Paper SI09

Lab Master List File Makes Easy

Clio Wu, FMD K&L Inc., Fort Washington, PA, USA

ABSTRACT

Clinical laboratory data plays a significant role in the clinical trials, the primary objectives for collecting lab data are to evaluate the safety of the patients in the trials and to demonstrate the efficacy of the investigational medications. As Lab data come from central and local laboratories with various format of units and lab parameters, converting lab data into a submission-ready SDTM LB dataset has always been a challenge for sponsors and their CRO partners.

This paper details a standard lab master list file and associated programming process to make the lab converting process easy, the project was initiated in year 2012 to resolve lab units and lab parameters related challenges where data collected from various lab vendors across multiple sponsors and TAs. This file is a living document to collect and process new identified lab units and new lab tests terms overtime to this day.

INTRODUCTION

When it comes to clinical laboratory data, it seems the challenges will never end, just when you think you have got it figured out, there is a new wrinkle – a new regulation, a new data source and a new data standard. We are today still trying to tackle the same or more complex issues or challenges we did 5 years, 10 years or even 20 years ago. Thus, it is no surprise that during the PharmaSUG 2019 Open Session, the PharmaSUG Family Feud Survey voted ADLB the least favorite ADaM data set to program.

So, what is lab data exactly? What are the challenges when working on the lab data?

Any test normally performed by a clinical laboratory is considered a lab test. Lab data is often referred to as the clinical laboratory data or the laboratory test results; CDISC SDTM LB is defined as the SDTM domain captures laboratory data collected on the CRF or received from a central provider or a third party vendor.

The major lab data challenges from the lab data end-to-end process include:

- Study Planning: what lab tests to be included? What laboratory to analyze the lab test? This is per sponsors protocol and study design, some sponsors reference the CDASH Library, others have their mandatory Lab Test List per Therapeutic Area (TA) standard and per Phase of the clinical trial.
- Data Collection:
    - Different types of EDC systems (e.g. Medidata Rave, InForm, Oracle Clinical, etc.)
    - Different types of labs eCRF and/or Data Transfer Specification (DTS) designs (e.g. good vs. bad CRF design, baseline eligibility, baseline stratification, central lab, local lab and unscheduled CRFs)
    - Variability in sources of the lab data (e.g. EDC CRF, Central Lab, Local lab, Pharmacodynamics and Specialty Lab data)
    - Variability in lab test names (e.g. same test name with different specimen type and test method)
    - Hematology differential test (e.g. mixture of absolute and relative results for one lab parameter)
- Data Tabulation:
    - Standardized lab test name and category
    - Standardized units (i.e. Need ONE standard unit per lab parameter)
    - Conversion of reported lab values to U.S. Conventional units or Système International (SI) units
    - Requirement for NCI Common Terminology Criteria for Adverse Events (CTCAE) toxicity grading
    - FDA Requirement for Logical Observation Identifiers Names and Codes (LOINC) Codes submission for studies that start after March 15, 2020 for NDAs, aNDAs and certain BLAs
- Statistical Analysis and TLF reporting (e.g. Bi-directional lab tests analysis and reporting, lab precision)

What can be done to make the data collection and tabulation process more harmonized to enable more efficient SDTM and ADaM programming?

This paper is to detail the Lab Master List File project initiated back in 2012 to address the issues and challenges associated with clinical lab data and its related data collection, tabulation and analysis process. The paper will demonstrate how Lab Master List File and supported process can help standardize and streamline the overall lab data end-to-end process.

LAB MASTER LIST FILE HISTORY AND OBJECTIVES

During a FDA NDA CDISC data submission package preparation for a major biotechnology sponsor company lead Oncology pivotal study, its supported Phase II/Phase III, outsourced clinical pharmacology (ClinPharm) and multiple legacy studies, study programming team have identified the lab data mapping inconsistency issue across studies and vendors, to facilitate future study lab data collection and mapping process, Lab Master List File project was initiated in year 2012 as a new process improvement initiative of developing global EDC, SDTM and ADaM standards, author was leading the standardization effort for the Oncology therapeutic area (TA), later this project was expanded to other TAs and other sponsors/vendors to maximize the lab data collection for the project.

The major objective of this project is to standardize the lab tests and related units collection for EDC, SDTM and ADaM data process, especially for the SDTM level LBCAT, LBTEST, LBTESTCD, LBSTRESU and lab units conversions process to streamline the overall lab data end-to-end process (Figure 1).

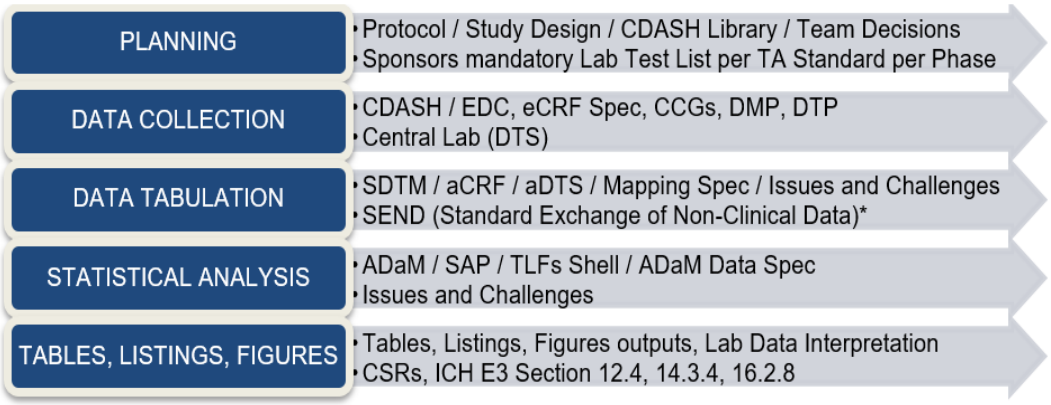

Figure 1: Lab Data End-to-End Process Flow

The development of end-to-end standards process is to avoid data “garbage in, garbage out” or “garbage in, -with intensive programming cleaning and mapping efforts-, then quality data out” scenarios. High-quality input data are fundamental to producing reliable and quality datasets, regardless how good the CDISC standards are, if the input source data used to create and implement them are bad, the datasets created are often questionable or required lot of workaround to rid of all data related issues, it is very resource- and time-consuming process that we have tried to avoid if possible. Very often sponsors decide not to clean or standardize the local lab data at the SDTM step if the study Statistical Analysis Plan (SAP) specifically design to use central lab data only to keep it simple for study analysis and reporting purposes, however, this approach will not help with or resolve the programming CDISC SDTM standardization and Pinnacle 21 compliance check issues, programming team still need to take extra effort to map and/or further explain in the clinical Study Data Reviewers Guide (cSDRG) why local lab data would not be standardized.

The concept of the project is rather simple, the ultimate objective for this project is to plan and design a global standard lab tool that can be used across TAs to resolve the identified lab data issues and to create high-quality submission-ready SDTM and ADaM lab data packages to be in compliance with FDA requirements and CDISC standards.

LAB MASTER LIST FILE DEVELOPMENT PROCESS

Even it seems to be a simple and effortless project, actually, it is an intensive and time-consuming effort during the initial development process. It began with a designated programming resource spent significant amount of time working on one particular TA to locate, load and filter the pooled lab data, search and reference extensive lab-related CDISC Controlled Terminology (CT), lab units conversion factors and recheck the new transferred lab data to ensure all lab terms and units are standardized. Here are some of the programming tasks involved during the initial development stage (Figure 2):

- Identify all source laboratory datasets to be pooled and standardized
- Identify all source lab variables used in the standardization for each dataset

- Filter the unique source lab test term and all available unique unit per lab test term
- Reference CDISC CT to map and standardize the available standard LBCAT, LBTEST, LBTESTCD, LBSPEC and LBMETHOD, or assign sponsor-defined or sponsor-provided LBTEST and LBTESTCD codelist if not available
- Reference and cross-check multiple World Wide Web (WWW) resources for documented lab units conversion factors, work with study clinicians to review and provide the missing or questionable ones if such information is not available
- Reference CTCAE to locate the lab test with toxicity grade and bi-directional analysis per defined version (e.g. CTCAE v4.03, v5.0)
- Identify the lab precision for reporting, work with statistician to acquire the standard process and formats if applicable
- Compile the Lab Master List File with all planned components and programmable layout, have study team independently validate, review and/or comment
- Prepare a guideline and programming implementation document to detail the purpose, scope and the procedure of using and implementing the file

Figure 2: Lab Master List File Process Flow

Once the initial Lab Master List File is validated for that particular TA, along with supported guideline and programming implementation document, it can be routed to study team for implementation.

It is also a continued programming effort as we have more studies, products and TAs data pooled in to maximum the magnitude and variabilities of source lab test terms and units. Figure 3 demonstrates the ongoing development strategies to expand the source lab data pool and maximum the collection of unique lab test names, units and unit conversion factors.

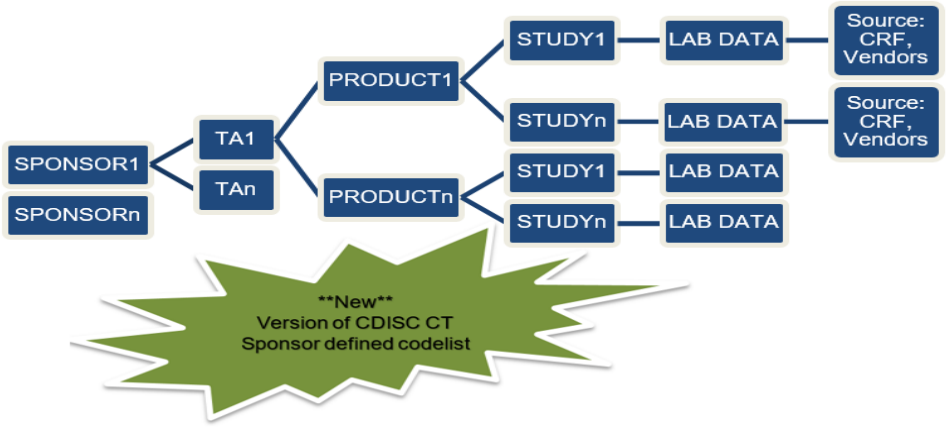

Figure 3: Lab Master List File continued effort to expand the source data pooled

LAB MASTER LIST FILE COMPONENTS AND FORMATS

The compiled Lab Master List File contains four major components for the standardization of the SDTM, ADaM and TLFs reporting process including:

1) Lab Test Conversion tab: lists all unique source CRF, local, central, specialty lab and pregnancy test names and their corresponding standardized SDTM lab test names, standardized lab SI units, the specimen type and test method (LBCAT, LBTEST, LBTESTCD, LBSTRESU, LBSPEC and LBMETHOD)
2) Lab Unit Conversion tab(s): lists all available source local, central lab tests units and their corresponding standardized SDTM lab test units (LBSTRESU), and the sponsors defined U.S. Conventional units and SI units conversion factors
3) Lab Test CTCAE Grading tab(s): list all available lab test CTCAE grading, the range criteria for each grade and the direction of the grading (i.e. increase or decrease), one tab per version (i.e. v4.03 and v5.0). Some lab tests have both increase and decrease directions.
4) Lab Precision tab: to identify precision to be displayed in formatted outputs for specific lab parameters for both U.S. Conventional and SI results

The standardized LB test name, unit and conversion factors will be revised periodically based on the most up-to-date version of CDISC CT, sponsor-defined codelist, SI units reference document, and the conversion equations for the reported lab units and U.S. Conventional/SI units.

To enhance the team review and programming process, the file is compiled using MS Excel file format (.xlsx) and converted to the SAS dataset format (.sas7bdat), one tab/dataset per component. These files are stored at a centralized secure server location that project/study team can reference and utilize based on their project/study and functional needs. This Lab Master List File will also be shared with global project team and programming vendors during the course of the TA/project/study lifecycle to streamline the end-to-end process. (Figure 4)

SOURCE LBTEST (LBTSTSP/TESTNAME)	SOURCE LBTESTCD (LBTSTSP)	SDTM LB.LBTEST	SDTM LB.LBTESTCD	SDTM LB.LBCAT	SDTM LB.LBSTRESU	SDTM LB.LBSPEC	SDTM LB.LBMETHOD
ALT (Alanine Aminotransferase)		Alanine	ALT	CHEMISTRY	U/L	BLOOD	
AMYLASE	AMYLASE	Amylase	AMYLASE	CHEMISTRY	U/L	BLOOD	
ABSOLUTE NEUTROPHIL COUNT	ANC (unit: other units)	Neutrophils	NEUT	HEMATOLOGY	10*9/L	BLOOD	
ABSOLUTE NEUTROPHIL COUNT	ANC (unit: %)	Neutrophils/Leukocyte	NEUTLE	HEMATOLOGY	%	BLOOD	
ASPARTATE	AST	Aspartate	AST	CHEMISTRY	U/L	BLOOD	
AST (Aspartate Aminotransferase)		Aspartate	AST	CHEMISTRY	U/L	BLOOD	
BASOPHILS	BASO (unit: other units)	Basophils	BASO	HEMATOLOGY	10*9/L	BLOOD	
BASOPHILS	BASO (unit: %)	Basophils/Leukocytes	BASOLE	HEMATOLOGY	%	BLOOD	
Basophils	(unit: %)	Basophils/Leukocytes	BASOLE	HEMATOLOGY	%	BLOOD	
DIRECT BILIRUBIN	BILDIR	Direct Bilirubin	BILDIR	CHEMISTRY	umol/L	BLOOD	
BILIRUBIN, TOTAL BILIRUBIN	BILI	Bilirubin	BILI	CHEMISTRY	umol/L	BLOOD	
Bilirubin (Total)		Bilirubin	BILI	CHEMISTRY	umol/L	BLOOD	
INDIRECT BILIRUBIN	BILIND	Indirect Bilirubin	BILIND	CHEMISTRY	umol/L	BLOOD	
PERIPHERAL BLASTS	BLASTS (unit: other units)	Blasts	BLAST	HEMATOLOGY	10*9/L	BLOOD	
PERIPHERAL BLASTS	BLASTS (unit: %)	Blasts/Leukocytes	BLASTLE	HEMATOLOGY	%	BLOOD	
BLOOD UREA NITROGEN	BUN	Urea Nitrogen	UREAN	CHEMISTRY	mmol/L	BLOOD	
BUN (Blood Urea Nitrogen)		Urea Nitrogen	UREAN	CHEMISTRY	mmol/L	BLOOD	

Lab Test Conversion tab

SOURCE LBTEST (LBTSTSP/TESTNAME)	SOURCE LBTESTCD (LBTSTSP)	SOURCE LAB UNIT (LBUNIT/LBORRESU)	LB.LBSTRESU	Conversion factor
BLOOD UREA NITROGEN	BUN	MMOL/L	mmol/L	1
BUN (Blood Urea Nitrogen)		mg/dL	mmol/L	0.357
CALCIUM	CA	MG/DL	mmol/L	0.25
CALCIUM	CA	MMOL/L	mmol/L	1
CALCIUM	CA	MEQ/L	mmol/L	1
Calcium		mg/dL	mmol/L	0.25
CHOLESTEROL, TOTAL	CHOL	MMOL/L	mmol/L	1
CHOLESTEROL, TOTAL	CHOL	MG/DL	mmol/L	0.02586
Cholesterol		mg/dL	mmol/L	0.02586
CHLORIDE	CL	MEQ/L	mmol/L	1
CHLORIDE	CL	MMOL/L	mmol/L	1
Chloride		mmol/L	mmol/L	1
CREATININE	CREAT	MCMOL/L	umol/L	1
CREATININE	CREAT	MG/DL	umol/L	88.4
CREATININE	CREAT	MG/L	umol/L	8.84
CREATININE	CREAT	MMOL/L	umol/L	1000
CREATININE	CREAT	UMMOL/L	umol/L	1
CREATININE	CREAT	UMOL/L	umol/L	1

Lab Unit Conversion tab(s)

PARAMCD	PARAMN	CONV_PREC	SI_PREC
WBC	1	1	1
HGB	2	1	1
HCT	3	0	2
NEUTLE	5	0	2
LYMLE	6	0	2
MONOLE	7	0	2
EOSLE	8	0	2
BASOLE	9	2	4
NEUT	10	0	0
LYM	11	0	0
MONO	12	0	0
EOS	13	0	0
BASO	14	0	0

Lab Precision tab

GRADE	RANGE1	LBTEST	LBTESTCD	RANGE2	LBSTRESU	DIR	TERM
1	1	Aspartate Aminotransferase	AST	3	*ULN	i	Aspartate aminotransferase increased
2	3	Aspartate Aminotransferase	AST	5	*ULN	i	Aspartate aminotransferase increased
3	5	Aspartate Aminotransferase	AST	20	*ULN	i	Aspartate aminotransferase increased
4	20	Aspartate Aminotransferase	AST	1.00E+11	*ULN	i	Aspartate aminotransferase increased
1	1	Bilirubin	BILI	1.5	*ULN	i	Blood bilirubin increased
2	1.5	Bilirubin	BILI	3	*ULN	i	Blood bilirubin increased
3	3	Bilirubin	BILI	10	*ULN	i	Blood bilirubin increased
4	10	Bilirubin	BILI	1.00E+11	*ULN	i	Blood bilirubin increased
1	lln	Calcium	CA	2	mmol/L	d	Hypocalcemia
2	2	Calcium	CA	1.75	mmol/L	d	Hypocalcemia
3	1.75	Calcium	CA	1.5	mmol/L	d	Hypocalcemia
4	1.5	Calcium	CA	0	mmol/L	d	Hypocalcemia
1	uln	Calcium	CA	2.9	mmol/L	i	Hypercalcemia
2	2.9	Calcium	CA	3.1	mmol/L	i	Hypercalcemia
3	3.1	Calcium	CA	3.4	mmol/L	i	Hypercalcemia
4	3.4	Calcium	CA	1.00E+11	mmol/L	i	Hypercalcemia
1	lln	Glucose	GLUC	3	mmol/L	d	Hypoglycemia
2	3	Glucose	GLUC	2.2	mmol/L	d	Hypoglycemia
3	2.2	Glucose	GLUC	1.7	mmol/L	d	Hypoglycemia
4	1.7	Glucose	GLUC	0	mmol/L	d	Hypoglycemia
1	uln	Glucose	GLUC	8.9	mmol/L	i	Hyperglycemia
2	8.9	Glucose	GLUC	13.9	mmol/L	i	Hyperglycemia
3	13.9	Glucose	GLUC	27.8	mmol/L	i	Hyperglycemia
4	27.9	Glucose	GLUC	1.00E+11	mmol/L	i	Hyperglycemia

Lab Test CTCAE Grading tab(s)

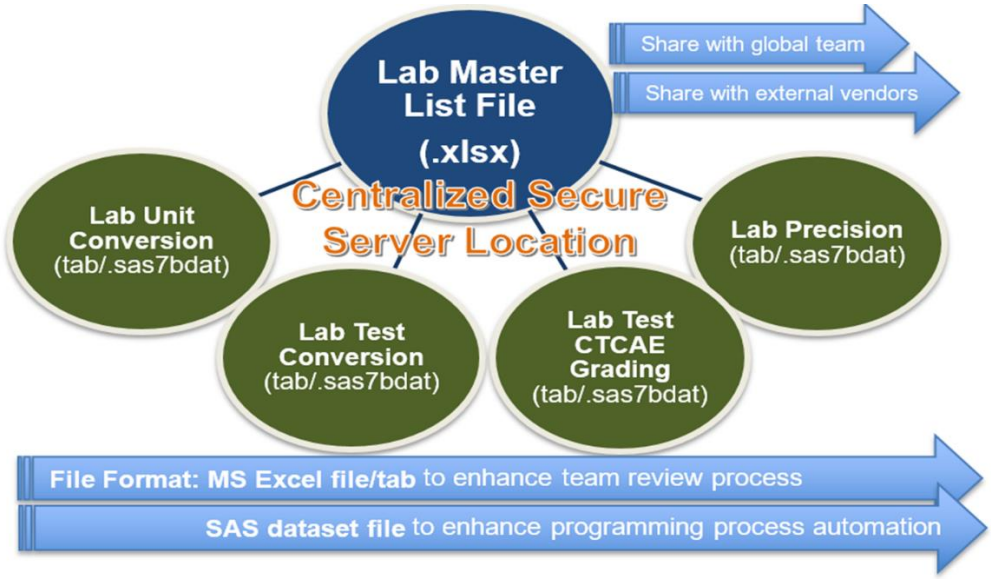

Figure 4: Lab Master List File Components and Formats

BIOMETRICS GROUP R&RS

As this is a jointed Biometrics effort to ensure the end-to-end process run smoothly and consistently to best utilize the compiled lab master list file. Here are the roles and responsibilities (R&Rs) for each Biometrics functional group to facilitate ongoing review process of the Lab Master List File.

- Lab Master List File Owner: to initiate the creation of Lab Master List File and update periodically to add in new lab test names, units, units conversion factors; to set up standard SDTM LB mapping rules, prepare supported guideline and programming implementation document; to serve as the owner and primary contact for updating this document, the file will be routed to the Biometrics groups for data cleaning, review, approval and/or implementation process.
- Study CDM: the lab master list file often contains missing or questionable labs tests or units entries, will need CDM to help query for data cleaning purposes; to safeguard the source data quality; to effectively communicate with EDC, central and local lab vendors to ensure lab data test names and units design and entries following this standards document
- Study Statistician: general review of the file, to follow this lab master list file when contribute to the Protocol and/or work on SAP and TLFs shells design
- Study Programmer: to implement this lab master list file during SDTM/ADaM/TLFs programming process; to provide new identified lab test names and units against the existing list file; to review and cross-check CDISC LBCAT, LBTEST, LBTESTCD, LBSTRESU, LBSPEC and LBMETHOD as per the most up-to-date version of CDISC CT, sponsors-defined codelist, SI units and unit conversion factors if applicable.

The Labs Master List File is to be considered as a living document and will be updated periodically when new lab test names are identified from existing or new sponsors, vendors, TAs and/or studies against the existing approved list.

LAB MASTER LIST FILE IMPLEMENTATION PROCESS

The Lab Master List File is designed to be implemented during the TA/project/study life cycle to ensure the consistency of data test names and units designs throughout the Protocol, eCRF, EDC, SDTM, ADaM, TLFs process.

To ensure TA/project/study team implement the process successfully, it is important to introduce this end-to-end lab standards at the initial TA/project/study kick-off (KO) meeting, to have TA/project/study team buy-in and committed early on to implement the standard CDM eCRF and vendor DTS designs, lab data entry and data collection, Programming SDTM/ADaM and Statistician SAP/TLFs shell process, and to implement at the internal and external vendors levels to facilitate the overall standard process.

PROGRAMMING CONSIDERATIONS

Here are some general considerations for programming group to enhance the implementation of the lab master list file and the supported process:

- Although SDTM and ADaM look quite simple, getting the structure, the relationships, and the metadata correct actually requires a considerable amount of planning and effort. And of course if your clinical trial has a few source data issues, then getting the conversion correct becomes even more complicated. It is best to start with CRFs that are compliant with SDTM standards. In order to do that, programming team need to be more proactively and actively involved in the CRF planning and design, EDC User Acceptance Test (UAT), SAP and TLFs shell process.

- And, what is the appropriate lab Standard Unit (LBSTRESU) to use? It is defined as the "Standardized unit used for LBSTRESC or LBSTRESN." What does the "Standardized Unit" really mean? Often there is a misconception between the "Standardized Unit" and "SI Unit". There are several types of units including Original units reported by the labs, U.S. Conventional units and SI units. Who or how to determine this standard unit?

  There is no standard FDA/CDISC requirements yet! Even though the FDA 2013 position paper recognized that eventually SI units is the one to use in the future, however, they also acknowledged that the majority of U.S. healthcare providers and FDA reviewers are trained using U.S. Conventional units. So as of today, it's totally up to sponsors, sponsors can define their own standard units for their submissions or their studies, this can be U.S. Conventional units or SI units, or per Protocol or CRF driven. Some sponsors have decided to keep the reported original units, U.S. Conventional units and SI units at SDTM either at the main LB or SUPPLB based on their analysis needs and noted in the cSDRG. This is to avoid the further never-end internal dissuasions regarding why favoring or choosing one unit than the other.

- As this lab master list file comes with SAS datasets and supported programming macro and utilities to enhance the programming efficiency, it is best to set up an automate process to incorporate these tools to streamline the lab tests, lab units standardization and conversion process. Same to the lab toxicity grading process if such information is not available at the source data.

KEEP EXTERNAL VENDORS IN THE LOOP

When come to this process improvement project, they are, of course, the obvious in-house key stakeholders - sponsors, TA/project/study team member and the Biometrics group that are directly impacted by this project. But undoubtedly, external lab or programming vendors are part of the key stakeholders as well, they are essential to the success of the project.

Therefore, it is important to have external vendors committed to the same lab standardization process to avoid the data "garbage in, garbage out", limit the variabilities in source data, and minimum the incompliance or inconsistency in submission data package issues.

IMPACTS ON LAB DATA PROCESS WITHIN BIOMETRICS GROUP

Figure 5 details the implementation of the Lab Master List File during the lab data end-to-end process and the impacts within the Biometrics group.

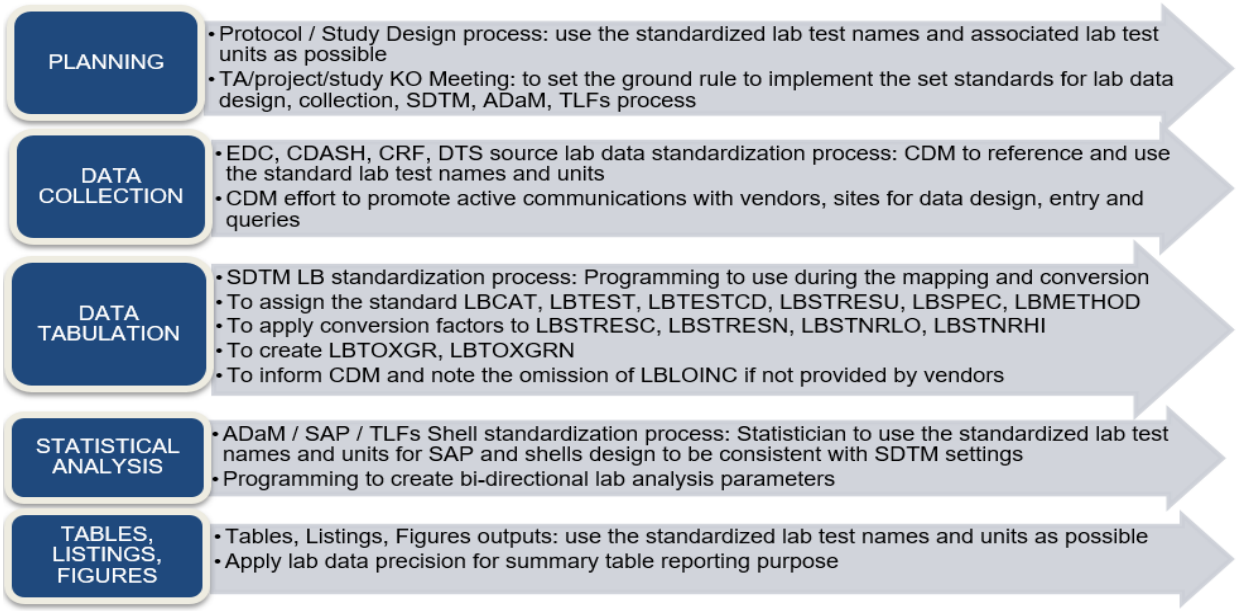

Figure 5: Lab Master List File Implementation during Lab Data End-to-End Process

BENEFIT OF USING LAB MASTER LIST FILE

Even though we have well established standard lab test names and lab test codes Controlled Terminology (CT) and the NCI CTCAE files in place, however, there are no standard libraries or documents to reference or cross-check for the correct unit per lab parameter and the correct unit conversion factors. Study programming team need to reference and cross-check multiple sources or documents for their programming needs, it is important and beneficial to have a simple centralized standard documentation or tool to put these information together that programming team can

better spend their time and resource doing the critical project tasks - as this seems to be an effortless task, but do take up programming time and effort repeatedly overtime.

The Lab Master List File is a one-stop-shop centralized tool for all lab data process needs. The file and the supported process have been implemented by several sponsors and their CRO partners to support their TA/project/study standard data design, data collection, data tabulation, data analysis and reporting process since the initiation of this project, it has supported multiple successful FDA NDA/BLA SDTM and ADaM Submission Packages.

With this centralized lab master list file tool and the supported process, the project team and especially the Biometrics group have significantly improved their overall efficiency, as we all follow the same set of lab data standards throughout the study life cycle and promote the effective and proactive communications within the study team and with the external lab data or programming vendors, to create more standardized, harmonized, accurate and high quality lab data package.

WHAT IS NEXT ON THE HORIZON

In January 2017, FDA formed a LOINC Working Group to address CDISC community concerns and develop recommendations for the submission of Logical Observation Identifiers Names and Codes (LOINC®) code. Based on the most current FDA Data Standard Catalog doc, sponsors are required to submit LOINC using SDTM LBLOINC for studies that start after March 15, 2020 for NDAs, aNDAs and certain BLAs.

Currently most sponsors leave this LBLOINC blank unless it is provided by the lab vendors, with this FDA requirement in mind, we need to be more proactive to have labs provide the LOINC regardless of the source of the lab data, central or local lab. Here are the recommendations for the lab data process flow:

- Data collection: CDM need to add this field to the EDC eCRF and/or DTS design to ensure LOINC is available at the source data
- Data Tabulation: If such LOINC codes are not available from the laboratory, the Working Group recommends that sponsors do not attempt to derive a mapping to LOINC as they create the SDTM files that can lead to serious data quality issues; and for any lab test where a LOINC code is not submitted, the reason of this omission should be noted in the cSDRG.

CONCLUSION

This Lab Master List File is a living document to collect and standardize new unique lab test names, units and unit conversion factors, list CTCAE grading for applicable lab tests and their associated grade criteria to support lab toxicity grading and bi-directional analysis. This is a jointed Biometrics team effort to standardize, harmonize and streamline the study laboratory data design, EDC and vendors' data entry, data collection, data tabulation, data analysis and reporting end-to-end process.

The Lab Master List File is not only a tool to standardize and streamline the overall end-to-end lab data process, it's also a one-stop-shop shared tool to make study team and programming job easier and more efficient when come to lab data process.

REFERENCES

1. CDISC SDTM v1.4 and SDTMIG v3.2
2. FDA. Study Data Technical Conformance Guide - Technical Specification Document. March 2019
3. FDA CDER/CBER. Position on Use of SI Units for Lab Tests. October 25, 2013
4. FDA. Recommendations for the Submission of LOINC® Codes in Regulatory Applications to the U.S. Food and Drug Administration. November 2017
5. Common Terminology Criteria for Adverse Events (CTCAE) v4.03: June 14, 2010. v5.0: November 27, 2017

CONTACT INFORMATION

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

Clio Wu
FMD K&L Inc.
1300 Virginia Dr., Ste 408,
Fort Washington, PA 19034
Email: clio.wu@klserv.com
www.klserv.com