PharmaSUG China 2025 - Paper DS-175

Navigating Unit Standardization in Clinical Trials: Bridging FDA and SI Unit Requirements through LB and LC Domains

Jialiang LIU, Sanofi

ABSTRACT

The standardization of laboratory data units in clinical trials remains a significant challenge due to differences in regulatory requirements, especially between the FDA's preference of conventional units (e.g., cells/mm³ for WBC) and the globally accepted SI units (e.g., ×10⁹/L for WBC). With the advent of mRNA vaccines and advances in vaccine technology, the LB (Laboratory) domain has become essential for vaccine and drug development, requiring strict unit standardization to meet regulatory rules and enable consistent data comparison. This article explores two approaches to address unit discrepancies are explored: vertical expansion (adding rows with a separate variable to differentiate SI and conventional units) and horizontal expansion (introducing parallel columns for different units), each with distinct advantages and limitations. The FDA's 2024 SDTCG v5.7[1] update added the LC (Laboratory Conventional) domain, requiring two datasets: LB for SI-unit lab results and LC for conventional-unit lab results, with matching observation counts and direct pairing via LBSEQ/LCSEQ keys. The article emphasizes collaborating with lab vendors to obtain direct dual-unit data, documenting conversion logic in Define-XML, and prioritizing early alignment with regulatory and statistical teams to ensure compliance, data accuracy, and transparency in vaccine development and clinical trial submissions.

KEYWORDS

Clinical Trials, SDTM, FDA, SI Units, Laboratory Data Harmonization, LB Domain, LC Domain, Unit Conversion, Data Standardization, mRNA Vaccines, Immunotherapy, Define-XML, CBER Guidelines

INTRODUCTION

The LB (Laboratory) domain, a core part of the SDTM Implementation Guide (SDTMIG)[2], is widely used in clinical trials to capture laboratory findings—particularly in areas like oncology, infectious diseases, and other therapeutic fields. In the past, vaccine studies focused more on safety endpoints such as Adverse Events (AEs) and Clinical Events (CEs), making the LB domain less critical. However, the rise of mRNA technology and immunotherapies has changed this, pushing vaccine research to pay closer attention to laboratory parameters like White Blood Cell (WBC) counts, electrolyte levels, and liver function tests. This shift has blurred the traditional boundaries between vaccine and drug development, making the LB domain an important part of clinical trials.

A critical challenge in this context is unit standardization, as regulatory agencies like the FDA and EMA often require different units for the same laboratory test. In vaccine studies submitted to the FDA, the "Guidance for Industry[3]" is typically referenced as a critical regulatory benchmark and also a key reference. As shown in the table below, this document provides clear grading criteria for both solicited reactions and lab tests. These guidelines are often included as an appendix in study protocols to align with regulatory expectations.

Focusing on the LB domain, a key challenge arises is that the FDA's "Guidance for Industry[3]" does not consistently use SI units[4]. For example, the FDA mandates that increases in white blood cell (WBC) counts be graded using cells/mm³, while the internationally recognized SI unit (×10⁹/L) is widely adopted globally. These differences highlight the need for careful unit alignment during data submission. Using mismatched units can lead to misinterpretation of safety signals and regulatory issues.

Laboratory Endpoint	**Unit** ***(FDA's unit)***	**Mild (Grade 1)**	**Moderate (Grade 2)**	**Severe (Grade 3)**	**SI unit**
WBC Increase	cells/mm3	10 800 – 15 000	15 001 – 20 000	≥ 20 001	10^9/L
Decrease in WBC	cells/mm3	2500 – 3500	1500 – 2499	≤ 1499	10^9/L

Hemoglobin (female)	gm/dL	10.5 – 10.9	9.0 – 10.4	< 9.0	g/L
Hemoglobin (female) change from baseline value	gm/dL	Any decrease – 1.5	1.6 – 2.0	≥ 2.1	g/L
Hemoglobin (male)	gm/dL	11.5 – 12.4	10.5 – 11.4	< 10.5	g/L
Hemoglobin (male) change from baseline value	gm/dL	Any decrease – 1.5	1.6 – 2.0	≥ 2.1	g/L
Decrease in Absolute Lymphocytes	cells/mm3	750 – 1000	500 – 749	< 500	10^9/L
Decrease in Absolute Neutrophils	cells/mm3	1500 – 2000	1000 – 1499	< 1000	10^9/L
Eosinophils	cells/mm3	650 – 1500	1501 – 5000	> 5000	10^9/L
Decrease in Platelets	cells/mm3	125 000 – 140 000	100 000 – 124 000	≤ 99 999	10^9/L
Creatine	mg/dL	1.5 to 1.7	1.8 to 2.0	≥ 2.1 or requires dialysis	umol/L

UNIT DISCREPANCIES AND REGULATORY REQUIREMENTS

The Center for Biologics Evaluation and Research (CBER) highlighted in SDTCG 4.7 (2021-03) that *"FDA may require laboratory data using conventional units for reviewing submissions and labeling. Sponsors should discuss with the review divisions what laboratory data should utilize conventional units prior to submission."* However, this guidance did not provide specific implementation strategies at the time. In response, the vaccine programming team internally evaluated two potential approaches to address this regulatory requirement.

Approach 1: Vertical Expansion (Row-Based Differentiation)

The first approach involves adding rows to the dataset and introducing a variable like DTYPE (similar to the BDS structure in ADAMIG[5]) to distinguish between unit categories (e.g., SI vs. conventional). This approach aligns with the longitudinal structure of the LB domain, where each row represents a single laboratory test result. For example:

LBTESTCD	LBTEST	LBORRES	LBORRESU	LBSTRESC	LBSTRESU	LBDTYPE
WBC	Leukocytes	4.95	10^9/L	4.95	10^9/L	SI
HGB	Hemoglobin	150	g/L	150	g/L	SI
WBC	Leukocytes	4.95	10^9/L	4950	cells/mm3	CONVENTIONAL
HGB	Hemoglobin	150	g/L	15	g/dL	CONVENTIONAL

Key Features:

- The first two rows represent original SI-unit results (copy form rawdata).
- The next two rows are converted values based on conventional units, used for CBER grading.
- The LBDTYPE variable explicitly labels the unit type (SI or CONVENTIONAL).

Advantages:

- Maintains the standard LB domain structure, where each row corresponds to a single observation.
- Explicitly separates raw data from derived values, enhancing traceability.
- Facilitates regulatory review by providing side-by-side comparisons of SI and conventional units.

Limitations:

- Increases the row count in datasets, potentially complicating data management.
- Requires careful alignment of derived variables (e.g., LBDTYPE) to avoid ambiguity.

Approach 2: Horizontal Expansion (Column-Based Unit Differentiation)

The second method avoids adding rows and instead introduces new variables to store SI and

conventional units in parallel columns. For example:

LBTESTCD	LBTEST	LBORRES	LBORRESU	LBSIRESC	LBSIRESU	LBCVRESC	LBCVRESU
WBC	Leukocytes	4.95	10^9/L	4.95	10^9/L	4950	cells/mm3
HGB	Hemoglobin	150	g/L	150	g/L	15	g/dL

Key Features:

- LBSIRESC/LBSIRESU: SI-unit results (SI = SI).
- LBCVRESC/LBCVRESU: Conventional-unit results (CV = Conventional).
- LBSTRESU: Standard unit (ST = Standard).

Advantages:

- Simpler data structure with no additional rows.
- Easily accommodates future regulatory requirements by horizontally adding new unit variables.

Limitations:

- Does not support conversion factor tracking (e.g., 1 cells/mm³ = 0.001 ×10⁹/L via a defined factor).
- May reduce readability for reviewers expecting a longitudinal structure.

PRACTICAL IMPLICATIONS AND RECOMMENDATIONS

Both approaches have distinct trade-offs:

- Vertical expansion (Approach 1) is methodologically sound, aligning with SDTM best practices for longitudinal data. It ensures clarity in unit differentiation and works well with ADaM for analysis.
- Horizontal expansion (Approach 2) offers operational simplicity, particularly when dealing with complex datasets or multi-regional trials with varying unit preferences. However, it sacrifices the ability to document conversion logic (e.g., mEq/L = mmol/L × valence/ratio).

The choice depends on regulatory needs and data usage:

- For FDA submissions requiring explicit SI/conventional comparisons, Approach 1 provides greater transparency.
- For global harmonization projects, Approach 2 may be preferable to streamline cross-study comparisons.

Ultimately, the programming team is best advised to collaborate closely with statistical and regulatory teams to meet both scientific and submission requirements. This includes pre-submission discussions with CBER to confirm unit preferences and document conversion methods in Define-XML.

FDA'S SDTCG 2024 UPDATE: LC DOMAIN INTRODUCTION

In March 2024, the FDA updated the SDTCG (Standards for Data Tabulation in Clinical Trials Guidance) to version 5.7, introducing the LC (Laboratory Conventional) domain for the first time. This concept was further refined in the latest SDTCG 6.0 (2025-03), which explicitly mandates:

"For clinical studies, submit two separate domains for lab results. The LB domain should contain SI units in LBSTRESU for the SI results in the LBSTRESC and LBSTRESN fields. An additional custom domain called LC, structured identically to LB, should contain conventional units in –STRESU for the results in conventional units in the –STRESC and –STRESN variables. It is ideal if both conventional and SI units come directly from the lab vendor. Submit the results of all tests obtained on subjects, including the

results from unscheduled tests or visits, and results obtained from local laboratories. Identify all reference ranges used for specific populations in the SDRG and ADRG."

This approach aligns closely with Approach 1 (vertical expansion with a DTYPE variable to differentiate units). However, the FDA does not rely on a specific variable (e.g., LBDTYPE) to distinguish SI and conventional units. Instead, it separates results by domain/dataset:

- The LB domain contains SI-unit results (e.g., mmol/L, ×10⁹/L).
- The LC domain contains conventional-unit results (e.g., mEq/L, mg/dL).

Both domains must have identical observation counts, with a one-to-one mapping between LBSEQ and LCSEQ. Below is an illustrative example:

LB domain

LBSEQ	LBTESTCD	LBTEST	LBORRES	LBORRESU	LBSTRESC	LBSTRESU
1	WBC	Leukocytes	4.95	10^9/L	4.95	10^9/L
2	HGB	Hemoglobin	150	g/L	150	g/L
3	PH	pH	7.5		7.5	

LC domain

LCSEQ	LCTESTCD	LCTEST	LCORRES	LCORRESU	LCSTRESC	LCSTRESU
1	WBC	Leukocytes	4.95	10^9/L	4950	cells/mm3
2	HGB	Hemoglobin	150	g/L	15	g/dL
3	PH	pH	7.5		7.5	

Key Observations:

- LBSEQ/LCSEQ = 1 or 2:
    - LBORRES/LCORRES originate from the same EDC (Electronic Data Capture) entry, so LBORRES/LCORRES and their units (LBORRESU/LCORRESU) are identical.
    - Differences in LBSTRESC/LBSTRESU and LCSTRESC/LCSTRESU reflect unit conversions for grading purposes (e.g., ×10⁹/L → cells/mm³ for WBC).
- LBSEQ/LCSEQ = 3:
    - pH has no unit, and for tests where SI and conventional units are equivalent, unit values in LB and LC domains are all empty.

The phrase "*structured identically to LB*" in the SDTCG emphasizes that LC must mirror LB's structure, including the same number of observations. This ensures consistency in data alignment and facilitates downstream analysis.

IS RELREC NECESSARY FOR LB/LC LINKAGE?

A key question is whether the RELREC (Relationship Record) variable is needed to link LB and LC domains. The author concludes that it is not necessary. The relationship between LB and LC is a modeled structure, similar to how DM (Demographics) and DC (Demographics for Multiple Participations) are aligned without RELREC. The LBSEQ/LCSEQ key ensures a direct one-to-one mapping, making RELREC redundant.

Sorting variables like USUBJID, LBTEST, and VISIT typically maintains consistent ordering, so RELREC

is not needed. Even if unit conversions (e.g., g/L → g/dL) affect sorting, the LBSEQ/LCSEQ alignment guarantees data correspondence. The SDTCG does not require RELREC for this use case, confirming its optional status.

The ADLB/ADLC domains in ADaM follow the same principles as LB/LC, ensuring consistent unit handling across SDTM and ADaM. Since the structure and logic are the same, this article does not repeat examples for ADLB/ADLC.

DATA QUALITY AND COMPLIANCE RECOMMENDATIONS

The SDTCG explicitly states: *"It is ideal if both conventional and SI units come directly from the lab vendor."* This underscores a critical data quality and compliance imperative: Prioritize direct vendor-provided data over post-hoc conversions in SDTM code. During study planning, collaborate closely with lab vendors to ensure dual-unit reporting (SI + conventional) for key tests (e.g., WBC, HGB, Creatinine). Additionally, document conversion logic and validation steps in Define-XML to ensure transparency and audit readiness.

CONCLUSION

This article provides a clear framework for addressing unit standardization challenges in clinical trials, especially regarding FDA and SI unit requirements. By adopting the LB/LC dual-domain approach, leveraging vertical/horizontal expansion strategies, and following SDTCG guidelines, SAS programmers can ensure regulatory compliance while maintaining scientific accuracy. The key takeaway is to collaborate early with lab vendors and regulatory teams to align unit preferences. This reduces submission risks and supports the development of safe, effective therapies in a data-driven regulatory environment.

REFERENCES

[1] U.S. Food and Drug Administration. (2025, March 27). Study Data Technical Conformance Guide - Technical Specifications Document (Guidance Document, Docket Number: FDA-2014-D-0092). Center for Drug Evaluation and Research. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/study-data-technical-conformance-guide-technical-specifications-document

[2] Clinical Data Interchange Standards Consortium (CDISC). (2021, November 29). Study Data Tabulation Model Implementation Guide (SDTMIG). https://www.cdisc.org/standards/foundational/sdtmig

[3] U.S. Food and Drug Administration (FDA). (2007, September). Guidance for industry: Toxicity grading scale for healthy adult and adolescent volunteers enrolled in preventive vaccine clinical trials (Center for Biologics Evaluation and Research). https://www.fda.gov/media/73679/download

[4] International Organization for Standardization (ISO). (2020). Quantities and units — Part 1: General (ISO/IEC 80000-1:2020). https://www.iso.org/standard/74598.html

[5] Clinical Data Interchange Standards Consortium (CDISC). (2019, March 1). Analysis Data Model (ADaM) (v2.1). https://www.cdisc.org/standards/foundational/adam