



Importing OpenClinica Data in SAS: It's Possible! — Our Experiences



Jules van der Zalm, OCS Consulting, the Netherlands



In this presentation...

- ▶ About OCS Consulting, our customer and the study
- ▶ OpenClinica and ODM
- ▶ Importing OpenClinica data in SAS
- ▶ Challenges we faced
- ▶ Conclusion

About OCS



- ▶ OCS Consulting
 - ▶ IT consultancy specialised in Business Intelligence and Data Analytics
 - ▶ Providing SAS consultancy since 1998
 - ▶ Relationship with customer since 2011

- ▶ Customer and study clinical trial
 - ▶ Manufacturer of medical devices
 - ▶ Cooperation with universities that use OpenClinica

OpenClinica and the ODM model



- ▶ Open source
- ▶ Free Community Edition, fully supported and licensed Enterprise Edition
- ▶ Data can be export to various formats such as SPSS and ODM XML

OpenClinica and the ODM model



- ▶ *“The Operational Data Model (ODM) is designed to facilitate the archive and interchange of the metadata and data for clinical research, its power being fully unleashed when data are collected from multiple sources.”*

- ▶ Most important metadata entities are:
 - ▶ Study
 - ▶ StudyEvent
 - ▶ FormDef
 - ▶ ItemGroupDef
 - ▶ ItemDef



Importing data in SAS

- ▶ PROC CDISC for subject data
- ▶ SAS XML Mapper and XMLV2 libname for meta data
- ▶ Bespoke SAS programs to generate analysis data sets



#1: One site per data file

- ▶ One XML file containing all the data
- ▶ StudyOID contains site reference
- ▶ One export per site
- ▶ Data is read and combined into “raw data sets”
- ▶ Risk: possible differences in meta data between sites

```
/* Perform a comparison of two datasets. If they are exactly  
the same, the result dataset _TMP_ will contains 0  
observations. */  
proc compare base=xmlmeta1.&dsname.  
             compare=xmlmeta&i..&dsname.  
             out =work._tmp_  
             noprint outnoequal;  
  
run;
```



#2: PROC CDISC and ODM v1.3

- ▶ PROC CDISC supports only ODM v1.2
- ▶ ODM v1.3 has parent-child relationship information

☐ Myocardial infarction, specify
Date of most recent Myocardial Infarction (MI) Day Month Year

- ▶ Access to SQL database with ODM v1.3 data
- ▶ Used ODM v1.2 for subject data, v1.3 for metadata



#3: Non-unique SASFieldNames

```
<ItemDef OID="I_EVENT_BASESUBCRT" Name="BaseSubCRT" DataType="integer" Length="1" SASFieldName="BaseSubC" Comment="Subject is indi
  <Question>
    <TranslatedText>
      Subject is indicated for CRT-D device according to current applicable ESC/AHA guidelines:
    </TranslatedText>
  </Question>
  <CodeListRef CodeListOID="CL_7378"/>
</ItemDef>
<ItemDef OID="I_EVENT_BASESUBCONTI" Name="BaseSubcont" DataType="integer" Length="1" SASFieldName="BaseSubc" Comment="Subject is on
  <Question>
    <TranslatedText>
      Subject is on continuous or uninterrupted infusion (inotropic) therapy for heart failure ( 2 stable infusions per week):
    </TranslatedText>
  </Question>
  <CodeListRef CodeListOID="CL_7378"/>
</ItemDef>
```

SASFieldName="BaseSubC"

SASFieldName="BaseSubc"



#3: Non-unique SASFieldNames

```
<ItemDef OID="I_EVENT_BASESUBCRT" Name="BaseSubCRT" DataType="integer" Length="1" SASFieldName="BaseSubC" Comment="Subject is indi
  <Question>
    <TranslatedText>
      Subject is indicated for CRT-D device according to current applicable ESC/AHA guidelines:
    </TranslatedText>
  </Question>
  <CodeListRef CodeListOID="CL_7378"/>
</ItemDef>
<ItemDef OID="I_EVENT_BASESUBCONT" Name="BaseSubcont" DataType="integer" Length="1" SASFieldName="BaseSubc" Comment="Subject is on
  <Question>
    <TranslatedText>
      Subject is on continuous or uninterrupted infusion (inotropic) therapy for heart failure ( 2 stable infusions per week):
    </TranslatedText>
  </Question>
  <CodeListRef CodeListOID="CL_7378"/>
</ItemDef>
```

	A	B
1	Original OID	New variable name
28	I_EVENT_BASESUBCRT	basesubc
29	I_EVENT_BASESUBCONT	basesub1



#3: Non-unique SASFieldNames

```
<ItemDef OID="I_EVENT_BASESUBCRT" Name="BaseSubCRT" DataType="integer" Length="1" SASFieldName="basesubc" Comment="Subject is indi
  <Question>
    <TranslatedText>
      Subject is indicated for CRT-D device according to current applicable ESC/AHA guidelines:
    </TranslatedText>
  </Question>
  <CodeListRef CodeListOID="CL_7378"/>
</ItemDef>
<ItemDef OID="I_EVENT_BASESUBCONT" Name="BaseSubcont" DataType="integer" Length="1" SASFieldName="basesub1" Comment="Subject is on
  <Question>
    <TranslatedText>
      Subject is on continuous or uninterrupted infusion (inotropic) therapy for heart failure ( 2 stable infusions per week):
    </TranslatedText>
  </Question>
  <CodeListRef CodeListOID="CL_7378"/>
</ItemDef>
```

	A	B
1	Original OID	New variable name
28	I_EVENT_BASESUBCRT	basesubc
29	I_EVENT_BASESUBCONT	basesub1



#4: Changing CRF's

- ▶ CRF's changed during programming
- ▶ Many hours spent re-programming and re-validating
- ▶ Essential to finalize CRF design prior to study start



#5: Changing SASFieldNames

- ▶ Out of approximately 500 variables, nearly 100 changed
- ▶ Reason? No one knows
- ▶ Two solutions
 - ▶ Use new SASFieldNames
 - ▶ Don't use SASFieldNames at all
- ▶ Updating programs using a helper-macro
- ▶ Alternative approach: automatically update programs



Database design

- ▶ One data set per ItemGroup
- ▶ Customer creates data sets by visit instead of domain
- ▶ Still more data sets than visits
- ▶ Close cooperation in the set up phase

Conclusion



- ▶ Exciting exercise from a programmers perspective
- ▶ Disappointed
- ▶ Generic, re-usable approach
- ▶ Hope to hear from more experienced users

Questions

