IFCC-WorldLab EuroMedLab Berlin 2011
IFCC WorldLab and EuroMedLab • Berlin 2011 • 15-19 May

STATISTICAL TEST FOR EQUIVALENCE IN ANALYSIS OF METHOD COMPARISON EXPERIMENTS: APPLICATION IN COMPARISON OF ANTI-MULLERIAN HORMONE ASSAYS

T Keller[1], SA Faye[2] and T Katzorke[3]

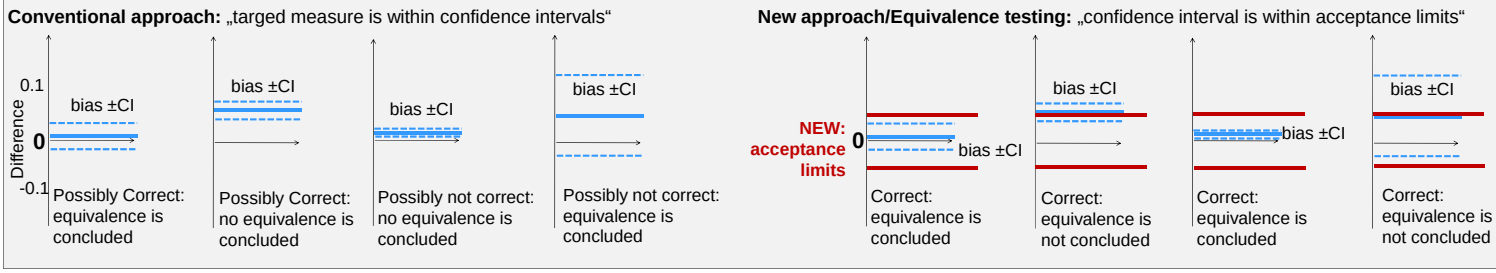
[1]ACOMED statistik, Leipzig, Germany, [2] Scientific Marketing, Beckman Coulter, Inc.,NYON, Switzerland and [3]Novum Clinic, Essen, Germany

Introduction

A considerable part of method comparison and bias estimation experiments are performed to *confirm* a distinct relationship, e.g. equivalence between two methods. Usually, the traditional null hypothesis of bias=0 is tested in statistical analysis and failure to reject the hypothesis (p>0.05) leads to the conclusion that both methods are equivalent. Because "absence of evidence is not evidence of absence", this approach is inappropriate[1]. False positive as well as false negative proof of relationship could follow as shown in Figure 1. Statistical testing of equivalence is the method of choice, whereby equivalence does not mean identity but variability within predefined acceptance limits. In equivalence testing simple statistical tests are used to show that deviations are within these limits. In the pharmaceutical industry method validation literature recommends the equivalence test[2] whereas, current guidelines in clinical chemistry and laboratory medicine[3] do not propose this type of verification.

In this study, equivalence testing is applied to a clinical chemistry scenario: A laboratory's wishing to verify a previously validated relationship between two different Anti-Mullerian hormone (AMH) assays (DSL AMH assay 10-14400 and AMH Gen II A79765 both from Beckman Coulter Inc) where the relationship relationship between AMH Gen II (y) and DSL AMH (x) assays was found to be y = 1.40 - 0.62[4] . The aim of this study was to demonstrate the equivalence of AMH GEN II and converted AMH DSL values ("corr. DSL") in the new dataset to confirm the previously derived conversion factor.

Fig.1: Principle of statistical deduction used in analysis of method comparison experiments when equivalence should be shown. Example of difference plot.



Methods (Statistics)

The statistical methods to estimate bias and its confidence interval are no different from the usual analyses performed in clinical chemistry laboratories for method comparison and vias estimation: Difference plots or regression methods (e.g. Deming and Passing-Bablok regressions) are applied, and the bias is estimated together with its 95%-confidence interval. However, these results are used in a different matter. The following steps have to be performed:

1. Limits of acceptance and the concentrations for which bias should be estimated have to be defined prospectively.
2. A statistical analysis of the method comparison experiment is performed as usual (Difference plots, Deming-regression, Passing-Bablok-regression). Bias ± confidence interval at decision point(s) is calculated. Software like Analyse-It® (see presented results) or Medcalc® can be used.
3. If the 95%-confidence interval (CI) of bias is within predefined acceptance limits, equivalence is shown with p<0.05.

Methods (Analytical)

196 women attending the Novum Fertility Clinic, Essen had AMH measured using both the DSL and AMH Gen II ELISA assays from Beckman Coulter Inc

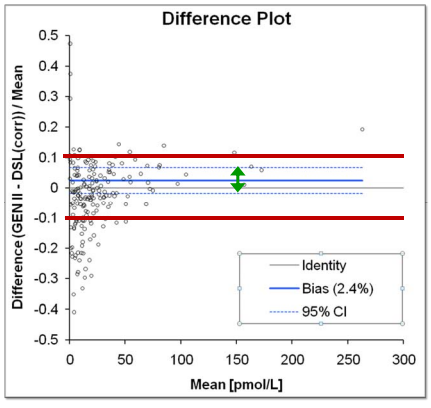
Results

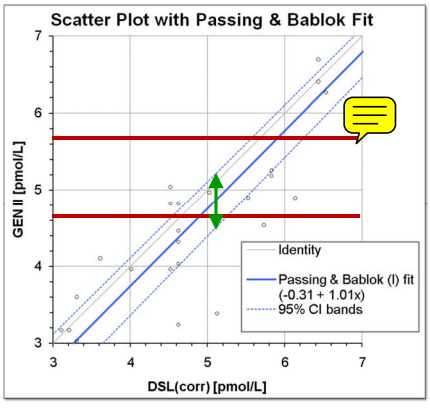
5.15 pmol/L, a clinically important decision limit for AMH was used as the point for statistical analysis. As acceptance limits ± 10% of the value of the x method was specified.

Two statistical different methodologies were used investigate equivalence as shown in Figures 1 and 2.

Fig.2 (left): Difference plot. Bias (blue line) is estimated as 2.4% and is within specifica-tion of ± 10%. Because 95%-CI of bias (green): -1.9 to 6.8% is completely within acceptance limits (red): -10% to 10%, equivalence is verified at 5% significance level.

Fig.3 (right): Passing-Bablok regression. Bias (blue line at DSL=5.15 pmol/L) is estimated -0.24 pmol/L and is within specification (± 0.51 pmol/L). However, because 95%-CI of bias (green): -0.60 to 0.11 pmol/L is not within acceptance limits (red) -0.51 to +0.51, equivalence can not be statistically verified at 5% significance level. This negative result is related to wider CI yielded by non parametric Passing-Bablok regression.





Conclusions

The recently derived conversion factor between the DSL AMH and AMH Gen II assays was confirmed in a new, independent study using equivalence testing applied to an Bland-Altman difference plot. The relationship was not verified using the Passing-Bablok regression, although the bias is within specification the confidence limits lie just outside. We believe that this is due to that Passing-Bablok regression, being a non parametric method, is not ideal for equivalence testing.

We believe that method validation experiments should use equivalence testing when a distinct relationship has to be confirmed. Not only is this statistically more correct but false positive and false negative conclusions are avoided and thus brings about a fundamental change in the conclusion – proof of hazard is replaced by proof of safety.
This methodology is also applicable for statistical analysis of a wide range of other method validation experiments (e.g. carry-over, robustness, commutability etc.)

References

1. Altman D and Bland M: (1995), Absence of evidence is not evidence of absence, BMJ 311: 485,
2. Ermer J & Miller J H McB (2005): Method Validation in Pharmaceutical Analysis. A Guide to Best Practice. WILEY-VCH Verlag GmbH & Co. KGaA,
3. CLSI EP9 A2IR (2010): Method comparison and bias estimation using patient samples. Clinical and Laboratory Standards Institute, Wayne, USA,
4. Wallace AM, Nelson SM, Faye SA, Fleming RA (2011): Multi Centre Evaluation of the new Beckman Coulter Anti-Müllerian Hormone (AMH) Immunoassay (AMH Gen II). Annals of Clinical Biochem 2011 (In Press)