

You Can't Spell **ADaM** without **Metadata**

Jeffrey Abolafia and Ryan Burns, Rho
Inc., Chapel Hill, NC

Disclaimer

The views and opinions expressed in the following are those of the individual presenter and should not be attributed to Rho, Inc., its directors, officers, employees, Special Interest Area Communities or affiliates, or any organization with which the presenter is affiliated.

Agenda

- Why Metadata
- Metadata Architecture
 - Project level
 - Datasets/variables/value
 - Results
- To Excel or **Not** to Excel: the interface
- Related Tools
- Conclusions

Why Metadata

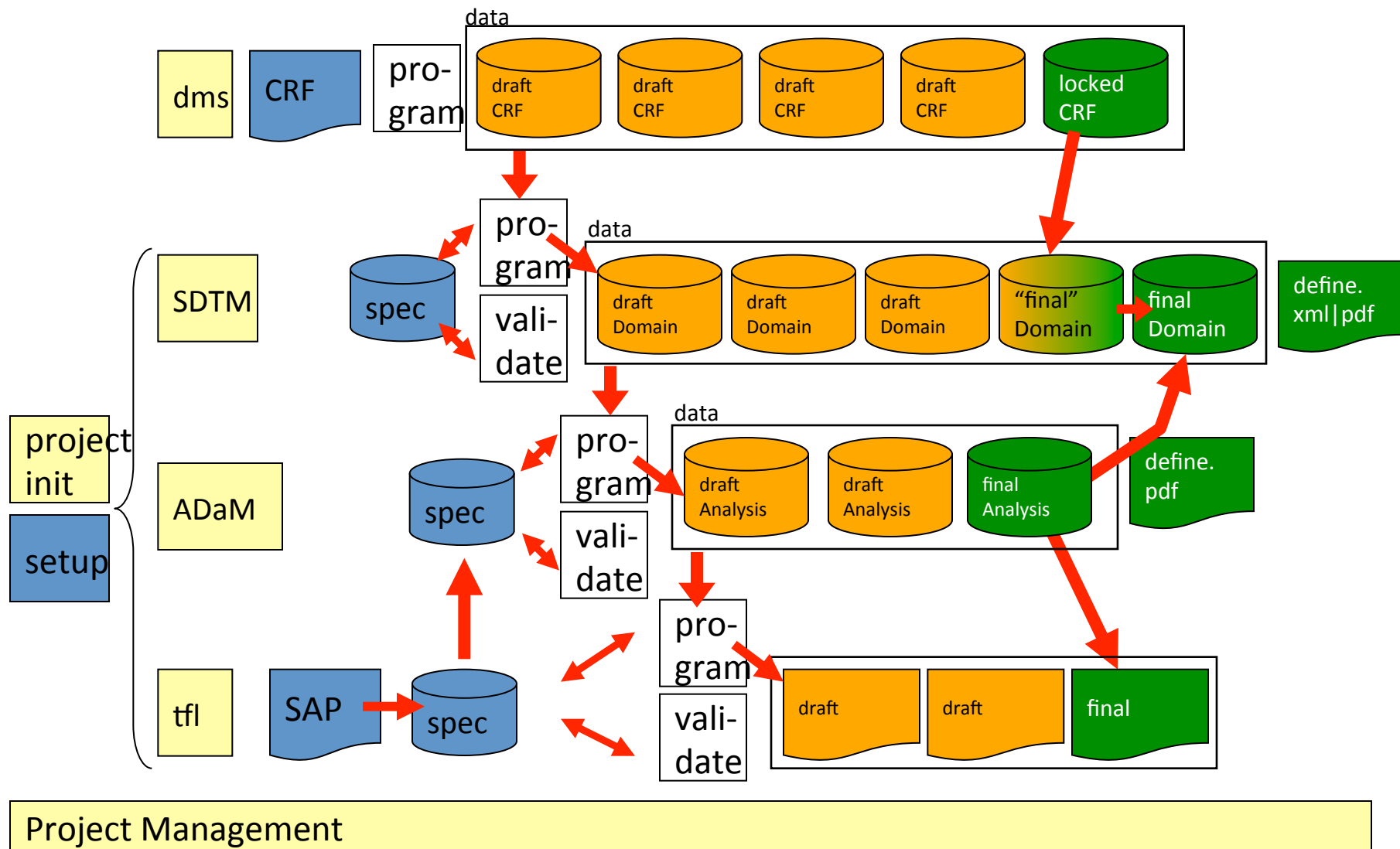
- Project complexity – many deliverables, detailed (and frequently-changing) data, display specifications.
- Many audiences
- This raises many management, quality, process issues: how can costs be controlled? Can we remove redundant tasks? How can we improve communication between project managers, analysts, and programmers? ...
- Address many of these issues by using metadata and metadata access tools throughout a project life cycle
- Result is significant cost reduction, and improved quality and reliability of output.

Why Metadata

- More than a tool to produce datasets, displays, and define
- A corporate asset
- Management/organizational tool
- Tool to control quality, costs, and time

Workflow

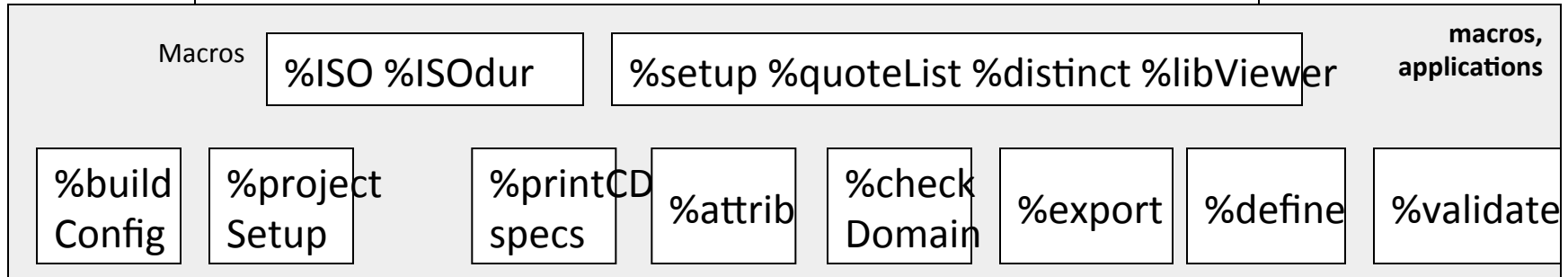
time →



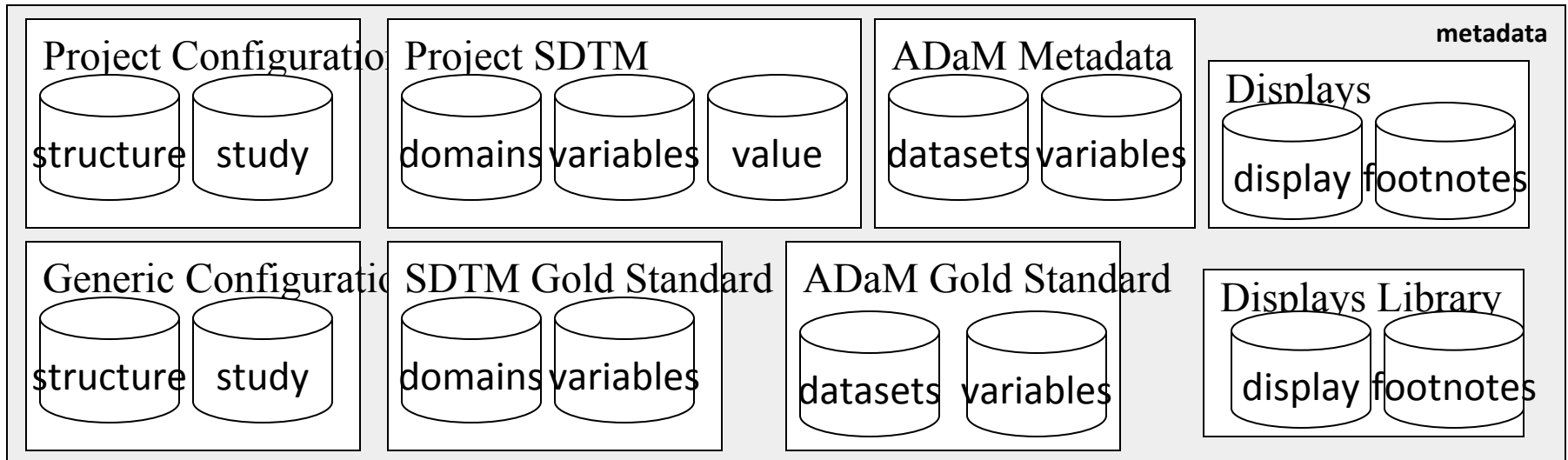
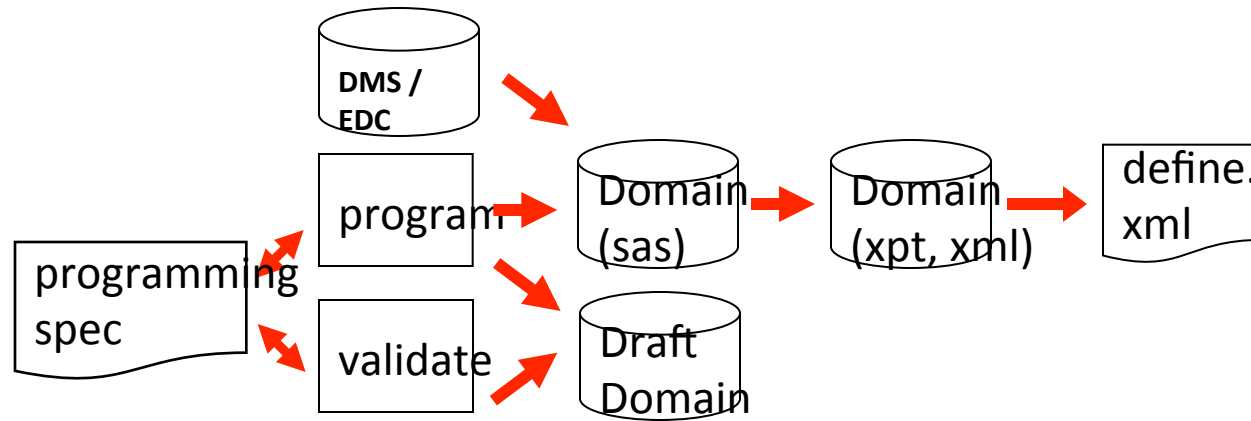
Goals of a Metadata System

- Facilitate creating ADaM Datasets / Displays
- Documentation / Traceability
- Define File
- Store ADaM (and other) Standards
- Gain efficiencies
- Improve quality
- Create uniform re-useable processes
- Repurposing
- Project Management / tracking

Metadata Architecture



- [-] CS02
 - [-] Analysis
 - [-] Clinical
 - [-] Data
 - [-] CDISC
 - [-] Clinical
 - [-] Derive
 - [-] Domain
 - [-] Randomize
 - [-] SampSize
 - [-] Sub_anl
 - [-] Doc
 - [-] Figures
 - [-] Format
 - [-] From
 - [-] HTML
 - [-] Listings
 - [-] Macros
 - [-] Metadata
 - [-] Analysis
 - [-] CDISC
 - [-] Config
 - [-] Domain
 - [-] Prog
 - [-] Analyses
 - [-] CDISC
 - [-] Clinical



Metadata Tables

- Common
 - Datasets
 - Variables
 - Value level
- Less Common
 - Controlled Terminology
 - Project Level
 - Results

Metadata Variables

- Obvious
 - Dataset and variable attributes
 - Displayed in define
 - Display metadata
- Less Obvious
 - Needed but not displayed in define
 - Needed for hyperlinking
 - Management/tracking
 - Improves efficiency

Project Level Metadata



Project Metadata

ref	macseq	fntseq	libname	directory	engine	readonly	engineOpts	submission	use
GLMACCFG	3	0		\submissions\macros\Config		0		0	1
GLMACSHELL	3	0		\submissions\macros\shell		0		0	1
GLMACSUB	3	0		\submissions\macros\sub		0		0	1
GLMACUTIL	3	0		\submissions\macros\util		0		0	1
GLMET	0	0	GLMET	\submissions\metadata		0		0	1
GLMETANL	0	0	GLMDANL	\submissions\metadata\variables.mdb		0		0	1
GLMETCDISC	0	0	GLMDCD	\submissions\metadata\SDTM311.mdb		0		0	1
GLMETCDSUPP	0	0	GLMDCDSU	\submissions\metadata\SDTM311supp.mdb		0		0	1
GLMETCFG	0	0	GLMDCFG	\submissions\metadata\Config.mdb		0		0	1
GLROOT	0	0		\submissions		0		0	1
GLSETUP	0	0		\submissions\prog\setupFiles		0		0	1
GLTEMP	0	0		\submissions\temp		0		0	1
GLTEMPL	0	0	TEMPLATE	\submissions\ODS		0		0	1
PRFMT	0	2	PRLIB	Format		0		0	1
PRHTML	0	0		HTML		0		0	1
PRMAC	2	0		Macros		0		0	1
PRMACANL	2	0		Macros\Analysis		0		0	1
PRMETCFG	0	0	CONFIG	MetaData\Config		0		0	1
PRMETCFGA	0	0	ACONFIG	MetaData\Config\config.mdb	ACCESS	0	dbmax_text=500	0	1
PRMETDOM	0	0	METADOM	MetaData\Domain		0		0	1
PRMETDOMA	0	0	AMETADOM	MetaData\Domain\variables.mdb	ACCESS	0	dbmax_text=100	0	1
PRPCFG	0	0		Prog\Config		0		0	1
PRPROG	0	0		Prog		0		0	1
PRPSUB	0	0		ProgSubmission		0		0	1
PRROOT	0	0		S:\submissions\TestStudy		0		0	1
STANL	0	0		Analysis		0		0	1
STANLVER	0	0		Analysis\Verify		0		0	1
STCLIN	0	0		Clinical		0		0	1
STDATCDISC	0	0	CDISC	Data\CDISC		0		0	1
STDATCLIN	0	0	RAW	Data\Clinical		0		0	1
STDATDER	0	0	DERIVE	Data\Derive		0		0	1
STDATDOM	0	0	DOMAIN	Data\Domain		0		0	1
STDATSDER	0	0	SDERIVE	Data\Sub_anl		0		0	1
STDOC	0	0		Doc		0		0	1
STFMT	0	1	LIBRARY	Format		0		0	1
STMAC	1	0		Macros		0		0	1
STMET	0	0	META	MetaData		0		0	1
STMETANL	0	0	METAANL	MetaData\Analysis		0		0	1
STMETANLA	0	0	AMETAANL	MetaData\Analysis\variables.mdb	ACCESS	0	dbmax_text=100	0	1
STMETCDISCA	0	0	AMETACD	MetaData\CDISC\SDTM.mdb	ACCESS	0	dbmax_text=100	0	1
STPCDISC	0	0		Prog\CDISC		0		0	1
STPCFG	0	0		Prog\Config		0		0	1
STPCLIN	0	0		Prog\Clinical		0		0	1
STPDER	0	0		Prog\Derive		0		0	1

Project Initiation/Setup Tool

- Copies generic configuration database to project directory
- Study configuration database can be customized
- Project directory structure is set up
- Generic programs are copied and to directory and customized for the project

Project Setup

Advantages of Using Metadata

- Automated project setup
- Setup can be done by all levels of personnel
- Setup is uniform for all projects
 - Same directory structure
 - Same libnames
 - Macro and format library search paths are set
- Generic programs are customized for the project
- Configuration database along with a setup macro is used for all subsequent programs
- Facilitate generating Define.xml

Project Level Metadata Define.xml

**Store project specific metadata needed for
Define.xml**

- Study name/description
- Protocol
- Originator (Sponsor)
- ADaM/SDTM version
- Name/location annotated CRF
- Name/location reviewer's guide
- Name/location style sheet

Datasets/Variables/Value Level Metadata



Interface

- Design is important
- Serve users, not the reverse
- Restrictions, reminders, suggestions
- Multiple users
- Audit trail
- Separate from database means flexibility
- Sharing between projects and standards

Interface


Metadata Import Utility
Help me get started
Main Menu

Sponsor
DEFINEXML5
DEPARTMENTS
DRAIS
EISAI
ELI LILLY
ENDTOEND

Study Title
CDISC Package
EndToEnd

Oracle Study ID Number
1356
1335

Dataset	Description	Structure
ADAE_L1	Contains the data for the Adverse Event Analyses	One record per subject per adverse event
ADAE_L2	Contains the data for the Adverse Event Analyses	One record per subject per adverse event
ADCM_L1	Contains the data for the Prior and Concomitant Medication Analyses	One record per subject per medication
ADCM_SDT	Contains the data for the Prior and Concomitant Medication Analyses	One record per subject per medication
ADEG_L1	Contains the data for Electrocardiogram Analyses	One record per subject per parameter per analysis visit
ADFG_L2	Contains the data for Electrocardiogram Analyses	One record per subject per parameter per analysis visit

Variables

Import	Target Dataset	Source	Short Name	Short Label	Long Name	Long Label
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AEBODSYS	Body System or Org		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AEDECOD	Dictionary-Derived		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AEEDT	End Date of Advers		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AEEDTF	End Date Imputati		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AEEDY	Relative End Day of		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AESDT	Start Date of Adver		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AESDTF	Start Date Imputat		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AESDY	Relative Start Day		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AESEQ	Sequence Number		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AESER	Serious Event		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AETERM	Reported Term for		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AOCCLFL	1st Occurrence of A		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AOCCLPFL	1st Occurrence of P		
☐		EndToEnd, EndToEnd, dataset [ADAE_L1]	AOCCLSFL	1st Occurrence of S		

Select: ☐ All ☐ None
Apply to all rows
View Selected
Import Variables

Target Dataset:

Dataset Level Metadata



DatasetName	dsSc	dsLabel	Variable	Subn	Structure	Class	Description	Input_notes	Output_note	Documentation	Keyfields	FileName	LastEditedBy
ADAE_L1	7.1	Adverse Events	Edit	Yes	One record per subject per adverse event	ADAE	Contains the data for the Adverse Event Analyses				USUBJID, AESEQ		Rob Woolson
ADAE_L2	7.2	Adverse Events	Edit	Yes	One record per subject per adverse event	ADAE	Contains the data for the Adverse Event Analyses				USUBJID, AESEQ		Rob Woolson
ADCM_L1	79.1	Concomitant Medications	Edit	Yes	One record per subject per medication	OTHER	Contains the data for the Prior and Concomitant Medication Analyses	The almost always variables. Datasets used: DERIVE.ADSL CDTM.L1			USUBJID, CMSEQ		Rob Woolson
ADCM_SDT	79.15	Concomitant Medications	Edit	Yes	One record per subject per medication	OTHER	Contains the data for the Prior and Concomitant Medication Analyses	The SDTM variables, just in case you need them. Datasets used: DERIVE.ADSL CDTM.L1			USUBJID, CMSEQ		Rob Woolson
ADEG_L1	80.1	Electrocardiogram Results	Edit	Yes	One record per subject per parameter per analysis visit	BDS	Contains the data for Electrocardiogram Analyses	The almost always variables Datasets used: DERIVE.ADSL CDTM.L1			USUBJID, PARAMCD, AVISIT		Rob Woolson
ADEG_L2	80.2	Electrocardiogram Results	Edit	Yes	One record per subject per parameter per analysis visit	BDS	Contains the data for Electrocardiogram Analyses	The second most common variables.			USUBJID, PARAMCD, AVISIT		Rob Woolson
ADEG_SDT	80.15	Electrocardiogram Results	Edit	Yes	One record per subject per parameter per analysis visit	BDS	Contains the data for Electrocardiogram Analyses	The SDTM variables, just in case you need them.			USUBJID, PARAMCD, AVISIT		Rob Woolson
ADLB_L1	81.1	Laboratory Results	Edit	Yes	One record per subject per parameter per analysis visit	BDS	Contains Laboratory Results.	The almost always variables Datasets used: DERIVE.ADSL CDTM.L1			USUBJID, PARAMCD, AVISIT		Rob Woolson
ADLB_L2	81.2	Laboratory Results	Edit	Yes	One record per subject per parameter per analysis visit	BDS	Contains Laboratory Results	The second most common variables.			USUBJID, PARAMCD, AVISIT		Rob Woolson
ADLB_SDT	81.15	Laboratory Results	Edit	Yes	One record per subject per parameter per analysis visit	BDS	Contains Laboratory Results.	The SDTM variables, just in case you need them.			USUBJID, PARAMCD, AVISIT		Rob Woolson
ADMH_L1	82.1	Medical History	Edit	Yes	One record per subject per body system per MedDRA preferred term per event start date	OTHER	Contains Medical History Results	The almost always variables. Datasets used: DERIVE.ADSL			USUBJID, MHBODYSYS, MHDECOD, MHSOTM		Rob Woolson

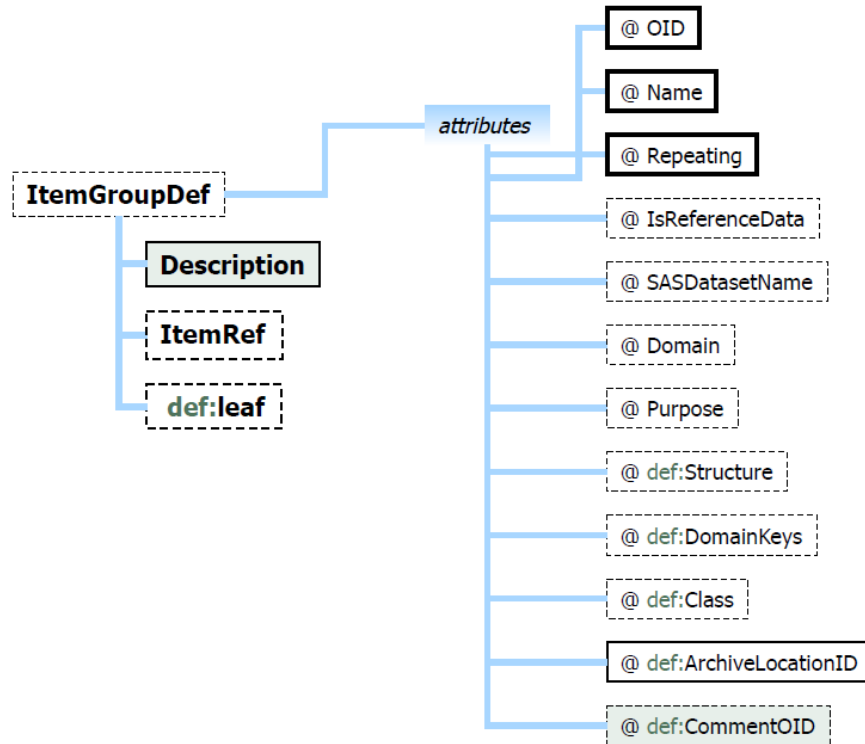
Dataset Level Metadata

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

Column Name	Data Type	Nullable	Data Default
DATASETNAME	VARCHAR2(255 BYTE)	Yes	
DSLABEL	VARCHAR2(40 BYTE)	Yes	
CLASS	VARCHAR2(40 BYTE)	Yes	
STRUCTURE	VARCHAR2(255 BYTE)	Yes	
USAGE	VARCHAR2(255 BYTE)	Yes	
KEYFIELDS	VARCHAR2(255 BYTE)	Yes	
FILENAME	VARCHAR2(255 BYTE)	Yes	
DOCUMENTATION	CLOB	Yes	EMPTY_CLOB()

Dataset Level Metadata



Column Name	Data Type	Nullable	Data Default
SASDATASETNAME	VARCHAR2(255 BYTE)	Yes	
REPEATING	NUMBER(1,0)	Yes	0
ISREFERENCEDATA	NUMBER(1,0)	Yes	0

Dataset Level Metadata

The CONTENTS Procedure

Data Set Name	ADAM.ADAE	Observations	3889
Member Type	DATA	Variables	52
Engine	V9	Indexes	0
Created	Tuesday, April 09, 2013 11:14:21 AM	Observation Length	1840
Last Modified	Tuesday, April 09, 2013 11:14:21 AM	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label	Adverse event data		
Data Representation	WINDOWS_32		
Encoding	wlatin1 Western (Windows)		

Column Name	Data Type	Nullable	Data Default
IDDATASETNAME	NUMBER(10,0)	No	
IDSTUDY	NUMBER(10,0)	No	
SUBMITDB	VARCHAR2(3 BYTE)	Yes	
INPUT_NOTES	CLOB	Yes	
OUTPUT_NOTES	CLOB	Yes	
PROGCOMPLETE	NUMBER(1,0)	Yes	0
PROGBY	VARCHAR2(255 BYTE)	Yes	
PROGDATETIME	DATE	Yes	
CREATINGPROGRAM	VARCHAR2(255 BYTE)	Yes	
INCLUDECOMMONVARS	VARCHAR2(3 BYTE)	Yes	
LASTEDITDBY	VARCHAR2(255 BYTE)	Yes	
LASTEDITDATETIME	DATE	Yes	

Variable & Value Level Metadata

Rho[®] Analysis Metadata Editor Oracle, v5.13 EndToEnd EndToEnd (Commercial) Variables for Dataset: ADPE_L1 Menu Export to Excel Print Datasets														
Dataset ADPE_L1														
idDatasetName	Sut	Sor	Name612	ParamID	Label612	Definition	ODMType	Type	Length	Format	VarsUsed	Codes	FDAdefinition	In/
ADPE_L1	Yes	1510	PARAM	*ALL*	Parameter	PE.PETEST	T	C	40		text		PE.PETEST. Analysis Parameter	
ADPE_L1	Yes	1511	PARAMCD	*ALL*	Parameter Code	PE.PETESTCD	T	C	8		text	[Insert code list]	PE.PETESTCD. Analysis parameter code.	
ADPE_L1	Yes	1512	PARAMN	*ALL*	Parameter (N)	[Useful for ordering and programmatic manipulation. There must be a one-to-one mapping with PARAM. Must be an integer.]	I	N	8		integer	[Insert code list]	Numeric representation for PARAM.	
ADPE_L1	Yes	1513	PARAMTYP	*ALL*	Parameter Type	[Populate as needed.]		C	20		text	(PARAMTYP)	Analysis parameter type.	
ADPE_L1	Yes	1530	AVAL	*ALL*	Analysis Value	ADPE.AVAL is the corresponding value (for subject and visit) of PE.PESTRESN when PE.PETESTCD=ADPE.PARAMCD	I	N	8		float		Analysis value for physical examination result for corresponding body system, derived from	
ADPE_L1	Yes	1531	AVALC	*ALL*	Analysis Value (C)	ADPE.AVAL is the corresponding value (for subject and visit) of PE.PESTRESN when PE.PETESTCD=ADPE.PARAMCD	T	C	1		text	N=Normal, A=Abnormal	Character analysis value for physical examination result for corresponding body system,	
ADPE_L1	Yes	1550	BASE	*ALL*	Baseline Value	ADPE.AVAL when ADPE.ABLFL = "Y"	I	N	8		float		Numeric baseline analysis value, derived from ADPE.AVAL on the baseline record flag.	
ADPE_L1	Yes	1551	BASEC	*ALL*	Baseline Value (C)	ADPE.AVALC when ADPE.ABLFL = "Y".	T	C	1		text	N=Normal, A=Abnormal	Character baseline analysis	

Record: 5 of 28
No Filter
Search

Variable & Value Level Metadata

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

Variable	Label	Type	Length / Display Format	Controlled Terms or Format	Source/Derivation/Comment
STUDYID	Study Identifier	text	12		Predecessor: ADSL.STUDYID
SITEID	Study Site Identifier	text	3		Predecessor: ADSL.SITEID
SITEGR1	Pooled Site Group 1	text	3		Predecessor: ADSL.SITEGR1
USUBJID	Unique Subject Identifier	text	11		Predecessor: ADSL.USUBJID

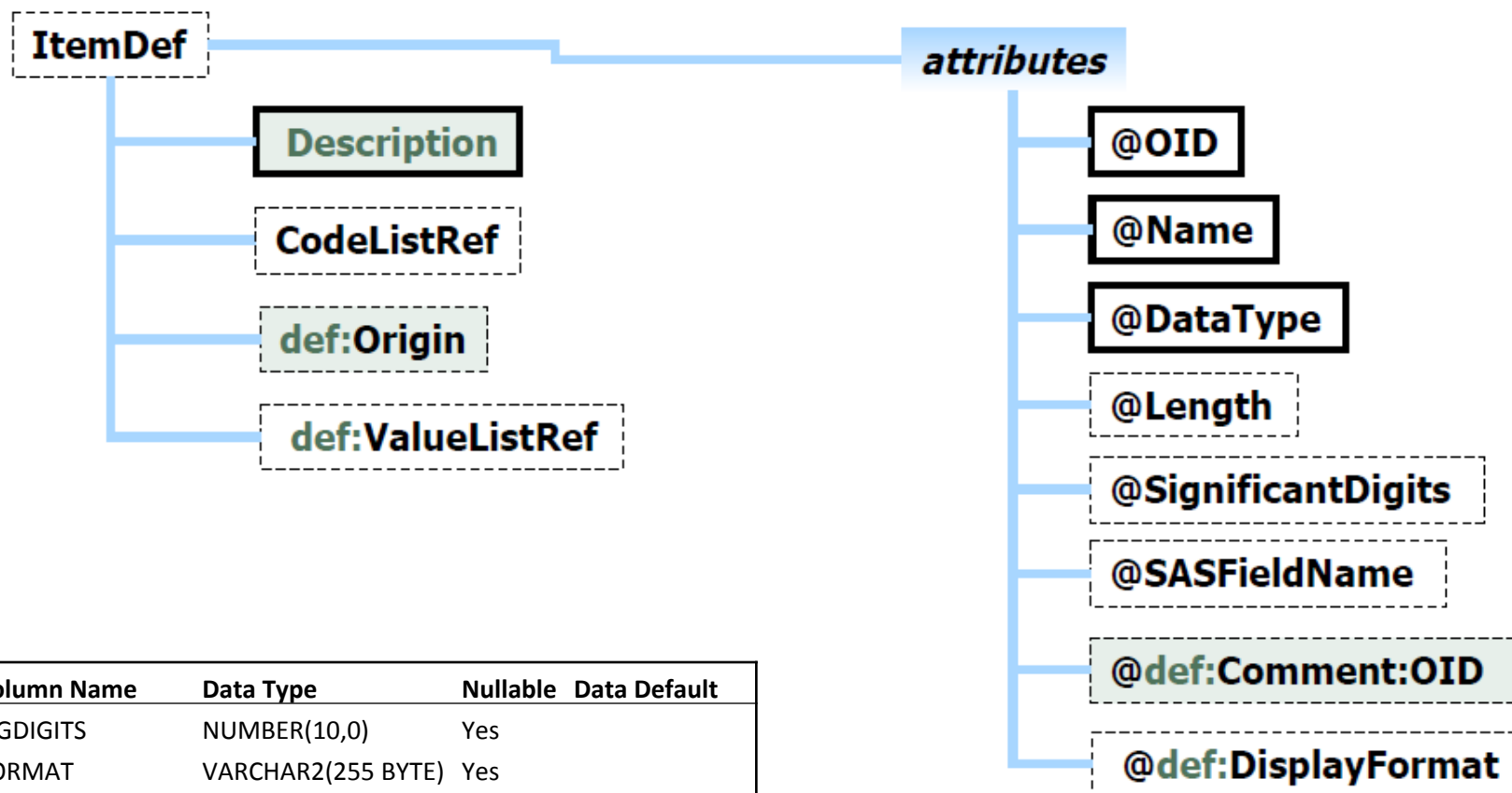
Column Name	Data Type	Nullable
VARNAME	VARCHAR2(32 BYTE)	Yes
VARLABEL	VARCHAR2(255 BYTE)	Yes
ODMTYPE	VARCHAR2(40 BYTE)	Yes
LENGTH	NUMBER(10,0)	Yes
IDCTNAMES	NUMBER(10,0)	Yes
ORIGIN	VARCHAR2(255 BYTE)	Yes
FDADEFINITION	CLOB	Yes
COMPALG	CLOB	Yes
WHERECLS	CLOB	Yes

Parameter Value List - ADQSADAS [QSSEQ]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
QSSEQ	PARAMCD NOTIN ("ACTOT" (Adas-Cog(11) Subscore))	integer	8		Predecessor	
QSSEQ	PARAMCD EQ ACTOT (Adas-Cog (11) Subscore)	integer	8		Assigned	Set QSSEQ to missing for post baseline records. Set to QS.QSSEQ where QS.VISIT=BASELINE and QS.QSTESTCD=ACTOT.

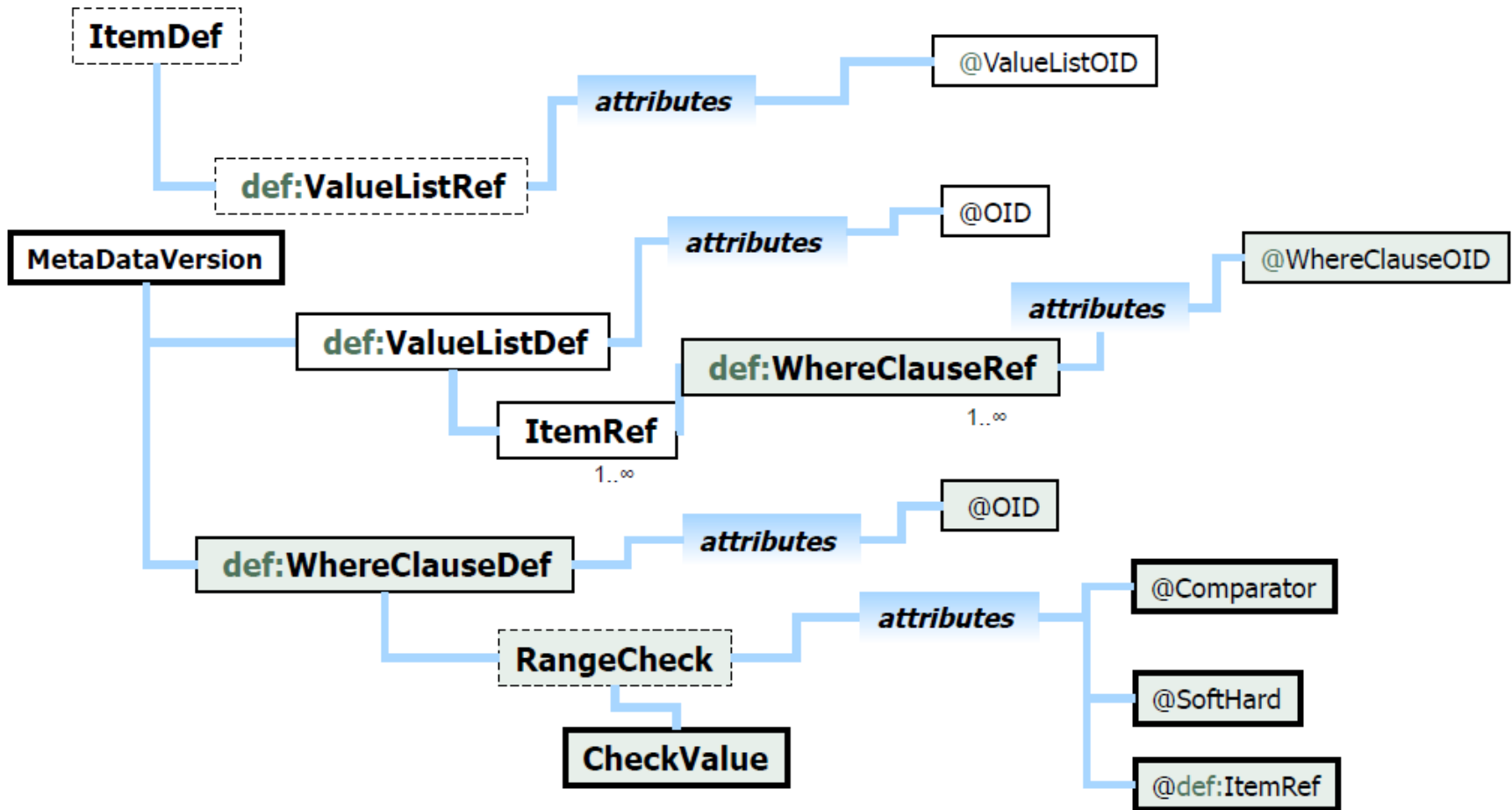
Variable & Value Level Metadata

ItemDef Element



Column Name	Data Type	Nullable	Data Default
SIGDIGITS	NUMBER(10,0)	Yes	
FORMAT	VARCHAR2(255 BYTE)	Yes	
SASVARNAME	VARCHAR2(8 BYTE)	Yes	

Variable & Value Level Metadata



Variable & Value Level Metadata

#	Variable	Type	Len	Format	Label
31	ABLFL	Char	1		ABLFL
23	ADT	Num	8	DATE9.	Analysis Date
22	ADY	Num	8	8.	Analysis Relative Day
9	AGE	Num	8	8.	Age
10	AGEGR1	Char	5		Pooled Age Group 1
11	AGEGR1N	Num	8	8.	Pooled Age Group 1 (N)
32	ANL01FL	Char	1		Analysis Record Flag 01
27	AVAL	Num	8		Analysis Value

Column Name	Data Type	Nullable	Data Default
IDVARIABLES	NUMBER(10,0)	No	
IDSTUDY	NUMBER(10,0)	No	
IDDATASETNAME	NUMBER(10,0)	Yes	
TYPE	VARCHAR2(32 BYTE)	Yes	
SUBMITDB	VARCHAR2(3 BYTE)	Yes	('Yes')
SORTORDER	FLOAT	Yes	999
DEFINITION	CLOB	Yes	
INALLDATASETS	NUMBER(1,0)	Yes	0
PROGCOMPLETE	NUMBER(1,0)	Yes	0
PROGBY	VARCHAR2(255 BYTE)	Yes	
PROGDATETIME	DATE	Yes	
LASTEDITDBY	VARCHAR2(255 BYTE)	Yes	
LASTEDITDATETIME	DATE	Yes	

Controlled Terminology Metadata

Rho[®]

Controlled Terminology Selection

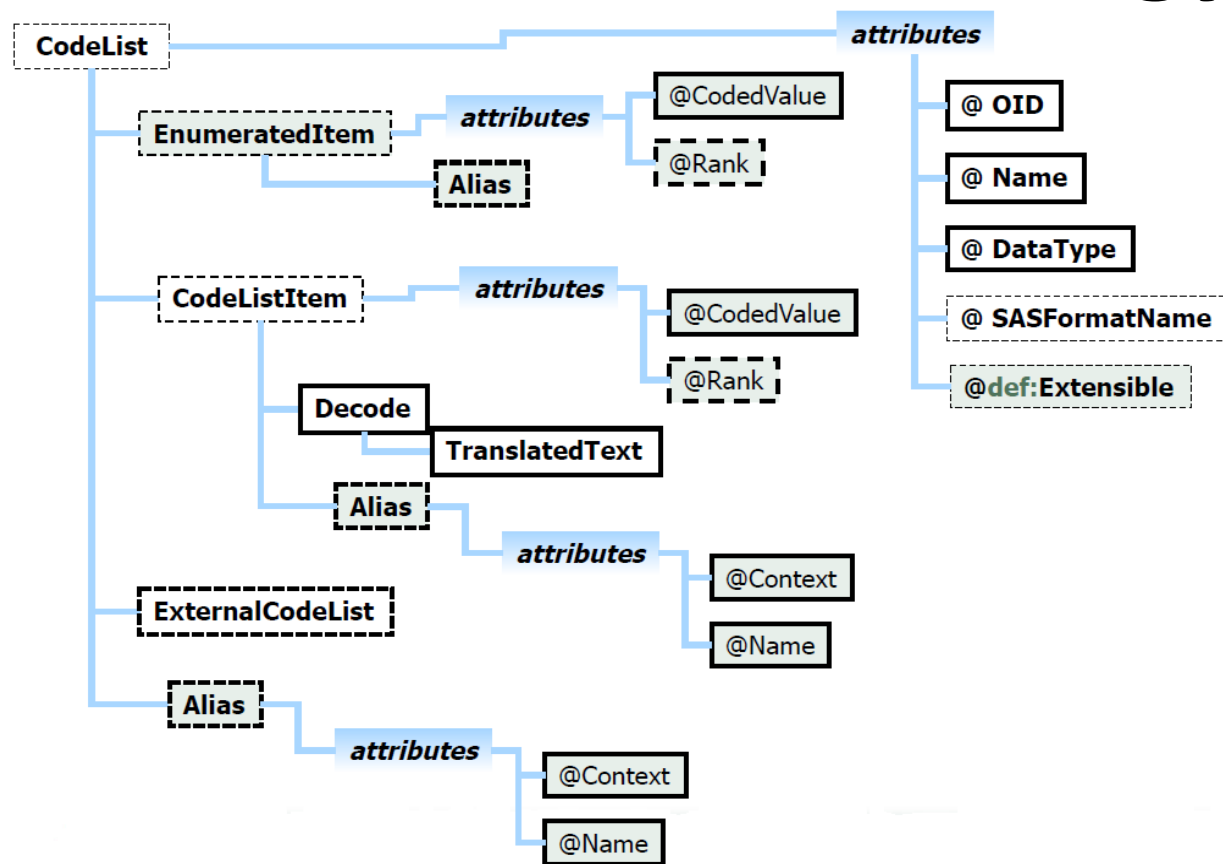
Close

CT: REL (Relationship) [Edit Description](#)

UseValue	SubmissionValue	OKtoChange	Name	Sort Order	Input Code
☒	UNKNOWN	Y	REL	0 99	
☒	NOT RELATED	Y	REL	1 0	
☒	UNLIKELY	Y	REL	2 1	
☒	POSSIBLE	Y	REL	3 2	
☒	PROBABLE	Y	REL	4 3	
☒	DEFINITE	Y	REL	5 4	
☒					

Record: 7 of 7 No Filter Search

Controlled Terminology Metadata



Sex [CL.SEX, C66731]

Permitted Value (Code)	Display Value (Decode)
F [C16576]	Female
M [C20197]	Male
U [C17998]	Unknown

Controlled Terminology Metadata

Lists

IDCTNAMES	LISTNAME	NCILISTCODE	LISTDESCRIPTION	EXTENSIBLE	LISTCUSTOM	LISTSYNONYM	LISTVERSION	NCITYPE	LISTNCIPREFTERM	LISTPREFTERM	FMTNAME
71	REL		Relationship	Y	1						

Values

IDCTVALUES	IDCTNAMES	CVALVALUE	CVALCODE	CVALSORTORDER	CVALCUSTOM	NCICODE	CVALDEFINITION	CVALNCIPREFTERM	CVALPREFTERM
18	71	UNKNOWN		6	1				
17	71	DEFINITE		5	1				
16	71	PROBABLE		4	1				
15	71	POSSIBLE		3	1				
14	71	UNLIKELY		2	1				
7	71	NOT RELATED		1	1				

Selections

IDCTUSED	IDSTUDY	IDCTVALUES	USEDPCODE	USEDPSORT
30	941	18	99	0
33	941	17	4	
34	941	16	3	
35	941	15	2	
29	941	14	1	
27	941	7	0	

Results Level Metadata



Results Level Metadata

- Documents results/analysis
- Not submitted in Define File by most sponsors
- Mostly used for efficient production of displays
- Same metadata be used for
 - Tracking
 - Project management
 - Validation reports

Results Level Metadata Display Production

Table EG_TAB Summary of 12-Lead Electrocardiogram Values by Treatment Group and Visit Population: Safety			
Characteristic Visit Statistic	Treatment A N=xx	Treatment B N=xx	All Treatments N=xx
Parameter 1 (unit)			
Baseline			
n	xx	xx	xx
Mean	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x
Range (Min, Max)	(xx, xx)	(xx, xx)	(xx, xx)
Visit 1			
n	xx	xx	xx
Mean	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x
Range (Min, Max)	(xx, xx)	(xx, xx)	(xx, xx)
Visit 2			
n	xx	xx	xx
Mean	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x
Range (Min, Max)	(xx, xx)	(xx, xx)	(xx, xx)
Etc.			
Suggested Footnotes: BL F01			
Dataset Required: [ADEG]			
Variables Required:			
Safety Population: [SAFFL='Y']			
Treatment: [TRT01P]			
Parameter/Unit: [PARAM]			
Baseline Value: [BASE]			
Value: [AVAL]			
Visit: [AVISIT]			
From TLF Library mock EG_TXB_01			

Metadata

Results Level Metadata Display Production

**PARAM /
AVISIT/ AVALC**

Table EG_TAB
Summary of 12-Lead Electrocardiogram Values by Treatment Group and Visit
Population: Safety

Characteristic Visit Statistic	TRTP	Treatment A N=xx	Treatment B N=xx	All Treatments N=xx
Parameter 1 (unit)				
Baseline				
n		xx	xx	xx
Mean		xx.x	xx.x	xx.x
SD		xx.xx	xx.xx	xx.xx
Median		xx.x	xx.x	xx.x
Range (Min, Max)		(xx, xx)	(xx, xx)	(xx, xx)
Visit 1				
n		xx	xx	xx
Mean		xx.x	xx.x	xx.x
SD		xx.xx	xx.xx	xx.xx
Median		xx.x	xx.x	xx.x
Range (Min, Max)		(xx, xx)	(xx, xx)	(xx, xx)
Visit 2				
n		xx	xx	xx
Mean		xx.x	xx.x	xx.x
SD		xx.xx	xx.xx	xx.xx
Median		xx.x	xx.x	xx.x
Range (Min, Max)		(xx, xx)	(xx, xx)	(xx, xx)
Etc.				

**ADaM
Annotation**

Suggested Footnotes: [BL F01](#)

Dataset Required: [ADEG]

Variables Required:

Safety Population: [SAFFL='Y']

Treatment: [TRT01P]

Parameter/Unit: [PARAM]

Baseline Value: [BASE]

Value: [AVAL]

Visit: [AVISIT]

From TLF Library mock EG_TXB_01

Results Level Metadata

Display Production

- **Display Level Metadata**
 - Contents – specific information about a single display
 - Structure – one record per display
- **Title / Footnote Metadata**
 - Contents – actual titles/footnotes used in displays
 - Structure – one record unique title/footnote
- **Project Level Metadata**
 - Contents
 - Project information
 - Library locations
 - Study design parameters including treatment groups
 - SAS Options/Client preferences
 - Structure – one record per study

Display Level Metadata

Display Production

Number	TableLetter	Title	Population	Population_Label	Selection_Criteria	Footnotes	Dataset
8.7.3.1	DM_TAA	Demographics Characteristics	ITTFL = 'Y'	Intent to Treat Population		ITT5, FT30, FT31	ADSL
8.7.3.2	AE_TAB	Adverse Events by Treatment Group	SAFFL = 'Y'	Safety Population		PPP5, FT30	ADAE
8.7.4.1	VS_TAB	Vital Sign Values by Treatment Group	SAFFL = 'Y'	Safety Population		FT1,FT2, FT3, SRC5	ADVS

Outcome	Treatment	TreatmentValues	TimeVariab	TimePoints	BaselineV	Analysis Variables
	TRT01P	1,2				AGE,WEIGHTB, HEIGHTB
	TRT01P	1,2				ANYAE, AEBODSYS, AEDECOD
AVAL	TRT01P	1,2	AVISIT	1,2,3,4,5	BASE	AVAL

Footnote Level Metadata Display Production

Values in the “Tables” sheet column “Footnotes” correspond to row values of “Footnote Code” in the “Footnotes” sheet

N
Footnotes
ITT3, FT30, FT31, FT31b, SR C3

Footnote codes are entered in the order in which the formatted footnote will appear in the table.

	A	B	C	D
1	Footnote Code	Population Label	Study ID Label	Footnote Text
2	ITT3	Intent to Treat Population	Study 03-103	Note: The Intent to Treat (modified) population consists of all subjects randomized and receiving at least one dose of study drug who also completed the baseline and at least one post-baseline assessment of corneal staining, conjunctival staining, and ocular symptoms.
13	SRC3		Study 03-103	Source: Study Report 03-103, Appendix 16.4
14	FT30			Note: Worst symptom is defined as the symptom with the highest (most severe) rating in the study eye at Visit 2, averaging across diary entries for days during the placebo run-in period.
15	FT31			Note: P-values correspond to the test of an overall treatment effect adjusting for site. Categorical variables were assessed via the Cochran-Mantel-Haenszel chi-square test and continuous variables were assessed via analysis of variance.
16	FT31b			Note: For treatment comparisons for race, categories are collapsed to Caucasian and Non-Caucasian.

A change to the footnote text will be reflected in all tables using the footnote. There are no changes to individual programs.

Project Level Metadata Display Production

- Name and description of study
- Location on network of study
- Location of study database, format, and macro libraries
- Other directory locations
- SAS Options
- Treatment group information
- Visit information
- Client specific presences for displays
 - Styles
 - formats

Generalized Program Display Production

Complete Calling Program (DM_TAA.SAS)

```
%inc 'S:\RHO\sponsor\project\prog\setup.sas' / nosource2 ;  
%setup(cad=yes);
```

Read Project level metadata to get project level information

```
%GetTable(tbl=TAA);
```

Read display metadata for rows related to table TAA. Write SAS macro variables, then run the shell associated with the table.

Results Level Metadata

Define.xml

- Used in updated Pilot Project
- If already using to produce displays, only several additional fields needed for Define file
- Adds significant value to submission

Table_14-3.12

Display	Table 14-3.12 Mean NPI-X total score from Weeks 4 through Week 24 - Windowed (Efficacy Population)
AnalysisResult	Descriptive statistics (n, mean, standard deviation, median, min, and max)
Analysis Parameter(s)	NPTOTMN=Mean NPI-X (9) Total (Week 4 to 24)
Analysis Variable(s)	AVAL BASE
Reason	pre-specified in SAP
Data References (incl. Selection Criteria)	ADQSNPIX [EFFFL="Y" and PARAMCD="NPTOTMN" and AVISIT="Weeks 4-24" and ANL01FL="Y"]
Documentation	SAP Section 10.2
Programming Statements	PROC UNIVARIATE data= ADQSNPIX (where =(EFFFL = 'Y' and PARAMCD = 'NPTOTMN' and AVISIT='Weeks 4-24' and ANL01FL= 'Y')); Var BASE AVAL; By TRTPN; Run;

“Updated SDTM/ADaM pilot Package”, online at www.cdisc.org/sdtm-adam-pilot-project

Results Level Metadata

Define.xml

Variable	Display Production	Define.xml
Display Identifier	Yes	Yes
Display Name	Yes	Yes
Analysis Result	No	Yes
Analysis Parameters	Yes	Yes
Analysis Variables	Yes	Yes
Reason	No	Yes
Dataset	Yes	Yes
Selection Criteria	Yes	Yes
Documentation	No	Yes
Programming Statements	Maybe	Yes

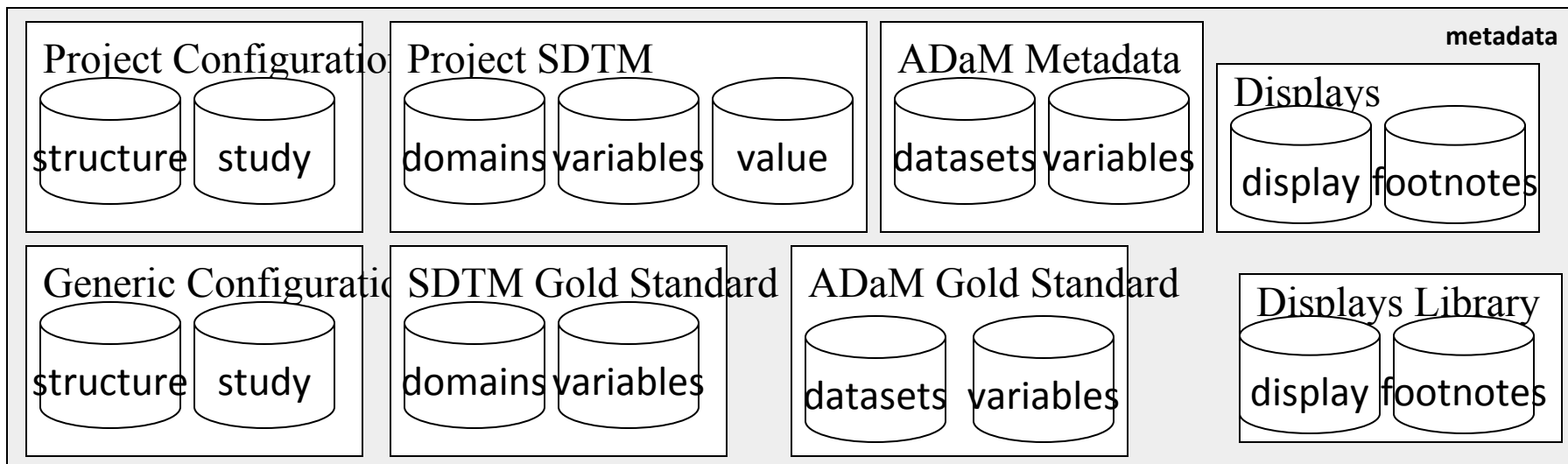
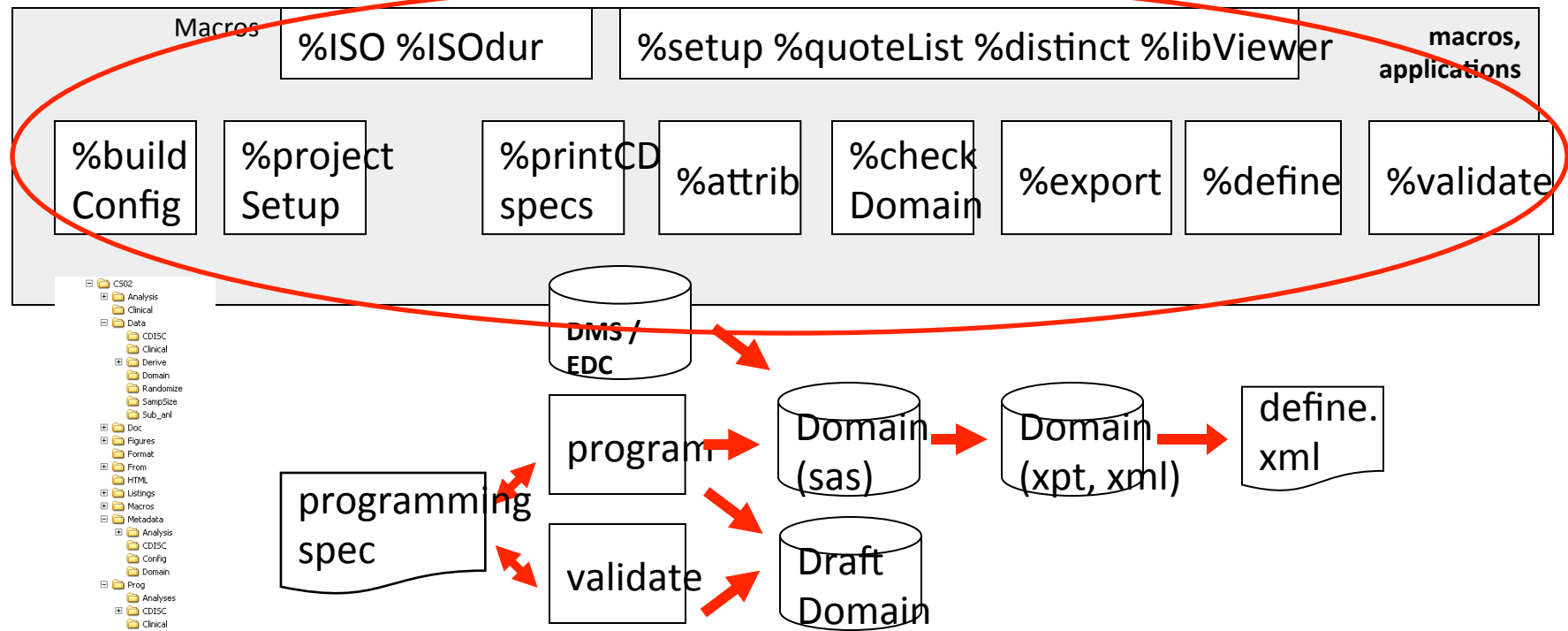
Results Level Metadata Management/tracking

Name	Number	Task	Status	Owner	Created	Started	Completed	Next Due Date	Instructions
TAA_DS_01	14.1.1	Create Program	Completed	kreece	08-Oct-2012 09:37:36 AM	11-Oct-2012 11:12:46 AM	11-Oct-2012 11:12:47 AM		
TAA_DS_01	14.1.1	Create Validation Program	Completed	pnguyen	08-Oct-2012 09:37:36 AM	16-Oct-2012 04:24:23 PM	16-Oct-2012 04:24:23 PM		
TAA_DS_01	14.1.1	Validate Cosmetics	Completed (Pass)	rwoolson	08-Oct-2012 09:37:36 AM	06-Dec-2012 09:50:05 AM	06-Dec-2012 09:50:18 AM		
TAA_DS_01	14.1.1	Validate Statistics	Completed (Pass)	pnguyen	08-Oct-2012 09:37:36 AM	19-Oct-2012 10:53:05 AM	19-Oct-2012 03:52:05 PM		
TAA_DS_01	14.1.1	QC	Superseded		08-Oct-2012 09:37:36 AM				
TAA_DS_01	14.1.1	Revise Program(1)	Completed	kreece	21-Nov-2012 12:39:13 PM	04-Dec-2012 09:42:05 AM	04-Dec-2012 09:42:05 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator
TAA_DS_01	14.1.1	Verify Revision(1)	Completed (Pass)	pnguyen	21-Nov-2012 12:39:13 PM	06-Dec-2012 09:43:05 AM	06-Dec-2012 09:43:05 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator
TAA_DS_01	14.1.1	Validate Cosmetics(1)	Completed (Pass)	rwoolson	21-Nov-2012 12:39:13 PM	06-Dec-2012 09:50:05 AM	06-Dec-2012 09:50:18 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator
TAA_DS_01	14.1.1	Validate Statistics(1)	Completed (Pass)	pnguyen	21-Nov-2012 12:39:13 PM	06-Dec-2012 09:43:08 AM	06-Dec-2012 09:43:08 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator
TAA_DS_01	14.1.1	QC(1)	Completed (Pass)	rwoolson	21-Nov-2012 12:39:13 PM	06-Dec-2012 09:50:28 AM	06-Dec-2012 09:50:36 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator

Results Level Metadata Management/tracking

Hide/Show Columns							CSV	PDF	Print
Search:									
Name	Links	Status	Title	Population	Group	Number	Validation Method		
TAD_DM_01 [Tasks]		Started	Demographic and Baseline Characteristics	Intent-to-Treat	Demographics	14.1.3	Double Programming		
TAD_DM_02 [Tasks]		Started	Demographic and Baseline Characteristics	Clinically Evaluable	Demographics	14.1.4	Double Programming		
TAD_DM_03 [Tasks]		Started	Demographic and Baseline Characteristics	Enrolled Safety	Demographics	14.1.5	Double Programming		
TBI_EX_01 [Tasks]		Started	Exposure to Final Dose by Final Dose Group	Enrolled Safety	Exposure and Compliance	14.3.5.5	Double Programming		
TBI_EX_02 [Tasks]		Started	Exposure to Final Dose by Final Dose Group	Randomized Safety	Exposure and Compliance	14.3.5.6	Double Programming		
TAG_EX_01 [Tasks]		Started	Exposure to Treatment	Enrolled Safety	Exposure and Compliance	14.3.5.1	Double Programming		
TAG_EX_02 [Tasks]		Started	Exposure to Treatment	Randomized Safety	Exposure and Compliance	14.3.5.2	Double Programming		
TAH_FD_01 [Tasks]		Started	Final Dose by Subgroups	Enrolled Safety	Exposure and Compliance	14.3.5.3	Double Programming		
TAH_FD_02 [Tasks]		Started	Final Dose by Subgroups	Randomized Safety	Exposure and Compliance	14.3.5.4	Double Programming		
TAX_AEO_01 [Tasks]		Started	Overview of Adverse Events During Double-Blind Phase	Randomized Safety	Adverse Events	14.3.1.3	Double Programming		
TAW_AEO_03 [Tasks]		Started	Overview of Adverse Events During Entire Study	Enrolled Safety	Adverse Events	14.3.1.4	Double Programming		
TAW_AEO_01 [Tasks]		Started	Overview of Adverse Events During Open-Label Phase	Enrolled Safety	Adverse Events	14.3.1.1	Double Programming		
TAW_AEO_02 [Tasks]		Started	Overview of Adverse Events During Open-Label Phase	Randomized Safety	Adverse Events	14.3.1.2	Double Programming		
TBE_VS_02 [Tasks]		Started	Potentially Clinically Significant (PCS) Vital Sign Values During Double-Blind Phase	Randomized Safety	Vital Signs	14.3.5.10	Double Programming		
TBE_VS_01 [Tasks]		Started	Potentially Clinically Significant (PCS) Vital Sign Values During Entire Study	Enrolled Safety	Vital Signs	14.3.5.9	Double Programming		

Metadata Tools



Metadata Tools

Among the *many* tools developed so far:

- Standardized study set up
- Standardized program startup (replaces hard-coded LIBNAMEs, etc.)
- Retrieve specifications (standards) and generate variable lists (ATTRIB, RENAME, etc.) statements
- Generate dataset creation spec in any format requested by a client (PDF, RTF, XML, etc.)
- Validate datasets conform to a given standard
- Compare datasets to metadata
- Create Export / “Define” files for FDA submission (perform data-metadata consistency checks)
- Validate define file
- Get display metadata

Metadata Tools: Define.xml

```
%inc 's:\submissions\macros\util\setup.sas' / nosource2 ;
```

```
%setup(project=S:\RHO\sponsor\project)
```

```
%let crf = &path\Doc\CRF\blankcrf\blankcrf.pdf;
```

```
%defineXMLprep4(type=ADAM, order=meta) /* ADAM or SDTM or CRT */
```

```
%defineXML(pdf=no, type=ADAM)
```

Conclusion

- Investing in a metadata system pays
- It's efficient
- Metadata can be repurposed throughout the life cycle of a project
- Metadata-centric systems provide a means to manage a variety of requirements
- Tedious programming task can be moved from programmers to admins

Conclusion

- Complementary metadata tools are essential to effectively leverage use of metadata
- The user interface for metadata entry is critical. If the users don't like it, they won't use it
- Continuously hold focus group sessions with end-users.
- End users are the best source of data for what's working and what's not working
- As the number of users and projects increase, consider and plan for upgrades to the backend database.

Metadata Resources

- **Managing the Change and Growth of a Metadata-based System**, *Jeffrey Abolafia, Frank Dilorio*
SAS Global Forum 2008
<http://www2.sas.com/proceedings/forum2008/128-2008.pdf>
- **From CRF Data to DEFINE.XML: Going “End to End” with Metadata**, *Frank C Dilorio, Jeffrey M Abolafia*,
PharmaSUG 2007
<http://www.lexjansen.com/pharmasug/2007/fc/FC04.pdf>
- **The Design and Use of Metadata: Part Fine Art, Part Black Art**, *Frank C. Dilorio, Jeffrey M. Abolafia*,
PharmaSUG 2006
<http://www.lexjansen.com/pharmasug/2006/TechnicalTechniques/TT10.pdf>
- **Controlling Controlled Terminology**
Ryan Burns, Rho
PhUSE 2012
<http://www.phusewiki.org/docs/2012/PAPERS/PP/PP24.pdf>

Thank you for attending our presentation

Send questions or comments to:

Jeff Abolafia (jeff_abolafia@rhoworld.com)

Ryan Burns (ryan_burns@rhoworld.com)

Acknowledgements

Frank Dilorio