解析結果メタデータの概要とSASによる作成方法の提案

高浪 洋平

武田薬品工業株式会社
クリニカルデータサイエンス部統計グループ

Overview and Creation of Analysis Results Metadata with SAS

Yohei Takanami

Biostatistics, Clinical Data Science Dept.
Takeda Pharmaceutical Company, Ltd.

要旨：

医薬品製造販売承認申請時の電子データ提出において、ADaMのDefine.xmlに含めて提出することが推奨されている解析結果メタデータの概要と実装時の留意事項、ならびにSASによる作成方法を提案する。

For e-Data submission for Japan NDA, Analysis Results Metadata (ARM) which is recommended to be included in ADaM Define.xml and submitted to PMDA. In this presentation, overview and considerations of ARM will be summarized and the creation process using SAS will be proposed.

キーワード：CDISC，ADaM，Analysis Results Metadata，Define.xml

Outline

- **Background**
  - Analysis Results Metadata for e-Data submission in Japan
- **Overview of Analysis Results Metadata**
  - Analysis Results Metadata in ADaM
  - Structure of Analysis Results Metadata
  - Considerations on Analysis Results Metadata
- **Creation of Analysis Results Metadata with SAS**
  - Excel Metadata File and SAS program to generate Analysis Results Metadata as a part of Define.xml

Outline

- **Background**
  - Analysis Results Metadata for e-Data submission in Japan
- **Overview of Analysis Results Metadata**
  - Analysis Results Metadata in ADaM
  - Structure of Analysis Results Metadata
  - Considerations on Analysis Results Metadata
- **Creation of Analysis Results Metadata with SAS**
  - Excel Metadata File and SAS program to generate Analysis Results Metadata as a part of Define.xml

Background

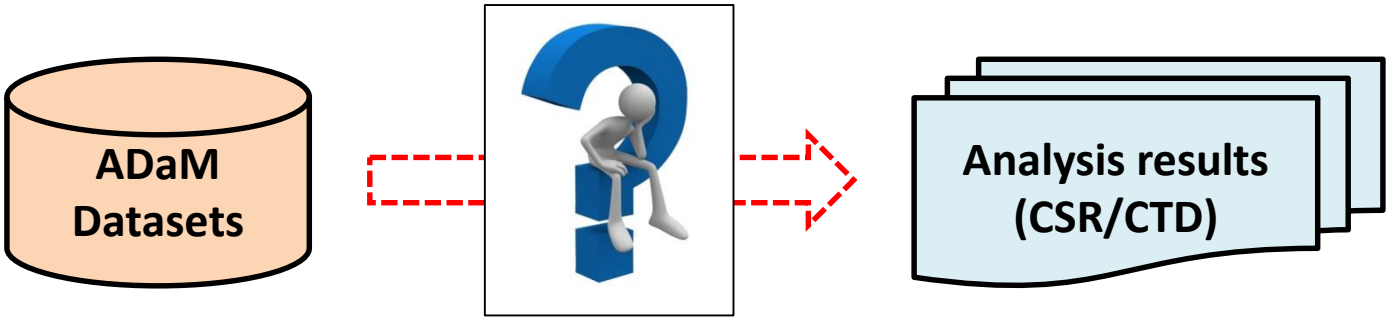
- New guidance and Technical Conformance Guide in Japan
  - PMDA will start to receive e-Data of clinical studies from Oct. 2016 (transition period is 3.5 years)
  - For mainly Ph-II, Ph-III and specific Ph-I/Clinical Pharmacology studies, sponsors have to submit the CDISC-compliant datasets (SDTM and ADaM), Define.xml, Reviewer's Guide and Annotated CRF
    - For Ph-I/CP studies, sponsors have to discuss the targeted studies and the contents of e-Data with PMDA at the consultation meetings
  - PMDA recommends sponsors to submit the Analysis Results Metadata (ARM) (Technical Conformance Guide).
    - Not for all analyses (e.g. primary and secondary analyses)
    - If it is difficult to prepare the ARM, sponsors need to prepare the other documents/specifications for the explanation of analyses

Outline

- Background
  - Analysis Results Metadata for e-Data submission in Japan
- **Overview of Analysis Results Metadata**
  - Analysis Results Metadata in ADaM
  - Structure of Analysis Results Metadata
  - Considerations on Analysis Results Metadata
- Creation of Analysis Results Metadata with SAS
  - Excel Metadata File and SAS program to generate Analysis Results Metadata as a part of Define.xml

What is the Analysis Results Metadata (ARM)?

- One of the Metadata in ADaM (V2.1)
  - Submitted to PMDA as a part of Define.xml for ADaM
    - Analysis Results Metadata V1.0 for Define.xml V2.0 was released in Jan., 2015
  - Provide traceability between analysis results and ADaM datasets (e.g. datasets, variables, selection criteria, programming code used in the analyses)
  - Reviewers can understand how the analysis results are created.



Traceability with ARM

- Between Analysis Results and ADaM datasets

Protocol: CDISCPILOT01 Page 1 of 1
Population: Efficacy

Table 14-3.11
ADAS Cog (11) - Repeated Measures Analysis of Change from Baseline to Week 24

	Placebo (N=79)	Xanomeline Low Dose (N=81)	Xanomeline High Dose (N=74)
LS Means (SE)	1.6 (0.49)	1.5 (0.52)	1.1 (0.56)
p-value(Xan - Placebo)		0.955	0.556
Diff of LS Means (SE)		-0.0 (0.70)	-0.4 (0.72)
95% CI		(-1.4;1.3)	(-1.9;1.0)
p-value(Xan High - Xan Low)			0.606
Diff of LS Means (SE)			-0.4 (0.75)
95% CI			(-1.9;1.1)

Description, Datasets, Variables, Selection criteria, Programming Code etc.

Metadata Field	*Definition of field*	Metadata
DISPLAY IDENTIFIER	*Unique identifier for the specific analysis display*	Table 14-3.11
DISPLAY NAME	*Title of display*	ADAS Cog (11) - Repeated Measures Analysis of Change from Baseline to Week 24
RESULT IDENTIFIER	*Identifies the specific analysis result within a display*	
PARAM	*Analysis parameter*	ADAS-Cog (11) Total Score
PARAMCD	*Analysis parameter code*	ACTOT11
ANALYSIS VARIABLE	*Analysis variable being analyzed*	CHG
REASON	*Rationale for performing this analysis*	Pre-specified in SAP
DATASET	*Dataset(s) used in the analysis.*	ADQSADAS
SELECTION CRITERIA	*Specific and sufficient selection criteria for analysis subset and / or numerator*	ITTFL='Y' and AVISITN GT 0 AND DTYPE NE 'LOCF' AND PARAMCD='ACTOT11'
DOCUMENTATION	*Textual description of the analysis performed*	SAP Section 10.1.1. Adjusted means for the change from baseline at week 24 and pairwise comparisons between treatment groups at Week 24 using a repeated measures model with treatment group (as class variable); site (as class variable); time; treatment*time interaction; baseline score and baseline*time interaction terms; and an unstructured covariance matrix. Efficacy data, observed cases data.
PROGRAMMING STATEMENTS	*The analysis syntax used to perform the analysis*	PROC MIXED; CLASS USUBJID SITEGR1 AVISITN TRTP; MODEL CHG = TRTP SITEGR1 AVISITN TRTP*AVISITN BASE BASE*AVISITN / OUTP=PRED DDFM=KR; REPEATED AVISITN / SUBJECT=USUBJID TYPE=UN; LSMEANS TRTP / DIFF CL; RUN;

ARM in the ADaM Define.xml

- Specific structure in Define.xml
  - Output Define.xml on a browser with XML stylesheet

Metadata Field	*Definition of field*	Metadata
DISPLAY IDENTIFIER	*Unique identifier for the specific analysis display*	Table 14-3.11
DISPLAY NAME	*Title of display*	ADAS Cog (11) - Repeated Measures Analysis of Change from Baseline to Week 24
RESULT IDENTIFIER	*Identifies the specific analysis result within a display*	
PARAM	*Analysis parameter*	ADAS-Cog (11) Total Score
PARAMCD	*Analysis parameter code*	ACTOT11
ANALYSIS VARIABLE	*Analysis variable being analyzed*	CHG
REASON	*Rationale for performing this analysis*	Pre-specified in SAP
DATASET	*Dataset(s) used in the analysis.*	ADQSADAS
SELECTION CRITERIA	*Specific and sufficient selection criteria for analysis subset and / or numerator*	ITTFL='Y' and AVISITN GT 0 AND DTYPE NE 'LOCF' AND PARAMCD='ACTOT11'
DOCUMENTATION	*Textual description of the analysis performed*	SAP Section 10.1.1. Adjusted means for the change from baseline at week 24 and pairwise comparisons between treatment groups at Week 24 using a repeated measures model with treatment group (as class variable); site (as class variable); time; treatment*time interaction; baseline score and baseline*time interaction terms; and an unstructured covariance matrix. Efficacy data, observed cases data.
PROGRAMMING STATEMENTS	*The analysis syntax used to perform the analysis*	PROC MIXED; CLASS USUBJID SITEGR1 AVISITN TRTP; MODEL CHG = TRTP SITEGR1 AVISITN TRTP*AVISITN BASE BASE*AVISITN / OUTP=PRED DDFM=KR; REPEATED AVISITN / SUBJECT=USUBJID TYPE=UN; LSMEANS TRTP / DIFF CL; RUN;

```
<arm:AnalysisResultDisplays>

  <arm:ResultDisplay OID="RD.Table_14-3.01" Name="Table 14-3.01">
    <Description>
      <TranslatedText xml:lang="en">Primary Endpoint Analysis: ADAS-Cog - Summary at Week 24 - LOCF (Efficacy Population)</T
    </Description>
    <def:DocumentRef leafID="LF.Table-14-3.01">
      <def:PDFPageRef PageRefs="2" Type="PhysicalRef"/>
    </def:DocumentRef>
    <arm:AnalysisResult
      OID="AR.Table_14-3.01.R.1"
      ParameterOID="IT.ADQSADAS.PARAMCD"
      AnalysisReason="SPECIFIED IN SAP"
      AnalysisPurpose="PRIMARY OUTCOME MEASURE">
      <Description>
        <TranslatedText xml:lang="en">Dose response analysis for ADAS-Cog changes from baseline</TranslatedText>
      </Description>
      <arm:AnalysisDatasets>
        <arm:AnalysisDataset ItemGroupOID="IG.ADQSADAS" >
          <def:WhereClauseRef WhereClauseOID="WC.Table_14-3.01.R.1.ADQSADAS" />
          <arm:AnalysisVariable ItemOID="IT.ADQSADAS.CHG"/>
        </arm:AnalysisDataset>
      </arm:AnalysisDatasets>
      <arm:Documentation>
        <Description>
        <TranslatedText xml:lang="en">Linear model analysis of CHG for dose response; using randomized dose (0 for placebo;
n SAS to produce p-value (from Type III SS for treatment dose).
        </TranslatedText>
        </Description>
        <def:DocumentRef  leafID="LF.SAP-SEC-10.1.1">
          <def:PDFPageRef PageRefs="4" Type="PhysicalRef"/>
        </def:DocumentRef>
      </arm:Documentation>
      <arm:ProgrammingCode Context="SAS version 9.2">
```

Table 14-3.01

Display	Table 14-3.01 Primary Endpoint Analysis: ADAS-Cog - Summary at Week 24 - LOCF (Efficacy Population)
Analysis Result	**Dose response analysis for ADAS-Cog changes from baseline**
Analysis Parameter(s)	ACTOT = Adas-Cog(11) Subscore
Analysis Variable (s)	CHG (Change from Baseline)
Analysis Reason	Pre-specified in SAP
Analysis Purpose	Primary Outcome Measure
Data References (incl. Selection Criteria)	ADQSADAS [PARAMCD = "ACTOT" and AVISIT = "Week 24" and EFFFL = "Y" and ANL01FL = "Y"]
Documentation	Linear model analysis of CHG for dose response; using randomized dose (0 for placebo; 54 for low dose; 81 for high dose) and site group in model. Used PROC GLM in SAS to produce p-value (from Type III SS for treatment dose). SAP Section 10.1.1

Programming Statements	[SAS version 9.2 or later]
	proc glm data = ADQSADAS;
	where EFFFL='Y' and ANL01FL='Y' and AVISIT='Week 24' and PARAMCD="ACTOT";
	class SITEGR1;
	model CHG = TRTPN SITEGR1;
	run;

Browser (with Stylesheet)

Structure of Analysis Results Metadata

- Analysis Results Metadata V1.0 for Define.xml V2.0

CDISC

Analysis Results Metadata Specification
Version 1.0
for
Define-XML Version 2

Prepared by
CDISC ADaM Metadata Sub-Team

[redacted]

Notes to Readers

- This is the specification for the Analysis Results Metadata extension of Version 2.0 of the CDISC Define-XML standard.
- This document is intended for companies and individuals involved in the analysis of clinical data that will be submitted to regulatory authorities.

Revision History

Date	Version	Summary of Changes
2015-01-27	1.0	Final version 1.0 incorporating all changes identified during the public comment period
2014-09-10	1.0 DRAFT	Draft for public review

Warranties Notice: see Appendix C

```
<?xml version="1.0" encoding="UTF-8"?>
<ODM xmlns="http://www.cdisc.org/ns/odm/v1.3"
  xmlns:def="http://www.cdisc.org/ns/def/v2.0"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:arm="http://www.cdisc.org/ns/arm/v1.0"
  ODMVersion="1.3.2" FileType="Snapshot" FileOID="CDISC-Sample"
  CreationDateTime="2014-03-28T11:07:23:00"
  Originator="CDISC ADaM Metadata Team">
  <Study OID="cdisc01">
    <GlobalVariables>
      <StudyName>CDISC Sample</StudyName>
      <StudyDescription>CDISC-Sample Data Definitions</StudyDescription>
      <ProtocolName>CDISC-Sample</ProtocolName>
    </GlobalVariables>
    <MetaDataVersion OID="MDV.CDISC01.ADaMIG.1.0.ADaM.2.1"
      Name="Study CDISC-Sample, Data Definitions"
      Description="Study CDISC01, Data Definitions"
      def:DefineVersion="2.0.0"
      def:StandardName="ADaM-IG"
      def:StandardVersion="1.0">
```

< Supplemental Data Definitions (def:SupplementalDoc) >
< Value Level Metadata (def:ValueListDef) >
< Where Clause Definitions (def:WhereClauseDef) >
< Domain Level Metadata (ItemGroupDef) >
< Variable Level Metadata (ItemDef) >
< Controlled Terminology Metadata (CodeList) >
< Computational Algorithms (MethodDef) >
< Comments (def:CommentDef) >
< Referenced Documents (def:leaf) >
< Analysis Results Metadata (arm:AnalysisResultDisplays) >

```
    </MetaDataVersion>
  </Study>
</ODM>
```

Structure of Analysis Results Metadata

- Tagsets under the <arm:AnalysisResultDisplays>

```
<arm:AnalysisResultDisplays>
  <arm:ResultDisplay>
    <Description>
      <TranslatedText>
    <def:DocumentRef>
      <def:PDFPageRef>
    <arm:AnalysisResult>
      <Description>
        <TranslatedText>
      <arm:AnalysisDatasets>
        <arm:AnalysisDataset>
          <def:WhereClauseRef>
          <arm:AnalysisVariable>
      <arm:Documentation>
        <Description>
          <TranslatedText>
         <def:DocumentRef>
          <def:PDFPageRef>
      <arm:ProgrammingCode>
        <arm:Code>
        <def:DocumentRef>
          <def:PDFPageRef>
```

Structure of Analysis Results Metadata

- <arm:ResultDisplay>
  - Corresponds to a table/figure in CSR

```
<arm:AnalysisResultDisplays>
   <arm:ResultDisplay OID="RD.Table_14-3.01" Name="Table 14-3.01">
      <Description>
         <TranslatedText xml:lang="en">Primary Endpoint Analysis: ADAS-Cog - Summary at
         Week 24 - LOCF (Efficacy Population)
         </TranslatedText>
      </Description>
      <def:DocumentRef leafID="LF.Table-14-3.01">
         <def:PDFPageRef PageRefs="49" Type="PhysicalRef"/>
      </def:DocumentRef>
      <arm:AnalysisResult>
      …
      </arm:AnalysisResult>
      <arm:AnalysisResult>
      …
      </arm:AnalysisResult>
   </arm:ResultDisplay>
   …
</arm:AnalysisResultDisplays>
```

Analsysis result display

Link to tables in CSR

Analysis result (multiple results can be included in one table/figure)

Structure of Analysis Results Metadata

- <arm:AnalysisResult>
  - Corresponds to an analysis result

Analysis Result Display

Analysis result (multiple results can be included in one table/figure)

```
<arm:ResultDisplay OID="RD.Table_14-3.01" Name="Table 14-3.01">
  ...
  <arm:AnalysisResult OID="AR.Table_14-3.01.R.1" ParameterOID="IT.ADQSADAS.PARAMCD"
    AnalysisReason="SPECIFIED IN SAP" AnalysisPurpose="PRIMARY OUTCOME MEASURE">
    <Description>
      <TranslatedText xml:lang="en">Dose response analysis for ADAS-Cog changes from baseline
      </TranslatedText>
    </Description>
    <arm:AnalysisDatasets>
    ...
    </arm:AnalysisDatasets>
    <arm:Documentation>
    ...
    </arm:Documentation>
    <arm:ProgrammingCode>
    ...
    </arm:ProgrammingCode>
  </arm:AnalysisResult>
  <arm:AnalysisResult>
  ...
  </arm:AnalysisResult>
</arm:ResultDisplay>
```

Datasets (multiple datasets can be described)

Link to the specification of an analysis

Analysis Program (code or link to program file)

Analysis Results Metadata (Summary) for Study CDISC-Sample

Table 14-3.01 Primary Endpoint Analysis: ADAS-Cog - Summary at Week 24 - LOCF (Efficacy Population)
Dose response analysis for ADAS-Cog changes from baseline
Pairwise comparisons to placebo for ADAS-Cog changes from baseline

ergent Serious Adverse Events by Treatment Group
s Adverse Events by Treatment Group

etail) for Study CDISC-Sample

Table 14-3.01

	Table 14-3.01 Primary Endpoint Analysis: ADAS-Cog - Summary at Week 24 - LOCF (Efficacy Population)
	Dose response analysis for ADAS-Cog changes from baseline
s)	PARAMCD = "ACTOT" (Adas-Cog(11) Subscore)
	CHG (Change from Baseline)
Analysis Reason	SPECIFIED IN SAP
	E MEASURE
	MCD = "ACTOT" and AVISIT = "Week 24" and EFFFL = "Y" and ANL01FL = "Y"]
	sis of CHG for dose response; using randomized dose (0 for placebo; 54 for low dose; 81 for high dose) and site group in model. Used to produce p-value (from Type III SS for treatment dose).
	SAP Section 10.1.1

Programming Statements	[SAS version 9.2]
	proc glm data = ADQSADAS;
	where EFFFL='Y' and ANL01FL='Y' and AVISIT='Week 24' and PARAMCD="ACTOT";
	class SITEGR1;
	model CHG = TRTPN SITEGR1;
	run;

Structure of Analysis Results Metadata

- <arm:AnalysisDatasets> and <arm:AnalysisDataset>
  - Corresponds to datasets, variables, selection criteria

```
<arm:AnalysisResultDisplays>
   …
   <arm:AnalysisDatasets def:CommentOID="COM.JOIN-ADSL-ADAE">
      <arm:AnalysisDataset ItemGroupOID="IG.ADAE">
         <def:WhereClauseRef WhereClauseOID="WC.Table_14-5.02.R.1.ADAE" />
         <arm:AnalysisVariable ItemOID="IT.ADAE.AEBODSYS"/>
         <arm:AnalysisVariable ItemOID="IT.ADAE.AEDECOD"/>
      </arm:AnalysisDataset>
      …
      <arm:AnalysisDataset ItemGroupOID="IG.ADSL">
      …
      </arm:AnalysisDataset>
   </arm:AnalysisDatasets>
…
</arm:AnalysisResultDisplays>
```

Datasets

Each dataset

Selection criteria
(OID for def:WhereClauseDef)

Variables used in the analysis
(OID for ItemDef)

Each dataset (for multiple datasets)

	…ous Adverse Events by Treatment Group
	…by Treatment Group
Analysis Parameter(s)	
Analysis Variable (s)	AEBODSYS (Body System or Organ Class) AEDECOD (Dictionary-Derived Term)
Analysis Reason	SPECIFIED IN SAP
Analysis Purpose	PRIMARY OUTCOME MEASURE
Data References (incl. Selection Criteria)	ADAE [TRTEMFL = "Y" and AESER = "Y"] ADSL [SAFFL = "Y"] Get denominators for percentages from ADSL and counts and numerators from ADAE. Join ADAE with ADSL based on the unique subject identifier (USUBJID) keeping only records in ADAE for the numerator.
Documentation	Unique count of subjects that experienced an Adverse Event by Preferred Term, System Organ Class, and Treatment Group and percentages based on the number of subjects in the safety population within each treatment group. The total number of times an event occurred was recorded by Preferred Term, System Organ Class, and Treatment Group. Fisher's exact test was used for treatment comparison of event rates. SAP Section 11.2
Programming Statements	[SAS version 9.2] at14-5-02.sas

If multiple def:WhereclauseRef are described, each criteria is combined using "or" operator

Structure of Analysis Results Metadata

- <arm:ProgrammingCode>
  - 1. Actual Code using <arm:Code>

Program Code/File

```
<arm:ProgrammingCode Context=" SAS version 9.2 or later">
   <arm:Code>
      proc glm data = ADQSADAS;
         where EFFFL='Y' and ANL01FL='Y' and AVISIT='Week 24' and PARAMCD="ACTOT";
         class TRTPN SITEGR1;
         model CHG = TRTPN SITEGR1 BASE;
         means TRTPN;
         lsmeans TRTPN / OM STDERR PDIFF CL;
      run;
   </arm:Code>
</arm:ProgrammingCode>
```

Actual Code

Programming Statements	[SAS version 9.2]
	
	proc glm data = ADQSADAS;
	where EFFFL='Y' and ANL01FL='Y' and AVISIT='Week 24' and PARAMCD="ACTOT";
	class TRTPN SITEGR1;
	model CHG = TRTPN SITEGR1 BASE;
	means TRTPN;
	lsmeans TRTPN / OM STDERR PDIFF CL;
	run;

  - 2. Link to Program file using <def:DocumentRef>

Definition of external file

```
<def:leaf ID="LF.at14-5-02.sas" xlink:href="../programs/at14-5-02-sas.txt">
   <def:title>at14-5-02.sas</def:title>
</def:leaf>
...
<arm:ProgrammingCode Context=" SAS version 9.2 or later">
   <def:DocumentRef leafID="LF.at14-5-02.sas" />
</arm:ProgrammingCode>
```

Program Code/File

Link to external file

Considerations on Analysis Results Metadata

- **Creation** of ARM in Define.xml
  - Need basic knowledge for XML structure
    - XML schema, well-formed /valid XML, XML stylesheet
  - Difficult to create ARM automatically without standardization
    - Several tagset patterns
      - e.g. multiple results/datasets/variables, multiple Whereclause elements
    - Define.xml Generator (Fujitsu) can generate Define.xml including ARM with Excel metadata
- **Validation**
  - At this point (Jun., 2015), OpenCDISC cannot validate the ARM
  - Need schema validation
- **Legacy Data Conversion**
  - If TLFs in CSR are created using legacy data, there will be no analysis programs using ADaM datasets
    - Sponsors may re-create the analysis programs/results to include them as ARM in Define.xml
    - For the validation, sponsors may compare the results with TLFs in CSR.

Outline

- **Background**
  - Analysis Results Metadata for e-Data submission in Japan
- **Overview of Analysis Results Metadata**
  - Analysis Results Metadata in ADaM
  - Structure of Analysis Results Metadata
  - Considerations on Analysis Results Metadata
- **Creation of Analysis Results Metadata with SAS**
  - Excel Metadata File and SAS program to generate Analysis Results Metadata as a part of Define.xml

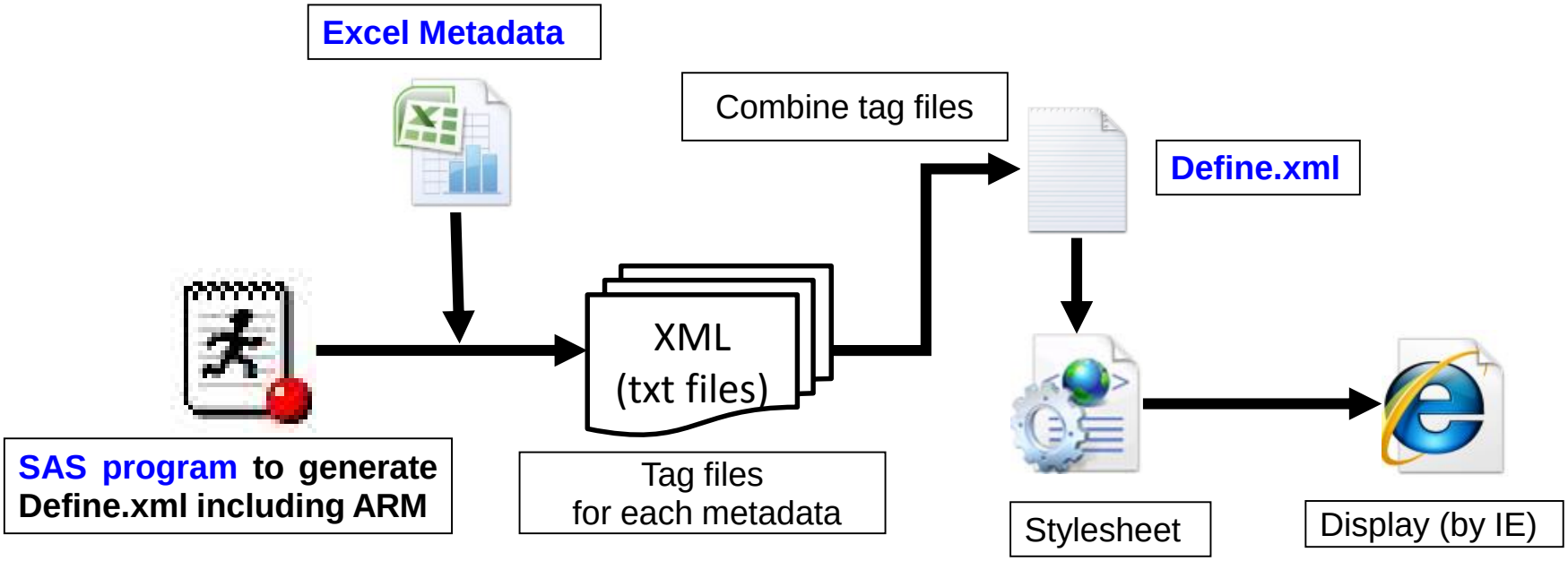
Creation of Analysis Results Metadata

- How to create Analysis Results Metadata
  - Tools
    - OpenCDISC, Define.xml Generator (Fujitsu) , Metadata Repository (MDR)
      - Latest Version of Define.xml Generator (Fujitsu) can generate ARM using Excel metadata and validate the Define.xml file using schema file
      - At this point (June 2015), OpenCDISC cannot generate and validate ARM
  - Programs
    - SAS or other programming languages
      - Read Excel metadata and generate using program
        - » Similar flow to the tools above (e.g. Define.xml Generator)
      - Need programming skill and understanding of the Define.xml structure including ARM

Creation of Analysis Results Metadata with SAS

- Excel metadata and SAS program to generate ARM
  - SAS program read the Excel Metadata and Generate the ADaM Define.xml including ARM

***Only the ARM generation part is covered in this presentation and the SAS program to generate the ARM is available on the web site**



Creation of Analysis Results Metadata with SAS

- Excel metadata for ARM (1)
  - Description of analyses and programming code or link to the program file

ID	Name	ResultDisplay	Analysis ID	Parameter OID	Reason	Purpose	Description	Documentation
Table_1.1	Table 1.1	Disease XXX Healing Rate at week 12 (FAS)	Table_1.1_1	ADEFF.PARAMCD	SPECIFIED IN SAP	PRIMARY OUTCOME MEASURE	Disease XXX Healing Rate at week 12 (FAS)	Wald test to compare the Disease XXX Healing Rate between treatment groups. FREQ procedure was used for this analysis.
Table_1.1	Table 1.1	Disease XXX Healing Rate at week 4, 8 (FAS)	Table_1.1_2	ADEFF.PARAMCD	SPECIFIED IN SAP	SECONDARY OUTCOME MEASURE	Disease XXX Healing Rate at week 4, 8 (FAS)	Calculate the Disease XXX Healing Rate at Week 4 and Week 8 using FREQ procedure
Table_2.1	Table 2.1	Summary of AE (Safety Analysis Set)	Table_2.1		SPECIFIED IN SAP	SAFETY OUTCOME MEASURE	Summary of AE (Safety Analysis Set)	Summary of TEAE

DocumentRef	ProgrammingCode	Version	ProgramID	href	Title
SAP	proc format ; value _TRTPF 1 = "Treat A" 2 = "Treat B" ; run ; proc freq data=ADEFF ; where FASFL = "Y" and AVISIT = "Week 12" and ANL01FL = "Y" ; table TRTPN*AVALC / riskdiff(equal var=sample) nocol nopercent ; format TRTPN _TRTPF. ; run ;	SAS version 9.2 or later			
SAP	proc format ; value _TRTPF 1 = "Treat A" 2 = "Treat B" ; run ; proc freq data=ADEFF ; where FASFL = "Y" and AVISIT in ("Week 4","Week 8") and ANL01FL = "Y" ; table TRTPN*AVALC / riskdiff(equal var=sample) nocol nopercent ; format TRTPN _TRTPF. ; run ;	SAS version 9.2 or later			
SAP		SAS version 9.2 or later	AEsas	AE.sas	AE.sas

Creation of Analysis Results Metadata with SAS

- Excel metadata for ARM (2)
  - Description of datasets, variables and selection criteria
    - Multiple datasets and variables can be specified

Analysis ID	WhereClauseOID	Analysis Variables	Dataset	Variable	Comparator	Value
Table_1.1_1	ADEFF.AVAL.HEAL	AVAL	ADSCOP	PARAMCD	EQ	SCHEAL
Table_1.1_1	ADEFF.AVAL.HEAL	AVAL	ADSCOP	FASFL	EQ	Y
Table_1.1_1	ADEFF.AVAL.HEAL	AVAL	ADSCOP	ANL01FL	EQ	Y
Table_1.1_1	ADEFF.AVAL.HEAL	AVAL	ADSCOP	AVISIT	EQ	Week 8
Table_1.1_2	ADEFF.AVAL.HEAL2	AVAL	ADSCOP	PARAMCD	EQ	SCHEAL
Table_1.1_2	ADEFF.AVAL.HEAL2	AVAL	ADSCOP	FASFL	EQ	Y
Table_1.1_2	ADEFF.AVAL.HEAL2	AVAL	ADSCOP	ANL01FL	EQ	Y
Table_1.1_2	ADEFF.AVAL.HEAL2	AVAL	ADSCOP	AVISIT	IN	Week2, Week 4
Table_2.1	ADSL.SAFFL		ADSL	SAFFL	EQ	Y
Table_2.1	ADAE.PT	AEDECOD,AEBODSYS	ADAE	AOCCPFL	EQ	Y
Table_2.1	ADAE.PT	AEDECOD,AEBODSYS	ADAE	SAFFL	EQ	Y

Creation of Analysis Results Metadata with SAS

- SAS program for ARM creation
  - Preprocessing and Read the Excel metadata using Excel libname engine
    - _MPATH and _CPATH are the full path for Excel metadata file and output text files (xml tagsets)

```
libname META excel "&_MPATH." SCAN_TEXTSIZE=YES ;
data RESULTS_METADATA ;
   set META.'Analysis Results Metadata$'n ;
   RESNO = _N_ ;
   if ID ne "" ;
run ;
data ARESM_WHR ;
   set META.'AResM Where$'n ;
   if WhereClauseOID ne "" ;
   WHRNO = _N_ ;
run ;

filename _AR      "&_CPATH.¥temp¥_AR.txt" ;
filename _ANPGMLF "&_CPATH.¥temp¥_ARPGM.txt" ;
filename _ARWHR   "&_CPATH.¥temp¥_ARWHR.txt" ;
proc sort data=RESULTS_METADATA ; by RESNO ID ; run ;
```

Metadata (1)

Metadata (2)

Output files:
1. Tagsets for main part of ARM
2. Link to Program files
3. WhereClause Definition, respectively

Creation of Analysis Results Metadata with SAS

- SAS program for ARM creation
  - Link to SAS programs

```
data _null_ ;
   set RESULTS_METADATA end=_EOF ;
   file _ANPGMLF ;
   if ProgramID ne "" then do ;
      put '      <def:leaf ID="LF.' ProgramID + (-1) '" xlink:href="' href + (-1) '">' ;
      put '        <def:title>' Title + (-1) '</def:title>' ;
      put '      </def:leaf>' ;
   end ;
run ;
```

Program ID and Path

File name

```
<def:leaf ID="LF.AEsas" xlink:href="AE.sas">
  <def:title>AE.sas</def:title>
</def:leaf>
```

Creation of Analysis Results Metadata with SAS

- SAS program for ARM creation
  - Generation of WhereClause Definition for ARM

```
proc sort data=ARESM_WHR ;
   by Analysis_ID WhereClauseOID WHRNO ;
run ;

data _null_ ;
   set ARESM_WHR end=_EOF ;
   by Analysis_ID WhereClauseOID WHRNO ;
   file _ARWHR ;

   if first.WhereClauseOID then put '      <def:WhereClauseDef OID="WC.' WhereClauseOID + (-1) '">' ;
   put '         <RangeCheck Comparator="' Comparator + (-1)
       '" SoftHard="Soft" def:ItemOID="IT.' Dataset + (-1) '.' Variable + (-1) '">' ;
   VAL_C = count(Value, ",") + 1 ;
   do I = 1 to VAL_C ;
      VAL = scan(Value, I, ",") ;
      put '            <CheckValue>' VAL + (-1) '</CheckValue>' ;
   end ;
   put '         </RangeCheck>' ;

   if last.WhereClauseOID then put '      </def:WhereClauseDef>' ;
run ;
```

Variables and Operator for WhereClause

For "IN" operator

Creation of Analysis Results Metadata with SAS

- SAS program for ARM creation
  - Generation of WhereClause Definition for ARM

```
<def:WhereClauseDef OID="WC.ADEFF.AVAL.HEAL">
  <RangeCheck Comparator="EQ" SoftHard="Soft" def:ItemOID="IT.ADEFF.PARAMCD">
    <CheckValue>HEAL</CheckValue>
  </RangeCheck>
  <RangeCheck Comparator="EQ" SoftHard="Soft" def:ItemOID="IT.ADEFF.FASFL">
    <CheckValue>Y</CheckValue>
  </RangeCheck>
  <RangeCheck Comparator="EQ" SoftHard="Soft" def:ItemOID="IT.ADEFF.ANL01FL">
    <CheckValue>Y</CheckValue>
  </RangeCheck>
  <RangeCheck Comparator="EQ" SoftHard="Soft" def:ItemOID="IT.ADEFF.AVISIT">
    <CheckValue>Week 12</CheckValue>
  </RangeCheck>
</def:WhereClauseDef>
....
```

Creation of Analysis Results Metadata with SAS

- SAS program for ARM creation
  - Main part of ARM

```
proc sort data=RESULTS_METADATA ;
   by Analysis_ID ;
run ;
data _AResM ;
   merge RESULTS_METADATA ARESM_WHR ;
   by Analysis_ID ;
run ;
proc sort data=_AResM ;
   where ID ne "" ;
   by ID Analysis_ID Dataset WhereClauseOID ;
run ;
data _null_;
   set _AResM end=_EOF ;
   by ID Analysis_ID Dataset WhereClauseOID ;
   file _AR lrecl=1000 ;

   ProgrammingCode = htmlencode(ProgrammingCode) ;

   if _N_ = 1 then put '      <arm:AnalysisResultDisplays>' ;
```

Marge the metadata (1) and (2)

Main part of ARM

Creation of Analysis Results Metadata with SAS

Start of ARM

By analysis results

Dataset, WhereClause and Analysis Variables

```
if first.ID then do ;
   put '        <arm:ResultDisplay OID="RD.' ID + (-1) '" ' 'Name="' Name + (-1) '">' ;
   put '          <Description>' ;
   put '            <TranslatedText xml:lang="en">' ResultDisplay + (-1) '</TranslatedText>' ;
   put '          </Description>' ;
end ;
if first.Analysis_ID then do ;
   put '          <arm:AnalysisResult OID="AR.'          Analysis_ID   + (-1) '"' ;
   put '                              ParameterOID="IT.' Parameter_OID + (-1) '"' ;
   put '                              AnalysisReason="'  Reason        + (-1) '"' ;
   put '                              AnalysisPurpose="' Purpose       + (-1) '">' ;
   put '            <Description>' ;
   put '              <TranslatedText xml:lang="en">' Description + (-1) '</TranslatedText>' ;
   put '            </Description>' ;
   put '            <arm:AnalysisDatasets>' ;
end ;
if first.Dataset then do ;
   put '              <arm:AnalysisDataset ItemGroupOID="IG.' Dataset + (-1) '" >' ;
   put '                <def:WhereClauseRef WhereClauseOID="WC.' WhereClauseOID   + (-1) '"/>' ;
   if Analysis_Variables ne "" then do ;
      _NVAL = count(Analysis_Variables,",") + 1 ;
      do I = 1 to _NVAL ;
         VARIABLE = strip(scan(Analysis_Variables,I,",")) ;
         put '                <arm:AnalysisVariable ItemOID="IT.' Dataset + (-1) '.' VARIABLE + (-1) '"/>' ;
      end ;
   end ;
   put '              </arm:AnalysisDataset>' ;
end ;
```

Creation of Analysis Results Metadata with SAS

```
if last.Analysis_ID then do ;
     put '            </arm:AnalysisDatasets>' ;
     put '            <arm:Documentation>' ;
     put '              <Description>' ;
     put '                <TranslatedText xml:lang="en">' Documentation + (-1) '</TranslatedText>' ;
     put '              </Description>' ;
     if DocumentRef ne "" then do ;
        put '               <def:DocumentRef leafID="LF.' DocumentRef + (-1) '">' ;
        put '               </def:DocumentRef>' ;
     end ;
     put '            </arm:Documentation>' ;
     if ProgrammingCode ne "" then do ;
        put '            <arm:ProgrammingCode Context="' Version + (-1) '">' ;
        put '              <arm:Code>' ;
        put ProgrammingCode ;
        put '              </arm:Code>' ;
        put '            </arm:ProgrammingCode>' ;
     end ;
     else if ProgramID ne "" then do ;
        put '            <arm:ProgrammingCode Context="' Version + (-1) '">' ;
        put '              <def:DocumentRef leafID="LF.' ProgramID + (-1) '"/>' ;
        put '            </arm:ProgrammingCode>' ;
     end ;
     put '          </arm:AnalysisResult>' ;
  end ;
```

Description and Documentation

Programming code or Link to program files

Creation of Analysis Results Metadata with SAS

- SAS program for ARM creation
  - Main part of ARM

```
<arm:AnalysisResultDisplays>
  <arm:ResultDisplay OID="RD.Table_1.1" Name="Table 1.1">
    <Description>
      <TranslatedText xml:lang="en">Disease XXX Healing Rate (FAS)</TranslatedText>
    </Description>
    <arm:AnalysisResult OID="AR.Table_1.1_1"
                        ParameterOID="IT.ADEFF.PARAMCD"
                        AnalysisReason="SPECIFIED IN SAP"
                        AnalysisPurpose="PRIMARY OUTCOME MEASURE">
      <Description>
        <TranslatedText xml:lang="en">Disease XXX Healing Rate at week 12 (FAS)</TranslatedText>
      </Description>
      <arm:AnalysisDatasets>
        <arm:AnalysisDataset ItemGroupOID="IG.ADEFF" >
          <def:WhereClauseRef WhereClauseOID="WC.ADEFF.AVAL.HEAL"/>
          <arm:AnalysisVariable ItemOID="IT.ADEFF.AVAL"/>
        </arm:AnalysisDataset>
      </arm:AnalysisDatasets>
.....
```

Creation of Analysis Results Metadata with SAS

- Output on a browser with a stylesheet

Analysis Results Metadata (Summary) for Study DRUG-XXX/STUDY-YYY

Table 1.1 Disease XXX Healing Rate (FAS)
- Disease XXX Healing Rate at week 12 (F
- Disease XXX Healing Rate at week 4, 8 (

Table 2.1 Summary of AE (Safety Analysis
- Summary of AE (Safety Analysis Set)

Analysis Results Metadata (Detail) for Study DRUG-XXX/STUDY-YYY

Table 1.1

Display	Disease XXX Healing Rate (FAS)
Analysis Result	Disease XXX Healing Rate at week 12 (FAS)
Analysis Parameter(s)	PARAMCD = "HEAL" (Healed/Unhealed of Disease XXX)
Analysis Variable (s)	AVAL (Analysis Value)
Analysis Reason	SPECIFIED IN SAP
Analysis Purpose	PRIMARY OUTCOME MEASURE
Data References (incl. Selection Criteria)	ADEFF [PARAMCD = "HEAL" and FASFL = "Y" and ANL01FL = "Y" and AVISIT = "Week 12"]
Documentation	Calculate the Disease XXX Healing Rate, proportion differance and their 95% CI between treatment groups and perform the Wald test for the comparison of treatment at Week 12. FREQ procedure was used for this analysis. SAP
Programming Statements	[SAS version 9.2 or later]

```
proc format ;
   value _TRTPF 1 = "Treat A" 2 = "Treat B" ;
run ;
proc freq data=ADEFF ;
   where FASFL = "Y" and AVISIT = "Week 12" and ANL01FL = "Y" ;
   table TRTPN*AVALC / riskdiff(equal var=sample) nocol nopercent ;
   format TRTPN _TRTPF. ;
```

Summary

- ARM in the e-Data submission package
  - PMDA recommends to submit the ARM as a part of ADaM Define.xml
    - Reviewers can understand how the analysis results are created
  - Need standardization to some extent
    - Several patterns (e.g. multiple datasets, variables, selection criteria, programming code)
  - Validation issue
    - OpenCDISC (V2.0) cannot validate ARM
  - Need re-creation/validation of programs/results for Legacy data conversion
  - Targeted analyses to be included in the ARM
    - Need the discussion/agreement with PMDA
- SAS is one of good tools for ARM generation
  - Need the standardization of input parameters (e.g. Excel metadata)
  - Need the basic knowledge of XML and Define.xml structure, but the programming code is not so complicated

References

- PMDA HP (e-Data submission)
  https://www.pmda.go.jp/review-services/drug-reviews/about-reviews/p-drugs/0003.html
- CDISC Define-XML Specification Version 2.0
  http://www.cdisc.org/define-xml
- CDISC ADaM (ADaM, ADaM IG and ARM are available)
  http://www.cdisc.org/adam
- tsClinical Define.xml Generator (Fujitsu)
  http://www.fujitsu.com/jp/solutions/industry/life-sciences/cdisctool/
- OpenCDISC
  http://www.opencdisc.org/
- Yohei Takanami (2013). Simple Tool for Creating ADaM Define.xml for Statisticians in Pharmaceutical Companies Using SAS and HTML Application with Excel Metadata File. CDISC 2013 Japan Interchange
- Yohei Takanami, Nobuo Funao (2012, Kohgakusha). Statistical Analysis Software - SAS.
- Chris Holland, Jack Shostak (2012). Implementing CDISC Using SAS: An End-to-End Guide. Sas Inst

Thank you for your attention