EP Evaluator ®

Method Comparison Report 1

For Regulatory Review

4/23/2018

Prepared for

John Swenson, MD
Medical Director

Prepared by

Jane Dow
Clinical Laboratory

Accepted by

_____

Signature

_____

Name / Title

_____

Date

EP Evaluator $^{®}$

Table of Contents

Accuracy and Linearity

	Assigned	Pct	N	Accuracy & Recovery			Linearity		
				Mean	% Rec	Status	Estimate	Residual	Status
CalKit-1	25.0	--	4	25.5	102.0	Pass	26.5	-1.0	Pass
CalKit-2	100.0	--	4	101.5	101.5	Pass	101.4	0.1	Pass
CalKit-3	250.0	--	4	249.5	99.8	Pass	251.3	-1.8	Pass
CalKit-4	400.0	--	4	407.8	101.9	Pass	401.1	6.6	Pass
CalKit-5	700.0	--	4	690.3	98.6	Pass	700.8	-10.6	Pass
CalKit-6	1000.0	--	4	984.5	98.5	Pass	1000.5	-16.0	Pass

See User's Specifications for Pass/Fail criteria.

Linearity Summary

	Overall	--
Slope	0.999	
Intercept	1.5	
Obs Err	1.0 mg/dl (conc) or	
	1.6%	
N	6	

LINEAR within SEa of 1.5 mg/dl (conc) or 2.5%

Experimental Results

CalKit-1	24	26	25	27
CalKit-2	101	102	100	103
CalKit-3	243	252	248	255
CalKit-4	400	410	409	412
CalKit-5	690	696	680	695
CalKit-6	970	990	985	993

X: Excluded from calculations
T: Exceeds Allowed Total Error

User's Specifications

Allowable Total Error	6 mg/dl (conc) or 10.0%
Systematic Error Budget	25%
Allowable Systematic Error	1.5 mg/dl (conc) or 2.5%

Supporting Data

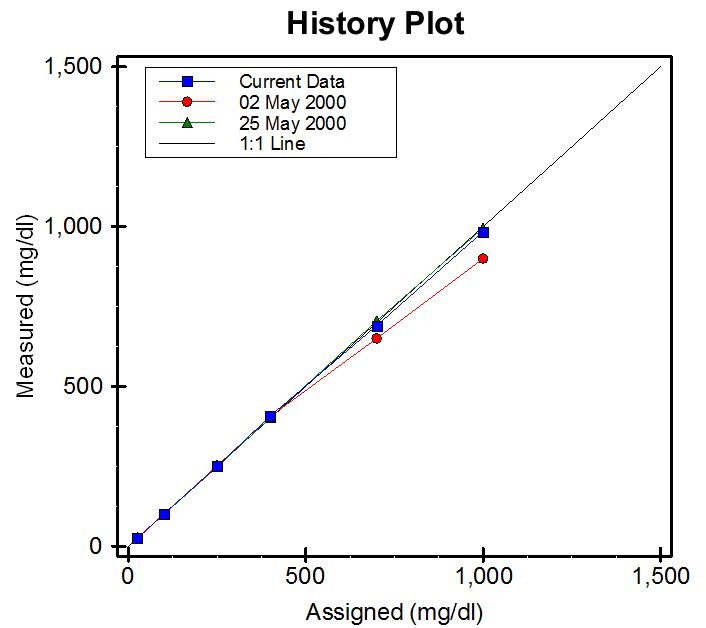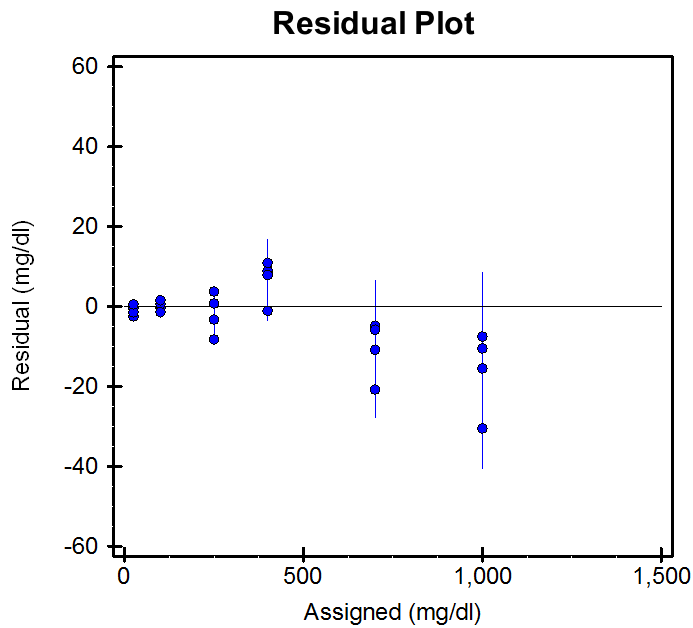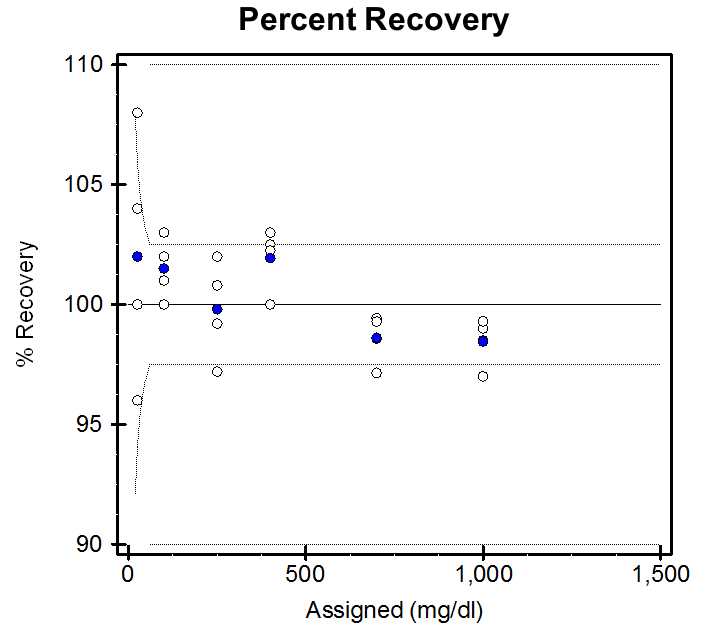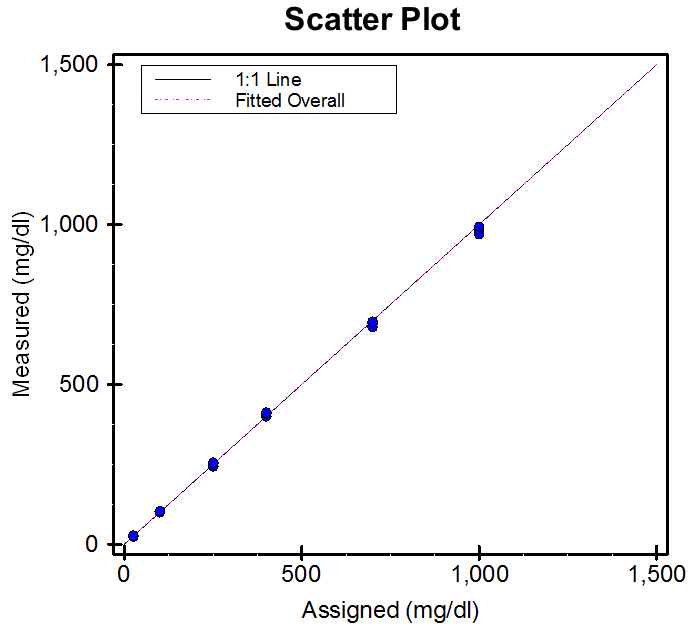| Analyst	Kate Doe
Date	01 Jun 2000
Value Mode	Pre-Assigned
Units	mg/dl
Controls	Krogers 123 exp 31 Jan 2024
Reagent	Krogers 159 exp 31 Jan 2024
Calibrators	CVS 789 exp 31 Jan 2024
Comment	comment 2

Evaluation of Results

The Accuracy and Linearity of GLUCOSE were analyzed on ASSAYER over a measured range of 25.5 to 984.5 mg/dl. Reportable Range was not evaluated in this experiment. This analysis assumes accurate assigned values. Allowable systematic error (SEa) was 1.5 mg/dl (conc) or 2.5%. The accuracy test PASSED. The maximum deviation for a mean recovery from 100% was 2.0%. 6 of 6 mean recoveries were accurate within the SEa. 24 of 24 results were accurate within the allowable total error (TEa) of 6 mg/dl (conc) or 10.0%. The results are LINEAR.

Accepted by: _____   _____

Signature   Date

EP Evaluator®

Product Strategy -- Data Innovations

GLUCOSE
Instrument ASSAYER

Accuracy and Linearity

Scatter Plot

Percent Recovery

Residual Plot

History Plot

EP Evaluator®

Product Strategy -- Data Innovations

Linearity

Report Interpretation Guide

In EP Evaluator, Linearity experiments are experiments that use specimens with defined concentrations. The Linearity module can also evaluate Calibration Verification, Accuracy, Reportable Range and Precision. This means that you can verify three of the four major CLIA '88 requirements with one properly designed experiment.

User-selectable options determine which of these analytical properties the report verifies. Also, the user may request Pass/Fail flags against a specific allowable error criterion, or he/she may simply report selected statistical measures.

Experiment Procedure: Replicate measurements are made on 3-11 specimens, with (known) concentrations spread across the reportable range. Ideally, the lowest and highest specimens should challenge the limits of the range.

Accuracy (or Recovery)

Definition: The ability to recover the correct amount of analyte present in the specimen.

Verification process: Accuracy can be verified only when the "correct" amount of analyte **(the Assigned Value)** is known. While it is possible to evaluate recovery using a single replicate, assaying 2 to 4 replicates provides a more reliable estimate.

Key statistic: Recovery = 100 x Measured Mean / Assigned Value

Reportable Range

Definition: As used in CLIA, Reportable Range refers to the Analytical Range or Assay Range -- the maximum range of values that can be assayed accurately without dilution. The CAP term "Analytical Measurement Range" (AMR) is a synonym for Reportable Range.

Verification Procedure: Reportable Range is verified if two conditions are met: 1) the assigned values of the lowest and highest specimens are within proximity limits of the Reportable Range limits, and 2) these two specimens are acceptably accurate.

Proximity Limits: Proximity Limits define how close the lowest and highest specimens must be to the Reportable Range limits.

Calibration Verification

Calibration Verification verifies whether a method is properly calibrated by verifying both Accuracy and Reportable Range. The report is titled to match the CAP and CLIA regulatory requirements.

CLIA requires a minimum of three specimens, each assayed in duplicate. Two specimens challenge the lower and upper limits of the reportable range. The third specimen is somewhere in between.

Linearity

Several definitions are in common use. Among them:

- **Traditional Linearity** (CAP Visual Inspection): Draw a scatter plot with assigned values on the X-axis and measured mean on the Y-axis. If it looks like a straight line, the method is linear.

- **Statistical Linearity** (CLSI EP6P and EP6-A): These procedures determine acceptability based on statistical significance (i.e., p-values) rather than medical significance. EP Evaluator does not compute Statistical Linearity.

- **Clinical Linearity:** The method is linear if it is possible to draw a straight line that passes within a user-defined allowable error of each specimen point.

Related concepts:

- **Best Fit Line:** If the user opts to verify Linearity, this line is obtained using the Clinical Linearity algorithm. Otherwise it is a regular linear regression line.

- **Outliers:** When verifying Linearity, the program first tries to determine an acceptable line using all specimens. If it fails, it then tries to find some subset of at least three specimens that are linear within allowable error. Specimens not in this acceptable subset are classified as outliers.

- **Slope and Intercept:** Coefficients of the Best Fit Line. The ideal slope is 1.00; the ideal intercept is zero.

- **Observed Error:** For Clinical Linearity, this is the minimum allowable error that could be defined for a data set and still have it be linear.

- **Standard Error of Estimate:** For regular regression, this measures the dispersion of the data points around the Best Fit Line.

- **Residual:** The difference between the best fit line and either an individual result or a mean measured value, depending on context.

- **RSQ:** The coefficient of determination r-squared (rsq) appears in the report's "Linearity Summary" or detail screen plot when enabled in the Preferences\Reports menu. It does not appear when "Clinical Linearity" is selected.

Precision

Definition: Ability to obtain the same result upon repeated measurement of a specimen.

Verification Process: Measure the specimen many times.

Linearity
Report Interpretation Guide

Compute the SD and CV, and verify that they are acceptably small. While 2-4 replicates are adequate for assessing accuracy, a minimum of 10 (and preferably 20 or more) is required to verify Precision.

The **Precision Index** is the ratio of SD to Allowable Random Error (defined below). The ideal -- and probably unattainable -- Precision Index is zero. A value of 1.00 indicates borderline acceptability. Any further increase in SD would exceed allowable error.

The **95% Confidence Interval** (CI) for the Precision Index indicates how much sampling variation might be expected. The CI narrows as the number of replicates increases.

Allowable Total Error (TEa), and the Error Budget

TEa states the laboratory's policy for how much error is medically (or administratively) acceptable. Regulatory requirements represent an upper limit. Example: the CLIA limit for Sodium is 4 mmol/L.

Total Error has two major components: **Systematic Error** (synonym Bias) and **Random Error** (synonym Imprecision). The **Error Budget** allocates a fraction of the Allowable Total Error for Systematic Error, and the remaining fraction for Random Error. Establishing an appropriate Error Budget allows the lab to control accuracy and precision separately, with reasonable confidence that Total Error will also remain in control. The recommended upper value for the Systematic Error Budget is 50%; for the Random Error Budget it is 25%.

Pass or Fail?

The program reports Pass/Fail for Accuracy and Linearity based on Allowable Systematic Error (SEa). Pass/Fail for Precision is based on Allowable Random Error (REa).

- A specimen passes Accuracy if its mean measured value is within SEa of the Assigned Value. Note that if the lowest standard has an assigned value of zero, then the TEa conc must not be blank or "0". Otherwise the experiment will fail.

- The experiment passes Linearity if it is possible to draw a straight line (on the scatter plot of mean measured value vs. assigned value) that passes within +/- SEa of each specimen point.

- A specimen passes Precision if SD does not exceed REa.

- The experiment passes Reportable Range if 1) the assigned values of the lowest and highest specimens are within proximity limits of the Reportable Range

Limits, and 2) these two specimens also pass accuracy.

Preliminary Report

The word PRELIMINARY printed diagonally across the report indicates that the data is incomplete, and the report is not acceptable as a final report. Some or all of the statistics may be missing.

The Linearity report is preliminary if there are less than three specimens.

Two Instrument Comparison

X Method XT   Y Method XS

Scatter Plot

Error Index
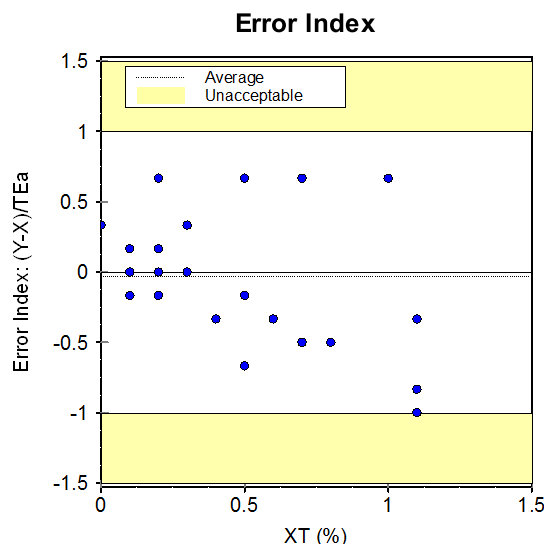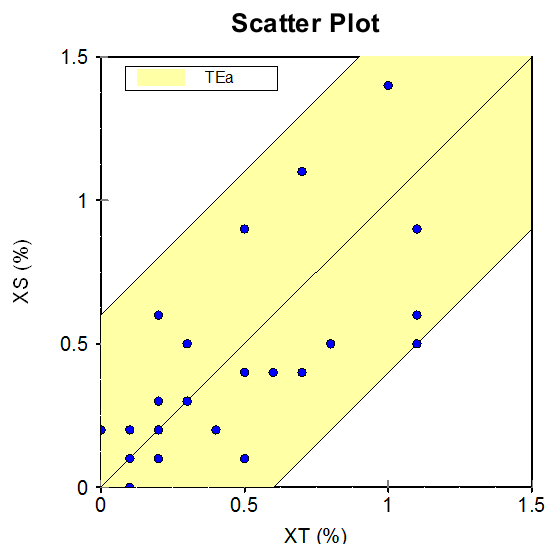
Evaluation of Results

Base was analyzed by methods XT and XS to determine whether the methods are equivalent within AllowableTotal Error of 0.6 % (conc) or 38.5%. 34 specimens were compared over a range of 0.0 to 1.1 %. The test Passed. The difference between the two methods was within allowable error for 34 of 34 specimens (100%). The average Error Index (Y-X)/TEa was -0.03, with a range of -1.00 to 0.67. The largest Error Index occured at a concentration of 1.1 %.

Key Statistics		Deming Regression Statistics	
Average Error Index	-0.03	**Y = Slope * X + Intercept**	
Error Index Range	-1.00 to 0.67	Correlation Coeff (R)	0.7109
Coverage Ratio	--	Slope	0.897 (0.647 to 1.147)
		Intercept	0.02 (-0.12 to 0.17)
Evaluation Criteria		Std Error Estimate	0.26
Allowable Total Error	0.6 % (conc) or 38.5%	N	34 of 35
Reportable Range	--		

Experiment Description

	X Method	Y Method
Expt Date	05 Feb 2018	05 Feb 2018
Result Ranges	0.0 to 1.1	0.0 to 1.4
Mean ± SD	0.44 ± 0.37	0.41 ± 0.34
Units	%	%
Analyst	Admin	Admin
Reagent Src		
Reagent Lot		
Reagent Exp		
Comment		

Accepted by: _____

Signature   Date

Two Instrument Comparison

X Method XT **Y Method XS**

Experimental Results

Specimen	X	Y	Error Index	Specimen	X	Y	Error Index	Specimen	X	Y	Error Index
S00001	0.1	0.2	0.17	S00013	0.1	0.2	0.17	S00025	0.5	0.9	0.67
S00002	0.2	0.3	0.17	S00014	0.4	0.2	-0.33	S00026	0.1	0	-0.17
S00003	0.1	0.1	0.00	S00015	0.3	0.3	0.00	S00027	0.7	1.1	0.67
S00004	0.2	0.2	0.00	S00016	0.2	0.1	-0.17	S00028	0	0.2	0.33
S00005	0.6	0.4	-0.33	S00017	0.3	0.5	0.33	S00029	1.1	0.6	-0.83
S00006	0.1	0.2	0.17	S00018	0.2	0.2	0.00	S00030	0.1	0.2	0.17
S00007	0.1	0.1	0.00	S00019	0.7	0.4	-0.50	S00031	1.1	0.9	-0.33
S00008	0.5	0.1	-0.67	S00020	0.1	0	-0.17	S00032	1.1	0.5	-1.00
S00009	0.2	0.3	0.17	S00021	0.7	1.1	0.67	S00033	0.5	0.4	-0.17
S00010	0.8	0.5	-0.50	S00022	0	0.2	0.33	S00034	0.3	0.3	0.00
S00011	0.2	0.6	0.67	S00023	1.1	0.9	-0.33	*Spec ID*	*X*	*X*	*Y*
S00012	1	1.4	0.67	S00024	1.1	0.5	-1.00				

Values with an "X" were excluded from the calculations.

EP Evaluator ®

Two Instrument Comparison
Report Interpretation Guide

The Two Instrument Comparison (2IC) Module. Two Modes are available in this module, and this Report Interpretation Guide includes a separate section for each mode.

- The regular (default) 2IC mode is a simple, straight-forward procedure for comparing two methods without using linear regression. The pass-fail rule is easy to understand: two methods X and Y are the same within allowable error (i.e., are clinically identical) if the difference between the two does not exceed Allowable Total Error (TEa) for each specimen. Regular 2IC requires at least 5 specimens.

- The optional CLSI EP26 mode is intended to assist the user in verifying the equivalency of new candidate lot numbers of reagents vs a comparative or reference lot number of the same test kit (same vendor part number). It follows the CLSI EP26 A Guideline: User Evaluation of Between-Reagent Lot Variation; Approved Guideline (2014), and requires at least 3 specimens

Report Interpretation Guide for regular 2IC

Clinical Equivalence vs. Statistical Equivalence

Traditional method comparison protocols test whether two methods are statistically equivalent. In contrast, 2IC tests whether the two methods are clinically equivalent. What's the difference?

Statistical Equivalence. Two methods are statistically equivalent if the difference between them is random. It does not matter how large the difference is.

The CLIA limit for Calcium is 1 mg/dL. However, two Calcium tests could differ by 3 mg/dL and still test as statistically identical -- if no bias is present. Suppose you constructed an experiment by arranging the specimens in increasing order by X, with a +3 mg/dL error at the even-numbered points and a -3 mg/dL error at the odd-numbered points. On average, the bias is zero. A regression-based comparison considers these methods "identical", even though the error is far greater than TEa.

Clinical Equivalence. Two methods are clinically equivalent if the difference between the two is less than allowable error for each specimen. It does not matter if there is a bias, as long as that bias is within allowable error.

For example, suppose Calcium for the Y method is, in every case, 0.2 mg/dL higher than for the X method. These methods are not statistically identical, because the difference is not random. However, they are clinically identical, since the difference is less than the allowable error of 1.0 mg/dL.

Data Quality

2IC gives a meaningful comparison with fewer data points than are required for traditional, regression-based procedures -- perhaps 5 to 10 specimens. However, it is important that the specimens cover the reportable range of the method, and include points near the Medical Decision Points.

Key Statistics and Evaluation Criteria

Allowable Total Error (TEa). TEa states the laboratory's policy for how much error is medically (or administratively) acceptable. Regulatory requirements represent an upper limit.

Examples: the CLIA limit for Sodium is 4 mmol/L; the CLIA limit for Glucose is 6 mg/dL or 10%, whichever is greater.

Error Index. The ratio of the difference Y-X to Allowable Total Error (evaluated at X). The Error Index is measured for each X-Y pair. An index greater than 1.00 or less than -1.00 is unacceptable -- it means the difference between the methods exceeds TEa. If an excessive number of specimens have an unacceptable Error Index (EI), the experiment fails. Excessive number of specimens occurs if the EI is unacceptable for at least one specimen if N < 20 or if the EI for more than 5% of the specimens is unacceptable when N >= 20.

Reportable Range. The maximum range of values that can be measured accurately without diluting the specimen. Synonyms: Analytical Measurement Range, Assay Range, Analytical Range.

Coverage Ratio. The percent of the Reportable Range covered by the analysis. The ideal is 100%.

Coverage is computed as 100 x (Xhi - Xlo) / (Rhi - Rlo). Rlo and Rhi are the lower and upper limits of the Reportable Range. Xhi is the smaller of maximum X and the upper limit of the range. Xlo is the larger of minimum X and the lower limit of the range.

Example: Suppose the Reportable Range is 100 to 300, and the range of X values measured is 200 to 350. There is a 50% overlap between the X-range and the Reportable Range. The coverage ratio is 100 x (300 - 200) / (300 - 100) = 50%.

Deming Regression Statistics. The report shows the slope, intercept, standard error of estimate, and correlation coefficient for reference only. They are not used to determine whether the experiment passes or fails. Note that the slope and intercept are computed from Deming regression, assuming that the two methods have comparable precision (i.e., the same representative SD). The coefficient of determination r-squared (rsq) appears in

EP Evaluator®

Two Instrument Comparison

Report Interpretation Guide

---

parenthesis when enabled in the Preferences\Reports menu.

Plots

Scatter Plot. The Scatter plot shows the data points, together with the 1:1 (Y=X) line with Allowable Error bounds around it. Vertical lines mark the Reportable Range and Medical Decision Points (if these values were input). The plot does NOT show the regression line.

Error Index Plot. The Error Index plot shows the Error Index. Points that fall in the shaded area have an unacceptable Error Index (> 1.00 or < -1.00).

Logarithmic Scale. When the data covers a wide range and has a tight cluster of points at the low end, the scatter plot may be shown on a logarithmic scale instead of a linear scale. Using a log scale gives better visual separation of the points at the low end.

When the plots are shown on a log scale, the regression fit is computed for the logs of the data points, and the Standard Error of Estimate (SEE) represents percent variation around the regression line. (SEE of 0.10 means 10% variation.)

Pass or Fail?

The objective is to determine whether the two methods are clinically equivalent. The experiment passes if the Error Index is acceptable for a suitable number of specimens. See Error Index above for a complete discussion.

Preliminary Report

The word PRELIMINARY printed diagonally across the report indicates that the data is incomplete, and the report is not acceptable as a final report. Some or all of the statistics may be missing.

The 2IC report is preliminary if there are less than 5 unexcluded data points.

Report Interpretation Guide for CLSI EP26

The EP Evaluator 2IC module includes an optional mode intended to assist the user in verifying the equivalency of new candidate lot numbers of reagents vs a comparative or reference lot number of the same test kit (same vendor part number). The optional mode follows the CLSI EP26 A Guideline: *User Evaluation of Between-Reagent Lot Variation; Approved Guideline (2014)*.

The CLSI EP26A Guideline incorporates a process to evaluate whether a new reagent lot number performs acceptably similar to a comparative lot number. EP Evaluator can use the three tables from Appendix A of the Guideline to derive a minimum number of specimens and rejection criteria for up to three medical decision point concntrations based on a specified critical difference (CD aka TEA), a requested statistical power, typical laboratory within run precision SD (Sr) and typical within reagent lot precision SD (Swrl) for each MDP concentration.

The tables were designed to balance the sensitivity and specificity for evaluation of lot-to-lot difference, so that there is no more than a 5% probability of detecting a difference when none exists. This is a type I error (false positive) where alpha (a) is defined to be <= 0.05. When there are two or three MDP levels being tested, the tables A2 and A3 distribute the alpha value of 0.05 (5%) equally among the number of levels. So for 2 levels, each level is assigned a probability of 2.5% that a difference is found when none exists. For 3 levels, each level is assigned a probability of 1.67%.

The statistical power that is achievable represents the percent probability that you can detect a difference as large as the CD with the number of samples required. This is a type II error (false negative) and the statistical power represents 1 - Beta, where Beta is the probability of NOT detecting the difference. The rejection criteria for pass/fail must use a percentage of the total CD in order to achieve the required statistical power with a small number of samples, and will vary from 55% to 90% of the CD. The smaller the percentage of the CD, the higher the statistical power will be. The tables provide statistical powers from 59% to 99%.

Parameters and Calculations

The following inputs are provided by the user:

- MDP: Medical Decision Points - defined concentrations of the measurand at which clinical decisions are often made. The EP26 mode requires at least one and up to three.

EP Evaluator®

Product Strategy -- Data Innovations

Two Instrument Comparison
Report Interpretation Guide

- Prox Lo, Prox High: the lowest and highest sample values to include in a single MDP target range.

- CD: Critical Difference: For the purposes of EP26 mode of the 2IC module, the CD For every included specimen is calculated from the TEa entry in the 2IC tab of the parameters screen.
Be aware that some sources of TEa may provide values that are inordinately large for this purpose as the values were proposed many years ago, when analytical technology was less robust. The most useful evaluation happens when the ratio of CD/Swrl is greater than 2.5.

- Swrl: Observed or historical within reagent lot SD at the MDP concentration.

- Sr: Within run SD at the MDP concentration.

- Statistical power (expressed as a percent in EE): represents the probability (95%) that the observed absolute mean difference between reagent lots will exceed the rejection limit when the true absolute difference is equal to the CD.

The program computes the following intermediate values for each MDP.

- CD (TEa) at the MDP concentration

- CD/Swrl ratio: As the CD/Swrl ratio increases, the performance of the method improves. When the table value of CD/Swrl is at the maximum value of the table, the number of data points needed to verify equivalency are at the minimum of one, and the best statistical power (in some cases greater than 95% can be achieved.

- Sr/Swrl ratio: the allowed ratios range from 0.3 to 1.0. When the ratio of within run precision compared to the historical or observed within reagent lot precision is near 1.0, fewer specimens are needed to achieve the desired statistical power. When the ratio is closer to 0.3, it means that the method exhibits more run to run or between day variability. Calibration variability, instrument maintenance or drift could be among the causative factors

The program can automatically use the lookup tables to create the following outputs for each MDP.

- **N:** Number of specimens required for each MDP at the requested statistical power. Note that there must be at least 3 total specimens across all of the MDPs.

- **% rejection limit (k):** The percentage factor selected from the tables for a particular set of CD/Swrl and SR/Swrl in order to achieve a specified power. The %

rejection limit multiplied by the CD yields the MDP rejection criterion (RejLimit), which is used to determine Pass/Fail

- Note: When values of CD/Swrl or Sr/Swrl fall between rows in the table, the row that provides the next lowest ratio is chosen, in order to ensure that the actual power achieved is greater than the value in the table.

User defined sample numbers and % rejection criteria. When the EP26 parameters screen box "Perform EP26 table calcs automatically" is unchecked, the user can define the number of specimens and the %rejection limit. Viewing the tables can be used as a tool to help balance the desired statistical power by reviewing the different %rejection limit factor choices.

Included data points: Data points are excluded from pass/fail calculations if they are not within the target range of any MDP level. Target ranges are implemented in EP Evaluator in order to assist the user in gathering specimens that are reasonably near the MDP concentrations. Excluded points appear in the graphs and in the results listing for information purposes.

Pass / Fail: Each MDP level will pass if the absolute mean difference is less than the rejection criteria. The entire experiment will pass if all MDPs pass.

Failure of the table to provide useful information. The process to balance a minimal number of specimens with a reasonably high statistical power to detect a true difference between reagent lots is dependent on the characteristics of the methods. For the most part, well performing methods will have a CD/Swrl ratio of 3 or higher and a Sr/Swrl ratio that is close to (but not greater than 1.0). It is usually possible to find a combination of a low number of samples and a power of 80% or higher in the tables. There are circumstances where the tables will fail to provide the desired information. Either the number of samples becomes too large to be practical or the statistical power cannot be achieved with the CD/Swrl ratio calculated from the user's inputs.In the EE tables, the largest number of samples that can be returned is 401 and values that are disallowed are indicated by a sample size of 1000. This usually happens when the imprecision of the test method is too close to the desired critical difference, or the Sr/ Swrl ratio is not close enough to 1.0.

In the laboratory, the typical imprecision characteristics are fairly well established, and while the user could adjust them slightly, the CLSI guideline suggests that either the desired power be reduced or the critical difference be increased depending on the impact to the clinical assessment at the MDP concentrations chosen. Section 9 of the CLSI EP26A guideline discusses some examples that illustrate these points.

EP Evaluator®

Product Strategy -- Data Innovations

Two Instrument Comparison
Report Interpretation Guide

EP26 mode Data Evalution

Experiment Detail Screen

- the "Error index" graph is a plot of (Y-X)/ (RejLimit) vs the observed X value.

The EP26 tab is present and provides a table of the experimental results grouped into the target ranges (proximity limits) for each defined MDP value. Each individual specimen displays its specID, current lot result (X), candidate lot result (Y), and the difference Y - X.

Calculations for each MDP group are displayed in bold at the bottom of its group and include: MDP concentration, the proximity range, the group mean difference, N (the number of included specimens in the group / the number of specimens specified), the CD calculated at the MDP value, the power achieved if provided by the tables, the rejection limit at the MDP conc, and Pass or Fail.

Each experiment reveals whether the experiment passed or Failed for all MDP values, and whether or not the tables were used.

Excluded specimens are listed at the end of the report

Report Summary

The report summary on the Overview screen grid and the printed report summary page will display "EP26" in the "Coverage" column and the "% Unacceptable column" will reveal the number and % of the MDP levels that fail.

Printed report: page one

The Y vs X scatter plot is not different than the classic 2IC mode. The yellow boundary area includes all data pairs within the defined TEA.

The error index chart plots [(Y-X)/%rejection limit for each point] versus the comparative X method value. The error index is measured for each X Y pair and calculated based on the rejection limit for the X value. Points in the shaded area have an error index greater than 1.00 or less than -1.00. Because the pass-fail criteria is based on the absolute mean difference for points included in an MDP target range, it is possible for an individual point to fail and yet the MDP level will still pass.

Regression statistics are not displayed.

Key statistics include:

Average error index: error index is calculated as Y-X / %rejectionlimit * CD

Error index range: from -1.0 to 1.0

Evaluation of Results:

When table values are used for all MDP levels and all MDP levels pass, the report will state "All specimens were within the target ranges for all MDPs and the reagent is verified for reporting of patient results."

Evaluation Criteria:

TEa: As entered in the regular 2IC parameters tab.

Reportable range: The low and high reportable range limits appear as dotted lines in the scatter plot and error index plot.

Printed report: page 2

A table of results includes 2 sections:

Section 1 displays each MDP value, Prox (the target range for each MDP), Swrl, Sr, Power requested, CD, CD/Swrl, N specimens included, the selected %rejection criteria.

Section 2 organizes the specimens in their defined MDP groups with content identical to the EP26 Results Table on the Details screen.

The last two rows indicate whether the experiment as a whole passes or fails, and whether the tables were used.

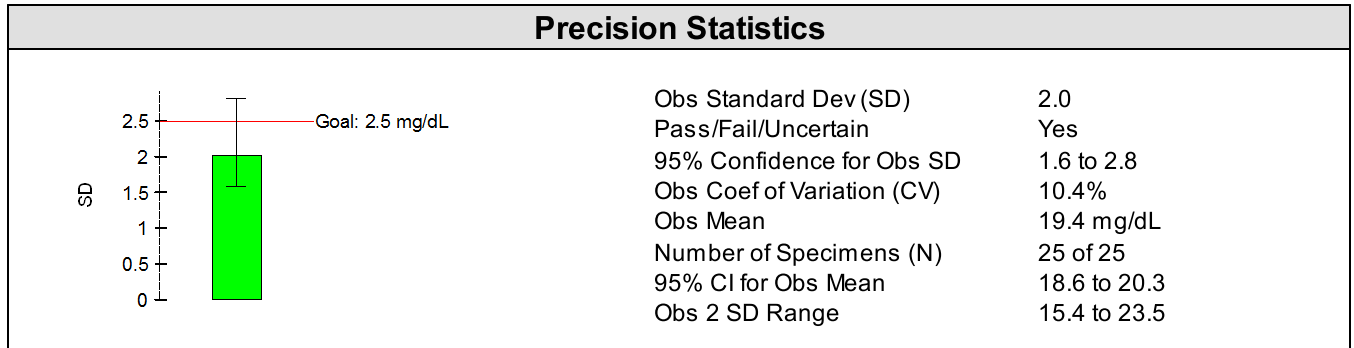Excluded specimens are listed at the end of the report.

Evaluation of reagent lot equivalency and Pass/Fail:

- If table values are used for all MDPs, the experiment passes if the absolute mean difference for all MDPs is less than their respective rejection criteria. And the statement that "the candidate reagent is verified for reporting of patient results appears on the report."

- If user defined values are used for any MDP for the number of samples or the % Rejection criteria, it is still possible for the experiment to pass, but the statement that "the candidate reagent is verified for reporting of patient results" does not appear.

- Preliminary: an experiment is marked Preliminary and Fail if there are fewer than 3 total included specimens.
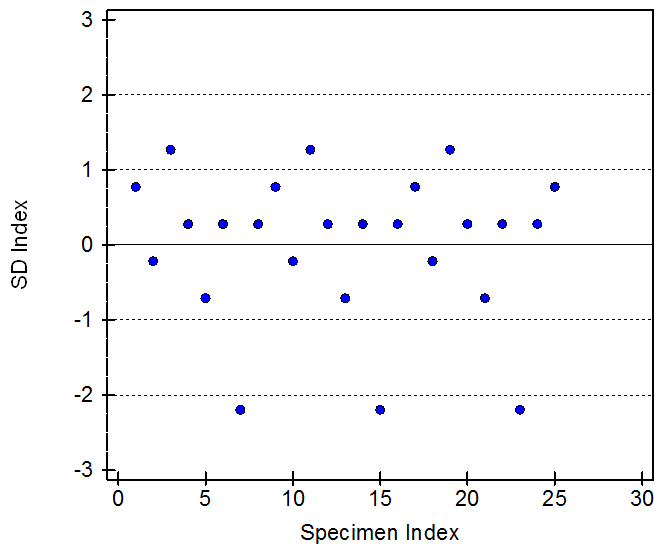
References

CLSI EP26 A Guideline: User Evaluation of Between-Reagent Lot Variation; Approved Guideline (2014).
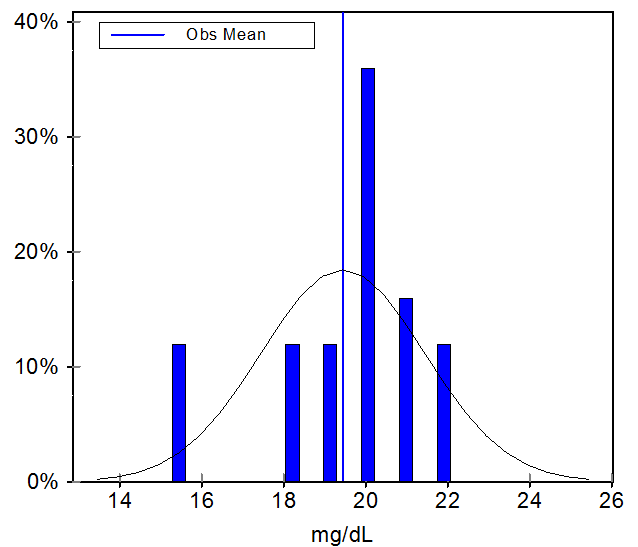
Simple Precision

Precision Statistics

Obs Standard Dev (SD)	2.0
Pass/Fail/Uncertain	Yes
95% Confidence for Obs SD	1.6 to 2.8
Obs Coef of Variation (CV)	10.4%
Obs Mean	19.4 mg/dL
Number of Specimens (N)	25 of 25
95% CI for Obs Mean	18.6 to 20.3
Obs 2 SD Range	15.4 to 23.5

Precision Plot

Histogram

Supporting Data

Analyst	Michael Doe
Expt Date	01 Jun 2000
Units	mg/dL
Target Mean	--
Target CV	--
Control Lot	Krogers 123 exp 31 Jan 2024
Reag Lot	Krogers 456 exp 31 Jan 2024
Cal Lot	CVS 789 exp 31 Jan 2024
Comment	Comment test

User's Specifications

Precision Verification Goal	TEa
Allowable Total Error	10.0 mg/dL (conc)
Random Error Budget	25%
Allowable Random Error	2.5 mg/dL (conc)

Accepted by: _____

_____ _____
Signature   Date

Simple Precision

Precision Data

Index	Results	Index	Results	Index	Results	Index	Results
1	21	8	20	15	15	22	20
2	19	9	21	16	20	23	15
3	22	10	19	17	21	24	20
4	20	11	22	18	19	25	21
5	18	12	20	19	22		
6	20	13	18	20	20		
7	15	14	20	21	18		

X: excluded from calculations

EP Evaluator ®

Product Strategy -- Data Innovations

Simple Precision
Report Interpretation Guide

CLIA and CAP require periodic verification of Precision. The Simple Precision experiment is a quick way to fulfill the requirement for repeatability using a "within-run" type of experiment. The procedure is to enter repeated measurements of a single sample on a single method, and compute simple statistics like mean, SD, and %CV. In addition, EP Evaluator Release 11.2 offers the following options:

1: Entry of one of three Precision Verification Goal Modes:

- TEA Based Allowable Random Error: entered as a percentage of the specified TEA.

- Vendor SD based: Entered as one within run SD at a specified concentration level.

- Vendor REa based: Entered as one within run SD or within run Percent or both.

When using RRE data entry techniques, the Precision Verification Goal Mode is defined in the Modules and Options form in Policies. When any type of goal is entered, the program will report that the test "passes" if the calculated SD does not exceed the goal. An option in File Preferences further enables the Pass/Fail criteria to include a designation for uncertain if the precision goal is within the 95% confidence limits around the Observed (OBS) SD.

2: Entry of a Target Mean value . Investigators often wish to compare the recovered mean with a target.

3: Show Histogram Result Distribution. If the "Show Histogram" option is enabled, then a **Histogram** displays a frequency distribution of recovered results centered around the observed mean. If "Show Target Range" is also enabled, then entry of a target mean is required and a user selectable SD target limit of 2 SD, 2.5 SD, 3 SD, or 3.5 SD will be drawn around the target mean.

Definitions

Precision: Ability to obtain the same result upon repeated measurement of a specimen - not necessarily the *correct* result, just the *same* result. Ability to get the correct result is *Accuracy.* Within-run precision is also known as *Repeatability,* which is the repeated measurement of a specimen under the same set of conditions.

Mean: Average value, computed by taking the sum of the results and dividing by the number of results.

Bias: The Target mean minus the observed mean.

%Bias: 100 times the Bias divided by the target mean.

Standard Deviation (SD): The primary measure of dispersion or variation of the individual results about the mean. In other words, a measure of Precision. For many analytes, SD varies with sample concentration. Example: For Glucose, an SD of 10 for a 400 mg/dL sample represents very good precision. For a 40 mg/dL sample, an SD of 10 represents very poor precision. The Simple Precision calculation for one SD is the standard calculation for sample SD - same as used in a pocket calculator.

Coefficient of Variation (CV): SD expressed as a percent of the mean. %CV = 100 times SD/mean. For analytes where error varies with concentration, CV is more likely to remain constant across the analytical range. Example: 2.5% CV for Glucose is good precision at any concentration. Target CV is calculated as 100 times the target SD / Observed mean.

Within-Run and Total SD: Manufacturers often publish these statistics in package inserts. Typically, they are computed from a more complex precision experiment that requires replicate measurements over several days. Total SD and within run SD from the complex precision experiment use Analysis of Variance (ANOVA) calculation techniques and are NOT calculated in the same way as EP Evaluator's Simple Precision SD.

Number of Specimens (N): A good precision study should include 20-50 replicates.

95% Confidence Intervals: These are calculated for both the observed mean and the observed SD. In a sense, this measures the "precision" of the mean or SD. It shows how much the observed mean or SD might vary if the precision experiment was repeated. The two-tailed 95% confidence interval (CI) is calculated from standard Chi-square tables. Its width depends on N and the observed mean or SD. The reliability of the precision estimate of the mean or SD improves as N increases.

Mathematically, it is possible to calculate SD and CV from only 3 replicates, but it is not a very reliable estimate. For example, suppose the true SD is 1.00. An estimate based on 3 replicates might easily be as low as 0.52 or as high as 6.29 (95% confidence). At N=20, the 95% confidence interval (CI) is much narrower: 0.76 to1.46. At N=50 it narrows further: 0.84 to 1.24.

Precision Verification Goal Modes:

Allowable Total Error (TEa), and the Error Budget

TEa states the laboratory's policy for how much error is medically (or administratively) acceptable. Regulatory requirements represent an upper limit of that specification. Example: the CLIA limit for Sodium is 4 mmol/L. The experimental Total Error has two major components: Systematic Error (synonym: Bias) and Random Error (synonym: Imprecision). The Error Budget for Total *allowable*

Simple Precision
Report Interpretation Guide

error allocates a fraction of the *Allowable* Total Error for Systematic Error (SEa), and the remaining fraction for Random Error (REa). Establishing an appropriate Error Budget allows the lab to evaluate and control accuracy and precision separately. Recommended ranges of values are 25-50% for the Systematic Error Budget and 16-25% for the Random Error Budget at one SD. The SP module uses the user entered TEa multiplied by the Random Error budget % to calculate the allowable random error (REa) specification at the observed mean.

Vendor Based Precision Goal

A vendor based precision goal allows the user to compare observed statistics with vendor recommendations for acceptable precision performance. Vendor performance goals can typically be found in the manufacturer's insert sheets for one or more levels. When using RRE data entry techniques, the vendor goals are extracted from user defined entries in the SP tab of the Analyte Settings form in Policies. Two types of vendor based precision goals are available for selection.
1. The SD based format requires entry of a vendor within-run precision SD. The concentration level at the vendor SD goal is optional and is not used in calculations. This is the original Vendor format used in EE and many manufacturers have discontinued this format for their claims, in favor of the REa based format.
2. The REa based format is analogous to the REa portion of the TEa. It requires entry of either a concentration component or a Percent component representing 1 SD, or both. When using RRE data entry techniques, the REa based format is defined for each class in Modules and Options, and the vendor goals are extracted from user defined entries in the SP tab of the Analyte Settings form.

Components of the Printed report

The components of the Simple Precision Report are the Summary page (when more than one experiment is selected for the report), Precision Statistics Table, Precision Plot, Histogram, Supporting Data Table, User's Specifications, and Precision Data Table.

Summary page:

The summary page displays the following experiment summary statistics:

- **"N of N":** non-excluded data points / total N Observations

- **Mean:** Obs mean / target mean. If optional Target mean is not entered, "Obs mean / -" is displayed unless "-" is suppressed in the File\Preferences\Reports.

- **SD:** Obs SD / target SD goal at the level of the Observed mean. If optional SD goal is not entered, "Obs SD / -" is displayed unless "-" is suppressed in File\Preferences\Reports.

- **%CV:** Obs %CV / target CV% at the level of the Observed mean. If optional SD goal is not entered, "%CV / -" is displayed unless "-" is suppressed in File\Preferences\Reports.

It is possible to change the Report Summary sort order from analyte name to sample name by checking the option in Preferences/Reports.

Charts and Tables Interpretation:

Precision Plot: A Levey-Jennings type chart of Standard Deviation Index (SDI) on the Y-axis vs. specimen index on the X-axis. Specimen index reflects the order in which the results were typed into the program. SDI = (Result - Obs mean) / SD, unless target mean is defined. If target mean is defined, then SDI = (result - target mean)/ target SD, and zero SDI would represent the target mean.

Precision Statistics Table: This table may include the following:

Calculated Statistics - always present:

- Obs Standard Dev (SD)

- 95% confidence for Obs SD

- Obs. Coefficient of Variation (CV)

- Observed Mean

- N

- 95% CI for Obs Mean

- Obs 2 SD range

Precision Goal: This chart appears in the Precision Statistics Table only when a Precision Verification Goal is provided. It shows the Observed SD and its 95% confidence interval in relation to the allowable random error goal. The wide colored bar represents the SD. The fence marks above and below the top of the SD bar represent the 95% confidence limits.

Pass/Fail/Uncertain?

A line for Pass/Fail/Uncertain appears when a Precision Verification Goal has been entered.
Yes means the experiment passes.
No means the experiment fails.
Uncertain means that more evaluation may be needed.

The objective is to determine whether precision is acceptable. The default determination that the test passes is

if the computed SD does not exceed the goal. When the designation Pass/Fail/Uncertain is enabled in Preferences, the test will pass only if the computed upper 95% CI around the computed SD does not exceed the goal.

Example: Suppose Allowable Total Error for Sodium is 4 mmol/L, with a Random Error Budget of 25%. Allowable Random Error is 25% x 4 or 1 mmol/L. The precision test passes when the computed SD is at or below 1 mmol/L.

- The report calls the test a "pass" if the SD (top of the wide colored bar) is less than or equal to the goal.

- The bar is green if the test passes, or red if it fails.

- If you consider all points within the 95% confidence interval to be statistically identical, you might consider the test at least marginally acceptable if the lower confidence limit does not exceed the goal. When enabled in Preferences, the designation **Uncertain** appears if the precision goal is within the 95% confidence limits around the observed SD. In this case, the colored bar is yellow.

- The ideal situation is when the upper 95% confidence limit does not exceed the goal.

Histogram: The Histogram chart appears when enabled in the parameters screen. Observed data point results are clustered around the Observed mean. If a target mean has been entered, optional SD target limits selectable from 1 SD to 3 SD are drawn around the target mean. Bias and % bias are provided in the experiment detail screen, but not shown on the report.

Supporting Data Table: This table always includes the analyst name, experiment date, and the analyte units. Optional data entries include Target mean, target range for the histogram, a Comment, and Reagent, Control or calibrator lot numbers.

User's Specifications: These only appear when a Precision Verification Goal Mode (PVGM) is selected and will include the type of PVGM selected, and the corresponding entries as described in the above section on PVGM.

Precision Data Table: This table includes the experiment's observed results and their index (order of entry). Preliminary Report

The word PRELIMINARY printed diagonally across the report indicates that the data is incomplete, and the report is not acceptable as a final report. Some or all of the statistics may be missing. The Simple Precision report is Preliminary if there are fewer than 3 unexcluded results.