Appl. Statist. (2005)
54, *Part* 1, *pp.* 63–76

Limits of quantitation for laboratory assays

Christopher Cox

National Institutes of Health, Bethesda, and University of Rochester Medical Center, USA

[Received May 2003. Revised November 2003]

Summary. A common problem with laboratory assays is that a measurement of a substance in a test sample becomes relatively imprecise as the concentration decreases. A standard solution is to establish lower limits for reliable measurement. A quantitation limit is a level above which a measurement has sufficient precision to be reliably reported. The paper proposes a new approach to defining the limit of quantitation for the case where a linear calibration curve is used to estimate actual concentrations from measured values. The approach is based on the relative precision of the estimated concentration, using the delta method to approximate the precision. A graphical display is proposed for the assessment of estimated concentrations, as well as the overall reliability of the calibration curve. Our research is motivated by a clinical inhalation experiment. Comparisons are made between the approach proposed and two standard methods, using both real and simulated data.

Keywords: Calibration; Delta method; Detection limit; Limit of quantitation; Linear regression

1. Introduction

Modern laboratory assays can make measurements at increasingly lower levels, and more sensitive procedures continue to be developed. It is widely recognized, however, that even for the best analytical procedures the relative uncertainty in the measurement typically increases as the amount of the test substance in the sample decreases. In this situation the analytical chemist may be faced with a result whose error is comparable with the reported value.

An important aspect of this uncertainty is that assays typically report a background level even when none of the test substance is actually present in the sample. It is standard practice to estimate and then to adjust for the background level, which will usually be specific to the instrument or procedure, as well as the test substance and perhaps even the analytical run. When a measured value (reported by the assay) is converted into an estimated concentration through the use of a calibration curve, the estimate of the background level is provided by the intercept of the curve, and the standard deviation of the intercept provides a measure of its precision.

To determine whether or not a measurement actually indicates the presence of the test substance, we compare the measured value with the background level. A measurement is not considered to represent an actual amount unless the measured value is sufficiently above the estimated background. A comparison with a background level leads to the concept of a detection limit. A detection limit is a level above which the substance can reliably be reported to be present in the sample.

At the same time to say that a substance is detectable in a sample does not mean that the level can be reliably measured, i.e. that the level is actually quantifiable. A quantitation limit is a

Address for correspondence: Christopher Cox, Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, Department of Health and Human Services, National Institutes of Health, Room 7B05, 6100 Executive Boulevard, Bethesda, MD 20892-7510, USA.
E-mail: coxchris@mail.nih.gov

level above which the measured value has sufficient precision to be reliably reported. Of course an estimated concentration can always be reported together with an appropriate measure of error, or a confidence interval. The problem of providing simultaneous confidence intervals for a series of estimated concentrations has received attention, and useful procedures have been developed (Mee *et al.*, 1991; Osborne, 1991). In many cases, however, it may be preferable simply to report the measured value, provided that reasonable precision can be assured (International Union of Pure and Applied Chemistry Commission on Analytical Nomenclature, 1995; American Chemical Society Committee on Environmental Improvement, 1980). Thus values that are quantifiable can be reported as measured concentrations, and, for example, used in summary values such as averages. Values that are not quantifiable can be reported together with appropriate measures of error.

This paper considers the case in which a linear calibration curve is used to estimate the true concentration from a measured value. A case-study is provided by the investigation that motivated this research. This was a clinical inhalation study of a complex organic compound for 1 h in healthy volunteers. The primary goal was to study the uptake (from the lung to the blood) and subsequent elimination of the inhaled dose, requiring repeated measurements of the substance in plasma. Blood samples were collected from a number of volunteers following exposure and were analysed by using an assay that was developed for the purpose. The assay utilized a linear calibration curve to estimate the concentration in the sample. Subjects were exposed in a specialized exposure chamber and could only be exposed one at a time. For this reason a separate calibration curve was used for each subject in the study. Following standard practice the analytical chemist wished to establish a limit of quantitation for each calibration curve. Two different methods were available, and the investigator wanted to know whether these methods were statistically sound, as well as how they compared with each other. A study of these alternative methods led to the development of a new approach.

In general various methods are available for the construction of the calibration curve. An alternative to the classical least squares approach is provided by inverse regression, in which the true concentration is predicted directly from the measured value (Krutchkoff, 1967). There has been considerable discussion concerning the merits of the two approaches, and others have been proposed as well. The setting that is considered here is a common setting in which the number of observations that are available for the construction of the curve is reasonably large, typically with multiple measurements for each known concentration, and the fit of the curve is good. In this situation the classical approach to estimation of the calibration curve appears to have several advantages (Shukla, 1972) and will be adopted in this discussion.

For the classical setting in which a linear calibration curve is used to estimate actual concentrations from measured values, many researchers have previously discussed the limit of quantitation (Clayton *et al.*, 1987; Hubaux and Vos, 1970; Dunne, 1995; Oppenheimer *et al.*, 1983; Gibbons *et al.*, 1992). Typically a limit is first defined on the measured scale Y, and then transformed to the concentration scale X by means of the evaluation (inverse regression) function. This type of approach ignores the additional variability that is involved in the final step. This paper proposes a new approach to the limit of quantitation when a linear calibration curve is used to estimate the actual concentration. The approach is based on the precision of the estimated concentration, rather than the measured value. A graphical display of the relative precision is proposed for assessing the reliability of estimated concentrations, as well as the overall reliability of the calibration curve. The approach is illustrated by using a case-study, which includes a comparison of the proposed approach with two standard approaches. Further comparisons are provided by additional examples and a simulation study.

2. A new approach to the limit of quantitation for a linear calibration curve

2.1. Limits of detection and quantitation

A calibration curve must be estimated from data, usually a set of known concentrations covering a desired range, which are prepared and then measured. Suppose that we have a series of n standardization samples, $(X_1, Y_1), \ldots, (X_n, Y_n)$, where the X-values represent known concentrations and corresponding to these are the measured Y-values from the instrument. Let α and $\beta > 0$ denote the true intercept and slope of the calibration line; the fitted regression line is $\hat{Y} = a + bX$. For a given measurement Y based on a test sample, the estimated concentration is given by the evaluation function $X = (Y - a)/b$. To set the stage we first briefly consider the limit of detection. Suppose that Y represents a measured value. Then, if $E(Y) > E(a) = \alpha$, it must be true that $\{E(Y) - \alpha\}/\beta > 0$, i.e. the true concentration is greater than 0. Therefore the definition of the detection limit can be based on the intercept of the regression line, a, together with its standard deviation SD_a. A definition that is commonly used for the detection limit is $Y_d = a + 3(s^2 + \mathrm{SD}_a^2)^{1/2}$, where s^2 is the residual mean square from the regression analysis.

The simplest approach to the limit of quantitation is by analogy with the detection limit, using a factor of 10 instead of 3, as above. Thus the quantitation limit would be defined by $Y_q = a + 10(s^2 + \mathrm{SD}_a^2)^{1/2}$, or the value can be transformed to the concentration scale by using the evaluation function. This definition is based on the intercept and not on the precision of the predicted value from the calibration curve. A conceptual improvement was provided by both Oppenheimer *et al.* (1983) and Gibbons *et al.* (1992), who defined the quantitation limit in terms of the relative precision of a measured value of Y predicted from a true concentration of X by using the calibration curve, i.e. Y_c is chosen to satisfy $\mathrm{SD}(Y_c) = cY_c$ and $X_c = (Y_c - a)/b$. This approach, however, ignores the additional uncertainty that is caused by the use of the evaluation function. A second difficulty is that the value of Y_c still depends on the intercept of the regression line. As illustrated below, the quantitation limit X_c therefore depends on the overall level of the measurements.

The approach can clearly be improved by working on the X-scale. For a given level of precision, $c \times 100\%$ $(0 < c < 1)$, the quantitation limit is defined as the smallest concentration X_q, for which the approximate standard deviation, as defined below, of the concentration X estimated from a reading Y equals cX. A value of $c = 0.1$ or $c = 0.2$ might commonly be chosen. Thus the approach is again based on relative precision, but on the scale of the actual concentrations. This approach should typically result in a larger value for the limit of quantitation than the method that was proposed by Oppenheimer *et al.* (1983) and Gibbons *et al.* (1992), since it includes the additional uncertainty that is involved in using the evaluation function.

2.2. A new limit of quantitation—an application of the delta method

To formulate the quantitation limit proposed we need a measure of precision for the estimated concentration X. It is well known that an asymptotic variance for $X = (Y - a)/b$ can be obtained from the delta method (Cox, 1998). For this application we estimate $E(Y - a)/E(b)$ by X, yielding the approximation

$$\mathrm{var}\left(\frac{Y-a}{b}\right) \approx \frac{s^2}{b^2}\left\{\frac{n+1}{n} + \frac{(X-\bar{X})^2}{SS_x}\right\} \tag{1}$$

where $\mathrm{SS}_x = \Sigma\,(X_i - \bar{X})^2$. Approximation (1) is valid under the assumption $g \sim 0$, where $g = t^2 s^2 / b^2 \mathrm{SS}_x$, and t is a percentile of the t-distribution with $n - 2$ degrees of freedom corresponding to the desired significance level (level of confidence). In other words the approximation is

justified by large sample theory. The relative precision is then $\mathrm{SD}(X)/X$, and the proposed limit of quantitation is defined by $\mathrm{SD}(X) = cX$.

A possible objection to the use of the delta method approximation as a measure of precision for the estimated concentration is the fact that the latter has infinite variance (Williams, 1969). As indicated above, we are considering the case where the slope of the regression line is highly significant, which avoids problems with the approximation. The main point, however, is that large sample theory still applies, and the variance of the limiting distribution that is provided by the delta method can be used for the construction of confidence limits and tests (Cox and Hinkley, (1974), section 9.2 (iii)). Thus its use as a measure of precision seems entirely appropriate.

The defining equation for the limit of quantitation yields a quadratic equation whose coefficients are given in Appendix A. This equation either has no real roots or, when the slope is highly significant, a unique positive root, which can be found by using the quadratic formula

$$X_q = \frac{-\bar{X} + [\bar{X}^2 + (c^2\hat{t}^2 - 1)\{\bar{X}^2 + \mathrm{SS}_x(n+1)/n\}]^{1/2}}{c^2\hat{t}^2 - 1} \tag{2}$$

where $\hat{t}^2 = b^2\mathrm{SS}_x/s^2$ is the square of the t-statistic for the slope, so that $g = t^2/\hat{t}^2$. By definition we have $\mathrm{SD}(X_q) = cX_q$. The value of X_q depends on the mean and variance of the X-values, as well as the sample size n and the ratio b/s. Multiplication of either the coefficients or the solution by $g = t^2/\hat{t}^2$ shows that X_q depends on the observed readings only through the value of g. Further insight into the nature of this dependence may be obtained by noting that

$$g = \frac{t^2 s^2}{b^2\mathrm{SS}_x} = \frac{t^2(\mathrm{SST} - \mathrm{SSR})/(n-2)}{\mathrm{SSR}} = \frac{t^2(1-R^2)}{(n-2)R^2} \tag{3}$$

where SSR denotes the sum of squares for regression and SST the corrected total sum of squares. This shows that g and therefore X_q depend on the data only through R^2. Note that g is small when R^2 is close to 1, or when n is large; in fact $g = o_p(1/n)$ (Cox, 1990). The new quantitation limit may be illustrated graphically by plotting $\mathrm{SD}(X)/X$ against X. This shows the relative precision of the estimated concentration as a percentage error and is useful for assessing values that are provided by the calibration curve as well as the curve itself.

For comparison with the approximation that is given by the delta method, consider upper and lower non-simultaneous prediction bands about the fitted calibration line, $Y = a + bX$. These are

$$Y \pm ts\left\{1 + \frac{1}{n} + \frac{(X-\bar{X})^2}{\mathrm{SS}_x}\right\}^{1/2}. \tag{4}$$

Note that the variance for X provided by the delta method (1) is simply the variance for Y in expression (4) divided by b^2. It is well known that an asymmetric confidence interval for the true concentration corresponding to a measured value of Y, estimated as $X = (Y-a)/b$, can be obtained by inverting these prediction bands. Draper and Smith (1981) gave the following expression for the resulting confidence limits, which are obtained as the roots of a quadratic equation:

$$\left.\begin{matrix} X_{\mathrm{U}} \\ X_{\mathrm{L}} \end{matrix}\right\} = X - \frac{(\bar{X}-X)g}{1-g} \pm \frac{(ts/b)\{(X-\bar{X})^2/\mathrm{SS}_x + (1-g)(n+1)/n\}^{1/2}}{1-g}. \tag{5}$$

The quadratic equation has two real roots which define a finite interval if the slope of the calibration curve is significantly different from 0, i.e. if $b/(s/\mathrm{SS}_x) > t$, or $g < 1$ (Draper and Smith, 1981; Brownlee, 1965). In the case of a calibration curve we would expect that typically $g \ll 1$.

As noted by a reviewer, the above derivation bears a close resemblance to Fieller's theorem. This result allows the computation of confidence limits for the ratio of two linear functions of the regression coefficients in a multiple-regression model (Zerbe, 1978). To see the connection with the calibration problem we must augment the standardization data set by adding an extra observation $(1, Y)$, and then include an extra parameter, $Y = 1\gamma$ in the model, so that $\hat{\gamma} = Y$. With this formulation and a suitable choice of the two linear functions of the three parameters, the confidence interval (5) is identical to the confidence interval that is obtained by application of Fieller's theorem. It is well known that under large sample regularity conditions the confidence limits (X_L, X_U) become symmetric with the approximate form (Cox, 1990) $X \pm t\,\mathrm{SD}(X)$. Brownlee (1965) stated that this is generally valid when $g < 0.1$. The symmetric confidence limits can be directly obtained from the confidence limits X_U and X_L by assuming that $g \sim 0$.

The assumption $g \sim 0$ can also be used to simplify further the expression for the approximate variance of X in equation (1) based on the delta method. Under this assumption the approximation

$$\mathrm{SD}(X) = \frac{s}{b}\left\{\frac{(\bar{X}-X)^2}{\mathrm{SS}_x} + \frac{n+1}{n}\right\}^{1/2} = \left\{\frac{g}{t^2}(\bar{X}-X)^2 + \frac{n+1}{n}\frac{s^2}{b^2}\right\}^{1/2} \approx \left(\frac{n+1}{n}\frac{s^2}{b^2}\right)^{1/2} \qquad (6)$$

yields an approximation for the limit of quantitation $X_q \approx X_g = \{(n+1)/n\}^{1/2} s/bc$. As shown in Appendix A, we actually have $X_g < X_q$, so X_g provides a lower bound for the limit of quantitation that is independent of the precision with which the calibration curve has been estimated. In the limit as $n \to \infty$ (under appropriate regularity conditions) we have $X_g \to_p \sigma/\beta c$. Under the same limiting conditions, we also have $\hat{t}^2 \approx \beta^2 \mathrm{SS}_x/\sigma^2 \to \infty$, so $X_q \to_p \sigma/\beta c$ as well. This parameter provides a simple description of the measurement process, allowing the comparison of different calibration curves. It expresses the variability on the appropriate (X-) scale, since if $X = (Y-\alpha)/\beta$ then $\mathrm{SD}_X = \mathrm{SD}_Y/\beta = \sigma/\beta$. The delta method approximation behaves in a similar way. Using this approximation it can also be seen that the limit of detection corresponds to a precision of approximately 33.3% (10/3 × 10%). Using arguments similar to those in Appendix A it can be shown that $X_q < \bar{X}$ if and only if $X_g < \bar{X}$.

2.3. Two standard approaches

We first want to consider the concentration corresponding to the limit of quantitation that was defined earlier in terms of a comparison of a new measurement with the intercept of the calibration curve. This is

$$Y_q = a + \frac{1}{c}(s^2 + \mathrm{SD}_a^2)^{1/2} = a + \frac{s}{c}\left(1 + \frac{1}{n} + \frac{\bar{X}^2}{\mathrm{SS}_x}\right)^{1/2}.$$

The estimated concentration corresponding to this limit is

$$X_y = \frac{Y_q - a}{b} = \frac{s}{bc}\left(1 + \frac{1}{n} + \frac{\bar{X}^2}{\mathrm{SS}_x}\right)^{1/2} = \frac{1}{c}\left(\frac{n+1}{n}\frac{s^2}{b^2} + \frac{g\bar{X}^2}{t^2}\right)^{1/2} = \frac{1}{c\hat{t}}\left(\bar{X}^2 + \mathrm{SS}_x\frac{n+1}{n}\right)^{1/2}. \qquad (7)$$

Clearly $X_y > X_g$, and neglecting g we have $X_y \approx X_g$. Under the assumption $X_y < 2\bar{X}$ it is shown in Appendix A that $X_q < X_y$. In addition we have $X_y/X_g = [n\bar{X}^2/\{(n+1)\mathrm{SS}_x\} + 1]^{1/2}$. This ratio is independent of the measurement data and can be made arbitrarily close to 1 by choice of the design points of the calibration assay. Thus, if prior information is available on the slope and residual error, the calibration run can be designed with a desired quantitation limit in mind.

Finally for comparison we also need to consider the quantitation limit that was proposed by Oppenheimer *et al.* (1983) and Gibbons *et al.* (1992). Following their approach we compute

the predicted value Y_c such that $\mathrm{SD}(Y_c)/Y_c = c$, and then define $X_c = (Y_c - a)/b$. This definition leads to a quadratic equation for Y_c, whose coefficients are also included in Appendix A. These coefficients show that the resulting value depends on the intercept of the calibration curve. We shall compare the proposed estimate with this alternative value by using a series of calibration data sets and see whether, as expected, the new limit of quantitation is indeed more conservative.

3. A case-study and three examples

3.1. A clinical inhalation study

As described in Section 1, a clinical inhalation study provided the motivation for the present research. The study actually involved a series of 10 separate calibration curves, each used for measurement of a series of post-exposure samples from a different subject. Both the measured value (the area ratio) and the actual value (the concentration ratio) were expressed as ratios, using an internal standardization process, and the scale of the measurements was between 0 and 1. With the exception of the first subject (10 concentrations), each standardization data set involved six different concentrations in the range 0.0–0.5. With the exception of the zero concentration, each was done in triplicate. Over half the samples were blanks (concentrations of 0), again with the exception of the first subject, which had only one blank and also had the smallest number of data points (28). Table 1 contains summary information for each of the 10 different calibration curves: it gives the number of points n, the mean of the standard concentrations X, the slope b and the values of R^2 and g for the regression line. In every case the number of points is fairly large, the goodness-of-fit is above 98.5% and the values of g are small, with smaller values of g corresponding to larger values of R^2. The means of the concentrations are all about 0.1. Slope estimates range from 1.90 to 2.50, so in terms of the ratio s/b the precision is approximately doubled on the concentration scale. The range of the residual standard deviation s was 0.01–0.04, with larger values of s generally corresponding to larger values of the slope.

Table 1. 10 calibration curves for different subjects S in a clinical inhalation study†

S	*Sample size* n	*Mean* $\bar{X}$	*Slope* b	R^2 *(%)*	g	*Detection limit* X_d	*Quantitation limit,* X_q		*Lower bound* X_g	*Standard approaches*	
							10%	*20%*		X_y	X_c
1	28	0.135	1.95	99.7	0.00043	0.026	0.086	0.043	0.086	0.087	0.067
2	38	0.090	2.58	99.0	0.0012	0.048	0.159	0.079	0.159	0.159	0.155
	35‡	0.055	2.29	99.8	0.00024	0.013	0.043	0.021	0.043	0.043	0.034
3	39	0.088	2.50	99.5	0.00060	0.034	0.114	0.057	0.113	0.114	0.111
4	36	0.095	2.50	99.4	0.00074	0.037	0.124	0.062	0.124	0.125	0.119
5	36	0.095	1.90	99.8	0.00022	0.021	0.068	0.034	0.068	0.069	0.054
6	36	0.095	2.46	98.8	0.0015	0.053	0.176	0.088	0.175	0.176	0.173
7	32	0.106	2.09	99.6	0.00059	0.032	0.107	0.054	0.107	0.108	0.089
8	33	0.103	2.26	98.7	0.0017	0.055	0.184	0.092	0.184	0.185	0.166
9	38	0.091	2.25	99.7	0.00036	0.026	0.087	0.044	0.087	0.088	0.065
10	38	0.093	2.19	99.3	0.00086	0.042	0.140	0.070	0.140	0.140	0.109

†The first six columns describe the regression line (for g see equation (3)). The transformed detection limit and two standard approaches are included for comparison with the proposed limit of quantitation (10% and 20% relative precision) and its (10%) lower bound.
‡Three outliers omitted.

Table 1 also includes the transformed detection limit $X_d = (Y_d - a)/b$ for each curve (using a multiple of 3), and both 10% ($c = 0.1$) and 20% ($c = 0.2$) quantitation limits X_q in equation (2). For comparison, the lower bound X_g (based on equation (6)) and the standard limits X_y in equation (7) and X_c are also included, all for a relative precision of 10%. For the second data set, it was noted that the curve was not perfectly linear, and the three points at the upper end of the curve, which were beyond the range of the actual samples, were dropped. The results with all the original data are included for comparison. For this data set, removing the three replicates at the highest concentration increased the value of R^2 and decreased the value of g considerably. The limit of quantitation is about a third of the original value.

As can be seen in Table 1, the condition $X_y < 2\bar{X}$ is satisfied by all 10 data sets, so $X_g < X_q < X_y$ in every case. As expected we have $X_c < X_q$, and in some instances the difference might be important in practice. For all 10 curves, both X_g and X_y agree quite well with the proposed quantitation limit X_q. Thus, in this study, the limit of quantitation depends primarily on the slope and residual error of the calibration curve, and not on the goodness of fit. The limits differ by a factor of at most 4, depending on the ratio s/b, and to a lesser extent the number of points on the curve. For six of the 10 curves the quantitation limit exceeds the mean of the standards. The 20% quantitation limits are about half the value of the 10% limits, and the detection limits are about 0.3 of the 10% quantitation limits, reflecting the approximate inverse proportionality to the constant c that can be seen in the expression for X_q, and more clearly for the lower bound X_g.

The results of this analysis indicate good agreement between the proposed limit of quantitation X_q and the standard limit X_y. This suggests that, from a practical point of view, either limit could be used in this study. Of course, since the new limit can be easily computed, there is no reason not to use it. As noted, the numbers of points for these curves are fairly large and the goodness of fit is quite good. To examine the agreement of the proposed limit with the two standards further, we consider some additional calibration data sets, all having smaller numbers of points.

3.2. Three examples from other studies

As a first example, we consider a data set that was used by Gibbons *et al.* (1992). They provided summary statistics ($b = 0.32$, $s = 0.06$, $\mathrm{SS}_x = 51.12$, $n = 16$, $\bar{X} = 3.55$ and $\bar{Y} = 1.25$) for the calculation of $Y_c = 0.64$ and the 10% limit of quantitation $X_c = 1.64$. These same summary values can be used to compute $R^2 = 99.0\%$, $g = 0.0032$, the proposed limit of quantitation $X_q = 1.98$, the limiting value $X_g = 1.93$ and the limit based on the intercept, $X_y = 2.15$. Thus we again have $X_c < X_q$, and $X_g < X_q < X_y$ as well.

The second example is taken from Brownlee (1965). This is a set of calibration data from a method for measuring blood flow, consisting of an inexpensive method calibrated against a gold standard. There are $n = 18$ points; the estimate (with the standard deviation in parentheses) of the intercept is $a = 15.35$ (56.06), and the slope is $b = 0.975$ (0.0251). The estimate of the residual error is $s = 56.39$. The value of R^2 is 98.9%, and $g = 0.00299$. Using a multiplier of 3, the limit of detection is $X_d = 244.8$. The 10% limit of quantitation is $X_q = 704.0$, and the 20% limit is 376.3. A plot of the regression line with 95% prediction bands is shown in Fig. 1(a), which also shows the estimated flow rate corresponding to a new measured value Y, together with a 95% confidence interval. The confidence limits for the 10% limit of quantitation are (550.1, 849.1); these are not entirely appropriate, although they provide an approximation.

The lower bound is $X_g = 594.4$. The reason that this does not agree very well with X_q appears to be that the term $(\bar{X} - X_q)^2$ is relatively large ($\bar{X} = 2165.0$). The quantitation limit that corresponds to the intercept is $X_y = 815.9$, which is much larger than X_q. Finally the limit of quantitation based on Oppenheimer *et al.* (1983) and Gibbons *et al.* (1992) is $X_c = 690.1 < 704.0 = X_q$.

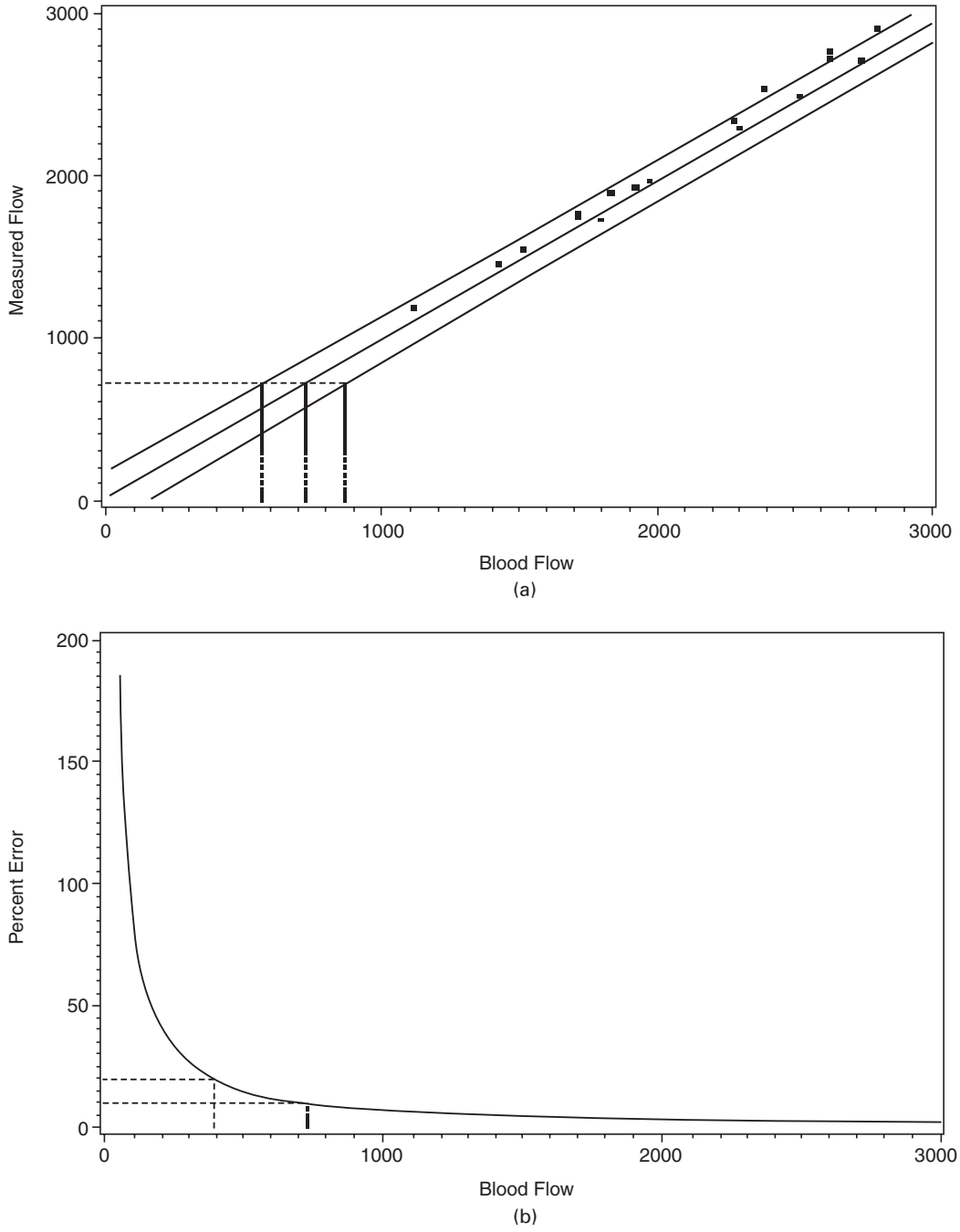

Fig. 1. Calibration curve comparing an indirect method for measuring blood flow with a direct method (Brownlee, 1965): (a) regression line with the data and 95% prediction bands and a predicted flow rate derived from a measured value, together with a 95% confidence interval; (b) relative error curve for this example together with 10% (704.0) and 20% (376.3) quantitation limits

Once again, the new limit of quantitation provides a good compromise. The 10% and 20% quantitation limits are shown graphically in the percentage error plot, Fig. 1(b). The plot was truncated so that values of the percentage error above 200% are not shown. The plot shows how the relative precision of the estimated blood flow decreases as the estimate decreases; the percentage error does not become large until the estimated flow rate is well below the range of the calibration data.

The final illustration was taken from Dunne (1995). This is an assay for cyclic adenosine monophosphate; measured values are reciprocals of radioactivity counts (times 10^3). The data consist of seven pairs ($n = 14$) of duplicate standard concentrations, in the range 0.0–8.0 (starting with 0.25, concentrations were doubled until 8.0). We have a (with standard deviation in parentheses) $= 0.175$ (0.00676) and $b = 0.0707$ (0.00193). The estimate of residual error is $s = 0.0193$. The value of R^2 is 99.1%, with $g = 0.00355$. The detection limit is $X_d = 0.868$. The 10% limit of quantitation is $X_q = 2.83$, and the 20% limit is 1.42. For the 10% limit we have $X_q > \bar{X} = 2.25$ and in fact the 10% quantitation limit is above the values for most of the standards. The lower bound is close to the limit of quantitation, $X_g = 2.83$, and the limit by using the intercept of the calibration curve is $X_y = 2.89$. Finally we have $X_c = 0.40$, which is actually smaller than the limit of detection. The percentage error plot provides a contrast with Fig. 1(b); in this case the overall reliability of the calibration curve in the range of the data is relatively poor. The error for an estimated value of 2.0 is approximately 15%, and for an estimate of 1.0 the percentage error is about 30%. The relative precision at X_c is very poor. The reason for the small value of X_c appears to be that the intercept of the calibration curve is relatively large. Indeed, since the numerator does not change, we can make the expression $\mathrm{SD}(Y_c)/Y_c$ as small as we please by adding a sufficiently large constant to all the measured values. Alternatively, if we subtract 0.1 from all the Y-values, we obtain $X_c = 1.77$.

4. Two simulation studies

At the suggestion of a reviewer a series of simulation studies was performed, first to verify the adequacy of the delta method approximation in the present setting and secondly to compare the proposed limit of quantitation with the two alternative approaches. Both simulation studies were structured primarily on the basis of the 10 different calibration curves in the case-study. Combinations of the following series of parameters were used: $\beta = 0.5, 1.0, 1.5$; $\alpha = 0.05, 0.10$; $\sigma = 0.01, 0.05, 0.10$. Sample sizes of $n = 15$ and $n = 25$ were used. For a given sample size, n values of the true concentration X_i were sampled from a uniform distribution on the unit interval. Measured values Y_i were sampled from the assumed regression line with normal errors.

To study the delta method approximation, the following computations were performed for each sample data set. First, the least squares estimates of the slope b and intercept a and the residual error were computed. Then, an additional observed value $Y = \alpha + \beta\bar{X}/4 + \varepsilon$ was generated, and from this the estimated concentration $X = (Y - a)/b$ and its standard deviation (1) were calculated. For a given configuration of X_i-values, this was repeated with 1000 different sets of Y_i-values. Finally the approximate standard deviations were averaged, and the sample standard deviation of the estimated concentrations was computed for comparison. Values in Table 2 are the means of 1000 samples of X_i-values. The results clearly show good agreement between the approximate standard deviation given by the delta method and the sample standard deviation of the estimated values X, even with a sample size as small as 15. The delta method typically underestimates the sample standard deviation slightly, and the approximation improves as the sample size increases.

Table 2. Simulation study of the delta method approximation to the standard deviation for a concentration estimated by the evaluation function†

σ^2	n	*Results for the following values of* β *and* α*:*											
		$\beta=0.5, \alpha=0.05$		$\beta=0.5, \alpha=0.10$		$\beta=1.0, \alpha=0.05$		$\beta=1.0, \alpha=0.10$		$\beta=1.5, \alpha=0.05$		$\beta=1.5, \alpha=0.10$	
		D	*MC*	*D*	*MC*	*D*	*MC*	*D*	*MC*	*D*	*MC*	*D*	*MC*
0.01	15	0.0215	0.0219	0.0215	0.0219	0.0107	0.0110	0.0107	0.0109	0.0072	0.0073	0.0072	0.0073
0.05	15	0.109	0.112	0.109	0.111	0.0539	0.0550	0.0539	0.0550	0.0359	0.0365	0.0359	0.0366
0.10	15	0.257	0.249	0.233	0.241	0.109	0.111	0.109	0.111	0.0721	0.0735	0.0721	0.0736
0.01	25	0.0209	0.0211	0.0209	0.0211	0.0104	0.0105	0.0104	0.0106	0.0070	0.0070	0.0070	0.0070
0.05	25	0.105	0.107	0.105	0.106	0.0523	0.0529	0.0523	0.0528	0.0348	0.0352	0.0348	0.0352
0.10	25	0.217	0.220	0.217	0.219	0.105	0.106	0.105	0.106	0.0699	0.0707	0.0699	0.0707

†The parameters for the simulations were the slope β, intercept α, error variance σ^2 and sample size n of the calibration curve: D, delta method; MC, Monte Carlo (descriptive) estimate.

Table 3. Comparison of limits of quantitation by using simulated data†

σ^2	n		*Limits for the following values of* β *and* α*:*					
			$\beta=0.5$, $\alpha=0.05$	$\beta=0.5$, $\alpha=0.10$	$\beta=1.0$, $\alpha=0.05$	$\beta=1.0$, $\alpha=0.10$	$\beta=1.5$, $\alpha=0.05$	$\beta=1.5$, $\alpha=0.10$
0.01	15	X_y	0.2234	0.2237	0.1118	0.1118	0.0745	0.0746
		X_q	0.2097	0.2099	0.1078	0.1078	0.0727	0.0727
		X_g	0.2023	0.2025	0.1013	0.1013	0.0675	0.0675
		X_c	0.1146	0.0211	0.0594	0.0111	0.0401	0.0076
0.05	15	X_y			0.4377	0.4379	0.3571	0.3569
		X_q			0.4048	0.4051	0.3297	0.3295
		X_g			0.4025	0.4028	0.3252	0.3250
		X_c			0.3607	0.3135	0.2978	0.2662
0.01	25	X_y	0.2138	0.2142	0.1070	0.1071	0.0713	0.0713
		X_q	0.2059	0.2062	0.1047	0.1048	0.0703	0.0703
		X_g	0.2015	0.2018	0.1009	0.1009	0.0672	0.0672
		X_c	0.1090	0.0133	0.0557	0.0068	0.0374	0.0045
0.05	25	X_y			0.4328	0.4326	0.3492	0.3497
		X_q			0.4120	0.4119	0.3321	0.3326
		X_g			0.4110	0.4108	0.3296	0.3301
		X_c			0.3670	0.3185	0.2997	0.2680

†For each simulation, the value of R^2 was required to exceed 0.98; empty cells indicate fewer than 500 acceptable simulations. The parameters for the simulations were the same as those for Table 2. The proposed limit is X_q, with lower bound X_g. For comparison X_y and X_c are standard methods.

For the second simulation study a single set of Y_i-values was generated for each set of sampled X_i-values, and the estimates X_y, X_q, X_g and X_c were computed. To be considered as simulated calibration data, the value of R^2 was required to exceed 0.98, which was the case for all the examples. For $\sigma = 0.1$ this resulted in rejection of all or nearly all the simulated data sets. The corresponding combinations of parameter values are clearly not realistic for calibration data, and this value was omitted. For each combination of parameters and sample size, a total of 10^5 data sets was generated. The averaged results in Table 3 show that the relationships between the four quantities were consistent with those that were found in the examples. In fact the condition $X_y < 2\bar{X}$ was satisfied for every acceptable sample so $X_g < X_q < X_y$ held in every case, not simply on average. In general X_q was closer to X_g than to X_y, but this was not always so. Values of X_c were generally smaller than X_g, although in some cases they exceeded X_y. The results illustrate the dependence of X_c on the intercept of the regression line, with the larger intercept having the smaller value of X_c. The results also show approximate proportionality between the quantities X_y, X_q and X_g and the ratio $\sigma/\beta c$.

5. Discussion

This paper proposes a new approach to defining the limit of quantitation, X_q, for a laboratory assay where a linear calibration curve is used to estimate actual concentrations from measured values. The new method is an improvement over previous approaches since it is based on the precision of the estimated concentration. Thus a value is quantifiable if it has adequate relative precision, a standard criterion for assessing the reliability of an assay. A simple standard approach (X_y) was shown to be a typically conservative approximation to the quantitation limit

proposed. Finally a lower bound X_g was provided that is independent of the goodness of fit of the calibration curve. This lower bound represents the smallest limit of quantitation that can be achieved for a given ratio $\sigma/\beta c$. Data from a case-study show that, when the number of points is relatively large, and the fit is good, the proposed limit of quantitation depends primarily on this ratio. Both real data examples and simulation studies show that the new approach provides a reasonable compromise between two standard approaches. The limit X_c can produce values that are very low, and it is sensitive to the intercept of the calibration line. As indicated in the case-study, the limit X_y can sometimes be used but is not guaranteed to be conservative, so the relative error can exceed the desired level. In both examples and simulations the difference between this limit and X_q would in some instances be of practical importance.

The new limit of quantitation depends on the goodness of fit (R^2) of the calibration curve. Thus the analytical chemist is rewarded in terms of the sensitivity of measurement for care in constructing the curve. This seems consistent with the use of the curve to estimate the actual concentration. At the same time errors in the measured values play a direct role. If the slope and error variance are regarded as fixed parameters of the calibration process, then the limit of quantitation can be made to approach the lower bound X_g by increasing the sample size, i.e. increasing the value of SS_x, since this will decrease the value of g. The value of X_g will in turn be small if the error variance of the calibration line is small.

The limit of quantitation is defined in terms of relative precision rather than absolute precision (International Union of Pure and Applied Chemistry Commission on Analytical Nomenclature, 1995). For most purposes this seems appropriate. There may be situations, however, where the absolute error is a more relevant criterion. An example of this is when the sole purpose of the analysis is to allow a comparison of a treated group of experimental subjects with a control. In this case the relative precision of the control values may be low compared with higher values in the treatment group, whereas the absolute precision, which would be relevant for the statistical comparison, might be relatively high.

The assessment of precision should include all the steps that are necessary to produce the final result. It would not be appropriate, for example, to adjust each of the estimated concentrations by multiplying by dilution factors that have themselves been measured with error. Rather the calibration curve should be constructed by using the final estimate as the independent variable.

Although in practice the value of X_q is simply used to determine the limit of acceptable precision for values reported by the assay, a statement of the precision of X_q itself may also be useful. A small sample confidence interval could be based on the fact that $g^{-1/2}$ has a non-central t-distribution. Alternatively we could compute a large sample standard deviation and corresponding confidence interval based on the independence and asymptotic normality of $(b, s_{y.x})$. Bootstrap resampling would seem to be preferable to either of these alternatives, although this would add considerably to the computational burden.

Acknowledgements

I thank Bob Gelein, Paul Morrow and Mark Utell for helpful discussions, and for the use of their data. The Joint Editor and two reviewers provided some helpful suggestions for revision of the manuscript.

Appendix A

The definition $\mathrm{SD}(X) = cX$ for the limit of quantitation X_q yields a quadratic equation with coefficients

$$a_q = c^2\hat{t}^2 - 1,$$

$$b_q = 2\bar{X},$$

$$c_q = -\{\bar{X}^2 + \mathrm{SS}_x(n+1)/n\}.$$

Since always $c_q < 0$, this equation either has no real roots or else, if $a_q > 0$ (when $\hat{t} > 1/c$, i.e. when g is small), has a unique positive root, which can be found by using the quadratic formula.

For the limit X_c, rewriting the results in the appendix of Gibbons *et al*. (1992), we obtain the following coefficients of the quadratic equation for Y_c:

$$a_y = c^2\hat{t}^2 - 1,$$

$$b_y = 2\bar{Y},$$

$$c_y = -\{\bar{Y}^2 + b^2\mathrm{SS}_x(n+1)/n\}.$$

These values should be compared with the coefficients for X_q above; they clearly reflect the fact that we are working on the Y-scale.

We next want to show that $X_q > X_g$, i.e. we want

$$X_q = \frac{-\bar{X} + [\bar{X}^2 + (c^2\hat{t}^2 - 1)\{\bar{X}^2 + \mathrm{SS}_x(n+1)/n\}]^{1/2}}{c^2\hat{t}^2 - 1} > \frac{1}{c}\left(\frac{\mathrm{SS}_x}{\hat{t}^2}\frac{n+1}{n}\right)^{1/2} = \left(\frac{n+1}{n}\right)^{1/2}\frac{s}{bc} = X_g.$$

Thus we want to show that

$$\frac{[\bar{X}^2 + (c^2\hat{t}^2 - 1)\{\bar{X}^2 + \mathrm{SS}_x(n+1)/n\}]^{1/2}}{c^2\hat{t}^2 - 1} > \frac{1}{c}\left(\frac{\mathrm{SS}_x}{\hat{t}^2}\frac{n+1}{n}\right)^{1/2} + \frac{\bar{X}}{c^2\hat{t}^2 - 1},$$

$$\frac{\bar{X}^2 + (c^2\hat{t}^2 - 1)\{\bar{X}^2 + \mathrm{SS}_x(n+1)/n\}}{(c^2\hat{t}^2 - 1)^2} > \frac{1}{c^2}\left(\frac{\mathrm{SS}_x}{\hat{t}^2}\frac{n+1}{n}\right) + \frac{\bar{X}^2}{(c^2\hat{t}^2 - 1)^2} + \frac{2\bar{X}}{c(c^2\hat{t}^2 - 1)}\left(\frac{\mathrm{SS}_x}{\hat{t}^2}\frac{n+1}{n}\right)^{1/2},$$

$$\frac{\bar{X}^2 + \mathrm{SS}_x(n+1)/n}{c^2\hat{t}^2 - 1} > \frac{1}{c^2}\left(\frac{\mathrm{SS}_x}{\hat{t}^2}\frac{n+1}{n}\right) + \frac{2\bar{X}}{c(c^2\hat{t}^2 - 1)}\left(\frac{\mathrm{SS}_x}{\hat{t}^2}\frac{n+1}{n}\right)^{1/2},$$

$$\frac{\bar{X}^2}{c^2\hat{t}^2 - 1} + \frac{\mathrm{SS}_x(n+1)/n}{c^2\hat{t}^2(c^2\hat{t}^2 - 1)} > \frac{2\bar{X}}{c(c^2\hat{t}^2 - 1)}\left(\frac{\mathrm{SS}_x}{\hat{t}^2}\frac{n+1}{n}\right)^{1/2},$$

$$c^2\hat{t}^2\bar{X}^2 + \mathrm{SS}_x(n+1)/n > 2c\hat{t}\bar{X}\{\mathrm{SS}_x(n+1)/n\}^{1/2}.$$

This inequality is a special case of $a^2 + b^2 > 2ab$. Now let us ask when $X_q < X_y$, i.e.

$$X_q = \frac{-\bar{X} + [\bar{X}^2 + (c^2\hat{t}^2 - 1)\{\bar{X}^2 + \mathrm{SS}_x(n+1)/n\}]^{1/2}}{c^2\hat{t}^2 - 1} < \frac{1}{c\hat{t}}\left(\bar{X}^2 + \mathrm{SS}_x\frac{n+1}{n}\right)^{1/2} = X_y.$$

Arguing as before, we obtain the following, equivalent, expressions:

$$\frac{\bar{X}^2 + \mathrm{SS}_x(n+1)/n}{c^2\hat{t}^2 - 1} < \frac{1}{c^2\hat{t}^2}\left(\bar{X}^2 + \mathrm{SS}_x\frac{n+1}{n}\right) + \frac{2\bar{X}}{c\hat{t}(c^2\hat{t}^2 - 1)}\left(\bar{X}^2 + \mathrm{SS}_x\frac{n+1}{n}\right)^{1/2},$$

$$\frac{\bar{X}^2 + \mathrm{SS}_x(n+1)/n}{c^2\hat{t}^2(c^2\hat{t}^2 - 1)} < \frac{2\bar{X}}{c\hat{t}(c^2\hat{t}^2 - 1)}\left(\bar{X}^2 + \mathrm{SS}_x\frac{n+1}{n}\right)^{1/2},$$

$$\bar{X}^2 + \mathrm{SS}_x(n+1)/n < 2c\hat{t}\bar{X}\{\bar{X}^2 + \mathrm{SS}_x(n+1)/n\}^{1/2}.$$

Finally we have $X_y < 2\bar{X}$. It was relatively easy to modify one of the examples so that this condition was not satisfied. Conversely if $X_y > 2\bar{X}$ then we can reverse the string of inequalities to obtain $X_q > X_y$.

References

American Chemical Society Committee on Environmental Improvement (1980) Guidelines for data acquisition and data quality evaluation in environmental chemistry. *Anal. Chem.*, **52**, 2242–2249.
Brownlee, K. A. (1965) *Statistical Theory and Methodology in Science and Engineering*, 2nd edn. New York: Wiley.
Clayton, C. A., Hines, J. W. and Elkins, P. O. (1987) Detection limits with specified assurance probabilities. *Anal. Chem.*, **59**, 2506–2514.
Cox, C. (1990) Fieller's theorem, the likelihood, and the delta method. *Biometrics*, **46**, 709–718.
Cox, C. (1998) Delta method. In *Encyclopedia of Biostatistics* (eds P. Armitage and T. Colton), pp. 1125–1127. New York: Wiley.
Cox, D. R. and Hinkley, D. V. (1974) *Theoretical Statistics*. London: Chapman and Hall.
Draper, N. R. and Smith, H. (1981) *Applied Regression Analysis*, 2nd edn. New York: Wiley.
Dunne, A. (1995) Decision and detection limits for linear homoscedastic assays. *Statist. Med.*, **14**, 1949–1959.
Gibbons, R. D., Garms, N. E., Jarke, F. H. and Stoub, K. P. (1992) Practical quantitation limits. *Chemometr. Intell. Lab. Syst.*, **12**, 225–235.
Hubaux, A. and Vos, G. (1970) Decision and detection limits for linear calibration curves. *Anal. Chem.*, **42**, 849–855.
International Union of Pure and Applied Chemistry Commission on Analytical Nomenclature (1995) Nomenclature in evaluation of analytical methods including detection and quantification capabilities. *Pure Appl. Chem.*, **67**, 1699–1723.
Krutchkoff, R. G. (1967) Classical and inverse regression methods of calibration. *Technometrics*, **9**, 425–439.
Mee, R. W., Eberhardt, K. R. and Reeve, C. P. (1991) Calibration and simultaneous tolerance intervals for regression. *Technometrics*, **33**, 211–219.
Oppenheimer, L., Capizzi, T. P., Weppelman, R. M. and Mehta, H. (1983) Determining the lowest limit of reliable assay measurement. *Anal. Chem.*, **55**, 636–643.
Osborne, C. (1991) Statistical calibration: a review. *Int. Statist. Rev.*, **59**, 309–336.
Shukla, G. K. (1972) On the problem of calibration. *Technometrics*, **14**, 547–553.
Williams, E. J. (1969) A note on regression methods in calibration. *Technometrics*, **11**, 189–192.
Zerbe, G. O. (1978) On Fieller's theorem and the general linear model. *Am. Statistn*, **32**, 103–105.