BivRegBLS : a new R package in method comparison studies with tolerance intervals and (correlated)-errors-in-variables regressions

Marion Berger[1] & Bernard G. Francq[2]

[1] Biostatistiques et Programmation, Sanofi, 371 rue du Professeur Joseph Blayac, 34184 Montpellier cedex 04, France, marion.berger@sanofi.com
[2] CMC Statistical Sciences (Technical R&D), GlaxoSmithKline, 89 rue de l'institut, 1330 Rixensart, Belgique, bernard.x.francq@gsk.com

Résumé. Lorsqu'on souhaite comparer deux méthodes X et Y, l'approche graphique Bland-Altman (ou graphe (M,D) pour Moyennes versus Différences) et ses intervalles d'agrément (AIs) sont fréquemment utilisés. Une autre approche intéressante est l'utilisation de régressions avec erreurs-sur-variables permettant une représentation dans un graphique classique (X,Y). Francq et Govaerts (2014) ont démontré que la régression Bivariate Least Squares est la régression la plus appropriée et la plus générale car elle inclut les situations gérées par le régression de Deming, la régression orthogonale, la régression des moindres rectangles ainsi que les régressions des moindres carrés verticaux et horizontaux. Francq et Govaerts (2016) ont aussi réconciliés les approches (M,D) et (X, Y) en proposant des intervalles de tolérance et des intervalles de prédiction ayant de bons taux de couverture pour remplacer les AIs dont les taux de couvertures chutent en présence de données répétées ou de valeurs aberrantes. Ils ont aussi démontré que les erreurs en (M, D) étaient corrélées. Les biais augmentent et les taux de couvertures s'effondrent si cette corrélation est ignorée. Ils ont proposé une nouvelle régression, la régression CBLS (Correlated Bivariate Least Square) pour laquelle les intervalles de prédiction et les intervalles de confiance ont d'excellents taux de couverture quel que soit le design. Un nouveau package R BivRegBLS a été développé pour permettre d'exécuter les régressions BLS et CBLS dans plusieurs situations (données non répétées, données répétées, design équilibré ou déséquilibré, variances homogènes ou hétérogènes). Ce package inclut un nouvel intervalle, l'intervalle généralisé, qui combine les concepts de confiance et de prédiction. L'utilisation des régressions BLS et CBLS est à encourager car l'une permet de visualiser la comparaison dans l'échelle d'origine tandis que l'autre comble les carences de l'approche Bland-Altman.

Mots-clés. Analyse des données, Biostatistique, Chimiométrie, Logiciels, Medecine, Epidemiologie, Qualité, Fiabilité, Tolérance, Régressions avec erreurs sur variables, BLS, BivRegBLS, R, études de comparaison de méthodes, Bland-Altman

Abstract. When comparing methods (say methods X and Y), the Bland–Altman plot (or (M,D) plot for Means versus Differences) with its Agreement Intervals (AIs) are frequently used. Another interesting approach is the use of errors-in-variables regressions in a classical (X,Y) plot. Francq and Govaerts (2014) demonstrated that the Bivariate Least Squares regression is the most suitable and general regression as it covers all the cases handled by the Deming regression, the orthogonal regression, the mean geometric regression, the vertical and horizontal ordinary least squares regressions. Francq and Govaerts (2016) also reconciled the (M,D) and (X,Y) approaches by proposing tolerance intervals and predictive intervals with good coverage probabilities to replace AIs which coverage probabilities drop in presence of replicated data or outliers. They also demonstrated that errors in (M,D) were correlated. Bias soar and coverage probabilities collapse if this correlation is ignored. They proposed a novel regression, CBLS regression (Correlated

Bivariate Least Square) where predictive intervals and confidence intervals have excellent coverage probabilities whatever the design. A new R package BivRegBLS has been developed to perform BLS and CBLS under various designs (unreplicated data, replicated data with balanced or unbalanced design, variance homogeneity or heterogeneity) including a new generalized interval that combines the confidence and the predictive concepts. The use of BLS and CBLS should be encouraged as one allows visualizing the comparison in the original unit scale, while the second fulfills the deficiencies of the Bland-Altman approach.

Keywords. Data analysis, Biostatistics, Chemometry, Software, Medecine, Epidemiology, Quality, Reliability, Tolerance, Errors in variables regressions, BLS, BivRegBLS, R, method comparison studies, Bland-Altman

Content

1 Introduction

The growing and intensive needs of researchers and laboratories lead to the development of new measurement methods which are faster, less expensive or more accurate than current methods. These methods should ideally give results comparable to a standard method. Two main methodologies for assessing equivalence in method-comparison studies are presented separately in the literature.

1.1 Bland-Altman approach

The first is the well-known Bland–Altman plot (or (M,D) plot) with its Agreement Intervals (AIs) presented by Bland and Altman (1999), where two methods (X and Y) are considered interchangeable if their differences are within an acceptance interval [-Δ, Δ] (defined a priori): |Y-X| < Δ. This has also been detailed by Francq and Govaerts (2016). The success of this approach is due to the importance of this topic applied in many fields from chemometrics, engineering, cutting-edge medicine, diagnostic tools, computing, robotic surgery, pharmacy, chemistry, and many other.

1.2 Errors-in-variables regression approach

The second approach is based on errors-in-variables regression in a classical (X,Y) plot and focuses on confidence intervals (CIs), whereby two methods are considered equivalent on average when providing similar measures on average: Y=X. This approach is presented by Francq and Govaerts (2014). During this talk, we will review and extend this approach by using an acceptance interval. The measurement methods are then considered equivalent on average if the regression confidence interval is within the acceptance interval; or for any sample if the tolerance interval is within the acceptance interval. Estimating a regression line in an (X,Y) plot is not straightforward as both axes are not error-free and the errors may be heteroscedastic (Figure 1). Many regression techniques are available and it has been shown by Francq and Govaerts (2014) that the BLS (Bivariate Least Square) regression is the most suitable and general regression (it includes the Deming Regression, the Orthogonal Regression and the Ordinary Least Square regressions).

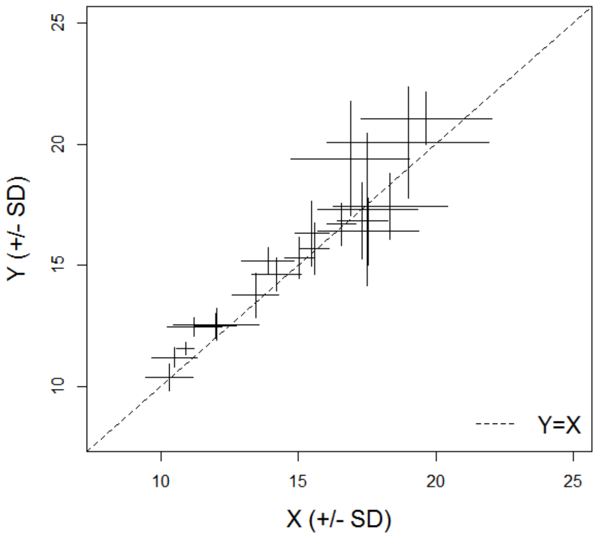

Figure 1 - Example of a data set with heteroscedasticity in a (X,Y) space obtained by the BivRegBLS R package

2 State-of-the-art

In a recent and novel paper, Francq and Govaerts (2016) reconcile these two approaches by proposing predictive and tolerance intervals in the (X,Y) space, and confidence intervals in the Bland-Altman plot. They show and explain that Tolerance Intervals (TIs) better suit the needs of the clinicians. TIs are easier to calculate and to interpret than AIs. Novel and robust TIs are provided, which are not influenced by outliers in comparison to the Bland-Altman agreement intervals. Moreover, the coverage probabilities of these TIs are excellent whatever the design of the study, while the AIs move away from their nominal levels with replicated data.

The AIs or TIs are drawn, by default, with two straight horizontal lines in a Bland-Altman plot with the averages (M=(X+Y)/2) assigned to the X-axis and their differences (D=Y-X) on the Y-axis. There is no clear solution when a pattern is observed (a proportional bias between the methods). Bland and Altman proposed to estimate a regression line, by default, with the Ordinary Least Square (OLS) method [1], and then to display two straight lines parallel to the OLS line (at 2 times the residual standard deviation). Carstensen (2010) proposed approximate predictive intervals based also on the OLS procedure, or MCMC simulations, or straight lines by neglecting the usual curvature term.

Francq and Govaerts (2016) demonstrate that modelling the differences against their averages in a Bland-Altman plot is more complicated than modelling directly the data in a (X,Y) plot. The measurement errors are, indeed, correlated in the (M,D) space. If this correlation is ignored, the biases soar considerably and the coverage probabilities collapse drastically up to 0%. Therefore, the Bland-Altman approach leads to substantial misunderstandings and misleading conclusions.

A new and promising correlated-errors-in-variables regression is then introduced: the CBLS regression (Correlated Bivariate Least Square). Novel predictive intervals (PIs) are introduced with excellent coverage probabilities whatever the design (unreplicated data, replicated data with balanced or unbalanced design).

3 BivRegBLS R Package

The new and promising R package BivRegBLS is then introduced and used to illustrate these new methodologies with real data. It will be explained how to model the data in the (X,Y) space with the BLS regression (Figure 2 left) or in the (M,D) space with the CBLS regression (Figure 2 right). It will be shown how to plot the results under different designs. A new generalized interval that combines the confidence and the predictive concepts will also be introduced.

4 Conclusion

The (correlated)-errors-in-variables regressions should be encouraged in method comparison studies, although the Bland–Altman approach is usually applied to avert their complexity. The Bland-Altman plot is, indeed, more complicated than the classical (X,Y) plot as the measurement errors are correlated. Additionally, Tolerance or predictive intervals are better than agreement intervals.

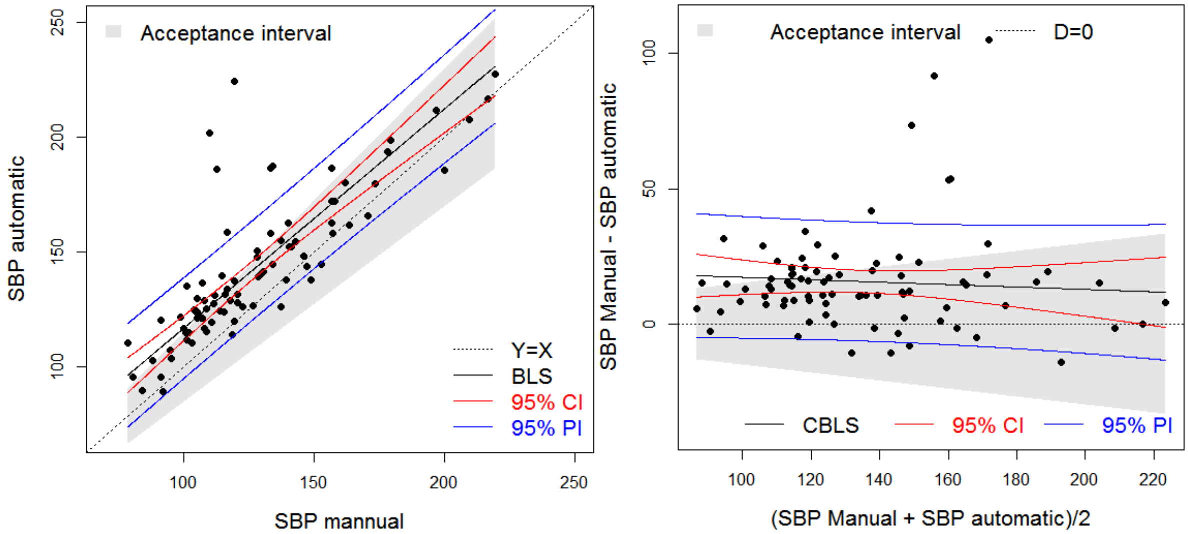

Figure 2 - The systolic blood pressure [1] in a (X,Y) plot on the left with the BLS regression, or in a (M,D) plot with the CBLS regression on the right; each with the 95% confidence intervals and predictive intervals, obtained by BivRegBLS R package

References

[1] JM Bland, DG Altman (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8:135–160.

[2] BG Francq, B Govaerts (2016). How to regress and predict in a Bland–Altman plot? Review and contribution based on tolerance intervals and correlated-errors-in-variables models. *Statistics in Medicine*, 35:2328-2358.

[3] BG Francq, BB Govaerts (2014). Measurement methods comparison with errors-in-variables regressions. From horizontal to vertical OLS regression, review and new perspectives. *Chemometrics and Intelligent Laboratory Systems*, 134:123–139.

[4] B Carstensen (2010). Comparing Clinical Measurement Methods: A Practical Guide. Wiley: Chichester, UK.