

Tackling Clinical Lab Data in Medical Device Environment

Juan Wu, Medtronic plc., Santa Rosa, CA

Min Lai, Medtronic plc., Santa Rosa, CA

ABSTRACT

Working with clinical laboratory data can be among the most challenging tasks for SAS[®] programmers when creating analysis datasets (AD) and tables, figures, and listings (TFL). Lab results are usually the largest dataset among clinical data. There are many different tests, and the same lab test can have different lab units. Many multi-centered medical device clinical studies use local labs, which result in the complication of lab units. Therefore the number of conversion factors adds up quickly. In order to tackle the complicated lab data, we developed a lab data manipulation process, which sets up a central lab test library and a lab unit conversion library to help improve proficiency in programming for AD and for TFL.

This paper will illustrate this process in detail. The sample lab test library and lab unit conversion library will be provided as well. We intend to reduce the complexity of coding and enable the trained clinical staff to maintain the lab test and lab unit libraries.

INTRODUCTION

The laboratory data is one of the most important parts of safety analysis in clinical trials for medical device companies. The laboratory data is often the largest dataset in a clinical trial. The nature of laboratory data, such as multiple visits, different lab tests, variability in lab units and their reference ranges for the same lab test (especially in international studies with local labs collecting the data), and dirty data, often poses many challenges for SAS programming teams. Hence it's critical to create an efficient approach to standardize the process of lab data manipulation, and to automate lab test and lab unit conversion in order to analyze and report lab results. Our goal is to bring the lab data processing closer to SDTM standards in medical device environment.

PROCESS FLOWCHART

The flow chart of the lab data manipulation process is displayed in Figure 1.

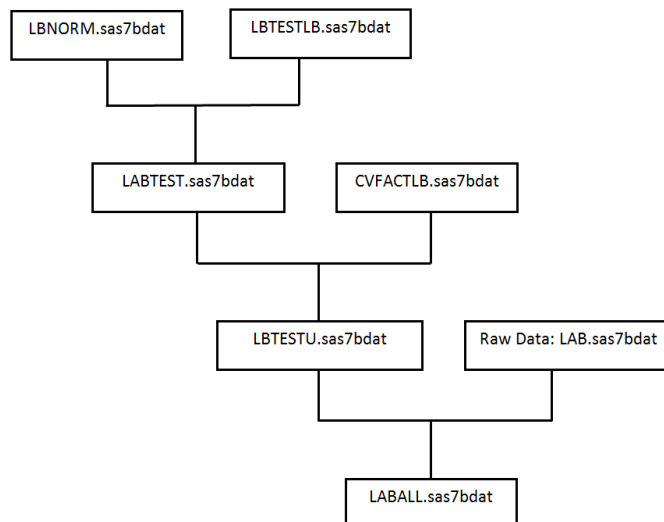


Figure 1: Lab Data Process Flowchart

LAB NORMAL DATASET

A lab normal dataset obtained from Oracle Clinical System is provided by data specialist for each specific study. The sample data definition table and the sample lab normal dataset are displayed in Tables 1 and 2.

VAR_NAME	VAR_TYPE	VAR_LEN	VAR_LABEL
LBSPEC	Char	6	Specimen Type
LBTESTCD	Char	14	Lab Test or Examination Short Name
LBCODE	Char	6	Lab Code
GENDER	Char	1	Gender
LBORRESU	Char	13	Lab Original Unit or Lab Unit Received
LBORNULO	Num	8	Reference Range Lower Limit in Original Unit
LBORNURI	Num	8	Reference Range Upper Limit in Original Unit
EFFSTDT	Num	8	Start Date of Effective Lab Unit
EFFENDDT	Num	8	End Date of Effective Lab Unit

Table 1: Sample Data Definition Table for Lab Normal

LBSPEC	LBTESTCD	LBCODE	GENDER	LBORRESU	LBORNULO	LBORNURI	EFFSTDT	EFFENDDT
BLOOD	HBA1C	L00456	F	%	.	6.1	1/1/1960	1/1/2100
BLOOD	HBA1C	L00456	M	%	.	6.1	1/1/1960	1/1/2100
BLOOD	HBA1C	L00627	F	%	4	6	1/1/1960	1/1/2100
BLOOD	HBA1C	L00627	M	%	4	6	1/1/1960	1/1/2100
PLASMA	PRA	L00456	F	mIU/L	.	.	1/1/1960	1/1/2100
PLASMA	PRA	L00456	M	mIU/L	.	.	1/1/1960	1/1/2100
PLASMA	PRA	L00627	F	mIU/L	4	37	1/1/1960	1/1/2100
PLASMA	PRA	L00627	M	mIU/L	4	37	1/1/1960	1/1/2100
SERUM	CA	L00456	F	mmol/L	2.14	2.5	1/1/1960	1/1/2100
SERUM	CA	L00456	M	mmol/L	2.14	2.5	1/1/1960	1/1/2100
SERUM	CA	L00627	F	mmol/L	2.15	2.6	1/1/1960	1/1/2100
SERUM	CA	L00627	M	mmol/L	2.15	2.6	1/1/1960	1/1/2100
SERUM	CK	L00456	F	U/L	40	200	1/1/1960	1/1/2100
SERUM	CK	L00456	M	U/L	60	285	1/1/1960	1/1/2100
SERUM	CK	L00627	F	U/L	30	170	1/1/1960	1/1/2100
SERUM	CK	L00627	M	U/L	30	190	1/1/1960	1/1/2100
SERUM	CREAT	L00456	F	umol/L	45	80	1/1/1960	1/1/2100
SERUM	CREAT	L00456	M	umol/L	60	105	1/1/1960	1/1/2100
SERUM	CREAT	L00627	F	umol/L	45	90	1/1/1960	1/1/2100
SERUM	CREAT	L00627	M	umol/L	60	110	1/1/1960	1/1/2100
SERUM	K	L00456	F	mmol/L	3.5	5	1/1/1960	1/1/2100
SERUM	K	L00456	M	mmol/L	3.5	5	1/1/1960	1/1/2100
SERUM	K	L00627	F	mmol/L	3.4	5	1/1/1960	1/1/2100
SERUM	K	L00627	M	mmol/L	3.4	5	1/1/1960	1/1/2100
SERUM	TROPONI	L00456	F	ng/L	.	30	7/2/2013	1/1/2100
SERUM	TROPONI	L00456	F	ug/L	.	0.03	1/1/1960	7/1/2013
SERUM	TROPONI	L00456	M	ng/L	.	30	7/2/2013	1/1/2100
SERUM	TROPONI	L00456	M	ug/L	.	0.03	1/1/1960	7/1/2013
SERUM	TROPONI	L00627	F	ng/L	0	16	11/12/2013	1/1/2100

SERUM	TROPONI	L00627	F	ug/L	0	0.04	1/1/1960	11/11/2013
SERUM	TROPONI	L00627	M	ng/L	0	26	11/12/2013	1/1/2100
SERUM	TROPONI	L00627	M	ug/L	0	0.04	1/1/1960	11/11/2013
URINE	CREAT	L00456	F	mmol/24h	2.5	19	1/1/1960	1/1/2100
URINE	CREAT	L00456	M	mmol/24h	3.5	23	1/1/1960	1/1/2100
URINE	CREAT	L00627	F	mmol/24h	5	16	1/1/1960	1/1/2100
URINE	CREAT	L00627	M	mmol/24h	9	18	1/1/1960	1/1/2100

Table 2: Sample Data for Lab Normal

The preprocessed lab normal dataset (LBNORM.sas7bat) is sorted by LBSPEC and LBTESTCD. And it's ready for merging with the lab test dataset. Some data cleaning steps need to be applied, but that is not the focus of this paper.

LAB TEST LIBRARY

The controlled terminology for the lab tests is part of the SDTM terminology, which is published on the NCI EVS (National Cancer Institute - Enterprise Vocabulary Services) website. The Lab test library generated in this paper is based on SDTM Lab test Package 3, and SDTM Implementation Guide Version 3.2. The lab test library is reviewed by clinical staff in a spreadsheet and imported into SAS by programmers. The following variables and labels are included in the sample lab test library dataset, which is shown in Table 3. The sample lab test library dataset is displayed in Tables 4. We will provide this lab test library to our database developer to standardize the lab test naming conversion.

VAR_Name	VAR_LABEL
LBCAT	Category for Lab Test
LBSCAT	Subcategory for Lab Test
LBTESTCD	Lab Test or Examination Short Name
LBTEST	Lab Test or Examination Long Name
LBSPEC	Specimen Type

Table 3: Variables and Labels in the Sample Lab Test Library

LBCAT	LBSCAT	LBTESTCD	LBTEST	LBSPEC
CHEMISTRY	General	ALBCREAT	Albumin/Creatinine Ratio	URINE
CHEMISTRY	General	BUN	Blood Urea Nitrogen	SERUM
CHEMISTRY	General	CA	Calcium	SERUM
CHEMISTRY	General	CL	Chloride	SERUM
CHEMISTRY	General	CHOL	Cholesterol	SERUM
CHEMISTRY	General	CREAT	Creatinine	SERUM
CHEMISTRY	General	CREAT	Creatinine	URINE
CHEMISTRY	General	CREATCLR	Creatinine Clearance	URINE
CHEMISTRY	General	GLUC	Glucose	SERUM
CHEMISTRY	General	GLUC	Glucose	URINE
CHEMISTRY	General	HBA1C	Hemoglobin A1C	BLOOD
CHEMISTRY	General	HDL	HDL Cholesterol	SERUM
CHEMISTRY	General	K	Potassium	SERUM
CHEMISTRY	General	LDL	LDL Cholesterol	SERUM
CHEMISTRY	General	TRIG	Triglycerides	SERUM
CHEMISTRY	General	SODIUM	Sodium	SERUM

CHEMISTRY	Enzymes	CK	Creatine Kinase	SERUM
CHEMISTRY	Enzymes	CKMB	Creatine Kinase MB	SERUM
CHEMISTRY	Enzymes	PRA	Plasma Renin Activity	PLASMA
CHEMISTRY	Hormones	INSULIN	Insulin	SERUM
CHEMISTRY	Peptides	CPEPTIDE	C-peptide	SERUM
CHEMISTRY	Peptides	NTBNP	NT-pro BNP	PLASMA
CHEMISTRY	Proteins	TROPONI	Troponin I	SERUM
CHEMISTRY	Proteins	TROPONT	Troponin T	SERUM

Table 4: Sample Lab Test Library

The lab test library dataset (LBTESTLB.sas7bat) is sorted by the same key variables (LBSPEC and LBTESTCD) as what's in the preprocessed lab normal dataset, and then those two datasets are merged together.

The sample SAS code will be provided as follows:

```
proc sort data=LBNORM; by LBSPEC LBTESTCD; run;
proc sort data=LBTESTLB; by LBSPEC LBTESTCD; run;

data LABTEST;
  merge LBNORM(in=a) LBTESTLB(in=b);
  by LBSPEC LBTESTCD;
  if a;
run;
proc sort; by LBSPEC LBTESTCD LBORRESU; run;
```

LAB UNIT CONVERSION LIBRARY

This is the key library. Lab data analysis requires programmer to remove the lab unit difference in the data before summarizing the results. This difference comes in during the data collection process. When a study uses multiple local labs and runs internationally, it is inevitable to face the fact that one lab test might be collected in multiple units with possibly different standards.

By creating a central Lab Unit Conversion Library, we hope to collect all possible units for different lab tests, and their corresponding conversion factors. Because it is captured in excel spreadsheet, the library can be maintained by non-statistical programmer and reviewed by trained clinical staff. The library is a living document, which can be added/removed/modified when needed. The library will be centrally controlled to maintain its stability and consistency.

The library has following key variables.

LBORRESU – Original unit or lab unit received

LBCVRESU – Conventional unit, U.S. standard

LBSTRESU – Standardized unit, per International Standard

SIFACTOR – Converting lab unit from the conventional units to the standardized unit

CVFACTOR – Converting lab unit from the standardized units to the conventional unit

Although a programmer can embed unit conversion in data step, but coding might be tedious and time consuming. The library enables the automated unit conversion process, reduces the repetitive effort across large amount of studies. It is a lot of effort to research and consolidate all the units in the library. The initial effort pays off when more lab data analyses use this library. Note that not all the units can be converted through conversion factors. Some additional data steps may be required for special units that cannot be converted via this method.

The sample lab unit conversion specification and the sample lab unit conversion library are displayed as follows in Tables 5 and 6, respectively.

VAR_NAME	VAR_TYPE	VAR_LEN	VAR_LABEL
LBSPEC	Char	6	Specimen Type
LBTESTCD	Char	8	Lab Test or Examination Short Name
LBORRESU	Char	10	Lab Original Unit or Lab Unit Received

LBCVRESU	Char	20	Lab Conventional Unit
LBSTRESU	Char	20	Lab Standard Unit
SIFACTOR	Num	8	Conversion Factor from Original Unit to Standard Unit
CVFACTOR	Num	8	Conversion Factor from Original Unit to Conventional Unit

Table 5: Sample Lab Unit Conversion Specification

LBTESTCD	LBSPEC	LBORRESU	LBCVRESU	LBSTRESU	SIFACTOR	CVFACTOR
ALBCREAT	URINE	mg/mmol	mg/g	mg/mmol	1	8.84
ALBCREAT	URINE	g/mol	mg/g	mg/mmol	1	8.84
ALBCREAT	URINE	ug/umol	mg/g	mg/mmol	1	8.84
ALBCREAT	URINE	mg/g	mg/g	mg/mmol	0.113	1
BUN	SERUM	mmol/l	mg/dl	mmol/l	1	2.801
BUN	SERUM	mg/dl	mg/dl	mmol/l	0.357	1
CA	SERUM	mmol/l	mg/dl	mmol/l	1	4.01
CA	SERUM	meq/l	mg/dl	mmol/l	0.5	2.005
CA	SERUM	mg/dl	mg/dl	mmol/l	0.2495	1
CL	SERUM	mmol/l	mg/dl	mmol/l	1	3.545
CL	SERUM	meq/l	mg/dl	mmol/l	1	3.545
CL	SERUM	mg/dl	mg/dl	mmol/l	0.2821	1
CREAT	SERUM	umol/l	mg/dl	umol/l	1	0.011
CREAT	SERUM	mg/dl	mg/dl	umol/l	88.402	1
CREAT	URINE	g/24h	g/24h	mmol/24h	8.84	1
CREAT	URINE	mmol/24h	g/24h	mmol/24h	1	0.113
GLUC	SERUM	mmol/l	mg/dl	mmol/l	1	18.02
GLUC	SERUM	mg/dl	mg/dl	mmol/l	0.0555	1
GLUC	URINE	mmol/l	mg/dl	mmol/l	1	18.02
GLUC	URINE	mg/dl	mg/dl	mmol/l	0.0555	1
HBA1C	BLOOD	%	%	%	1	1
K	SERUM	mmol/l	mg/dl	mmol/l	1	0.910
K	SERUM	meq/l	mg/dl	mmol/l	1	0.910
K	SERUM	mg/dl	mg/dl	mmol/l	0.2557	1
CK	SERUM	u/l	u/l	u/l	1	1
CK	SERUM	1ukat/l	u/l	u/l	58.8235	58.8235
PRA	PLASMA	uiu/ml	miu/l	miu/l	1	1
PRA	PLASMA	miu/l	miu/l	miu/l	1	1
INSULIN	SERUM	miu/l	miu/l	pmol/l	0.144	1
INSULIN	SERUM	uiu/ml	miu/l	pmol/l	0.144	1
INSULIN	SERUM	pmol/l	miu/l	pmol/l	1	0.95
TROPONI	SERUM	ng/l	ng/l	ng/l	1	1
TROPONI	SERUM	ug/l	ng/l	ng/l	1000	1000
TROPONI	SERUM	ng/ml	ng/l	ng/l	1000	1000

Table 6: Sample Lab Unit Conversion Library

The dataset for lab unit conversion library is sorted by LBSPEC, LBTESTCD, LBORRESU, and then merged with the dataset after combining the lab test library dataset and lab normal dataset by LBSPEC, LBTESTCD, and LBORRESU. Only the records in the combined dataset are kept. After this step, the resulting dataset is sorted by LBCODE, LBTESTCD, and GENDER.

The sample SAS code will be provided as follows:

```
proc sort data=cvfactlb; by LBSPEC LBTESTCD LBORRESU; run;

data labtestu;
  merge LABTEST (in=c) cvfactlb(in=d);
  by LBSPEC LBTESTCD LBORRESU;
  if c;
run;
proc sort; by LBCODE LBSPEC LBTESTCD GENDER; run;
```

LAB DATA FROM OC DATABASE

The last step is to combine central conversion library, lab normal with lab results collected from each different site. The actual results can be extracted from central clinical databases, or data files provided by each lab.

Merging combined library datasets with lab data from OC database by LBCODE, LBSPEC, LBTESTCD, and GENDER. Now all the required variables including lab normal, conversion factors, lab test results, date of lab specimen collection, the start and end dates of the effective lab unit, et al are all in the same dataset for creating analysis datasets (AD) and tables, figures, and listings (TFL).

The sample SAS code will be provided as follows:

```
proc sort data=lab; by LBCODE LBSPEC LBTESTCD GENDER; run;

data alllab;
  merge lab (in=a) labtestu(in=b);
  by LBCODE LBSPEC LBTESTCD GENDER;
  if a;
  if EFFSTDT <= LABDT <= EFFENDDT;
run;
```

CHALLENGES

When working with the clinical laboratory data, challenges consistently exist. There are unpredictable issues such as misspellings in units, multiple entries on the same test item, partial data. Therefore SAS programmers have to detect unexpected data issues and do the data clean process after retrieving lab normal data from central clinical database. These tasks are very study specific.

Moreover, test unit effective date should be considered when any given laboratory changes their test system at any time. For example, if the assays for certain lab test (Troponin-I) change on a specific date (12-Nov-2013), as a result, the unit for this lab would have to be changed from $\mu\text{g/L}$ to ng/L as well. In this case, we have to keep two different lab units for the same lab and put constraints on the date of specimen collection between the start and end dates of the effective lab unit when merging the final lab test unit library dataset with the preprocessed raw lab data.

CONCLUSION

The lab test and lab conversion libraries have been generated, and a standard process of tackling lab data has been developed for selected clinical trials in Medtronic plc.

The generation of the lab test and lab unit conversion libraries reduced lab test, lab unit or lab unit conversion issues, and made the lab data processing more efficiently, especially for a multi-center, internationally run clinical trial with local labs.

REFERENCES

- "CDSIC Standards and Implementation." 2014. Available at <http://www.cdisc.org/standards-and-implementations>.
- "CDISC Terminology." NCI Enterprise Vocabulary Services (EVS), Dictionaries, and Partnerships. Available at <http://www.cancer.gov/cancertopics/cancerlibrary/terminologyresources/cdisc>.

ACKNOWLEDGMENTS

Appreciation goes to Clinical Team for confirming the accuracy of the lab unit conversion and Data Management Team for the maintenances of the Lab Test Library. Therefore SAS programming team can implement the process in our own practice.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Juan Wu, M.S.
Enterprise: Medtronic plc.
Address: 3576 Unocal Place
City, State ZIP: Santa Rosa, CA 95403
Work Phone: 707-591-2276
E-mail: juan.wu@medtronic.com

Name: Min Lai, M.S.
Enterprise: Medtronic plc.
Address: 3576 Unocal Place
City, State ZIP: Santa Rosa, CA 95403
Work Phone: 707-543-5028
E-mail: min.lai@medtronic.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.