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 <u>and</u> 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

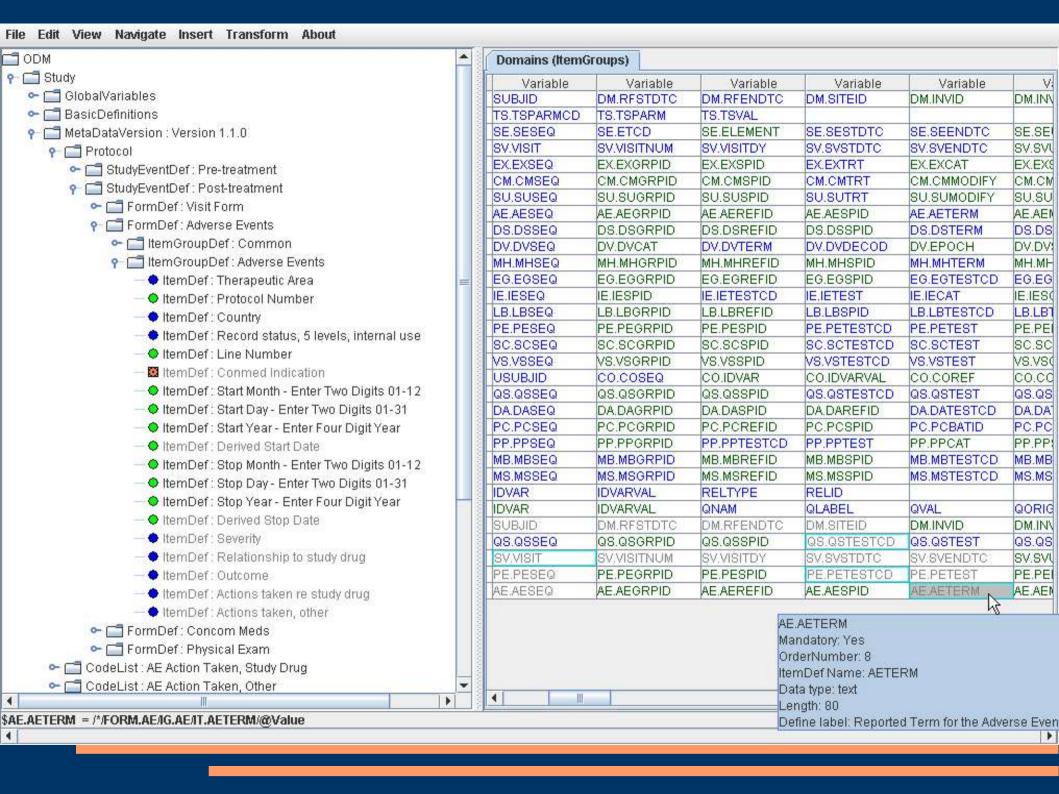
- Problem and use cases defined by TMF e.V.
 - The German Telematics Platform for Networks in Medical Research
- Contest started in 2005
- Contract won by XML4Pharma in 2006
- SDTMWandler and SDTM-ETL

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 studyspecific 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



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

FormDef: Adverse Events 🛅 ltemGroupDef : Common ItemGroupDef : Adverse Events ItemDef: Therapeutic Area ItemDef: Protocol Number — • ItemDef : Country ItemDef: Record status, 5 levels, internal use. - 💚 ltemDef : 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	T
3 - Severe	MILD - MILD	l l
4 - Life Threatening MISSING/INVALID VALUE	MODERATE - MODERATE SEVERE - SEVERE MISSING VALUE NULL	

The generated mapping script:

```
The Transformation Script-
if ($0DMVALUE == 1) {
   $AE.AESEV = 'MILD':
 } elsif ($ODMVALUE == 2) 🤾
   $AE.AESEV = 'MODERATE';
 } elsif ($ODMVALUE == 3) 🤾
   $AE.AESEV = 'SEVERE';
 } elsif ($ODMVALUE == 4) 🤾
   $AE.AESEV = 'SEVERE';
                                           Ш
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 = 11;
```

Full implementation of SDTM-IG

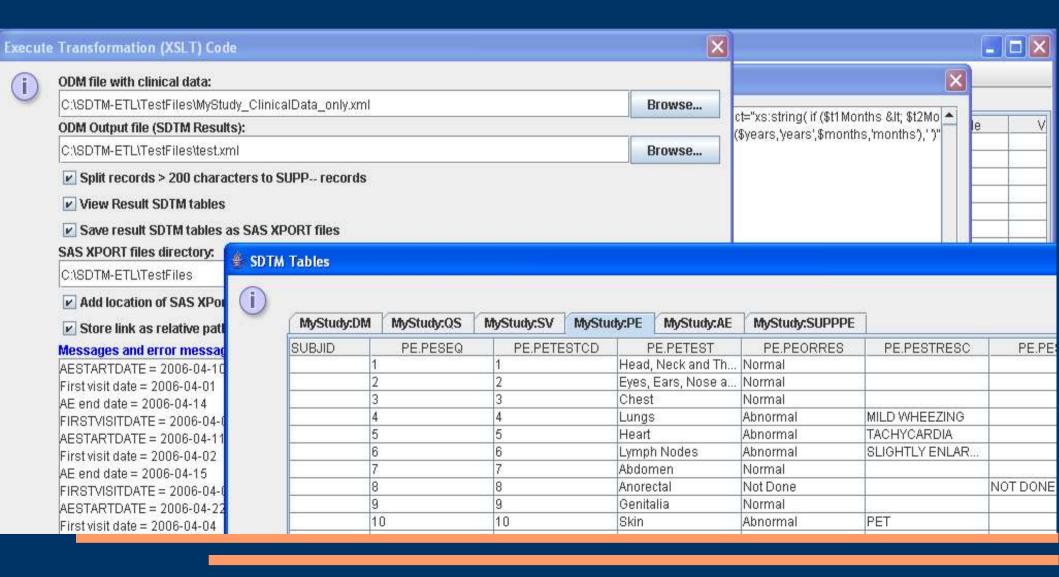
SDTM CDISC Note for Variable AEENRF



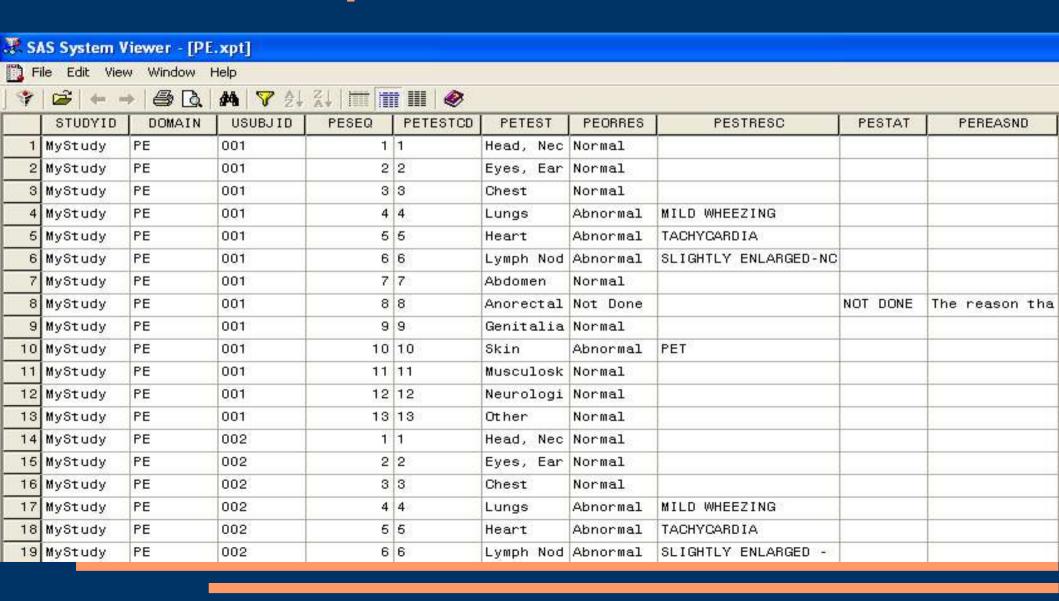


Identifies the end of the event as being BEFORE, DURING, DURING/AFTER or AFTER the sponsor-defined reference period. The sponsor-defined reference period is a continuous period of time defined by a discrete starting point and a discrete ending point. Typically, this period is defined by the start (RFSTDTC) 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



Execution of mapping: SDTM Records and SAS Transport files



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!

