Brit. J. Pharmacol. (1948), 3, 328.

A MODIFICATION OF THE METHOD OF DALE AND LAIDLAW FOR STANDARDIZATION OF POSTERIOR PITUITARY EXTRACT

BY

PAMELA HOLTON

From the Department of Pharmacology, Oxford

(Received June 28, 1948)

The activity of a posterior pituitary extract is still generally determined by the method introduced by Dale and Laidlaw (1912). They used the uterus of a virgin guinea-pig, and the assay consisted in setting limits to the potency of the test solution in terms of the standard solution by the usual ABBA grouping of doses. The method suffered from three defects.

1. Suitable guinea-pigs were relatively scarce.
2. The assay often required many hours.
3. The error of the method was about 20 per cent (Gaddum, 1938).

A different method of assay was described by Thompson (1944). He used the relationship between the dose of pituitary and the fall in blood pressure in the chicken which had been investigated by Smith and Vos (1943). Thompson obtained results with very small errors and was able to assess the reliability of each result from the experimental data. He claimed that the method was rapid and easy to perform. There is, however, one drawback to the method. It is possible that the active principle responsible for the fall in blood pressure is not identical with that which makes the uterus contract. Thompson himself found a significant difference between potency on the blood pressure and potency on the guinea-pig uterus in 1 out of 17 parallel determinations and suggested that the blood pressure method cannot always be relied upon to estimate the oxytocic activity.

It seemed best, therefore, to return to the original method of standardization using the isolated uterus as a test preparation. Various modifications of the original method have recently been introduced. Morrell, Allmark, and Bachinski (1940) used 8 strips of muscle which were cut from the same uterus and suspended in the bath together. They showed that the percentage of strips responding was related to the log of a dose of pituitary put into the bath. By this measurement of "all or nothing" responses Morrell *et al.* obtained more accurate estimations of oxytocic activity than previous workers, but they made no statement of the reliability of each estimate. One estimate of potency had an error of 35 per cent. If this was associated with a high standard error for that particular determination the observer would have been warned that the estimate was unreliable, but it was included in a table of results all of which were considered satisfactory. As Bliss (1941) pointed out, "a determination of potency should always include an estimate of its error, computed as an integral part of the assay." Gaddum (1933) provided the basis for a biological assay which included an estimate of the precision of the result and showed that his formulae applied equally well to measured responses and to quantal data. Schild (1942) published an account of a null hypothesis assay conducted on statistically sound principles. His method has been followed almost completely to obtain the results in this paper.

METHODS

A modification of the classical method of Dale and Laidlaw was adopted.

The rat's uterus was used as test preparation since rats are cheaper and more easily obtained than guinea-pigs. Moreover, Garcia de Jalon, Bayo Bayo, and Garcia de Jalon (1945) have shown that most rats' uteri do not contract spontaneously in Locke's solution in which the calcium and glucose concentrations are ¼ and ½ respectively of the usual. One horn of the uterus was suspended in a 10-ml. bath in modified Locke's solution. One end was attached to an

isotonic and linear lever of Schild's (1947) design. The lever was so arranged that the load on the uterus was about 1.2 g. and the contractions were magnified four times. It was equipped with a glass frontal writing point and its angular excursion was limited to 30° each side of the horizontal.

The bath was supplied with oxygen containing 5 per cent CO_2; this was to eliminate possible changes of *p*H, but the CO_2 is probably unnecessary when the Locke's solution is replaced at short intervals.

Non-pregnant white rats were used. Their weights varied from 120 to 200 g. No record was kept of the position in the oestrus cycle of each rat. Table I shows the variation which may be expected within a group of rats.

TABLE I

Assay No.	No. of rats	Spontaneous contractions	Remarks
1, 2, 3, 9, 10	4	No	Satisfactory assays
5, 6, 7	2	Yes, overcome	
4	1	Yes, overcome	Unsatisfactory assays
8	1	Yes	
—	1	Yes	Useless

It will be seen that 6 rats were good, 2 poor, and 1 useless out of a total of 9. In practice the rats in assays 4 and 8 would have been discarded after a short time. The uterus in assay 4 produced only 12 contractions. In assay 8 the uterus had so much spontaneous activity that it was obviously unreliable.

Performance of an assay

1. The rat is killed by a blow on the head and bled out. The uterus is removed and one horn is suspended in the bath and attached to the lever. The temperature must be constant at about 32° C.

2. Doses of pituitary should be given without delay at regular intervals of 3 or 4 min. Two doses must be found such that the contraction for the higher dose is at least twice as great as that for the lower dose. Dose ratios of 4:3, 3:2, 8:5, and 2:1 have been used (4:3 is the most common). In order to ensure linearity between response and log dose it is best to use contractions below 80 per cent of maximal (see *Tests for Linearity* below).

If spontaneous contractions are troublesome they may sometimes be overcome by lowering the temperature or by reducing the time between successive doses. If the uterus is very insensitive (i.e., will not respond to 0.05 unit) the sensitivity may sometimes be increased by raising the temperature.

3. The strength of the unknown must be guessed by matching it with the standard. The error of the assay is much smaller if a good guess is made.

4. The assay is now continued exactly as described by Schild (1942) for histamine assays. Four doses are used, two of standard and two of unknown. The ratio of high dose to low dose should be the same for standard and unknown. Each dose is given once in each group of four doses and its position within the group is decided by chance. Fig. 1 shows the record obtained in one assay. The drum was turned on 15 sec. before the dose was given. The uterus

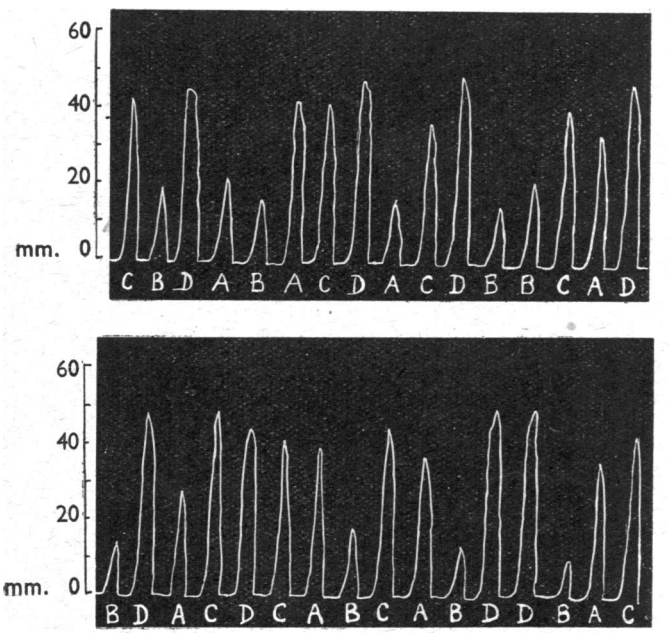

FIG. 1.—An assay of pituitary. The record shows thirty-two contractions of a rat's uterus. The contractions are responses to four different doses of pituitary A, B, C, and D, each of which is given once in each group of 4 contractions. There are eight groups in all. A = 0.05 units, B = 0.04 units, C = 0.064 units, and D = 0.08 units. B and C were treated as "standard." A and D were treated as "unknown." A : D = B : C = 5 : 8. Estimate of unknown: standard = 1.25. True value unknown: standard $= \frac{0.05}{0.04}$ = 1.25. A dose was put into the bath every 4 min. and washed out after 45 sec. Weight of rat 140 g. Temperature 34°–36° C. Load on uterus 1.3 g. Total experimental time 4 hr.

contracted for 45 sec., after which the drum was stopped and the bath was refilled with Locke's solution. A and D were dilutions of the unknown, and B and C were dilutions of the standard; C was stronger than B in the proportion 8 to 5. The ratio of D to A is the same as that of C to B.

TABLE II

Heights of contractions in mm.$=y$

	Dose units	log dose $=x$	1	2	3	4	5	6	7	8	Sum
D			48	50	52	51	49	45	52	51	398
C	0.064	0.806	43	43	38	42	48	43	44	44	345
A			21	43	17	35	28	39	37	35	255
B	0.04	0.602	21	18	16	25	16	19	14	12	141
Group totals			133	154	123	153	141	146	147	142	1139

RESULTS

The results of this assay are shown in Table II.

In Fig. 2 the mean height of contraction in response to each dose is plotted against the log dose. The "null hypothesis" states that there is no difference between the standard solution and the unknown solution, and so the standard log doses are used in plotting the points for the unknown. Since the unknown and standard contained the same active principle the two lines joining the mean responses to high and low doses of each are parallel. The lines are the regression lines relating height of contraction to log dose, and they were drawn by eye to be the parallel lines that fitted the points most nearly. The regression coefficient b is the slope of each line. An estimate of the potency of the unknown solution can be made by using the two regression lines. The horizontal distance (M) between the lines is the difference between log doses producing the same effect, and it is therefore the log of the ratio of unknown to standard. By elementary geometry

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

where $\bar{y}_u$ and $\bar{y}_s$ are the mean responses to unknown and standard respectively; if b is known M may be calculated.

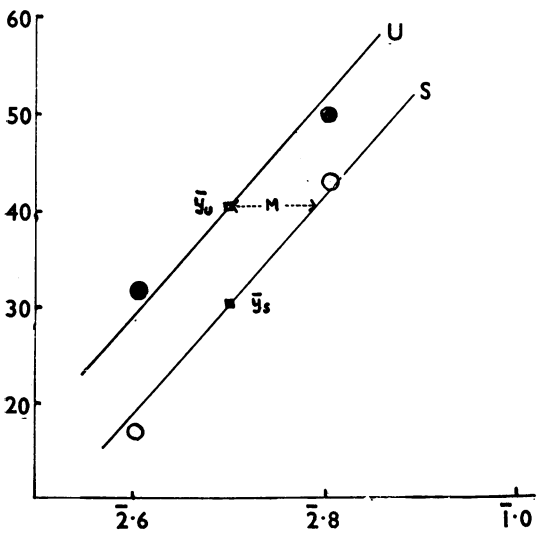

FIG. 2.—Ordinates: mean height of contraction in mm. Abscissae: $\log_{10}$ dose in units of pituitary. White circles are the mean responses for high and low doses of standard, black circles are the mean responses for high and low doses of unknown. The black squares $\bar{y}$ and $\bar{y}_u$ are the means of all the responses to standard and unknown. S and U are the regression lines relating response to log dose of standard and unknown. The lines were drawn by eye. M is the distance between the lines. M = log unknown − log standard = log $\frac{\text{unknown.}}{\text{standard}}$ Graphically $M = 0.09$. $\therefore \frac{\text{unknown}}{\text{standard}}$ = antilog 0.09 = 1.23.

An estimate of potency of a preparation does not give much information about the true potency unless it is accompanied by a statement of its error. The simple graphical measurement of M as described above does not enable us to estimate the reliability of the estimate M; for this we need to calculate the standard error of M. In order to do this M should be calculated rather than determined graphically, and its standard error must also be calculated. The methods of calculation and analysis of variance given by Schild are simple and quick. An example is given below. When the method of assay is slightly varied, for instance, to include three dose levels of standard and of unknown, more complex methods must be used. These may be found in any standard textbook of statistics, but are given particularly clearly by Finney (1947).

Calculation

In order to determine M we must calculate the numerator and denominator of the expression

$$\frac{\bar{y}_u - \bar{y}_s}{b}$$

1. From Table II $\bar{y}_u = \frac{398 + 255}{\frac{1}{2}n}$

and $\bar{y}_s = \frac{345 + 141}{\frac{1}{2}n}$

where n is the total number of responses; $n = 32$.

2. When only 4 doses are used the equation for b is:

$$b = \frac{\text{sum of responses to high doses} - \text{sum of responses to low doses}}{\frac{1}{2}n \times (\log \text{high dose} - \log \text{low dose})}$$

With the data from Table II

$$b = \frac{(398 + 345) - (255 + 141)}{\frac{1}{2}n \times d}$$

where d = log high dose − log low dose
= log 0.064 − log 0.04
= $\bar{2}.806 - \bar{2}.602$
= 0.204

The value of b will be needed later. It is

$$b = \frac{398 + 345 - 255 - 141}{16 \times 0.204}$$
$$= 106.3$$

For calculation of M, however, it is simpler to use the above expressions, noticing that $\frac{1}{2}n$ occurs in numerator and denominator of the expression for M and so disappears.

$$M = \frac{[(398 + 255) - (345 + 141)] \times 0.204}{(398 + 345) - (255 + 141)}$$
$$= \frac{167 \times 0.204}{347}$$
$$= 0.09811$$

$$\frac{\text{Unknown}}{\text{Standard}} = R = \text{antilog } M = 1.25.$$

Since this assay was conducted with two dilutions, "unknown" and "standard," of the same pituitary extract, the true value of the ratio $\frac{\text{"unknown"}}{\text{standard}} = \rho$ was known. The "unknown" doses of D and A were 0.08 unit and 0.05 unit respectively. Hence

$$\rho = \frac{D}{C} = \frac{A}{B} = \frac{0.08}{0.064} = 1.25$$

In order to make a statement about the reliability of the estimate R, the standard error (s_M) of M must be calculated. This is obtained from the standard deviation (s_y) of a single observation (y). The standard deviation of a single observation—the error of the assay—can only be obtained through an analysis of variance which segregates variation due to known causes from that due to the variability of the preparation.

Table III shows that the sums of squares of deviations attributable to the variation

1. between groups (variation of the preparation with time)

$$= \tfrac{1}{4}(133^2 + 154^2 + 123^2 + \text{etc.}) - \frac{1139^2}{32} = 182$$

2. between standard and unknown

$$= \frac{(398 + 255 - 345 - 141)^2}{32} = 872$$

3. between high doses and low doses (regression)

$$= \frac{(398 + 345 - 255 - 141)^2}{32} = 3763$$

4. between the slopes of the two regression lines (deviations from parallelism)

$$= \frac{(398 + 141 - 345 - 255)^2}{32} = 116$$

The sum of squares of deviations of all the observations from the common mean = total sum of squares

$$= (48^2 + 50^2 + 52^2 + \text{etc.}) - \frac{1139^2}{32}$$
$$= 5536$$

When all the sums of squares due to known causes have been subtracted from the total the remainder is due to the error of the assay. This remainder, divided by its degrees of freedom, gives the variance of a single observation according to the usual formula for standard deviation. The standard deviation of a single observation is the square root of the variance. In the example the residual sum of squares of deviations = 5536 − 182 − 872 − 3763 − 116 = 603. This had 21 degrees of freedom because from the total number (32 − 1) 7 must be subtracted for the 8 different groups and one each for the other three known sources of variation.

$$\text{The variance } s_y^2 = \frac{603}{21} = 28.71$$

TABLE III

Source of variation	Sum of squares	Degrees of freedom	Variance	F	P
1. Groups ..	182	7	26	1.1	>0.05
2. Standard and unknown ..	872	1	872	30	<0.01
3. Regression ..	3763	1	3763	131	<0.01
4. Deviation from parallelism..	116	1	116	4	>0.05
5. Error ..	603	21	28.71		
Total ..	5536	31			

F is the variance ratio, e.g., for "Deviation from parallelism" $F = \frac{116}{28.7} = 4$. A table of F for $n_1 = 1$ and $n_2 = 21$ (where n_1, n_2 are the degrees of freedom for "Deviations from parallelism" and for "Error" variance respectively) gives $P > 0.05$. This means that there is a proba-

bility of more than 5 per cent that this deviation from parallelism would occur by chance and so the slopes of the two regression lines are not significantly different. The values given in the column P are the probabilities that such variation would occur by chance. Thus there is less than a one per cent probability that so great a variation between standard and unknown would have occurred had the unknown not been different from the standard. The unknown is said to be significantly different from the standard. Similarly, the regression is significantly different from zero. The variation between groups is not greater than would be expected by chance in more than one in twenty experiments. (Here $F = \frac{28{\cdot}7}{26} = 1.1$ because $F = \frac{\text{larger mean square}}{\text{smaller mean square}}$ and hence $n_1 = 21$ and $n_2 = 7$).

The standard error of M

The standard error of the estimate of M may be obtained from Schild's formula:

$$s_M^2 = \frac{4s_y^2}{nb^2}\left(\frac{M^2}{d^2} + 1\right)$$

In the example

$$s_M^2 = \frac{4 \times 28.71}{32 \times (106.3)^2}\left[\frac{(0.09811)^2}{(0.2041)^2} + 1\right]$$

$$= \frac{114.84 \times 1.2312}{361600}$$

$$= 0.0003909$$

$$\therefore s_M = \sqrt{0.0003909} = 0.01976$$

The fiducial limits of M

The limits of the value of M are obtained from adding and subtracting the standard error of M multiplied by t. The value of t depends on the degrees of freedom associated with the error sum of squares. In the example there were 21 degrees of freedom, and so t is obtained from the table under $n = 21$. For a probability level of 0.05, $t_{21} = 2.08$.

$$s_M \times 2.080 = 0.0421$$

$$\therefore M \pm s_M \times 2.080 = 0.0981 \pm 0.0421$$

Hence the fiducial limits ($P = 0.05$) for M are 0.0560 and 0.1402

The limits for R are :

$$R_1 = \text{antilog } 0.0560 = 1.14$$
$$R_2 = \text{antilog } 0.1402 = 1.38$$

That is to say, there is only one chance in twenty that these results would have been obtained if the true value of the unknown lay outside the limits 1.14 – 1.38.

Now $1.38 - 1.14 = 0.24$ and $\frac{0{\cdot}24}{2} = 0.12$

$\therefore$ the approximate range of results is

$$R \pm 0.12 = 1.25 \pm 0.12$$

Hence the limit of error is $\frac{0.12}{1.25} \times 100 = 9.6\%$.

(This assumes that R_1 and R_2 are equidistant from R, which is not strictly true since M_1 and M_2 are equidistant from M, which is on a logarithmic scale, but the assumption is true enough for an approximate calculation of the limits of error.)

TABLE IV

Assay	Time	n	ρ	R	R_1	R_2	E	s_y/b
1	5	36	1.25	1.22	0.97	1.55	2.6	0.145
2	4½	32	1.25	1.25	1.14	1.38	0	0.050
3	5	32	1.14	1.10	1.06	1.13	3.6	0.018
4	2	12	0.80	0.91	0.81	1.02	14	0.044
5	4	24	1.08	1.10	0.96	1.25	1.9	0.062
6	3½	24	1.07	1.09	1.02	1.18	1.25	0.034
7	3	24	1.07	1.05	0.96	1.15	1.25	0.044
8	8	24(-4)	0.91	1.06	0.77	1.56	15	0.131
9	2½	20	1.00	0.99	0.94	1.04	0.85	0.022
10	2	20	1.07	1.01	0.96	1.06	5.8	0.022

Table IV includes all of a recent series of 10 assays in order to show the variation which may be expected from a group of 9 rats. Eight out of these 10 assays were considered satisfactory. Two assays (4 and 8) were unsatisfactory and in practice the results from them would have been subjected to further tests. The uterus gave only 12 contractions in assay 4 and then lost all sensitivity. In assay 8 the uterus had a great deal of spontaneous activity and the contractions were obviously unreliable. The assays which are bracketed were performed on the two horns of the uterus from the same rat.

In assays 4–10 the true value of the ratio of unknown to standard (ρ) was unknown to the observer until after the results were obtained. The columns R, R_1 and R_2 show the estimate of ρ and the two fiducial limits (R_1 and R_2) calculated as described above for $P = 0.05$. It will be noticed that ρ falls within the range $R_1 - R_2$ except in assays 3, 4, and 10, where it is just outside the range. Recalculation of R_1 and R_2 according to Schild's equation for the exact limits, which is given in a footnote to his paper, did not alter the values. It must be borne in mind, however, that the statistical analysis is based on the assumption that all measurements are absolutely

accurate and the calculation of error does not include unavoidable errors in pipetting which occur during the dilution of the test solutions. Hence, in order to make a reliable statement about the potency of an unknown, the fiducial limits should be extended by 4 per cent (this allows for 1 per cent error for each dilution). The figures in the column "Time" are the times in hours required for the experimental part of each assay and E is the actual percentage error in the estimate of potency.

The ratio of the standard deviation of a single response to the slope of the regression line is shown in the last column of Table IV (s_y/b). It is an inverse measure of the usefulness of the preparation and for a satisfactory assay (s_y/b) should not be much greater than 0.05.

Gaddum's calculation

Gaddum (1938) calculated the standard error of pituitary assays performed by the method of Dale and Laidlaw. He used only those assays which were considered satisfactory and obtained a value for the standard error of 7.73 per cent, which gave fiducial limits at 19.9 per cent ($P = 0.01$). If his calculation is repeated with the results of the assays in Table IV, excluding assays 4 and 8 for reasons given above and using ρ instead of his mean estimated potency, the standard error is 2.84 per cent and fiducial limits are at 9.4 per cent (for $P=0.01$ $t_7=3.499$), which is about half the value found by Gaddum for the original Dale and Laidlaw method.

The dose-response relationship

The calculations and statistical analysis of the results are based on the assumption that the relation between response and log dose is exactly linear. In fact, the curve relating response to log dose is sigmoid and approximates to linearity over part of the response range. Clearly it should be the aim of the observer to confine the responses of the uterus to this linear part of the range. Since the region of linearity varies from one preparation to another it must be left to the observer to decide its position, but it has been found that for most uteri any responses less than 80 per cent of the maximal can be used.

Tests for linearity

Assay No. 1 was exceptional in that 6 different doses (3 standard and 3 unknown) were used. Here a test of linearity was possible without knowing the true value of the "unknown." It showed that there was no significant departure from linearity. The calculations for this assay were slightly more complex than for the 4-dose assays and were performed as described by Finney (1947). The 6-dose design is not recommended for routine work because the experimental procedure is more complicated and the time interval between successive doses may be too short for the additional diluting needed for 6 doses.

It should be realized, however, that the 4-dose design suffers from the defect that no test of linearity can be applied from the data of a single assay. The worker is safeguarded to a certain extent by the test for parallelism. The variance due to deviations from parallelism would be significant if there were any serious departure from linearity and if M were not negligible.

Table V shows the variance ratios (F) for deviations from parallelism. An F value smaller than that for $P = 0.05$ indicates that the deviation was not significant. In assay 9 the deviation from parallelism was just significant. In this assay the strengths of the unknown and standard were exactly the same owing to a lucky initial guess of the potency. Hence it was not possible to calculate the deviation from linearity as was done by Schild. However, the estimate of the potency of the "unknown" in assay 9 was a good one, which is in accordance with Schild's conclusion that the method is relatively insensitive to deviations from parallelism. The last two columns in Table V show the highest and the lowest contractions as recorded on the drum. The maximum contraction was generally not recorded, but general experience with the rat uterus preparation indicates that most uteri are capable of 100 mm. contraction with the same apparatus, and so the figures in the table are approximate percentages of the maximum.

TABLE V

Assay	Deviation from parallelism F	$F(P=0.05)$	$F(P=0.01)$	Highest contraction mm.	Lowest contraction mm.
1	26	249		65	2
2	14	244		52	16
3	10	244		73	3
4	34	249		61	3
5	10	244		32	15
6	1.5	244		35	11
7	2	4.3		49	1
8	1.5	4.75		85	0
9	5.2	4.75	9.33	80	17
10	50	244		96	11

SUMMARY

1. Schild's null hypothesis method was applied to the assay of posterior pituitary extract according to Dale and Laidlaw.

2. A rat's uterus was used as the test preparation.

3. One assay is described and the result is calculated by Schild's method.

4. Six rats out of a total of nine could be used for satisfactory assays.

5. Eight satisfactory assays were performed. The mean percentage error was 2.16. Fiducial limits for the estimate of potency were calculated. The mean experimental time of an assay was $3\frac{1}{4}$ hours.

This work was done at the suggestion of Prof. J. H. Burn. My thanks are due to Dr. D. J. Finney and his staff for help in the calculations.

REFERENCES

Bliss, C. I. (1941). *Ind. Eng. Chem.*, **12**, 84. Quoted by S. A. Thaler (1946). *Vitamins and Hormones*, **4**, 318.
Dale, H. H., and Laidlaw, P. P. (1912). *J. Pharmacol.*, **4**, 75.
Finney, D. J. (1947). *J. Hyg.*, **45**, 397.
Gaddum, J. H. (1933). *Med. Res. Council Special Report*, Series No. 183.
Gaddum, J. H. (1938). *Quart. J. Pharm.*, **11**, 697.
Garcia de Jalon, P., Bayo Bayo, J., and Garcia de Jalon, M. (1945). *Farmacoterapia Actual*, No. 11, 313.
Morrell, C. A., Allmark, M. G., and Bachinski, W. M. (1940). *J. Pharmacol.*, **70**, 440.
Schild, H. O. (1942). *J. Physiol.*, **101**, 115.
Schild, H. O. (1947). *Brit. J. Pharmacol.*, **2**, 189.
Smith, R. B., Jr., and Vos, B. J., Jr. (1943). *J. Pharmacol.*, **78**, 72.
Thompson, R. E. (1944). *J. Pharmacol.*, **80**, 373.