Journal of Pharmaceutical & Biomedical Analysis
Vol. 2, Nos 3/4, pp. 395–402, 1984
Printed in Great Britain

0731–7085/84 $3.00 + 0.00
© 1984 Pergamon Press Ltd.

General approach to handling nonuniform variance in assay calibration

LEON AARONS

Pharmacy Department, University of Manchester, Manchester M13 9PL, UK

Abstract: A general method for handling nonuniform variance in data from assay calibrations is discussed. Calibration data from the analysis of ibuprofen and aspirin by high-performance liquid chromatography was analysed by the traditional least squares method; nonuniform error variance was found to be significant. Weighted least squares analysis overcomes the problem of nonuniform variance but relies on good estimates of the error variance. The method of extended least squares, a maximum likelihood method, is described which incorporates handling of the weighting in the regression analysis. The extended least squares method produces accurate and precise estimates of the parameters of the calibration and allows precise estimates of the variance of future predictions, provided that a sufficient number of calibrators are used.

Keywords: *Nonuniform variance; assay calibration; linear least squares; extended least squares.*

Introduction

As most assay calibrations are linear the usual method of determining the 'best' calibration line is the method of ordinary linear least squares (OLS). In the application of OLS it is assumed that the data are randomly and normally distributed about the fitted line and that the error variance is constant. If the variance is not constant, OLS will be inefficient since the precision of the estimates of the slope and intercept will be poor although, on average, these estimates will be unbiased [1].

Weighted least squares (WLS), in which the squared deviation of each datum point from the fitted line is weighted by the reciprocal of the variance of that datum point, is the appropriate method to use when the variance in the data is not constant. There are two problems associated with WLS. First, the variance in the data is not generally known and, even when as many as five or six replicates are made, the estimate of the variance is poor. Second, the variance, if determined, is only known for the standards used in the calibration; consequently, variances at intermediate values have to be interpolated. The purpose of the present study was to investigate a more systematic approach to handling nonuniform variance (presented in part at the British Pharmaceutical Conference, Edinburgh, 1982 [2]). The results have been applied to calibration data for ibuprofen and aspirin obtained by high-performance liquid chromatography (HPLC).

Experimental Methods and Results

The ibuprofen assay was a reversed-phase HPLC method. The mobile phase was pumped through a 100 × 4.6-mm i.d. stainless-steel column packed with 5-μm Spherisorb ODS (Phase Separations Ltd, UK) with a Waters M6000A pump. The Pye Unicam LC3 variable wavelength detector was set at 240 nm. The mobile phase was propan-2-ol–methanol–0.1% v/v acetic acid (40:5:55 v/v/v). All reagents were of AnalaR grade.

Calibration standards (n = 30) containing ibuprofen at concentrations of 5–45 mg/l were prepared in 0.05 M phosphate buffer (pH 7.4). A 0.5-ml quantity of methanol containing phenylbutazone (internal standard; 20 mg/l) was added to 0.5 ml standard solution and, after mixing, a 20-μl aliquot was injected on to the column using a Rheodyne model 7125 valve injector. Calibration graphs of the peak–height ratio of ibuprofen to phenylbutazone against ibuprofen concentration were linear over the range of concentrations studied (r = 0.9909; n = 30). All replicate standards (m = 6) were independently prepared.

The aspirin assay was also a reversed-phase HPLC method using a 250 × 4.6 mm i.d. ODS column (Zorbax, Du Pont). The model 440 Waters UV detector was set at 229 nm. The mobile phase was methanol–0.1 M phosphate buffer (pH 2.7) (52:48 v/v); flow rate, 1 ml/min.

Calibration standards (n = 24) containing aspirin at concentrations of 0.05–10 mg/l were prepared in human plasma. To 1 ml of sample was added the internal standard solution (4-methoxybenzoic acid to give a concentration of 5 mg/l) and 0.5 ml of sulphuric acid (0.5 M); the mixture was extracted with 10 ml diethyl ether–hexane (1:1 v/v) and back-extracted into 0.05 M phosphate buffer (pH 7.4). The pH of the aqueous extract was adjusted to about pH 2 to minimize aspirin hydrolysis, and 100 μl of the resulting solution (0.5 ml) was injected on to the column. All replicate standards (m = 3) were independently prepared and the resultant calibration graph of peak–height ratio against concentration was linear over the concentration range studied (r = 0.9986; n = 24).

Data Analysis

Background

As noted in the Introduction, the major problem with WLS concerns the proper assignment of the weights. To overcome this problem it is possible to fit the experimentally determined variances to a model and to calculate weights for WLS using this model. However, this method would require two regressions to be performed, one for the variance model and one for the calibration; the method could be applied only when the experimental variances were known.

A more systematic approach to the problem is to use the method of extended least squares (ELS) suggested by Sheiner and Beal [3]. ELS is a maximum likelihood method which attempts to determine both a structural model and a variance model. The objective function that is minimized in order to obtain the parameters, c, of the calibration is

$$\Theta\,(c,\xi,y) = \sum_{i=1}^{N} \{\frac{[y_i - f(c,x_i)]^2}{v(c,\xi,y_i)} + \ln\,[v(c,\xi,y_i)]\}. \qquad (1)$$

Minimizing equation (1) is equivalent to maximizing the likelihood with a normal distribution of errors [4]. The first term in equation (1) resembles a weighted sum of squares. The term $f(c,x_i)$ is the structural model which in the present work is a linear function:

$$y = c_0 + c_1 x \tag{2}$$

where y is the peak–height ratio at concentration x. The function $v(c,\xi,y)$ is the variance model which contains additional parameters, ξ. The purpose of modelling the variance is simply to account for changes in the variance. The variance parameters are of no interest in themselves and the function $v(c,\xi,y)$ should be chosen to be the simplest possible function that leads to accurate and precise estimates of the structural parameters, c. It should be noted that if equation (1) were minimized without the log term, this would force the variance to be infinite and hence indeterminate.

Maximum likelihood estimators are known to be biased for small sample sizes and this bias may obviate any gain in efficiency. Raab [5] examined the problem of estimating a variance function for immunoassay data and found that unbiased and efficient estimates could be obtained by the use of a modified likelihood function. In order to assess the consistency and efficiency of the ELS estimates of the structural parameters, simulation studies were performed.

Four experimental designs were used:

design a x = 0.2, 0.4, 0.6, 0.8, 1.0
design b x = 0.2, 0.2, 1.0, 1.0
design c x = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0
design d x = 0.1, 0.1, 0.1, 0.1, 0.1, 1.0, 1.0, 1.0, 1.0, 1.0.

For each design exact data were generated using the relationship $y = x$ and random noise was added. The random variation was taken from a normal distribution of mean equal to zero and variance equal to $(0.01x^2 + 0.0001)$, which corresponds to a RSD of about 10%. The constant term, 0.0001, can be viewed as a constant background noise from the detector.

Two variance models were investigated:

$$v = \xi_0 \exp(\xi_1 \hat{y}) \tag{3}$$

$$v = \xi_0 + \xi_1 (\hat{y})\xi_2, \tag{4}$$

where $\hat{y}$ is the predicted value of the peak–height ratio [obtained from equation (2)] rather than the experimental value. The two models gave virtually identical results and only those for equation (3) are presented here.

Equation (1) was minimized using a modified Newton method [6]. Asymptotic standard errors were obtained from an approximate covariance matrix calculated using a method devised by Gauss [7].* 100 simulations were performed for each design. A summary of the results of these studies is presented in Table 1.

The structural parameters, c_0, c_1, were accurately determined in all four designs.

* A FORTRAN 77 listing of the program is available from the author upon request.

Table 1
Results of simulation studies

Design		$\hat{c}_0$†	$\hat{c}_1$	$\hat{se}(c_0)$‡	$\hat{se}(c_1)$	Sensitivity§ ‖OLS	‖ELS
a	Experimental*	−0.009 (0.032)	1.028 (0.110)	0.021	0.053	0.353	0.102
	Theoretical	0	1	0.033	0.086	0.058	
b	Experimental	−0.002 (0.025)	1.013 (0.103)	0.015	0.053	0.358	0.073
	Theoretical	0	1	0.026	0.091	0.058	
c	Experimental	0.001 (0.013)	1.002 (0.045)	0.015	0.057	0.338	0.098
	Theoretical	0	1	0.015	0.051	0.058	
d	Experimental	−0.002 (0.009)	1.006 (0.046)	0.0076	0.046	0.393	0.062
	Theoretical	0	1	0.0086	0.051	0.058	

* Experimental results are the means of 100 simulations.
† $\hat{c}_0$, $\hat{c}_1$ are the estimated intercept and slope; figures in brackets are S.D.s.
‡ $\hat{se}(c_0)$, $\hat{se}(c_1)$ are the estimated standard errors of the intercept and slope.
§ Sensitivity is defined as the concentration that gives rise to a 20% RSD.
‖ OLS = ordinary least squares; ELS = extended least squares.

However, as expected, the standard errors of these parameters were underestimated by the ELS method. The theoretical values for the standard errors were obtained from the exact covariance matrix, calculated using the known variance–covariance matrix of the errors [8]. It can be seen that the problem is worse for designs a and b which have fewer data points. With 10 data points the standard error estimates are relatively close to the theoretical values. It is also worth noting that the 'optimal' designs, b and d, produce more precise estimates of c_0.

Perhaps more important than the structural parameters themselves is the precision and accuracy with which the calibration can be used for future predictions. Of particular interest is the assay sensitivity as determined from the calibration.

Following Liteanu and Rica [9] the assay sensitivity is defined as that concentration which gives rise to a particular RSD of prediction, chosen in this instance to be 20%. The variance of a prediction, x_0, from the calibration plot is given by equation (5):

$$v(x_0) = V\left[\frac{(y_0 - \bar{y})}{c_1}\right], \tag{5}$$

where $\bar{y}$ is the weighted mean for the calibration standards and y_0 is the observed peak–height ratio [10] (see also Oppenheimer *et al.* [11]). The RSD will be 20% when

$$\frac{\sqrt{[V(x_0)]}}{x_0} = 0.2. \tag{6}$$

Equations (5) and (6) are solved simultaneously to calculate x_0.

The ELS method gives good estimates of the assay sensitivity particularly when compared with the OLS estimator (Table 1). Better prediction is obtained with 10 calibrators and with the optimal designs b and d.

Application to ibuprofen

The results of reproducibility studies are presented in Table 2. The variance in peak–height ratio at a concentration of 45 mg/l is about 25 times greater than at 5 mg/l. The data were analysed using the method of OLS. An inspection of the resultant residual plot, which represents the difference between observed and calculated values, clearly illustrates the nonuniform character of the error variance (Fig. 1a). This trend was confirmed using the F_k-test [12]:

$$F_k = \frac{\text{SS last } k \text{ residuals}}{\text{SS first } k \text{ residuals}},$$

where the numerator is the sum of squares of the k residuals at the highest concentration, and the denominator is the corresponding sum of squares at the lowest concentration. The most efficient value of k for this data set is six and the calculated value of F_6 (26.4) is much greater than the critical value (8.47) at a confidence level of 99%. Hence the variance in peak–height ratio increases significantly with ibuprofen concentration.

Consequently, the data were re-analysed using the method of WLS, in which each squared residual was weighted by the reciprocal of the estimated variance (Table 2). The parameter estimates and their standard errors are presented in Table 3. The OLS and WLS estimates of the slope, c_1, are virtually identical. However, the precision of the

Table 2
Reproducibility studies for ibuprofen assay

Concentration (mg/l)	5.1	10.1	20.2	30.2	46.5
Mean peak–height ratio (n = 6)	0.1083	0.1977	0.4278	0.5916	0.9050
Variance	0.00018	0.00023	0.00050	0.0021	0.0049
RSD	13.4	7.9	5.0	7.8	8.5

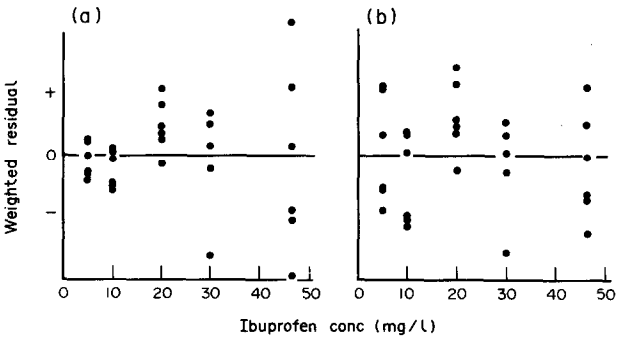

Figure 1
Residual plots for ibuprofen calibration by OLS analysis and ELS analysis. The weighted residual is equal to the residual divided by the square root of the estimated variance of the data: $[y_i - f(c,x_i)]/\sqrt{[v(c,\xi,y_i)]}$. For OLS the weighted residual is the same as the residual.

Table 3
Analysis of ibuprofen calibration data

	Method*		
	OLS	WLS	ELS
c_0†	0.0141 (0.0138)	0.0057 (0.0067)	0.0072 (0.0062)
c_1‡	0.0193 (0.0005)	0.0199 (0.0005)	0.0197 (0.0005)

* OLS = ordinary least squares; WLS = weighted least squares; ELS = extended least squares.
† c_0 = intercept. ‡c_1 = slope; figures in brackets are standard errors.

OLS estimate of the intercept, c_0, is significantly worse than the WLS estimate. This is due to the failure of the OLS method to account for the reduction in error variance at low concentrations. The lack of precision in the OLS calibration is more evident in the calculation of assay sensitivity. The calculated sensitivity, 11.1 mg/l, is clearly an overestimate, given that the RSD in the peak–height ratio at an ibuprofen concentration of 5 mg/l is 13.4%. It is not possible to compute the assay sensitivity (as defined above) for the WLS method, since the variance is only known for the standards used in the calibration.

The data were next analysed using the ELS method [variance model: equation (4)]. Estimates of c_0 and c_1 (Table 3) were very similar to the OLS and WLS results. Comparison of the residual distributions from the OLS and ELS analyses (Fig. 1) shows that the ELS distribution is much more uniform. More significantly, the ELS estimate of sensitivity, 3.4 mg/l, is much more consistent with the author's knowledge of the assay. The improvement in prediction obtained with ELS over OLS is shown in Fig. 2, where the RSD of prediction is plotted against concentration for the two methods. The line at 20% determines the assay sensitivity.

Application to aspirin

The aspirin data were analysed using the OLS method and a trend can be seen in the resulting residual plot (Fig. 3). However, the trend is different from that seen in the ibuprofen data. The data seem to break into two distinct regions. Above an aspirin concentration of approximately 2 mg/l the error variance is much greater than below this concentration ($F_6 = 30.8$). Within each region the variance is approximately constant. There is no apparent reason for this 'break'. A change in the detector scale or a change in volumetric equipment between the two regions could give rise to such a phenomenon.

The data were re-analysed using the ELS method with the following variance model:

$$v = \xi_1 [1 - H(x_1)] + \xi_2 H(x_1), \tag{7}$$

where $H(x_1) = 0$ for $x \leqslant x_1$, $H(x_1) = 1$ for $x > x_1$ and $\xi_2 > \xi_1$.

As with the ibuprofen calibration, the ELS estimates of the calibration parameters, $c_0 = 2.1 \times 10^{-8}$ and $c_1 = 0.771$, were not very different from the OLS estimates ($c_0 = 3.5 \times$

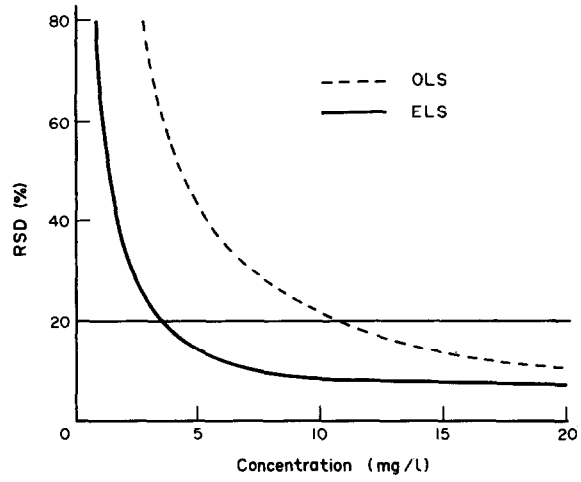

Figure 2
Calculated precision of prediction for ibuprofen calibration.

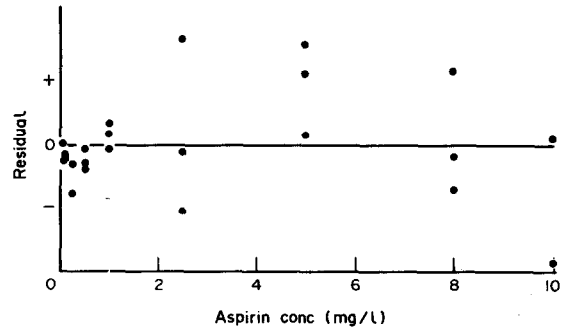

Figure 3
Residual plot for aspirin calibration: OLS analysis.

10^{-2} and $c_0 = 0.756$). However, the ELS estimate of sensitivity, 0.38 mg/l, was much smaller than the OLS estimate, 1.2 mg/l.

Discussion

A major advantage of OLS is its computational simplicity. Many small calculators are now pre-programmed to perform OLS analysis. Furthermore, even when the variance in the data is not uniform and particularly when the data are relatively precise as in the examples discussed here, the parameters may be well determined by OLS. However, OLS poorly estimates the variance associated with the fitted calibration line, particularly with the lower concentrations. Consequently, the OLS estimate of assay sensitivity is generally far too large. In theory WLS should overcome the shortcomings of OLS, but in practice it is difficult to obtain good estimates of the variance in the data. Moreover, it is not possible to estimate the variance at concentrations other than those of the standards used in the calibration and this makes the calculation of sensitivity difficult.

The ELS method resolves most of these problems. Weighting is automatically handled by the variance model and standards do not necessarily have to be replicated to determine the weighting function. However, it is advisable to conduct preliminary experiments to determine the nature of the variance model. Parameter bias can be reduced by using more calibrators. The simulations suggest that with 10 calibrators, accurate and precise estimates are produced of both the regression parameters and future predictions based on the calibration. The major disadvantage of the ELS method is that it requires much more computation than does OLS. However, with the proliferation of microcomputers and appropriate software this problem is becoming less important. It is now quite feasible to build ELS into a standard calibration software package.

Acknowledgements: I wish to thank Margaret Richards and David Leitch for performing the assays and supplying me with the experimental details of those assays. I would also like to thank Lewis Sheiner for supplying me with a copy of the program ELSFIT that was used to perform the ELS analysis and for many useful suggestions.

References

[1] J. S. Garden, D. G. Mitchell and W. N. Mills, *Anal. Chem.* **52**, 2310–2315 (1980).
[2] L. Aarons, *J. Pharm. Pharmacol.* **34**, 86P (1982).
[3] L. B. Sheiner and S. L. Beal, *J. Pharmacokin. Biopharm.* **8**, 553–571 (1980).
[4] Y. Bard, *Nonlinear Parameter Estimation*, pp. 61–71. Academic Press, New York (1974).
[5] G. M. Raab, *Appl. Statist.* **30**, 32–40 (1981).
[6] L. B. Sheiner, *ELSFIT User's Manual*. A Technical Report of the Division of Clinical Pharmacology, University of California, San Francisco (1981).
[7] Y. Bard, *Nonlinear Parameter Estimation*, pp. 96–99. Academic Press, New York (1974).
[8] N. R. Draper and H. Smith, *Applied Regression Analysis*, Chap. 2, pp. 70–140. John Wiley, New York (1966).
[9] C. Liteanu and I. Rica, *Statistical Theory and Methodology of Trace Analysis*, Chap. 8, pp. 342–344. Ellis Horwood, Chichester (1980).
[10] L. Aarons, *Analyst* **106**, 1249–1254 (1981).
[11] L. Oppenheimer, T. P. Capizzi, R. M. Weppelman and H. Mehta, *Anal. Chem.* **55**, 638–643 (1983).
[12] L. Endrenyi and H. F. Kwong, in *Kinetic Data Analysis* (L. Endrenyi, Ed.), pp. 89–103. Plenum Press, New York (1981).

[*First received for review 15 August 1983; revised manuscript received 28 October 1983*]