J. Physiol. (1942) 101, 115–130 547.781.5:612.01

# A METHOD OF CONDUCTING A BIOLOGICAL ASSAY ON A PREPARATION GIVING REPEATED GRADED RESPONSES ILLUSTRATED BY THE ESTIMATION OF HISTAMINE

By H. O. SCHILD

*From the Department of Pharmacology, University College, London*

(*Received* 21 *January* 1942)

Biological assays on isolated preparations are generally based on the principle of 'bracketing' doses of the unknown with doses of the standard until their effects are matched. This method does not lend itself readily to statistical analysis. Many workers do indeed attempt to estimate the accuracy of their assays by performing preliminary experiments with solutions of known composition. Such preliminary experiments are, however, inefficient and also apt to be misleading, as conditions often do not approximate to those of a real assay. It is thus preferable to deduct the error of an assay from the data of the experiment itself.

Special difficulties arise if the object of the experiment is to test whether two solutions have the same activity. As the two solutions are tested repeatedly chance variations occur, and owing to the lack of criteria for dealing with these the result frequently becomes more dubious the more the experiment is prolonged. It is in this type of experiment involving the setting up and testing of a 'null hypothesis' [Fisher, 1937], that statistical methods are most useful, since they provide a definite answer, provided that the question is put in the right terms and the experiment designed on sound lines.

The object of this paper is to describe a method of conducting a biological assay on a single preparation in such a way that a valid null hypothesis may be set up and the accuracy of the result may be estimated from the data of the experiment itself. The method has been applied to the assay of histamine on the guinea-pig's gut. The design is based on a simple plan used in field experiments on adjacent plots [Fisher, 1938]. The statistical argument has been largely adapted from the work of

Gaddum [1933] and Bliss & Marks [1939 *a*, *b*]. The performance of the assay and its statistical analysis are discussed in detail, and it is hoped that readers who are not acquainted with statistical methods will have no difficulty in following the main argument and performing the test.

## Methods

The experiments were done on preparations of isolated gut from guinea-pigs. Most assays were performed at 28–32° C. At this temperature no spontaneous contractions of the intestinal strip occurred.

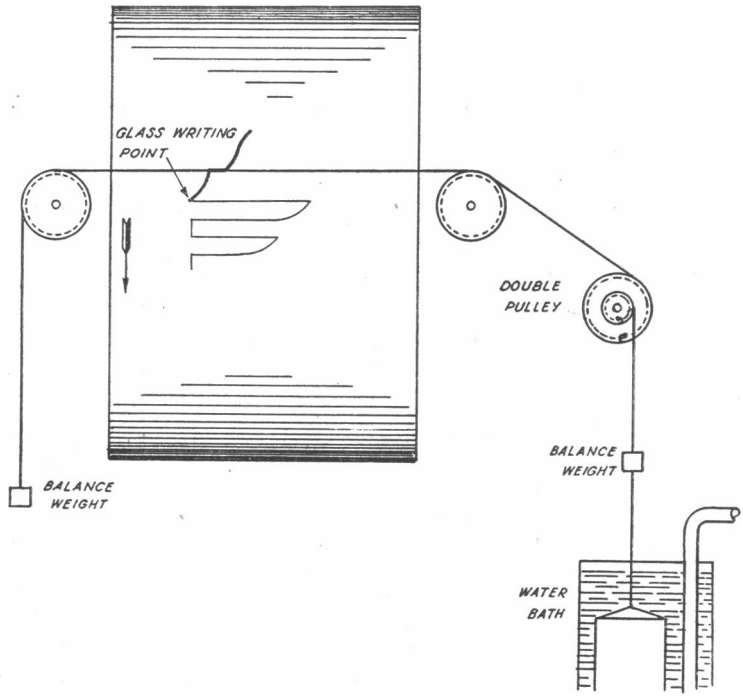


**Fig. 1. Pulley system for linear recording of smooth muscle contraction.**

A frontal writing lever was used in earlier experiments, but was discarded later owing to its relatively large error in recording at high angles of excursion. Instead, a pulley system of recording, shown in Fig. 1, was adopted; this provided a faithful record of the intestinal movements. A fine glass frontal writing point is attached to a horizontally moving silk thread which is kept taut by means of two small balance weights. Light vulcanite pulleys are used, and if magnification is desired it can be obtained by means of a double pulley as shown in Fig. 1. In the present

experiments, however, a single pulley without magnification was used in its place.

The bath volume was 25 c.c. and that of the test solutions added to the bath usually 1 c.c. Solutions were added at intervals of 3 min.

### Performance of the assay

The assay is based on the assumption that, over the range of concentrations used, the contraction of the gut increases linearly with the logarithm of the dose.[1] It is carried out with the aid of only four doses, two of the standard and two of the unknown. They are chosen in a preliminary test and should fall within the limits of 10 and 90% of the maximum effect. The ratio of activity of the two doses of the unknown must be the same as that of the two doses of the standard, and the logarithm ($d$) of this ratio should preferably be at least twice the logarithm ($M$) of the ratio of activity of unknown and standard.

A suitable constant volume of test solution is added to the bath at regular intervals, the number and order of determinations being established at the outset according to the following scheme. The total number of determinations depends on the accuracy required, but it has to be a multiple of four, and every group of four consecutive determinations must contain each dose once. It is essential that within 'groups' doses should follow each other in random succession, which may be determined by means of random numbers or some physical process of randomization.

The logarithm of the ratio of potencies is given by

$$M=\frac{\bar{y}_u-\bar{y}_s}{b},$$

where $\bar{y}_u-\bar{y}_s$ is the difference between the mean responses to unknown and standard, and $b$ is the slope of the regression line plotted against log dose of standard.

If $S(y_s)_1$ denotes the sum of all the effects (heights of recorded excursion) due to the larger dose of the standard and $S(y_s)_2$, $S(y_u)_1$, $S(y_u)_2$ represent corresponding sums of effects of the smaller dose of the standard and larger and smaller dose of the unknown, and if $N$ is the number of groups,

$$\bar{y}_u-\bar{y}_s=\frac{S(y_u)_1+S(y_u)_2-S(y_s)_1-S(y_s)_2}{2N} \quad \text{and} \quad b=\frac{S(y_u)_1+S(y_s)_1-S(y_u)_2-S(y_s)_2}{2Nd}.$$

Defining $S(y_u)_1+S(y_u)_2-S(y_s)_1-S(y_s)_2=A$

and $S(y_u)_1+S(y_s)_1-S(y_u)_2-S(y_s)_2=B$

[1] Any other function of the dose giving an approximately linear relationship between dose and effect over a given range can be substituted for the logarithm without essential modification in the method of assay.

the expression for $M$ becomes

$$M = \frac{A}{B}\, d.$$

*Graphical presentation.* Fig. 2 illustrates an experiment in which two known solutions of histamine—termed for convenience 'standard' and

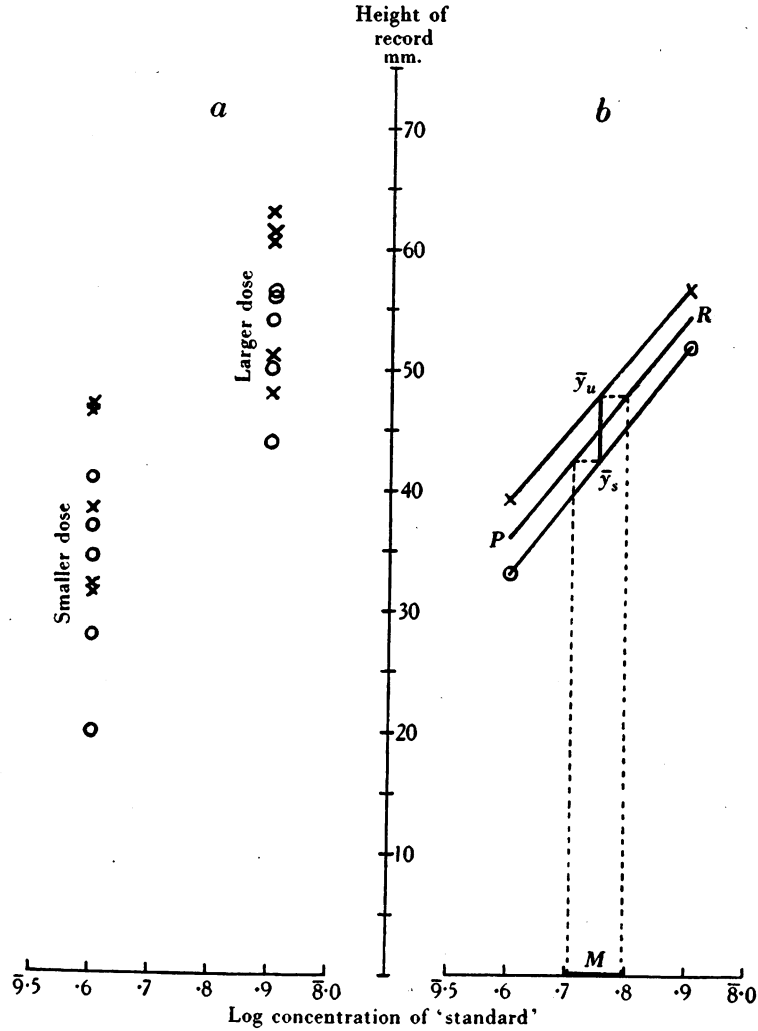


Fig. 2*a*. Effects of two doses of 'standard' (○) and two doses of 'unknown' (×). Ratio of activities 4: 5; $d = 0{\cdot}30103$; $N = 5$.

Fig. 2*b*. Graphical determination of $M$. $PR$ is the average regression.

'unknown'—were compared. The ratio of activity of the two solutions was 4 : 5, and the concentration ratio of doses was 1 : 2 ($d = 0{\cdot}30103$). The effects are plotted against log concentration of standard.

Fig. 2*a* shows the scatter of results and the overlapping of effects due to standard and unknown. In Fig. 2*b* the mean effects of the four doses have been computed, and with their aid and the use of the mean regression line *M* is determined. The slope of *PR*, the mean regression line, is an average of the estimated slopes for standard and unknown.

Graphically *M* works out at about 0·09, this being the logarithm of 1·23, the estimated difference of activity is approximately 23 %. The data of this experiment will be worked out in detail in the following sections.

## Outline of the statistical analysis

Fig. 3*a* illustrates the sequence of injections and the size of single responses in the above experiment, and Fig. 3*b* once more the mean effects of the four doses. The first object of the statistical analysis is to find out whether these mean effects differ from each other significantly, compared with the experimental error. In the simplest type of experiment the experimental error would be constituted by the variations in response to repeated tests with the same dose of histamine. In the present experiment, involving grouping, determination of the experimental error is somewhat more complex.

Fig. 3*a* shows that the mean 'group' response varies considerably in the course of the experiment, indicating marked changes in the sensitivity of the preparation. The effect of these variations in sensitivity has been largely eliminated from the experimental comparisons by the method of grouping which ensures that each dose is given at various levels of sensitivity and thus provides a well-balanced mean estimate for each dose. It is essential, however, that the differences between groups should be eliminated not only from the experimental comparison but also from the estimate of error, by the methods of the analysis of variance described in the following section. As a result, the estimate of the experimental error is reduced to the same value as if the mean sensitivity of groups had not changed in the course of the experiment. This is illustrated by Figs. 3*c* and 3*d*. The former shows the varying effects produced in the course of the assay by the same dose of histamine, the latter the same effects after eliminating the variations between groups. The adjusted effects are much more homogeneous and show a marked reduction of the variation ascribable to experimental error.

Having thus reduced the experimental error, the next step consists in extracting the maximum amount of useful information from con-

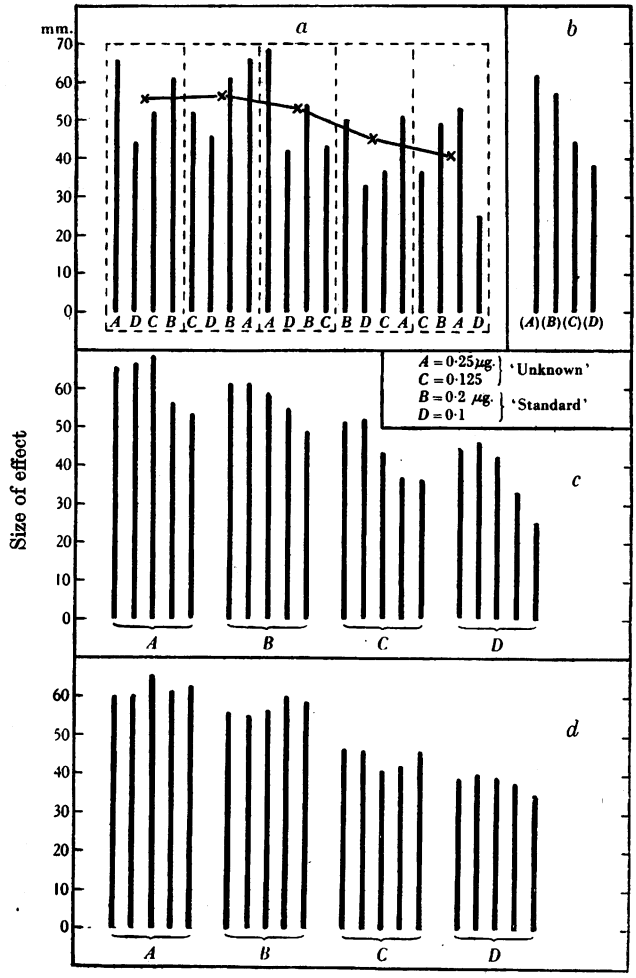


Fig. 3. Data from same experiment as in Fig. 2.

Fig. 3*a*. Order of tests and height of responses. Crosses indicate the mean response in successive groups.

Fig. 3*b*. Mean response to two doses of 'standard' and two doses of 'unknown'.

Fig. 3*c*. Successive responses to the same dose of histamine.

Fig. 3*d*. Values from Fig. 3*b* corrected to represent the effects that would have been obtained if the mean sensitivity of groups had not changed in the course of the assay.

trasting the mean effects of the four doses as presented in Fig. 2*b*. Three 'independent comparisons' can be made, and the significance

of each contrast may be assessed by relating it to the experimental error.

The effects of the two doses of the standard may be contrasted as a group with those of the two doses of the unknown. If variation between the two groups is significantly greater than the experimental error, it may be concluded without further assumptions that the two solutions differ in activity.

Secondly, the effects of the larger doses of both standard and unknown may be contrasted with those of the smaller doses. This is, in fact, a test for regression, since unless the larger dose produces a significantly greater effect than the smaller dose, no estimate of the regression coefficient and consequently no quantitative estimate of activity can be made.

Lastly, the sum of the effects of the larger dose of the standard and the smaller dose of the unknown may be contrasted with the sum of the other two effects. This test is a measure of parallelism, since if the two sums are equal the regression lines must be parallel. It cannot, of course, be expected, owing to chance variations, that the mean regression lines of standard and unknown should be perfectly parallel. If, however, the deviations from parallelism are *significantly* greater than the experimental error great caution must be used in the interpretation of results.

### *The analysis of variance*

The analysis of variance is 'a simple arithmetical procedure by means of which the results may be arranged and presented in a single compact table, which shows both the structure of the experiment and the relevant results in such a way as to facilitate the necessary tests of their significance' [Fisher, 1937]. It is essential, with the present method of assay, to compute an analysis of variance for each experiment, since it provides the error component for determining the limits of accuracy of the assay and leads to the various tests of significance outlined in the preceding section.

A typical analysis of variance computed from the data of the histamine assay previously quoted is shown in Table 2. The variate (Table 1) is the recorded maximum height of contraction produced by the addition of 1 c.c. of histamine solution to the bath. Table 2 shows that in the analysis five distinct sources of variation have been isolated. For each source of variation an expression called the *sum of squares* (of deviations from the mean) is computed, which divided by the appropriate *degrees of freedom* (*df*) yields a *mean square*. The ratio of two mean squares in conjunction with the degrees of freedom from which they are derived affords a test of significance.

TABLE 1. Effects of four doses of histamine applied in five successive randomized groups

Height of response in ½ mm.

| Dose | Groups 1 | 2 | 3 | 4 | 5 | Sum |
|---|---|---|---|---|---|---|
| 0·25 μg. ($U_1$) | 131 | 132 | 136 | 112 | 106 | 617 |
| 0·2 μg. ($S_1$) | 122 | 122 | 118 | 110 | 98 | 570 |
| 0·125 μg. ($U_2$) | 103 | 104 | 87 | 74 | 73 | 441 |
| 0·1 μg. ($S_2$) | 89 | 92 | 84 | 66 | 50 | 381 |
| Sum | 445 | 450 | 425 | 362 | 327 | 2009 |

TABLE 2. Analysis of variance of histamine assay

| Source of variation | Sum of squares | Degrees of freedom | Mean square |
|---|---|---|---|
| Between groups | 2976·7 | 4 | 744·18* |
| Between 'standard' and 'unknown' | 572·45 | 1 | 572·45* |
| Regression | 6661·25 | 1 | 6661·25* |
| Deviation from parallelism | 8·45 | 1 | 8·45 |
| Error | 330·1 | 12 | 27·51 |
| Total | 10548·95 | 19 | |

* Highly significant.

The computations leading to the test of significance will be briefly described. The computational scheme is adapted from Snedecor (1938).

Given that $S_1, S_2, \ldots, S_N$ = each sum of items in a group of four responses (sum of each column of Table 1), and $S = S_1 + S_2 + \ldots + S_N$, the following items are computed and then summarized as shown in Table 2:

(1) The correction term $= C = S^2/4N = 2009^2/4\ (5) = 201804{\cdot}05$.

(2) The sum of the squares of all items

$$= S(X^2) = 131^2 + \ldots + 50^2 = 212353.$$

(3) The 'total' *sum of squares*

$$= S(X^2) - C = 212353 - 201804{\cdot}05 = 10548{\cdot}95.$$

(4) The *sum of squares* for groups

$$= (S^2_1 + S^2_2 + \ldots + S^2_N)/4 - C = (445^2 + \ldots + 327^2)/4 - 201804{\cdot}05 = 2976{\cdot}7;$$

and the corresponding mean square by division by $(N-1)$, the corresponding degrees of freedom: $2976{\cdot}7/(5-1) = 744{\cdot}175$.

(5) The *sum of squares* for

(*a*) variation between standard and unknown

$$= A^2/4N = (617 + 441 - 570 - 381)^2/4\ (5) = 572{\cdot}45;$$

(*b*) regression

$$= B^2/4N = (617 + 570 - 441 - 381)^2/4\ (5) = 6661{\cdot}25;$$

(*c*) deviation from parallelism

$$= [S(y_u)_1 + S(y_s)_2 - S(y_u)_2 - S(y_s)_1]^2/4N$$
$$= (617 + 381 - 570 - 441)^2/4N = 8{\cdot}45.$$

For each of these three sources of variation, only a single degree of freedom is available and their mean square is thus numerically equal to their sum of squares.

(6) The *sum of squares* for error

= total − (groups + standard *v*. unknown + regression + parallelism)
= 10548·95 − (2976·7 + 572·45 + 6661·25 + 8·45) = 330·1,

and the corresponding mean square by division by $(3N-3)$ the corresponding degrees of freedom: $330{\cdot}1/(15-3) = 27{\cdot}51$.

A useful partial check of computations is afforded by the expression $(S^2(y_u)_1 + S^2(y_u)_2 + S^2(y_s)_1 + S^2(y_s)_2)/N - C$, which must be equal to the sum total of the three sums of squares with a single degree of freedom. Thus $(617^2 + 570^2 + 441^2 + 381^2)/5 - 201804{\cdot}05$ must be equal to

$$572{\cdot}45 + 6661{\cdot}25 + 8{\cdot}45.$$

In fact, both expressions add up to 7242·15.[1]

The test of significance is made by relating each mean square thus computed to the error mean square. The ratio $F$ = larger mean square/smaller mean square is formed, and the simple value of $F$ thus obtained is compared with a tabulated value of $F$. If the sample value exceeds the tabulated value for the 5 % level of probability of $F$ it is likely to occur less than once out of twenty times by chance and is said to be significant; similarly, if it exceeds the 1 % level it is said to be highly significant. The numerical value of $F$ depends not only on the required level of probability but also on the degrees of freedom from which the two mean

[1] If, in the course of the assay, a wrong dose is given by mistake, or some other accident occurs, the missing item ($X$) can be supplied with the aid of a formula proposed by Allen and Wishart and Yates [quoted from Snedecor, 1938]. Adapted to the present purpose the formula is

$$X = \frac{4D + NG - S}{3N - 3},$$

where $D$ = the sum of effects produced by the same dose as the missing effect, $G$ = the sum of effects produced in the same group as missing effect, and $N$ and $S$ retain their previous significance. The value of $X$ is entered in the table as the missing response, and the analysis of variance proceeds as usual with this one modification that the degrees of freedom for error are reduced by unity.

Thus assuming that in Table 1 item 87 from column 3, row 3 were missing,

$$X = \frac{4\,(441 - 87) + 5\,(425 - 87) - 2009}{3\,(5) - 3} = 91,$$

and the degrees of freedom for error are reduced to 11.

squares forming $F$ are derived. Thus if it is desired to find the value of $F$ derived from $n_1$ and $n_2$ degrees of freedom for a given level of probability, the appropriate table of $F$ by Snedecor [1938] (or the corresponding table of $e^{2z}$ by Fisher & Yates [1938]) is entered at the column headed $df = n_1$, and the required value is found at the row headed $df = n_2$.

The analysis of variance in Table 2 yields the following $F$-values:

$F$ for variation between standard and unknown $= \frac{572 \cdot 45}{27 \cdot 51} = 20 \cdot 81$ (the 1 % point of $F$ for $n_1 = 1df$ and $n_2 = 12df$ is 9·33).

$F$ for regression $= \frac{6661 \cdot 25}{27 \cdot 51} = 242 \cdot 13$ (1 % point of $F = 9 \cdot 33$).

$F$ for deviations from parallelism $= \frac{27 \cdot 51}{8 \cdot 45} = 3 \cdot 26$ (the 5 % point of $F$ for $n_1 = 12df$ and $n_2 = 1df$ is 243·9).

$F$ for groups $= \frac{744 \cdot 18}{27 \cdot 51} = 27 \cdot 05$ (the 1 % point of $F$ for $n_1 = 4df$ and $n_2 = 12df$ is 5·41).

It may be concluded that there is a highly significant difference in activity between standard and unknown, and a highly significant regression between the smaller and the larger dose, making a quantitative estimate of activity possible. Deviations from parallelism are very slight, in fact the corresponding mean square is smaller than the error mean square, though not significantly smaller. Lastly, the high $F$ value for groups is a justification of the experimental design, showing as it does highly significant variations between groups.

*The limits of error of the estimate of M*

$s_M$, the standard error of $M = (\bar{y}_u - \bar{y}_s)/b$, may be computed from the expression

$$s_M = 2\sigma d \sqrt{N} \frac{\sqrt{(A^2 + B^2)}}{B^2},$$

where $\sigma$ is the square root of the error mean square in the analysis of variance, and the other terms retain their previous significance. The formula is derived on the assumption that $\bar{y}_u - \bar{y}_s$ and $b$ are uncorrelated.

The standard error of a quotient whose numerator and denominator are uncorrelated is

$$s_{x/y} = \frac{x}{y} \sqrt{\left( \frac{s_x^2}{x^2} + \frac{s_y^2}{y^2} \right)}.$$

If $\bar{y}_u - \bar{y}_s = A/2N$, $b = B/2Nd$, $s_{(\bar{y}_u - \bar{y}_s)} = \sigma/\sqrt{N}$ and $s_b = \sigma/\sqrt{N}\, d$, the above expression for $s_M$ is obtained. It is equivalent to that given by Bliss & Marks [1939*b*].

The $P$ 0·99 limits of error of the assay are constituted by $M \pm s_M t$.

The value of $t$ in this expression is obtained from a table of $t$ [Fisher, 1938] for the 1 % level of significance and $(3N - 3)$ degrees of freedom, the same number as for error in the analysis of variance.

In the numerical example

$$M=\frac{A}{B}d=\frac{107}{365}\times 0{\cdot}30103=0{\cdot}08825$$

and $$s_M=2\sqrt{(27{\cdot}51)}\times 0{\cdot}30103\times\sqrt{5}\times\frac{\sqrt{(107^2+365^2)}}{365^2}=0{\cdot}02016.$$

Since the value of $t$ for the 1 % level and $12df$ is 3·055,

$$M-3{\cdot}055s_M=0{\cdot}02666 \quad \text{and} \quad M+3{\cdot}055s_M=0{\cdot}14984$$

constitute the $P$ 0·99 limits of error of the assay. Taking the antilogarithm of these numbers and multiplying by 100 the estimate of activity is 122·6 %, and the limits of error are 106·3 and 141·2 %. The true activity, 125 %, is well within the computed limits of error.[1]

In interpreting the expression for $s_M$ it might profitably be transformed to

$$s_M=\frac{\sigma}{\sqrt{N}\,b}\sqrt{\left(\frac{M^2}{d^2}+1\right)}.$$

In this expression the ratio $\sigma/b$ is an absolute measure of the variability of the preparation. Provided it remains constant the standard error diminishes with the square root of $N$, the number of groups in the assay. There are thus two factors limiting the accuracy obtainable. One is the total number of responses that can be elicited, and the other is the constancy of the preparation. If towards the end of an experiment the variability of the preparation increases, any further prolonging of the assay may well increase rather than decrease $s_M$.

$s_M$ is reduced by any decrease in the value of the quotient $M/d$. In practice, provided that $M/d$ is not greater than 0·5, any further reduction of the ratio will not markedly alter the value of $s_M$.

The following relationship exists between the limits of error for $M$ as derived from $s_M$, and the variance ratio test (F test) assessing the significance of the difference between standard and unknown. When $M=0$ the result of the two tests is identical, when, however, $M>0$ the F test is more discriminating. This is due to the fact that the F test is not affected by variations in the slope of the regression line.

[1] Dr I. O. Irwin has pointed out to me that these limits of error are only approximate since the $t$ distribution is not strictly applicable to $s_M$ in view of the error in $b$. He suggests computing the exact fiducial limits from an expression equivalent in our notation to

$$\frac{ABd}{B^2-R}\pm\frac{2\sigma td}{B^2-R}\sqrt{[N(A^2+B^2-R)]},$$

where $R=4t^2\sigma^2N$, and the other terms retain their previous significance. Derivations of similar expressions will be found in Bliss [1935] and Fieller [1940]. It will be seen that the formula differs from the previous one by the introduction of the term ($R$). In our example the two methods yield almost identical results, the fiducial limits by the above formula working out at 106·9 and 142·7 %, but more important differences may arise if the slope is not well determined.

### Assays with solutions of known composition

The results of a series of assays with known concentrations of histamine are shown in Table 3. Every assay comprises various tests of significance as well as an estimate of potency and the $P$ 0·99 limits of error.

In every experiment the estimated potency was well within the computed $P$ 0·99 limits of error. These varied considerably from one assay to the other, ranging from $-8{\cdot}2$ and $+8{\cdot}8$ to $-27{\cdot}7$ and $+32{\cdot}8$ %. These differences in the error range are largely due to a change in the numerical value of the ratio $\sigma/b$, measuring variability (last col. of Table 3). In two extreme experiments differences in variability were such that fourteen tests on one preparation would have been needed to furnish the amount of information provided by a single test on the other preparation.

The $F$ values testing the difference in activity between 'standard' and 'unknown' are highly significant in all assays where solutions differed by 15 % or more. When concentrations differed by only 10 % the results were less definite. In two such experiments (Exps. 7 and 10) the estimated differences of activity were 6·3 and 8·6 %. Statistical analysis, however, showed that differences as great or greater than those found in Exp. 7 would have occurred by chance nearly five out of one hundred times, and differences as great or greater than in Exp. 10 almost twenty out of one hundred times. The differences are thus in neither assay highly significant and barely significant only in Exp. 7. Possibly a highly significant result would have been obtained by further prolonging the assay.

#### *Deviations from parallelism and linearity*

The assay in its present form includes a test of deviation from parallelism but no test of deviation from linearity of regression. If such a test were desired it would be necessary to determine more than two points on each regression line.[1] The test for deviation from parallelism gives, however, an indirect indication of deviation from linearity which is sufficiently stringent for the present purpose.

The test of departure from parallelism is related to the test for quadratic regression, indeed, the two tests are numerically equal when $M = \frac{1}{2}d$. It is thus cubic regression which is more likely to lead to error, and the test for parallelism would be inadequate if the regression line had a pronounced sigmoid shape.

In the present series of assays with solutions of known composition, tests of deviation from both parallelism and linearity may be made, the

[1] A general discussion of assays involving more than two points on the regression line is provided by Bliss & Marks [1939 *a*, *b*].

TABLE 3. Assays with histamine solutions of known composition

| Exp. no. | Method* of recording | Temp. °C. | Activity of 'unknown' in terms of 'standard' % | No. of groups (*N*) | Test of significance† (F) for: Difference in activity of two solutions | Test of significance† (F) for: Deviations from parallelism of regression lines | Estimated activity % | Limits of error‡ of estimate % | $\sigma/b$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | f.l. | 28 | 200 | 8 | H | O | 186·3 | 162·5–213·6 | 0·05458 |
| 2 | f.l. | 28 | 200 | 4 | H | H | *205·9* | *179·4–236·3* | 0·02863 |
| 3 | f.l. | 35·5 | 150 | 8 | H | O | 143·8 | 133·2–155·1 | 0·01929 |
| 4 | f.l. | 36 | 125 | 8 | H | O | 128·2 | 120·8–145·5 | 0·02044 |
| 5 | f.l. | 30 | 125 | 5 | H | O | 122·6 | 106·3–141·2 | 0·04326 |
| 6 | l.r. | 32 | 150 | 5 | H | O | 145·5 | 131·8–160·7 | 0·02762 |
| 7§ | l.r. | 32 | 110 | 10 | S | O | 106·3 | 100·2–112·9 | 0·03992 |
| 8 | l.r. | 32 | 125 | 3 | H | O | 133·0 | 106·7–165·8 | 0·04152 |
| 9 | l.r. | 32 | 115 | 6 | H | H | *114·9* | *106·3–124·1* | 0·02738 |
| 10 | l.r. | 32 | 110 | 6 | O | O | 108·6 | | 0·05268 |
| 11§ | l.r. | 32 | 120 | 3 | H | O | 117·3 | 109·1–126·1 | 0·01444 |

* f.l. = frontal righting lever; l.r. = linear method of recording.
† H = highly significant, $P > 0{\cdot}99$; S = significant, $P > 0{\cdot}95$; O = not significant, $P < 0{\cdot}95$.
‡ $P$ 0·99 limits of error are indicated in all experiments except no. 7, where $P$ 0·95 limits are given.
§ One missing value interpolated.

latter by treating the two doses of 'standard' and 'unknown' as forming a sequence of four different concentrations of the same solution. Such tests are presented in Table 4. The test of departure from linear regression was done according to standard methods [Snedecor, 1938, p. 317].

TABLE 4. Deviations from parallelism and linearity

| Exp. no. | $F$ for deviations from parallelism | $F$ for deviations from linearity | 5% level of $F$ | 1% level of $F$ |
|---|---|---|---|---|
| 1 | 0·03 | 1·55 | 4·32 | — |
| 2 | 30·85 | 31·40 | 5·99 | 13·74 |
| 3 | 1·45 | 3·81 | 4·32 | — |
| 4 | 1·06 | 1·58 | 5·12 | — |
| 5 | 0·31 | 0·77 | 4·75 | — |
| 6 | 0·05 | 0·67 | 4·75 | — |
| 7 | 0·29 | 1·69 | 4·22 | — |
| 8 | 3·81 | 4·76 | 5·99 | — |
| 9 | 9·23 | 9·26 | 4·54 | 8·68 |
| 10 | 0·36 | 0·71 | 4·54 | — |
| 11 | 0·95 | 2·30 | 6·61 | — |

Only in two experiments of the series (Exps. 2 and 9) deviations from parallelism occurred, and in both instances they were associated with significant deviations from linearity. In the rest of the assays there were no significant deviations from either. Thus in all experiments except two there was no reason to assume that the regression lines differed from linearity within the range of the experiment. Possible reasons for departure from linearity were in one experiment the use of a frontal writing lever at a high angle of excursion, and in the other the application of a dose producing an effect greater than 90% of the maximum.

Tentative estimates of potency in these experiments, treating data as if regression were linear, gave results which were surprisingly close to the true value. This suggests that the assay is relatively insensitive to deviations from parallelism.

Besides non-linearity of regression various other factors may cause deviation from parallelism, such as differential deterioration of histamine during the period of assay,[1] failure to dilute standard and unknown equally in preparing the second dose of each and possibly qualitative differences between standard and unknown.

### *The effect of grouping*

An indication of the importance of eliminating gradual changes in sensitivity taking place in the course of the assay is provided by the $F$ value for groups as computed in the analysis of variance. Table 5 shows

[1] This is probably due to bacterial action and may be prevented by boiling up solutions briefly.

TABLE 5. Significance of variations between groups

| Exp. no. | $F = \frac{\text{M.sq. for groups}}{\text{M.sq. for error}}$ | Degrees of freedom $n_1$ | Degrees of freedom $n_2$ | 5% level of $F$ | 1% level of $F$ |
|---|---|---|---|---|---|
| 1 | 16·03 | 7 | 21 | 2·49 | 3·65 |
| 2* | 4·28 | 3 | 6 | 4·76 | 9·78 |
| 3 | 2·37 | 7 | 21 | 2·49 | 3·65 |
| 4 | 10·40 | 3 | 9 | 3·86 | 6·99 |
| 5 | 27·05 | 4 | 12 | 3·26 | 5·41 |
| 6 | 2·06 | 4 | 12 | 3·26 | 5·41 |
| 7 | 2·40 | 9 | 26 | 2·27 | 3·17 |
| 8 | 12·69 | 2 | 6 | 5·14 | 10·92 |
| 9 | 2·94 | 5 | 15 | 2·90 | 4·56 |
| 10 | 1·25 | 5 | 15 | 2·90 | 4·56 |
| 11† | 34·17 | 2 | 5 | 5·79 | 13·27 |

* Latin square arrangement.
† A somewhat longer time interval occurred between the first and second group.

that in the majority of experiments this value is significant, proving that the variations in sensitivity are real. The direction of these changes is illustrated in Fig. 4, which comprises all the significant results. Perhaps the most common feature is an initial rise in sensitivity followed by a gradual decrease.

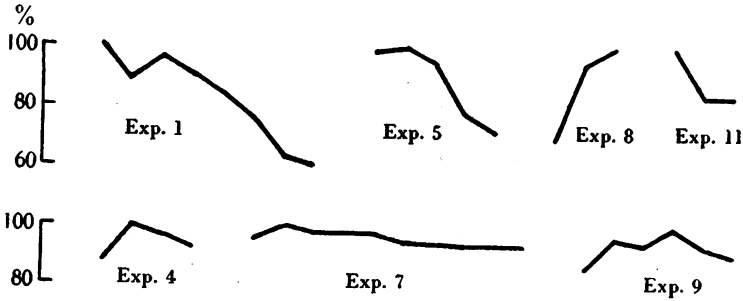


Fig. 4. Mean response in successive 'groups' in terms of maximum response.

In spite of irregularities a slight linear trend is usually discernible, and it is possible that a further reduction of the experimental error could be effected by other restrictions in design or the analysis of covariance. In one experiment (Exp. 2) in which a Latin square arrangement was used in order to equalize the order of tests within groups a substantial mean square was segregated for 'order of injections'. When, however, two further $4 \times 4$ squares were appended in the same experiment (this part of the assay has been omitted in the text since some errors occurred) the mean square for order of injections became less than the error mean square.

### *Independence of $\bar{y}_u - \bar{y}_s$ and $b$*

In deriving the expression for $s_M$, the standard error of the ratio $(\bar{y}_u - \bar{y}_s)/b = M$, it was assumed that numerator and denominator of the fraction are independent. This assumption holds only if the slope of the

regression line does not alter in the course of the assay, since any real variation, as distinct from sampling variation, of the slope entails corresponding variations of the differences between effects. If the slope varied there should thus be significant correlation between mean difference of effects and slope in successive groups, or between successive values of $(y_u)_1+(y_u)_2-(y_s)_1-(y_s)_2$ and $(y_u)_1+(y_s)_1-(y_u)_2-(y_s)_2$.

The strength of this correlation has been measured in each assay. In no instance did the correlation coefficient attain the 5 % level of significance. When the correlation coefficients from all assays (except Exps. 2, 4 and 11) were pooled by means of the $z$ transformation (Fisher, 1938) for 31 degrees of freedom, a non-significant negative correlation of $r=-0{\cdot}048$ was obtained. On the available evidence there is thus no reason to assume that $\bar{y}_u-\bar{y}_s$ and $b$ are correlated or that a change of slope of the regression line occurs in the course of the assay.

## Summary

1. A method is described for conducting a biological assay on an isolated preparation in such a way that a null hypothesis may be adequately tested and the error of the assay may be estimated from the data of the experiment itself.

2. The method is applicable in its present form, if there is a linear relationship between log dose and effect over a given range, and if the slope of the regression line does not alter in the course of the experiment.

3. The method has been applied to the assay of histamine on the guinea-pig's intestine. In assays with solutions of known composition the $P$ 0·99 limits of error ranged in different experiments from −8·2 and +8·8 % to −27·7 and +32·8 %.

I am indebted to Dr I. O. Irwin and Prof. J. H. Gaddum for reading and criticizing the manuscript. The work was aided by a grant of the T. Smythe Hughes Medical Research Fund.

## REFERENCES

Bliss, C. I. [1935]. *Ann. appl. Biol.* **22**, 134.

Bliss, C. I. & Marks, H. P. [1939 *a*, *b*]. *Quart. J. Pharm.* **12**, 82, 182.

Fieller, E. C. [1940]. *J. Roy. Statist. Soc.* Suppl. **7**, 1.

Fisher, R. A. [1937]. *The Design of Experiments*, 2nd ed. London and Edinburgh: Oliver and Boyd.

Fisher, R. A. [1938]. *Statistical Methods for Research Workers*, 7th ed. Edinburgh and London: Oliver and Boyd.

Fisher, R. A. & Yates, F. [1938]. *Statistical Tables*. Edinburgh and London: Oliver and Boyd.

Gaddum, J. H. [1933]. *Spec. Rep. Ser. Med. Res. Coun., Lond.*, no. 183.

Snedecor, G. W. [1938]. *Statistical Methods*. Ames, Iowa: Collegiate Press, Inc.