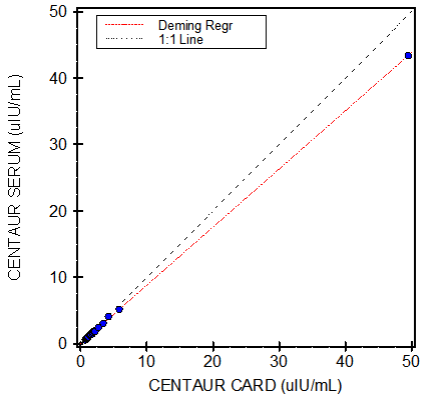
Prepared for: Chemistry -- New Century Diagnostics

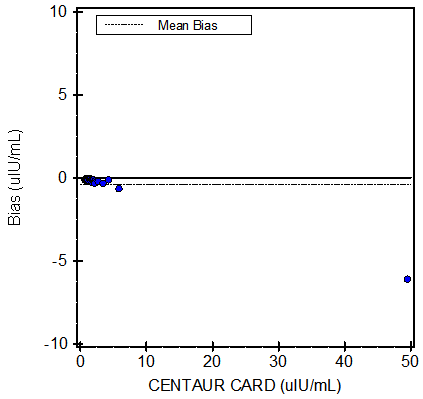
Alternate (Quantitative) Method Comparison

X Method: CENTAUR CARD **Y Method: CENTAUR SERUM**

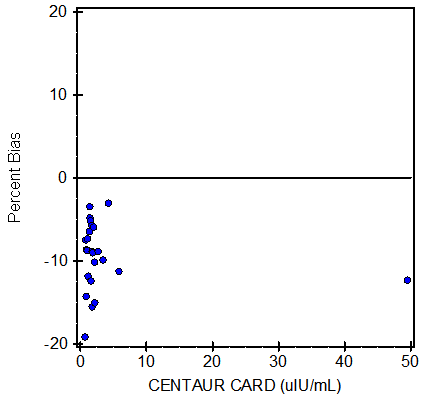
Scatter Plot



Bias



Percent Bias



Regression Analysis

	Deming	Regular
Slope:	0.876 (0.872 to 0.880)	0.876 (0.872 to 0.880)
Intercept:	0.066 (0.025 to 0.106)	0.066 (0.025 to 0.106)
Std Err Est:	0.091	0.091

95% Confidence Intervals are shown in parentheses

Supporting Statistics

Corr Coef (R):	0.9999	SubRange Bounds:	None
Bias:	-0.406	Points (Plotted/Total):	25/25
X Mean ± SD:	3.796 ± 9.589	Outliers:	Not Tested
Y Mean ± SD:	3.390 ± 8.397	Scatter Plot Bounds:	None
Std Dev Diffs:	1.195		

Experiment Description

	X Method	Y Method
Expt Date:	18 Jun 2011	18 Jun 2011
Rep SD:	1	1
Result Ranges:	0.73 to 49.49	0.59 to 43.38
Units:	uIU/mL	uIU/mL
Analyst:	CC	CC
Comment:		

Accepted by: _____ _____

Signature Date

Prepared for: Chemistry -- New Century Diagnostics

Alternate (Quantitative) Method Comparison

X Method: CENTAUR CARD **Y Method: CENTAUR SERUM**

Experimental Results

	Spec ID	Results X	Results Y	Bias	Calc'd Y	SEE Factor
1	S00001	1.53	1.45	-0.08	1.406	0.5
2	S00002	1.88	1.71	-0.17	1.712	0.0
3	S00003	1.45	1.38	-0.07	1.336	0.5
4	S00004	1.61	1.41	-0.20	1.476	-0.7
5	S00005	1.54	1.46	-0.08	1.414	0.5
6	S00006	49.49	43.38	-6.11	43.405	-0.3
7	S00007	1.43	1.38	-0.05	1.318	0.7
8	S00008	3.43	3.09	-0.34	3.069	0.2
9	S00009	4.24	4.11	-0.13	3.779	3.6
10	S00010	2.19	1.86	-0.33	1.984	-1.4
11	S00011	2.16	1.94	-0.22	1.957	-0.2
12	S00012	2.02	1.9	-0.12	1.835	0.7
13	S00013	1.8	1.64	-0.16	1.642	0.0
14	S00014	0.92	0.84	-0.08	0.871	-0.3
15	S00015	5.85	5.19	-0.66	5.189	0.0
16	S00016	2.7	2.46	-0.24	2.430	0.3
17	S00017	0.73	0.59	-0.14	0.705	-1.3
18	S00018	0.8	0.74	-0.06	0.766	-0.3
19	S00019	1.03	0.94	-0.09	0.968	-0.3
20	S00020	1.09	1.01	-0.08	1.020	-0.1
21	S00021	1.8	1.52	-0.28	1.642	-1.3
22	S00022	1.39	1.3	-0.09	1.283	0.2
23	S00023	1.18	1.04	-0.14	1.099	-0.7
24	S00024	0.91	0.78	-0.13	0.863	-0.9
25	S00025	1.73	1.63	-0.10	1.581	0.5

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

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 that it is statistically identical to the method currently in use.

The statistical tool used is linear regression. Bottom line -- 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 almost always underestimates the true slope, sometimes by a very 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.

All Regression Lines on the EP Evaluator graphs are Deming Regression Lines. 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.

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: While Passing-Bablok provides confidence intervals for the slope and intercept, it does not give confidence intervals for predicted Medical Decision Points. This is a serious deficiency if a primary objective of the study is to evaluate equivalence of the MDPs. Passing-Bablok is also computationally intensive, particularly for large N, and and it may be unreliable for very small N. EP Evaluator does not show Passing-Bablok statistics when N<10 or N>250.

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 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. Also, the regression coefficients are suspect.)

An outlier is, by definition, a rare occurrence. If the mathematical algorithm excludes more than 5% of the data points, the report is stamped PRELIMINARY. This indicates that the automatic procedure has failed. The user should disable automatic outlier detection, and exclude outliers manually if necessary.

Interpreting your Results

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

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 lots of error, the width will be greater and will result in 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.

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 estimates are expected to fall within the confidence interval.

Medical Decision Point Analysis: A Medical Decision Point is an an analyte concentration at which medical decisions change. If the concentration is to one side of the MDP, one decision is make; 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. 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

Alternate (Quantitative) Method Comparison

Report Interpretation Guide

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 "small", as defined by 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.

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 rather than by a mathematical algorithm.