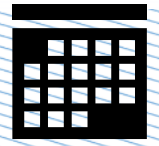
Define-XML v2.1 and ADaM



17-Jul-2020

Monika Kawohl

Coincidentally…

FDA Data Standards Catalog v6.4 (07-17-2020) - Supported and Required Standards

For full description of column headings, see Instr.& Column Descriptions tab

Use	Data Exchange Standard	Exchange Format	Standards Development Organization (SDO)	Supported Version	Supported Implementation Guide Version	FDA Center(s)	Date Support Begins (MM/DD/YYYY)	Date Support Ends (MM/DD/YYYY)	Date Requirement Begins (MM/DD/YYYY)	Date Requirement Ends (MM/DD/YYYY)	Statutory, Regulatory, or Guidance Authority	Information Sources
Study data definition	Define	XML	CDISC	1.0	N/A	CDER, CBER	Ongoing	03/15/2018	12/17/2016 [1] 12/17/2017 [2]	03/15/2018	Standardized Study Data	CDISC.org - Define-XML
Study data definition	Define	XML	CDISC	2.0	N/A	CDER, CBER	08/07/2013		12/17/2016 [1] 12/17/2017 [2]		Standardized Study Data	CDISC.org - Define-XML
Study data definition	Define	XML	CDISC	2.1	N/A	CDER, CBER	03/15/2021		03/15/2022 [1] 03/15/2023 [2]		Standardized Study Data	CDISC.org - Define-XML
Structured Product Labeling	SPL Image	JPG, XML	FDA and NLM	3.02	N/A	CDER, CBER, CVM	03/21/2012		3/21/2012 [7]		Structured Product Labeling (SPL) Implementation Guide with Validation Procedures (Section 3.2.12, 3.2.19, and Table 4)	SPLIMAGE File Specification v3.02

Notes:	
[1]	For NDAs, ANDAs, and certain BLAs. See section II.A of the Providing Regulatory Submissions In Electronic Format — Standardized Study Data guidance document
[2]	For certain INDs. See section II.A of the Providing Regulatory Submissions In Electronic Format - Standardized Study Data
[3]	Providing Regulatory Submissions in Electronic Format - Content of Labeling
[4]	The requirements of the electronic labeling rule can be found in § 314.50(l) for NDAs, § 314.94(d) for ANDAs, § 601.14(b) for BLAs, and § 314.81(b) for annual reports to marketing applications. The effective date of the rule was June 8, 2004.
[5]	For NDAs, ANDAs, and BLAs. See section B of the Providing Regulatory Submissions in Electronic Format - Certain Human Pharmaceutical Product Applications and Related Submission Using the eCTD Specifications guidance document
[6]	For Commercial INDs and Master Files. See section B of the Providing Regulatory Submissions in Electronic Format - Certain Human Pharmaceutical Product Applications and Related Submission Using the eCTD Specifications guidance document
[7]	The pill image is optional, but it is highly recommended to address numerous potential patient safety issues. Also see Federal Register Volume 76, Number 99, and Federal Register Volume 79, Number 182 as background.
[8]	" Ongoing"- Prior to the posting of a FDA Data Standards Catalog in 2011, FDA supported a number of data standards and there was no specific date when support began for some of these standards. As a result, the support status of "Ongoing" was entered for "Date Support Begins".

Agenda

Introduction

- ADaM example in the Define-XML v2.1 package

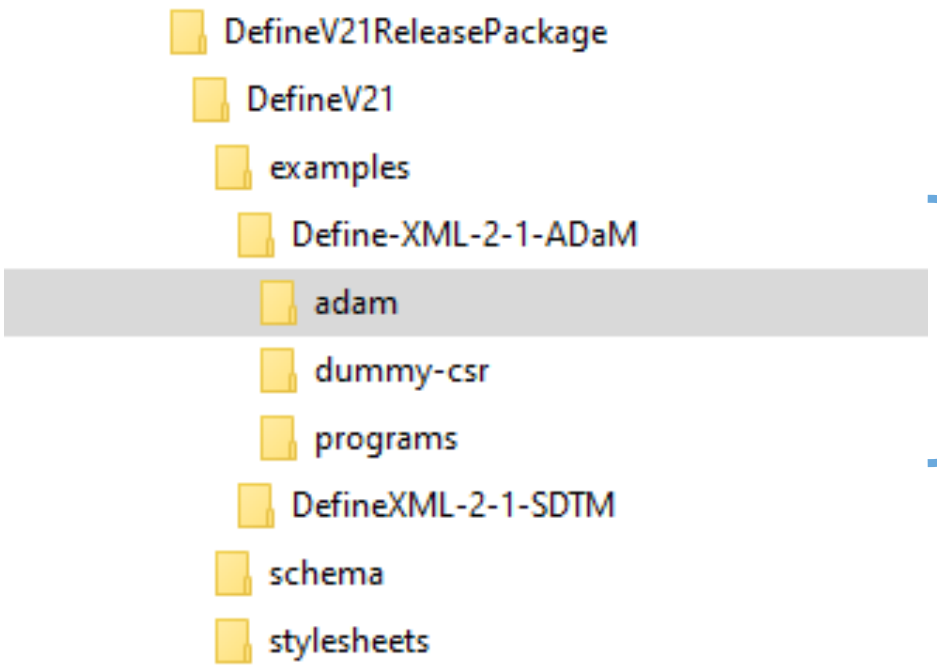
Relevant Updates

- Context
- Applicable Standards
- Subclass
- Define-XML CT
- Origin (Type and Source)
- More efficient references for external documents

Updates to check out for SDTM

- Origin
- HasNoData

ADaM Example in Define-XML v2.1



ADaM define.xml example from the ARM-for-Define-XML package updated according to
- ADaMIG v1.1
- Define-XML v2.1

Look and feel similar as enhanced Define-XML v2.0 stylesheet developed by a PHUSE WG

Context and Applicable Standards

CDISC-Sample

- Supplemental Documents
  Standards
- Analysis Results Metadata
- Datasets
- Controlled Terminology
- Methods

Expand all VLM

Collapse all VLM

Date/Time of Define-XML document generation: 2018-11-15T11:01:00
Define-XML version: 2.1.0
Define-XML Context: Submission
Stylesheet version: 2019-02-11

Study Name	CDISC-Sample
Study Description	CDISC-Sample Data Definition
Protocol Name	CDISC-Sample
Metadata Name	Study CDISC-Sample, Data Definitions
Metadata Description	Study CDISC-Sample, Data Definitions

Standards for Study CDISC-Sample

Standard	Type	Status	Documentation
ADaMIG 1.1	IG	Final	
CDISC/NCI ADaM 2017-09-29	CT	Final	ADaM specific CT is applicable for a few variables only.
CDISC/NCI SDTM 2018-06-29	CT	Final	SDTM CT is applicable for variables copied from SDTM to ADaM.

Dataset Metadata Referencing the Applicable Standard

Datasets

Dataset	Description	Class - SubClass	Structure	Purpose	Keys	Documentation	Location
ADSL [ADaMIG 1.1]	Subject-Level Analysis	SUBJECT LEVEL ANALYSIS DATASET	one record per subject	Analysis	STUDYID, USUBJID	Screen Failures are excluded since they are not needed for this study analysis. See referenced dataset creation program and ADRG adsl.sas Analysis Data Reviewer's Guide [6]	adsl.xpt
ADQSADAS [ADaMIG 1.1]	ADAS-Cog Analysis	BASIC DATA STRUCTURE	One record per subject per parameter per analysis visit per analysis date	Analysis	STUDYID, USUBJID, PARAMCD, AVISIT, ADT	See referenced dataset creation program and ADRG adqsadas.sas Analysis Data Reviewer's Guide [Section2.1]	adqsadas.xpt
ADAE [ADaMIG 1.1]	Adverse Events Analysis Dataset	OCCURRENCE DATA STRUCTURE • ADVERSE EVENT	one record per subject per adverse event	Analysis	STUDYID, USUBJID, AETERM, ASTDT, AESEQ	See SAS program adae.sas	adae.xpt

CTs Referencing the Applicable Standards

Unit - AWU [*C71620*] [CDISC/NCI SDTM 2018-06-29]

Permitted Value (Code)
DAYS [*C25301*]

```
<CodeList OID="CL.AWU"
          Name="Unit - AWU"
          DataType="text"
          def:StandardOID="STD.CT.02">
  <EnumeratedItem CodedValue="DAYS">
    <Alias Context="nci:ExtCodeID" Name="C25301"/>
  </EnumeratedItem>
  <Alias Context="nci:ExtCodeID" Name="C71620"/>
</CodeList>
<CodeList OID="CL.BMICAT"
          Name="Body Mass Index Category"
          DataType="text"
          def:IsNonStandard="Yes">
```

Body Mass Index Category [Non Standard]

Permitted Value (Code)	Display Value (Decode)	Rank
<25	Normal	1
25-<30	Overweight	2
>30	Obese	3

Date Imputation Flag [*C81223*] [CDISC/NCI ADaM 2017-09-29]

Permitted Value (Code)	Display Value (Decode)
D [*C81212*]	Day Imputed: Day is imputed
M [*C81211*]	Month Imputed: Month and day are imputed
Y [*C81210*]	Year Imputed: Entire date (year, month and day) is imputed

New Information: SubClass

Datasets

Dataset	Description	Class - SubClass	Structure	Purpose	Keys	Documentation	Location
ADSL [ADaMIG 1.1]	Subject-Level Analysis	SUBJECT LEVEL ANALYSIS DATASET	one record per subject	Analysis	STUDYID, USUBJID	Screen Failures are excluded since they are not needed for this study analysis. See referenced dataset creation program and ADRG adsl.sas Analysis Data Reviewer's Guide [6]	adsl.xpt
ADQSADAS [ADaMIG 1.1]	ADAS-Cog Analysis	BASIC DATA STRUCTURE	One record per subject per parameter per analysis visit per analysis date	Analysis	STUDYID, USUBJID, PARAMCD, AVISIT, ADT	See referenced dataset creation program and ADRG adqsadas.sas Analysis Data Reviewer's Guide [Section2.1]	adqsadas.xpt
ADAE [ADaMIG 1.1]	Adverse Events Analysis Dataset	OCCURRENCE DATA STRUCTURE • ADVERSE EVENT	one record per subject per adverse event	Analysis	STUDYID, USUBJID, AETERM, ASTDT, AESEQ	See SAS program adae.sas	adae.xpt

Unfortunately not a valid example anymore…

To account for this:
Define-XML v2.1
Attribute def:Class
becomes an element
in Define-XML v2.1

```
<def:Class Name="OCCURRENCE DATA STRUCTURE">
    <def:SubClass Name="ADVERSE EVENT"/>
</def:Class>
```

Extract from a CDISC Webinar

The ADaM Use Case for SubClass

Monika Kawohl

Background

- ADaM defines 3 Standard Data Structures

ADaM Standard Data Structures	Identified in Define-XML via a Class value of …
ADSL	SUBJECT LEVEL ANALYSIS DATASET
BDS	BASIC DATA STRUCTURE
OCCDS	OCCURRENCE DATA STRUCTURE

  - Other analysis datasets which adhere to the ADaM fundamental principles and naming conventions, but are not ADSL, BDS or OCCDS, are specified as Class="ADAM OTHER"
- For a given study analysis, there are typically multiple ADaM datasets of Class BDS or OCCDS
- ADaM does not define standard dataset names (exception: ADSL)

Use Cases

- There is currently no way to identify specific applications of these classes, e.g.,
  - BDS dataset for time-to-event (TTE) analysis
  - OCCDS dataset for adverse event analysis

Sample TTE dataset

STUDYID	USUBJID	PARAMCD	PARAM	AVAL	ADT	STARTDT	CNSR	EVNTDESC	…
ABC	ABC_101	TTDEATH	Time to Death (days)	15	15JAN2018	01JAN2018	0	Death	

Subclasses adhere to the definition of the "Parent" Class but have additional requirements and rules to be considered.

Sample adverse event analysis dataset

STUDYID	USUBJID	AESEQ	AEDECOD	AEBODSYS	ASTDT	AENDT	AESER	…
ABC	ABC-102	5	Headache	Nervous system	02JAN2018	03JAN2018	Y	

➢ **Subclasses facilitate validation**

Usage of SubClass

- Note: There is no requirement that a SubClass is defined for every ADaM dataset.
- SubClass is considered metadata → need a standard place for it in Define-XML
- Controlled Terminology CT(s) for valid values
  - Dependent on "parent" Class
  - Collaboration between ADaM team and Define-XML team
  - CTs may grow over time
    - ADaM documents in the pipeline may have use cases for SubClass, too.
- Work in progress
  - ADaM Define-XML v2.1 example including SubClass
  - Definition of SubClass in a suitable ADaM document

Extract from a CDISC Webinar

CTs Requested Regarding SubClass

New Codelist Name: SCLBDS (this CT should be non-extensible and should be added to the new Define-XML CT to be created)
New Codelist Description: Subclass of Standard Data Stucture/General Observation Class BASIC DATA STRUCTURE

CDISC Submission Value	CDISC Synonym(s)	CDISC Definition	Reason for this request and how it will be used
TIME-TO-EVENT	Time-to-Event, TTE	A specific application of the ADaM BASIC DATA STRUCTURE (BDS) for Time-to-Event ADaM datasets. The value of "TIME-TO-EVENT" can only be used in the context of time-to-event ADaM datasets of Class "BASIC DATA STRUCTURE".	The introduction of this subclass will facilitate identification of Time-to-Event datasets generated according to the document "CDISC ADaM Basic Data Structure for Time-to-Event Analysis" via metadata (define.xml). Identification of this subclass of a BASIC DATA STRUCTURE will enable specification of validations rules specific to this subclass.

New Codelist Name: SCLOCCDS (this CT should be non-extensible and should be added to the new Define-XML CT to be created)
New Codelist Description: Subclass of Standard Data Stucture/General Observation Class OCCURRENCE DATA STRUCTURE

CDISC Submission Value	CDISC Synonym(s)	CDISC Definition	Reason for this request and how it will be used
ADVERSE EVENTS	Adverse Events	A specific application of the ADaM OCCURENCE DATA STRUCTURE (OCCDS) for ADaM adverse events datasets. The value of "ADVERSE EVENTS" can only be used in the context of ADaM datasets of Class "OCCURRENCE DATA STRUCTURE".	The introduction of this subclass will facilitate identification of ADaM adverse events datasets generated according to the document "CDISC ADaM Structure for Occurrence Data (OCCDS)" via metadata (define.xml). Identification of this subclass of an OCCURRENCE DATA STRUCTURE will enable specification of validations rules specific to this subclass.

Side Tour: Define-XML CT

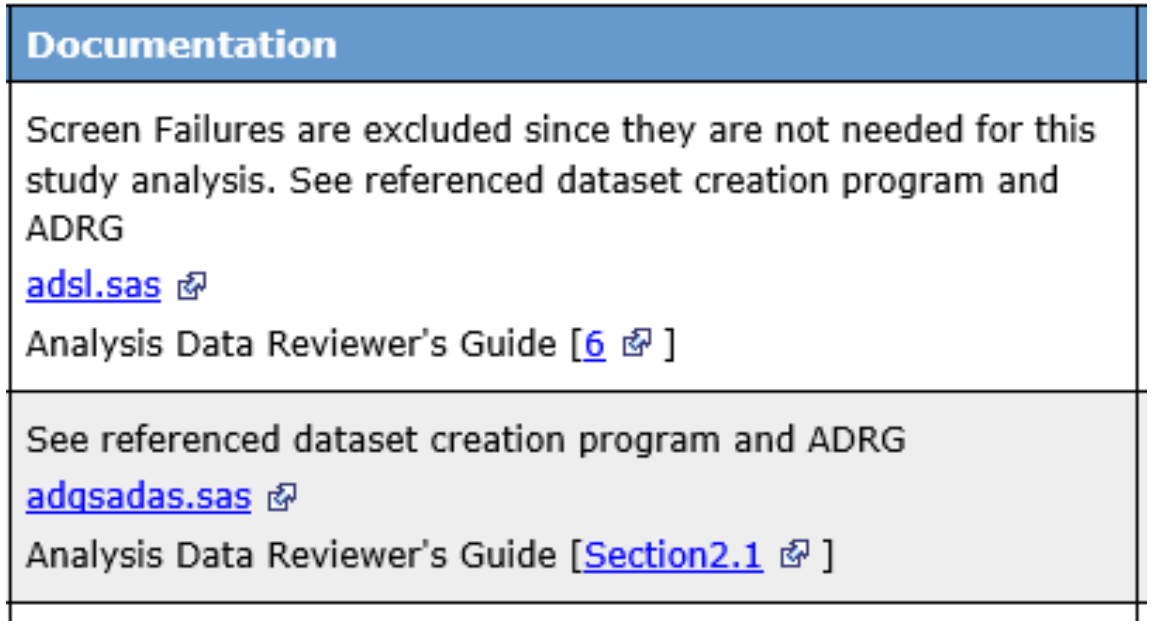
Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition	NCI Preferred Term
C165635		No	ADaM Basic Data Structure Subclass	BDSSC	ADaM Basic Data Structure Subclass	Terminology relevant to the subclasses of the ADaM basic data structure.	CDISC Define-XML ADaM Basic Data Structure Subclass Terminology
C172452	C165635		ADaM Basic Data Structure Subclass	NON-COMPARTMENTAL ANALYSIS	NCA	A dataset containing data that is used for non-compartmental analyses.	Non-Compartmental Analysis Dataset
C165637	C165635		ADaM Basic Data Structure Subclass	TIME-TO-EVENT	TTE	A dataset containing data that is used for Time-to-Event analyses.	Time-to-Event Dataset
C165636		No	ADaM Integrated Basic Data Structure Subclass	BDSISC	ADaM Integrated Basic Data Structure Subclass	Purpose of a specific analysis result described in ADaM analysis results metadata.	CDISC Define-XML ADaM Integrated Basic Data Structure Subclass Terminology
C165637	C165636		ADaM Integrated Basic Data Structure Subclass	TIME-TO-EVENT	TTE	A dataset containing data that is used for Time-to-Event analyses.	Time-to-Event Dataset
C117745		Yes	Analysis Purpose	ANLPURP	Analysis Purpose	Purpose of a specific analysis result described in Define-XML analysis results metadata.	CDISC Define-XML Analysis Purpose Terminology
C98724	C117745		Analysis Purpose	EXPLORATORY OUTCOME MEASURE	Exploratory Outcome Measure	The outcome measure(s) that is part of a pre-specified analysis plan used to evaluate the exploratory endpoint(s) associated with exploratory study objective(s) and/or any other measures, excluding post-hoc measures, that are a focus of the study. (After clinicaltrials.gov)	Exploratory Outcome Measure
C98772	C117745		Analysis Purpose	PRIMARY OUTCOME MEASURE	Primary Outcome Measure	The outcome measure(s) of greatest importance specified in the protocol, usually the one(s) used in the power calculation, to evaluate the primary endpoint(s) associated with the primary study objective(s). (After Clinicaltrials.gov)	Primary Outcome Measure
C98781	C117745		Analysis Purpose	SECONDARY OUTCOME MEASURE	Secondary Outcome Measure	The outcome measure(s) that is part of a pre-specified analysis plan used to evaluate the secondary endpoint(s) associated with secondary study objective(s) and/or used to evaluate any measure(s) ancillary to the primary or secondary endpoint(s). (After Clinicaltrials.gov).	Secondary Outcome Measure
C117744		Yes	Analysis Reason	ANLREAS	Analysis Reason	Reason for reporting a specific analysis result described in Define-XML analysis results metadata.	CDISC Define-XML Analysis Reason Terminology
C117750	C117744		Analysis Reason	DATA DRIVEN		The analysis was triggered by findings in the data.	Data Driven Analysis
C117751	C117744		Analysis Reason	REQUESTED BY REGULATORY AGENCY		The analysis has been requested by a regulatory agency.	Analysis Requested by Regulatory Agency
C117752	C117744		Analysis Reason	SPECIFIED IN PROTOCOL		The analysis is specified in a protocol.	Analysis Specified in Protocol
C117753	C117744		Analysis Reason	SPECIFIED IN SAP		The analysis is specified in a statistical analysis plan.	Analysis Specified in Statistical Analysis Plan
C172331		Yes	CDISC Controlled Terminology Standard Type	CTSTDTYP	CDISC Controlled Terminology Standard Type	Terminology relevant to the classification of the CDISC controlled terminology standard described in the Define-XML document.	CDISC Define-XML CDISC Controlled Terminology Standard Type Terminology
C81222	C172331		CDISC Controlled Terminology Standard Type	ADaM		The terminology subset that includes terms pertaining to the Clinical Data Interchange Standards Consortium (CDISC) Analysis Data Model (ADaM).	CDISC ADaM Terminology
C77527	C172331		CDISC Controlled Terminology Standard Type	CDASH		The terminology subset that includes terms relevant to the Clinical Data Interchange Standards Consortium (CDISC) Clinical Data Acquisition Standards Harmonization (CDASH) group.	CDISC CDASH Terminology
C165634	C172331		CDISC Controlled Terminology Standard Type	DEFINE-XML		The terminology subset that includes terms relevant to the Clinical Data Interchange Standards Consortium (CDISC) Define-XML standard.	CDISC Define-XML Terminology
C66830	C172331		CDISC Controlled Terminology Standard Type	SDTM		The terminology subset that includes terms pertaining to the Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM).	CDISC SDTM Terminology
C77526	C172331		CDISC Controlled Terminology Standard Type	SEND		The terminology subset that includes terms relevant to the Clinical Data Interchange Standards Consortium (CDISC) Standard for Exchange of Non-	CDISC SEND Terminology

ReadMe | Def-XML Terminology 2020-06-26

https://www.cancer.gov/research/resources/terminology/cdisc#define-xml-terminology

More efficient external document references

```
<def:CommentDef OID="COM.ADSL">
  <Description>
    <TranslatedText xml:lang="en">Screen Failures are excluded since they are not needed for this study analysis. See referenced dataset creation program and ADRG</TranslatedText>
  </Description>
  <def:DocumentRef leafID="LF.ADSL.PGM"/>
  <def:DocumentRef leafID="LF.ADRG">
    <def:PDFPageRef PageRefs="6" Type="PhysicalRef"/>
  </def:DocumentRef>
</def:CommentDef>
<def:CommentDef OID="COM.ADQSADAS">
  <Description>
    <TranslatedText xml:lang="en">See referenced dataset creation program and ADRG</TranslatedText>
  </Description>
  <def:DocumentRef leafID="LF.ADQSADAS.PGM"/>
  <def:DocumentRef leafID="LF.ADRG">
    <def:PDFPageRef PageRefs="Section2.1" Type="NamedDestination"/>
  </def:DocumentRef>
</def:CommentDef>
```



Additionally available now: Attribute Title for element def:PDFPageRef

Compare respective documentation needs in Define-XML v2.0

```
<def:CommentDef OID="COM.ADSL">
  <Description>
   <TranslatedText>Screen Failures are excluded since they are not needed for this study analysis. See Analysis Data Reviewer's Guide, page 6.</TranslatedText>
  </Description>
  <def:DocumentRef leafID="LF.ADRG">
   <def:PDFPageRef PageRefs="6" Type="PhysicalRef"/>
  </def:DocumentRef>
 </def:CommentDef>
<def:CommentDef OID="COM.ADQSADAS">
  <Description>
   <TranslatedText>See referenced dataset creation program and Analysis Data Reviewer's Guide, Section 2.1</TranslatedText>
  </Description>
  <def:DocumentRef leafID="LF.ADQSADAS.PGM"/>
  <def:DocumentRef leafID="LF.ADRG">
   <def:PDFPageRef PageRefs="Section2.1" Type="NamedDestination"/>
  </def:DocumentRef>
 </def:CommentDef>
```

Documentation
Screen Failures are excluded since they are not needed for this study analysis. See Analysis Data Reviewer's Guide, page 6. Analysis Data Reviewer's Guide
See referenced dataset creation program and Analysis Data Reviewer's Guide, Section 2.1 adqsadas.sas Analysis Data Reviewer's Guide

Origin (Type) => Origin (Type + Source)

ADQSADAS (ADAS-Cog Analysis) - [ADaMIG 1.1] Location: adqsadas.xpt

Variable	Where Condition	Label / Description	Type	Length or Display Format	Controlled Terms or ISO Format	Origin / Source / Method / Comment
STUDYID		Study Identifier	text	12		Predecessor: ADSL.STUDYID
ADY		Analysis Relative Day	integer	3		Derived (Source: Sponsor) ADY = ADT - TRTSDT + 1, if ADT>=TRTSDT. ADY = ADT - TRTSDT, if ADT<TRTSDT.
ADT		Analysis Date	integer	date9.		Derived (Source: Sponsor) SAS date from QS.QSDTC
PARAM		Parameter	text	100	ADAS-Cog Parameter [15 Terms]	Assigned (Source: Sponsor)
PARAMCD		Parameter Code	text	8	ADAS-Cog Parameter Code [15 Terms]	Assigned (Source: Sponsor)

More complex for SDTM

SDTM Origin – Type and Source Combinations

Type	Source				Notes
	Subject	Investigator	Vendor	Sponsor	
Collected	ePro	CRF	Lab data, ECG	X	This term should be used for clinical data that were actually observed or recorded by a person or received from an instrument; it should not be used for data that have been interpreted, calculated, or derived from other information.
Derived	X	X	Lab data, ECG	SDTM	Derivation examples include calculations performed during data collection (e.g., --DY). Other derivation examples: calculations within ePRO (e.g., questionnaire section scores) and calculations within EDC (e.g., BMI, BSA).
Assigned	X	X	Adjudicator	SDTM	Examples of this include third-party attributions by an adjudicator, coded terms that are supplied as part of a coding process, and values that are set independently of any subject-related data values in order to complete SDTM fields such as DOMAIN and --TESTCD
Protocol	X	X	X	SDTM	An example would be VSPOS (Vital Signs Position), which could be specified in the protocol and be provided by other means (e.g. CRF, eDT).
Predecessor	X	X	X	X	Use when a value is an exact copy of another value in an SDTM dataset.

Attribute def:HasNoData (SDTM)

- Mark with def:HasNoData="Yes" and explain via a Comment if data are missing
  - Dataset

XX [Non Standard] **[No Data]**	X Findings	FINDINGS	One record per finding per visit per subject	Tabulation	STUDYID, USUBJID, XXTESTCD, XXDTC, VISITNUM	Special domain contingent on rare conditions observed.	

  - Variable

XSORRESU **[No Data]**	Original Units	text	20	Units for S Findings Results • "g/dL" = "g/dL" • "mg/dL" = "mg/dL"	Collected (Source: Vendor) Planned Numeric tests were not performed.

  - Value-level, e.g.,
    QNAM value for "Other specify" is not available in the respective SUPP dataset because option other has never been selected

Thank You!

Questions?

Contact Details

Monika Kawohl

Principal Statistical Programmer

E-mail monika.kawohl@mainanalytics.de
Internet www.mainanalytics.de