1. Introduction

Concordance analysis is a comparison between the results generated by two laboratories in which the assay methods used by each laboratory are considered subject to random measurement errors. Here neither of the laboratories involved is considered the "gold standard". Otherwise, it becomes a calibration problem.

The importance of the method stemmed from the fact that pharmaceutical company manages or partnerships with laboratories scattered around the world or outsources the testing of their clinical samples. By successfully demonstrating concordance between laboratories, clinical trial can be tested in multiple laboratories throughout a clinical program.

2. Materials and Methods

A typical design of concordance analysis requires a panel of 30-60 samples covering the range of the assay. It is suggested that the panel of the samples used should consist of low, medium, and high concentration samples. It is also suggested to include samples less than the lower limit of quantitation (LLOQ) of the assay for the purpose of verifying of the assay in both laboratories. Ideally, in each laboratory, qualified technician shall execute 3-6 independent runs for each sample of the panel.

The comparison between the assays was performed based upon the geometric mean titers (GMT) of each sample. The GMT result for an individual sample was obtained by averaging the results of each run. To appropriately compute the GMTs, precautious step shall be followed for those results that could not be quantitated. Individual titers listed below the LLOQ in each laboratory shall be assigned half of the LLOQ. If the computed GMT has result below the LLOQ, it shall be reported '<LLOQ'. For concordance analysis samples with GMTs below the LLOQ shall be used to assess sensitivity of the assay and not included in the estimation procedure.

2.1. Estimation of intercept and concordance slope

The test results between two laboratories were compared and the intercept and concordance slope was estimated using the errors in variables model. The estimates were calculated under the assumption that the two assay procedures have comparable variability (λ=1) as the methods were validated in each laboratory. The following model, which naturally models comparison studies and is known as errors-in-variables model was used (Tan and Iglewicz, 1999):

$$\begin{cases} X_i = \xi_i + \delta_i \\ Y_i = \eta_i + \varepsilon_i \\ \eta_i = \alpha + \beta\xi_i \end{cases} \qquad i = 1, 2, \ldots, N,$$

where X_i and Y_i denote the *log* of measurements (GMT) in both laboratories of the ith sample, respectively. Due to the nature of the assay, log base 2 is used for quantal assay while log base 10 is used for continuous assay; however, any base could be used. ξ_i and η_i represent the unobservable population parameter ("true values") of X_i and Y_i, respectively. In the previous model the measurement errors, δ_i and ε_i are assumed to be bivariate Gaussian with mean zero and variances σ^2.

The concordance intercept and slope were estimated as $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$, and

$$\hat{\beta} = \frac{S_{yy} - S_{xx} + \sqrt{(S_{yy} - S_{xx})^2 + 4S_{xy}^2}}{2S_{xy}},$$

respectively, where $S_{xx} = \sum_{i=1}^{N}(X_i - \bar{X})^2$, $S_{xy} = \sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})$, and $S_{yy} = \sum_{i=1}^{N}(Y_i - \bar{Y})^2$.

In the case the two methods don't have comparable variability, the difference in variability can be included in the slope estimation procedure using the precision ratio λ defined by $\lambda = \sigma_\varepsilon^2 / \sigma_\delta^2$, where σ_ε^2 and σ_δ^2 are the precision parameters for the two methods or assays. The equation for the estimation of the slope becomes:

$$\hat{\beta} = \frac{S_{yy} - \lambda S_{xx} + \sqrt{(S_{yy} - \lambda S_{xx})^2 + 4\lambda S_{xy}^2}}{2S_{xy}}.$$

The variance of $log_a(\hat{\beta})$ was estimated using the sampling properties of principal components. Here the parameter *a* represents the base of logarithm transformation. An equivalent estimate of the slope, $\hat{\beta}$, is given by e_{12}/e_{11}, where e_{11} and e_{12} are the components of the first eigenvector (**e₁**) of the variance/covariance matrix of X_i and Y_i. The variance/covariance matrix of e_{11} and e_{12} is given by:

$$\frac{\lambda_1 \lambda_2}{(\boldsymbol{n}-1)(\lambda_2 - \lambda_1)} \boldsymbol{e}_2 \boldsymbol{e}_2^T,$$

where λ_k $(k = 1, 2)$ is the *k* eigenvalue and $\mathbf{e_2}$ is the second eigenvector of the variance/covariance matrix of X_i and Y_i. Thus, the approximate variance of $log_a(\hat{\beta})$ is obtained using the delta method:

$$\text{var}\{\log_a(\hat{\beta})\} = \text{var}\left\{\log_a\left(\frac{e_{12}}{e_{11}}\right)\right\} = \left(\frac{1}{\ln(a)}\right)^2 \left[\frac{\text{var}(e_{12})}{e_{12}^2} + \frac{\text{var}(e_{11})}{e_{11}} - 2\frac{\text{cov}(e_{11}, e_{12})}{e_{11} \times e_{12}}\right]$$

The 95% confidence limits for the concordance slope was computed as

$$a^{\left[\log_a(\hat{\beta}) \pm t_{\alpha/2, N-2}\sqrt{\text{var}\{\log_a(\hat{\beta})\}}\right]}.$$

The variance of the concordance intercept was estimated using the delta method as

$var(\hat{\alpha}) = \frac{s^2}{n} + \bar{X}^2 \hat{\beta}^2 \left[\frac{s_{\hat{\beta}}^2}{\hat{\beta}^2} + \frac{s_{\bar{X}}^2}{\bar{X}^2} - 2\rho \frac{s_{\hat{\beta}} s_{\bar{X}}}{\hat{\beta}\bar{X}}\right]$, where ρ is the correlation between the results generated by both laboratories, s^2 is the regression mean square error (MSE), $s_{\bar{X}}$ is the

standard error of the mean, $s_{\hat{\beta}}$ is the standard error of the slope,

and $var\left\{log_a(\hat{\beta})\right\} \cong \frac{var(\hat{\beta})}{\hat{\beta}^2(\ln a)^2}$. The 95% confidence limits were given by

$$\hat{\alpha} \pm t_{\alpha/2,N-2}\sqrt{var(\hat{\alpha})}\,.$$

2.2. Agreement

Agreement shall be estimated using the constant bias according to the formula 100% x $(a^{\overline{D}} - 1)$, where $\overline{D}$ is the mean difference of the log_a -transformed results of the 2 laboratories.

3. Results

Hypothetical examples were used to illustrate the proposed statistical methodology. When a company operates laboratories around the world, it is appropriate to compare results generated by any two laboratories performing the same test. If test results are comparable clinical testing can be conducted in either laboratory. For each assay, a panel of 30 to 60 samples was selected based upon the reported titer and sent to the laboratories that would perform the assay. Each sample was tested in 3 to 6 independent runs by qualified technicians in each laboratory. The GMT of the 3 to 6 independent runs was used for statistical analysis.

3.1. Example 1

In this example, radioimmunoassay (RIA) test results generated by two laboratories that tested the same panel of samples were compared to assess the comparability of results.

The concordance slope estimated was 1.03 and the calculated 95% confidence interval for was 1.00 -1.07. A suggested acceptance criterion for the concordance slope was "the 95% confidence interval for concordance slope is within 0.8 – 1.25", which was verified here. The agreement between the results generated by the two laboratories for the radioimmunoassay was estimated using the constant bias. The estimate (95% confidence interval) was –11.31 (-17.26, -4.94). The acceptance of this result will depend on the level of bias accepted by the laboratory and historically observed difference between the laboratories. The result obtained was within the variability of the test and should be acceptable. The 4-fold rise was 4.17 with 95% CI of (3.89, 4.39), which showed that the 4-fold rise in Lab 1 was not significantly different from 4-fold rise in Lab 2 as the 95% CI contained 4.

The statistical linear relationship between the results generated by both laboratories is displayed graphically in Figure 1, in log_{10} scale. The line of perfect concordance (slope=1 and intercept=0) is represented by the dashed blue line. The parameter estimates are given in Table 1.

Figure 1: Statistical linear relationship of radioimmunoassay measured in two laboratories

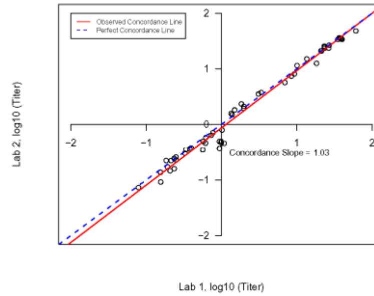

Table 1: Concordance parameter estimates for Radioimmunoassay tests

Parameter	Estimate	95% LCL	95% UCL
Concordance Slope	1.03	1.00	1.07
Intercept	-0.06	-0.10	-0.03
Agreement (%)	-11.31	-17.26	-4.94
Four-Fold Rise	4.17	3.98	4.39

3.2. Example 2

The statistical linear relationship of two laboratories testing quantal assay is displayed graphically in Figure 2. Due to the nature of the assay, the GMT results were $\log_2$-transformed prior to statistical analysis. The concordance slope estimated was 1.00 with the 95% confidence interval 0.95- 1.05. However, a constant bias was observed across the range of the assay.

The Agreement estimated was 87.10% with 95% confidence interval of (70.85%, 104.89%). These results translate in fold difference estimate of 1.87 with 95% confidence interval of (1.71, 2.05). This shows that on average the results generated by the 2 laboratories are within a single two-fold dilution.

Figure 2: Statistical Linear Relationship between two Laboratories for Functional Assay

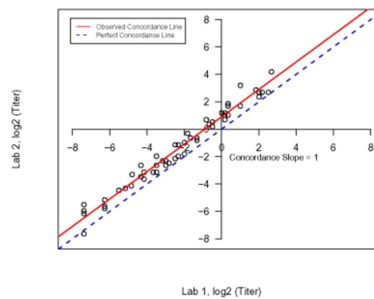

Table 2: Parameter Estimates

Parameter	Estimate	95% LCL	95% UCL
Concordance Slope	1.00	0.95	1.05
Intercept	0.90	0.62	1.18
Agreement (%)	87.10	70.85	104.89
Four-Fold Rise	4.00	3.89	4.29

4. Conclusion

In general, the agreement measurement can be used when the interest is in assessing the constant bias between the two laboratories. The fold rise is mostly useful when increased in response is of interest.
For the comparison of serology results produced in different laboratories, the traditional statistical model (linear regression, etc) can easily misrepresent the data due to the violations of assumptions. Statistical linear relationship established by concordance model (errors in variables) shall be used as the alternative in clinical program to bridge the serology data between the laboratories. A 'perfect concordance' (example 3.1) indicates that individual result can be used directly in a clinical program regardless the laboratory where the sample has been tested. If the concordance analysis shows two laboratories have significant constant bias (example 3.2), the concordance equation can be used to reassess clinical thresholds.

References

1. Tan C. and Iglewicz B. (1999) Measurement-Methods Comparisons and Linear Statistical relationship, Technometrics Vol. 41, No. 3.Tan, C and Iglewicz, B. (1999), *Measurement-Methods Comparisons and Linear Statistical relationship*, Technometrics Vol. 41, No. 3.
2. U.S. Food and Drug Administration (1992), "Bioavailability and Bioequivalence Requirements," in *U.S. Code of Federal Regulations (21, Chap. 320)*, Washington, DC; U.S. Government Printing Office.
3. Schofield, Tim (2000). *Assay Validation*. Encyclopedia of Biopharmaceutical Statistics. New York: Marcel Dekker, pp. 21-30.
4. Casella, G. and Berger, R. L. (1990), *Statistical Inference*, Pacific Grove, CA: Wadsworth.
5. Amelia Dale Horne (1998), *Statistics, Use in Immunology*, Division of Biostatistics and Epidemiology, CBER, US FDA, Rockville, MD, Encyclopedia of Immunology, London: Academic Press.