

Laboratory Test Verification and Validation Toolkit

Clinical laboratory testing should be accurate and reliable to ensure correct diagnosis and treatment. Laboratories that perform testing on human specimens for the purpose of diagnosis, treatment or mitigation of a disease or disease state must demonstrate that new instruments and tests have acceptable performance prior to implementation. Studies that demonstrate the performance characteristics of the instrument or assay must be documented and accessible for external inspections and assessments. This process is known as verification or validation.

The purpose of this toolkit is to assist laboratories in determining the difference between a validation and a verification, when each should be performed, and to provide guidance on how to perform a verification or validation. This toolkit will help laboratories plan, implement, and analyze data associated with new testing methods (FDA-approved tests and FDA-modified or laboratory developed tests) and technology implementation. In addition, the toolkit serves as a guide to help determine if the quality requirements for establishment of a new test system have been met. These recommendations should be followed to ensure a new or modified method, analyte or instrumentation (hereto referred to as test method) has acceptable performance.

This toolkit is the result of collective knowledge of subject matter experts and contains guidance for performing verifications and validations that will be compliant with requirements defined by the Clinical Laboratory Improvement Amendments (CLIA). Many of the described processes may be useful for other regulatory requirements. In addition to the templates and examples for CLIA verifications and validations, a list of additional resources for CAP, ISO, FSIS and FDA regulations and references to TNI documents is included.

This document combines all eight sections of the toolkit.

Download individual sections from:
aphl.org/VV-Toolkit

Inside the Toolkit

Verification and Validation 101	3	Safety Considerations and Risk Assessments	26
Verification/Validation Process Checklist ..	6	Cost Analysis and Budget	29
Obtaining Appropriate Test Samples	9	Reference Documents	31
Qualitative Assays.....	11	Glossary & Acronyms.....	33
Quantitative Assays	16	References	36
Related Processes.....	20		

Table of Contents

Verification and Validation 101	3	Related Processes	20
Verification or Validation Process Determination	3	Sample or Analyte Stability and Integrity	20
Required Performance Characteristics Determination ..	3	Bridging Studies and Addendums	20
General Steps of a Verification or Validation Plan	5	Instrumentation Verification or Validation	21
Verification/Validation Process Checklist ..	6	Reagent Lot-to-lot Comparison	22
Process Checklist	6	Software Upgrades	24
Additional Process Resources	8	Additional Related Processes Resources	25
Obtaining Appropriate Test Samples	9	Safety Considerations and Risk Assessments	26
Sample Types	9	Risk Assessment Considerations	26
Sample Numbers	10	Risk Assessment Tables	27
Sample Volume	10	Frequency for Reviewing a Risk Assessment	28
Qualitative Assays	11	Hazardous Waste Management	28
Controls	12	Additional Risk Assessment Resources	28
Accuracy	12	Cost Analysis and Budget	29
Precision	13	Labor Costs	29
Analytical Sensitivity	14	Materials Costs	29
Analytical Specificity	14	Site Preparation Costs	30
Reportable Range	15	Maintenance Costs	30
Reference Range	15	Reference Documents	31
Additional Qualitative Assay Resources	15	Templates	31
Quantitative Assays	16	Examples	31
Calibration and Statistical Calculations	16	Additional Resources	32
Quantitative Controls	17	Glossary & Acronyms	33
Accuracy	17	CLSI Harmonized Terminology Database	33
Precision	18	Glossary	33
Analytical Sensitivity	18	Acronyms	35
Analytical Specificity	18	References	36
Reportable Range	18		
Reference Range	19		
Ongoing Verification of Quantitative Assays	19		

Verification and Validation 101

Laboratory testing compliance requirements are defined by the Clinical Laboratory Improvement Amendments (CLIA) regulations in [Title 42 of Code of Federal Regulations Section 493, 42 CFR 493](#).¹ The (CLIA) regulatory requirements related to establishment and verification of performance specifications of clinical test systems prior to reporting patient test results are found in [Section 493.1253](#).²

CLIA defines a 'test system' as instructions and all the instrumentation, equipment, reagents, and supplies needed to perform an assay or examination and generate test results. As explained by CLIA in their [Survey Procedures and Interpretive Guidelines](#),³ a clinical test system verification or validation may be required for:

- Any test system first used in a laboratory to measure a new analyte
- A test currently performed by a laboratory but on a new test system
- An analyte added to a system currently used by the laboratory to perform other testing,
- A modification to a test system already being used in the laboratory (e.g., different specimen type or specimen volume) or
- Multiple instruments used to perform the same tests.

The implementation of a new or modified test method to the laboratory test menu involves a step-by-step process to demonstrate the performance of the test and is required for a CLIA Certificate of Compliance (CoC) or Certificate of Accreditation (CoA). There are two different processes:

- **Verification:** The one-time process by which a laboratory determines that an unmodified US Food and Drug Administration (FDA) cleared or approved test performs according to the manufacturer's specifications when used as directed.
- **Validation:** The process used to confirm with objective evidence that a laboratory-developed test (LDT) or modified FDA-cleared or approved test method or instrument system delivers reliable results for the intended application.

Laboratories that operate under a CoA are subject to the accrediting body regulatory requirements such as College of American Pathologists (CAP) which, in some cases, may be more stringent than CLIA. In these instances, the laboratory should ensure compliance by confirming these requirements with the accrediting organization.

Verification or Validation Process Determination

The choice of which process is followed depends on the new or modified method that will be implemented, which will dictate the type of performance characteristics to evaluate. **Table 1** shows whether a verification or validation is recommended based on the category of test. The FDA maintains [searchable databases](#) to determine if a medical device is approved, cleared or authorized.⁴

Required Performance Characteristics Determination

A range of performance characteristics may be required to be verified or validated (see the **Glossary (page 33)** for definitions):

- Accuracy
- Precision (reproducibility)
- Analytical sensitivity
- Analytical specificity
- Reportable range/intervals (normal values)
- Reference range
- Any other performance characteristic required for test performance

Requirements may depend on intended use, type of assay, applicable manufacturer's studies and EUA instructions. **Table 2** outlines the performance characteristics generally required for test systems; more information about requirements is provided in the **Qualitative Assays (page 11)** and **Quantitative Assays (page 16)** sections.

There are other variables in a verification or validation process such as the sample type(s), sample number(s) (see **Obtaining Appropriate Test Samples (page 9)**) and whether the test method is a **qualitative** or **quantitative assay**, which are discussed in other sections of this toolkit.

Table 1. Recommended Process Based on Categories of Laboratory Tests

Method Type	Definition	Recommended Process
FDA Approved	The device has been approved through the Premarket Approval (PMA) Process , which evaluates the safety and effectiveness of Class III medical devices.	Verification
FDA Cleared	The device has been cleared as a substantially equivalent to a legally marketed device through Section 510(k) of the Food, Drug and Cosmetic Act.	Verification
FDA Authorized (EUA)	During a declared public health emergency, a device that is neither authorized nor cleared, has been evaluated by FDA through the Emergency Use Authorization (EUA) process and found acceptable for use to prevent serious or life-threatening diseases when no alternatives exist.	Verification*
FDA Modified	Any modification to an FDA Approved or Cleared test. Modifications can include the intended use, sample types, patient age, collection device, etc. These modifications require that the method is evaluated as a High Complexity laboratory developed test.	Validation
Laboratory Developed Test (LDT)	A test used for analyzing samples that is: 1. Performed by the clinical laboratory that developed the test and 2. Is neither FDA approved nor FDA cleared, or 3. Is an FDA approved or FDA cleared test that has been modified. This may include analyte specific reagents (ASR) or adoption of another laboratory's LDT or non-cleared or approved test. These tests are considered high complexity.	Validation

Table 2. General Summary of Performance Characteristics Required for Test Systems

Test Type	Accuracy	Precision	Analytical Sensitivity	Analytical Specificity	Reportable Range	Reference Range
FDA Approved	Required	Required			Required	Required
FDA Cleared	Required	Required			Required	Required
FDA Modified or LDT	Required	Required	Required	Required	Required	Required
FDA Authorized (EUA)**	Required	Required			Required	Required

* FDA authorized methods require minimal evaluation of performance, typically accuracy and precision. If an EUA expires, a complete validation is required if the method is not approved or cleared.

** Requirements may vary depending on EUA. Performance characteristics will be defined by the EUA Instructions for Use and the laboratory director. Learn more about implementing a test under emergency use conditions at [cfsi.org/standards/products/method-evaluation/documents/ep43/](https://www.fda.gov/standards/products/method-evaluation/documents/ep43/)

General Steps of a Verification or Validation Plan

After determining whether a verification or validation is needed and which performance characteristics must be verified, the following are the general steps of a verification or validation plan:

1. Develop a plan or proposal

- Reason for study
- Safety Considerations
- Methodology
- Acceptance criteria
- Data analysis and acceptance plan

2. Plan approval by laboratory director

3. Initiate plan

4. Analyze

5. Re-evaluate and modify, if necessary

6. Complete testing and create a summary report

7. Complete additional supporting documents

- SOPs
- Training
- LIMS updates
- Other

8. Summary report review and approval by laboratory director

9. Test implementation

Details on these steps are available in other sections of this toolkit.

Verification/Validation Process Checklist

This section of the Verification and Validation Toolkit provides a checklist that walks users through test verification or validation plan development, plan initiation, creation of the testing and summary report and test implementation. The steps below will assist in determining if a verification or validation needs to be performed, and the events or steps that should occur for implementation of a new test method.⁵

This section also includes **Additional Process Resources (page 8)** with editable templates and examples related to the verification/validation process, including an [editable version of this checklist](#).

Process Checklist

Choose a Verification or Validation Process

Refer to **Verification and Validation 101 (page 3)** to determine which approach is appropriate.

Develop a Plan/Proposal

The following is a general outline for a test verification or validation plan or proposal:

I. Introduction and Reason for the Study

- May include the purpose and scope, definitions, literature review, staff roles and responsibilities, and location of laboratory (for laboratories operating under multi-site certificates)
- Verification of FDA-cleared test, modification of FDA-cleared test, or validation of new LDT, etc.
- Statement regarding internal review board (IRB) approval requirement or request needed
- Whether there is an established SOP or associated job aids, or if they need to be created
- How test results will be used (screening, diagnostic, confirmatory or monitoring)

II. Safety Considerations

- Consider the inclusion of the laboratory's safety officer during the planning stages
- Biorisk assessment
- Chemical risk assessment
- Hazardous waste management

III. Methodology for Each Performance Characteristic Evaluated

- Which performance characteristics will be evaluated is dependent on qualitative or quantitative testing (see **Qualitative Assays (page 11)** and **Quantitative Assays (page 16)** sections)
- Number and type of samples (See **Obtaining Appropriate Test Samples (page 9)**)
- Testing procedure
- Whether the test will be run on multiple instruments
- Controls (including calibrators for quantitative testing)
- Whether an **Individualized Quality Control Plan (IQCP)** is appropriate
- Gold standard comparator
- Environmental and storage conditions
- Limiting factors indicated in FDA EUA or approval, or as defined per manufacturer

IV. Acceptance Criteria

V. Data Analysis and Evaluation Plan

Get Plan Approval from Laboratory Director

Additional individuals may need to pre-approve the plan prior to the laboratory director, such as quality assurance officers or unit supervisors or managers.

Initiate Verification/Validation Plan

- Ensure there are sufficient quantities of reagents, consumables and personnel time prior to executing the verification or validation plan.
- Ongoing evaluation of data obtained during the verification or validation should occur to determine if the planned testing method needs modification. If modifications are identified, it is critical that any changes to the proposed plan be documented, technically justified, reviewed and approved.
 - Are the acceptance criteria being met?
 - Are the samples appropriate?
 - Are there any limitations to collecting the appropriate sample amounts?
 - Are there matrix issues?
 - Are there cross reactivity or interference issues?
 - Will there need to be additional testing performed beyond what was proposed in the plan?

Complete Testing

Create Testing and Summary Report

The following is a general outline for a test verification or validation summary report:

I. Overall Data Summary

- May include overall summary of data and acceptance criteria results
- Statement of any modifications made to the validation plan with rationale

II. Performance Data Results

- Results with supporting traceability to repeat the testing under conditions as close to the original as possible
- Deviations to sample size and acceptance criteria from the plan that were made along with a justification. If deviations were made, a statement of impact or variance should be included.
- Calculations of the performance characteristics
- Explanation of discrepancies
- Instrument-to-instrument comparison (if applicable)
- Limitations

III. Conclusion

- Statement regarding the acceptability of the method, its fitness for use, and any clinical claims that will be made (if applicable).
- Recommendations of changes that need to be made to the test process based on the evaluation.
- If the study showed that the test method was not acceptable or could not be properly performed, documentation of the corrective action steps taken and its approval by the responsible laboratory leadership.

Complete Associated Documents

- Pertinent SOP or SOP edits
- Individualized Quality Control Plan (IQCP), if applicable
- Training documentation and competency assessments
- Related equipment functionality and maintenance documentation
- Any job aids or other documentation necessary to routinely perform testing
- LIMS updates and verification

Get Summary Report Approval from Laboratory Director

Approval of the verification or validation summary report by quality assurance officers, unit supervisors, or managers may be necessary prior to laboratory director approval.

Test Implementation

Once the verification or validation summary report is approved at the appropriate level, the laboratory can move on to implementation activities.

The following considerations should be made, and if applicable be included in the verification or validation plan:

- Adding the procedure to the document control system and to the test menu
- Adding new equipment to the laboratory's maintenance and calibration plan (if not already completed)
- Training staff members
- Updating scope of service documents, as applicable
- Delegation statement if responsibilities are delegated by laboratory director
- Scheduling proficiency testing or comparison studies
- Updating the laboratory quality assurance plan and quality control processes to include continuous monitoring
- Communicating implementation plan to submitters
- Confirming electronic patient test result output with submitters

Additional Process Resources

Templates

- [Editable Verification/Validation Process Checklist](#)
- [Method Verification/Validation Plan Approval Checklist](#)
- [Verification Plan Template](#)
- [Validation Plan Template](#)
- [Verification Summary Report Template](#)
- [Validation Summary Report Template](#)

Examples

- [Employee Training Verification Checklist Example](#)
- [Guidelines for Verification and Validation of Laboratory Methods](#) (Minnesota)
- [Method Validation-Verification Summary Report Example](#) (Fairfax County, VA)
- [Method Verification Template Example](#) (Indiana)
- [Validation and Verification SOP Example](#) (Texas)
- [Validation Plan Example](#) (Washington)
- [Validation Summary Report Example](#) (Washington)
- [Verification Plan Example](#) (Washington)

Obtaining Appropriate Test Samples

This section of the toolkit provides information on sample types, sample numbers and sample volume.

Sample Types

Selection, sourcing or creation of the experimental samples plays a critical role in the verification or validation study of new, modified or laboratory-developed test systems and instrumentation. These samples will serve as the gold-standard comparator for validations or verifications. Sources of samples can be commercially available calibrators/calibration or quality control materials with known values, proficiency testing materials that have established values, or previously tested patient specimens with established values.

Potential clinical samples must be sufficiently characterized, or contrived samples thoughtfully designed, to allow for predictions or expectations of their performance. Poorly chosen samples could increase the rate of false positives or negatives leading to a failure in assay validation, verification, or result in misleading performance specifications. Poorly characterized samples could also introduce unknown variables or unexpected complications that compromise any derived claims based on the experimental data.

General considerations for choosing sample types:

- Examine manufacturer's statements of intended use, claims, approved specimen types and collection devices, recommended specimen preparation/handling, instructions on operating the assay, storing essential materials and any restrictions or limitations. Manufacturers' claims often exclude certain patient populations; attempts to include indicated manufacturer exclusions or clear omissions require validation and samples must be sourced from the population.
- Qualitative assay sample selection is dependent on the presence or absence of the target analyte. Knowledge of the target analyte concentration is not essential, but efforts should be made to acquire samples that span the detection range (e.g., low, medium, or high positives). Alternatively, limiting the range to clinically relevant concentrations is a viable option.
- Quantitative assay sample selection requires knowledge of the target analyte concentration. Reportable data is quantified, and an acceptance criterium can be imposed at the distinct stages of assessment (i.e., accuracy, precision, etc.).
- Clinical samples are the preferred sample type to simulate the conditions during routine testing. These samples should be well characterized including patient demographics and sample or analyte integrity. Refer to **Related Processes (page 20)** to assess sample or analyte stability and integrity.
- Contrived samples may be utilized when clinical samples are unavailable or not suitable for the type of testing being performed. Creation of contrived samples includes spiking, or introducing, the target analyte into a matrix that is initially known not to contain the target analyte. This matrix may be an individual or homogenized (pooled) clinical sample, buffer, medium or other inert substances.

Sample Numbers

The total number of samples tested for a given study is dependent upon the evaluator's intended use of the method, the recommended statistical analysis, and the sample population. The Clinical and Laboratory Standards Institute (CLSI) has published method evaluation standards that provide explanations and instructions for evaluation of test method performance characteristics such as accuracy and precision. [CLSI EP09c](#)⁶ recommends analyzing a minimum of 40 samples by both test and reference method. Ultimately, sample number is at the discretion of the CLIA Laboratory Director, and 40 samples may not be appropriate depending on the evaluator's intended use. For studies utilizing smaller sample sizes, greater statistical implications will be seen when outliers or non-correlating results are present, requiring a smaller margin of error to achieve an anticipated confidence interval. The larger the sample size, the more confident one can be that the results are reflective of the population.

Sample numbers may vary based on a number of scenarios, including:

- Category of laboratory test performed
 - FDA-approved or cleared or authorized verifications may state suggested sample numbers, but often require fewer samples than a validation.
 - FDA modified and laboratory-developed tests require validations and often larger sample sizes than a verification.
 - EUA tests often provide guidance for verification testing regarding sample size in the instructions for use (IFU) documentation.
- **Qualitative Assays (page 11) vs. Quantitative Assays (page 16)**
- Difficult to acquire specimens (e.g., rare positives, rare organisms, etc.)
- Costly test reagents (e.g., whole genome sequencing)
- Instrument verification or validation (see **Related Processes (page 20)**)
- Reagent comparisons (e.g., discontinued products) (see **Related Processes (page 20)**)

Sample Volume

Sufficient sample volume is a major consideration when selecting samples for a verification or validation. The laboratorian should consider the volume needed to evaluate the assay including replicate testing, use of multiple instruments and the potential need for repeat testing. Insufficient sample volume for repeat testing and/or the repeat of discrepant results could lead to insufficient investigation into any method error and, in turn, result in insufficient data to validate or verify a testing method. If comparison studies are being included as part of the validation process, the laboratory should take into consideration sample volume requirements for any additional test systems involved.

If the test method design calls for multiple replicates and instrumentation, it may be difficult to acquire sufficient volumes from clinical samples. Therefore, contrived or alternative samples may be ideal in these situations as volume issues can typically be avoided if sufficient large batches are prepared or purchased.

Qualitative Assays

This section of the toolkit provides information on the selection and frequency of quality controls and determining performance characteristics (accuracy, precision, sensitivity, specificity, and reportable and reference ranges) for qualitative assays. Find checklist examples in **Additional Qualitative Assay Resources (page 15)**.

Qualitative assays are methods that provide only two categorical results (i.e., positive or negative; present or absent; reactive or nonreactive; yes or no). Some qualitative assays have no numerical value associated with the result whereas other assays are labeled as qualitative because one of only two results is reported (i.e., positive or negative) even though a numerical value is derived. The overall objective of qualitative assays is to recognize the presence or absence of an analyte. In qualitative assays, the cutoff value is defined as the threshold above which the result is reported as positive and below which the result is reported as negative.

Clinical laboratory uses for qualitative assays are described as screening, diagnostic, confirmatory or monitoring.⁷ The utility of a given assay is determined based on the sensitivity and specificity, predictive values, and the prevalence of disease or condition in the population tested.

- **Screening Methods:** Use to test a population subset for the presence or absence of an analyte or agent.
- **Diagnostic Methods:** Use clinical suspicion of a particular disease or condition to guide testing. Both screening and diagnostic assays should have high sensitivity; lower specificity is tolerated if a confirmatory test is available, and the results are low consequence.
- **Confirmatory Methods:** Follow screening or diagnostic test results and enable clinicians to establish a diagnosis with testing that is designed to be specific, sometimes at the expense of sensitivity, and have a high **Positive Predictive Value (PPV)**.

The type of verification or validation for qualitative assays is dependent on clearance or approval from a regulatory entity (i.e., **FDA Cleared** or **FDA Approved**). In general, refer to the assay IFU documentation to determine the number and type of samples to use for the verification or validation. Additional guidance recommendations are provided in this toolkit.

Table 3. Summary of Performance Characteristics Required Depending on Qualitative Test Type²

Test Type	Accuracy	Precision	Analytical Sensitivity	Analytical Specificity	Reportable Range	Reference Range
FDA Approved	Required	Required			Required	Required
FDA Cleared	Required	Required			Required	Required
FDA Modified or LDT	Required	Required	Required	Required	Required	Required
FDA Authorized (EUA)*	Required	Required			Required	Required

* Requirements may vary depending on EUA. Performance characteristics will be defined by the EUA IFU and the laboratory director. Instructions should include requirements for verification or validation.

Controls

Positive and negative controls should be chosen such that they provide expected results when the test is functioning properly. Control design near the cutoff value can detect more errors; however, can lead to rejection of a test run that does not have significant errors. To determine an optimal set of controls, use the provided guidance by the manufacturer, stable commercial or clinical controls, or perform a precision experiment to understand the imprecision of the assay (refer to [CLSI EP12⁷](#) for performing an imprecision experiment). Per [CLIA §493.1256](#),² a laboratory must not use control materials outside the patient reportable range. Control samples not containing the analytes or substances to be controlled are not acceptable as control material.

For most qualitative assays, it is acceptable to perform a negative and positive quality control daily, while other testing methods may require more frequent testing of controls on a per run basis (check the IFU per method). Verification or validation of an assay can help determine the performance of controls and ascertain the frequency required for addition of controls to a given assay. Depending on the assay, the laboratory could customize its QC plan using an **IQCP**.^{8,9} If controls fail to produce the expected results, the run must be rejected, and the failure should be investigated to identify the cause.

Accuracy

For qualitative assays, accuracy studies should validate if the test method detects the presence or absence of the analyte. Sources vary on recommended number of samples to test for accuracy. If there is no guidance from the IFU, [CLSI EP09c⁶](#) suggests 40 total samples (20 each of positive and negative value) should be tested. This number is a minimum suggestion, and an assessment should be conducted to determine if more samples are necessary to increase the statistical relevance. In addition, the clinical impact and repercussions or consequences to the patient of a false negative or false positive result is critical in determining the number of samples for testing. If fewer than 40 samples are all that can be obtained, prior approval by the laboratory director or quality assurance officer should be obtained.

Accuracy testing should be performed over a minimum course of five days to simulate a range of conditions over which samples would normally be run.

2 x 2 Contingency Table

The table in **Figure 1** can be used to calculate the estimated sensitivity, specificity, total accuracy, PPV and negative predictive value (NPV). Results should correlate with an expected total accuracy of $\geq 95\%$ agreement with the reference method. Westgard QC has a [2 x 2 contingency calculator](#).*

Kappa Coefficient

If an imperfect standard is being used for the verification or validation, the overall agreement of the assay can be calculated using the Kappa coefficient. The Kappa coefficient is a measurement of the degree of agreement between the methods above what is expected by chance alone.

Use the formulas described below in **Figure 2** in association with the values from **Figure 1**. An online calculator is available at [graphpad.com](#).**

Understanding Kappa:

- A Kappa of, or approaching, one indicates that there is very good agreement
- A Kappa approaching zero indicates that the agreement is no better than chance.
- A negative Kappa means that the agreement is worse than chance.

* www.westgard.com/qualitative-test-clinical-agreement.htm

** www.graphpad.com/quickcalcs/kappa1.cfm

Figure 1. 2 x 2 Contingency Table

		Diagnostic Accuracy Criteria		
		Positive	Negative	Total
Method X	Positive	# true positive (TP)	# false positive (FP)	TP + FP
	Negative	# false negative (FN)	# true negative (TN)	FN + TN
	Total	TP + FN	FP + TN	N

Estimated Sensitivity

$$= 100 \times \frac{TP}{TP + FN}$$

Estimated Specificity

$$= 100 \times \frac{TN}{FP + TN}$$

Positive Predictive Value (PPV)

$$= 100 \times \frac{TP}{TP + FP}$$

Negative Predictive Value (NPV)

$$= 100 \times \frac{TN}{FN + TN}$$

Total Accuracy

$$= 100 \times \frac{TP + TN}{TP + FP + FN + TN}$$

Figure 2. Calculating the Kappa coefficient

$A = TP + \frac{FP}{N}$ <hr/> $B = TP + \frac{FN}{N}$	$Pr(e) = (A \times B) + [(1-A) \times (1-B)]$ <hr/> $Pr(a) = 100 \times \frac{(TP + TN)}{N}$ <i>(overall percent agreement)</i>	Kappa (K) $= \frac{Pr(a) - Pr(e)}{1 - Pr(e)}$
---	---	---

Precision

Samples for precision should be near the high and low cutoff values to provide the best estimation of error at medically relevant decision levels. A minimum of one positive and one negative sample is recommended with a total of 10-30 measurements. Precision testing should be performed over a course of multiple days using more than one laboratorian to demonstrate reproducibility. The experimental design consists of the following precision measurements:

- **Intra-assay** (within run): Same samples run multiple times on the same run and day
- **Inter-assay** (between run): Same samples run in different runs on the same day or different days, and preferably by a different laboratorian.

If any of the aforementioned precision measurements are not applicable to a given assay, discuss with the quality assurance officer (QAO) or laboratory directory to determine feasibility or requirements of testing. The manufacturers' statements of precision should be used as a minimum performance requirement. Alternatively, if numerical data are available and standard deviation can be calculated, the coefficient of variation (CV) can be used to express the precision and repeatability of an assay.

CV Calculation:¹⁰ The coefficient of variation is the ratio of the standard deviation to the mean. CV is expressed as a percentage. The ideal CV is <15%, and generally should not exceed 20%.

Figure 3. Calculating the Coefficient of Variation

$$CV \% = 100 \times \frac{\text{Standard Deviation}}{\text{Mean}}$$

Imprecision of the method can be analyzed using concentrations near the cutoff. However, it is not appropriate to measure the imprecision of qualitative assays with low-negative or high-positive samples since these values are usually too far away in analyte concentration from the medical decision point. Details regarding how to perform a qualitative method precision experiment to understand imprecision of an assay can be found in [CLSI EP12](#) section 8.3.⁷

Analytical Sensitivity

Analytical sensitivity is referred to as 'limit of detection studies.' Limit of detection (LoD) seeks to define the lowest concentration of an analyte in a matrix that can be consistently detected. For LDT's, this measurement must be established during the method validation. For other assays, the manufacturer has completed LoD studies.

Analytical sensitivity for qualitative assays can be challenging to complete and will not always provide a definitive quantity as the LoD. Some qualitative methods will have a measurable value (i.e., cycle threshold, optical density, titer, colony forming unit, etc.), or measurand, that is subsequently used to determine the qualitative test result. In these cases, [CLSI EP17:A2](#)¹¹ recommends making serial dilutions of a sample with a known measurand content in replicate. The samples should be run in duplicate or triplicate over three days with a recommended minimum of 20 measurements for each sample concentration to verify a manufacturer claim, and 60 measurements to establish the LoD. However, the exact number of samples to use should be determined on a case-by-case basis with input from the QAO, supervisors, or the laboratory director. If the laboratory wants to establish a precise LoD, the laboratorian calculates and plots the hit rate for each dilution, which is defined as the total number of positive results divided by the total number of replicates, using regression modeling with hit rate on the y-axis and measurand dose on the x-axis. The laboratorian then selects the hit rate that corresponds to detection of the analyte in the majority of samples (i.e., 95%) as the LoD. A probit fit analysis using computer software can be used to easily perform this calculation. A [detailed description of probit analysis](#) is found on the Westgard QC website.¹²

Westgard offers an alternative and simpler approach to address LoD in qualitative assays that have a measurand that is assumed to be continuous along a range of concentrations. If the assay cutoff is known, the laboratory can test samples that are expected to fall below and above the cutoff in replicates of 20. Next, the number of positive results are evaluated. No more than 5% of replicate results below the cutoff value should be positive. Conversely, at least 95% of replicate results above the cutoff value should be positive. These calculations correspond to the 95% confidence interval.

Analytical Specificity

Analytical specificity is the evaluation of cross-reactivity by testing a panel of similar, potentially interfering organisms, substances, or analytes to assess constant systematic error. For LDT's, this measurement must be established during the method validation. For other assays, the manufacturer has completed analytical specificity studies.

When determining analytical specificity, the test agents or substances should include as many organisms or analytes as possible that may be found in the relevant test sample or that cause the same symptoms as the target agent. Consider potential sources of variability that could affect the assay (i.e., matrix composition, lot-to-lot variability, temperature, etc.) and include them in the design of the verification or validation. The recommended number of samples to use is between three to five, containing each of the potentially interfering or cross-reactive test organisms, analytes or substances to test. Results should correlate with an expected value $\geq 95\%$. If cross-reactivity is observed, assay conditions may need to be adjusted or reevaluated. In instances where cross-reactivity cannot be eliminated, it must be noted as a limitation of the assay. An inhibition control may need to be included in assay runs where inhibition is prone and needs to be monitored (i.e., molecular assays direct from specimens).

Reportable Range

Reportable range refers to the range of diagnostic results that will be reported. Depending on the assay method used, the reportable range could be a binary result (positive or negative), non-binary result (positive, negative, indeterminate or invalid), the LoD, cutoff value or the 95% confidence interval. The result outcomes are stated in the verification or validation plan and may be adjusted for the final report based on data from the verification or validation. The reportable range should be included in the final SOP.

Reference Range

Reference range is the typical result expected in a healthy population that does not have the condition for which the test is performed. No samples are tested to determine the reference range. Instead, the expected result for a healthy population is stated within the verification or validation plan, report and final SOP. The laboratory may use the manufacturer's reference range provided it is appropriate for the laboratory's patient population. If the manufacturer has not provided reference ranges appropriate for the laboratory's patient population, the laboratory may use published reference range(s).

Additional Qualitative Assay Resources

- [Example of Microbiology MALDI-TOF Validation Supplemental Checklist](#)
- [Example of Microbiology NAAT Checklist](#)
- [NGS Method Validation Plan Template](#)
- [NGS Method Validation Summary Report Template](#)

Quantitative Assays

Quantitative assays are methods that provide a numerical value to the submitter and therefore have additional requirements. Those additional requirements include additional verification or validation data, calibration (may use a standard curve), and the use of quantitative controls.

For quantitative tests, the manufacturer's limits of detection, linearity, reportable range and precision must all be validated or verified by the lab. **Table 4** shows the updated requirements for quantitative tests.

Table 4. Summary of Performance Characteristics Required Depending on Qualitative Test Type²

Test Type	Accuracy	Precision	Analytical Sensitivity	Analytical Specificity	Reportable Range	Reference Range
FDA Approved	Required	Required			Required	Required
FDA Cleared	Required	Required			Required	Required
FDA Modified or LDT	Required	Required	Required	Required	Required	Required
FDA Authorized (EUA)*	Required	Required			Required	Required

Calibration and Statistical Calculations

Calibration is establishing, under specified conditions, the relationship between reagent system or instrument response and the corresponding concentration of an analyte. If the quantitative test is FDA-approved or cleared, calibration must be carried out according to the manufacturer's instructions. If calibrating an FDA-modified or LDT test, ensure that high quality, matrix appropriate materials are used that will provide ideal target values. Materials may need to be purchased from qualified vendors.

Through the verification process, the laboratory defines the frequency for calibration performance as well as the type, number, and concentration of calibration materials used to monitor, detect error, and evaluate method performance. The frequency for calibration performance must not be less than the frequency specified in the manufacturer's instructions.

Calibration requirements vary by assay and depend on the level of FDA approval, but a minimum is typically duplicate calibration runs. FDA approved or cleared tests may allow for a single run. Lower concentrations of calibrators may require more replicates due to the potential for greater variability. Use calibration data to perform regression analysis. Examples are listed below, but the requirements for each method may vary.

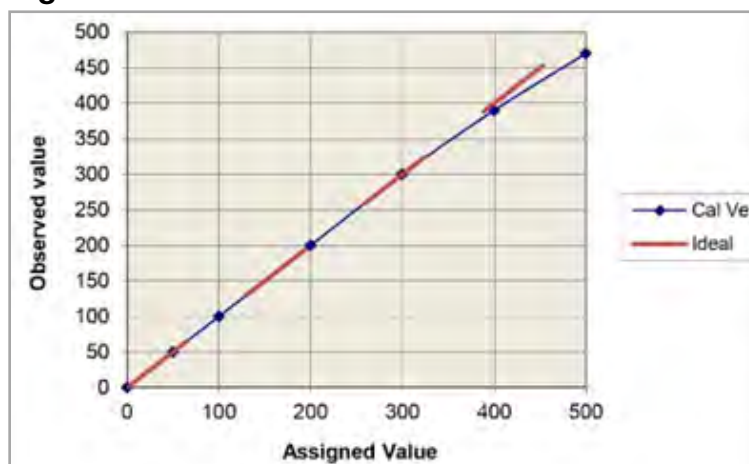
* Requirements may vary depending on EUA. Performance characteristics will be defined by the EUA IFU and the laboratory director. Instructions should include requirements for verification or validation.

Graphical Data Assessment

Westgard QC describes a graphical data assessment, seen in **Figure 4**.¹³ Plot the measurement results on the y-axis vs the assigned values on the x-axis. Draw a 45° line of identity, then draw a “point-to-point” line for the measurement results. Compare the two lines.

- **The coefficient of determination (R^2)** is a measure of how well the regression model fits the observed data. R^2 should be close to 1 (i.e., 0.991 is a very close fit).
- **The coefficient of variation (CV)** can be used to express the precision and repeatability of an assay. The ideal CV is <10-15% and generally should not exceed 20%.
- Calculating the slope is another measure of comparability. The slope should be close to 1.

Figure 4. Calibration Verification



Coefficient of Variation

CV Calculation:¹⁰ The CV is the ratio of the standard deviation to the mean. CV is expressed as a percentage. The ideal CV is <15%, and generally should not exceed 20%. Use the formula in **Figure 3 (page 13)** to calculate the CV.

Method Validation Data Analysis Tool Kit

Westgard QC offers a Method Validation Data Analysis Tool Kit and an online Paired-Data Calculator.¹⁴ This calculator can be used with data from a comparison of methods experiment to calculate linear regression statistics (slope, y-intercept, and standard deviation about the regression line, $s_{y/x}$), and the correlation coefficient (r , Pearson product moment correlation coefficient); t-test statistics (average difference between two methods or $\text{bias}_{\text{meas}}$; SD_{diff} , standard deviation of the differences between the two methods). It can also be used to provide a “comparison plot” that shows the test method results on the y-axis versus the comparative method results on the x-axis, as well as a “difference plot” that displays the difference between the test minus comparative results on the y-axis versus the comparative method result on the x-axis.

Quantitative Controls

Quantitative controls are required for verification or validation and for testing runs following approval of the method. The minimum recommendation is a low positive, high positive, and a negative control. Depending on the assay, the laboratory could customize its QC plan using an IQCP.^{8,9} If controls fail to produce the expected results, the run must be rejected, and the failure should be investigated to identify the cause.

Accuracy

This guidance follows the accuracy requirements for qualitative assays; however, the accuracy study must consider the quantitative results and span the analytical measurement range (AMR). Appropriate samples include commercially available calibrators/calibration or quality control materials with known values, proficiency testing materials that have established values, or previously tested patient specimens with established values. Contrived samples may be used if the appropriate sample type or quantity is not available. Sample mixtures may also be used to achieve the appropriate quantitative value for testing.

Precision

Precision for quantitative assays includes the same structure (inter- and intra- assay measurements) and calculations (%CV) as qualitative tests. Runs should include samples with at least two levels of analyte concentrations. Prior to starting, criteria for identifying and handling outliers should be established to ensure any operational problems do not distort the data. If the test does not routinely include runs, four samples should be run as two sets of pairs at different times on the same day. Those results should be treated as if they were two results resulting from the same run. If the test includes a run that can be completed multiple times in a single day, two runs (separated by at least two hours) with two samples (run in duplicate) should be analyzed with at least one quality control sample and 10 patient samples (if possible) each day. These should be repeated for 20 days with the order of the test materials and quality control samples changed each run or day.¹⁵

Analytical Sensitivity

Analytical sensitivity establishes the limit of detection for a quantitative assay. Dilutions of a known concentration should be tested until the method or instrument no longer detects the presence of the analyte within a matrix. If the laboratory is evaluating a modified test system, it is acceptable to use the manufacturer's stated lower limit if it can be shown that the modification had no effect on the lower limit threshold. Once the LoD for an LDT is established or when verifying a manufacturer's limit of detection claim, it is recommended to test replicates of a number of samples with concentrations both below and above the limit of detection. [CLSI EP17-A2¹¹](#) recommends testing four low positives and four blank samples. The samples should be run in duplicate or triplicate over three days with a recommended minimum of 20 measurements for each sample concentration to verify a manufacturer claim, and 60 measurements to establish the LoD. However, the exact number of samples to use should be determined on a case-by-case basis with input from the QAO, supervisors, or the laboratory director. The LoD should be defined as the concentration in which the assay can distinguish positive from negative samples 95% of the time.

Analytical Specificity

Analytical specificity is the evaluation of cross-reactivity by testing a panel of similar, potentially interfering organisms, substances, or analytes to assess constant systematic error. The test agents or substances should include as many organisms or analytes as possible that may be found in the relevant test sample or that cause the same symptoms as the target agent. Consider potential sources of variability that could affect the assay (i.e., matrix composition, lot-to-lot variability, temperature, etc.) and include them in the design of the verification or validation. The recommended number of samples to use are three to five containing each of the potentially interfering or cross-reactive test organisms, analytes or substances to test. Results should correlate with an expected value $\geq 95\%$. If cross-reactivity is observed, assay conditions may need to be adjusted or reevaluated. In instances where cross-reactivity cannot be eliminated, it must be noted as a limitation of the assay. An inhibition control may need to be included in assay runs where inhibition is prone and needs to be monitored (i.e., molecular assays direct from specimens).

Reportable Range

Reportable range refers to the range of values that can be accurately measured by the test. Calibrators may be used to establish reportable range. For quantitative tests, the manufacturer's reportable range must be verified by the testing laboratory. No quantitative value can be reported that falls outside of the validated range. Other considerations for quantitative assays are to consider how to report numerical results in order to provide the most clinically meaningful information (i.e., log scale vs. integers; 5.30–5 vs 200,000–100,000). Also, some assays may include a cutoff or threshold value that correlates to disease while many do not.

Reference Range

Reference intervals or reference ranges are the span of results expected for a group of healthy individuals who are disease or analyte free. These are results that are considered “normal” for a healthy individual. Reference values for an infectious disease test may not be applicable. However, for values reported in quantitative assays, such as blood chemistry assays, a set of known healthy samples may need to be tested to establish the reference intervals. It is acceptable to use manufacturer or published reference intervals provided they are appropriate for the laboratory’s patient population. If this is done, the verification/validation summary should document the source of the reference intervals and an explanation of their evaluation for appropriateness against the laboratory’s patient population. If, after testing commences, it is determined that the established reference interval is no longer appropriate, it needs to be reevaluated, and documentation of the reevaluation retained with the original verification/validation records.

Ongoing Verification of Quantitative Assays

Calibration Verification

Calibration verification must be performed every six months (or more frequently if specified in the manufacturer’s instructions)¹⁶ and:

- When changes are made to the assay, instrument or overall test system
- When there are major shifts in QC results
- When proficiency testing is inconsistent
- When quality assurance activities indicate discrepant results

There are exceptions to calibration verification requirements:

- Instruments that are factory or manufacturer calibrated do not require calibration verification.
- If the test system’s calibration procedure includes three or more levels of calibration material, and includes a low, mid, and high value, and is performed at least once every six months, then the requirement for calibration verification is also met.

The laboratory defines acceptance or rejection criteria.

Analytical Measurement Range Verification^{7,17,18}

- Analytical measurement range (AMR) verification must be performed every six months and when changes are made to the assay.
- If calibration includes low, midpoint and high values spanning the AMR, additional testing may not be required.
- Control materials may be used.

Additional Quantitative Assay Resources

- [APHL CLIA-compliant Analytical Method Validation Plan and Template for LRN-C Laboratories](#)
- [APHL CRO Breakpoint Implementation Toolkit](#)
- [CLSI 2023 Breakpoint Implementation Toolkit](#)

Related Processes

Sample or Analyte Stability and Integrity

Integrity of the sample analyte should be scrutinized, especially regarding potential causes of instability. The stability of a sample can be assessed at the pre-analytical and analytical phases of testing with the acceptable parameters for sample stability determined during a validation and confirmed for a verification.

Creating an established written policy for sample stability is the end goal when selecting verification or validation samples for pre-analytical and analytical testing.

Pre-analytical sample criteria to consider are:

- Proper shipping and storage
- Appropriate collection devices
- Receipt within the acceptable testing timeframe.

Temperatures of acceptance for shipping, receipt and storage can be assessed during a validation by including samples within a temperature range.

Analytical sample criteria include:

- Assessing the expiration of samples (were the samples received after the acceptable timeframe for testing)
- Minimum volume requirements
- Effects of temperature fluctuations on sample integrity
- Specimen suitability requirements (pH, turbidity, specific gravity, etc.).

For example, molecular extracts or eluates often have limitations on freeze-thaw cycles; therefore, selecting samples that are nearing the allowable freeze-thaw limit is not advisable. Failure to fully assess sample integrity during a verification or validation may result in misguided result analysis and potential diagnostic testing errors.

Bridging Studies and Addendums

Modifications to an approved assay (FDA cleared or approved, EUA authorized, approved LDT) must be validated or verified to demonstrate equivalent performance to the unmodified authorized method. These assay modifications may be referred to as bridging studies or addendums and are meant to demonstrate that the modifications do not negatively affect the result or clinical interpretation of the assay result.

Bridging Studies

Introduced by the FDA in 1998 in the context of drug and biological compound registrations, the concept of bridging studies was originally defined as a study to extrapolate the foreign efficacy data and/or safety data to a new region. The term bridging study has since been adopted in clinical laboratory validation testing to bridge a new component (i.e., reagent, equivalent compound) into a modified assay by establishing equivalent performance between parallel testing of samples with the new and original components. Bridging studies are generally more applicable with EUA assays where it has been deemed acceptable by the regulatory authority (i.e., FDA approval for SARS-CoV-2 molecular assays). The need for a verification or validation is dependent on the specific language in the EUA IFU and your laboratory director. For example, the FDA recommended bridging studies for SARS-CoV-2 molecular assays could consist of testing three-fold serial dilutions of viral material in a pooled respiratory sample matrix in triplicate until a positivity rate of <100% was achieved. If the resulting LoD was the same as the unmodified authorized test LoD, then the two methods are considered to have equivalent performance. Important note: some IFU documents will not accept any modifications; therefore, a bridging study would not be applicable and an LDT validation would be needed.

Addendums

CLSI defines an addendum as an “ancillary report with additional information that expands or clarifies the original final diagnosis but does not change it. Examples include adding information derived from additional diagnostic studies, recuts from specimen blocks, or consultations with experts.” Suppose diagnostic real-time PCR testing demonstrated that the positive PCR control being used resulted in unclear result interpretation due to clinical samples demonstrating low relative fluorescence (RF) for a specific target gene. Confirmatory testing showed that the low RF is due to a specific point mutation where the probe anneals; parallel testing of a new PCR control containing that point mutation clarifies result interpretation. This study would be included as an addendum to the SOP.

Instrumentation Verification or Validation

New Instrument

When bringing online a new instrument or test system, reference the **Qualitative Assays (page 11)** section for performance characteristics that must be considered for verification or validation. This applies even when the new instrument is the same as one already in use. The manufacturer’s recommendations must be reviewed and either accepted or a justification for a deviation must be provided, noting this may lead to the FDA-modified path. Instrument maintenance and troubleshooting must be included in the protocols for establishment of the new method.

Additional Instrument

For a test system or instrument in which there is already an approved method, and the aim is to add an additional test system or instrument of the same make or model, an additional verification or validation must be performed. It is up to the discretion of the laboratory director to determine the amount of testing to be involved in these studies, but the same criteria must be met as if bringing on the assay for the first time. Some criteria may be satisfied by the original verification or validation and a smaller study could be carried out to satisfy the requirement. If the intention is to bring on multiple new instruments, the studies may be divided across instruments as long as each instrument is involved in all parts of the study (accuracy, precision, reportable range) and the serial numbers are identified with the correlating data.

Instrument Move or Repairs

Any movement or repair of an instrument requires additional verification to ensure the alteration did not affect the test performance on that instrument. These studies are typically abbreviated when compared to a new method verification or validation and are up to the discretion of the laboratory director.

Instrument to Instrument

If the laboratory uses more than one instrument or method to test for a given analyte, the instruments or methods must be checked against each other at least twice a year for comparability of results. This is typically aimed to be performed every six months. A subset of samples (QC and/or patient samples) may be compared on all test systems for a given analyte.^{2,16}

Reagent Lot-to-lot Comparison

It is important that each new reagent is verified upon receipt prior to being put into routine use to report patient results.¹⁹ Reagents are validated by the manufacturer and must meet quality standards before release. However, there are several factors that could affect reagent performance once received by the laboratory that include:

- Changes in reagent component materials
- Instability of a component in the reagent
- Reagents compromised in transportation and/or storage
- Incorrect calibration of the new reagent lot.

Differences in performance between current and new or candidate reagent lots hold a potential risk for patient results, which needs to be assessed for all reagent lot changes. Assessment can occur before, or concurrent with, initial use of the candidate reagent lot for patient testing. It is important that verification of candidate reagent lots is not emergent; therefore, ensuring sufficient stock of the previously verified or current reagent lot is essential.

There are no universal acceptance criteria for defining the degree of difference in laboratory results between clinical samples that will influence the clinical decision. The laboratory director must determine the limits of acceptable differences with the primary objective that both the current and candidate reagent lots yield the same clinical interpretation of results.

Note: [CLSI document EP26](#)¹⁹ Appendix A-D provides calculations and statistical considerations to help determine the number of samples needed for reagent lot comparisons and rejection limits.

Defining When to Perform Comparisons

Reagent Comparisons

Good laboratory practice would be to evaluate each new lot and shipment of reagents received. However, the following circumstances **do not** require a full evaluation of a new shipment or candidate reagent lot:

- A new shipment received in the laboratory that has already been evaluated and deemed acceptable for testing with clinical samples. Since the reagent lot was previously determined to not have a significant lot-to-lot difference with clinical samples, any potential shipping damage will be reliably detected using the QC samples. Therefore, each new shipment of the same reagent lot may be verified by testing QC samples only.
- Affiliated laboratories using the same reagent lots and vendor instruments do not need to repeat verification using clinical samples since an initial verification of consistency between current and candidate reagent lots would have been performed. However, all affiliated laboratories need to check each new shipment of reagents for shipping damage using QC samples for verification.
- If there are several of the same instruments used in a single laboratory, verification of the candidate reagent lot only needs to be performed with clinical samples using a single instrument, not all instruments used in the testing.

Methodology Comparisons

Careful attention should be given to laboratories who test the same analyte with different methodologies. In particular, if a laboratory employs more than one methodology to test for an analyte, verification studies should be conducted on both methods. This is true for quantitative and qualitative test procedures.

Examples of this would include the following scenarios.

- A laboratory routinely utilizes a moderate test kit on an analyte that is reported qualitatively. The laboratory decides to employ a back-up test kit due to shortages or other reasons. The laboratory validates and begins testing using the replacement test kit. The laboratory wishes to utilize both test kits when the original test kit is back in stock. In this case, a comparative validation would be warranted.

- The laboratory routinely conducts moderate qualitative kit testing using chromatographic immunoassay as well as moderate instrument testing on the same analyte using automated chemiluminescence. In this case, a comparative validation would be required. If the qualitative result is determined by a numerical value (i.e., concentration of the analyte in the specimen), the method comparison should seek to correlate with these quantitative values. The laboratory should also define the relationship between the two methods during the evaluation process.

Sample Type, Number and Volume for Comparison

The optimal samples for reagent comparison testing are native clinical samples appropriately collected using validated methods, properly stored, and transported under appropriate conditions. Reagent lot-to-lot comparisons should not be performed using QC samples only since observed differences between current and candidate lots could be attributed to a matrix effect (influence of sample property independent of analyte presence) with no change in patient sample measurements. Inversely, it is also possible to have no change in the QC result, but patient results are significantly impacted by reagent lot-to-lot differences. In addition, manufacturer supplied QC material may be optimized to perform correctly with each new reagent lot underscoring any lot-to-lot reagent effect on clinical samples. Therefore, performance of new reagent lots should be evaluated using both patient and QC samples with few exceptions.

Sample Type

- Native clinical samples collected using validated methods and sources.
- Concentrations of the clinical samples should span the measurement interval, or medical decision of interest. If a range cannot be obtained, sample concentrations at two key medical decision points (true positive and true negative) is sufficient. Typically, concentrations near the QC concentrations are clinically relevant.
- Residual clinical samples can be used. It is important to re-measure the clinical samples using both the current and candidate reagent lots, as close in time as possible, to determine any alterations in concentration or activity that are affected by time or storage conditions.
- QC samples are included with clinical samples in each reagent lot verification.

The use of sample material other than native or pooled clinical samples (described in **Sample Volume** below) should only be considered when the aforementioned sample types are unavailable. This could include samples with stability limitations. In these instances, the material used for external evaluation of performance (proficiency testing [PT], external quality assessment, QC material, internal challenge samples) may be considered.

Sample Number

A minimum of three clinical samples is recommended to reduce any potential bias of a single sample. The number of replicate measurements per sample increases the power and reduces error. However, increasing replicate measurements alone, rather than testing more samples, increases the risk that reagent lot or patient sample interactions will not be evaluated.

Sample Volume

- Estimate twice the combined measurement procedure plus dead volume. It is recommended to have sufficient sample volume to allow for repeat testing.
- Samples can be pooled if it is not feasible or practical to obtain sufficient volume per sample. Pools need to be tested using both current and candidate reagent lots, and previous individual sample results cannot be used. A minimum number of unique comparisons should be three (same as individual native samples).

Reagent Comparison Failures

If there are any clinically significant differences in patient results between the current and candidate reagent lots, the candidate reagent lot should not be used until the discrepancy is resolved. Initially, the instrument should be calibrated using controls to minimize any differences that may be attributed to calibration-to-calibration variability. More robust evaluation of the current and candidate reagent lots can be performed using a minimum of 40 clinical samples that span the measurement interval. However, this approach may not be feasible, and the manufacturer should be contacted for follow-up investigation.

If the candidate reagent demonstrates a significant shift in the QC values only, and acceptable performance with the clinical samples, then the QC target values should be updated to reflect the changed performance with the candidate reagent lot. QC material performance may change with different reagent lots, which are typically artifacts of the interaction of the reagent with an altered sample matrix of the QC material. Note: Some local regulatory requirements may restrict any changes to QC target values, which would require a more extensive reagent comparison.

Software Upgrades

Software upgrades for clinical laboratory instruments, or process and reporting management at all levels, is a complex process.²⁰ Any software upgrades that will negatively affect the institution or patient results should be carefully and thoroughly considered before proceeding. In particular, the level and extent of software upgrade testing necessary is largely dependent on the total risk associated with failure: the greater the potential harm associated with failure, the more extensive planning and testing required.

Prior to implementation of software upgrades (whether using new software programs or simply updating to a new version of the current software), the following factors should be considered at minimum:

- Features and functions
- Cost
- Service and upgrade terms and policies
- Hardware requirements
- Performance characteristics (database access and size)
- Compatibility with other institutional systems (i.e., laboratory information system [LIS])
- Data interchange formats supported (e.g., third party applications, Health Level 7 [HL7] standards)
- Data security.

The process of software testing may include boundary testing, input testing and stress testing:

- **Boundary testing** is the use of test cases generated using the extremes of the input value range (i.e., maximum, minimum, within the limits, expected or typical values, and out-of-range [error] values).
- **Input testing** assesses the entry of data input function values including valid, invalid, null and zero, and incorrect data types.
- **Stress testing** is a more general type of focused, thorough testing used to evaluate the stability of a given system, including testing beyond normal operational expectations of a system in order to observe how failure conditions are handled.

Software version upgrades that improve the overall functionality of the system (i.e., enable better visualization, more detailed report summaries, improve the operator experience, etc.) may justify a verification whereas version upgrades that add significant features that could affect the quality or accuracy (e.g., MALDI-TOF database updates including new reference library organisms) necessitate more comprehensive validation of the new system. The software vendor may release technical bulletins or alerts that recommend or require a software version upgrade. Follow instructions from the vendor to correctly upgrade the software package(s). It is important to validate or verify the software upgrade prior

to implementation in the laboratory workflow: ensure the previous software version is available for comparison testing until the software upgrade or new software has been approved.

Verification of software updates uses previous data or known controls to run through the system verifying all the potential result outcomes. The verification should ensure that boundary and input testing are performed to ensure that there is a range of valid input and output values obtained. An example is a qualitative molecular test with Ct value cutoffs of <35 (positive), 36-40 (indeterminate), >40 (negative), and control failure (invalid). The verification should ensure that Ct values within each respective range yield the correct end result, which should match the previous software version. If the end results do not match, contact the software vendor to troubleshoot and/or perform a more extensive validation of the system.

Validation of the software upgrade should follow the key elements described in **Qualitative Assays (page 11)** using the manufacturer or vendor data or documentation as the comparator. A validation will likely encompass an entire workflow focus, which would include both wet-lab and dry-lab testing. For example, a database upgrade for the MALDI-TOF would require culturing standardized microorganisms to spot onto the MALDI target plate followed by system comparison of the microorganism mass spectrum with the database library of reference mass spectra. If the end results do not match the vendor specifications, contact the software vendor to troubleshoot.

The implementation of a new LIS would be another example of a software upgrade requiring a validation. The clinical implication of improper LIS functionality has high risk associated with failure. Therefore, extensive boundary, stress, and input testing methods should be applied to test cases performed in an off-line static testing environment and during parallel testing once the system is 'live.' The length of 'live' parallel testing should be determined based on risk associated with failure and decided on by the QAO and laboratory director.

The system upgrade validations or verifications are deemed acceptable once all relevant measures of performance can be demonstrated. The major features and functions of the system should have been sufficiently challenged and all test cases or samples should have passed. Failed test cases should be repeated and reviewed to determine if the error is obvious and can be rectified. Any failed cases or unexplained conditions need to be analyzed and determined to be noncritical for the system upgrade to pass.

Additional Related Processes Resources

Templates

- [Bridging, Addendum and Extension Studies Template](#)
- [Reagent Comparison QC Studies Template](#)
- [Software Update SOP and Template](#)

Example

- [Analytical Instrument Verification Example](#)

Safety Considerations and Risk Assessments

Prior to working with a new biological or chemical agent or validation of a new test, safety should be considered in order to protect the health of employees.²¹ Safety risk assessments are a systematic procedure for identifying and managing hazards in all stages of the testing process (pre-analytical, analytical, and post-analytical) and should be performed for both biological and chemical and risks. Risk assessments allow the laboratorians to evaluate the work environment, laboratory processes, equipment and available protective measures.

While this step is not required for a verification or validation, inclusion of the laboratory's safety professional during the planning stages of a verification or validation demonstrates commitment to the safety culture of a laboratory. It is the best opportunity to protect employee health in a preventive—rather than reactive—way. Making recommendations in advance allows for effective mitigation procedures and training needs.

Risk Assessment Considerations

Biological Considerations

- Pathogenicity of the agent of interest
- Route of Transmission
- Agent Stability
- Infectious Dose
- Immune status of worker
- Agent Concentration
- Agent Volume
- Splash Potential
- Aerosol Generation
- Percutaneous Hazard
- Biosafety Laboratory Level
- Work Practices
- PPE (Head, Body, Respiratory)
- Viability Study
- Exposure Control Plan
- Vaccination Availability
- Treatment Availability

Chemical Considerations

- Sample Matrix
- Known or Suspected Chemical Hazards (Grade, Concentration, Manufacturer, MSDS)
- Splash Potential
- Health Hazard
- Flammability
- Reactivity
- Oxidizing Potential
- Corrosion Ability
- Environmental Hazards
- Chemical Incompatibilities
- Chemical Storage
- Carcinogen
- Reproductive Toxin
- Toxicity
- Route of Exposure
- Action Levels and Permissible Limits
- Hazardous Operations
- Physical or Equipment Hazards
- Engineering Controls
- Containment Resources
- PPE (Head, Body, Respiratory)
- Waste Management and Disposal

Other Considerations

Some laboratories find it useful to assess the probability and severity of identified risk. Probability refers to the likelihood of an identified risk occurring, while severity refers to the magnitude of the potential consequences if a risk is not appropriately mitigated. The probability and severity assessment helps the laboratory to prioritize the risks and mitigation strategies in a strategic way.

Risk Assessment Tables

Tables 5–7 can be used to assess the risk level associated with each hazard identified in the pre-analytical, analytical and post-analytical stages.

- **Table 5** is used to rate the likelihood of the hazard occurring.
- **Table 6** is used to rate the consequence if the hazard were to occur.
- **Table 7** is used to identify the initial risk level of each hazard based on the likelihood and consequence determined from Tables 5 and 6. For example, if your hazard likelihood is likely and the hazard consequence is moderate then the initial risk level is high per the risk assessment matrix.

A **Residual Risk Rating** is determined after mitigating the initial risk with control and protection procedures. It requires analysis of the initial risk rating and mitigating factors to determine if the residual risk rating is lower, higher or the same.

Table 5: Likelihood of Hazard Occurrence²²

Likelihood	Hazard Likelihood Description
Rare	Will only occur in exceptional circumstances.
Unlikely	Not likely to occur within the foreseeable future.
Possible	May occur within the foreseeable future, sporadic exposure is possible.
Likely	Likely to occur within the foreseeable future, routine exposure is likely.
Highly Likely	Almost certain to occur within the foreseeable future, consistent exposure is highly likely.

Table 6: Consequence of Hazard Occurrence²³

Consequence	Hazard Consequence Description
Insignificant	No treatment required
Minor	Minor injury requiring First Aid treatment (e.g., minor cuts, bruises, bumps)
Moderate	Injury requiring medical treatment or lost time
Major	Serious injury (injuries) requiring specialist medical treatment or hospitalization
Critical	Loss of life, permanent disability or multiple serious injuries

Table 7: Risk Assessment Matrix²⁴

		Hazard Consequence				
		Insignificant	Minor	Moderate	Major	Critical
Hazard Likelihood	Rare	Low	Low	Low	Medium	Medium
	Unlikely	Low	Low	Medium	Medium	High
	Possible	Low	Medium	High	High	High
	Likely	Low	Medium	High	High	Extreme
	Highly Likely	Medium	Medium	High	Extreme	Extreme

Frequency for Reviewing a Risk Assessment

Risk assessments should be reviewed regularly to ensure they are up to date and comply with any changes in regulations or work environment.²³

- The laboratory must review the risk assessment effectiveness at least annually. Risk assessment evaluations may be required more frequently if multiple instances of deviation from established quality thresholds appear in the test system.
- The evaluation must include a review of all components listed in the risk assessment (specimen, testing personnel, environment, reagents, test system) and the QC Plan. It must indicate whether the risk assessment has been effective, and if not, what adjustments are necessary to consistently assure quality.
- Following any quality failures, or when there have been significant changes in any aspect of the original risk assessment, the laboratory must re-evaluate the plan and adjust, if necessary.
- When a quality failure occurs, the laboratory must determine the cause of failure and its impact on patient care, and make any necessary adjustments to the risk assessment.

Hazardous Waste Management

In the performance of testing, hazardous biological, chemical or radiological waste may be generated. The laboratory should have a plan in place which specifies the proper disposal of these wastes. When performing a validation or verification, the laboratory should review the possible generation of hazardous wastes and ensure they are compliant with the laboratory's hazardous waste management plan.

Additional Risk Assessment Resources

[Biological Risk Assessment Template for All New Assays](#)

Cost Analysis and Budget

Testing must occur within an approved budget to ensure fiscal responsibility and sustainability. A cost analysis, or summary of expenses, is a useful tool to determine and compare the cost per test prior to verification or validation of a new or changing method.

There are a number of ways to perform a cost analysis and this section provides only basic cost accounting guidance. Costs can be categorized as direct, indirect, fixed and variable (see **Glossary (page 33)** for definitions). When implementing new testing or modifying current testing, start-up costs must also be considered.

The basic cost per test can be determined using the cost of instrumentation, direct materials and direct labor. For annual budgeting planning, it may be helpful to estimate the yearly spend on a new test system or method and additional information may be needed, such as laboratory information management costs, maintenance costs, site-preparation costs, and depreciation costs (purchased equipment). Within this document, only site preparation costs and yearly maintenance costs will be considered.

Labor Costs

The cost to perform a test per unit of time is called labor cost. The time to perform a test includes all phases of testing—pre-analytical, analytical and post-analytical—and should include all employees involved in the direct production of the actual test result. Other than the hands-on testing time, consider any sample or instrument preparation and maintenance as well as result review and reporting. Estimate labor costs with the equations in **Figure 5**.

Figure 5. Estimation of labor costs

Labor Costs

$$= \frac{\text{Salary with Fringe/Year}}{\text{Number of Tests/Year}}$$

or

$$= \frac{\text{Salary with Fringe}}{1 \text{ Year}} \times \frac{1 \text{ Year}}{2080 \text{ Hours}} \times \frac{1 \text{ Hour}}{60 \text{ Minutes}} \times \frac{\# \text{ Minutes}}{1 \text{ Test}}$$

Materials Costs

Any reagents and consumables including proficiency testing materials used in the performance of the test are considered direct materials. Laboratorians should consider at least the following: price per kit or total expense for all reagents; number of tests per kit or total reagent volume; controls and calibrators per analysis; consumables; and it may be necessary to try and predict retests. See **Figure 6** for an estimation equation.

Figure 6. Estimation of materials costs

Material Costs

$$= \frac{(\text{Cost of Reagents} + \text{Cost of Consumables})}{\text{Number of Tests}}$$

Site Preparation Costs

The purchase of new or updated instrumentation may require site preparation. Expenses related to site preparation include work area renovation or any utilities required for proper operation, including electrical, plumbing, ventilation and air-conditioning. This is often a one-time expense that should be considered part of the start-up costs.

Maintenance Costs

Instrumentation requires routine maintenance to ensure optimal operation for its expected lifetime. This maintenance can be both expected and unexpected. While unexpected maintenance is hard to predict, the user may speak to laboratories using the same equipment or the vendor to create a realistic estimate for unexpected events. Estimated costs related to equipment maintenance can be incorporated into the per test costs or yearly expenses. Consideration for maintenance costs should include not only the instrumentation and materials, but also supporting systems and their associated maintenance costs as well (e.g., LIS service agreements, temperature monitoring systems).

The laboratory can more effectively plan and estimate routine preventive maintenance activities over a period of time and this information should be included in the annual budget. Consider any consumables needed for daily, weekly, monthly, or other routine maintenance. Maintenance contracts may be purchased from the vendor and can also be included as an annual cost.

Reference Documents

Templates

Below are a collection of downloadable/editable resources to help implement a test verification or validation:

- [Biological Risk Assessment for All New Assays Template](#)
- [Breakpoint Implementation Toolkit, CRO 2021](#) (APHL)
- [Breakpoint Implementation Toolkit, MIC 2023](#) (CLSI)
- [Bridging, Addendum, and Extension Studies Template](#)
- [CLIA-Compliant Analytical Method Validation Plan and Template FOR LRN-C Laboratories](#) (APHL)
- [Method Verification Validation Plan Approval Checklist](#)
- [NGS Method Validation Plan Template](#) (NGS Quality Initiative)
- [NGS Method Validation Summary Report Template](#) (NGS Quality Initiative)
- [Printable Checklist for the Verification or Validation Process](#)
- [Reagent Comparison QC Studies Template](#)
- [Software Update SOP and Template](#)
- [Validation Plan Template](#)
- [Validation Summary Report Template](#)
- [Verification Plan Template](#)
- [Verification Summary Report Template](#)

Examples

See examples of these principles in action:

- [Analytical Instrument Verification Example](#)
- [Employee Training Verification Checklist Example](#)
- [Guidelines for Verification and Validation of Laboratory Methods](#) (Minnesota)
- [Method Validation-Verification Summary Report Example](#) (Fairfax County, VA)
- [Method Verification Template Example](#) (Indiana)
- [Microbiology MALDI-TOF Validation Supplemental Checklist](#) (New York CLEP)
- [Microbiology NAAT Checklist](#) (New York CLEP)
- [NGS Method Validation SOP Example](#) (NGS Quality Initiative)
- [Validation and Verification SOP Example](#) (Texas)
- [Validation Plan Example](#) (Washington)
- [Validation Summary Report Example](#) (Washington)
- [Verification Plan Example](#) (Washington)

Additional Resources

The following resources provide additional information for verification or validation of analytes or test systems, some of which are not covered by CLIA, such as food or environmental testing.

- [Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics](#) (B. Magnusson and U. Örnemark (eds.) 2nd ed. 2014)
- [Planning and Reporting Method Validation Studies – Supplement to Eurachem Guide on the Fitness for Purpose of Analytical Methods](#) (V. Barwick (ed.), 2019)
- CAP [All Common Checklist. Items: COM.40250, COM.40350, COM.40475](#) (2020)
- [Validation of Laboratory-Developed Molecular Assays for Infectious Diseases](#) (EM Burd, 2010)
- CLSI [Preliminary Evaluation of Quantitative Clinical Laboratory Measurement Procedures; Approved Guideline - Third Edition](#) (2014)
- CLSI [User Verification of Performance for Precision and Trueness; Approved Guideline – Third Edition](#) (2014)
- FSIS [Guidance for Test Kit Manufacturers, Laboratories: Evaluating the Performance of Pathogen Test Kit Methods](#) (2010)
- FDA [Guidelines for the Validation of Analytical Methods for the Detection of Microbial Pathogens in Foods and Feeds, 3rd Edition](#) (2019)
- FDA [Guidelines for the Validation of Chemical Methods in Food, Feed, Cosmetics, and Veterinary Products, 3rd Edition](#) (2019)
- FDA [Guidelines for the Validation of Microbiological Methods for the FDA Foods Program, 3rd Edition](#) (2019)
- ISO/IEC 17025:2017 [General Requirements for the Competence of Testing and Calibration Laboratories](#) (2017)
- TNI Standard Environmental Laboratory Sector, Volume 1. [Management and Technical Re-quirements for Laboratories Performing Environmental Analysis](#). Contains seven modules:
 1. Proficiency Testing
 2. Quality Systems General Requirements
 3. Quality Systems for Asbestos Testing
 4. Quality Systems for Chemical Testing
 5. Quality Systems for Microbiological Testing
 6. Quality Systems for Radiological Testing
 7. Quality Systems for Toxicity Testing

Glossary & Acronyms

CLSI Harmonized Terminology Database

The Clinical and Laboratory Standards Institute (CLSI) has compiled a [Harmonized Terminology Database*](#) of internationally accepted terminology used in laboratory sciences and related health organizations. This tool is publicly available to encourage broad acceptance and usage of internationally accepted terminology in the laboratory community.

Glossary

- **Accuracy**
An analytical performance measurement that assesses the ability of a method to produce correct results, as compared to a reference standard. Diagnostic sensitivity (known absence for target analyte) and diagnostic specificity (known presence of target analyte) are used.
- **Analytical Measurement Range**
The range of analyte values that a method can directly measure on the specimen without any dilution, concentration, or other pretreatment not part of the usual assay process.
- **Analytical Sensitivity (Limit of Detection)**
The lowest concentration, or amount of an analyte, that can be measured and distinguished from a blank (i.e., minimum detection limits).
- **Analytical Specificity (Interfering Substances)**
The ability of an instrument or test system to measure only the intended organism or substance. Tests the ability to discriminate between the target analyze and other related, but non-target analytes (i.e., cross-reactivity, interfering substances).
- **Coefficient of Variation (CV)**
A measure of relative precision. It is calculated as 100 times the standard deviation, divided by the mean, and expressed as a percentage.
- **Control—High Positive**
A sample, preferably matrix-matched, to evaluate the ability of a laboratory to identify the target of interest near the upper limit of the reportable range, if applicable.
- **Control—Low Positive**
A sample, preferably matrix-matched, to evaluate the ability of a laboratory test to identify the target of interest near the lower limit of the reportable range or near the cutoff.
- **Control—Negative**
A sample, preferably matrix-matched, that lacks the target of interest and used to evaluate the ability of a test not to detect the target when it is not present.
- **Cutoff Value**
In qualitative assays, the cutoff is defined as the threshold above which the result is reported as positive and below which the result is reported as negative.
- **Direct Costs**
Specific costs traceable to the test produced. These can be fixed or variable. Examples include supplies, reagents, consumables, labor, instrument costs, standards and controls.
- **Emergency Use Authorization (EUA)**
An EUA is a mechanism that enables the FDA to facilitate the availability and use of medical countermeasures during declared public health emergencies.
- **False Negative**
A negative result incorrectly ascribed to a positive sample.
- **False Positive**
A positive result incorrectly ascribed to a negative sample.
- **FDA Approved**
The device has been approved through the Premarket Approval process.

* htd.clsi.org/

- **FDA Authorized**
The device has been reviewed by FDA through the EUA mechanism.
- **FDA Cleared**
The device has been cleared as a substantially equivalent device through Section 510(k) of the Food, Drug and Cosmetic Act.
- **FDA Modified**
Any modification to an FDA approved or cleared test. The modification should be handled as a laboratory developed test.
- **Fixed Costs**
A cost that remains constant regardless of workload and within a specific range of activity. Examples include Labor, rent, equipment depreciation, equipment/instrumentation.
- **Gold Standard**
Any standardized clinical assessment, method, procedure, intervention or measurement of known validity and reliability which is generally taken to be the best available, against which new tests or results and protocols are compared.
- **High Complexity**
The most complex testing category assigned to a test by the FDA, based on seven scored criteria. CLIA requirements for laboratories will vary based on the assigned complexity of a test with more stringent requirements for high complexity testing.
- **Indirect Costs**
Any cost that is not assigned to the direct production of a test but contributes to the adequate provision of the work environment. Examples include supervisory salaries, quality assurance, education, travel, administrative costs, building maintenance, security and training.
- **Individualized Quality Control Plan (IQCP)**
The Clinical Laboratory Improvement Amendments (CLIA) Quality Control (QC) procedure for an alternate QC option allowed by 42CFR493.1250. The guidance and concepts for IQCP are a formal representation and compilation of many things laboratories already do to ensure quality test results. IQCP permits the laboratory to customize its QC plan according to test method and use, environment, and personnel competency while providing for equivalent quality testing.
- **Laboratory Developed Test**
Test developed wholly, or in part, by the performing laboratory. This may include analyte specific reagents (ASR) or adoption of another laboratory's LDT or non-cleared or approved test.
- **Limit of Detection (LoD)**
In quantitative and qualitative measurement procedures, the lowest concentration of analyte that can be consistently detected (typically, in $\geq 95\%$ of samples tested under routine medical laboratory conditions and in a defined type of sample).
- **Matrix Effect**
The influence of sample property independent of analyte presence.
- **Negative Predictive Value (NPP)**
The probability that a negative result accurately indicates the absence of the analyte or specific disease.
- **Positive Predictive Value (PPV)**
The probability that a positive result accurately indicates that the analyte or specific disease is present.
- **Precision**
An analytical performance measurement that assesses the closeness of agreement between independent results of measurements obtained under stipulated conditions. Assesses the inherent random error of a test system to determine how close two or more repeated measurements are to each other, regardless of accuracy.
- **Premarket Approval (PMA) Process**
This is the FDA process of scientific and regulatory review to evaluation the safety and effectiveness of Class III Medical Devices.
- **Premarket Notification 510(k)**
Before marketing a device in the US intended for human use that does not require a PMA must submit a 510(k) to FDA (unless otherwise exempt).
- **Reference Range**
The typical result (qualitative) or range of values (quantitative) expected in a non-diseased population that do not have the condition for which the test is performed, including variation due to type of specimen and demographic variables such as age and sex, as applicable.

- **Reportable Range**

The span of test result values over which the laboratory can establish or verify the accuracy of the instrument or test system measurement response. For a qualitative test, the reportable range could be the limit of detection, the cutoff value, or the 95% confidence interval.

- **True Negative**

A negative result that correctly reflects the condition of a sample.

- **True Positive**

A positive result that correctly reflects the condition of a sample.

- **Validation**

The process used to confirm with objective evidence that a laboratory-developed test (LDT) or modified FDA-cleared or approved test method or instrument system delivers reliable results for the intended application.

- **Variable Costs**

A cost that varies with changes in test volume. Examples include reagents and consumables. Labor costs can sometimes be variable when significant increases or decreases in test volume occur, but typically labor will be a fixed cost.

- **Verification**

The one-time process by which a laboratory determines that an unmodified FDA-cleared or approved test performs according to the manufacturer's specifications when used as directed.

Acronyms

AMR: Analytical Measurement Range

ANSI: [American National Standards Institute](#)

APHL: [Association of Public Health Laboratories](#)

ASR: Analyte Specific Reagents

CAP: [College of American Pathologists](#)

CFU: Colony Forming Units

CLIA: Clinical Laboratory Improvement Amendments

CLSI: [Clinical and Laboratory Standards Institute](#)

CoA: CLIA Certificate of Accreditation

CoC: CLIA Certificate of Compliance

CV: Coefficient of Variation

EUA: Emergency Use Authorization

FDA: US Food and Drug Administration

FSIS: [USDA Food Safety and Inspection Service](#)

HL7: Health Level 7 Standards for Electronic Transfer of Data

IFU: Instructions for Use

IQCP: Individualized Quality Control Plan

ISO: [International Organization for Standardization](#)

LIS: Laboratory Information System

LoD: Limit of Detection

MSDS: Material Safety Data Sheet

PPE: Personal Protective Equipment

PPV: Positive Predictive Value

PT: Proficiency Testing

QAO: Quality Assurance Officer

QC: Quality Control

RF: Relative Fluorescence

TCID50: Median Tissue Culture Infectious Dose

TNI: [The NELAC Institute](#)

USDA: [US Department of Agriculture](#)

References

- 1 Code of Federal Regulations. 42 CFR 493 Laboratory Requirements. National Archives and Records Administration. Accessed Oct 1, 2023: www.ecfr.gov/current/title-42/chapter-IV/subchapter-G/part-493
- 2 Code of Federal Regulations. 42 CFR 493.1253 Standard: Establishment and verification of performance specifications. National Archives and Records Administration; 2003. Accessed Oct 1, 2023: www.ecfr.gov/current/title-42/chapter-IV/subchapter-G/part-493/subpart-K/subject-group-ECFRc96dae4380f6ed/section-493.1253
- 3 CMS CLIA. State Operations Manual Appendix C - Survey Procedures and Interpretive Guidelines for Laboratories and Laboratory Services, Rev. 166, 02-03-17. Accessed Jan 15, 2024: www.cms.gov/regulations-and-guidance/guidance/manuals/downloads/som107ap_c_labpdf.pdf
- 4 FDA. Are There “FDA Registered” or “FDA Certified” Medical Devices? How Do I Know What Is FDA Approved? Accessed Oct 8, 2023: www.accessdata.fda.gov/scripts/cdrh/devicesatfda/index.cfm
- 5 Bankowski MJ, Cankovic M, Dunlap J, Furtado LV, Gong J, Huard T, et al. Molecular Diagnostic Assay Validation. Update to the 2009 AMP Molecular Diagnostic Assay Validation White Paper. 2014. Accessed Oct 1, 2023: www.amp.org/AMP/assets/File/resources/201503032014AssayValidationWhitePaper.pdf?pass=29
- 6 CLSI. Measurement Procedure Comparison and Bias Estimation Using Patient Samples, 3rd ed., CLSI guideline EP09c, Wayne, PA: Clinical and Laboratory Standards Institute; 2018. Available from: <https://clsi.org/standards/products/method-evaluation/documents/ep09/>
- 7 CLSI. Evaluation of Qualitative, Binary Output Examination Performance, 3rd ed. CLSI guideline EP12. Clinical and Laboratory Standards Institute, 2023. Available from: clsi.org/standards/products/method-evaluation/documents/ep12/
- 8 CMS CLIA. State Operations Manual Appendix C - Survey Procedures and Interpretive Guidelines for Laboratories and Laboratory Services, Rev. 166, 02-03-17. IQCP is on page 198. Accessed Jan 15, 2024 from: www.cms.gov/regulations-and-guidance/guidance/manuals/downloads/som107ap_c_labpdf.pdf
- 9 CDC Laboratory Quality Tools and Resources. Individualized Quality Control Plan (IQCP). Accessed Oct 15, 2023 from: www.cdc.gov/labquality/iqcp.html
- 10 Bio-Rad UnityWeb. Coefficient of Variation. Accessed October 15, 2023 from: unityweb.qcnet.com/Documentation/Help/UnityWeb/402.htm
- 11 CLSI. Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline—Second Edition. CLSI document EP17-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2012. Available from: clsi.org/standards/products/method-evaluation/documents/ep17/
- 12 Westgard QC. James O. Westgard, Sten A. Westgard. Probit Analysis 1: Practical Application to Determine Limit of Detection. August 2020. Accessed October 1, 2023 from: www.westgard.com/probit-part-one.htm
- 13 Westgard QC. James O. Westgard. Calibration Verification: Defining Criteria for Acceptable Performance. May 2016. Accessed October 1, 2023 from: www.westgard.com/cal-verification-criteria.htm
- 14 Westgard QC. Method Validation Data Analysis Tool Kit. Accessed October 15, 2023 from: www.westgard.com/mvtools.htm
- 15 CLSI. Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline—Third Edition. CLSI document EP05-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2014. Available from: clsi.org/standards/products/method-evaluation/documents/ep05/
- 16 CMS CLIA. Calibration and Calibration Verification Brochure. Accessed Jan 14, 2024 from: www.cms.gov/files/document/clia-brochure-calibration-and-calibration-verification-april-2006.pdf

- 17 Validation and Implementation of Quantitative Molecular Assays, Morgan A. Pence April 10, 2019, Molecular Diagnostics 2019, Sponsored by DiaSorin Molecular LLC. Accessed October 1, 2023 from: www.youtube.com/watch?v=kON_toug60Q
- 18 CMS CLIA. Calibration and Calibration Verification Brochure #3. Accessed October 1, 2023 from: www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/downloads/6065bk.pdf
- 19 CLSI. User Evaluation of Acceptability of a Reagent Lot Change. 2nd ed. CLSI Guideline EP26. Clinical & Laboratory Standards Institute. 2022. Available from: clsi.org/standards/products/method-evaluation/documents/ep26/
- 20 CLSI. Laboratory Instruments and Data Management Systems: Design of Software User Interfaces and End-User Software Systems Validation, Operation, and Monitoring, 2nd Edition. CLSI document AUTO13-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2003. Available from: clsi.org/standards/products/automation-and-informatics/documents/auto13/
- 21 Safeopedia. Safety Risk Assessment. Accessed October 15, 2023 from: www.safeopedia.com/definition/731/safety-risk-assessment
- 22 CDC Center for Surveillance, Epidemiology and Laboratory Services. Reynolds L Salerno. Risk Assessment: The Foundation of Every Good Biorisk Management System. March 2, 2020. Accessed January 15, 2024 from: www.cdc.gov/safelabs/docs/CLEARED-Salerno-Biosafety-Symposium-Risk-Assessment-FINAL.pdf
- 23 World Health Organization. Laboratory Biosafety Manual (pages 16 and 25). 3rd Edition. Geneva. 2004. Accessed January 14, 2024 from: www.who.int/publications/i/item/9789240011311
- 24 Salerno, R, and Gaudioso, J. (2015) Laboratory Biorisk Management: Biosafety and Biosecurity, CRC Press.



Association of Public Health Laboratories

The Association of Public Health Laboratories (APHL) works to strengthen laboratory systems serving the public's health in the US and globally. APHL's member laboratories protect the public's health by monitoring and detecting infectious and foodborne diseases, environmental contaminants, terrorist agents, genetic disorders in newborns and other diverse health threats.

7700 Wisconsin Avenue, Suite 1000 Bethesda, MD 20814 | 240.485.2745 | www.aphl.org

A special acknowledgment to the following individuals who contributed to the design and development of the toolkit: Kathy Ross, Deb Severson, Laurie Gregg, Bonita Bryant, Leann Covington, Rajesh Rarmar, Kim Smith, Lydia Mikhail, Mary Bonifas, Ana Gross, Jennifer Dale, Colleen Courtney, Kara Mitchell, Marilyn Bibbs-Freeman, David Silva, Latricia Lewis, Berihun Taye, Jackie (John) Lyle, Cindy Wong, Tabatha East, Marie Earley (CDC), Susie Zanto (Laboratory Solutionz), Adom Yusuf (APHL), Andrea Wright (APHL), Anne Gaynor (APHL), Ashley Gibbs (APHL), Afifa Alkozai (APHL), Tina Su (APHL), Sarah Buss (APHL), Abigail Raymer (APHL), Kelly Wroblewski (APHL), Madeline Rooney (APHL) and Lorelei Kurimski (APHL).

© Copyright 2024, Association of Public Health Laboratories. All Rights Reserved.

This publication was supported by Cooperative Agreements #N600E000104, 100% funded by the US Centers for Disease Control and Prevention (CDC). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of CDC or the US Department of Health and Human Services.