Journal of Pharmaceutical and Biomedical Analysis
16 (1998) 1107–1115

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

Bioanalytical calibration curves: variability of optimal powers between and within analytical methods[1]

Ebi Kalahi Kimanani *, Jean Lavigne

Department of Biometrics and Pharmacokinetics R & D, Phoenix International Life Sciences, 2305 Cohen Street, Ville St. Laurent, Montreal, Quebec H4R 2N6, Canada

Received 13 December 1996; received in revised form 6 February 1997

Abstract

The objective of this study was to investigate the variability of optimal power models in contrast to common regression models within and between analytical methods, as well as the frequency of outlier rejection. This was done by fitting the power model to calibration curve data using the minimum sum of squared residuals as a curve selection criterion. The jacknife percent deviation was used for detecting outliers. The data were obtained from 2087 analytical batches for 91 projects using various analytical techniques. The most frequent regression model varied between analytical techniques while the median and interquartile range of the optimal powers were stable. Outlier rejection is highest in GC and LCMS in which the Wagner (Quadratic, log-log) is the most frequent model. These results suggest that the greatest source of variability in the ideal transformation may not be the analytical technique but other within-lab sources. Outlying values may be due to these other sources of variability as suggested by the outlier rejection profile. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Bioanalytical; Calibration curves; Optimal power

1. Introduction

A Box-Cox-type [1] power transformation method has been proposed as objective criteria to fit and cross-validate bioanalytical calibration curves [2–4]. For each set of Standard Curve data (concentrations and responses), the suggested transformation involves fitting a linear regression between scaled responses raised to a certain power and scaled concentrations raised to the same power. Details of this are presented in [2] and can be found in the original paper by Box and Cox [1]. The optimal power is determined empirically as the value that minimises the sum of scaled squared residuals and will therefore depend on the specific set of concentrations and responses. As such, optimal powers will vary between analytical techniques and between analyses within techniques. The extent of this variability could have implications on analytical technique stabiIity.

* Corresponding author. Tel.: + 1 151 4330033; fax: + 1 151 4338861.

[1] Presented at the Analysis and Pharmaceutical Quality Section of the Eleventh Annual American Association of Pharmaceutical Scientists Meeting, October 1996, Seattle, Washington, USA.

PII S0731-7085(97)00063-0

In this article, two related questions pertinent to the power model are addressed. The first of these is whether for any specific dataset, the estimated optimal power is related to the true variance weight of that data. Given this information, the second question is what inferences can be drawn regarding the variability, and more generally, the distribution of the optimal powers that arise from experimental data between and within analytical techniques. The first question is investigated through simulations, in which the true variance weight is known and a relationship is sought between this weight and the optimal power estimated. The second question is investigated through extensive data analysis of various analyses across analytical techniques. The results are tabulated and plotted.

2. Methodology

2.1. A review of the power model

Given standard responses (Y) and concentrations (X), the power model transforms Y and X according to the transformation equation below and fits a linear regression between the transformed data according to the linear regression equation also given below.

$$y^{(\lambda)} = \begin{cases} \dfrac{y^{\lambda}-1}{\lambda \dot{y}^{\lambda-1}}, & \text{if } \lambda \neq 0 \\ \dot{y} \;\; \ln y, & \text{if } \lambda = 0 \end{cases} \quad \text{(Transformation equation)}$$

$$y^{(\lambda)} = a + bx^{(\lambda)} + \epsilon \quad \text{(Regression equation)}$$

where λ is a power value and $\dot{y}$ is the geometric mean of the responses. This power value can be any real number. The optimal power for a given dataset is determined as the value at which the sum of scaled squared deviations is minimised. This value is obtained by searching empirically within some interval called the power search range using the scaled residual sum of squares as the selection criteria and an appropriate numerical algorithm with a preset precision. The numerical algorithm that was used (Appendix A) allows one to start with as wide an interval as possible and converges to the right value quickly with a high precision.

After the optimal power is obtained, outlying standards are flagged using the 20-20 Jack%Dev Rule [2]. Using this rule, a standard is an outlier if its jacknife percent deviation exceeds 20% and omitted from the calibration curve if the total number of outliers is, at most, 20% of the original number of standards. The curve is then cross-validated using six (6) Quality Control (QC) samples, with duplicate values at low, medium and high concentrations. The curve is accepted if the back-calculated values of the QCs pass the 20-15-10: 4/6 rule [5]. Thus, at least one of the percent deviations of the QCs at low, medium, high concentration is less than or equal to 20, 15, and 10%, respectively, and at least 4 out of the 6 QCs passed.

2.2. Simulation

The primary objective of the simulation was to establish the relationship between observed optimal powers from the proposed method and the true variance weight, hence outlier detection or cross-validation using QC samples was not performed at this stage. As well, the relationship between the currently used regression models and the true underlying model was investigated. To do this, standard curve data were simulated according to the Simulation Al gorithm in the Appendix B. The variance of the data was proportional to $\text{Concentration}^{\text{Weight}}$. Thus the term "true variance weight" is a surrogate for the weight value in $\text{Concentration}^{\text{Weight}}$ and the underlying variance was controlled by simulating data with weight = 0 (linear, un-weighted), 0.25, 0.5, 0.75, 1.00 (linear, weighted 1/Concentration), 1.25, 1.50, 1.75, 2.00 (linear, weighted $1/\text{Concentration}^2$).

Once the data were simulated, the linear regression equation was fit to the data. The power search range was (− 16, 16) with a precision of 0.0001, thus the optimal power obtained was correct to four decimal places. The numerical al-

Table 1
Median R^2 values over 500 simulations in each weight parameter and selection frequency of regression (Fq)

Weight	0		0.25		0.5		0.75		1		1.25		1.5		1.75		2	
Regress	R^2	Fq	R^2	Fq	R^2	Fq	R^2	Fq	R^2	Fq	R^2	Fq	R^2	Fq	R^2	Fq	R^2	Fq
Lin	0.9993	287	0.9987	280	0.9964	255	0.9925	198	0.9854	110	0.9590	45	0.9190	16	0.8621	4	0.7797	3
Lin $1/X$	0.9941	1	0.9938	8	0.9916	13	0.9891	51	0.9864	112	0.9703	116	0.9460	74	0.9127	43	0.8606	17
Lin $1/X^2$	0.8624	0	0.8532	0	0.8505	0	0.8587	0	0.8804	0	0.8713	0	0.8567	0	0.8443	1	0.8161	6
Quad	0.9993	210	0.9987	208	0.9965	212	0.9929	188	0.9863	150	0.9612	81	0.9220	51	0.8710	24	0.7924	20
Quad $1/X$	0.9936	2	0.9933	2	0.9912	10	0.9890	30	0.9866	38	0.9705	61	0.9475	50	0.9161	27	0.8624	18
Quad $1/X^2$	0.8528	0	0.8421	0	0.8413	0	0.8540	0	0.8737	1	0.8669	0	0.8555	0	0.8459	1	0.8143	3
Log-log	0.9673	0	0.9636	0	0.9605	4	0.9617	7	0.9651	38	0.9649	91	0.9499	145	0.9416	182	0.9213	253
Wagner	0.9699	0	0.9657	2	0.9646	7	0.9660	26	0.9677	51	0.9665	106	0.9521	165	0.9445	218	0.9235	180
Power	0.9993		0.9987		0.9966		0.9940		0.9912		0.9821		0.9665		0.9559		0.9381	

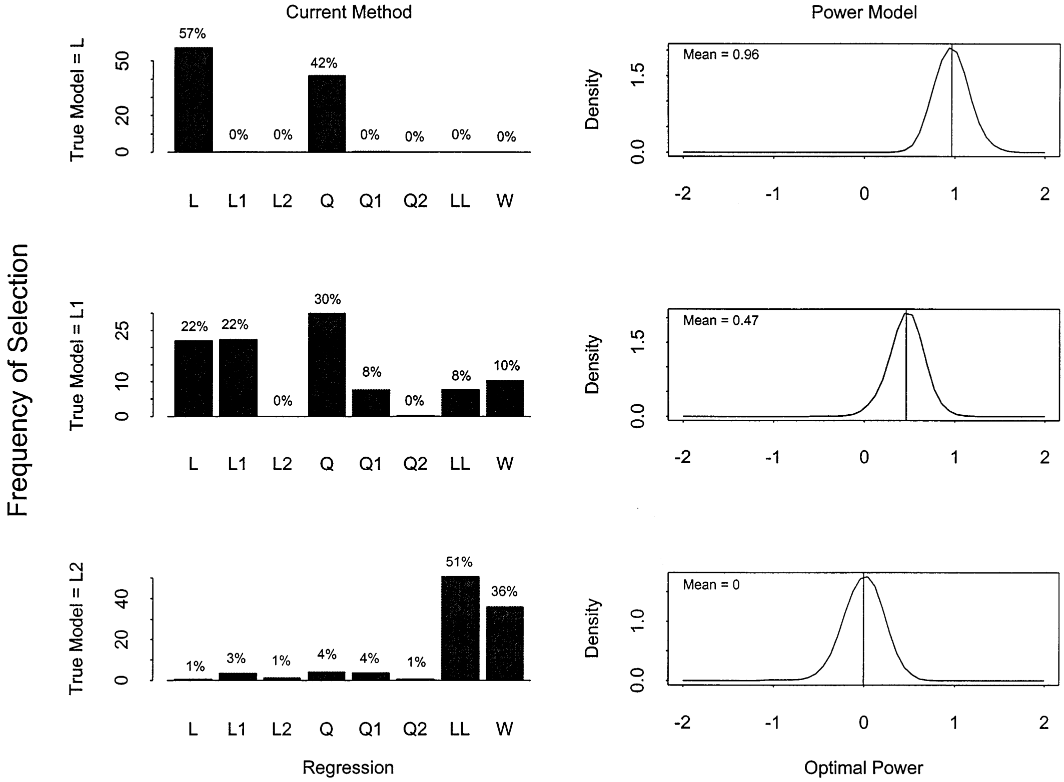

Fig. 1. Curve selection density distribution using the power model and using the current method (L = Linear, Q = Quadratic, LL = Log-Log, W = Wagner, 1 = Weighted $1/X$, 2 = Weighted $1/X^2$).

gorithm used is given in Appendix A. For each set of simulation parameters, and in particular, for each true variance weight, 500 datasets were simulated. For each of these datasets, an optimal power and corresponding R^2 value were obtained, and the mean optimal power and R^2 were tabulated. As well, the current procedure was used to select a regression model. The frequency distribution of regression models over the 500 cases was tabulated. The results are given in Table 1 and Fig. 1 and Fig. 2.

2.3. Data analysis

To analyse the real data, the complete power model was fit, as reviewed in Section 2.1. Thus, the optimal power was obtained, outlying standards flagged according to the 20-20 Jack%Dev Rule and the curve cross-validated using the QC acceptance rule.

A total of 2087 standard curve data from 91 projects from four different analytical techniques, High Pressure Liquid Chromatography (HPLC), Gas Chromatography (GC), Gas Chromatography with Mass Spectometry (GCMS), Liquid Chromatography with Mass Spectrometry (LCMS), were analysed. The proposed power method was used as well as the current method. In the current method, a regression type was selected from a set of eight (8), these being: Linear, Linear-Weighted $1/X$, Linear-Weighted $1/X^2$, Quadratic, Quadratic-Weighted $1/X$, Quadratic-Weighted $1/X^2$, Log-Log and Wagner using the R^2 as a selection criteria. The Log-Log model is a linear regression between log(Response) and log(Concentration). The Wagner is a log-log with a

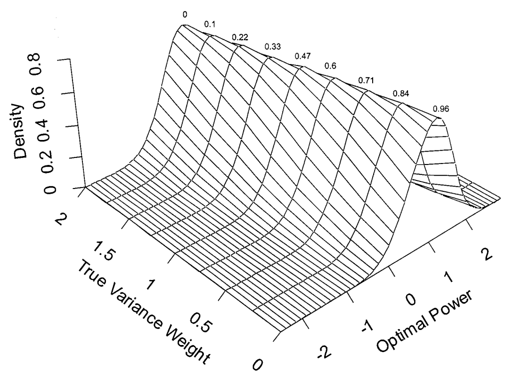

Fig. 2. A perspective plot of the density distribution of optimal power as a function of the true variance weight.

quadratic term. The results are given in Table 2 and Table 3 and Figs. 3–8.

3. Results

All the R^2 values were adjusted for the number of parameters in the regression.

3.1. Simulation results

Fig. 1 shows a histogram of the frequency of selection of the eight (8) regression-models (using the current method) when the underlying true model is Linear Unweighted or Weighted $1/X$ or $1/X^2$. In this plot, L = Linear, L1 = Linear-Weighted $1/X$, L2 = Linear Weighted $1/X2$, Q = Quadratic, Q1 = Quadratic-Weighted $1/X$, Q2 = Quadratic-Weighted i $1/X2$, LL = Log-Log and W = Wagner. The density plot of the estimated optimal powers is also given.

When the true model is an unweighted linear, the most frequent regression model picked by the current method using R^2 (as a criterion) is Linear with a frequency of 57%. However, the Quadratic is also picked with a similar frequency of 42%. Thus, unless a test for the significance of the quadratic term is included in the model selection, there will be ambiguity of choice between the two models. To contrast this, the power model gives optimal powers that appear normally distributed around the true value of 1 (mean = 0.96, 90% confidence interval-(0.70,1.24)).

When the true model is a linear-weighted $1/X$, there is a distribution of choice between the regression models with the Quadratic being the most frequently selected and the Linear $1/X$ having a similar selection frequency as the Linear. The R^2 choice is even more ambiguous in this case. On the other hand, the optimal powers are normally distributed around 0.47, with a 90% confidence interval of (0.15, 0.72).

When the true model is a Linear weighted $1/X^2$, the true regression is picked only 1% of the time. The most frequent models in this case are the Log-log and the Wagner. Again without a significance test for the quadratic term, it is difflcult to evaluate the advantage of the Wagner over the Log-log. The optimal powers are normally distributed around the true value of 0 (mean = 0.00, 90% confidence interval = (− 0.34, 0.32)).

Table 1 summarises the mean R^2 and the selection frequencies of the regression models and the Power model for all the simulated true models. In each case, the mean R^2 from the power model is at least as good as and often higher than that from the corresponding regression model even

Table 2
A summary of the optimal powers in each analytical techniques

Analytical technique	Projects	Runs	Optimal Power				
			1st Qu.	Median	Mean	3rd Qu.	IQR
GC	23	386	−0.08	0.04	0.03	0.17	0.25
GCMS	22	577	−0.11	0.05	0.05	0.21	0.32
HPLC	22	663	−0.09	0.08	0.07	0.27	0.36
SCIEX	24	461	−0.11	0.04	0.01	0.17	0.28
Total	91	2087	−0.10	0.05	0.04	0.21	0.31

Table 3
A summary of optimal powers in each group of runs in which the current method selected a given regression type

Regression	Runs	Optimal Power				
		1st Qu.	Median	Mean	3rd Qu.	IQR
Linear	1270	−0.11	0.07	0.06	0.24	0 35
Quadratic	107	−0.12	0.03	0.03	0.18	0.31
Wagner	710	−0.08	0.04	0.02	0.16	0.24
Total	2087	−0.10	0.05	0.04	0.21	0.31

though the R^2 is the selection criteria in the latter procedure and not in the former.

A perspective plot of the probability density plots of the optimal powers for all the underlying true models is shown in Fig. 2. This plot demonstrates firstly that the estimated optimal powers are sharply normally dirtributed around the true power and that this true power is distinct and linearly related to the true variance weight. From the values obtained the relationship between variance Weight and true Optimal Power is estimated as

$$\text{Weight} = 2.05 \times (0.96 - \text{Optimal power})$$

(Specificity equation)

The specificity of the power model is thus illustrated.

These results show that the power model is highly specific and gives better fits than the current model. They also show that the most frequently selected regression model using R^2 is not necessarily the true model.

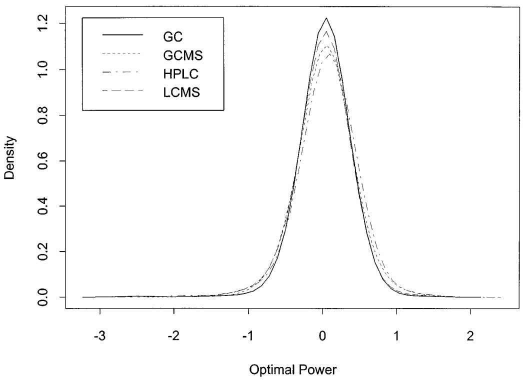

Fig. 3. Probability density distribution of optimal power in each analytical technique.

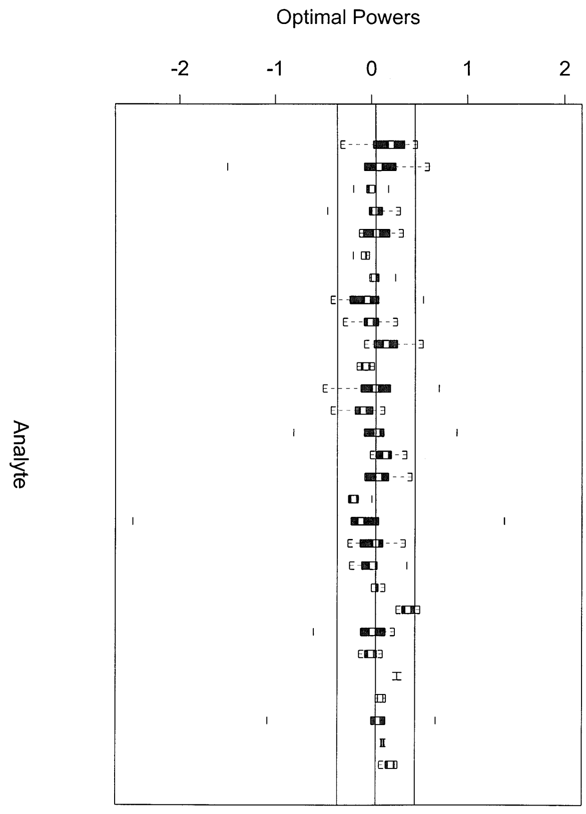

Fig. 4. Boxplot of optimal power models for each analyte in the GC analytical technique. Horizontal lines correspond to 5th, 50th and 95th percentiles, the median is 0.044.

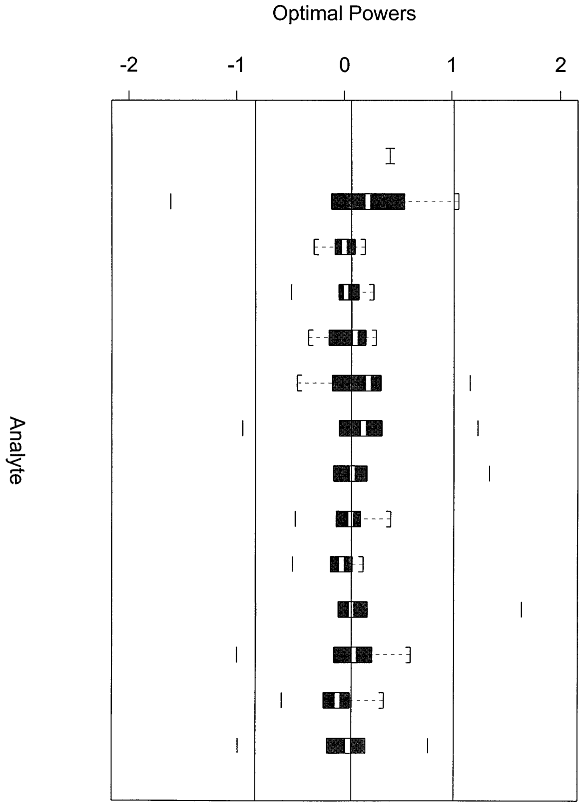

Fig. 5. Boxplot of optimal power models for each analyte in the GCMS analytical technique. Horizontal lines correspond to 5th, 50th and 95th percentiles, the median is 0.066.

4. Data analysis results

4.1. Variability between analytical techniques

Table 2 gives a summary of the optimal power in each of the analytical techniques. In all four techniques, the medians and means of these values are all around 0.1, when rounded to 1 dp (decimal place). As well, the interquartile ranges are similar (0.3, correct to 1dp). From the Specificity Equation above, this suggests that firstly the true variance weights from the data across the four techniques vary normally and equally around $1/X^2$. The suggested normal distribution is even more apparent in Fig. 3 which shows the density plots of the optimal powers from the four techniques.

4.2. Variability between analytes within techiniques

Figs. 4–7 contains box plots of the analyte-within analytical technique optimal powers. These are all evenly scattered around the analytical technique-median values implying a stable variation within each technique.

4.3. Variability between typical regression types

Table 3 is a summary of optimal powers in each of the regression types that are typically used for fitting calibration curves, such as Linear, Linear, weighted $1/X$ or $1/X^2$, Quadratic with weights $1/X$ or $1/X^2$, Log-log and Wagner. In this table, Linear (or Quadratic) includes weighted $1/X$ and

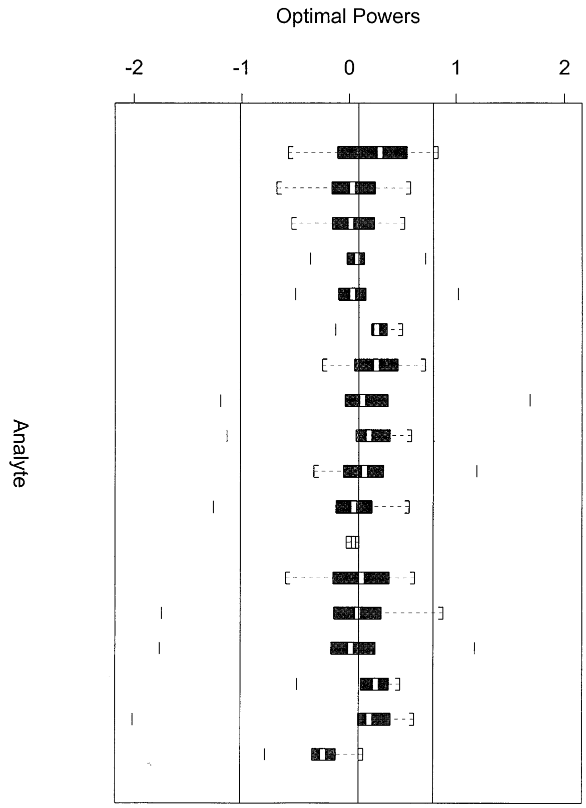

Fig. 6. Boxplot of optimal power models for each analyse in the HPLC analytical technique. Horizontal lines correspond to 5th, 50th and 95th percentiles, the median is 0.09.

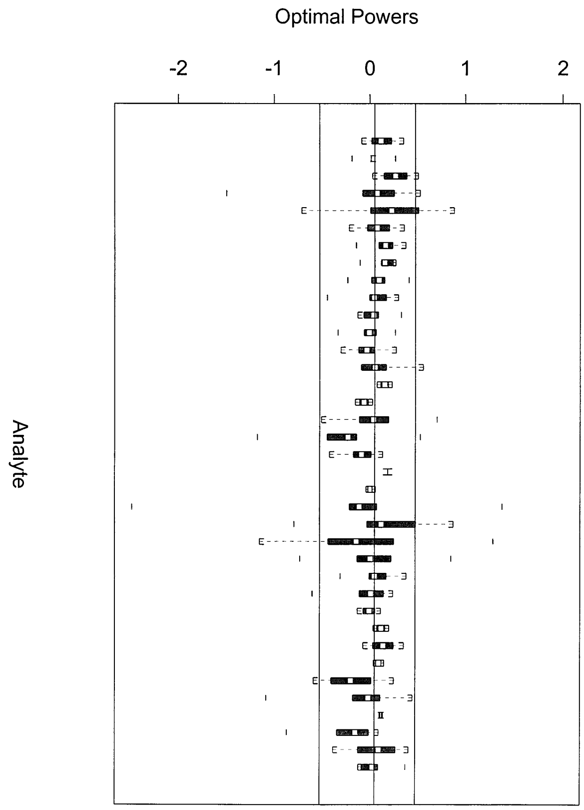

Fig. 7. Boxplot of optimal power models for each analyse in the LCMS analytical technique. Horizontal lines correspond to 5th, 50th and 95th percentiles, the median is 0.051.

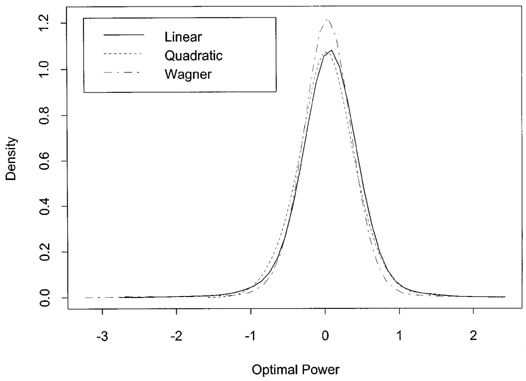

Fig. 8. Probability density distribution of optimal power by regression Curve.

weighted $1/X^2$. The optimal power values in the Quadratic and Wagner regression types are slightly smaller suggesting a slightly bigger variance weight in the data due to these regression types. However, all the median and interquartile range seem to be stable around 0 again implying an underlying true weight of $1/X^2$. Fig. 8 shows density plots of the optimal powers in each of the regression types and demonstrates more clearly the stability of distribution.

5. Discussion of results and conclusion

The results indicate that the optimal power transformation is robust between and within analytical techniques and between regression types. The currently observed variability in the choice of regression may therefore be more due to other sources of variability (for example the ambiguity of R^2) than the inherent true transformation. Further, the apparent robustness of the proposed power method implies that the power transformation can be used without imposing bounds on the expected power value.

In conclusion, it has been demonstrated that the power approach provides objective criteria for fitting calibration curves. The robustness of the power transformation method has been demonstrated and suggests that the currently observed variability between regression types is not necessarily due to inherently different regressions but may be due to other sources of variability between and within analytical techniques. The power model has also been demonstrated to be highly specific to the underlying true variance weight through the specificity equation.

Acknowledgements

The authors would like to acknowledge the contributions from bioanalytical departments at Phoenix International Life Sciences without whose full cooperation this project would not have been performed.

Appendix A. Power Search Algorithm

The algorithm which was used to find the value of the optimal power is given below.

1. Set the pivot $\lambda_0 = 0$ and evaluate its sum of scaled squared deviation, SSR_0.
2. Set the search range and precision. This is done by assigning a value to the parameter step and setting the maximum number of iterations, n. Step = 8 and $n = 16$ coresponds to a range of $(-16,16)$ and precision of 0.0001.
3. Set $\lambda_{-1} = \lambda_0$-Step and $\lambda_1 = \lambda_0 +$ Step.
4. Evaluate sum of scaled squared deviation for λ_{-1}, and λ_1 (SSR_{-1} and SSR_1).
5. Set λ_0 and SSR_0 to be the one which minimizes the sum of scaled squared deviations from λ_{-1}, λ_0 and λ_1.
6. If $i = n$ return λ_0 and SSR_0, the optimal power and its corresponding sum of scaled squared deviation.
7. Set $i = i + 1$, Step = Step/2 and goto 3.

Appendix B. Simulation algorithm

In this section the algorithm which was used to generate values of drug concentration and responses is presented.

B.1. Simulate Concentration

- Set a maximum concentration value MaxConc as a uniform random number between 100 and 1000 (this can be scaled up or down)
- Choose 10 concentration values as follows: (1, 1, 2, 5, 15, 50, 75, 90, 100, and 100%) of MaxConc. The LLOQ as well as the maximum value are duplicated as is often done in practice.

.

B.2. Simulate Regression Slope

- Generate the maximum response, MaxResp, from a log Normal distribution such that mean (MaxResp) = 1 and probability of MaxResp > 4 is 0.001. To do this, a log-normal distribution with mean = -0.1186 and S.D. = 0.4870 was used.
- Generate Intercept as a uniform variate between -0.01 and 0.01.
- Then the slope is given by:

$$\text{Slope} = (\text{MaxResp} - \text{Intercept})/\text{MaxConc}$$

B.3. Simulate Responses

$$\text{Response} = \text{Intercept} + \text{Slope} \times \text{Concentration} + (\text{Concentration}/10)^{\text{wgt}} \times \sigma^2$$

where σ^2 is a random uniform variate between 0.0001 and 0.0003 and the simulated weight is wgt.

References

[1] G.E.P. Box, D.R. Cox, An analysis of transformations, J. Royal Stat. Soc. B26 (1964) 211–252.

[2] Kimanani, E. Bioanalytical Calibration curves: Proposal for statistical criteria, J. Pharm. Biomed., (1998) 1117–1124.

[3] E. Kimanani, M. Mihailovici, J. Gaudreault, J. Lavigne, R. Lalonde, A statistical procedure to fit and detect outliers in calibration curves, Pharm. Sci. 13 (9) (1996) 22.

[4] E. Kimanani, M. Mihailovici, J. Lavigne, J. Gaudreault, Bioanalytical Calibration curves, Variability of optimal powers across different aanalytical techniques and apnalytes, J. Pharm. Sci. 13 (9) (199) 22.

[5] R.O. Kringle, An assessment of the 4-6-20 rule for acceptance of analytical runs in bioavailability, bioequivalence and pharamacokinetic studies, Pharm. Res. 11 (4) (1994) 556–560.