STATISTICS IN MEDICINE, VOL. 9, 1463–1473 (1990)

ESTIMATION OF THE LINEAR RELATIONSHIP BETWEEN THE MEASUREMENTS OF TWO METHODS WITH PROPORTIONAL ERRORS

KRISTIAN LINNET

The Department of Clinical Chemistry, KK 7642, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, DK-2100 Copenhagen, Denmark

SUMMARY

The linear relationship between the measurements of two methods is estimated on the basis of a weighted errors-in-variables regression model that takes into account a proportional relationship between standard deviations of error distributions and true variable levels. Weights are estimated by an iterative procedure. As shown by simulations, the regression procedure yields practically unbiased slope estimates in realistic situations. Standard errors of slope and location difference estimates are derived by the jackknife principle. For illustration, the linear relationship is estimated between the measurements of two albumin methods with proportional errors.

INTRODUCTION

When the measurements of the same compound using two different methods are to be compared, it is commonplace to estimate a linear relationship between the measurements. If the slope of the line deviates significantly from unity, a proportional systematic error exists between the measurements of the two methods. A systematic proportional error may arise if the concentrations of the calibrator of one of the methods deviate by a common factor from those of the other method. A constant systematic error is present if the line does not pass through the origin (0, 0). For example, such an error may be caused by insufficient background correction in photometry or counts of radioactivity for one of the methods. Various statistical methods are available for estimation of a linear relationship.[1–3] A review of method comparison studies published in '*Clinical Chemistry*' over the first three months of 1988 revealed that ordinary least-squares regression analysis was used in 46 out of 50 cases. Deming's procedure[4] was applied in the remaining studies. Ordinary least-squares regression analysis presupposes that one of the methods is without random measurement error, and that the standard deviation of the error distribution of the other method is constant throughout the measurement range.[1] These assumptions are almost never fulfilled, because both methods are usually subject to random measurement errors, and the standard deviations of the error distributions are more likely to be proportional to the variable level than constant.[5,6] Violation of the first assumption causes a downwards bias of the slope estimate, the magnitude of which depends on the relation between the dispersion of the errors and the dispersion of the true values of the variables. In practice, the problem is more acute when the dispersion of the true values of the variables is small. Table I displays the ranges of serum concentrations of a series of clinical chemical compounds, expressed as 95 per cent reference

0277–6715/90/121463–11$05.50

Received January 1989
Revised May 1990

Table I. 95 per cent reference intervals for serum concentrations of some common clinical chemical analytes

	Units	95 per cent reference interval		Ratio
		Lower limit	Upper limit	(upper/lower)
Electrolytes				
Sodium	mmol/l	136	146	1·1
Calcium	mmol/l	2·10	2·55	1·2
Potassium	mmol/l	3·2	4·5	1·4
Metabolites				
Glucose	mmol/l	3·9	5·8	1·5
Creatinine	μmol/l	44	106	2·4
Urea	mmol/l	1·2	3·0	2·5
Enzymes				
Lactate dehydrogenase	U/l	45	90	2·0
Amylase	U/l	25	125	5·0
Creatine kinase	U/l	12	80	6·7
Proteins				
Albumin	μmol/l	550	810	1·5
Transferrin	μmol/l	22	49	2·2
Immunoglobulin M	mg/l	400	3450	8·6
Hormones				
Triiodothyronine	nmol/l	1·9	3·0	1·6
Estradiol	pmol/l	40	1800	45·0
Follicle stimulating hormone	IU/l	4	250	62·5

intervals for healthy subjects.[7] Electrolytes, some metabolites, and the serum protein albumin have limited ranges with ratios between upper and lower interval limits varying from 1·1 to 1·5. Although the ranges may be expanded somewhat by inclusion of specimens from diseased subjects, a slope bias problem may remain in connection with ordinary least-squares regression analysis. For example, suppose that the linear relationship between the measurements of two albumin methods is estimated by ordinary least-squares regression analysis and a range ratio of three is assumed with a uniform distribution of true values on the interval. This situation may be accomplished by sampling specimens from equal groups of healthy and diseased subjects. If, in addition, proportional measurement error with coefficient of variation of 0·05 is assumed for both methods then simulations show that the slope bias is $-0{\cdot}030$ for single measurements and $-0{\cdot}014$ for duplicate measurements of each specimen by both methods. The problem is overcome by using a regression model that takes into account random measurement errors for both methods. The appropriate procedure, which is called a functional or structural relationship model or an errors-in-variables regression method in the statistical literature,[2,8] has been termed in clinical chemistry Deming's method.[4] However, the method is based on the assumption of constant measurement errors, and, accordingly, is not optimal for the situation with proportional errors. The efficiency of Deming's method will be low, when estimating the linear relationship between substances such as hormones, which may have very wide distributions of serum concentrations (Table I). Furthermore, estimation of standard errors of parameter estimates[2] will not be correct. Special procedures[9,10] have been introduced for situations with heteroscedastic error distributions, but these methods are very complicated and not well-suited for clinical chemical problems. In this paper, a weighted modification of Deming's method is presented. The method is efficient for situations with proportional measurement errors, and standard errors are

correctly estimated by the jackknife method. The assumptions of the method are checked using a plot of residuals.

ESTIMATION OF THE LINEAR RELATIONSHIP BETWEEN THE MEASUREMENTS OF TWO METHODS WITH CONSTANT ERRORS

We distinguish between true (X_i) and observed (x_i) values for a compound measured by a chemical method. 'True' should here be understood as the average value of an indefinite number of repeated measurements of a specimen by the method, that is the expected value. Given two measurement methods for a compound, we have:

$$x_i = X_i + \varepsilon_i$$
$$y_i = Y_i + \delta_i.$$

The error terms ε_i and δ_i are supposed to be independent Normal variables with expected values zero. In clinical chemistry, the standard deviations of the distributions of the error terms are known as the analytical standard deviations. On the basis of a series of k paired measurements (x_i, y_i) from the two methods, a linear relationship is estimated by

$$\hat{Y}_i = a_0 + b\hat{X}_i = a + b(\hat{X}_i - \bar{X}). \tag{1}$$

The circumflex denotes *estimates* of the true values.

Suppose that the error variances $V(\varepsilon)$ and $V(\delta)$ are constant, and that their ratio $\lambda = V(\varepsilon)/V(\delta)$ is known. A set of k paired observations (x_i, y_i) are available for estimation of the linear relationship. Using a least-squares approach,[11] the sum of squares

$$S = \sum_{i=1}^{k} [(x_i - \hat{X}_i)^2 + \lambda(y_i - \hat{Y}_i)^2]$$

is minimized. This corresponds to minimizing the sum of squared deviations from the line at an angle determined by λ. For comparison, in an ordinary least-squares regression analysis, the sum of squared deviations in the vertical direction is minimized. In order to obtain a solution, the following sums are computed for all k pairs:[11]

$$u = \sum (x_i - \bar{x})^2, \qquad q = \sum (y_i - \bar{y})^2, \qquad p = \sum (x_i - \bar{x})(y_i - \bar{y}).$$

It can be shown that the slope estimate is

$$b = \{(\lambda q - u) + \sqrt{[(u - \lambda q)^2 + 4\lambda p^2]}\}/2\lambda p$$

and

$$a = \bar{y}.$$

In the equation for the line, $\bar{x}$ is substituted for $\bar{X}$. This is Deming's regression method,[4] also known as a structural or functional relationship model.[2] 'Structural' is for the case where the true values X and Y are random variables, whereas in a functional relationship X and Y are supposed fixed. The distinction is analogous to the two situations in ordinary least-squares regression analysis.[1] In practical situations, the distinction is of minor importance, because the computations are identical for the two cases.

The null hypothesis of identity that $Y = X$, is tested by the t-tests:

$$t = (b - 1)/\widehat{\mathrm{SE}}(b)$$
$$t = (a - \bar{x})/\widehat{\mathrm{SE}}(a - \bar{x}).$$

The latter t-test is a test of location difference which is independent of the test of the slope. This represents a more simple statistical analysis than testing whether the intercept deviates from zero.[1] Computation of the standard errors by the jackknife method is explained in a later section.

ESTIMATION OF THE LINEAR RELATIONSHIP BETWEEN THE MEASUREMENTS OF TWO METHODS WITH PROPORTIONAL ERRORS

If the error variances are assumed proportional to the squares of the average of the true values, then

$$V(\varepsilon_i) = f_x^2\,[(X_i + Y_i)/2]^2$$
$$V(\delta_i) = f_y^2\,[(X_i + Y_i)/2]^2.$$

Under this assumption, $\lambda = V(\varepsilon_i)/V(\delta_i) = f_x^2/f_y^2$ is constant. A weighted modification of the slope estimation procedure is to minimize the sum of squares

$$S_w = \sum_{i=1}^{k} [w_i(x_i - \hat{X}_i)^2 + \lambda w_i(y_i - \hat{Y}_i)^2]$$

where the weights

$$w_i = 1/[(X_i + Y_i)/2]^2$$

are inversely proportional to the error variances, and

$$\hat{Y}_i = a + b(\hat{X}_i - \bar{X}_w).$$

Denoting

$$\bar{x}_w = \sum_i w_i x_i \Big/ \sum_i w_i, \qquad \bar{y}_w = \sum_i w_i y_i \Big/ \sum_i w_i,$$

$$u_w = \sum_i w_i(x_i - \bar{x}_w)^2, \qquad q_w = \sum_i w_i(y_i - \bar{y}_w)^2,$$

$$p_w = \sum_i w_i(x_i - \bar{x}_w)(y_i - \bar{y}_w),$$

the slope estimate is

$$b = \{(\lambda q_w - u_w) + \sqrt{[(u_w - \lambda q_w)^2 + 4\lambda p_w^2]}\}/2\lambda p_w$$

and

$$a = \bar{y}_w.$$

In the equation for the line $\bar{x}_w$ is substituted for $\bar{X}_w$.

SIMULATIONS

The unweighted and weighted estimation procedures were evaluated by a simulation which mimics the problem of estimating a linear relationship between the measurements of two albumin methods. According to Table I, the measurement range extends from 550 to 810 μmol/l for 95 per cent of a healthy population. Liver disease causes a decrease of the serum concentrations. Thus, supposing that specimens from both diseased and healthy subjects are included, we may have

Table II. Simulated estimation of the slope of the linear relationship between two albumin methods using alternative Deming models. 5000 simulation runs for each case

f_x	f_y	λ		Unweighted	True weight	Estimated weight	Iterative estimated weight
0·05	0·05	1	$\bar{b}$	1·000	1·000	1·000	1·001
			SE(b)	0·025	0·021	0·021	0·021
0·05	0·15	$1/3^2$	$\bar{b}$	1·001	1·000	0·990*	1·000
			SE(b)	0·057	0·048	0·049	0·048
0·15	0·05	3^2	$\bar{b}$	1·001	1·001	1·012*	1·001
			SE(b)	0·058	0·049	0·050	0·048

* Different from 1 ($P < 0·001$)

specimens with concentrations varying from about 250 to 800 μmol/l, giving a range ratio of about 3. In order to estimate λ, we suppose that duplicate measurements are performed by each method, and that the mean values enter the statistical computations. Thus the k pairs $(\bar{x}_i = (x_{1i} + x_{2i})/2, \bar{y}_i = (y_{1i} + y_{2i})/2)$ are used for the simulation. The constant f_x^2 is estimated by

$$\hat{f}_x^2 = \left\{ \sum_i (x_{1i} - x_{2i})^2 / [(\bar{x}_i + \bar{y}_i)/2]^2 \right\} \Big/ 2k$$

with a similar expression for $\hat{f}_y^2$, and finally

$$\hat{\lambda} = \hat{f}_x^2 / \hat{f}_y^2.$$

Pseudo-random Normal numbers were generated by the NAG routine GO5DDF[12] according to the stated model in which the null hypothesis $Y = X$ was supposed. A uniform distribution of true values is assumed. This generated two random numbers x_{1i} and x_{2i} for X_i and two y_{1i} and y_{2i} for Y_i. The results of 5000 simulation runs, assuming $k = 50$, are summarized in Table II. For $f_x = f_y = 0·05$, both the unweighted and the weighted method yield unbiased slope estimates, but the standard error of the slope estimated by the weighted procedure is smallest. The shown standard errors are the standard deviations of the distributions of b for the simulation runs. The same pattern is observed for $\lambda = 1/3^2$ and 3^2, respectively, and where the coefficient of variation for one of the methods has been assumed equal to 0·15.

Although the weighted Deming method appears to be a straightforward procedure, there is a problem concerning the weights. In the simulation example the true weights

$$w_i = [(X_i + Y_i)/2]^{-2}$$

were used. In real situations, only *estimates* of the weights are available:

$$\hat{w}_i = [(\bar{x}_i + \bar{y}_i)/2]^{-2}$$

because the true values X_i and Y_i are unknown. Column 3 of Table II shows that the weighted Deming method with *estimated* weights yields biased slope estimates for λ values different from unity.

A possible way to resolve the problem is to try to get better estimates of the true values than those provided by the directly measured values. We may obtain an initial estimate of the linear relationship by using the unweighted method, that is by setting $w_i = 1$. According to Mandel,[11] we may derive estimates of the true values $(\hat{X}_i; \hat{Y}_i)$ for each point $(\bar{x}_i, \bar{y}_i)$ by a projection onto the

line. Denoting

$$d_i = \bar{y}_i - (a_0 + b\bar{x}_i)$$

we have

$$\hat{X}_i = \bar{x}_i + \hat{\lambda} b d_i/(1 + \hat{\lambda} b^2)$$

$$\hat{Y}_i = \bar{y}_i - d_i/(1 + \hat{\lambda} b^2).$$

We may now re-estimate the line using the weights

$$\hat{w}_i = [(\hat{X}_i + \hat{Y}_i)/2]^{-2}$$

and λ is also re-estimated by substituting $\hat{X}_i$ for $\bar{x}_i$ and $\hat{Y}_i$ for $\bar{y}_i$ in the formulae for $\hat{f}_x^2$ and $\hat{f}_y^2$. It is usual to iterate the procedure to obtain even better estimates of the true values. It turned out that the slope estimate usually converged towards a target value with successive differences of less than 10^{-4} within 3 to 4 iterations. Column 4 of Table II shows that this iterative method yields slope estimates with a negligible bias for the given situation, and that the standard error is almost identical to that of the true weight based method. Thus, this method seems suitable for estimation of the linear relationship between the two albumin methods under the given conditions.

The iterative method was also evaluated by a further series of simulations. These studies showed that the method was subject to a small bias of the slope estimate, the magnitude of which depended on the sample size and the ratio between the coefficients of variation of the error and true value distributions (the coefficient of variation of the error distribution refers to the error distribution of the means of duplicate measurements). In practice, these ratios are unlikely to exceed 0·25. For simulation with ratios not exceeding 0·25, slope biases of 0·0056 (sample size 20) to 0·0012 (sample size 100) were recorded. Simulation examples with fixed as well as random true values were also performed. The true values were assumed distributed according to Normal or log-Normal distributions and λ ranged from 10^{-2} to 10^{2}. The observed decrease of bias with increasing sample size suggests that the method is asymptotically unbiased. Further, the almost identical standard errors of the slope estimates based on true and iteratively estimated weights, respectively, suggests that the iterative method is asymptotically efficient (Table II).

The relative advantage of the weighted over the unweighted approaches depends on the degree of heteroscedasticity. Given a uniform distribution of true variables on the interval [1, 5], the variance of the slope estimate of the weighted approach equals 0·60 times that of the unweighted method for a sample size of 100. Simulations show that this relation is 0·49 for the interval [1, 10], dropping to 0·25 for the interval of [1, 100].

STANDARD ERRORS OF ESTIMATES

Assuming homoscedastic, Normally distributed measurement errors, expressions have been derived for the asymptotic standard error of the slope estimate.[2,8,13] However, it is difficult to derive a formula for the standard deviation, even if the Normality of error distributions are presumed, for the heteroscedastic case. Therefore, a non-parametric alternative such as the jackknife method seems appropriate.[13-15] We compute b^i = the slope based on the subset $\{(x_j; y_j)\}$ which does not contain $(x_i; y_i)$.

$$b_i = kb - (k-1)b^i \qquad \text{(the } i\text{the pseudovariate, } i = 1, \ldots, k)$$

where b is the slope estimated from the complete data set.

$$\tilde{b} = \sum_{i=1}^{k} b_i/k \qquad \text{(the jackknifed estimator)}$$

$$V_j(b) = \sum_{i=1}^{k} (b_i - \tilde{b})^2/(k-1)$$

$$\widehat{\mathrm{SE}}(b) = \sqrt{(V_j(b)/k)}.$$

Notice that this method allows for the uncertainty of the λ estimate, for example, based on duplicate measurements, because λ estimates based on the subsets are used for calculation of the pseudo-variates. In a similar way we can obtain a jackknife estimate of $\widehat{\mathrm{SE}}(a - \bar{x}_w)$.

The jackknife method for estimation of standard errors was evaluated by simulation assuming the model of the previously mentioned albumin example. Given a nominal significance level of 0·05 for test of slope deviation from unity and location difference, we expect 100 rejections out of 2000 simulation runs under the null hypothesis. The actual numbers of rejections were 103 and 100 for slope deviation and location difference, respectively. The overall number of rejections of the null hypothesis was 199. Other simulation examples also revealed good agreement between observed and expected rejection frequencies. Thus, the jackknife method seems to operate satisfactorily.

Finally, a variance ratio test for the regression model may be considered. If λ is known in advance and duplicate measurements are performed, independent estimates of the proportionality factors of the error variances can be obtained. From the dispersion of $\bar{x}_i$ and $\bar{y}_i$ about the line, we have[11]

$$\hat{f}_x'^2 = \{\lambda/[(k-2)(1+\lambda b^2)]\} \sum_i d_i^2/[(\hat{X}_i + \hat{Y}_i)/2]^2$$

$$\hat{f}_y'^2 = \hat{f}_x'^2/\lambda$$

where d_i is the vertical distance of a point from the line. The ratios between these estimates and those founded on the duplicate measurements, as described in the previous section ($\hat{f}_x^2/2$ and $\hat{f}_y^2/2$; the denominator 2 is used because we are dealing with $\bar{x}_i$ and $\bar{y}_i$) can be related to the critical value of the variance ratio statistic, with $(k-2)$ degrees of freedom for the numerator and k for the denominator. In most real examples, however, λ is not known in advance, but it is estimated from duplicate measurements. Under this condition, the two variance estimates are not independent, and the variance ratio test is not exact.

AN EXAMPLE FROM CLINICAL CHEMISTRY

In the previous sections, the problem of estimating a linear relationship between the measurements of two albumin methods has served as a model for the simulation studies. Now a concrete data example is considered. In the laboratory, two methods are available for determination of serum concentrations of albumin. Both methods are based on an antigen–antibody reaction, but they differ with respect to various technical details. A study is undertaken to evaluate whether there is any systematic difference between the measurements by the two methods. It is considered that a slope deviation of $\Delta\beta = 0{\cdot}1$ should be detected with a high probability, say 0·95 or higher. A type I error of 0·025 is chosen, because together with the same type I error for the test of location difference, the overall type I error becomes 0·05,. As mentioned previously, coefficients of

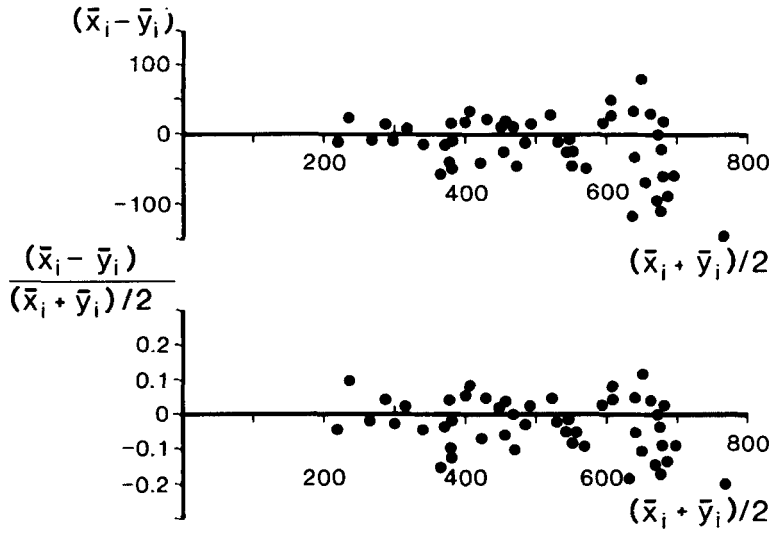

Figure 1. Absolute (top) and relative (bottom) differences between albumin measurements by two methods plotted against the average levels. The unit on the abscissa is μmol/l

variation of about 0·05 are expected for the error distributions, or $0{\cdot}05/\sqrt{2} = 0{\cdot}035$ for the means of duplicate measurements. Given these data together with an assumption of a uniform distribution over the interval 250–800 μmol/l, the power is determined as 0·98 for a sample size of 50, using jackknife standard errors from a 1000 runs. This suggests that approximately 50 observations are reasonable for this study. Duplicate measurements of 49 serum specimens by each method were actually carried out. Plots of absolute and relative differences between the measurements by the two methods against the average levels suggested proportional errors (Figure 1). Plots of differences between duplicate measurements by each method pointed in the same direction (Figure 2). These plots suggest that the weighted estimation procedure is appropriate. For comparison, both an unweighted and weighted regression analysis were performed. The equation for the line estimated by unweighted procedure is

$$\begin{aligned}\hat{Y}_i &= 515{\cdot}95 + 1{\cdot}0877\,(\hat{X}_i - 499{\cdot}1429)\\ &= -26{\cdot}99 + 1{\cdot}0877\,\hat{X}_i\end{aligned}$$

and by the weighted method:

$$\begin{aligned}\hat{Y}_i &= 417{\cdot}35 + 1{\cdot}0713\,(\hat{X}_i - 407{\cdot}37)\\ &= -19{\cdot}06 + 1{\cdot}0713\,\hat{X}_i.\end{aligned}$$

In this particular example, there is only a small difference between the two types of estimates, and so only the line estimated by the weighted method is shown in Figure 3, together with the line of identity. The estimates for f_x^2, f_y^2 and λ were 0·001681, 0·005641 and 0·298, respectively. These correspond to coefficients of variation of 0·041 and 0·074.

A plot of weighted residuals

$$r_i = \text{sign}\,[\hat{w}_i(\bar{x}_i - \hat{X}_i)^2 + \hat{w}_i\hat{\lambda}(\bar{y}_i - \hat{Y}_i)^2]^{1/2}$$

against $(\hat{X}_i + \hat{Y}_i)/2$ is displayed in Figure 4. The sign of r_i is identical with the sign of the vertical

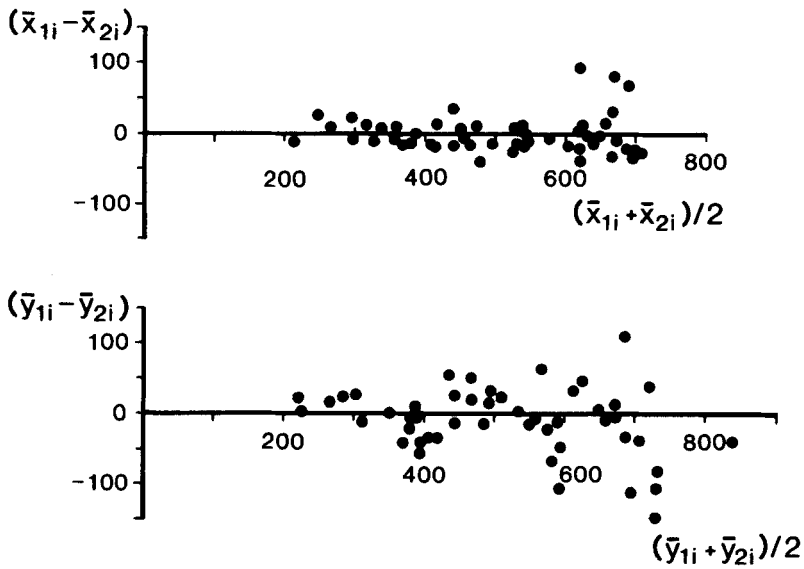

Figure 2. Absolute difference between duplicate measurements by the methods plotted against the means of the measurements

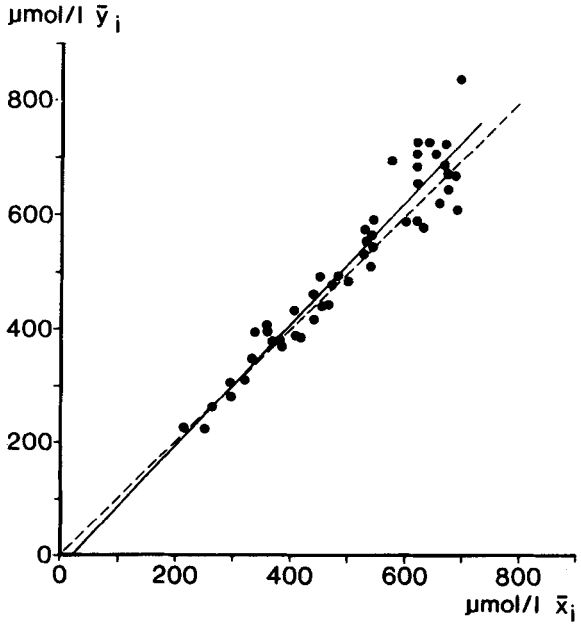

Figure 3. $(\bar{x}_i, \bar{y}_i)$ plot for the albumin example with the estimated regression line using the weighted Deming method (——). The line of identity is $Y = X$ (– – –)

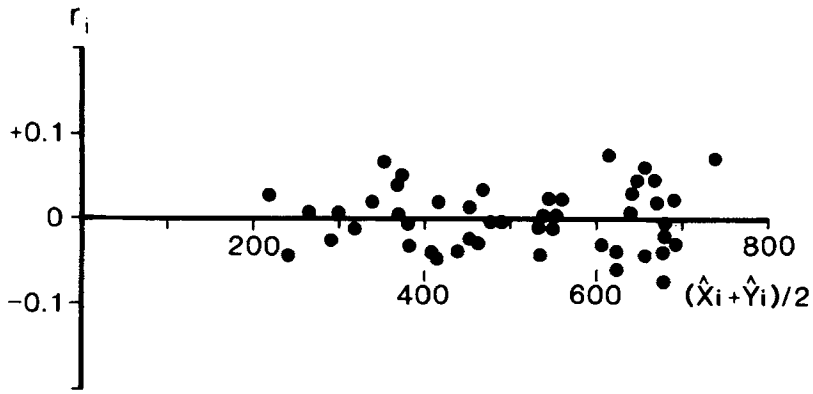

Figure 4. A plot of weighted residuals (r_i) against $(\hat{X}_i + \hat{Y}_i)/2$

distance d_i from the line. A fairly homogeneous scatter is observed, supporting the hypothesis of a linear relationship between X and Y with a proportional measurement error.

We proceed with tests for slope deviation from unity and location difference, using jackknife standard errors, then

$$t = (b - 1)/\widehat{\mathrm{SE}}(b) = 0{\cdot}0713/0{\cdot}0399 = 1{\cdot}79$$

and

$$t = (a - \bar{x}_w)/\widehat{\mathrm{SE}}(a - \bar{x}_w) = (417{\cdot}35 - 407{\cdot}37)/4{\cdot}8201 = 2{\cdot}07.$$

Given a significance level of 0·025 for each test, corresponding to $t = 2{\cdot}32$, we do not reject the null hypothesis of identity of the methods.

DISCUSSION

Estimation of the linear relationship between the measurements of two methods, both of which are subject to errors that depend on the variable levels, is a common problem in clinical chemistry and other fields of medicine. Although the focus has been on proportional measurement errors, other relationships might also be considered, particularly for immunoassays. Computer programs have been developed for estimation of variance functions on the basis of maximum likelihood theory.[16-18] In order to preserve the condition that λ is constant, however, one has to operate with a common form of the functions for both assays.

In this presentation, focus has been on clinical chemical problems, in which replicate measurements are available, so that the error variances and their ratio can be estimated. In clinical medicine, replicate measurements are not always attainable. If technical considerations make a variance ratio of one plausible, the analysis may be carried out on this basis. A plot of the differences $(y_i - x_i)$ against $(x_i + y_i)/2$ may suggest whether a weighted or an unweighted analysis is appropriate.

Comparison of methods on the basis of an estimated linear relationship yields more information than is achieved by merely testing for a location difference.[19,20] In clinical chemistry, one is *not* just interested in testing the null hypothesis of identity. If the null hypothesis is rejected, a recalibration of one of the methods is frequently performed in order to eliminate systematic

measurement differences. For this purpose, estimation of both slope and location difference is necessary.

REFERENCES

1. Snedecor, G. W. and Cochran, W. G. (eds) *Statistical Methods, 6*. Iowa State University Press, Ames, Iowa, 1967, Chapter 6.
2. Kendall, M. G. and Stuart, A. *The Advanced Theory of Statistics*, Charles Griffin & Co., London, 1961, Vol. 2, Chapter 29.
3. Hald, A. *Statistical Theory with Engineering Applications*, Wiley, New York, 1952, Chapter 18.
4. Cornbleet, P. J. and Gochman, N. 'Incorrect least-squares regression coefficients in method-comparison analysis', *Clinical Chemistry*, **25**, 432–438 (1979).
5. Ross, J. W. and Fraser, M. D. 'The effect of analyte and analyte concentration upon precision estimates in clinical chemistry', *American Journal of Clinical Pathology*, **66**, 193–205 (1976).
6. Kay, C., Nix, A. B. J., Kemp, K. W., Rowlands, R. J., Richards, G., Groom, G. V., Griffiths, K. and Wilson, D. W. 'Dose-interpolation of immunoassay data: uncertainties associated with curve-fitting', *Statistics in Medicine*, **5**, 183–192 (1986).
7. Tietz, N. W. (ed.) *Textbook of Clinical Chemistry*, W. B. Saunders & Co., Philadelphia, 1986, Chapter 20.
8. Kelly, G. E. 'Use of the structural equations model in assessing the reliability of a new measurement technique', *Applied Statistics*, **34**, 258–263 (1985).
9. Sprent, P. 'A generalized least-squares approach to linear functional relationships (with discussion)', *Journal of the Royal Statistical Society, Series B*, **28**, 278–297 (1966).
10. Barnett, V. D. 'Fitting straight lines–The linear functional relationship with replicated observations', *Applied Statistics*, **19**, 135–144 (1970).
11. Mandel, J. *The Statistical Analysis of Experimental Data*, Wiley, New York, 1964, Chapter 12.
12. Numerical Algorithms Group (NAG) *Fortran Library Manual*, Numerical Algorithms Group Ltd., Oxford, 1984.
13. Kelly, G. E. 'The influence function in the errors in variables problem', *The Annals of Statistics*, **12**, 87–100 (1984).
14. Brillinger, D. R. 'Discussion of paper by P. Sprent', *Journal of the Royal Statistical Society, Series B*, **28**, 294 (1966).
15. Wu, C. F. J. 'Jackknife, bootstrap and other resampling methods in regression analysis (with discussion)', *The Annals of Statistics*, **14**, 1261–1295 (1986).
16. Sadler, W. A. and Smith, M. H. 'A computer program for variance function estimation, with particular reference to immunoassay data', *Computers and Biomedical Research*, **20**, 1–11 (1987).
17. Sadler, W. A., Smith, M. H. and Legge, H. M. 'A method for direct estimation of imprecision profiles, with reference to immunoassay data', *Clinical Chemistry*, **34**, 1058–1061 (1988).
18. Davidian, M. and Carroll, R. J. 'Variance function estimation', *Journal of the American Statistical Association*, **82**, 1079–1091 (1987).
19. Altman, D. G. and Bland, J. M. 'Measurement in medicine: the analysis of method comparison studies', *Statistician*, **32**, 307–317 (1983).
20. Bland, J. M. and Altman, D. G. 'Statistical methods for assessing agreement between two methods of clinical measurement', *Lancet*, **i**, 307–310 (1986).