Paper SM02

Decoding NCI-CTCAE: Smarter Lab Grading Through Metadata and Automation

Zulfat Kapadia, Precision for Medicine, Gujarat, India

ABSTRACT

The NCI Common Terminology Criteria for Adverse Events (CTCAE) plays a vital role in standardizing the grading of adverse events, especially laboratory abnormalities. With the release of CTCAE v5.0, the inclusion of baseline values in toxicity grading has introduced new layers of complexity, requiring more advanced and adaptive approaches.

This presentation explores the key challenges in implementing CTCAE guidance, including the integration of baseline data, handling bi-directional grading, and ensuring consistency across evolving standards. We will showcase the development of robust metadata frameworks and reusable SAS macros that streamline and automate the grading process.

By fostering cross-functional collaboration among clinicians, statisticians, and statistical programmers, this approach ensures accurate interpretation of CTCAE guidelines, thus enhancing data integrity, regulatory compliance, and patient safety in global clinical development.

Keywords: CTCAE v5.0, SAS automation, metadata-driven programming, laboratory toxicity grading.

INTRODUCTION

The Common Terminology Criteria for Adverse Events (CTCAE) is an internationally accepted standard for the classification and severity grading of adverse events (AEs) observed in clinical trials, especially in oncology. Developed and maintained by the National Cancer Institute (NCI), CTCAE facilitates consistent and comprehensive safety monitoring by providing a common language to clinicians, researchers, and regulators. Version 5 of CTCAE, released in 2017 with ongoing updates, represents the most current and detailed iteration, expanding the list of AEs, refining severity criteria, and incorporating new event definitions to align with evolving therapeutic modalities.

CTCAE v5 is a pivotal standard employed in clinical trials to systematically categorize and grade adverse events (AEs), including treatment-emergent abnormalities, on a severity scale from Grade 1 (mild) to Grade 5 (death related to the event). The grades indicate the increasing severity of the abnormality or symptom:

- Grade 1 – Mild; asymptomatic or mild symptoms; intervention not indicated
- Grade 2 – Moderate; minimal, local, or noninvasive intervention indicated
- Grade 3 – Severe or medically significant but not immediately life-threatening; hospitalization may be required
- Grade 4 – Life-threatening consequences; urgent intervention required
- Grade 5 – Death related to the adverse event.

This updated version introduces additional complexity by incorporating baseline measurement status into laboratory toxicity grades differing for patients based on whether baseline values were normal or abnormal. Such enhancements facilitate more accurate and clinically relevant toxicity assessment, improving patient safety monitoring and regulatory reporting.

Traditional approaches that rely on embedding study-specific logic directly into programs are not only error-prone but also difficult to maintain and scale across studies. These methods often lead to inconsistencies, increased validation burden, and reduced transparency.

Implementation of laboratory toxicity grading under CTCAE v5 presents notable challenges due to the intricacy of grading criteria, including multiple thresholds and bi-directional comparisons with baseline and reference ranges. To operationalize these criteria efficiently and reproducibly, a metadata-driven SAS macro approach is advantageous over error-prone hardcoding. This approach encapsulates the complex CTCAE grading rules as metadata delineating parameter-specific grading thresholds, directions of worsening, and baseline adjustments. The SAS macros dynamically interpret this metadata to apply toxicity grades to laboratory data programmatically, facilitating automation, consistent application, and ease of maintenance amid evolving CTCAE versions.

We propose a metadata-driven SAS macro framework for CTCAE v5 laboratory toxicity grading that supports baseline-dependent and uni-/bi-directional grading complexities. The framework utilizes standardized metadata inputs

representing CTCAE thresholds and logical conditions, allowing clinical trial programmers to automate laboratory toxicity grading accurately within SAS environments, thus enhancing data quality and timeliness in safety reporting workflows.

This metadata-driven SAS macro approach not only enhances data quality assurance and programming efficiency but also supports regulatory compliance through robust, traceable grading procedures. It represents a significant methodological advancement in clinical safety, providing a foundation for improved clinical trial reporting in oncology and other therapeutic areas.

OVERVIEW OF CTCAE V5.0 LAB TOXICITY GRADING

CTCAE v5.0 encompasses a total of 837 MedDRA lowest-level terms (LLTs), spanning 26 system organ classes (SOCs). Among these, 56 adverse event (AE) terms explicitly incorporate laboratory test results into their grading criteria. Notably, 39 of these terms utilize upper and/or lower limits of normal (ULN/LLN) as part of their grading logic. This highlights the critical role of laboratory data in AE assessment and underscores the need for precise, automated grading mechanisms that can accommodate such complexity within clinical trial safety workflows.

Laboratory toxicity grade reflects both magnitude and direction of abnormality:

- Increases (e.g., ALT, AST, creatinine)
- Decreases (e.g., hemoglobin, calcium)
- Bidirectional markers (e.g., sodium)

Example: CTCAE Grading Logic for “Creatinine Increased”

Grade	Criteria (ULN = Upper Limit of Normal)	Interpretation
1	>ULN - 1.5 x ULN	Mild elevation
2	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	Moderate
3	>3.0 x baseline; >3.0 - 6.0 x ULN	Severe
4	>6.0 x ULN	Life threatening

These criteria are dynamically applied based on metadata mapping within SAS.

CHALLEGES IN IMPLEMENTING LABORATORY TOXICITY GRADING

The implementation of CTCAE (Common Terminology Criteria for Adverse Events) version 5.0 for laboratory-based toxicity grading introduces several technical and clinical complexities. These challenges stem from the variability in grading rules, the contextual nature of some criteria, and dependencies on baseline or institutional reference data. The key challenges are summarized below:

1. **Rule Complexity**
   CTCAE v5.0 grading criteria differ across laboratory analytes, and the structure of these rules is not standardized. Each analyte may follow unique thresholds, conditions, or units, necessitating analyte-specific programming and validation. This lack of uniformity significantly increases implementation and maintenance complexity.

2. **Floating Point Precision**
   Grading calculations often rely on numerical thresholds that can be affected by floating-point precision issues. Minor rounding differences may lead to incorrect grade assignment near threshold boundaries. Careful handling of numerical comparisons, consistent rounding methods, and appropriate tolerance settings are required to ensure reproducible results.

3. **Clinical Context and Conditional Rules**
   Certain CTCAE criteria rely on clinical context rather than numeric thresholds, for example, conditions described as “life-threatening consequences” or “requiring urgent intervention.” These context-dependent criteria are difficult to automate and may require clinician input or rule-based decision support logic to ensure accurate grading.

4. **Baseline Dependence**
   Some grading criteria are baseline-dependent, defined relative to an individual’s pre-treatment value (e.g., “>3× baseline”). The absence of a reliable baseline or upper limit of normal (ULN) introduces ambiguity in determining grades. Handling missing or incomplete baseline/reference data adds an additional layer of complexity to automated systems.

5. **Overlapping Ranges**
   In certain cases, the CTCAE thresholds for different grades overlap, resulting in ambiguity when an analyte value falls within intersecting ranges. Implementations must define and apply consistent rule hierarchy or tie-breaking logic to ensure deterministic grade assignment.

6. **Bi-directional Grading**
   Many analytes have toxicity criteria for both increased and decreased values (e.g., sodium, potassium, glucose). Implementing bi-directional rules requires careful structuring to ensure that both directions are evaluated independently and that only the appropriate grade is assigned for a given deviation.

Implementing CTCAE v5.0 for laboratory toxicity grading demands fine-grained rule design, robust numeric handling, and context-aware logic. Automated systems must be flexible enough to manage analyte-specific nuances while maintaining transparency, auditability, and clinical interpretability.
Metadata-driven SAS macros and standardized ADaM structures (e.g., ADLB) provide robust frameworks for accurate and scalable implementation, ensuring consistency across multi-site or multi-lab clinical trials.

METADATA FRAMEWORK FOR TOXICITY GRADING

A metadata structure for a toxicity grading system is a structured representation of how the framework encodes rules, definitions, and relationships necessary to automate CTCAE-based laboratory toxicity grading. It determines how each data element (such as lab test, baseline, grading thresholds, MedDRA mapping) is captured, controlled, and used for deriving toxicity grades.

Below is the screenshot of metadata structure which will be used to derive laboratory toxicity grading.

CTCAE5Term	LBCAT	LBTEST	LBTESTCD	LBSTRESU	ATOXGR	ATOXGRN	LowOperator	LowLimitAbsolute	LowLimitRelative
Anemia	HEMATOLOGY	Hemoglobin	HGB	g/L	Grade 1	1	GE	100	
Anemia	HEMATOLOGY	Hemoglobin	HGB	g/L	Grade 2	2	GE	80	
Anemia	HEMATOLOGY	Hemoglobin	HGB	g/L	Grade 3	3			
Hemoglobin increased	HEMATOLOGY	Hemoglobin	HGB	g/L	Grade 1	1	GT		ULN
Hemoglobin increased	HEMATOLOGY	Hemoglobin	HGB	g/L	Grade 2	2	GT		(20+ULN)
Hemoglobin increased	HEMATOLOGY	Hemoglobin	HGB	g/L	Grade 3	3	GT		(40+ULN)

CTCAE5Term	ATOXGR	ATOXGRN	HighOperator	HighLimitAbsolute	HighLimitRelative	BASEFLAG	HIGHFLAG	LOWFLAG
Anemia	Grade 1	1	LT		LLN			Y
Anemia	Grade 2	2	LT	100				Y
Anemia	Grade 3	3	LT	80				Y
Hemoglobin increased	Grade 1	1	LE		(20+ULN)		Y	
Hemoglobin increased	Grade 2	2	LE		(40+ULN)		Y	
Hemoglobin increased	Grade 3	3					Y	

CTCAE5Term: MedDRA LLT	**LOWLIMITABSOLUTE**: Starting numeric thresholds not relative to ULN or LLN
LBCAT: Lab Category (Based on CDISC CT)	**LOWLIMITRELATIVE**: Start numeric thresholds relative to ULN or LLN
LBTEST: Lab Test (Based on CDISC CT)	**HIGHOPERATOR**: End operator i.e., LT/LE
LBTESTCD: Lab Test Code (Based on CDISC CT)	**HIGHLIMITABSOLUTE**: End numeric thresholds not relative to ULN or LLN
LBSTRESU: Lab Standard Units (Based on CDISC CT)	**HIGHLIMITRELATIVE**: End numeric thresholds relative to ULN or LLN
ATOXGR: Toxicity Grade	**BASEFLAG**: indicates if baseline values influence grade determination
ATOXGRN: Numeric value of Toxicity Grade	**HIGHFLAG**: determines whether increases trigger grading
LOWOPERATOR: Start operator i.e., GT/GE	**LOWFLAG**: determines whether decreases trigger grading

This metadata structure ensures automated, transparent, and reproducible toxicity grading across CTCAE versions by integrating lab data, thresholds, and derivation logic within a unified metadata repository.
A metadata-driven design centralizes grading logic definitions and allows SAS programs to operate generically across multiple datasets and study protocols. The core advantage of this approach is the ability to modify grading criteria without changing SAS code, ensuring reproducibility, increasing auditability and compliance with CDISC ADaM standards.

METADATA-DRIVEN SAS MACRO FRAMEWORK / IMPLEMENTATION

The macro is a comprehensive SAS-based utility designed to automate the grading of laboratory test results according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. This macro is designed to systematically assign toxicity grades to laboratory test results based on both absolute and relative grading thresholds defined in CTCAE. This macro facilitates the transformation of raw lab data into clinically meaningful toxicity grades (Grade 0 to 4), which are essential for safety monitoring and regulatory reporting in clinical trials. It accepts several parameters, including the input dataset, output dataset, the variable to be graded (typically the numeric lab result), lab category, and the lower and upper limits of normal (LLN and ULN).

The macro begins by preparing a reference dataset from the CTCAE grading metadata. This involves duplication and identification of grading rules that apply to high and low value thresholds for each lab test. It then filters the input lab data to retain only those records with non-missing values and valid units, and merges this with baseline values when available. This baseline information is crucial for grading rules that depend on changes from baseline rather than absolute thresholds.

Next, the macro dynamically calculates the grading thresholds for each lab test. These thresholds can be absolute (e.g., a fixed value) or relative (e.g., 1.5×ULN, +BASE), and the macro uses string parsing and mathematical operations to evaluate these expressions. Once the thresholds are established, the macro compares each lab result against the appropriate grading criteria using logical operators such as greater than or equal to (GE), less than (LT), and so on.

After determining which grading criteria are met, the macro assigns the corresponding CTCAE grade and captures the associated CTCAE term. If no grade is applicable but the lab value falls within the normal range, a Grade 0 is assigned.

Finally, the macro then consolidates the matched CTCAE terms into a single descriptive variable (ATOXDSC) and determines the final toxicity grade (ATOXGR). It also generates additional variables to indicate whether the abnormality was due to a high or low value (HIGHRANGE, LOWRANGE) and provides separate fields for low and high range grades and descriptions (ATOXGRL, ATOXGRH, ATOXDSCL, ATOXDSCH). Numeric versions of the grades (ATOXGRLN, ATOXGRHN) are also derived for downstream statistical analysis. This macro significantly reduces manual effort, ensures consistency with CTCAE standards, and enhances the reproducibility and auditability of lab data grading in clinical research.

The design is fully metadata-driven, supporting bidirectional grading, complex rules, and multiple CTCAE versions without macro code changes. All lab toxicity variables are derived as per ADaMIG v1.3 as shown below screenshot.

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
ATOXGR	Analysis Toxicity Grade	Char		Perm	Toxicity grade of AVAL or AVALC for analysis; may be based on SDTM --TOXGR or an imputed or assigned value.
ATOXGRN	Analysis Toxicity Grade (N)	Num		Perm	Numeric representation of ATOXGR. There must be a one-to-one relationship between ATOXGRN and ATOXGR within a parameter. ATOXGRN cannot be present unless ATOXGR is also present. When ATOXGR and ATOXGRN are present, then on a given record, either both must be populated or both must be null.
ATOXGRL	Analysis Toxicity Grade Low	Char		Perm	Low toxicity grade of AVAL or AVALC for analysis; may be based on SDTM --TOXGR or an imputed or assigned value. Used to assess when a subject's lab value falls within the low toxicity range.
ATOXGRLN	Analysis Toxicity Grade Low (N)	Num		Perm	Numeric representation of ATOXGRL. There must be a one-to-one relationship between ATOXGRLN and ATOXGRL within a parameter. ATOXGRLN cannot be present unless ATOXGRL is also present. When ATOXGRL and ATOXGRLN are present, then on a given record, either both must be populated or both must be null.
ATOXGRH	Analysis Toxicity Grade High	Char		Perm	High toxicity grade of AVAL or AVALC for analysis; may be based on SDTM --TOXGR or an imputed or assigned value. Used to assess when a subject's lab value falls within the high toxicity range.
ATOXGRHN	Analysis Toxicity Grade High (N)	Num		Perm	Numeric representation of ATOXGRH. There must be a one-to-one relationship between ATOXGRHN and ATOXGRH within a parameter. ATOXGRHN cannot be present unless ATOXGRH is also present. When ATOXGRH and ATOXGRHN are present, then on a given record, either both must be populated or both must be null.

Source: CDISC Analysis Data Model Implementation Guide Version 1.3 (Final) 2021-11-29

KEY COMPONENTS

- Metadata storage of grading criteria for each lab test, versioned by CTCAE
- Automated macro logic to interpret and apply metadata rules
- Support for directional grading (e.g., increase, decrease)
- Integration with ADaM clinical data structures
- Handling baseline and normal ranges flexibly

KEY CHALLENGES

- Complexity of Metadata Design: Precisely capturing CTCAE criteria, including directionality and version differences, in metadata tables requires deep domain understanding and meticulous design.
- Baseline Definition: Setting meaningful baselines impacts grading outcomes; variations in baseline definitions across trials need flexible macro customization.
- Version Control: Managing multiple CTCAE versions in metadata without code changes requires robust organization and validation of metadata.

- Exceptions Handling: Some toxicity criteria or trial-specific rules may not fit well into generic metadata-driven logic, requiring manual interventions.
- Data Quality Dependence: Accuracy depends on high-quality input data like ULN, LLN, and baseline flags, making preprocessing critical.

ADVANTAGES OF METADATA-DRIVEN SAS MACRO APPROACH

- Reusability: Single macro framework applies to multiple lab tests and CTCAE versions by updating metadata rather than code.
- Maintainability: Metadata updates can be made without rewriting SAS code, reducing bugs and simplifying validation.
- Consistency: Centralized metadata ensures consistent grading rules across datasets and studies, enhancing data quality.
- Efficiency: Automating complex grading logic reduces programming effort and accelerates study timelines.
- Scalability: Easily extensible to new lab tests or CTCAE criteria changes with minimal disruption.
- Transparency: Metadata files provide a clear, reviewable record of grading criteria for regulatory submissions and audits.

MAIN MACRO CALL

The main call of the macro includes parameters for:

- IN: name of the input dataset (e.g., lab data)
- OUT: name of the output dataset with graded results
- VAR: Variable to be graded (e.g., LBSTRESN, AVAL)
- CAT: Variable to consider for Lab Category (e.g., LBCAT, PARCAT)
- COND: Lab Category to be considered
- LOW: Variable to consider for lower limit of normal (LLN)
- HIGH: Variable to consider for upper limit of normal (ULN)

Here is an example of macro call,

```
%CTCAE (
        in   = lb,
        out  = ctcae_lb,
        var  = aval,
        cat  = lbcat,
        cond = CHEMISTRY,
        low  = anrlo,
        high = anrhi
        );
```

IMPLEMENTATION HURDLES AND SCALABLE STRATEGIES

We encountered a wide spectrum of challenges in implementing automated grading of laboratory toxicities using the CTCAE v5.0, the grading framework has become significantly more complex, demanding a deeper understanding of both clinical context and data structure. What initially seemed like a straightforward task, assigning grades based on numeric thresholds that quickly evolved into a multifaceted problem involving baseline-dependent logic, overlapping grading criteria, categorical lab results, and clinical assessments.

CTCAE grading scenarios can be broadly categorized into three types: (1) those based solely on numeric thresholds, (2) those requiring both numeric and clinical assessment, and (3) those dependent on baseline status. Each category presented unique implementation challenges for automation and required tailored mitigation strategies to ensure accurate, reproducible, and regulatory-compliant grading.

SCENARIO 1. NUMERIC THRESHOLDS ONLY

Some CTCAE terms rely solely on numeric thresholds for grading, making them ideal for automation. An example is Hemoglobin increased as shown in Table 1. These thresholds can be programmed directly using metadata based on defined numeric values.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin increased	Increase in >0 - 2 g/dL	Increase in >2 - 4 g/dL	Increase in >4 g/dL	-

Table 1. Example of grading scenario for Hemoglobin increased

SCENARIO 2. CLINICAL ASSESSMENT - WHEN DATA ALONE ISN'T ENOUGH

Some lab parameters require both numeric evaluation and clinical judgment. These are more complex, introduce complexity into automation workflows and necessitate collaboration between programmers, clinicians, and sponsors.

CASE 1. MUTUALLY EXCLUSIVE NUMERIC THRESHOLDS WITH CLINICAL INPUT

In Case 1, such as in the grading of Activated Partial Thromboplastin Time (aPTT) prolonged as shown in Table 2, Grades 1 and 2 are defined purely by numeric multiples of the upper limit of normal (ULN), Grade 3 includes clinical condition of "bleeding", which cannot be determined from lab data alone. While numeric thresholds can be programmed, clinical input is required to finalize the grade. This necessitates collaboration with sponsors or clinicians to interpret ambiguous cases and ensure accurate grading. Here first approach would be to derive grading solely on numeric thresholds. In this instance, the sponsor deemed the bleeding insignificant or advised to link any reported adverse events. This decision must be clearly documented and communicated to all stakeholders to ensure transparency and consistency.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4
Activated partial thromboplastin time prolonged	>ULN - 1.5 x ULN	>1.5 - 2.5 x ULN	>2.5 x ULN; bleeding	-

Table 2. Example of grading scenario for Activated Partial Thromboplastin Time (aPTT) prolonged

CASE 2. OVERLAPPING NUMERIC THRESHOLDS

Another challenge arises when grading criteria includes overlapping numeric thresholds based on both the upper limit of normal (ULN) and the patient's baseline value. This is exemplified by Creatinine increased shown in Table 3, where Grade 2 is defined as ">1.5–3.0 × baseline or >1.5–3.0 × ULN," and Grade 3 as ">3.0 × baseline or >3.0–6.0 × ULN."

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4
Creatinine increased	>ULN - 1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	>3.0 x baseline; >3.0 - 6.0 x ULN	>6.0 x ULN

Table 3. Example of grading scenario for Creatinine increased

In such cases, a single lab result may satisfy multiple grading criteria. To resolve this ambiguity, the conservative grading principle is applied, assigning the higher grade to avoid underestimating potential toxicity. This approach aligns with regulatory expectations and ensures patient safety. However, it also requires careful documentation and validation to ensure that the logic is consistently applied across datasets and trials.

CASE 3. OVERLAPPING THRESHOLDS WITH CLINICAL INPUT -THE GRAY ZONE

Some grading criteria include both overlapping numeric thresholds and clinical qualifiers. Hyperuricemia is a representative example shown in Table 4. Grade 1 is defined as ">ULN without physiologic consequences," while Grade 3 is ">ULN with physiologic consequences." The distinction between these grades depends on the presence of clinical symptoms such as gout, renal dysfunction, or other metabolic complications.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4
Hyperuricemia	>ULN without physiologic consequences	-	>ULN with physiologic consequences	Life-threatening consequences

Table 4. Example of grading scenario for Hyperuricemia

Since physiological consequences are not typically captured in lab data, clinical input is essential. If such consequences are not documented, the conservative approach is to assign Grade 3. Alternatively, sponsors may define rules to identify related adverse events (AEs) and downgrade the grade if no relevant AEs are found. This scenario underscores the importance of integrating clinical data with laboratory results to ensure accurate grading.

CASE 4. CATEGORICAL LAB RESULTS - MAPPING THE UNSTRUCTURED

Certain lab parameters, such as Proteinuria as shown in Table 5, are reported using categorical or character-based values (e.g., 1+, 2+, 3+), rather than numeric measurements. CTCAE v5.0 provides grading criteria that map these categories to specific grades. For example, 1+ proteinuria corresponds to Grade 1, while 4+ corresponds to Grade 3.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4
Proteinuria	1+ proteinuria; urinary protein ≥ULN - <1.0 g/24 hrs	Adult: 2+ and 3+ proteinuria; urinary protein 1.0 - <3.5 g/24 hrs; Pediatric: Urine P/C (Protein/Creatinine) ratio 0.5 - 1.9	Adult: Urinary protein >=3.5 g/24 hrs; 4+ proteinuria; Pediatric: Urine P/C (Protein/Creatinine) ratio >1.9	-

Table 5. Example of grading scenario for Proteinuria

These character-based results are often not standardized across labs and may not be consistently structured in datasets. While such results are generally not graded programmatically. However, if required by the sponsor or for analysis purposes, these categorical results can be assigned grades in accordance with CTCAE guidance.

SCENARIO 3. DYNAMIC LOGIC REQUIRED - BASELINE-DEPENDENT GRADING

CTCAE v5.0 introduces grading rules that depend on baseline status. For example, “Blood bilirubin increased” shown in Table 6 is graded differently depending on whether the baseline was normal or abnormal.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4
Blood bilirubin increased	>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal

Table 6. Example of grading scenario for Blood bilirubin increased

If the baseline is above the upper limit of normal (ULN), it becomes the new reference point for grading. If the baseline is below ULN, the standard ULN is used. In cases where baseline data is missing, it is assumed to be normal unless otherwise indicated. Similarly, if ULN is missing, the baseline can be used for grading. These rules require dynamic logic in programming and highlight the importance of metadata in defining grading criteria.

An example of above logic implementation is presented in the screenshot below.

USUBJID	PARAMCD	AVISITN	AVAL	BASE	ABLFL	ANRLO	ANRHI	ANRIND	ATOXDSCH	ATOXGRH
002	BILI	10	14	.		.	21	NORMAL	Blood bilirubin increased	Grade 0
002	BILI	20	21	21	Y	.	21	NORMAL	Blood bilirubin increased	Grade 0
002	BILI	50	12	21		.	21	NORMAL	Blood bilirubin increased	Grade 0
002	BILI	100	24	21		.	21	HIGH	Blood bilirubin increased	Grade 1
002	BILI	180	23	21		.	21	HIGH	Blood bilirubin increased	Grade 1
002	BILI	240	15	21		.	21	NORMAL	Blood bilirubin increased	Grade 0

SCENARIO 4. TWO SIDES OF THE SAME COIN - BI-DIRECTIONAL TOXICITY GRADING

Certain laboratory parameters, such as Hemoglobin, necessitate bi-directional toxicity grading to capture both decreases and increases relative to clinical norms or baseline values. For instance, shown in table 7, “Anemia” is graded based on hemoglobin levels falling below the lower limit of normal (LLN), while “Hemoglobin increased” is assessed based on elevations above the participant’s baseline.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4
Anemia	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hgb <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80g/L	Hgb <8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	Life-threatening consequences; urgent intervention indicated
Hemoglobin increased	Increase in >0 - 2 g/dL	Increase in >2 - 4 g/dL	Increase in >4 g/dL	-

Table 7. Example of bi-directional grading scenario for Hemoglobin

Implementing such bi-directional grading requires the definition of direction-specific operator logic and conditional rules that dynamically apply the appropriate grading criteria. This ensures a comprehensive and clinically accurate

assessment of laboratory toxicities, in alignment with regulatory standards such as the ADaM Implementation Guide (ADaMIG) v1.3.

The ADLB dataset should include separate variables for low and high toxicity grading. For low hemoglobin values, the variables ATOXGRL and ATOXDSCL are used to store toxicity grade and description, such as Grade 1 Anemia. For high values, ATOXGRH and ATOXDSCH capture the grade and description for Hemoglobin increased. Baseline grades are stored in BTOXGRL and BTOXGRH.

USUBJID	PARAMCD	PARAM	AVISITN	AVAL	ABLFL	ANRLO	ANRHI	ATOXDIR	ATOXGR	ATOXGRL	ATOXDSCL	ATOXGRH	ATOXDSCH
002	HGB	Hemoglobin (g/L)	10	114		110	161		Grade 0	Grade 0	Anemia	Grade 0	Hemoglobin increased
002	HGB	Hemoglobin (g/L)	20	117	Y	110	161		Grade 0	Grade 0	Anemia	Grade 0	Hemoglobin increased
002	HGB	Hemoglobin (g/L)	50	115		110	161		Grade 0	Grade 0	Anemia	Grade 0	Hemoglobin increased
002	HGB	Hemoglobin (g/L)	100	115		110	161		Grade 0	Grade 0	Anemia	Grade 0	Hemoglobin increased
002	HGB	Hemoglobin (g/L)	130	121		110	161		Grade 0	Grade 0	Anemia	Grade 0	Hemoglobin increased
002	HGB	Hemoglobin (g/L)	180	108		110	161	L	Grade 1	Grade 1	Anemia	Grade 0	Hemoglobin increased
002	HGB	Hemoglobin (g/L)	240	104		110	161	L	Grade 1	Grade 1	Anemia	Grade 0	Hemoglobin increased
002	HGB	Hemoglobin (g/L)	300	102		110	161	L	Grade 1	Grade 1	Anemia	Grade 0	Hemoglobin increased
002	HGB	Hemoglobin (g/L)	360	108		110	161	L	Grade 1	Grade 1	Anemia	Grade 0	Hemoglobin increased
005	HGB	Hemoglobin (g/L)	10	161		130	175		Grade 0	Grade 0	Anemia	Grade 0	Hemoglobin increased
005	HGB	Hemoglobin (g/L)	20	160	Y	130	175		Grade 0	Grade 0	Anemia	Grade 0	Hemoglobin increased
005	HGB	Hemoglobin (g/L)	50	.		.	.						
005	HGB	Hemoglobin (g/L)	100	158		130	175		Grade 0	Grade 0	Anemia	Grade 0	Hemoglobin increased
005	HGB	Hemoglobin (g/L)	180	177		130	175	H	Grade 1	Grade 0	Anemia	Grade 1	Hemoglobin increased

As shown in above screenshot, a participant with a hemoglobin value of 104 g/L is assigned to Grade 1 Anemia (ATOXGRL = Grade 1) and Grade 0 Hemoglobin increased (ATOXGRH = Grade 0). Conversely, a participant with a value of 177 g/L would be graded as Grade 1 Hemoglobin increased and Grade 0 Anemia. This dual grading ensures that both directions of toxicity are evaluated independently and accurately.

In some implementations, a stacked format may be used where each record represents a single direction of toxicity. In such cases, ATOXGR is used to store the grade and ATOXDIR indicates the direction ("LOW" or "HIGH"). This format is useful for simplifying reporting and visualization, but care must be taken to preserve traceability and ensure accurate derivation logic.

This implementation aligns with ADaMIG v1.3 by maintaining one record per lab test per participant per visit, using analysis-ready variables, and ensuring traceability to SDTM source data. Programming logic must apply CTCAE grading rules separately for low and high values. Automation using macros or scripts is recommended to ensure consistency and scalability across studies, especially when handling large volumes of lab data with potential bi-directional toxicity.

CONCLUSION

Automating CTCAE v5.0 laboratory toxicity grading is a nuanced endeavor, far from a one-size-fits-all solution. It demands a deliberate fusion of rule-based programming, robust metadata architecture, and close clinical collaboration. By adopting a conservative grading philosophy and designing a flexible, metadata-driven framework, we have successfully streamlined the toxicity grading process while upholding the highest standards of accuracy, transparency, and regulatory compliance.

The implementation demonstrates that a scalable and reproducible approach can effectively address the multifaceted challenges of CTCAE v5.0, including baseline-dependent logic, bi-directional grading, and context-sensitive clinical rules. The framework's adaptability allows for seamless integration of future CTCAE updates and study-specific requirements, without the need to alter core programming logic.

Beyond technical efficiency, this methodology fosters cross-functional synergy between clinical and statistical teams, ensuring that both clinical relevance and analytical precision are preserved. As clinical trials continue to evolve in complexity and scale, such intelligent automation strategies will be instrumental in advancing safety analytics, accelerating data-driven decision-making, and supporting high-quality, patient-centric research.

REFERENCES

NIH National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0. November 27, 2017
https://dctd.cancer.gov/research/ctep-trials/for-sites/adverse-events/ctcae-v5-5x7.pdf

CDISC Analysis Data Model Implementation Guide Version 1.3 (Final) 2021-11-29
https://www.cdisc.org/standards/foundational/adam/adamig-v1-3

Xiaoyin (Sherry) Zhong. “Get to the Bottom of Lab Toxicity Grading: Challenges and Implementation of CTCAE Version 5” PHUSE 2020
https://www.lexjansen.com/phuse-us/2020/dh/DH16.pdf

Keith Shusterman, Mario Widel. “Implementing Laboratory Toxicity Grading for CTCAE Version 5” PharmaSUG 2019
https://www.pharmasug.org/proceedings/2019/BP/PharmaSUG-2019-BP-128.pdf

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Surabhi Ahluwalia, Hemangini Gurlani, Anitha Priyadharshini, Brian Beus, Harshil Modi, Ryan Kurz, and Bob Hull for their valuable insights and contributions throughout the development of this work. I also extend my thanks to the PFM leadership team for their continued support and endorsement, which have been instrumental in bringing this initiative to fruition.

RECOMMENDED READING

NIH National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) Version 6.0. July 22, 2025
https://dctd.cancer.gov/research/ctep-trials/for-sites/adverse-events/ctcae-v6.pdf

CONTACT INFORMATION

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

Zulfat Kapadia
Precision for Medicine
12th floor, A wing, Krish Cubical,
Sindhu Bhavan Marg, Thaltej,
Ahmedabad, Gujarat 380059
kapadia.zulfat@precisionformedicine.com