

Exploring the Possibility of Automating the SDTM Subject Visits (SV) Domain

Dony Sabu

Associate Lead - Clinical Data Analytics, Genpro Research

Dony.sabu@genproresearch.com

Akhil Aniyar

Statistical Programmer II, Genpro Research

Akhil.aniyar@genproresearch.com

Limna Salim

India Head – Biometrics & Operations, Genpro Research

Limna.salim@genproresearch.com

SV DOMAIN

- It is a required domain as part of SDTM submission.
- It contains information for each visit for a subject including start and end date for a particular visit.
- It takes input from TV dataset to map the VISIT and VISITNUM information.
- Derivation is consistent across studies.

SV MACRO WORKFLOW

- Finding required raw datasets and appropriate variables.
- Obtaining earliest and latest dates (SVSTDTC/SVENDTC) corresponding to each visits.
- Directly maps the corresponding VISITNUM and VISITDY from TV domain for scheduled visits.
- Handles the final hurdle, that is the derivation of VISITNUM for Unplanned visits.
- Derives the SVSTDY and SVENDY variables corresponding to SVSTDTC and SVENDTC.

SV MACRO PARAMETERS

Macro Parameter	Description	Examples
Inplib (Required)	Input library which contains of input/raw datasets necessary for SV domain generation.	RAW or EDC
Subjid (Required)	Variable name in which subject identifier is specified in all input datasets	SUBJECT or PATIENT_ID
Outlib (Required)	The library in which the final SV output should be placed.	SDTM or SDS
vis1 (Required)	Variable which contains visit information in the input dataset which we use to derive the planned visit variable (VISIT) in SV.If the visit variable is different across datasets, we use vis2 and vis3 macro parameters to assign these variables. Note that, there can be only 1 visit variable corresponding to each dataset.	instancename
vis2 (Optional)	If there are multiple visit variables from input dataset that we use to derive VISIT along with vis1, we use vis2 macro parameter to assign the other visit parent variable.	foldername
vis3 (Optional)	If there are multiple visit variables from input dataset that we use to derive VISIT, along with vis1 and vis2, we use vis3 macro parameter to assign the other visit parent variable.	instancerepeatnumber
Incr (Required)	This is the increment of VISITNUM for unscheduled visits. For instance consider VISITNUM of a scheduled visit is 101, and if we are planning to assign the visit number of an unscheduled visit just after the scheduled visit as 101.01 (101 + 0.01), then incr = 0.01.	0.1 or 0.01 or 0.001

SV MACRO PARAMETERS

Macro Parameter	Description	Examples
exc_dt (Required)	The date variables that need to be excluded from the input datasets for deriving SVSTDTC and SVENDTC. When we consider different input datasets, there are some dates we never consider to derive SV dates such as system generated variables.	MAXUPDATED, RECORDDATE, DEATHDAT, UNBLNDAT..etc.
exc_data (Required)	List of datasets located in input library specified in “inplib” which doesn’t have any role in generation of SV dataset either to derive VISIT or SVSTDTC/SVENDTC.	AE, CM, MH, CP etc.
UplIndes (Optional)	Specify the dataset which contains unplanned visits description. Make sure that the dataset should be in inplib library. Note that, macro parameter only accept name of single dataset with unplanned visits description. If you don’t want to map SVUPDES in SV, then keep this parameter as missing.	Unsc or unpldes or unspc etc.
Byval (Optional)	List of variables which we need to merge unplanned visit description and that variables should be common to other raw datasets too. Note that, variables passing to above macro parameters vis1,vis2 or vis3 will be preferable.	Instancename instancerepeatnumber
desvar1 (Optional)	Variable from the dataset specified in parameter uplIndes, which contains unplanned visits description.	unsrsn
desvar2 (Optional)	Variable from the dataset specified in parameter uplIndes, which contains unplanned visits description for any other specified values.	unplspect

WORKING OF REQUIRED_DATA1.XLS FILE

Raw_datasets tab

LIBNAM	MEMNAME	NAME	TYPE	LABEL	datasetnam	flag1
RAW	DA1	DADAT	1	Date Of Assessment	'RAW.DA1'	
RAW	DA1	DADAT_DD	1	Date Of Assessment Day	'RAW.DA1'	exclude var
RAW	DA1	DADAT_INT	1	Date Of Assessment Interpolated	'RAW.DA1'	exclude var
RAW	DA1	DADAT_MM	1	Date Of Assessment Month	'RAW.DA1'	exclude var
RAW	DA1	DADAT_YYYY	1	Date Of Assessment Year	'RAW.DA1'	exclude var

Visit_fmt tab

INSTANCENAME	VISIT_
008 Day 1	Day 1
008 Day 11	Day 11 - Telephone Contact
008 Day 15	Day 15
008 Day 18	Day 18 - Telephone Contact
008 Day 22	Day 22
008 Day 29/Early Withdrawal	Day 29/Early Withdrawal
008 Day 4	Day 4 - Telephone Contact

SV MACRO – MAPPING OF VISIT AND VISITNUM

- Visit values from Raw data are transformed to standard visit values using visit_fmt tab in required_data1.xls
- Merge with SDTM TV dataset using the standard visit values
- Map VISITNUM and VISITDY for scheduled visits from TV dataset
- Sort the dataset based on Subject, SVSTDTC and SVENDTC so that actual order of unscheduled visit will be obtained
- Derive the VISITNUM for unscheduled visit by incrementing 0.01 with the preceding scheduled VISITNUM value (incr macro parameter can be used if necessary).

SV MACRO – MAPPING OF SVUPDES AND STUDY DAY

- Macro assumes that unscheduled visit description captured in a single dataset
- Macro variables UpIndes, byval, desvar1 and desvar2 can be utilized to map unscheduled visit description
- Merging with SDTM DM dataset to obtain RFSTDTC and SVSTDY, SVENDY can be calculated

SV MACRO – SAMPLE CALL

```
%sv(inplib=raw ,  
    Subjid=subject ,  
    outlib=sdtm ,  
    vis1=instancename,  
    vis2=foldername,  
    vis3=instancerepeatnumber,  
    incr=0.01,  
    exc_dt=("MAXUPDATED","RECORDDATE"),  
    exc_data = %str(AE, CM),  
    uplIndes=unsc,  
    byval=instancename instancerepeatnumber,  
    desvar1= unsrsn,  
    desvar2 = unplspect);
```

SV MACRO – FINAL DATASET

STUDYID	DOMAIN	USUBJID	VISITNUM	VISIT	VISITDY	SVSTDTC	SVENDTC	SVSTDY	SVENDY	SVUPDES
TRID-02	SV	TRID-02-801-030	1	Screening	.	2020-07-30	2020-07-31	-21	-20	
TRID-02	SV	TRID-02-801-030	2	Baseline	.	2020-08-20	2020-08-20	1	1	
TRID-02	SV	TRID-02-801-030	3	Day 1	1	2020-08-20	2020-08-20	1	1	
TRID-02	SV	TRID-02-801-030	4	Day 4 - Telephone Contact	4	2020-08-24	2020-08-24	5	5	
TRID-02	SV	TRID-02-801-030	5	Day 8	8	2020-08-27	2020-08-27	8	8	
TRID-02	SV	TRID-02-801-030	6	Day 11 - Telephone Contact	11	2020-08-31	2020-08-31	12	12	
TRID-02	SV	TRID-02-801-030	7	Day 15	15	2020-09-03	2020-09-03	15	15	
TRID-02	SV	TRID-02-801-030	8	Day 18 - Telephone Contact	18	2020-09-08	2020-09-08	20	20	
TRID-02	SV	TRID-02-801-030	8.01	Unscheduled 8.01	.	2020-09-08	2020-09-08	20	20	Other-KIT REPLACEMENT
TRID-02	SV	TRID-02-801-030	9	Day 22	22	2020-09-10	2020-09-10	22	22	
TRID-02	SV	TRID-02-801-030	10	Day 29/Early Withdrawal	29	2020-09-16	2020-09-16	28	28	
TRID-02	SV	TRID-02-801-030	11	Safety Follow-up	36	2020-09-24	2020-09-24	36	36	
TRID-02	SV	TRID-02-801-030	12	SAE Follow-up - Telephone Contact	59	2020-10-13	2020-10-13	55	55	

LIMITATIONS

- As the macro does not use the information from specification it contains numerous parameters.
- The input date variables should be in numeric type.
- If there are input datasets with visit information captured not on the primary visit variable, this should be handled separately. For example, assume INSTANCENAME is the variable having visit information captured across datasets. But in one dataset, INSTANCENAME contains the form name and actual visit is captured in another variable. It becomes difficult to understand which variable needs to be considered for derivation and those scenarios will be not be handled by the macro.
- For variable SVUPDES, the macro handles the case when unplanned visits description is collected only on one dataset.

CONCLUSION

- Macrotizing SV dataset is a great idea as the derivation is standard across studies
- Saves time and effort.
- Improve efficiency and quality.
- As the macro is built by simple SAS codes, any programmers with basic knowledge in macros can make changes as per their need if required.

THANK YOU

Any Questions ?