Pergamon

*Journal of Pharmaceutical & Biomedical Analysis*, Vol. 12, No. 4, pp. 523–531, 1994
Elsevier Science Ltd
Printed in Great Britain

# Development and application of a calibration regression routine in conjunction with linear and non-linear chromatographic detector responses

JOHN L. BURROWS* and KEN V. WATSON

*Bioanalytical Department, Hoechst Pharmaceutical Research Laboratories, Walton Manor, Walton, Milton Keynes, Bucks MK7 7AJ, UK*

**Abstract**: Linear regression is widely used in the calibration of chromatographic assays even in conjunction with some chromatographic detectors which show significant non-linearity in their response characteristics. A calibration routine, based upon the curve $y = ax\ln x + bx + c$ is presented which describes the non-linear behaviour of some chromatographic systems, including electron capture, nitrogen–phosphorus and UV photometric detectors, and gives comparable results to weighted linear regression with assays showing linear concentration versus response relationships. The ratio of the coefficients $a$ and $b$ in the equation allows quantification of the deviation from linearity and provides a more sensitive indicator of linearity than the correlation coefficient often quoted with linear regression.

**Keywords**: *Assay calibration; non-linear regression; estimation of curvature.*

## Introduction

The accuracy of the results from a chromatographic analysis is dependent upon the calibration curve adequately describing the actual concentration versus response characteristics of the assay. Regressions are widely used to define these calibration curves in mathematical terms now that the use of computer based integrators has become widespread and there is no longer any need to perform the associated calculations by hand. However, in any calibration there will be differences between the experimentally derived calibration function and the actual response which will lead to a concentration dependent bias in the assay.

One cause of these differences is due to the normal experimental errors related to the calibrators which give rise to an uncertainty in the coefficients of the regression line and so results in a degree of variability between individual calibration curves. This variability can be minimized by optional use of weighting factors as well as the number and distribution of calibrators and it is this aspect of calibration with linear regressions which has been concentrated on in the literature [1–7].

A second source of bias in an assay can be due to the mathematical function used to describe the calibration curve which may only be an approximation to the actual concentration versus response characteristics of the assay. Linear calibration functions are employed in the majority of chromatographic assays that appear in the literature. However, it has been reported that some degree of non-linearity in the sensitivity, i.e. response per unit concentration of analyte, is present in some detectors used in high-performance liquid chromatography (HPLC) and the non-linearity of the response from electron capture detectors (ECD) used in gas chromatography (GC) can be quite marked [8–10]. In such cases a concentration dependent bias will appear in the estimated values if a linear calibration is used.

Regressions based upon non-linear equations such as $y = ax^2 + bx + c$, $y = ax^n$, $\ln(y) = a[\ln(x)]^2 + b\ln(x) + c$ have appeared in the literature, as well as more complex polynomial expressions, but often lack critical evaluation [11–13]. In order to obtain optimum results the regression should not only match the actual shape of the response curve but also include appropriate weighting factors to compensate for the error distribution if it is non-homoscedastic. Non-linear calibration routines are often included as part of the capabilities of

* Author to whom correspondence should be addressed.

chromatographic data systems but may only be suitable for assays covering a small dynamic range and exhibiting a marked degree of curvature. As these regressions are often only available in their unweighted forms they usually impart no improvement in the accuracy of estimated data when compared with weighted linear regressions in assays covering wide dynamic ranges and showing only small deviations from linearity. Indeed it will be shown in this study that some compensation for non-linearity using the weighted linear regression, $y = bx + c$ is provided in the estimates of the coefficients $b$ and $c$.

Despite improvements made with the variable-frequency constant-current mode of operation non-linear response still remains a characteristic of the ECD used in GC [14, 15]. Figure 1 shows a plot of the detector sensitivity versus the logarithm of the concentration taken from an unpublished GC assay using ECD for a drug in plasma. It can be seen that there appears to be a linear correlation between the two parameters and similar results have been reported elsewhere [9]. A study concerning the GC ECD analysis of polychlorinated biphenyls investigated the suitability of various calibration functions and concluded that second-order curves ($y = ax^2 + bx + c$) fitted the data best [10]. However, the raw data also appears to exhibit a linear relationship of the form shown in Fig. 1.

Ultra-violet photometric detectors used in HPLC are generally considered to show excellent linearity of response versus concentration. However, a critical evaluation of such detectors revealed that the sensitivity can fall by as much as 5% over a 10-fold increase in analyte concentration though still be within the accepted criteria for linearity [8]. Again a linear relationship was observed between the sensitivity and the logarithm of the concentration.

Similar correlations have also been observed in this laboratory with assays involving either a UV detector in conjunction with HPLC or a nitrogen–phosphorus detector (NPD) and GC, except that in these cases the sensitivity increased with concentration as can be seen in the example of a GC assay in Fig. 2.

In all these examples the sensitivity (response/concentration, $y/_x$) can be defined in terms of the concentration ($x$) according to:

$$y/_x = a\ln x + b,$$

hence

$$y = ax\ln x + bx,$$

where $a$ and $b$ are constants. If allowance is also made for an intercept ($c$) at zero concentration the equation becomes:

$$y = ax\ln x + bx + c.$$

It is the evaluation of the equations relating to this calibration function in simulated and real analytical situations which is described in this communication.

## Equations Describing the Regression $y = ax\ln x + bx + c$

Estimates of the three coefficients $a$, $b$ and $c$ describing the calibration curve can be obtained using the following expressions:

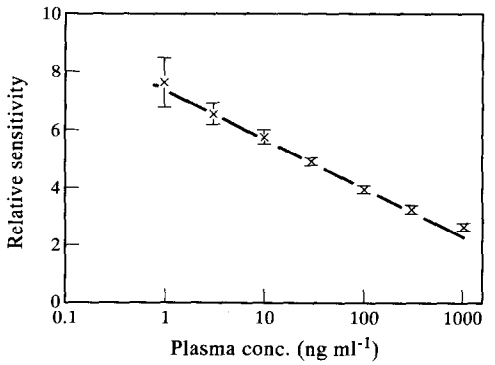


**Figure 1**
The change in sensitivity (normalized peak height ratio) of an ECD as a function of plasma drug concentration in a GC assay (mean of six calibration curves).

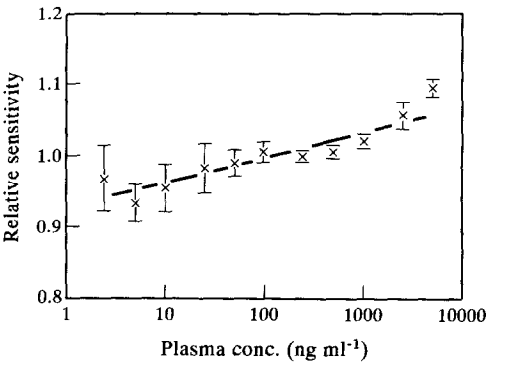


**Figure 2**
The change in sensitivity (normalized peak height ratio) of a nitrogen specific detector as a function of plasma drug concentration in a GC assay (intra-assay data of six determinations).

$$a = \frac{T1}{T4}, \; b = \frac{T2}{T4} \text{ and } c = \frac{T3}{T4},$$

where the terms $T1$–$T4$ are given by:

$T1 = \sum wxy\ln x(\sum wx)^2 + \sum wx^2\ln x\sum wxy\sum w + \sum wx\ln x\sum wx^2\sum wy - \sum wxy\ln x\sum w\sum wx^2 - \sum wx^2\ln x\sum wx\sum wy - \sum wx\ln x\sum wx\sum wxy;$

$T2 = (\sum wx\ln x)^2\sum wxy + \sum wx\sum wy\sum wx^2(\ln x)^2 + \sum wx^2\ln x\sum wxy\ln x\sum w - \sum wxy\sum w\sum wx^2(\ln x)^2 - \sum wx^2\ln x\sum wx\ln x\sum wy - \sum wx\ln x\sum wxy\ln x\sum wx;$

$T3 = (\sum wx^2\ln x)^2\sum wy + \sum wx^2\sum wx\ln x\sum wxy\ln x + \sum wx\sum wx^2(\ln x)^2\sum wxy - \sum wx^2\sum wy\sum wx^2(\ln x)^2 - \sum wx^2\ln x\sum wx\ln x\sum wxy - \sum wx\sum wx^2\ln x\sum wxy\ln x;$

$T4 = (\sum wx)^2\sum wx^2(\ln x)^2 + (\sum wx\ln x)^2\sum wx^2 + \sum w(\sum wx^2\ln x)^2 - \sum wx^2\ln x\sum wx\ln x\sum wx - \sum wx\sum wx^2\ln x\sum wx\ln x - \sum w\sum wx^2\sum wx^2(\ln x)^2.$

A weighting factor ($w$) was applied to the square of the residuals when deriving the equations above. In many assays, covering a wide dynamic range, the precision is proportional to the concentration ($x$) and in these cases it is appropriate to apply a weighting factor of $1/x^2$ [1]. Substituting this value into the equations introduces the term $n$ (number of calibrators) and simplifies the terms $T1$–$T4$ as follows:

$T1 = \sum(y\ln x)/x(\sum^1/x)^2 + \sum\ln x\sum^y/x\sum^1/x^2 + n\sum(\ln x)/x\sum^y/x^2 - n\sum(y\ln x)/x\sum^1/x^2 - \sum\ln x\sum^1/x\sum^y/x^2 - \sum(\ln x)/x\sum^1/x\sum^y/x;$

$T2 = [\sum(\ln x)/x]^2\sum^y/x + \sum^1/x\sum^y/x^2\sum(\ln x)^2 + \sum\ln x\sum(y\ln x)/x\sum^1/x^2 - \sum^y/x\sum^1/x^2\sum(\ln x)^2 - \sum\ln x\sum(\ln x)/x\sum^y/x^2 - \sum(\ln x)/x\sum(y\ln x)/x\sum^1/x;$

$T3 = (\sum\ln x)^2\sum^y/x^2 + n\sum(\ln x)/x\sum(y\ln x)/x + \sum^1/x\sum(\ln x)^2\sum^y/x - n\sum^y/x^2\sum(\ln x)^2 - \sum\ln x\sum(\ln x)/x\sum^y/x - \sum^1/x\sum\ln x\sum(y\ln x)/x;$

$T4 = (\sum^1/x)^2\sum(\ln x)^2 + n[\sum(\ln x)/x]^2 + \sum^1/x^2(\sum\ln x)^2 - \sum\ln x\sum(\ln x)/x\sum^1/x - \sum^1/x\sum\ln x\sum(\ln x)/x - n\sum^1/x^2\sum(\ln x)^2.$

It is in this form that the regression has been used and evaluated.

The slope of the line $y = ax\ln x + bx + c$ at any concentration $x$, i.e. the sensitivity, is given by:

$$\frac{dy}{dx} = a(1 + \ln x) + b.$$

Hence the change in sensitivity between two concentrations $x_1$ and $x_2$ is the difference in slope at these two concentrations. For convenience, if $x_2 = 10x_1$ and the result is expressed as a percentage of the underlying linear slope ($b$), the degree of non-linearity, or curvature, over a 10-fold increase in concentration can be described by:

$$\text{Curvature} = \frac{100}{b}[a(1 + \ln x_2) + b - a(1 + \ln x_1) - b]$$

$$= \frac{230a}{b}\,\%.$$

Once a calibration curve has been established it is normal analytical practice to estimate the concentrations of the unknowns ($x$), from the measured responses ($y$), by rearrangement of the regression equation. Unfortunately, there appears to be no direct transformation of the equation $y = ax\ln x + bx + c$ to enable direct calculation of $x$ given values of $y$. However, values can be obtained by the Newton–Raphson iteration procedure [6], described below, which rapidly converges to a consistent value after approximately three iterations.

Consider the equation $y = ax\ln x + bx + c$ and let the successive approximations to the solution be $x = Z^{[j]}$ (where $j = 0, 1, 2 \ldots$ and represents the number of approximations). The Newton–Raphson method allows the following estimate to be made for $Z^{[j+1]}$ from the previous estimated value $Z^{[j]}$:

$$Z^{[j+1]} = Z^{[j]} - \frac{f(Z^{[j]})}{f^1(Z^{[j]})},$$

where the functions of $Z$ are given by

$$f(Z^{[j]}) = aZ^{[j]}\ln Z^{[j]} + bZ^{[j]} + c - y$$

$$f^1(Z^{[j]}) = a(1 + \ln Z^{[j]}) + b.$$

An initial value $z^{[0]}$ of $y/b$ is used and the interaction is terminated when:

$$\frac{|Z^{[j+1]} - Z^{[j]}|}{|Z^{[j]}|} < 0.00001.$$

One problem which was identified with this approach is that, at concentrations approaching zero, negative values of $Z^{[j+]}$ can be generated which cause the calculation to fail on the subsequent pass because of the terms ln $Z^{[j]}$ in the equation. This was solved by substitution of small positive values when the estimates were negative. The iteration can therefore be used to determine estimates between zero and the lowest calibrator (limit of quantification) although such extrapolation can be subject to errors with any regression and values so obtained should not be used in pharmacokinetic calculations [17]. Negative values of $x$ cannot be determined with this equation but these have no real meaning within analytical chemistry and thus do not present a limitation on its use.

## Methods

*Simulations*

Computational errors were checked using two ideal concentration ($x$) versus response ($y$) curves. A simple linear relationship ($y = 10x + 1$) was used for the first and this was modified for the second so as to introduce a deviation in the sensitivity, over a 1000-fold concentration range, of ±10% in accordance with the model previously described. [$(2x\ln x)/(3\ln 10) + 9x + 1$, such that as $x \rightarrow 1$ the equation reduces to $y = 9x + 1$ and as $x \rightarrow 1000$ reduces to $y = 11x + 1$]. Twelve pairs of $x$, $y$ values corresponding to points distributed along the curves were entered into the regression equations described and estimates of $x$ back calculated using the iteration. Single precision (seven-digit) numbers were used throughout.

The calibration functions were tested further using Monte Carlo simulations [8] again using data based on the two ideal curves. This time random errors were introduced such that individual values of the response, $y_1$, at each concentration, $x$, were generated according to $y_i = y + Gs$, where $G$ is a gaussian or normally distributed random number (mean and standard deviation = 0 ± 1) and $s$ is the standard deviation of the measured response at a given concentration and was defined as $s = 0.17 + 0.03x$. This precision profile was chosen to represent an assay with an intra-run relative standard deviation (RSD) of about ±3% over most of its working range, increasing to ±20% at the limit of quantification, i.e. when $x = 1$. Ten simulated analytical runs were performed for each concentration versus response curve, each using 12 calibrators. The mean inter-assay calibration bias was determined over the assay range by comparison of the individual calibration curves with the actual response curves as previously described [8]. In addition, the mean calibration bias, defined as the mean of the moduli of the differences between the actual and estimated concentrations, was also calculated as this parameter gave a better estimate of the likely bias associated with an individual calibration curve. Simulations were also performed with the same raw data using weighted linear and quadratic regressions (weighting factor = $1/x^2$) for comparison.

## Real Analyses

Two unpublished assays of drugs in plasma were selected to test the calibration routine under real analytical conditions. A GC ECD assay with a marked degree of non-linearity (see Fig. 1) was chosen as the first, and an HPLC UV assay exhibiting a high degree of linearity of the response was chosen as the second. Fortified samples of blank human plasma were analysed in place of unknowns on six separate occasions for each assay and new calibration curves were generated for each run using sets of calibrators prepared independently from the fortified blanks. Estimates of the inter-assay precision and accuracy were made from the results obtained with the new regressions, weighted linear regression and weighted quadratic regression for comparison using the same raw chromatographic data, i.e. peak height ratios compared with an internal standard.

## Results and Discussion

No computational errors (<0.001%) were found in the estimates of the concentration obtained by back calculations using data points lying on the two curves in the simulation, demonstrating the validity of the iterative calculation even when wide dynamic ranges are involved, and the computation limited to single precision numbers.

The precision profile chosen for the Monte Carlo simulations was similar to that observed in many assays used in this laboratory, and the equation describing precision in terms of a limiting value plus a concentration dependent

**Table 1**
Monte Carlo simulations comparing three weighted regressions with the linear concentration versus response relationship $y = 10x + 1$ using 4 calibrators in triplicate* (10 simulations)

| | $y = bx + c$† | | $y = ax\ln x + bx + c$‡ | | $y = ax^2 + bx + c$§ | |
|---|---|---|---|---|---|---|
| Concentration ($x$) | Per cent inter-assay calibration bias (mean ± SD) | Per cent mean calibration error | Per cent inter-assay calibration bias (mean ± SD) | Per cent mean calibration error | Per cent mean calibration error | Per cent mean calibration error |
| 1 | −0.7 ± 8.6 | 7.1 | −0.8 ± 8.7 | 7.2 | −0.9 ± 8.7 | 7.3 |
| 2 | −0.4 ± 4.2 | 3.3 | −0.6 ± 3.9 | 3.0 | −0.5 ± 4.2 | 3.3 |
| 5 | −0.2 ± 1.8 | 1.5 | −0.4 ± 1.7 | 1.3 | −0.2 ± 1.8 | 1.4 |
| 10 | −0.1 ± 1.2 | 0.9 | −0.4 ± 1.5 | 1.2 | −0.2 ± 1.5 | 1.1 |
| 20 | −0.1 ± 1.1 | 0.8 | −0.3 ± 1.4 | 1.2 | −0.1 ± 1.5 | 1.1 |
| 50 | −0.1 ± 1.1 | 0.8 | −0.2 ± 1.2 | 0.9 | −0.1 ± 1.6 | 1.2 |
| 100 | −0.1 ± 1.1 | 0.8 | −0.1 ± 1.1 | 0.8 | −0.1 ± 1.5 | 1.2 |
| 200 | −0.1 ± 1.1 | 0.8 | 0.0 ± 1.1 | 0.8 | −0.1 ± 1.4 | 1.0 |
| 500 | −0.1 ± 1.1 | 0.8 | 0.1 ± 1.2 | 0.9 | −0.1 ± 1.0 | 0.8 |
| 1000 | −0.1 ± 1.1 | 0.8 | 0.2 ± 1.4 | 1.0 | 0.0 ± 1.2 | 1.0 |

*$3 \times (1, 10, 100, 1000)$.
†$b = 10.01 \pm 0.11$, $c = 1.06 \pm 0.89$.
‡$a = -(1.15 \pm 5.59) \times 10^{-2}$, $b = 10.06 \pm 0.32$, $c = 1.00 \pm 1.08$.
§$a = (-0.10 \pm 2.34) \times 10^{-4}$, $b = 10.01 \pm 0.18$, $c = 1.06 \pm 0.95$.

**Table 2**
Monte Carlo simulations comparing three weighted regressions with the curved concentration versus response relationship $y = 0.2895297x\ln x + 9x + 1$ using four calibrators in triplicate* (10 simulations)

| | $y = bx + c$† | | $y = ax\ln x + bx + c$‡ | | $y = ax^2 + bx + c$§ | |
|---|---|---|---|---|---|---|
| Concentration ($x$) | Per cent inter-assay calibration bias (mean ± SD) | Per cent mean calibration error | Per cent inter-assay calibration bias (mean ± SD) | Per cent mean calibration error | Per cent mean calibration bias (mean ± SD) | Per cent mean calibration error |
| 1 | 0.0 ± 7.5 | 6.2 | −0.8 ± 8.6 | 7.0 | −0.5 ± 7.9 | 6.5 |
| 2 | −4.9 ± 3.7 | 5.0 | −0.6 ± 3.7 | 2.9 | −3.3 ± 3.7 | 4.0 |
| 5 | −6.4 ± 1.6 | 6.4 | −0.4 ± 1.7 | 1.3 | −3.5 ± 1.7 | 3.6 |
| 10 | −5.8 ± 1.1 | 5.8 | −0.3 ± 1.5 | 1.2 | −2.5 ± 1.4 | 2.6 |
| 20 | −4.6 ± 1.0 | 4.6 | −0.3 ± 1.4 | 1.1 | −1.1 ± 1.5 | 1.5 |
| 50 | −2.4 ± 1.0 | 2.4 | −0.2 ± 1.2 | 0.9 | 0.9 ± 1.6 | 1.4 |
| 100 | −0.6 ± 1.1 | 1.0 | −0.1 ± 1.1 | 0.8 | 2.3 ± 1.5 | 2.4 |
| 200 | 1.2 ± 1.1 | 1.4 | 0.0 ± 1.0 | 0.8 | 3.1 ± 1.4 | 3.1 |
| 500 | 3.7 ± 1.1 | 3.7 | 0.1 ± 1.1 | 0.9 | 2.6 ± 1.0 | 2.6 |
| 1000 | 5.7 ± 1.1 | 5.7 | 0.2 ± 1.3 | 1.0 | −0.2 ± 1.1 | 0.9 |

*3 × (1, 10, 100, 1000).
†$b = 10.41 \pm 0.11$, $c = 0.41 \pm 0.81$.
‡$a = 0.279 \pm 0.056$, $b = 9.06 \pm 0.32$, $c = 1.00 \pm 0.99$.
§$a = (1.01 \pm 0.23) \times 10^{-3}$, $b = 10.01 \pm 0.18$, $c = 2.96 \pm 0.87$.
Note: $2/(3 \ln 10) = 0.2895297$.

term has also been reported elsewhere [19, 20]. Comparison of the Monte Carlo simulations for the linear concentration versus response curve (Table 1) shows that use of the new regression adds between 0.0 and 0.4% to the estimates of the mean calibration bias of the results depending upon the concentration but has negligible effect on the inter-assay component of the precision. An increase in the mean calibration bias was expected due to the greater uncertainty in the determination of three coefficients in the new equation rather than just two coefficients in the linear regression using the same data. However, the differences were less than those estimated when a less optimal distribution of calibrators ($2 \times$ 1,4,16,60,250,1000) was used for the weighted linear regression. Increasing the number of calibrators relative to those used in linear regression would also decrease the uncertainty.

Once a degree of curvature (7.4% per decade increase in concentration) was introduced into the response function then a significant improvement in the accuracy of the estimated results could be seen in the new regression compared with weighted linear regression (Table 2). Virtually no difference was observed in the performance of the new regression in the simulations between curved and linear concentration versus response profiles. It is interesting to note that although the slope of the concentration versus response curve changed by ±10%, from 9 at $x$ = 1–11 to $x$ = 1000, over the assay range in the simulations the calibration bias was limited to about ±6% in the estimates from weighted linear regression, some compensation for curvature being provided by the estimates of the coefficients $b$ and $c$ in the regression.

The GC ECD assay evaluated showed a marked degree of non-linearity and the validation results (Table 3) clearly demonstrate that the new regression describes the response characteristics much better than linear or quadratic regressions. Values of the mean accuracy lie within 96–104% of the target value over most of the assay range and include experimental errors in addition to those due to the calibration function. A plot of the residuals from the mean data from the six calibration curves also shows the new regression to fit the data within ±5% over a dynamic range of 1000 (Fig. 3). This fit can be further improved by omission of the highest calibration point which shows significant deviation from the curve in Fig. 1 because of loss of chromatographic resolution between the drug and internal standard at this concentration. An estimate of the deviation from linearity from the mean estimates of the coefficients $a$ and $b$ shows that the sensitivity of the ECD fell by 22% per decade increase in concentration. This value of curvature contrasts greatly with the impression created by a correlation coefficient of 0.9969 obtained by application of a linear regression to the same data.

Comparison of the validation results from an HPLC UV assay of a drug in plasma (Table 4) shows similar errors using either weighted

**Table 3**
Comparison of the accuracy and precision of the results from a GC ECD assay using the weighted regressions $y = bx + c$ and $y = ax\ln x + bx + c$ and $y = ax^2 + bx + c$

| Drug added (ng ml$^{-1}$) | Estimated assay bias (%)* | | |
|---|---|---|---|
| | I | II | III |
| 1.03 | −25.9 ± 10.1 | −9.4 ± 6.2 | −18.3 ± 8.4 |
| 2.59 | 21.4 ± 6.7 | −3.7 ± 3.7 | 10.7 ± 5.2 |
| 5.13 | 37.2 ± 4.2 | 3.7 ± 2.2 | 20.6 ± 2.6 |
| 10.4 | 31.9 ± 3.5 | 2.1 ± 1.3 | 14.4 ± 2.0 |
| 26.0 | 19.3 ± 4.6 | 0.6 ± 3.0 | 3.5 ± 2.9 |
| 51.9 | 7.3 ± 2.1 | −1.6 ± 1.7 | −6.3 ± 1.2 |
| 104 | −4.3 ± 2.2 | −2.5 ± 2.4 | −14.7 ± 2.8 |
| 260 | −19.1 ± 1.6 | −2.6 ± 2.6 | −23.3 ± 2.1 |
| 521 | −30.9 ± 2.8 | −3.3 ± 3.6 | −26.6 ± 3.5 |

*Mean and standard deviation of six independent analytical batches each using 14 calibrators ($2 \times$ 1, 3, 10, 31, 102, 307, 1022 ng ml$^{-1}$).

I: $y = bx + c$, $b = (2.013 \pm 0.040) \times 10^{-2}$, $c = (2.120 \pm 0.459) \times 10^{-2}$.

II: $y = ax\ln x + bx + c$, $a = (-3.436 \pm 0.352) \times 10^{-3}$, $b = (3.582 \pm 0.178) \times 10^{-2}$, $c = (2.948 \pm 4.244) \times 10^{-3}$; curvature = $230a/b = -22\%$ per decade increase in concentration.

III: $y = ax^2 + bx + c$, $a = (-1.244 \pm 0.117) \times 10^{-5}$, $b = (2.363 \pm 0.057) \times 10^{-2}$, $c = (1.654 \pm 0.438) \times 10^{-2}$.

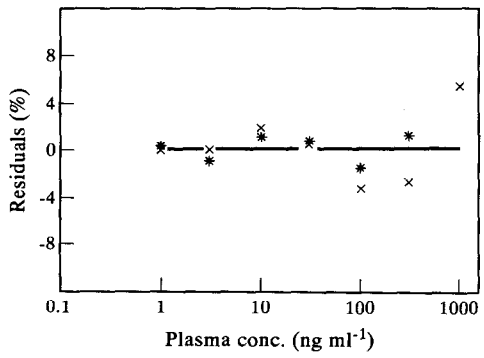


**Figure 3**
Residuals plotted against plasma drug concentration for a non-linear GC assay with ECD using the regression $y = ax\ln x + bx + c$ [range 1–1000, (x); range 1–300, (*)].

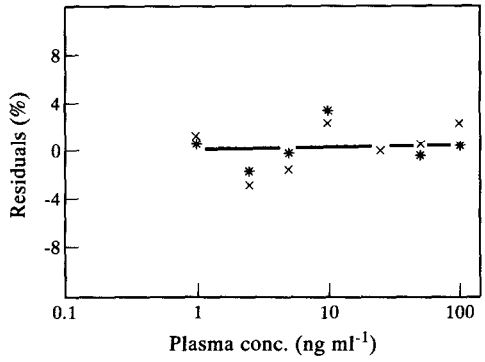


**Figure 4**
Residuals plotted against plasma drug concentration for a linear HPLC assay with UV detection comparing the weighted regressions for $y = bx + c$, (x) against $y = ax\ln x + bx + c$, (*).

**Table 4**
Comparison of the accuracy and precision of the results from a HPLC UV assay using the weighted regressions $y = bx + c$, $y = ax\ln x + bx + c$ and $y = ax^2 + bx + c$

| | Estimated assay bias (%)* | | |
|---|---|---|---|
| Drug added (ng ml$^{-1}$) | I | II | III |
| 1.0 | −5.0 ± 10.5 | −5.0 ± 10.5 | −5.0 ± 10.5 |
| 2.5 | 0.3 ± 6.4 | −1.1 ± 5.9 | −0.4 ± 7.1 |
| 5.0 | 0.6 ± 2.1 | 0.6 ± 3.0 | 0.6 ± 3.0 |
| 10.0 | 2.2 ± 1.9 | 1.8 ± 1.9 | 1.8 ± 1.9 |
| 25.1 | −0.1 ± 1.6 | −0.2 ± 1.6 | −0.3 ± 1.7 |
| 50.2 | 0.8 ± 2.6 | 0.9 ± 2.6 | 0.8 ± 2.6 |
| 100 | 0.9 ± 1.6 | 1.1 ± 1.4 | 1.1 ± 1.4 |

*Mean and standard deviation of six independent analytical batches each using 12 calibrators (4 × 1, 10, 100, ng ml$^{-1}$).
I: $y = bx + c$, $b = (3.020 \pm 0.032) \times 10^{-2}$, $c = (8.320 \pm 2.31) \times 10^{-3}$.
II: $y = ax\ln x + bx + c$, $a = (-2.06 \pm 3.94) \times 10^{-4}$, $b = (3.047 \pm 0.090) \times 10^{-2}$, $c = (8.04 \pm 3.05) \times 10^{-3}$; curvature = $230a/b = -1.6\%$ per decade increase in concentration.
III: $y = ax^2 + bx + c$, $a = (-1.69 \pm 4.85) \times 10^{-6}$, $b = (3.030 \pm 0.047) \times 10^{-2}$, $c = (8.21 \pm 2.59) \times 10^{-3}$.

linear or the new regression. A plot of normalized residuals of the mean responses of a calibration curve also shows little difference in the fit of the three regressions to the data (Fig. 4). An estimate of the degree of non-linearity was made from the six calibration curves using the new regression which showed a decrease in sensitivity of 1.6% per decade increase in concentration, which was not significantly different from zero, thus demonstrating the high degree of linearity of the assay.

It is worth making a brief comparison of this new regression with the more familiar quadratic equation, $y = ax^2 + bx + c$, sometimes used to compensate for non-linearity. The influence of the term $ax^2$ is only significant over a short range at the highest concentrations, whereas the term $ax\ln x$ describes a departure from linearity which is the same over each order of magnitude of the assay range. Obviously, the choice of regression will be dependent upon the characteristics of the assay in question, but in the example given, using ECD the term $ax\ln x$ plainly describes the non-linear behaviour of the assay better than the term $ax^2$. Although the quadratic equation allows direct calculation of unknown concentrations via the equation $x = (-b + [b^2 - 4a(c - y)]½)/2a$ there are limitations associated with this expression. Firstly, large errors can arise if the regression is used to describe a linear function and the coefficient $a$ is very close to zero. Secondly, it was not possible to back calculate the concentrations of the highest calibrators of the ECD assay as the poor fit of the regression at this level resulted in the term $4a(c - y)$ becoming greater than $b^2$.

The non-linear concentration versus response model described in this communication was based upon many such observations of assays used in this laboratory. However, this type of behaviour may not be solely due to the response characteristics of the detector. The previously reported non-linearity of HPLC UV photometric detectors [8] has not been seen in the current UV or fluorescent instrumentation

in this laboratory possibly due to improvements in instrument technology, but a similar apparent increase in sensitivity has been attributed to the poor chromatographic properties of one analyte. In GC the fall in sensitivity observed with ECD is due to the detector as is the increase in sensitivity of about 5% per decade increase in concentration of the NPD. Another cause of such non-linearity may be due to the hardware or software of chromatographic data systems. Whatever the reason, this type of response characteristic is not unusual and, because it is reproducible, can be compensated for by the regression described here.

The new regression has been presented primarily as a way to compensate for the type of non-linearity observed in some chromatographic systems. However, a case can be made for its more general use as a replacement for linear regression. Extra information about linearity is provided in a way that is readily understood by the analyst and is a worthwhile addition to the information concerning the sensitivity and intercept given by straightforward linear regression, whereas a correlation coefficient close to 1, though often quoted as evidence of linearity, has no real meaning in analytical calibration [21]. The increase in uncertainty in definition of individual calibration curves compared with linear regression has been shown to be small and, in association with other experimental errors, would be difficult to detect in practice. One area in which the new regression may be useful is in monitoring the ageing of the rubidium ceramic beads of the NPD in GC. Normally the degree of non-linearity associated with this detector is <5% per decade increase in concentration, but recently values in the range 10–20% have been observed in this laboratory. Experience may show that the change in linearity of this detector can be correlated with operational history and recoating of the ceramic bead with the rubidium salt.

Although the mathematics are more complex than those associated with linear regression the use of computer based data processing renders these transparent to the analyst and the accuracy of the calculations has been demonstrated over a wide dynamic range despite the use of single precision numbers. In this work a weighting factor of $1/x^2$ was used throughout in conjunction with the regressions but the equations can be modified to accommodate other weighting factors if these are felt appropriate. Initial work with this regression has been encouraging and it is being incorporated into assays in which the distribution pattern of the normalized residuals is consistently in agreement with the pattern predicted from simulations assuming that the response is described by the equation $y = ax\ln x + bx + c$. Even when the degree of non-linearity is small, the use of the new regression may lead to better quality control of assays by providing a sensitive monitor of changes in linearity which could be predictive of column or detector failure due to ageing.

*Acknowledgements* — The author would like to acknowledge the skills of the analysts within the Bioanalytical Department of Hoechst, particularly Keith Jolley, whose data provided much evidence for the concentration versus response model presented and some of which is included in this communication.

## References

[1] J.S. Garden, D.G. Mitchell and W.N. Mills, *Anal. Chem.* **52**, 2310–2315 (1980).
[2] L. Aarons, *Analyst* **106**, 1249–1254 (1981).
[3] L. Aarons, *J. Pharm. Biomed. Anal.* **2**, 395–402 (1984).
[4] L. Aarons, S. Toon and M. Rowland, *J. Pharmacol. Meth.* **17**, 337–346 (1978).
[5] R. Wolters and G. Kateman, *J. Chemometrics* **3**, 329–342 (1989).
[6] A.M. McLean, D.A. Ruggirello, C. Banfield, M.A. Gonzalez and M. Bialer, *J. Pharm. Sci.* **79**, 1005–1008 (1990).
[7] J.N. Miller, *Analyst* **116**, 3–14 (1991).
[8] C.A. Dorschel, J.L. Akmanis, J.E. Oberholtzer, F.V. Warren Jr and B.A. Bidlingmeyer, *Anal. Chem.* **61**, 951–968 (1989).
[9] H.T. Karnes and C. March, *J. Pharm. Biomed. Anal.* **9**, 911–918 (1991).
[10] E. Storr-Hansen, *J. Chromatogr.* **558**, 375–391 (1991).
[11] M.F. Delaney, *J. Liq. Chromatogr.* **3**, 164–268 (1985).
[12] R.P.W. Scott, *Liquid Chromatography Detectors*, pp. 13–19. Elsevier, New York (1986).
[13] P.K. Wilkinson, J.G. Wagner and A.J. Sedman, *Anal. Chem.* **47**, 1506–1510 (1975).
[14] J.J. Sullivan and C.A. Burgett, *Chromatographia* **8**, 176–179 (1975).
[15] C.F. Poole, *J. HRC CC* **5**, 454–471.
[16] G.A. Korn and T.M. Korn, in *Mathematical Handbook for Scientists and Engineers* (2nd edn), pp. 719. McGraw-Hill, New York (1968).
[17] A.C. Cartwright *et al.*, *Drug Inf. J.* **23**, 471–482 (1991).
[18] J.L. Burrows, *J. Pharm. Biomed. Anal.* **11**, 523–531 (1993).
[19] M. Thompson, *Analyst* **107**, 1169–1180 (1982).
[20] J. Agterdenbos, *Anal. Chim. Acta* **108**, 315–323 (1979).
[21] J.S. Hunter, *J. Assoc. Off. Anal. Chem.* **64**, 574–582 (1981).

*[Received for review 14 June 1993; revised manuscript received 4 October 1993]*