# How to Validate Quantitative Laboratory Assays

Anthony A. Killeen, MD, PhD
Professor and Vice-Chair
Dept. of Laboratory Medicine & Pathology
University of Minnesota



## Disclosures

- None



## Most clinical lab assays are quantitative

- Lab personnel should be familiar with the basic requirements of how to validate assays to comply with CLIA and good lab practice



## What is required for assay validation for FDA-Cleared Assays?

- **§ 493.1253 Standard: Establishment and verification of performance specifications.**
- (a) *Applicability.* Laboratories are not required to verify or establish performance specifications for any test system used by the laboratory before April 24, 2003.
- (b)(1) *Verification of performance specifications.* Each laboratory that introduces an unmodified, FDA-cleared or approved test system must do the following before reporting patient test results:
- (i) Demonstrate that it can obtain performance specifications comparable to those established by the manufacturer for the following performance characteristics:
- (A) Accuracy.
- (B) Precision.
- (C) Reportable range of test results for the test system.
- (ii) Verify that the manufacturer's reference intervals (normal values) are appropriate for the laboratory's patient population.



## Non-FDA-Cleared Assays

- (2) *Establishment of performance specifications.* Each laboratory that modifies an FDA-cleared or approved test system, or introduces a test system not subject to FDA clearance or approval (including methods developed in-house and standardized methods such as text book procedures), or uses a test system in which performance specifications are not provided by the manufacturer must, before reporting patient test results, establish for each test system the performance specifications for the following performance characteristics, as applicable:
- (i) Accuracy.
- (ii) Precision.
- (iii) Analytical sensitivity.
- (iv) Analytical specificity to include interfering substances.
- (v) Reportable range of test results for the test system.
- (vi) Reference intervals (normal values).
- (vii) Any other performance characteristic required for test performance.



## For all assays, we are also required to comply with

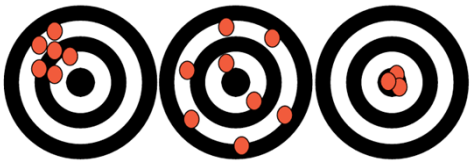
- **§ 493.1255 Standard: Calibration and calibration verification procedures.**
- **§ 493.1256 Standard: Control procedures.**

## What are Precision and Trueness?

- Precision:
  - Closeness of agreement of measured quantity values between replicate measurements
  - Usually expressed as the imprecision
    - Most commonly the coefficient of variation (SD/mean)
- Trueness:
  - The closeness of agreement between the *mean quantity value of a large (infinite) number of replicates* and the reference quantity value



## The Bull's-Eye Analogy Does This Make Sense?



Precise, but not accurate | Accurate, but not precise | Precise and accurate



## Accuracy vs. Trueness

- Accuracy refers to the difference between a single measurement and the true value
- Trueness refers to the mean difference between many replicated measurements and the true value



## For FDA-cleared methods, precision is *verified*, bias (inaccuracy) is *established*

- It would be very unusual for a manufacturer to offer a method with a bias that a user would need to verify!
- Claims of precision need to be *verified*



## Precision over what time period?

- Time, calibration, operator, and equipment can change
- Short-term: "repeatability" (all variables the same)
  - Formerly "within-run imprecision"
- Long-term: "within-lab imprecision": defined time and operators, same facility and equipment, but reagent lots/calibrators may change
  - Formerly "total imprecision"
- Within-lab imprecision cannot be less than the repeatability



## Example of Manufacturer Precision Claim (Siemens Cholesterol IFU) Advia 1200

***Precision*** [7]

Each sample was assayed 2 times per run, 2 runs per day, for at least 20 days. Precision estimates were computed according to CLSI document EP05-A2, *Evaluation of Precision Performance of Quantitative Measurement Methods*; Approved Guideline.[7]

Data contained in this section represents typical performance for ADVIA Chemistry systems. Your laboratory data may differ from these values.

**Conversion factor: mg/dL x 0.0259 = mmol/L**

## Manufacturer Precision Claim Example (Siemens, cholesterol)

***ADVIA 1200***

| | | Within-Run | | Total | |
|---|---|---|---|---|---|
| **Specimen Type** | **Level** | **SD** | **CV (%)** | **SD** | **CV (%)** |
| | | **Common Units (mg/dL)** | | | |
| Serum | 146 | 2.4 | 1.7 | 4.6 | 3.1 |
| Serum | 259 | 3.1 | 1.2 | 5.5 | 2.1 |
| Serum | 350 | 4.2 | 1.2 | 7.1 | 2.0 |



## Manufacturer AMR claim examples (Siemens, cholesterol)

***Analytical Range***

This method is linear from 0 – 675 mg/dL (0 – 17.48 mmol/L) for serum and plasma.

Siemens has validated an automatic rerun condition for this method that extends the reportable range up to 3375 (87.41 mg/dL) on the ADVIA 1200, and up to 1350 mg/dL (34.97 mmol/L) on the ADVIA 1650/1800/2400 systems for serum and plasma.



## Manufacturer's calibration statement (Siemens, cholesterol) from IFU

The ADVIA CHOL method is traceable to the CDC reference method, which uses reference materials from the National Institute of Standards and Technology (NIST), via patient sample correlation. See the correlation data in System Correlation for the relationship. Assigned values of Siemens Chemistry Calibrator are traceable to this standardization.



## Procedure to follow to verify precision claims (CLSI EP15-A3)

- Use **at least** 2 samples at different concentrations, chosen carefully
  - Patient samples (pooling is OK)
  - Clinical decision points if possible, or normal/abnormal, or
  - Concentrations near the manufacturer's package insert
- 5 replicates for each sample per run on each of 5 days
- Total of 25 data points per sample
- Schedule testing so different operators use the system
- Review data in real time
- Exclude outliers where there is justification to do so



## Procedure points from CLSI EP15-A

- If QC fails, reject the run
- Doing a few more than 5 runs is OK
- Runs do not have to be on sequential days



## Data analysis: Outliers

- Testing for possible outliers is performed by Grubb's test (recommended by CLSI EP15-A)
- With a 5x5 experiment, Grubb's test identifies samples that are 3.1 SD beyond the mean (see CLSI EP15-A for more detail)

## Data Analysis: Calculating Imprecision

- Recommended approach is to use Analysis of Variance (ANOVA)
- ANOVA partitions variation into component sources
  - Here, we're interested in within-run (replication) and between-run (within-lab)



## Some Data for Sodium (mmol/L) (for example only)

| | Run 1 | Run 2 | Run 3 | Run 4 | Run 5 |
|---|---|---|---|---|---|
| Replicate 1 | 140 | 140 | 141 | 139 | 141 |
| Replicate 2 | 141 | 139 | 139 | 139 | 141 |
| Replicate 3 | 142 | 139 | 140 | 140 | 140 |
| Replicate 4 | 141 | 138 | 141 | 139 | 141 |
| Replicate 5 | 140 | 140 | 139 | 138 | 140 |



## ANOVA Output (Stata)

```
                    Analysis of Variance
    Source      SS          df        MS          F     Prob > F
------------------------------------------------------------------------
Between groups 13.04         4         3.26       5.09     0.0054
Within groups  12.8          20         .64
------------------------------------------------------------------------
    Total      25.84         24        1.077
```

- $Variance_{within}$ = 0.64
- $Variance_{between}$ = (3.26 – 0.64)/5 = 0.524



## Expressing the components of variation

- Calculate the standard deviations (*S*)
  - $S_R = \sqrt{(VW)}$ = 0.80 mmol/L (within-run or repeatability)
  - $S_B = \sqrt{(VB)}$ = 0.72 mmol/L (between-run)
  - $S_{WL} = \sqrt{(VW + VB)}$ = 1.08 mmol/L (within-lab or total)
- Convert to CV% by dividing by the grand mean (139.9 mmol/L)
  - $CV_R$ = 0.57%
  - $CV_B$ = 0.51%
  - $CV_{WL}$ = 0.78%
- Compare these to the manufacturer claims of within run and total precision



## Manufacturer Claim for Accuracy

***ADVIA 1200***

| Specimen Type | Comparison System (x) | N | Regression Equation | Sy.x | r | Sample Range |
|---|---|---|---|---|---|---|
| Serum | ADVIA 1650 | 300 | y = 1.05x + 1.1 | 4.2 | 0.998 | 33 – 672 mg/dL |
| | | | y = 1.05x + 0.03 | 0.11 | 0.998 | 0.85 – 17.48 mmol/L |
| Plasma* | ADVIA 1200 (serum) | 30 | y = 0.99x - 2.1 | 2.1 | 0.999 | 130 – 275 mg/dL |
| | | | y = 0.99x - 0.06 | 0.05 | 0.999 | 3.36 – 7.13 mmol/L |
| Serum | Reference Method | 46 | y = 1.01x + 0.0 | 11.1 | 0.997 | 127 – 669 mg/dL |
| | | | y = 1.01x + 0.001 | 0.29 | 0.997 | 3.29 – 17.32 mmol/L |

*lithium heparin



## How to Assess Trueness

- Run samples of known concentration with the test method and calculate the bias
- Remember, trueness is the difference between the mean of many (theoretically an infinite number) of replicates of the sample and the true value
- Samples that can be used include
  - Reference materials with values assigned by reference method
  - PT/EQA materials that have been tested by comparable methods
  - Spiked samples (e.g., drugs in serum/urine)
  - Patient samples if the true value is known (e.g., reference value assisgnment)
  - Composition of material is important (matrix effect)
- Calculations: see EP15-A3 for details
  - Statistical consideration (do 95% of measurements include the true value?)
  - Is any bias clinically significant?

## What is required for assay validation for FDA-Cleared Assays?

- **§ 493.1253 Standard: Establishment and verification of performance specifications.**
- (a) *Applicability.* Laboratories are not required to verify or establish performance specifications for any test system used by the laboratory before April 24, 2003.
- (b)(1) *Verification of performance specifications.* Each laboratory that introduces an unmodified, FDA-cleared or approved test system must do the following before reporting patient test results:
- (i) Demonstrate that it can obtain performance specifications comparable to those established by the manufacturer for the following performance characteristics:
- (A) Accuracy.
- (B) Precision.
- (C) Reportable range of test results for the test system.
- (ii) Verify that the manufacturer's reference intervals (normal values) are appropriate for the laboratory's patient population.



## What is the Reportable Range?

- Two concepts
  1. The values that a method can report without any special treatment of a sample (e.g., dilution or concentration)
     - Defined by the analytical measurement range (AMR)
  2. The range of values that a lab will report an actual value (as opposed to "less than" or "greater than")
     - Defined by the lab itself (e.g., any hCG more than 100,000 U/ml will be reported as "> 100,000 U/mL")



## Reference Interval Verification

- For FDA-cleared methods, the lab must verify that the manufacturer's suggested reference interval is appropriate for the patient population
  - Often a weak link in the practice of Lab Medicine
- At least 20 healthy patient samples should be used and no more than 2 should fall outside the proposed interval
- Literature references can be used in some cases e.g., national cholesterol or glucose guidelines, or CSF protein (impractical to perform a formal reference study). Must document!



## Ongoing Plan to Monitor a New Assay

- Calibration
- Calibration Verification
- Analytic Measurement Range (AMR)
- Enroll in PT/EQA, or have a plan for validation of results at least 6 monthly



## Calibration

- Process that establishes the relationship between reagent system/instrument response and the corresponding concentration/activity values of an analyte
- If the calibration changes, *patient results will also change*
- Calibrators ideally should be traceable to a reference method to ensure accuracy and be matrix-appropriate



## Calibration Verification

- **Confirms current calibration settings remain valid for test system**
- **"*Trueness*" assumes there is a value that the instrument should report for a specific sample**
  - **Calibration establishes this assignment; calibration verification shows that this is still true**
  - **Controls do not usually come with assigned values valid for the instrument *unless the manufacturer proves these values***

## Method/matrix-appropriate materials and target values suitable for Cal Ver

- Calibrators used to calibrate analytical measurement system (different lot)
- Materials provided by instrument/method vendor for calibration verification
- Previously tested unaltered patient/client specimens
- Primary or secondary standards or reference materials with method-appropriate matrix characteristics and target values
- Third party general purpose reference materials if commutable
- PT material with method-appropriate matrix characteristics and target values



## Analytic Measurement Range (AMR) and AMR verification

| AMR | AMR Verification |
|---|---|
| **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** | **Process of confirming that the assay system will correctly recover the concentration or activity of the analyte over the AMR** |

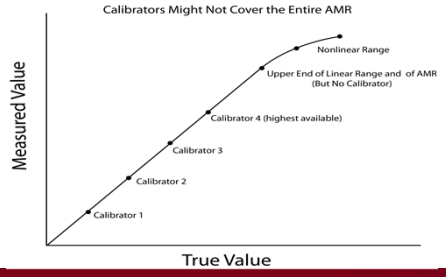


## "Linearity" and the AMR

- AMR verification is accomplished by demonstrating a linear relationship for an appropriate set of samples that cover the AMR
- A plot of measured results for an analyte obtained across the AMR vs. expected concentrations or concentration relationships (or expected activity or activity relationships) in a set of samples should show a linear relationship



## Linearity and the AMR






## Frequency of AMR verification and Calibration Verification

- **When a method is implemented**
- **At least every 6 months; as necessary**
  - **A change of reagent lots for chemically or physically active or critical components**
  - **If QC fails to meet established criteria**
  - **After major maintenance or service**
  - **When recommended by the manufacturer**



## Matrix and target value appropriate materials suitable for AMR verification

- **Linearity material of appropriate matrix (e.g., CAP CVL Survey-based material)**
- **Previously tested patient/client specimens, altered by admixture with other specimens, dilution, spiking in known amounts of an analyte, or other technique**
- **Previously tested patient/client specimens, unaltered**
- **Primary or secondary standards or reference materials**
- **Calibrators used to calibrate the analytic measurement system from a different lot than used for calibration**
- **Control materials with method specific target values**

## How many samples are required for AMR verification?

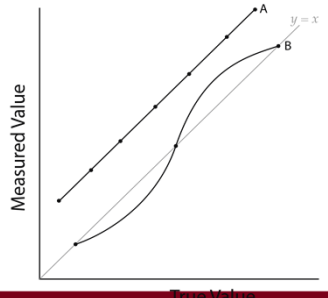
3 4 5

*Minimum number of samples to verify the AMR is 3 (low, mid-point, and high)*





## Non-Linearity may be missed if too few points








## Closeness of sample concentrations or activities to the AMR Upper/Lower limits

- **Use samples that are near upper and lower limits of the AMR**
- **Factors to consider:**
  - **Expected analytic imprecision near the limits**
  - **Clinical impact of errors near the limits**
  - **Availability of test specimens near the limits**





## Analyte concentrations outside the AMR

- May be measured if samples are diluted (or concentrated)
- Diluent should be provided by manufacturer or shown to be suitable
- Remember that dilution may introduce errors
- For certain analytes, laboratories may set limits of dilution and report extremely high values as “greater than” the highest measured value





## Measuring results beyond the AMR

- May report below the AMR *lower limit* by:
  - Concentrating the sample
    - Amicon concentrator
    - Extraction and re-suspension
  - Increasing the ratio of sample to reagent
    - Altering the programming of the instrument





## Measuring results beyond the AMR (cont’d)

- **May report above AMR *upper limit* by:**
  - **Decreasing the ratio of sample to reagent**
  - **Diluting the sample before analysis**
- **Often manufacturer provides information/mechanism for modification**
  - **Autodilution/autoconcentration**
  - **Dilution protocol**
  - **Concentration protocol**

## Control Procedures

- Manufacturer's instructions for controls must be followed
- Purpose is to monitor accuracy and precision of test performance
- General rule is that *at least 2 controls* at different levels must be run each day of patient testing for quantitative assays
- Acceptable tolerances (ranges) must be defined
- There must be a system of reviewing control results and acting immediately on control results outside the acceptable limits
  - Westgard rules are commonly used for monitoring



## Summary of Key Points for Validating a New Assay

- CLIA requires validation/verification of the following for FDA-cleared tests
  - Accuracy
  - Precision
  - Reportable range
  - Reference interval
  - Plan for Calibration/Calibration Verification
  - Plan for Controls