Pergamon

0731-7085(95)01310-5

Journal of Pharmaceutical & Biomedical Analysis, Vol. 13, No. 4/5, pp. 385–394, 1995

0731-7085/95 $9.50 + 0.00

Cross-validation of bioanalytical methods between laboratories*

MARY T. GILBERT,† IRINA BARINOV-COLLIGON and JOY R. MIKSIC

Rhône-Poulenc Rorer, 500 Arcola Road, Collegeville, PA 19426, USA

Abstract: Increased reliance on pharmacokinetic studies in regulatory submissions emphasizes the need for cross-validating bioanalytical methods between different laboratories to allow comparison of data. Globalization of pharmaceutical development results in a greater need to define cross-validation standards. A strategy for performing cross-validation experiments using prepared biological samples of known concentration and "real" samples from clinical trials is presented. The statistical techniques used to compare data sets and establish acceptability of the assays are illustrated by practical examples.

Keywords: *Validation; cross-validation; bioanalytical chemistry; comparison studies.*

Introduction

There has been a considerable emphasis in recent years on the standardization of validation procedures for bioanalytical assays. An international conference focused on Analytical Methods Validation: Bioavailability, Bioequivalence and Pharmacokinetic Studies in Washington, DC in December 1990 [1, 2]. Less attention has been paid to the topic of comparison studies. This issue has become increasingly more important with the advent of "global" development projects resulting in clinical studies and bioanalytical programs being conducted in more than one location. Additionally, as resources become stretched and fast-track development increases in importance, more pharmaceutical companies are contracting out assay work.

Various texts (e.g. [3, 4]) present the statistical procedures used to compare two methods or laboratories. Schemes for evaluating comparison studies have been presented for clinical chemistry assays [5, 6] and analytical chemistry methods [7]. Westgard and Hunt [8] while examining the potential of the various statistical tests for determining errors in comparison studies, concluded that linear regression probably provides the most useful information. More rigorous linear regression models which consider the effect of errors associated with both variables have recently been applied to the comparison of analytical methods [9].

Bioanalytical methods employed in pharmacokinetic and toxicokinetic studies present particular problems for comparative studies. The assays are often performed over several orders of magnitude and the concentration dependence of the assay variance may become significant. The statistical tests are limited by the relatively small amount of subject sample available. Also, historical data banks are not generally available for statistical examination.

Methods

Data sets

A simulated data set of concentration values was generated and manipulated to mimic error types. Systematic errors were simulated by the addition of a fixed or relative concentration to each reference value. Random errors were produced in the data set by using a normal distribution of random numbers generated by Microsoft Excel, 4.0.

Inter-laboratory cross-validations were performed by analysing both prepared controls and samples collected from subjects in clinical studies. The samples were analysed in two

* Presented at the Fifth International Symposium on Pharmaceutical and Biomedical Analysis, Stockholm, Sweden, September 1994.

† Author to whom correspondence should be addressed.

different laboratories by similar analytical methods.

Each experimental data set was examined using linear regression and a paired *t*-test to determine the various types of errors in the comparative data and to establish criteria for evaluating the laboratories. Statistical calculations were performed in Microsoft Excel, 4.0. Weighted linear regression was performed using SAS, 6.08. Data obtained from replicate analyses of standards or controls performed during the validation of the assay in each laboratory were used to determine the assay reproducibility.

Cross-validation scheme

A suggested scheme for performing a cross-validation experiment is presented in Fig. 1. Although the same procedure could apply to the cross-validation of one assay in two different laboratories (determination of inter-laboratory precision) or to the comparison of two different methods for the determination of the same analyte (verification of accuracy), only the former process will be discussed.

The first step is to establish and validate the assay in both laboratories according to accepted standards of accuracy, precision and specificity [1, 2]. Obtaining an accurate estimate of precision across the concentration range is necessary to determine appropriate statistical tests. A thorough review of all laboratory documentation and SOPs related to the assay procedure should be undertaken to ensure consistency in preparation of standards, including matrix sources, calculation methods and data rounding techniques.

At this point a decision has to be made about

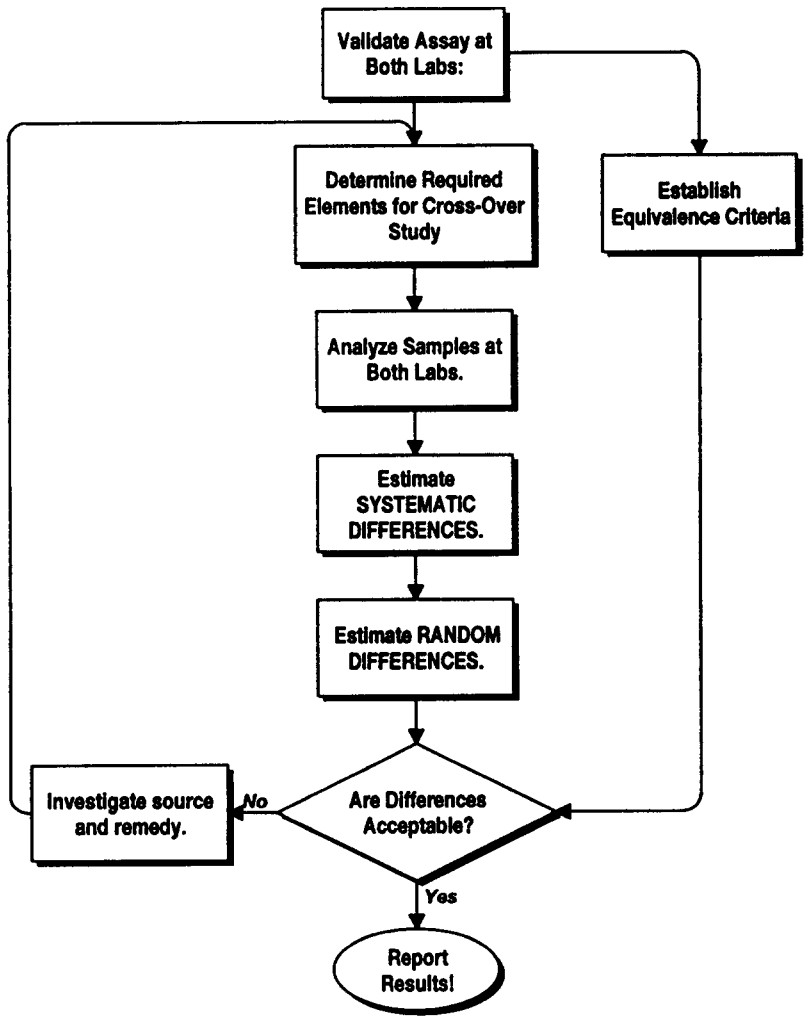

Figure 1
A proposed scheme for performing cross-validations between assays.

what constitutes acceptable reproducibility based on the performance of the assay in each laboratory, and the requirements of the study. An estimation of the expected inter-laboratory precision can be made from the precision measured during the validation in each individual laboratory.

Duplicate sets of samples are then assayed simultaneously. The sample sets should include both samples of known concentration (controls), prepared in the appropriate matrix, and a series of subject samples obtained during a study. The sample concentrations should be chosen to cover the calibration range of the assay, or the full range expected to be encountered during a study. The number of samples necessary to achieve adequate power to determine a difference between two methods can be calculated based on the method variability and statistical probability [10]. Generally, for the comparison of two assays with similar precision, a sample size of 20–30 samples is sufficient to detect a difference of 1 RSD% at the 0.05 significance level with a power of 90–95%.

Once the samples have been analysed the data sets are compared and the differences between the two determinations evaluated by the methods described below.

Data analysis — systematic error

Linear regression is the simplest and most common tool used to assess systematic error. The technique is applicable to the determination of both constant (fixed) systematic error and proportional (relative) systematic error. Non-weighted linear regression is often suitable as a first approximation. However, it is more appropriate to use weighted linear regression in cases where variance increases across the concentration range [11]. In chromatographic methods the increase in variance is approximately proportional to the concentration, indicating a weight of $1/x$ is appropriate.

Either laboratory may be arbitrarily chosen as the reference, but, if the assay precision is significantly different in the two laboratories, the more precise laboratory should be used as the reference. The data generated by the test laboratory are compared to those produced by the reference laboratory using least squares or linear regression analysis. If the only differences between the determinations in the two laboratories are caused by random errors

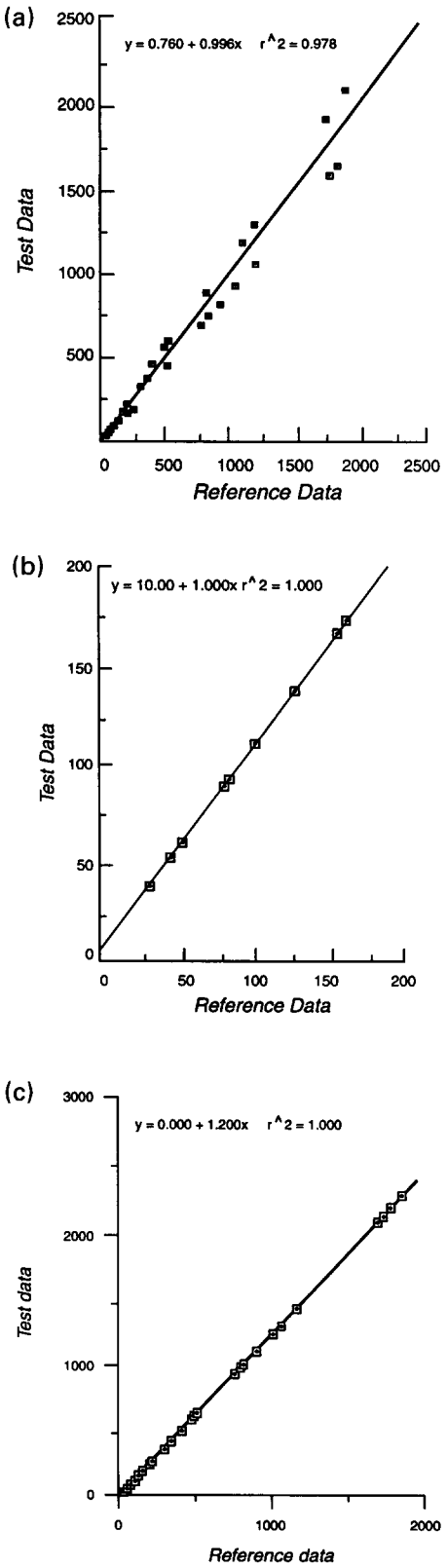

Figure 2
Representative linear regression plots illustrating the effects of different errors: (a) random error; (b) constant error; (c) proportional error.

associated with the measurements, the linear regression analysis will produce a straight line passing through the origin, with a slope of 1, as shown in Fig. 2(a). The data points will be evenly scattered about the 45° line. Any deviation from this situation indicates the presence of systematic errors. However, these errors must be evaluated to determine if they are significant.

A non-zero intercept gives a measure of the *constant error*. The simulated data shown in Fig. 2(b) show a fixed bias of 10 ng ml^{-1}. The intercept obtained from such a regression analysis is only an estimate of the true intercept. A confidence interval for the true intercept is given by the following equation

$$CI = b \pm t\,(SE),$$

where b is the calculated intercept, t is taken from the t table with $(n - 2)$ degrees of freedom at the appropriate significance level (commonly, $\alpha = 0.05$), SE is the standard error of the intercept, obtained from the regression analysis, and n is the number of data points.

The existence of a fixed bias is indicated if $b = 0$ does not fall within the confidence interval.

Alternatively, a significance test on the intercept can be performed by computing

$$t = b/SE.$$

If the value calculated for the t statistic is greater than the critical value found in the t table with $(n - 2)$ degrees of freedom, then a fixed bias exists.

If the presence of a fixed bias is detected, the cause of the interference should be investigated and, if possiblie, eliminated prior to repeating the comparison experiment. Once a constant error has been identified and quantified it is always possible to adjust the test data set to remove this bias prior to proceeding further with the analysis.

The deviation of the slope from 1 indicates the presence of a *proportional error* as shown in Fig. 2(c). Once again the confidence interval for the true slope should be calculated from

$$CI = m \pm t\,(SE),$$

where m is the calculated slope, t is obtained from the table with $(n - 2)$, degrees of freedom, and SE is the standard error of the slope, obtained from the regression analysis.

A proportional error exists if $m = 1$ does not lie within this confidence interval.

The significance of the difference of the calculated slope from unity can also be tested using a t test, where

$$t = (m - 1.00)/SE.$$

If the calculated value of the t statistic is greater than the appropriate value in the t table, the existence of a proportional error is indicated.

The use of least squares linear regression may, however, produce an underestimation of the slope and intercept. Ordinary linear regression applies to the situation where only the dependent variable (Y) is subject to error and the independent variable (X) consists of constant values. In a cross-validation experiment both the reference data (X) and the test data (Y) are subject to measurement error and ordinary least-squares may be inappropriate. In this case, the measurement error model [9] would give better estimates of the slope and intercept. If outliers are present in the data, least squares based regression techniques can be unreliable and a robust linear regression technique may be more appropriate [9].

Data analysis — random error

It is common to look at the differences between the paired individual concentration values to examine the *random error*. The standard deviation of the differences (SD_d) provides a measure of the random error. However, when working with a typical bioanalytical assay that covers a wide concentration range, the significance of this random error may be difficult to interpret. The simulated data set shown in Table 1 contains a random error of approximately 10%. The random error, estimated from SD_d, is 51.8 ng ml^{-1}. An error of this magnitude is considered acceptable at the higher concentration levels but would be unacceptable at the lower end of the calibration curve.

Since SD_d is affected by systematic errors, it will not give an accurate estimate of the random error in the presence of a large systematic error. In this case, the standard error of the y-estimate (S_y) obtained from the linear regression analysis, provides a better estimate of the random error since it is un-

Table 1
Estimation of random error

Ref.	Test	Diff.	% Diff.
1175	1351	176	15.0
1182	1159	−23	−1.9
1081	1033	−48	−4.4
1026	1090	64	6.3
419	467	48	11.0
777	821	44	5.6
522	631	109	20.8
507	527	20	3.9
200	183	−17	−8.5
504	502	−2	−0.4
514	523	9	1.8
501	538	37	7.4
486	469	−17	−3.4
1820	1940	120	6.6
1892	1844	−48	−2.6
1764	1861	97	5.5
1735	1750	15	0.8
916	929	13	1.4
810	746	−64	−7.9
829	938	109	13.1
345	373	28	8.1
210	241	31	14.6
156	163	7	4.3
223	258	35	15.7
301	288	−13	−4.3
127	146	19	14.7
205	168	−37	−18.2
198	226	28	14.3
162	160	−2	−1.1
101	97	−4	−4.0
82.5	81.9	−0.6	−0.7
79.3	78.1	−1.2	−1.6
51.2	55.4	4.2	8.1
43.0	43.9	0.9	2.1
Mean		21.6	3.6
SD		51.8	8.4
t		2.44	2.51
	t critical =	2.0322	

affected by the presence of systematic errors [8].

A more meaningful estimate of the random error is obtained by calculating a difference normalized to the reference concentration (% difference). The standard deviation (8.4% in Table 1) provides a useful measure of random error across the concentration range. In addition, examination of the normalized differences allows observation of possible outliers. A significance test such as Dixon's [12] can be used to determine whether or not the data point is a true outlier and can, therefore, be excluded from the data set.

A common statistical procedure used to compare two sets of data and to determine whether the difference between them is significant is the paired *t* test. The *t* statistic is calculated from the mean of the differences (d) and the standard error of the differences ($SD_d/\sqrt{n}$)

$$t = \frac{d}{SD_d}\sqrt{n}.$$

The value obtained for the *t* statistic is compared to the critical *t* found in the tables for a two-sided test at the 5% significance level ($\alpha = 0.05$) with $(n - 1)$ degrees of freedom.

Since *t* is determined from the ratio of the mean difference (bias) and the SD_d (random error), it is possible to obtain a low value for *t* when both constant error and random error are large. Similarly, it is possible to obtain a significant value for *t* when both types of error are small. In the data set shown in Table 1, a random error was simulated by multiplying each reference value by a random number from a normal distribution. The largest difference between any pair of values is 20.8%. The mean difference (21.6 ng ml^{-1}, 3.6%) and SD_d (51.8 ng ml^{-1}, 8.4%) are both small, and the data sets would, probably, be considered comparable. The resultant *t* is, however, greater than *t* critical (2.03). It is, therefore, recommended that the *t* statistic should not be used as an acceptance or rejection criterion, although the components of the statistic may provide useful information about the errors present.

Acceptance criteria

Acceptability is, to a certain extent, a matter of judgement based on the purpose for the assay comparison. More stringent criteria would be required if data from both assays were to be combined in a statistical analysis. Such a situation may occur in a population pharmacokinetic study. Slightly less strict criteria may be tolerated where data sets are not expected to be combined although cross-study comparisons may be performed.

Once the magnitude of the errors has been estimated criteria should be applied to determine if these are within acceptable limits. A decision on whether or not the differences are reasonable can be made by considering the precision of the assay in the two laboratories.

A combined standard deviation is obtained by adding the variances of the two assays

$$\text{combined SD} = \sqrt{(SD_1)^2 + (SD_2)^2}.$$

Using the inter-assay precision data obtained

during the validation of the assays, it is possible to calculate the expected combined standard deviation. Since the variance increases with concentration, this value should be calculated at several different concentrations, sufficient to cover the calibration range.

The expected combined assay precision at any concentration level, can be back calculated from the standard deviation and the concentration

$$\text{combined \% RSD} = [\text{SD}_{\text{combined}}/\text{conc.}] \times 100.$$

The calculated combined assay precisions for two assays with individual precisions from 5 to 20% are reported in Table 2.

Table 2
Predicted combined % RSD for two assays of known precision

Assay A % RSD	Assay B % RSD			
	5	10	15	20
5	7.1	11.2	15.8	20.6
10	11.2	14.1	18.0	22.4
15	15.8	18.0	21.2	25.0
20	20.6	22.4	25.0	28.3

As a first approximation, if the estimated random error is less than or equal to the combined assay precision the cross-validation can be considered successful.

The combined standard deviation can also be used to calculate the expected reproducibility factor, r

$$r = 2\sqrt{2}\text{SD}_{\text{combined}},$$

where 2 is an approximation for t, and $\sqrt{2}$ is $\sqrt{n}$, the number of assays.

At least 95% of the duplicate assay determinations should agree within $\pm$ r for a successful cross-validation. Once again, r should be calculated across the concentration range to take account of the increase in variance with concentration.

This is an approximation to the reproducibility limit, R, which can be statistically determined from the within laboratory standard deviation (s_w) and the between-laboratory standard deviation (s_b) [3, 7]

$$R = 2 \times \sqrt{2} \times S_R,$$

where

$$S_R = \sqrt{(s_w)^2 + (s_b)^2}.$$

R is defined as the maximum tolerable difference (with 95% confidence) between two individual determinations in two different laboratories. This latter procedure is normally more applicable to a situation where the assay is performed in several laboratories.

Results and Discussion — Application of the Procedures

Case 1

The data in Table 3 were obtained during a cross-validation experiment. The samples were assayed by an HPLC method that had been independently validated in each laboratory to required standards [1, 2], over a calibration range of 25–2500 ng ml^{-1}. Unweighted linear regression analysis of the data indicates an intercept of 11.80 ng ml^{-1} with a standard error of 18.80 ng ml^{-1}. The calculated confidence interval of -26.92 to 50.52 ng ml^{-1} contains zero so the intercept is not considered to be significant. Use of weighted linear regression analysis estimates the intercept at 12.26, but with a much narrower confidence interval (3.64, 20.87 ng ml^{-1}) which does not contain zero. Application of the errors in variables regression model proposed by Roy [11] calculates the intercept as insignificant (9.25 ng ml^{-1}; -25.92, 44.42).

The slope is, however, found to be significantly different from 1 by all three methods, since 1 is not contained within the confidence interval. These data indicate the presence of a systematic proportional error of between 22 and 27%.

A critical area to examine when a proportional error is detected is the preparation of standards. Particular attention should be paid to corrections that were made during the preparation of standard stock solutions such as corrections for the purity of the reference standard material or corrections for salt content. Comparison of stock standard solutions and batches of reference standard materials may help identify a problem.

Investigation in this case, traced part of the error to the preparation of standards. Two different lots of reference standard material had been used, and corrections were not made for adsorption of water in one of the labora-

Table 3
Comparison of data between laboratories — identification of proportional error

Lab. A value (ng ml^{-1})	Lab. B Value (ng ml^{-1})
4047	3050
3804	2900
3763	2950
2139	1900
952	773
470	408
328	271
221	177
103	91.3
57.1	46.6
52.0	48.2
42.9	37.9
41.9	35.4
2365	1750
3121	2160
2810	2070
1410	1010
932	650
486	346
351	253
155	134
107	89.2
78.2	71
45.5	44.3
37.6	72.1
30.8	34.1
27.3	28.5

Regression statistics

	LS	WLS ($1/x$)
Multiple R	0.9973	09964
R square	0.9946	0.9928
Standard error	77.2652	74.4279

		Coefficients	Standard error	t Statistic	Lower 95%	Upper 95%
LS*	Intercept	11.803	18.803	0.628	−26.924	50.529
	Slope	0.754	0.011	−22.190	0.731	0.776
WLS† ($1/x$)	Intercept	12.258	4.201	2.918	3.638	20.879
	Slope	0.753	0.013	−19.290	0.727	0.779
X-err‡	Intercept	9.253	17.139	0.540	−25.914	44.419
	Slope	0.755	0.011	−23.288	0.734	0.777

Observations	27
Degrees of freedom	25
t-critical =	2.05183

*Unweighted least squares regression.
†Weighted least squares regression.
‡Errors in variables regression.

Table 4
Comparison of data obtained in two laboratories for the analysis of plasma samples by similar HPLC methods — linear regression analysis

Lab. A value (ng ml^{-1})	Lab. B Value (ng ml^{-1})		Regression statistics					
1.18	0.992			LS	WLS ($1/x$)			
1.18	0.994							
1.08	0.966		Multiple R	0.99253	0.9933			
1.03	0.984		R square	0.98511	0.9867			
0.419	0.385		Standard error	0.06431	0.0600			
0.777	0.651							
0.522	0.480							
0.507	0.436							
0.200	0.195			Coefficients	Standard error	t Statistic	Lower 95%	Upper 95%
0.504	0.480							
0.514	0.491	LS*	Intercept	−0.403	0.030	−1.418	−0.106	0.021
0.501	0.463		Slope	0.979	0.028	−0.740	0.919	1.039
0.486	0.412	WLS† ($1/x$)	Intercept	−0.018	0.021	−0.870	−0.061	0.025
1.82	1.84		Slope	0.952	0.026	−1.831	0.897	1.007
1.89	1.85	X-err‡	Intercept	−0.050	0.030	−1.654	−0.113	0.013
1.76	1.59		Slope	0.986	0.028	−0.481	0.927	1.046
1.74	1.74							
0.916	0.901		Observations		20			
0.810	0.733		Degrees of freedom		18			
0.829	0.831		t-critical =		2.1009			

*Unweighted least squares regression.
†Weighted least squares regression.
‡Errors in variables regression.

tories. After the appropriate corrections were made a successful cross-validation was performed.

Case 2

Data obtained from another cross-validation experiment are presented in Table 4. The calibration range of the HPLC assay was 0.025–20 μg ml^{-1} in this case, but the drug concentrations measured in the study samples did not exceed 2 μg ml^{-1}. The assay accuracy and precision were each within 10% in both laboratories (15% at the minimum quantifiable limit in laboratory A).

An unweighted linear regression analysis shows a slope of 0.979 that is found not to be significantly different from one, using either the *t*-statistic ($t = 0.74$) or the 95% confidence interval about the slope (0.919–1.039). There is, therefore, no significant proportional error in this data set. Similarly the statistical tests for significance of the intercept, indicate that the intercept of −0.043 μg ml^{-1} is not significantly different from 0 and no significant fixed bias is determined between the two sets of data.

Use of weighted linear regression (Table 4) changes the estimate for the slope (0.952) and intercept (−0.018 ng ml^{-1}) but does not affect the conclusion about the significance of the errors. Similarly, application of the error in variables regression method indicates that the systematic errors are not significant.

The simplest approach for viewing the data to examine the random error is to prepare a frequency distribution of relative differences. In this method the number of pairs of samples that fall within certain percentages of each other are tabulated. The more sample pairs that fall in the lower part of the frequency distribution, the better is the agreement between the two assays. If any pairs fall well outside the acceptable limit an outlier test [12] can be used to check if the value can be rejected. A frequency distribution of this data set is shown in Table 5. Fifty per cent of the sample pairs fall within 5% of each other, and all data pairs are within 16.2%.

The calculation of the combined assay standard deviation and precision from the individual assay precisions is presented in Table 6. The observed random error of 5.8% is within the calculated inter-assay precision of 5.7–15.5%. The data are plotted in Fig. 3. The lines representing the reference $\pm r$ are included. All 20 values fall within the expected reproducibility criterion.

The assays have been shown to produce

Table 5
Determination of random error

	Diff. (A–B) (μg ml^{-1})	% Diff. (A–B)/A		Frequency %			
				0–5	5–10	10–15	15–20
	0.183	15.6					x
	0.188	15.9					x
	0.115	10.6				x	
	0.042	4.1		x			
	0.034	8.1			x		
	0.126	16.2					x
	0.042	8.0			x		
	0.071	14.0				x	
	0.005	2.5		x			
	0.024	4.8		x			
	0.023	4.5		x			
	0.038	7.6			x		
	0.074	15.2				x	
	−0.020	−1.1		x			
	0.042	2.2		x			
	0.174	9.9			x		
	−0.005	−0.3		x			
	0.015	1.6		x			
	0.077	9.5		x			
	−0.002	−0.2		x			
Mean	0.062	7.4	Total number	10	4	3	3
SD	0.064	5.8		50%	20%	15%	15%
t	4.385	5.7					
t-critical	2.093						

Table 6
Calculation of inter-assay standard deviation and precision

	0.025 μg ml^{-1}		0.400 μg ml^{-1}		10.0 μ ml^{-1}	
	A	B	A	B	A	B
Mean	0.0260	0.0255	0.417	0.380	10.1	10.8
SD	0.0037	0.0011	0.034	0.010	0.417	0.384
Var	1.37×10^{-5}	1.28×10^{-6}	0.001	9.59×10^{-5}	0.174	0.148
SD_{comb}	0.0039		0.036		0.568	
% RSD_{comb}	15.5		8.9		5.7	

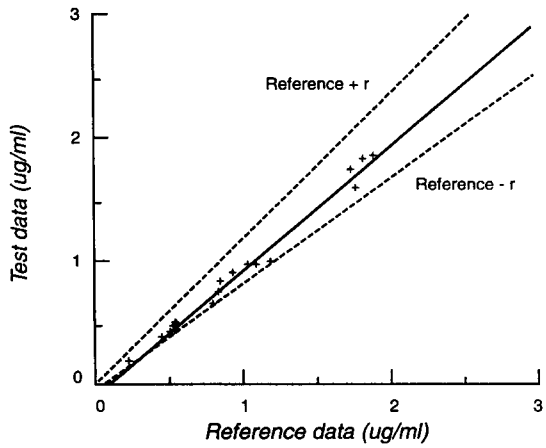

Figure 3
Linear regression plot of data from Laboratory B (test data) against data from Laboratory A (reference data). The calculated reproducibility limits are plotted as reference ±*r*.

comparable results using all these acceptance criteria.

Conclusion

Analysis of 20–30 samples in two different laboratories generally provides sufficient data to allow evaluation of the systematic and random errors associated with the two assays using simple statistical techniques.

The primary method of choice is least squares linear regression since the procedure is generally accessible and it provides good estimates for the intercept and slope. If the appropriate software is available a weighted linear regression should be used since this corrects for the change in variance throughout the calibration range. More rigorous methods that take account of the errors in both variables and are not affected by outliers should provide the best estimate of the systematic errors but these methods require sophisticated programs that are not routinely available in the bioanalytical laboratory.

Random error should be examined by calculating the normalized differences between the paired determinations. A frequency distribution can be a useful tool for viewing these differences. The expected random error can be calculated from the precision of the two individual assays, and acceptance criteria can be established accordingly. The judgement of acceptability is made based on the established criteria and the proposed use of the data.

Determination of systematic errors should lead to the development of standardized procedures between sites within a company, and with contract facilities. This will ultimately result in smaller differences between assays.

Acknowledgement — The authors would like to acknowledge Dr Eunhee Hwang of the Biostatistics Deparment for help with some of the statistical aspects of this paper.

References

[1] V.P. Shah, K.K. Midha, S. Dighe, I.J. McGilveray, J.P. Skelly, A. Yacobi, T. Layloff, C.T. Viswanathan, C.E. Cook, R.D. McDowell, K.A. Pittman and S. Spector, *J. Pharm. Sci.* **81**, 309–312 (1992).
[2] V.P. Shah, K.K. Midha, S. Dighe, I.J. McGilveray, J.P. Skelly, A. Yacobi, T. Layloff, C.T. Viswanathan, C.E. Cook, R.D. McDowell, K.A. Pittman and S. Spector, *Pharm. Res.* **9**, 588–592 (1992).
[3] R. Caulcutt and R. Boddy, *Statistics for Analytical Chemists*. Chapman and Hall, London (1989).
[4] D.L. Massart, A. Dijkstra and L. Kaufman, *Evaluation and Optimization of Laboratory Methods and Analytical Procedures*, Chapter 3. Elsevier, Amsterdam (1978).
[5] R.N. Barnett, *Am. J. Clin. Pathol.* **43**, 562–569 (1965).
[6] R.N. Barnett and W.J. Youden, *Am. J. Clin. Pathol.* **54**, 454–462 (1970).
[7] W. Horwitz, *Food Add. Contam.* **10**, 61–69 (1993).
[8] J.O. Westgard and M.R. Hunt, *Clin. Chem.* **19**, 49–57 (1973).
[9] T. Roy, *J. Pharm. Biomed. Anal.* **12**, 1265–1269 (1994).
[10] D. Mazzo and M. Connolly, *Pharm. Res.* **9**, 601–606 (1992).
[11] M. Davidian and P.D. Haaland, *Chemo. Intell. Lab. Sys.* **9**, 231–248 (1990).
[12] W.J. Dixon, *Biometrics* **9**, 74–89 (1953).

[*Received for review 21 September 1994; revised manuscript received 22 November 1994*]