

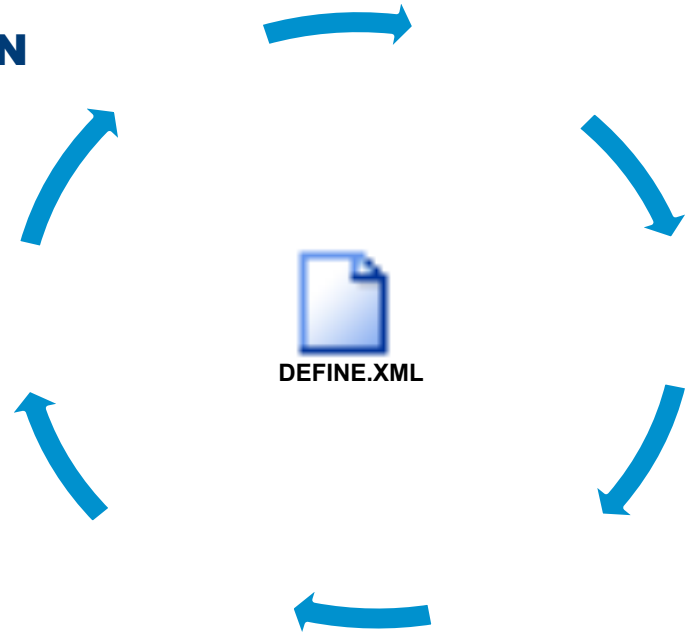
ROUND TRIP TICKET – USING THE DEFINE.XML FILE TO SEND AND RECEIVE YOUR STUDY SPECIFICATIONS

JULIE MADDOX, SAS INSTITUTE



DEFINE.XML CRT-DDS CASE REPORT TABULATION DATA DEFINITION

- Machine-readable metadata describing the format and content of the submission datasets
- Can be used as a template to extract(read) metadata for a new study, or to transmit(write) metadata about a completed study



DEFINE.XML ROUND TRIP - CREATE DEFINE FILE



Sponsor

- Creates define.xml based on study specifications



DEFINE.XML

DEFINING THE STUDY SPECIFICATIONS

Sponsor

- Identify CDISC standard and version
- Identify the domains
- Identify code lists
- Determine the computational algorithms
- Identify the value-level metadata

DEFINING THE STUDY SPECIFICATIONS

Domain/Table Metadata

ItemGroupDef

Datasets for Study CDISC01						
Dataset	Description	Class	Structure	Purpose	Keys	Location
EG	ECG Test Results	Findings	One record per ECG observation per visit per subject	Tabulation	STUDYID, USUBJID, EGTESTCD, EGDTC, VISITNUM	eg.xpt

```
<ItemGroupDef OID="EG" Name="EG" Repeating="Yes" IsReferenceData="No"
Purpose="Tabulation" def:Label="ECG Test Results" def:Structure="One record per ECG
observation per visit per subject" def:DomainKeys="STUDYID, USUBJID, EGTESTCD, EGDTC,
VISITNUM" def:Class="Findings" def:ArchiveLocationID="Location.EG">
....
  <def:leaf ID="Location.EG" xlink:href="eg.xpt">
    <def:title>eg.xpt</def:title>
  </def:leaf>
</ItemGroupDef>
```

DEFINING THE STUDY SPECIFICATIONS

Column Metadata

ItemDef

Variable	Label	Type	Controlled Terms or Format	Origin	Role	Comment
STUDYID	Study Identifier	text		Protocol	IDENTIFIER	
DOMAIN	Domain Abbreviation	text		Assigned	IDENTIFIER	
USUBJID	Unique Subject Identifier	text		Derived	IDENTIFIER	Concatenation of STUDYID.SUBJID
EGTEST	ECG Test or Examination Name	text		CRF Page 12	SYNONYM QUALIFIER	
EGPOS	ECG Position of Subject	text		Protocol	RECORD QUALIFIER	Equal to SUPINE
EGORRES	Result or Finding in Original Units	text	EGRESU	CRF Page 12	RESULT QUALIFIER	

```

<ItemDef OID="EG.EGSEQ" Name="EGSEQ" DataType="integer" Length="2" Origin="Derived"
Comment="Sequential number identifying records within each USUBJID and generated using
the key sequence of the domain." def:Label="Sequence Number"/>
<ItemDef OID="EG.EGSPID" Name="EGSPID" DataType="text" Length="6" Origin="Assigned"
Comment="ECG parameter ordering variable" def:Label="Sponsor-Defined Identifier"/>
<ItemDef OID="EG.EGTESTCD" Name="EGTESTCD" DataType="text" Length="6" Origin="Assigned"
Comment="" def:Label="ECG Test or Examination Short Name">
  <def:ValueListRef ValueListOID="ValueList.EG.EGTESTCD"/>
/ItemDef>

```

DEFINING THE STUDY SPECIFICATIONS

Code Lists

CodeList

NYNA, Reference Name (NYNA)	
Code Value	Code Text
Y	YES
N	NO
NA	NOT APPLICABLE
NY, Reference Name (NY)	
Code Value	Code Text
Y	YES
N	NO
U	UNKNOWN

```
<CodeList ID="NYNA" Name="NYNA" DataType="text">
  <CodeListItem CodedValue="N" def:Rank="2">
    <Decode>
      <TranslatedText xml:lang="en">NO</TranslatedText>
    </Decode>
  </CodeListItem>
  <CodeListItem CodedValue="NA" def:Rank="3">
    <Decode>
      <TranslatedText xml:lang="en">NOT APPLICABLE</TranslatedText>
    </Decode>
  </CodeListItem>
  <CodeListItem CodedValue="Y" def:Rank="1">
    <Decode>
      <TranslatedText xml:lang="en">YES</TranslatedText>
    </Decode>
  </CodeListItem>
</CodeList>
```

DEFINING THE STUDY SPECIFICATIONS

Computation Method

ComputationMethod

Computational Algorithms (COMPMETHOD.QTCB)	
Reference Name	Computation Method
COMPMETHOD.QTCB	$QTcB = QT \text{ interval} / \text{square root of } (60 / \text{heart rate})$
Computational Algorithms (COMPMETHOD.QTCF)	
Reference Name	Computation Method
COMPMETHOD.QTCF	$QTcF = QT \text{ interval} / \text{cubic root of } (60 / \text{heart rate})$

```
<def:ComputationMethod OID="COMPMETHOD.QTCB">QTcB = QT interval / square  
root of (60 / heart rate)</def:ComputationMethod>  
<def:ComputationMethod OID="COMPMETHOD.QTCF">QTcF = QT interval / cubic  
root of (60 / heart rate)</def:ComputationMethod>
```

DEFINING THE STUDY SPECIFICATIONS

Value Level Metadata

ValueListDef

Value Level Metadata (ValueList.EG.EGTESTCD)						
Source Variable	Value	Label	Type	Controlled Terms or Format	Origin	Comment
EGTESTCD	INTP	Interpretation	text	NABCLIN	CRF Page 12	
EGTESTCD	PRMEAN	Summary (Mean) PR Duration	integer		CRF Page 12	
EGTESTCD	QRS DUR	Summary (Mean) QRS Duration	integer		CRF Page 12	
EGTESTCD	QTcB	QTcB - Bazett's Correction Formula	float		DERIVED	See Computational Method: COMPMETHOD.QTcB
EGTESTCD	QTcF	QTcF - Fridericia's Correction Formula	float		DERIVED	See Computational Method: COMPMETHOD.QTcF

```

<def:ValueListDef OID="ValueList.EG.EGTESTCD">
  <ItemRef ItemOID="EG.EGTESTCD.PRMEAN" OrderNumber="4" Mandatory="No"/>
  <ItemRef ItemOID="EG.EGTESTCD.QTcB" OrderNumber="6" Mandatory="No"/>
</def:ValueListDef>

<ItemDef OID="EG.EGTESTCD.PRMEAN" Name="PRMEAN" DataType="integer" Length="3"
Origin="CRF Page 12" Comment="" def:Label="Summary (Mean) PR Duration"/>
<ItemDef OID="EG.EGTESTCD.QTcB" Name="QTcB" DataType="float" SignificantDigits="1"
Length="5" Comment="" Origin="DERIVED" def:ComputationMethodOID="COMPMETHOD.QTcB"
def:Label="QTcB - Bazett's Correction Formula"/>

```

DEFINING THE STUDY SPECIFICATIONS

Annotated Case Report Form

AnnotatedCRF

VS=Vital Signs

CDISC Study: CDISC01	Assessment Date: VSDTC ____/____/____
VITAL SIGNS	
VSTEST	VSPOS
Height . ☐ cm ☐ in VSORRES / VSORRESU when VSTESTCD = HEIGHT	Sitting Blood Pressure VSORRES / VSORRESU when VSTESTCD = SYSBP, DIABP / mmHg Systolic Diastolic
Weight . ☐ kgs ☐ lbs VSORRES / VSORRESU when VSTESTCD = WEIGHT	Radial Pulse Rate bpm VSORRES / VSORRESU when VSTESTCD = PULSE
☐ Small Body Frame Size ☐ Medium ☐ Large VSORRES when VSTESTCD = FRMSIZE	

```
<def:AnnotatedCRF>
  <def:DocumentRef leafID="blankcrf"/>
</def:AnnotatedCRF>
<def:leaf ID="blankcrf" xlink:href="blankcrf.pdf">
  <def:title>Annotated Case Report Form</def:title>
</def:leaf>
```

CREATING THE DEFINE.XML FILE



DEFINE.XML

CREATING THE DEFINE.XML FILE

SAS CLINICAL STANDARDS TOOLKIT

Collect metadata
into tables



DEFINE.XML

- Create and populate the SAS datasets containing the study specification metadata
 - %cst_createdsfromtemplate* creates empty table shells

Five metadata tables to capture study specifications

- Source_Study – Study metadata
- Source_Tables – Domain table metadata
- Source_Columns – Domain column metadata
- Source_Values – Domain value-level metadata
- Source_Documents – study documentation metadata

SOURCE_STUDY

sasref	definedocumentname	studyname	studydescription	protocolname	standard	standardversion	formalstandardname	formalstandardversion
SRCDATA	define.xml	CDISC01	CDISC Test Study	CDISC01	CDISC-SDTM	3.1.3	CDISC-SDTM	3.1.2

Study Metadata

- Study Name, Description
- Protocol
- CDISC data standard/version

SOURCE_TABLES

sasref	table	label	class	xmlpath	xmltitle	structure	purpose	keys	date	standard	standardversion
SRCDATA	EG	ECG Test Results	Findings	eg.xpt	eg.xpt	One record per ECG observation per visit per subject	Tabulation	STUDYID USUBJID EGTESTCD EGDTC VISITNUM	2011-06-15	CDISC-SDTM	3.1.3
SRCDATA	EX	Exposure	Interventions	ex.xpt	ex.xpt	One record per constant dosing interval per subject	Tabulation	STUDYID USUBJID EXSTDTC EXENDTC EXTRT EXDOSE	2011-06-15	CDISC-SDTM	3.1.3

Table Metadata

- Table Name, Label
- Keys
- Structure
- Purpose
- CDISC Standard and Version
- Transport File Name

SOURCE_COLUMNS

SASref	table	column	label	order	type	length	displayformat	xmlatatype	xmlcodelist	origin	role	algorithm	standard	standardversion
SRCDATA	AE	STUDYID	Study Identifier	1	C	7		text		Protocol	IDENTIFIER		CDISC-SDTM	3.1.3
SRCDATA	AE	DOMAIN	Domain Abbreviation	2	C	2		text		Assigned	IDENTIFIER		CDISC-SDTM	3.1.3
SRCDATA	AE	USUBJID	Unique Subject Identifier	3	C	14		text		Derived	IDENTIFIER		CDISC-SDTM	3.1.3
SRCDATA	AE	AESEQ	Sequence Number	4	N	1		integer		Derived	IDENTIFIER		CDISC-SDTM	3.1.3
SRCDATA	AE	AESPID	Sponsor-Defined Identifier	5	C	4		text		CRF Page 21	IDENTIFIER		CDISC-SDTM	3.1.3
SRCDATA	AE	AETERM	Reported Term for the Adverse Event	6	C	25		text		CRF Page 21	TOPIC		CDISC-SDTM	3.1.3
SRCDATA	AE	AEMODIFY	Modified Reported Term	7	C	9		text		Assigned	SYNONYM QUALIFIER		CDISC-SDTM	3.1.3
SRCDATA	AE	AEDECOD	Dictionary-Derived Term	8	C	18		text	ADVERSE EVENT DICTIONARY	Assigned	SYNONYM QUALIFIER		CDISC-SDTM	3.1.3

Column Metadata

- Column Name, Label
- Order
- Variable Length, Type
- Code List
- Origin
- Computational Algorithm

SOURCE_VALUES

SASref	table	column	value	label	order	type	length	displayformat	xmldatatype	xmlcodelist	core	origin	role	term	algorithm
SRCDATA	EG	EGTESTCD	INTP	Interpretation	3	C	8		text	NABCLIN		CRF Page 12			
SRCDATA	EG	EGTESTCD	PRMEAN	Summary (Mean) PR Duration	4	N	3		integer			CRF Page 12			
SRCDATA	EG	EGTESTCD	QRS DUR	Summary (Mean) QRS Duration	5	N	3		integer			CRF Page 12			
SRCDATA	EG	EGTESTCD	QTcB	QTcB - Bazett's Correction Formula	6	N	5		float			DERIVED			QTcB = QT interval / square root of (60 / heart rate)
SRCDATA	EG	EGTESTCD	QTcF	QTcF - Fridericia's Correction Formula	7	N	5		float			DERIVED			QTcF = QT interval / cubic root of (60 / heart rate)

Value-level Metadata

- Column Name
- Value , Description
- Order
- Code List
- Origin
- Computational Algorithm

sasref	doctype	documentref	href	title	standard	standardversion
SRCDATA	CRF		blankcrf.pdf	Annotated Case Report Form	CDISC-SDTM	3.1.3

Document Metadata

- Title
- Reference
- Document type

CREATING THE DEFINE.XML FILE

SAS CLINICAL STANDARDS TOOLKIT

Collect metadata in
tables

Execute
`%crtds_sdtmtodefine`



DEFINE.XML

CREATING THE DEFINE.XML

Execute **%crtdds_sdtmtodefine** macro to create SAS representation of CRT-DDS data sets

```
%crtdds_sdtmtodefine(  
  _cstOutLib=crtddsLib,  
  _cstSourceTables=source_tables,  
  _cstSourceColumns=source_columns,  
  _cstSourceValues=source_values,  
  _cstSourceDocuments=source_documents,  
  _cstSourceStudy=source_study  
);
```

CREATING THE DEFINE.XML FILE

SAS CLINICAL STANDARDS TOOLKIT

Collect metadata in
tables

Execute
`%crtds_sdtmtodefine`

Execute
`%crtds_write`



DEFINE.XML

CREATING THE DEFINE.XML

Execute **%crtds_write** macro to create DEFINE.XML from the SAS CRT-DDS data sets

```
%crtds_write(_cstCreateDisplayStyleSheet=1);
```

COLLECTING THE DEFINE.XML METADATA

SAS CLINICAL DATA INTEGRATION

The image displays two overlapping SAS EG Properties dialog boxes. The background dialog is the 'General' tab, showing the dataset name 'EG' and description 'ECG Test Results'. The foreground dialog is the 'Keys' tab, showing a list of keys and a details table.

EG Properties (Background - General Tab)

- Name: EG
- Description: ECG Test Results

EG Properties (Foreground - Keys Tab)

Keys:

- Primary Key**
 - EG_Primary
- Unique Keys**
- Foreign Keys**

Details:

#	Name	Description
1	STUDYID	Study Identifier
2	USUBJID	Unique Subject Ident...
3	EGTESTCD	ECG Test or Examina...
4	VISITNUM	Visit Number
5	EGTPTREF	Time Point Reference
6	EGTPTNUM	Planned Time Point N...

EG Properties (Right - General Tab)

- Domain Name: ECG Test Results
- Data Standard: SDTM312
- Study/Submission: sdtnjm (Study)

Properties:

Name	Value
Archive Filename	../transport/eg.xpt
Archive Title	ECG Test Results SAS transport file
Classification	FINDINGS
Identifier	EG
Is Custom	false
Purpose	TABULATION
Structure	One record per ECG observation per time p
Template Identifier	EG
Type	SDTM

COLLECTING THE DEFINE.XML METADATA

SAS CLINICAL DATA INTEGRATION

EG Properties

Physical Storage		Options	Notes	Extended Attributes	Authorization	Clinical
General		Columns	Indexes	Keys	Parameters	
#	Name	Description	Type	Length	Inform	
1	STUDYID	Study Identifier	Character	40	(None)	
2	DOMAIN	Domain Abbreviation	Character	8	(None)	
3	USUBJID	Unique Subject Identifier	Character	40	(None)	
4	EGSEQ	Sequence Number	Numeric	8	(None)	
5	EGGRPID	Group ID	Character	40	(None)	
6	EGREFID	ECG Reference ID	Character	40	(None)	
7	EGSPID	Sponsor-Defined Identifier	Character	40	(None)	
8	EGTESTCD	ECG Test or Examination Short Name	Character	8	(None)	
9	EGTEST	ECG Test or Examination Name	Character	40	(None)	
10	EGCAT	Category for ECG	Character	40	(None)	
11	EGSCAT	Subcategory for ECG	Character	40	(None)	
12	EGPOS	ECG Position of Subject	Character	40	(None)	
13	EGORRES	Result or Finding in Original Units	Character	200	(None)	
14	EGORRESU	Original Units	Character	40	(None)	
15	EGSTRESC	Character Result/Finding in Std Format	Character	200	(None)	
16	EGSTRESN	Numeric Result/Finding in Standard Units	Numeric	8	(None)	
17	EGSTRESU	Standard Units	Character	40	(None)	
18	EGSTAT	Completion Status	Character	40	(None)	
19	EGREASND	Reason ECG Not Performed	Character	200	(None)	
20	EGXFN	ECG External File Name	Character	200	(None)	

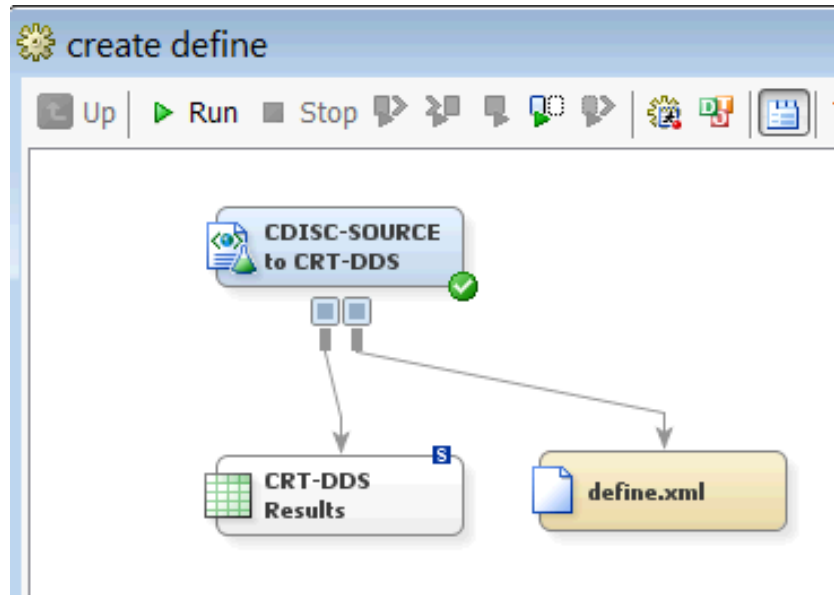
EGTESTCD Properties

Notes	Extended Attributes	Clinical
Label:		
ECG Test or Examination Short Name		
Properties:		
Name	Value	
Contributes to Key	true	
Core	REQ	
Display Format		
Method		
Origin		
Prefix by Identifier	true	
Qualifiers	UPPERCASE	
Role	TOPIC	
Term	(EGTESTCD)	
Type	SDTM	
XML Codelist	EGTESTCD	
XML Type	text	

CREATING THE DEFINE.XML

SAS CLINICAL DATA INTEGRATION

Execute job using the CDISC-SOURCE to CRT-DDS transformation to produce the define.xml



CDISC-SOURCE to CRT-DDS Properties

Precode and Postcode		Parameters	Notes	Extended Attributes	
General	Tables	Generation	Reports	Options	Code

Header comment:
Produced from SAS data using SAS Clinical Data Integration

Output encoding: UTF-8 Log level: Warning

Define.pdf
☒ Create define.pdf file
☐ Create internal PDF hyperlinks

Output Stylesheet
☐ No output stylesheet
☒ Use default output stylesheet
☐ Use a custom output stylesheet
Custom stylesheet location: Browse...

SDTM Datasets for Study CDISC01

Dataset	Description	Class	Structure	Purpose	Keys	Lo
TA	Trial Arms	Trial Design	One record per planned Element per Arm	Tabulation	STUDYID, ARMCD, TAETORD	ta.
TE	Trial Elements	Trial Design	One record per planned Element	Tabulation	STUDYID, ETC	te.
TI	Trial Inclusion/Exclusion Criteria	Trial Design	One record per I/E criterion	Tabulation	STUDYID, IETESTCD	ti.
TS	Trial Summary	Trial Design	One record per trial summary parameter value	Tabulation	STUDYID, TSPARMCD, TSSEQ	ts.
TV	Trial Visits	Trial Design	One record per planned Visit per Arm	Tabulation	STUDYID, VISITNUM, ARMCD	tv.
DM	Demographics	Special Purpose	One record per subject	Tabulation	STUDYID, USUBJID	dm.
SE	Subject Elements	Special Purpose	One record per actual Element per subject	Tabulation	STUDYID, USUBJID, SESTDTC, SEENDTC, TAETORD, ETC	se.

- + Annotated Case Report Form
- + Reviewers Guide
- + SDTM Datasets
 - Trial Arms (TA)
 - Trial Elements (TE)
 - Trial Inclusion/Exclusion
 - Trial Summary (TS)
 - Trial Visits (TV)
 - Demographics (DM)
 - Subject Elements (SE)
 - Subject Visits (SV)
 - Concomitant Medications
 - Exposure (EX)
 - Adverse Events (AE)
 - Disposition (DS)
 - Medical History (MH)
 - Drug Accountability (DA)
 - ECG Test Results (EG)

DEFINE.XML ROUND TRIP - EXTRACT METADATA



Sponsor

- Creates define.xml based on study specifications



**Data Management
Team (CRO)**

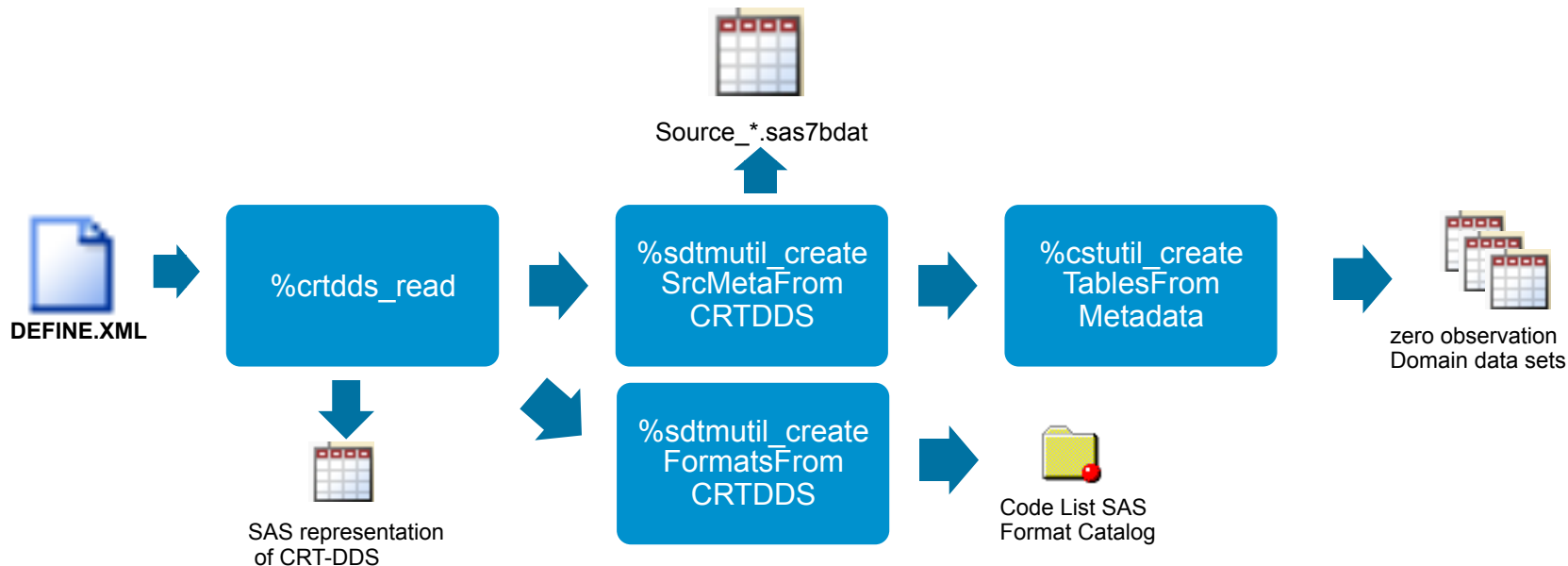
- Extracts metadata from define.xml file



DEFINE.XML

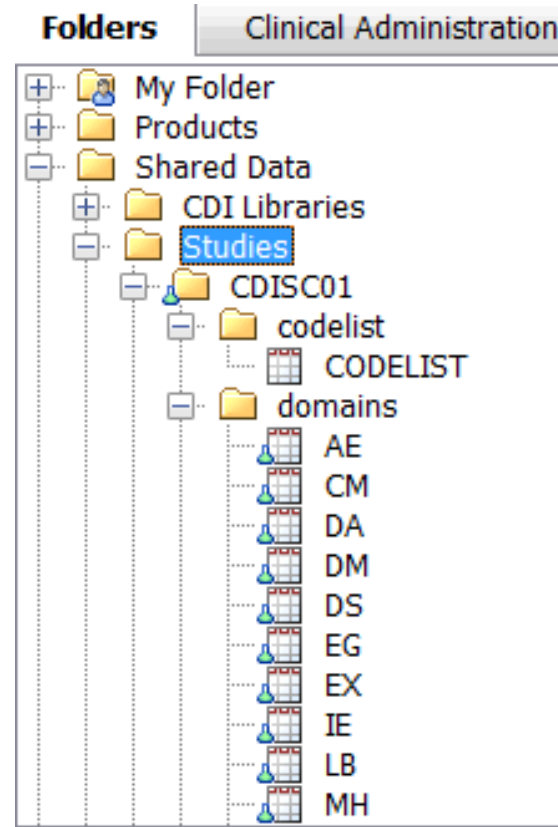
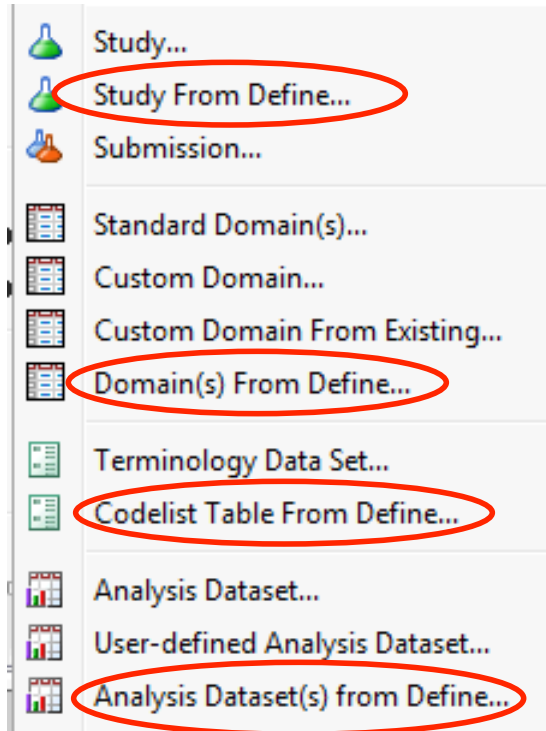
READING THE DEFINE.XML

SAS CLINICAL STANDARDS TOOLKIT

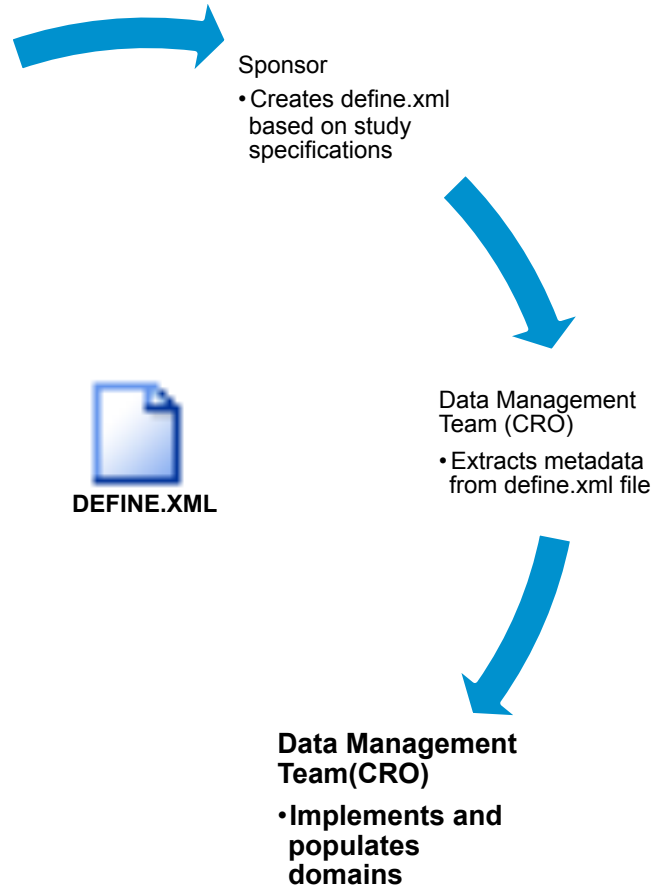


READING THE DEFINE.XML

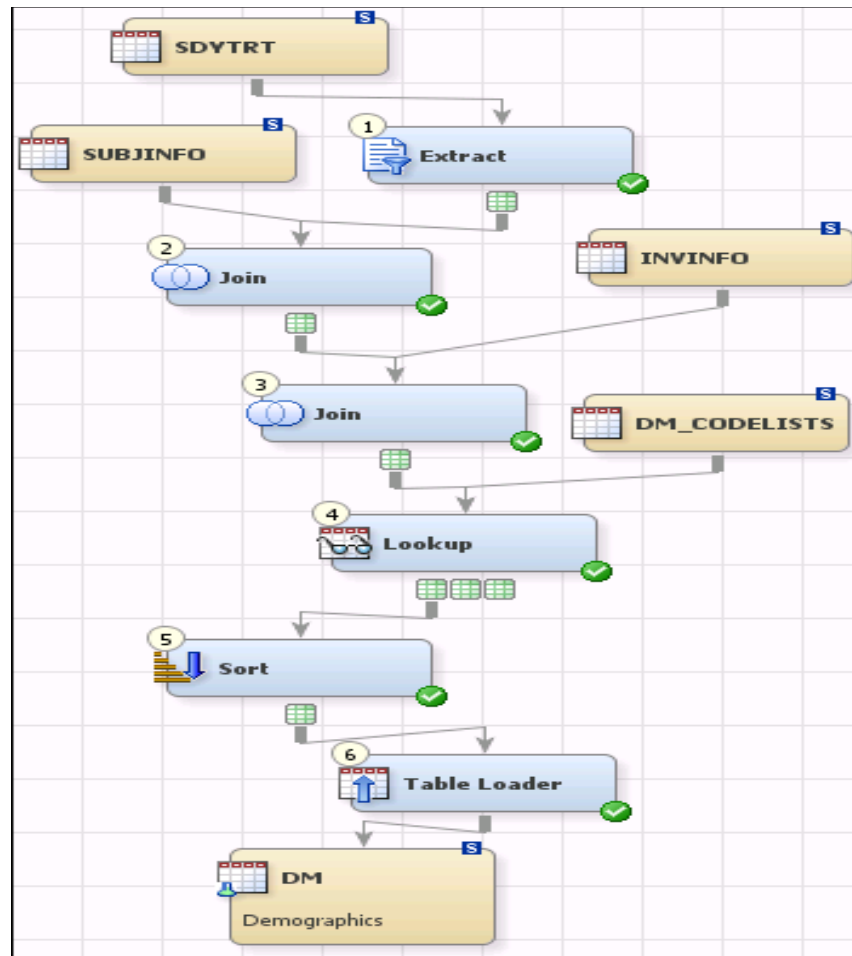
SAS CLINICAL DATA INTEGRATION



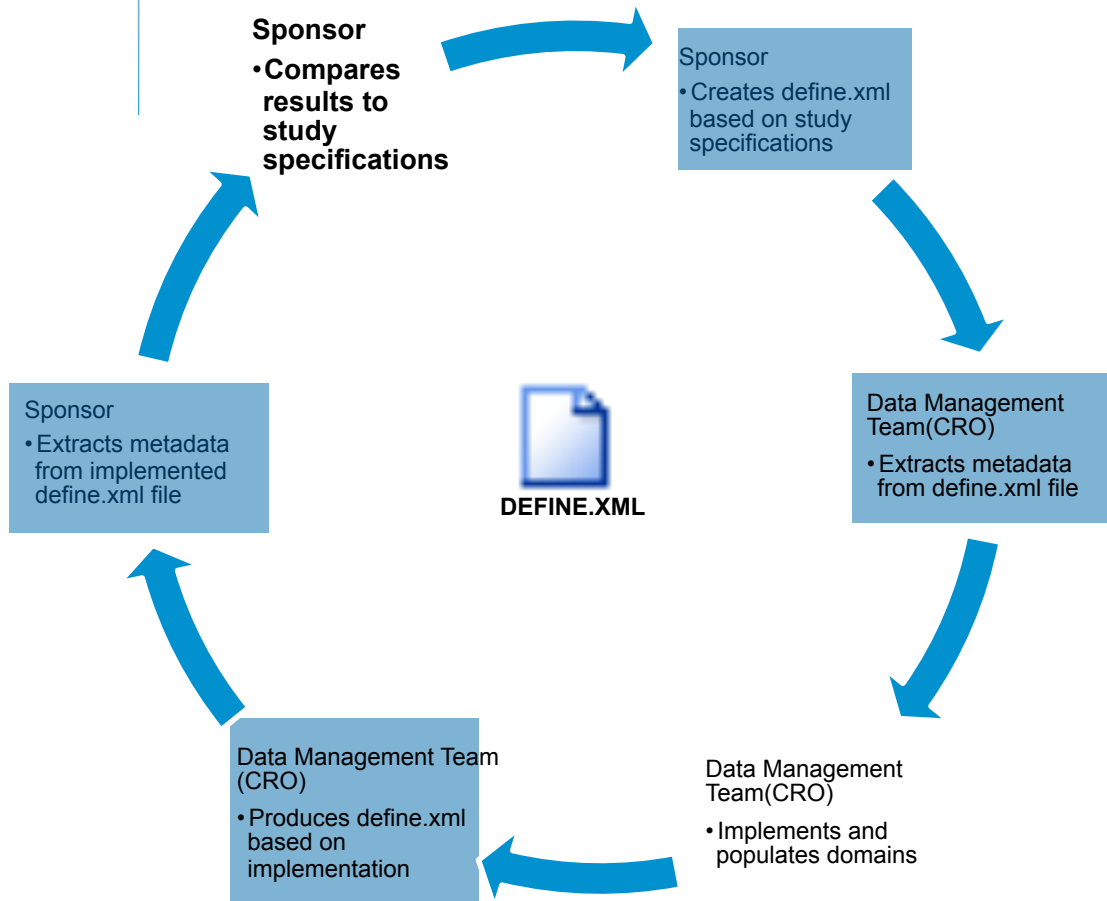
DEFINE.XML ROUND TRIP – POPULATE DOMAINS



IMPLEMENT AND POPULATE DOMAINS

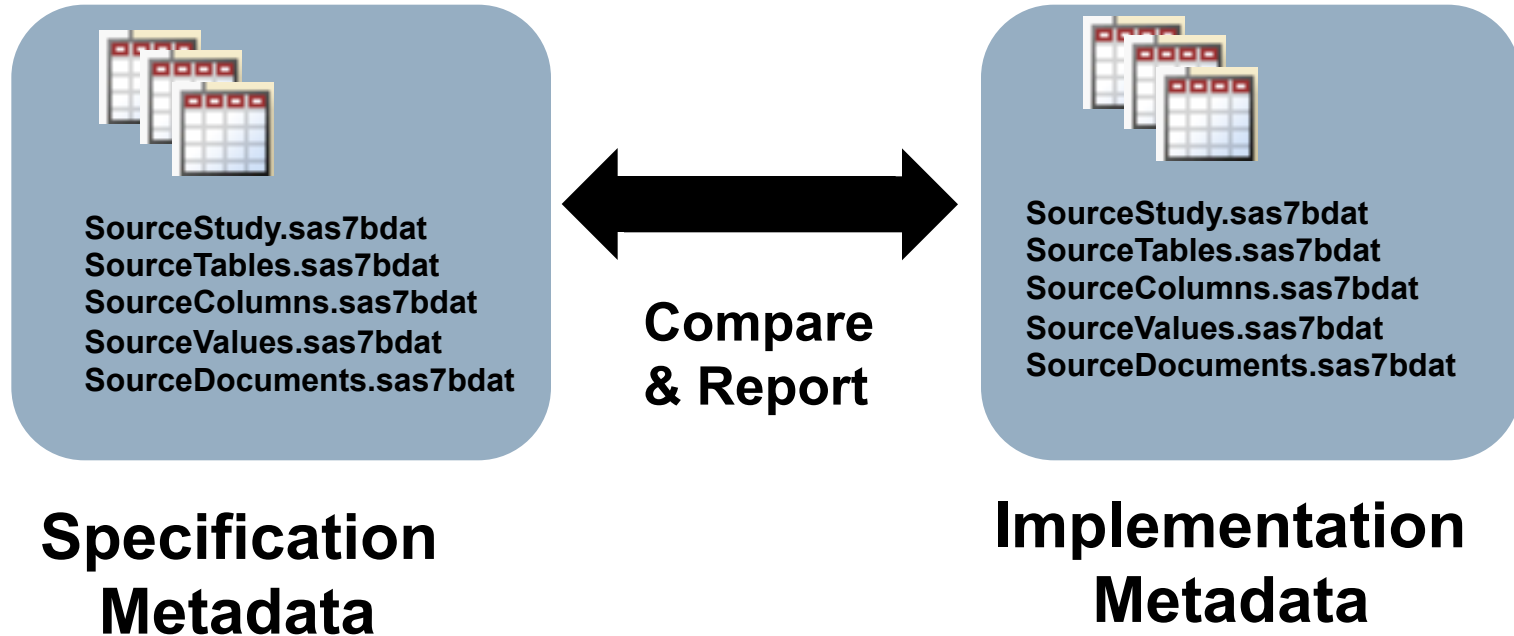


DEFINE.XML ROUND TRIP

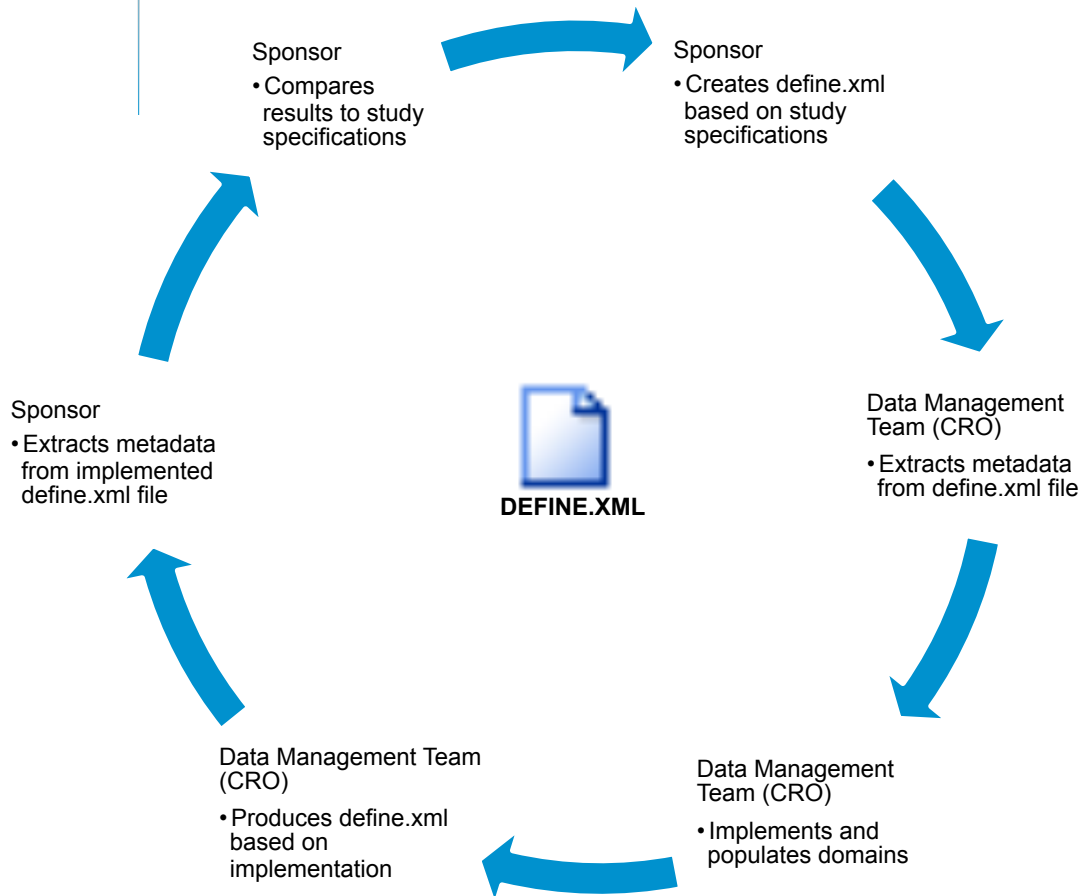


COMPARE AND REPORT

SPECIFICATION VS IMPLEMENTATION



DEFINE.XML ROUND TRIP



SUMMARY

DEFINE.XML

- Contains metadata representation of the study submission datasets
 - Domains
 - Code lists
 - Computational algorithms
 - Value-level metadata
 - CRF, supplemental documentation
- Read and Write Define.xml using SAS Clinical Standards Toolkit or SAS Clinical Data Integration
- Use define.xml to exchange metadata and compare the implementation to the specification

QUESTIONS ??



**THE
POWER
TO KNOW®**