ICON
A Symbol of Excellence

Clinical Research

A Data Drive Approach for Generating Define.xml v2.0 using ADaM Specifications and ADaM Datasets

Agenda

- ADaM CoE in ICON
- Introduction
- Define.xml Content Overview
- Pre-requirements
- Process and Validation
- General Information

ADaM CoE in ICON

- ICON is committed to innovation and standardization. Our CoE is focused on leveraging our significant global expertise in CDISC standards from CDASH through eSubmission.

Overview of ADaM CoE:

- Streamlined end-to-end processes (Consistency/Accuracy/Timeliness)
- In-house tools
  - Standard ADaM Dataset Generation
  - Report Generation
  - Submission Packages
- Highly trained Programmers and Biostatisticians

Introduction

- Case Report Tabulation Data Definition Specification format in XML format or Define.xml or Define.pdf (v1.0 only)
- Required document when doing CDISC based electronic submission
- Provides information to help understand the collected data (SDTM) and data used for analysis (ADaM)
- Latest version is Define.xml v2.0

Content Overview

- Define.xml provides metadata for
  - Datasets:
    - Name, Description, Class, Structure, Purpose, Keys, Location Documentation
  - Variables:
    - Name, Label, Type, Length/Display Format, Controlled Terms Source/Derivation/Comment
  - Parameter Value Level Metadata
  - Controlled Terminology
  - Analysis Derivations
  - Comments
  - External Documentation (ADRG)

Pre-requirements

- Define Stylesheet (define2-0-0.xls) for define.xml v2.0

  OR

  define.css, define1-0-0.xls and define-extension.xsd for define.xml v1.0

- ADaM Datasets
  - Trimmed ADaM Datasets
  - SAS Transport Files using Trimmed ADaM Datasets
- ADaM ADS Specifications in required format

Pre-requirements

- Documentation
  - ADRG
  - SAS Codes of ADaM Datasets
  - Other documentation that is being referred in ADS Specifications

Process

- ADaM Datasets
  - Trimming of ADaM Datasets
    - Character variable length adjusted as per data values
    - Use ads_trimming.sas to trim ADaM Datasets
      - Trimmed ADS
      - CSV file with names of ADS > 5 GB
    - Datasets > 5 GB will be divided into two or more smaller datasets with size <= 5 GB.
    - Generate SAS Transport Files using PROC COPY

- ADaM ADS Specifications
  - Formatting of ADaM ADS Specifications

Process – ADaM ADS Specifications

ICON
A Symbol of Excellence

Clinical Research

- Study Information:

Parameter	Description
StudyName	111-P-222
StudyDescription	Define for Study XYZ
ProtocolName	Define for Study XYZ
StandardName	ADaM-IG
StandardVersion	1.0
Language	en

Process – ADaM ADS Specifications

- Analysis Metadata:

Dataset	Description	Structure	Purpose	Class	KeyVariables
ADSL	Subject Level Analysis	One record per screened subject	Analysis	SUBJECT LEVEL ANALYSIS DATASET	USUBJID
ADVS	Vital Signs	One record per subject per parameter per visit	Analysis	BASIC DATA STRUCTURE	USUBJID,PARAMCD,AVISITN

Repeating	Documentation	Document
No	Refer to Analysis data reviwer's guide for futher information	LF.ADRG; LF.ADSL.PGM;
Yes	SAP, Protocol	LF.ADVS.PGM;

Process – ADaM ADS Specifications

ICON
A Symbol of Excellence
Clinical Research

- Variable Metadata:

Dataset	Variable	Label	Type	Length	Sigdigit	Key	Format	CodeList
ADSL	STUDYID	Study Identifier	text	15				
ADSL	SITEID	Study Site Identifier	text	25				
ADSL	INVID	Investigator Identifier	text	25				
ADSL	INVNAM	Investigator Name	text	200				
ADSL	USUBJID	Unique Subject Identifier	text	25		1		
ADSL	SUBJID	Subject Identifier for the Study	text	25				
ADSL	RANDNO	Randomization Number	integer	8				
ADSL	BRTHDTC	Date/Time of Birth	datetime	19				
ADSL	AGE	Age	integer	8				
ADSL	AGEU	Age Units	text	6				AGEU

Process – ADaM ADS Specifications

- Variable Metadata:

Origin	Source	ValueList	Mandotary	FormalExpression	Document
Predecessor	DM.STUDYID		Yes		
Predecessor	DM.SITEID		Yes		
Predecessor	DM.INVID		Yes		
Predecessor	DM.INVNAM		No		
Predecessor	DM.USUBJID		Yes		
Predecessor	DM.SUBJID		No		
Derived	DS.DSREFID where DS.DSSCAT="RANDOMIZATION"		No		
Predecessor	DM.BRTHDTC		No		
Predecessor	DM.AGE		Yes		
Predecessor	DM.AGEU		Yes		

Process – ADaM ADS Specifications

ICON
A Symbol of Excellence

Clinical Research

- ValueLevel Metadata:

Dataset	Variable	Label	Type	Length	Sigdigit	Format	Origin	PARAMCD	Source	Document
ADVS	PARAM	Parameter	text	200			Derived	VSALL	PARAM= trim(left(VS.VSTEST))	
ADVS	PARAM	Parameter	text	200			Derived	BMI	PARAM= trim(left(VS.VSTEST))	
ADVS	PARAM	Parameter	text	200			Derived	DIABP	PARAM= trim(left(VS.VSTEST))||" ("||trim(left(VS.VSSTRESU))||")"	
ADVS	PARAM	Parameter	text	200			Derived	HEIGHT	PARAM= trim(left(VS.VSTEST))||" ("||trim(left(VS.VSSTRESU))||")"	
ADVS	PARAM	Parameter	text	200			Derived	PULSE	PARAM= trim(left(VS.VSTEST))||" ("||trim(left(VS.VSSTRESU))||")"	
ADVS	PARAM	Parameter	text	200			Derived	SYSBP	PARAM= trim(left(VS.VSTEST))||" ("||trim(left(VS.VSSTRESU))||")"	
ADVS	PARAM	Parameter	text	200			Derived	WEIGHT	PARAM= trim(left(VS.VSTEST))||" ("||trim(left(VS.VSSTRESU))||")"	
ADVS	PARAMN	Parameter (N)	integer	8			Derived	SYSBP	if PARAMCD="SYSBP" then PARAMN=1;	
ADVS	PARAMN	Parameter (N)	integer	8			Derived	DIABP	if PARAMCD="DIABP" then PARAMN=2;	
ADVS	PARAMN	Parameter (N)	integer	8			Derived	PULSE	if PARAMCD="PULSE" then PARAMN=3;	
ADVS	PARAMN	Parameter (N)	integer	8			Derived	HEIGHT	if PARAMCD="HEIGHT" then PARAMN=4;	
ADVS	PARAMN	Parameter (N)	integer	8			Derived	WEIGHT	if PARAMCD="WEIGHT" then PARAMN=5;	
ADVS	PARAMN	Parameter (N)	integer	8			Derived	BMI	if PARAMCD="BMI" then PARAMN=6	
ADVS	PARAMN	Parameter (N)	integer	8			Derived	VSALL	if PARAMCD="VSALL" then PARAMN=100;	

Process – ADaM ADS Specifications

ICON
A Symbol of Excellence
Clinical Research

- Controlled Terminology:

Codelist	Name	Datatype	Code	Decode	Rank	OrderNumber	CodeListNum	Codenum
AGEU	Age Unit	text	YEARS			1	C66781	C29848
SEX	Sex	text	M			1	C66731	C20197
SEX	Sex	text	F			2	C66731	C16576
RACE	Race	text	AMERICAN INDIAN OR ALASKA NATIVE			1	C74457	C41259
RACE	Race	text	ASIAN			2	C74457	C41260
RACE	Race	text	BLACK OR AFRICAN AMERICAN			3	C74457	C16352
RACE	Race	text	NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER			4	C74457	C41219
RACE	Race	text	WHITE			5	C74457	C41261
AVISIT	AVISIT	text	Screening			1		
AVISIT	AVISIT	text	Day -2 to -1			2		
AVISIT	AVISIT	text	Day 1 to 3			3		
AVISIT	AVISIT	text	Day 8			4		
AVISITN	AVISITN	integer	1	Screening		1		
AVISITN	AVISITN	integer	2	Day -2 to -1		2		
AVISITN	AVISITN	integer	3	Day 1 to 3		3		
AVISITN	AVISITN	integer	4	Day 8		4		
PARAM_ADVS	PARAM_ADVS	text	All Vital Signs Data			7		
PARAM_ADVS	PARAM_ADVS	text	Body Mass Index			6		
PARAM_ADVS	PARAM_ADVS	text	Diastolic Blood Pressure (mmHg)			2		
PARAM_ADVS	PARAM_ADVS	text	Height (cm)			4		
PARAM_ADVS	PARAM_ADVS	text	Pulse Rate (BEATS/MIN)			3		
PARAM_ADVS	PARAM_ADVS	text	Systolic Blood Pressure (mmHg)			1		
PARAM_ADVS	PARAM_ADVS	text	Weight (kg)			5		

Process

- Run define.sas tool to generate define.xml v2.0

Process

ICON
A Symbol of Excellence
Clinical Research

- Validate define.xml
  - Run Pinnacle 21
    - All the issues/errors/warnings identified should be resolved prior to final package or a valid explanation should be provided explaining why these errors can't be resolved.
  - Check define.xml Source/Derivation/Comment section
  - Check ALL the external documentation hyper-links
  - Perform all the checks given in ADaM Electronic Submission Material Checklist

General Information

- Define.sas tool refer to ADaM datasets for
  - The order of variables in a particular ADaM dataset
  - The variables attributes
  - Controlled Terminology of Derived variables

Useful Links

- Location of Reference Material:
  - http://cdisc.org/define-xml
  - https://www.pinnacle21.com/

Thank you