Method Validation and Verification

Lauren N. Pearson, DO, MPH
Laboratory Director, University of Utah Health Sciences Center Clinical Laboratories

AUGUST 2020
## Objectives

<table>
<thead>
<tr>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify the difference between method validation and method verification</td>
</tr>
<tr>
<td>Describe the studies required to document method performance</td>
</tr>
<tr>
<td>Interpret method performance data and statistical data outcomes</td>
</tr>
</tbody>
</table>
Outline

Context and definitions
Regulatory requirements
Studies required for analytical verification and analytical validation
Context and Definitions
Validation versus Verification

Validation
» Establishing the performance specifications of a new diagnostic tool such as a new test, laboratory developed test or modified method

Verification
» A one-time process to determine performance characteristics of a test before use in patient testing
Why Evaluate a Method?

• Document initial performance:

  » Reference when troubleshooting problems

  » Quality assurance - to ensure results

  » Helpful for clinical consultations

  » Meet regulatory requirements
Laboratory Regulations

• General and open to some interpretation

• Direct what must be done, not “how” it is accomplished
U.S. Test Categorization

• Determined during FDA pre-market approval

• Waived testing
  » Approved for home and point-of-care use
  » “Low risk of patient mismanagement if performed incorrectly”

• Non-waived testing
  » Moderate Complexity
  » High Complexity e.g. LDTs
  » Modified Tests

Regulatory Requirements
Waived Tests

Labs have only 3 requirements!

• Pay biennial fee (every 2 years) for CLIA certificate renewal
• Follow manufacturers instructions for use
• Allow the laboratory to be inspected
  – Generally, for cause (patient complaint)
  – Random state survey
  – Periodic inspections not required!

Note: No method evaluation required
Nonwaived Tests
Method “Validation” to CLIA

Moderate Complexity
» Precision
» Accuracy
» Reportable Range
» Reference Range

High Complexity
» Precision
» Accuracy
» Reportable Range
» Reference Range(s)
» Analytical Sensitivity (LOD)
» Analytical Specificity

• Mnemonic: PARR

Establish calibration and control procedures
Other performance criteria

Test Modifications

Any change in the intended use or change to an assay that could affect performance:

- Different sample matrix (urine in a serum assay)
- Promoting different use (screen vs diagnosis)
- Type of analysis (qualitative vs quantitative)
- Incubation times and temperatures
- Sample or reagent dilution
- Using different calibration material or set-point
- Change or eliminating a procedural step
Analytical Verification/Validation

• Laboratories are required to perform analytical verification or validation of each nonwaived test, method, or instrument system before use in patient testing
  – Regardless of when it was first introduced by the laboratory
  – Includes instruments of the same make and model and temporary replacement (loaner) instruments

• There is no exception for analytical verification or validation of tests introduced prior to a specific date

• The laboratory must retain records as long as the method is in use and for at least two years after discontinuation
How to Meet the Regulations

• There is no one right way
• Consensus - CLSI protocols
• Literature - do what others have done
• Manufacturer’s recommendations
• Balance between cost and what is reasonable
Validation Studies
Accuracy

Bias to a “reference” method
- Absolute
- Relative
Accuracy Studies

Method comparison
- Carefully select “reference” method
- Curate high quality samples with a range of analyte concentration
- Analyze >40 specimens by both test and reference method
  - Best to analyze in duplicate over a period of many days

Data analysis
- Scatter plot of data
- Calculate regression statistics and estimate bias
- Compare results with claims or internal criteria to judge acceptability

CLSI EP-09
Fig. 1. Result comparison plot for combined Architect PCT results.
Result comparison plot for combined Architect PCT results 0 to 130.0 ng/mL (n = 192). Pas Bab, Passing-Bablok regression. Slope = 0.834 (0.808–0.857); intercept = −0.013 (−0.027 to −0.010).

192 specimens
2 lots of reagent and calibrator
Special Considerations

- Medical decision points
- Clinically relevant cutoffs
Fig. 2. Result comparison, bias, and percent bias plots for Architect vs Kryptor PCT assays.

Result comparison (A), bias (B), and % bias (C) plots for Architect vs Kryptor PCT assays limited to Architect values of 0 to 3.0 ng/mL (n = 149). Pas Bab, Passing–Bablok regression. Slope = 0.792 (0.769–0.821); intercept = −0.008 (−0.019 to 0.001).
Table 2. Agreement between Kryptor PCT and Architect PCT assays at 0.10, 0.25, 0.50, and 2.00 ng/mL.\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th>Positive agreement, %</th>
<th>Negative agreement, %</th>
<th>Total agreement, % (95% CI)</th>
<th>κ, % (95% CI)</th>
<th>Predicted MDC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10 ng/mL</td>
<td>92.4</td>
<td>96.8</td>
<td>93.3 (88.1–96.3)</td>
<td>81.4 (70.3–95.5)</td>
<td>0.07 (0.06–0.08)</td>
</tr>
<tr>
<td>0.25 ng/mL</td>
<td>90.3</td>
<td>100.0</td>
<td>93.3 (88.1–96.3)</td>
<td>85.2 (76.3–94.0)</td>
<td>0.19 (0.18–0.20)</td>
</tr>
<tr>
<td>0.50 ng/mL</td>
<td>78.9</td>
<td>97.4</td>
<td>88.6 (88.5–92.8)</td>
<td>76.9 (66.6–87.3)</td>
<td>0.39 (0.38–0.40)</td>
</tr>
<tr>
<td>2.00 ng/mL</td>
<td>42.9</td>
<td>100.0</td>
<td>94.6 (89.8–97.3)</td>
<td>57.6 (29.0–86.2)</td>
<td>1.57 (1.55–1.59)</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Agreement between Kryptor PCT and Architect PCT assays limited to PCT values of 0 to 3.0 ng/mL (n = 149, n = 78 with PCT ≤0.5 ng/mL on Kryptor, n = 71 with PCT >0.5 ng/mL on Kryptor).
Precision

- Within-run (Intra-assay)
- Between-run
- Day-to-day (Inter-assay or total)
Precision Studies

• Selection of appropriate material

• Verification study
  – 5 x 5 study design

• Full precision study
  ▪ Within run
    – 20 consecutive replicates/single run
  ▪ Total
    – 2 replicates/concentration level/run
    – 2 runs/day x 20 days

• Data analysis
  – Calculate mean, SD, and CV
  – Compare results with claims or internal criteria to judge acceptability
3 controls run twice per day in duplicate x for 20 days

### Table 1. Assay precision for Architect A and Architect B (n = 80).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Architect A</th>
<th></th>
<th>Architect B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low QC</td>
<td>Medium QC</td>
<td>High QC</td>
<td>Low QC</td>
</tr>
<tr>
<td>Mean, ng/mL</td>
<td>0.206</td>
<td>1.989</td>
<td>70.468</td>
<td>0.209</td>
</tr>
<tr>
<td>SD within run, ng/mL</td>
<td>0.006</td>
<td>0.042</td>
<td>1.593</td>
<td>0.005</td>
</tr>
<tr>
<td>CV within run, %</td>
<td>2.9</td>
<td>2.1</td>
<td>2.3</td>
<td>2.4</td>
</tr>
<tr>
<td>SD total, ng/mL</td>
<td>0.007</td>
<td>0.055</td>
<td>1.962</td>
<td>0.007</td>
</tr>
<tr>
<td>CV total precision, %</td>
<td>3.4</td>
<td>2.8</td>
<td>2.8</td>
<td>3.3</td>
</tr>
</tbody>
</table>

* QC, quality control.
Utility of Precision Data

• Future troubleshooting

• Clinical queries about significant change

• Setting QC ranges
Reportable Range

Includes:

- Analytical measurement range (AMR)
  - Range of values that an instrument can report directly without alteration or pretreatment of the sample

- Clinically reportable range (CRR)
  - Range of values that can be reported with alteration of the sample
    - Medical director discretion
Range for Reporting Patient Samples

• May use the AMR
• May modify AMR to create expanded range (CRR)
  » Must document that modifications to the sample and method produce reliable results
    ▪ Verify every 6 months
  » Must be verified or established before patient testing begins
  » Must establish reportable limits (undiluted) and maximum dilution
Reportable Range Studies

• Linearity Study
• 5+ concentrations of analyte throughout range
  • Spike low sample with known amount of analyte
  • Dilute high sample with a blank
  • Mix high and low sample to create a curve
• Standard reference materials
• Commercial linearity products
• Two replicates at each level
• Data analysis
  o Evaluate linear fit with XY plot
  o Calculate slope and intercept

CLSI EP-6
Pooled patient serum sample and calibrator A mixed to get 6 sample concentrations.

Run in triplicate on each Architect instrument.

Both instruments demonstrated linearity of the assay consistent with manufacturer’s claims.
Reference Intervals (RI)

• Labs are not required to establish their own
  » Good practice is to verify that RI is appropriate for patient population

• Can use previously established RI or create a new one
  » Discretion of the medical director

• Transfer of a RI is acceptable if test subject population and methodology are the same or comparable
  » Verified by testing N ≥ 20 samples
  » If ≤ 2 outside limits, then accept
Establishing a RI

• Typically the central 95% of the values for the study population

• Considerations
  • Exclusion criteria
  • Partitioning
  • Pre-analytical considerations
  • Specimen handling and storage
  • Special or unique patient populations
Protocol for Full RI Study

- Establish selection criteria for individuals
- Establish a list of interferences or sources of biological variability
- Decide on appropriate number of individuals based on desired confidence limits (e.g. n=120)
- Collect and analyze specimens
- Evaluate data using histogram to evaluate distribution

CLSI EP28A3C
Samples from 20 apparently healthy donors into PST and SST tubes

Donor exclusion criteria

Samples had PCT concentrations of 0.01 to 0.03 ng/mL confirming the manufacturer’s claims
Analytical Sensitivity

• Establishes the analytical sensitivity (lower detection limit) of the assay

• For modified FDA-cleared/approved tests or laboratory-developed tests (LDTs)
Analytical Sensitivity Studies

• Acquire measurements from multiple, independent blank and low-level samples or pools of samples
  » At least four samples of each type
  » Can dilute or spike samples to provide low level samples at desired analyte levels
  » Low level sample around assumed LoD
  » Obtain a series of replicate results

• Data analysis
  » Parametric or nonparametric statistical methods
Limit of Blank determination:
Calibrator A (concentration 0 ng/mL) analyzed 10 times on each instrument
Calibrator B (concentration 0.1 ng/mL) analyzed 3 times on each instrument

Limit of Quantitation determination:
8 calibrator samples including 4 low level concentrations analyzed over 10 days

Results were equivalent to the manufacturer’s claims
Analytical Specificity (Interferences)

• Refers to the ability of a test or procedure to correctly identify or quantify an entity in the presence of interfering or cross-reactive substances.

• For modified FDA-cleared/approved tests or laboratory-developed tests (LDTs)
Interfering Substances (IFS)

• Interference- a significant difference in test result because of another component of the sample

• Interfering substance- a substance that causes the measurement to be inaccurate

• Can cause a concentration dependent difference in the test
• Manufacturers screen for IFS during method development
Identifying Error from IFS

Quantify effects by performing paired difference study:
- Pairs of test samples
- One with potential IFS, the other without
- Measure analyte of interest
- Calculate differences
- May be performed with patient samples

CLSI EP07
Summary

• Regulations require performance verification prior to patient testing
• Precision, accuracy, reportable range and reference interval must be evaluated, at a minimum, for all nonwaived tests before patient use
• No “one size fits all” approach to validation/verification
References


• HCFA 42 CFR; Final Rule; Medicare, Medicaid and CLIA Programs; Regulations implementing Clinical Laboratories Improvement Amendments of 1988 (CLIA. Fed. Regist. 57;7001-288 (28 February 1992)

• College of American Pathologists Commission on Laboratory Accreditation. Inspection Checklist. 2016 CAP, Northfield, IL.
