**Bland Altman 1999 Research Notes**

Agreement between methods of measurement can be quantified using observations upon the same subject, made simultaneously by each measurement method.

A lack of agreement , even to a minor degree, is inevitable between measurement methods. What matter is the degree of disagreement.

**Limits of Agreement**

Bland and Altman’s approach is motivated by the need for a metric that is easy for medical researchers to estimate and to interpret. Furthermore they intend their approach to be implementable using basic statistical software.

The **inter-method bias** (often referred to simply as “bias”) is simply the mean of the case-wise differences.

The **95% limits of agreement** are estimated by inter-method bias (i.e. the mean of the case-wise differences) plus or minus 1.96 ( often approximated by 2 with minimal loss of accuracy) times the standard deviation of the case-wise differences. These limits provide an interval within which 95% of the case-wise differences are expected to lie.

**Assumption of normality of case-wise differences.**

Calculation of the limits of agreement relies on the assumption that the case-wise differences are normally distributed.

The calculation removes a lot of the variation between subjects, leaving measurement error, which is likely to be normally distributed. (**Bland and Altman 1999** remark that this assumption is easy to check using a **normal plot**.)

**Confidence intervals for Limits of Agreement.**

It is unclear on how these confidence intervals aid in the decision of whether two methods of agreement.

As noted by several authors, it must be assumed that neither method can be assumed to be without error (**Bland Altman 1999**, for example)

**O’Brien Blood Data (JSR)**

A data set often used for demonstration purposes is a tabulation of systolic blood pressure data from a study in which simultaneous measurements were made by two experienced observers ( simply referred to as “J” and “R”) using a sphygmomanometer and by a semi-automatic blood pressure monitor (denoted as “S”).

* Inter-method Bias **\bar{d} = -16.29** mmHg
* Standard deviation of Case-wise differences **S\_d = 19.61** mmHg

These limits of agreement define the range within most( i.e. 95%) differences between measurements by the two methods would lie.

**Precision of estimated limits of agreement**

***Calculation of standard Errors and Confidence intervals of the limits of agreement.***

* Where **n** is the sample size, the variance of the **inter-method bias** is estimated by **S\_d^2/n** .
* The variance of S\_d is approximated as S\_d^2/2(n-1)

The inter-method bias, and the variance of case-wise differences are assumed to be independent**.**

Bland and Altman show that the variance of the limits of agreement to be

**\left( {1 \over n} + \frac{1.96^2}{2(n-1)} \right) S^2\_d**

Unless **n** is small, this is approximated by **1.71^2 {S^2\_d \over n}.**

**Relationship between difference and magnitude**

It is assumed that the **inter-method bias** and **standard deviation of the case-wise differences** are constant throughout the range of measurement.

The inter-method bias may be found to be approximately proportional to the range of measurements. In such a case this relationship would be evident from the **Bland-Altman plot**.

**Logarithmic Transformation**

**Section 3.1 of Bland Altman 1999**

**Nadler Hurley Data set**

**Regression approach for non-uniform differences**

**Milk (Triglyceride / Gerber) data**

This data set contains the estimated fat content of human milk (g/100ml) estimated by two different methods, an enzymic procedure which we shall refer to as the “ triglyceride method” and the standard Gerber method.

**Regression based approach (Bland Altman 1999)**

Bland and Altman 1999 propose a regression based approach.

**\hat{D} = b\_o + b\_1 A**

If **b\_1** is not significant, the **\hat{D}** is equal to **\bar{d}.**

**Disjoint Regression (O’Brien 2012)**

Separating the data set into two : one set containing observations where the case-wise difference is greater than the inter-method bias, and the other containing the observations where the case-wise difference is less.

Separate regression analyses are performed for both data sets, where the case-wise mean acts as the predictor variable. The tests of significance for the slopes determine whether or not the assumption of constant variance is met.

Importantly, as these two test are a family of tests that are test jointly, appropriate multiple comparison procedures must be employed. **(Benjamini et al)**

Reliant on Multiple Comparison Procedures (FWER)

**Non parametric variation of the Limits of Agreement approach.**

The British Hypertension Society (BHS) protocol for evaluating these machines recommended a simple non-parametric method

**Inappropriate approaches to MCS**

* Structural equation
* Intraclass correlation