13-METHOD VALIDATION

Prepared by: Yusra Othman /Director/Supervisor-Chem **Date:** May/28/2024
signature/title

Reviewed by: Jordan Dillard /Instructor **Date:** June 25 2024
signature/title

Approved by: Sanford N. Bauch, M.D /Chairman **Date:** July 7 2024
signature/title

ANNUAL REVIEW:

REVIEWED	Sanford N. Bauch, M.D	July-16-2025
	signature/title	Date
REVIEWED		
	signature/title	Date
REVIEWED		
	signature/title	Date
REVIEWED		
	signature/title	Date
REVIEWED		
	signature/title	Date
REVIEWED		
	signature/title	Date

REVISED		
	signature/title	Date/Page/Paragraph
REVISED		
	signature/title	Date/Page/Paragraph
REVISED		
	signature/title	Date/Page/Paragraph
REVISED		
	signature/title	Date/Page/Paragraph
REVISED		
	signature/title	Date/Page/Paragraph

SUPERSEDES: Procedure titled _____

1. **Purpose:**

   1.1To define the required studies needed to validate laboratory tests.

2. **Policy:**

   2.1 This policy apply to all routine tests (Qualitative and Quantitative tests). However, for sections with different scope of tests, they should follow national/International guidelines matching the test.

2.2 Most methods have been well studied by the manufacturers and copies of the relevant documentation or certificates of validation should be obtained and retained, usually available with the method package insert. For most purposes it will only be necessary to perform a Method Comparison (correlation) study against the current method if found correlated, then reference ranges will be adopted. If laboratory is new never provide patient testing before, then manufacturer provided reference range will be adopted with referral to the population used to establish the reference ranges, every effort will be made to verify it locally if able to get volunteers (staff member) or samples from nearby blood bank. Some reference ranges adopt the published well studied by expert with reference to disease like, glucose, A1C, Lipids, 25 OH Vit D and microalbumin/creatinine ratio. Replication (precision) study and a verification of the reportable range (linearity) study, however, where methods have been developed "in-house" or a method has been modified, more extensive evaluation is required.

2.3 If an instrument is moved, the laboratory is responsible for determining that the method performance specifications are not affected by the relocation process or any changes due to the new environment (eg,refer to the manufacturer's manual regarding critical requirements, such as set-up limitations, environmental conditions, The laboratory must follow manufacturer's instructions for instrument set up, maintenance, and system verification. If instruments or equipment are moved, the laboratory must perform appropriate function checks to ensure that they were not adversely affected by the relocation process or changes due to the new environment. This does not apply to portable equipment used following the manufacturer's instructions.

3. Definitions/Abbreviations:

3.1 CLIA: Clinical Laboratory Improvement Amendment
3.2 QC : Quality Control
3.3 EQC : External Quality Control
3.4 LLD : Lower Limit of Detection
3.5 BLD : Biological Limit of Detection

4. Scope:

4.1 Before performing a method validation it is necessary to:

4.1.1 List the desired characteristics of the test
4.1.2 Determine the suitability of the method/instrument for a particular situation
4.1.3 List the characteristics of the methodology (principle, calibration, QC requirements, etc.)
4.1.4 Determine how well the method performs e.g. range, precision, recovery, accuracy, interference etc.

4.2 The following should be considered and carefully defined:

4.2.1 Sample type
4.2.2 Sample volumes
4.2.3 Speed of analysis
4.2.4 Reliability

4.2.5 Specificity
4.2.6 Reagent stability
4.2.7 Equipment required
4.2.8 Complexity
4.2.9 Normal range
4.2.10 Cost
4.2.11 Level of automation
4.2.12 Units reported

4.3 The Validation Plan: Consists of 4 phases

4.3.1**Familiarization period**

4.3.1.1 Establish working procedure
4.3.1.2 Check working range of test using range of samples
4.3.1.3 Check calibration
4.3.1.4 Check detection limit

4.3.2 Preliminary Validation.

4.3.2.1 Perform within-run precision/replication study
4.3.2.2 Perform interference study (manufacturer studies are acceptable for FDA approved test
4.3.2.3 Perform recovery study (only if Method Comparison not done or difficult to do, usually by comparing the measured value to the accepted value, this is achieved running quality controls and compare the achieved mean with the target manufacturer mean or peer group mean if available)
4.3.2.4 Assess analytical acceptability/Linearity/Reportable range
4.3.2.5 Sensitivity, lower limit of measurement or Limit of quantification. The manufacturer provided will be adopted for FDA approved test

4.3.3 Final Validation.

4.3.3.1 Perform total precision/replication study
4.3.3.2 Comparison of methods study
4.3.3.3 Assess reportable range
- Analytical Measurement Range (AMR)

4.3.3.4 Verify reference interval/reference range
4.3.3.5 Perform carry over studies, if common prob is used for sampling
4.3.3.6 Document studies
4.3.3.7 The report of the validation results should be sign by laboratory director, with statement that its find to be accurate for patients testing

4.3.4 Implementation.

4.3.4.1 Select QC procedure
4.3.4.2 Write Operation policy or short instruction and refer to manufacturer operation manual
4.3.4.3 Train analysts
4.3.4.4 Introduce method
4.3.4.5 Monitor routine performance

5. Procedure: THE VALIDATION STUDIES

The performance of the device/method during the evaluation period must be confirmed using quality control samples.

- A single lot of reagents and calibration materials may be used during the evaluation, but introducing several lots of these materials might give a better estimate of performance.
- All date, data, QC results, calibration data and reagent records must be recorded and included and maintain as part of the evaluation and kept 2 years after discontinuation of the current instrument/assy.

5.1 REPLICATION OR PRECISION STUDY.

5.1.1 Must be performed on all new or modified methods, kits and instruments.
5.1.2 Estimated PRECISION or RANDOM ERROR
5.1.3 Estimates variation due to pippetting, timing, different operators, instability etc
5.1.4 Ideally variation should be small.
5.1.5 Test the same material (within:10 is acceptable between run 15-20 replications.
5.1.6 As much as possible use sample matrix close to that of expected patients
5.1.7 Calculate mean, standard deviation and coefficient of variation
5.1.8 Later check performance over a 20-day period will be preferred as this includes all components of variability.

5.1.1 Calculations: Use a statistical/scientific calculator or MS Excel spreadsheet.

5.1.2 Acceptance Criteria:

- Judgment of acceptability depends on allowable analytical error. It is recommended that the CLIA criteria for allowable total error (TE) are used.
- Long term (20 day) imprecision is defined as less that ½ -1/3 of CLIA total error or ½ of intra individual variation. CAP/API return is a good source as well. Having target CV not to exceed CAP CV at the normal range for a particular assay
- Short-term precision (within-run or within-day) should be less than

½ to 1/3 of the allowable total error or ½ intra individual variation having target CV not to exceed CAP average CV of all samples for a particular assay

5.2 COMPARISON OF METHODS.

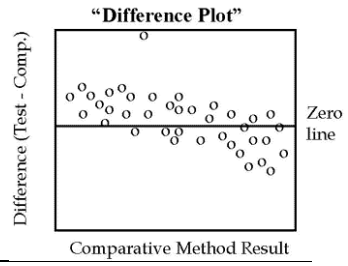
- Performed for all new methods in comparison to old in use method for established laboratory
- Performed to estimate inaccuracy or systematic error. Samples are analysed by the new method and a comparative method, then error is estimated from observed differences. The comparative method is usually the "current" method or reference lab using same methodology or manufacturer provided samples for comparison
- Where possible a "reference method", which is traceable to standard reference materials, should be used for comparison.

5.2.1 METHOD:

- A minimum of 40 patient samples is preferable, but 20 well distributed over the linearity range will do. CAP/API proficiency samples which were stored appropriately may be analyzed and compared between the 2 current compare methods or compare against the method peer mean as reported by CAP/API. This provides a better coverage of the entire working range of the method and spectrum of diseases expected on practice. Twenty carefully selected samples will provide better information than 100 randomly selected samples not covering the method reportable range. If duplicate are not performed then specimens showing large method differences should be retested
- Specimen stability is a factor and samples should be tested by the two methods within the stability claimed by the manufacturer.
- Data should be graphed and visually inspected for discrepancies. All calculations and graphs should be electronically incorporated in a worksheet.

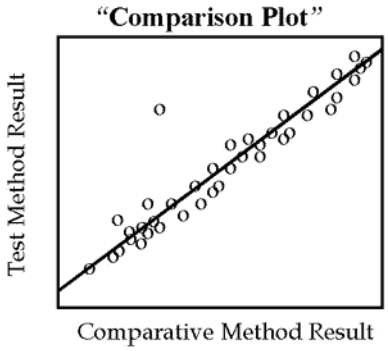
5.2.2 GRAPHS:

- A **Difference Plot:** Highlights large differences between the methods and shows which results need to be repeated. Plot the reference method results on the x-axis and the difference between the test minus the reference method results on the y-axis. The differences should scatter evenly around the zero line. Outliers should be repeated.



- **A Comparison Plot:**

Used where one-to-one agreement is not expected, e.g. enzyme assays. Plot the test results on the y-axis against the reference method results on the x-axis. Draw a best-fit line. Discrepant results are then visibly identified. The correlation equation of $Yc = a + bXc$ will be the best representing the correlation study



5.2.3 CALCULATIONS:

Numerical estimates of error can only be obtained from statistics. For results covering a wide range, such as glucose and cholesterol, linear regression is preferred. This allows estimation of systematic error at more than one medical decision level. Programs such as Excel, psmile source excel sheet, EP-evaluator will provide linear regression calculation of the **slope (b)** and the **y-intercept (a)** and the SD of the points around the best-fit line.

The **Systematic error (SE)** for a given medical concentration (X_c) is determined by calculating the corresponding y-value (Y_c) from the regression line, then taking the difference between Y_c and Xc as follows:

$Yc = a + bXc$

$SE = Yc - Xc$

e.g. In a cholesterol method comparison the regression line is $Y = 0.05 + 1.03X$

i.e. the y- intercept is 0.05 mmol/l and the slope is 1.03

The Y value corresponding to a medical decision level of 5.2 mmol/l would be 5.4 mmol/l

The **correlation coefficient (r)** is useful for assessing whether the data range is wide enough to provide good estimates of the slope and intercept, rather than judging method acceptability. When *r* is less than 0.95 it would be better to collect additional data to expand the concentration range.

5.2.4 ACCEPTANCE CRITERIA.

This is complicated by the fact that any individual test result is subjected to random error in addition to systematic error. Thus the Total Error (TE) is the Systematic Error (SE) plus the Random Error (RE). This is calculated as follows:

$$\mathbf{TE}_{\mathbf{calc}} = \mathbf{SE} + \mathbf{RE}$$

$$\mathbf{TE}_{\mathbf{calc}} = \mathbf{bias}_{\mathbf{meas}} + \mathbf{3s}_{\mathbf{meas}}$$

Where s_{meas} is the SD from the replication study and $bias_{meas}$ is the difference of $Y_c - X_c$ from the regression statistics.

Total error is then compared with the CLIA criteria for acceptable performance and to CAP criteria when CLIA is not available. If both not available, clinical director will identify the acceptable criteria.

R: 0.71 to 0.9 are strong correlations; and correlations > 0.9 are excellent

Bias best to be +/- 5% of the mean of sample used for comparison

Slope: between +/-1

5.3 LINEARITY OR REPORTABLE RANGE STUDY.

5.3.1 To assess the lowest and highest test results those are reliable and reportable.
5.3.2 CLIA recommends that laboratories verify manufacturer's claims in this regard for tests of high complexity.
5.3.3.CAP or other reliable source provided linearity material is acceptable to assess linearity
5.3.4 A series of samples of known concentration is tested and plotted on the y-axis against the known values on the x-axis.
5.3.4 Different levels or concentrations should be used 3 to 5 level
5.3.5 Standard solutions, patient dilutions or diluted pooled sera, can be used, diluted to cover the range compare achieved measured to expected after dilution.
5.3.5 Follow manufacturer's instruction regarding suitable diluent, e.g. saline, water or bovine albumin etc.

Prepare dilutions as below:

Dilution Preparation:

- Use 2 pools: one near the lower limit and one near the expected upper limit
- Label these **Pool 1** and **Pool 5** respectively.
- Prepare **Mixture 2** [75/25] – 3 part Pool 1 plus 1 part Pool 5.
- Prepare **Mixture 3** [50/50] – 2 part Pool 1 plus 2 parts Pool 5.
- Prepare **Mixture 4** [25/75] - 1 part Pool 1 plus 3 parts Pool 5.
- Perform 3 to 4 replicates on each mixture, including the original pools.

Evaluation

- Plot the mean of the measured values on the y-axis against the known values on the x-axis.
- Visually inspect for a linear relationship.
- Linear regression analysis can be performed with $R > 0.95$ but it is generally accepted that reportable range can be adequately determined using the preceding method.
- Record all results, conclusions and the estimated reportable range.

5.4 DETECTION LIMIT STUDY:

- **Known as**: sensitivity, analytical sensitivity, quantitative detection limit, functional sensitivity, limit of detection, and limit of quantitation.

- **Optional:** need only be performed where this data is not available from manufacturer/supplier of method and if non FDA approved test.
- Estimates lowest measurable concentration of analyte.
- Obviously important in drug testing.
- Application for tumor markers e.g. PSA
- Good laboratory practice indicates that detection limits need to be verified when relevant.

Preparation:

- Prepare two different samples: a "blank" with zero conc. of the analyte under test, and a "spiked" sample with a low conc. of the analyte.
- **Blank:** Common to use a zero standard from a set of calibrators.
  Spiked: The amount of analyte added should be within the expected detection range.
- Twenty replicate measurements are recommended though for some tests 10 measures may be more cost-effective and practical.
- Perform study over 1 day if focus is on the 'blank'. If the interest is with the spiked samples then a longer day-to-day study is indicated, i.e. 10 measurements over 10 days.

Quantity to be Estimated :

- **Lower Limit of detection (LLD):**
  The mean of the blank sample plus 2 or 3 times the SD of the blank. It is preferable to use raw values such as absorbence to calculate the mean and SD

and then convert to concentration, to eliminate the effect of rounding off. LLD or "limit of absence" is the measurement results possible when the actual concentration is zero.

- **Calculation: LLD = mean_{blk} + $2s_{blk}$** based on 10 replicates.
- **Biological Limit of Detection (BLD):**
  Estimated as the LLD plus 2 to 3 times the SD of the spiked sample. BLD is defined as that concentration which is really different from a zero value.
- **Calculation: BLD = LLD + $2s_{spike}$**
- **LOQ: Lower limit of Quantification: The lowest quantity to measure which can be achieved with CV < 25%, 10-5 replicate measurement, clcultae, mean, SD and CV**

Calculated Example:

The blank and spiked samples are the zero and 10ug /L standards.

Both samples analysed 10 times and the mean and SD in raw units (eg OD or absorbence) calculated.

Blank: mean = 1000 units, SD = 100 units

Spiked: mean = 2000 units SD = 200 units

Spiked – Blank = 2000 – 1000 = 1000 units

LLD = 2 SD (blank) = 200 units or 2 ug/L

2 SD (spike) = 400 units or 4 ug/L

BLD = LLD + $2s_{spike}$ = 2 ug/L + 4 ug/L = 6 ug/L

5.5 INTERFERENCE STUDY

- **Only necessary when methods have been modified by the laboratory.**
- Estimates the effects of specific materials on the accuracy of the method.
- The results of 2 aliquots of a sample are compared after one has had a quantity of the suspected interfering substance added. Test substances based on manufacturer's claims as well as common interferences such as bilirubin, haemolysis, lipaemia etc.

Method:

- Prepare 2 aliquots of each patient or pooled sera.
- Add an appropriate volume/concentration of the interfering substance (I) to one of the samples; add an equal volume of appropriate diluent (dil) to the other.
- Analyze each aliquot in duplicate and determine the mean concentration of analyt in each. Use Excel spreadsheets, where possible to perform calculations.
- Compare the results for any differences in values.
  Calculated Example:

	Duplicates	
Sample A (+I)	110 mg/dl	112 mg/dl
Sample A (+dil)	98 mg/dl	102 mg/dl

Calculate average concentration:

Sample A (+I) 111 mg/dl

Sample A (+dil) 100 mg/dl

Calculate difference of paired samples:

Sample A = 11 mg/dl

Calculate the average difference for all specimens tested at given level of interference

e.g. 11 mg/dl

Acceptance Criteria:

Compare the error due to interference to the allowable error for the test (CLIA data). If the allowable error is 10% and the upper reference range is 100 mg/dl then an observed interference of 11 mg/dl is greater than the allowable error and the method performance is unacceptable.

5.6 RECOVERY STUDY

- Using Quality control by comparing achieved mean to manufacturer mean or peer group mean if available.
- only when Method Comparison cannot be performed due to unavailability of another method or lab never provide test before.
- Performed to estimate proportional **systematic error** whose magnitude increases with the concentration of analyte.
- Also can be tested the same way as to the interference study except the analyte to be measured is added rather than an interfering substance. A standard or calibration solution of the analyte under measurement is usually added.

Method:

- No more than 0.1 ml of standard solution is added to 0.9 ml of patient sample.
- Pipetting accuracy is critical.
- Add sufficient concentration of standard solution to reach the next decision level for the test. The standard solution needs to be 10 times the desired level of addition for the dilution of 1 in 10 used.
- Perform duplicate analyses. Where any drug metabolites might affect recovery then samples from many different patients should be tested.

Calculations:

- Use a spreadsheet to perform calculations.
- Recovery is expressed as percentage. Ideal is 100% and a recovery of 95% represents a proportional error of 5%.
  - Using calcium as an example:
  - If 0.1 ml of a 5 mmol/l standard is added to 1 ml serum then the amount added is
    5 x (0.1/1.1) = 0.46 mmol/l
  - Tabulate the replicates:

Sample A (added)	2.85	2.90	= avg.	2.88 mmol/l
Sample A (diln)	2.43	2.48	= avg.	2.45 mmol/l

Difference = 0.43 mmol/l

Recovery = difference/amount added X 100

Recovery = 0.43/0.46 X 100 = 93.4%

Proportional error = 100 – Average recovery

= 100 – 93.4 = 6.6%

Acceptance Criteria:

Based on the CLIA total allowable error for the analyte, where this is given as a percentage, this is compared to the proportional error calculated as above. Where the TE is given as a concentration then the TE should be converted to a percentage of the center of the reference range.

e.g. for calcium CLIA TE is 1 mg/dl. At the center of the reference range, 10 mg/dl, this represents 10%. Thus a measured error of 6.6% meets the criteria for acceptability of the method.

5.7 Carry Over study

This will be performed to validate the instrument probe is clean and will not carry any material from previous sample or reagent. For reagent the carryover will be tested as the manufacturer recommendation. However, for sample carry over, run several replicate of low samples after several replicate of high samples The carryover is less than the error limits which bases upon three times the Low-Low SD, usually not to exceed 2%

5.8 Validation for intermittent testing and Qualitative methods

5.8.1 Intermittent Testing:

When a test is put back into service after being stop for more than 6 months, the following requirements must be met:

1. PT or alternative assessment sample (e.g. previous CAP/API or comparison with other CLIA/CAP accredited laboratory) performed prior to restarting patient testing.
2. Method performance specifications verified, precision, within and between, and accuracy

3. Competency assessed for analysis prior to restarting patient if the test handling is different from daily other routine testing.

5. 8.2: Method validation for qualitative:

1. The validation study for qualitative assay will be the same as for quantitative for the assessment of precision analytical specificity and the carry over. Linearity with AMR and CRR as well as analytical sensitivity are Not applicable for qualitative tests. However the threshold value need to be verified when the test first introduce and biannually; if the test have quantitative results interpretive as negative and positive as per the threshold. However, if results are positive and negative, and the test performed as per manufacturer without any modification or it is POCT and used exactly as per the manufacturer recommendation, the threshold will be verified only upon introduction of the test. The total error allowed (TEA) is the responsibility of the section chief to identify, CAP return report is of great help as a reference .
2. **Accuracy** is the true value of a substance being measured. Verification of accuracy is the process of determining that the test system is producing true, valid results.
    a. Determine your comparison or reference method
        i. The comparison method must be previously validated.
        ii. The comparison method must be currently performing successfully on EQA.
        iii. Comparison to a method currently in use in the laboratory is preferred if this method meets the above criteria.
        iv. Samples with known values, such as proficiency testing samples or commercial standards, may be used as the reference method.
        v. Examples of possible reference methods for qualitative tests include in-house or external quantitative methods, use of a reference lab, and correlation with another lab using a previously validated method, use of EQA samples or other commercially prepared reference material with known values.
    b. Sample Criteria
        i. A minimum of 10 samples for each expected result. For example, if a test method gives results of "Positive/Negative", the accuracy study must include 10 known positives and 10 known negatives.
    c. Testing:

        Two levels of quality control must be run each day that testing is performed, not including controls internal to the kit cartridge/testing device.

    d. Acceptability criteria:

The performance of qualitative tests is most commonly described in terms of sensitivity and specificity. The table below is a contingency table that compares the results of a qualitative test with the outcome of the diagnostic accuracy criteria. The entry in each cell of the table represents the number of specimens corresponding to the labels in the margins.

Method being Validated	Diagnostic Sensitivity and Specificity (Results from Comparison Study)		**Total**
	Positive	Negative	
Positive	***# true positive (TP)***	***# false positive (FP)***	***TP+FP***
Negative	***# false negative (FN)***	***# true negative (TN)***	***FN+TN***
Total	***TP+FN***	***FP+TN***	***N***

i. Calculate the estimated Diagnostic Sensitivity(True positive rate) = 100 x [TP/(TP+FN)]
ii. Calculate the estimated Diagnostic Specificity(True negative rate) = 100 x [TN/(FP+TN)]
iii. Calculate the percent Positive Agreement (Positive Predictive Value) =100 x TP/(TP+FP)
iv. Calculate the percent Negative Agreement (Negative Predictive Value) =100 x TN/(TN+FN)
v. Compare the results calculated above with the manufacturer's stated claims for Sensitivity, Specificity and Agreement found in the test kit package insert.
vi. Results must be equal to, or greater than, the manufacturer's claims for the method to be considered accurate. Or achieved > 90% specificity and sensitivity.

5.9 Reference Ranges Verification:

Reference range verification will be tested by selecting 20 normal individual, e.g Laboratory staff or Blood donor, if 18 out of 20 must be within the manufacturer provided, if not then additional of another 2 sample if still out then establishment of the reference ranges will be a must. However it is not required for tests like glucose, cholesterol, enzymes, hormones, drugs etc. where the reference range is more relevant for interpretation or international guidelines are published and followed.

5.10 Final Approval:

- The Section Head/Consultant/ Chief in Charge of the Laboratory performing the validation should sign approval of the relevant worksheets.
- Summaries the findings of the validation/evaluation.

8.0 References:

8.1 Basic Quality Assurance and Quality Control in the Clinical Laboratory, A/ Wayne Bruce, Little Brown & Company, 1984.
8.2 Preliminary Evaluation of Quantitative Clinical Laboratory Methods, EP10-A, NCCLS, Vol. 17, No.18.

8.3 Evaluation of Precision Performance of Clinical Chemistry Devices;
Approved Guideline, EP5-A, NCCLS Vol.17, No.18.
8.4 Method Validation, James O. Westgard:http://www.westgard.com/lesson20.htm
8.5 CLIA Requirements for Analytical Quality:http://www.westgard.com/clinical.htm
NB: The web sites above served as the principal references for this document and provide excellent sources of further detail and practical examples.
Many chart from www.psmils.org is very valuable, method comparison, linearity and precision.