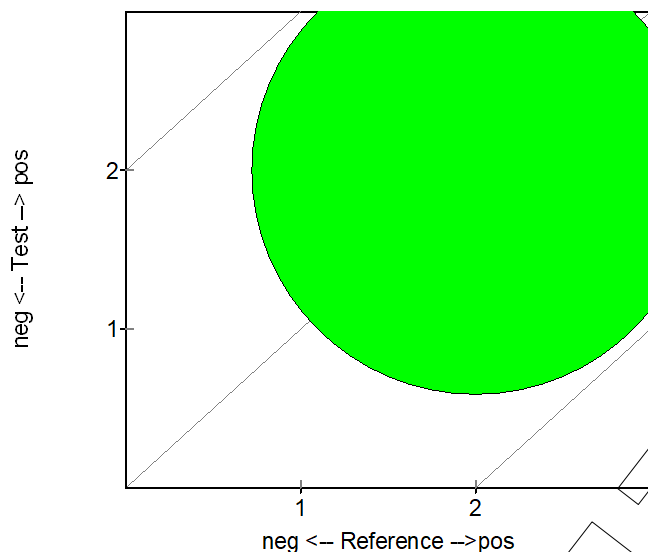## Qualitative Method Comparison

**Ref Method  veneus EDTA T=72**                    **Test Method  HC LiHep T=72**



### Statistical Analysis

(Comparison of two Laboratory Methods)

| | |
|---|---|
| Agreement | 100,0% (67,6 to 100,0%) |
| Positive Agreement | 100,0% |
| Negative Agreement | -- |

95% confidence interval calculated by the "Score" method.

McNemar Test for Symmetry:

| | |
|---|---|
| Test < Reference | 0  (0,0%) |
| Test > Reference | 0  (0,0%) |
| Symmetry test PASSES | |

Cohen's Kappa          100,0% (100,0 to 100,0%)

Kappa is the portion of agreement above what is expected by chance
The rule of thumb is that Kappa > 75% indicates "high" agreement.
We would like to see VERY high agreement (close to 100%).

### Statistical Summary

| | Negative Reference | Positive Reference | Total |
|---|---|---|---|
| **Negative Test** | -- | -- | 0 |
| **Positive Test** | -- | 8 | 8 |
| **Total** | 0 | 8 | 8 |

Number excluded or missing: 0

### Legend

| | Reference | Test |
|---|---|---|
| 1 | Negative (N) | Negative (N) |
| 2 | Positive (P) | Positive (P) |

### Experiment Description

| | Reference Method | Test Method |
|---|---|---|
| Analyst: | Jan | Jan |
| Date: | 11 dec 2019 | 11 dec 2019 |
| Comment | | |

Accepted by: _____        _____

Signature                                              Date

# EP Evaluator ®

Clinical Chemistry -- Labonovum Lab Services

## Qualitative Method Comparison

**Ref Method  veneus EDTA T=72**                              **Test Method  HC LiHep T=72**

### Experimental Results

| Spec ID | Ref | Test | Spec ID | Ref | Test | Spec ID | Ref | Test |
|---------|-----|------|---------|-----|------|---------|-----|------|
| 1 | p | p | 22 | p | p | 49 | p | p |
| 11 | p | p | 29 | p | p | 7 | p | p |
| 17 | p | p | 44 | p | p | | | |

X:Excluded  F:No Agreement

# Qualitative Method Comparison
## Report Interpretation Guide

## Qualitative Method Comparison

The purpose of this experiment is to compare two methods that report results as Positive/Negative. QMC is useful in the following circumstances:

- To determine the ability of an instrument to detect a pathological condition.

  In this case, the reference method is a definitive diagnosis, or **Gold Standard.** The key statistics are **Sensitivity** and **Specificity,** and the proportion of **False Positives** and **False Negatives.**

- To compare the performance of two laboratory instruments, perhaps when a new instrument is introduced into the lab.

  Here both the reference method and the test method are subject to experimental error. There is no reason to assume the reference method is correct, or even that it is better than the test method. In fact, when the reference method is an old method and the test method is a new method, it is quite possible that the new method is the better method.

  When comparing two laboratory methods, the key statistic is **Agreement** -- do the two methods give the same results?

## Semi-Quantitative Method Comparison

**Purpose:** to compare methods that return non-quantitative results. Example: a study comparing a dipstick method for urinary protein to a quantitative method. The results for the quantitative method are divided into groups corresponding to the dipstick categories. This approach can be used to compare semi-quantitative tests with quantitative tests. The key statistic used in semi-quantitative comparison is **Agreement.**

## Definitions

**Gold Standard:** A method that is absolutely correct.

**True Positive (TP):** A specimen that tests positive by both methods.

**True Negative (TN):** A specimen that tests negative by both methods.

**False Negative (FN):** A specimen that is negative by the test method and positive by the reference method, with the implicit assumption that the reference method is correct.

**False Positive (FP):** A specimen that is positive by the test method and negative by the reference method, again with the implicit assumption that the reference method is correct.

**Sensitivity:** The probability that the test will be positive in a population in which everyone should test positive. The ideal sensitivity is 100%. Applicable only when the reference method is a Gold Standard.

**Specificity:** The probability that the test will be negative in a population in which everyone should test negative. The ideal specificity is 100%. Applicable only when the reference method is a Gold Standard.

**Prevalence:** The frequency of positives in a given population -- possibly some population other than the one evaluated in the method comparison experiment.

**Predictive Value:** These statistics vary with prevalence, and are applicable only when the reference method is a Gold Standard.

- **Predictive Value Positive:** Probability that a positive result accurately shows the presence of the underlying condition.

- **Predictive Value Negative:** Probability that a negative result accurately shows the absence of the underlying condition.

**Agreement:** The percent of total cases in which the two methods give the same result. Related statistics for qualitative tests:

- **Positive Agreement** is the percent of cases that match when the reference method is positive: $TP/(TP+FN)$.

- **Negative Agreement** is the percent of cases that match when the reference method is negative: $TN/(TN+FP)$.

**Cohen's Kappa:** Similar to Agreement, but adjusted for the probability that the two methods agree by chance. Kappa ranges from -100% to 100%. A value of zero indicates random agreement. A value of 100% indicates perfect agreement. It is desirable for Kappa to be well above 75%.

**McNemar Test for Symmetry:** A test for bias -- whether one method is consistently larger than the other. If the number of cases where X>Y is equal (within random error) to the number of cases X<Y, the method is unbiased, and the symmetry test passes. If most of the differences between X and Y occur when X>Y (or when X<Y), the symmetry test fails.

## 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.

This report is preliminary if there are less than 20 unexcluded results.

## Qualitative Method Comparison

### Report Interpretation Guide

### Chart Interpretation

The Bubble Chart is the equivalent of a scatter plot for non-quantitative data. Green circles on the central diagonal represent points of agreement. The size (area) of the circle is proportional to number of specimens. The ideal chart has only green circles.

Yellow and red circles represent points of disagreement.

- For a qualitative comparison, red indicates false positives, and yellow indicates false negatives.

- For a semiquantitative comparison, yellow circles represent cases where the reference and test are in adjacent levels. Red circles are cases where the test and reference are two or more levels apart.

### References

1. CLSI Document EP12-A. User protocol for evaluation of qualitative test performance; Approved guideline. CLSI, 940 West Valley Road, Suite 1400, Wayne, PA 19087-1898 USA, 2002. (References to this document will be to CLSI:EP12)