variables bound by non-extensible codelists?

PHUSE Team Response: 07 June 2017

It is recommended to review SDTMIG (v3.1.2, v3.1.3, or v3.2) Section 4.1.3 Coding and Controlled Terminology Assumptions. Furthermore, please also review the existing CDISC Controlled Terminology (CT) and CDISC Questionnaires, Ratings & Scales (QRS) supplements, as well as the related details found on the QRS page (see reference below).

<https://www.cdisc.org/qrs>

Question

What are the best practices for creating CT for/representing questionnaire responses?

PHUSE Team Response: 07 June 2017

It is recommended to review SDTMIG (v3.1.2, v3.1.3, or v3.2) Section 4.1.3 Coding and Controlled Terminology Assumptions. Furthermore, please also review the existing CDISC Controlled Terminology (CT) and CDISC Questionnaires, Ratings &

SDTM/ADaM IG Nuances

Scales (QRS) supplements, as well as the related details found on the QRS page (see reference below).

Additional References:

<https://www.cdisc.org/qrs>

Question

What is the general recommendation/approach for generating/submitting custom domains (e.g. non-standard CDISC SDTM domains) to regulatory agencies?

PHUSE Team Response: 12 September 2017

According to the CDISC SDTM IG version 3.2, a sponsor should submit the domain datasets that were actually collected (or directly derived from the collected data) for a given study. Decisions on what data to collect should be based on the scientific objectives of the study, rather than what is present in SDTM. Note that any data collected and to be submitted in an analysis dataset must also appear in the tabulation dataset.

Both PMDA and FDA allow the creation/submission of custom domains if the study data does not fit into a standard SDTM domain; however, a custom domain may only be created if the data are different in nature and do not fit into an existing published domain (e.g. standard SDTM, Therapeutic Area Standards)*.

NOTE: When assessing the need for a custom domain, also consider storing data in supplemental qualifiers (SUPP--) or findings about (FA--) domains. Helpful references on when to use findings about or supplemental qualifiers are present in the CDSIC SDTM IG ("When to Use Findings About", "How to Determine where data belong in SDTM Compliant Data Tabulations" and the Supplemental Qualifiers section). Another reference is the PHUSE Paper "Findings About".

The overall process for creating a custom domain is clearly explained in the SDTM IG. It must always be based on one of the three SDTM general observation classes (interventions, events or findings).

Custom domains must be clearly described in the cSDRG/SDRG, and specifically, PMDA prefers to be consulted beforehand when considering storing data in a custom domain.

Source for FDA:

[Study Data Technical Conformance Guide](#)

Source for PMDA:

SDTM/ADaM IG Nuances

Revision of Technical Conformance Guide on Electronic Study Data Submissions

Source for CDISC:

CDISC SDTM IG

Source for PHUSE:

Findings about "Findings About"