

# Chapter 1

## Appendix G : Agreement Criteria

Roy (2009) proposes a suite of hypothesis tests for assessing the agreement of two methods of measurement, when replicate measurements are obtained for each item, using a LME approach. (An item would commonly be a patient).

Two methods of measurement can be said to be in agreement if there is no significant difference between in three key respects. Firstly, there is no inter-method bias between the two methods, i.e. there is no persistent tendency for one method to give higher values than the other.

Secondly, both methods of measurement have the same within-subject variability. In such a case the variance of the replicate measurements would consistent for both methods. Lastly, the methods have equal between-subject variability.

For the mean measurements for each case, the variances of the mean measurements from both methods are equal.

Roy (2009) sets out three criteria for two methods to be considered in agreement. Firstly that there be no significant bias. Second that there is no difference in the between-subject variabilities, and lastly that there is no significant difference in the within-subject variabilities. Roy (2009) further proposes examination of the the overall variability by considering the second and third criteria be examined jointly. Should both the second and third criteria be fulfilled, then the overall variabilities of both methods would be equal.

Barnhart et al. (2007) describes the sources of disagreement as differing population means, different between-subject variances, different within-subject variances between two methods and poor correlation

between measurements of two methods.

Roy (2009) considers two methods to be in agreement if three conditions are met.

1. no significant bias, i.e. the difference between the two mean readings is not "statistically significant",
2. high overall correlation coefficient,
3. the agreement between the two methods by testing their repeatability coefficients.

Roy (2009) demonstrates a LME model specification, and a series of tests that look at each of these agreement criteria individually. If two methods of measurement lack agreement, the specific reason or reasons for this lack of agreement can be identified.

# Bibliography

- Barnhart, H., M. Haber, and L. Lin (2007). An overview of assessing agreement with continuous measurements. *Journal of Biopharmaceutical Statistics* 17, 529–569.
- Roy, A. (2009). An application of the linear mixed effects model to ass the agreement between two methods with replicated observations. *Journal of Biopharmaceutical Statistics* 19, 150–173.