EP Evaluator®

-- Labonovum B.V.

Alternate (Quantitative) Method Comparison

X Method Venous EDTA 2u   **Y Method** HemCol EDTA 120u

Scatter Plot — Deming Regr, 1:1 Line; HemCol EDTA 120u (10^9/L) vs Venous EDTA 2u (10^9/L)

Bias — Mean Bias; Bias (10^9/L) vs Venous EDTA 2u (10^9/L)

Percent Bias — Percent Bias vs Venous EDTA 2u (10^9/L)

Regression Analysis

	Deming	Regular
Slope	0,889 (0,750 to 1,028)	0,828 (0,691 to 0,965)
Intercept	0,085 (-0,009 to 0,179)	0,124 (0,032 to 0,216)
Std Err Est	0,078	0,077

95% Confidence Intervals are shown in parentheses

Supporting Statistics

Corr Coef (R)	0,9225	SubRange Bounds	None
Bias	0,014 (2,213 %)	Points (Plotted/Total)	29/30
X Mean ± SD	0,639 ± 0,219	Outliers	None
Y Mean ± SD	0,653 ± 0,197	Scatter Plot Bounds	None
Std Dev Diffs	0,085		

Experiment Description

	X Method	**Y Method**
Expt Date	14 feb 2018	14 feb 2018
Rep SD	1	1
Result Ranges	0,28 to 1,26	0,32 to 1,15
Units	10^9/L	10^9/L
Reagent	--	--
Calibrators	--	--
Analyst	R. Huisman	R. Huisman
Comment		

Accepted by: _____   _____

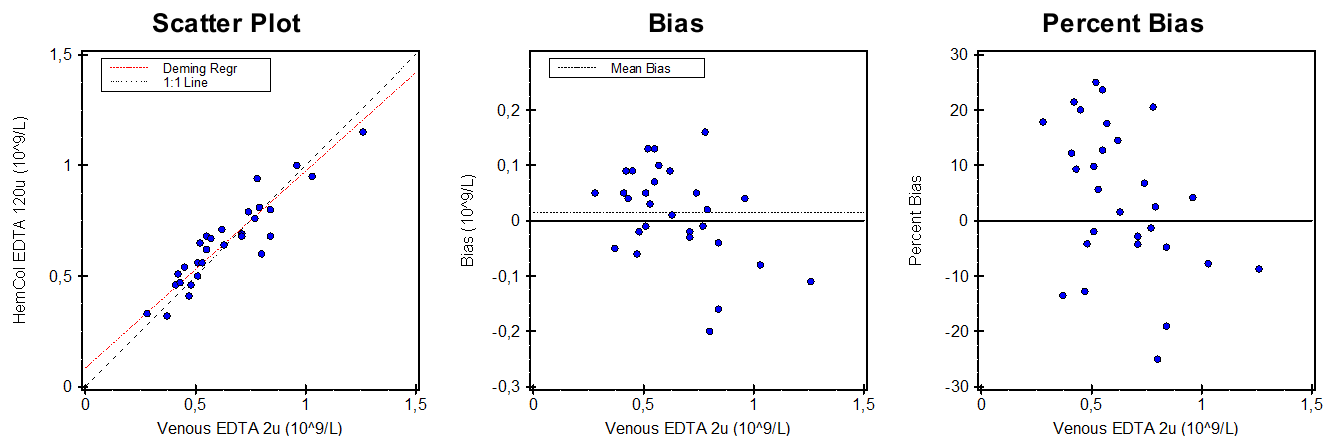Signature   Date

Alternate (Quantitative) Method Comparison

X Method Venous EDTA 2u   **Y Method HemCol EDTA 120u**

Experimental Results

Specimen	X	Y	Bias	Specimen	X	Y	Bias	Specimen		X	Y	Bias
S00001	0,71	0,69	-0,02	S00011	0,96	1,00	0,04	S00021		0,71	0,68	-0,03
S00002	0,51	0,50	-0,01	S00012	0,41	0,46	0,05	S00022		1,26	1,15	-0,11
S00003	0,77	0,76	-0,01	S00013	0,55	0,68	0,13	S00023		0,78	0,94	0,16
S00004	0,37	0,32	-0,05	S00014	0,74	0,79	0,05	S00024		0,43	0,47	0,04
S00005	0,28	0,33	0,05	S00015	0,79	0,81	0,02	*S00025*	*X*	*2,54*	*2,62*	
S00006	0,42	0,51	0,09	S00016	0,55	0,62	0,07	S00026		0,84	0,68	-0,16
S00007	0,52	0,65	0,13	S00017	0,57	0,67	0,10	S00027		0,84	0,80	-0,04
S00008	1,03	0,95	-0,08	S00018	0,51	0,56	0,05	S00028		0,62	0,71	0,09
S00009	0,53	0,56	0,03	S00019	0,45	0,54	0,09	S00029		0,48	0,46	-0,02
S00010	0,63	0,64	0,01	S00020	0,8	0,60	-0,20	S00030		0,47	0,41	-0,06

Values with an "X" were excluded from the calculations. Outliers "O" were also excluded.

Alternate (Quantitative) Method Comparison

X Method Venous EDTA 2u   **Y Method HemCol EDTA 120u**

Scatter Plot

Bias

Percent Bias
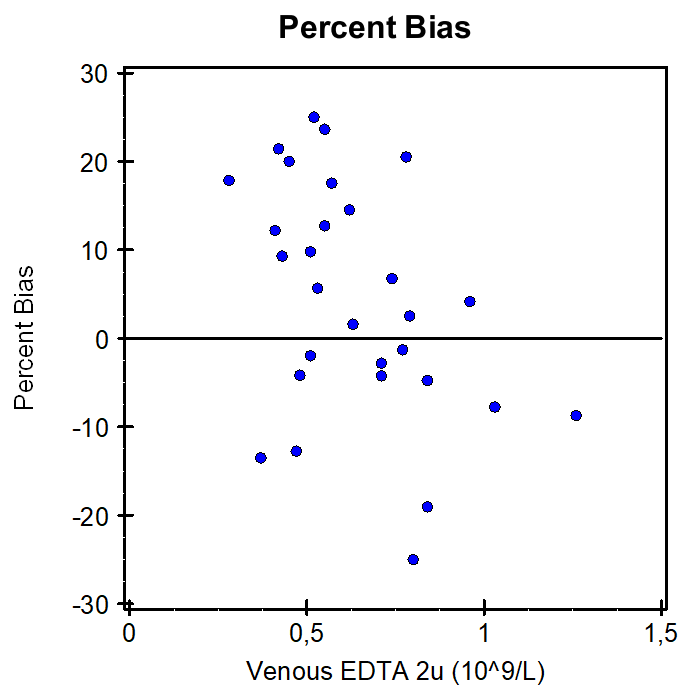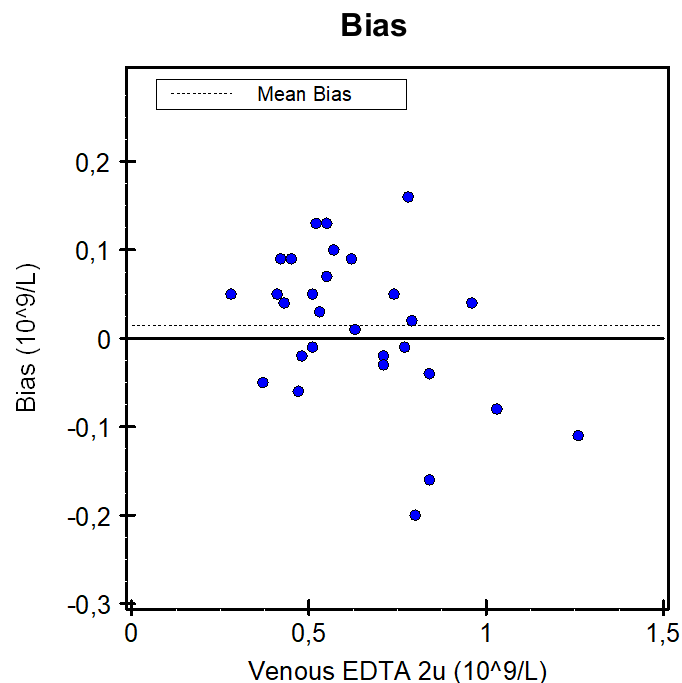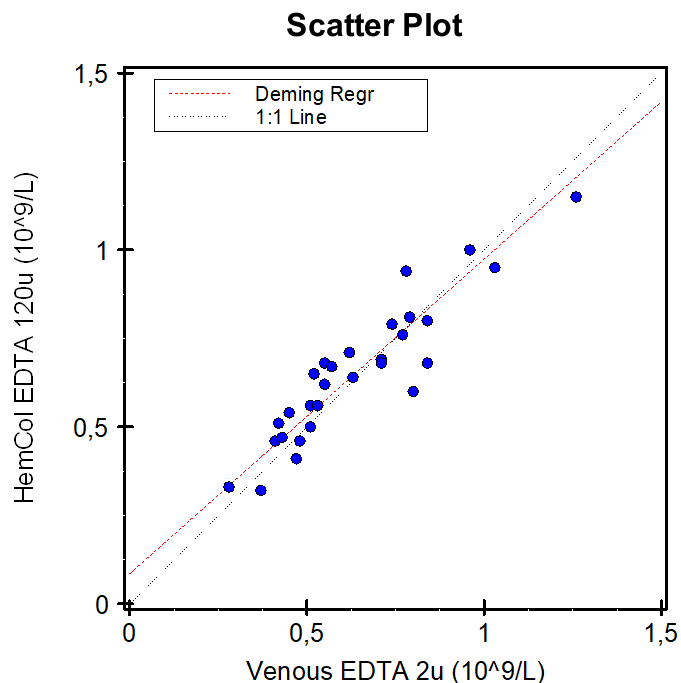
This page contains a larger, working copy of the same graphs that appear on page 1.

EP Evaluator ®

-- Labonovum B.V.

Alternate (Quantitative) Method Comparison
Report Interpretation Guide

There are many reasons for doing method comparison studies. Perhaps the most common:

- To determine the relationship of the Medical Decision Points (MDPs) of an old method with those of a new method. In other words, "Can I continue to use the same MDPs with the new method?"
- To validate a new method being brought into the lab, by demonstrating the statistical relationship to the method currently in use.

The statistical tool used is linear regression. The methods can be considered statistically identical if:

- The slope is 1.00 (within 95% confidence)
- The intercept is 0.00 (within 95% confidence)
- The predicted Y MDPs are equal to the X MDPs (within 95% confidence)

Not all regressions are method comparisons. This Report Interpretation assumes that X and Y are alternative methods for measuring the same quantity, and that the purpose of the experiment is to determine whether X is statistically identical to Y. If the purpose is to predict weight as a function of height, or to predict APTT levels from Heparin levels, some of the interpretive comments may not apply.

Regression Approaches

The report shows at least two, and (optionally) three sets of regression coefficients.

Regular Regression: This is the ordinary least squares regression line commonly provided in spreadsheets and general statistical software. It is shown only to provide a familiar frame of reference; it is not used to estimate Medical Decision Points. The problem with using regular regression to compare methods is that it assumes the X method is measured with no random error -- not very likely for clinical laboratory results. Regular regression often underestimates the true slope, sometimes by a significant amount.

Deming Regression: This approach assumes that both the X and Y methods are subject to measurement error. In theory, a **Representative SD** (precision estimate) is input for each method. In practice, only the ratio of the two precisions affects the calculation. If exact precisions are unknown, entering 1.0 for both Representative SDs says "these methods have about the same precision", and gives reasonable results in most cases.

Several studies have shown that Deming Regression is the best approach to use when the two methods are expected to be identical, and the data is well-distributed and free of outliers. It can, however, be seriously affected by outliers. EP Evaluator provides the option to automatically exclude extreme outliers, or the user can exclude them manually.

The default Regression Lines on the AMC graphs are Deming Regression Lines. A Passing Bablok regression line can be enabled in Preferences\calculations. When MDPs are estimated by linear regression, Deming linear regression is used.

Passing-Bablok Regression: Passing-Bablok regression is a non-parametric regression technique developed specifically to be resistant to outliers. Non-parametric statistical calculations make no assumptions about the shape of the population distribution, while parametric approaches assume that the population has a Gaussian distribution.

Main strengths: There is no need to exclude perceived outliers, either manually or automatically. Like Deming, it does not assume that X is free from error. Comparative studies show that it performs about as well as Deming Regression in most cases, and better than Deming when outliers are present.

Main weaknesses: Passing-Bablok is computationally intensive, particularly for large N, and it may be unreliable for very small N. EP Evaluator does not show Passing-Bablok statistics when N<10 or N>500.

Removing Outliers

An outlier is a point so far from the others as to arouse suspicion that it was generated by a different mechanism. Some common causes: typing a number with the decimal point in the wrong place, analyzing the wrong sample, or entering incorrect specimen identification. The best way to deal with an outlier is to (manually) determine its cause and correct it. Another option is to use a statistical procedure to remove outliers automatically.

EP Evaluator uses a somewhat complex iterative algorithm to identify outliers. The goal is to eliminate points whose distance from the regression line exceeds 10 times the Standard Error of Estimate (SEE), where SEE is computed not from the full data set, but from the data set with outliers excluded. When outliers are included, the SEE is over-stated, and the regression coefficients are suspect.

An outlier is, by definition, a rare occurrence. If more than 5% of the points are excluded, either by the mathematical algorithm or the user, the report is stamped PRELIMINARY. It is then incumbent on the user to re-evaluate both methods to determine their suitability.

Interpreting your Results

When interpreting a method comparison report, there are two areas which must be addressed:

EP Evaluator ®
-- Labonovum B.V.

Alternate (Quantitative) Method Comparison
Report Interpretation Guide

- First, is the QUALITY OF THE DATA adequate to accurately draw conclusions?
- Second, what conclusions can be drawn from those data?

These issues MUST be addressed in this order. If the data quality is not adequate, then any additional conclusions drawn from those data may well be wrong.

Data Quality Statistics

The most important elements of a good method comparison study are a reasonable N (number of x-y pairs) and a good distribution of results. Generally a good experiment will include 30 to 50 specimens with their results distributed more or less evenly across the method's reportable range.

Results Range: The minimum and maximum values of X and Y. It is inappropriate to draw conclusions outside the range of data studied. When evaluating MDPs, it is important to include data points that cover the full range of MDPs.

Result Range Analysis: This (optional) table shows how the X values are distributed within the range. A relatively even distribution is desirable. If 99% of the values are at the low end and 1% are at the high end, with none in the middle, the regression slope is almost totally determined by the handful of high points.

Points (Plotted/Total): More commonly called N, the number of x-y pairs in the regression. "Plotted" is the number on which calculations are based. The difference between Plotted and Total is points that were excluded, either manually or by the automatic outlier removal procedure. CLSI considers N=40 to be the minimum for a good method comparison study. Increasing N improves the quality up to a point, but a good distribution of data is much more important than a large N.

Correlation Coefficient (R): R generally corresponds to the width of an ellipse drawn around the data. The narrower the ellipse relative to its length, the higher R will be. If there is large amount of error, the width of the ellipse will increase which will in turn cause a lower R.

R ranges from -1 to 1. Zero means there is absolutely no relationship. +1 or -1 means there is a perfect relationship, and a very high-quality regression. An R of 1.000 could be achieved just as easily with a slope and intercept of 1.000 and 0.0 as with a slope and intercept of 0.5 and 400 respectively. In other words, *it specifies the degree of correlation, not the degree to which the two methods match.*
.

In a method comparison setting, R has special significance:

- A small R may be a sign that the Results Range is inadequate. Adding samples to increase the range of X will improve both the R value, and the quality of the study.
- If R is less than a user-selectable cutoff value (0.90, 0.95, or 0.975), regression is not used to evaluate Medical Decision Points. Instead, they are evaluated by the method of Partitioned Biases, as explained in more detail, below. The coefficient of determination r-squared (RSQ) is also provided in parenthesis when enabled in the Preferences\Reports menu.

Interpreting the Regression Statistics

Assuming that the quality of the data is adequate, you may proceed to interpreting the results.

Slope, Intercept, and their Confidence Intervals: When two methods are statistically identical, the 95% confidence interval for the slope includes 1.00, and the 95% confidence interval for the intercept includes 0.0.

Example: If the 95% CI for the slope is 0.92 to 1.02, 1.00 is included in the interval. However, if the 95% CI is 0.82 to 0.92, 1.00 is not included in the interval.

If the experiment were repeated with different data, the slope and intercept would be a bit different. But 95% of such experiments are expected to fall within the confidence interval.

Medical Decision Point Analysis: A Medical Decision Point is an analyte concentration at which medical decisions change. If the concentration is to one side of the MDP, one decision is made; if on the other side of the MDP, a different decision is made. For example, Fasting Plasma Glucose above 126 mg/dL (7 mmol/L) indicates hyperglycemia which, if confirmed, establishes a diagnosis of diabetes. For obvious reasons, it is particularly important that the two methods agree at the MDPs.

When the two methods are statistically identical, the 95% Confidence Interval for each Y MDP includes the corresponding X MDP.

Standard Error of Estimate (SEE): measures the spread of the x-y data around the linear regression line. If both methods have the same constant precision SD across the full analytical range, SEE should be about 1.4 times the precision SD.

Bias, and its Relationship with Regression

Bias is the difference Y-X. The **Bias Plot** is a scatter plot with X on the x-axis, and Y-X on the y-axis. The ideal bias plot would have all points falling exactly on the zero line.

Alternate (Quantitative) Method Comparison
Report Interpretation Guide

That is unlikely to occur in practice, because both X and Y are measured with some random error. A good bias plot is centered on the zero-line, and forms an envelope of approximately constant width about it.

Constant Bias is present when Y is consistently greater than (or less than) X by a constant amount. The bias plot forms a constant-width envelope around the average bias line instead of the zero line. The regression intercept measures constant bias. In fact, if the slope is exactly 1.000, the regression intercept is equal to the average bias.

Proportional Bias is present when Y differs from X in a way that is proportional to X. For example, Y may be consistently 5% higher than X instead of 5 units higher. On the Bias plot, the points center around an upward or downward-sloping line instead of a horizontal line. The regression slope is a measure of proportional bias.

The **Method of Partitioned Biases** comes into play when R is below the cutoff value (0.90, 0.95, or 0.975). In this situation, the Bias Plot is divided into three segments, with the same number of points in each segment. It is assumed that bias is approximately constant within each segment. This segmented structure provides an estimate of bias and its 95% confidence interval at the Medical Decision Points. The method of Partitioned Bias is an alternative approach used to define medical decision points. If it is used then slope, etc. are not used.

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. Causes:

- Less than 3 unexcluded x-y pairs.
- More than 5% of points are outliers.
- Excluding outliers reduced the range of X by more than 50%. The range of X is a significant aspect of data quality, and it should be confirmed by the analyst in this case, rather than by a mathematical algorithm.