VOLUME 45, NO. 4 DECEMBER 1947

THE ADJUSTMENT OF BIOLOGICAL ASSAY RESULTS FOR VARIATION IN CONCOMITANT OBSERVATIONS

BY D. J. FINNEY

Lecturer in the Design and Analysis of Scientific Experiment, University of Oxford

(With 1 Figure in the Text)

1. INTRODUCTION

The measured responses of the individual test subjects used in a biological assay sometimes show an appreciable correlation (positive or negative) with the values of a concomitant variate which is itself unaffected by the drug or other substance under assay. For example, in the assay of insulin by the rabbit method, the decrease in blood sugar during the experimental period is affected not only by the dose of insulin injected but also by the level of blood sugar before injection. Again, in the assay of parathyroid extract by injection into dogs and measurement of the subsequent changes in the serum calcium, the responses produced by any specified dose may be correlated with the weights of the dogs. The precision of such an assay will be increased if the effects of this type of variation, irrelevant to the dose-response relationship which is the basis of the assay, can be eliminated. Pharmacologists and other users of assay techniques have frequently adopted arbitrary adjustments for the removal of concomitant variation. For example, when the response is the weight of some organ of the test subject, this may be expressed as a proportion of the total body weight and the proportion taken as a response metameter in the statistical analysis; similarly, in the insulin assay, the fall in blood sugar may be expressed as a percentage of the initial value. Alternatively, doses may be expressed 'per unit of body weight' instead of on an absolute scale.

An adjustment of this kind involves, by implication, an assumption relating to the form of the dependence of the response on the concomitant variate, and in fact assumes that the dependence is one of direct proportionality. Bliss & Marks (1939*a*, *b*; see especially pp. 97–100) have pointed out that there is often no sound reason for accepting this assumption, and have stated clearly the arguments in favour of preferring a method of adjustment derived by covariance analysis from the internal evidence of the data. They have described a system of calculation for use when the response is linearly related to the logarithm of the dose, and have illustrated it by reference to numerical results for an insulin assay using two doses of a standard and two of a test preparation; the percentage fall in blood sugar of injected rabbits is adjusted for inequalities in the initial level of blood sugar. In a later paper, Bliss (1940) has discussed a similar analysis for an assay of vitamin D by means of the ash content of the femur of rats, and has shown how allowance can be made for the effect of inequalities in the weight of organic matter in the bones, a quantity unaffected by vitamin D. Bliss & Rose (1940) have given examples, from assays of more complex design, showing how the serum calcium of dogs used in the assay of parathyroid extracts may be adjusted for inequalities in the initial body weights of the dogs. Fieller (1940), as part of a very full account of insulin assays using cross-over designs, has also discussed the adjustment of percentage fall in blood sugar by means of a covariance analysis on initial blood sugar. He was able to estimate a single regression coefficient from the combined evidence of many assays, and to use this in adjusting the responses for each assay. Providing that the regression coefficient does not vary significantly from one assay to another, this procedure possesses all the advantages emphasized by Bliss as appertaining to the adjustment determined from the data instead of by arbitrary choice; at the same time, Fieller's estimate of the adjustment will be more precise than any from a single assay, so that the complications which adjustment introduces into the assessment of errors and fiducial limits will be less important.

The analysis proposed by Bliss and others can be used whatever the number of doses tested, and whether or not they are equally spaced on the scale of the dose metameter, though the published examples involve only two or three doses of the standard and test preparations; it may be extended so as to cover simultaneous adjustment for several concomitant variates. In the form given by Bliss, however, the method is not exactly that of orthodox covariance analysis, as it takes no account of the correlation between dose and the concomitant

variate. By hypothesis, this variate is unaffected by the dose, so that the correlation should be small and attributable entirely to random causes; nevertheless, it will have a sampling value different from zero. Analysis according to the generally accepted procedure for covariance work is a little more laborious, since it requires consideration of the multiple regression of response on dose and the concomitant variate. The extra labour should be small once the routine has become familiar, and, indeed, the operations will be more natural for those accustomed to the covariance technique. For most practical purposes, however, the refinement in the analysis will make little difference and need seldom be undertaken by those who find difficulty in it.

The correct method does become rather tedious if full tests of linearity and parallelism of regression lines are needed. Though these qualities are usually essential to the validity of an assay, the significance of deviations from perfect linearity and parallelism can often be assessed satisfactorily from inspection of the data, or from tests based on unadjusted responses; unless the adjustments are large, their effect on the tests of validity is unlikely to be important. Nevertheless, the full method of testing will be described here, as some account of it should be placed on record.

The reader of this paper will be assumed to have some familiarity with the analysis of variance and with the estimation of multiple linear regression equations, as well as with ordinary biological assay analysis. To have given a full account of calculations for these standard statistical techniques would have made the paper inordinately long; details may be found in text-books of statistical science, such as those of Fisher (1946) and Snedecor (1946).

2. THE ANALYSIS OF UNADJUSTED RESPONSES

The calculations for the estimation of the potency of a test preparation relative to a standard, using an adjustment for a concomitant variate, follow closely the method of covariance analysis familiar to statisticians. There are, however, special features, particularly in the assessment of the precision of the estimate, which can be shown with reference to a numerical example more conveniently than by formula. For this purpose, data on a prolactin assay, from the Research Division of Glaxo Laboratories, will be used. The writer has been concerned only with the statistical analysis of the data, but he has been informed by Mr A. L. Bacharach that, in the assay procedure that was adopted, 'Suitable doses of the sample to be assayed and of the standard preparation are injected into groups of pigeons, as uniform as possible in respect of weight, age, and general appearance, and preferably of an inbred or "pure" line, the doses being given on six successive days by subcutaneous injection. The resultant enlargement of the crop gland is established for each animal separately, after it has been killed on the seventh day of the test, by weighing the gland emptied of "pigeon milk", cleaned of adventitious tissue, and preserved in 70% ethanol for 24 hr. Except in minor details, the test follows the well-known procedure originally laid down by Riddell, and described in more detail by Folley, Turner and others.'

The assay was based upon the crop-gland weights, and the concomitant variate for which adjustment was to be made was the body weight of the pigeons at the beginning of the assay. In this section, as a preliminary to the discussion of methods of adjustment, the analysis of the unadjusted responses will be described. Twenty-four pigeons were used in the assay, four on each of three doses of a standard prolactin preparation and four on each of three doses of a test preparation. Seven of the birds were female and the remainder male, but preliminary examination of the results showed clearly that sex differences in response were negligible, and, in order to avoid further complication of this account, sex has been ignored. Allowance for sex differences could in fact be made by an extension of the analysis, introducing a second concomitant variate which takes the value 1 for males and 0 for females. The body weights (u) and crop-gland weights (y) are shown in Table 1.

Table 1

(*a*) *Body weights of pigeons* (u)
(in units of 10 g.)

Dose of standard preparation (i.u.)			Dose of test preparation (mg.)		
1·25	2·50	5·00	0·125	0·250	0·500
49*	49	49	51*	48	45*
53	53	53	51	51	52
44	46*	41*	50	48*	50
49	51*	43	52	50	53
195	199	186	204	197	200

(*b*) *Crop-gland weights of pigeons* (y)
(in units of $\frac{1}{10}$ g.)

Dose of standard preparation (i.u.)			Dose of test preparation (mg.)		
1·25	2·50	5·00	0·125	0·250	0·500
38*	53	85	28*	48	60*
39	102	144	65	47	130
48	81*	54*	35	54*	83
62	75*	85	36	74	60
187	311	368	164	223	333

Pigeons marked * were ♀, the remainder ♂.

https://doi.org/10.1017/S0022172400014078 Published online by Cambridge University Press

If the mean response at each dose level is plotted against the logarithm of the dose, little experience of biological assay is needed in order to be satisfied that the data will be adequately fitted by two parallel regression lines, or in other words that neither deviations from linearity nor deviations from parallelism approach significance. Hence important conditions for the validity of the assay (Finney, 1948) appear to be satisfied. Since the doses for each preparation are in the ratio of 1 : 2 : 4, the arithmetic is simplified by the adoption of a logarithmic dose metameter based upon logarithms to base 2, and with arbitrary units of log dose x, such that its values are -1 for the low, 0 for the intermediate, and 1 for the high dose of either preparation. Mean responses are plotted against this scale in Fig. 1.

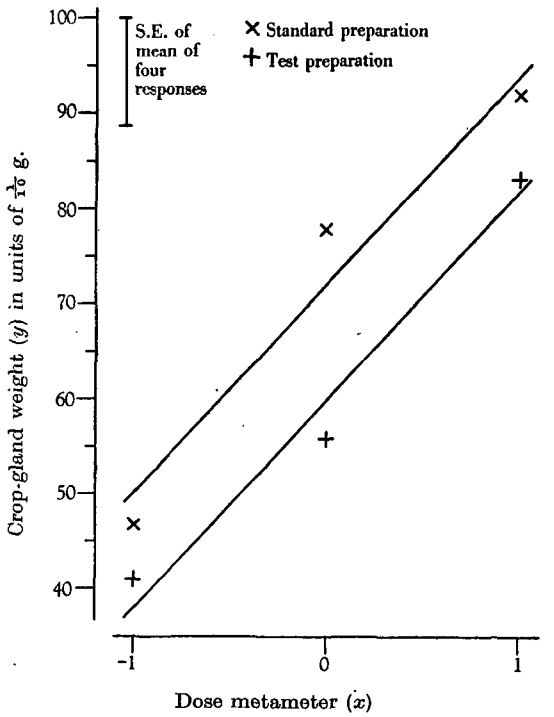

Fig. 1. Relationship between the mean crop-gland weights from Table 1 (*b*) and the dose metameter, without adjustment for body weight.

Bliss & Marks (1939*b*) and Bliss (1940) have shown how polynomial coefficients may be used in assays with equally spaced log doses, and especially in four-point and six-point designs, for expediting the calculations needed in exact tests of linearity and parallelism. Though inspection of Table 1 and Fig. 1 here suffices to show that these tests are not required, the complete calculations will be given for comparison with those for the adjusted responses discussed in § 5. The coefficients for this assay are listed in Table 2, and the subsequent calculations may be considered as the estimation of a linear regression of the response y, on x_1, x_2, x_3 and x_4; x_1 is identical with the dose-metameter x, and the regression coefficient on x_1 is used in the calculation of potency. A significant regression on x_2 would indicate that the linear regressions on log dose for the two preparations were of different slope. A significant regression on x_3 would indicate a departure from linearity as evidenced by the combined results for the two preparations, and a significant regression on x_4 would show an inequality in the deviations from linearity of the two series. Any of these results would cast grave doubts on the validity of the assay, or at least on its analysis in terms of a linear regression on log dose (Finney, 1948).

Table 2. *Table of polynomial coefficients for tests of assay validity*

Dose	x_1	x_2	x_3	x_4
Standard				
1·25	−1	1	1	−1
2·50	0	0	−2	2
5·00	1	−1	1	−1
Test				
0·125	−1	−1	1	1
0·250	0	0	−2	−2
0·500	1	1	1	1

Since x_1, x_2, x_3 and x_4 have been chosen so as to be mutually orthogonal, the regressions on each may be calculated independently of the others. The sums of squares and the sums of products with y are required for each x; these should be the 'within preparations' components with 22 degrees of freedom, but, since the mean value of any x is the same (zero) for each preparation, the 'between-preparations' component is zero for each of these sums of squares and products. The sums are:

$$Sx_1^2 = 16, \quad Sx_1y = 350,$$
$$Sx_2^2 = 16, \quad Sx_2y = -12,$$
$$Sx_3^2 = 48, \quad Sx_3y = -16,$$
$$Sx_4^2 = 48, \quad Sx_4y = 118.$$

Each variate is responsible for a portion $[Sxy]^2/Sx^2$ of the 'within-preparations' sum of squares for y. The analysis of variance of the values of y in Table 1 (*b*) may now be completed in the form shown in Table 3.

The non-significance of the regressions on x_2, x_3 and x_4, apparent from Fig. 1, is confirmed by Table 3. Since the analysis has been cast in this form, in strictness the residual mean square with 18 degrees of freedom should be used as the estimate of error variance; any amalgamation of other mean squares which have been examined and found non-significant with the residual introduces a risk of

https://doi.org/10.1017/S0022172400014078 Published online by Cambridge University Press

bias into the error variance. Here, however, the complete analysis was really unnecessary, and for purposes of comparisons with results to be obtained in subsequent sections an error variance will be formed by combining the x_2, x_3 and x_4 components with the residual; this gives

$$s^2 = \tfrac{10686}{21}$$
$$= 508{\cdot}9 \text{ with 21 degrees of freedom.} \quad (1)$$

Table 3. *Analysis of variance for unadjusted crop-gland weights*

	D.F.	Sum of squares	Mean square
Between preparations	1	888	—
Regression on x_1	1	7,656	7,656
Regression on x_2	1	9	9
Regression on x_3	1	5	5
Regression on x_4	1	290	290
Residual	18	10,382	576·8
Total	23	19,230	

From the sums of squares and products already quoted, the regression coefficient of y on x_1 (or x) is

$$b = \tfrac{350}{16}$$
$$= 21{\cdot}875.$$

Hence the two regression lines are

$$Y_s = \bar{y}_s + b\,(x - \bar{x}_s)$$
$$= 72{\cdot}167 + 21{\cdot}875x,$$

and
$$Y_t = 60{\cdot}000 + 21{\cdot}875x,$$

since for the mean log dose $\bar{x}_s = \bar{x}_t = 0$. The estimate of relative potency is derived from M, the difference between x-values which give equal values to the expected responses Y_s and Y_t. Hence

$$M = \bar{x}_s - \bar{x}_t - \frac{\bar{y}_s - \bar{y}_t}{b} \quad (2)$$
$$= -\frac{12{\cdot}167}{21{\cdot}875}$$
$$= -0{\cdot}5562.$$

Application of the usual formula for the variance of this estimate (see, for example, Finney, 1948) gives

$$V(M) = \frac{s^2}{b^2}\left[\frac{1}{6} + \frac{(M - \bar{x}_s + \bar{x}_t)^2}{16}\right] \quad (3)$$
$$= \frac{508{\cdot}9}{(21{\cdot}875)^2}[0{\cdot}1667 + 0{\cdot}0193]$$
$$= 0{\cdot}1978,$$

and therefore

$$M = -0{\cdot}5562 \pm 0{\cdot}4447. \quad (4)$$

As is well known, however, the standard error calculated in this way can safely be used in deriving fiducial limits for M only if b is very large compared with *its* standard error. A convenient criterion is provided by

$$g = t^2 V(b)/b^2 \quad (5)$$
$$= \frac{(2{\cdot}080)^2 \times 508{\cdot}9}{16 \times (21{\cdot}875)^2}$$
$$= 0{\cdot}288;$$

t is the deviate corresponding to the probability level (here 5 %) selected for the fiducial limits, and may be read from Fisher & Yates's (1947) Table III with the number of degrees of freedom appropriate to s^2. If g is almost zero (say less than 0·1), expression (6) below will be practically unaffected by g, and multiplication of t by the standard error derived from equation (3) will give the width of the fiducial interval on either side of M. If the assay is to provide any useful estimate, g must be less than 1, for g greater than 1 would imply that the regression coefficient of response on x was not significantly different from zero. The exact formula for the fiducial limits is

$$M + \frac{g}{1-g}(M - \bar{x}_s + \bar{x}_t) \pm \frac{st}{b\,(1-g)}\sqrt{\left(\frac{1-g}{6} + \frac{(M - \bar{x}_s + \bar{x}_t)^2}{16}\right)} \quad (6)$$
$$= -0{\cdot}5562 - 0{\cdot}2250 \pm \frac{2{\cdot}080}{21{\cdot}875 \times 0{\cdot}712}\sqrt{70{\cdot}228}$$
$$= -0{\cdot}7812 \pm 1{\cdot}1192$$
$$= 0{\cdot}3380 \text{ and } -1{\cdot}9004;$$

formulae equivalent to (6) have been given in earlier publications, notably by Fieller (1940) and Irwin (1943), but the writer finds the form given here the most convenient for calculation, especially as it shows clearly whether or not g is important. From the relationship between the x-scale and the true dose levels in Table 1, the value of M in equation (4) is seen to correspond to a potency for the test preparation of

$$10 \times 2^{-0{\cdot}5562} = 6{\cdot}80 \text{ i.u./mg.} \quad (7)$$

Similarly, the fiducial limits just calculated correspond to 12·6 and 2·68 i.u./mg.

3. ADJUSTMENT BY PROPORTIONALITY

Inspection of Table 1 suggests that the crop-gland weights of birds receiving the same dose are correlated with the body weights. A common, and perhaps a natural, method of taking account of correlation of this kind so as to increase the precision of the assay is to express the weight of the organ used for the assay as a proportion of the body weight and to perform the subsequent calculations on the proportions instead of on the absolute weights. It might seem more logical to express the gland weights as proportions of final, rather than initial,

https://doi.org/10.1017/S0022172400014078 Published online by Cambridge University Press

body weights, but there is a danger that the sensitivity of the new response metameter to changes in dose would be reduced if body weight were itself affected by the substance under assay. In this example, body weight showed no appreciable effect of prolactin dosage, but the general practice of using initial weights which necessarily cannot have been influenced by subsequent treatment (assuming, of course, that an unbiased randomization procedure has been adopted in allocating pigeons to doses) is to be preferred. For arithmetical convenience, the proportions may be calculated to give y (in Table 1) as a percentage of u; the response metameters so obtained are given in Table 4.

Table 4. *Proportional crop-gland weights* ($100y/u$)

Dose of standard preparation (i.u.)			Dose of test preparation (mg.)		
1·25	2·50	5·00	0·125	0·250	0·500
78	108	173	55	100	133
74	192	272	127	92	250
109	176	132	70	112	166
127	147	198	69	148	113
388	623	775	321	452	662

The figures in Table 4 may be analysed exactly as were those for y in § 2. The analysis of variance is given in Table 5; again there are clearly no difficulties about the validity of the assay. Equations (2) and (3) lead to

$$M = -0{\cdot}6429 \pm 0{\cdot}3818, \qquad (8)$$

which, as was to be expected, is a rather more precise estimate than that in equation (4). The variance s^2, is now relatively smaller (1568 if the x_2, x_3 and x_4 components are again included with the error), so that g, calculated from equation (5) at the 5% probability level, takes the lower value 0·205. The formula (6) now gives 0·0973 and $-1{\cdot}7147$ for the fiducial limits of M. Calculation as in equation (7) leads to 6·40 i.u./mg. as the estimated potency of the test preparation, with fiducial limits at 10·7 and 3·05 i.u./mg.

4. ADJUSTMENT BY COVARIANCE

The adjustment of responses by expressing them as proportions of the concomitant variate may often be very effective in improving the precision of an assay. Nevertheless, as has been pointed out in the Introduction to this paper, it involves an arbitrary assumption about the nature of the relationship between the two variates which may often be hard to justify; the proportionality adjustment might indeed lower the precision of assays in which it happened to be inappropriate, even though there was a marked correlation between the response and the concomitant variate (an obvious example is the situation in which the correlation is negative). A further disadvantage of the method is that its application is restricted to a single concomitant variate, though assays may occur in which simultaneous adjustment for two or more variates is needed.

Bliss and others have given adequate reasons for preferring a method of taking account of concomitant variation which estimates the appropriate adjustment from the assay data themselves, instead of determining it arbitrarily. They recommend the calculation of a covariance analysis for y, the response, and u, the concomitant variate. The error regression of y on u is then taken as the estimated effect on individual responses of unit change in u. This figure is used for the adjustment of the mean response at each dose level to the basis of a mean value for u, and the relationship between the adjusted mean response and the dose metameter x, is then used for the assay just as were the unadjusted responses in § 2; the estimation of error, of course, is complicated by this procedure. The method, in fact, entails first the estimation of the regression of y on u, and then the estimation of the regression of adjusted values of y on x; a procedure more in accordance with statistical practice, and a more satisfactory way of studying the relationship between x, u and y, is to estimate a multiple regression equation of y on x and u simultaneously. This may require slightly more laborious calculations than Bliss's method, though the difference is not great unless detailed validity tests are necessary. On the other hand, the estimation of error and the assessment of fiducial limits is accomplished more satisfactorily by the method proposed here. If there were no correlation between u and x, the two methods would be identical; in practice, even though u is a pre-treatment measurement which cannot be correlated with dose in the whole population, there will be sampling variation from zero correlation within any assay, and it is desirable that this be taken into account.

Table 5. *Analysis of variance for proportional crop-gland weights*

	D.F.	Sum of squares	Mean square
Between preparations	1	5,133	—
Regression on x_1	1	33,124	33,124
Regression on x_2	1	132	132
Regression on x_3	1	0	0
Regression on x_4	1	547	547
Residual	18	32,244	1,791
Total	23	71,180	

https://doi.org/10.1017/S0022172400014078 Published online by Cambridge University Press

In the prolactin assay, preliminary examination makes clear that the data are satisfactorily fitted by two parallel straight lines. After the analysis of variance shown in Table 3, there is no real need for exact tests of linearity and parallelism on adjusted responses; unless the adjustment altered the responses very considerably, or very much reduced the residual variance s^2, the conclusions from that table could not be altered. As the estimation of potency is much simplified when these tests can be ignored, the calculations will first be given on that assumption. They are begun by the preparation of a table of analysis of variance and covariance for x, u and y, only the 'between-preparations' component being removed from the total. The figures required for the analysis of y^2 are already available in Table 3, and the other columns are completed in a similar manner, using either the squares of the variates or the products of corresponding values of two variates. The calculations present no special features, and their result is shown in Table 6.

Table 6. *Analysis of variance and covariance for x, u and y*

	D.F.	(x^2)	(xu)	(u^2)	(xy)	(uy)	(y^2)
Between preparations	1	0	0	18	0	−128	888
Within preparations	22	16	−13	234	350	531	18,342
Total	23	16	−13	252	350	403	19,230

The regression coefficients of y on x and u are determined from the 'within-preparations' components in Table 6, and are the solutions of the equations

$$\left.\begin{aligned} 16b - 13b_u &= 350, \\ -13b + 234b_u &= 531. \end{aligned}\right\} \qquad (9)$$

Since the variances of the regression coefficients will be required, these equations should be solved by Fisher's 'c-multiplier' or 'inverse matrix' method (1946, § 29). Full explanation of this can be found in many standard works on statistical methods; it requires the construction of an inverse matrix V, whose first row is formed by the solution of equations (9) with the numbers 1, 0 replacing 350, 531 on the right-hand side and whose second row is formed by the solutions with 0, 1 on the right-hand side. Simple arithmetic gives this matrix as

$$V = \begin{pmatrix} 0{\cdot}065455 & 0{\cdot}003636 \\ 0{\cdot}003636 & 0{\cdot}004476 \end{pmatrix}, \qquad (10)$$

and the regression coefficients are then obtained by multiplying each row of the matrix in turn by the right-hand sides of equations (9):

$$\begin{aligned} b &= 350 \times 0{\cdot}065455 + 531 \times 0{\cdot}003636 \\ &= 24{\cdot}840, \end{aligned}$$

$$\begin{aligned} b_u &= 350 \times 0{\cdot}003636 + 531 \times 0{\cdot}004476 \\ &= 3{\cdot}649. \end{aligned}$$

Now the regression on x and u jointly removes an amount

$$350b + 531b_u = 10{,}632$$

from the 'within-preparations' sum of squares, as compared with $350^2/16$ or 7656 removed by a regression on x alone. The difference between these amounts is a sum of squares with 1 degree of freedom, representing the additional portion of the within-preparations sum of squares attributable to variations in u, after account has been taken of the component due to x. Hence, as is shown by Table 7, the inclusion of u in the regression equation effects a significant reduction in the residual sum of squares, and improvement in the precision of the assay may result from the elimination of variations due to u.

The regression equation for either preparation is

$$Y = \bar{y} + b_u\,(u - \bar{u}) + b\,(x - \bar{x}),$$

where $\bar{y}$, $\bar{x}$ and $\bar{u}$ are mean values for the twelve doves used for the preparation. The dose metameter has been so chosen that $\bar{x}$ is zero for each preparation, and the two regression equations are therefore

Table 7. *Analysis of variance for test of significance of body-weight regression*

	D.F.	Sum of squares	Mean square
Regression on x	1	7,656	
Addition for inclusion of u	1	2,976	2,976
Regression on x and u	2	10,632	
Between preparations	1	888	
Residual	20	7,710	385·5
Total	23	19,230	

$$\begin{aligned} Y_s &= 72{\cdot}1667 + b_u\,(u - 47{\cdot}3333) + bx \\ &= -104{\cdot}20 + 3{\cdot}649u + 24{\cdot}840x, \end{aligned}$$

and

$$\begin{aligned} Y_t &= 60{\cdot}0000 + b_u\,(u - 50{\cdot}0833) + bx \\ &= -122{\cdot}75 + 3{\cdot}649u + 24{\cdot}840x. \end{aligned}$$

The relative potency is derived from M, the difference between x-values which, for equal values of u, give equal values of Y. The general expression for M is

$$M = \bar{x}_s - \bar{x}_t - \frac{(\bar{y}_s - \bar{y}_t) - b_u\,(\bar{u}_s - \bar{u}_t)}{b}; \qquad (11)$$

https://doi.org/10.1017/S0022172400014078 Published online by Cambridge University Press

this here gives

$$M = -\frac{12{\cdot}1667 + 1{\cdot}7500b_u}{b}$$
$$= -\frac{18{\cdot}552}{24{\cdot}840}$$
$$= -0{\cdot}7469,$$

which corresponds to a potency estimate of 5·96 i.u./mg., a figure obtained as in equation (7).

In order to express the precision of this estimate, formulae rather more complicated than for the unadjusted responses are required, since they must take into account the covariance between b and b_u indicated by equation (10). From Table 7 the residual variance per response, after the fitting of the regression equation, is

$$s^2 = 385{\cdot}5.$$

Multiplication of the diagonal elements of the matrix V, equation (10), by s^2 gives the variances of b and b_u, and a similar multiplication of the remaining element gives the covariance. Write now

$$v_{11} = \text{variance of } [(\bar{y}_s - \bar{y}_t) - b_u(\bar{u}_s - \bar{u}_t)]$$
$$= \frac{s^2}{n_s} + \frac{s^2}{n_t} + (\bar{u}_s - \bar{u}_t)^2 V(b_u), \quad (12)$$

where n_s and n_t are the total numbers of individuals used for the two preparations, in this example 12 for each;

$$v_{12} = \text{covariance of } b \text{ and } [(\bar{y}_s - \bar{y}_t) - b_u(\bar{u}_s - \bar{u}_t)]$$
$$= -(\bar{u}_s - \bar{u}_t) \text{ covariance } (b, b_u); \quad (13)$$

and

$$v_{22} = V(b). \quad (14)$$

The variance of M may then be written

$$V(M) = \frac{1}{b^2}[v_{11} + 2Mv_{12} + M^2 v_{22}]; \quad (15)$$

for the more general expression of equation (11), with $\bar{x}_s$ and $\bar{x}_t$ unequal, M in equation (15) must be replaced by $(M - \bar{x}_s + \bar{x}_t)$. For the data under discussion,

$$v_{11} = 385{\cdot}5 \times [\tfrac{1}{12} + \tfrac{1}{12} + (\tfrac{21}{12})^2 \times 0{\cdot}004476]$$
$$= 69{\cdot}53,$$
$$v_{12} = -385{\cdot}5 \times (-\tfrac{21}{12}) \times 0{\cdot}003636$$
$$= 2{\cdot}45,$$
$$v_{22} = 385{\cdot}5 \times 0{\cdot}065455$$
$$= 25{\cdot}23;$$

substitution in equation (15) gives

$$V(M) = [69{\cdot}53 - 2 \times 0{\cdot}7469 \times 2{\cdot}45 + (0{\cdot}7469)^2 \times 25{\cdot}23]/617{\cdot}03$$
$$= \frac{79{\cdot}94}{617{\cdot}03}$$
$$= 0{\cdot}1296,$$

and therefore

$$M = -0{\cdot}7469 \pm 0{\cdot}3600. \quad (16)$$

As in § 2, however, the standard error cannot be used in assigning fiducial limits to M unless the quantity defined as g in equation (5) is small; g, of course, must now be recalculated according to the numerical results of this section, taking $V(b)$ from equation (14) and using $t = 2{\cdot}086$ (5 % level and 20 degrees of freedom). The exact expression for the limits, of which (6) is a particular case, may be obtained by means of a theorem stated by Fieller (1944; Fisher, 1946, § 26·2). A convenient form for calculation is practically the same as that for another purpose given elsewhere by the writer (Finney, 1948, equation (32)):

$$M + \frac{g}{1-g}\left[M + \frac{v_{12}}{v_{22}}\right] \pm \frac{t}{b(1-g)} \times \sqrt{\left[v_{11} + 2Mv_{12} + M^2 v_{22} - g\left(v_{11} - \frac{v_{12}^2}{v_{22}}\right)\right]}, \quad (17)$$

in which, at every appearance except the first, M must be replaced by $(M - \bar{x}_s + \bar{x}_t)$ when the two values of $\bar{x}$ are unequal. For the example

$$g = 0{\cdot}178,$$

and the 5 % fiducial limits are therefore

$$-0{\cdot}7469 + \frac{0{\cdot}178}{0{\cdot}822}\left(-0{\cdot}7469 + \frac{2{\cdot}45}{25{\cdot}23}\right) \pm \frac{2{\cdot}086}{0{\cdot}822 \times 24{\cdot}840} \times \sqrt{\left[79{\cdot}94 - 0{\cdot}178 \times \left(69{\cdot}53 - \frac{2{\cdot}45^2}{25{\cdot}23}\right)\right]},$$

or $-0{\cdot}0476$ and $-1{\cdot}7276$, corresponding to 9·68 and 3·02 i.u./mg. These limits are a little narrower than for the adjustment by proportionality discussed in § 3; the three estimates, and their fiducial limits, are summarized in Table 8.

Table 8 shows a clear, though not very great, advantage for the covariance adjustment by comparison with the unadjusted results; adjustment by proportionality has been almost as good. Examination of other data has indicated that, even though there is a significant regression on the concomitant variate, the benefit of making allowance for it may be nullified by low precision in the estimation of the adjustment, and by the fact that the partial regression coefficient of response on dose chances to be lower than the regression coefficient when u is ignored. In these circumstances, of course, no arbitrary adjustment, by proportionality or otherwise, is likely to prove any better. Providing that the regression on u is significant, adjustment can scarcely lead to any appreciable *loss* of precision; that, however, is insufficient to justify the more troublesome calculations! The present finding, taken in conjunction with that of Bliss (1940) for

https://doi.org/10.1017/S0022172400014078 Published online by Cambridge University Press

other data, suggests that the gain in precision is often less than might be expected merely from consideration of the reduction in s^2 caused by allowance for the concomitant variate. This conclusion, of course, is applicable whether the computations are carried out in the form recommended here or in that used by Bliss and others. Further experience, from analyses of other assays, is needed before any general recommendation can be made about the circumstances in which covariance adjustment is likely to be profitable.

In all the analyses made on these data, the variance per response has been assumed to be independent of the dose or the magnitude of the response. Inspection of Table 1 suggests that, in fact, the responses are more variable at high doses than at low, and allowance for this should perhaps be made by using weighted rather than unweighted regressions, assigning greater weight to the less variable responses. The data are used in this paper only in illustration of covariance technique, and for clarity it seemed preferable to avoid the additional complication of weighted regressions.

5. TESTS OF ASSAY VALIDITY

In § 2 have been described tests of significance for deviations from linearity and parallelism of the regression lines fitted to the unadjusted data. When the logarithms of the doses tested are equally spaced, these tests are made particularly simple by the use of polynomial coefficients. A concomitant variate, such as body weight in the example under discussion, destroys some of the simplicity, and full analysis involves a rather laborious procedure. However, it is required only when inspection of the data, or a preliminary graphical analysis, leaves some doubt of the assay validity, and it need not be adopted as a routine.

In the prolactin assay, the analysis summarized in Table 3 has shown that, for the unadjusted crop-gland weights, the deviations do not approach significance. Those who are experienced in the analysis of assays probably need no further analysis to satisfy them that the adjustment for u is insufficient to upset this result, but if exact tests are required the multiple regression on x_1, x_2, x_3, x_4 and u must be investigated; u is not orthogonal with the other four, and calculation from Tables 1 and 5 gives

$$Sx_1u = -13, \quad Sx_2u = 5, \quad Sx_3u = -7, \quad Sx_4u = 27.$$

Hence the equations for estimating the regression coefficients are

$$\left.\begin{aligned} 16b_1 \qquad\qquad\qquad\qquad -13b_u &= 350, \\ 16b_2 \qquad\qquad\qquad +5b_u &= -12, \\ 48b_3 \qquad\qquad -7b_u &= -16, \\ 48b_4 \quad +27b_u &= 118, \\ -13b_1 + 5b_2 - 7b_3 + 27b_4 + 234b_u &= 531, \end{aligned}\right\} \quad (18)$$

Table 8. *Estimates of potency and fiducial limits*
(in i.u./mg.)

	Adjustment for body weight		
	None	Proportionality	Covariance
Estimated potency of test preparation	6·80	6·40	5·96
5 % limits	2·68–12·6	3·05–10·7	3·02–9·68
5 % limits as % of mean	39–185	48–167	51–162
1 % limits	1·14–16·9	1·81–13·1	1·95–11·7
1 % limits as % of mean	17–249	28–205	33–196

the solutions of which are*

$$\left.\begin{aligned} b_1 &= 24{\cdot}840, \quad b_4 = 0{\cdot}406, \\ b_2 &= -1{\cdot}890, \quad b_u = 3{\cdot}649, \\ b_3 &= 0{\cdot}199. \end{aligned}\right\} \quad (19)$$

The sum of the products of the regression coefficients with the right-hand sides of equations (18), 10,699, is the sum of squares attributable to the five-variate regression. Table 7 may now be extended to the form of Table 9, in order to test the significance of b_2, b_3 and b_4. The sum of squares for the 3 degrees of freedom is obviously too small to contain any significant component. If necessary, any one of the regression coefficients could be examined separately by omission of those parts of equations (18) relating to the other two, and repetition of the calculations, leading to a sum of squares with 1 degree of freedom for this coefficient alone. Had the variances of the five partial regression coefficients been required, equations (18) would have been solved by the inverse matrix method, but usually b_2, b_3 and b_4 are required only for tests of significance and no other interest attaches to their values; if any one of them were significant, it would indicate either that the assay

* That b_1 and $\bar{b}_u$ are identical with the b and b_u of § 4 must be attributed to coincidence.

https://doi.org/10.1017/S0022172400014078 Published online by Cambridge University Press

was invalid or that the metameters used needed modification in order to give linearity (Finney, 1948). If it were customary to make the full analysis for every assay, the residual mean square in Table 9 (as in Table 3) would naturally be used as s^2, but, since this procedure should seldom be needed, values of s^2 obtained without the inclusion of x_2, x_3 and x_4 in the regression equation have been used for comparative purposes in this paper.

Table 9. *Analysis of variance for validity tests with adjusted crop-gland weights*

	D.F.	Sum of squares	Mean square
Regression on x_1	1	7,656	
Addition for inclusion of u	1	2,976	2,976
Regression on x_1 and u	2	10,632	
Linearity and parallelism	3	67	22·3
Regression on x_1, x_2, x_3, x_4, u	5	10,699	
Between preparations	1	888	
Residual	17	7,643	449·6
Total	23	19,230	

In an assay based on a greater number of doses of each preparation, these still being equally spaced on the logarithmic scale, the same procedure of expressing components of non-parallelism and non-linearity in terms of orthogonal coefficients can be followed (Bliss & Marks, 1939*a*, *b*). If the log doses are not equally spaced, the tests become more laborious; standard methods of covariance analysis still serve to determine whether the linear partial regression coefficients on log dose are significantly different for the two preparations, and a test of linearity could be put in the form of a test of significance of the regression on the square of the log dose. Fortunately, so complete an analysis will seldom be needed; experience and discretion will enable a decision to be reached without exact tests in many assays, and will allow attention to be focused on the features most open to suspicion in others.

6. SUMMARY

When individual responses in a biological assay show considerable variation associated with the values of a concomitant variate, covariance analysis may be used in order to adjust the mean responses and to improve the precision of the assay. Usually this is preferable to the choice of an adjustment which involves an arbitrary assumption about the effect of variations in the concomitant variate on the measured response. Published accounts of the process are open to certain theoretical objections, though they may be sufficiently exact for most practical purposes.

The present paper describes a method of calculating the relative potency, and its precision, which may be a little more laborious, but which is in full accord with standard statistical practice. The computations are illustrated on data from a prolactin assay by the pigeon crop-gland technique, in which the final crop-gland weight showed a positive correlation with the body weight at the start of the assay. The results are compared with those obtained either from the unadjusted crop-gland weights or from these weights expressed as proportions of body weights. The covariance method leads to a more precise estimate of the potency of the test preparation than do either of the others; there is evidence, however, that the increase in precision will not necessarily be large unless the correlation between the response and the concomitant variate is very close.

In a final section, the full statistical tests of assay validity in the covariance analysis are described; these are lengthy, and fortunately are required only when the validity is in considerable doubt.

The methods of adjustment have been described in this paper with respect to an assay depending upon parallel regression lines of responses on the logarithms of doses. They may be adapted for use with 'slope-ratio' assays (Bliss, 1946; Finney, 1945, 1948; Wood & Finney, 1946), in which the regression of response on dose itself is linear. So far the need for adjusting for concomitant variation in these assays seems not to have arisen, and discussion of computational details may be postponed until the need is felt.

I am indebted to Mr A. L. Bacharach and to Glaxo Laboratories Ltd. for the results of the prolactin assay used as the numerical illustration of the methods of this paper. I wish also to acknowledge valuable suggestions from Mr Bacharach and from Dr E. C. Wood on the argument of the paper.

REFERENCES

Bliss, C. I. (1940). Factorial design and covariance in the biological assay of vitamin D. *J. Amer. Statist. Ass.* **35**, 498–506.

Bliss, C. I. (1946). An experimental design for slope-ratio assays. *Ann. Math. Statist.* **17**, 232–7.

Bliss, C. I. & Marks, H. P. (1939*a*). The biological assay of insulin. I. Some general considerations directed to increasing the precision of the curve relating dosage and graded response. *Quart. J. Pharm.* **12**, 82–110.

Bliss, C. I. & Marks, H. P. (1939*b*). The biological assay of insulin. II. The estimation of drug potency from a graded response. *Quart. J. Pharm.* **12**, 182–205.

Bliss, C. I. & Rose, C. L. (1940). The assay of parathyroid extract from the serum calcium of dogs. *Amer. J. Hyg.* **31** A, 79–98.

Fieller, E. C. (1940). The biological standardization of insulin. *J. Roy. Statist. Soc. Suppl.* **7**, 1–64.

Fieller, E. C. (1944). A fundamental formula in the statistics of biological assay, and some applications. *Quart. J. Pharm.* **17**, 117–23.

Finney, D. J. (1945). The microbiological assay of vitamins: The estimate and its precision. *Quart. J. Pharm.* **18**, 77–82.

Finney, D. J. (1948). The principles of biological assay. *J. Roy. Statist. Soc. Suppl.* **9**, 46–91.

Fisher, R. A. (1946). *Statistical Methods for Research Workers*, 10th ed. Edinburgh: Oliver and Boyd.

Fisher, R. A. & Yates, F. (1947). *Statistical Tables for Biological, Agricultural and Medical Research*, 3rd ed. Edinburgh: Oliver and Boyd.

Irwin, J. O. (1943). On the calculation of the error of biological assays. *J. Hyg., Camb.*, **43**, 121–8.

Snedecor, G. W. (1946). *Statistical Methods*. Ames, Iowa: The Collegiate Press, Inc.

Wood, E. C. & Finney, D. J. (1946). The design and statistical analysis of micro-biological assays. *Quart. J. Pharm.* **19**, 112–27.

(*MS. received for publication* 2. II. 47.—Ed.)