

Moving your ADaM Metadata into Define 2.0

Geoffrey Mann and the ADaM
Metadata Team

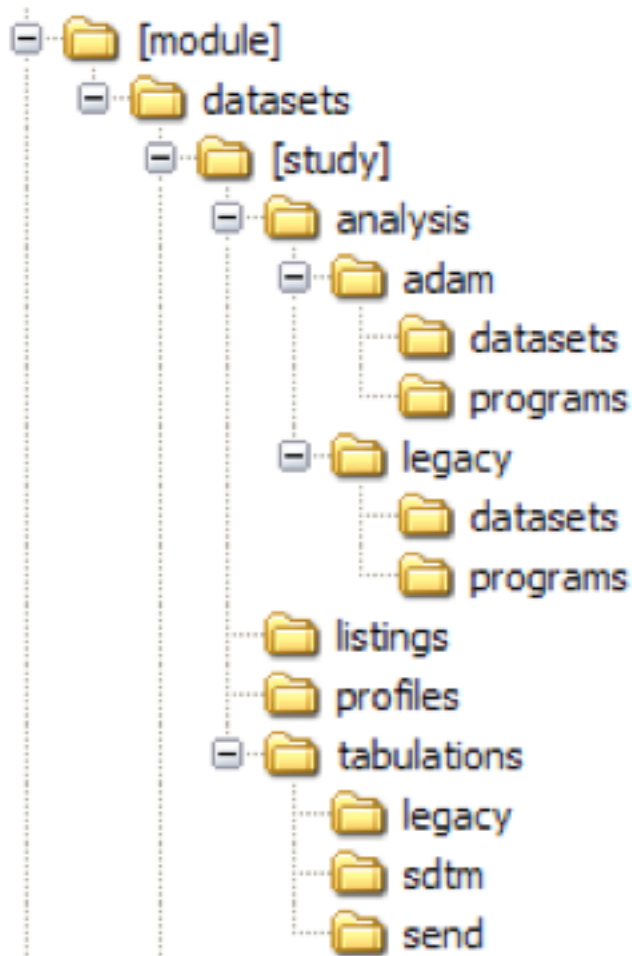
ADaM Metadata Team Members

- Monika Kawohl
- Jeff Abolafia
- Lin Yan
- Kim Minkalis
- Randall Austin
- Terek Peterson
- Stephen Hamburg
- Geoffrey Mann
- Jian Chen
- Karin Lapann
- Helieh Ghandehari
- Ken Stoltfus
- Dana Soloff
- Steven Hege
- Mike Willis
- **Lex Jansen (XML Tech)**
- **Sally cassells (XML Tech)**

ADaM Metadata Submission Components

- eCTD Module 5 Structure
- Submission datasets
- Submission programs
- Submission results
- **define.xml**
- **Analysis Results Metadata Schema Extension**
- **Standard Stylesheet**
- Optional Reviewers' Guide

ADaM Metadata Submission Guidelines Document



- Guidance for compiling the eCTD module 5 “analysis” folder.
- Create example submission components for ADaM data and metadata.
- Illustrate acceptable practices and formats that sponsors should incorporate into their submissions.

Dataset Metadata

ADaM-MSG

Analysis Data Model Document

Table 5.1.1 Analysis Dataset Metadata Fields

Analysis Dataset Metadata Field	Description
DATASET NAME	The file name of the dataset, hyperlinked to the corresponding analysis dataset variable descriptions (i.e., the data definition table) within the define file.
DATASET DESCRIPTION	A short descriptive summary of the contents of the dataset
DATASET LOCATION	The folder and filename where the dataset can be found, ideally hyperlinked to the actual dataset (i.e., XPT file)
DATASET STRUCTURE	The level of detail represented by individual records in the dataset (e.g., “One record per subject,” “One record per subject per visit,” “One record per subject per event”).
KEY VARIABLES OF DATASET	A list of variable names that parallels the structure, ideally uniquely identifies and indexes each record in the dataset.
CLASS OF DATASET	Identification of the general class of the dataset using the name of the ADaM structure (i.e., “ADSL,” “BDS”) or “OTHER” if not an ADaM-specified structure
DOCUMENTATION	Description of the source data, processing steps, and analysis decisions pertaining to the creation of the dataset. Software code of various levels of functionality and complexity, such as pseudo-code or actual code fragments may be provided. Links or references to external documents (e.g., protocol, statistical analysis plan, software code) may be used.

Table 5.1.1.1 Analysis Dataset Metadata

Table 5.1.1.1 Analysis Dataset Metadata for the ADQSADAS Analysis Dataset²

Dataset Name	Dataset Description	Dataset Location	Dataset Structure	Key Variables of Dataset	Class of Dataset	Documentation
<i>filename of the dataset.</i>	<i>short summary of contents of the dataset</i>	<i>where the dataset can be found</i>	<i>the level of detail in the dataset</i>	<i>variable names that parallel the structure</i>	<i>general class of the dataset using controlled terminology</i>	<i>links or references to documentation re how the dataset was created</i>
ADQSADAS	Data for the ADAS-Cog (11) Analyses	<u>adqsadas.xpt</u>	one record per subject per parameter per analysis visit	USUBJID, PARAMCD, AVISIT	BDS	DSADQSADAS.SAS, <u>Section 14.11</u> of SAP for detailed ADAS-Cog scoring algorithm

ADaM Dataset Metadata Cheat Sheet

ADaM 2.1

Define 1.0

Stylesheet

Dataset Name

Name

Dataset

Dataset Description

def:Label

Description

Dataset Location

def:ArchiveLocationID

Location

Dataset Structure

def:Structure

Structure

Key Variables of Dataset

def:DomainKeys

Keys

Class of Dataset

def:Class

Class

Documentation

Comment

Documentation

Purpose

Purpose

No Purpose*

ADaM Dataset Metadata Cheat Sheet

ADaM 2.1

Define 2.0

Stylesheet

Dataset Name

Name

Dataset

Dataset Description

def:Label

Description

Dataset Location

def:ArchiveLocationID

Location

Dataset Structure

def:Structure

Structure

Key Variables of Dataset

def:DomainKeys

Keys

Class of Dataset

def:Class

Class

Documentation

def:CommentOID

Documentation

Purpose

Purpose

No Purpose*

Dataset Metadata (Define 1.0)

```
<ItemGroupDef OID="ADLBHY"
```

```
  Name="ADLBHY"
```

```
  Repeating="Yes"
```

```
  IsReferenceData="No"
```

```
  Purpose="Analysis"
```

```
  def:Label="Analysis Dataset Lab Hy's Law"
```

```
  def:Structure="one record per subject per parameter per  
    analysis visit"
```

```
  def:DomainKeys="USUBJID, PARAMCD, AVISIT, LBSEQ"
```

```
  def:Class="BDS"
```

```
  def:ArchiveLocationID="Location.ADLBHY"
```

```
  Comment="See Data Guide Section 2.1"
```

```
>
```

Dataset Metadata (Define 2.0)

```
<ItemGroupDef OID="ADLBHY"  
  Name="ADLBHY"  
  Repeating="Yes"  
  IsReferenceData="No"  
  Purpose="Analysis"  
  def:Structure="one record per subject per parameter per analysis visit"  
  def:DomainKeys="USUBJID, PARAMCD, AVISIT, LBSEQ"  
  def:Class="BDS"  
  def:ArchiveLocationID="Location.ADLBHY"  
  <Description>  
    <TranslatedText xml:lang="en">Analysis Dataset Lab Hy's Law</  
TranslatedText>  
  </Description>  
  def:CommentOID="COM.ADLBHY"  
>
```

Dataset Metadata (Define 2.0)

```
<def:CommentDef OID="COM.ADLBHY">  
  <Description>  
    <TranslatedText xml:lang="en">See referenced  
    documents:</TranslatedText>  
  </Description>  
  <def:DocumentRef leafID="LF.ADPGM_ADLBHY"/>  
  <def:DocumentRef leafID="LF.DataGuide">  
    <def:PDFPageRef PageRefs="Section2.1"  
    Type="NamedDestination"/>  
  </def:DocumentRef>  
</def:CommentDef>
```

ADaM Variable Metadata

ADaM-MSG

Analysis Data Model Document

Table 5.2.1 Analysis Variable Metadata Fields

Analysis Variable Metadata Field	Description
DATASET NAME	The file name of the analysis dataset
VARIABLE NAME	The name of the variable
VARIABLE LABEL	A brief description of the variable
VARIABLE TYPE	The variable type. Valid values are as defined in the Case Report Tabulation Data Definition Specification Standard (e.g., in version 1.0.0 they include “text,” “integer,” and “float”)
DISPLAY FORMAT	The variable display information (i.e., the format used for the variable in a tabular or graphical presentation of results). It is suggested that the syntax be consistent with the format terminology incorporated in the software package used for analysis (e.g., \$16 or 3.1 if using SAS).
CODELIST / CONTROLLED TERMS	A list of valid values or allowable codes and their corresponding decodes for the variable. The field can include a reference to an external codelist (identified by name and version) or a hyperlink to a list of the values in the codelist/controlled terms section of the define file.
SOURCE / DERIVATION	Provides details about the variable’s lineage – what was the predecessor, where the variable came from in the source data (SDTM or other analysis dataset) or how the variable was derived. This field is used to identify the immediate predecessor source and/or a brief description of the algorithm or process applied to that source and can contain hyperlinked text that refers readers to additional information. The source / derivation can be as simple as a two level name (e.g., ADSL.AGEGR) identifying the data file and variable that is the source of the variable (i.e., a variable copied with no change). It can be a simple description of a derivation and the variable used in the derivation (e.g., “categorization of ADSL.BMI”). It can also be a complex algorithm, where the element contains a complete description of the

Stylesheet Example (Define 2.0)

Analysis Variable Metadata

Sample Performance Score Analysis Dataset (ADSAMPLE)					adsample.xpt
Variable	Label	Type	Length/ Display Format	Controlled Terms or Format	Source/Derivation/Comment
STUDYID	Study Identifier	text	3		Predecessor: ADSL.STUDYID

Analysis Data Model Document

Table 5.2.2.1 Analysis Variable Metadata

Dataset Name	Parameter Identifier	Variable Name	Variable Label	Variable Type	Display Format	Codelist / Controlled Terms	Source / Derivation
<i>file name of the analysis dataset</i>	<i>PARAMCD or *ALL* or *DEFAULT*</i>	<i>name</i>	<i>description</i>	<i>type</i>	<i>display information</i>	<i>valid values or codes and decodes</i>	<i>where the variable came from in the source data or how the variable was derived</i>
ADQSADAS	*ALL*	STUDYID	Study Identifier	text	\$12		ADSL.STUDYID
ADQSADAS	*ALL*	SITEID	Study Site Identifier	text	\$3		ADSL.SITEID
ADQSADAS	*ALL*	SITEGR1	Pooled Site Group 1	text	\$3		ADSL.SITEGR1
ADQSADAS	*ALL*	USUBJID	Unique Subject Identifier	text	\$11		ADSL.USUBJID
ADQSADAS	*ALL*	AVISIT	Analysis Visit	text	\$19	Baseline, Week 8, Week 16, Week 24	If ADQSADAS.ITTRFL='Y' then AVISIT is the name of the analysis visit; if ADQSADAS.ITTRFL=blank then AVISIT=blank. 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	*ALL*	VISIT	Visit Name	text	\$19		QS.VISIT

ADaM Variable Metadata

- ADaM: *Dataset Name*, Name, Label, Type, *Display Format*, Codelist/Controlled Terms, *Source/Derivation*
- SDTM: Name, Label, Type, *Key*, *Length*, Codelist/Controlled Terms, Derivation, *Origin*, *Role*, *Comments*

ADaM Variable Metadata Cheat Sheet

ADaM 2.1

Define 1.0

Stylesheet

Variable Name

Name

Variable

Variable Label

def:Label

Label

Variable Type

DataType

Type

Display Format

def:DisplayFormat

Display Format

Codelist/Controlled Terms

CodeListRef

Code List/
Controlled Terms

Source/Derivation

Comment

Source/
Derivation/
Comments

def:Origin

Length

Length

Origin="Derived"

Length

No Role

ADaM Variable Metadata Cheat Sheet

ADaM 2.1

Variable Name

Variable Label

Variable Type

Display Format

Codelist/Controlled Terms

Source/Derivation

Define 2.0

Name

Description

DataType

def:DisplayFormat

CodeListRef

MethodDef

Type=Computation

Description

def:Origin Type="..."

Length

Stylesheet

Variable

Label

Type

Length/Display Format

Controlled Terms or Format

Source/

Derivation/

Comments

Length/Display Format

Origin="Derived"

Length

No Role

Define 1.0 Example: Analysis Variable Metadata

```
<ItemRef ItemOID="ADLBC.BASE" OrderNumber="16"  
Mandatory="No"/>
```

```
<ItemDef OID="ADLBC.BASE" Name="BASE"  
DataType="float" Length="8" SignificantDigits="5"  
Origin="Derived" Comment="ADLBC.LBSTRESN when  
ADLBC.ABLFL=Y" def:Label="Baseline Value">  
</ItemDef>
```

Define 2.0 Example: Analysis Variable Metadata

```
<ItemRef ItemOID="ADLBC.BASE" OrderNumber="16"  
Mandatory="No" MethodOID="MT. ADLBC.BASE"/>
```

```
<ItemDef OID="ADLBC.BASE" Name="BASE"  
DataType="float" SignificantDigits="5" Length="8"  
def:DisplayFormat="8.1">
```

```
<Description>
```

```
<TranslatedText xml:lang="en">Baseline Value</  
TranslatedText>
```

```
</Description>
```

```
<def:Origin Type="Derived"/>
```

```
</ItemDef>
```

Define 2.0 Example: Analysis Variable Metadata, cont.

```
<MethodDef Type="Computation" OID="MT.  
ADLBC.BASE" Name="Algorithm for ADLBC.BASE">  
  <Description>  
    <TranslatedText xml:lang="en">AVAL where  
AVISIT=&quot;BASELINE&quot;</TranslatedText>  
  </Description>  
</MethodDef>
```

Analysis Results Metadata (ARM)

ADaM-MSG

Analysis Results Metadata

Analysis Results Metadata Field	Description
DISPLAY IDENTIFIER	A unique identifier for the specific analysis display (such as a table or figure number)
DISPLAY NAME	Title of display, including additional information if needed to describe and identify the display (e.g., analysis population)
RESULT IDENTIFIER	Identifies the specific analysis result within a display. For example, if there are multiple p-values on a display and the analysis results metadata specifically refers to one of them, this field identifies the p-value of interest. When combined with the display identifier provides a unique identification of a specific analysis result.
PARAM	The analysis parameter in the BDS analysis dataset that is the focus of the analysis result. Does not apply if the result is not based on a BDS analysis dataset.
PARAMCD	Corresponds to PARAM in the BDS analysis dataset. Does not apply if the result is not based on a BDS analysis dataset.
ANALYSIS VARIABLE	The analysis variable being analyzed
REASON	The rationale for performing this analysis. It indicates when the analysis was planned (e.g., "Pre-specified in Protocol," "Pre-specified in SAP," "Data Driven," "Requested by Regulatory Agency") and the purpose of the analysis within the body of evidence (e.g., "Primary Efficacy," "Key Secondary Efficacy," "Safety"). The terminology used is sponsor defined. An example of a reason is "Primary Efficacy Analysis as Pre-specified in Protocol."
DATASET	The name of the dataset used to generate the analysis result. In most cases, this is a single dataset. However, if multiple datasets are used, they are all listed here.
SELECTION CRITERIA	Specific and sufficient selection criteria for analysis subset and / or numerator – a complete list of the variables and their values used to identify the records selected for the analysis. Though the syntax is not ADaM-specified, the expectation is that the information could easily be included in a WHERE clause or something equivalent to ensure selecting the exact set of records appropriate for an analysis. This information is required if the analysis does not include every record in the analysis dataset.
DOCUMENTATION	Textual description of the analysis performed. This information could be a text description, pseudo code, or a link to another document such as the protocol or statistical analysis plan, or a link to an analysis generation program (i.e., a statistical software program used to generate the analysis result). The contents of the documentation metadata element contains depends on the level of detail required to describe the analysis itself, whether or not the sponsor is providing a corresponding analysis generation program, and sponsor-specific requirements and standards. This documentation metadata element will remain free form, meaning it will not become subject to a rigid structure or controlled terminology.
PROGRAMMING STATEMENTS	The software programming code used to perform the specific analysis. This includes, for example, the model statement (using the specific variable names) and all technical specifications needed for reproducing the analysis (e.g., covariance structure). The name and version of the applicable software package should be specified either as part of this metadata element or in another document, such as a Reviewer's Guide (see Appendix B for more information about a Reviewer's Guide).

Analysis Results Metadata Cheat Sheet

ADaM 2.1

Define 1.0

Stylesheet

	adamref:AnalysisResultDisplays	
	ResultDisplay	
Display Identifier	DisplayIdentifier	
Display Name	DisplayLabel	Display
	adamref:AnalysisResults	
Result Identifier	adamref:ResultIdentifier	
Reason	Reason	Reason
	adamref:ParameterList	Analysis Parameter(s)
	adamref:Parameter	
Param	Param	ParamCD=Param
ParamCD	ParamCD	
Analysis Variable	adamref:AnalysisVariable	Analysis Variable(s)
Dataset	adamref:AnalysisDataset	Data References (incl. Sel Crit)
Selection Criteria	adamref:SelectionCriteria	
Documentation	adamref:Documentation	Documentation
Programming Statements	adamref:ProgrammingCode	Programming Statements

Display	Table 14-1.01 Summary of Populations
AnalysisResult	
Analysis Parameter(s)	
Analysis Variable(s)	ITTFL SAFFL EFFFL COMP24FL DISCONFL
Reason	pre-specified in SAP
Data References (incl. Selection Criteria)	ADSL
Documentation	SAP Section 9.1
Programming Statements	<pre>proc freq data = pilot.adsl; tables (ittfl saffl efffl comp24fl disconfl)*trt01pn / missing; run;</pre>

Define 1.0/2.0

```
<adamref:AnalysisResultDisplays>
```

```
  <adamref:ResultDisplay DisplayIdentifier="Table_14-1.01" OID="Table_14-1.01"  
  DisplayLabel="Summary of Populations" leafID="Table_14-1.01">
```

```
    <adamref:AnalysisResults OID="Table_14-1.01.R.1" Reason="pre-specified in SAP">
```

```
      <adamref:AnalysisVariable ItemOID="ADSL.ITTFL"/>
```

```
      <adamref:AnalysisDataset>
```

```
        <ItemGroupRef ItemGroupOID="ADSL" Mandatory="No"/>
```

```
      </adamref:AnalysisDataset>
```

```
      <adamref:Documentation leafID="SAP_SEC_9.1"/>
```

```
      <adamref:ProgrammingCode>
```

```
        <def:ComputationMethod OID="Table_14-1.01.R.1">
```

```
          proc freq data = pilot.adsl;
```

```
          tables (ittfl saffl efffl comp24fl disconf1)*trt01pn / missing;
```

```
          run;
```

```
        </def:ComputationMethod>
```

```
      </adamref:ProgrammingCode>
```

```
    </adamref:AnalysisResults>
```

```
  </adamref:ResultDisplay>
```

```
</adamref:AnalysisResultDisplays>
```

Parameter Value-Level Metadata: ADaM Basic Data Structure (BDS)

ADaM-MSG

Analysis Data Model Document

5.2.1 Analysis Parameter Value-Level Metadata

Analysis Variable Metadata Field	Description
PARAMETER IDENTIFIER	Contains either 1) the value of PARAMCD that identifies the analysis parameter to which the variable metadata applies; or 2) controlled terminology to indicate groupings of analysis parameters: *ALL* - Used when the variable metadata applies to all analysis parameters in the dataset. *DEFAULT* - Used when the variable metadata applies to all analysis parameters in the dataset except for those specifically listed within the metadata.

Analysis Data Model Document

Table 5.2.2.1 Analysis Variable Metadata

Dataset Name	Parameter Identifier	Variable Name	Variable Label	Variable Type	Display Format	Codelist / Controlled Terms	Source / Derivation
<i>file name of the analysis dataset</i>	<i>PARAMCD or *ALL* or *DEFAULT*</i>	<i>name</i>	<i>description</i>	<i>type</i>	<i>display information</i>	<i>valid values or codes and decodes</i>	<i>where the variable came from in the source data or how the variable was derived</i>
ADQSADAS	*ALL*	STUDYID	Study Identifier	text	\$12		ADSL.STUDYID
ADQSADAS	*ALL*	SITEID	Study Site Identifier	text	\$3		ADSL.SITEID
ADQSADAS	*ALL*	SITEGR1	Pooled Site Group 1	text	\$3		ADSL.SITEGR1
ADQSADAS	*ALL*	USUBJID	Unique Subject Identifier	text	\$11		ADSL.USUBJID
ADQSADAS	*ALL*	AVISIT	Analysis Visit	text	\$19	Baseline, Week 8, Week 16, Week 24	If ADQSADAS.ITTRFL='Y' then AVISIT is the name of the analysis visit; if ADQSADAS.ITTRFL=blank then AVISIT=blank. 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	*ALL*	VISIT	Visit Name	text	\$19		QS.VISIT

Background: ADaM Basic Data Structure (BDS)

- BDS can contain multiple analysis parameters.
- In BDS, the variable PARAM contains a unique description for every analysis parameter included in that dataset. This is similar to LBTEST or VSTEST in SDTM.
- Each value of PARAM identifies a set of one or more rows in the dataset.

Define 1.0 Example: Parameter Value-Level Metadata

Source Variable	Where PARAMCD=	Where PARAM=	Type	Length	Display Format	Code List / Controlled Term	Source/Derivation/Comments
CHG	CK	Creatine Kinase (U/L)	integer	8			AVAL-BASE
CHG	CL	Chloride (mmol/L)	integer	8			AVAL-BASE
CHG	CREAT	Creatinine (umol/L)	float	8	6.2		AVAL-BASE

Define 1.0 Example: Parameter Value-Level Metadata

```
<ItemDef OID="ADLBC.CHG"  
  Name="CHG"  
  DataType="float"  
  Length="8"  
  SignificantDigits="5"  
  Origin="Derived"  
  Comment=" "  
  def:Label="Change from Baseline"  
>  
  <def:ValueListRef ValueListOID="ValueList.ADLBC.CHG"/>  
</ItemDef>
```

Define 1.0 Example: Parameter Value-Level Metadata

```
<def:ValueListDef OID="ValueList.ADLBC.CHG">
```

```
...
```

```
<ItemRef ItemOID="ADLBC.CHG.CREAT"  
  OrderNumber="299" Mandatory="No"/>
```

```
...
```

```
</def:ValueListDef>
```

Define 1.0 Example: Parameter Value-Level Metadata

```
<ItemDef OID="ADLBC.CHG.CREAT"  
  Name="CREAT"  
  DataType="float"  
  Length="8"  
  SignificantDigits="2"  
  Origin="Derived"  
  Comment="AVAL-BASE"  
  def:Label="Creatinine (umol/L)"  
  def:DisplayFormat="6.2"  
>
```

Define 2.0 Example: Parameter Value-Level Metadata

Value Level Metadata (VL.ADSAMPLE.AVAL)					
Variable	Where	Type	Length/ Display Format	Controlled Terms or Format	Source/Derivation/Comment
AVAL	PARAMCD EQ SCORE1	integer			QS.QSSTRESN where QS.QSTESTCD="SCORE1"
AVAL	PARAMCD EQ SCORE2	integer			QS.QSSTRESN where QS.QSTESTCD="SCORE2"
AVAL	PARAMCD EQ TOTSCORE	float	8.1		Derived from Score 1 and Score 2. Details are described in the Data Guide (Page 1) Data Guide (dataguide.pdf)

Define 2.0 Example: Parameter Value-Level Metadata

```
<ItemDef OID="IT.ADSAMPLE.AVAL" Name="AVAL"  
  DataType="float" SignificantDigits="1" Length="8"  
  def:CommentOID="COM.ADSAMPLE.AVAL"  
  def:DisplayFormat="8.1">  
  <Description>  
    <TranslatedText xml:lang="en">Analysis Value</  
TranslatedText>  
  </Description>  
  <def:ValueListRef ValueListOID="VL.ADSAMPLE.AVAL"/>  
</ItemDef>
```

Define 2.0 Example: Parameter Value-Level Metadata

```
<def:ValueListDef OID="VL.ADSAMPLE.AVAL">
  <ItemRef ItemOID="IT.ADSAMPLE.AVAL.SCORE1" OrderNumber="1"
    Mandatory="No" MethodOID="MT.ADSAMPLE.AVAL.SCORE1">
    <def:WhereClauseRef WhereClauseOID="WC.ADSAMPLE.AVAL.SCORE1"/>
  </ItemRef>
  <ItemRef ItemOID="IT.ADSAMPLE.AVAL.SCORE2" OrderNumber="2"
    Mandatory="No" MethodOID="MT.ADSAMPLE.AVAL.SCORE2">
    <def:WhereClauseRef WhereClauseOID="WC.ADSAMPLE.AVAL.SCORE2"/>
  </ItemRef>
  <ItemRef ItemOID="IT.ADSAMPLE.AVAL.TOTSCORE" OrderNumber="3"
    Mandatory="No" MethodOID="MT.ADSAMPLE.AVAL.TOTSCORE">
    <def:WhereClauseRef WhereClauseOID="WC.ADSAMPLE.AVAL.TOTSCORE"/>
  </ItemRef>
</def:ValueListDef>
```

Define 2.0 Example: Parameter Value-Level Metadata

```
<def:WhereClauseDef OID="WC.ADSAMPLE.AVAL.TOTSCORE">  
  <RangeCheck SoftHard="Soft"  
    def:ItemOID="IT.ADSAMPLE.PARAMCD" Comparator="EQ">  
      <CheckValue>TOTSCORE</CheckValue>  
    </RangeCheck>  
</def:WhereClauseDef>
```

Define 2.0 Example: Parameter Value-Level Metadata

```
<MethodDef Type="Computation"
OID="MT.ADSAMPLE.AVAL.TOTSCORE" Name="Algorithm for Total
Score">
  <Description>
    <TranslatedText xml:lang="en">Derived from Score 1 and Score 2.
Details are described in the Data Guide (Page 1)</TranslatedText>
  </Description>
  <def:DocumentRef leafID="LF.DataGuide">
    <def:PDFPageRef PageRefs="1" Type="PhysicalRef"/>
  </def:DocumentRef>
</MethodDef>
```