

***Bridging the gap between
ODM and SDTM:
Tools for creating
SDTM datasets from ODM data***

Jozef Aerts
XML4Pharma



The problem

- How can SDTM data sets and define.xml be created from operational data (ODM)
 - In a very-user friendly way (drag-and-drop & wizards)
 - Without needing to write SAS macros
 - And create and populate an SDTM database
-
-

The history

- Problem and use cases defined by TMF e.V.
 - The German Telematics Platform for Networks in Medical Research
- Contest started in 2005
- Contract won by XML4Pharma in 2006
- SDTMWandler and SDTM-ETL



Key concepts and features

- **define.xml template** containing standard SDTM variables for all subject-related domains
 - Containing all published **CDISC Controlled Terminology** as define.xml CodeLists
 - **ODM-SDTM mapping** stored within define.xml
 - Mapping scripts transformed into XSLT
 - Transformation execution into **SDTM-XML** and **SAS Transport**
 - SQL statement generation for creation and population of **SDTM database**
 - **Near-submission-ready define.xml** file
-
-

Workflow

- Load **ODM metadata**
 - Validate ODM
 - Load **define.xml template** and instantiate to study-specific **SDTM domains**
 - **Generate mappings** by
 - Drag and drop
 - Mapping wizards
 - Auto-generation of mapping scripts
 - Test mappings on clinical data
 - **Execute** mappings to generate **SDTM data sets**
 - Generate SQL for generation and population of **SDTM database**
-
-

ODM

- Study
 - GlobalVariables
 - BasicDefinitions
 - MetaDataVersion : Version 1.1.0
 - Protocol
 - StudyEventDef : Pre-treatment
 - StudyEventDef : Post-treatment
 - FormDef : Visit Form
 - FormDef : Adverse Events
 - ItemGroupDef : Common
 - ItemGroupDef : Adverse Events
 - ItemDef : Therapeutic Area
 - ItemDef : Protocol Number
 - ItemDef : Country
 - ItemDef : Record status, 5 levels, internal use
 - ItemDef : Line Number
 - ItemDef : Conmed Indication
 - ItemDef : Start Month - Enter Two Digits 01-12
 - ItemDef : Start Day - Enter Two Digits 01-31
 - ItemDef : Start Year - Enter Four Digit Year
 - ItemDef : Derived Start Date
 - ItemDef : Stop Month - Enter Two Digits 01-12
 - ItemDef : Stop Day - Enter Two Digits 01-31
 - ItemDef : Stop Year - Enter Four Digit Year
 - ItemDef : Derived Stop Date
 - ItemDef : Severity
 - ItemDef : Relationship to study drug
 - ItemDef : Outcome
 - ItemDef : Actions taken re study drug
 - ItemDef : Actions taken, other
 - FormDef : Concom Meds
 - FormDef : Physical Exam

Domains (ItemGroups)					
Variable	Variable	Variable	Variable	Variable	V.
SUBJID	DM.RFSTDTC	DM.RFENDTC	DM.SITEID	DM.INVID	DM.INV
TS.TSPARMCD	TS.TSPARM	TS.TSVAL			
SE.SESEQ	SE.ETCD	SE.ELEMENT	SE.SESTDTC	SE.SEENDTC	SE.SE
SV.VISIT	SV.VISITNUM	SV.VISITDY	SV.SVSTDTC	SV.SVENDTC	SV.SV
EX.EXSEQ	EX.EXGRPID	EX.EXSPID	EX.EXTRT	EX.EXCAT	EX.EX
CM.CMSEQ	CM.CMGRPID	CM.CMSPID	CM.CMTRT	CM.CMMODIFY	CM.CM
SU.SUSEQ	SU.SUGRPID	SU.SUSPID	SU.SUTRT	SU.SUMODIFY	SU.SU
AE.AESEQ	AE.AEGRPID	AE.AEREFID	AE.AESPID	AE.AETERM	AE.AE
DS.DSSEQ	DS.DSGRPID	DS.DSREFID	DS.DSSPID	DS.DSTERM	DS.DS
DV.DVSEQ	DV.DVCAT	DV.DVTERM	DV.DVDECOD	DV.EPOCH	DV.DV
MH.MHSEQ	MH.MHGRPID	MH.MHREFID	MH.MHSPID	MH.MHTERM	MH.MH
EG.EGSEQ	EG.EGGRPID	EG.EGREFID	EG.EGSPID	EG.EGTESTCD	EG.EG
IE.IESEQ	IE.IESPID	IE.IETESTCD	IE.IETEST	IE.IECAT	IE.IE
LB.LBSEQ	LB.LBGRPID	LB.LBREFID	LB.LBSPID	LB.LBTESTCD	LB.LB
PE.PESEQ	PE.PEGRPID	PE.PESPID	PE.PETESTCD	PE.PETEST	PE.PE
SC.SCSEQ	SC.SCGRPID	SC.SCSPID	SC.SCTESTCD	SC.SCTEST	SC.SC
VS.VSSEQ	VS.VSGRPID	VS.VSSPID	VS.VSTESTCD	VS.VSTEST	VS.VS
USUBJID	CO.COSEQ	CO.IDVAR	CO.IDVARVAL	CO.COREF	CO.CO
QS.QSSEQ	QS.QSGRPID	QS.QSSPID	QS.QSTESTCD	QS.QSTEST	QS.QS
DA.DASEQ	DA.DAGRPID	DA.DASPID	DA.DAREFID	DA.DATESTCD	DA.DA
PC.PCSEQ	PC.PCGRPID	PC.PCREFID	PC.PCSPID	PC.PCBATID	PC.PC
PP.PPSEQ	PP.PPGRPID	PP.PPTESTCD	PP.PPTEST	PP.PPCAT	PP.PP
MB.MBSEQ	MB.MBGRPID	MB.MBREFID	MB.MBSPID	MB.MBTESTCD	MB.MB
MS.MSSEQ	MS.MSGRPID	MS.MSREFID	MS.MSSPID	MS.MSTESTCD	MS.MS
IDVAR	IDVARVAL	RELTYPE	RELID		
IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG
SUBJID	DM.RFSTDTC	DM.RFENDTC	DM.SITEID	DM.INVID	DM.INV
QS.QSSEQ	QS.QSGRPID	QS.QSSPID	QS.QSTESTCD	QS.QSTEST	QS.QS
SV.VISIT	SV.VISITNUM	SV.VISITDY	SV.SVSTDTC	SV.SVENDTC	SV.SV
PE.PESEQ	PE.PEGRPID	PE.PESPID	PE.PETESTCD	PE.PETEST	PE.PE
AE.AESEQ	AE.AEGRPID	AE.AEREFID	AE.AESPID	AE.AETERM	AE.AE

AE.AETERM
 Mandatory: Yes
 OrderNumber: 8
 ItemDef Name: AETERM
 Data type: text
 Length: 80
 Define label: Reported Term for the Adverse Event

\$AE.AETERM = /*/FORM.AE/IG.AE/IT.AETERM/@Value

Suggesting ODM items for SDTM Variables (case 1:1)

- 📁 FormDef : Adverse Events
 - 🔑 📁 ItemGroupDef : Common
 - 🔑 📁 ItemGroupDef : Adverse Events
 - ● ItemDef : Therapeutic Area
 - ● ItemDef : Protocol Number
 - ● ItemDef : Country
 - ● ItemDef : Record status, 5 levels, internal use
 - ● ItemDef : Line Number
 - 🛠 ItemDef : Conmed Indication
 - ● ItemDef : Start Month - Enter Two Digits 01-12
 - ● ItemDef : Start Day - Enter Two Digits 01-31
 - ● ItemDef : Start Year - Enter Four Digit Year

Auto-generation of mapping scripts: Codelist mapping

CodeList mapping between ODM "AE Severity" and SDTM "Severity/Intensity S..."

ODM CodeList Item	SDTM CodeList Item
1 - Mild	MILD - MILD
2 - Moderate	MILD - MILD
3 - Severe	MILD - MILD
4 - Life Threatening	MODERATE - MODERATE
MISSING/INVALID VALUE	SEVERE - SEVERE
	MISSING VALUE
	NULL

The generated mapping script:

The Transformation Script

```
if ($ODMVALUE == 1) {  
  $AE.AESEV = 'MILD';  
}  
elseif ($ODMVALUE == 2) {  
  $AE.AESEV = 'MODERATE';  
}  
elseif ($ODMVALUE == 3) {  
  $AE.AESEV = 'SEVERE';  
}  
elseif ($ODMVALUE == 4) {  
  $AE.AESEV = 'SEVERE';  
}
```

Scripting Language Functions

+	-	*	/
xpath	usubjid	investigator	site
substring	substring-before	substring-after	concat

Easy-to-learn scripting language allowing complicated mappings

The Transformation Script

```
$START = $AE.AESTDTC;  
$END = $AE.AEENDTC;  
if(string-length($START) > 0 and string-length($END) > 0) {  
    $DIFF = datediff($AE.AEENDTC,$AE.AESTDTC);  
} else {  
    $DIFF = '';  
}  
if($DIFF != '') {  
    $AE.AEDUR = concat('P',$DIFF,'D');  
} else {  
    $AE.AEDUR = '';  
}
```

Full implementation of SDTM-IG

SDTM CDISC Note for Variable AEENRF



Identifies the end of the event as being BEFORE, DURING, DURING/AFTER or AFTER the sponsor-defined reference period. The sponsor-defined reference period is a continuous period of time defined by a discrete starting point and a discrete ending point. Typically, this period is defined by the start (RFSTDTC) and end (RFENDTC) of the trial. Sponsors should define the reference period in the study metadata. Events that are ongoing at the end of the reference period should have a value of AFTER for this variable. If information such as "PRIOR", "ONGOING", or "CONTINUING" was collected, this information should be translated into AEENRF.

Execution of mapping: SDTM Records and SAS Transport files

Execute Transformation (XSLT) Code

ODM file with clinical data:
C:\SDTM-ETL\TestFiles\MyStudy_ClinicalData_only.xml **Browse...**

ODM Output file (SDTM Results):
C:\SDTM-ETL\TestFiles\test.xml **Browse...**

Split records > 200 characters to SUPP-- records

View Result SDTM tables

Save result SDTM tables as SAS XPORT files

SAS XPORT files directory:
C:\SDTM-ETL\TestFiles

Add location of SAS XPORT files

Store link as relative path

Messages and error messages

AESTARTDATE = 2006-04-10
First visit date = 2006-04-01
AE end date = 2006-04-14
FIRSTVISITDATE = 2006-04-01
AESTARTDATE = 2006-04-11
First visit date = 2006-04-02
AE end date = 2006-04-15
FIRSTVISITDATE = 2006-04-01
AESTARTDATE = 2006-04-22
First visit date = 2006-04-04

SDTM Tables

MyStudy:DM	MyStudy:QS	MyStudy:SV	MyStudy:PE	MyStudy:AE	MyStudy:SUPPPE		
SUBJID	PE.PESEQ	PE.PETESTCD	PE.PETEST	PE.PEORRES	PE.PESTRESC	PE.PE	
	1	1	Head, Neck and Th...	Normal			
	2	2	Eyes, Ears, Nose a...	Normal			
	3	3	Chest	Normal			
	4	4	Lungs	Abnormal	MILD WHEEZING		
	5	5	Heart	Abnormal	TACHYCARDIA		
	6	6	Lymph Nodes	Abnormal	SLIGHTLY ENLAR...		
	7	7	Abdomen	Normal			
	8	8	Anorectal	Not Done			NOT DONE
	9	9	Genitalia	Normal			
	10	10	Skin	Abnormal	PET		

Execution of mapping: SDTM Records and SAS Transport files

SAS System Viewer - [PE.xpt]

File Edit View Window Help

	STUDYID	DOMAIN	USUBJID	PESEQ	PETESTCD	PETEST	PEORRES	PESTRESC	PESTAT	PEREASND
1	MyStudy	PE	001	1	1	Head, Nec	Normal			
2	MyStudy	PE	001	2	2	Eyes, Ear	Normal			
3	MyStudy	PE	001	3	3	Chest	Normal			
4	MyStudy	PE	001	4	4	Lungs	Abnormal	MILD WHEEZING		
5	MyStudy	PE	001	5	5	Heart	Abnormal	TACHYCARDIA		
6	MyStudy	PE	001	6	6	Lymph Nod	Abnormal	SLIGHTLY ENLARGED-NC		
7	MyStudy	PE	001	7	7	Abdomen	Normal			
8	MyStudy	PE	001	8	8	Anorectal	Not Done		NOT DONE	The reason tha
9	MyStudy	PE	001	9	9	Genitalia	Normal			
10	MyStudy	PE	001	10	10	Skin	Abnormal	PET		
11	MyStudy	PE	001	11	11	Musculosk	Normal			
12	MyStudy	PE	001	12	12	Neurologi	Normal			
13	MyStudy	PE	001	13	13	Other	Normal			
14	MyStudy	PE	002	1	1	Head, Nec	Normal			
15	MyStudy	PE	002	2	2	Eyes, Ear	Normal			
16	MyStudy	PE	002	3	3	Chest	Normal			
17	MyStudy	PE	002	4	4	Lungs	Abnormal	MILD WHEEZING		
18	MyStudy	PE	002	5	5	Heart	Abnormal	TACHYCARDIA		
19	MyStudy	PE	002	6	6	Lymph Nod	Abnormal	SLIGHTLY ENLARGED -		

Other features

- Export of near-submission-ready define.xml
 - Tabular HTML view of underlying define.xml
 - Automated generation of SQL to generate SDTM database
 - Automated generation of SQL to populate SDTM database with SDTM records
-
-

Conclusions

- Mapping made very user-friendly
 - But you still need to know your study
 - SDTM annotation in ODM helps !
 - **Domain attribute in ItemGroupDef**
 - **SDSVarName in ItemDef**
 - But you still need to understand SDTM
 - “CDISC Notes” implemented in software
 - Amount of training required: 1 day
-
-

Thank you !

