Concordance Between Laboratories Testing Clinical Samples and Its Impact on Vaccine Booster Thresholds

Eloi P. Kpamegan, Ph.D.^{a,†}; Lingyi Zheng, MS^a

^a Sanofi Pasteur, One Discover Drive, Swiftwater PA 18370, USA;

[†] Currently at Novavax Inc., 22 Firsfield Road, Gaithersburg, MD 20878, USA

Concordance study between two laboratories testing clinical samples is important to a clinical program to assess the comparability of test results, especially when the two laboratories performed the same test at different locations. Assays that measure immune response should not be compared using the simple linear regression model; instead literature publications recommend an “errors-in-variables” model that accounts for the variability in these assays (Horne, 1998). In this paper, concordance slope and intercept was estimated using the statistical linear relationship (Tan and Iglewicz, 1999) for various tests between laboratories. The confidence bounds for the concordance line were used to assess the impact on vaccine booster thresholds when a consistent bias was observed between the results generated by two laboratories.

Key words: Concordance slope, errors-in-variables, statistical linear relationship

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 samples can be tested in multiple laboratories throughout a clinical program.

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 with concentrations or titers should consist of low, medium, and high concentration or titer samples. It is also suggested to include samples less than the lower limit of quantitation (LLOQ) of the assay for the purpose of verifying that negative samples are negatives in both laboratories. Ideally, a qualified technician in each laboratory shall execute 3-6 independent runs for each sample in the panel.

The comparison between the assays was performed based upon the geometric mean titer (GMT) of each sample. For individual titers listed below the LLOQ in each laboratory, an assigned GMT of half of the LLOQ

Corresponding author: Eloi P. Kpamegan, Ph.D., Sanofi Pasteur, One Discover Drive, Swiftwater PA 18370, USA.

is used. However, a computed GMT result below the LLOQ shall be reported as ‘<LLOQ’. For concordance analysis samples with GMTs below the LLOQ shall be used to assess sensitivity of the assay and should not be included in the estimation procedure.

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 ($\lambda = 1$) 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):

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

where X_i and Y_i denote the *log* of measurements (GMT) in both laboratories of the i^{th} 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

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}, \text{ and } \hat{\beta} = \frac{S_{yy} - S_{xx} + \sqrt{(S_{yy} - S_{xx})^2 + 4S_{xy}^2}}{2S_{xy}},$$

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 do not 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 (e_1) of the variance/covariance matrix of X_i and Y_i . The variance/covariance matrix of e_{11} and e_{12} is given by:

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

where λ_k ($k = 1, 2$) is the k eigenvalue and 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:

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

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

$$\text{var}(\hat{\alpha}) = \frac{s^2}{n} + \bar{X}^2 \hat{\beta}^2 \left[\frac{s_{\hat{\beta}}^2}{\hat{\beta}^2} + \frac{s_{\bar{X}}^2}{\bar{X}^2} - 2\rho \frac{s_{\hat{\beta}} s_{\bar{X}}}{\hat{\beta} \bar{X}} \right],$$

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

$$\text{var}\{\log_a(\hat{\beta})\} \cong \frac{\text{var}(\hat{\beta})}{\hat{\beta}^2 (\ln a)^2}.$$

The 95% confidence limits were given by

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

Agreement

Agreement shall be estimated using the constant bias according to the formula $100\% \times (a^{\bar{D}} - 1)$, where

$\bar{D}$ is the mean difference of the $\log_a$ -transformed results of the 2 laboratories.

***f*-fold Rise**

It is common in the vaccine business to measure the antibody concentration prior and after vaccination. The *f*-fold rise is the ratio antibody concentration pre- and post-vaccinations. Given the concordance slope $\hat{\beta}$, an *f*-fold-rise in response in Laboratory 1 is equivalent to $f^{\hat{\beta}}$ -fold rise in response in Laboratory 2 with 95% confidence interval given by

$$f^{\exp\left[\ln(\hat{\beta}) \pm t_{0.95, n-2} \sqrt{\text{Var}[\ln(\hat{\beta})]}\right]}.$$

To show that the fold-rise in response in laboratory 2 is given by $f^{\hat{\beta}}$ per *f*-fold increase in laboratory 1, let x_{pre} and x_{post} denote pre-vaccination and post-vaccination titers for a sample tested in Laboratory 1, and the fold change in response (i.e., $\frac{x_{post}}{x_{pre}}$), by f . Also let y_{pre} and y_{post} denote pre-vaccination and

post-vaccination titers for a sample tested in Laboratory 2. Given the concordance estimates for slope and intercept, $\hat{\beta}$ and $\hat{\alpha}$, respectively, we have from the concordance analysis that

$$\log_a(y_{pre}) = \hat{\alpha} + \hat{\beta} \times \log_a(x_{pre}) \quad \text{and} \quad \log_a(y_{post}) = \hat{\alpha} + \hat{\beta} \times \log_a(x_{post}) \quad . \quad \text{Therefore,}$$

$$\log_a(y_{post}) - \log_a(y_{pre}) = \hat{\alpha} + \hat{\beta} \times \log_a(x_{post}) - (\hat{\alpha} + \hat{\beta} \times \log_a(x_{pre})) \quad , \quad \text{which implies}$$

$$\frac{y_{post}}{y_{pre}} = \left(\frac{x_{post}}{x_{pre}} \right)^{\hat{\beta}} = f^{\hat{\beta}}.$$

Confidence Bounds for the Expected Laboratory 2 Titers Given Laboratory 1 Titers

The concordance equation was used to obtain the 95% confidence interval for the predicted Laboratory 2 titers given various values of corresponding Laboratory 1 titers.

Assume that the estimated statistical linear relationship in log scale is defined by:

$$\hat{Y}_{lab2}(t_{lab1}) = \hat{\alpha} + \hat{\beta} \times t_{lab1}$$

$\hat{Y}_{lab2}(t_{lab1})$ is the estimated predicted log-transformed titer in the Laboratory 2 given a corresponding log-transformed titer t_{lab1} in Laboratory 1.

The 95% confidence interval of the predicted log-transformed titers given that of Laboratory 1 is given by the following equation:

$$\hat{Y}_{lab2}(t_{lab1}) \pm t_{\alpha/2, (n-2)} \times \sqrt{\text{var}(\hat{Y}_{lab2}(t_{lab1}))}$$

where

$$\text{var}[\hat{Y}_{lab2}(t_{lab1})] = \frac{s^2}{n} + \hat{\beta}^2(t_{lab1} - \bar{X}_{lab1})^2 \left[\frac{s_{\hat{\beta}}^2}{\hat{\beta}^2} + \frac{s_{\bar{X}_{lab1}}^2}{(t_{lab1} - \bar{X}_{lab1})^2} - 2\rho \frac{s_{\hat{\beta}} s_{\bar{X}_{lab1}}}{\hat{\beta}(t_{lab1} - \bar{X}_{lab1})} \right].$$

$\bar{X}_{lab1}$ is the mean of the log base 10 (or log base 2) transformed Laboratory 1 results and s^2 is the error mean square.

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. In this paper, a panel of 30-60 sixty samples was selected based upon the reported titer and sent to two different laboratories to perform the same assay. Each sample was tested in 3-6 independent runs by qualified technicians in each laboratory. The GMT of the 3-6 independent runs was used for statistical analysis.

Example 1

In this example, the comparability of the radioimmunoassay (RIA) test results generated by two laboratories on the same panel of samples was assessed.

The concordance slope estimated was 1.03 and the calculated 95% confidence interval (CI) was 1.00 to 1.07. The suggested acceptance criterion for the concordance slope; i.e., “the 95% (CI) for concordance slope is within 0.8 and 1.25”, was verified. The agreement between the results generated by the two laboratories for the radioimmunoassay was estimated using the constant bias. The estimate (95% CI) was -11.31 (-17.26, -4.94). The acceptance of this result was dependent on the level of bias accepted by the laboratory and historically observed difference between the laboratories. In this case, the result obtained was within the variability of the test and was acceptable. The 4-fold rise was 4.17 with 95% CI of (3.89, 4.39), which showed that the 4-fold rise in Laboratory 1 was not significantly different from 4-fold rise in Laboratory 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 log₁₀ 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.

Example 2

The statistical linear relationship of two laboratories testing using a quantal assay is displayed graphically in Figure 2. Due to the nature of the assay, the GMT results were log₂-transformed prior to statistical analysis. The concordance slope estimated was 1.00 with the 95% CI of (0.95, 1.05). However, a constant bias was observed across the range of the assay.

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

Determination of Clinical Threshold Based on Confidence Bounds of the Concordance Line

Laboratory 2 from Example 2 shows a consistent bias across the range of the assay. If clinical testing

needs to be conducted in Laboratory 2, the 95% confidence bound for the concordance line can be used to determine the new threshold. The 95% CI for the expected Laboratory 2 titers given Laboratory 1 titers is given in Table 3. The upper and lower 95% CI of the expected Laboratory 2 results given Laboratory 1 results were used to determine 95% CI for the expected ratio of Laboratory 2 result over Laboratory 1's result. The 95% CI expected ratios are given in the last two columns of Table 3.

The minimum and the maximum values estimated for the 95% CIs of the expected ratio shown in the last 2 columns of Table 3 are 1.70 and 2.74, respectively. These values were used as multiplicative constants to the previously determined cut-off level from the historical data (1.28 IU/mL). When applied for the assay in Example 2, the computed cut off values were 2.2 and 3.5. The quantal assay is based upon serial dilutions and provides quantal results. Since the only possible normal reported result between 2.2 and 3.5 was 2.56, the cut-off value of 2.56 is proposed.

In the case the normal reported value does not fall within these numbers, the closest value to the intervals can be used if accompanied with additional confirmatory analyses and is supported by clinicians.

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

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

Table 3

95% Confidence Bounds Estimated and 95% Confidence Intervals for the 2-fold Relationship Between Laboratory 1 and Laboratory 2

Sample ID	Lab 1 GMT of Replicates (N=6)	Lab 2 GMT of Replicates (N=6)	Expected Lab 2 GMT (95% CI)	Standard Error	Ratio of 95% LCL over Lab 1	Ratio of 95% UCL over Lab 1
1	0.011	0.018	0.024 (0.019, 0.030)	0.1688	1.71	2.73
2	0.02	0.101	0.043 (0.035, 0.053)	0.1546	1.73	2.66
3	0.04	0.113	0.085 (0.070, 0.103)	0.1388	1.75	2.58
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
57	0.16	0.453	0.334 (0.287, 0.389)	0.1100	1.79	2.43
58	1.28	2.032	2.605 (0.325, 2.919)	0.0819	1.82	2.28
59	2.56	9.123	5.165 (4.624, 5.769)	0.0796	1.81	2.25
60	10.24	18.25	20.311(17.972, 22.953)	0.0881	1.76	2.24

Conclusion

In general, the agreement measurement can be used when the primary interest is to assess the constant bias between two laboratories. The fold rise measurement is more useful when an increased response is the primary 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

- Tan C. and Iglewicz B. (1999) Measurement-Methods Comparisons and Linear Statistical relationship, *Technometrics* Vol. 41, No. 3.
- 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.
- Schofield, Tim (2000). *Assay Validation*. Encyclopedia of Biopharmaceutical Statistics. New York: Marcel Dekker, pp. 21-30.
- Casella, G. and Berger, R. L. (1990), *Statistical Inference*, Pacific Grove, CA: Wadsworth.
- Amelia Dale Horne (1998), *Statistics, Use in Immunology*, Division of Biostatistics and Epidemiology, CBER, US FDA, Rockville, MD, Encyclopedia of Immunology, London: Academic Press.