ADaM Structure for Occurrence Data (OCCDS) Implementation Guide

Version 1.1 (Final)

Developed by the
CDISC Analysis Data Model Team

Notes to Readers

- This is the final Version 1.1 of the ADaM Structure for Occurrence Data (OCCDS) Implementation Guide.
- This structure is based on the Analysis Data Model (ADaM) v2.1 and the ADaM Implementation Guide (ADaMIG) v1.2.

Revision History

Date	Version
2021-11-29	1.1 Final
2016-02-16	1.0 Final

See Appendix D for representations and warranties, limitations of liability, and disclaimers.

CONTENTS

1 Introduction

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.

1.1.1 Background

The ADaM Data Structure for Adverse Event Analysis, prepared by the CDISC ADaM Team and released in May 2012, described analysis related to occurrence data—but only adverse event analysis data. In the first version of the ADaM Structure for Occurrence Data (OCCDS), released in 2014, the scope was expanded to include more than just adverse event data.

1.1.2 Principles

This document describes analysis methods using the OCCDS. Specifically, the data and analysis described in this document must meet these criteria:

1. There is no need for AVAL or AVALC. Occurrences are counted in analysis, and there are typically 1 or more records for each occurrence assessment.
2. A dictionary is often used for coding the occurrence and typically includes a well-structured hierarchy of categories and terminology. Remapping this hierarchy to BDS variables PARAM and generic *CAT variables would lose the structure and meaning of the dictionary. Per the SDTM Implementation Guide (SDTMIG) v3.3 (available at https://www.cdisc.org/standards/foundational/sdtmig), a dictionary is expected for adverse events and concomitant medications and recommended for medical history. Although not as common, clinical events, procedures, and substance use may also be coded. (Data for a particular study that could have been coded but was not should use this structure because analysis results are similar, and this will allow analysis programming to work the same way—for example, medical history data might be coded in one study and not coded in another, and yet the analysis tables look very similar.)
3. The data content is typically not modified for analysis purposes. In other words, there is no need for analysis versions of the variables that hold the dictionary hierarchy or category terms.

This does not mean that all categorical data are appropriate for OCCDS. More standard categorical data that would never be mapped to a hierarchical dictionary, such as questionnaire responses, fit nicely in BDS and should not use OCCDS.

Typically, findings data fit nicely into BDS, and events and interventions fit nicely into OCCDS. However, this is not always the case: Exposure data, from an interventions SDTM structure, is quite often analyzed in BDS because that analysis does not simply count records, although there could be an OCCDS intermediate dataset used to help derive those BDS summary parameters. In all cases, it is the combination of input data and analysis needs that determines the dataset structure required.

The structure presented in this document is built on the nomenclature of the SDTMIG v3.3 standard for collected data, and adds attributes, variables, and data structures required for statistical analyses. The primary source domain for the structure is the SDTM domain plus the corresponding Supplemental Qualifier dataset. Many additional variables are added from the ADaM Subject-Level Analysis Dataset (ADSL).

In this document, the analysis datasets described are required when SDTM data are not sufficient to support all analyses. Whether an analysis dataset is needed is left up to the producer (see ADaM Section 4.1.1, The Number and Content of Analysis Datasets; https://www.cdisc.org/standards/foundational/adam). If an analysis dataset is needed, and it meets the criteria listed above, it should use OCCDS.

The dataset and variable naming conventions and the dataset structure described in this document should be followed.

The structure for the occurrence analysis dataset is usually 1 record per each record in the corresponding SDTM domain. Examples of when the number of records in the analysis dataset would not match the number in SDTM include:

- SDTM data contain screen failures but screen failures are not analyzed. In this case, the screen failure records are not needed in the analysis dataset.
- SDTM data contain --OCCUR = "N" records but these are not needed for analysis or denominators. In this case, --OCCUR = N records may be excluded from the OCCDS analysis dataset (**Note:** This example does not apply to the ADVERSE EVENT subclass defined in Section 3.1.2, SubClass ADVERSE EVENT).
- The topic (e.g., an adverse event, concomitant medication) spans several treatment periods and needs to be counted in each. Based on the analysis need, a separate row might be required for each treatment period spanned and analyzed.
- An adverse event needs to be analyzed along multiple coding paths. In this case, a row would be needed for each coding path analyzed. An alternate solution, if multiple coding paths are not needed together, would be to put records for each coding path into a separate analysis dataset.

This does not preclude producers from creating additional datasets for other analyses, or from using a different structure if needed for analysis (e.g., time-to-event of adverse events of special interest).

Datasets using the OCCDS structure are assigned a dataset metadata class value of "OCCURRENCE DATA STRUCTURE". See the Define-XML documentation (available at https://www.cdisc.org/standards/data-exchange/define-xml) for additional information regarding CLASS/SUBCLASS.

1.1.3 What's New

The following table highlights some of the updates between this version and OCCDS Version 1.0.

Table 1.1.3.1 Highlights of OCCDS v1.1 Updates

Section/Example	What's New
Principles	The metadata class value of OCCURRENCE DATA STRUCTURE has added a SubClass of ADVERSE EVENT, starting with the release of Define-XML v2.1.
Section 3.2.2 Section 3.2.11 Section 10: Example 7	• Addition of SRCDOM and SRCSEQ as row identifier variables (Table 3.2.2.2) when data are stacked from multiple SDTM datasets • Introduction of a U prefix, as a solution for OCCDS dataset created from stacking rows from multiple SDTM datasets. This prefix stands for "unmodified" and indicates no modifications made to the content of the SDTM variables.
Section 3.2.5 Table 3.2.5.3 Table 3.2.5.4 Section 7: Example 4	• Addition of TREMxxFL, TRTEMwFL, ONTRxxFL, and ONTRTwFL when there are multiple periods or other analysis needs for multiple treatment emergent or on treatment flags • Addition of TREM01FL, TREM02FL, and TREM03FL to Example 4
Section 3.2.9 Table 3.2.9.1	Addition of ADECODy variable (Analysis Dictionary-Derived Term y) for sponsors/clinicians to collapse and/or group preferred terms under a common term to investigate potential safety concerns
Section 9: Example 6	Solution for OCCDS created from multiple SDTM datasets where 1 qualifier changes over time and affects the analysis
Section 14: Example 11	An example of an OCCDS dataset with derived records where data-point traceability is not feasible and value-level metadata is needed

For more details, see Appendix C, Revision History.

1.2 Points to Consider When Interpreting this Document

In reviewing the metadata and examples in this document, some points to consider include:

- **Ordering of variables:** Within this document, no specific ordering of variables within the illustrated datasets is applied. ADaM v2.1 states that ideally the ordering of the variables in the analysis dataset follows a logical ordering (not simply alphabetic), but does not provide a specific recommendation for the ordering of the variables. Within this document, the author of each example applied their own logical ordering. Although there is not an across-example consistency of ordering of variables, within an example the ordering of the variables within the illustrated analysis dataset matches the order of the variables as presented in the associated metadata.
- **Identification of source dataset:** When identifying the source dataset for a variable, the immediate predecessor is used, as described in ADaM. For example, in ADSL the source is identified as DM.SUBJID in the analysis variable metadata. When SUBJID is used in the occurrence analysis dataset, the source is identified as ADSL.SUBJID.
- **Analysis-ready:** The occurrence analysis dataset should be "analysis-ready," meaning it should contain all of the variables needed for the specific analysis, so that the analysis can be replicated by performing the actual statistical test without first having to manipulate data. "Analysis-ready" does not mean that a formatted display can be generated in a single statistical procedure. For typical occurrence analyses, unique subject counts are derived by running a standard statistical procedure (e.g., SAS PROC, S-PLUS function) on the occurrence analysis dataset, whereas denominator counts can be derived from ADSL.
- **Examples for illustration only:** The examples in this document are only intended as illustrations and should not be viewed as a statement of the standards themselves. In addition, the examples are intended to illustrate content and not appearance; it is understood that there are many different ways that data can be displayed. This document does not cover display formats.
- **Display of metadata for illustration only:** Although the metadata elements have been defined in ADaM, how the metadata are displayed is a function of the mechanism used to display the content. The presentation formats used in this document are for the purposes of illustration of content only, and are not intended to imply any type of display standard or requirement. Additionally, the metadata examples just include the metadata necessary to understand the respective example datasets. Refer to Define-XML v2.1 (available at https://www.cdisc.org/standards/data-exchange/define-xml) for additional information (e.g., variable length and origin) required when building a valid define.xml file according to the Define-XML standard.
- **Analysis results metadata:** Analysis results metadata have not been included for any examples in this document. As stated in ADaM, analysis results metadata are not required. However, best practice is that they be provided to assist the consumer by identifying the critical analyses; providing links between results, documentation, and datasets; and documenting the analyses performed.
- **Examples not all-inclusive regarding variables:** The examples describe some of the key variables and records that would be included in the dataset. They are not intended to illustrate every possible variable that might be included in the analysis dataset (e.g., core variables required for subgroup analyses are not included in all illustrations).
- **Source/Derivation column:** The algorithms provided in the Source/Derivation column are for illustration purposes only and are not intended to imply universally accepted definitions or derivations of variables. Algorithms are producer-defined and dependent on trial and analysis design.
- **Endorsement of vendors or products:** As with other ADaM documents, references to specific vendor products are examples only and therefore should not be interpreted as an endorsement of these vendors or products.

1.3 Conventions Used in this Document

Throughout this document the terms “producer” and “consumer” are used to refer to the originator/sender/owner/sponsor of the data and the user/reviewer/recipient of the data, respectively. These terms are used to simplify the document, and are not intended to imply that these examples only apply to analysis datasets in the context of electronic submissions to regulatory agencies.

2 Data Analysis and Coding

2.1 Statistical Analysis

The most frequently used method for the comparison between treatment groups of data in OCCDS is the summarization of the number of subjects with at least 1 occurrence of a term. These counts and related percentages are presented at different levels of the dictionary hierarchy, when the hierarchy exists. The denominator used for calculating percentages is often determined by a population flag, such as the total number of subjects at risk or the number of subjects exposed to treatment. Note that some subjects in the population may not have any records, and therefore these subjects would not be represented in the SDTM domain nor the corresponding OCCDS analysis dataset. Thus, the denominators usually need to be obtained from the ADaM subject-level analysis dataset (ADSL) rather than directly from the occurrence analysis dataset.

ADaM primarily covers the creation of an analysis dataset needed for the presentation of frequencies and percentages. However, the example analysis datasets presented in this document could be used to construct more in-depth analysis datasets, even in a different structure such as a time-to-event analysis dataset of adverse events of interest in the ADaM Basic Data Structure. For ADaM data structures other than OCCDS, see ADaMIG v1.2 (available at https://www.cdisc.org/standards/foundational/adam).

2.2 Dictionary Coding

Data are often collected in textual or "verbatim" content, a short description of an event or intervention generally written in free text on the case report form (CRF). Verbatim content is then processed through a coding dictionary so that similar verbatim content is grouped together by classifying them into a hierarchy of medical granularity.

The Medical Dictionary for Regulatory Activities (MedDRA; https://www.meddra.org) and WHO Drug Dictionary (https://www.who-umc.org/) have become widely recognized as a global standard for the coding of adverse events and medications. Examples of other coding dictionaries include WHO Adverse Reaction Terminology (WHO-ART; see https://www.who-umc.org/), no longer actively maintained; the International Classification of Disease (ICD; http://www.who.int/classifications/icd/en/); and Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART; see http://www.nlm.nih.gov/research/umls/), which was replaced by MedDRA but can still be found in older studies.

The coding dictionary is characterized by classifying each verbatim into a hierarchy of medical granularity. For example, if the verbatim content recorded was "stomach virus," the COSTART coding hierarchy would place this event in the Body as a Whole system, in the General subcategory for this body system, and with the preferred term of "Flu Syndrome." If using MedDRA v23.1, for example, this verbatim content would result in a system organ class (SOC) of "Infections and infestations" and a preferred term (PT) of "Gastroenteritis viral."

When using coding dictionaries, it is recommended that coding rules and guidelines be developed by the producer prior to the classification of terminology. The process of coding verbatim terms with a dictionary is outside the scope of this document. The objective of coding guidelines is to promote medical accuracy and consistency when using the controlled vocabulary of the dictionary. This consistency will support a variety of downstream analysis needs, such as when events need to be recoded to integrate data from 2 or more clinical studies.

2.2.1 Recoding of Occurrence Data

In some situations, multiple study reports are created for a single study. For example, an initial study report may be created at the time of the primary analysis for the primary efficacy endpoint. If subjects are followed for safety, a second report may be created years later so that long-term safety data can be incorporated. At that time, there may be a desire to update the coding dictionary so that all content is coded using the most recent version of a dictionary.

It should be noted that a more common scenario involving recoding occurrence data is when data are recoded for an integrated analysis and submitted to a regulatory agency for marketing approval. However, neither the current version of the ADAMIG nor this document fully covers integration of multiple studies. The ADaM Team is developing a document to address integration of multiple studies. Some of the suggestions included here for handling multiple dictionaries may be revised after this integration document is released.

2.3 Adverse Events

The safety evaluation of a clinical trial includes the analysis of adverse events, and that analysis is typically done using OCCDS. The definition of an *adverse event*, per the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) E2A guideline,[1] is:

> Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment.

Important attributes include the level of severity of the adverse event (mild, moderate, or severe), whether it is considered to be related to the study product (yes or no), and whether it is considered serious (yes or no). Of particular importance in the analysis of adverse events is the definition of *treatment-emergent*. The ICH E9 guidance document[2] defines *treatment-emergent* as an event that emerges during treatment having been absent pretreatment, or worsens relative to the pretreatment state. Operationally, classifying adverse events as treatment-emergent will utilize, in part, the start or worsening date of the event relating to the trial or treatment start. Other important attributes of adverse events include the action taken in response to the event and whether the event led to permanent discontinuation of the investigational product.

2.4 Concomitant Medications Data

Concomitant medications data can be coded to a hierarchy such as WHO Drug and summarized by medication and/or ingredient within class.

2.5 Pre-specified Data

In some cases, data can be gathered on a CRF that contains a pre-specified category and a checkbox to indicate whether the subject had the event, condition, or treatment. This information is stored as described in the SDTMIG (available at https://www.cdisc.org/standards/foundational/sdtmig). --PRESP is used to indicate that the term is pre-specified; --OCCUR is either "Y" or "N" (or null if --STAT is populated) to indicate whether the subject did or did not have the event, condition, or treatment. --TERM or --TRT will have a known and finite set of values, so these values may be adequate to use as a summarization category. Often these pre-specified terms are grouped into categories at collection using the --CAT and --SCAT variables, creating additional levels of summarization categories. In this situation, data are analyzed by variables such as --TERM, --TRT, --CAT and --SCAT, and dictionary coding might not be necessary. For example, medical history and clinical events data are often captured in this way.

Note that this pre-specified data option works in a different way for adverse events because the SDTMIG does not permit the use of variable AEOCCUR. In other words, all records in an SDTM Adverse Event dataset must correspond to an actual occurrence of the event; pre-specified events that do not occur are captured in another dataset.

2.6 Combining Spontaneous and Pre-specified Occurrences

It is technically feasible to apply the same coding dictionary to both collected and pre-specified data and combine these data for analysis. Whether to do so is a statistical judgment that should be carefully considered and described in programming specifications. A pre-specified question on a CRF makes it more likely to receive data, and would therefore increase the frequency in a summary. In deciding whether to pool pre-specified and spontaneous data, the statistician should consider the way data were gathered and weigh the possibility of over-reporting pre-specified data. The statistician should also carefully consider and describe the correct denominator for percentages.

If data are pooled in this way, take care that non-occurring data (--OCCUR = N) are properly excluded from the analysis.

2.7 Other Data

Other similar data, with or without coding hierarchy, can be summarized similarly. For an example of summarizing without a coding hierarchy, see Section 13, Example 10: Analysis of Medical History Pre-specified Events.

Some examples of other data that can also be summarized using OCCDS include:

- Clinical events, when collected by category and not mapped to a dictionary, but summarized in a similar way as adverse events
- Protocol violations, when summarized by counting subjects with violations within each category
- Laboratory data containing National Cancer Institute Common Toxicity Criteria (NCI-CTC)[3] information that has been coded with MedDRA and summarized as laboratory events. When these events are summarized like adverse events, an extension of the adverse event examples that are shown in this document can be used.

In all cases, OCCDS should be used when a summary of the hierarchy is done, counting the number of subjects at each level of the hierarchy. Alternatively, the ADaM Basic Data Structure could be used for counting when there is not a hierarchy, when the terms are counted rather than the subjects, and when variables such as AVAL and PARAM are appropriate to include.

3 ADaM Metadata

As described in ADaM Section 4.1.2, Analysis Dataset and Variable Naming Conventions (available at https://www.cdisc.org/standards/foundational/adam), variables that are copied from SDTM must have the same variable name, label, values, and meaning as in SDTM. Because the OCCDS can be used for adverse events, concomitant medications, and other occurrence data, metadata shown in this section reference different SDTM domains. For clarity, the following conventions are used:

- When referring to the 2-letter prefix in variable names, the standard convention is to use "--", as described in SDTM v1.7 (available at https://www.cdisc.org/standards/foundational/sdtm) Section 2.2, The General Observation Classes.
- The "--" convention was intended for variable names, not domain names, and "--" is difficult to read in the documentation for SDTM domain names. This document uses the convention of "XX" to represent a domain name, as was done in the ADaM Examples in Commonly Used Statistical Analysis Methods document (https://www.cdisc.org/standards/foundational/adam/).
- Variable labels that differ depending on SDTM domain are referenced by a note at the end of the relevant table.

Take care when creating actual metadata to replace "--", "XX", and generic variable labels with the actual 2-letter domain code and label from the SDTM.

3.1 Dataset Metadata

Typically, the analysis data structure for an OCCDS dataset is specified as follows:

Table 3.1.1 Data Structure

Data Structure Name	Data Structure Description	Class of Dataset	SubClass of Dataset	CDISC Notes
OCCDS	Occurrence Data Structure	OCCURRENCE DATA STRUCTURE		Generally these are 1 record per record in SDTM domain (optional: per coding path, per Analysis Period and/or Phase. See Section 1.1, Purpose, for examples of when the analysis data structure might not be one record per record in SDTM domain.)
AE	Occurrence Data Structure Adverse Event	OCCURRENCE DATA STRUCTURE	ADVERSE EVENT	Datasets in the SubClass ADVERSE EVENT must have a Class of OCCURRENCE DATA STRUCTURE and meet all the principles of that class. The SDTM input dataset for the ADVERSE EVENT SubClass is always AE, with some additional information from SUPPAE, FA, and ADSL. See Section 3.1.2, SubClass ADVERSE EVENT, for more details.

The Data Structure Name, Data Structure Description, and CDISC Notes are intended to provide information to assist producers in preparing their datasets and are not intended to be metadata submitted in define.xml.

3.1.1 Example Define-XML v2.1 ADaM Datasets

Beginning with Define-XML v2.1 (available at https://www.cdisc.org/standards/data-exchange/define-xml), dataset metadata can include SubClass in addition to Class. SubClass has controlled terminology, and one of the allowed terms is "ADVERSE EVENT". Table 3.1.1.1 shows how the SubClass of "ADVERSE EVENT" is used within the ADaM dataset metadata.

Table 3.1.1.1 Example Define-XML v2.1 ADaM Datasets

Dataset	Description	Class - SubClass	Structure	Purpose	Keys	Documentation	Location
ADSL	Subject-Level Analysis	SUBJECT LEVEL ANALYSIS DATASET	One record per subject	Analysis	STUDYID, USUBJID	Screen Failures are excluded since they are not needed for this study analysis. See referenced dataset creation program and ADRG adsl.sas Analysis Data Reviewer's Guide [6]	adsl.xpt
ADQSADAS	ADAS-Cog Analysis	BASIC DATA STRUCTURE	One record per subject per parameter per analysis visit per analysis date	Analysis	STUDYID, USUBJID, PARAMCD, AVISIT, ADT	See referenced dataset creation program and ADRG adqsadas.sas Analysis Data Reviewer's Guide [Section2.1]	adqsadas.xpt
ADAE	Adverse Events Analysis Dataset	OCCURRENCE DATA STRUCTURE • ADVERSE EVENT	One record per subject per adverse event	Analysis	STUDYID, USUBJID, AETERM, ASTDT, AESEQ	See SAS program adae.sas	adae.xpt

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

3.2 Variable Metadata

As stated previously, OCCDS is different from the ADaM Basic Data Structure (BDS). There is no PARAM nor AVAL, for example. However, some of the variables described for BDS in ADaMIG v1.2 (available at https://www.cdisc.org/standards/foundational/adam) can be used in OCCDS, as shown below. Likewise, some variables defined below may be used in non-OCCDS ADaM dataset structures, such as TRTEMFL if needed.

The more standardized variables commonly occurring in the ADaM OCCDS are described here in tabular format. In general, include all variables from the SDTM dataset and corresponding supplemental qualifiers that are needed for analysis or traceability. For traceability when copying variables from SUPPQUAL, it is recommended to use variable names that exactly match the corresponding SUPPQUAL.QNAM values. Additional study- or therapeutic-specific variables may be added as needed but should follow the standard variable-naming conventions described in ADaMIG v1.2 Section 3.3, ADaM Basic Data Structure (BDS) Variables. For example, variables with the 2-letter SDTM prefix are most commonly those that are copied from the SDTM or transposed SUPPQUAL dataset, or the numeric version of the SDTM variable, but not analysis versions of SDTM variables. Choose variable names with care to prevent unintended conflicts with standard names.

As described in ADaM, the 3 rightmost columns of the table (i.e., Core, SubClass ADVERSE EVENT Core, CDISC Notes) provide information about the variables to assist users in preparing their datasets. These columns are not meant to be metadata submitted in define.xml. The Core column, as defined in ADaMIG v1.2, describes whether a variable is required (Req), conditionally required (Cond), or permissible (Perm) in the OCCURRENCE DATA STRUCTURE Class. The SubClass ADVERSE EVENT Core column describes whether a variable is required (Req), conditionally required (Cond), permissible (Perm), or not used (Not used) in the ADVERSE EVENT SubClass. The CDISC Notes column provides more information about the variable. In addition, the Type column is being used to define whether the variable is character (Char) or numeric value (Num). More specific information will be provided in metadata.

SubClass ADVERSE EVENT Variables

OCCDS variables are described later in this section. Variables in an OCCDS dataset that is of SubClass ADVERSE EVENT can include:

- Any variable copied unchanged from ADSL
- Any variable copied unchanged from SDTM AE
  - SubClass ADVERSE EVENT should include all SDTM AE variables with the SDTM core value of Required or Expected.
- Any variable copied or derived from SDTM FA (see Section 9, Example 6: Analysis of Adverse Events that Change over Time)
  - When FA is used as input to the ADVERSE EVENT SubClass dataset, a unique identifier variable, such as FASEQ or FASPID, is required for traceability.
- Any variable copied from content in SDTM SUPPAE
  - It is recommended to use SUPPAE QNAM as the ADVERSE EVENT SubClass dataset variable name, SUPPAE QLABEL as the ADVERSE EVENT SubClass dataset variable label, and SUPPAE QVAL as the ADVERSE EVENT SubClass variable content.
- Derived variables, such as numeric timing variables, TRTEMFL and other indicators, analysis versions of descriptive variables, and other variables derived using the naming conventions described in Section 3.2.11, User-specified Variable Naming Conventions

3.2.1 ADSL Variables

Merge any subject-level variables needed for analysis or reference.

Be aware that only subjects with a record in an input SDTM domain would have a record in the analysis dataset. For this reason, it is recommended that population indicators and denominator counts for percentages be derived from ADSL and not from the occurrence analysis dataset.

3.2.2 Identifier Variables

Table 3.2.2.1 includes the identifier variables from SDTM.

Table 3.2.2.1 OCCDS Identifier Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
STUDYID	Study Identifier	Char			Req	Req	XX.STUDYID
USUBJID	Unique Subject Identifier	Char			Req	Req	XX.USUBJID
SUBJID	Subject Identifier for the Study	Char			Perm	Perm	ADSL.SUBJID
SITEID	Study Site Identifier	Char			Perm	Perm	ADSL.SITEID

In addition to the above identifiers, row identifiers are included to support datapoint traceability. The combination of dataset or domain name and the relevant sequence number, when used in conjunction with USUBJID, provide datapoint traceability. There are 2 row identifier options: --SEQ or the combination of SRCDOM + SRCSEQ.

Commonly, rows in an OCCDS dataset are created from a single SDTM domain. In this case, --SEQ provides all of the row identifier (and thus datapoint traceability) information: The 2-letter prefix of the variable name is the SDTM domain name, and the content is the relevant sequence number.

Less commonly, rows in an OCCDS dataset are from multiple SDTM domains or from 1 or more ADaM datasets. Either the --SEQ variable from each domain can be added, or the variables SRCDOM and SRCSEQ can be used to identify the input rows.

Note: SRCDOM and SRCSEQ are described in ADaMIG Section 3.3.9, Datapoint Traceability Variables, as part of BDS datapoint traceability. The BDS data-point traceability variable SRCVAR is not included in OCCDS; SRCVAR is used to hold the name of the primary variable used to derive AVAL or AVALC in BDS, but AVAL and AVALC are not applicable to OCCDS. (For additional guidance, see the ADaMIG; https://www.cdisc.org/standards/foundational/adam).

Table 3.2.2.2 Row Identifier Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
--SEQ	Sequence Number	Num			Cond	Req	XX.--SEQ Identifies the sequence number in SDTM domain XX that is the source for this row. This variable would be copied unchanged from the SDTM domain XX. Required for traceability back to SDTM when rows in the dataset are created from a single SDTM domain. This may be missing for derived rows. For SubClass ADVERSE EVENT, use AESEQ, copied from SDTM AE.AESEQ.
SRCDOM	Source Data	Char			Perm	Not used	Identifies the name of the domain or dataset that is the source for this row. If the source data is a supplemental qualifier in SDTM, this variable will contain the value of RDOMAIN in SUPP-- or SUPPQUAL. Used when rows in the OCCDS dataset are from multiple SDTM domains or from one or more ADaM datasets. Not to be used in conjunction with --SEQ. This may be missing for derived rows. For SubClass ADVERSE EVENT, AESEQ is used rather than SRCDOM and SRCSEQ.
SRCSEQ	Source Sequence Number	Num			Perm	Not used	Identifies the sequence number that is the source for this row. If SRCDOM is a SUPPQUAL, then this variable will contain the sequence number of the relevant related domain record. Used when rows in the OCCDS dataset are from multiple SDTM domains or from one or more ADaM datasets. Not to be used in conjunction with --SEQ. This may be missing for derived rows. For SubClass ADVERSE EVENT, AESEQ is used rather than SRCDOM and SRCSEQ.
ASEQ	Analysis Sequence Number	Num			Perm	Perm	Sequence number given to ensure uniqueness of subject records within an ADaM dataset. ASEQ is useful for traceability when the OCCDS dataset is used as input to another ADaM dataset and the --SEQ variable is not included or unique. ASEQ is described in more detail in ADaMIG v1.2, Section 3.3.1, Identifier Variables for BDS Datasets.

For the ADVERSE EVENT SubClass, when data from FA is used in the ADVERSE EVENT SubClass, a unique identifier variable, such as FASEQ or FASPID, is also required for traceability.

3.2.3 Dictionary Coding and Categorization Variables

Dictionary coding and categorization variables provided in the SDTM should be included as needed for analysis, review, or traceability. Common coding variables are provided in this section. If other coding variables are included in the SDTM and pertinent for analysis, these should be included in ADaM using a similar naming convention as shown here. For any public versioned dictionary, the metadata for each coding variable should include both the name and version of the dictionary.

Common Dictionary Coding Variables for MedDRA

MedDRA coding is typically used for adverse events and medical history. Copy the needed MedDRA terms and codes from SDTM to the analysis dataset.

For the SubClass ADVERSE EVENT, it is required that all levels of terms for the primary path in the MedDRA hierarchy—System Organ Class (SOC), High Level Group Term (HLGT), High Level Term (HLT), Lowest Level Term (LLT), and Preferred Term (PT)—be included. For other OCCDS datasets, these variables are recommended (they are frequently useful in further analyses of events), but not required.

Table 3.2.3.1 MedDRA Dictionary Coding Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
--TERM	Reported Term	Char			Req	Req	Copied from XX.--TERM This variable label differs depending on the SDTM domain. See SDTM v1.7 Section 2.2.2, The Events Observation Class, and SDTMIG v3.3 Section 6.2, Models for Events Domains, for details.
--DECOD	Dictionary-Derived Term	Char		MedDRA	Cond	Req	Copied from XX.--DECOD This variable is typically one of the primary variables used in an analysis and would be brought in from the SDTM domain. Equivalent to the MedDRA PT. All other SDTM domain variables and supplemental qualifiers needed for analysis or traceability should also be included. Include the dictionary version in the metadata. Conditional on whether coded and used for analysis. Required for adverse event data.
--BODSYS	Body System or Organ Class	Char		MedDRA	Cond	Req	Copied from XX.--BODSYS Include the dictionary version in the metadata. Conditional on whether coded and used for analysis. Required for adverse event data.
--BDSYCD	Body System or Organ Class Code	Num		MedDRA	Perm	Req	Copied from XX.--BDSYCD or the supplemental qualifier This would be copied from the SDTM domain XX or supplemental qualifier dataset. Include the dictionary version in the metadata.
--LLT	Lowest Level Term	Char		MedDRA	Cond	Req	Copied from XX.-- LLT or the supplemental qualifier Include the dictionary version in the metadata. Conditional on whether coded and used for analysis.
--LLTCD	Lowest Level Term Code	Num		MedDRA	Perm	Req	Copied from XX.--LLTCD or the supplemental qualifier Include the dictionary version in the metadata.
--PTCD	Preferred Term Code	Num		MedDRA	Perm	Req	Copied from XX.--PTCD or the supplemental qualifier Include the dictionary version in the metadata.
--HLT	High Level Term	Char		MedDRA	Cond	Req	Copied from XX.--HLT or the supplemental qualifier Include the dictionary version in the metadata. Conditional on whether used for analysis.
--HLTCD	High Level Term Code	Num		MedDRA	Perm	Req	Copied from XX.--HLTCD or the supplemental qualifier Include the dictionary version in the metadata.
--HLGT	High Level Group Term	Char		MedDRA	Cond	Req	Copied from XX.--HLGT or the supplemental qualifier Include the dictionary version in the metadata. Conditional on whether used for analysis.
--HLGTCD	High Level Group Term Code	Num		MedDRA	Perm	Req	Copied from XX.--HLGTCD or the supplemental qualifier Include the dictionary version in the metadata.
--SOC	Primary System Organ Class	Char		MedDRA	Cond	Req	Copied from XX.--SOC or the supplemental qualifier Include the dictionary version in the metadata. Conditional on whether a secondary SOC was used for analysis.
--SOCCD	Primary System Organ Class Code	Num		MedDRA	Perm	Req	Copied from XX.--SOCCD or the supplemental qualifier Include the dictionary version in the metadata.

Note: MedDRA allows a term to be represented in more than 1 SOC (e.g., primary and, in some cases, secondary coding paths). See Section 8, Example 5: MedDRA Secondary Path, for an example layout for a possible way to handle the need to report on secondary paths along with primary paths.

Common Dictionary Coding Variables for WHO Drug

WHO Drug coding is typically used for concomitant medications. Copy to the analysis dataset the needed WHO Drug terms and codes from SDTM CM and SUPPCM. The variables shown in Table 3.2.3.2 are intended for a single WHO Drug coding path.

Table 3.2.3.2 WHO Drug Dictionary Coding Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
CMTRT	Reported Name of Drug, Med, or Therapy	Char			Req	Not Used	CM.CMTRT
CMDECOD	Standardized Medication Name	Char		WHO Drug	Cond	Not Used	CM.CMDECOD This is typically one of the primary variables used in CM analysis and would be copied from the SDTM CM domain. Include the dictionary version in the variable metadata. Conditional on whether coded and used for analysis.
CMCLAS	Medication Class	Char			Perm	Not Used	CM.CMCLAS Include the dictionary version in the metadata.
CMCLASCD	Medication Class Code	Char			Perm	Not Used	CM.CMCLASCD Include the dictionary version in the metadata.
ATCy	ATC Level y Text	Char		WHO Drug	Cond	Not Used	Corresponds to the ATC Level Text for WHO Drug. Include the dictionary version in the variable metadata. Conditional, based on analysis at multiple levels (y)
ATCyCD	ATC Level y Code	Char		WHO Drug	Cond	Not Used	Corresponds to the ATC Level Code for WHO Drug. Include the dictionary version in the variable metadata. Conditional, based on analysis at multiple levels (y)

Other Categorization Variables

When categories are used for the intended analysis, instead of or in addition to MedDRA or WHO Drug, these generic categorization variables are commonly used:

Table 3.2.3.3 Other Categorization Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
--CAT	Category	Char			Perm	Perm	Copied from XX.--CAT This variable label differs depending on the SDTM domain. See SDTM v1.7 Section 2.2, The General Observation Classes, and SDTMIG v3.3 Section 6, Domain Models Based on the General Observation Classes, for details.
--SCAT	Subcategory	Char			Perm	Perm	Copied from XX.--SCAT This variable label differs depending on the SDTM domain. See SDTM v1.7 Section 2.2, The General Observation Classes, and SDTMIG v3.3 Section 6, Domain Models Based on the General Observation Classes, for details.
ACATy	Analysis Category y	Char			Perm	Perm	Category used in analysis. May be derived from --CAT and/or --SCAT. Examples include records of special interest like prohibited medications, concomitant medications taken during an infusion reaction, growth factors, antimicrobial medications, and other such categories not defined elsewhere or present in SDTM domains.

3.2.4 Timing Variables

Timing variables are copied from the SDTM and derived within ADaM. Common timing variables are provided in this section. If other timing variables are collected in SDTM and pertinent for analysis, these should be included in ADaM. Additional timing variables, such as those for period start and end dates, can be included. For more details on timing variables, see ADaMIG Section 3.3.3, Timing Variables for BDS Datasets.

Table 3.2.4.1 Timing Variables

Variable Name	Variable Label	Type	Codelist[a]	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
--STDTC	Start Date/Time of Observation	Char		ISO 8601	Cond	Req	Copied from XX.–STDTC Conditional on whether start date is pertinent for study and is populated in SDTM This variable label differs depending on the SDTM domain. See SDTM v1.7 Section 2.2.5, Timing Variables for All Classes, and SDTMIG v3.3 Section 6, Domain Models Based on the General Observation Classes for details.
ASTDT	Analysis Start Date	Num			Cond	Req	Created from converting XX.--STDTC from character ISO 8601 format to numeric date format, applying imputation rules if specified in the statistical analysis plan (SAP) or metadata Conditional on whether start date is pertinent for study and is populated in SDTM
ASTTM	Analysis Start Time	Num			Cond	Cond	Created from converting XX.--STDTC from character ISO 8601 format to numeric time format, applying imputation rules if specified in the SAP or metadata Conditional on whether start time is pertinent for study and is populated in SDTM
ASTDTM	Analysis Start Datetime	Num			Cond	Cond	Created from converting XX.--STDTC from character ISO 8601 format to numeric datetime format, applying imputation rules if specified in the SAP or metadata Conditional on whether start datetime is pertinent for study and is populated in SDTM
ASTDTF	Analysis Start Date Imputation Flag	Char	(DATEFL)		Cond	Cond	The level of imputation of analysis start date. Imputation flags are described in ADaMIG v1.2 Section 3.1.3, Date and Time Imputation Flag Variables. Conditional on whether ASTDT (or the date part of ASTDTM) was imputed
ASTTMF	Analysis Start Time Imputation Flag	Char	(TIMEFL)		Cond	Cond	The level of imputation of analysis start time. Imputation flags are described in ADaMIG v1.2 Section 3.1.3, Date and Time Imputation Flag Variables. Conditional on whether ASTTM (or the time part of ASTDTM) was imputed
--ENDTC	End Date/Time of Observation	Char		ISO 8601	Cond	Req	Copied from XX.--ENDTC Conditional on whether end date is pertinent for study and is populated in SDTM This variable label differs depending on the SDTM domain. See SDTM v1.7 Section 2.2.5, Timing Variables for All Classes, and SDTMIG v3.3 Section 6, Domain Models Based on the General Observation Classes, for details.
AENDT	Analysis End Date	Num			Cond	Req	Created from converting XX.--ENDTC from character ISO 8601 format to numeric date format, applying imputation rules if specified in the SAP or metadata Conditional on whether end date is pertinent for study and is populated in SDTM
AENTM	Analysis End Time	Num			Cond	Cond	Created from converting XX.--ENDTC from character ISO 8601 format to numeric time format, applying imputation rules if specified in the SAP or metadata Conditional on whether end time is pertinent for study and is populated in SDTM
AENDTM	Analysis End Datetime	Num			Cond	Cond	Created from converting XX.--ENDTC from character ISO 8601 format to numeric datetime format, applying imputation rules if specified in the SAP or metadata Conditional on whether end datetime is pertinent for study and is populated in SDTM
AENDTF	Analysis End Date Imputation Flag	Char	(DATEFL)		Cond	Cond	The level of imputation of analysis end date. Imputation flags are described in the ADaMIG v1.2 Section 3.1.3, Date and Time Imputation Flag Variables. Conditional on whether AENDT (or the date part of AENDTM) was imputed
AENTMF	Analysis End Time Imputation Flag	Char	(TIMEFL)		Cond	Cond	The level of imputation of analysis end time. Imputation flags are described in the ADaMIG v1.2 Section 3.1.3, Date and Time Imputation Flag Variables. Conditional on whether AENTM (or the date part of AENDTM) was imputed

Variable Name	Variable Label	Type	Codelist[a]	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
ASTDY	Analysis Start Relative Day	Num			Cond	Req	The number of days from an anchor date (not necessarily DM.RFSTDTC) to ASTDT Example derivation: ASTDT - ADSL.TRTSDT + 1 if ASTDT >= ADSL.TRTSDT, else ASTDT - ADSL.TRTSDT if ASTDT< ADSL.TRTSDT This variable may instead be copied from --STDY. Conditional on whether analysis start relative day is pertinent to the study
--STDY	Study Day of Start of Observation*	Num			Perm	Cond	Copied from XX.--STDY ASTDY may differ from --STDY due to date imputation and the option in ADaM to use a reference date other than SDTM's RFSTDTC. Including XX.--STDY in addition to ASTDY adds traceability. For SubClass ADVERSE EVENT, conditional on whether the AESTDY variable is in the SDTM AE dataset
AENDY	Analysis End Relative Day	Num			Perm	Req	The number of days from an anchor date (not necessarily DM.RFSTDTC) to AENDT Example derivation: AENDT - ADSL.TRTSDT + 1 if AENDT >= ADSL.TRTSDT, else AENDT - ADSL.TRTSDT if AENDT < ADSL.TRTSDT This variable may instead be copied from --ENDY.
--ENDY	Study Day of End of Observation*	Num			Perm	Cond	Copied from XX.--ENDY AENDY may differ from --ENDY due to date imputation and the option in ADaM to use a reference date other than SDTM's RFSTDTC. Including XX.--ENDY in addition to AENDY adds traceability. For SubClass ADVERSE EVENT, conditional on whether the AEENDY variable is in the SDTM AE dataset
ADURN	Analysis Duration (N)	Num			Perm	Perm	Derive from ASTDT (or ASTDTM) and AENDT (or AENDTM).
ADURU	Analysis Duration Units	Char	(UNIT)		Cond	Cond	Conditional on whether ADURN is included
--DUR	Duration of XX	Char		ISO 8601	Perm	Cond	Copied from XX.--DUR Because --DUR is a collected field and ADURN is derived, the values will often differ. Including XX.--DUR in addition to ADURN adds traceability. For SubClass ADVERSE EVENT, conditional on whether the AEDUR variable is in the SDTM AE dataset
APERIOD	Period	Num			Perm	Cond	APERIOD is a record-level timing variable that represents the analysis period within the study associated with the record for analysis purposes. The value of APERIOD (if populated) must be one of the xx values found in the ADSL TRTxxP variables. See ADaMIG v1.2 Section 3.3.3, Timing Variables for BDS variables for more information on this variable. For SubClass ADVERSE EVENT, conditional on whether multiple analysis periods are defined for the study
APERIODC	Period (C)	Char			Perm	Perm	Text characterizing to which period the record belongs. One-to-one map to APERIOD.
APHASE	Phase	Char			Perm	Perm	APHASE is a categorization of timing within a study, for example a higher-level categorization of APERIOD or an analysis epoch. For example, APHASE could describe spans of time for SCREENING, ON TREATMENT, and FOLLOW-UP. See the ADaM Implementation Guide v1.2 for more information on this variable.

[a]Codelists in parenthesis are the names of CDISC Controlled Terminology.

3.2.5 Indicator Variables

Although some indicator variables can be copied from the SDTM, others are derived within ADaM. If indicator variables other than those shown in this section are included in the SDTM and pertinent for analysis, these should be copied to ADaM. Other indicator analysis variables that are needed for analysis can also be added.

Table 3.2.5.1 SDTM Indicator Variables

Variable Name	Variable Label	Type	Codelist[a]	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
--OCCUR	XX Occurrence	Char	(NY)		Cond	Not used	Copied from XX.--OCCUR Conditional on whether this content is pertinent for analysis and is populated in SDTM SDTM does not allow variable AEOCCUR, so this variable is not available to include in ADaM.
--PRESP	XX Pre-Specified	Char	(NY)		Cond	Cond	Copied from XX.--PRESP Conditional on whether this content is pertinent for analysis and is populated in SDTM

[a]Codelists in parenthesis are the names of CDISC Controlled Terminology.

Table 3.2.5.2 OCCDS Indicator Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
ANLzzFL	Analysis Flag zz	Char	Y		Cond	Cond	The ANLzzFL flag is useful in many circumstances; an example is when more than 1 coding path is included for analysis, in which case separate analysis flags could be used to denote primary coding path or the records used for analysis from each coding path. A codelist of Y, N, null may be used as described in ADaMIG v1.2 Section 3.3.8, Indicator Variables for BDS Datasets. This variable is conditional on whether analysis records flags are needed for analysis.

With adverse events and concomitant medications, typically indicator flags are also assigned based on the timing of the analysis record in relation to the study. Tables 3.2.5.3, 3.2.5.4, and 3.2.5.5 include some common indicator flags for these types of data.

Table 3.2.5.3 Adverse Events Indicator Variables

Variable Name	Variable Label	Type	Codelist[a]	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
TRTEMFL	Treatment Emergent Analysis Flag	Char	Y		Cond	Req	Treatment-emergent flag as defined for analysis Example derivation: If ADSL.TRTSDT<=ASTDT<=ADSL.TRTEDT + x days then TRTEMFL="Y" The number *x* in this derivation is defined by the producer and often incorporates the known half-life of the drug. For datasets other than SubClass ADVERSE EVENT, this variable is conditional on whether the concept of treatment emergent is a key feature of the analysis.
TREMxxFL	Treatment Emergent Period xx Flag	Char	Y		Cond	Cond	This variable is required if there are multiple periods where treatment emergence is a key feature of the analysis for each period. If TREMxxFL is included, TRTEMFL is defined as the overall treatment-emergent flag.

Variable Name	Variable Label	Type	Codelist[a]	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
TRTEMwFL	Treatment Emergent Analysis w Flag	Char	Y		Perm	Perm	This variable is used if there are other analysis needs (e.g., different cut-offs) where treatment emergence is a key feature of the analysis. If TREMwFL is included, TRTEMFL is defined as the overall treatment-emergent flag.
AETRTEM	Treatment Emergent Flag	Char	(NY)		Perm	Cond	Treatment-emergent flag from SDTM, if available. See SDTMIG v3.3, Appendix C2, Supplemental Qualifiers Name Codes, for more information. Derivation: SUPPAE.QVAL where QNAM="AETRTEM" TRTEMFL may differ from AETRTEM due to different definitions, date imputation, and other analysis rules. Including AETRTEM in addition to TRTEMFL will add traceability. For SubClass ADVERSE EVENT, conditional on whether the AETRTEM variable is in the SDTM AE dataset and populated

[a]Codelists in parenthesis are the names of CDISC Controlled Terminology.

Table 3.2.5.4 Concomitant Medications Indicator Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
ONTRTFL	On Treatment Record Flag	Char	Y		Cond	Perm	Character indicator of whether the observation occurred while the subject was on treatment. A codelist of Y, N, null may be used as described in ADaMIG Section 3.3.8, Indicator Variables for BDS Datasets. Example derivation: If ADSL.TRTSDT <= ASTDT <= ADSL.TRTEDT then ONTRTFL = "Y" This variable is conditional on whether the concept of on treatment is a feature of the study and used in analysis.
ONTRxxFL	On Treatment Period xx Flag	Char	Y		Perm	Perm	This variable is used if there are multiple periods where on treatment is a key feature of the analysis for each period. If ONTRxxFL is included, ONTRTFL is defined as the overall on-treatment flag.
ONTRTwFL	On Treatment Record w Flag	Char	Y		Perm	Perm	This variable is used if there are other analysis needs (e.g., different cut-offs) where on treatment is a key feature of the analysis. If ONTRTwFL is included, ONTRTFL is defined as the overall on-treatment flag.

Table 3.2.5.5 Adverse Event and Concomitant Medications Indicator Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
PREFL	Pre-treatment Flag	Char	Y		Cond	Cond	Character indicator of whether the observation occurred before the subject started treatment Example derivation: If ASTDT < ADSL.TRTSDT then PREFL="Y" This variable is conditional on whether the concept of pre-treatment is a feature of the study and used in analysis.
FUPFL	Follow-up Flag	Char	Y		Cond	Cond	Character indicator of whether the observation occurred while the subject was on follow-up Example derivation: If ASTDT > ADSL.TRTEDT then FUPFL="Y" This variable is conditional on whether the concept of follow-up is a feature of the study and used in analysis.

3.2.6 Occurrence Flag Variables

Occurrence flags can be used to prepare data for analysis. They are typically created by sorting the data in the required order and then flagging the first treatment-emergent record. The use of the word "first" in this section does not necessarily mean chronological, although that is an option. The more common occurrence flags and a structure for additional flags are shown in Tables 3.2.6.1 and 3.2.6.2.

Table 3.2.6.1 OCCDS Occurrence Flag Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
AOCCFL	1st Occurrence within Subject Flag	Char	Y		Perm	Perm	Character indicator for the first occurrence of any event/intervention/finding within the subject Example derivation: Sort the data in the required order and flag the first treatment emergent record for each subject.
AOCCPFL	1st Occurrence of Preferred Term Flag	Char	Y		Perm	Perm	Character indicator for the first occurrence of the preferred term within the subject Example derivation: Sort the data in the required order and flag the first treatment emergent record for each --DECOD for each subject.
AOCCIFL	1st Max Sev./Int. Occurrence Flag	Char	Y		Perm	Perm	Character indicator for the first occurrence of the event/intervention/finding with the maximum severity/intensity within the subject Example derivation: Sort the data in the required order and flag the first treatment emergent record for maximum severity for each subject.
AOCCPIFL	1st Max Sev./Int. Occur Within PT Flag	Char	Y		Perm	Perm	Character indicator for the first occurrence of the maximum severity/intensity within the subject and preferred term Example derivation: Sort the data in the required order and flag the first treatment emergent record for maximum severity within preferred term for each subject.
AOCCzzFL	1st Occurrence of ...	Char	Y		Perm	Perm	Additional flag variables as needed for analysis. Derivation rules for these flags need to be described in the metadata.

Table 3.2.6.2 MedDRA Occurrence Flag Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
AOCCSFL	1st Occurrence of SOC Flag	Char	Y		Perm	Perm	Character indicator for the first occurrence of the system organ class within the subject Example derivation: Sort the data in the required order and flag the first treatment-emergent record for each body system for each subject.
AOCCSIFL	1st Max Sev./Int. Occur Within SOC Flag	Char	Y		Perm	Perm	Character indicator for the first occurrence of the maximum severity/intensity within the subject and system organ class Example derivation: Sort the data in the required order and flag the first treatment-emergent record for maximum severity within body system for each subject.

3.2.7 Treatment/Dose Variables

The treatment variable used for analysis must be included. Typically, this would be TRTP, TRTA, TRTxxP, or TRTxxA. See ADaMIG v1.2 Section 3.2, ADSL Variables, and Section 3.3.2, Record-Level Treatment and Dose Variables for BDS Datasets, for more details on these variables (available at https://www.cdisc.org/standards/foundational/adam/). Additional dosing variables may also be included.

Table 3.2.7.1 Treatment/Dose Variables

Variable Name	Variable Label	Type	Codelist[a]	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
DOSEON	Treatment Dose at Record Start	Num			Perm	Perm	Dose received at the point in time of the record start date Example derivation: Obtained from EX.EXDOSE where --STDTC falls between the values of EX.EXSTDTC and EX.EXENDTC
DOSCUMA	Cumulative Actual Treatment Dose	Num			Perm	Perm	Cumulative actual study drug dosage at the point in time of the record start date
DOSEU	Treatment Dose Units	Char	(UNIT)		Cond	Cond	The units associated with DOSEON and/or DOSCUMA. Conditional on whether DOSEON and/or DOSCUMA are included.

[a]Codelists in parenthesis are the names of CDISC Controlled Terminology.

3.2.8 Descriptive Variables

Variables that describe the record are often used in analysis. Include these and any other SDTM variables if used in analysis. If the analysis version of the variable differs from the version in the SDTM, additional variables must be added using the conventions below and described in Section 3.2, Variable Metadata.

Table 3.2.8.1 lists some common descriptive variables that are often included in ADAE. Any other SDTM variables should be included as appropriate (e.g. AEOUT, AESDTH).

Table 3.2.8.1 Adverse Event Descriptive Variables

Variable Name	Variable Label	Type	Codelist[a]	Controlled Terms[b]	Core	SubClass ADVERSE EVENT Core	CDISC Notes
--SER	Serious Event	Char	(NY)		Perm	Req	XX.--SER
--SEV	Severity/Intensity	Char	(AESEV) or (SEVRS)		Perm	Cond	XX.--SEV For SubClass ADVERSE EVENT, conditional on whether the --SEV variable is in the SDTM AE dataset. Note that either --SEV or --TOXGR should be included in SDTM.
--SEVN	Severity/Intensity (N)	Num	1, 2, 3		Perm	Perm	Code XX.--SEV to numeric Low intensity should correspond to low value
ASEV	Analysis Severity/Intensity	Char		*	Perm	Perm	Apply imputation rules for missing severity of adverse events as specified in the SAP or metadata. May change case of text, such as from all uppercase in --SEV to mixed case in ASEV.
ASEVN	Analysis Severity/Intensity (N)	Num	1, 2, 3		Perm	Perm	Code ASEV to numeric Low intensity should correspond to low value
SEVGRy	Pooled Severity Group y	Char		*	Perm	Perm	Pooled grouping of AE severity for analysis (e.g., mild/moderate, severe)
SEVGRyN	Pooled Severity Group y (N)	Num		*	Perm	Perm	Code SEVGRy to numeric Low intensity should correspond to low value
--REL	Causality	Char		*	Perm	Req	XX.--REL

Variable Name	Variable Label	Type	Codelist[a]	Controlled Terms[b]	Core	SubClass ADVERSE EVENT Core	CDISC Notes
--RELN	Causality (N)	Num		*	Perm	Perm	Code XX.--REL to numeric Low relation should correspond to low value
AREL	Analysis Causality	Char		*	Perm	Perm	Apply imputation rules for missing causality of study drug as specified in the SAP or metadata. May change case of text, such as from all uppercase in --REL to mixed case in AREL.
ARELN	Analysis Causality (N)	Num		*	Perm	Perm	Code AREL to numeric
RELGRy	Pooled Causality Group y	Char		*	Perm	Perm	Pooled grouping of causality of study drug for analysis (e.g. related, not related)
RELGRyN	Pooled Causality Group y (N)	Num		*	Perm	Perm	Code of RELGRy to numeric Low relation should correspond to low value
--TOXGR	Standard Toxicity Grade	Char		*	Perm	Cond	XX.--TOXGR For SubClass ADVERSE EVENT, conditional on whether the --TOXGR variable is in the SDTM AE dataset. Note that either --SEV or --TOXGR should be included in SDTM.
--TOXGRN	Standard Toxicity Grade (N)	Num		*	Perm	Perm	Code --TOXGR to numeric Low toxicity should correspond to low value
ATOXGR	Analysis Toxicity Grade	Char		*	Perm	Perm	Toxicity grade for analysis. May be based on --TOXGR or an imputed or assigned value. May change case of text, such as from all uppercase in --TOXGR to mixed case in ATOXGR.
ATOXGRN	Analysis Toxicity Grade (N)	Num		*	Perm	Perm	Code ATOXGR to numeric Low toxicity should correspond to low value
TOXGGRy	Pooled Toxicity Grade Group y	Char		*	Perm	Perm	Pooled grouping of toxicity grade for analysis
TOXGGRyN	Pooled Toxicity Grade Group y (N)	Num		*	Perm	Perm	Code of TOXGGRy to numeric Low toxicity should correspond to low value
--ACN	Action Taken with Study Treatment	Char	(ACN)		Perm	Cond	XX.--ACN Required if XX.--ACN is present and populated

[a]Codelists in parenthesis are the names of CDISC Controlled Terminology.
[b]Asterisk indicates variable may be subject to producer-defined controlled terminology.

Medical history data typically does not contain descriptive variables. If needed for analysis, use variables as shown in Table 3.2.8.1 for adverse events, replacing the prefix "AE" with "MH".

Table 3.2.8.2 lists some common descriptive variables that are often included in ADCM. Any other SDTM variables should be included as appropriate.

Table 3.2.8.2 Concomitant Medications Descriptive Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	CDISC Notes
--STAT	Completion Status	Char			Perm	XX.--STAT
--INDC	Indication	Char			Perm	XX.--INDC
--DOSE	Dose per Administration	Num			Perm	XX.--DOSE
--DOSFRM	Dose Form	Char			Perm	XX.--DOSFRM
--DOSRGM	Intended Dose Regimen	Char			Perm	XX.--DOSRGM
--ROUTE	Route of Administration	Char			Perm	XX.--ROUTE

Concomitant medication descriptive variables are not used in the ADVERSE EVENT SubClass.

3.2.9 Standardized MedDRA Query Variables

Standardized MedDRA Queries (SMQs; see https://www.meddra.org/standardised-meddra-queries)[4] are becoming increasingly common in clinical trial safety evaluations, particularly when known or suspected safety issues are associated with experimental compounds. In addition, customized queries (CQs) are often used to modify an SMQ or identify records of special interest. Table 3.2.9.1 lists variables used to identify SMQs and CQs, where *zz* is replaced with a zero-padded 2-digit integer (01-99) for each SMQ or CQ of interest. This ordering can be based on importance or some other producer-defined criteria. It is recommended that ordering be consistent across studies within a development program, but it is recognized that there may be situations where this is not possible or practical.

Table 3.2.9.1 Standardized MedDRA Query Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms	Core	SubClass ADVERSE EVENT Core	CDISC Notes
SMQzzNAM	SMQ zz Name	Char			Cond	Cond	The Standardized MedDRA Query name. Would be blank for terms that are not in the SMQ. Therefore this variable could be blank for all records if no terms within the study were included in the SMQ. Conditional on whether SMQ analysis is done
SMQzzCD	SMQ zz Code	Num			Perm	Perm	The standardized MedDRA queries number code
SMQzzSC	SMQ zz Scope	Char	BROAD, NARROW		Cond	Cond	The search strategy for SMQs can be narrow or broad. The preferred terms that are narrow in scope have high specificity for identifying events of interest, whereas the broad terms have high sensitivity. By definition, all narrow terms are also considered within the broad scope. Therefore, to summarize all broad terms, terms with either narrow ***or*** broad would be considered. Will be null for terms that do not meet the criteria. Conditional on whether SMQ analysis is done
SMQzzSCN	SMQ zz Scope (N)	Num	1, 2		Perm	Perm	Will be null for terms that do not meet the criteria
CQzzNAM	Customized Query zz Name	Char			Cond	Cond	The CQ name or name of the adverse event of special interest category based on a grouping of terms. Would be blank for terms that are not in the CQ. Conditional on whether CQ analysis is done Examples: “DERMATOLOGICAL EVENTS” “CARDIAC EVENTS”, “IARS (INFUSION ASSOCIATED REACTIONS)”
ADECODy	Analysis Dictionary-Derived Term y	Char			Perm	Perm	The terms used for the analysis when combining multiple customized query or multiple standardized MedDRA queries and the original MedDRA dictionary terms under 1 variable Although designed for MedDRA queries, this variable could be used for other OCCDS analysis needs.

3.2.10 Original or Prior Coding Variables

Keeping original coding variables is not common, but there are a couple of instances where it might be helpful:

- When a study is mapped to one version of a mapping dictionary for an interim analysis and another for final analysis
- When studies using different versions of a mapping dictionary are pooled together for an integrated analysis

Because the dictionary-coded values may change with each version, counts for study tables can differ between versions for the same study. Therefore, keeping the original variables provides traceability to original (or prior) analysis(es).

The variables described below provide traceability to original (or prior) analysis(es). The suffix "w" represents an integer (1-9) corresponding to a previous version. Include the dictionary name and version as part of the metadata for each variable.

These variable names at this time are recommendations only. Additional prior or original variables needed for traceability may be added following the same conventions. There is an ADaM subteam working on integration, and this group may create different naming conventions for that type of analysis.

Table 3.2.10.1 Original or Prior MedDRA Coding Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms[a]	Core	SubClass ADVERSE EVENT Core	CDISC Notes
DECDORGw	PT in Original Dictionary w	Char		MedDRAw	Perm	Perm	Original preferred term coding of XX.--TERM using MedDRA or other dictionary version X.X.
BDSYORGw	SOC in Original Dictionary w	Char		MedDRAw	Perm	Perm	Original body system coding of XX.--TERM using MedDRA or other dictionary version X.X.
HLGTORGw	HLGT in Original Dictionary w	Char		MedDRAw	Perm	Perm	Original HLGT coding of XX.--TERM using MedDRA or other dictionary version X.X
HLTORGw	HLT in Original Dictionary w	Char		MedDRAw	Perm	Perm	Original HLT coding of XX.--TERM using MedDRA or other dictionary version X.X.
LLTORGw	LLT in Original Dictionary w	Char		MedDRAw	Perm	Perm	Original LLT coding of XX.--TERM using MedDRA or other dictionary version X.X.
LLTNORGw	LLT Code in Original Dictionary w	Char		MedDRAw	Perm	Perm	Original LLT code of XX.--TERM using MedDRA or other dictionary version X.X.

[a]For each version of an external dictionary, a different reference name must be used. The individual reference names will point to a dedicated section in the data definition file where all external dictionaries used in the analysis are listed, including dictionary name and version.

Table 3.2.10.2 Original or Prior WHO Drug Coding Variables

Variable Name	Variable Label	Type	Codelist	Controlled Terms[a]	Core	SubClass ADVERSE EVENT Core	CDISC Notes
DECDORGw	Standardized Med Name in Orig Dict w	Char		WHODRUGw	Perm	Perm	Original standardized medication name of CM.CMTRT using WHO Drug version X.X
CLASORGw	Medication Class in Orig Dictionary w	Char		WHODRUGw	Perm	Perm	Original medication class of CM.CMTRT using WHO Drug version X.X
CLCDORGw	Medication Class Code in Orig Dict w	Char		WHODRUGw	Perm	Perm	Original medication class code of CM.CMTRT using WHO Drug version X.X
ATyCORGw	ATC Level y Code in Orig Dictionary w	Char		WHODRUGw	Perm	Perm	Original ATC Level y code of CM.CMTRT using WHO Drug version X.X
ATyTORGw	ATC Level y Text in Orig Dictionary w	Char		WHODRUGw	Perm	Perm	Original ATC Level y text of CM.CMTRT using WHO Drug version X.X

[a]For each version of an external dictionary, a different reference name must be used. The individual reference names will point to a dedicated section in the data definition file where all external dictionaries used in the analysis are listed, including dictionary name and version.

3.2.11 User-specified Variable Naming Conventions

ADaM allows for user-specified variable names when creating additional variables not specified in this document or the ADaMIG. The ADaMIG has variable name fragments and other conventions, which are often needed with ADSL and BDS datasets. OCCDS datasets may have fewer derived variables, but there are some conventions and fragments described here to help build easily understood variable names.

Many variables from SDTM data are copied to ADaM with no change to the variable name, label, or content. ADaM follows a principle of harmonization (same name, same meaning, same values) as described in the ADaMIG Section 3.1.1, General Naming Conventions. This means that when an SDTM variable name is in an ADaM dataset, the values within that variable are exactly the same as the SDTM data.

Many times, a modified version of an SDTM variable is needed, such as when missing data is imputed. As shown in Section 3.2.8, Descriptive Variables, when creating an analysis version of an SDTM descriptive variable, the prefix "A" (for Analysis) replaces the 2-letter SDTM domain code. This same convention can be used to create analysis versions of variables not specified in Section 3.2.8. For example, if changes needed to be made to AEBODSYS for analysis, the variable name ABODSYS (with a prefix of "A" instead of "AE") would convey that it is the analysis version of the AEBODSYS variable, and that the value of ABODSYS may not be the same as AEBODSYS on some records.

When variable content remains unmodified but the name of the SDTM variable cannot be used in the analysis dataset, the prefix "U" (for "unmodified") can be used instead of the 2-letter SDTM domain code. For example, if combining records from Adverse Events (AE) and Medical History (MH) into an analysis dataset, it would likely not be analysis-ready to have body system information spread across 2 variables, AEBODSYS on the rows from the AE dataset and MHBODSYS on the rows from the MH dataset. Instead, a single variable that contains the value from AEBODSYS or MHBODSYS is needed. Using the "A" prefix and calling the variable ABODSYS implies that it is an analysis version of a variable, and that the value of ABODSYS may not be the same as AEBODSYS and MHBODSYS on some records. If no changes are made to AEBODSYS and MHBODSYS, using instead a "U" prefix allows for a single column called UBODSYS but makes it clear that no modifications to the content of the SDTM variables were made. An example with variables using the "U" prefix can be found in Section 10, Example 7: Analysis of Adverse Events from Multiple Input Domains.

3.3 Other Metadata

Because OCCDS does not use parameters, there is typically no need for value-level metadata. However value level metadata may be useful in some cases, as illustrated in Section 14, Example 11: Analysis of Protocol Deviations.

The other type of ADaM metadata which may be included is the analysis results metadata. The CDISC Define-XML v2.1 and Analysis Results Metadata (ARM) v1.0 for Define-XML v2.0 release packages (available at https://www.cdisc.org/standards/data-exchange/define-xml) both contain examples of how to represent analysis results metadata.

4 Example 1: Analysis of Treatment-emergent Adverse Event

The basic summary of adverse event frequencies described in ICH Guideline E3[5] Sections 12.2.2 and 14.3.1 should be used to display frequencies in treatment and control groups.

This example displays a simple summary of all treatment-emergent adverse events. The example is based on a 2-treatment parallel design study. The display summarizes (1) the number of subjects in each treatment group in whom the adverse event occurred and (2) the rate of occurrence in each treatment group.

4.1 Analysis Display Example Layout - Example 1

Figure 4.1.1 Example of Summary of Treatment-emergent Adverse Events[a]

Summary of Treatment-emergent Adverse Events by
System Organ Class and Preferred Term
Analysis Population: Safety

SYSTEM ORGAN CLASS Preferred Term	Treatment A (N = xxx) n (%)	Treatment B (N = xxx) n (%)
Number of subjects reporting at least one adverse event	x (x.x)	x (x.x)
BLOOD AND LYMPHATIC SYSTEM DISORDERS		
At least one event	x (x.x)	x (x.x)
Anaemia	x (x.x)	x (x.x)
…	x (x.x)	x (x.x)
CARDIAC DISORDERS		
At least one event	x (x.x)	x (x.x)
Angina pectoris	x (x.x)	x (x.x)
Coronary artery disease	x (x.x)	x (x.x)
Ventricular tachycardia	x (x.x)	x (x.x)
Myocardial infarction	x (x.x)	x (x.x)
…	x (x.x)	x (x.x)
<Other SOCs and PTs>		

Page 1 of x

N = Safety subjects, i.e., subjects who received at least one dose of study drug
n = Number of subjects reporting at least one treatment emergent adverse event
% = n / N * 100
Adverse events are presented by descending frequency within Treatment B
System organ classes and preferred terms are coded using MedDRA version x.x.

[a]The style of the display of the results of an analysis is determined by the producer. The example is intended to illustrate content, not appearance.

4.2 Sample ADaM Variable Metadata - Example 1

This example describes an adverse events ADaM dataset named ADAE. ADAE is not a required dataset name.

Table 4.2.1 Example of ADaM Variable Metadata

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADAE	STUDYID	Study Identifier	text			AE.STUDYID	
ADAE	USUBJID	Unique Subject Identifier	text			AE.USUBJID	
ADAE	AESEQ	Sequence Number	integer			AE.AESEQ	
ADAE	AETERM	Reported Term for the Adverse Event	text			AE.AETERM	
ADAE	AEDECOD	Dictionary-Derived Term	text		MedDRA	AE.AEDECOD MedDRA Version XX.X	
ADAE	AEBODSYS	Body System or Organ Class	text		MedDRA	AE.AEBODSYS MedDRA Version XX.X	
ADAE	TRTEMFL	Treatment Emergent Analysis Flag	text	Y			If ADSL.TRTSDT <= ASTDT<=(ADSL.TRTEDT +14) then TRTEMFL="Y"
ADAE	PREFL	Pre-treatment Flag	text	Y			If ASTDT < ADSL.TRTSDT then PREFL="Y"
ADAE	FUPFL	Follow-up Flag	text	Y			If ASTDT > ADSL.TRTEDT+14 then FUPFL="Y"
ADAE	AESTDTC	Start Date/Time of Adverse Event	date		ISO 8601	AE.AESTDTC	
ADAE	ASTDT	Analysis Start Date	integer				<Producer will insert derivation here>
ADAE	ASTDTF	Analysis Start Date Imputation Flag	text	D, M, Y			If start date is completely missing or missing the year then ASTDTF="Y" Else if start date has month missing then ASTDTF="M" Else if start date has day missing then ASTDTF="D"
ADAE	AEENDTC	End Date/Time of Adverse Event	date		ISO 8601	AE.AEENDTC	
ADAE	AENDT	Analysis End Date	integer				<Producer will insert derivation here>
ADAE	AENDTF	Analysis End Date Imputation Flag	text	D, M, Y			If end date is completely missing or missing the year then AENDTF="Y" Else if end date has month missing then AENDTF="M" Else if end date has day missing then AENDTF="D"
ADAE	AESER	Serious Event	text	Y, N		AE.AESER	
ADAE	APHASE	Phase	text	PRE-TREATMENT, TREATMENT, FOLLOW-UP			If ASTDT<ADSL.TRTSDT, then APHASE="PRE-TREATMENT" Else if ASTDT > ADSL.TRTEDT + 14 days then APHASE="FOLLOW-UP", Else APHASE="TREATMENT"
ADAE	AESEV	Severity/Intensity	text	MILD, MODERATE, SEVERE		AE.AESEV	
ADAE	ASEV	Analysis Severity/Intensity	text	Mild, Moderate, Severe			If AE.AESEV="MILD" then ASEV="Mild" Else if AE.AESEV="MODERATE" then ASEV="Moderate" Else if AE.AESEV is equal to "SEVERE" or Severity/Intensity is missing then ASEV="Severe"

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADAE	ASEVN	Analysis Severity/Intensity (N)	integer	1, 2, 3			Map ASEV to ASEVN in the following manner: "Mild" = 1 "Moderate" = 2 "Severe" = 3
ADAE	AEREL	Causality	text	NOT RELATED, UNLIKELY RELATED, POSSIBLY RELATED, PROBABLY RELATED, DEFINITELY RELATED		AE.AEREL	
ADAE	RELGR1	Pooled Causality Group 1	text	Not Related, Related			If AE.AEREL is equal to "NOT RELATED" or "UNLIKELY RELATED" then RELGR1="Not Related" Else if AE.AEREL is equal to "POSSIBLY RELATED" or "PROBABLY RELATED" or "DEFINITELY RELATED" or Causality is missing then RELGR1="Related"
ADAE	RELGR1N	Pooled Causality Group 1 (N)	integer	0, 1			Map RELGR1 to RELGR1N in the following manner: "Not Related" = 0 "Related" = 1
ADAE	SAFFL	Safety Population Flag	text	Y, N		ADSL.SAFFL	
ADAE	AOCCFL	1st Occurrence within Subject Flag	text	Y			Subset ADAE to Treatment Emergent Adverse Events (TRTEMFL="Y") Sort by Subject (USUBJID), Analysis Start Date (ASTDT), and Sequence Number (AESEQ) and flag the first record (set AOCCFL="Y") within each Subject
ADAE	AOCCSFL	1st Occurrence of SOC Flag	text	Y			Subset ADAE to Treatment Emergent Adverse Events (TRTEMFL="Y") Sort by Subject (USUBJID), System Organ Class (AEBODSYS), Analysis Start Date (ASTDT), and Sequence Number (AESEQ) and flag the first record (set AOCCSFL ="Y") within each Subject and SOC
ADAE	AOCCPFL	1st Occurrence of Preferred Term Flag	text	Y			Subset ADAE to Treatment Emergent Adverse Events (TRTEMFL="Y") Sort by Subject (USUBJID), System Organ Class (AEBODSYS), Preferred Term (AEDECOD) Analysis Start Date (ASTDT), and Sequence Number (AESEQ) and flag the first record (set AOCCPFL ="Y") within each Subject, SOC, and PT
ADAE	TRTA	Actual Treatment	text	Drug A, Drug B		ADSL.TRT01A	
ADAE	TRTAN	Actual Treatment (N)	integer	1, 2		ADSL.TRT01AN	Drug A = 1 Drug B = 2
ADAE	TRTSDT	Date of First Exposure to Treatment	integer			ADSL.TRTSDT	
ADAE	TRTEDT	Date of Last Exposure to Treatment	integer			ADSL.TRTEDT	
ADAE	AGE	Age	integer			ADSL.AGE	
ADAE	AGEGR1	Pooled Age Group 1	text	<65, >=65		ADSL. AGEGR1	
ADAE	SEX	Sex	text	M, F		ADSL.SEX	

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADAE	RACE	Race	text	BLACK OR AFRICAN AMERICAN, AMERICAN INDIAN OR ALASKA NATIVE, ASIAN, NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER, WHITE		ADSL.RACE	

4.3 Sample ADaM Data - Example 1

Table 4.3.1 illustrates the adverse events analysis dataset (ADAE) defined in Section 4.2, Sample ADaM Variable Metadata - Example 1. The ADAE dataset in this example was designed to support some standard subsets and/or classifications of treatment-emergent adverse events including seriousness, severity, and relationship to study drug. The example describes some of the key variables and records that would be included in the dataset.

Key points to note in the example are:

1. The producer of the dataset chose to use a record-level actual treatment variable (TRTA) populated with the same value across all rows in the dataset rather than a subject-level treatment variable (TRT01A). For a parallel design, either TRTA or TRT01A could be used as the actual treatment identifier. The producer interpreted TRTA as the treatment associated with the record for analysis display purposes and populated the pre-treatment records with treatment even though subjects had not yet received treatment at that time.
2. Variables such as AESEQ, AETERM, and AESTDTC are copied in from the SDTM AE domain to provide data-point traceability.
3. Variables such as AEBODSYS, AEDECOD, AESER, AESEV, and AEREL are copied in from the SDTM AE domain for analysis purposes.
4. ASTDT is the AE timing variable used for analysis. Other timing variables (e.g., AENDT, ASTDTF, AENDTF, AESTDTC, AEENDTC, TRTSDT, TRTEDT) are supportive variables for metadata traceability.
5. The addition of ASEV and RELGR1 allows for the imputation of missing severity and grouping and imputation of relationship to study drug as specified in the statistical analysis plan (SAP).
6. The occurrence flags (AOCC*FL) are permissible. The main purpose of these flags is to facilitate data-point traceability between records in the dataset and unique counts in the summary displays. In addition, if a time-to-event (TTE) analysis is built using adverse events, the flags provide a crucial link between the summary records in the TTE BDS and the source of the records in ADAE.
7. The core variables of AGE, AGEGR1, SEX, and RACE are included in ADAE to facilitate subgroup analyses and data review.

Table 4.3.1 Sample ADaM Data[a]

Row	STUDYID	USUBJID	AESEQ	AETERM	AEDECOD	AEBODSYS	TRTEMFL	PREFL	FUPFL	AESTDTC	ASTDT	ASTDTF	AEENDTC	AENDT	AENDTF	AESER	APHASE	AESEV	ASEV	ASEVN	AEREL	RELGR1	RELGR1N	SAFFL	AOCCFL	AOCCSFL	AOCCPFL	TRTA	TRTAN	TRTSDT	TRTEDT	AGE	AGEGR1	SEX	RACE
1	XYZ	XYZ-001-001	1	HEADACHE	Headache	Nervous system disorders		Y		2006-01	01JAN2006	D	2006-01-22	22JAN2006		N	PRE-TREATMENT	MILD	Mild	1	NOT RELATED	Not Related	0	Y				Drug A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN
2	XYZ	XYZ-001-001	2	CHRONIC BACK PAIN	Back pain	Musculoskeletal and connective tissue disorders		Y		2006-01-21	21JAN2006		2006-01-28	28JAN2006		N	PRE-TREATMENT	MODERATE	Moderate	2	NOT RELATED	Not Related	0	Y				Drug A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN
3	XYZ	XYZ-001-001	3	NOSE BLEEDING RIGHT NOSTRIL	Epistaxis	Respiratory, thoracic and mediastinal disorders		Y		2006-01-22	22JAN2006		2006-01-22	22JAN2006		N	PRE-TREATMENT	MILD	Mild	1	NOT RELATED	Not Related	0	Y				Drug A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN
4	XYZ	XYZ-001-001	4	PROBLEMS OF HYPOTENSION	Hypotension	Vascular disorders	Y				23JAN2006	Y		15MAY2006	Y	N	TREATMENT	MILD	Mild	1	POSSIBLY RELATED	Related	1	Y	Y	Y	Y	Drug A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN
5	XYZ	XYZ-001-001	5	HEADACHE	Headache	Nervous system disorders	Y			2006-01-24	24JAN2006		2006-01	31JAN2006	D	N	TREATMENT	MODERATE	Moderate	2	PROBABLY RELATED	Related	1	Y		Y	Y	Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
6	XYZ	XYZ-001-001	6	HEADACHE	Headache	Nervous system disorders	Y			2006-02	01FEB2006	D	2006-02-05	05FEB2006		N	TREATMENT	SEVERE	Severe	3	PROBABLY RELATED	Related	1	Y				Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
7	XYZ	XYZ-001-001	7	LOOSE STOOL	Diarrhoea	Gastrointestinal disorders	Y			2006-03-05	05MAR2006		2006-03-06	06MAR2006		N	TREATMENT		Severe	3	DEFINITELY RELATED	Related	1	Y		Y	Y	Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
8	XYZ	XYZ-001-001	8	ABDOMINAL DISCOMFORT	Abdominal discomfort	Gastrointestinal disorders	Y			2006-03-05	05MAR2006		2006	15MAY2006	M	N	TREATMENT	MODERATE	Moderate	2	DEFINITELY RELATED	Related	1	Y			Y	Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
9	XYZ	XYZ-001-001	9	DIARRHEA	Diarrhoea	Gastrointestinal disorders	Y			2006-03-17	17MAR2006		2006-03-18	18MAR2006		N	TREATMENT	MODERATE	Moderate	2	DEFINITELY RELATED	Related	1	Y				Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
10	XYZ	XYZ-001-001	10	ABDOMINAL FULLNESS DUE TO GAS	Abdominal distension	Gastrointestinal disorders	Y			2006-03-17	17MAR2006		2006-03-19	19MAR2006		N	TREATMENT	MILD	Mild	1	DEFINITELY RELATED	Related	1	Y			Y	Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
11	XYZ	XYZ-001-001	11	NAUSEA (INTERMITTENT)	Nausea	Gastrointestinal disorders	Y			2006-04-20	20APR2006		2006-04-22	22APR2006		N	TREATMENT	MILD	Mild	1	PROBABLY RELATED	Related	1	Y			Y	Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
12	XYZ	XYZ-001-001	12	WEAKNESS	Asthenia	General disorders and administration site conditions	Y			2006-05-17	17MAY2006		2006-05-20	20MAY2006		N	TREATMENT	MILD	Mild	1	POSSIBLY RELATED	Related	1	Y		Y	Y	Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
13	XYZ	XYZ-001-001	13	HEADACHE	Headache	Nervous system disorders	Y			2006-05-20	20MAY2006		2006-05-22	22MAY2006		N	TREATMENT	MILD	Mild	1	UNLIKELY RELATED	Not Related	0	Y				Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
14	XYZ	XYZ-001-001	14	HEADACHE	Headache	Nervous system disorders	Y			2006-05-23	23MAY2006		2006-06-27	27JUN2006		N	TREATMENT	MILD	Mild	1	UNLIKELY RELATED	Not Related	0	Y				Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
15	XYZ	XYZ-001-001	15	HYPOTENSIVE	Hypotension	Vascular disorders	Y			2006-05-21	27MAY2006		2006-05-25	29MAY2006		Y	TREATMENT	SEVERE	Severe	3	UNLIKELY RELATED	Not Related	0	Y				Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN
16	XYZ	XYZ-001-001	16	HEADACHE	Headache	Nervous system disorders			Y	2006-06-01	01JUN2006		2006-06-01	01JUN2006		N	FOLLOW-UP	MILD	Mild	1	UNLIKELY RELATED	Not Related	0	Y				Drug A	1	23JAN 2006	15MAY 2006	54	<65	M	ASIAN

[a]Variables ending in suffix DTC are character Date/Time fields in ISO 8601 format. Variables ending in DT are numeric dates, here shown using SAS date format date9. Other numeric date formats can be used, but care should be taken with newer date formats which might not be understood by all statistical packages.

5 Example 2: Analysis of Hemorrhages (SMQ) among Treatment-emergent Adverse Events by Sex

This example demonstrates how to incorporate Standardized MedDRA Queries (SMQs) into an adverse event analysis data set. In this example, an SMQ for hemorrhages is being used. This particular SMQ is hierarchical with only narrow-scope terms, including terms referring to different types of hemorrhage, hematoma, bleeding, and so on. (For a full description of SMQs, see the *Introductory Guide for Standardized MedDRA Queries*.[4])

Key points to note in the example are:

1. The exact name of the SMQ being used in this example is "Haemorrhages (SMQ)". This precise terminology is used throughout the displays.
2. This particular SMQ contains only narrow scope terms. However, in order to illustrate best practice, the scope is also specified when a reference is made to the SMQ. Although redundant in this particular case, it is important to show which scope is being used when providing SMQ-based summaries because the scope can have a profound effect on the percentage of subjects who meet certain SMQ criteria.

5.1 Analysis Display Example Layouts - Example 2

Note: The following examples are intended to illustrate content, not appearance. The style of the display of the results of an analysis is determined by the producer.

Figure 5.1.1 Example SMQ Summary

Summary of Haemorrhages (SMQ) (Narrow Scope) Adverse Events by
Sex and Actual Treatment Group
Analysis Population: Safety

	Gender n (%)			
	Females		Males	
Preferred Term	B (N=281)	A (N=166)	B (N=297)	A (N=158)
Any Haemorrhages (SMQ) (Narrow Scope) Event	36 (12.8)	48 (28.9)	26 (8.8)	31 (19.6)
Cerebral haemorrhage	11 (3.9)	15 (9.0)	6 (2.0)	13 (8.2)
Conjunctival haemorrhage	0	1 (0.6)	0	0
Ecchymosis	1 (0.4)	0	0	0
Epistaxis	0	1 (0.6)	0	0
Extradural haematoma	1 (0.4)	0	1 (0.3)	1 (0.6)
Gastrointestinal haemorrage	10 (3.6)	4 (2.4)	8 (2.7)	6 (3.8
Haematuria	1 (0.4)	2 (1.2)	0	3 (1.9)
Haemoptysis	1 (0.4)	1 (0.6)	0	0
Haemorrhage	1 (0.4)	2 (1.2)	0	0
Infusion site haemorrhage	1 (0.4)	4 (2.4)	2 (0.7)	2 (1.3)
Melaena	0	0	0	1 (0.6)
Petechiae	0	1 (0.6)	0	0
Subarachnoid haemorrhage	14 (5.0)	24 (14.5)	12 (4.0)	11 (7.0)
Subdural haemorrhage	2 (0.7)	2 (1.2)	0	0

Figure 5.1.2 Example of SMQ Mosaic Plot

Mosaic Plot of Haemorrhages (SMQ) (Narrow Scope) Preferred Terms
by Sex and Actual Treatment Group
Analysis Population: Safety

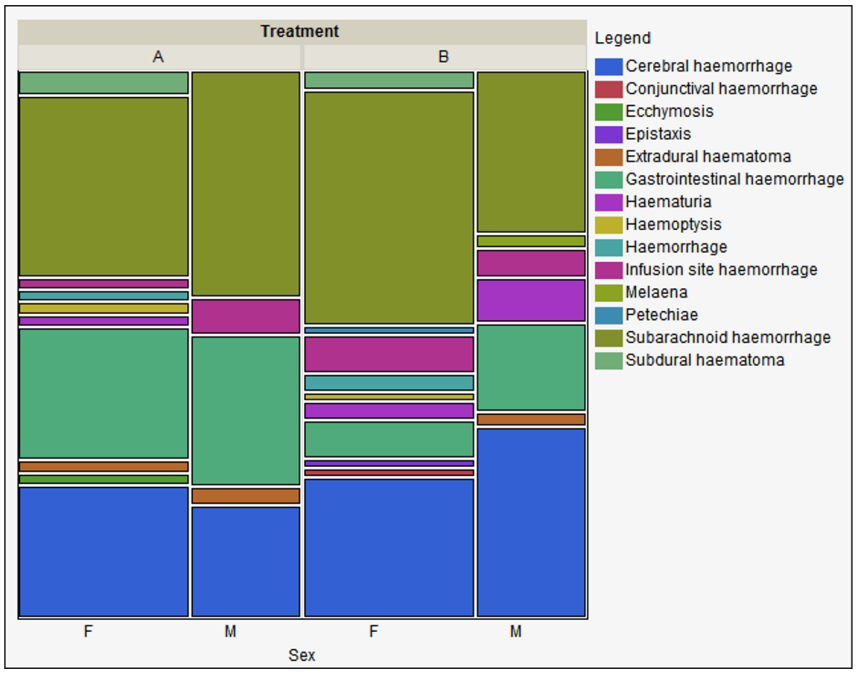
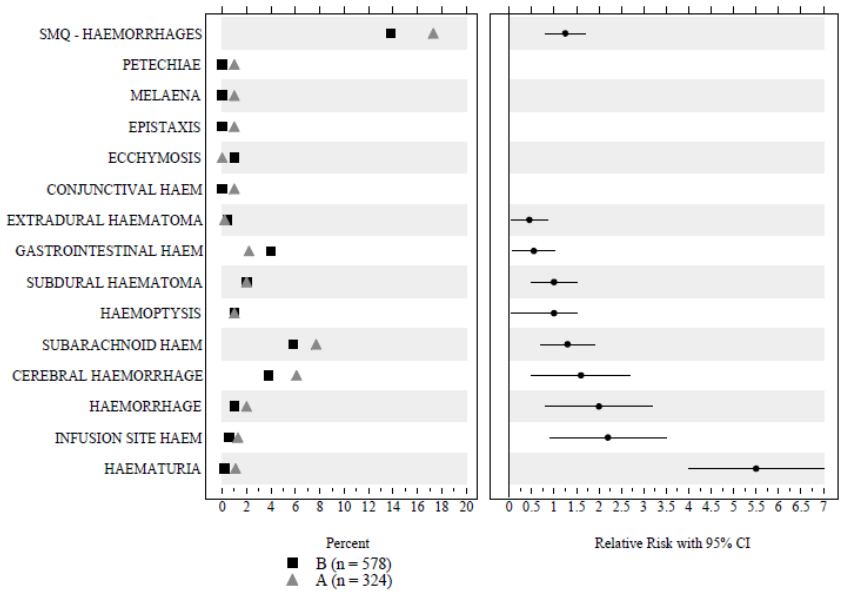


Figure 5.1.2 Example of SMQ Relative Risk Summary

Summary of Haemorrhages (SMQ) (Narrow Scope) Preferred Terms
Sorted by Relative Risk
Analysis Population: Safety



5.2 Sample ADaM Variable Metadata - Example 2

This example describes an adverse events ADaM dataset named ADAE. ADAE is not a required dataset name. In Table 5.2.1, 4 variables relate to the primary SMQ of interest (hemorrhage terms): SMQ01CD, SMQ01NAM SMQ01SC, and SMQ01SCN. The first 2 of these variables, SMQ01CD and SMQ01NAM, contain the numeric code and name for the SMQ from MedDRA. The next 2 variables, SMQ01SC and SMQ01SCN, are character and numeric variables, respectively, that indicate not only whether or not the given AE meets the criteria for the given SMQ, but also whether the term meets the SMQ's broad or narrow scope (the *SC* suffix is for "scope").

Table 5.2.1 Example of ADaM Variable Metadata

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation	Comments
ADAE	USUBJID	Unique Subject Identifier	text			AE.USUBJID		
ADAE	AETERM	Reported Term for the Adverse Event	text			AE.AETERM		
ADAE	AEDECOD	Dictionary-Derived Term	text		MedDRA	AE.AEDECOD		
ADAE	AEBODSYS	Body System or Organ Class	text		MedDRA	AE.AEBODSYS		
ADAE	ASTDT	Analysis Start Date	integer				<Producer will insert derivation here>	
ADAE	AEPTCD	Preferred Term Code	integer		MedDRA	AE.AEPTCD		
ADAE	SMQ01CD	SMQ 01 Code	integer	20000039			Set equal to the SMQ code, if the AEPTCD is included in this SMQ.	
ADAE	SMQ01NAM	SMQ 01 Name	text	Haemorrhage terms (excl. laboratory terms) (SMQ)				Set equal to the controlled term, if the AEPTCD is included in this SMQ.
ADAE	SMQ01SC	SMQ 01 Scope	text	NARROW				For this given SMQ, all scopes are Narrow.
ADAE	SMQ01SCN	SMQ 01 Scope (N)	integer	2			If SMQ01SC="NARROW" then SMQ01CN=2	

5.3 Sample ADaM Data - Example 2

Table 5.3.1: Sample ADaM Data Showing SMQ Variables[a]

Row	USUBJID	AETERM	AEDECOD	AEBODSYS	ASTDT	AEPTCD	SMQ01CD	SMQ01NAM	SMQ01SC	SMQ01SCN
1	0092017	SCLERAL BLEED RIGHT EYE	Scleral haemorrhage	Eye disorders	09JUN2009	10050508	20000039	Haemorrhage terms (excl laboratory terms) (SMQ)	NARROW	2
2	0112012	BRUISING OF LEFT UPPER ARM	Contusion	Injury, poisoning and procedural complications	27AUG2008	10050584	20000039	Haemorrhage terms (excl laboratory terms) (SMQ)	NARROW	2
3	0112012	BRUISING TO LEFT WRIST	Contusion	Injury, poisoning and procedural complications	22AUG2007	10050584	20000039	Haemorrhage terms (excl laboratory terms) (SMQ)	NARROW	2
4	0112013	NAUSEA	Nausea	Gastrointestinal disorders	16JUN2010	10028813				
5	0112014	NOSE BLEEDING	Epistaxis	Respiratory, thoracic and mediastinal disorders	22NOV2009	10015090	20000039	Haemorrhage terms (excl laboratory terms) (SMQ)	NARROW	2
6	0122006	EPISTAXIS	Epistaxis	Respiratory, thoracic and mediastinal disorders	06NOV2009	10015090	20000039	Haemorrhage terms (excl laboratory terms) (SMQ)	NARROW	2

[a]Variables ending in DT are numeric dates, here shown using SAS date format date9. Other numeric date formats can be used, but care should be taken with newer date formats which might not be understood by all statistical packages.

6 Example 3: Analysis of Peripheral Sensory Neuropathy (PSN) Adverse Events by Severity and Cumulative Dose Exposure

Some institutions and organizations use standardized coding guidelines (e.g., National Cancer Institute, ACTG's antiviral therapeutic areas) for reporting of adverse events. These scales may be based upon variables as collected on AE CRFs, such as a grading scheme based upon severity (AESEV/AESEVN). Other guidelines may be so objective that some variables—for example, drug relatedness (AEREL/AERELN)—are not captured.

In this example, the AE analysis dataset is used to summarize the frequency of peripheral sensory neuropathy (PSN) by cumulative dose exposure in an oncology study. PSN was collected on the study CRF at each cycle and at each 6-month follow-up, and graded as:

- 0 = None
- 1 = Asymptomatic; loss of deep tendon reflexes or paresthesia
- 2 = Moderate symptoms; limiting instrumental activities of daily living (ADL)
- 3 = Severe symptoms; limiting self-care ADL
- 4 = Life-threatening consequences; urgent intervention indicated
- 5 = Death

All these PSN events were coded to "Paresthesia".

6.1 Analysis Display Example Layout - Example 3

Figure 6.1.1 Example of Adverse Event Summary by Cumulative Dose Exposure[a]

Summary of Cumulative Dose Quartiles to First Onset for PSN
by Severity Grade
Analysis Population: Intent-to-Treat

Cumulative dose	Number of patients Exposed	PSN grade			
		Number (%) of patients with grade ≥ 1	Number (%) of patients with grade ≥ 2	Number (%) of patients with grade ≥ 3	Number (%) of patients with grade 4 or 5
Total number of patients with PSN		x (x.x)	x (x.x)	x (x.x)	x (x.x)
1st quartile	N	x (x.x)	x (x.x)	x (x.x)	x (x.x)
2nd quartile	N	x (x.x)	x (x.x)	x (x.x)	x (x.x)
3rd quartile	N	x (x.x)	x (x.x)	x (x.x)	x (x.x)
4th quartile	N	x (x.x)	x (x.x)	x (x.x)	x (x.x)
Median cumulative dose to first onset (mg/m^2)		X	X	X	X

[a]The style of the display of the results of an analysis is determined by the producer. This example is intended to illustrate content, not appearance.

6.2 Sample ADaM Variable Metadata - Example 3

This example describes an Adverse Events ADaM dataset named ADAE. ADAE is not a required dataset name.

Table 6.2.1: Sample ADaM Variable Metadata for Selected Variables

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADAE	USUBJID	Unique Subject Identifier	text			AE.USUBJID	
ADAE	ITTFL	Intent-to-Treat	text	Y, N		ADSL.ITTFL	

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
		Population Flag					
ADAE	AEDECOD	Dictionary-Derived Term	text		MedDRA	AE.AEDECOD	
ADAE	AETOXGR	Standard Toxicity Grade	text	1, 2, 3, 4, 5		AE.AETOXGR	
ADAE	AETOXGRN	Standard Toxicity Grade (N)	integer	1, 2, 3, 4, 5			Code AE.AETOXGR to numeric
ADAE	DOSCUMA	Cumulative Actual Treatment Dose	float				Total all values of EX.EXDOSE for the subject up to the start of the adverse event
ADAE	DOSEU	Treatment Dose Units	text	mg		EX.EXDOSEU	
ADAE	DOSCUMG1	Cumulative Dose Group 1	text	1st Quartile, 2nd Quartile, 3rd Quartile, 4th Quartile			Missing if DOSCUMA=0, else DOSCUMG1 = 1st Quartile if DOSCUMA is in the 1st quartile, 2nd Quartile if in the 2nd quartile, 3rd Quartile if in the 3rd quartile and 4th Quartile if in the 4th quartile.

6.3 Sample ADaM Data - Example 3

Key points to note in the example are:

1. Variable DOSCUMG1 is not a standard variable. It has been added for analysis purposes and uses the naming conventions from the ADaMIG.
2. This is a simple example to only illustrate the cumulative dose variables that can be added to ADAE. It does not include additional variables that would also be needed for analysis, such as a flag to indicate the first occurrence of PSN.
3. Rows 3 and 7 include 2 patients who had no dose of study drug at the time of PSN and would not be included in the table.

Table 6.3.1 Sample ADaM Data Showing Cumulative Dose Variables

Row	USUBJID	ITTFL	AEDECOD	AETOXGR	AETOXGRN	DOSCUMA	DOSEU	DOSCUMG1
1	101-002	Y	PARESTHESIA	3	3	247.06	mg	1st Quartile
2	101-003	Y	PARESTHESIA	2	2	674.02	mg	3rd Quartile
3	101-005	Y	PARESTHESIA	1	1	0	mg	
4	101-006	Y	PARESTHESIA	2	2	900.00	mg	4th Quartile
5	101-008	Y	PARESTHESIA	4	4	493.30	mg	2nd Quartile
6	101-010	Y	PARESTHESIA	3	3	894.29	mg	4th Quartile
7	101-012	Y	PARESTHESIA	1	1	0	mg	

7 Example 4: Analysis of Treatment-emergent Adverse Events in a Cross-Over Interaction Study

This example is a phase 1, open-label, 3-period cross-over study. Subjects are treated for 7 days within each period, with a 7-day wash-out between periods. In each period, subjects receive 1 of 3 treatments (A, B, or A + B combined) in order of the sequence to which they are randomized. Treatment-emergent adverse events were defined as adverse events that occurred or worsened from the start of the treatment period through 72 hours after the end of the treatment period. Non-treatment emergent adverse events were those that occurred before the first treatment period or more than 72 hours after the end of the treatment period until the start of the next treatment period. Post-treatment emergent adverse events were those that occurred more than 72 hours after the last treatment period.

7.1 Analysis Display Example Layout - Example 4

Figure 7.1.1 Example of Adverse Event Summary for a Cross-over Study[a]

Summary of Treatment-emergent Adverse Events by System Organ Class and
Preferred Term and Treatment Group
Analysis Population: Safety

	Treatment A (N = xxx)		Treatment B (N = xxx)		Treatment A + B (N = xxx)	
SYSTEM ORGAN CLASS **Preferred Term**	**n (%)**	**No. of events**	**n (%)**	**No. of events**	**n (%)**	**No. of events**
Any TEAE	x (x.x)	x	x (x.x)	x	x (x.x)	x
GASTROINTESTINAL DISORDER	x (x.x)	x	x (x.x)	x	x (x.x)	x
Nausea	x (x.x)	x	x (x.x)	x	x (x.x)	x
Constipation	x (x.x)	x	x (x.x)	x	x (x.x)	x
Vomiting	x (x.x)	x	x (x.x)	x	x (x.x)	x
Diarrhoea	x (x.x)	x	x (x.x)	x	x (x.x)	x
INFECTIONS AND INFESTATIONS	x (x.x)	x	x (x.x)	x	x (x.x)	x
Pharyngitis	x (x.x)	x	x (x.x)	x	x (x.x)	x
NERVOUS SYSTEM DISORDERS	x (x.x)	x	x (x.x)	x	x (x.x)	x
Headache	x (x.x)	x	x (x.x)	x	x (x.x)	x
Dizziness	x (x.x)	x	x (x.x)	x	x (x.x)	x
Syncope	x (x.x)	x	x (x.x)	x	x (x.x)	x
<Other SOCs and PTs>						

TEAE = treatment emergent adverse event
N = Safety subjects, i.e., subjects who received at least one dose of study drug in that particular period
n = Number of subjects reporting at least one treatment emergent adverse event
% = n / N * 100
Adverse events are presented by descending frequency of SOC and PT within SOC within Treatment A+B
System organ classes and preferred terms are coded using MedDRA version x.x.

[a]The style of the display of the results of an analysis is determined by the producer. This example is intended to illustrate content, not appearance.

7.2 Sample ADaM Variable Metadata - Example 4

This example describes an adverse events ADaM dataset named ADAE. ADAE is not a required dataset name.

Table 7.2.1 Sample ADaM Variable Metadata for Selected Variables

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADAE	USUBJID	Unique Subject Identifier	text			AE.USUBJID	
ADAE	TRTA	Actual Treatment	text	Treatment A, Treatment B, Treatment A+B			ADSL.TRT01A if in the first period, ADSL.TRT02A if in the second period, or ADSL.TRT03A if in the third period
ADAE	TRTAN	Actual Treatment	integer	1, 2, 3			Code TRTA to numeric Treatment A = 1 Treatment B = 2 Treatment A+B = 3
ADAE	SAFFL	Safety Population Flag	text	Y, N		ADSL.SAFFL	
ADAE	AEBODSYS	Body System or Organ Class	text		MedDRA	AE.AEBODSYS	
ADAE	AEDECOD	Dictionary-Derived Term	text		MedDRA	AE.AEDECOD	
ADAE	ASTDTM	Analysis Start Datetime	integer				Converting AE.AESTDTC from character ISO8601 format to numeric date format, applying producer defined imputation rules
ADAE	ASTDTF	Analysis Start Date Imputation Flag	text	D, M, Y			The level of imputation done for the start date (D if day was imputed, M if month was imputed, Y if year was imputed)
ADAE	ASTTMF	Analysis Start Time Imputation Flag	text	M, H			The level of imputation done for the start time (H if hour was imputed, M if minutes were imputed)
ADAE	TRTEMFL	Treatment Emergent Analysis Flag	text	Y			If ADSL.TR01SDTM LE ASTDTM LE (ADSL.TR01EDTM+72 hours) or ADSL.TR02SDTM LE ASTDTM LE (ADSL.TR02EDTM+72 hours) or ADSL.TR03SDTM LE ASTDTM LE (ADSL.TR03EDTM+72 hours) then TRTEMFL=Y
ADAE	TREM01FL	Treatment Emergent Period 01 Flag	text	Y			Y if ADSL.TR01SDTM LE ASTDTM LE (ADSL.TR01EDTM+72 hours)
ADAE	TREM02FL	Treatment Emergent Period 02 Flag	text	Y			Y if ADSL.TR02SDTM LE ASTDTM LE (ADSL.TR02EDTM+72 hours)
ADAE	TREM03FL	Treatment Emergent Period 03 Flag	text	Y			Y if ADSL.TR03SDTM LE ASTDTM LE (ADSL.TR03EDTM+72 hours)
ADAE	PREFL	Pre-treatment Flag	text	Y			If TRTEMFL ^="Y" and FUPFL^="Y" then PREFL="Y"
ADAE	FUPFL	Follow-up Flag	text	Y			if ASTDTM GT (ADSL.TR03EDTM+72 hours) then FUPFL="Y"
ADAE	ASTDY	Analysis Start Relative Day	integer				Date portion of ASTDTM - date portion of ADSL.TRT01SDTM + 1 day if date portion of ASTDTM is on or after date portion of TRT01SDTM, else date portion of ASTDTM - date portion of ADSL.TR01SDTM if date portion of ASTDTM precedes date portion of TR01SDTM
ADAE	EPOCH	Epoch	text	RUN-IN, FIRST TREATMENT, FIRST WASHOUT, SECOND		AE.EPOCH	

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
				TREATMENT, SECOND WASHOUT, THIRD TREATMENT, FOLLOW-UP			
ADAE	APHASE	Phase	text	RUN-IN, FIRST TREATMENT, FIRST WASHOUT, SECOND TREATMENT, SECOND WASHOUT, THIRD TREATMENT, FOLLOW-UP			If ASTDTM < ADSL.TR01SDTM then APHASE="RUN-IN", else ifADSL.TR01SDTM LE ASTDTM LE (ADSL.TR01EDTM+72 hours) then APHASE ="FIRST TREATMENT", else if (ADSL.TR01EDTM+72 hours) < ASTDTM< ADSL.TR02SDTM then APHASE="FIRST WASHOUT", etc.
ADAE	APERIOD	Period	integer	1, 2, 3			If TR01SDTM LE ASTDTM LE (TR01EDTM+72 hours) then APERIOD=1, else if TR02SDTM LE ASTDTM LE (TR02EDTM+72 hours) then APERIOD=2, else if TR03SDTM LE ASTDTM LE (TR03EDTM+72 hours) then APERIOD=3
ADAE	APERIODC	Period (C)	text	PERIOD 01, PERIOD 02, PERIOD 03			If APERIOD=1 then APERIODC="PERIOD 01", else if APERIOD=2 then APERIODC="PERIOD 02", else if APERIOD=3 then APERIODC="PERIOD 03"
ADAE	TR01SDTM	Datetime of First Exposure in Period 01	integer			ADSL.TR01SDTM	
ADAE	TR01EDTM	Datetime of Last Exposure in Period 01	integer			ADSL.TR01EDTM	
ADAE	TR02SDTM	Datetime of First Exposure in Period 02	integer			ADSL.TR02SDTM	
ADAE	TR02EDTM	Datetime of Last Exposure in Period 02	integer			ADSL.TR02EDTM	
ADAE	TR03SDTM	Datetime of First Exposure in Period 03	integer			ADSL.TR03SDTM	
ADAE	TR03EDTM	Datetime of Last Exposure in Period 03	integer			ADSL.TR03EDTM	

7.3 Sample ADaM Data - Example 4

Table 7.3.1 illustrates the adverse events analysis dataset (ADAE) defined in Section 7.2, Sample ADaM Variable Metadata - Example 4.

Key points to note in the example are:

1. The SDTM variable EPOCH was kept for traceability and to illustrate the differences between this variable and APHASE and APERIOD.
2. Treatment start and end datetimes for each period were kept and used to calculate APERIOD, TRTEMFL, and TREMxxFL variables. Another option is to use ADSL variables relating to period start and end datetimes (APxxSDTM and APxxEDTM). However, this latter option would not work when defining different periods for efficacy and for safety.
3. The producer of the dataset chose to populate APERIOD as an analysis period where the wash-out and follow-up periods were not populated for APERIOD. The same applied for the record-level actual treatment variable (TRTA), which was left missing for records not associated with a treatment. However, this is left up to the producer.
4. Row 5 indicates an adverse event that occurred in the follow-up epoch, post-treatment emergent, and not related to any analysis period or treatment.

5. Row 8 indicates an adverse event that occurred in the follow-up epoch, but within the third treatment phase and analysis period and associated with treatment A + B.

Table 7.3.1 Sample ADaM Data[a]

Row	USUBJID	TRTA	TRTAN	SAFFL	AEBODSYS	AEDECOD	ASTDTM	ASTDTF	ASTTMF	TRTEMFL	TREM01FL	TREM02FL	TREM03FL	PREFL	FUPFL	ASTDY	EPOCH	APHASE	APERIOD	APERIODC	TR01SDTM	TR01EDTM	TR02SDTM	TR02EDTM	TR03SDTM	TR03EDTM
1	101-001	A	1	Y	GASTROINTESTINAL DISORDERS	VOMITING	05MAY08:16:00:00		M	Y	Y					5	FIRST TREATMENT	FIRST TREATMENT	1	PERIOD 01	01MAY08:10:05:00	07MAY08:09:10:10	15MAY08:08:15:00	21MAY08:10:30:00	20MAY08:13:50:00	03JUN08:07:20:00
2	101-001	B	2	Y	INFECTIONS AND INFESTATIONS	PHARYNGITIS	16MAY08:06:42:00			Y		Y				16	SECOND TREATMENT	SECOND TREATMENT	2	PERIOD 02	01MAY08:10:05:00	07MAY08:09:10:00	15MAY08:08:15:00	21MAY08:10:30:00	29MAY08:13:50:00	03JUN08:07:20:00
3	101-001	A+B	3	Y	NERVOUS SYSTEM DISORDERS	HEADACHE	01JUN08:15:30:00			Y			Y			32	THIRD TREATMENT	THIRD TREATMENT	3	PERIOD 03	01MAY08:10:05:00	07MAY08:09:10:00	15MAY08:08:15:00	21MAY08:10:30:00	29MAY08:13:50:00	03JUN08:07:20:00
4	101-001	A+B	3	Y	NERVOUS SYSTEM DISORDERS	CONSTIPATION	02JUN08:07:15:00			Y			Y			33	THIRD TREATMENT	THIRD TREATMENT	3	PERIOD 03	01MAY08:10:05:00	07MAY08:09:10:00	15MAY08:08:15:00	21MAY08:10:30:00	29MAY08:13:50:00	03JUN08:07:20:00
5	101-001			Y	INFECTIONS AND INFESTATIONS	ORAL HERPES	07JUN08:08:00:00								Y	38	FOLLOW-UP	FOLLOW-UP			01MAY08:10:05:00	07MAY08:09:10:00	15MAY08:08:15:00	21MAY08:10:30:00	29MAY08:13:50:00	03JUN08:07:20:00
6	101-002			Y	VASCULAR DISORDERS	HYPOTENSION	25MAY08:13:20:00							Y		26	SECOND WASHOUT	SECOND WASHOUT			30APR08:12:05:00	06MAY08:08:32:00	14MAY08:11:55:00	20MAY08:08:10:00	26MAY08:15:40:00	01JUN08:09:13:00
7	101-002	A+B	3	Y	NERVOUS SYSTEM DISORDERS	HEADACHE	27MAY08:22:10:00			Y			Y			28	THIRD TREATMENT	THIRD TREATMENT	3	PERIOD 03	30APR08:12:05:00	06MAY08:08:32:00	14MAY08:11:55:00	20MAY08:08:10:00	26MAY08:15:40:00	01JUN08:09:13:00
8	101-002	A+B	3	Y	NERVOUS SYSTEM DISORDERS	HEADACHE	02JUN08:22:10:00			Y			Y			34	FOLLOW-UP	THIRD TREATMENT	3	PERIOD 03	30APR08:12:05:00	06MAY08:08:32:00	14MAY08:11:55:00	20MAY08:08:10:00	26MAY08:15:40:00	01JUN08:09:13:00

[a]Variables ending in DTM are numeric datetimes, here shown using SAS format datetime16. Other numeric datetime formats can be used, but care should be taken with newer formats which might not be understood by all statistical packages.

8 Example 5: MedDRA Secondary Path

In MedDRA, a collected term can be mapped along more than 1 path, as shown in Figure 5.1.

Whenever more than 1 path is possible, there is always a primary coding path plus 1 or more secondary paths. When a secondary path will be used for analysis, SDTMIG v3.3 allows for capture of both a primary and secondary system organ class (SOC), as described in the CDISC Notes column for these variables:

- **AEBODSYS**: "Dictionary derived. Body system or organ class used by the sponsor from the coding dictionary (e.g., MedDRA). When using a multi-axial dictionary such as MedDRA, this should contain the SOC used for the sponsor's analyses and summary tables which may not necessarily be the primary SOC."
- **AESOC**: "Dictionary-derived text description of the primary System Organ Class. Will be the same as AEBODSYS if the primary SOC was used for analysis."

As with other SDTM variables, these are typically copied from the SDTM to ADaM and used directly in the occurrence analysis.

This section describes different ways to handle multiple coding paths and provides an example for how to create a single analysis dataset with 2 different coding paths.

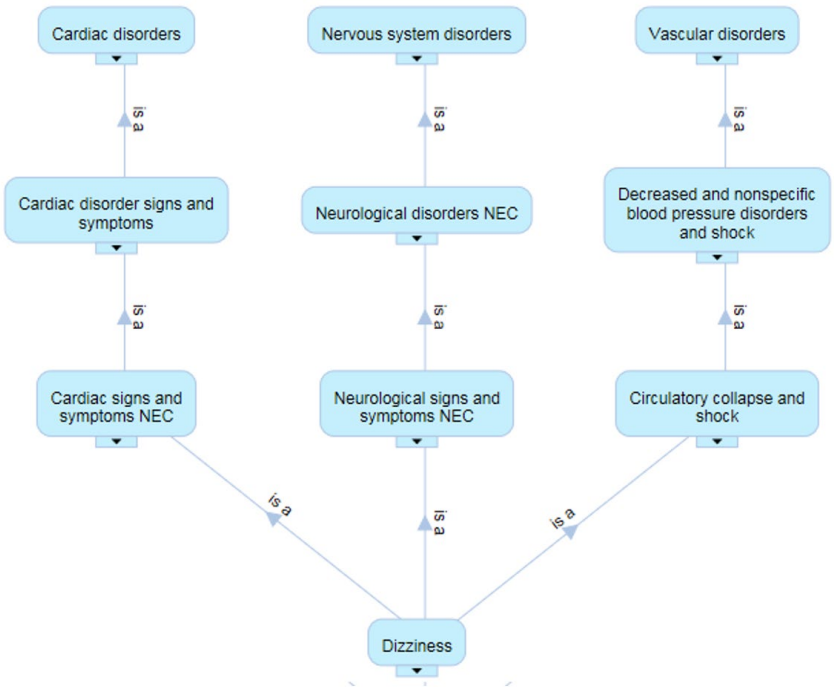
Typically, adverse event analysis includes only the primary coding path. However, some indications also perform analysis on the secondary path. For example, in a study of brain cancer, a headache might need to be analyzed according to a secondary path that attributes this to the cancer. This is why the SDTM has the option of including 2 different paths.

When a secondary path is used, often the analysis need is:

- One set of tables showing the primary path, such as in Section 4, Example 1: Analysis of Treatment-emergent Adverse Event, using the primary coding path for analysis.
- Another separate set of tables allowing for a secondary coding path to be chosen for analysis.

In addition, there may be an analysis need where both primary and secondary coding paths are summarized in the same table. The remainder of this section describes this analysis dataset.

Figure 8.1 Possible MedDRA Coding Paths for Term "Dizziness"



8.1 Analysis Display Example Layout - Example 5

The analysis need is to produce the following table:

Figure 8.1.1 Example Layout for Adverse Event Primary and Secondary Path Summary[a]

Table 14.2.8.3
Treatment Emergent AEs by Primary and Secondary SOCs, Preferred Term
(Population: Safety Subjects)

System Organ Class Preferred Term	Placebo (n=xxx) Primary SOC n (%)	Placebo (n=xxx) Primary + Secondary SOC n (%)	Drug X (n=xxx) Primary SOC n (%)	Drug X (n=xxx) Primary + Secondary SOC n (%)
CARDIAC DISORDERS	0 (x.x)	0 (x.x)	0 (x.x)	1 (x.x)
Dizziness[b]	0 (x.x)	0 (x.x)	0 (x.x)	1 (x.x)
ENDOCRINE DISORDERS	0 (x.x)	1 (x.x)	2 (x.x)	2 (x.x)
Autoimmune thyroiditis[a]	0 (x.x)	0 (x.x)	1 (x.x)	1 (x.x)
Thyroid atrophy[a]	0 (x.x)	1 (x.x)	1 (x.x)	1 (x.x)
NERVOUS SYSTEM DISORDERS	2 (x.x)	2 (x.x)	1 (x.x)	1 (x.x)
Dizziness[a]	2 (x.x)	2 (x.x)	1 (x.x)	1 (x.x)
VASCULAR DISORDERS	0 (x.x)	1 (x.x)	0 (x.x)	0 (x.x)
Dizziness[b]	0 (x.x)	1 (x.x)	0 (x.x)	0 (x.x)

SOC = System Organ Class
[a] Preferred term comes from primary system organ class path
[b] Preferred term comes from secondary system organ class path

[a]The style of the display of the results of an analysis will be determined by the producer. The example is intended to illustrate content not appearance.

8.2 Sample SDTM AE Data - Example 5

As described in the introduction to this section, the SDTM variables AEBODSYS and AESOC are both included in SDTM AE data, and each represents a different coding path used for analysis. In Table 8.2.1, note that AEBODSYS and AESOC are the same on some rows but different on others. When only a primary path is to be used, the values of AEBODSYS and AESOC are the same. When a secondary path is to be used, AEBODSYS and AESOC are different. For the purpose of this example, the adverse event "dizziness" was coded to different MedDRA coding paths in different subjects based on further information available.

Table 8.2.1 Sample SDTM AE Data for Selected Variables

Row	STUDYID	USUBJID	AESEQ	AEDECOD	AEBODSYS	AESTDTC	AESOC
1	XYZ	XYZ-1-001	1	Autoimmune thyroiditis	Endocrine disorders	2008-05-13	Endocrine disorders
2	XYZ	XYZ-1-001	2	Dizziness	Cardiac disorders	2008-06-13	Nervous system disorders
3	XYZ	XYZ-2-002	1	Dizziness	Vascular disorders	2008-09-13	Nervous system disorders
4	XYZ	XYZ-3-003	1	Thyroid atrophy	Endocrine disorders	2008-09-13	Endocrine disorders
5	XYZ	XYZ-4-004	1	Dizziness	Nervous system disorders	2008-09-09	Nervous system disorders

8.3 Sample ADaM Data - Example 5

In this example, a single analysis dataset ADAE, with SubClass ADVERSE EVENT was created with additional rows added for each coding path to facilitate the analysis shown in Section 8.1, Analysis Display Example Layout - Example 5.

Key points to note in the example are:

1. The analysis record flag variables can be used to differentiate between primary and secondary path records. In this example, ANL01FL is used to identify the primary coding path and ANL02FL is used to identify the secondary coding path.
2. Rows 2 and 3 represent a single adverse event of dizziness from the SDTM, coded to 2 system organ classes: nervous system disorders (primary SOC) and cardiac disorders (original coding). Both rows in ADAE have the same value of AESEQ.

3. Rows 4 and 5 represent a single adverse event of dizziness, coded to 2 SOCs: nervous system disorders (primary SOC) and vascular disorders (original coding). Both rows in ADAE have the same value of AESEQ.
4. Rows 6 and 7 use only 1 path, so no additional records are necessary.
5. AEBODSYS and AESOC are unchanged from the SDTM. The new variable ASOC is added as an analysis version of the body system to facilitate the analysis. ASOC is not a required name.
6. The purpose of this example is not to state how MedDRA secondary paths are to be handled, only to provide an example.

Table 8.3.1 Sample ADaM ADAE Data for Selected Variables

Row	USUBJID	TRTA	AESEQ	AEDECOD	AEBODSYS	AESOC	ASOC	ANL01FL	ANL02FL
1	XYZ-1-001	Drug X	1	Autoimmune thyroiditis	Endocrine disorders	Endocrine disorders	Endocrine disorders	Y	
2	XYZ-1-001	Drug X	2	Dizziness	Cardiac disorders	Nervous system disorders	Cardiac disorders		Y
3	XYZ-1-001	Drug X	2	Dizziness	Cardiac disorders	Nervous system disorders	Nervous system disorders	Y	
4	XYZ-2-002	Placebo	1	Dizziness	Vascular disorders	Nervous system disorders	Vascular disorders		Y
5	XYZ-2-002	Placebo	1	Dizziness	Vascular disorders	Nervous system disorders	Nervous system disorders	Y	
6	XYZ-3-003	Drug X	1	Thyroid atrophy	Endocrine disorders	Endocrine disorders	Endocrine disorders	Y	
7	XYZ-4-004	Placebo	1	Dizziness	Nervous system disorders	Nervous system disorders	Nervous system disorders	Y	

9 Example 6: Analysis of Adverse Events that Change over Time

The structure of the Adverse Events (AE) domain is 1 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 1 record that covers an adverse event from start to finish. Alternatively, if there is a need to evaluate adverse events at a more granular level, a sponsor may submit a new record when severity, causality, or seriousness changes or worsens. One option is to submit 1 record per adverse event per subject and then to submit changes over time in severity, causality, or seriousness in a separate dataset based on the Findings About Events and Interventions (FA) model. Sponsors may represent findings data in a single FA dataset or split the FA domain into separate datasets, following the guidance in the SDTMIG. In this example, 1 record that covers an adverse event from start to finish is recorded in AE; changes in severity are captured in a split FAAE dataset. In this example, the statistical analysis plan (SAP) and protocol require reporting the highest severity of treatment emergent adverse events, so the 2 domains (AE and FAAE) need to be merged together to create 1 record for each change in severity within event.

9.1 Analysis Display Example Layouts - Example 6

This example displays a simple summary of all treatment-emergent adverse events by maximum severity. The example is based on a 2-treatment parallel design study. The display summarizes (1) the number of subjects in each treatment group in which the adverse event occurred, counted once under maximum severity and once for the total; and (2) the rate of occurrence in each treatment group overall and by maximum severity.

There are multiple ways to present this data. Two common approaches are shown in Figures 9.1.1 and 9.1.2.

Figure 9.1.1 Example Layout for Adverse Events Summary by Maximum Severity - Option 1

Summary of Treatment-Emergent Adverse Events by Maximum Severity Safety Population			
System Organ Class Preferred Term	Maximum Severity	Treatment 1 (N=XXX) n(%)	Treatment 2 (N=XXX) n(%)
Number of subjects reporting treatment-emergent adverse events	Mild	xx (xx.x%)	xx (xx.x%)
	Moderate	xx (xx.x%)	xx (xx.x%)
	Severe	xx (xx.x%)	xx (xx.x%)
	Total	xx (xx.x%)	xx (xx.x%)
[System Organ Class Class #1]	Mild	xx (xx.x%)	xx (xx.x%)
	Moderate	xx (xx.x%)	xx (xx.x%)
	Severe	xx (xx.x%)	xx (xx.x%)
	Total	xx (xx.x%)	xx (xx.x%)
[Preferred Term #1]	Mild	xx (xx.x%)	xx (xx.x%)
	Moderate	xx (xx.x%)	xx (xx.x%)
	Severe	xx (xx.x%)	xx (xx.x%)
	Total	xx (xx.x%)	xx (xx.x%)

Figure 9.1.2 Example Layout for Adverse Events Summary by Maximum Severity - Option 2

Summary of Treatment-Emergent Adverse Events by Maximum Severity Safety Population								
	Treatment 1 (N=XXX)				Treatment 2 (N=XXX)			
System Organ Class Preferred Term	Mild n(%)	Moderate n(%)	Severe n(%)	Total n(%)	Mild n(%)	Moderate n(%)	Severe n(%)	Total n(%)
Number of subjects reporting treatment-emergent adverse events	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
[System Organ Class Class #1]	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
[Preferred Term #1]	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)

9.2 Sample SDTM Domains - Example 6

In the AE dataset, for a given AE record, 1 record that covers the adverse event from start to finish is recorded in the AE domain and the worst case reported overall for a given incident is recorded in AESEV. Subject ABC-001 had 3 distinct events, hypotension from 25MAR2017 until 25APR2017 with highest severity of Severe; dizziness from 27MAR2017 until 29MAR2017 with highest severity of Mild; and hypotension again from 15MAY2017 until 21MAY2017 with highest severity of Moderate.

Table 9.2.1 Sample AE Data

Row	USUBJID	AESEQ	AEGRPID	AETERM	AEDECOD	AEBODSYS	AESEV	AESTDTC	AEENDTC	EPOCH
1	ABC-001	101	AE1	LOW BLOOD PRESSURE	Hypotension	Vascular disorders	SEVERE	2017-03-25	2017-04-25	SCREENING
2	ABC-001	102	AE2	DIZZY	Dizziness	Nervous system disorders	MILD	2017-03-27	2017-03-29	TREATMENT
3	ABC-001	103	AE3	LOW BLOOD PRESSURE	Hypotension	Vascular disorders	MODERATE	2017-05-15	2017-05-21	TREATMENT
4	ABC-002	101	AE1	HEART FAILURE	Cardiac failure	Cardiac disorders	SEVERE	2017-03-21	2017-03-25	TREATMENT

If an adverse event changes in severity over time during an incident, then all records capturing this will be mapped to the FAAE domain. The severity will be recorded in FAORRES/FASTRESC, start date in FADTC and end date in FAENDTC for each episode.

Table 9.2.2 Sample FAAE Data

Row	USUBJID	FASEQ	FAOBJ	FAGRPID	FATESTCD	FATEST	FAORRES	FASTRESC	FADTC	FAENDTC	EPOCH
1	ABC-001	201	Hypotension	AE1	SEV	Severity	MILD	MILD	2017-03-25	2017-03-31	SCREENING
2	ABC-001	202	Hypotension	AE1	SEV	Severity	SEVERE	SEVERE	2017-04-01	2017-04-10	TREATMENT
3	ABC-001	203	Hypotension	AE1	SEV	Severity	MODERATE	MODERATE	2017-04-11	2017-04-25	TREATMENT

The AE domain is related to FAAE by USUBJID, where the value of FAGRPID = AEGRPID. The value in FAOBJ matches the value in AEDECOD, but multiple incidents can occur within a subject in AE as seen with the events of hypotension for subject ABC-001 in the above example. The FAGRPID = AE1 in FAAE links the changes in severity back to the first incident of hypotension for subject ABC-001, having AEGRPID = AE1.

9.3 Sample ADaM Variable Metadata - Example 6

This example describes an adverse events ADaM dataset named ADAEFA with SubClass ADVERSE EVENT. ADAEFA is not a required dataset name. ADAEFA is used to demonstrate that the input data is from both AE and FA.

Table 9.3.1 Sample ADaM Variable Metadata for Selected Variables

Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
USUBJID	Unique Subject Identifier	text			AE.USUBJID	
AESEQ	Sequence Number	integer			AE.AESEQ	
FASEQ	Sequence Number	integer				FA.FASEQ on records where AE and FAAE are joined by Subject, Group ID (AEGRPID/FAGRPID), and Topic (AEDECOD/FAOBJ)
ACAT1	Analysis Category 1	text	ADVERSE EVENTS SUMMARY, SEVERITY DETAILS			Create a record for each record in AE with ACAT1="ADVERSE EVENTS SUMMARY" If records exist in FAAE where AEGRPID is equal to FAGRPID within subject, join together by USUBJID, AEGRPID with FAGRPID, and AEDECOD with FAOBJ and for each record set ACAT1="SEVERITY DETAILS"
AEGRPID	Group ID	text			AE.AEGRPID	
AETERM	Reported Term for the Adverse Event	text			AE.AETERM	
AEDECOD	Dictionary-Derived Term	text		MedDRA	AE.AEDECOD	
AEBODSYS	Body System or Organ Class	text		MedDRA	AE.AEBODSYS	
AESTDTC	Start Date/Time of Adverse Event	datetime		ISO 8601	AE.AESTDTC	
AEENDTC	End Date/Time of Adverse Event	datetime		ISO 8601	AE.AEENDTC	
FADTC	Date/Time of Clinical Finding	datetime		ISO 8601		FA.FADTC on records where AE and FAAE are joined by Subject, Group ID (AEGRPID/FAGRPID), and Topic (AEDECOD/FAOBJ)
FAENDTC	End Date/Time of Clinical Finding	datetime		ISO 8601		FA.FAENDTC on records where AE and FAAE are joined by Subject, Group ID (AEGRPID/FAGRPID), and Topic (AEDECOD/FAOBJ)
ASTDT	Analysis Start Date	integer				<sponsor defined> Example derivation: Populate with numeric version of FADTC if ACAT1="SEVERITY DETAILS". Populate with the numeric version of AESTDTC if ACAT1="ADVERSE EVENTS SUMMARY".
AENDT	Analysis End Date	integer				<sponsor defined> Example derivation: Populate with numeric version of FAENDTC if ACAT1="SEVERITY DETAILS". Populate with the numeric version of AEENDTC if ACAT1= "ADVERSE EVENTS SUMMARY".
AESEV	Severity/Intensity	text	MILD, MODERATE, SEVERE		AE.AESEV	
SEV	Severity/Intensity	text	MILD, MODERATE, SEVERE			FAORRES where FATESTCD="SEV" on records where FAGRPID is equal to AEGRPID within a subject
ASEV	Analysis Severity/Intensity	text	Mild, Moderate, Severe			Populate with SEV if ACAT1="SEVERITY DETAILS". Populate with AESEV if ACAT1= "ADVERSE EVENTS SUMMARY".

Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ASEVN	Analysis Severity/Intensity (N)	text	1=Mild, 2=Moderate, 3=Severe			If ASEV="Mild" then ASEVN=1 Else if ASEV="Moderate" then ASEVN=2 Else if ASEV="Severe" then ASEVN=3
TRTSDT	Date of First Exposure to Treatment	integer			ADSL.TRTSDT	
TRTEMFL	Treatment Emergent Analysis Flag	text	Y=Yes			<sponsor defined> Example derivation: Assume TRTEMFL="Y" if ADSL.TRTSDT is not missing and ASTDT is not less than ADSL.TRTSDT.
AOCCIFL	1st Max Sev./Int. Occurrence Flag	text	Y=Yes			Flag the maximum severity of a TEAE per subject. Subset records to treatment-emergent adverse events. Sort by subject, descending analysis severity/intensity (N), and sequence number, and flag the first record per subject.
AOCCPIFL	1st Max Sev./Int. Occur Within PT Flag	text	Y=Yes			Flag the maximum severity of a TEAE per preferred term per subject. Subset records to treatment-emergent adverse events. Sort by subject, preferred term, descending analysis severity/intensity (N), and sequence number, and flag the first record per preferred term.

9.4 Sample ADaM Data - Example 6

Table 9.4.1 is an illustration of the adverse event analysis dataset (ADAEFA) defined in Section 9.3, Sample ADaM Variable Metadata - Example 6.

Key points to note in the example are:

1. This is a simple example to illustrate the creation of ADAEFA from merging 2 SDTM domains (AE and FAAE). It does not include all variables that are needed for analysis, such as treatment (TRTA/TRT01A) or covariates (AGEGR1, SEX, RACE).
2. The ACAT1 variable indicates the type of record: an overall AE record or a severity detail record.
3. For subject ABC-001, the event of hypotension increased from mild (row 2) to severe (row 3) after the start of treatment. If only data from the AE domain were used (row 1), this severe event would not have been considered treatment-emergent. Utilizing the information in FAAE, this event is now correctly identified as a severe treatment-emergent adverse event.
4. Variables such as AESEQ, FASEQ, and AEGRPID are kept to enhance traceability between ADAEFA and AE and FAAE. Both AESEQ and FASEQ are populated on all records where the changes in severity over the time frame of a single episode have been provided. On these records the value of AEGRPID is equal to FAGRPID and AEDECOD is equal to FAOBJ, so FAGRPID and FAOBJ are not needed in the dataset.
5. This idea could be extended if additional and multiple qualifiers (e.g., causality, seriousness, toxicity grading) were captured in FAAE. If multiple qualifiers exist in FAAE, a suitable 2-step approach is to first copy records from FAAE, then create additional derived rows based on the overlap of timing variables between qualifiers that change over time. For records copied from FAAE for a subject having multiple qualifiers for a given event, ACAT1 would be used to denote the type of record (CAUSALITY DETAILS, SERIOUSNESS DETAILS, TOXICITY GRADE DETAILS) and for traceability, not analysis. Additional derived records (ACAT1 = "DERIVED" or "CAUSALITY + SERIOUSNESS") would be added considering the changes in time of all of the qualifiers in FAAE; occurrence flags (AOCCzzFL) would be used to flag the correct record to use for a given analysis.

Table 9.4.1 Sample Adverse Events Data[a]

Row	USUBJID	AESEQ	FASEQ	ACAT1	AEGRPID	AEDECOD	AEBODSYS	AESEV	SEV	AESTDTC	AEENDTC	FADTC	FAENDTC	ASTDT	AENDT	ASEV	ASEVN	TRTSDT	TRTEMFL	AOCCIFL	AOCCPIFL
1	ABC-001	101		ADVERSE EVENTS SUMMARY	AE1	Hypotension	Vascular disorders	SEVERE		2017-03-25	2017-04-25			25MAR2017	25APR2017	Severe	3	26MAR2017			
2	ABC-001	101	201	SEVERITY DETAILS	AE1	Hypotension	Vascular disorders	SEVERE	MILD	2017-03-25	2017-04-25	2017-03-25	2017-03-31	25MAR2017	31MAR2017	Mild	1	26MAR2017			
3	ABC-001	101	202	SEVERITY DETAILS	AE1	Hypotension	Vascular disorders	SEVERE	SEVERE	2017-03-25	2017-04-25	2017-04-01	2017-04-10	01APR2017	10APR2017	Severe	3	26MAR2017	Y	Y	Y
4	ABC-001	101	203	SEVERITY DETAILS	AE1	Hypotension	Vascular disorders	SEVERE	MODERATE	2017-03-25	2017-04-25	2017-04-11	2017-04-25	11APR2017	25APR2017	Moderate	2	26MAR2017	Y		
5	ABC-001	102		ADVERSE EVENTS SUMMARY	AE2	Dizziness	Nervous system disorders	MILD		2017-03-27	2017-03-29			27MAR2017	29MAR2017	Mild	1	26MAR2017	Y		Y
6	ABC-001	103		ADVERSE EVENTS SUMMARY	AE3	Hypotension	Vascular disorders	MODERATE		2017-05-15	2017-05-21			15MAY2017	21MAY2017	Moderate	2	26MAR2017	Y		
7	ABC-002	101		ADVERSE EVENTS SUMMARY	AE1	Cardiac failure	Cardiac disorders	SEVERE		2017-03-21	2017-03-25			21MAR2017	25MAR2017	Severe	3	01MAR2017	Y	Y	Y

[a]Variables ending in suffix DTC are character date/time fields in ISO 8601 format. Variables ending in DT are numeric dates, here shown using SAS date format date9. Other numeric date formats can be used, but care should be taken with newer date formats which might not be understood by all statistical packages.

10 Example 7: Analysis of Adverse Events from Multiple Input Domains

Hypoglycemia is a clinical event that involves an abnormally diminished blood glucose concentration. It can produce a variety of symptoms and effects, but the principal problems arise from an inadequate supply of glucose to the brain, resulting in impairment of function. The occurrence of hypoglycemia is usually identified by measuring a person's blood or plasma glucose concentration and assessing the symptoms typical of hypoglycemia, which include but are not limited to palpitations, tremor, hunger, sweating, dizziness, behavioral changes, difficulty thinking, and/or frank confusion.

In the diabetes study used for this example, hypoglycemia events and symptoms are captured on a CRF separately from adverse events and are mapped to the SDTM Clinical Events (CE) domain. In this example, the protocol and SAP instruct that the hypoglycemic events themselves and signs and symptoms of hypoglycemia are to be summarized on some tables with data from the Adverse Event (AE) domain. The focus of this example is to show how to construct an OCCDS dataset by stacking multiple SDTM domains to meet that analysis need. This example study is a randomized parallel design study consisting of 2 treatment groups (i.e., standard of care, study drug versus standard of care alone).

10.1 Sample SDTM Domains - Example 7

The data in this example are adverse events for subject ABC-001. Not all variables collected have been included in the example.

Table 10.1.1 Sample AE Data

Row	STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AEDECOD	AEBODSYS	AESER	AEACN	AEOUT	AETOXGR	AESTDTC	AEENDTC
1	ABC	AE	ABC-001	2	HEADACHE	Headache	Nervous system disorders	N	DOSE NOT CHANGED	RECOVERED/RESOLVED	2	2017-10-21	2017-10-23
2	ABC	AE	ABC-001	1	STREP THROAT	Pharyngitis streptococcal	Infections and infestations	N	DOSE NOT CHANGED	RECOVERED/RESOLVED	3	2017-11-05	2017-11-15
3	ABC	AE	ABC-001	3	BODY ACHES	Pain	General disorders and administration site conditions	N	DOSE NOT CHANGED	RECOVERED/RESOLVED	2	2017-11-17	2017-11-22
4	ABC	AE	ABC-001	4	SHORTNESS OF BREATH	Dyspnoea	Respiratory, thoracic and mediastinal disorders	Y	DRUG INTERRUPTED	RECOVERED/RESOLVED	3	2017-11-25	2017-11-30

The following data are for 2 clinical events for subject ABC-001, who had 2 episodes of hypoglycemia: one occurring on 25OCT2017 at 7:30 AM and the other occurring on 02NOV2017 at 5:35 PM. The page (CRF) would only be triggered if the subject had a hypoglycemic event (see rows 1 and 9). The type of hypoglycemic event—severe, documented symptomatic, asymptomatic, probable, or pseudo—is captured in CETERM. Seven symptoms (sweating, tremors/trembling, dizziness, cognitive impairment, loss of consciousness, convulsion/seizure, coma) are pre-specified on the CRF with an occurrence question. If answered "Yes" (OCCUR="Y"), then additional information is collected about the sign or symptom. CEGRPID is used to group the episode of hypoglycemia with the symptoms. The first episode had associated sweating, tremors, and dizziness; the second had dizziness, cognitive impairment, and loss of consciousness. Qualifiers such as CESER, CEACN, CEREL, CEOUT, CETOXGR, CESTDTC, and CEENDTC are populated only on the hypoglycemic episodes and the symptoms that occurred during those episodes.

Table 10.1.2 Sample CE Data

Row	STUDYID	DOMAIN	USUBJID	CEGRPID	CESEQ	CETERM	CEDECOD	CECAT	CEPRESP	CEOCCUR	CEBODSYS	CESER	CEACN	CEREL	CEOUT	CETOXGR	CESTDTC	CEENDTC
1	ABC	CE	ABC-001	HYPO1	1	PROBABLE SYMPTOMATIC HYPOGLYCEMIC EVENT	Hypoglycaemia	HYPO EVENTS			Hepatobiliary disorders	N	DOSE NOT CHANGED	RELATED	RECOVERED/RESOLVED	1	2017-10-25T07:30	2017-10-25T08:10
2	ABC	CE	ABC-001	HYPO1	2	SWEATING	Hyperhidrosis	HYPO EVENTS	Y	Y	General disorders and administration site conditions	N	DOSE NOT CHANGED	RELATED	RECOVERED/RESOLVED	1	2017-10-25T07:20	2017-10-25T08:05
3	ABC	CE	ABC-001	HYPO1	3	TREMORS/TREMBLING	Tremor	HYPO EVENTS	Y	Y	Nervous system disorders	N	DOSE NOT CHANGED	RELATED	RECOVERED/RESOLVED	1	2017-10-25T07:10	2017-10-25T08:05
4	ABC	CE	ABC-001	HYPO1	4	DIZZINESS	Dizziness	HYPO EVENTS	Y	Y	Nervous system disorders	N	DOSE NOT CHANGED	RELATED	RECOVERED/RESOLVED	2	2017-10-25T07:05	2017-10-25T08:05

Row	STUDYID	DOMAIN	USUBJID	CEGRPID	CESEQ	CETERM	CEDECOD	CECAT	CEPRESP	CEOCCUR	CEBODSYS	CESER	CEACN	CEREL	CEOUT	CETOXGR	CESTDTC	CEENDTC
5	ABC	CE	ABC-001	HYPO1	5	COGNITIVE IMPAIRMENT	Cognitive Disorder	HYPO EVENTS	Y	N	Psychiatric disorders							
6	ABC	CE	ABC-001	HYPO1	6	LOSS OF CONSCIOUSNESS	Loss of Consciousness	HYPO EVENTS	Y	N	Nervous system disorders							
7	ABC	CE	ABC-001	HYPO1	7	CONVULSIONS/SEIZURES	Convulsion	HYPO EVENTS	Y	N	Congenital, familial and genetic disorders							
8	ABC	CE	ABC-001	HYPO1	8	COMA	Coma	HYPO EVENTS	Y	N	Nervous system disorders							
9	ABC	CE	ABC-001	HYPO2	9	DOCUMENTED SYMPTOMATIC HYPOGLYCEMIC EVENT	Hypoglycaemia	HYPO EVENTS			Hepatobiliary disorders	Y	DRUG INTERRUPTED	POSSIBLY RELATED	RECOVERED/RESOLVED	3	2017-11-02T17:35	2017-11-02T18:20
10	ABC	CE	ABC-001	HYPO2	10	SWEATING	Hyperhidrosis	HYPO EVENTS	Y	N	General disorders and administration site conditions							
11	ABC	CE	ABC-001	HYPO2	11	TREMORS/TREMBLING	Tremor	HYPO EVENTS	Y	N	Nervous system disorders							
12	ABC	CE	ABC-001	HYPO2	12	DIZZINESS	Dizziness	HYPO EVENTS	Y	Y	Nervous system disorders	N	DRUG INTERRUPTED	POSSIBLY RELATED	RECOVERED/RESOLVED	2	2017-11-02T17:25	2017-11-02T18:10
13	ABC	CE	ABC-001	HYPO2	13	COGNITIVE IMPAIRMENT	Cognitive Disorder	HYPO EVENTS	Y	Y	Psychiatric disorders	Y	DRUG INTERRUPTED	POSSIBLY RELATED	RECOVERED/RESOLVED	3	2017-11-02T17:40	2017-11-02T18:05
14	ABC	CE	ABC-001	HYPO2	14	LOSS OF CONSCIOUSNESS	Loss of Consciousness	HYPO EVENTS	Y	Y	Nervous system disorders	Y	DRUG INTERRUPTED	POSSIBLY RELATED	RECOVERED/RESOLVED	3	2017-11-02T17:45	2017-11-02T18:01
15	ABC	CE	ABC-001	HYPO2	15	CONVULSIONS/SEIZURES	Convulsion	HYPO EVENTS	Y	N	Congenital, familial and genetic disorders							
16	ABC	CE	ABC-001	HYPO2	16	COMA	Coma	HYPO EVENTS	Y	N	Nervous system disorders							

10.2 Sample ADaM Variable Metadata - Example 7

This example describes an Adverse Events (AE) ADaM dataset named ADAECE. ADAECE is not a required dataset name. ADAECE is used to demonstrate that the input data is from both AE and Clinical Events (CE). A name other than ADAE is suggested because the analysis dataset contains events in addition to those captured in the AE domain, and therefore this dataset is not of SubClass ADVERSE EVENT.

Table 10.2.1 Sample ADaM Variable Metadata

Variable Name	Variable Label	Variable Type	Codelists	Controlled Terms	Source	Derivation
STUDYID	Study Identifier	text			AE.STUDYID CE.STUDYID	
USUBJID	Unique Subject Identifier	text			AE.USUBJID CE.USUBJID	
SAFFL	Safety Population Flag	text	Y=Yes, N=No		ADSL.SAFFL	
TRT01A	Actual Treatment	text	SOC+SD=Standard of Care + Study Drug, SOC=Standard of Care		ADSL.TRT01A	
TRTSDT	Date of First Exposure to Treatment	integer			ADSL.TRTSDT	

Variable Name	Variable Label	Variable Type	Codelists	Controlled Terms	Source	Derivation
TRTSDTM	Datetime of First Exposure to Treatment	integer			ADSL.TRTSDTM	
SRCDOM	Source Data	text				Set to "AE" if record is from AE dataset Set to "CE" if record is from CE dataset and CEPRESP is missing or (CEPRESP="Y" and CEOCCUR="Y")
SRCSEQ	Source Sequence Number	integer				Set to AE.AESEQ if record is from AE dataset Set to CE.CESEQ if record is from CE dataset and CEPRESP is missing or (CEPRESP="Y" and CEOCCUR="Y")
UTERM	Reported Term	text			AE.AETERM CE.CETERM	
UDECOD	Dictionary-Derived Term	text		MedDRA	AE.AEDECOD CE.CEDECOD	
ACAT1	Analysis Category 1	text	ADVERSE EVENTS, HYPO EVENTS			If record is from CE then ACAT1=CE.CECAT (HYPO EVENTS) If record is from AE then ACAT1="ADVERSE EVENTS"
CEPRESP	Clinical Event Pre-Specified	text	Y=Yes		CE.CEPRESP	
UBODSYS	Body System or Organ Class	text		MedDRA	AE.AEBODSYS CE.CEBODSYS	
USER	Serious Event	text	Y=Yes, N=No		AE.AESER CE.CESER	
UREL	Causality	text	DEFINITELY RELATED, POSSIBLY RELATED, PROBABLY RELATED, RELATED, NOT RELATED, UNLIKELY RELATED		AE.AEREL CE.CEREL	
ARELGR1	Pooled Causality Group 1	text	Related, Not Related			If UREL in ("DEFINITELY RELATED" "POSSIBLY RELATED" "PROBABLY RELATED" "RELATED") then ARELGR1="Related". Else if UREL in ("NOT RELATED" "UNLIKELY RELATED") then ARELGR1="Not Related".
UTOXGR	Toxicity Grade	text	1, 2, 3, 4, 5		AE.AETOXGR CE.CETOXGR	
UACN	Action Taken with Study Treatment	text	DOSE NOT CHANGED, DOSE REDUCED, DRUG INTERRUPTED, DRUG WITHDRAWN, NOT APPLICABLE		AE.AEACN CE.CEACN	
UOUT	Outcome	text	FATAL, NOT RECOVERED/NOT RESOLVED, RECOVERED/RESOLVED		AE.AEOUT CE.CEOUT	
USTDTC	Start Date/Time of Event	datetime		ISO 8601	AE.AESTDTC CE.CESTDTC	
UENDTC	End Date/Time of Event	datetime		ISO 8601	AE.AEENDTC CE.CEENDTC	
ASTDT	Analysis Start Date	integer				<Producer will insert derivation here> For example: Convert date part of AESTDTC/CESTDTC to numeric date variable OR Date part of AESTDTC/CESTDTC. If full date is present, convert to numeric. If day is missing but year and month correspond with treatment start year and month, then set day to the start day of treatment. Otherwise assume the first of the month. If day and month are missing but year corresponds with treatment start year, then set month and day to treatment start month and day. Otherwise assume January 1. If start date is completely missing do not impute.
ASTDTM	Analysis Start Datetime	integer				<Producer will insert derivation here> For example: If time is provided, then convert AESTDTC/CESTDTC to a numeric datetime variable
TRTEMFL	Treatment Emergent Analysis Flag	text	Y=Yes			<Producer will insert derivation here> For example: Assume TRTEMFL="Y", unless proven that event is not treatment emergent. If both the start datetime of the adverse event and treatment are present and populated, and start datetime of adverse event is prior to start datetime of treatment (MISSING<ASTDTM<TRTSDTM), then set TRTEMFL to

Variable Name	Variable Label	Variable Type	Codelists	Controlled Terms	Source	Derivation
						NULL. If either the start datetime of the adverse event or the start datetime of treatment is missing, and both the start date of the adverse event and treatment are present and and populated, and start date of adverse event is prior to start date of treatment (MISSING<ASTDT<TRTSDT), then set TRTEMFL to NULL. If start date of adverse event is missing, but end datetime or end date is present and prior to start datetime or start date of treatment, then set TRTEMFL to NULL.

10.3 Sample ADaM Data - Example 7

Table 10.3.1 is an illustration of the adverse events analysis dataset (ADAECE) defined in Section 10.2, Sample ADaM Variable Metadata - Example 7.

Key points to note in the example are:

1. Only clinical events (CE) that occur for subject ABC-001 and are considered to be candidates for adverse events are included in ADAECE. For example, the actual hypoglycemic episodes (Table 10.1.2, rows 1 and 9) and corresponding symptoms (rows 2-4 and 12-14, respectively) are copied into ADAECE in Table 10.3.1 as rows 1-4 and 5-8. Note that symptoms that did not occur (OCCUR = "N"), such as cognitive impairment on row 5, were not copied into ADAECE.
2. Because the study is a parallel design, the subject-level treatment variable TRT01A is kept from ADSL to be used for the analysis rather than the creation of row-level treatment variable TRTA, as is compliant with ADAMIG Version 1.2.
3. Because multiple SDTM domains are stacked together, SRCSEQ and SRCDOM are used to point to the domain and record where the data came from. See Section 3.2.2, Identifier Variables, for a discussion of identifier variables and traceability in OCCDS.
4. As seen in preceding examples, many variables are copied directly into the OCCDS structure from SDTM domains without modification. With stacking, the U* variable allows users to preserve the copy feature yet stack the same type data into the same column. U* can only be used for a direct copy (the U* indicates “unmodified”). Any changes in type, casing, or harmonization requires the creation of A* variables, as seen with ARELGR1. See Section 3.2.11, User-specified Variable Naming Conventions, for a discussion about use of the A* and U* prefixes.
5. ACAT1 is populated to indicate type of data. CEGRPID could also be included to link symptoms together, but was not required for analysis.
6. Because this is just an example, not all variables needed for analysis have been included.

Table 10.3.1 Sample ADaM Variable Metadata[a]

Row	STUDYID	USUBJID	SAFFL	TRT01A	TRTSDT	TRTSDTM	SRCDOM	SRCSEQ	UTERM	UDECOD	ACAT1	CEPRESP	UBODSYS	USER	UREL	ARELGR1	UTOXGR	UACN	UOUT	USTDTC	UENDTC	ASTDT	ASTDTM	TRTEMFL
1	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	CE	1	PROBABLE SYMPTOMATIC HYPOGLYCEMIC EVENT	Hypoglycaemia	HYPO EVENTS		Hepatobiliary disorders	N	RELATED	Related	1	DOSE NOT CHANGED	RECOVERED/RESOLVED	2017-10-25T07:30	2017-10-25T08:10	25OCT2017	25OCT17:07:30:00	Y
2	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	CE	2	SWEATING	Hyperhidrosis	HYPO EVENTS	Y	General disorders and administration site conditions	N	RELATED	Related	1	DOSE NOT CHANGED	RECOVERED/RESOLVED	2017-10-25T07:20	2017-10-25T08:05	25OCT2017	25OCT17:07:20:00	Y
3	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	CE	3	TREMORS/TREMBLING	Tremor	HYPO EVENTS	Y	Nervous system disorders	N	RELATED	Related	1	DOSE NOT CHANGED	RECOVERED/RESOLVED	2017-10-25T07:10	2017-10-25T08:05	25OCT2017	25OCT17:07:10:00	Y
4	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	CE	4	DIZZINESS	Dizziness	HYPO EVENTS	Y	Nervous system disorders	N	RELATED	Related	2	DOSE NOT CHANGED	RECOVERED/RESOLVED	2017-10-25T07:05	2017-10-25T08:05	25OCT2017	25OCT17:07:05:00	Y
5	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	CE	9	DOCUMENTED SYMPTOMATIC HYPOGLYCEMIC EVENT	Hypoglycaemia	HYPO EVENTS		Hepatobiliary disorders	Y	POSSIBLY RELATED	Related	3	DRUG INTERRUPTED	RECOVERED/RESOLVED	2017-11-02T17:35	2017-11-02T18:20	02NOV2017	02NOV17:17:35:00	Y
6	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	CE	12	DIZZINESS	Dizziness	HYPO EVENTS	Y	Nervous system disorders	N	POSSIBLY RELATED	Related	2	DRUG INTERRUPTED	RECOVERED/RESOLVED	2017-11-02T17:25	2017-11-02T18:10	02NOV2017	02NOV17:17:25:00	Y
7	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	CE	13	COGNITIVE IMPAIRMENT	Cognitive Disorder	HYPO EVENTS	Y	Psychiatric disorders	Y	POSSIBLY RELATED	Related	3	DRUG INTERRUPTED	RECOVERED/RESOLVED	2017-11-02T17:40	2017-11-02T18:05	02NOV2017	02NOV17:17:40:00	Y
8	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	CE	14	LOSS OF CONSCIOUSNESS	Loss of Consciousness	HYPO EVENTS	Y	Nervous system disorders	Y	POSSIBLY RELATED	Related	3	DRUG INTERRUPTED	RECOVERED/RESOLVED	2017-11-02T17:45	2017-11-02T18:01	02NOV2017	02NOV17:17:45:00	Y
9	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	AE	1	HEADACHE	Headache	ADVERSE EVENTS		Nervous system disorders	N	NOT RELATED	Not Related	2	DOSE NOT CHANGED	RECOVERED/RESOLVED	2017-10-21	2017-10-23	21OCT2017		

Row	STUDYID	USUBJID	SAFFL	TRT01A	TRTSDT	TRTSDTM	SRCDOM	SRCSEQ	UTERM	UDECOD	ACAT1	CEPRESP	UBODSYS	USER	UREL	ARELGR1	UTOXGR	UACN	UOUT	USTDTC	UENDTC	ASTDT	ASTDTM	TRTEMFL
10	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	AE	2	STREP THROAT	Pharyngitis streptococcal	ADVERSE EVENTS		Infections and infestations	N	NOT RELATED	Not Related	3	DOSE NOT CHANGED	RECOVERED/RESOLVED	2017-11-05	2017-11-15	05NOV2017		Y
11	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	AE	3	BODY ACHES	Pain	ADVERSE EVENTS		General disorders and administration site conditions	N	NOT RELATED	Not Related	2	DOSE NOT CHANGED	RECOVERED/RESOLVED	2017-11-17	2017-11-22	17NOV2017		Y
12	ABC	ABC-001	Y	SOC + SD	24OCT2017	24OCT17:10:05:00	AE	4	SHORTNESS OF BREATH	Dyspnoea	ADVERSE EVENTS		Respiratory, thoracic and mediastinal disorders	Y	NOT RELATED	Not Related	3	DRUG INTERRUPTED	RECOVERED/RESOLVED	2017-11-25	2017-11-30	25NOV2017		Y

[a]Variables ending in suffix DTC are character date/time fields in ISO 8601 format. Variables ending in DT are numeric dates, here shown using SAS date format date9. Variables ending in DTM are numeric datetimes, here shown using SAS format datetime16. Other numeric datetime formats can be used, but care should be taken with newer formats which might not be understood by all statistical packages.

11 Example 8: Analysis of Concomitant Medications

This example displays a simple summary of all concomitant medications. The example is based on a 2-treatment parallel design study. The display summarizes (1) the number of patients in each treatment group who took a concomitant medication and (2) the rate of occurrence in each treatment group. In this example, analysis results metadata have not been included. As stated in the ADaMIG, analysis results metadata are not needed or even advisable for every analysis included in a clinical study report or submission.

11.1 Analysis Display Example Layout - Example 8

Figure 11.1.1 Example Layout for Concomitant Medications Summary[a]

Summary of Concomitant Medications by Medication Class and Medication Name
Analysis Population: Safety

Medication Class/Medication Name	Treatment A (N=4)	Treatment B (N=5)	Total (N=9)
Any Concomitant Medication	4 (100.0%)	4 (80.0%)	8 (88.9%)
ANALGESICS	2 (50.0%)	2 (40.0%)	4 (44.4%)
PARACETAMOL	2 (50.0%)	2 (40.0%)	4 (44.4%)
ANTIBACTERIALS FOR SYSTEMIC USE	1 (25.0%)	1 (20.0%)	2 (22.2%)
AMOXICILLIN	1 (25.0%)	1 (20.0%)	2 (22.2%)
ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS	1 (25.0%)	2 (40.0%)	3 (33.3%)
IBUPROFEN	1 (25.0%)	2 (40.0%)	3 (33.3%)
DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	0	2 (40.0%)	2 (22.2%)
MONTELUKAST	0	1 (20.0%)	1 (11.1%)
SALBUTAMOL	0	2 (40.0%)	2 (22.2%)
NASAL PREPARATIONS	2 (50.0%)	0	2 (22.2%)
FLUTICASONE PROPIONATE	2 (50.0%)	0	2 (22.2%)
PSYCHOANALEPTICS	1 (25.0%)	0	1 (11.1%)
SERTRALIN	1 (25.0%)	0	1 (11.1%)

[a]The style of the display of the results of an analysis is determined by the producer. The example is intended to illustrate content, not appearance.

11.2 Sample ADaM Variable Metadata - Example 8

This example describes a Concomitant Medications ADaM dataset named ADCM. ADCM is not a required dataset name.

Table 11.2.1 Example of ADaM Variable Metadata

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADCM	STUDYID	Study Identifier	text			CM.STUDYID	
ADCM	USUBJID	Unique Subject Identifier	text			CM.USUBJID	
ADCM	CMSEQ	Sequence Number	integer			CM.CMSEQ	

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADCM	CMTRT	Reported Name of Drug, Med or Therapy	text			CM.CMTRT	
ADCM	CMMODIFY	Modified Reported Name	text			CM.CMMODIFY	
ADCM	CMDECOD	Standardized Medication Name	text		WHODRUG		CM.CMDECOD WHO Drug Dictionary March 2019
ADCM	ATC1CD	ATC Level 1 Code	text		WHODRUG		ATC Level 1 Code WHO Drug Dictionary March 2019
ADCM	ATC2CD	ATC Level 2 Code	text		WHODRUG		ATC Level 2 Code WHO Drug Dictionary March 2019
ADCM	ATC3CD	ATC Level 3 Code	text		WHODRUG		ATC Level 3 Code WHO Drug Dictionary March 2019
ADCM	ATC1	ATC Level 1 Text	text		WHODRUG		ATC Level 1 Text WHO Drug Dictionary March 2019
ADCM	ATC2	ATC Level 2 Text	text		WHODRUG		ATC Level 2 Text WHO Drug Dictionary March 2019
ADCM	ATC3	ATC Level 3 Text	text		WHODRUG		ATC Level 3 Text WHO Drug Dictionary March 2019
ADCM	AOCCFL	1st Occurrence within Subject Flag	text	Y			<Producer will insert derivation here>
ADCM	AOCC01FL	First Occurrence of ATC Level 1 Flag	text	Y			<Producer will insert derivation here>
ADCM	AOCC02FL	First Occurrence of ATC Level 2 Flag	text	Y			<Producer will insert derivation here>
ADCM	AOCC03FL	First Occurrence of ATC Level 3 Flag	text	Y			<Producer will insert derivation here>
ADCM	AOCCPFL	1st Occurrence of Preferred Term Flag	text	Y			<Producer will insert derivation here>
ADCM	CMINDC	Indication	text			CM.CMINDC	
ADCM	CMDOSFRM	Dose Form	text	TABLET		CM.CMDOSFRM	
ADCM	CMDOSE	Dose per Administration	text			CM.CMDOSE	
ADCM	CMDOSU	Dose Units	text	mg		CM.CMDOSU	
ADCM	CMDOSFRQ	Dosing Frequency Per Interval	text	ONCE, PRN, QD, QID		CM.CMDOSFRQ	
ADCM	CMROUTE	Route of Administration	text	ORAL		CM.CMROUTE	
ADCM	CMSTDTC	Start Date/Time of Medication	date		ISO 8601	CM.CMSTDTC	
ADCM	ASTDT	Analysis Start Date	integer				<Producer will insert derivation here>
ADCM	ASTDTF	Analysis Start Date Imputation Flag	text	D, M, Y			If start date is completely missing or missing the year then ASTDTF="Y" Else if start date has month missing then ASTDTF="M" Else if start date has day missing then ASTDTF="D"
ADCM	CMENDTC	End Date/Time of Medication	date		ISO 8601	CM.CMENDTC	
ADCM	AENDT	Analysis End Date	integer				<Producer will insert derivation here>
ADCM	AENDTF	Analysis End Date Imputation Flag	text	D, M, Y			If end date is completely missing or missing the year then AENDTF="Y" Else if end date has month missing then AENDTF="M" Else if end date has day missing then AENDTF="D"
ADCM	CMENRF	End Relative to Reference Period	text		ONGOING	CM.CMENRF	
ADCM	ONTRTFL	On-Treatment Flag	text	Y			<Producer will insert derivation here>
ADCM	PREFL	Pre-treatment Flag	text	Y			<Producer will insert derivation here>
ADCM	SAFFL	Safety Population Flag	text	Y, N		ADSL.SAFFL	
ADCM	TRTA	Actual Treatment	text	Drug A, Drug B		ADSL.TRT01A	

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADCM	TRTAN	Actual Treatment (N)	integer	1, 2			ADSL.TRT01AN Drug A = 1 Drug B = 2
ADCM	TRTSDT	Date of First Exposure to Treatment	integer			ADSL.TRTSDT	
ADCM	TRTEDT	Date of Last Exposure to Treatment	integer			ADSL.TRTEDT	
ADCM	AGE	Age	integer			ADSL.AGE	
ADCM	AGEGR1	Pooled Age Group 1	text	<65, >=65		ADSL.AGEGR1	
ADCM	SEX	Sex	text	M, F		ADSL.SEX	
ADCM	RACE	Race	text	ASIAN		ADSL.RACE	

11.3 Sample ADaM Data - Example 8

Table 11.3.1 is an illustration of the Concomitant Medications analysis dataset (ADCM) defined in Section 11.2, Sample ADaM Variable Metadata - Example 8. The ADCM dataset illustrated in this example was designed to support some standard subsets and/or classifications of concomitant medications. The example describes some of the key variables and records that would be included in the dataset.

Key points to note in the example are:

1. The producer of the dataset chose to use the record-level actual treatment variable (TRTA) populated with the same value across all rows in the dataset rather than the subject-level treatment variable (TRT01A). For a parallel design, either TRTA or TRT01A could be used as the actual treatment identifier. The producer interpreted TRTA as the treatment associated with the record for analysis display purposes and populated the baseline records with treatment even though subjects had not yet received treatment at that time.
2. Variables such as CMSEQ, CMTRT, and CMSTDTC are copied in from the SDTM CM domain to provide data-point traceability.
3. Variables such as CMDECOD are copied in from the SDTM CM domain for analysis purposes.
4. ASTDT and AENDT are the CM timing variables used for analysis. Other timing variables (e.g., ASTDTF, AENDTF, CMSTDTC, CMENDTC, TRTSDT, TRTEDT) are supportive variables for metadata traceability.
5. The occurrence flags AOCCFL, AOCCPFL, AOCC01FL, AOCC02FL, and AOCC03FL are permissible, and not required. The main purpose of these flags is to facilitate data-point traceability between records in the dataset and unique counts in the summary displays. In addition, if a time-to-event analysis is built off of concomitant medications, the flags provide a crucial link between the summary records in the TTE Basic Data Structure and the source of the records in ADCM.
6. The core variables of AGE, AGEGR1, SEX, and RACE are included in ADCM to facilitate subgroup analyses.

Table 11.3.1 Sample ADCM Data[a]

Row	STUDYID	USUBJID	CMSEQ	CMTRT	CMMODIFY	CMDECOD	ATC1CD	ATC1	ATC2CD	ATC2	ATC3CD	ATC3	AOCCFL	AOCCPFL	AOCC01FL	AOCC02FL	AOCC03FL	CMINDC	CMDOSFRM	CMDOSE	CMDOSU	CMDOSFRQ	CMROUTE	CMSTDTC	ASTDT	CMENDTC	AENDT	CMENRF	ONTRTFL	PREFL	SAFFL	TRTA	TRTAN	TRTSDT	TRTEDT	AGE	AGEGR1	SEX	RACE
1	ABC	ABC-001	1	TYLENOL	TYLENOL	PARACETAMOL	N	NERVOUS SYSTEM	N02	ANALGESICS	N02B	OTHER ANALGESICS AND ANTIPYRETICS	Y	Y	Y	Y	Y	HEADACHE	TABLET	100	mg	ONCE	ORAL	2011-01-02	02JAN2011	2011-01-02	02JAN2011			Y	Y	Drug A	1	23JAN2011	15MAY2011	54	<65	M	ASIAN
2	ABC	ABC-001	2	TYLENOL	TYLENOL	PARACETAMOL	N	NERVOUS SYSTEM	N02	ANALGESICS	N02B	OTHER ANALGESICS AND ANTIPYRETICS						HEADACHE	TABLET	100	mg	ONCE	ORAL	2011-01-04	04JAN2011	2011-01-04	04JAN2011			Y	Y	Drug A	1	23JAN2011	15MAY2011	54	<65	M	ASIAN
3	ABC	ABC-001	3	TYLENOL	TYLENOL	PARACETAMOL	N	NERVOUS SYSTEM	N02	ANALGESICS	N02B	OTHER ANALGESICS AND ANTIPYRETICS						HEADACHE	TABLET	100	mg	ONCE	ORAL	2011-01-10	10JAN2011	2011-01-10	10JAN2011			Y	Y	Drug A	1	23JAN2011	15MAY2011	54	<65	M	ASIAN
4	ABC	ABC-001	4	TYLENOL	TYLENOL	PARACETAMOL	N	NERVOUS SYSTEM	N02	ANALGESICS	N02B	OTHER ANALGESICS AND ANTIPYRETICS						HEADACHE	TABLET	100	mg	ONCE	ORAL	2011-01-15	15JAN2011	2011-01-15	15JAN2011			Y	Y	Drug A	1	23JAN 2011	15MAY2011	54	<65	M	ASIAN
5	ABC	ABC-001	5	CONTAC MS	CONTAC MS	CONTAC MS	N	NERVOUS SYSTEM	N02	ANALGESICS	N02B	OTHER ANALGESICS AND ANTIPYRETICS		Y				COLD	TABLET	200	mg	ONCE	ORAL	2011-01-17	17JAN2011	2011-01-17	17JAN2011			Y	Y	Drug A	1	23JAN2011	15MAY2011	54	<65	M	ASIAN
6	ABC	ABC-001	6	FLONASE	FLONASE	FLUTICASONE PROPIONATE	R	RESPIRATORY SYSTEM	R01	NASAL PREPARATIONS	R01A	DECONGESTANTS AND OTHER NASAL PREPARATIONS FOR TOP		Y	Y	Y	Y	COUGH	TABLET	50	mg	QD	ORAL	2009-02-01	01FEB2009			ONGOING	Y	Y	Y	Drug A	1	23JAN2011	15MAY2011	54	<65	M	ASIAN
7	ABC	ABC-002	1	ROBITUSSIN COUGH	ROBITUSSIN	NOVAHISTINE DMX	R	RESPIRATORY SYSTEM	R05	COUGH AND COLD PREPARATIONS	R05FA	COUGH SUPPRESSANTS AND EXPECTORANTS, COMBINATIONS		Y		Y	Y	INFECTION	SUSPENSION	500	mg	QID	ORAL	2011-03-01	01MAR2011	2011-03-15	15MAR2011			Y	Y	Drug B	2	10MAY2011	25NOV201	54	<65	M	ASIAN
8	ABC	ABC-002	2	MOTRIN	MOTRIN	IBUPROFEN	M	MUSCULO-SKELETAL SYSTEM	M01	ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS	M01A	ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS, NON-S		Y	Y	Y	Y	LEG PAIN	TABLET	500	mg	PRN	ORAL	2011-05-14	14MAY2011	2011-06-01	01JUN2011		Y		Y	Drug B	2	10MAY2011	25NOV2011	54	<65	M	ASIAN
9	ABC	ABC-002	3	IBUPROFEN	IBUPROFEN	IBUPROFEN	M	MUSCULO-SKELETAL SYSTEM	M01	ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS	M01A	ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS, NON-S						ARTHRITIS	TABLET	250	mg	QD	ORAL	2011-06-10	10JUN2011			ONGOING	Y		Y	Drug B	2	10MAY2011	25NOV2011	54	<65	M	ASIAN
10	ABC	ABC-003	1	ZOLLOFT	ZOLOFT	SERTRALIN	N	NERVOUS SYSTEM	N06	PSYCHOANALEPTICS	N06A	ANTIDEPRESSANTS		Y	Y	Y	Y	ANXIETY	TABLET	50	mg	QD	ORAL	2001-03				ONGOING	Y	Y	Y	Drug A	1	16JUN2011	03JAN2012	54	<65	M	ASIAN

[a]Variables ending in suffix DTC are character date/time fields in ISO 8601 format. Variables ending in DT are numeric dates, here shown using SAS date format date9. Other numeric date formats can be used, but care should be taken with newer date formats which might not be understood by all statistical packages.

12 Example 9: Analysis of Medical History Mapped to MedDRA

The basic summary of medical history frequencies described in ICH Guideline E3[5] Sections 12.2.2 and 14.3.1 should be used to display frequencies in treatment and control groups.

This example displays a simple summary of all spontaneously reported medical history. The example is based on a 2-treatment parallel design study. The display summarizes (1) the number of subjects in each treatment group who had a given medical history event and (2) the rate of occurrence in each treatment group.

12.1 Analysis Display Example Layout - Example 9

The count and percentage of unique subjects per classification group may be based on any of the classification variables. In this example, the count and percentage of unique subjects in Figure 12.1.1 would be summarized by the variables MHSCAT, MHBODSYS, and MHDECOD. The table also summarizes the number of subjects who had any medical history event (i.e., the row Any Medical History). The denominator counts (the (N=) in column headings) are taken from the ADSL dataset and are based on the count of subjects in the population of interest. Note that not all subjects in the population of interest will necessarily have data in the MH dataset.

This presentation is analogous to the logic typically used for adverse event summaries.

Figure 12.1.1 Example Layout of Medical History Summary[a]

Summary of General Medical History Events
Safety Population

HISTORY CATEGORY Body System Event	Active Drug (N=4)	Placebo (N=5)	Total (N=9)
Any Medical History	3 (75.0%)	3 (60.0%)	6 (66.7%)
GASTROINTESTINAL	1 (25.0%)	0	1 (11.1%)
Gastrointestinal disorders	1 (25.0%)	0	1 (11.1%)
Abdominal pain	1 (25.0%)	0	1 (11.1%)
Gastroesophageal reflux disease	1 (25.0%)	0	1 (11.1%)
Nausea	1 (25.0%)	0	1 (11.1%)
HEMATOLOGICAL/LYMPHATIC	1 (25.0%)	0	1 (11.1%)
Blood and lymphatic system disorders	1 (25.0%)	0	1 (11.1%)
Anaemia	1 (25.0%)	0	1 (11.1%)
RESPIRATORY	3 (75.0%)	3 (60.0%)	6 (66.7%)
Immune system disorders	2 (50.0%)	3 (60.0%)	5 (55.6%)
Seasonal allergy	2 (50.0%)	3 (60.0%)	5 (55.6%)
Infections and infestations	1 (25.0%)	2 (40.0%)	3 (33.3%)
Upper respiratory tract infection	1 (25.0%)	2 (40.0%)	3 (33.3%)
Respiratory, thoracic and mediastinal disorders	1 (25.0%)	3 (60.0%)	4 (44.4%)
Asthma	1 (25.0%)	0	1 (11.1%)
Dyspnoea	0	3 (60.0%)	3 (33.3%)

[a]The style of the display of the results of an analysis is determined by the producer. The example is intended to illustrate content, not appearance.

12.2 Sample ADaM Variable Metadata - Example 9

This example describes a Medical History ADaM dataset named ADMH. ADMH is not a required dataset name.

Table 12.2.1 Example of ADaM Variable Metadata

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADMH	STUDYID	Study Identifier	text			MH.STUDYID	
ADMH	USUBJID	Unique Subject Identifier	text			MH.USUBJID	
ADMH	MHSEQ	Sequence Number	integer			MH.MHSEQ	
ADMH	MHCAT	Category for Medical History	text			MH.MHCAT	

Dataset Name	Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
ADMH	MHSCAT	Sub Category for Medical History	text			MH.MHSCAT	
ADMH	MHDECOD	Dictionary-Derived Term	text			MH.MHDECOD	
ADMH	MHBODSYS	Body System or Organ Class	text			MH.MHBODSYS	
ADMH	MHTERM	Reported Term for the Medical History	text			MH.MHTERM	
ADMH	MHSTDTC	Start Date/Time of Medication	datetime		ISO 8601	MH.MHSTDTC	
ADMH	ASTDT	Analysis Start Date	integer				From MH.MHSTDTC, converted to SAS date. Any derivations to derive partial start dates are applied here and listed in comments.
ADMH	ASTTM	Analysis Start Time	integer				From MH.MHSTDTC, converted to SAS time
ADMH	ASTDTM	Analysis Start Datetime	integer				From MH.MHSTDTC, converted to SAS datetime
ADMH	MHENDTC	End Date/Time of Medication	datetime		ISO 8601	MH.MHENDTC	
ADMH	AENDT	Analysis End Date	integer				From MH.MHENDTC, converted to SAS date. Any derivations to derive partial start dates are applied here and listed in comments.
ADMH	AENTM	Analysis End Time	integer				From MH.MHENDTC, converted to SAS time
ADMH	AENDTM	Analysis End Datetime	integer				From MH.MHENDTC, converted to SAS datetime
ADMH	MHENRF	End Relative to Reference Period	text			MH.MHENRF	

12.3 Sample ADaM Data - Example 9

Table 12.3.1 is an illustration of the Medical History analysis dataset (ADMH) defined in Section 12.2, Sample ADaM Variable Metadata - Example 9. The ADMH dataset illustrated in this example was designed to support some standard subsets and/or classifications of medical histories. The example describes key variables and records that would be included in the dataset.

The example data are assumed to be gathered on a CRF that contains a set of defined categories. A subject may or may not have had any significant medical history in any of the categories on the form. There is a record in the SDTM MH dataset for each symptom or condition listed on the form; subjects with no recorded medical history may not appear in this file. The MHCAT variable indicates the type of CRF page on which data were gathered, and MHSCAT indicates the CRF category. MHTERM is the symptom term that was recorded; MHDECOD and MHBODSYS are taken from matching the text in MHTERM with a coding dictionary (in this case MedDRA). The date variables indicate the beginning and end timing of the medical history event.

Key points to note in the example are:

1. Variables such as MHSTDTC and MHENDTC are copied in from the SDTM MH domain to provide data-point traceability.
2. Variables such as MHSCAT, MHDECOD, and MHBODSYS are copied in from the SDTM MH domain for analysis purposes.
3. ASTDT and AENDT are the timing variables used for analysis. Timing variable MHENRF is a supportive variable for metadata traceability.

4. This is a simple example to only illustrate variables that are relevant to ADMH. It does not include all variables that could be needed for analysis or all indicated in the metadata in Section 12.2, Sample ADaM Variable Metadata - Example 9. For example, it does not include variables such as severity of the MH event.
5. For this analysis, the subject-level treatment variable (TRT01A or TRT01P, not shown) would be appropriate. Record-level treatment variables would not be needed because the data are gathered prior to start of study treatment.
6. The TRTEMFL (treatment-emergent flag), ONTRTFL (on-treatment flag), FUPFL (follow-up flag), and PREFL (pre-treatment flag) are not included in this analysis file because this dataset will only be used for baseline summaries. (Note that these variables could be defined if there was an analysis purpose that called for them.) Similarly, occurrence flags are permissible but not required unless needed for a specific analysis purpose.
7. Core variables (e.g., AGE, RACE, and SEX) would typically be added to the dataset but are not shown in this example.

Table 12.3.1 Sample Medical History Data for Spontaneously Reported Events[a]

Row	USUBJID	MHTERM	MHDECOD	MHBODSYS	MHCAT	MHSCAT	MHSTDTC	ASTDT	MHENDTC	AENDT	MHENRF
1	ABC-001	ANEMIA	Anaemia	Blood and lymphatic system disorders	MEDICAL HISTORY	HEMATOLOGICAL/ LYMPHATIC	2010-02-01	01FEB2010			ONGOING
2	ABC-001	GERD	Gastroesophageal reflux disease	Gastrointestinal disorders	MEDICAL HISTORY	GASTROINTESTINAL	2011-01-04	04JAN2011	2011-01-04	04JAN2011	
3	ABC-001	NAUSEA	Nausea	Gastrointestinal disorders	MEDICAL HISTORY	GASTROINTESTINAL	2011-01-10	10JAN2011	2011-01-10	10JAN2011	
4	ABC-001	SPLEEN PAIN	Abdominal pain	Gastrointestinal disorders	MEDICAL HISTORY	GASTROINTESTINAL	2011-01-15	15JAN2011	2011-01-15	15JAN2011	
5	ABC-001	ASTHMA	Asthma	Respiratory, thoracic and mediastinal disorders	MEDICAL HISTORY	RESPIRATORY	2011-01-17	17JAN2011	2011-01-17	17JAN2011	
6	ABC-002	SEASONAL ALLERGIES	Seasonal allergy	Immune system disorders	MEDICAL HISTORY	RESPIRATORY	2011-05-14	14MAY2011	2011-06-01	01JUN2011	

[a]Variables ending in suffix DTC are character date/time fields in ISO 8601 format. Variables ending in DT are numeric dates, here shown using SAS date format date9. Other numeric date formats can be used, but care should be taken with newer date formats which might not be understood by all statistical packages.

13 Example 10: Analysis of Medical History Pre-specified Events

For the example in this section, data are gathered on a CRF that contains a pre-specified category (i.e, diabetes history), including a checkbox to indicate whether the subject had this condition. Diabetes history is not coded.

Analysis of the number of subjects with and without pre-specified events is an option for medical history. This option does not have a counterpart in adverse events analysis, because the AE domain does not allow for the collection of pre-specified events with AEOCCUR = "N".

13.1 Analysis Display Example Layout - Example 10

The data are analyzed here by counting the number of unique subjects per treatment group (MHCAT, MHTERM, and MHOCCUR). The values of MHOCCUR are formatted from "Y" and "N" to more readable values (e.g., "Y = Reported History") for presentation. Here it is assumed that the ADMH file is merged with ADSL to ensure that all safety subjects are identified. However, there may be some safety subjects in ADSL who do not occur in ADMH. (This situation can occur due to missing CRF data.)

Note: The style of the display of the results of an analysis are determined by the producer. These examples illustrate content, not appearance.

Figure 13.1.1 Example Layout of Medical History Summary

Summary of Diabetes History Events
Safety Population

Diabetes History Category	Active Drug (N=4)	Placebo (N=4)	Total (N=8)
DIABETES HISTORY			
DIABETES MELLITUS[1]			
N=No History	3 (75.0%)	2 (100.0%)	5 (83.3%)
Y=Reported History	1 (25.0%)	0	1 (16.7%)
Diabetes History Not Available[1]	0	2	2

[1] Population counts in the column header include all subjects in the safety population. Percentages are based on the number of safety subjects in each treatment group for whom diabetes history data are available. The 'No Reported History' counts are based on subjects with the 'No' box checked on Medical History CRF page xxx.

The choice of denominator is based on statistical judgment and should be clearly described in the programming specifications. The choice of denominator should also be clearly identified somewhere on the report (e.g., in the title, via a footnote).

In this example, the denominator is based on only the subjects in the population of interest who have records in ADMH with MHCAT = "DIABETES HISTORY".

An alternative analysis would be to base the denominator on the number of subjects in the population (typically defined by the number of subjects with appropriate population flags in ADSL), as shown in Figure 13.1.2.

Figure 13.1.2 Alternate Example Layout of Medical History Summary

Summary of Diabetes History Events
Safety Population

Diabetes History Category	Active Drug (N=4)	Placebo (N=4)	Total (N=8)
DIABETES HISTORY			
DIABETES MELLITUS			
N=No Reported History	3 (75.0%)	2 (50.0%)	5 (62.5%)
Y=Reported History	1 (25.0%)	0 (0.0%)	1 (12.5%)
Unknown[1]	0 (0.0%)	2 (50.0%)	2 (25.0%)

[1] Population counts in the column header include all subjects in the safety population. Percentages are based on the number of safety subjects in each treatment group, whether they had diabetes history data or not. The 'Unknown' counts are based on subjects who did not have a Medical History CRF page xxx.

13.2 Sample ADaM Variable Metadata - Example 10

This example describes a Medical History ADaM dataset named ADMH. ADMH is not a required dataset name.

Table 13.2.1 Sample ADaM Variable Metadata

Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
STUDYID	Study Identifier	text			MH.STUDYID	
USUBJID	Unique Subject Identifier	text			MH.USUBJID	
MHSEQ	Sequence Number	integer			MH.MHSEQ	
MHCAT	Category for Medical History	text			MH.MHCAT	
MHTERM	Reported Term for the Medical History	text			MH.MHTERM	
MHSTDTC	Start Date/Time of Medical History Event	datetime		ISO 8601	MH.MHSTDTC	
ASTDT	Analysis Start Date	integer				From MHSTDTC, converted to SAS date. Any derivations to derive partial start dates are applied here and listed in comments.
ASTTM	Analysis Start Time	integer				From MH.MHSTDTC, converted to SAS time
ASTDTM	Analysis Start Datetime	integer				From MH.MHSTDTC, converted to SAS datetime
MHENDTC	End Date/Time of Medical History Event	datetime		ISO 8601	MH.MHENDTC	
AENDT	Analysis End Date	integer				From MHENDTC, converted to SAS date. Any derivations to derive partial start dates are applied here and listed in comments.
AENTM	Analysis End Time	integer				From MHENDTC, converted to SAS time
AENDTM	Analysis End Datetime	integer				From MHENDTC, converted to SAS datetime
MHPRESP	Medical History Event Pre-Specified	text	Y			MH.MHPRESP Is Med Hx event from pre-specified CRF page
MHOCCUR	Medical History Occurrence	text	N, Y			MH.MHOCCUR Did subject have the event, Y or N
MHENRTPT	End Relative to Reference Time Point	text	ONGOING		MH.MHENRTPT	
MHENTPT	End Reference Time Point	text	SCREENING		MH.MHENTPT	

13.3 Sample ADaM Data - Example 10

Table 13.3.1 is an illustration of the Medical History (Pre-specified Events) analysis dataset (ADMH) defined in Section 13.2, Sample ADaM Variable Metadata - Example 10. The ADMH dataset illustrated in this example was designed to support some standard subsets and/or classifications of medical histories. The example describes key variables and records that would be included in the dataset.

Key points to note in the example are:

1. This is a simple example to only illustrate the ADMH Pre-specified Events. It does not include all variables that could also be needed for analysis or all indicated in metadata.
2. The dataset is prepared to support analysis of pre-specified events by populating the variables MHCAT, MHTERM, MHPRESP, and MHOCCUR.
3. The MHCAT variable indicates the type of CRF page where the data were gathered, and MHPRESP is "Y" to indicate that the term is pre-specified. MHOCCUR is either "Y" or "N" to indicate whether the subject did or did not have the event.
4. MHTERM is the symptom term. Because the MHTERM variable for pre-specified events will have a known and finite set of values, these values are used here as a summarization category.
5. In this example, MHDECOD and MHBODSYS are not used. Instead, MHCAT is used to categorize the data.
6. The date variables indicate the beginning and end timing of the medical history event (if any) and are null on records that do not indicate an event.

Table 13.3.1 Sample Medical History Data for Pre-specified Events[a]

Row	USUBJID	MHSEQ	MHTERM	MHCAT	MHPRESP	MHOCCUR	MHSTDTC	MHENDTC	MHENRTPT	MHENTPT
1	ABC-001	6	DIABETES MELLITUS	DIABETES HISTORY	Y	N				
2	ABC-002	1	DIABETES MELLITUS	DIABETES HISTORY	Y	N				
3	ABC-003	1	DIABETES MELLITUS	DIABETES HISTORY	Y	Y	2001-03		ONGOING	SCREENING
4	ABC-004	3	DIABETES MELLITUS	DIABETES HISTORY	Y	N				
5	ABC-005	4	DIABETES MELLITUS	DIABETES HISTORY	Y	N				
6	ABC-006	6	DIABETES MELLITUS	DIABETES HISTORY	Y	N				
7	ABC-007	5	DIABETES MELLITUS	DIABETES HISTORY	Y	N				
8	ABC-008	6	DIABETES MELLITUS	DIABETES HISTORY	Y	N				
9	ABC-009	1	DIABETES MELLITUS	DIABETES HISTORY	Y	N				

[a]Variables ending in suffix DTC are character date/time fields in ISO 8601 format.

14 Example 11: Analysis of Protocol Deviations

Many OCCDS datasets can be created more or less directly from SDTM data. However, it is not a requirement that all OCCDS data be sourced directly from records in an SDTM domain. Instead, records may be derived or come from multiple SDTM domains or ADaM datasets.

This example displays a summary of all protocol deviations or violations, many of which must be derived. The example is based on a 2-treatment parallel study design. The display summarizes the number of patients who had a protocol deviation or violation in each treatment group. In this example, analysis results metadata have not been included.

14.1 Analysis Display Example Layout - Example 11

The data are analyzed here by counting the number of unique subjects per treatment group.

Figure 14.1.1 Example Layout of Protocol Deviations/Violations Summary

Summary of Subjects With Protocol Deviations/Violations
Safety Population

Protocol Deviation/Violation	Drug A (N=xxx)	Drug B (N=xxx)
Subjects with at least one major protocol deviation/violation	xx (xx.x%)	xx (xx.x%)
Average of Less than 70 mm on the Visual Analog Scale of Worst Knee Pain Intensity at Baseline	xx (xx.x%)	xx (xx.x%)
Less than 4 pain scores recorded in diary on Days -7 through Day -1 (Baseline)	xx (xx.x%)	xx (xx.x%)
Less than 28 Days of Treatment	xx (xx.x%)	xx (xx.x%)
Less than 80% Overall Compliance with Study Drug	xx (xx.x%)	xx (xx.x%)
Developed withdrawal criteria but not withdrawn	xx (xx.x%)	xx (xx.x%)
Entered but did not satisfy criteria	xx (xx.x%)	xx (xx.x%)
Received a disallowed concomitant treatment	xx (xx.x%)	xx (xx.x%)
etc…		
Subjects with at least one minor protocol deviation/violation	xx (xx.x%)	xx (xx.x%)
Visit that falls greater than +/-3 days out of window	xx (xx.x%)	xx (xx.x%)
Skipped or missed assessment/procedure	xx (xx.x%)	xx (xx.x%)
At least one skipped or missed dose	xx (xx.x%)	xx (xx.x%)
etc...		
Note: Subjects may appear in more than one category.		

The analysis above can be further split out by site/center/investigator to look for issues within these categories.

14.2 Sample ADaM Variable Metadata - Example 11

This example describes a Protocol Deviations ADaM dataset named ADPDEV. ADPDEV is not a required dataset name. The dataset name ADDV was not used because data come from not only DV, but also the SDTM Inclusion/Exclusion Criteria Not Met (IE) domain, other ADaM datasets (ADSL, ADPAIN, ADVS, and ADEG), and some completely derived records (see Table 14.2.1).

The ADaM datasets ADPDEV and ADSL are used to create the table described in Section 14.1, Analysis Display Example Layout - Example 11.

The sample ADPDEV dataset shown in Table 14.2.2 contains only a subset of variables needed for the analysis and is for illustration purposes only. Additional variables need to be created to make the datasets submission-ready. In this table, information is provided to highlight the various source data utilized to create this analysis dataset.

Table 14.2.1 Sample Source Data for ADPDEV

Dataset	Label	Description	Variables/Parameters of Interest
ADPAIN	Knee Pain Intensity Analysis Dataset	Dataset contains 2 parameters of interest: 1 parameter is the daily worst knee pain scores (intensity) captured in source data and the other is the derived average weekly worst knee pain intensity. Analysis windows for 7-day periods are created from baseline through week 16, based on start date of treatment. If 4 or more daily pain scores fall within a window, an average is created. Analysis is performed on these "averages" and primary/sensitivity analyses are performed with focus on week 16.	PARAMCD="DKNEEPI" (Daily Worst Knee Pain Intensity) PARAMCD="AKNEEPI" (Average Worst Knee Pain Intensity)
ADSL	Subject Level Analysis Dataset	ADSL subject-level trial experience variables	TRTDURD (Total Treatment Duration (Days)) TRCMP (Treatment Compliance (%))
DV	Protocol Deviations	The intent of the domain is to capture protocol violations and deviations during the course of the study and will store only those criteria violated by or deviated from by the subject and not a response to each violation or deviation. Usually these are deviations that occur after the subject has been randomized or received the first treatment.	Protocol Deviation Term (DVTERM), Protocol Deviation Coded Term (DVDECOD), and Start Datetime of Deviation (DVSTDTC)
IE	Inclusion/Exclusion Criteria Not Met	The intent of the domain model is to only collect those criteria that cause the subject to be in violation of the inclusion/exclusion criteria, not a response to each criterion.	IETEST (Inclusion/Exclusion Criterion)
ADVS ADEG	Vital Signs Analysis Dataset ECG Analysis Dataset	1 record per subject per parameter per analysis visit per analysis datetime Vital Signs and ECG are scheduled at each site visit biweekly	Analysis Relative Day (ADY), Analysis Window Target (AWTARGET), Analysis Window Beginning Timepoint (AWLO), Analysis Window Ending Timepoint (AWHI),

Although this example follows an OCCDS structure, it has data from multiple sources of varying structure and derived records. This adds additional complexity to the metadata. Variable metadata alone is not sufficient to clearly describe the source or derivation of one record as compared to another. Unlike BDS, there is no inherent variable like PARAM to identify a set of 1 or more rows in the dataset and easily create value-level metadata. This example uses SRCDOM to group records from the same source and create value-level metadata for ATERM, ADECOD, and ASTDT. This works fairly well here as records from 1 source have at most 2 derivations that are fairly straightforward. For more complex examples, the source/derivation column for value-level metadata could reference a table in the ADRG for specific details.

Table 14.2.2 Sample ADaM Variable Metadata

Variable Name	Variable Label	Variable Type	Codelist	Controlled Terms	Source	Derivation
STUDYID	Study Identifier	text			ADSL.STUDYID	
USUBJID	Unique Subject Identifier	text			ADSL.USUBJID	
ASTDT	Analysis Start Date	integer				See Value-level Metadata below, split out by SRCDOM
ATERM	Analysis Term	text				See Value-level Metadata below, split out by SRCDOM
ADECOD	Analysis Standardized Term	text				See Value-level Metadata below. split out by SRCDOM
ACAT1	Analysis Category 1	text	MINOR PROTOCOL DEVIATION, MAJOR PROTOCOL DEVIATION			Assign as "MINOR PROTOCOL DEVIATION" or "MAJOR PROTOCOL DEVIATION" based on Section XX of SAP
AVISIT	Analysis Visit	text				AVISIT from ADVS, ADEG and ADPAIN
AVISITN	Analysis Visit (N)	integer				AVISITN from ADVS, ADEG and ADPAIN
SRCDOM	Source Data	text				Populate with the dataset name respective to the source of each observation
SRCSEQ	Source Sequence Number	integer				Populate with the sequence number from each source dataset.

Table 14.2.3 Sample ADaM Value-level Metadata

Variable	Where	Type	Codelist	Controlled Terms	Source/Derivation/Comment
ATERM	SRCDOM="DV"	text			DV.DVTERM
ATERM	SRCDOM="IE"	text			IE.IETEST
ATERM	SRCDOM="ADSL"	text			Evaluation of 2 potential deviations in ADSL: 1. For Treatment Duration: If missing<TRTDURD<28 then ATERM="Subject has only (TRTDURD) day/days of treatment" 2. For Treatment Compliance: If missing< TRCMP<80 then ATERM="Subject was (TRCMP) % compliant with study drug"
ATERM	SRCDOM="ADPAIN"	text			Evaluation of 2 potential deviations in ADPAIN: 1. For Pain not meeting criteria: Subject has record where PARAMCD="AKNEEPI" (Average Worst Knee Pain Intensity) and AVISIT="Baseline" and missing<AVAL<70 then ATERM="Subject has an average of (AVAL) mm on the Visual Analog Scale of Worst Knee Pain Intensity" 2. For Number of Assessments not meeting criteria: Get records where PARAMCD="DKNEEPI" (Daily Worst Knee Pain Intensity) and −7 ≤ ADY≤ −1. If count of these records is less than 4, ATERM="Subject has only COUNT daily worst knee pain scores recorded in the source data between days -7 to days -1"
ATERM	SRCDOM in("ADVS" "ADEG")	text			Stack ADVS and ADEG keeping Subject (USUBJID), Analysis Visit (N) (AVISITN), Analysis Visit (AVISIT), Analysis Relative Day (ADY), Analysis Window Target (AWTARGET), Analysis Window Beginning Timepoint (AWLO), Analysis Window Ending Timepoint (AWHI) where subject has an assessment. Get rid of duplicate observations by these variables. If missing<ADY<AWLO-3 or ADY>AWHI+3 then ATERM="Visit Out of (+/-3 Day) Window"
ADECOD	SRCDOM="DV"	text			DV.DVDECOD
ADECOD	SRCDOM="IE"	text			ADECOD="Entered but did not satisfy criteria"
ADECOD	SRCDOM="ADSL"	text			Evaluation of 2 potential deviations in ADSL: 1. For Treatment Duration: If missing<TRTDURD<28 then ADECOD="Less than 28 Days of Treatment" 2. For Treatment Compliance: If missing< TRCMP<80 then ADECOD="Less than 80% Overall Compliance with Study Drug" **Note:** The decodes used in example are sponsor-defined controlled terminology.
ADECOD	SRCDOM="ADPAIN"	text			Evaluation of 2 potential deviations in ADPAIN: 1. For Pain not meeting criteria: Subject has record where PARAMCD="AKNEEPI" (Average Worst Knee Pain Intensity) and AVISIT="Baseline" and missing<AVAL<70 then ADECOD="Average of Less than 70 mm on the Visual Analog Scale of Worst Knee Pain Intensity at Baseline" 2. For Number of Assessments not meeting criteria: Get records where PARAMCD="DKNEEPI" (Daily Worst Knee Pain Intensity) and −7 ≤ ADY≤ −1. If count of these records is less than 4, ADECOD="Less than 4 pain scores recorded in diary on Days --7 through Day -1 (Baseline)"
ADECOD	SRCDOM in("ADVS" "ADEG")	text			Stack ADVS and ADEG keeping Subject (USUBJID), Analysis Visit (N) (AVISITN), Analysis Visit (AVISIT), Analysis Relative Day (ADY), Analysis Window Target (AWTARGET), Analysis Window Beginning Timepoint (AWLO), Analysis Window Ending Timepoint (AWHI) where subject has an assessment. Get rid of duplicate observations by these variables. If missing<ADY<AWLO-3 or ADY>AWHI+3 then ADECOD="Visit Out of (+/-3 Day) Window"
ASTDT	SRCDOM="DV"	text			DV.DVSTDTC converted to numeric
ASTDT	SRCDOM="IE"	text			Date of Screening Visit (IE.IEDTC) converted to numeric
ASTDT	SRCDOM="ADSL"	text			Set to TRTSDT
ASTDT	SRCDOM="ADPAIN"	text			Leave NULL
ASTDT	SRCDOM in ("ADVS" "ADEG")	text			Set to ADT

14.3 Sample ADaM Data - Example 11

Table 14.3.1 is an illustration of the Protocol Deviations analysis dataset (ADPDEV) defined in Section 12.2, Sample ADaM Variable Metadata - Example 9.

Key points to note in the example are:

1. This is a simple example to illustrate the ADPDEV protocol deviations and some other violations. It does not include all variables that could also be needed for analysis or all indicated in the metadata.
2. The ACAT1 variable indicates the category of deviation; these values are used as a summarization category.
3. The variables SRCDOM and SRCSEQ help identify the source datasets as well as traceability to SDTM or ADaM datasets. See Section 3.2.2, Identifier Variables, for a discussion of identifier variables and traceability in OCCDS.
4. The A* rather than the U* prefix is used in the creation of this dataset because some rows and variables are derived. As previously noted, the U* variable is only used when a value is a direct copy of a value from two or more SDTM variables with the same root and a different domain prefix. See Section 3.2.11, User-specified Variable Naming Conventions, for a discussion about use of the A* and U* prefixes.
5. SRCSEQ is not populated for the ADSL records because this dataset is 1 record per subject.
6. Note that SRCSEQ is populated on the ADPAIN record for subject ABC-101 because this points back to the record in ADPAIN where PARAMCD = "AKNEEPI", ASEQ = 8, and AVAL = 65. SRCSEQ is not populated on the ADPAIN record for subject ABC-102 as a count of records where PARAMCD="DKNEEPI" and $-7 \leq ADY \leq -1$ (multiple records falling within a certain window of time) is used to determine if this criteria is met.

There is flexibility in both the presence and population of timing variables due to the multiple sources, varying structure of input datasets, and deriving of rows. For example, row 1 has the analysis visit copied from the record in ADPAIN where subject = ABC-101 and ASEQ = 8. Note that ASTDT was left NULL because the record represents an average and ADT is not populated in ADPAIN. Row 3 captures a treatment duration deviation and thus is not associated with a single visit; this results in a NULL value for AVISIT and AVISITN even though the start of the deviation (ASTDT) is populated with the start of treatment.

In this example, deviations are categorized using ACAT1 as deviations could come from multiple datasets.

Table 14.3.1 Sample Deviations Data[a]

Row	USUBJID	SRCSEQ	SRCDOM	ACAT1	ATERM	ADECOD	ASTDT	AVISIT	AVISITN
1	ABC-101	8	ADPAIN	MAJOR PROTOCOL DEVIATION	Subject has an average of 66 mm on the Visual Analog Scale of Worst Knee Pain Intensity	Average of Less than 70 mm on the Visual Analog Scale of Worst Knee Pain Intensity at Baseline		Baseline	0
2	ABC-102		ADPAIN	MAJOR PROTOCOL DEVIATION	Subject has only 3 daily worst knee pain scores recorded in the source data between days -7 to days -1	Less than 4 pain scores recorded in diary on Days --7 through Day -1 (Baseline)		Baseline	0
3	ABC-104		ADSL	MAJOR PROTOCOL DEVIATION	Subject has only 26 days of treatment	Less than 28 Days of Treatment	10JAN2017		
4	ABC-112		ADSL	MAJOR PROTOCOL DEVIATION	Subject was 75% compliant with study drug	Less than 80% Overall Compliance with Study Drug	23JAN2017		
5	ABC-221	33	ADVS	MINOR PROTOCOL DEVIATION	Visit Out of (+/-3 Day) Window	Visit Out of (+/-3 Day) Window	12FEB2017	Week 4	4
6	ABC-221	21	ADEG	MINOR PROTOCOL DEVIATION	Visit Out of (+/-3 Day) Window	Visit Out of (+/-3 Day) Window	16MAR2017	Week 8	8
7	ABC-301	1	IE	MAJOR PROTOCOL DEVIATION	Male or female adult, 40 to 75 years of age, inclusive, at the time of screening	Entered but did not satisfy criteria	05JAN2017		
8	ABC-402	1	DV	MINOR PROTOCOL DEVIATION	Subject ABC-402 was missing a radiographic assessment and Kellgren-Lawrence classification of osteoarthritis in target knee (Grade 2-3)	Skipped or missed assessment/procedure	20MAR2017		

[a]Variables ending in DT are numeric dates, here shown using SAS date format date9. Other numeric date formats can be used, but care should be taken with newer date formats which might not be understood by all statistical packages.

15 Appendices

Appendix A: Glossary and Abbreviations

The following abbreviations and terms are used in this document. Additional definitions can be found in the CDISC Glossary (available at https://www.cdisc.org/standards/glossary).

ACTG	AIDS Clinical Trial Group; resource for antiviral therapeutic areas
ADaM	Analysis Data Model
ADaMIG	ADaM Implementation Guide
ADSL	(ADaM) Subject-Level Analysis Dataset
Adverse event	Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment
Analysis-ready	A dataset that contains all of the variables needed for the specific analysis, so that actual statistical test can be performed without first having to manipulate data
BDS	(ADaM) Basic Data Structure
CDISC	Clinical Data Interchange Standards Consortium
CQ	Customized MedDRA Query
Consumer	The user/reviewer/recipient of the data
CRF	Case report form (sometime *case record form*); a printed, optical, or electronic document designed to record all required information to be reported to the sponsor for each trial subject
CTCAE	Common Terminology Criteria for Adverse Events Dictionary
Dataset	A collection of structured data in a single file
Domain	A collection of data points related by a common topic, such as adverse events or demographics
HLGT	(MedDRA) High Level Group Term
HLT	(MedDRA) High Level Term
ICD	International Classification of Disease
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
LLT	(MedDRA) Lowest Level Term
MedDRA	Medical Dictionary for Regulatory Activities; global standard medical terminology designed to supersede other terminologies (e.g., COSTART, ICD9) used in the medical product development process
NCI	National Cancer Institute
NCI-CTC	NCI Common Toxicity Criteria
OCCDS	(ADaM) Structure for Occurrence Data
Occurrence analysis	The counting of subjects with a given record or term; often includes a structured hierarchy of dictionary coding categories
Producer	The originator/sender/owner/sponsor of the data
Protocol deviation	A variation from processes or procedures defined in a protocol. Deviations usually do not preclude the overall evaluability of subject data for either efficacy or safety, and are often acknowledged and accepted in advance by the sponsor. Good clinical practice recommends that deviations be summarized by site and by category as part of the report of study results so that the possible importance of the deviations to the findings of the study can be assessed (cf. *protocol violation;* see also ICH E3[5]).
Protocol violation	A significant departure from processes or procedures required by a protocol. Violations often result in data that are not deemed evaluable for a per-protocol analysis, and may require that subjects who violate the protocol be discontinued from the study (cf. *protocol deviation*).
PSN	Peripheral sensory neuropathy
PT	(MedDRA) Preferred Term
SAP	Statistical analysis plan
SDTM	Study Data Tabulation Model

SDTMIG	Study Data Tabulation Model Implementation Guide
SMQ	Standardized MedDRA Query
SOC	(MedDRA) System Organ Class
Treatment-emergent	An event that emerges during treatment having been absent pre-treatment, or worsens relative to the pre-treatment state
TTE	Time-to-event (analysis)
WHO	World Health Organization
WHO-ART	WHO Adverse Reaction Terminology
WHO Drug	WHO Drug Dictionary

Appendix B: References

1. European Medicines Agency. *ICH Topic E 2 A. Clinical Safety Data Management: Definitions and Standards for Expedited Reporting. Step 5. Note for Guidance on Clinical Safety Data Management: Definitions and Standards for Expedited Reporting* (CPMP/ICH/377/95). 1995; European Medicines Agency. Accessed October 21, 2019. https://www.ema.europa.eu/en/documents/
2. European Medicines Agency. *ICH Topic E9. Statistical Principles for Clinical Trials. Step 5. Note for Guidance on Statistical Principles for Clinical Trials* (CPMP/ICH/363/96). 1998; European Medicines Agency. Accessed October 21, 2019.https://www.ema.europa.eu/en/documents/
3. Common terminology criteria for adverse events (CTCAE). National Cancer Institute Division of Cancer Treatment and Diagnosis. Accessed November 25, 2019. https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm
4. MedDRA. *Introductory Guide for Standardised MedDRA Queries* (SMQs) Version 22.1. 2019; ICH. Accessed November 19, 2020. https://admin.new.meddra.org/sites/default/files/guidance/file/000356_smq_intguide_22_1.pdf
5. European Medicines Agency. *ICH Topic E3. Structure and Content of Clinical Study Reports. Step 5. Note for Guidance on Structure and Content of Clinical Study Reports* (CPMP/ICH/137/95). 1996; European Medicines Agency. Accessed October 21, 2019. https://www.ema.europa.eu/en/documents/

Appendix C: Revision History

This section lists all changes in the OCCDS document from Version 1.0 (Final).

Category/Section	Type	Description
Title	Update	Added Implementation Guide to the document title.
General	Update	Updated text: referring to SDTM variables using Date/Time vs when referring ADaM variables using datetime.
General	Update	Examples were reordered and expanded from previous OCCDS version.
Introduction	Update	Update introduction to include new text regarding the update from OCCDS V1.0 to V1.1. Dropped reference to ADAE document.
Introduction	Update	Added another example where records in SDTM wouldn't match the analysis dataset (i.e. --OCCUR=N records may be excluded).
3.2.2	Update	Added new Table 3.2.2.2 Row Identifier Variables to include ASEQ and SRC variables and moved --SEQ from Table 3.2.2.1 to Table 3.2.2.2.
3.2.4	Update	Changed --STDTC to conditional for consistency with --ENDTC Updated label text for ADaM datetime variables (ASTDTM, AENDTM) in Table 3.2.4.1 Updated notes for ASTDY/AENDY and imputation date and time variables to be consistent with latest ADaMIG text Updated notes that ADURU is only conditional on whether ADURN is included
3.2.5	Update	Removed a sentence in TRTEMFL and added TREMxxFL, TRTEMwFL, ONTRxxFL, and ONTRTwFL.
3.2.8	Update	Tables 3.2.8.1 and 3.2.8.2 were updated to use more generic variable names, without the AE or CM prefix.
3.2.9	Update	Table 3.2.9.1 added row for ADECODy.
3.2.11	New	Added text discussing user-specified variable naming conventions.
Example 2 Section 5.1.1	Update	Percentages and number of patients in the safety population were updated to be consistent.
Example 3	Update	Changed the variable name for Cumulative Dose Group 1 from DOSCMGR1 to DOSCUMG1 in Table 6.2.1. Updated variable metadata and sample data for DOSCUMG1 to match display layout.
Example 4	Update	Added the TREMxxFL variables to the metadata and data.
Example 5	Update	Removed reference to creating two separate datasets, one for each coding path.
Example 6	New	Added occurrence analysis dataset from multiple sources: merging (analysis of adverse events that change over time).
Example 7	New	Added occurrence analysis dataset from multiple sources: stacking (analysis of adverse events from multiple input domains).
Example 8	Update	Changed reference of adverse events to concomitant medications to match example.
Example 9	Update	Changed reference of adverse events to medical history to match example.
Example 11	New	Added analysis of protocol deviations example.

Appendix D: Representations and Warranties, Limitations of Liability, and Disclaimers

CDISC Patent Disclaimers

It is possible that implementation of and compliance with this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken with respect to the existence or validity of any claim or of any patent rights in connection therewith. CDISC, including the CDISC Board of Directors, shall not be responsible for identifying patent claims for which a license may be required in order to implement this standard or for conducting inquiries into the legal validity or scope of those patents or patent claims that are brought to its attention.

Representations and Warranties

"CDISC grants open public use of this User Guide (or Final Standards) under CDISC's copyright."

Each Participant in the development of this standard shall be deemed to represent, warrant, and covenant, at the time of a Contribution by such Participant (or by its Representative), that to the best of its knowledge and ability: (a) it holds or has the right to grant all relevant licenses to any of its Contributions in all jurisdictions or territories in which it holds relevant intellectual property rights; (b) there are no limits to the Participant's ability to make the grants, acknowledgments, and agreements herein; and (c) the Contribution does not subject any Contribution, Draft Standard, Final Standard, or implementations thereof, in whole or in part, to licensing obligations with additional restrictions or requirements inconsistent with those set forth in this Policy, or that would require any such Contribution, Final Standard, or implementation, in whole or in part, to be either: (i) disclosed or distributed in source code form; (ii) licensed for the purpose of making derivative works (other than as set forth in Section 4.2 of the CDISC Intellectual Property Policy ("the Policy")); or (iii) distributed at no charge, except as set forth in Sections 3, 5.1, and 4.2 of the Policy. If a Participant has knowledge that a Contribution made by any Participant or any other party may subject any Contribution, Draft Standard, Final Standard, or implementation, in whole or in part, to one or more of the licensing obligations listed in Section 9.3, such Participant shall give prompt notice of the same to the CDISC President who shall promptly notify all Participants.

No Other Warranties/Disclaimers. ALL PARTICIPANTS ACKNOWLEDGE THAT, EXCEPT AS PROVIDED UNDER SECTION 9.3 OF THE CDISC INTELLECTUAL PROPERTY POLICY, ALL DRAFT STANDARDS AND FINAL STANDARDS, AND ALL CONTRIBUTIONS TO FINAL STANDARDS AND DRAFT STANDARDS, ARE PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, WHETHER EXPRESS, IMPLIED, STATUTORY, OR OTHERWISE, AND THE PARTICIPANTS, REPRESENTATIVES, THE CDISC PRESIDENT, THE CDISC BOARD OF DIRECTORS, AND CDISC EXPRESSLY DISCLAIM ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR OR INTENDED PURPOSE, OR ANY OTHER WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, FINAL STANDARDS OR DRAFT STANDARDS, OR CONTRIBUTION.

Limitation of Liability

IN NO EVENT WILL CDISC OR ANY OF ITS CONSTITUENT PARTS (INCLUDING, BUT NOT LIMITED TO, THE CDISC BOARD OF DIRECTORS, THE CDISC PRESIDENT, CDISC STAFF, AND CDISC MEMBERS) BE LIABLE TO ANY OTHER PERSON OR ENTITY FOR ANY LOSS OF PROFITS, LOSS OF USE, DIRECT, INDIRECT, INCIDENTAL, CONSEQUENTIAL, OR SPECIAL DAMAGES, WHETHER UNDER CONTRACT, TORT, WARRANTY, OR OTHERWISE, ARISING IN ANY WAY OUT OF THIS POLICY OR ANY RELATED AGREEMENT, WHETHER OR NOT SUCH PARTY HAD ADVANCE NOTICE OF THE POSSIBILITY OF SUCH DAMAGES.

Note: The CDISC Intellectual Property Policy can be found at
http://www.cdisc.org/system/files/all/article/application/pdf/cdisc_20ip_20policy_final.pdf