Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Pharmaceutical and Biomedical Analysis 41 (2006) 219–227
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

# A new approach to evaluate regression models during validation of bioanalytical assays

T. Singtoroj [a,1], J. Tarning [b,1], A. Annerberg [a], M. Ashton [b], Y. Bergqvist [c], N.J. White [a,d], N. Lindegardh [a,d,*], N.P.J. Day [a,d]

[a] *Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand*
[b] *Department of Pharmacology, Sahlgrenska Academy at Göteborg University, Sweden*
[c] *Dalarna University College, Borlänge, Sweden*
[d] *Nuffield Department of Clinical Medicine, Centre for Tropical Medicine, University of Oxford, Oxford, UK*

Received 6 August 2005; received in revised form 4 November 2005; accepted 6 November 2005
Available online 5 December 2005

**Abstract**

The quality of bioanalytical data is highly dependent on using an appropriate regression model for calibration curves. Non-weighted linear regression has traditionally been used but is not necessarily the optimal model. Bioanalytical assays generally benefit from using either data transformation and/or weighting since variance normally increases with concentration. A data set with calibrators ranging from 9 to 10 000 ng/mL was used to compare a new approach with the traditional approach for selecting an optimal regression model. The new approach used a combination of relative residuals at each calibration level together with precision and accuracy of independent quality control samples over 4 days to select and justify the best regression model. The results showed that log–log transformation without weighting was the simplest model to fit the calibration data and ensure good predictability for this data set.


*Keywords:* Bioanalytical assays; Calibration; Transformation; Weighting; Piperaquine; Linear regression; Liquid chromatography

## 1. Introduction

In recent years great efforts has been made to standardize international validation procedures for bioanalytical assays. Different European and American authorities such as the FDA, ICH and ISO continuously develop validation guidelines and directives about experimental design and data evaluation in the field of bioanalytical method validation [1–4]. A first attempt at harmonization and standardisation was the conference held in Washington in 1990 to discuss what a validation of bioanalytical methods should consist of, i.e. which analytical parameters (bias, precision, etc.) need to be documented to validate a method. The resulting Washington Conference Report and publications related to the conference are generally viewed as the basis for bioanalytical method validation [5,6]. However, the usefulness of some of the recommendations is questionable, particularly given the lack of advice for the practical execution of a validation study. In the light of this critique, a new SFSTP (Société Francaise des Sciences et Techniques Pharmaceutiques) committee was founded in 1995 to develop guidance for validation of bioanalytical methods. The SFSTP validation guide of chromatographic methods for drug bioanalysis was published in 1999 by Hubert et al. and illustrated the same year by Chiap et al. [7,8]. The guide has recently been updated by the introduction of the concept of an accuracy profile [9]. The accuracy profile utilises a "*β-expectation tolerance interval*" to visually discriminate between acceptable and non-acceptable regression models during pre-validation. The "*β-expectation tolerance interval*" is constructed using estimates of the bias and the standard deviation of the intermediate precision obtained from validation standards or back calculated concentrations of calibration standards analysed in replicate series [10–15]. The concept and content of the two validation phases (i.e. pre-validation and validation) is substantially covered in the literature [10–12,15–18]. Boulanger et al. state that: "During the 'pre-validation', the model to be used

* Corresponding author at: Wellcome Unit, Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok 10400, Thailand.
*E-mail address:* niklas@tropmedres.ac (N. Lindegardh).
[1] Authors have contributed equally to this work.

0731-7085/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2005.11.006

as calibration curve will be identified and the quality of fit will be assessed only at this stage. The experiments proposed are designed to consistently evaluate the adequacy of the model. In the second phase, called 'validation', the objective is to mimic the routine practice that is envisaged. The model will be used as is – the parameters will of course be estimated based on the new data – and no more investigation specific to the quality of fit will be conducted, the same way it should be carried out during routine. In this second step, the experiments are designed to focus on the estimation of the bias and precision of the method, not on the calibration curve." [18]. The present paper suggests a new approach to choose an optimal regression model. Instead of fixate the regression model during pre-validation the final choice should be based on all available data from the validation phase.

A good regression model is the foundation for accurate and reproducible quantification over the whole calibration range. A linear model is commonly preferred since the complexity increases with the use of non-linear regression. FDA guidelines state that: "Standard curve fitting is determined by applying the simplest model that adequately describes the concentration–response relationship using appropriate weighting and statistical tests for goodness of fit" [1]. These requirements sound very clear and straightforward. However, complying with the stipulations might in reality not be so simple. The simplest and most commonly used parameter to define the degree of association between two variables as a straight line is denoted by the coefficient of determination ($r^2$). Many analysts depend entirely on the value of $r^2$ being greater than 0.99 as an acceptance criterion when evaluating regression model and linearity. However, $r^2$ alone is not adequate to demonstrate linearity since $r^2$ values above 0.999 can be achieved even when the data show signs of curvature [19].

The most common approach to fit a calibration curve to data points ($x$, $y$) is by ordinary linear regression (OLR) using least squares calculation. This approach presupposes that each data point in the range has a constant absolute variance (i.e. homoscedastic data). Most bioanalytical assays usually have to cover a broad concentration range and the variance is more likely to increase with concentration (i.e. heteroscedastic data) [19–24]. A consequence of using OLR is that deviations at high concentrations will influence the regression line more than deviations at low concentrations. Thus the use of OLR with heteroscedastic data will lead to impaired accuracy despite an acceptable $r^2$ value, particularly at the lower end of the concentration range [23].

All bioanalytical assays could benefit from a regression model more complex than OLR. Alternative models include weighted linear regression (WLR) and/or data transformations [20–25]. These models will normally generate a better curve fit (i.e. smaller sum of residuals and random scatter in residual plots) than OLR. They will also minimise time-dependent variation (i.e. minimise variation in slope and intercept for standard curves obtained over several days) and increase accuracy over the whole concentration range.

Traditionally the regression model is chosen in the pre-validation phase by evaluating 3–5 series of calibration curves and comparing the total sum of residuals for each tested regression model [17,24,26,27]. Some reports have also incorporated predictability by looking at the accuracy of independent quality control (QC) samples before choosing the final model [25].

We propose a strategy that will enable the analyst to choose the regression model that gives the optimal overall performance over time. This approach is based on parameter ranking of data generated during several days (4 days in the present paper) to mimic the actual conditions during routine bioanalysis instead of only one day of pre-validation data. The curve fit was evaluated by minimising the residuals at each calibration level rather than just the total sum of residuals. Accuracy and precision were also incorporated for three independent QC levels during several days of analysis before the final regression model was chosen. Nineteen different regression models were evaluated using data obtained during the validation of a liquid chromatographic assay for piperaquine (PQ) quantification in urine using a 1000-fold concentration range (9–10 000 ng/mL) [28].

## 2. Experimental

### 2.1. Background

#### 2.1.1. Homoscedasticity

The first step during an evaluation of regression models should include a test for homoscedasticity. The two most common ways to evaluate homoscedasticity are to conduct an $F$-test (i.e. test for significant difference in variance) or to visually examine a residual versus concentration plot [24]. If the variance is constant (i.e. homoscedastic data) over the calibration range the residual versus concentration plot should show residuals randomly distributed around the $x$-axis [21]. In the $F$-test the experimental $F$-value ($F_{exp}$) is expressed as the ratio between the variance at the lowest and at the highest concentration in the calibration range as proposed by the International Organization for Standardization [29]. If the $F_{exp}$ value is greater than the tabulated $F$-value ($F_{tab}$) at a chosen confidence level the variances are significantly different (i.e. the data are heteroscedastic) [24,30–33].

#### 2.1.2. Ordinary and weighted linear regression (OLR/WLR)

OLR assumes homoscedasticity and associates the dependent variable $y$ with the independent variable $x$. The regression line is constructed so as to minimise the squared sum of the vertical distance (sum of squared residuals, SSR) between the observations and the constructed regression line [24]. One method of dealing with heteroscedastic data is to apply a weighted regression model. The principle of weighting is to give more importance to data points with a low variance and less importance to data points with high variance. Weighted models are particularly suitable for assays where the relative standard deviation (R.S.D.) is constant (i.e. S.D. increases proportional to concentration) throughout the concentration range. An optimal weighted model will balance the regression line to generate an evenly distributed error throughout the calibration range. The most commonly used weights are the empirical weights $1/x$, $1/x^2$ and $1/x^{1/2}$.