

Define-XML v2.0 or What Happened to Value Level Metadata?

Monika Kawohl

Principal Statistical Programmer

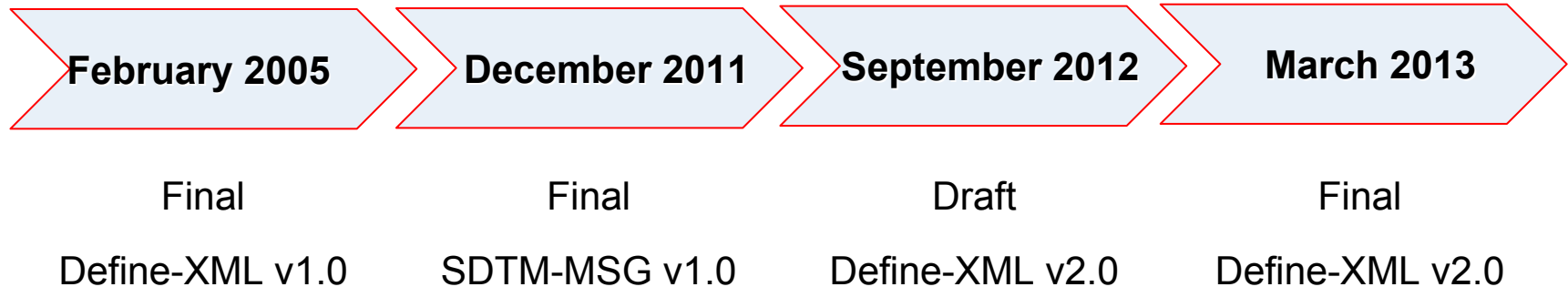
Accovion



Overview

- ✓ Define-XML History
- ✓ Define-XML v2.0 Package
- ✓ Illustration of Major Changes
 - With Focus on Value Level Metadata
- ✓ Getting Started with Define-XML v2.0
- ✓ Future Plans
- ✓ Closing Remarks

Define-XML History



✓ Define-XML v1.0 → Draft Define-XML v2.0

- “Major Overhaul”, Higher Complexity
- Change to Latest ODM Version
- Applicable for SDTM, **ADaM**, SEND

✓ Draft Define-XML v2.0 → Final Define-XML v2.0

- Schema: Minor Changes Only
- Stylesheet: Quite Some Changes

Define-XML v2.0 Package

✓ Introduction – Package Overview

✓ Specifications

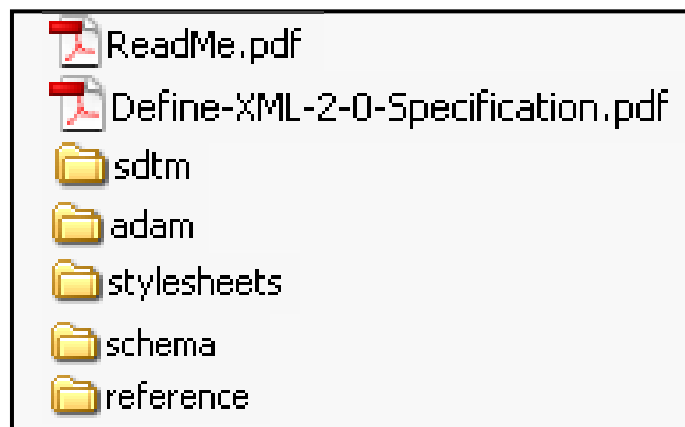
- General Specifications with Illustrations for the Broad Public
- Technical Specifications for Implementers/Tool Developers

✓ Samples for SDTM and ADaM

✓ Updated Stylesheet applicable for SDTM+ADaM

- Stable Back Button
- No Supportive Files Needed (e.g., .css, .gif)
- Can be Printed ☺

✓ Schema and References (i.e., Technical Details)



ADaM-IG 1.0
Analysis Data Reviewer's Guide
• Analysis Datasets
• Parameter Value Level Meta
• Controlled Terminology
• Analysis Derivations
• Comments

Date of document generation: 2013-03-03T14:47:00
Stylesheet version: 2013-03-04

Analysis Datasets for Study CDISC-Sample (ADaM-IG 1.0)

Dataset	Description	Class	Structure	Purpose	Keys	Location	Documentation
ADSL	Subject-Level Analysis	SUBJECT LEVEL ANALYSIS DATASET	one record per subject	Analysis	USUBID	adsl.xpt	Screen Failures are excluded since they are not needed for this study analysis
ADQSDAD	ADaM-Cos Analysis	BASIC DATA STRUCTURE	One record per subject per parameter per analysis visit per analysis date	Analysis	USUBID, PARAMCD, AVISIT, ADT	adqsdad.xpt	See referenced dataset creation program and Analysis Data Reviewer's Guide, Section 2.3. adqsdad.xpt Analysis Data Reviewer's Guide

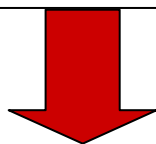
Go to the [top](#) of the define.xml

Subject-Level Analysis (ADSL) [Location: [adsl.xpt](#)]

Variable	Label	Type	Length / Number of Values	Controlled Terminology	Source/Derivation/Comment
STUDID	Study Identifier	text	12		Predecessor: DM.STUDID
USUBID	Unique Subject Identifier	text	11		Predecessor: DM.USUBID
SUBID	Subject Identifier for the Study	text	4		Predecessor: DM.SUBID
SITEID	Study Site Identifier	text	3		Predecessor: DM.SITEID
SITEGR1	Pooled Site Group 1	text	3		Derived: refer to SAP, Section 7.1 - if not pooled then SITEGR1=SITEID. If pooled, SITEGR1 will be 900
ARM	Description of Planned Arm	text	20 ("Placebo", "Nonrandomized Low Dose", "Nonrandomized High Dose")		Predecessor: DM.ARM

Define-XML Technology Sample

```
...
<!-- ***** -->
<!-- ItemGroupDef Definitions Section (Datasets and first set of variable properties) -->
<!-- ***** -->
<ItemGroupDef OID="IG.ADSL" Name="ADSL" SASDatasetName="ADSL" Repeating="No" IsReferenceData="No"
  Purpose="Analysis" def:Structure="one record per subject"
    def:Class="SUBJECT LEVEL ANALYSIS DATASET" def:CommentOID="COM.ADSL" def:ArchiveLocationID="LF.ADSL">
      <Description>
        <TranslatedText xml:lang="en">Subject-Level Analysis</TranslatedText>
      </Description>
      <ItemRef ItemOID="IT.ADSL.STUDYID" OrderNumber="1" Mandatory="No"/>
      <ItemRef ItemOID="IT.ADSL.USUBJID" OrderNumber="2" Mandatory="No" KeySequence="1"/>
      ...
      <def:leaf ID="LF.ADSL" xlink:href="adsl.xpt">
        <def:title>adsl.xpt</def:title>
      </def:leaf>
    </ItemGroupDef>
  ...
```



stylesheet **define-2-0-0.xsl** which controls the define rendition in a browser

Dataset	Description	Class	Structure	Purpose	Keys	Location	Documentation
ADSL	<u>Subject-Level Analysis</u>	SUBJECT LEVEL ANALYSIS DATASET	one record per subject	Analysis	USUBJID	<u>adsl.xpt</u>	Screen Failures are excluded since they are not needed for this study analysis

Overview

- ✓ Define-XML History
- ✓ Define-XML v2.0 Package
- ✓ **Illustration of Major Changes**
 - **With Focus on Value Level Metadata**
- ✓ Getting Started with Define-XML v2.0
- ✓ What Does the CDISC Define-XML Team Do Now?
- ✓ Closing Remarks

SDTM Value Level Metadata in Define-XML v1.0

Vital Signs Dataset (VS)						vs.xpt
Variable	Label	Type	Controlled Terms or Format	Origin	Role	Comment
VSTESTCD	Vital Signs Test Short Name	text		Assigned	TOPIC	



Value Level Metadata (ValueList.VS.VSTESTCD)						
Source Variable	Value	Label	Type	Controlled Terms or Format	Origin	Comment
VSTESTCD	DIABP	Diastolic Blood Pressure	integer		CRF Page 11	
VSTESTCD	FRMSIZE	Body Frame Size	text	SIZE	CRF Page 11	
VSTESTCD	HEIGHT	Height	float		CRF Page 11	
VSTESTCD	PULSE	Pulse Rate	integer		CRF Page 11	
VSTESTCD	SYSBP	Systolic Blood Pressure	integer		CRF Page 11	
VSTESTCD	WEIGHT	Weight	float		CRF Page 11	

- ✓ "Describing VS after a PROC TRANSPOSE"
- Value of --TESTCD → Variable
 - Value of --TEST → Label
 - Value of --ORRES → Variable Value

Parameter Value Level Metadata in ADaM v2.1

Dataset Name	Parameter Identifier	Variable Name	Variable Label	Variable Type	Display Format	Codelist / Controlled Terms	Source / Derivation
ADQSADAS	*ALL*	USUBJID	Unique Subject Identifier	text	\$11		ADSL.USUBJID
ADQSADAS	*ALL*	AVISIT	Analysis Visit	text	\$19	Baseline, Week 8, Week 16, Week 24	Refer to Section 8.2 of the SAP for a detailed description of the windowing algorithm used to determine the analysis visit based on ADQSADAS.ADY
ADQSADAS	PARAMCD	PARAMCD	Parameter Code	text	\$8	ACITM01, ACITM02, ACITM03, ACITM04, ACITM05, ACITM06, ACITM07, ACITM08, ACITM09, ACITM10, ACITM11, ACITM12, ACITM13, ACITM14, ACTOT	Corresponds to PARAM
ADQSADAS	*DEFAULT*	AVAL	Analysis Value	float	3.0		When ADQSADAS.PARAMCD indicates an item score (rather than a total score), AVAL is the corresponding value (for subject and visit) of QS.QSSTRESN when QS.QSTESTCD = ADQSADAS.PARAMCD
ADQSADAS	ACTOT	AVAL	Analysis Value	float	3.0		Sum of ADAS scores for items 1, 2, 4, 5, 6, 7, 8, 11, 12, 13, and 14, see SAP section 14.2 for details on adjusting for missing values
ADQSADAS	*DEFAULT*	DTYPE	Derivation Type	text	\$4		Not applicable, therefore blank
ADQSADAS	ACTOT	DTYPE	Derivation Type	text	\$4	LOCF	DTYPE = 'LOCF' when the value of ADQSADAS.AVAL (and thus the entire record) has been imputed using the LOCF algorithm, blank otherwise.

Value Level Metadata Requirements for V2.0

✓ ADaM

- Support for ADaM "Parameter Identifier"
 - Allow to Describe **Any** Variable's Metadata Depending on PARAMCD not just AVAL
 - Allow to Specify a Variable's Metadata for a Set of PARAMCDs (*DEFAULT*)

✓ SDTM

- Allow to Identify a Specific Test or Measurement at One Glance
 - --TESTCDs not Necessarily Uniquely Describe a Specific Measurement (e.g., LBTESTCD="GLUC")
 - Alternative Solution for **"Nesting"** (e.g., Linking from --CAT to --SPEC to --METHOD to --TESTCD)

✓ Develop Something Generic that Works for Both SDTM and ADaM

SDTM Value Level Metadata in Define-XML v2.0

Vital Signs (VS) [Location: [vs.xpt](#)]

Variable	Label	Key	Type	Length	Controlled Terms or Format	Origin	Derivation/Comment
VSORRES	Result or Finding in Original Units		text	30		CRF Page 11	



Value Level Metadata - VS [[VSORRES](#)]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
VSORRES	VSTESTCD EQ DIABP (Diastolic Blood Pressure)	integer	2		CRF Page 11	
VSORRES	VSTESTCD EQ FRMSIZE (Body Frame Size)	text	6	["LARGE", "MEDIUM", "SMALL"] < Size >	CRF Page 11	
VSORRES	VSTESTCD EQ HEIGHT (Height)	float	5.1		CRF Page 11	
VSORRES	VSTESTCD EQ PULSE (Pulse Rate)	integer	2		CRF Page 11	
VSORRES	VSTESTCD EQ SYSBP (Systolic Blood Pressure)	integer	3		CRF Page 11	
VSORRES	VSTESTCD EQ WEIGHT (Weight)	float	5.1		CRF Page 11	

✓ **"Where"** - Describing a Variable's Metadata after Subsetting the Dataset as Specified

SDTM Value Level Metadata in Define-XML v2.0

Previously a Use Case For "Nesting"

Value Level Metadata - LB [LBORRES]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
LBORRES	LBTESTCD EQ BILI (Bilirubin) AND LBCAT EQ CHEMISTRY AND LBSPEC EQ BLOOD	float	3		eDT	
LBORRES	LBTESTCD EQ BUN (Blood Urea Nitrogen) AND LBCAT EQ CHEMISTRY AND LBSPEC EQ BLOOD	float	4		eDT	
LBORRES	LBTESTCD EQ GLUC (Glucose) AND LBCAT EQ CHEMISTRY AND LBSPEC EQ BLOOD	float	3		eDT	
LBORRES	LBTESTCD EQ GLUC (Glucose) AND LBCAT EQ URINALYSIS AND LBSPEC EQ URINE AND LBMETHOD EQ DIPSTICK	text	8		eDT	
LBORRES	LBTESTCD EQ GLUC (Glucose) AND LBCAT EQ URINALYSIS AND LBSPEC EQ URINE AND LBMETHOD EQ QUANT	text	8		eDT	

- ✓ Understand that there are 3 Different Types of Glucose Measurements and How to Select the Specific Test or Records of Interest

ADaM Parameter Value Level Metadata in Define-XM V2.0

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

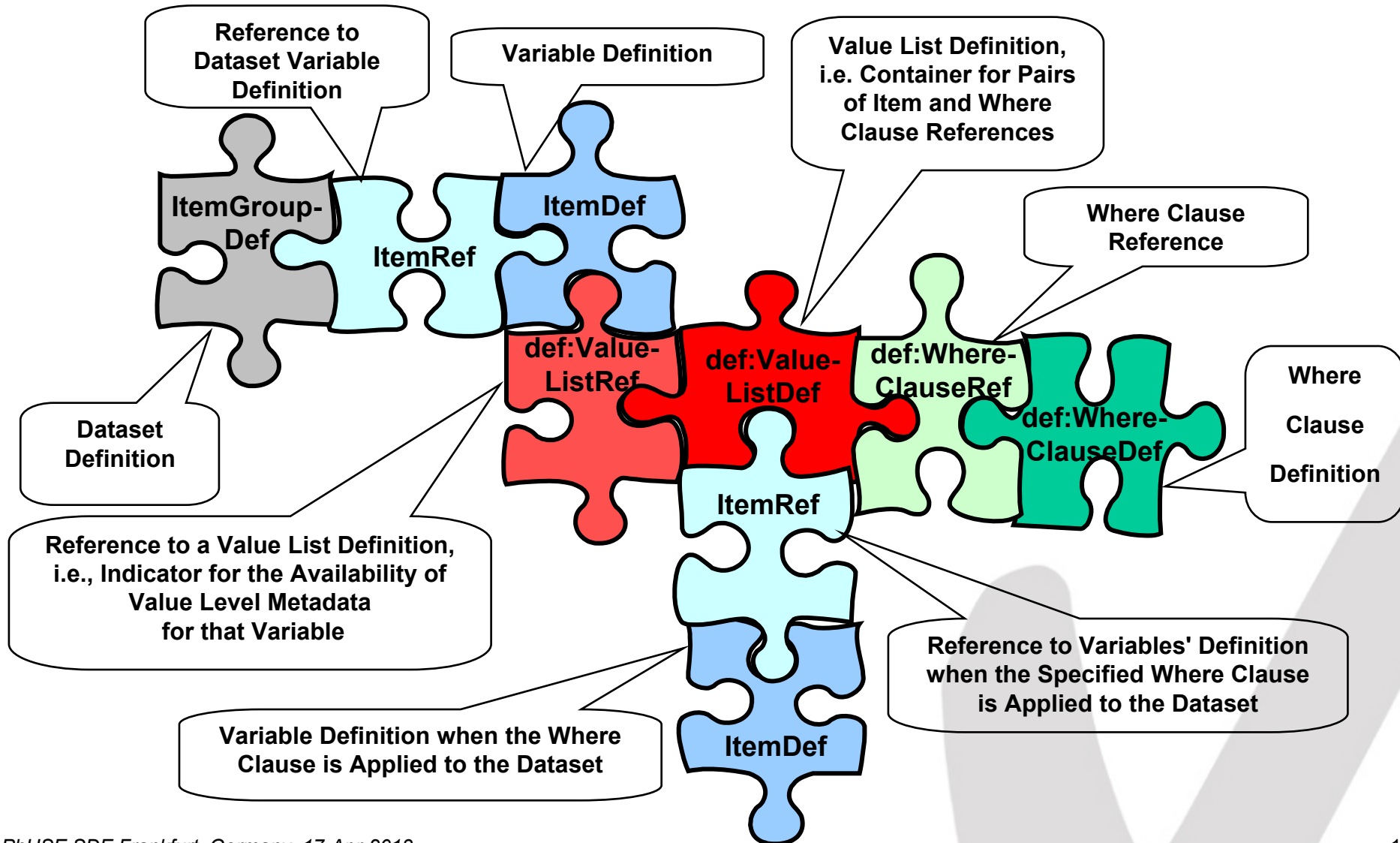
Variable	Label	Type	Length / Display Format	Controlled Terms or Format	Source/Derivation/Comment
PARAM	Parameter	text	100	PARAM_ADQSADAS	Assigned:
PARAMCD	Parameter Code	text	8	PARAMCD_ADQSADAS	Assigned:
AVAL	Analysis Value	integer	8		Derivations are described per parameter in the parameter value level metadata



Parameter Value List - ADQSADAS [AVAL]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
AVAL	PARAMCD NOTIN ("ACTOT" (Adas-Cog(11) Subscore))	integer	8		Derived	QS.QSSTRESN where QSTESTCD=PARAMCD
		Parameter Identifier: *DEFAULT*				
AVAL	PARAMCD EQ ACTOT (Adas-Cog(11) Subscore)	integer	8		Derived	Sum of ADAS scores for items 1, 2, 4, 5, 6, 7, 8, 11, 12, 13, and 14, see Analysis Data Reviewers Guide (Page 3) for details on adjusting for missing values. Analysis Data Reviewer's Guide (analysis-data-reviewers-guide.pdf)
		Parameter Identifier: ACTOT				

"Value Level Metadata Behind the Scenes"



Value Level Metadata – A Different View: "Slices"

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

Where PARAMCD NOTIN("ACTOT" (Adas-Cog(11) Subscore))

Variable	Label	Type	Length/ Display Format	Controlled Terms or Format	Source/Derivation/Comment
STUDYID	Study Identifier	text	12		Predecessor: ADSL.STUDYID
...					
AVAL	Analysis Value	integer	8		Derived: QS.QSSTRESN where QSTESTCD=PARAMCD
DTYPE	Derivation Type	text	10		Assigned: Value: null

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

Where PARAMCD EQ "ACTOT" (Adas-Cog(11) Subscore)

Variable	Label	Type	Length/ Display Format	Controlled Terms or Format	Source/Derivation/Comment
STUDYID	Study Identifier	text	12		Predecessor: ADSL.STUDYID
...					
AVAL	Analysis Value	integer	8		Derived: Sum of ADAS scores for items 1, 2, 4, 5, 6, 7, 8, 11, 12, 13, and 14, see Analysis Data Reviewers Guide (Page 3) for details on adjusting for missing values. Analysis Data Reviewer's Guide (analysis-data-reviewers-guide.pdf)
DTYPE	Derivation Type	text	10	["LOCF" = "Last Observation Carried Forward"] < Derivation Type >	Assigned: Value: LOCF denotes that the LOCF imputation method was used to impute the value for the given parameter and analysis visit.

Controlled Terminology: Enumerated Items vs. Code Lists

Vital Signs Test Name [CL.VSTEST, C67153]

Permitted Value (Code)
Body Frame Size [C49680]
Diastolic Blood Pressure [C25299]
Height [C25347]
Pulse Rate [C49676]
Systolic Blood Pressure [C25298]
Weight [C25208]

Vital Signs Test Code [CL.VSTESTCD, C66741]

Permitted Value (Code)	Display Value (Decode)
DIABP [C25299]	Diastolic Blood Pressure
FRMSIZE [C49680]	Body Frame Size
HEIGHT [C25347]	Height
PULSE [C49676]	Pulse Rate
SYSBP [C25298]	Systolic Blood Pressure
WEIGHT [C25208]	Weight

For CDISC Defined Controlled Terminologies (CT) Specify:

- **CT Long Name**
- NCI C-Codes for **CT** and **CT Values**

Links to External Documents – SDTM Sample

Variable	Label	Key	Type	Length	Controlled Terms or Format	Origin	Derivation/Comment
EGDRVFL	Derived Flag		text	1	["N" = "No", "U" = "Unknown", "Y" = "Yes"] < No Yes Response Subset >	Derived	EGDRVFL = "Y" for derived EGTESTCDs QTCB and QTCF. Null otherwise. Complex Algorithms (complexalgorithms.pdf)

<ItemRef ItemOID="IT.EG.EGDRVFL" OrderNumber="15" Mandatory="No" MethodOID="MT.EGDRVFL" />

<ItemDef OID="IT.EG.EGDRVFL" Name="EGDRVFL" DataType="text" Length="1" SASFieldName="EGDRVFL">

<Description>

<TranslatedText xml:lang="en">Derived Flag</TranslatedText>

</Description>

<CodeListRef CodeListOID="CL.NY" />

<def:Origin Type="Derived" />

</ItemDef>

<MethodDef Type="Computation" OID="MT.EGDRVFL" Name="Algorithm to derive EGDRVFL">

<Description>

<TranslatedText xml:lang="en">Equal to Y for derived EGTESTCDs QTCB and QTCF</TranslatedText>

</Description>

<def:DocumentRef leafID="LF.ComplexAlgorithms">

<def:PDFPageRef PageRefs="EG" Type="NamedDestination" />

</def:DocumentRef>

</MethodDef>

<def:leaf ID="LF.ComplexAlgorithms" xlink:href="complexalgorithms.pdf">

<def:title>Complex Algorithms</def:title>

</def:leaf>

Links to External Documents – ADaM Sample

Analysis Datasets for Study CDISC-Sample (ADaM-IG 1.0)

Dataset	Description	Class	Structure	Purpose	Keys	Location	Documentation
ADSL	Subject-Level Analysis	SUBJECT LEVEL ANALYSIS DATASET	one record per subject	Analysis	USUBJID	adsl.xpt	Screen Failures are excluded since they are not needed for this study analysis
ADQSADAS	ADAS-Cog Analysis	BASIC DATA STRUCTURE	One record per subject per parameter per analysis visit per analysis date	Analysis	USUBJID, PARAMCD, AVISIT, ADT	adqsadas.xpt	See referenced dataset creation program and Analysis Data Reviewer's Guide, Section 2.1 adqsadas.sas Analysis Data Reviewer's Guide

In Print Preview:

Analysis Datasets for Study CDISC-Sample (ADaM-IG 1.0)

Dataset	Description	Class	Structure	Purpose	Keys	Location	Documentation
ADSL	Subject-Level Analysis	SUBJECT LEVEL ANALYSIS DATASET	one record per subject	Analysis	USUBJID	adsl.xpt	Screen Failures are excluded since they are not needed for this study analysis
ADQSADAS	ADAS-Cog Analysis	BASIC DATA STRUCTURE	One record per subject per parameter per analysis visit per analysis date	Analysis	USUBJID, PARAMCD, AVISIT, ADT	adqsadas.xpt <adqsadas.xpt>	See referenced dataset creation program and Analysis Data Reviewer's Guide, Section 2.1 adqsadas.sas <../programs/adqsadas-sas.txt> Analysis Data Reviewer's Guide <analysis-data-reviewers-guide.pdf>

Getting Started with Define-XML v2.0

A valid *define.xml* file must

- Properly reference versions of the CDISC standards.
- Be well formed and conform to the XML schemas.
- Meet all of the requirements documented in this specification.

✓ Familiarize with the New Package

- To Understand the Requirements (XML Schema, Business Rules)
 - e.g., Valid Data Type Values (text, integer, float, ...)
 - e.g., for Each Variable with Origin=Derived Provide a MethodDef

✓ Design Metadata Specifications

- To Capture all Required Details
- Could be an Excel Workbook

✓ Create a Tool

- To Create the define.xml File
- Could be a SAS Program

What Do the Define-XML/ADaM Metadata Teams Do Now?

- ✓ Define-XML V2.0 Implementation Guide
 - e.g., Stylesheet to Support the "Slices" View
- ✓ Define-XML V2.0 Validation Rules
- ✓ ADaM Results Metadata Extension for Define-XML v2.0
 - Based on Updated CDISC SDTM/ADaM Pilot Submission Package

Display	Figure 14-1 Time to Dermatologic Event by Treatment Group
AnalysisResult	treatment group comparison for time to event variables
Analysis Parameter(s)	TTDE=Time to First Dermatological Event
Analysis Variable(s)	AVAL CNSR
Reason	pre-specified in SAP
Data References (incl. Selection Criteria)	ADTTE [SAFFL="Y"]
Documentation	SAP Section 11.2 Kaplan-Meier estimates and log-rank analysis of time to first dermatological adverse event, safety population
Programming Statements	<pre>proc lifetest data = adtte (where=(saffl="Y")) plots=s; id usubjid; strata trtan; time aval*cnsr(1); test trtan; run;</pre>

Closing Remarks

- ✓ Define-XML v2.0 not yet Formally Accepted by the FDA
- ✓ More Complex than Define-XML v1.0

BUT

- ✓ It Includes Elements and Functionality Previously Missed, Especially for ADaM
- ✓ Clear Definitions, Less Room for Interpretation
- ✓ Harmonized Approach for SDTM+ADaM as far as Possible
- ✓ SDTM+ADaM Examples and a Modernized Stylesheet

Recommendation: Let's Start Making Use of It!

Thank You!

Questions?

Contact Information:

Monika Kawohl
Principal Statistical Programmer
Accovion GmbH
Softwarecenter 3
35037 Marburg
Germany

Phone: +49-6421-94849-20
Fax: +49-6421-94849-61
Email: monika.kawohl@accovion.com