GUI Method of Mapping SDTM from Raw Data

Elisa Natale
Martina Garlet

CROS NT, Verona, Italy

cros nt

Introduction

- Purpose

  Improving and speeding up the **mapping of SDTM from raw data**, as the standard CDASH are not frequently used

- How to reach the purpose

  By using **Excel** as GUI support of SAS programming

- Why Excel

  More user-friendly interface, easier and faster to compile (expecially for a junior programmer) and gives less opportunity for programming errors

Overall Process

1. Complete the mapping information for every SDTM domain into an Excel tool
2. Writing and run a SAS program for each domain:

   2.1 Recall of macro %createSDTM(domain=);

   2.2 Eventually fix the temporary dataset “union”;

   2.3 Recall of the final macro %createDATASET.

The Excel tool

Section A: Domain reference

Name of the domain (ex: QS, QS AE, all (key word))

Section B: Variables to create information

Name of the SDTM variables (ex: *STUDYID, DOMAIN...*)

Their format (*Char* or *Num*)

Section C: Input information

Name of the dataset (ex: *libname.dsname* or *dsname*)

Selection (ex: *where CRIT='Y';*)

Section D: Mapping codes

List of variables to transpose (When variable is 'TRANSPOSE VAR')

Otherwise: Algorithms, Variables, String or just 'same' (key word)

The Excel tool

	A	B	C	D	E	F	G
1	Dataset	Variable	Type				
2	all	STUDYID	Char	"&Protocol_Number."			
3	QS AE CM	USUBJID	Char	USUBJID=compress("&Protocol_Number."||"-"||ecode);			
4	VS DM	USUBJID	Char	USUBJID=compress("&Protocol_Number."||"-"||subj);			
5	QS	Data		data.visit1	data.visit5	data.form	data.form
6	QS	Subgroup				where prteevst=1;	where prteevst=2;
7	QS	TRANSPOSE VAR	Num	v1prteeP,v1prteeS, v1prteeU	v5prteeP v5prteeS	prtee11--prtee2B4	prtee11--prtee2B4
8	QS	DOMAIN	Char	"QS"	same	same	same
9	QS	QSTESTCD	Char	qstestcd=put(_name_,$qstcd.);	same	qstestcd=put(_name	same
10	QS	QSTEST	Char	qstest=put(QSTESTCD,$qst.);	same	qstest=put(QSTESTC	same
11	QS	QSCAT	Char	"PRTEE QUESTIONNAIRE"	same	same	same
12	QS	QSSCAT	Char	"PRTEE SCORE"	same	"PRTEE ITEM"	same
13	QS	QSORRES	Char	qsorres=strip(col1);	same	same	same
14	QS	QSSTRESC	Char	qsstresc=strip(col1);	same	same	same
15	QS	QSBLFL	Char	"Y"	""	"Y"	""
16	QS	VISITNUM	Num	"1"	"5"	"1"	"5"
17	QS	VISIT	Char	"VISIT 1"	"VISIT 5"	"VISIT 1"	"VISIT 5"
18	QS	QSDTC	Char	%Crea_date(d=v1dt_dd,m=v1dt	%Crea_date(d=v5dt	%Crea_date(d=prtee	%Crea_date(d=prte
19	VS	Data		input1	input1	input2	
20	VS	Subgroup		where formName ne "Neonatal	where formName eq "Neonatal Vital Signs" and vsperf in ("Y'		
21	VS	TRANSPOSE VAR	Num	hr, syssu, diasu, sysst ,diast, wt,	hrn sys dia rr	weigh headcir apgar	
22	VS	USUBJID	Char	SITEID1=substr(siteID,1,4); usub	same	same	
23	VS	DOMAIN	Char	"VS"	same	same	
24	VS	VSSCAT	Char	""	""	"NEONATAL ASSESSMENTS"	
25	VS	VSTESTCD	Char	VSTESTCD=put(_NAME_,$VSTES	same	same	

A B C D D

The Program

Example of a program:

```
%createSDTM(domain=SC);
/*Optional AD HOC steps*/
%createDATASET (datashell=SC, dataset=Union);
```

Variable xxSEQ:

Normally the program computes it automatically. If the variable is not needed just edit the call of the macro this way:

```
%createSDTM(domain=DM, seq=N);
```

The Program

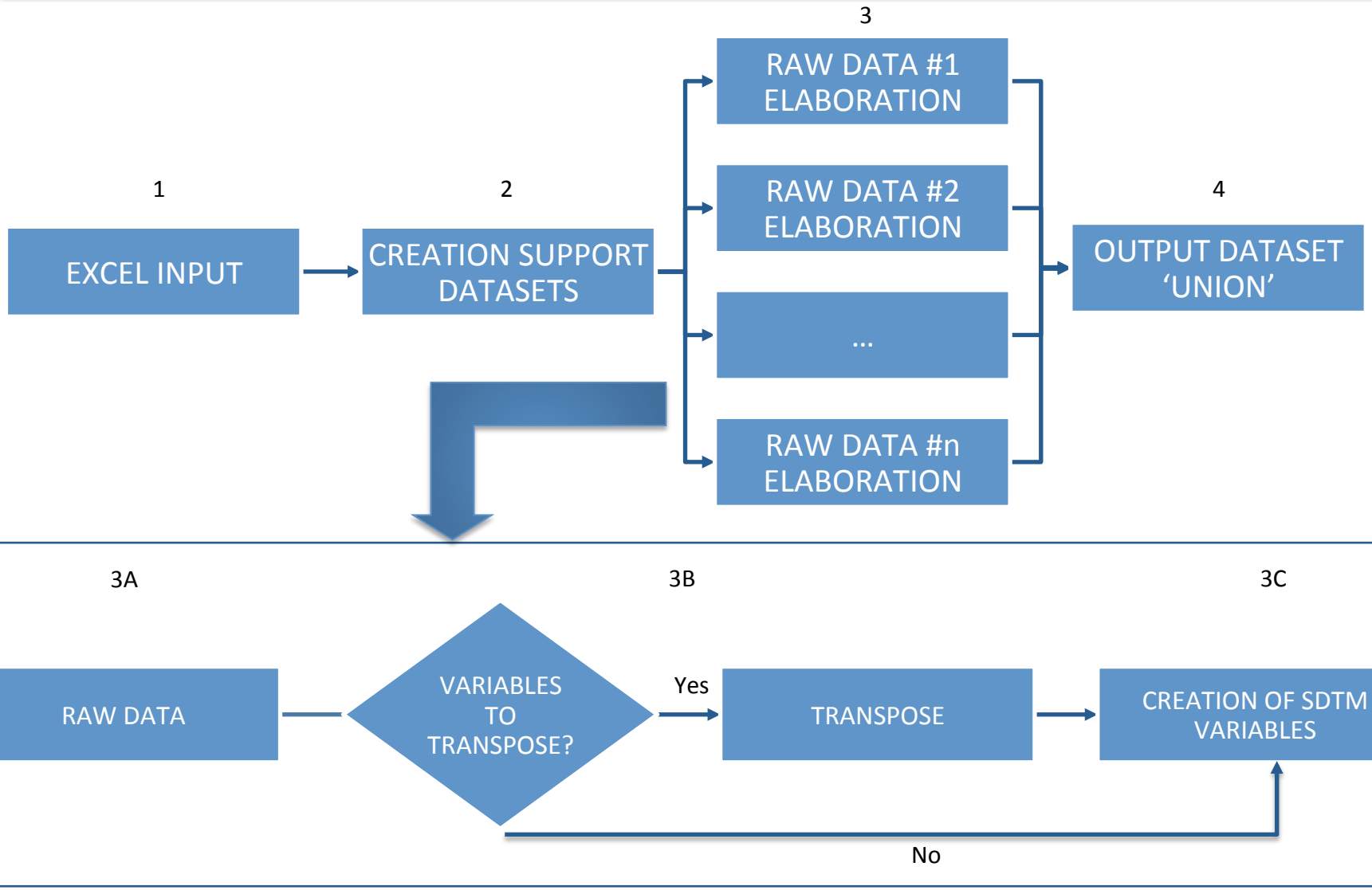
```
 Run on      : 03FEB17[12:51]
 Sas Version: 9.4


0 Errors
Dataset Union created from 8 input datasets

NOTE: There were 1 observations read from the data set TEMP.ERRSUM.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds
```

If no errors, then the draft dataset can be checked and finalized

The Macro %createSDTM

1. The import of the Excel dataset;
2. The creation of support datasets;
3. Work on each of the raw data separately:
   - 3a. The import into temporary library of the raw dataset;
   - 3b. The transpose of the dataset, if necessary;
   - 3c. The creation of the SDTM variables;
4. The union of all the modified input dataset into one, ‘Union’.

The Macro %createSDTM



Limits

1. It doesn't work with Trial Domains and Supp Domains
2. Selection of records (Ex: IE domain)
3. The presence of variables in the raw data with the same name of SDTM variables to create
4. Hard to add new records

Solutions

1. Not required
2. Editing the dataset Union
3. The program automatically renames the variable in the raw data by adding an underscore before it
4. Uncommon for SDTM, possible with ad hoc space, maybe not to use this tool

Conclusion

- This process covers many situations
  - Many input datasets
  - Both structure of input datasets (either if they need to be transposed or not)
  - Every mapping way is possible (algorithms, string...)
- More efficient
  - Less code to write → less likely to make mistakes
  - Easier for a Junior programmer (more about SDTM theory)


Questions