define.xml - it's all about the Metadata

Lex Jansen
Software Developer
SAS

lex.jansen@sas.com

Agenda

- define.xml - background
- define.xml - what is it
- define.xml - content
- define.xml - data model
- define.xml - end-to-end

define.xml - background

define.xml - background

- **July 2004** – FDA adds Study Data Specifications v1.0 to draft eCTD Guidance. This specification references the CDISC SDTM for data tabulation datasets

Electronic Common Technical Document (eCTD)

- Draft Guidance for Industry on Providing Regulatory Submissions in Electronic Format--Human Pharmaceutical Applications and Related Submissions. (Posted 8/28/2003)
  - *Federal Register* Notice [TXT] [PDF]
  - The Draft Guidance

Specifications

- eCTD Backbone Files Specification for Module 1
- eCTD Backbone File Specification for Modules 2 through 5
- eCTD Backbone File Specification for Study Tagging Files
- FDA eCTD Table of Contents Headings and Hierarchy
- Study Data Specifications New!! (Posted 7/21/2004)

define.xml - background

- **March 2005** – Study Data Specifications v1.1: Updates Specifications for Data Set Documentation
  - data definitions
  - annotated case report forms (CRFs)
- *"The specification for the data definitions for datasets provided using the CDISC SDTM is included in the Case Report Tabulation Data Definition Specification* (**define.xml**) *developed by the CDISC define.xml Team"*

define.xml - background

- As of **January 1, 2008**: follow the eCTD guidance and document submitted data by including data definition tables (**define.xml**) and annotated case report forms (blankcrf.pdf)

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<?xml-stylesheet type="text/xsl" href="define1-0-0.xsl"?>
<ODM xmlns="http://www.cdisc.org/ns/odm/v1.2"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
     xmlns:xlink="http://www.w3.org/1999/xlink"
     xmlns:def="http://www.cdisc.org/ns/def/v1.0"
     xsi:schemaLocation="http://www.cdisc.org/ns/odm/v1.2 define1-0-0.xsd"
     FileOID="Studycdisc01"
     ODMVersion="1.2" FileType="Snapshot"
     CreationDateTime="2007-04-09T12:24:09">
     - <Study OID="cdisc01">
       - <GlobalVariables>
            <StudyName>CDISC01</StudyName>
            <StudyDescription>CDISC01 Test Study.</StudyDescription>
            <ProtocolName>CDISC01</ProtocolName>
       </GlobalVariables>
       <MetaDataVersion  OID="CDISC.SDTM.3.1.1"
            Name="Study CDISC01, Data Definitions"
            Description="Study CDISC01, Data Definitions"
            def:DefineVersion="1.0.0"
            def:StandardName="CDISC SDTM"
            def:StandardVersion="3.1.1">
...
   <ItemGroupDef OID="MH"
         Name="MH" Repeating="Yes" IsReferenceData="No"
         Purpose="Tabulation" def:Label="Medical History"
         def:Structure="One record per medical history event per subject"
         def:DomainKeys="STUDYID, USUBJID, MHCAT, MHTERM, MHSTDTC"
         def:Class="EVENTS"
         def:ArchiveLocationID="Location.MH">
     <ItemRef ItemOID="STUDYID" OrderNumber="1" Mandatory="Yes"
        Role="IDENTIFIER" RoleCodeListOID="RoleCodeList" />
   ...

     <def:leaf ID="Location.MH" xlink:href="mh.xpt">
           <def:title>mh.xpt</def:title>
     </def:leaf>
   </ItemGroupDef>
```

define.xml - background

- As of **January 1, 2008**: follow the eCTD guidance and document submitted data by including data definition tables (**define.xml**) and annotated case report forms (blankcrf.pdf)

```
<ItemGroupDef OID="MH"
      Name="MH" Repeating="Yes" IsReferenceData="No"
      Purpose="Tabulation" def:Label="Medical History"
      def:Structure="One record per medical history event per subject"
      def:DomainKeys="STUDYID, USUBJID, MHCAT, MHTERM, MHSTDTC"
      def:Class="EVENTS"
      def:ArchiveLocationID="Location.MH">
  <ItemRef ItemOID="STUDYID" OrderNumber="1" Mandatory="Yes"
     Role="IDENTIFIER" RoleCodeListOID="RoleCodeList" />
...

  <def:leaf ID="Location.MH" xlink:href="mh.xpt">
            <def:title>mh.xpt</def:title>
  </def:leaf>
</ItemGroupDef>
```

define.xml - background

- **May 2011** – FDA CDER **Common Data Standards Issue Document**, Version 1.0, May 2011
- *"A properly functioning* ***define.xml*** *file is an important part of the submission of electronic datasets and should not be considered optional. As a transition step, CDER prefers that sponsors submit both the define.pdf and define.xml formats. CDER will advise when it is ready to only receive define.xml"*
- *"Additionally, sponsors should make certain that every data variable's codelist, origin, and derivation is clearly and easily accessible from the define file. An insufficiently documented define file is a common deficiency that reviewers have noted."*

define.xml - what is it

define.xml – what is it

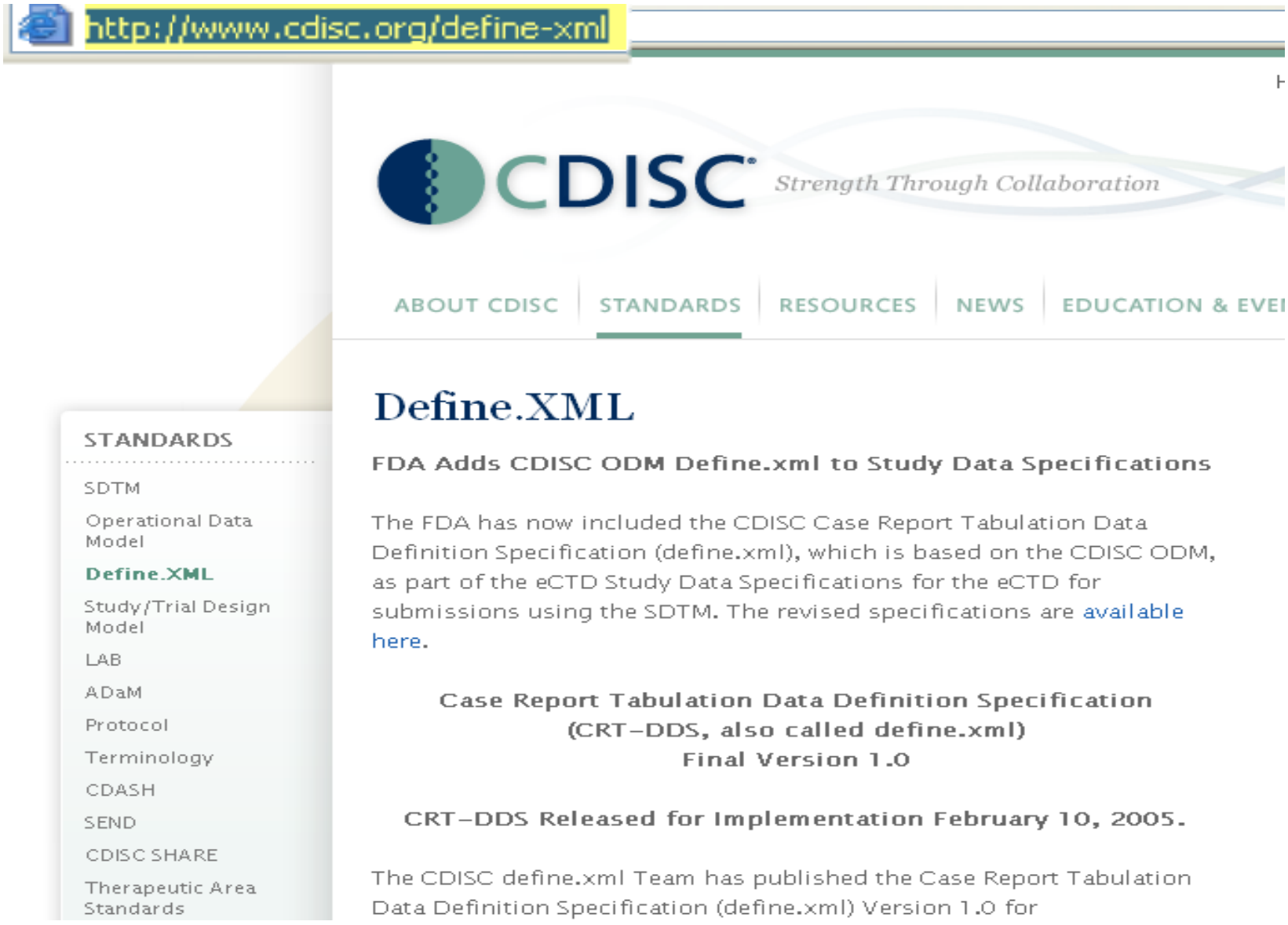
- Case Report Tabulation Data Specification (CRT-DDS, or **define.xml**): Production version: **1.0.0**
  CRT-DDS **1.0.0** is the only production version right now
- *"This specification defines the* **metadata** *structures that are to be used to describe the Case Report Tabulation datasets and variables in a manner that meets or exceeds the minimum FDA requirements."*

define.xml – what is it

- Extension of the CDISC Operational Data Model (**ODM**), an **XML** specification to facilitate the archival and interchange of the data and metadata for clinical research
- Maintained by CDISC's **XML Technologies Team**
- New define.xml version 2 in development with additional metadata support for SDTM and ADaM (based on ODM 1.3.1)

  (→ *CDISC Interchange in October*)

define.xml – what is it

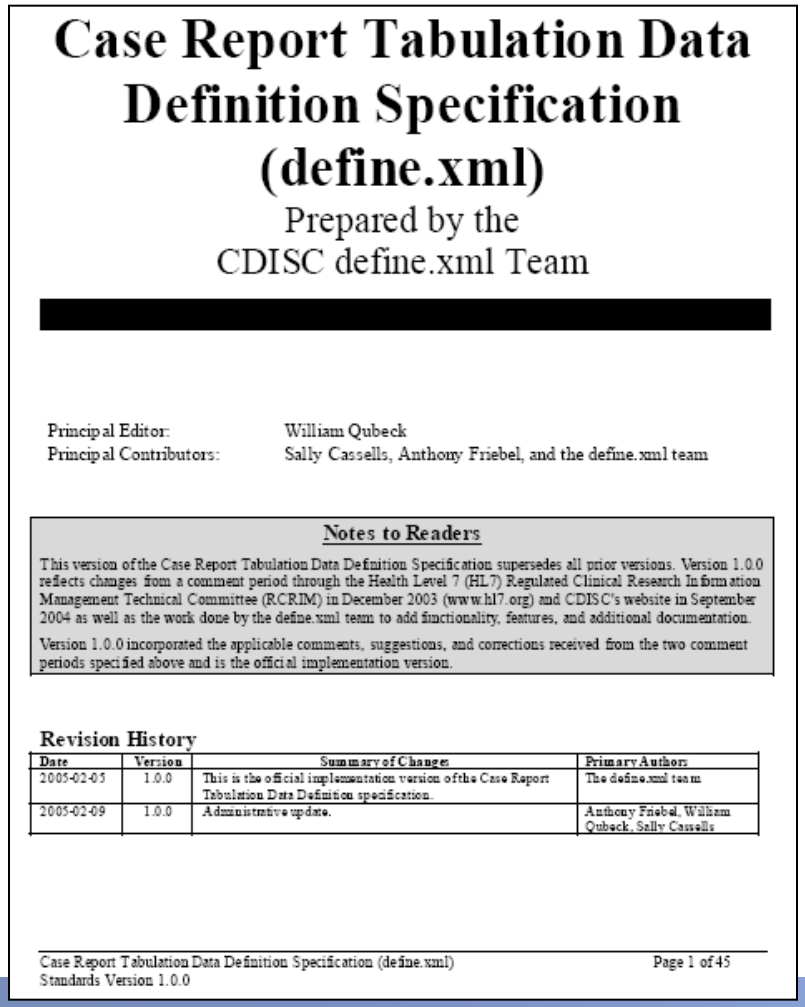


define.xml – what is it

The specifications

http://www.cdisc.org/define-xml

http://www.cdisc.org/odm



Case Report Tabulation Data Definition Specification (define.xml)

Prepared by the CDISC define.xml Team

[redacted]

Principal Editor: William Qubeck
Principal Contributors: Sally Cassells, Anthony Friebel, and the define.xml team

Notes to Readers

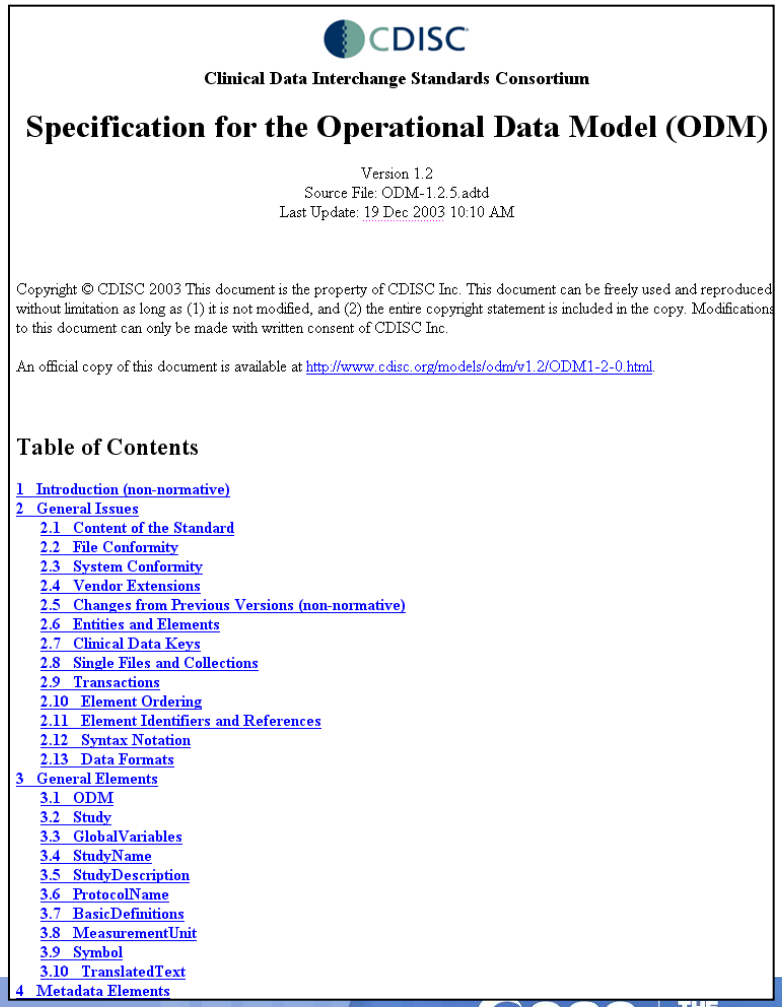
This version of the Case Report Tabulation Data Definition Specification supersedes all prior versions. Version 1.0.0 reflects changes from a comment period through the Health Level 7 (HL7) Regulated Clinical Research Information Management Technical Committee (RCRIM) in December 2003 (www.hl7.org) and CDISC's website in September 2004 as well as the work done by the define.xml team to add functionality, features, and additional documentation.

Version 1.0.0 incorporated the applicable comments, suggestions, and corrections received from the two comment periods specified above and is the official implementation version.

Revision History

Date	Version	Summary of Changes	Primary Authors
2005-02-05	1.0.0	This is the official implementation version of the Case Report Tabulation Data Definition specification.	The define.xml team
2005-02-09	1.0.0	Administrative update.	Anthony Friebel, William Qubeck, Sally Cassells

Case Report Tabulation Data Definition Specification (define.xml) Page 1 of 45
Standards Version 1.0.0



CDISC

Clinical Data Interchange Standards Consortium

Specification for the Operational Data Model (ODM)

Version 1.2
Source File: ODM-1.2.5.adtd
Last Update: 19 Dec 2003 10:10 AM

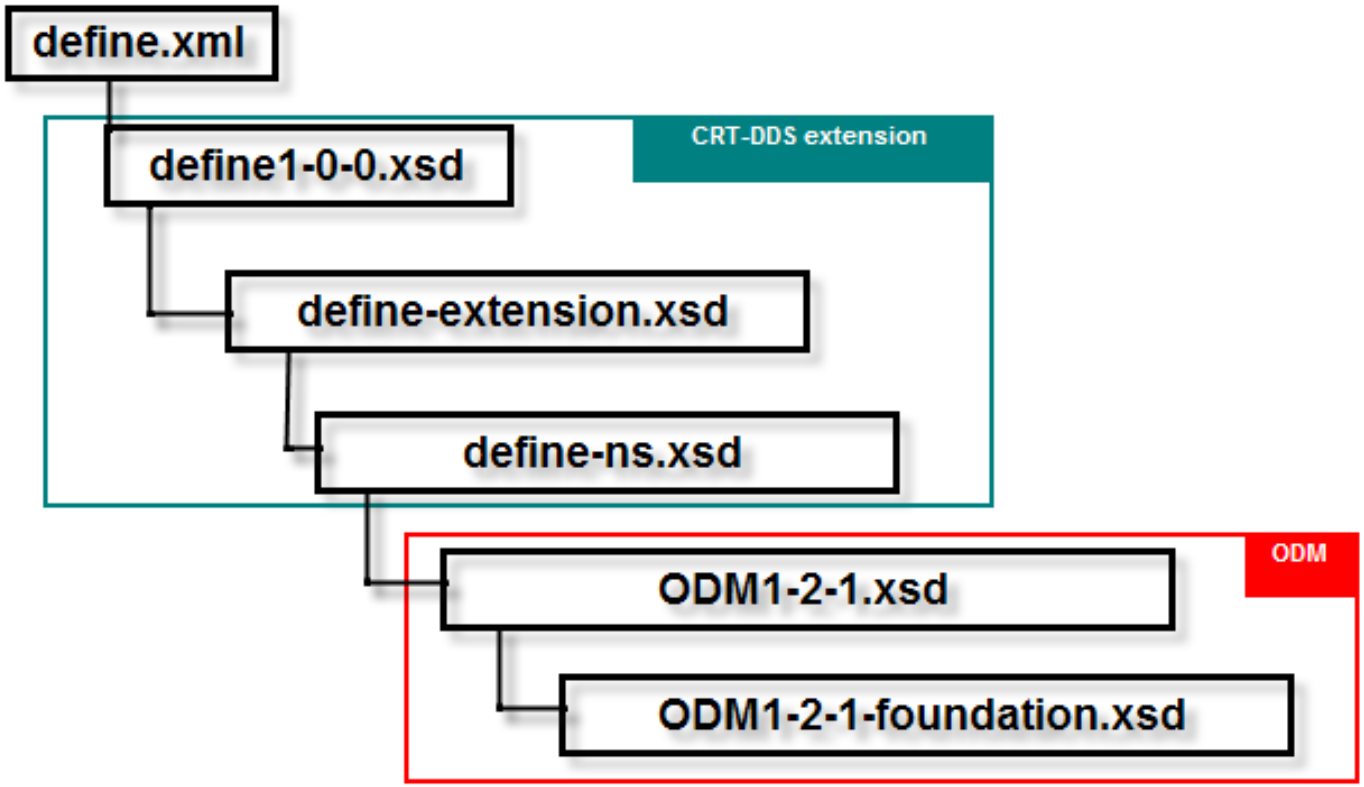
Copyright © CDISC 2003 This document is the property of CDISC Inc. This document can be freely used and reproduced without limitation as long as (1) it is not modified, and (2) the entire copyright statement is included in the copy. Modifications to this document can only be made with written consent of CDISC Inc.

An official copy of this document is available at http://www.cdisc.org/models/odm/v1.2/ODM1-2-0.html.

Table of Contents

define.xml – what is it

XML schema definitions (XSD) describe the structure of the define.xml



define.xml – what is it

http://www.cdisc.org/define-xml

Watch for the upcoming

"Metadata Submission Guidelines"

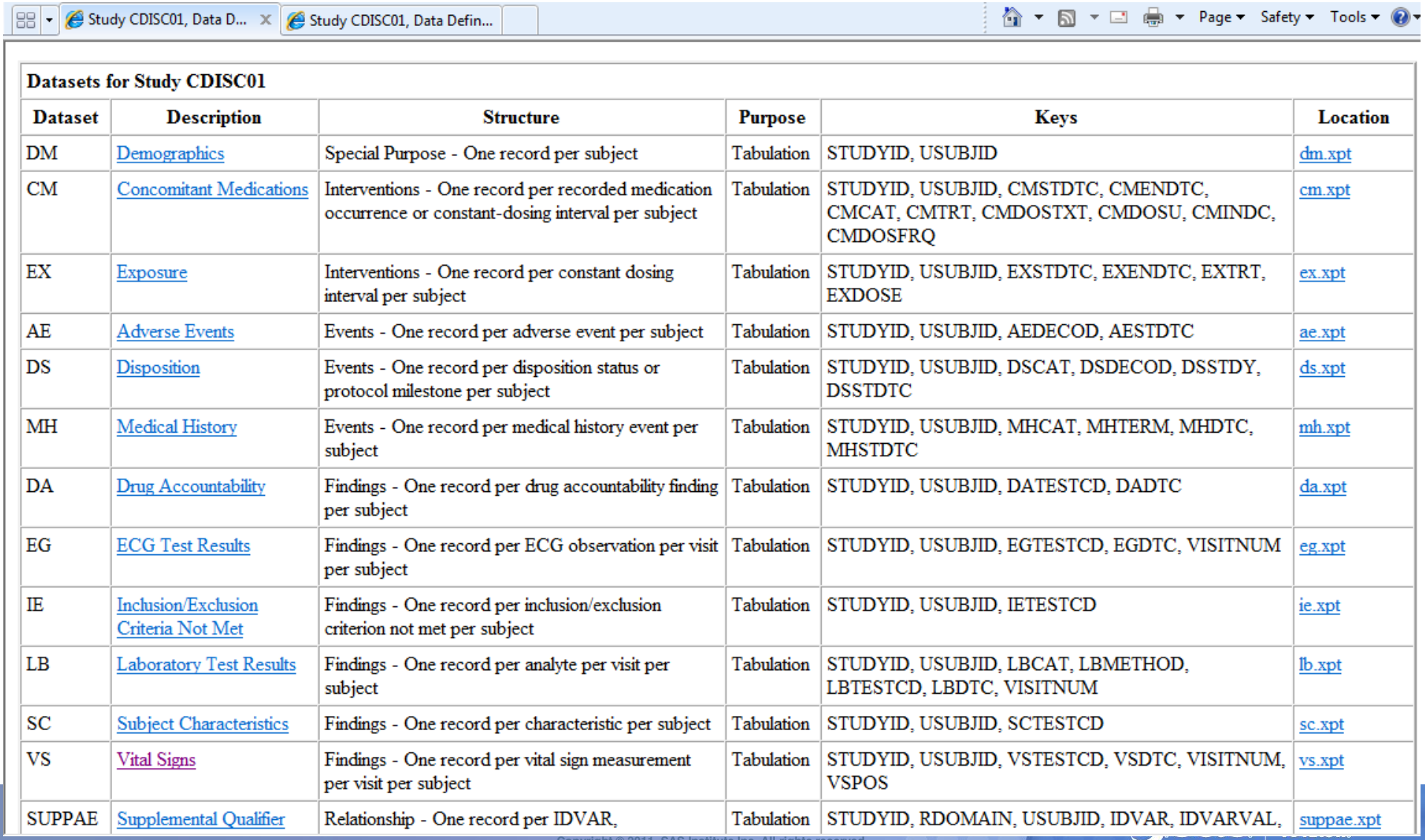
define.xml – what is it

- define.xml contains **metadata** and is **machine** readable
- define.xml becomes **human** readable with a **stylesheet**

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<?xml-stylesheet type="text/xsl" href="define1-0-0.xsl"?>
<ODM
 xmlns="http://www.cdisc.org/ns/odm/v1.2"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:xlink="http://www.w3.org/1999/xlink"
 xmlns:def="http://www.cdisc.org/ns/def/v1.0"
 xsi:schemaLocation="http://www.cdisc.org/ns/odm/v1.2 define1-0-0.xsd"
 FileOID="Study1234"
 ODMVersion="1.2"
 FileType="Snapshot"
 CreationDateTime="2004-07-28T12:34:13-06:00">
<Study OID="1234">
 <GlobalVariables>
  <StudyName>1234</StudyName>
  <StudyDescription>1234 Data Definition</StudyDescription>
  <ProtocolName>1234</ProtocolName>
 </GlobalVariables>
 <MetaDataVersion OID="CDISC.SDTM.3.1.0"
  Name="Study 1234, Data Definitions"
  Description="Study 1234, Data Definitions"
  def:DefineVersion="1.0.0"
  def:StandardName="CDISC SDTM"
  def:StandardVersion="3.1.0">
  <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>
  <def:SupplementalDoc>
   <def:DocumentRef leafID="SupplementalDataDefinitions"/>
  </def:SupplementalDoc>
  <def:leaf ID="SupplementalDataDefinitions"
   xlink:href="supplementaldatadefinitions.pdf">
   <def:title>Supplemental Data Definitions Document</def:title>
  </def:leaf>
```

define.xml – what is it

define.xml becomes **human** readable with an **XSL stylesheet**

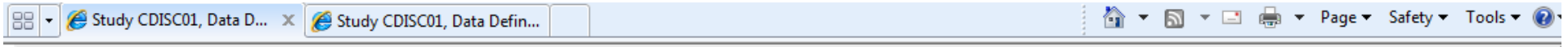


Datasets for Study CDISC01

Dataset	Description	Structure	Purpose	Keys	Location
DM	Demographics	Special Purpose - One record per subject	Tabulation	STUDYID, USUBJID	dm.xpt
CM	Concomitant Medications	Interventions - One record per recorded medication occurrence or constant-dosing interval per subject	Tabulation	STUDYID, USUBJID, CMSTDTC, CMENDTC, CMCAT, CMTRT, CMDOSTXT, CMDOSU, CMINDC, CMDOSFRQ	cm.xpt
EX	Exposure	Interventions - One record per constant dosing interval per subject	Tabulation	STUDYID, USUBJID, EXSTDTC, EXENDTC, EXTRT, EXDOSE	ex.xpt
AE	Adverse Events	Events - One record per adverse event per subject	Tabulation	STUDYID, USUBJID, AEDECOD, AESTDTC	ae.xpt
DS	Disposition	Events - One record per disposition status or protocol milestone per subject	Tabulation	STUDYID, USUBJID, DSCAT, DSDECOD, DSSTDY, DSSTDTC	ds.xpt
MH	Medical History	Events - One record per medical history event per subject	Tabulation	STUDYID, USUBJID, MHCAT, MHTERM, MHDTC, MHSTDTC	mh.xpt
DA	Drug Accountability	Findings - One record per drug accountability finding per subject	Tabulation	STUDYID, USUBJID, DATESTCD, DADTC	da.xpt
EG	ECG Test Results	Findings - One record per ECG observation per visit per subject	Tabulation	STUDYID, USUBJID, EGTESTCD, EGDTC, VISITNUM	eg.xpt
IE	Inclusion/Exclusion Criteria Not Met	Findings - One record per inclusion/exclusion criterion not met per subject	Tabulation	STUDYID, USUBJID, IETESTCD	ie.xpt
LB	Laboratory Test Results	Findings - One record per analyte per visit per subject	Tabulation	STUDYID, USUBJID, LBCAT, LBMETHOD, LBTESTCD, LBDTC, VISITNUM	lb.xpt
SC	Subject Characteristics	Findings - One record per characteristic per subject	Tabulation	STUDYID, USUBJID, SCTESTCD	sc.xpt
VS	Vital Signs	Findings - One record per vital sign measurement per visit per subject	Tabulation	STUDYID, USUBJID, VSTESTCD, VSDTC, VISITNUM, VSPOS	vs.xpt
SUPPAE	Supplemental Qualifier	Relationship - One record per IDVAR,	Tabulation	STUDYID, RDOMAIN, USUBJID, IDVAR, IDVARVAL,	suppae.xpt

define.xml – what is it

define.xml becomes **human** readable with an **XSL stylesheet**

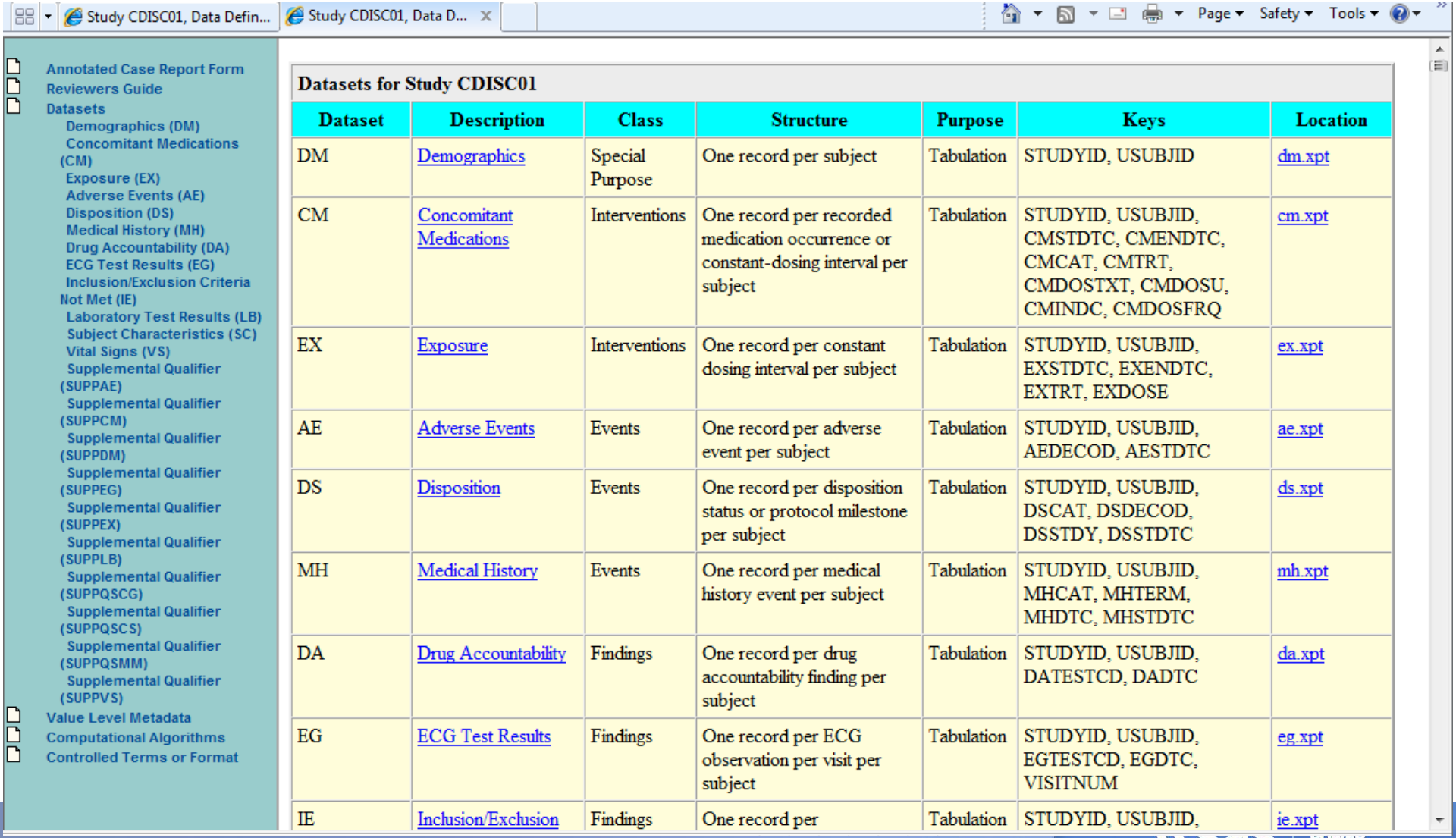


Vital Signs Dataset (VS)

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
VSSEQ	Sequence Number	integer		Derived	IDENTIFIER	Sequential number identifying records within each USUBJID and generated using the key sequence of the domain.
VSTESTCD	Vital Signs Test Short Name	text		Assigned	TOPIC	
VSTEST	Vital Signs Test Name	text		CRF Page 11	SYNONYM QUALIFIER	
VSPOS	Vital Signs Position of Subject	text		CRF Page 11	RECORD QUALIFIER	
VSORRES	Result or Finding in Original Units	text		CRF Page 11	RESULT QUALIFIER	
VSORRESU	Original Units	text	WUNIT	CRF Page 11	VARIABLE QUALIFIER	
VSSTRESC	Character Result/Finding in Std Format	text		Derived	RESULT QUALIFIER	Data collected in non-standard units (i.e. lbs, inches) is converted using standard conversion factors to standard units.
VSSTRESN	Numeric Result/Finding in Standard Units	float		Derived	RESULT QUALIFIER	When VSSTRESC contains numeric data the numeric equivalent is derived in VSSTRESN
VSSTRESU	Standard Units	text	VSRESU	Derived	VARIABLE QUALIFIER	Standard units consistent with CDISC controlled terminology
VSBLFL	Baseline Flag	text	NY	Derived	RECORD QUALIFIER	Safety subjects only: VSBLFL=Y for last non missing record on or before the first dose date (RFSTDTC)

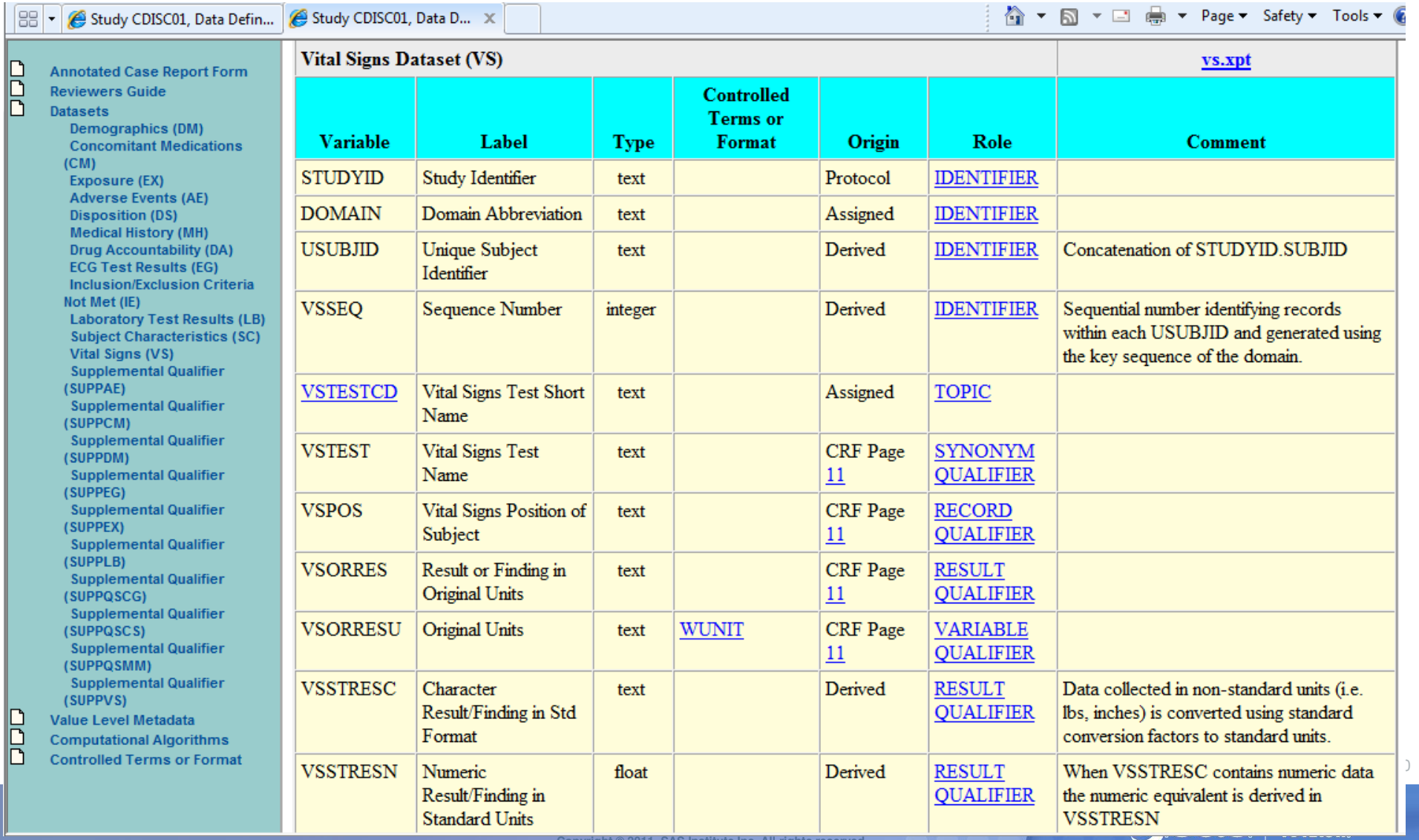
define.xml – what is it

… and looks even fancier with a different **stylesheet**



define.xml – what is it

… and looks even fancier with a different **stylesheet**



Study CDISC01, Data Defin... | Study CDISC01, Data D... | Page | Safety | Tools

- Annotated Case Report Form
- Reviewers Guide
- Datasets
  - Demographics (DM)
  - Concomitant Medications (CM)
  - Exposure (EX)
  - Adverse Events (AE)
  - Disposition (DS)
  - Medical History (MH)
  - Drug Accountability (DA)
  - ECG Test Results (EG)
  - Inclusion/Exclusion Criteria Not Met (IE)
  - Laboratory Test Results (LB)
  - Subject Characteristics (SC)
  - Vital Signs (VS)
  - Supplemental Qualifier (SUPPAE)
  - Supplemental Qualifier (SUPPCM)
  - Supplemental Qualifier (SUPPDM)
  - Supplemental Qualifier (SUPPEG)
  - Supplemental Qualifier (SUPPEX)
  - Supplemental Qualifier (SUPPLB)
  - Supplemental Qualifier (SUPPQSCG)
  - Supplemental Qualifier (SUPPQSCS)
  - Supplemental Qualifier (SUPPQSMM)
  - Supplemental Qualifier (SUPPVS)
- Value Level Metadata
- Computational Algorithms
- Controlled Terms or Format

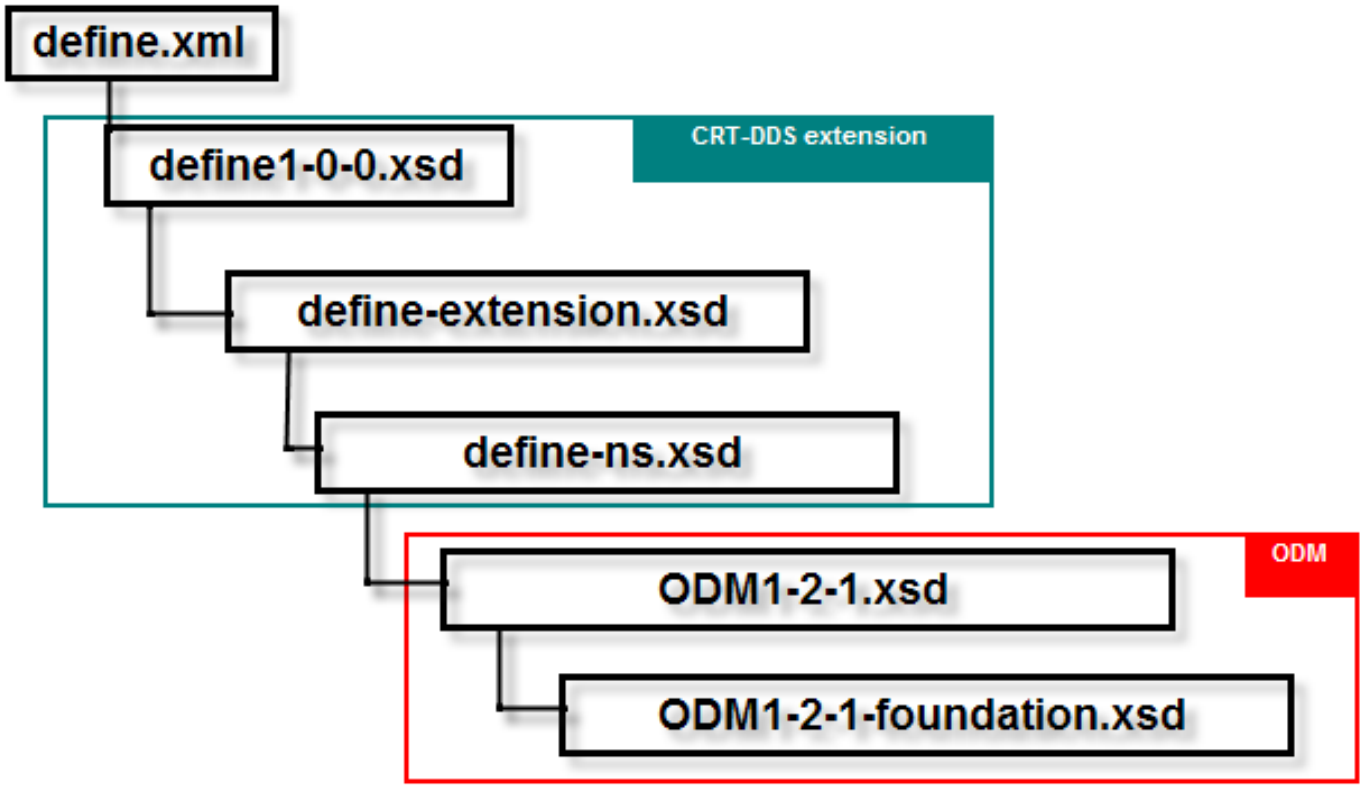
Vital Signs Dataset (VS) — vs.xpt

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
VSSEQ	Sequence Number	integer		Derived	IDENTIFIER	Sequential number identifying records within each USUBJID and generated using the key sequence of the domain.
VSTESTCD	Vital Signs Test Short Name	text		Assigned	TOPIC	
VSTEST	Vital Signs Test Name	text		CRF Page 11	SYNONYM QUALIFIER	
VSPOS	Vital Signs Position of Subject	text		CRF Page 11	RECORD QUALIFIER	
VSORRES	Result or Finding in Original Units	text		CRF Page 11	RESULT QUALIFIER	
VSORRESU	Original Units	text	WUNIT	CRF Page 11	VARIABLE QUALIFIER	
VSSTRESC	Character Result/Finding in Std Format	text		Derived	RESULT QUALIFIER	Data collected in non-standard units (i.e. lbs, inches) is converted using standard conversion factors to standard units.
VSSTRESN	Numeric Result/Finding in Standard Units	float		Derived	RESULT QUALIFIER	When VSSTRESC contains numeric data the numeric equivalent is derived in VSSTRESN

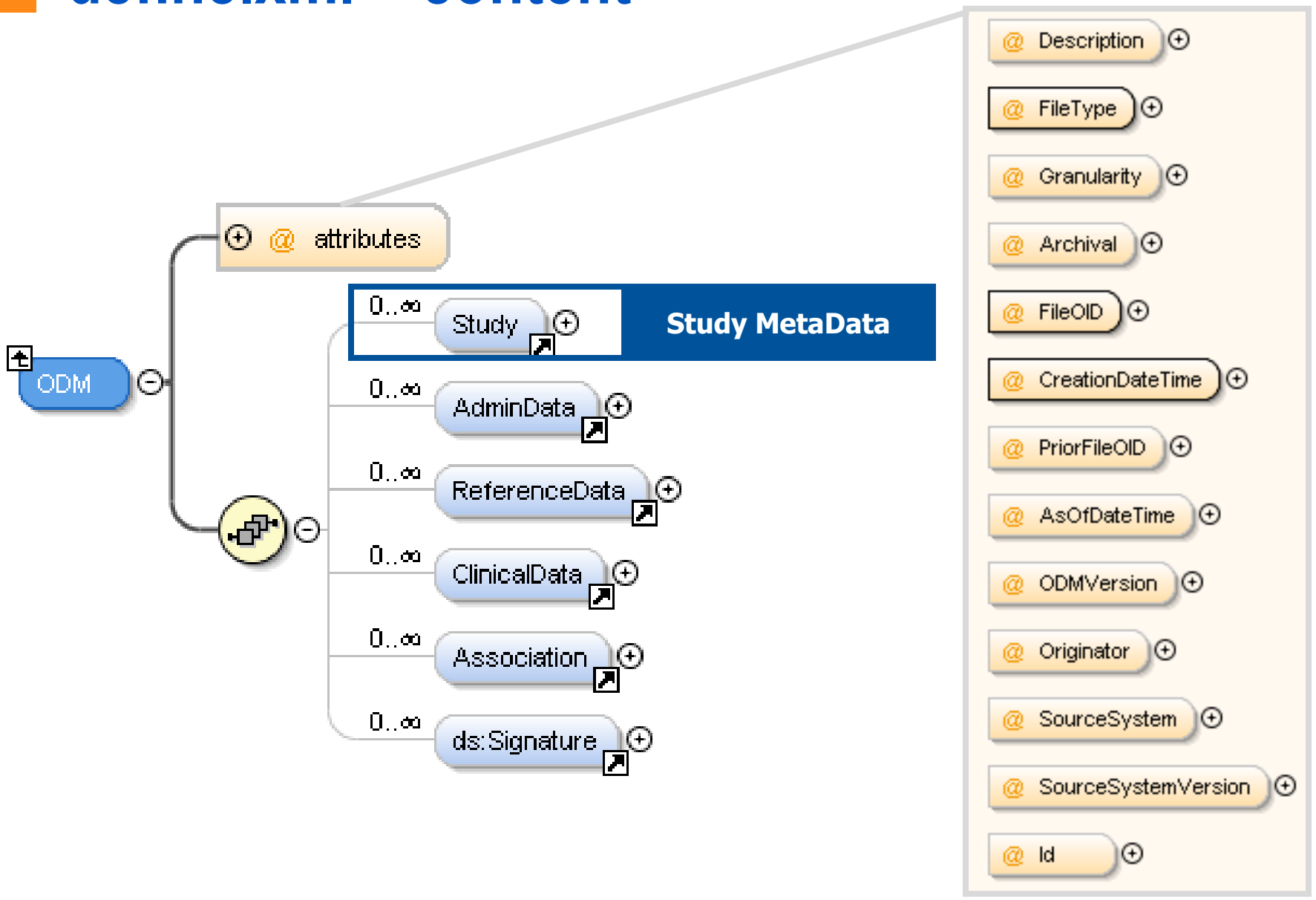
define.xml - content

define.xml – content

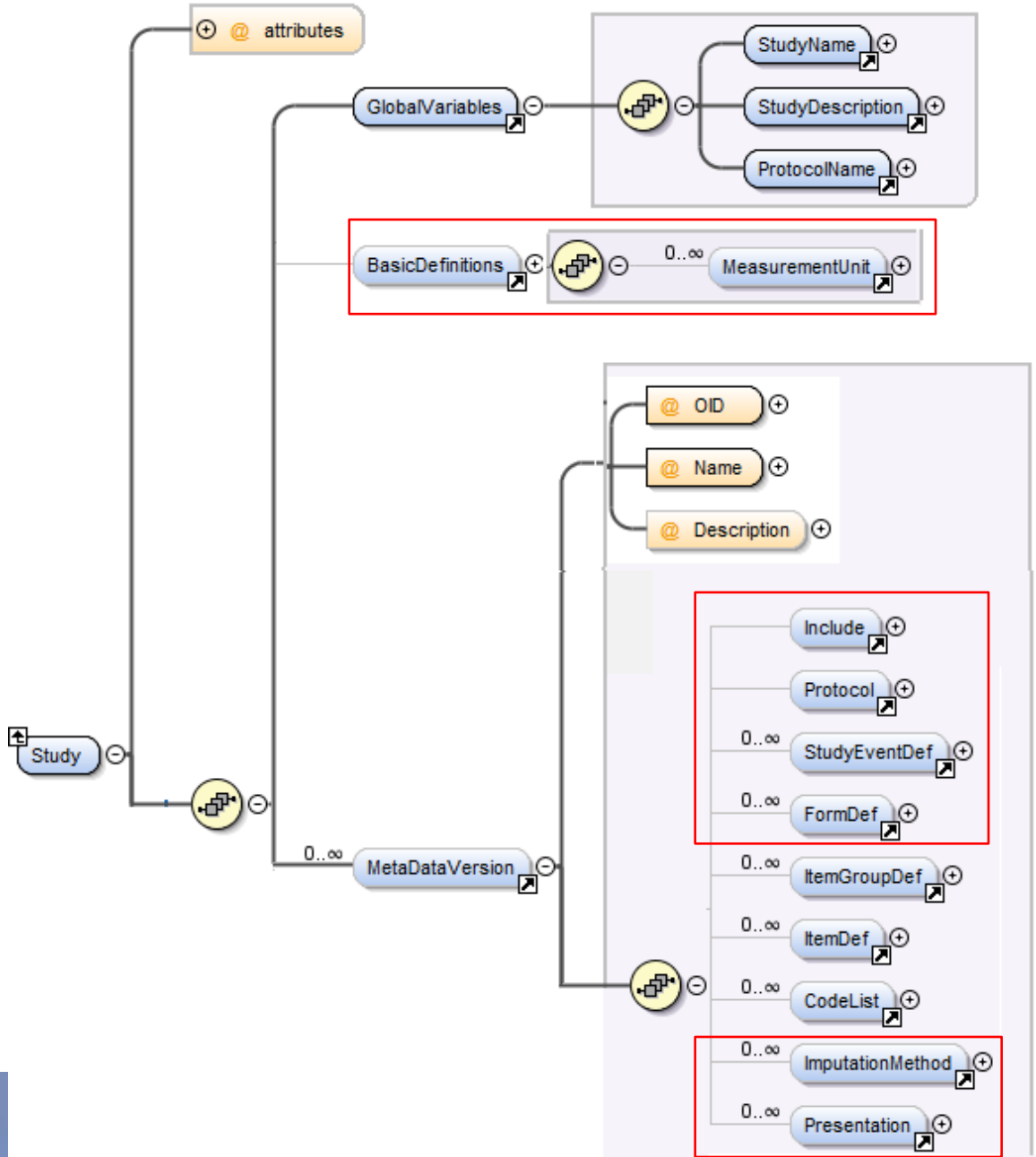
define.xml schema adds elements and attributes to the ODM schema



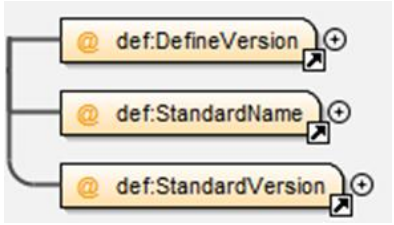
define.xml – content

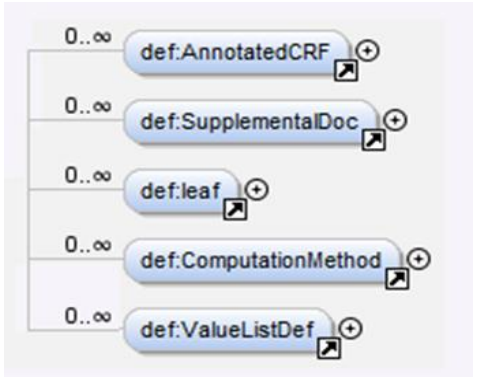


define.xml – content

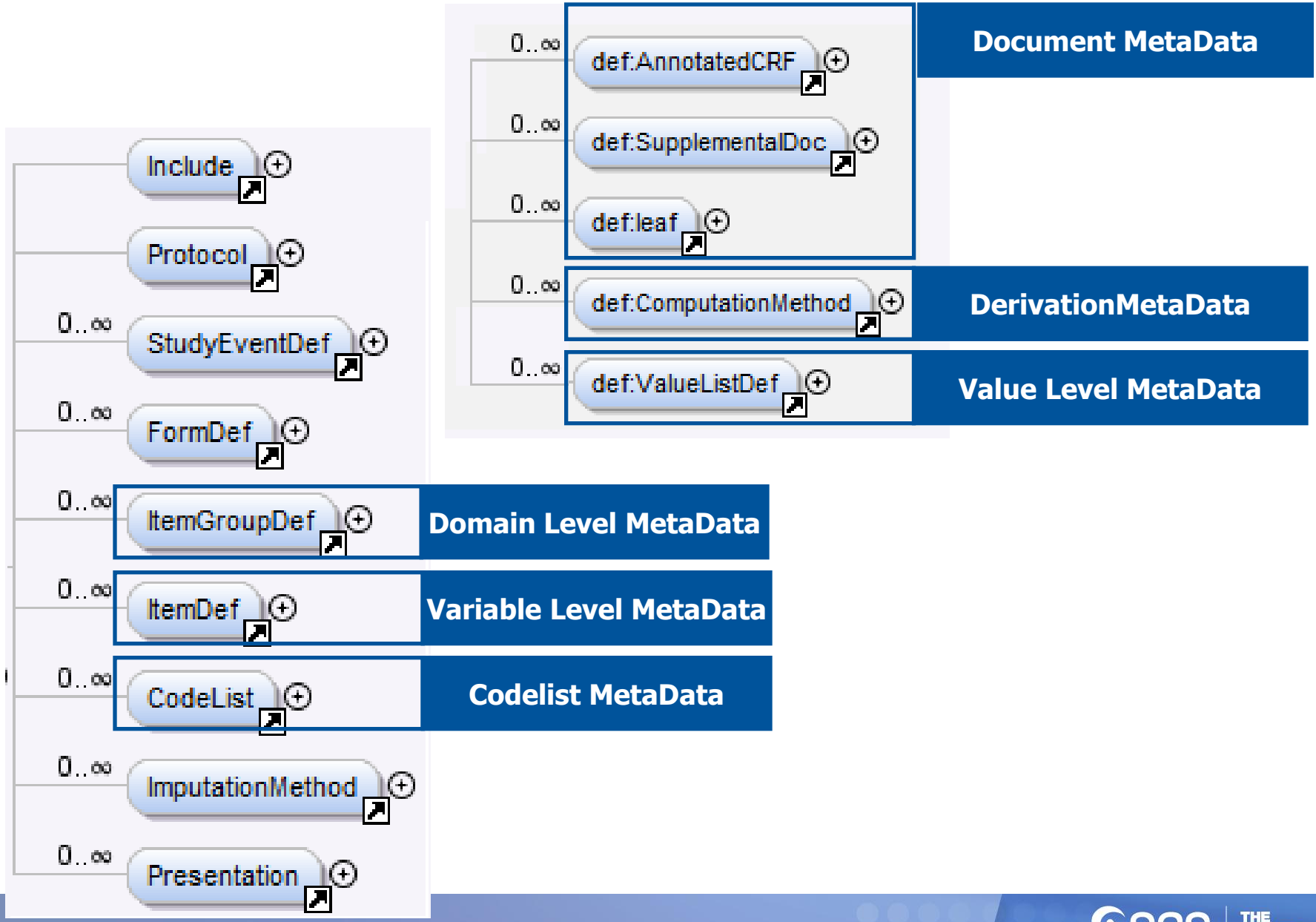


define.xml adds

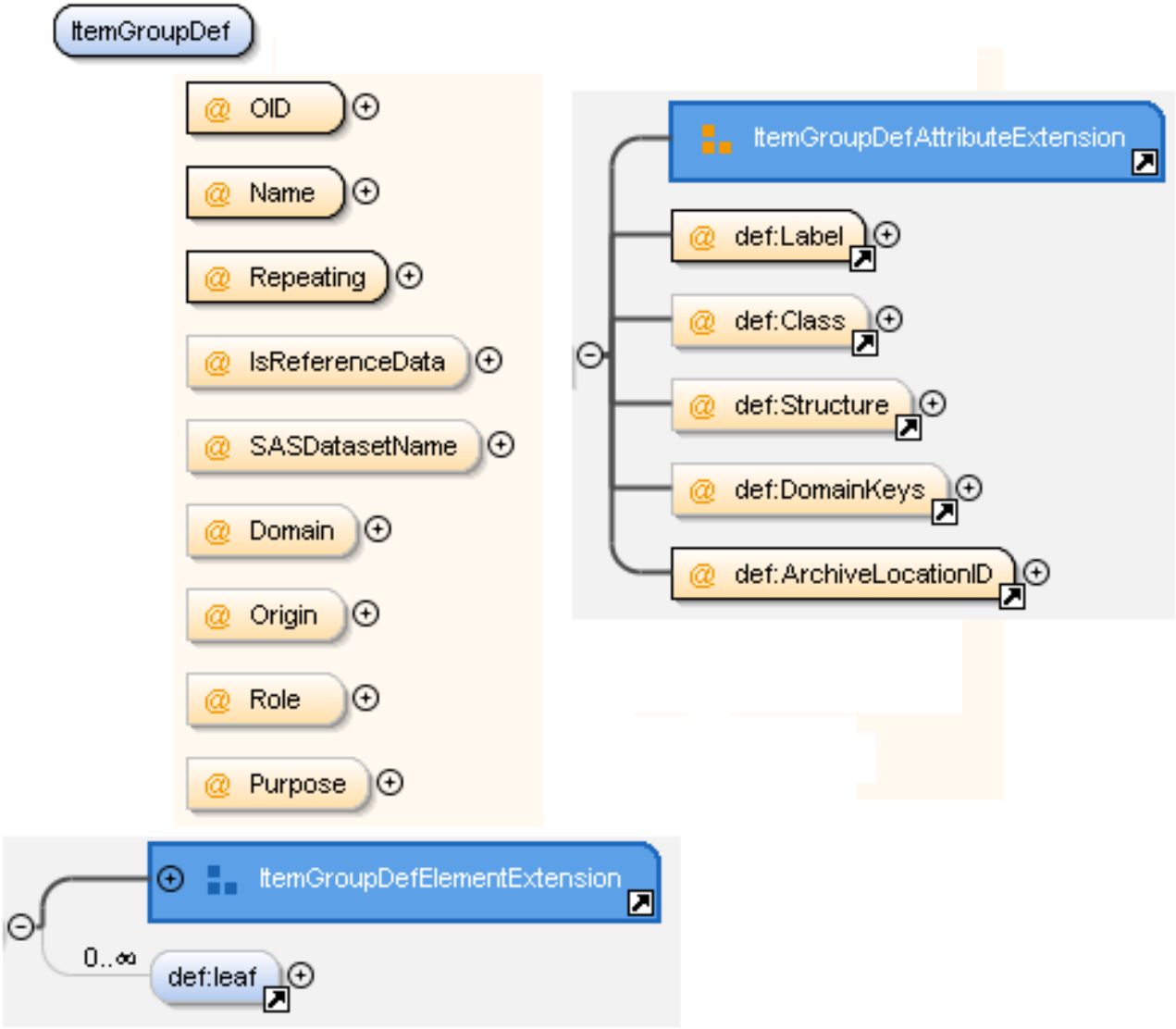




define.xml – MetadataVersion elements



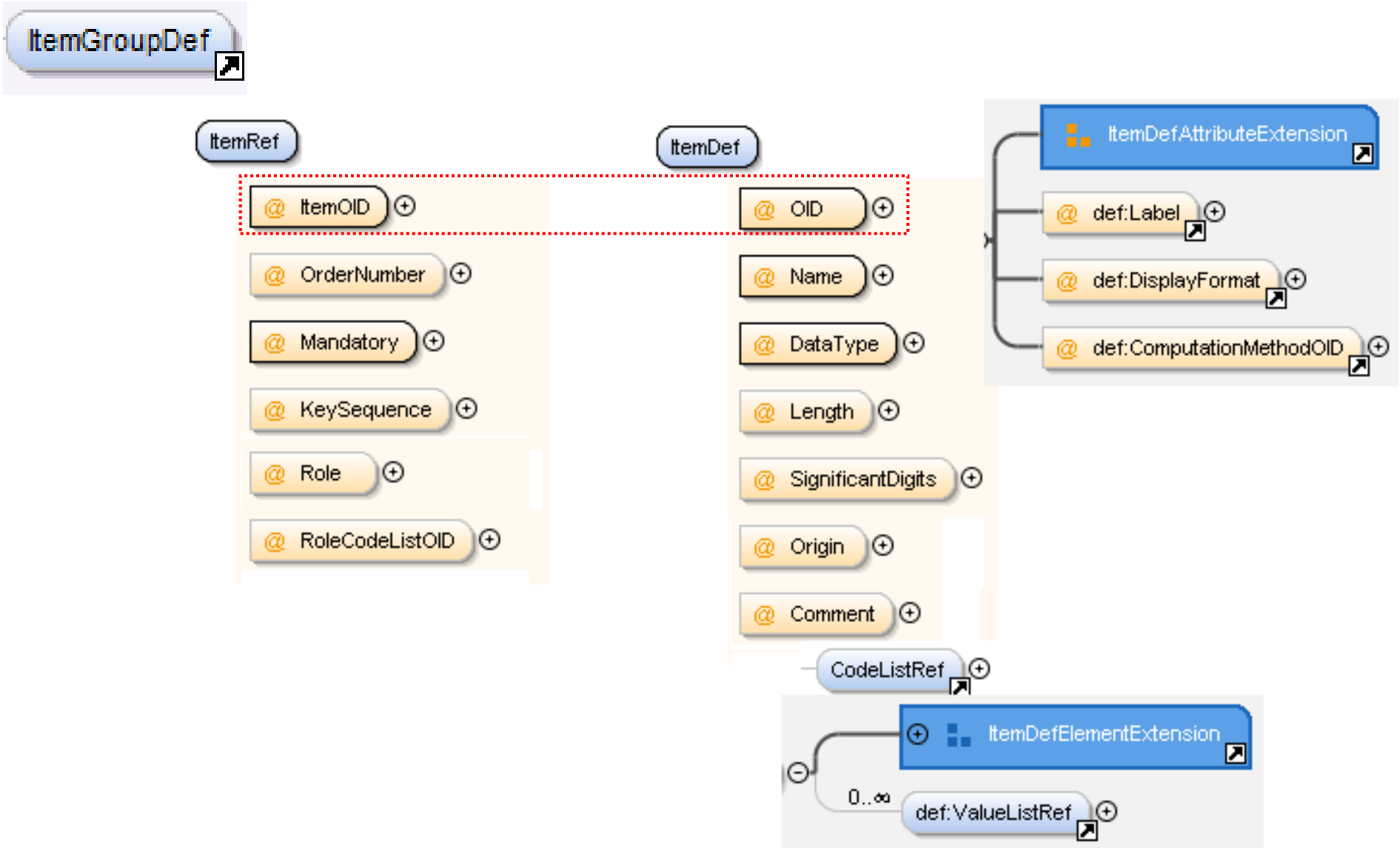
define.xml – Domain level metadata



define.xml – Domain level metadata

```
<ItemGroupDef
      OID="IG.MH"  Name="MH" Repeating="Yes"
      IsReferenceData="No"
      SASDatasetName="MH"
      Purpose="Tabulation"
      def:Label="Medical History"
      def:Structure="One record per medical history event per subject"
      def:DomainKeys="STUDYID, USUBJID, MHCAT, MHTERM, MHSTDTC"
      def:Class="EVENTS"
      def:ArchiveLocationID="Location.MH">

  <ItemRef ItemOID="I.STUDYID"
           OrderNumber="1" Mandatory="Yes" Role="IDENTIFIER" />
  <ItemRef ItemOID="I.DOMAIN"
           OrderNumber="2" Mandatory="Yes" Role="IDENTIFIER" />
  <ItemRef ItemOID="I.USUBJID"
           OrderNumber="3" Mandatory="Yes" Role="IDENTIFIER" />
  ...
  <ItemRef ItemOID="I.MHTERM"
           OrderNumber="8" Mandatory="Yes" Role="TOPIC" />
  ...

  <def:leaf ID="Location.MH" xlink:href="mh.xpt">
           <def:title>mh.xpt</def:title>
  </def:leaf>
</ItemGroupDef>
```

define.xml – Variable level metadata



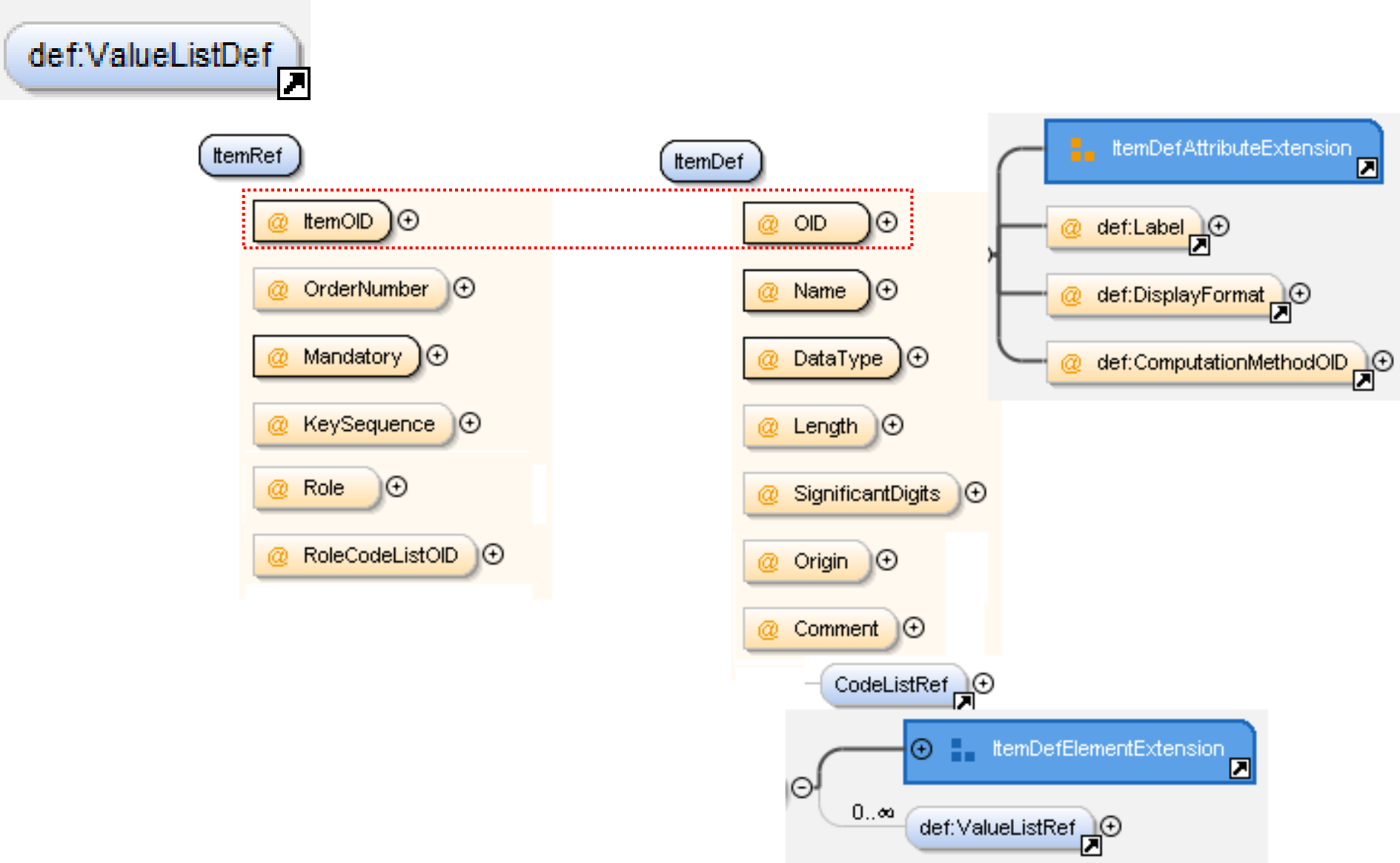
define.xml – Variable level metadata

```
<ItemRef ItemOID="I.VSTESTCD" OrderNumber="6" Mandatory="Yes"
      Role="TOPIC" RoleCodeListOID="RoleCodeList" />

<ItemRef ItemOID="I.VSBLFL" OrderNumber="14" Mandatory="No"
      Role="RECORD QUALIFIER" RoleCodeListOID="RoleCodeList" />
```

Watch for CRT-DDS V2 !

```
<ItemDef OID="I.VSTESTCD"
   Name="VSTESTCD" DataType="text"
   Length="8" Origin="CRF Page 9"
   def:Label="Vital Signs Test Short Name">
      <def:ValueListRef ValueListOID="VL.VSTESTCD" />
</ItemDef>
  ...

<ItemDef OID="I.VSBLFL" Name="VSBLFL" DataType="text" Length="1"
   Origin="DERIVED" Comment="Derivation of baseline: Safety patients only: VSBLFL=Y
   for last non missing record on or before the first dose date (RFSTDTC)."
   def:Label="Baseline Flag"
   def:ComputationMethodOID="CM.VSBLFL">
  <CodeListRef CodeListOID="CL.NY" />
</ItemDef>
```

define.xml – Value level metadata



define.xml – Value level metadata

Watch for CRT-DDS V2 !

```
<ItemDef OID="I.VSTESTCD"
    Name="VSTESTCD" DataType="text"
    Length="8" Origin="CRF Page 9"
    def:Label="Vital Signs Test Short Name">
      <def:ValueListRef ValueListOID="VL.VSTESTCD" />
</ItemDef>
```

```
<def:ValueListDef OID="VL.VSTESTCD">
    <ItemRef ItemOID="I.VSTESTCD.HEIGHT" OrderNumber="1" Mandatory="No" />
    <ItemRef ItemOID="I.VSTESTCD.WEIGHT" OrderNumber="2" Mandatory="No" />
    ...
</def:ValueListDef>
```

```
<ItemDef OID="I.VSTESTCD.HEIGHT"
    Name="HEIGHT" DataType="integer" Length="3"
    Origin="CRF Page 18" def:Label="Height" />

<ItemDef OID="I.VSTESTCD.WEIGHT"
    Name="WEIGHT" DataType="float" SignificantDigits="1" Length="5"
    Origin="CRF Page 18" def:Label="Weight" />
```

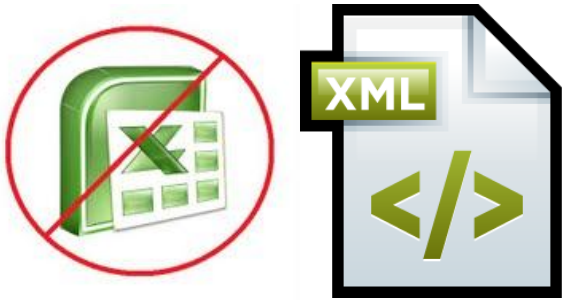
define.xml – Codelist metadata

Watch for CRT-DDS V2 !

```
<CodeList OID="ACN" Name="ACN" DataType="text">
  <CodeListItem CodedValue="DOSE NOT CHANGED">
    <Decode>
      <TranslatedText>DOSE NOT CHANGED</TranslatedText>
    </Decode>
  </CodeListItem>
```

```
<CodeList OID="CL.AEDICT " Name="ADVERSE EVENT DICTIONARY" DataType="text">
  <ExternalCodeList Dictionary="MEDDRA" Version="8.0" />
</CodeList>
```

CDISC Controlled Terms now downloadable in ODM XML !



define.xml – Derivation metadata

```
<def:ComputationMethod OID="CM.VSNRIND">
 if .&lt; VSSTRESN &lt; VSSTNRLO
      then VSNRIND=&apos;LOW&apos;;
      else if  VSSTRESN &gt; VSSTNRHI
            then VSNRIND=&apos;HIGH&apos;;
            else if VSSTNRLO&lt;=VSSTRESN&lt;=VSSTNRHI
                  then VSSTRESN=&apos;NORMAL&apos;;
</def:ComputationMethod>
```

Computational Algorithms (CM.VSNRIND)

Reference Name	Computation Method
CM.VSNRIND	if .< VSSTRESN < VSSTNRLO then VSNRIND='LOW'; else if VSSTRESN > VSSTNRHI then VSNRIND='HIGH'; else if VSSTNRLO<=VSSTRESN<=VSSTNRHI then VSSTRESN='NORMAL';

define.xml – Document metadata

```
<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>

<def:SupplementalDoc>
        <def:DocumentRef leafID="SupplementalDataDefinitions" />
</def:SupplementalDoc>

<def:leaf ID="SupplementalDataDefinitions"
      xlink:href="supplementaldatadefinitions.pdf">
      <def:title>Supplemental Data Definitions Document</def:title>
</def:leaf>
```

define.xml - data model

define.xml – data model

- How will you be maintaining all of this metadata?
- Traditionally: Excel spreadsheets
- Problems:
  - Version control, auditing, access control, data quality, impact analysis, scalability ……
- … Excel is no database or metadata registry
- Excel spreadsheets can multiply fast

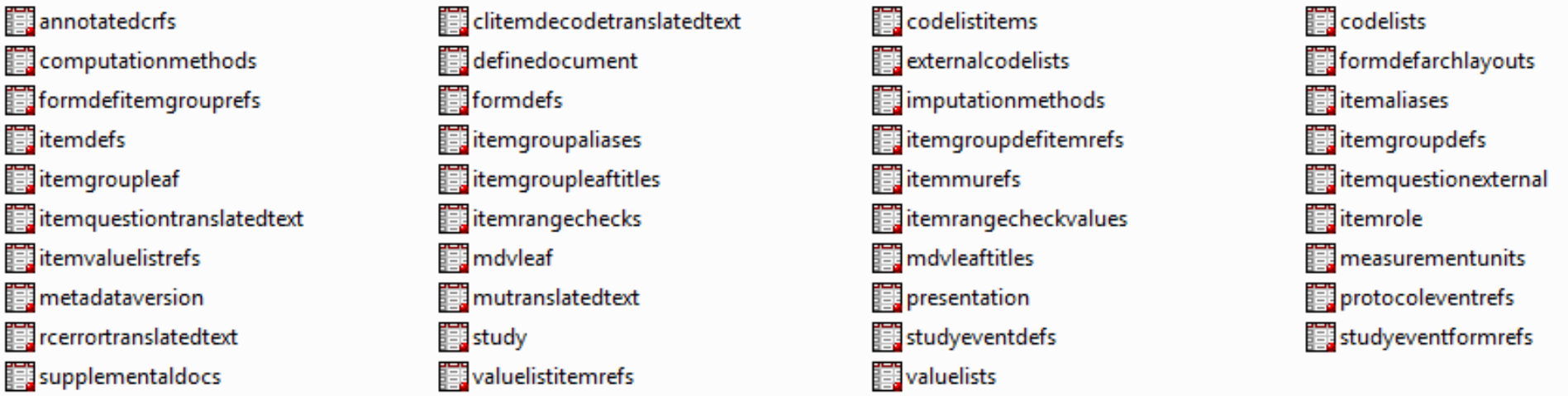
define.xml – data model

- define.xml has a deep hierarchy
- define.xml contains many relations

```
<ItemGroupDef OID="MH"
      Name="MH" Repeating="Yes" IsReferenceData="No"
      Purpose="Tabulation" def:Label="Medical History"
      def:Structure="One record per medical history event per subject"
      def:DomainKeys="STUDYID, USUBJID, MHCAT, MHTERM, MHSTDTC"
      def:Class="EVENTS" def:ArchiveLocationID="Location.MH">
  <ItemRef ItemOID="STUDYID" OrderNumber="1" Mandatory="Yes"
     Role="IDENTIFIER" RoleCodeListOID="RoleCodeList" />
...

   <def:leaf ID="Location.MH" xlink:href="mh.xpt">
            <def:title>mh.xpt</def:title>
   </def:leaf>
</ItemGroupDef>

...

<ItemDef OID="STUDYID" Name="STUDYID" DataType="text" Length="7"
         Origin="CRF Page 3" def:Label="Study Identifier" />
```

define.xml – data model

- SAS Clinical Standards Toolkit has a data model that represents the define.xml in 39 SAS data sets
- 20 of these *typically* used for define.xml
- Patterned to match the XML element and attribute structure of the define.xml file
- XML element → table
  XML attribute → column



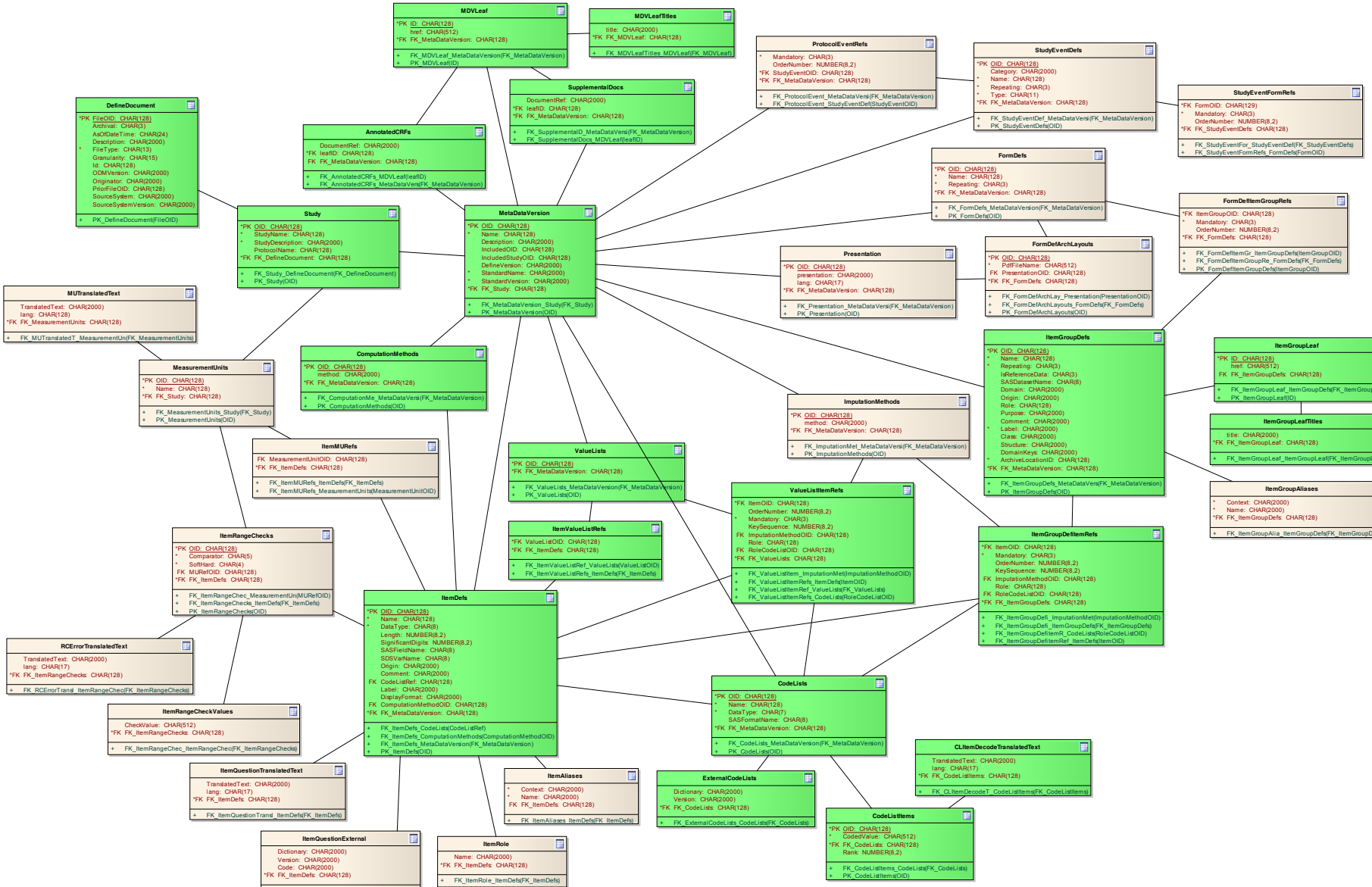
define.xml – data model

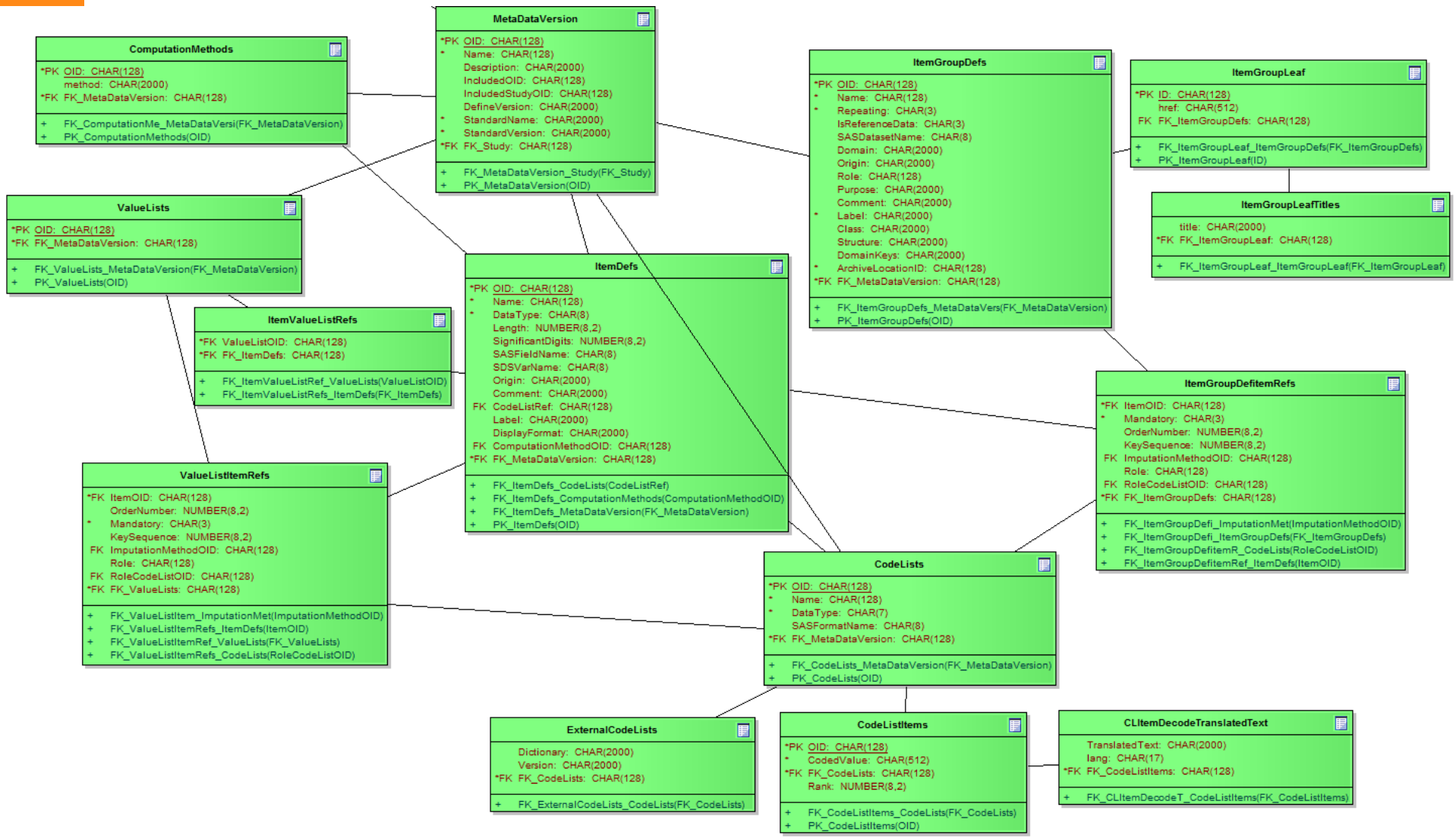
```
<ItemDef OID="COL10" Name="BRTHDTC" DataType="text" Length="64"
         SASFieldName="BRTHDTC" def:Label="Date/Time of Birth"/>
<ItemDef OID="COL11" Name="AGE" DataType="float"
          Length="8" SASFieldName="AGE" def:Label="Age"/>
<ItemDef OID="COL12" Name="AGEU" DataType="text"
         Length="10" SASFieldName="AGEU" def:Label="Age Units">
         <CodeListRef CodeListOID="AGEU"/>
</ItemDef>
```

VIEWTABLE: _svwork.Itemdefs

	OID	Name	DataType	Length	CodeListRef	label
10	COL10	BRTHDTC	text	64		Date/Time of Birth
11	COL11	AGE	float	8		Age
12	COL12	AGEU	text	10	AGEU	Age Units

define.xml – data model

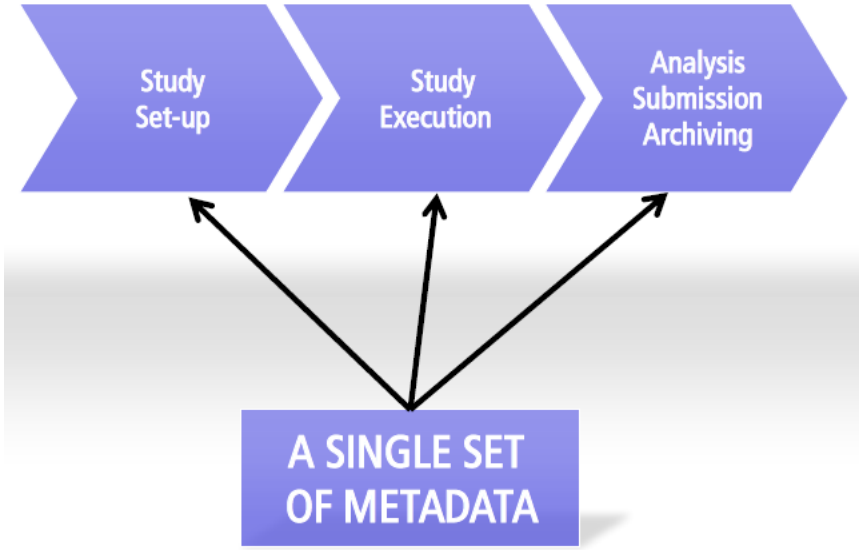




define.xml - end-to-end

define.xml – end-to-end

- Common practice: define.xml being created based on the SAS submission dataset
- Think of the potential when this metadata is part of a single set of metadata throughout the process
- Metadata can drive the process
- define.xml is then just the publishing of metadata



Picture courtesy of Philippe Verplancke

Questions

SAS Institute Inc.
World Headquarters
SAS Campus Drive
Cary, NC 27513 USA
Tel: +1 919 677 8000
Fax: +1 919 677 4444
www.sas.com

Lex Jansen
Software Developer
Health and Life Sciences
Tel: +1 919 531 9860 Ext 19860

E-mail: Lex.Jansen@sas.com

www.sas.com