Analysis Data Model (ADaM) Examples of Traceability

Version 1.0 (Final)

Developed by the
CDISC Analysis Data Model Team

Notes to Readers

- This is the final Version 1.0 of the Analysis Data Model (ADaM) Examples of Traceability.
- This document is based on the principles, structures, and standards described in the CDISC Analysis Data Model Version 2.1 and Implementation Guide v1.2.

Revision History

Date	Version
2022-05-12	1.0 Final

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

CONTENTS

1 Introduction

1.1 Purpose

The Analysis Data Model Implementation Guide (ADaMIG) Version 1.2 (available at https://www.cdisc.org/standards/foundational/adam) states in Section 2.2:

> To assist review, ADaM datasets and metadata must clearly communicate how the ADaM datasets were created. The verification of derivations in an ADaM dataset requires having at hand the input data used to create the ADaM dataset. A CDISC-conformant submission includes both SDTM and ADaM datasets; therefore, it follows that the relationship between SDTM and ADaM must be clear. This requirement highlights the importance of traceability between the analyzed data (ADaM) and its input data (SDTM).

Traceability is built by clearly establishing the path between an element and its immediate predecessor. The full path is traced by going from one element to its predecessors, then on to their predecessors, and so on, back to the SDTM datasets, and ultimately to the data collection instrument.

The objective of this document is to provide examples of traceability using the Analysis Data Model (ADaM). See ADaM Version 2.1, ADAMIG Version 1.2, and the ADaM Structure for Occurrence Data (OCCDS) Version 1.0 (https://www.cdisc.org/standards/foundational/adam) for required background about ADaM and the ADaM data structures.

1.2 The Need for Traceability in ADaM

Clinical studies are conducted to test the safety and effectiveness of new drugs and therapies. Data collected from study participants is analyzed, and the results submitted to regulatory agencies and released to the public. To ensure the results are robust and verifiable, the steps by which the collected data is processed into the analysis results should be clearly documented. This documentation is known as *traceability*.

Why is this important? Suppose a new vaccine shows promise, as illustrated in Table 1.2.1.

Table 1.2.1. Sample Efficacy Table

Table xx.x Primary Efficacy Endpoint (ITT Population)

	Drug n (%) (N=8000)	Control n (%) (N=8000)	Odds Ratio	P-Value
Occurrence of primary study disease at 2 years	8 (0.1%)	64 (0.8%)	0.1241	< 0.0001

This result table shows the study drug reduced the occurrence rate of disease by >80% over the control, a clear improvement. Due to the importance of this analysis in the submission, a reviewer decides to verify the computation of this result and starts by reviewing the provided traceability documentation. The define.xml analysis results metadata in Table 1.2.2 shows 1 type of metadata reviewers may find useful (the Define.xml Specification is available at https://www.cdisc.org/standards/data-exchange/define-xml).

Table 1.2.2. Sample Analysis Results Metadata

Display	Table xx.x Primary Efficacy Endpoint
Analysis result	Occurrence of Primary Study Disease at 2 Years
Analysis parameter(s)	PARAMCD = "PRI" (Primary Efficacy Endpoint)
Analysis variable(s)	AVAL (Analysis Value)
Analysis reason	SPECIFIED IN SAP
Analysis purpose	PRIMARY OUTCOME MEASURE
Data references (incl. selection criteria)	ADEF [PARAMCD = "PRI" and ITTFL = "Y"]
Documentation	SAP Section 4.1
Programming statements	[SAS Version 9.4] proc freq data=adef(where=(ittfl='Y' and paramcd='PRI')); table trt01pn*aval; exact or; run;

The metadata identifies the analysis dataset used as ADEF along with the necessary subset conditions. The SAS code snippet allows quick verification of table values using submitted data. This is an example of *metadata traceability*. Table 1.2.3 shows a slice of the dataset ADEF.

Table 1.2.3. Sample ADEF Records

USUBJID	SRCDOM	SRCSEQ	PARAMCD	PARAM	AVAL	AVALC	ITTFL
XYZ-01-001	PF	2	PRI	Primary Efficacy Endpoint	0	DISEASE	Y
XYZ-01-002	LB	52	PRI	Primary Efficacy Endpoint	0	DISEASE	Y
XYZ-01-003			PRI	Primary Efficacy Endpoint	1	NO DISEASE	Y
XYZ-01-004			PRI	Primary Efficacy Endpoint	1	NO DISEASE	Y

After verifying that the odds ratio and *P* value are calculated correctly using the provided ADaM dataset, the reviewer continues to verify that the parameter itself is derived correctly. To facilitate this part of the review, Table 1.2.4 provides variable metadata for the ADEF dataset.

Table 1.2.4. Variable Metadata for ADEF

ADEF Variable Metadata

Name	Variable Label	Variable Metadata
USUBJID	Unique Subject Identifier	ADSL.USUBJID
SRCDOM	Source Data	If AVAL=0, identify whether the corresponding record is from PF or LB SDTM domain
SRCSEQ	Source Sequence	If AVAL=0, copy over the corresponding PFSEQ or LBSEQ value from the corresponding record
PARAMCD	Parameter Code	Set to "PRI"
PARAM	Parameter	Set to " Primary Efficacy Endpoint"
AVAL	Analysis Value	If subject has a biopsy record in PF where PFTEST="BIOMARKER 1" and PFSTRESC="PRESENT" then set AVAL=0. Else if subject does not have any biopsy records in PF and has an enzyme record in LB where LBTEST="ENZYME A" and LBSTRESC="POSITIVE" then set AVAL=0. (note: if a biopsy record is present, do not check enzyme test records) Otherwise set AVAL=1 Refer to SAP section 4.1 for more details
AVALC	Analysis Value (C)	If AVAL=0 then set AVALC="DISEASE" If AVAL=1 then set AVALC="NO DISEASE"
ITTFL	Intent-To-Treat Population Flag	ADSL.ITTFL

Table 1.2.4 provides the derivation of the primary efficacy parameter in the variable AVAL, as another example of metadata traceability. In addition, when study disease is identified for a subject, the variables SRCDOM and SRCSEQ identify the exact record in the SDTM that led to this determination. This is an example of *data point traceability*. With this, the data lineage from the efficacy table to SDTM is complete, allowing a reviewer to fully reproduce the presented results table.

In the case that the traceability between analysis results and SDTM is incomplete, the reviewer may have to decrypt the code in submitted analysis programs or hold review question-and-answers cycles with the sponsor, both of which take time. Providing both metadata and data point traceability enables quick and efficient reviews of analysis data and results and thus is a cornerstone of a quality submission.

1.3 Points to Consider When Interpreting this Document

In reviewing the metadata and examples in this document, consider the following:

- **Optimum number of analysis datasets.** As stated in ADaM (https://www.cdisc.org/standards/foundational/adam), one goal in creating analysis datasets is to have the optimum number of analysis datasets needed to perform the various analyses. A single dataset can support multiple analyses. Examples 2.1, General ADSL Traceability, 2.2, General BDS Traceability, 2.3, General OCCDS Traceability, and 2.4, Traceability with Parameters from Multiple Input Datasets, illustrate a single analysis dataset used to support multiple statistical analyses. It should also be noted that the same analysis dataset can be used to generate descriptive statistics as illustrated in Example 2.4.
- **Ordering of variables.** Within this document, no specific ordering of variables within the illustrated datasets is applied. ADaM v2.1 states that the ordering of the variables in the analysis dataset should follow a logical ordering (not simply alphabetic), but does not provide a specific recommendation for ordering 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 in 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 v2.1. For example, in the ADaM Subject-level Analysis Dataset (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.
- **Parameter value-level metadata.** Throughout this document, parameter value-level metadata are included for ADaM Basic Data Structure (BDS) analysis datasets, as required in variable-level metadata for a BDS analysis dataset (as defined in ADaM). ADSL analysis variable metadata do not include parameter value-level metadata, as ADSL is a single record per subject structure. The OCCDS does not include the concept of parameters, so it also has no parameter-level metadata.
- **Display format metadata element.** It should be noted that the metadata element, display format, may be the attributed format of the variable contained in the analysis dataset or the display format for the associated statistical table. A good example of this is in Table 2.7.3.3 regarding AVAL. It is possible for display format to vary within the same variable per Value-Level Metadata and differ between the dataset variable and statistical output.
- **Submission-ready metadata.** Although much of the metadata described in this document could be used for submission, this document does not attempt to convey how to submit. For example, it doesn't speak to which directory in the FDA folder structure a look-up table belongs. Submission rules are developed and maintained by regulatory agencies, and questions about these rules should be directed to the regulatory agency that is responsible for that document.
- **Analysis-ready.** The 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. In addition to required variables such as subject identifiers and treatment variables, the critical variables included in the analysis dataset will depend on the specific nature of the disease or indication, the analyses planned in the protocol, and the statistical analysis plan (SAP), and may include:
  - Baseline values
  - Stratification or grouping variables
  - Selection flags (e.g., population flag)
  - Predictor variables (also known as explanatory variables, independent variables, or covariates)
  - Response variables (also known as dependent variables)
  - Supportive variables for complex predictors and/or responses
  - Supportive variables to facilitate traceability

Refer to the ADaMIG (https://www.cdisc.org/standards/foundational/adam/) for descriptions of issues to be considered in designing and constructing analysis datasets and the associated metadata. Examples of these issues include (but are not limited to):

- Identification of baseline
- Facilitating subgroup analyses
- Inclusion of transformed analysis values
- Deriving composite/compound parameters
- Identifying imputation of missing values
- Deriving a new variable (column) vs. a new record
- Inclusion of additional records and/or variables for sensitivity analyses and/or future analyses

1.4 Caveats and Disclaimers

- **Examples are for illustration only.** Note that 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 is for illustration of content only.** Although the metadata elements have been defined in ADaM v2.1 (https://www.cdisc.org/standards/foundational/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. In addition, the metadata examples include only 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 are not included.** Analysis results metadata are not included for any examples in this document. As stated in ADaM v2.1, 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 are not meant to be 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; for example, core variables required for subgroup analyses are not included in all illustrations.
- **The Source/Derivation Column algorithms are for illustration only.** 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.
- **No 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.

1.5 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 Examples

2.1 General ADSL Traceability

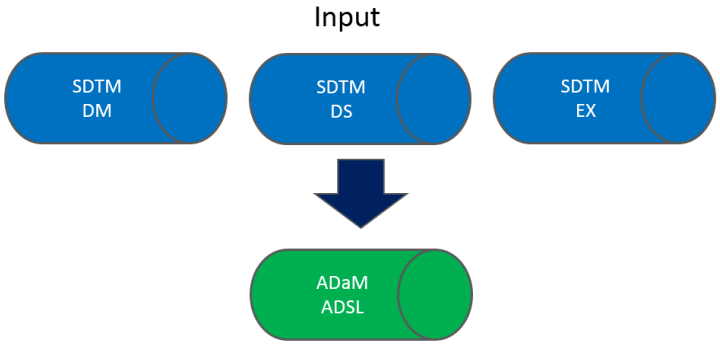
This example shows the traceability for common ADSL variables.

2.1.1 Analysis Need 1

The ADSL dataset includes important subject-level variables both to support analysis and to fulfill certain core variable requirements from regulatory agencies. Common ADSL variables include some copied from SDTM and others derived within the ADSL dataset.

2.1.2 Data Flow 1

Figure 2.1.2.1. Example General ADSL Data Flow



2.1.3 Traceability Metadata 1

Table 2.1.3.1 shows an example of variable-level metadata traceability for each ADSL variable. In this example, AAGEGR1 was created to serve the analysis of grouping subjects by age categories. The traceability from AAGEGR1 all the way back to the SDTM variable DM.BRTHDTC was shown in several steps:

1. deriving AAGE as the actual age used for the grouping,
2. creating BRTHDT to calculate AAGE, and
3. keeping the predecessor of DM.BRTHDTC to show how it was imputed to BRTHDT.

Table 2.1.3.1. Variable-level Metadata

ADSL Variable Metadata

Variable Name	Variable Metadata
STUDYID	Predecessor: DM.STUDYID
USUBJID	Predecessor: DM.USUBJID
SUBJID	Predecessor: DM.SUBJID
SITEID	Predecessor: DM.SITEID
SEX	Predecessor: DM.SEX
RACE	Predecessor: DM.RACE
AGE	Predecessor: DM.AGE
AGEU	Predecessor: DM.AGEU
BRTHDTC	Predecessor: DM.BRTHDTC
BRTHDT	Derived: Numeric version of DM.BRTHDTC. If only month and year are collected, impute day to 15; else if only year is collected, impute month to 07 and day to 01; else if missing, do not impute.
BRTHDTF	Derived: If only day is imputed, set to 'D'; else if both day and month are imputed, set to 'M'. Missing when no imputation is done.
RANDDT	Derived: Numeric version of DS.DSSTDTC when DS.DSTERM = 'RANDOMIZED'. If any part of the date is missing, do not impute.
AAGE	Derived: YRDIF(BRTHDT, RANDDT, 'AGE'). Missing if either date is missing.
AAGEGR1	Derived: If AAGE is missing then AAGEGR1 is missing; else if AAGE <41 then set to "< 41"; else if AAGE < 61 then set to "41-60"; else set to "61 or older".
ARM	Predecessor: DM.ARM

Variable Name	Variable Metadata
ARMCD	Predecessor: DM.ARMCD
TRT01P	Derived: Set to the first component of DM.ARM before "-". Leave as null if DM.ARMCD = "SCRNFAIL' or "NOTASSGN".
TRT02P	Derived: If there are two components in DM.ARM separated by "-", set to the second component of DM.ARM. Otherwise, leave as null.
TRTSEQP	Derived: Set to ADSL.TRT01P || "-" || ADSL.TRT02P. If ADSL.TRT02P is null, set to ADSL.TRT01P.
TRTSDT	Derived: Numeric version of the earliest EX.EXSTDTC. If any part of the date is missing, do not impute.
TRTEDT	Derived: Numeric version of the last EX.EXENDTC. If any part of the date is missing, do not impute.
TR01SDT	Derived: Numeric version of the earliest EX.EXSTDTC when EX.EPOCH="DOUBLE-BLIND TREATMENT". If any part of the date is missing, do not impute.
TR01EDT	Derived: Numeric version of the last EX.EXENDTC when EX.EPOCH="DOUBLE-BLIND TREATMENT". If any part of the date is missing, do not impute.
TR02SDT	Derived: Numeric version of the earliest EX.EXSTDTC when EX.EPOCH="OPEN-LABEL TREATMENT". If any part of the date is missing, do not impute.
TR02EDT	Derived: Numeric version of the last EX.EXENDTC when EX.EPOCH="OPEN-LABEL TREATMENT". If any part of the date is missing, do not impute.

2.1.4 Input and Analysis Data 1

The following example shows the input SDTM domains, including Demographics (DM), Disposition (DS), and Exposure (EX), and how they are used to create ADSL. Only variables needed for illustration are shown in the examples.

Table 2.1.4.1. Sample DM Data

Row	STUDYID	USUBJID	SUBJID	SITEID	SEX	RACE	AGE	AGEU	BRTHDTC	ARM	ARMCD
1	ABC123	ABC12301001	001	01	M	WHITE		YEARS	1958-12	Drug A - Drug B	AB
2	ABC123	ABC12301002	002	01	F	ASIAN	40	YEARS	1975-05-10	Placebo - Drug B	PB
3	ABC123	ABC12302003	003	02	M	WHITE	53	YEARS	1963-09-03	Drug A	A

Table 2.1.4.2. Sample DS Data

Row	USUBJID	DSTERM	DSSTDTC
1	ABC12301001	RANDOMIZED	2016-05-17
2	ABC12301002	RANDOMIZED	2016-02-07
3	ABC12302003	RANDOMIZED	2016-10-25

Table 2.1.4.3. Sample EX Data

Row	USUBJID	EXSEQ	EXTRT	EXDOSE	EXDOSEU	EXSTDTC	EXENDTC	EPOCH
1	ABC12301001	1	Drug A	50	mg	2016-05-24	2016-07-22	DOUBLE-BLIND TREATMENT
2	ABC12301001	2	Drug B	100	mg	2016-08-01	2017-01-30	OPEN-LABEL TREATMENT
3	ABC12301002	1	Placebo	0	mg	2016-02-15	2016-04-06	DOUBLE-BLIND TREATMENT
4	ABC12301002	2	Drug B	100	mg	2016-04-25	2016-10-28	OPEN-LABEL TREATMENT
5	ABC12302003	1	Drug A	50	mg	2016-11-01	2016-11-29	DOUBLE-BLIND TREATMENT

Table 2.1.4.4. Sample ADSL Data

Row	STUDYID	USUBJID	SUBJID	SITEID	SEX	RACE	AGE	AGEU	BRTHDTC	BRTHDT	BRTHDTF	RANDDT	AAGE	AAGEGR1	ARM	ARMCD	TRT01P	TRT02P	TRTSEQP	TRTSDT	TRTEDT	TR01SDT	TR01EDT	TR02SDT	TR02EDT
1	ABC123	ABC12301001	001	01	M	WHITE		YEARS	1958-12	15DEC1958	D	17MAY2016	57	41-60	Drug A - Drug B	AB	Drug A	Drug B	Drug A - Drug B	24MAY2016	30JAN2017	24MAY2016	22JUL2016	01AUG2016	30JAN2017
2	ABC123	ABC12301002	002	01	F	ASIAN	40	YEARS	1975-05-10	10MAY1975		07FEB2016	40	<41	Placebo - Drug B	PB	Placebo	Drug B	Placebo - Drug B	15FEB2016	28OCT2016	15FEB2016	06APR2016	25APR2016	28OCT2016
3	ABC123	ABC12302003	003	02	M	WHITE	53	YEARS	1963-09-03	03SEP1963		25OCT2016	53	41-60	Drug A	A	Drug A		Drug A	01NOV2016	29NOV2016	01NOV2016	29NOV2016		

2.1.5 Other Uses 1

Variable-level metadata is useful for variables that are copied or derived from easily referenced data.

Because ADSL is 1 record per subject, there is no opportunity to include variables such as sequence number to provide data point traceability. For many ADSL variables, such as those mentioned in this section, variable-level traceability is sufficient. Section 2.9, Using an Intermediate Dataset for ADSL Traceability, provides an example that includes an intermediate dataset prior to ADSL in order to provide additional derivation and traceability information.

2.2 General BDS Traceability

The BDS structure provides the ability to group analysis by parameters as well as by timepoints. BDS datasets often contain additional rows and columns to support needed analyses, as well as to show traceability.

The BDS dataset structure contains 1 analysis value (AVAL or AVALC) per record, and descriptor variables to enable analysis. The BDS structure contains 1 or more records per subject, per analysis parameter, per analysis timepoint. In situations where there is no analysis timepoint, the structure is 1 or more records per subject per analysis parameter (see the ADaMIG; https://www.cdisc.org/standards/foundational/adam/).

In practice, the BDS can be used to analyze data from an SDTM domain, such as Laboratory Test Results (LB), Vital Signs (VS), and ECG Test Results (EG). It can also be used to analyze data from multiple SDTM domains for complex analysis, such as for time-to-event, efficacy, and derived safety analysis.

2.2.1 Analysis Need 2

One of the safety tables requested for a study is summary statistics of vital signs measurements at baseline, week 24, week 48, and week 52, along with change from baseline summary statistics for post-baseline visit. Table 2.2.1.1 provides a sample shell. To support this table, an analysis dataset ADVS will be created.

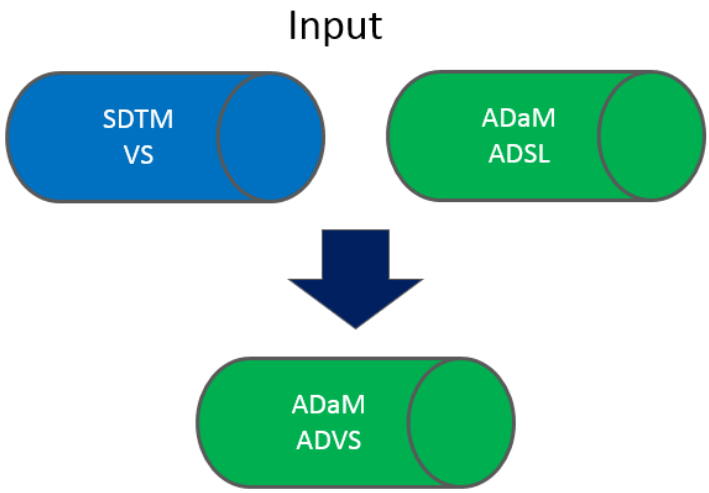
Table 2.2.1.1. Sample Table Shell

Table x.y.z Summary of Vital Signs Safety Population				
	Treatment A		Treatment B	
	(N=xxx)		(N=xxx)	
Weight (kg)	Measured Values	Change from Baseline	Measured Values	Change from Baseline
Baseline				
n Mean (SD) Median Q1, Q3 Min, Max	xx xxx.x (xx.xx) xxx.x xxx.x, xxx.x xxx, xxx		xx xxx.x (xx.xx) xxx.x xxx.x, xxx.x xxx, xxx	
...				
Week 24				
n Mean (SD) Median Q1, Q3 Min, Max	xx xxx.x (xx.xx) xxx.x xxx.x, xxx.x xxx, xxx	xx xxx.x (xx.xx) xxx.x xxx.x, xxx.x xxx, xxx	xx xxx.x (xx.xx) xxx.x xxx.x, xxx.x xxx, xxx	xx xxx.x (xx.xx) xxx.x xxx.x, xxx.x xxx, xxx

2.2.2 Data Flow 2

General ADVS data flow where ADSL is merged with VS to create ADVS.

Figure 2.2.2.1. Example General BDS Data Flow



2.2.3 Traceability Metadata 2

Table 2.2.3.1. Dataset Metadata for ADVS

ADVS Dataset Metadata

Dataset Name	Dataset Description	Dataset Structure	Class of Dataset
ADVS	Vital Signs Analysis Dataset	One record per subject per visit per parameter	BASIC DATA STRUCTURE

Table 2.2.3.2. Variable Metadata for ADVS

ADVS Variable Metadata

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Source/Derivation/Comment
STUDYID	Study Identifier	text		VS.STUDYID
USUBJID	Unique Subject Identifier	text		VS.USUBJID
TRTA	Actual Treatment	text		ADSL.TRT01A
TRTSDT	Date of First Exposure to Treatment	integer		ADSL.TRTSDT
PARAM	Parameter	text		VS.VSTEST (VS.VSSTRESU) and create derived parameter "Log10(Weight (kg))" based on Weight (kg) parameter
PARAMCD	Parameter Code	text		VS.VSTESTCD and "L10WT" for derived parameter "Log10(Weight (kg))"
ADT	Analysis Date	integer		Derived: Numeric representation of VS.VSDTC
AVISIT	Analysis Visit	text		propcase(VS.VISIT)
AVISITN	Analysis Visit (N)	integer	AVISITN	Assigned based on Codelist
VSSEQ	Sequence Number	integer		VS.VSSEQ
VSSTRESN	Numeric Result/Finding in Standard Units	integer		VS.VSSTRESN
ABLFL	Baseline Record Flag	text	["Y" = "Yes"] <Y-NULL>	Derived: Per SAP, baseline is defined as the last non-missing value before the first dose of study drug. ABLFL="Y" for last non missing assessment obtained before the drug administration (ADT<ADSL.TRTSDT), per usubjid, paramcd
AVAL	Analysis Value	float		VS.VSSTRESN and for derived parameter "Log10(Weight (kg))" take log10 of VS.VSSTRESN
BASE	Baseline Value	float		Derived: AVAL where ABLFL="Y"
CHG	Change from Baseline	float		Derived: AVAL-BASE, populate only for post-baseline records

2.2.4 Input and Analysis Data 2

Table 2.2.4.1. Sample VS Data

Row	STUDYID	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSSTRESN	VSSTRESU	VISITNUM	VISIT	VSDTC
1	XYZ	XYZ-001-001	1164	WEIGHT	Weight	99	kg	1	Screening	2018-03-19
2	XYZ	XYZ-001-001	1165	WEIGHT	Weight	101	kg	2	Run-In	2018-03-26
3	XYZ	XYZ-001-001	1166	WEIGHT	Weight	100	kg	3	Baseline	2018-04-16
4	XYZ	XYZ-001-001	1167	WEIGHT	Weight	94	kg	4	Week 24	2018-09-30
5	XYZ	XYZ-001-001	1168	WEIGHT	Weight	92	kg	5	Week 48	2019-03-17
6	XYZ	XYZ-001-001	1169	WEIGHT	Weight	95	kg	6	Week 52	2019-04-14

Table 2.2.4.2. Sample ADVS Data

Row	STUDYID	USUBJID	TRTA	TRTSDT	PARAM	PARAMCD	ADT	AVISIT	AVISITN	VSSEQ	VSSTRESN	ABLFL	AVAL	BASE	CHG
1	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Weight (kg)	WEIGHT	2018-03-19	Screening	-4	1164	99		99	100	
2	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Weight (kg)	WEIGHT	2018-03-26	Run-In	-2	1165	101		101	100	
3	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Weight (kg)	WEIGHT	2018-04-16	Baseline	0	1166	100	Y	100	100	
4	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Weight (kg)	WEIGHT	2018-09-30	Week 24	24	1167	94		94	100	-6
5	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Weight (kg)	WEIGHT	2019-03-17	Week 48	48	1168	92		92	100	-8
6	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Weight (kg)	WEIGHT	2019-04-14	Week 52	52	1169	95		95	100	-5
7	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Log10(Weight (kg))	LOGWT	2018-03-19	Screening	-4	1164	99		1.9956	2	
8	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Log10(Weight (kg))	LOGWT	2018-03-26	Run-In	-2	1165	101		2.0043	2	
9	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Log10(Weight (kg))	LOGWT	2018-04-16	Baseline	0	1166	100	Y	2	2	
10	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Log10(Weight (kg))	LOGWT	2018-09-30	Week 24	24	1167	94		1.9731	2	-0.0269
11	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Log10(Weight (kg))	LOGWT	2019-03-17	Week 48	48	1168	92		1.9638	2	-0.0362
12	XYZ	XYZ-001-001	Treatment B	16-Apr-19	Log10(Weight (kg))	LOGWT	2019-04-14	Week 52	52	1169	95		1.9777	2	-0.0223

Note: In the transformed parameter, using the VSSEQ value enables quickly identifying the source SDTM record. Also, by keeping the VSSTRESN value, the computation for AVAL can be quickly verified.

2.2.5 Other Uses 2

This BDS ADVS example demonstrated how each BDS record can retain SDTM variables to identify source SDTM records and help verify derivations. The BDS data structure within ADaM is extremely flexible; this document contains many examples of how BDS datasets can support analysis while retaining traceability to SDTM (and ADaM) source records.

2.3 General OCCDS Traceability

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 for occurrence analysis dataset is usually 1 record per record in the corresponding SDTM domain.

2.3.1 Analysis Need 3

The basic summary of adverse event frequencies described in Section 12.2.2 of the ICH Guideline E3[1] report 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.

Analysis display example layout:

Table 2.3.1.1. Summary of Treatment-emergent Adverse Events

Table x.x Summary of Treatment Emergent Adverse Events by System Organ Class and Preferred Term Analysis Population Safety		
System Organ Class	Treatment A	Treatment B
Preferred Term	(N=xxx)	(N=xxx)
	n (%)	n (%)
Number of subjects reporting at least one treatment-emergent 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)
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)
<Other SOCs and PTs>		

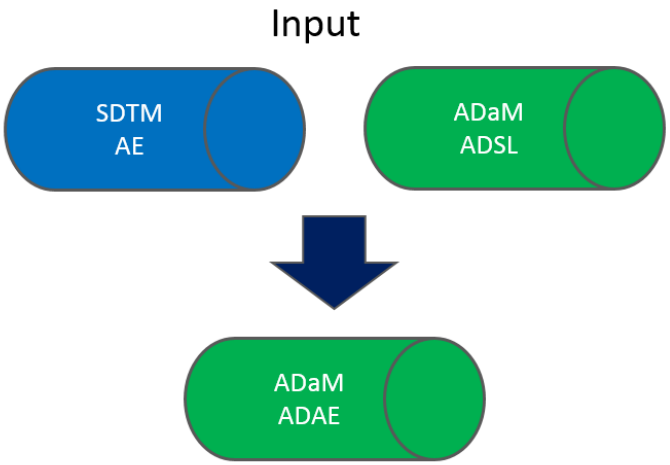
N = Safety subjects, i.e., subjects who received at least 1 dose of study drug
n = Number of subjects reporting at least 1 treatment-emergent adverse event
% = *n* / *N* * 100

In this example, the source of adverse events occurrence data for oncology is to be analyzed; the variables derived to describe the actual treatment at the time an adverse event took place have also been illustrated.

2.3.2 Data Flow 3

General ADAE data flow can be created by merging ADaM.ADSL and SDTM.AE. Figure 2.3.2.1 illustrates the concept.

Figure 2.3.2.1. Example Data Flow



2.3.3 Traceability Metadata 3

Table 2.3.3.1. Dataset Metadata for ADAE

ADAE Dataset Metadata

Dataset Name	Dataset Description	Dataset Structure	Class of Dataset
ADAE	Adverse Events Analysis Dataset	One record per record in the SDTM AE domain (USUBJID AETERM ASTDT AENDT AESEQ).	OCCURRENCE DATA STRUCTURE

Table 2.3.3.2. Variable Metadata for ADAE

ADAE Variable Metadata

Name	Variable Label	Variable Metadata
STUDYID	Study Identifier	AE.STUDYID
USUBJID	Unique Subject Identifier	AE.USUBJID
AESEQ	Sequence Number	AE.AESEQ
AETERM	Reported Term for the Adverse Event	AE.AETERM
AEDECOD	Dictionary-Derived Term	AE.AEDECOD MedDRA[2] Version 11.1
AEBODSYS	Body System or Organ Class	AE.AEBODSYS MedDRA Version 11.1
TRTEMFL	Treatment Emergent Analysis Flag	If ADSL.TRTSDT <= ASTDT<=(ADSL.TRTEDT +14) then TRTEMFL="Y"
AESTDTC	Start Date/Time of Adverse Event	AE.AESTDTC

Name	Variable Label	Variable Metadata
ASTDT	Analysis Start Date	Numeric version by converting AE.AESTDTC from character ISO8601 format to SAS format, applying imputation rules as specified in the SAP or metadata.
ASTDTF	Analysis Start Date Imputation Flag	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"
AEENDTC	End Date/Time of Adverse Event	AE.AEENDTC
AENDT	Analysis End Date	Numeric version by converting AE.AEENDTC from character ISO8601 format to SAS format, applying imputation rules as specified in the SAP or metadata.
AENDTF	Analysis End Date Imputation Flag	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"
AESER	Serious Event	AE.AESER
APHASE	Phase	If ASTDT<ADSL.TRTSDT, then APHASE="PRE-TREATMENT" Else if ASTDT > ADSL.TRTEDT + 14 days then APHASE="FOLLOW-UP", Else APHASE="TREATMENT"
AESEV	Severity/Intensity	AE.AESEV
ASEV	Analysis Severity/Intensity	If AE.AESEV="MILD" then ASEV="Mild" Else if AE.AESEV="MODERATE" then ASEV="Moderate" Else if AE.AESEV="SEVERE" or Severity/Intensity is missing then ASEV="Severe"
ASEVN	Analysis Severity/Intensity (N)	Map ASEV to ASEVN in the following manner: Mild = 1 Moderate = 2 Severe = 3
AEREL	Causality	AE.AEREL
RELGR1	Pooled Causality Group 1	If AE.AEREL="NOT RELATED" or "UNLIKELY RELATED" then RELGR1="Not Related" Else if AE.AEREL="POSSIBLY RELATED" or "PROBABLY RELATED" or "DEFINITELY RELATED" or Causality is missing then RELGR1="Related"
RELGR1N	Pooled Causality Group 1 (N)	Map RELGR1 to RELGR1N in the following manner: Not Related = 0 Related = 1
SAFFL	Safety Population Flag	ADSL.SAFFL
AOCCFL	1st Occurrence within Subject Flag	Looking at only TRTEMFL="Y" records Sort by Subject (USUBJID), Analysis Start Date (ASTDT), and Sequence Number (AESEQ) and flag the first record (set AOCCFL="Y") within each Subject
AOCCSFL	1st Occurrence of SOC Flag	Looking at only TRTEMFL="Y" records 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
AOCCPFL	1st Occurrence of Preferred Term Flag	Looking at only TRTEMFL="Y" records 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
TRTA	Actual Treatment	ADSL.TRT01A
TRTAN	Actual Treatment (N)	ADSL.TRT01AN Treatment A = 1 Treatment B = 2
TRTSDT	Date of First Exposure to Treatment	ADSL.TRTSDT
TRTEDT	Date of Last Exposure to Treatment	ADSL.TRTEDT
AGE	Age	ADSL.AGE
AGEGR1	Pooled Age Group 1	ADSL.AGEGR1
SEX	Sex	ADSL.SEX
RACE	Race	ADSL.RACE

2.3.4 Input and Analysis Data 3

This ADAE SAS dataset is the implementation of the ADAE metadata and each variable value here should reflect the definitions in Section 2.3.3, Traceability Metadata 3. The variables STUDYID, USUBJID, AESEQ, AETERM, AEDECOD, AEBODSYS, AESTDTC, AEENDTC, AESER, AESEV, and AEREL are copied from and offer traceability to the SDTM AE dataset.

Table 2.3.4.1. ADAE Sample Records

Row	STUDYID	USUBJID	AESEQ	AETERM	AEDECOD	AEBODSYS	TRTEMFL	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		2006-01	01JAN2006	D	2006-01-22	22JAN2006		N	PRE-TREATMENT	MILD	Mild	1	NOT RELATED	Not Related	0	Y				Treatment A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN
2	XYZ	XYZ-001-001	2	CHRONIC BACK PAIN	Back pain	Musculoskeletal and connective tissue disorders		2006-01-21	21JAN2006		2006-01-28	28JAN2006		N	PRE-TREATMENT	MODERATE	Moderate	2	NOT RELATED	Not Related	0	Y				Treatment A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN
3	XYZ	XYZ-001-001	3	NOSE BLEEDING RIGHT NOSTRIL	Epistaxis	Respiratory, thoracic and mediastinal disorders		2006-01-22	22JAN2006		2006-01-22	22JAN2006		N	PRE-TREATMENT	MILD	Mild	1	NOT RELATED	Not Related	0	Y				Treatment 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	Treatment 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	POSSIBLY RELATED	Related	1	Y		Y	Y	Treatment A	1	23JAN2006	15MAY2006	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				Treatment A	1	23JAN2006	15MAY2006	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	Treatment A	1	23JAN2006	15MAY2006	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	Treatment A	1	23JAN2006	15MAY2006	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				Treatment A	1	23JAN2006	15MAY2006	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	Treatment A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN
11	XYZ	XYZ-001-001	11	NAUSEA (INTERMITTENT)	Nausea	Gastrointestinal disorders	Y	2006-04-20	20APR2006		2006-04-22	2APR2006		N	TREATMENT	MILD	Mild	1	PROBABLY RELATED	Related	1	Y			Y	Treatment A	1	23JAN2006	15MAY2006	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	Treatment A	1	23JAN2006	15MAY2006	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				Treatment A		23JAN2006	15MAY2006	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				Treatment A	1	23JAN2006	15MAY2006	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				Treatment A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN
16	XYZ	XYZ-001-001	16	HEADACHE	Headache	Nervous system disorders		2006-06-01	01JUN2006		2006-06-01	01JUN2006		N	FOLLOW-UP	MILD	Mild	1	UNLIKELY RELATED	Not Related	0	Y				Treatment A	1	23JAN2006	15MAY2006	54	<65	M	ASIAN

2.3.5 Other Uses 3

The OCCDS ADAE dataset in Table 2.3.4.1 demonstrates how each occurrence record can be traced back to SDTM source through the use of variable metadata as shown in Table 2.3.3.1, and through the use of data point traceability in variables such as AESEQ. The final dataset is ADaM ADAE. Other uses can be expanded by adding new variables to this ADAE if the analysis requires, for example, to add toxicity grade flag by period (as long as the added variable naming convention follows the ADaMIG).

2.4 Traceability with Parameters from Multiple Input Datasets

During review of the statistical analysis plan and table and figure shells, it may indicate that a particular endpoint(s) may come from multiple data sources. This example illustrates when an ADaM dataset comes from multiple SDTM domains. Although Figure 2.4.2.1 shows the 3 SDTM domains as the inputs to an ADaM dataset, another ADaM dataset could be used as an input.

2.4.1 Analysis Need 4

The use of multiple input datasets for the creation of an ADaM dataset is dependent on the analysis needs. For example, if time to an event is needed, then it is possible that the data can be captured in multiple domains as shown in the data flow. ADaMIG Section 4.4 includes an example of a time-to-event for hypertension analysis dataset with the input from several SDTM domains. The example is expanded in order to illustrate how traceability can be maintained. In the example, an analysis parameter defines the study day of a hypertension event as the earliest study day among the following events:

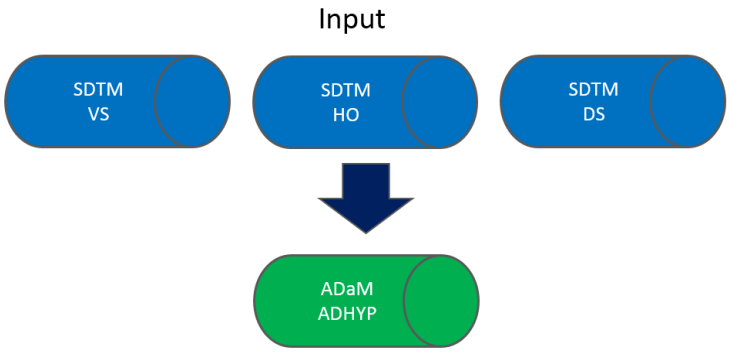
- Hospitalization
- Systolic blood pressure exceeds 140
- Diastolic blood pressure exceeds 90

If the subject did not have the event, then the subject is censored based on final study disposition. Figure 2.4.2.1 illustrates the 3 input datasets used to create the hypertension time-to-event dataset.

2.4.2 Data Flow 4

Data for this example time-to-event analysis dataset (ADHYP) is from at least SDTM Healthcare Encounters (HO), Vital Signs (VS), and Disposition (DS). Figure 2.4.2.1 demonstrates that data flow. It is possible for data to come from additional input sources, but for this illustration we are only looking at input data that is used to derive the particular parameter(s) of interest. The diagram does not illustrate any data coming from ADSL.

Figure 2.4.2.1. Multiple Input Data Flow



Note: In the data flow, only the datasets that will be used in the creation of the parameters and analysis value are included. ADSL and/or other ADaM datasets that contain treatment variables and other covariates may also be included as input in the creation of ADHYP. These datasets do not have an effect on the derivation of each parameter and are not included.

2.4.3 Traceability Metadata 4

Although the primary focus of the analysis is the hypertension event, the sub-events (i.e., hospitalization, diastolic blood pressure > 90, systolic blood pressure > 140) are also included to support traceability and to allow for future analysis by sub-event if necessary.

Table 2.4.3.1. Variable Metadata for ADHYP

ADHYP Variable Metadata

Variable Name	Variable Label	Type	Codelist/Controlled Terms	Variable Metadata
STUDYID[a]	Study Identifier	Char		xx.STUDYID that corresponds to input dataset used for USUBJID.
USUBJID	Unique Subject Identifier	Char		VS.USUBJID if subject had either systolic blood pressure > 140 and/or diastolic blood pressure > 90 HO.USUBJID if subject is hospitalized DS.USUBJID if subject did not have hospitalization or hypertension

Variable Name	Variable Label	Type	Codelist/Controlled Terms	Variable Metadata
Treatment Variable[a]	Treatment Variable Label	Char		Treatment/Treatment Group that will be used in the analysis. At least one treatment variable needs to be included in the dataset.
PARAM	Parameter	Char	Time to First Hospital Admission (day); Time to First DBP > 90 (day); Time to First SBP > 140 (day); Time to Hypertension Event (day)	Time to Hypertension Event parameter is used for the analysis of time to hypertension. The other 'Time to' parameters specified in the codelist are sub-events.
PARAMCD	Parameter Code	Char	HOSPADM; DBP; SBP; HYPEREVT	Create one record for each PARAMCD even if the subject did not have the event.
AVAL	Analysis Value	Num		See parameter-level metadata below
STARTDT[a]	Time-to-Event Origin Date for Subject	Num		Date of first treatment. ADSL.TRTSDT
ADT[a]	Analysis Date	Num		See parameter-level metadata below
CNSR	Censored	Num	1; 0	See parameter-level metadata below
EVNTDESC	Event or Censoring Description	Char		See parameter-level metadata below
SRCDOM	Source Data	Char	HO; VS; DS	See parameter-level metadata below
SRCVAR	Source Variable	Char	HOSTDY; VSDY; DSSTDY	See parameter-level metadata below
SRCSEQ	Source Sequence Number	Num		The sequence number --SEQ of the row in the input dataset identified in the SRCDOM that relates to the analysis value being derived.

[a]Variables excluded from the sample input data (see Section 2.4.4, Input and Analysis Data 4).

Table 2.4.3.2. Parameter-level Metadata for ADHYP

Parameter Value List - ADHYP

Variable	Where	Controlled Terms / Formats	Source / Derivation / Comment
ADT	PARAMCD = 'HOSPADM'		If subject was admitted to the hospital, then ADT = HOSTDTC converted to numeric date. Otherwise use completion or discontinuation date ADSL.EOSDT.
ADT	PARAMCD = 'DBP'		If subject had diastolic blood pressure > 90, then ADT = VSDTC converted to numeric date. Otherwise use completion or discontinuation date ADSL.EOSDT.
ADT	PARAMCD = 'SBP'		If subject had systolic blood pressure > 140, then ADT = VSDTC converted to numeric date. Otherwise use completion or discontinuation date ADSL.EOSDT.
ADT	PARAMCD = 'HYPEREVT'		Use sub-events determine the earliest event time and set AVAL accordingly.
AVAL	PARAMCD = 'HOSPADM'		DERIVED: If subject was admitted to the hospital, then AVAL = HO.HOSTDY. Otherwise set to the study day based on completion or discontinuation date ADSL.EOSDT
AVAL	PARAMCD = 'DBP'		DERIVED: If subject had diastolic blood pressure > 90, then AVAL = VS.VSDY. Otherwise set to the study day based on completion or discontinuation date ADSL.EOSDT
AVAL	PARAMCD = 'SBP'		DERIVED: If subject had systolic blood pressure > 140, then AVAL = VS.VSDY. Otherwise set to the study day based on completion or discontinuation date ADSL.EOSDT
AVAL	PARAMCD = 'HYPEREVT'		DERIVED: If subject had a sub-event (records where PARAMCD = HOSPADM, DBP, SBP), then set to minimum of the AVAL values among the sub-event records where CNSR= 0. Otherwise set to the study day based on completion or discontinuation date ADSL.EOSDT.
CNSR	PARAMCD = 'HOSPADM'	1; 0	DERIVED: If subject was not admitted to the hospital, then CNSR = 1. Otherwise, CNSR = 0.
CNSR	PARAMCD = 'DBP'	1; 0	DERIVED: If subject never had diastolic blood pressure > 90, then CNSR = 1. Otherwise, CNSR = 0.
CNSR	PARAMCD = 'SBP'	1; 0	DERIVED: If subject never had systolic blood pressure > 140, then CNSR = 1. Otherwise, CNSR = 0.
CNSR	PARAMCD = 'HYPEREVT'	1; 0	DERIVED: If all of the sub-events (HOSPADM, DBP, SBP) had CNSR = 1, then CNSR = 1. Otherwise, CNSR = 0.
EVNTDESC	PARAMCD = 'HOSPADM'		DERIVED: If subject was admitted to the hospital, then EVNTDESC="FIRST HOSPITAL ADMISSION". Otherwise if DS.DSDECOD="COMPLETED" then EVNTDESC="COMPLETED THE STUDY". Otherwise EVNTDESC = DS.DSDECOD, which is the discontinuation reason.
EVNTDESC	PARAMCD = 'DBP'		DERIVED: If subject had diastolic blood pressure > 90, then EVNTDESC="FIRST DBP > 90". Otherwise if DS.DSDECOD="COMPLETED" then EVNTDESC="COMPLETED THE STUDY". Otherwise EVNTDESC = DS.DSDECOD, which is the discontinuation reason.
EVNTDESC	PARAMCD = 'SBP'		DERIVED: If subject had systolic blood pressure > 140, then EVNTDESC ="FIRST SBP > 140". Otherwise if DS.DSDECOD="COMPLETED" then EVNTDESC="COMPLETED THE STUDY". Otherwise EVNTDESC = DS.DSDECOD, which is the discontinuation reason.
EVNTDESC	PARAMCD = 'HYPEREVT'		DERIVED: If at least one of the sub-events (HOSPADM, DBP, SBP) had EVNTDESC that indicated "FIRST…", then EVNTDESC="HYPERTEN. EVENT". Otherwise if DS.DSDECOD="COMPLETED" then EVNTDESC="COMPLETED THE STUDY". Otherwise EVNTDESC = DS.DSDECOD, which is the discontinuation reason.
SRCDOM	PARAMCD = 'HOSPADM'	HO; DS	Assigned: If subject was admitted to the hospital, then SRCDOM="HO". Otherwise SRCDOM="DS".
SRCDOM	PARAMCD = 'DBP'	VS; DS	Assigned: If subject had diastolic blood pressure > 90, then SRCDOM="VS". Otherwise SRCDOM="DS".

Variable	Where	Controlled Terms / Formats	Source / Derivation / Comment
SRCDOM	PARAMCD = 'SBP'	VS; DS	Assigned: If subject had systolic blood pressure > 140, then SRCDOM="VS". Otherwise SRCDOM="DS".
SRCDOM	PARAMCD = 'HYPEREVT'	HO; VS; DS	DERIVED: Using sub-events determine the earliest event time and set SRCDOM accordingly. Otherwise SRCDOM="DS".
SRCVAR	PARAMCD = 'HOSPADM'	HOSTDY; DSSTDY	Assigned: If subject was admitted to the hospital, then SRCVAR="HOSTDY". Otherwise, SRCVAR="DSSTDY".
SRCVAR	PARAMCD = 'DBP'	VSDY; DSSTDY	Assigned: If subject had diastolic blood pressure > 90, then SRCVAR="VSDY". Otherwise, SRCVAR="DSSTDY".
SRCVAR	PARAMCD = 'SBP'	VSDY; DSSTDY	Assigned: If subject had systolic blood pressure > 140, then SRCVAR="VSDY". Otherwise, SRCVAR="DSSTDY".
SRCVAR	PARAMCD = 'HYPEREVT'	HOSTDY; VSDY; DSSTDY	DERIVED: Using sub-events determine the earliest event time and set SRCVAR accordingly. Otherwise, SRCVAR="DSSTDY".

2.4.4 Input and Analysis Data 4

As indicated in the data flow, HO, VS and DS datasets are 3 inputs that are needed for the creation of the analysis dataset (ADHYP).

Table 2.4.4.1. Data as Found in SDTM VS Dataset

Row	USUBJID	VISITNUM	VSSEQ	VSDTC	VSDY	VSTESTCD	VSSTRESN
1	2010	1	22	2004-08-05	1	SYSBP	115
2	2010	1	23	2004-08-05	1	DIABP	75
3	2010	2	101	2004-08-12	8	SYSBP	120
4	2010	2	102	2004-08-12	8	DIABP	90
5	2010	3	207	2004-08-19	15	SYSBP	135
6	2010	3	208	2004-08-19	15	DIABP	92
7	2010	4	238	2004-08-25	21	SYSBP	138
8	2010	4	239	2004-08-25	21	DIABP	95
9	3082	1	27	2004-09-08	1	SYSBP	120
10	3082	1	28	2004-09-08	1	DIABP	80
11	3082	2	119	2004-09-15	8	SYSBP	125
12	3082	2	120	2004-09-15	8	DIABP	84

Table 2.4.4.2. Data as Found in SDTM HO Dataset

ROW	USUBJID	HOSEQ	HOTERM	HODECOD	HOSTDTC	HOENDTC	HOSTDY	HOENDY
1	2010	99	HOSPITAL	HOSPITAL	2004-08-13	2004-08-15	9	11
2	2010	199	HOSPITAL	HOSPITAL	2004-08-20	2004-08-22	16	18

Table 2.4.4.3. Data as Found in SDTM DS Dataset

ROW	USUBJID	DSSEQ	DSSTDTC	DSSTDY	DSDECOD	DSTERM
1	2010	25	2004-08-05	1	RANDOMIZED	Subject Randomized
2	2010	301	2004-08-26	22	COMPLETED	Subject Completed
3	3082	20	2004-09-08	1	RANDOMIZED	Subject Randomized
4	3082	130	2004-09-17	10	COMPLETED	Subject Completed

Utilizing the SRCDOM, SRCVAR, and SRCSEQ variables in ADHYP enables tracing each record back to the source dataset.

Table 2.4.4.4. ADaM dataset ADHYP

ROW	USUBJID	PARAM	PARAMCD	AVAL	CNSR	EVNTDESC	SRCDOM	SRCVAR	SRCSEQ
1	2010	Time to First Hospital Admission (day)	HOSPADM	9	0	FIRST HOSPITAL ADMISSION	HO	HOSTDY	99
2	2010	Time to First DBP>90 (day)	DBP	15	0	FIRST DBP>90	VS	VSDY	208
3	2010	Time to First SBP>140 (day)	SBP	22	1	COMPLETED THE STUDY	DS	DSSTDY	301
4	2010	Time to Hypertension Event (day)	HYPEREVT	9	0	HYPERTEN. EVENT	HO	HOSTDY	99
5	3082	Time to First Hospital Admission (day)	HOSPADM	10	1	COMPLETED THE STUDY	DS	DSSTDY	130
6	3082	Time to First DBP>90 (day)	DBP	10	1	COMPLETED THE STUDY	DS	DSSTDY	130
7	3082	Time to First SBP>140 (day)	SBP	10	1	COMPLETED THE STUDY	DS	DSSTDY	130
8	3082	Time to Hypertension Event (day)	HYPEREVT	10	1	COMPLETED THES TUDY	DS	DSSTDY	130

Note that Row 1 can be traced back to the HO data (HOSEQ = 99); the value of Row 2 can be traced back to VS data (VSSEQ = 208).

2.4.5 Other Uses 4

If an ADaM dataset is used as input to another ADaM dataset, then SRCDOM, SRCVAR, and SRCSEQ should be used as data-point traceability (because the use of --SEQ is not possible). In the ADaM dataset that is used as input, ASEQ should be incorporated and this value will be used to populate SRCSEQ. See Section 2.8, Using an Intermediate Dataset for BDS Traceability, for a TTE example with SRC* variables pointing to an ADaM dataset.

2.5 Traceability When Multiple Input Datasets Are Stacked to Create OCCDS

There are instances when it is necessary for multiple SDTM domains to be used as inputs to a single OCCDS dataset. This example will address the scenario when rows from multiple input datasets are stacked together to form a single analysis dataset for occurrence analyses.

2.5.1 Analysis Need 5

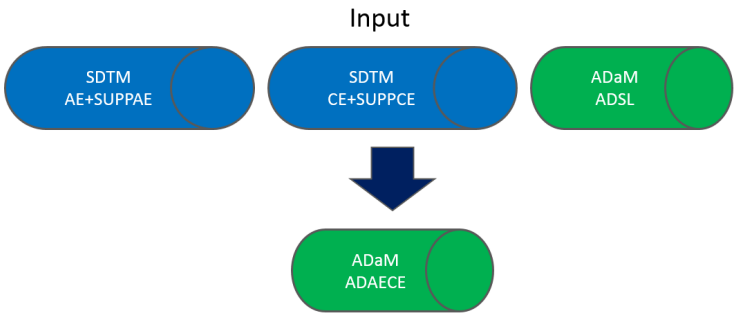
In a double-blind parallel-design study, the performance of infusions containing the study drug is compared against infusions containing the standard of care. Two different eCRFs captured related information about local infusion site reactions (SDTM CE) and spontaneous infusion reactions adverse events (SDTM AE). For the safety summary, the study statistician decided that both sources of data would be combined and reported.

Table 2.5.1.1. Sample Table Shell

Summary of All Adverse Events by System Organ Class, MedDRA Preferred Term and Treatment Group Safety Population		
System Organ Class MedDRA Preferred Term	Study Drug (N=XXX)	Standard of Care (N=XXX)
Subjects with at least one Adverse Event	xx (xx.x)	xx (xx.x)
Blood and lymphatic system disorders	xx (xx.x)	xx (xx.x)
Anaemia deficiencies	xx (xx.x)	xx (xx.x)
Lymphadenopathy	xx (xx.x)	xx (xx.x)

2.5.2 Data Flow 5

Figure 2.5.2.1. Data Flow



In this example AE and CE SDTM domains plus their corresponding supplemental qualifier dataset are combined and used as input to the OCCDS ADAECE dataset. Additional needed variables are taken from the ADSL.

2.5.3 Traceability Metadata 5

Table 2.5.3.1. Dataset Metadata for ADAECE

ADAECE Dataset Metadata

Dataset	Description	Class	Structure	Keys	Purpose
ADAECE	Adverse/Clinical Events Analysis Dataset	OCCURRENCE DATA STRUCTURE	One record per subject per combined preferred term per start datetime	STUDYID, USUBJID, UDECOD, ASTDTM, ASTDT	Analysis

Table 2.5.3.2. Variable Metadata for ADAECE

ADAECE Variable Metadata

Variable Name[a]	Variable Label	Codelists	Variable Metadata
STUDYID	Study Identifier	XYZ	ADSL.STUDYID
USUBJID	Unique Subject Identifier		ADSL.USUBJID
SAFFL	Safety Population Flag	'Y'='Yes' 'N'='No'	ADSL.SAFFL
TRTA	Actual Treatment	SOC+SD=Standard of Care + Study Drug SOC=Standard of Care	ADSL.TRT01A
TRTSDT	Date of First Exposure to Treatment	yymmdd10.	ADSL.TRTSDT
TRTSDTM	Datetime of First Exposure to Treatment	datetime20.	ADSL.TRTSDTM
SRCDOM	Source Data		Set to "AE" if record is from AE dataset. Set to "CE" if record is from CE dataset.
SRCSEQ	Source Sequence Number		Set to AE.AESEQ if record is from AE dataset. Set to CE.CESEQ if record is from CE dataset.
ACAT1	Analysis Category 1	ADVERSE EVENTS LOCAL INFUSION SITE REACTIONS	If record is from AE then ACAT1="ADVERSE EVENTS" Else ACAT1="LOCAL INFUSION SITE REACTIONS"
UTERM	Reported Term		AE.AETERM if record is from AE dataset CE.CETERM if record is from CE dataset
UDECOD	Dictionary-Derived Term	MedDRA	AE.AEDECOD if record is from AE dataset CE.CEDECOD if record is from CE dataset
UBODSYS	Body System or Organ Class	MedDRA	AE.AEBODSYS if record is from AE dataset CE.CEBODSYS if record is from CE dataset
USTDTC	Start Date/Time of Event	ISO8601	AE.AESTDTC if record is from AE dataset CE.CESTDTC if record is from CE dataset
UENDTC	End Date/Time of Event	ISO8601	AE.AEENDTC if record is from AE dataset CE.CEENDTC if record is from CE dataset
ASTDT	Analysis Start Date	yymmdd10.	For example: Date part of AESTDTC. 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 1st. If start date is completely missing do not impute.
ASTDTF	Analysis Start Date Imputation Flag	DATEF	If start date has month missing then ASTDTF="M" Else if start date has day missing then ASTDTF="D"
ASTDTM	Analysis Start Date/Time	datetime20.	For example: Convert AESTDTC to a numeric datetime variable
AREL	Analysis Causality		If record is from AE then AREL=AE.AEREL converted to proper case Else if record is from CE then AREL="Definitely Related"
ARELGR1	Pooled Causality Group 1	Related Not Related	If AREL in('Definitely Related' 'Possibly Related' 'Probably Related') then ARELGR1="Related". Else if AREL in('Not Related' 'Unlikely Related') then ARELGR1="Not Related".
UTOXGR	Toxicity Grade	'1'='Grade 1' '2'='Grade 2' '3'='Grade 3' '4'='Grade 4' '5'='Grade 5'	Set to AE.AETOXGR if record is from AE dataset Set to CE.CETOXGR if record is from CE dataset
UACN	Action Taken with Study Treatment		Set to AE.AEACN if record is from AE dataset Set to CE.CEACN if record is from CE dataset
CEPRESP	Clinical Event Pre-Specified	'Y'='Yes'	CE.CEPRESP The Pre-specified Flag is copied in to support ad-hoc analyses involving comparisons between pre-specified and spontaneous events.
TRTEMFL	Treatment Emergent Analysis Flag	'Y'='Yes'	For example: Assume TRTEMFL="Y" unless proven that event is not treatment emergent If both the Start Date/Time of the Adverse Event and Treatment are present and populated and Start Date/Time of Adverse Event is prior to Start Date/Time of Treatment (MISSING<ASTDTM<TRTSDTM) then set TRTEMFL to NULL. If either the Start Date/Time of the Adverse Event or the Start Date/Time 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 Date/Time or End Date is present and prior to Start Date/Time or Start Date of Treatment then set TRTEMFL to NULL.

[a]For the use of variables with a U prefix and SRCDOM/SRCSEQ, see OCCDS v1.1 (https://www.cdisc.org/standards/foundational/adam/).

Variables with a U* prefix (i.e., UTERM, UBODSYS, UDECOD, USTDTC, UENDTC, UTOXGR, UACN) have been added in order to stack unmodified content from multiple existing SDTM domains with the same root variables into one field. For example, both CE and AE have the root variable BODSYS (Body System or Organ Class) in the same dictionary version. Because AEBODSYS and CEBODSYS are stacked without modification for analysis, the

variable UBODSYS was created to denote this. The U* prefix variable-naming convention extends beyond variables used for analysis; it can be used for traceability variables such as USTDTC and UENDTC.

SRCVAR is not included because it holds the name of the primary variable used to derive AVAL or AVALC in BDS datasets, so is not applicable here.

2.5.4 Input and Analysis Data 5

Table 2.5.4.3 is an illustration of the adverse events analysis dataset (ADAECE) defined by the preceding metadata.

Key points to note in this example are:

1. Only CE events that occurred for subject XYZ-001-001 and were considered to be candidates for adverse events were included in ADAECE. For example, rows 1 and 4 in the sample CE data did not occur (CEOCCUR ="N"), so the data were not included.
2. Because multiple SDTM domains are stacked together, SRCSEQ and SRCDOM may be used in OCCDS datasets to point to the domain and record where the data came from. These variables are mentioned in OCCDS v1.1.
3. In ADAECE, the U* variables allows users to preserve the same type of data from multiple SDTM domains in the same column. U* can only be used for a direct copy as the "U" indicates “unmodified.” Any changes in type, casing, harmonization or imputation requires the creation of A* variables (as seen with ARELGR1 and AREL).
4. Note that ACAT1 was populated to indicate the type of data for traceability purposes and was not required for analysis.

Table 2.5.4.1. Sample AE Data

ae.xpt

Row	STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AEDECOD	AEBODSYS	AESER	AEACN	AEREL	AETOXGR	AESTDTC	AEENDTC
1	XYZ	AE	XYZ-001-001	1	FEVER	Pyrexia	General disorders and administration site conditions	N	DRUG INTERRUPTED	PROBABLY RELATED	3	2014-02-15T20:15	2014-02-17T05:01
2	XYZ	AE	XYZ-001-001	2	CHILLS	Chills	General disorders and administration site conditions	N	DRUG INTERRUPTED	POSSIBLY RELATED	2	2014-02-15	2014-02-17
3	XYZ	AE	XYZ-001-001	3	HEADACHE	Headache	Nervous system disorders	N	DOSE NOT CHANGED	POSSIBLY RELATED	1	2014-02	2014-02-17T20:40
4	XYZ	AE	XYZ-001-001	4	LOW NEUTROPHILS	Neutropenias	Blood and lymphatic system disorders	N	DOSE NOT CHANGED	POSSIBLY RELATED	2	2014-04-14T09:21	2014-06-12T08:30
5	XYZ	AE	XYZ-001-001	5	DIARRHEA	Diarrhea	Gastrointestinal disorders	N	DOSE NOT CHANGED	PROBABLY RELATED	1	2014-05-15	2014-05-16
6	XYZ	AE	XYZ-001-001	6	PNEUMONIA	Pneumonia	Infections and infestations	Y	DOSE REDUCED	POSSIBLY RELATED	3	2014-05-13	2014-05-15
7	XYZ	AE	XYZ-001-001	7	NAUSEA	Nausea	Gastrointestinal disorders	N	DOSE NOT CHANGED	PROBABLY RELATED	1	2014-07-12T14:00	2014-07-13T22:00

Table 2.5.4.2. Sample CE Data

ce.xpt

Row	STUDYID	DOMAIN	USUBJID	CESEQ	CETERM	CEDECOD	CECAT	CEPRESP	CEOCCUR	CEBODSYS	CEACN	CETOXGR	CESTDTC	CEENDTC
1	XYZ	CE	XYZ-001-001	1	PAIN AT THE INFUSION SITE	Pain	LOCAL INFUSION SITE REACTIONS	Y	N	General disorders and administration site conditions				
2	XYZ	CE	XYZ-001-001	2	REDNESS AT THE INFUSION SITE	Skin erythema	LOCAL INFUSION SITE REACTIONS	Y	Y	Skin and subcutaneous tissue disorders	DOSE NOT CHANGED	2	2014-02-15T10:05	2014-02-15T18:00
3	XYZ	CE	XYZ-001-001	3	SWELLING AT THE INFUSION SITE	Edema peripheral	LOCAL INFUSION SITE REACTIONS	Y	Y	General disorders and administration site conditions	DOSE NOT CHANGED	1	2014-02-15T10:30	2014-02-15T18:35
4	XYZ	CE	XYZ-001-001	4	RASH AT THE INFUSION SITE	Rash	LOCAL INFUSION SITE REACTIONS	Y	N	Skin and subcutaneous tissue disorders				

Table 2.5.4.3. Sample ADAECE Data

Row	STUDYID	USUBJID	SAFFL	TRTA	TRTSDT	TRTSDTM	SRCDOM	SRCSEQ	ACAT1	UTERM	UDECOD	UBODSYS	USTDTC	UENDTC	ASTDT	ASTDTF	ASTDTM	AREL	ARELGR1	ATOXGR	UACN	CEPRESP	TRTEMFL
1	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	CE	2	LOCAL INFUSION SITE REACTIONS	REDNESS AT THE INFUSION SITE	Skin erythema	Skin and subcutaneous tissue disorders	2014-02-15T10:05	2014-02-15T18:00	15FEB2014		15FEB2014:10:05	Definitely Related	Related	2	DOSE NOT CHANGED	Y	Y
2	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	CE	3	LOCAL INFUSION SITE REACTIONS	SWELLING AT THE INFUSION SITE	Edema peripheral	General disorders and administration site conditions	2014-02-15T10:30	2014-02-15T18:35	15FEB2014		15FEB2014:10:30	Definitely Related	Related	1	DOSE NOT CHANGED	Y	Y
3	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	AE	1	ADVERSE EVENTS	FEVER	Pyrexia	General disorders and administration site conditions	2014-02-15T20:15	2014-02-17T05:01	15FEB2014		15FEB2014:20:15	Probably Related	Related	3	DRUG INTERRUPTED		Y
4	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	AE	2	ADVERSE EVENTS	CHILLS	Chills	General disorders and administration site conditions	2014-02-15	2014-02-17	15FEB2014			Possibly Related	Related	2	DRUG INTERRUPTED		Y
5	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	AE	3	ADVERSE EVENTS	HEADACHE	Headache	Nervous system disorders	2014-02	2014-02-17T20:40	15FEB2014	D		Possibly Related	Related	1	DOSE NOT CHANGED		Y
6	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	AE	4	ADVERSE EVENTS	LOW NEUTROPHILS	Neutropenias	Blood and lymphatic system disorders	2014-04-14T09:21	2014-06-12T08:30	14APR2014		14APR2014:09:21	Possibly Related	Related	2	DOSE NOT CHANGED		Y
7	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	AE	5	ADVERSE EVENTS	DIARRHEA	Diarrhea	Gastrointestinal disorders	2014-05-15	2014-05-16	15MAY2015			Probably Related	Related	1	DOSE NOT CHANGED		Y
8	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	AE	6	ADVERSE EVENTS	PNEUMONIA	Pneumonia	Infections and infestations	2014-05-13	2014-05-15	13MAY2014			Possibly Related	Related	3	DOSE REDUCED		Y
9	XYZ	XYZ-001-001	Y	SOC + SD	15FEB2014	15FEB2014:10:05	AE	7	ADVERSE EVENTS	NAUSEA	Nausea	Gastrointestinal disorders	2014-07-12T14:00	2014-07-13T22:00	12JUL2014		12JUL2014:14	Probably Related	Related	1	DOSE NOT CHANGED		Y

2.5.5 Other Uses 5

This example demonstrates the "stacking" concept with multiple SDTM datasets being combined in an OCCDS dataset to support a complex derivation. The concept of stacking records from multiple sources and maintaining data point traceability was first introduced in BDS. The difference between stacking in BDS and OCCDS datasets is that although both use SRCDOM and SRCSEQ to identify the source record, BDS may also use the SRCVAR variable to identify the source variable related to AVAL/AVALC.

2.6 Traceability When Multiple Datasets Are Merged

Some analyses involve merging and subsetting of data that is quite straightforward. Examples of this include source data where the records have variables built in to allow for merging or linking, such as TRLNKID relating assessment records in the Tumor/Lesion Results (TR) domain with an identification record in the Tumor/Lesion Identification (TU) domain or MBGRPID linking findings in Microbiology Specimen (MB) with an organism in the Microbiology Susceptibility (MS) domain. Sometimes datasets have a set of keys that allow them to be merged together (e.g., subject, analysis visit number, analysis visit, date). In these cases a brief description of the merge and/or subset in the documentation section of the dataset-level metadata in the define.xml and/or ADaM Reviewer's Guide (ADRG) may be sufficient for the user of the data to understand the origin. In other circumstances, merging of data is not so clear cut and data needs to be joined together and slotted by way of windows or algorithms. The creation of one analysis dataset may rely on derived variables from another analysis dataset. This could create a dependency of one dataset on the other, which is acceptable as long as no circular dependencies exist. This could result in the same variables being created twice (permanent in one dataset, intermediate in the other)—which is acceptable if they are created in exactly the same manner and any changes applied to the creation in both datasets. This example focuses on maintaining traceability in instances where the merges and subsets are not well defined.

2.6.1 Analysis Need 6

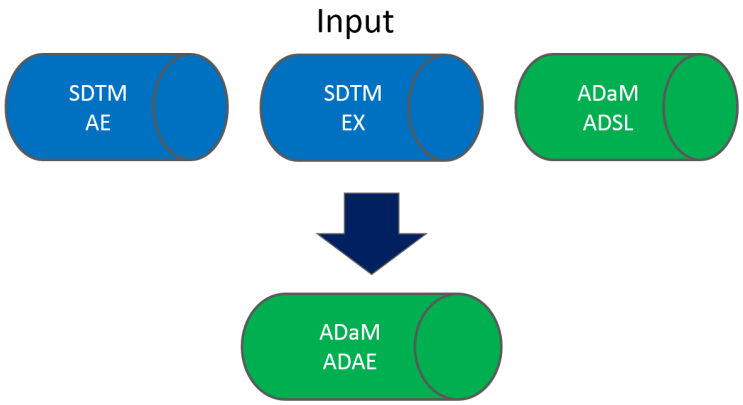
The focus of this example is the identification of infusion-associated reactions as a subset of Adverse Events. In this example, an adverse event is considered an infusion reaction (1) if it occurred on the day of an infusion (on or after the start of the infusion, if time of event is captured) or within 24 hours of the end of an infusion and (2) has been identified by the clinical team as qualifying as this type of an event. A lookup table, provided by the clinical team, is used to flag events considered as candidates for infusion reactions, and this list is also included in the ADAE section of the ADRG (see also Section 2.11, Traceability When Using a Look-up Table). In this example, the incidence of infusion-associated reactions occurring in ≥5% of patients is to be summarized by all grades and grade ≥3. The count and percentage of subjects with Infusion reactions will be summarized in the overall adverse event summary table, along with the count and percentage of subjects with an event associated with the first infusion (300 mg) and the second infusion (1000 mg).

Table 2.6.1.1. Sample Table Shell

Incidence of All Adverse Infusion Reactions Occurring in ≥5% of the Subjects (Safety Population)		
Adverse Infusion Reaction	Drug Z (N=xxx)	
	All Grades n (%)	Grade ≥3 n (%)
Pneumonia[a]	x (x.x)	x (x.x)
Pyrexia	x (x.x)	x (x.x)
Cough	x (x.x)	x (x.x)
Diarrhea	x (x.x)	x (x.x)
…..		
Rash[b]	x (x.x)	x (x.x)
…..		
Sepsis[c]	x (x.x)	x (x.x)
…..		
[a] Includes pneumonia, lung infection, lobar pneumonia, and bronchopneumonia. [b] Includes rash, rash macular, and rash vesicular. [c] Includes sepsis, neutropenic sepsis, bacteremia, and septic shock.		

2.6.2 Data Flow 6

Figure 2.6.2.1. Data Flow



There are multiple approaches to building the OCCDS dataset to be used for this analysis. The approach demonstrated here is to start with the SDTM datasets, create ASTDTM/AENDTM/AVISIT/INFNUM from EX, and use them in combination with a custom query on Adverse Events of Special Interest (AESIs) to derive the events of interest in ADAE. An important feature of this approach is that it does not reference ADEX. This allows an ADEX to be created downstream from this ADAE dataset and to make use of the identified infusion-associated reactions in its derivations.

2.6.3 Traceability Metadata 6

Table 2.6.3.1. Variable-level Metadata for ADAE

ADAE Variable Metadata

Variable Name	Variable Label	Codelist	Variable Metadata
STUDYID	Study Identifier	'XYZ'	ADSL.STUDYID
USUBJID	Unique Subject Identifier		ADSL.USUBJID
AESEQ	Sequence Number		AE.AESEQ
AETERM	Reported Term for the Adverse Event		AE.AETERM
AEDECOD	Dictionary-Derived Term	MedDRA	AE.AEDECOD
CQ01NAM	Customized Query 01 Name	'IARS (INFUSION ASSOCIATED REACTIONS)'	Set to 'IARS (INFUSION ASSOCIATED REACTIONS)' if deemed a potential Infusion Associated Reaction by Sponsor's Clinical Team. A list of all unique preferred terms (AEDECOD) by system organ class (AEBODSYS) were reviewed by the clinical team. This information is included in a look-up table (LUT) and merged in using AEDECOD. (The LUT not included in this example, please see 2.11 for an example on LUT)
AEBODSYS	Body System or Organ Class	MedDRA	AE.AEBODSYS
AESTDTC	Start Date/Time of Adverse Event	ISO8601	AE.AESTDTC
AEENDTC	End Date/Time of Adverse Event	ISO8601	AE.AEENDTC
ASTDTM	Analysis Start Datetime	datetime20.	If start date/time of adverse event is present (AESTDTC) then convert to numeric. If start date is present but time is missing and the start date coincides with the start date of an infusion then set the start date of the adverse event to the start date and time of the infusion. If the start date is present and time is missing and start date does not coincide with the start date of the infusion, then set the time to 00:00. If day is missing but at least one infusion occurs within the same month as the event, then set to the earliest date/time of the infusion within that month. If not, assume the first of the month and assume a time of 00:00. There are no records with only year populated for the start of the adverse event or records where the start date is completely missing.
ASTTMF	Analysis Start Time Imputation Flag	TIMEFL 'H'='Hours Imputed' 'M'='Minutes Imputed' 'S'='Seconds Imputed'	If start time of adverse event is completely missing or missing the hour then ASTTMF="H" Else if start time of adverse event has minutes missing then ASTTMF="M"
INFWSDTM	Infusion Window Start Datetime	datetime20.	For TEAEs, the datetime of the most recent infusion (EX.EXSTDTC on or before ASTDTM), converted to numeric

Variable Name	Variable Label	Codelist	Variable Metadata
INFWEDTM	Infusion Window End Datetime	datetime20.	EX.EXENDTC of the records used to populate INFWSDTM, converted to numeric + 24 hours
AVISIT	Analysis Visit	'CYCLE 1 DAY 1' 'CYCLE 1 DAY 8' 'CYCLE 2 DAY 1' 'CYCLE 3 DAY 1' 'CYCLE 4 DAY 1' 'CYCLE 5 DAY 1' 'CYCLE 6 DAY 1'	For treatment emergent adverse events, AVISIT is the visit corresponding with the infusion associated with this adverse event. If the start date/time of the adverse event falls on or after the start of infusion X but before the start of the subsequent infusion X+1, then AVISIT is populated with the visit corresponding to Infusion X. For the subject's last infusion, if the adverse event falls within 30 days of the last infusion, associate that Adverse Event with that Infusion. If greater than 30 days, leave variable NULL.
INFNUM	Infusion Number	1, 2, 3, 4, 5, 6, 7	For treatment emergent adverse events, INFNUM is the Infusion Number (Count of infusions up to and including this infusion) corresponding with the infusion associated with this adverse event. If the start date/time of the adverse event falls on or after the start of infusion X but before the start of the subsequent infusion X+1, then INFNUM is populated with the count of infusions received up to and including Infusion X.
IW24HRFL	Infusion 24 Hour Window Flag	'Y'='Yes'	If the event occurs during or within 24 hours of an infusion then set IW24HRFL="Y". [If (INFWSDTM<=ASTDTM<=INFWEDTM) then IW24HRFL="Y"]
IARSFL	Infusion Associated Reaction Flag	'Y'='Yes'	If the event occurs during or within 24 hours of an infusion (INFWEDTM<=ASTDTM<=INFWEDTM) and is considered to be a potential infusion associated reaction then set IARSFL="Y". [If (IW24HRFL='Y') and CQ01NAM="IARS (INFUSION ASSOCIATED REACTIONS)" then IARSFL="Y"]

2.6.4 Input and Analysis Data 6

Key points to note in this example are:

1. AVISIT is used to associate or relate adverse events with an infusion. This variable may be used to look at adverse events associated with a particular infusion (e.g. first infusion) or to see if infusion reactions increase in intensity over time. The producer may choose to use a custom variable such as INFVISIT when adverse events are to be slotted into analysis visits based on the protocol/SAP that extend beyond infusions.
2. Note that the Customized Query 01 Name (CQ01NAM) variable is created using a look-up table (see Section 2.11, Traceability When Using a Look-up Table, for further details).
3. Note that Infusion 24 Hour Window Flag (IW24HRFL) is set to Y if the analysis start datetime (includes imputations) of the adverse event occurs during the infusion or up to 24 hours after the end of the infusion (after datetime imputations have taken place if necessary). A custom flag was chosen here as options such as CRITy/CRITyFL and ATPT/ATPTN are unclear based on CDISC Notes and current examples in the ADaMIG. The pair of variables CRITy/CRITyFL originally apply to a criterion within a parameter. Although ATPT can contain derived analysis time window names, it is not clear if populating ATPTREF="START OF INFUSION" and ATPT="FROM START OF INFUSION UNTIL END + 24 HOURS" on records that met the windowing criteria would be appropriate.
4. The sponsor defined flag IARSFL (Infusion Associated Reaction Flag) was included to flag records that matched a clinical term and fell within the time interval for an infusion reaction. ANLzzFL was discussed as an option to be used for this flag but was not based on the definition and current examples in the ADaMIG. IARSFL="Y" when IW24HRFL="Y" and CQ01NAM is populated so it is not necessary for record selection. Most likely AOCCzzFL will need to be included to make the data analysis-ready as subjects may have multiple infusion reactions at the different levels of summarization.

Table 2.6.4.1. Sample AE

ae.xpt

Row	STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AEDECOD	AEBODSYS	AESTDTC	AEENDTC
1	XYZ	AE	XYZ-001-001	1	REDNESS AT THE INFUSION SITE	Skin erythema	Skin and subcutaneous tissue disorders	2014-02-15T10:05	2014-02-15T18:00
2	XYZ	AE	XYZ-001-001	2	SWELLING AT THE INFUSION SITE	Edema peripheral	General disorders and administration site conditions	2014-02-15T10:30	2014-02-15T18:35
3	XYZ	AE	XYZ-001-001	3	FEVER	Pyrexia	General disorders and administration site conditions	2014-02-15T20:15	2014-02-17T05:01
4	XYZ	AE	XYZ-001-001	4	CHILLS	Chills	General disorders and administration site conditions	2014-02-17	2014-02-18
5	XYZ	AE	XYZ-001-001	5	HEADACHE	Headache	Nervous system disorders	2014-03-16T05:35	2014-03-17T20:40
6	XYZ	AE	XYZ-001-001	6	LOW NEUTROPHILS	Neutropenias	Blood and lymphatic system disorders	2014-04-14T09:21	2014-06-12T08:30
7	XYZ	AE	XYZ-001-001	7	DIARRHEA	Diarrhea	Gastrointestinal disorders	2014-05-15	2014-05-16
8	XYZ	AE	XYZ-001-001	8	NAUSEA	Nausea	Gastrointestinal disorders	2014-05-13	2014-05-15
9	XYZ	AE	XYZ-001-001	9	NAUSEA	Nausea	Gastrointestinal disorders	2014-07-12T14:00	2014-07-13T22:00

Table 2.6.4.2. Sample EX

ex.xpt

Row	STUDYID	DOMAIN	USUBJID	EXSEQ	EXTRT	EXDOSE	EXDOSU	EXDOSFRM	EXDOSFRQ	EXROUTE	VISITNUM	VISIT	EXSTDTC	EXENDTC
1	XYZ	EX	XYZ-001-001	1	DRUG Z	300	mg	SOLUTION	ONCE	INTRAVENOUS	101	CYCLE 1 DAY 1	2014-02-15T09:35	2014-02-15T14:15
2	XYZ	EX	XYZ-001-001	2	DRUG Z	1000	mg	SOLUTION	ONCE	INTRAVENOUS	108	CYCLE 1 DAY 8	2014-02-23T08:10	2014-02-23T14:20
3	XYZ	EX	XYZ-001-001	3	DRUG Z	1000	mg	SOLUTION	ONCE	INTRAVENOUS	201	CYCLE 2 DAY 1	2014-03-16T09:01	2014-03-16
4	XYZ	EX	XYZ-001-001	4	DRUG Z	1000	mg	SOLUTION	ONCE	INTRAVENOUS	301	CYCLE 3 DAY 1	2014-04-14T09:21	2014-04-14T13:51
5	XYZ	EX	XYZ-001-001	5	DRUG Z	1000	mg	SOLUTION	ONCE	INTRAVENOUS	401	CYCLE 4 DAY 1	2014-05-13T21:30	2014-05-14T01:45
6	XYZ	EX	XYZ-001-001	6	DRUG Z	1000	mg	SOLUTION	ONCE	INTRAVENOUS	501	CYCLE 5 DAY 1	2014-06-12T08:30	2014-06-12T14:00
7	XYZ	EX	XYZ-001-001	7	DRUG Z	1000	mg	SOLUTION	ONCE	INTRAVENOUS	601	CYCLE 6 DAY 1	2014-07-11T07:03	2014-07-11T13:29

Table 2.6.4.3. Sample ADAE

Row	STUDYID	USUBJID	AESEQ	AETERM	AEDECOD	CQ01NAM	AEBODSYS	AESTDTC	AEENDTC	ASTDTM	ASTTMF	INFWSDTM	INFWEDTM	AVISIT	INFNUM	IW24HRFL	IARSFL
1	XYZ	XYZ-001-001	1	REDNESS AT THE INFUSION SITE	Skin erythema	IARS (INFUSION ASSOCIATED REACTIONS)	Skin and subcutaneous tissue disorders	2014-02-15T10:05	2014-02-15T18:00	15FEB2014:10:05:00		15FEB2014:09:35:00	16FEB2014:14:15:00	CYCLE 1 DAY 1	1	Y	Y
2	XYZ	XYZ-001-001	2	SWELLING AT THE INFUSION SITE	Edema peripheral	IARS (INFUSION ASSOCIATED REACTIONS)	General disorders and administration site conditions	2014-02-15T10:30	2014-02-15T18:35	15FEB2014:10:30:00		15FEB2014:09:35:00	16FEB2014:14:15:00	CYCLE 1 DAY 1	1	Y	Y
3	XYZ	XYZ-001-001	3	FEVER	Pyrexia	IARS (INFUSION ASSOCIATED REACTIONS)	General disorders and administration site conditions	2014-02-15T20:15	2014-02-17T05:01	15FEB2014:20:15:00		15FEB2014:09:35:00	16FEB2014:14:15:00	CYCLE 1 DAY 1	1	Y	Y
4	XYZ	XYZ-001-001	4	CHILLS	Chills	IARS (INFUSION ASSOCIATED REACTIONS)	General disorders and administration site conditions	2014-02-17	2014-02-18	17FEB2014:00:00:00	H	15FEB2014:09:35:00	16FEB2014:14:15:00	CYCLE 1 DAY 1	1		
5	XYZ	XYZ-001-001	5	HEADACHE	Headache	IARS (INFUSION ASSOCIATED REACTIONS)	Nervous system disorders	2014-03-16T05:35	2014-03-17T20:40	16MAR2014:05:35:00		23FEB2014:08:10:00	24FEB2014:14:20:00	CYCLE 1 DAY 8	2		
6	XYZ	XYZ-001-001	6	LOW NEUTROPHILS	Neutropenias		Blood and lymphatic system disorders	2014-04-14T09:21	2014-06-12T08:30	14APR2014:09:21:00		14APR2014:09:21:00	15APR2014:13:51:00	CYCLE 3 DAY 1	4	Y	
7	XYZ	XYZ-001-001	7	DIARRHEA	Diarrhea	IARS (INFUSION ASSOCIATED REACTIONS)	Gastrointestinal disorders	2014-05-15	2014-05-16	15MAY2014:00:00:00	H	13MAY2014:21:30:00	15MAY2014:01:45:00	CYCLE 4 DAY 1	5	Y	Y
8	XYZ	XYZ-001-001	8	NAUSEA	Nausea	IARS (INFUSION ASSOCIATED REACTIONS)	Gastrointestinal disorders	2014-05-13T21:30	2014-05-15	13MAY2014:21:30:00		13MAY2014:21:30:00	15MAY2014:01:45:00	CYCLE 4 DAY 1	5	Y	Y
9	XYZ	XYZ-001-001	9	NAUSEA	Nausea	IARS (INFUSION ASSOCIATED REACTIONS)	Gastrointestinal disorders	2014-07-12T14:00	2014-07-13T22:00	12JUL2014:14:00:00		11JUL2014:07:03:00	12JUL2014:13:29:00	CYCLE 6 DAY 1	7		

2.7 Traceability When Adding a Row to a BDS Dataset

It is not uncommon to need to derive an analysis value from multiple rows from a preceding SDTM dataset. The ADaM basic dataset structure variable DTYPE is available to indicate when a new derived row has been added to a dataset and to define how the analysis value (AVAL) was derived. The example demonstrates measuring electrocardiogram (ECG) values in triplicate at each time point in a study; the average of these triplicate values is used in the analysis.

2.7.1 Analysis Need 7

In this example, the analysis requirement is to summarize the average of the triplicate ECG interval values (AVAL) as well as change from baseline (CHG), where baseline is defined as the average of the triplicate ECG intervals collected prior to the first administration of study drug. This summary will be performed by analysis visit.

Table 2.7.1.1. Sample Table Shell

Dataset=ADEG

Table x.x.x
Actual Values and Change from Baseline in ECG Parameters by Time point
(Safety Analysis Set) SAFFL

PARAM
AVISIT

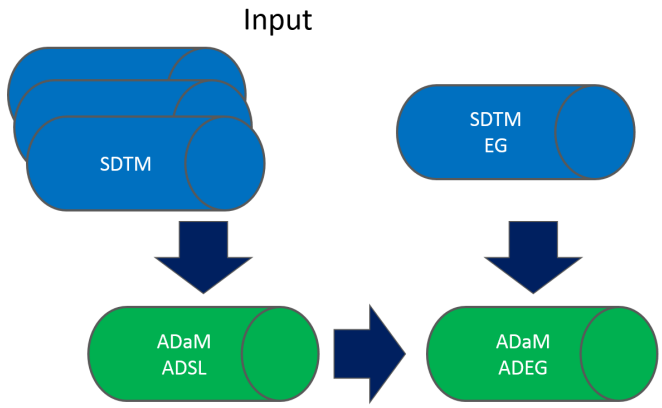
Parameter Time point	Statistic	TRTA: Placebo (N=xx)		Active xx mg (N=xx)		Active yy mg (N=xx)	
		Actual Value (AVAL)	Change from Baseline (CHG)	Actual Value	Change from Baseline	Actual Value	Change from Baseline
QTcF Interval (msec)							
Baseline	n	xx		xx		xx	
	Mean (SD)	xx.x (xx.xx)		xx.x (xx.xx)		xx.x (xx.xx)	
	Median	xx.x		xx.x		xx.x	
	Min, Max	xx, xx		xx, xx		xx, xx	
Visit 1							
	N	xx	xx	xx	xx	xx	xx
	Mean (SD)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)
	Median	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
Visit 2	n	xx	xx	xx	xx	xx	xx
	Mean (SD)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)
	Median	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
	Min, Max	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx	xx, xx
. . .							

The average of triplicate or available ECG measurements collected at each nominal time point are used for analysis.
Cross-reference: Listing 8.3.1

Program: xxxxxxx.sas, Output: xxxxxxx.rtf, Generated on: DDMONYYYY xx:xx Page x of y
Programming Notes:
Note [1]: Repeat for all ECG parameters and time points including Day 28, but NOT Day 28/EDV.
Note [2]: Present ECG parameters ordered as follows: Heart rate, PR interval, QRS duration, QTcB, QTcF.

2.7.2 Data Flow 7

Figure 2.7.2.1. Data Flow



2.7.3 Traceability Metadata 7

Table 2.7.3.1. Dataset-level Metadata for ADEG

ADEG Dataset Metadata

Dataset	Dataset Description	Data Structure	Class of Dataset
ADEG	Electrocardiogram Analysis Dataset	One record per subject, parameter, analysis visit, reference ID, derivation type	BASIC DATA STRUCTURE

Table 2.7.3.2. Variable Metadata for ADEG

ADEG Variable Metadata

Variable Name	Variable Label	Code List/Controlled Terminology	Source/Derivation/Comment
STUDYID	Study Identifier		Predecessor: ADSL.STUDYID
USUBJID	Unique Subject Identifier		Predecessor: ADSL.USUBJID
EGSEQ	Sequence Number		Predecessor: EG.EGSEQ Set to missing for derived rows.
EGREPNUM	Repetition Number		Predecessor: EG.EGREPNUM Set to missing for derived rows.
PARAM	Parameter		Derived: Create PARAM value in format "EGTESTCD (EGSTRESU)" for example: "Heart Rate (beats/min)" "PR interval (msec)" "QRS interval (msec)" "QTCB Interval (msec)" "QTCF Interval (msec)"
VISIT	Visit Name		Predecessor: EG.VISIT
AVISIT	Analysis Visit		Derived: Derivation is explained in analysis data reviewer's guide, Section 7.5.2 Analysis Data Reviewer's Guide
EGDTC	Date/Time of ECG		Predecessor: EG.EGDTC Set to missing for derived rows.
BASE	Baseline Value		Derived: For post-baseline records it is value of AVAL for each subject and parameter where ABLFL = Y
AVAL	Analysis Value		See Value-Level Metadata
CHG	Change from Baseline		Derived: AVAL – BASE. It is populated for all post-baseline records.
DTYPE	Derivation Type	["AVERAGE" = "Average"] <Derivation Type>	Derived: Value is AVERAGE for created records added for each visit as the average of the triplicate values collected for each parameter.
ABLFL	Baseline Record Flag	"Y"="Yes"	Derived: For each subject and parameter the baseline flag is set to Y for the last record prior to treatment start where DTYPE="AVERAGE"
TRTA	Actual Treatment	Placebo Active 20mg Active 40mg <Actual Treatment>	Assigned: Value of ADSL.TRT01A for a particular subject. Populate for baseline and post-baseline analyzed records
SAFFL	Safety Population Flag	["N"="No", "Y"="Yes"] <No Yes Response>	Predecessor: ADSL.SAFFL

Table 2.7.3.3. Value-level Metadata for ADEG

Parameter Value List - ADEG [AVAL]

Variable	Where	Type	Length/Display Format	Controlled Terms/ Formats	Source/Derivation/Comment
AVAL	DTYPE='AVERAGE'	Float	5.1		DERIVED: Average of the triplicate values collected at each visit for the parameter.
AVAL	DTYPE Not Equal 'AVERAGE'	Integer	3		Predecessor: EG.EGSTRESN

2.7.4 Input and Analysis Data 7

Table 2.7.4.1. EG Sample Records

Row	USUBJID	EGSEQ	EGREPNUM	EGTESTCD	EGSTRESN	EGSTRESU	EGBLFL	VISIT	EGDTC
1	XYZ-1001	1	1	QTCFAG	385	msec		SCREENING	2016-02-24T07:50:16
2	XYZ-1001	2	2	QTCFAG	399	msec		SCREENING	2016-02-24T07:52:59
3	XYZ-1001	3	3	QTCFAG	396	msec	Y	SCREENING	2016-02-24T07:56:07
4	XYZ-1001	4	1	QTCFAG	384	msec		VISIT 2	2016-03-08T09:45:11
5	XYZ-1001	5	2	QTCFAG	393	msec		VISIT 2	2016-03-08T09:48:07
6	XYZ-1001	6	3	QTCFAG	388	msec		VISIT 2	2016-03-08T09:51:04
7	XYZ-1001	7	1	QTCFAG	385	msec		VISIT 3	2016-03-22T10:45:03
8	XYZ-1001	8	2	QTCFAG	394	msec		VISIT 3	2016-03-22T10:48:07
9	XYZ-1001	9	3	QTCFAG	402	msec		VISIT 3	2016-03-22T10:51:05
10	XYZ-1002	1	1	QTCFAG	399	msec		SCREENING	2016-02-22T07:55:02
11	XYZ-1002	2	2	QTCFAG	410	msec		SCREENING	2016-02-22T07:58:05
12	XYZ-1002	3	3	QTCFAG	392	msec	Y	SCREENING	2016-02-22T08:01:06
13	XYZ-1002	4	1	QTCFAG	401	msec		VISIT 2	2016-03-06T09:50:04
14	XYZ-1002	5	2	QTCFAG	407	msec		VISIT 2	2016-03-06T09:53:51
15	XYZ-1002	6	3	QTCFAG	400	msec		VISIT 2	2016-03-06T09:56:21
16	XYZ-1002	7	1	QTCFAG	412	msec		VISIT 3	2016-03-24T10:50:07
17	XYZ-1002	8	2	QTCFAG	414	msec		VISIT 3	2016-03-24T10:53:08
18	XYZ-1002	9	3	QTCFAG	402	msec		VISIT 3	2016-03-24T10:56:05

Table 2.7.4.2. ADEG Sample Records

Row		USUBJID	EGSEQ	EGREPNUM	PARAM	VISIT	AVISIT	EGDTC	BASE	AVAL	CHG	DTYPE	ABLFL	TRTA	SAFFL
1		XYZ-1001	1	1	QTcF Interval (msec)	SCREENING	Baseline	2016-02-24T07:50:16		385					Y
2		XYZ-1001	2	2	QTcF Interval (msec)	SCREENING	Baseline	2016-02-24T07:52:59		399					Y
3		XYZ-1001	3	3	QTcF Interval (msec)	SCREENING	Baseline	2016-02-24T07:56:07		396					Y
4		XYZ-1001			QTcF Interval (msec)	SCREENING	Baseline			393.3		AVERAGE	Y	Placebo	Y
5		XYZ-1001	4	1	QTcF Interval (msec)	VISIT 2	Visit 2	2016-03-08T09:45:11	393.3	384	-9.3			Placebo	Y
6		XYZ-1001	5	2	QTcF Interval (msec)	VISIT 2	Visit 2	2016-03-08T09:48:07	393.3	393	-0.3			Placebo	Y
7		XYZ-1001	6	3	QTcF Interval (msec)	VISIT 2	Visit 2	2016-03-08T09:51:04	393.3	388	-5.3			Placebo	Y
8		XYZ-1001			QTcF Interval (msec)	VISIT 2	Visit 2		393.3	388.3	-5.0	AVERAGE		Placebo	Y
9		XYZ-1001	7	1	QTcF Interval (msec)	VISIT 3	Visit 3	2016-03-22T10:45:03	393.3	385	-8.3			Placebo	Y
10		XYZ-1001	8	2	QTcF Interval (msec)	VISIT 3	Visit 3	2016-03-22T10:48:07	393.3	394	0.7			Placebo	Y
11		XYZ-1001	9	3	QTcF Interval (msec)	VISIT 3	Visit 3	2016-03-22T10:51:05	393.3	402	8.7			Placebo	Y
12		XYZ-1001			QTcF Interval (msec)	VISIT 3	Visit 3		393.3	393.7	0.3	AVERAGE		Placebo	Y
13		XYZ-1002	1	1	QTcF Interval (msec)	SCREENING	Baseline	2016-02-22T07:58:05		399					Y
14		XYZ-1002	2	2	QTcF Interval (msec)	SCREENING	Baseline	2016-02-22T07:58:05		410					Y
15		XYZ-1002	3	3	QTcF Interval (msec)	SCREENING	Baseline	2016-02-22T08:01:06		392					Y
16		XYZ-1002			QTcF Interval (msec)	SCREENING	Baseline			400.3		AVERAGE	Y	Active 20mg	Y
17		XYZ-1002	4	1	QTcF Interval (msec)	VISIT 2	Visit 2	2016-03-06T09:50:04	400.3	401	0.7			Active 20mg	Y
18		XYZ-1002	5	2	QTcF Interval (msec)	VISIT 2	Visit 2	2016-03-06T09:53:51	400.3	407	6.7			Active 20mg	Y
19		XYZ-1002	6	3	QTcF Interval (msec)	VISIT 2	Visit 2	2016-03-06T09:56:21	400.3	400	-0.3			Active 20mg	Y
20		XYZ-1002			QTcF Interval (msec)	VISIT	Visit 2		400.3	402.7	2.3	AVERAGE		Active 20mg	Y
21		XYZ-1002	7	1	QTcF Interval (msec)	VISIT 3	Visit 3	2016-03-24T10:50:07	400.3	412	11.7			Active 20mg	Y
22		XYZ-1002	8	2	QTcF Interval (msec)	VISIT 3	Visit 3	2016-03-24T10:53:08	400.3	414	13.7			Active 20mg	Y
23		XYZ-1002	9	3	QTcF Interval (msec)	VISIT 3	Visit 3	2016-03-24T10:56:05	400.3	402	1.7			Active 20mg	Y
24		XYZ-1002			QTcF Interval (msec)	VISIT 3	Visit 3		400.3	409.3	9.0	AVERAGE		Active 20mg	Y

2.7.5 Other Uses 7

This example demonstrates how to maintain traceability when creating new records in ADaM BDS datasets. The traceability is both metadata-driven (i.e., variable-level metadata defining how the new record is derived) and data-driven (e.g., maintaining the source records and variables from SDTM in the ADEG dataset, such as EGSEQ). If this dataset is used as an input to another analysis dataset then ASEQ will need to be included. There are numerous use cases similar to this example, including:

- An endpoint analysis where new records are created for an analysis visit of Endpoint using a derivation type (DTPYE) of LOCF or WOCF
- Creating a composite endpoint
- Interpolating missing values (when not using mixed-modeling methods)

2.8 Using an Intermediate Dataset for BDS Traceability

The use of an intermediate dataset can help to gather all the necessary components needed for the creation of analysis endpoints in one location.

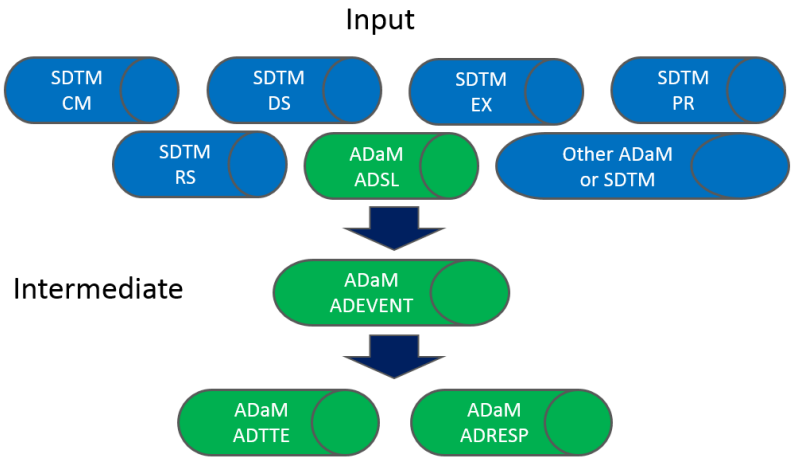
2.8.1 Analysis Need 8

For complex derivations of analysis endpoints, it may be necessary and beneficial to identify and capture all possible events associated with the derivation for a patient in a single dataset. These events could be an assessment date, protocol violation, prohibited medication, or other intervention. The creation of an intermediate dataset aids in the review and understanding of the analysis endpoints that it supports. Using the example found in the Therapeutic Area Data Standards User Guide for Breast Cancer (available at https://www.cdisc.org/standards/therapeutic-areas/breast-cancer) Section 5.3, an intermediate dataset that supports both the time-to-event dataset and the best response dataset is created.

2.8.2 Data Flow 8

In this following example, the intermediate dataset (ADEVENT) contains all events associated with the derivation of 2 different analysis datasets: time-to-event (ADTTE) and best overall response (ADRESP). Both the ADTTE and ADRESP datasets are based on the RECIST assessment. However, ADTTE is also based on other elements such as date of randomization, the use of prohibited medication, or disposition. Figure 2.8.2.1 illustrates the various input datasets that can be used to create the intermediate dataset, ADEVENT.

Figure 2.8.2.1. Intermediate Dataset Data Flow



Note: For this data flow, "Other SDTM or ADaM" datasets pertain to any dataset that will contain an event that is necessary for the derivation of the various analysis endpoints. This is not an implication that all SDTM or ADaM data need to be incorporated into one dataset; only the ones used for the derivations should be included.

2.8.3 Traceability Metadata 8

Table 2.8.3.1. Variable Metadata for ADEVENT

ADEVENT Variable Metadata

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Variable Metadata
STUDYID	Study Identifier	Char		xx.STUDYID that corresponds to input dataset used for USUBJID
USUBJID	Unique Subject Identifier	Char		xx.USUBJID that corresponds to the input dataset used for the assignment of AVAL
Treatment Variable[a]	Treatment Variable Label	Char		Treatment/Treatment Group that will be used in the analysis. At least one treatment variable needs to be included in the dataset.
ASEQ	Analysis Sequence Number	Num		Sequential number for associating a record number in the ADEVENT dataset. A unique number per subject, per parameter, per parameter qualifier, per analysis start date.
ASTDT	Analysis Start Date	Num		The date that the event occurred is the corresponding --DTC variable for each PARAMCD converted to numeric date format. RS.RSDTC when PARAMCD = "ASSESSI" or "ASSESSC" DS.DSSTDTC when PARAMCD = "DISPOSIT" AE.AESTDTC or MH.MHSTDTC or DV.DVSTDTC or CM.CMSTDTC, or PR.PRSTDTC or some other source data for an event which prevents further assessments when PARAMCD = "EVENT".
ASTDY	Analysis Start Relative Day	Num		The number of days from randomization to the date of the reported event. ASTDT – ADSL.RANDDT + 1
PARAM	Parameter	Char	DISPOSITION; ASSESSMENT - INVESTIGATOR; ASSESSMENT - CENTRAL; TREATMENT; EVENT	Set using PARAMCD "DISPOSIT"="DISPOSITION" "ASSESSI"="ASSESSMENT - INVESTIGATOR" "ASSESSC"="ASSESSMENT - CENTRAL" "TRTM"="TREATMENT" "EVENT"="EVENT"
PARAMCD	Parameter Code	Char	DISPOSIT; ASSESSI; ASSESSC; TRTM; EVENT	If RECIST assessment then PARAMCD = "ASSESS" For investigator based tumor response assessments, append "I" to the PARAMCD. For central imaging tumor response assessments, append "C" to the PARAMCD. If disposition event then PARAMCD = "DISPOSIT" If study treatment then PARAMCD = "TRTM" If event that is a protocol violation or prevents further assessments then PARAMCD = "EVENT".
AVALC	Analysis Value (C)	Char		Reported Assessment associated with the ASTDT.
SRCDOM	Source Data	Char		See parameter-level metadata (Table 2.8.3.4)
SRCVAR	Source Variable	Char		See parameter-level metadata (Table 2.8.3.4)
SRCSEQ	Source Sequence Number	Num		The sequence number --SEQ or ASEQ of the row in the dataset identified in the SRCDOM that relates to the analysis value being derived.
ANL01FL	Analysis Flag 01 Use in TTE	Char	Y	Identifies whether the event can be used in time-to-event analysis. If assessment is prior to baseline or after a censoring event then they are not included.

[a]Variables are not included in the sample input data (see Section 2.8.4, Input and Analysis Data 8).

Table 2.8.3.2. Variable Metadata for ADTTE

ADTTE Variable Metadata

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Variable Metadata
STUDYID	Study Identifier	Char		ADEVENT.STUDYID
USUBJID	Unique Subject Identifier	Char		ADEVENT.USUBJID
Treatment Variable[a]	Treatment Variable Label	Char		Treatment/Treatment Group that will be used in the analysis. At least one treatment variable needs to be included in the dataset.
PARAM	Parameter	Char	Progression-free Survival - Investigator; Progression-free Survival - Central; Overall Survival; Duration of Response	Set to value in the codelist that corresponds to PARAMCD.
PARAMCD	Parameter Code	Char	PFSI; PFSC; OS; DOR	If progression-free survival, then set to "PFS". For investigator based tumor response assessments, append 'I' to the PARAMCD. For central imaging tumor response assessments, append 'C' to the PARAMCD. If overall survival, then set to "OS". If duration of response, then set to "DOR".
AVAL	Analysis Value	Num		The number of days from the reference start date to the date of the assessment, disposition, or event recorded. See parameter-level metadata below

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Variable Metadata
CNSR	Censored	Num	1; 0	Censoring Value. Set to 1 if value is censored based on rules. See parameter-level metadata (Table 2.8.3.4)
EVNTDESC	Event or Censoring Description	Char	DOCUMENTED PROGRESSION; DEATH; DISEASE PROGRESSED; CENSORED AT TIME OF LAST ASSESSMENT	See parameter-level metadata (Table 2.8.3.4)
SRCDOM[a]	Source Data	Char	ADEVENT	SRCDOM = "ADEVENT"
SRCVAR[a]	Source Variable	Char	ASTDY	If PARAMCD = "DOR", then SRCVAR is null. Otherwise SRCVAR = "ASTDY"
SRCSEQ	Source Sequence Number	Num		If PARAMCD = "DOR", then SRCSEQ is null. Otherwise SRCSEQ = ADEVENT.ASEQ

[a]Variables are not included in the sample input data (see Section 2.8.4, Input and Analysis Data 8).

Table 2.8.3.3. Variable Metadata for ADRESP

ADRESP Variable Metadata

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Variable Metadata
STUDYID	Study Identifier	Char		ADEVENT.STUDYID
USUBJID	Unique Subject Identifier	Char		ADEVENT.USUBJID
Treatment Variable[a]	Treatment Variable Label	Char		Treatment/Treatment Group that will be used in the analysis. At least one treatment variable needs to be included in the dataset.
PARAM	Parameter	Char	Best Overall Response - Investigator; Best Overall Response - Central	Set based on PARAMCD "BORI"="Best Overall Response - Investigator" "BORC"="Best Overall Response - Central"
PARAMCD	Parameter Code	Char	BORI; BORC	For best overall response, set to "BOR". For investigator based tumor response assessments, append "I" to the PARAMCD. For central imaging tumor response assessments, append "C" to the PARAMCD.
AVAL	Analysis Value	Num	1; 2; 3; 4	Numerical ranking of the results. Where PARAMCD in ("BORI" "BORC") If AVALC = "CR" then AVAL = 1 If AVALC = "PR" then AVAL = 2 If AVALC = "SD" then AVAL = 3 If AVALC = "PD" then AVAL = 4
AVALC	Analysis Value (C)	Char	CR; PR; SD; PD	Where PARAMCD in ("BORI" "BORC") ADEVENT.ANL01FL = "Y" and ADEVENT.PARAMCD contains "ASSESS", then set to the best response where ranking from best to worst is "CR", "PR", "SD", "PD". Note that the suffix "I" or "C" must match between PARAMCD.
SRCDOM[a]	Source Data	Char	ADEVENT	SRCDOM = "ADEVENT"
SRCVAR[a]	Source Variable	Char	AVALC	SRCVAR = "AVALC"
SRCSEQ	Source Sequence Number	Num		ADEVENT.ASEQ

[a]Variables are not included in the sample input data (see Section 2.8.4, Input and Analysis Data 8).

Table 2.8.3.4. Value-level Metadata

Parameter Value List - ADEVENT, ADTTE

Dataset	Variable Name	Where	Type	Derivation/Comment
ADEVENT	SRCDOM	PARAMCD in ("ASSESSI" "ASSESSC")	Char	Set to "RS"
ADEVENT	SRCDOM	PARAMCD = "DISPOSIT"	Char	Set to "DS"
ADEVENT	SRCDOM	PARAMCD = "TRTM"	Char	Set to "EX"
ADEVENT	SRCDOM	PARAMCD = "EVENT"	Char	Set to "AE", "MH", "DV", "CM" or "PR" based on whether the source of ASTDT is AE.AESTDTC, MH.MHSTDTC, DV.DVSTDTC, CM.CMSTDTC, or PR.PRSTDTC
ADEVENT	SRCVAR	PARAMCD in ("ASSESSI" "ASSESSC")	Char	Set to "AVALC"
ADEVENT	SRCVAR	PARAMCD = "DISPOSIT"	Char	Set to "DSDECOD"
ADEVENT	SRCVAR	PARAMCD = "TRTM"	Char	Set to "EXTRT"
ADEVENT	SRCVAR	PARAMCD = "EVENT"	Char	Set to "AEDECOD", "MHTRT", "DVDECOD", "CMTRT" or "PRTRT" based on whether the source of AVALC is AE.AEDECOD, MH.MHTRT, DV.DVDECOD, CM.CMTRT, or PR.PRTRT

Dataset	Variable Name	Where	Type	Derivation/Comment
ADTTE	AVAL	PARAMCD in ("PFSI" "PFSC")	Num	ADEVENT.ASTDY when ADEVENT.ANL01FL = "Y" and ADEVENT.PARAMCD contains "ASSESS" and disease progressed, or when censored, use maximum ADEVENT.ASTDY where ADEVENT.ANL01FL = "Y" and ADEVENT.PARAMCD = "DISPOSIT". Note that the suffix "I" or "C" must match between PARAMCD.
ADTTE	AVAL	PARAMCD = "OS"	Num	ADEVENT.ASTDY when ADEVENT.ANL01FL = "Y" and ADEVENT.PARAMCD = "EVENT" and ADEVENT.AVALC ="'DEATH", or when censored, use maximum ADEVENT.ASTDY where ADEVENT.ANL01FL = "Y" and ADEVENT.PARAMCD = "DISPOSIT".
ADTTE	AVAL	PARAMCD = "DOR"	Num	Time from best response (CR or PR) to when disease progressed
ADTTE	CNSR	PARAMCD in ("PFSI" "PFSC")	Num	If disease did not progress, then CNSR = 1. Otherwise CNSR = 0.
ADTTE	CNSR	PARAMCD = "OS"	Num	If subject did not die, then CNSR = 1. Otherwise CNSR = 0.
ADTTE	CNSR	PARAMCD = "DOR"	Num	If after having best response the disease did not progress, then CNSR = 1. Otherwise CNSR = 0.
ADTTE	EVNTDESC	PARAMCD in ("PFSI" "PFSC")	Char	If CNSR = 0, then EVNTDESC = "DOCUMENTED PROGRESSION". Otherwise, if CNSR = 1, then EVNTDESC = "CENSORED AT TIME OF LAST ASSESSMENT".
ADTTE	EVNTDESC	PARAMCD = "OS"	Char	If CNSR = 0, then EVNTDESC = "DEATH". Otherwise, if CNSR = 1, then EVNTDESC = "CENSORED AT TIME OF LAST ASSESSMENT".
ADTTE	EVNTDESC	PARAMCD = "DOR"	Char	If CNSR = 0, then EVNTDESC = "DISEASE PROGRESSED". Otherwise, if CNSR = 1, then EVNTDESC = "CENSORED AT TIME OF LAST ASSESSMENT".

2.8.4 Input and Analysis Data 8

Table 2.8.4.1. SDTM CM Data

Row	STUDYID	USUBJID	CMSEQ	CMTRT	CMSTDTC
1	ABC-123	ABC-123-001	1	TAMOXIFEN	2014-03-31

Table 2.8.4.2. SDTM DS Data

Row	STUDYID	USUBJID	DSSEQ	DSTERM	DSDECOD	DSSTDTC
1	ABC-123	ABC-123-001	1	INFORMED CONSENT	INFORMED CONSENT OBTAINED	2013-12-03
2	ABC-123	ABC-123-001	2	RANDOMIZED	RANDOMIZED	2013-12-29
3	ABC-123	ABC-123-002	1	INFORMED CONSENT	INFORMED CONSENT OBTAINED	2013-10-25
4	ABC-123	ABC-123-002	2	RANDOMIZED	RANDOMIZED	2013-11-10

Table 2.8.4.3. SDTM EX Data

Row	STUDYID	USUBJID	EXSEQ	EXTRT	EXSTDTC	EXENDTC
1	ABC-123	ABC-123-001	1	DRUG A	2014-01-01	2014-01-01
2	ABC-123	ABC-123-001	2	DRUG A	2014-03-30	2014-03-30
3	ABC-123	ABC-123-002	1	DRUG B	2013-11-13	2013-11-13
4	ABC-123	ABC-123-002	2	DRUG B	2013-12-29	2013-12-29

Table 2.8.4.4. SDTM PR Data

Row	STUDYID	USUBJID	PRSEQ	PRTR	PRSTDTC
1	ABC-123	ABC-123-002	1	LUMPECTOMY	2013-12-14

Table 2.8.4.5. SDTM RS Data

Row	STUDYID	USUBJID	RSSEQ	RSTESTCD	RSSTRESC	RSEVAL	RSDTC
1	ABC-123	ABC-123-001	1	OVRLRESP	PD	INVESTIGATOR	2013-12-30
2	ABC-123	ABC-123-001	2	OVRLRESP	SD	CENTRAL	2013-12-31
3	ABC-123	ABC-123-001	3	OVRLRESP	SD	INVESTIGATOR	2014-01-21
4	ABC-123	ABC-123-001	4	OVRLRESP	SD	CENTRAL	2014-01-22
5	ABC-123	ABC-123-001	5	OVRLRESP	PR	INVESTIGATOR	2014-02-13
6	ABC-123	ABC-123-001	6	OVRLRESP	PR	CENTRAL	2014-02-14
7	ABC-123	ABC-123-001	7	OVRLRESP	PR	INVESTIGATOR	2014-03-06
8	ABC-123	ABC-123-001	8	OVRLRESP	PR	CENTRAL	2014-03-07
9	ABC-123	ABC-123-001	9	OVRLRESP	PD	INVESTIGATOR	2014-03-28
10	ABC-123	ABC-123-001	10	OVRLRESP	PD	CENTRAL	2014-03-29
11	ABC-123	ABC-123-002	1	OVRLRESP	PD	INVESTIGATOR	2013-11-11
12	ABC-123	ABC-123-002	2	OVRLRESP	PD	CENTRAL	2013-11-12
13	ABC-123	ABC-123-002	3	OVRLRESP	SD	INVESTIGATOR	2013-12-01
14	ABC-123	ABC-123-002	4	OVRLRESP	SD	CENTRAL	2013-12-02
15	ABC-123	ABC-123-002	5	OVRLRESP	PR	INVESTIGATOR	2013-12-27
16	ABC-123	ABC-123-002	6	OVRLRESP	PR	CENTRAL	2013-12-28

Table 2.8.4.6. ADaM ADEVENT Data

Row	STUDYID	USUBJID	ASEQ	ASTDT	ASTDY	PARAM	PARAMCD	AVALC	ANL01FL	SRCDOM	SRCVAR	SRCSEQ
1	ABC-123	ABC-123-001	1	2013-12-29	-4	DISPOSITION	DISPOSIT	RANDOMIZED		DS	DSDECOD	2
2	ABC-123	ABC-123-001	2	2013-12-30	-2	ASSESSMENT - INVESTIGATOR	ASSESSI	PD	Y	RS	RSSTRESC	1
3	ABC-123	ABC-123-001	3	2013-12-31	-1	ASSESSMENT - CENTRAL	ASSESSC	SD	Y	RS	RSSTRESC	2
4	ABC-123	ABC-123-001	4	2014-01-01	1	TREATMENT	TRTM	DRUG A	Y	EX	EXTRT	1
5	ABC-123	ABC-123-001	5	2014-01-21	20	ASSESSMENT - INVESTIGATOR	ASSESSI	SD	Y	RS	RSSTRESC	3
6	ABC-123	ABC-123-001	6	2014-01-22	22	ASSESSMENT - CENTRAL	ASSESSC	SD	Y	RS	RSSTRESC	4
7	ABC-123	ABC-123-001	7	2014-02-13	44	ASSESSMENT - INVESTIGATOR	ASSESSI	PR	Y	RS	RSSTRESC	5
8	ABC-123	ABC-123-001	8	2014-02-14	45	ASSESSMENT - CENTRAL	ASSESSC	PR	Y	RS	RSSTRESC	6
9	ABC-123	ABC-123-001	9	2014-03-06	65	ASSESSMENT - INVESTIGATOR	ASSESSI	PR	Y	RS	RSSTRESC	7
10	ABC-123	ABC-123-001	10	2014-03-07	66	ASSESSMENT - CENTRAL	ASSESSC	PR	Y	RS	RSSTRESC	8
11	ABC-123	ABC-123-001	11	2014-03-28	87	ASSESSMENT - INVESTIGATOR	ASSESSI	PD	Y	RS	RSSTRESC	9
12	ABC-123	ABC-123-001	12	2014-03-29	88	ASSESSMENT - CENTRAL	ASSESSC	PD	Y	RS	RSSTRESC	10
13	ABC-123	ABC-123-001	13	2014-03-30	89	TREATMENT	TRTM	DRUG A	Y	EX	EXTRT	2
14	ABC-123	ABC-123-001	14	2014-03-31	90	EVENT	EVENT	TAMOXIFEN		CM	CMTRT	1
15	ABC-123	ABC-123-002	1	2013-11-10	-3	DISPOSITION	DISPOSIT	RANDOMIZED		DS	DSDECOD	2
16	ABC-123	ABC-123-002	2	2013-11-11	-2	ASSESSMENT - INVESTIGATOR	ASSESSI	PD	Y	RS	RSSTRESC	1
17	ABC-123	ABC-123-002	3	2013-11-12	-1	ASSESSMENT - CENTRAL	ASSESSC	PD	Y	RS	RSSTRESC	2
18	ABC-123	ABC-123-002	4	2013-11-13	1	TREATMENT	TRTM	DRUG B	Y	EX	EXTRT	1
19	ABC-123	ABC-123-002	5	2013-12-01	19	ASSESSMENT - INVESTIGATOR	ASSESSI	SD	Y	RS	RSSTRESC	3
20	ABC-123	ABC-123-002	6	2013-12-02	20	ASSESSMENT - CENTRAL	ASSESSC	SD	Y	RS	RSSTRESC	4
21	ABC-123	ABC-123-002	7	2013-12-14	32	EVENT	EVENT	LUMPECTOMY		PR	PRTRT	1
22	ABC-123	ABC-123-002	8	2013-12-27	45	ASSESSMENT - INVESTIGATOR	ASSESSI	PR		RS	RSSTRESC	5
23	ABC-123	ABC-123-002	9	2013-12-28	46	ASSESSMENT - CENTRAL	ASSESSC	PR		RS	RSSTRESC	6
24	ABC-123	ABC-123-002	10	2013-12-29	47	TREATMENT	TRTM	DRUG B		EX	EXTRT	2

Table 2.8.4.7. ADaM ADTTE Data

Row	STUDYID	USUBJID	PARAM	PARAMCD	AVAL	CNSR	EVNTDESC	SRCSEQ
1	ABC-123	ABC-123-001	Progression-free Survival - Investigator	PFSI	87	0	DOCUMENTED PROGRESSION	11
2	ABC-123	ABC-123-001	Progression-free Survival - Central	PFSC	88	0	DOCUMENTED PROGRESSION	12
3	ABC-123	ABC-123-001	Overall Survival	OS	89	1	CENSORED AT TIME OF LAST ASSESSMENT	13
4	ABC-123	ABC-123-001	Duration of Response	DOR	44	0	DISEASE PROGRESSED	
5	ABC-123	ABC-123-002	Progression-free Survival - Investigator	PFSI	19	1	CENSORED AT TIME OF LAST ASSESSMENT	5
6	ABC-123	ABC-123-002	Progression-free Survival - Central	PFSC	20	1	CENSORED AT TIME OF LAST ASSESSMENT	6
7	ABC-123	ABC-123-002	Overall Survival	OS	1	1	CENSORED AT TIME OF LAST ASSESSMENT	4

Table 2.8.4.8. ADaM ADRESP Data

Row	STUDYID	USUBJID	PARAM	PARAMCD	AVAL	AVALC	SRCSEQ
1	ABC-123	ABC-123-001	Best Overall Response - Investigator	BORI	2	PR	7
2	ABC-123	ABC-123-001	Best Overall Response - Central	BORC	2	PR	8
3	ABC-123	ABC-123-002	Best Overall Response - Investigator	BORI	3	SD	5
4	ABC-123	ABC-123-002	Best Overall Response - Central	BORC	4	SD	6

Note: The use of the intermediate dataset ADEVENT and sorting it by ADT makes it easy to review all events of interest in order and determine why a specific value was used for an endpoint or why a subject was censored. For example, patient ABC-123-002 shows a partial response in the Disease Response and Clin Classification (RS) domain (Table 2.8.4.5). However, in ADTTE (Table 2.8.4.7) rows 5 and 6, the patient is censored and in ADRESP (Table 2.8.4.8) rows 3 and 4 the best response is indicated to be stable disease. Referring back to the intermediate dataset, ADEVENT (Table 2.8.4.6), it is evident on row 21 that the patient had a LUMPECTOMY, which is a prohibited medical procedure. Therefore, any assessments after the procedure would not be eligible for the determination of the endpoint.

2.8.5 Other Uses 8

The intermediate dataset allows reviewers to consider the impact on the analysis if an alternate value is selected. In addition, it allows for the analysis endpoint to use a different date and/or value to perform sensitivity analyses.

2.9 Using an Intermediate Dataset for ADSL Traceability

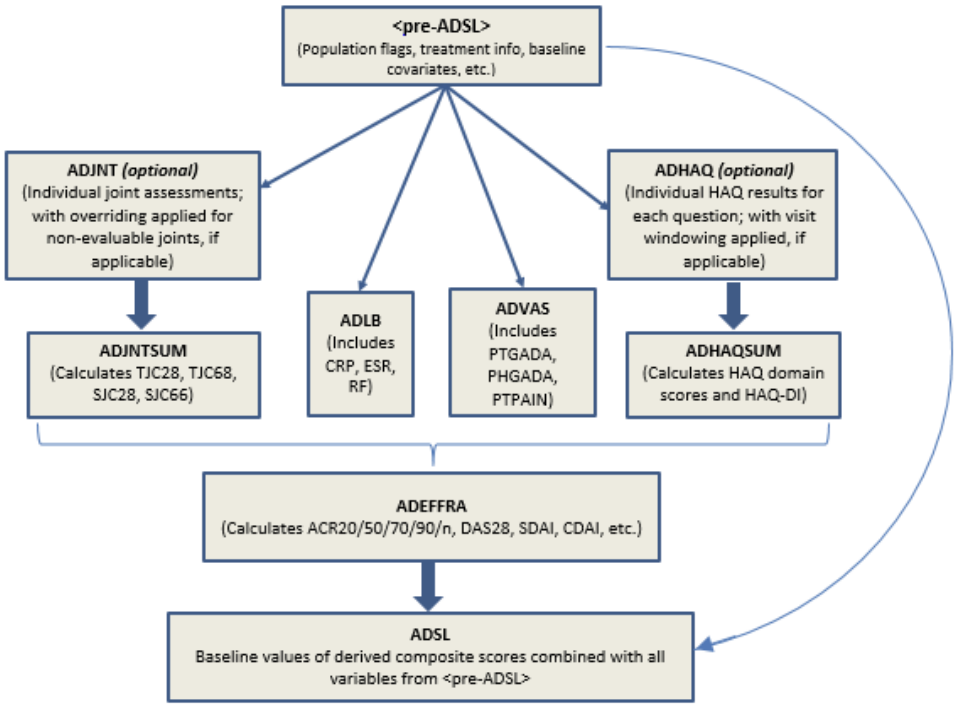
Sometimes it is difficult to create ADSL first, due to complex derivations that are difficult to present clearly and with sufficient traceability in a 1-record-per-subject dataset structure. In such cases, it is possible to create supportive analysis datasets before creating ADSL.

2.9.1 Analysis Need 9

This example describes a possible approach for the order of creation of the analysis datasets typically used in rheumatoid arthritis (RA) studies, as described in the RA Therapeutic Area User Guide (TAUG; available at https://www.cdisc.org/standards/therapeutic-areas/). Other than ADSL, the dataset names used in this flow chart are meaningful examples of ADaM-compliant names, but are not being proposed as a standard naming convention. These dataset names are used for reference in the discussion in subsequent sections. The intermediate dataset ADEFFRA is created to hold several highly derived baseline values typically found in RA studies which are used for subgroup analyses, and thus need to ultimately be stored in ADSL.

2.9.2 Data Flow 9

Figure 2.9.2.1. Intermediate Dataset Data Flow



As an initial step, <pre-ADSL> is created. The dataset name would begin with the letters "AD" to be ADaM-compliant. This dataset contains a subset of the variables found in the final ADSL, including population flags and treatment information, but without the RA-specific baseline values to be calculated later. <pre-ADSL> takes the place of ADSL as the source of ADSL variables for all other ADaM datasets created and ADSL is created at the end of this process.

It is the sponsor's decision whether to include <pre-ADSL> in a regulatory submission. If <pre-ADSL> is not submitted, it cannot be used as a reference in metadata.

After <pre-ADSL>, there are 2 datasets that are optional: ADJNT (the analysis datasets for tender/swollen joints) and ADHAQ (Health Assessment Questionnaire). The dataset ADJNT, if present, would be used to create ADJNTSUM, which contains the derived summary joint scores. Similarly, ADHAQ would be used to create ADHAQSUM, which would contain the 8 subdomain scores and the overall HAQ-DI score. Reasons why it may be advantageous to create these intermediate datasets include imputation or the need to perform significant visit reassignment. In cases where ADJNT or ADHAQ is a simple transformation of the source SDTM data into an ADaM BDS dataset, they may be unnecessary.

The ADaM dataset (ADEFFRA) is created from joint count calculations (ADJNTSUM), HAQ-DI score (ADHAQSUM), CRP and/or ESR values (ADLB), and VAS scores (ADVAS). Composite endpoints, such as ACRx scores and the DAS28 scores, can be derived from these data. It is recommended that as much as possible is harmonized across those four predecessor datasets to simplify the combining of the four datasets. For example, applying the same visit windowing algorithms and consistent use of analysis flags within each of those 4 predecessor ADaM datasets will facilitate combining by USUBJID and AVISIT as the first step towards the derivation of the composite scores for each visit.

2.9.3 Input and Analysis Data 9

Here are some sample datasets that could be produced in order to obtain the final ADSL dataset following the preceding process flow.

Table 2.9.3.1. Sample Pre-ADSL

adslpre.xpt

Row	STUDYID	USUBJID	TRTSDT	TRTEDT	TRT01P	ITTFL	SAFFL
1	ABC	10-001	01Feb2017	18Jul2017	<Active Drug> 100mg	Y	Y

In the following tables, the baseline values from ADJNTSUM in **bold** have been transposed into subject-level variables ending in -BL and added to both ADEFFRA and the final ADSL.

Table 2.9.3.2. Sample ADJNTSUM

adjntsum.xpt

Row	STUDYID	USUBJID	PARAMCD	ADT	AVISIT	AVAL	ABLFL	BASE	CHG	PCHG
1	ABC	10-001	SJC28	28Jan2017	Baseline	**17**	Y	17		
2	ABC	10-001	SJC28	28Feb2017	Week 4	16		17	-1	-5.88
3	ABC	10-001	SJC28	28Mar2017	Week 8	14		17	-3	-17.65
4	ABC	10-001	SJC28	25Apr2017	Week 12	12		17	-5	-29.41
5	ABC	10-001	SJC28	23May2017	Week 16	9		17	-8	-47.06
6	ABC	10-001	SJC28	20Jun2017	Week 20	8		17	-9	-52.94
7	ABC	10-001	SJC28	18Jul2017	Week 24	5		17	-12	-70.59
8	ABC	10-001	TJC28	28Jan2017	Baseline	**15**	Y	15		
9	ABC	10-001	TJC28	28Feb2017	Week 4	14		15	-1	-6.67
10	ABC	10-001	TJC28	28Mar2017	Week 8	12		15	-3	-20.00
11	ABC	10-001	TJC28	25Apr2017	Week 12	10		15	-5	-33.33
12	ABC	10-001	TJC28	23May2017	Week 16	8		15	-7	-46.67
13	ABC	10-001	TJC28	20Jun2017	Week 20	6		15	-9	-60.00
14	ABC	10-001	TJC28	18Jul2017	Week 24	4		15	-11	-73.33
15	ABC	10-001	SJC66	28Jan2017	Baseline	**46**	Y	46		
16	ABC	10-001	SJC66	28Feb2017	Week 4	44		46	-2	-4.35
17	ABC	10-001	SJC66	28Mar2017	Week 8	42		46	-4	-8.70
18	ABC	10-001	SJC66	25Apr2017	Week 12	42		46	-4	-8.70
19	ABC	10-001	SJC66	23May2017	Week 16	35		46	-11	-23.91
20	ABC	10-001	SJC66	20Jun2017	Week 20	29		46	-17	-36.96
21	ABC	10-001	SJC66	18Jul2017	Week 24	22		46	-24	-52.17
22	ABC	10-001	TJC68	28Jan2017	Baseline	**39**	Y	39		
23	ABC	10-001	TJC68	28Feb2017	Week 4	38		39	-1	-2.56
24	ABC	10-001	TJC68	28Mar2017	Week 8	35		39	-4	-10.26
25	ABC	10-001	TJC68	25Apr2017	Week 12	28		39	-11	-28.21
26	ABC	10-001	TJC68	23May2017	Week 16	25		39	-14	-35.90
27	ABC	10-001	TJC68	20Jun2017	Week 20	20		39	-19	-48.72
28	ABC	10-001	TJC68	18Jul2017	Week 24	17		39	-22	-56.41

Table 2.9.3.3. Sample ADEFFRA

adeffra.xpt

Row	STUDYID	USUBJID	PARAMCD	AVISIT	AVALC	SJC28BL	TJC28BL	SJC66BL	TJC68BL	ITTFL
1	ABC	10-001	ACR20	Week 4	N	**17**	**15**	**46**	**39**	Y
2	ABC	10-001	ACR20	Week 8	N	**17**	**15**	**46**	**39**	Y
3	ABC	10-001	ACR20	Week 12	N	**17**	**15**	**46**	**39**	Y

Row	STUDYID	USUBJID	PARAMCD	AVISIT	AVALC	SJC28BL	TJC28BL	SJC66BL	TJC68BL	ITTFL
4	ABC	10-001	ACR20	Week 16	Y	17	15	46	39	Y
5	ABC	10-001	ACR20	Week 20	Y	17	15	46	39	Y
6	ABC	10-001	ACR20	Week 24	Y	17	15	46	39	Y
7	ABC	10-001	ACR50	Week 4	N	17	15	46	39	Y
8	ABC	10-001	ACR50	Week 8	N	17	15	46	39	Y
9	ABC	10-001	ACR50	Week 12	N	17	15	46	39	Y
10	ABC	10-001	ACR50	Week 16	N	17	15	46	39	Y
11	ABC	10-001	ACR50	Week 20	N	17	15	46	39	Y
12	ABC	10-001	ACR50	Week 24	Y	17	15	46	39	Y
13	ABC	10-001	ACR70	Week 4	N	17	15	46	39	Y
14	ABC	10-001	ACR70	Week 8	N	17	15	46	39	Y
15	ABC	10-001	ACR70	Week 12	N	17	15	46	39	Y
16	ABC	10-001	ACR70	Week 16	N	17	15	46	39	Y
17	ABC	10-001	ACR70	Week 20	N	17	15	46	39	Y
18	ABC	10-001	ACR70	Week 24	N	17	15	46	39	Y
19	ABC	10-001	ACR90	Week 4	N	17	15	46	39	Y
20	ABC	10-001	ACR90	Week 8	N	17	15	46	39	Y
21	ABC	10-001	ACR90	Week 12	N	17	15	46	39	Y
22	ABC	10-001	ACR90	Week 16	N	17	15	46	39	Y
23	ABC	10-001	ACR90	Week 20	N	17	15	46	39	Y
24	ABC	10-001	ACR90	Week 24	N	17	15	46	39	Y

Table 2.9.3.4. Sample ADSL

adsl.xpt

Row	STUDYID	USUBJID	TRTSDT	TRTEDT	TRT01P	ITTFL	SAFFL	SJC28BL	TJC28BL	SJC66BL	TJC68BL
1	ABC	10-001	01Feb2017	18Jul2017	<Active Drug> 100mg	Y	Y	17	15	46	39

2.10 Traceability When Multiple Analysis Variables Are Needed on the Same Row

In cases of statistical modeling which feature multiple dependent and/or independent variables, statistical software requires all analysis variables to be in the same record for processing. In many cases, this requirement can be met by a BDS dataset using provided variables AVAL, BASE, and CHG, along with subject-level covariates from ADSL. However, this method is limited to 1 analysis variable per row as an intended feature of the BDS standard. The following example shows a way to support multiple analysis variables on 1 row and still maintain the ADaM principle of traceability.

The approach demonstrated is to first create a BDS dataset named ADQS, making use of traceability built into the BDS standard to explain the origin, derivation, imputation, and any other complexity behind each analysis value. The values of PARAMCD and PARAM are designed with the intention of using them as the variable name and label in a subsequent wide format dataset. The BDS dataset is finally transposed into a wide format (class of ADAM OTHER) dataset named ADQST to support statistical analysis and review.

This concept of creating a BDS and transposing is not new; it has been described in the ADaM Examples in Commonly Used Statistical Analysis Methods (available at https://www.cdisc.org/standards/foundational/adam/), Example 6 (Multivariate Analysis of Variance). Text in that section describes how a BDS dataset would need to be transposed in order to be truly analysis-ready.

2.10.1 Analysis Need 10

In this example, the scores of a motor function questionnaire are to be analyzed. There are multiple scores collected at baseline and after 1 month of treatment. Subtotal scores for upper and lower body mobility need to be derived. An exploratory correlation analysis is to be run on the change from baseline of the individual question scores grouped by treatment, and the derived subtotal scores. In order for the dataset to meet the ADaM fundamental principle of being analysis-ready for these correlations, and to meet additional requirements by the study statistician, the dataset must be in a horizontal form.

Table 2.10.1.1 shows a shell of the correlation analysis table to be produced.

Table 2.10.1.1. Correlation Table Shell

Table XX.X

Exploratory Correlation Analysis of Questionnaire Scores

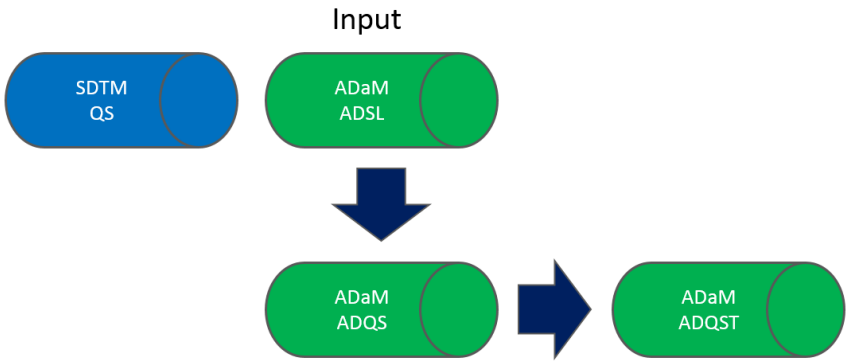
ITT Population

Change from Baseline at Month 1	Drug A (N=xx) n (P-value)	Drug B (N=xx) n (P-value)	Total (N=xx) n (P-value)
Correlation of Upper Body Score versus			
Lower Body	xx (x.xxx)	xx (x.xxx)	xx (x.xxx)
Question 1	xx (x.xxx)	xx (x.xxx)	xx (x.xxx)
Question 2	xx (x.xxx)	xx (x.xxx)	xx (x.xxx)
Question 3	xx (x.xxx)	xx (x.xxx)	xx (x.xxx)
Correlation of Lower Body Score versus			
Question 4	xx (x.xxx)	xx (x.xxx)	xx (x.xxx)
Question 5	xx (x.xxx)	xx (x.xxx)	xx (x.xxx)
Question 6	xx (x.xxx)	xx (x.xxx)	xx (x.xxx)

2.10.2 Data Flow 10

In Figure 2.10.2.1, ADaM dataset ADQS uses input data from the SDTM Questionnaires (QS) dataset and the ADSL; the ADaM dataset ADQST is derived directly from ADQS.

Figure 2.10.2.1. Example Data Flow When Transposing BDS



2.10.3 Traceability Metadata 10

Dataset-level, variable-level, and parameter-value-level metadata are shown in the following tables. The variable metadata for ADQS provides traceability to the source SDTM data variables and describes the process of deriving new derived records:

Table 2.10.3.1. Variable Metadata for ADQS

ADQS Variable Metadata

Variable Name	Where Condition	Variable Metadata
STUDYID		QS.STUDYID
USUBJID		QS.USUBJID
TRTP		ADSL.TRT01P
AVISIT		QS.VISIT
PARAMCD		Keep QS records where QS.QSCAT='MOTOR FUNCTION QUESTIONNAIRE' and QS.VISIT in ("BASELINE" "MONTH 1"), set PARAMCD=QS.QSTESTCD and PARAM=QS.QSTEST. For each visit, create 2 derived parameters with PARAMCD/PARAM values as: UPPER = Upper Body Motor Score, and LOWER = Lower Body Motor Score
PARAM		See PARAMCD
AVAL	PARAMCD not in ("UPPER" "LOWER")	QS.QSSTRESN

Variable Name	Where Condition	Variable Metadata
	PARAMCD in ("UPPER" "LOWER")	Where PARAMCD="UPPER", the sum of the scores for questions 1-3 Where PARAMCD="LOWER", the sum of the scores for questions 4-6 If any scores are missing, do not impute sum
ABLFL		Set to Y where AVISIT=BASELINE
BASE		AVAL value from the record where ABLFL=Y, populate for post-baseline records only
CHG		AVAL-BASE
QSSEQ		QS.QSSEQ
QSCAT		QS.QSCAT

Table 2.10.3.2. Dataset Metadata for ADQST

ADQST Dataset Metadata

Dataset	Description	Class	Structure	Description
ADQST	Transposed ADQS	ADAM OTHER	One record per subject	This dataset is derived from ADQS by transposing CHG by USUBJID, using the values of PARAMCD as new variable names and the values of PARAM as new variable labels

The variable metadata for ADQST is relatively simple, describing the transpose process and providing the predecessor origins for each variable:

Table 2.10.3.3. Variable Metadata for ADQST

ADQST Variable Metadata

Variable Name	Variable Label	Variable Metadata
STUDYID	Study Identifier	ADQS.STUDYID
USUBJID	Unique Subject Identifier	ADQS.USUBJID
TRTP	Planned Treatment	ADQS.TRTP
AVISIT	Analysis Visit	ADQS.VISIT Only keep records for Month 1
S01	Score 1	ADQS.CHG where PARAMCD="S01"
S02	Score 2	ADQS.CHG where PARAMCD="S02"
S03	Score 3	ADQS.CHG where PARAMCD="S03"
S04	Score 4	ADQS.CHG where PARAMCD="S04"
S05	Score 5	ADQS.CHG where PARAMCD="S05"
S06	Score 6	ADQS.CHG where PARAMCD="S06"
UPPER	Upper Body Score	ADQS.CHG where PARAMCD="UPPER"
LOWER	Lower Body Score	ADQS.CHG where PARAMCD="LOWER"

2.10.4 Input and Analysis Data 10

The following is an example of ADQS data. Only a few variables are included here to illustrate the example.

Table 2.10.4.1. Intermediate Data Example ADQS

Row	STUDYID	USUBJID	TRTP	AVISIT	PARAMCD	PARAM	AVAL	ABLFL	BASE	CHG	QSSEQ	QSCAT
1	XYZ	XYZ-001	DRUG A	BASELINE	S01	Score 1	40	Y			1	MOTOR FUNCTION QUESTIONNAIRE
2	XYZ	XYZ-001	DRUG A	MONTH 1	S01	Score 1	55		40	15	2	MOTOR FUNCTION QUESTIONNAIRE
3	XYZ	XYZ-001	DRUG A	BASELINE	S02	Score 2	30	Y			3	MOTOR FUNCTION QUESTIONNAIRE
4	XYZ	XYZ-001	DRUG A	MONTH 1	S02	Score 2	40		30	10	4	MOTOR FUNCTION QUESTIONNAIRE
5	XYZ	XYZ-001	DRUG A	BASELINE	S03	Score 3	45	Y			5	MOTOR FUNCTION QUESTIONNAIRE
6	XYZ	XYZ-001	DRUG A	MONTH 1	S03	Score 3	40		45	-5	6	MOTOR FUNCTION QUESTIONNAIRE
7	XYZ	XYZ-001	DRUG A	BASELINE	S04	Score 4	20	Y			7	MOTOR FUNCTION QUESTIONNAIRE
8	XYZ	XYZ-001	DRUG A	MONTH 1	S04	Score 4	30		20	10	8	MOTOR FUNCTION QUESTIONNAIRE
9	XYZ	XYZ-001	DRUG A	BASELINE	S05	Score 5	50	Y			9	MOTOR FUNCTION QUESTIONNAIRE
10	XYZ	XYZ-001	DRUG A	MONTH 1	S05	Score 5	45		50	-5	10	MOTOR FUNCTION QUESTIONNAIRE
11	XYZ	XYZ-001	DRUG A	BASELINE	S06	Score 6	40	Y			11	MOTOR FUNCTION QUESTIONNAIRE
12	XYZ	XYZ-001	DRUG A	MONTH 1	S06	Score 6	40		40	0	12	MOTOR FUNCTION QUESTIONNAIRE
13	XYZ	XYZ-001	DRUG A	BASELINE	UPPER	Upper Body Score	115	Y				
14	XYZ	XYZ-001	DRUG A	MONTH 1	UPPER	Upper Body Score	135		115	20		

Row	STUDYID	USUBJID	TRTP	AVISIT	PARAMCD	PARAM	AVAL	ABLFL	BASE	CHG	QSSEQ	QSCAT
15	XYZ	XYZ-001	DRUG A	BASELINE	LOWER	Lower Body Score	110	Y				
16	XYZ	XYZ-001	DRUG A	MONTH 1	LOWER	Lower Body Score	115		110	5		

Note: In the sample data for ADQS shown in Table 2.10.4.1, records that originate from SDTM have a value in QSSEQ and records which are derived have no value in QSSEQ. Including variable QSSEQ enables identifying the exact source record from QS that was used for the row in ADQS.

The following is an example of ADQST data, where the values of PARAMCD and PARAM from Table 2.10.4.1 have become new variable names and labels, respectively.

Table 2.10.4.2. Transposed Data Example ADQST

Row	STUDYID	USUBJID	TRTP	AVISIT	S01	S02	S03	S04	S05	S06	UPPER	LOWER
1	Study Identifier	Unique Subject Identifier	Planned Treatment	Analysis Visit	Score 1	Score 2	Score 3	Score 4	Score 5	Score 6	Upper Body Score	Lower Body Score
2	XYZ	XYZ-001	DRUG A	MONTH 1	15	10	-5	10	-5	0	20	5
3	XYZ	XYZ-002	DRUG B	MONTH 1	0	5	20	15	5	5	25	25
4	XYZ	XYZ-003	DRUG A	MONTH 1	30	10	15	20	25	30	55	75
5	XYZ	XYZ-004	DRUG B	MONTH 1	-5	0	-10	0	5	5	-15	10
6	XYZ	XYZ-005	DRUG A	MONTH 1	10	0	5	-10	-5	0	15	-15
7	XYZ	XYZ-006	DRUG B	MONTH 1	10	5	0	0	5	5	15	10

The sample data for ADQST shown in Table 2.10.4.2 supports the needs of the statistical analysis, and through the dataset, variable, and parameter metadata it is possible to trace each analysis value to a specific record in ADQS, and from there to the source SDTM records. It is important to note that if ADQS was not produced and only ADQST provided, the traceability between source and analysis data would be lost.

2.10.5 Other Uses 10

This example demonstrated how each data point in a wide multiple analysis variables dataset can be traced back across derivations to its SDTM source through the use of variable metadata, and data point traceability provided by the BDS standard. The dataset ADQST is an ADaM OTHER class dataset because it needs more than 1 analysis value on the same row and lacks the required variables needed for other classes. It is possible to build this dataset into a BDS or OCCDS class dataset by adding the required variables. In such a case, the transposed parameters could serve as traceability to a complex derivation.

2.11 Traceability When Using a Look-up Table

When creating an analysis dataset, most information needed is available in or derived from SDTM variables. However, there are times when using a look-up table (LUT) can provide additional content, avoid lengthy "if-then-else" programming statements, and reduce data-entry errors. This example provides a use case for LUTs, sample data and metadata, and inclusion of the LUT in the submission as a misc. file. Please note the ADRG is another common alternative location to provide LUTs.

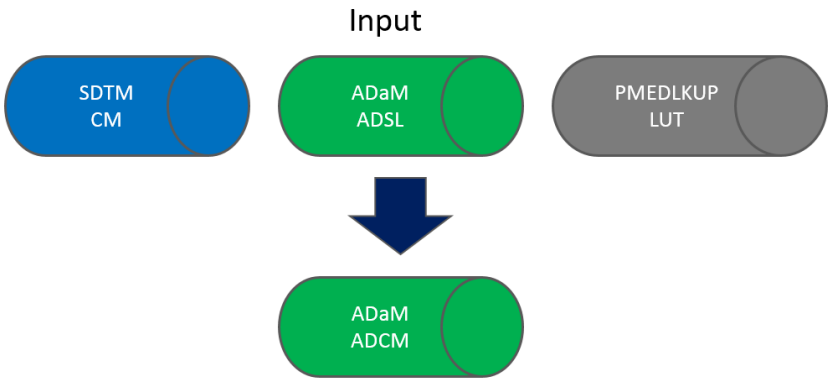
2.11.1 Analysis Need 11

In this example, prohibited medication is to be summarized. The protocol specifies certain classes of medications that are prohibited while the participant is on study. The number and percentage of subjects who had any prohibited medication during the study is summarized by the protocol-specified category.

2.11.2 Data Flow 11

In Figure 2.11.2.1, ADaM dataset ADCM uses as input data from the SDTM CM dataset, the ADSL, and the look-up table named PMEDLKUP.

Figure 2.11.2.1. Sample LUT Data Flow



To facilitate analysis, the study team reviewed the concomitant medications list and identified prohibited medication classes and their corresponding ATC codes, which are stored in look-up table PMEDLKUP with columns CATEGORY and CMCLASCD as prohibited medication class and ATC code, respectively. A single ATC class code for each concomitant medication is provided in CMCLASCD in the SDTM CM domain in this example, and is used to link with the look-up table to identify whether the medication belongs to a prohibited medication class.

2.11.3 Traceability Metadata 11

Table 2.11.3.1. Dataset Metadata for ADCM

ADCM Dataset Metadata

Dataset Name	Dataset Description	Dataset Structure	Class of Dataset
ADCM	Concomitant Medications Analysis Dataset	One record per subject per occurrence dataset created from merging CM, ADSL, and look-up table PMEDLKUP	OCCURRENCE DATA STRUCTURE

Variable-level metadata for ADCM is shown in Table 2.11.3.2. It provides traceability to the source SDTM and ADaM data variables, and describes the process of deriving ACAT1 using the look-up table.

Table 2.11.3.2. Variable Metadata for ADCM

ADCM Variable Metadata

Variable Name	Variable Label	Variable Metadata
STUDYID	Study Identifier	Predecessor: ADSL.STUDYID
USUBJID	Unique Subject Identifier	Predecessor: ADSL.USUBJID
CMSEQ	Sequence Number	Predecessor: CM.CMSEQ
CMDECOD	Standardized Medication Name	Predecessor: CM.CMDECOD
CMCLASCD	Medication Class Code	Predecessor: CM.CMCLASCD
ACAT1	Analysis Category 1	Derived: Populate by merging SDTM.CM with the look-up table, PMEDLKUP, by CMCLASCD. Set to the value of CATEGORY from the look-up table if available. Leave as null otherwise.

2.11.4 Input and Analysis Data 11

The following is an example of ADCM data, and the input data such as CM, ADSL, and the lookup table PMEDLKUP. Only a few variables are included here to illustrate the example.

Table 2.11.4.1. Example CM Data

cm.xpt

Row	STUDYID	DOMAIN	USUBJID	CMSEQ	CMDECOD	CMCLASCD
1	ABCD	CM	ABCD011001	1	METHYLPREDNISOLONE	H02AB
2	ABCD	CM	ABCD011001	2	ALPHARIX	J07BB
3	ABCD	CM	ABCD011001	3	BECLOMETHASONE	A07EA
4	ABCD	CM	ABCD011002	1	AMOXICILLINE	J01CA
5	ABCD	CM	ABCD011002	2	AZATHIOPRINUM	L04AX
6	ABCD	CM	ABCD021003	1	ADALIMUMAB	L04AB
7	ABCD	CM	ABCD021003	2	PREVENAR	J07AL

Table 2.11.4.2. Example Look-up Table PMEDLKUP

CATEGORY	CMCLASCD
Corticosteroid	A01AC
Corticosteroid	D07AC
Corticosteroid	H02AB
Corticosteroid	C05AA
Corticosteroid	A07EA
Corticosteroid	R01AD
Thiopurines	L01BB
Thiopurines	L04AX
Insulins	A10AD

Following US FDA guidance, look-up table PMEDLKUP could be submitted as a PDF (pmedlkup.pdf) or SAS transport file (pmedlkup.xpt) in the datasets/[study]/misc folder within the submission package.

Table 2.11.4.3. Example ADCM Data

Row	STUDYID	USUBJID	CMSEQ	CMDECOD	CMCLASCD	ACAT1
1	ABCD	ABCD011001	1	METHYLPREDNISOLONE	H02AB	Corticosteroid
2	ABCD	ABCD011001	2	ALPHARIX	J07BB	
3	ABCD	ABCD011003	3	BECLOMETHASONE	A07EA	Corticosteroid
4	ABCD	ABCD011002	1	AMOXICILLINE	J01CA	
5	ABCD	ABCD011002	2	AZATHIOPRINUM	L04AX	Thiopurines
6	ABCD	ABCD021003	1	ADALIMUMAB	L04AB	
7	ABCD	ABCD021003	2	PREVENAR	J07AL	

2.12 Complex Traceability Example

In cases where the parameter of interest for a statistical analysis is heavily derived, it may be impractical to create an ADaM dataset that supports this analysis directly from SDTM. This example demonstrates how an ADaM dataset containing a heavily derived parameter can be created from other ADaM datasets, with clear traceability for each data point.

The parameter of interest here is tumor lysis syndrome (TLS), a group of metabolic abnormalities that can occur as a complication during the treatment of cancer. During TLS, a large amounts of tumor cells are killed (lysed) at the same time by the treatment, releasing their contents into the bloodstream. This is a potentially fatal complication; if there is a risk, patients can be closely monitored before, during, and after their course of chemotherapy for TLS. The detection of TLS involves checking both lab results and adverse events; if specific criteria are met in a narrow time frame, TLS occurred.

The approach to derive TLS demonstrated here is to first create typical ADAE and ADLB datasets from the SDTM AE and LB datasets, respectively. In addition to what is needed in ADLB and ADAE to support other safety analyses, variables are included to identify records that might signal TLS. Afterward, an ADTLS dataset pulls these identified ADAE and ADLB records together and makes the determination whether TLS occurred. A derived parameter in ADTLS records the result of this check.

Note: The derivations in this example have been simplified from their originals for the purpose of clarity and the definitions of TLS are not intended to be guidance for deriving TLS in general.

2.12.1 Analysis Need 12

The analysis of TLS in this example comprises 2 parts: laboratory tumor lysis syndrome (LTLS) and clinical tumor lysis syndrome (CTLS).

LTLS was defined as 2 of the following lab results occurring simultaneously (within 24 hours of each other) in a time window starting from 3 days before to 7 days after chemotherapy:

- Uric acid > 8 mg/dL
- Potassium > 6 meq/L
- Phosphate > 4.5 mg/dL
- Calcium < 7 mg/dL

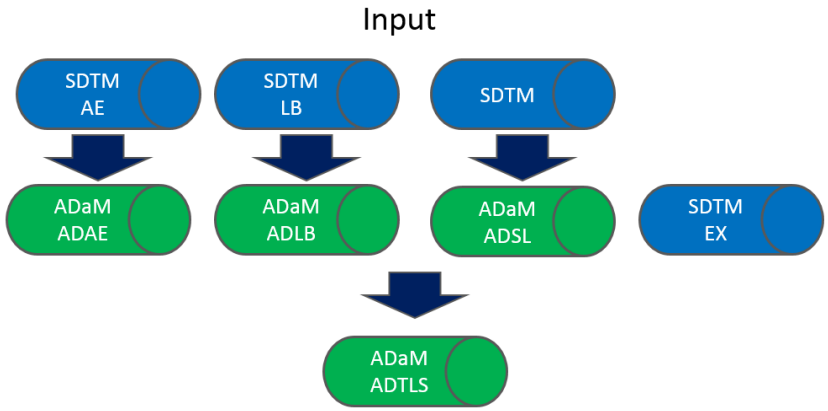
CTLS was defined as the presence of LTLS plus 1 of the following lab or adverse events occurring within a time window starting from the same day as the earlier of the LTLS lab events to 1 day after:

- Increased serum creatinine (> 1.5 x ULN)
- Cardiac arrhythmia or sudden death
- Seizure
- Any symptomatic hypocalcemia

2.12.2 Data Flow 12

In Figure 2.12.2.1, ADaM datasets ADAE and ADLB are derived from the SDTM AE and LB datasets to support normal safety analysis. They also contain elements that will simplify downstream TLS derivations. The ADaM dataset ADTLS will use both ADAE and ADLB as well as Exposure (EX) as input data.

Figure 2.12.2.1. Example Data Flow for Complex Derivation



2.12.3 Traceability Metadata 12

The following tables provide dataset-level, variable-level, and parameter-value-level metadata.

The variable metadata for ADLB provide traceability to the source SDTM records using LBSEQ, support downstream TLS derivations by evaluating relevant criteria with variables CRIT1 and CRIT1FL, and anticipate the need to trace to this dataset by including ASEQ. **Note:** Standard variables such as baseline, change from baseline, normal range, toxicity grade, and other common variables have been omitted from this example for simplicity.

Table 2.12.3.1. Variable Metadata for ADLB

ADLB Variable Metadata

Variable Name	Variable Label	Where Condition	Variable Metadata
STUDYID	Study Identifier		ADSL.STUDYID
USUBJID	Unique Subject Identifier		LB.USUBJID
LBSEQ	Sequence Number		LB.LBSEQ
ASEQ	Analysis Sequence Number		Sort data by USUBJID, TRTAN, AVISITN, PARAMN. For each USUBJID, set the first record to ASEQ = 1 and increment ASEQ by 1 for each record within USUBJID.
TRTA	Actual Treatment		ADSL.TRT01A
ADTM	Analysis Datetime		LB.LBDTC converted to numeric datetime
PARAM	Parameter		Concatenation of LB.LBTEST with LB.LBSTRESU
PARAMCD	Parameter Code		LB.LBTESTCD
AVAL	Analysis Value		LB.LBSTRESN
CRIT1	Analysis Criterion 1	PARAMCD="URATE"	set to "> 8.0 mg/dL"
		PARAMCD="PHOS"	set to "> 4.5 mg/dL"
		PARAMCD="K"	set to "> 6 mEq/L"
		PARAMCD="CACR"	set to "< 7.0 mg/dL"
		PARAMCD="CREAT"	set to "> 1.5 ULN"
CRIT1FL	Criterion 1 Evaluation Result Flag	PARAMCD="URATE"	if AVAL>8 then set "Y"; else null
		PARAMCD="PHOS"	if AVAL>4.5 then set "Y"; else null
		PARAMCD="K"	if AVAL>6 then set "Y"; else null
		PARAMCD="CACR"	if AVAL<7 then set "Y"; else null
		PARAMCD="CREAT"	if AVAL>1.5*ANRHI then set "Y"; else null Note: ANRHI (upper limit of normal ULN) is not displayed in this example

The variable metadata for ADAE provide traceability to the source SDTM records using AESEQ, support downstream TLS derivations by identifying AESIs relevant to TLS with variable CQ01NAM, and anticipate the need to trace to this dataset by including ASEQ. **Note:** Standard variables such as treatment emergence, AE severity, outcome, toxicity grade, and other common variables have been omitted from this example for simplicity.

Table 2.12.3.2. Variable Metadata for ADAE

ADAE Variable Metadata

Variable Name	Variable Label	Variable Metadata
STUDYID	Study Identifier	ADSL.STUDYID
USUBJID	Unique Subject Identifier	AE.USUBJID
TRTA	Actual Treatment	ADSL.TRT01A
AESEQ	Sequence Number	AE.AESEQ
ASEQ	Analysis Sequence Number	Sort data by USUBJID, ASTDTM, AEDECOD. For each USUBJID, set the first record to ASEQ = 1 and increment ASEQ by 1 for each record within USUBJID.
AETERM	Reported Term	AE.AETERM
AEDECOD	Dictionary-Derived Term	AE.AEDECOD
AEBODSYS	Body System or Organ Class	AE.AEBODSYS
ASTDTM	Analysis Start Date/Time	AE.AESTDTC converted to numeric datetime
AENDTM	Analysis End Date/Time	AE.AEENDTC converted to numeric datetime
CQ01NAM	Customized Query 01 Name	Set to "AESI for CTLS" if AEDECOD value is in the following list: Myopathy Laryngospasm Bronchospasm Oliguria ... Note: For a longer list, consider referencing a look-up table instead of listing individual terms here. See example in 2.11 for more details.

The dataset ADTLS pulls in TLS-related records from ADLB and ADAE, checks the times of these events against visit windows derived from cycle exposure dates from EX, and evaluates whether LTLS or CTLS have occurred. The metadata for ADTLS provides traceability to the source ADLB and ADAE records using SRCDOM and SRCSEQ, and newly created derived records with PARAMCD = "LTLS" and "CTLS" include vertical record traceability using variables ADTM and VISIT.

Table 2.12.3.3. Variable Metadata for ADTLS

ADTLS Variable Metadata

Variable Name	Variable Label	Where Condition	Variable Metadata
STUDYID	Study Identifier		ADSL.STUDYID
USUBJID	Unique Subject Identifier		ADSL.USUBJID
SRCDOM	Source Data		Copy records from ADLB where CRIT1FL="Y" and set SRCDOM="ADLB". Copy records from ADAE where CQ01NAM is not NULL and set SRCDOM="ADAE".
SRCSEQ	Source Sequence Number		ASEQ value from dataset indicated in SRCDOM
TRTA	Actual Treatment		ADSL.TRT01A
PARCAT1	Parameter Category 1	SRCDOM IN ("ADAE" "ADLB")	Set to "TLS Events"
		PARAMCD IN ("LTLS" "CTLS")	Set to "TLS Evaluation"
PARAMCD	Parameter Code	SRCDOM EQ "ADLB"	ADLB.PARAMCD
		SRCDOM EQ "ADAE"	Set to "AE"
		PARCAT1 EQ "TLS Evaluation"	Create two records for each subject, with PARAMCD values of "LTLS" and "CTLS"
PARAM	Parameter		One to one mapping of PARAMCD based on controlled terminology
AVALC	Analysis Value	SRCDOM EQ "ADLB"	Set to "Y"
		SRCDOM EQ "ADAE"	ADAE.AEDECOD
		PARAMCD EQ "LTLS"	If there are two distinct records with PARAMCD value in ("URATE" "K" "PHOS" "CACR") and AWSDTM not NULL and ADTM values 24 hours or less apart, then set AVALC to "Y", otherwise "N"
		PARAMCD EQ "CTLS"	If the subject is evaluated positive for LTLS, then take the date from ADTM. If a record with PARAMCD value in ("AE" "CREAT") with ADTM that falls on the same day of LTLS or day + 1, then set AVALC to "Y", otherwise "N"
ADTM	Analysis Datetime	SRCDOM EQ "ADLB"	ADLB.ADTM
		SRCDOM EQ "ADAE"	ADAE.ASTDTM
		PARCAT1 EQ "TLS Evaluation"	If AVALC="Y", then retain ADTM with the earliest datetime from the records that satisfy the TLS criteria
AVISIT	Analysis Visit	PARCAT1 EQ "TLS Events"	EX.VISIT, populate from EX record selected by AWSDTM derivation
		PARCAT1 EQ "TLS Evaluation"	If AVALC="Y", then retain VISIT value from the same record retained for ADTM
AWSDTM	Analysis Window Start Date/Time		For records where SRCDOM is in ("ADLB" "ADAE"), if ADTM falls between a cycle interval described by [EX.EXSTDTC-3days, EX.EXENDTC+7days], then use this record to populate EXSDTM, EXEDTM, AWSDTM, AWEDTM, and AVISIT. Populate AWSDTM with EXSDTM-3
AWEDTM	Analysis Window End Date/Time		See AWSDTM, EXEDTM+7
EXSDTM	Exposure Start Date/Time		See AWSDTM, numeric value of EX.EXSTDTC Note: If event does not fall into a cycle interval, then EXSDTM will remain null
EXEDTM	Exposure End Date/Time		See AWSDTM, numeric value of EX.EXENDTC Note: If event does not fall into a cycle interval, then EXEDTM will remain null

2.12.4 Input and Analysis Data 12

The following is an example of ADLB data. Only selected variables and events are shown. The actual ADAE dataset would have variables to support safety analysis (e.g., AESEV), and events other than those evaluated for TLS.

Table 2.12.4.1. Complex Traceability Example ADLB

Row	STUDYID	USUBJID	LBSEQ	ASEQ	TRTA	ADTM	PARAMCD	PARAM	AVAL	CRIT1	CRIT1FL
1	AZY300	AZY300-101-001	117	125	Drug A	15FEB2016:7:01	URATE	Uric Acid (mg/dL)	8.3	> 8.0 mg/dL	Y
2	AZY300	AZY300-101-001	118	126	Drug A	15FEB2016:7:01	PHOS	Phosphate (mg/dL)	4.9	> 4.5 mg/dL	Y
3	AZY300	AZY300-101-001	119	127	Drug A	15FEB2016:7:01	K	Potassium (mEq/L)	3.8	> 6 mEq/L	
4	AZY300	AZY300-101-001	120	128	Drug A	15FEB2016:7:01	CACR	Corrected Calcium (mg/dL)	8.7	< 7.0 mg/dL	
5	AZY300	AZY300-101-001	122	129	Drug A	15FEB2016:7:01	CREAT	Creatinine (mg/dL)	2.0	> 1.5 ULN	Y
6	AZY300	AZY300-101-003	133	143	Drug B + C	28FEB2016:14:14	URATE	Uric Acid (mg/dL)	4.0	> 8.0 mg/dL	
7	AZY300	AZY300-101-003	134	144	Drug B + C	28FEB2016:14:14	PHOS	Phosphate (mg/dL)	4.1	> 4.5 mg/dL	
8	AZY300	AZY300-101-003	135	145	Drug B + C	28FEB2016:14:14	K	Potassium (mEq/L)	3.6	> 6 mEq/L	
9	AZY300	AZY300-101-003	136	146	Drug B + C	28FEB2016:14:14	CACR	Corrected Calcium (mg/dL)	6.5	< 7.0 mg/dL	Y
10	AZY300	AZY300-101-003	138	147	Drug B + C	28FEB2016:14:14	CREAT	Creatinine (mg/dL)	2.1	> 1.5 ULN	Y
11	AZY300	AZY300-101-003	157	171	Drug B + C	02MAR2016:9:42	URATE	Uric Acid (mg/dL)	4.1	> 8.0 mg/dL	
12	AZY300	AZY300-101-003	158	172	Drug B + C	02MAR2016:9:42	PHOS	Phosphate (mg/dL)	5.0	> 4.5 mg/dL	Y
13	AZY300	AZY300-101-003	159	173	Drug B + C	02MAR2016:9:42	K	Potassium (mEq/L)	4.2	> 6 mEq/L	
14	AZY300	AZY300-101-003	160	174	Drug B + C	02MAR2016:9:42	CACR	Corrected Calcium (mg/dL)	8.8	< 7.0 mg/dL	
15	AZY300	AZY300-101-003	162	175	Drug B + C	02MAR2016:9:42	CREAT	Creatinine (mg/dL)	2.2	> 1.5 ULN	Y
16	AZY300	AZY300-101-024	155	183	Drug A	14MAR2016:16:16	URATE	Uric Acid (mg/dL)	5.2	> 8.0 mg/dL	
17	AZY300	AZY300-101-024	156	184	Drug A	14MAR2016:16:16	PHOS	Phosphate (mg/dL)	3.8	> 4.5 mg/dL	
18	AZY300	AZY300-101-024	157	185	Drug A	14MAR2016:16:16	K	Potassium (mEq/L)	6.3	> 6 mEq/L	Y
19	AZY300	AZY300-101-024	158	186	Drug A	14MAR2016:16:16	CACR	Corrected Calcium (mg/dL)	7.0	< 7.0 mg/dL	
20	AZY300	AZY300-101-024	160	187	Drug A	14MAR2016:16:16	CREAT	Creatinine (mg/dL)	0.9	> 1.5 ULN	
21	AZY300	AZY300-101-024	177	211	Drug A	15MAR2016:07:36	URATE	Uric Acid (mg/dL)	5.3	> 8.0 mg/dL	
22	AZY300	AZY300-101-024	178	212	Drug A	15MAR2016:07:36	PHOS	Phosphate (mg/dL)	3.9	> 4.5 mg/dL	
23	AZY300	AZY300-101-024	179	213	Drug A	15MAR2016:07:36	K	Potassium (mEq/L)	6.0	> 6 mEq/L	
24	AZY300	AZY300-101-024	180	214	Drug A	15MAR2016:07:36	CACR	Corrected Calcium (mg/dL)	6.5	< 7.0 mg/dL	Y
25	AZY300	AZY300-101-024	181	215	Drug A	15MAR2016:07:36	CREAT	Creatinine (mg/dL)	0.9	> 1.5 ULN	
26	AZY300	AZY300-101-344	205	232	Drug B + C	22APR2016:07:07	URATE	Uric Acid (mg/dL)	3.5	> 8.0 mg/dL	
27	AZY300	AZY300-101-344	206	233	Drug B + C	22APR2016:07:07	PHOS	Phosphate (mg/dL)	3.1	> 4.5 mg/dL	
28	AZY300	AZY300-101-344	207	234	Drug B + C	22APR2016:07:07	K	Potassium (mEq/L)	4.5	> 6 mEq/L	
29	AZY300	AZY300-101-344	208	235	Drug B + C	22APR2016:07:07	CACR	Corrected Calcium (mg/dL)	8.5	< 7.0 mg/dL	
30	AZY300	AZY300-101-344	210	236	Drug B + C	22APR2016:07:07	CREAT	Creatinine (mg/dL)	0.7	> 1.5 ULN	
31	AZY300	AZY300-101-411	203	221	Drug B + C	03APR2016:12:12	URATE	Uric Acid (mg/dL)	9.0	> 8.0 mg/dL	Y
32	AZY300	AZY300-101-411	204	222	Drug B + C	03APR2016:12:12	PHOS	Phosphate (mg/dL)	3.5	> 4.5 mg/dL	
33	AZY300	AZY300-101-411	205	223	Drug B + C	03APR2016:12:12	K	Potassium (mEq/L)	4.1	> 6 mEq/L	
34	AZY300	AZY300-101-411	206	224	Drug B + C	03APR2016:12:12	CACR	Corrected Calcium (mg/dL)	6.8	< 7.0 mg/dL	Y
35	AZY300	AZY300-101-411	208	225	Drug B + C	03APR2016:12:12	CREAT	Creatinine (mg/dL)	0.7	> 1.5 ULN	

In the preceding ADLB, the variables CRIT1 and CRIT1FL support downstream TLS derivations. Specifically, labs results that meet TLS criteria are flagged and can be easily extracted from this dataset with filter CRIT1FL="Y". The variable LBSEQ provides traceability to SDTM source records, and the variable ASEQ allows other ADaM datasets to trace to this dataset. Only a few variables are presented here.

The following is an example of ADAE data. Only selected variables and parameters are shown. The actual ADLB dataset would have variables to support safety analysis (e.g., AVAL, CHG), and parameters other than those evaluated for TLS.

Table 2.12.4.2. Complex Traceability Example ADAE

Row	STUDYID	USUBJID	TRTA	AESEQ	ASEQ	AETERM	AEDECOD	AEBODSYS	ASTDTM	AENDTM	CQ01NAM
1	AZY300	AZY300-101-024	Drug A	15	13	MYOPATHY (MUSCLE CRAMPS, STIFFNESS, AND SPASM)	Myopathy	Musculoskeletal and connective tissue disorders	15MAR2016:08:02	18MAR2016:14:35	AESI for CTLS
2	AZY300	AZY300-101-344	Drug B + C	10	9	LARYNGOSPASM (20 SEC SPASMS OF VOCAL CORDS BLOCKING BREATHING)	Laryngospasm	Respiratory, thoracic and mediastinal disorders	22APR2016:09:22	29APR2016:13:44	AESI for CTLS
3	AZY300	AZY300-101-411	Drug B + C	6	5	BRONCHOSPASM (NARROWING/CONSTRICTING OF AIRWAYS)	Bronchospasm	Immune system disorders	03APR2016:12:33	04APR2016:11:59	AESI for CTLS
4	AZY300	AZY300-101-422	Drug A	3	3	DECREASE IN URINE OUTPUT	Oliguria	Renal and urinary disorders	25MAY2016:14:14	31MAY2016:12:37	AESI for CTLS

In the preceding ADAE, the variables CQ01NAM supports downstream TLS derivations. Specifically, adverse events that are of special interest in TLS derivations are flagged for easy extraction from this dataset with filter CQ01NAM="AESI for CTLS". The variable AESEQ provides traceability to SDTM source records, and the variable ASEQ allows other ADaM datasets to trace to this dataset.

The sample EX data shown below serves as the last piece of input to the ADTLS dataset. For each treatment cycle, the infusion start and end times are recorded. Only a few variables and records relevant to the example are displayed.

Table 2.12.4.3. Complex Traceability Example EX

Row	STUDYID	USUBJID	EXSEQ	VISIT	EXSTDTC	EXENDTC
1	AZY300	AZY300-101-001	31	CYCLE 3 DAY 1	2016-02-15T09:25	2016-02-15T14:44
2	AZY300	AZY300-101-003	31	CYCLE 3 DAY 1	2016-03-01T08:33	2016-03-01T12:55
3	AZY300	AZY300-101-024	41	CYCLE 4 DAY 1	2016-03-15T08:55	2016-03-15T13:17
4	AZY300	AZY300-101-344	51	CYCLE 5 DAY 1	2016-04-22T08:14	2016-04-22T13:06
5	AZY300	AZY300-101-411	41	CYCLE 4 DAY 1	2016-03-20T07:58	2016-03-20T12:22
6	AZY300	AZY300-101-422	61	CYCLE 6 DAY 1	2016-05-15T08:15	2016-05-15T13:32

The following is an example of ADTLS data. To produce this dataset, analysis window start and end date/times are first calculated from EX infusion records. TLS-related records are then subset from ADLB and ADAE and the records are pulled into ADTLS under PARCAT1="TLS Events". If the lab date/time or AE start date/time fall into an analysis window, the corresponding exposure and windowing variables are populated. These records are then evaluated to determine if LTLS or CTLS occurred.

Table 2.12.4.4. Complex Traceability Example ADTLS

Row	STUDYID	USUBJID	SRCDOM	SRCSEQ	TRTA	PARCAT1	PARAMCD	PARAM	AVALC	ADTM	AVISIT	EXSDTM	EXEDTM	AWSDTM	AWEDTM
1	AZY300	AZY300-101-001	ADLB	125	Drug A	TLS Events	URATE	Hyperuricemia (> 8.0 mg/dL)	Y	15FEB2016:7:01	CYCLE 3 DAY 1	15FEB2016:09:25	15FEB2016:14:44	12FEB2016:09:25	22FEB2016:14:44
2	AZY300	AZY300-101-001	ADLB	126	Drug A	TLS Events	PHOS	Hyperphosphatemia (> 4.5 mg/dL)	Y	15FEB2016:7:01	CYCLE 3 DAY 1	15FEB2016:09:25	15FEB2016:14:44	12FEB2016:09:25	22FEB2016:14:44
3	AZY300	AZY300-101-001	ADLB	129	Drug A	TLS Events	CREAT	Elevated Creatinine (> 1.5 ULN)	Y	15FEB2016:7:01	CYCLE 3 DAY 1	15FEB2016:09:25	15FEB2016:14:44	12FEB2016:09:25	22FEB2016:14:44
4	AZY300	AZY300-101-001			Drug A	TLS Evaluation	LTLS	Laboratory Tumor Lysis Syndrome	Y	15FEB2016:7:01	CYCLE 3 DAY 1				
5	AZY300	AZY300-101-001			Drug A	TLS Evaluation	CTLS	Clinical Tumor Lysis Syndrome	Y	15FEB2016:7:01	CYCLE 3 DAY 1				
6	AZY300	AZY300-101-003	ADLB	146	Drug B + C	TLS Events	CACR	Hypocalcemia (< 7.0 mg/dL)	Y	28FEB2016:14:14	CYCLE 3 DAY 1	01MAR2016:08:33	01MAR2016:12:55	27FEB2016:08:33	08MAR2016:12:55
7	AZY300	AZY300-101-003	ADLB	147	Drug B + C	TLS Events	CREAT	Elevated Creatinine (> 1.5 ULN)	Y	28FEB2016:14:14	CYCLE 3 DAY 1	01MAR2016:08:33	01MAR2016:12:55	27FEB2016:08:33	08MAR2016:12:55
8	AZY300	AZY300-101-003	ADLB	172	Drug B + C	TLS Events	PHOS	Hyperphosphatemia (> 4.5 mg/dL)	Y	02MAR2016:9:42	CYCLE 3 DAY 1	01MAR2016:08:33	01MAR2016:12:55	27FEB2016:08:33	08MAR2016:12:55
9	AZY300	AZY300-101-003	ADLB	175	Drug B + C	TLS Events	CREAT	Elevated Creatinine (> 1.5 ULN)	Y	02MAR2016:9:42	CYCLE 3 DAY 1	01MAR2016:08:33	01MAR2016:12:55	27FEB2016:08:33	08MAR2016:12:55
10	AZY300	AZY300-101-003			Drug B + C	TLS Evaluation	LTLS	Laboratory Tumor Lysis Syndrome	N						
11	AZY300	AZY300-101-003			Drug B + C	TLS Evaluation	CTLS	Clinical Tumor Lysis Syndrome	N						
12	AZY300	AZY300-101-024	ADLB	185	Drug A	TLS Events	K	Hyperkalemia (> 6.0 mEq/L)	Y	14MAR2016:16:16	CYCLE 4 DAY 1	15MAR2016:08:55	15MAR2016:13:17	12MAR2016:08:55	22MAR2016:13:17
13	AZY300	AZY300-101-024	ADLB	214	Drug A	TLS Events	CACR	Hypocalcemia (< 7.0 mg/dL)	Y	15MAR2016:07:36	CYCLE 4 DAY 1	15MAR2016:08:55	15MAR2016:13:17	12MAR2016:08:55	22MAR2016:13:17
14	AZY300	AZY300-101-024	ADAE	13	Drug A	TLS Events	AE	AESI for CTLS	Myopathy	15MAR2016:08:02	CYCLE 4 DAY 1	15MAR2016:08:55	15MAR2016:13:17	12MAR2016:08:55	22MAR2016:13:17
15	AZY300	AZY300-101-024			Drug A	TLS Evaluation	LTLS	Laboratory Tumor Lysis Syndrome	Y	14MAR2016:16:16	CYCLE 4 DAY 1				
16	AZY300	AZY300-101-024			Drug A	TLS Evaluation	CTLS	Clinical Tumor Lysis Syndrome	Y	14MAR2016:16:16	CYCLE 4 DAY 1				
17	AZY300	AZY300-101-344	ADAE	9	Drug B + C	TLS Events	AE	AESI for CTLS	Laryngospasm	22APR2016:09:22	CYCLE 5 DAY 1	22APR2016:08:14	22APR2016:13:06	19APR2016:08:14	29APR2016:13:06
18	AZY300	AZY300-101-344			Drug B + C	TLS Evaluation	LTLS	Laboratory Tumor Lysis Syndrome	N						
19	AZY300	AZY300-101-344			Drug B + C	TLS Evaluation	CTLS	Clinical Tumor Lysis Syndrome	N						
20	AZY300	AZY300-101-411	ADLB	221	Drug B + C	TLS Events	URATE	Hyperuricemia (> 8.0 mg/dL)	Y	03APR2016:12:12					
21	AZY300	AZY300-101-411	ADLB	224	Drug B + C	TLS Events	CACR	Hypocalcemia (< 7.0 mg/dL)	Y	03APR2016:12:12					
22	AZY300	AZY300-101-411	ADAE	5	Drug B + C	TLS Events	AE	AESI for CTLS	Bronchospasm	03APR2016:12:33					
23	AZY300	AZY300-101-411			Drug B + C	TLS Evaluation	LTLS	Laboratory Tumor Lysis Syndrome	N						
24	AZY300	AZY300-101-411			Drug B + C	TLS Evaluation	CTLS	Clinical Tumor Lysis Syndrome	N						
25	AZY300	AZY300-101-422			Drug A	TLS Events	AE	AESI for CTLS	Oliguria						
26	AZY300	AZY300-101-422			Drug A	TLS Evaluation	LTLS	Laboratory Tumor Lysis Syndrome	N						
27	AZY300	AZY300-101-422			Drug A	TLS Evaluation	CTLS	Clinical Tumor Lysis Syndrome	N						

In the preceding ADTLS, it is straightforward to determine the source of the "TLS Events" records by examining the SRCDOM and SRCSEQ records. If the conditions for LTLS or CTLS criteria are met, the ADTM and visit values are retained to make it easy to identify the records that led to the assessment. If no TLS event occurred, it is possible to quickly review all TLS-related records and verify whether the assessment was performed correctly.

2.12.5 Other Uses 12

Within current ADaM traceability variables, SRCDOM and SRCSEQ variables allow any one data point to be referenced in the derivation of a record. This example shows a case where derivations refer to 2 or more records, from 1 or 2 different datasets. To make this derivation understandable and traceable, relevant records were first pulled into one dataset and the final derivation is performed as a derived record. See Section 2.5, Traceability When Multiple Input Datasets Are Stacked to Create OCCDS, and Section 2.8, Using an Intermediate Dataset for BDS Traceability, for more examples on the using SRC* variables.

3 Appendices

Appendix A: Glossary and Abbreviations

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

ADaM	Analysis Dataset Model
ADaMIG	ADaM Implementation Guide
ADRG	ADaM Reviewer's Guide
ADSL	(ADaM) Subject-level Analysis Dataset
AVAL	Analysis value
BDS	(ADaM) Basic Data Structure
CDISC	Clinical Data Interchange Standards Consortium
Controlled terminology	A set of standard value lists that are used throughout the clinical research process, from data collection through analysis and submission
CTLS	Clinical tumor lysis syndrome
ECG	Electrocardiogram
eCRF	Electronic case report form/case record form
Define-XML	CDISC standard for transmitting metadata that describes any tabular dataset structure
FDA	(US) Food and Drug Administration
LTLS	Laboratory tumor lysis syndrome
LUT	Look-up table
OCCDS	(ADaM) Structure for Occurrence Data
SAP	Statistical analysis plan
SDTM	Study Data Tabulation Model
SDTMIG	SDTM Implementation Guide
TAUG	Therapeutic area user guide
TLS	Tumor lysis syndrome

Appendix B: References

1. European Medicines Agency. *ICH Topic E3. Structure and Content of Clinical Study Reports*. https://tinyurl.com/32zvfs45
2. Medical Dictionary for Regulatory Activities. http://www.meddramsso.com/

Appendix C: 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