# THE ASSAY OF DIGITALIS.*,[1]

## I. Criteria for Evaluating Various Methods Using Frogs.

BY LLOYD C. MILLER, CHESTER I. BLISS[2] AND HERBERT A. BRAUN.

An important problem in bioassay is to minimize the variations which seem to be inherent in the responses of living organisms. In the bioassay of digitalis, using frogs, numerous modifications in the technical procedure have been advocated toward this end. Length of the period of observation, weight, sex and species of the frogs, volume and alcoholic content of the injected material, environmental temperature, all these and other factors have been varied by the many who have worked in this field.

Adoption of certain restrictions may possibly bring about a reduction in the variations associated with the present official method. All will agree that such adoption should be based upon a critical examination of the results of well-planned experiments. With the preparation of the twelfth revision of the United States Pharmacopœia scheduled to begin next year, it seems likely that the assay of digitalis will be the subject for considerable investigation in the near future. At present there exist no generally accepted criteria by which the results of such studies can be evaluated objectively. The purpose of this paper is to present certain criteria which should meet this need and it is hoped that their availability will encourage the designing of experiments in a manner that will permit their application.

As a first prerequisite it is suggested that all investigators test their proposed modification of the method on some standard material, preferably a uniform preparation of powdered digitalis. In view of the adequacy[3] of the supply of U. S. P. XI Reference Digitalis Powder, this is probably the best one for all to use. The comparability of the results on the same standard for each particular proposal will furnish the first step in bringing the data of all laboratories to a common meeting-ground. A second prerequisite is still more important to guarantee fully the desired result and that is to plan each experiment so that it will be self-contained. Only in this way can the experimental error be calculated satisfactorily and the results evaluated objectively. To supply this essential internal check, each modification in procedure should be tested at more than one dosage level of any given material. Thus if the variation in the amount of alcohol in the test material is suspected of influencing the assay, tests should be run in parallel with two or more concentrations of alcohol, each at two or more dosage levels. This will permit the construction of a dosage-effect curve for each concentration of alcohol tested and

* Presented before the Scientific Section, A. PH. A., Atlanta meeting, 1939.

[1] From the Division of Pharmacology, Food & Drug Administration, U. S. Department of Agriculture, Washington, D. C.

[2] Consultant.

[3] Prof. E. Fullerton Cook, Chairman of the U. S. P. XI Revision Committee, has indicated that there is ample U. S. P. XI Reference Digitalis Powder available; less than one-fifth of the original supply has been distributed in the three years since its adoption.

(Note added in press.)—At the conference on the Assay of Digitalis held at Atlanta, Ga., August 23, 1939, it was decided that the unusually high potency and possibly other characteristics of the U. S. P. XI Reference Digitalis Powder made it less suitable for use in collaborative investigations than a composite powder more nearly like the average digitalis of commerce. Such a composite is being prepared for the projected collaborative U. S. P. study.

will afford a means of estimating the experimental error associated with the comparison. It cannot be assumed that this error will remain constant, even from one day to the next, so that no serious comparisons should ever be attempted on the basis of this assumption.

When comparative data have been obtained, as, for example, on two modifications of the procedure, a method of computation is needed that will extract from the results all of the relevant information. In the following sections a method of computation is outlined by which certain informative characteristics or parameters can be obtained from the data and the criteria to be presented involve comparisons of these parameters. If the data of all investigations are such as to permit drawing these comparisons, the result will be a tremendous simplification in the task of coördinating the results and deriving from them objective evidence on which decisions regarding the U. S. P. XII monograph for the assay of digitalis can be based.

## THE DETERMINATION OF RELATIVE POTENCY.

The classical work of Trevan (1) forms a starting point for the consideration of a suitable procedure. It is rather generally known from his study that an S-shaped curve is obtained when the percentage of frogs showing positive effects of digitalis is plotted against the dosage. Less generally known is the fact that the same data can be transformed into a straight line. This transformation involves plotting the logarithm of the dose against an expression of the percentage effect in units derived from the normal frequency curve. These units were first suggested in a form suitable for computation by Gaddum (2) ("N. E. D.") and later by Bliss (3) in the somewhat more convenient "probits." Tables of probits may be looked upon in the same light as tables of logarithms, which are so familiar that most people use them without concern as to their derivation. This fact may encourage those who hesitate to adopt the percentage-to-probit transformation prior to becoming acquainted with its theoretical basis, which has been described elsewhere (4). The determination of the dosage-effect curve from small numbers of animals has been treated in detail in a recent publication (5). The computation of the basic curve may be reviewed by means of a numerical example.

*Experimental Details.*—Male frogs weighing from 17 to 30 Gm. were distributed into groups of five in the assay tank which was maintained at 19.6 ± 0.2° C. By preliminary trials, two series of three doses of digitalis were selected such that their logarithms were equally spaced, one or two digit numbers. They were prepared from a tincture of U. S. P. XI Reference Digitalis Powder (0.0745 Gm. per cc.) and assigned to the groups of frogs at random so that six groups, or thirty frogs, received each dose from each series. One series of doses was contained in a volume of 0.01 cc. per Gm. and was injected intramuscularly, dividing the dose between the two thighs as suggested by Dooley and Higley (6). The second series, contained in 0.02 cc. per Gm., was injected into the ventral lymph sac. One-half of the frogs receiving each series of doses were pithed and examined one hour after injection, while the remaining frogs were pithed and examined after

TABLE I.—INFLUENCE OF ROUTE OF INJECTION ON POTENCY.

U. S. P. XI Reference Standard Digitalis Given Intramuscularly and by Lymph Sac in One-Hour Method.

| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|
| Route. | Dose, *Cc./Kg.* | Log-Dose. *x.* | Result. | % | Empirical Probit. | Expected Probit. *Y.* | Corrected Probit. *y.* | Weight. *w.* | Products. *wx.* | *wy.* |
| Lymph Sac | 5.62 | .75 | 2/15 | 13.3 | 3.889 | 3.93 | 3.891 | 6.22 | 4.6650 | 24.20202 |
| | 7.08 | .85 | 5/15 | 33.3 | 4.569 | 4.53 | 4.570 | 8.81 | 7.4885 | 40.26170 |
| | 8.91 | .95 | 8/15 | 53.3 | 5.084 | 5.13 | 5.084 | 9.49 | 9.0155 | 48.24716 |
| Intra-Muscular | 2.82 | .45 | 2/15 | 13.3 | 3.889 | 3.86 | 3.892 | 5.86 | 2.6370 | 22.80712 |
| | 3.55 | .55 | 5/15 | 33.3 | 4.569 | 4.64 | 4.571 | 9.11 | 5.0105 | 41.64181 |
| | 4.47 | .65 | 10/15 | 66.7 | 5.431 | 5.41 | 5.430 | 8.98 | 5.8370 | 48.76140 |

an interval of four hours. Thus the complete experiment or "assay" involved four dosage-effect curves, showing the reaction after one hour to digitalis injected in the lymph sac and injected intramuscularly, and after four hours to the same two modes of injection. For present purposes only the one-hour results will be considered.

*Computation of the Dosage-Effect Curve.*—The data and the first stages of the computation are given in Table I. Column 1 indicates the route of administration of the doses of digitalis listed in column 2, the logarithms of which ($x$) are in column 3. Column 4 lists the fractions representing in the denominators the total number of frogs injected and in the numerators the number showing systolic standstill at the time of observation. The results have been changed to percentages in column 5 and to the probits corresponding to these percentages in column 6. The next step in the procedure is to plot the corresponding entries in columns 3 and 6 against each other on cross-section paper (Fig. 1). Each set of points can be fitted by eye with a straight line (not shown in the figure) which represents the first approximation to the relation between dose and effect under the two conditions of the experiment, namely, intramuscular and lymph sac injection. Each such line is the best graphic estimate of the relation that one would expect to find if an infinitely large number of similar frogs could be used in a single test. Provisionally, then, it can be looked upon as the *expected* relationship, and for each dose there is a corresponding *expected* probit.[1] The expected probits ($Y$, column 7) are used to enter appropriate tables[2] (7, 8) to obtain both the corrected probit ($y$, column 8) and the weight ($w$, column 9), which also depends on the number of frogs used, to be assigned to each observation. In columns 10 and 11, the respective products of $wx$ and $wy$ are tabulated.

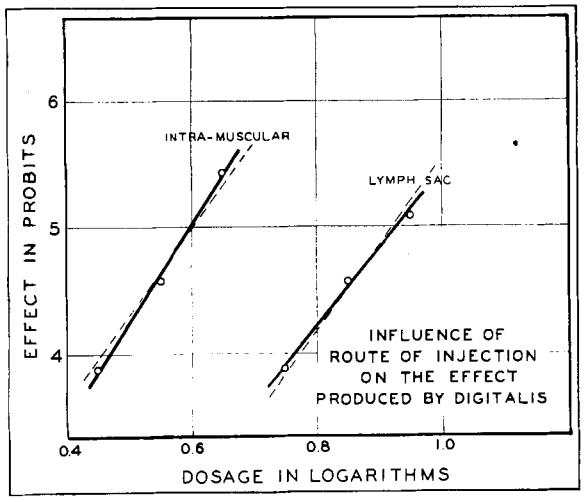


Fig. 1.

From these values it is possible to compute the first calculated approximation, which is usually sufficient, by formulas which are reproduced for convenience in Table II together with the numerical terms computed from the data in Table I.

The parameters listed in Table II are the basic units defining each dosage-effect curve and will be used subsequently in several combinations. The position of the best-fitting straight line for any single series is determined by the weighted mean probit ($\bar{y}$) at the weighted mean log-dose ($\bar{x}$), through which point the line will pass with a slope given by $b$. These computed curves have been plotted as solid lines in Fig. 1. If the experiment is conducted so that the frogs receiving the different doses are really equivalent in susceptibility, the observed probits should differ from the computed line only by chance. This is checked by the last term in Table II, $\chi^2$ ("chi-square"), which shows whether or not the variation is too great to be considered due to errors of sampling. When $\chi^2$ exceeds the limit for $P = 0.05$ (9) there is less than one chance in twenty that the frogs used at the different dosages were really equivalent and not much importance can be attached to the results of so erratic a test. With the proper experimental precautions, however, $\chi^2$ usually

[1] The expected probits may be read directly from the graph if only three or four doses are used in the experiment. With a larger number of doses it is probably more rapid to calculate the expected probits as suggested elsewhere (5).

[2] Expanded tables giving the corrected probits in terms of the expected probits with a minimum of interpolation for all possible results in groups of twenty animals or less have been prepared in mimeograph form and are available, upon request, from the Division of Pharmacology, Food and Drug Administration.

TABLE II.—PARAMETERS OF DATA IN TABLE I.

| Parameter. | Formula. | For Lymph Sac Curve. "Standard" | For Intramuscular Curve. "Unknown" |
|---|---|---|---|
| $S(w)$ | (Sum of Col. 9) | 24.52 | 23.95 |
| $S(wx)$ | (Sum of Col. 10) | 21.1690 | 13.4845 |
| $\bar{x}$ | $\frac{S(wx)}{S(w)}$ | $\frac{21.1690}{24.52} = .8633361$ | .5630271 |
| $S(wy)$ | (Sum of Col. 11) | 112.71088 | 113.21033 |
| $\bar{y}$ | $\frac{S(wy)}{S(w)}$ | $\frac{112.71088}{24.52} = 4.5966917$ | 4.7269449 |
| $[wx^2]$ | $S(wx^2) - \bar{x}\,S(wx)$ | $18.42870 - .8633361(21.1690) = .152738$ | .144336 |
| $[wxy]$ | $S(wxy) - \bar{y}\,S(wx)$ | $98.208762 - 4.5966917(21.1690) = .901395$ | 1.120621 |
| $b$ | $\frac{[wxy]}{[wx^2]}$ | $\frac{.901395}{.152738} = 5.901577$ | 7.763974 |
| $[wy^2]$ | $S(wy^2) - \bar{y}\,S(wy)$ | $523.454590 - 4.5966917(112.71088) = 5.357424$ | 8.745435 |
| $\chi^2$ | $[wy^2] - b[wxy]$ | $5.357424 - 5.901577(.901395) = .037772$ | .044963 |

will fall within the sampling error as required for a valid assay. The equations in the following sections are based upon the assumption that this, in fact, is true. Chi-square not only tests this assumption critically but provides as well a useful check on the arithmetic of the whole computation, since mistakes often lead to excessively large or to impossible, negative values of $\chi^2$.

*The Measurement of Relative Potency.*—Up to this point each individual curve has been computed separately and the explanation has been relatively brief, since a full description of the theory and computation is already in the literature (5). The individual results must be combined to obtain an expression of the most probable value of the potency of the unknown sample or experimental factor (as in the above example) relative to that of the standard; the calculation involved in these steps may be given in more detail. Gaddum (2) first proposed a formula for this expression, designating it as "M," and defining it as

$$M = \log \frac{\text{potency of first preparation (unknown)}}{\text{potency of second preparation (standard)}}$$

Thus $M$ is the logarithm of the ratio of potencies. Remembering that potencies are inversely proportional to the doses producing equivalent biological effects, it may be seen that $M$ has its simplest form when exactly equal effects have been obtained, in which case

$M$ = weighted mean log-dose of standard − weighted mean log-dose of unknown = $\bar{x}_S - \bar{x}_U$, where the subscripts "$S$" and "$U$" refer, respectively, to the standard and the unknown sample, or may represent two variations in technique with the same material.

Corresponding to the weighted mean log-dosages, $\bar{x}_S$ and $\bar{x}_U$, the weighted mean effects in probits for standard and unknown may be designated as $\bar{y}_S$ and $\bar{y}_U$, respectively. In the above derivation of $M$ it was assumed that the latter were equal, *i. e.*, that equal biological effects had been obtained. It seldom happens, however, that equivalent doses of two preparations produce exactly the same biological effect even if one has been fortunate enough to select exactly equivalent doses. When $\bar{y}_S$ is not equal to $\bar{y}_U$, a correction must be introduced for this inequality. It will be recognized that $\bar{y}_S - \bar{y}_U$ represents the difference in *effect* whereas $M$ represents a difference in *log-dose*, so that to correct $M$ for the difference in effect it is necessary to convert $\bar{y}_S - \bar{y}_U$ into units of log-dose.

The form of this conversion depends upon whether the two samples or experimental conditions are qualitatively similar, so that their respective dosage-effect curves do not differ significantly in slope. If the samples differ qualitatively and the dosage-effect curves are not parallel, the log-ratio of potencies, $M$, will depend upon the probit at which the two curves are compared. In this case some given level of effect, such as five probits (50 per cent effect), must be selected for purposes of comparison. Such comparisons will have, as a consequence, only limited validity

However, when the two curves have substantially the same slope, so that two parallel lines can be drawn through the two sets of points as shown by the broken lines in Fig. 1, comparisons of potency are valid at all levels of effect. The significance of any difference in slope may be tested by computing $\chi^2$ for the difference, $\chi_b^2$, which is given by the formula

$$\chi_b^2 = \frac{(b_S - b_U)^2}{\frac{1}{[wx^2]_S} + \frac{1}{[wx^2]_U}} \tag{1}$$

when the observed probits agree with their respective curves within the sampling error. When $\chi_b^2$ exceeds 3.84, the dosage-effect curves for the two samples differ more than would be expected by chance and caution should be used in comparing them by $M$. If a lower value of $\chi_b^2$ is obtained, it usually may be assumed that the samples are qualitatively similar and that the two parallel lines will represent a better estimate of the true slope than either line considered alone. This latter condition has been found in the great majority of the tests in this laboratory and is the only one considered below.

The common or combined slope, $b_c$, of the two parallel lines can be derived from terms already calculated as

$$b_c = \frac{S[wxy]}{S[wx^2]} \tag{2}$$

Since $b_c$, by definition, measures the change in probit effect associated with a unit change in log-dose, it follows that its reciprocal, $1/b_c = s_c$, gives the change in log-dose corresponding to one probit. This value, which has been termed $\lambda$ by Gaddum (2), is the population standard deviation of the just effective log-dose and may be used in converting $\bar{y}_S - \bar{y}_U$ into log-dose units in arriving at the general form for $M$ as

$$M = \bar{x}_S - \bar{x}_U - s_c(\bar{y}_S - \bar{y}_U) \tag{3}$$

where the sign of $M$ is negative for samples less potent than the standard and positive for samples stronger than the standard.

One should not be satisfied with a mere statement of the most probable value of the potency of a sample but should require that its precision be known as well, so that the reliability of the assay can be judged. This information is furnished by computing the standard error of $M$, $s_M$, which is given with sufficient accuracy by the formula

$$s_M = s_c\sqrt{\frac{1}{S(w)_S} + \frac{1}{S(w)_U} + \frac{(\bar{y}_S - \bar{y}_U)^2 s_c}{S[wxy]}} \tag{4}$$

Like $M$, $s_M$ is in logarithmic units of dosage so that the quantities $M + s_M$ and $M - s_M$ give the range within which $M$ may be expected to fall in two out of three assays repeated under similar conditions. Few bioassayists are willing to be wrong as often as one-third of the time so that they usually adopt the more conservative range of $M \pm 2s_M$ which reduces the odds to less than one in twenty. It is worthy of note that the still more stringent limits of one in one hundred have been adopted for biological assays by the British Pharmacopœia (1936 Addendum) as indicated by the following quotation:

"ERRORS OF BIOLOGICAL ASSAYS. In expressing the limits of error of biological assays the term 'limits of error ($P = 0.99$)' is used. The statements of the errors of these assays are based on the convention that, for practical purposes, a probability of 0.99 is equivalent to certainty. In other words, it has been estimated that the result of the assay will be within the stated limits 99 times out of every 100 times that the assay is made. These limits are given as percentages of the true result. Thus, the statement 'limits of error ($P = 0.99$) 95 and 105 per cent.' means that it has been estimated that in 99 assays out of 100 the result will be greater than 95 per cent., and less than 105 per cent., of the true result.

"If the error of the test, or its logarithm, is normally distributed, the stated limits of error correspond to the range covered by $\pm 2.576$ times the standard deviation."

Although in treating the data by this method all computations and statistical tests of significance are made in terms of logarithms, these are transformed to original units for final statements of relative potency. The most probable estimate of the relative potency of two samples or of a single sample when assayed by two different techniques is given by the antilogarithm of $M$,

which is multiplied by 100 if percentage potency is preferred. At odds of twenty-one in twenty-two this ratio is established within the limits given by the antilogarithm of $M + 2s_M$ and of $M - 2s_M$, or if the more stringent limits of ninety-nine in one hundred adopted by the British Pharmacopœia are preferred, $s_M$ is multiplied by 2.576 instead of by 2. From the very nature of the logarithmic transformation, the upper limit in original units will differ more from the most probable value than will the lower limit. If the error is relatively small, the inequality may not be large and an approximate average error may be more useful as a descriptive term. As shown by Cochran (10), this is given by the formula

$$s.\ e.\ \text{of relative potency} = 2.3026\ s_M\ (\text{antilog. } M) \qquad (5)$$

The computation of the log-ratio of potencies and its error may be illustrated by substituting in the above equations the parameters of the numerical example listed in Table II. To determine whether the two dosage-effect curves differ significantly in slope, equation (1) is used to compute

$$\chi_b^2 = \frac{(5.902 - 7.764)^2}{\frac{1}{0.1527} + \frac{1}{0.1443}} = 0.256.$$

Since $\chi_b^2$ is considerably less than 3.84, the two curves may be considered as parallel within the sampling error and their slopes may be combined by equation (2) to obtain

$$b_c = \frac{0.9014 + 1.1206}{0.15274 + 0.14434} = 6.8064 \text{ and } s_e = 0.14692.$$

This combined slope has been used in drawing the parallel, broken lines in Fig. 1, their positions being fixed by the means $\bar{x}_S$, $\bar{y}_S$ and $\bar{x}_U$, $\bar{y}_U$. The log-ratio of potencies may now be computed from equation (3) as

$$M = 0.86334 - 0.56303 - 0.14692\ (4.59669 - 4.72694) = 0.31945$$

in which, as in Table II, injection into the lymph sac has been considered as the "standard" and into the thigh muscles as the "unknown." Graphically, $M$ may be represented as the horizontal distance between the two parallel lines in Fig. 1.

The standard error of the log-ratio of potencies is given by equation (4) as

$$s_M = 0.14692 \sqrt{\frac{1}{24.52} + \frac{1}{23.95} + \frac{(4.59669 - 4.72694)^2\ 0.14692}{2.0220}} = \pm 0.04253.$$

If, under these conditions, digitalis had the same potency by both routes of injection, $M$ would not differ significantly from 0 (which corresponds to 100 per cent) *i. e.*, $M$ would be less than twice its standard error. In the present case, however, the ratio of $M/s_M$ exceeds 2 considerably so that there is much less than one chance in twenty that the potency was identical under both experimental procedures. In original units, the ratio of the intramuscular potency to the lymph sac potency was equal to the antilogarithm of 0.31945 or to 2.087, *i. e.*, to 208.7 per cent. The approximate standard error of the result, which corresponds to odds of two in three, was determined by means of equation (5) as $\pm 2.3026\ (0.04253)\ (208.7) = \pm 20.4$ per cent. At odds corresponding to twice the standard error the observation was established within the limits of 171.5 and 253.8 per cent. Despite the statistical significance of the result of this single comparison, it was desirable to reduce the error to narrower limits by replication (see Table IV). The general aspects of reducing the standard error are discussed in the next section.

### METHODS FOR INCREASING THE PRECISION OF $M$.

The final objective of a bioassay is to determine $M$ with a known level of precision as indicated by $s_M$. It is often impossible to obtain the desired precision with a single assay, and even when it is possible, one might prefer to continue an experiment over several independent component tests to broaden the basis for his conclusions. Methods for increasing precision, therefore, are of several types. One is to attain the smallest error in each individual assay by following an efficient design. Another is to draw upon the relevant experience of other assays through the medium of a standard curve. A third and most important method is to combine the evidence

from several assays to obtain a mean $M$ with a reduced error. These and related problems may be considered next.

*The reduction of $s_M$ by the Design of the Individual Assay.*—Since the result of an assay is measured in terms of $M$, it follows that the smaller the standard error of $M$, $s_M$, the greater is its reliability. A study of the various components in the equation for $s_M$ indicates how to plan assays so as to minimize this error. Two possibilities are open to the experimenter. The first and more obvious technique is to increase the number of animals used in a given test. Since $w$, the weight assigned to each observation, is proportional to the number of animals upon which it is based and since $w$ occurs in the denominator of each fraction in equation (4), $s_M$ decreases proportionately to the increase in the square root of the number of individuals. Thus the addition of ten frogs on each dose will effect a proportionately greater reduction in $s_M$ if the original number is ten than if it is twenty, and soon a limit is reached where other methods of increasing precision become more profitable. In general, it is better not to enlarge a single test beyond the point where the frogs can be handled easily but rather to repeat the assay several times independently.

The second possibility is to plan the experiment so as to make the most efficient use of each animal, a problem which may be approached from several angles. The information in a given probit effect and hence its weight, $w$, depends not only upon the number of animals but also upon a weighting coefficient determined from the probit expected at each dose. Since the weighting coefficient is largest at an expected probit of five (50 per cent effect) and diminishes both above and below this point, $S(w)_S$ and $S(w)_U$ decrease as the dosage interval is enlarged. The use of too small a dosage interval, on the other hand, is objectionable because it may so reduce $S[wxy]$ that the third term under the square root contributes unduly to the error. A narrower interval also reduces the reliability of $s_c$. With either two or three doses in a dosage-effect curve, these two opposing tendencies are balanced most satisfactorily when the expected effects of the high and low doses are from 0.9 to 1.0 probit above and below 5. Assuming that it is somewhat impractical to use more than three doses of each preparation in a given assay, the only choice lies between a two or three-dose technique if individual dosage-effect curves are desired. With forty frogs available for each preparation, an assay in which twenty frogs are used on each of two doses may be expected to have an average $s_M$ slightly larger than if the forty frogs were distributed between three doses with fifteen, ten and fifteen frogs on the low, intermediate and high doses, respectively. The three-dose arrangement would be preferred, therefore, quite apart from other advantages discussed later.

The precision of an assay as measured by $s_M$ is also dependent upon the difference in the mean probits for the two curves, $\bar{y}_S - \bar{y}_U$. When the potency of the unknown has been estimated from preliminary experiments and the test planned so that $\bar{y}_U$ differs from $\bar{y}_S$ only by chance variation in sampling, the component in equation (4) containing $(\bar{y}_S - \bar{y}_U)$ frequently averages less than 2 per cent of the total of the three terms beneath the radical or square root sign. But if the two samples differ significantly in the effects they produce, this component can contribute heavily to the error. For this reason alone, an assay in which there is a wide gap between the assumed and the observed potency of the unknown is never as precise as when the experiment confirms the assumption. When an assay is repeated, therefore, the assumption for each successive trial should be the combined value of $M$ from all earlier tests so as to reduce $\bar{y}_S - \bar{y}_U$ to within the sampling error.

A third factor in $s_M$ is the reciprocal of the slope, $s_c$. Since $s_M$ is directly proportional to $s_c$, the steeper the slope the more precise the assay. If dosages are selected which give the most efficient levels of effect as described above, the terms beneath the radical sign are independent of the slope. Then the steeper the slope the smaller is the standard error of the assay even though the reliability of the estimate of the slope diminishes. When an experiment is planned specifically for a comparison of slopes, the most efficient procedure is to use only two dosage levels, spaced so that they give 7–10 and 90–93 per cent of positive effects.

*The Question of a Standard Curve.*—It will be noted that the above design does not depend upon a standard curve but that the slope of such a curve is determined independently in each assay. If forty frogs are used on the standard and forty on the unknown—equal numbers on each preparation being the most efficient allocation of a given amount of experimental material—the combined slope has a standard error not less than 20 per cent of its observed value. It may be contended that a standard curve, such as is used by many bioassayists, avoids this error since

it is based upon a large number of animals and in consequence has a high reliability. This would be true, however, only if it were known that the predetermined slope of the standard curve were applicable to the assay in question, a fact which can be determined with certainty only by obtaining the slope under the conditions of the assay and comparing it with that of the standard curve. The slope of the dosage-effect curve may vary between laboratories and with the season. Until considerably more is known about the variations in the dosage-effect relation, assays should be planned so that they are self-contained and not dependent upon an assumed slope.

The practice of employing more than one dose in each assay has the advantage, moreover, that the data on the standard from separate assays can always be combined, either currently or subsequently, if statistical test shows that the individual determinations are, in fact, parallel. In this way one can arrive at an improved estimate of the slope of the characteristic dosage-effect curve for the standard and if experience demonstrates its advisability, a "standard curve" can be built up eventually which will have a much broader basis than if established as the result of an isolated series of tests. The combined slope of such a curve may be computed from equation (2) which may be used to combine the data from any number of assays having parallel dosage-effect curves. To determine whether they are sufficiently parallel to justify combination the following chi-square test may be applied:

$$\chi_b^2 = S\{[wx^2]_1 (b_1 - b_c)^2\} \tag{6}$$

where subscript "1" refers to each individual result. Hence, whenever a "standard curve" would be valid, this procedure is always available for increasing the precision of $M$ and $s_M$.

*Combining the Results of Replicated Assays.*—Although no single assay may suffice, the log-ratio of potencies can be determined to any reasonable level of precision by sufficient replication. The results of independent assays seldom have identical standard errors and therefore are not of equal value, so that in combining such results each should be given a weight proportional to the information it contains. The information in a parameter has been shown by Fisher to equal the reciprocal of its variance, a principle which has been used in weighting the individual observations when computing the dosage-effect curve for any given assay. In the present case the information in each assay as a whole is equal to $1/(s_M)^2$, which is used as a weight in computing the weighted mean as

$$\overline{M} = \frac{S(w_M M)}{S(w_M)} \tag{7}$$

where $w_M = 1/(s_M)^2$. This equation is identical in principle and in form with those for computing the weighted mean log-dose ($\bar{x}$) or weighted mean probit ($\bar{y}$) (Table II).

All of the individual $M$ values combined in a mean log-ratio of potencies should agree with one another within their sampling errors, especially when assaying an unknown sample of digitalis. $\chi^2$ can be used to test whether the separate estimates of $M$ are mutually consistent by an equation analogous to that used for $\chi^2$ in Table II, namely,

$$\chi_M^2 = S(w_M M^2) - \overline{M}S(w_M M) \tag{8}$$

which will have one less degree of freedom than the number of independent values of $M$ entering into the equation. If all determinations are homogeneous within the limits of $\chi_M^2$ or an equivalent test, the standard error of the mean $\overline{M}$ is given by the equation

$$s_{\overline{M}} = \sqrt{\frac{1}{S(w_M)}} \tag{9}$$

But if $\chi_M^2$ indicates a significant degree of heterogeneity, as sometimes occurs, the standard error given in equation (9) must be increased to include the variation between the component $M$'s, so that then

$$s_{\overline{M}} = \sqrt{\frac{\chi_M^2}{(N-1)S(w_M)}}, \tag{10}$$

where $N$ is the number of individual $M$'s entering into the mean.

CRITERIA FOR SELECTING AN ASSAY TECHNIQUE.

Having in mind the quantitative basis for the all-or-none assay, we may now consider which criteria show the greatest promise for discriminating between alternate techniques. Of the several parameters described above, four suffice to characterize the assay procedure.

The first of these is the chi-square test given as the last item in Table II. It is used here to determine whether the points on the plot of probit *vs.* log-dose diverge so far from linearity that it is inadvisable to fit them with a straight line. Since it is possible to draw only one straight line through two points, at least three doses must be used before there can be any divergence which is measurable by $\chi^2$ and the fact that otherwise $\chi^2$ cannot be computed is another reason for distributing the number of available frogs over three or more doses. In cases where $\chi^2$ indicates a divergence from linearity, the whole experimental procedure should be scrutinized searchingly for possible causes of heterogeneity, such as a change in the source and storage conditions of the frogs, failure to assign the frogs at random to the different dosages, unequal exposure to some influence during the assay, etc. Thus $\chi^2$ is a measure of the homogeneity of the frogs between dosages and of the experimental conditions. A procedure for which the chi-square test continues to indicate heterogeneity despite all possible precautions should be viewed unfavorably in comparison with alternatives which yield more homogeneous data under similar conditions. Thus the first criterion for a good assay procedure is a value for $\chi^2$ which is consistently within the error of sampling.

The second parameter characterizing any assay is $b$, the slope of the dosage-effect curve. As a measure of the change in effect per unit change in log-dose, it is evident that any assay technique yielding a high value for $b$ deserves consideration. Burn (11) has also emphasized this point. The fact that different laboratories have reported different slopes shows that the slope can be increased experimentally and any such change in technique would lead to an immediate improvement in precision. In searching for a procedure giving a steep dosage-effect relation, the constancy of a given slope from one assay to the next must not be overlooked. This can be checked by computing $\chi_b^2$ as given by equation (1) or (6) and by relating the variation in $b$ to other factors in the experiment. The second criterion suggested for a good assay technique is a consistent and relatively high value for $b$.

The third characteristic by which procedure is to be judged is the magnitude of $s_M$, the standard error of the log-ratio of potencies for an individual assay as given by equation (4). Although partly a corollary of the second criterion, the most precise method is clearly that which yields the lowest value of $s_M$. For this factor to have its full diagnostic value, assays should be designed with the precautions discussed in an earlier section on the reduction of $s_M$. A consistently low value of $s_M$ is then a suitable third criterion for a good assay technique.

The fourth requirement is that the procedure shall give reproducible results. When an assay is replicated, the successive values of $M$ should agree with one another within the sampling error, as measured by $\chi_M^2$ computed from equation (8) or its equivalent. In principle this is by no means a new criterion since it has served as the acid test for recommending many procedures in the past. One of the more important advantages of the present quantitative approach is that it provides unequivocal, objective standards for judging whether or not successive assays are in agreement.

In proposing these criteria, it is anticipated that they will find their greatest usefulness in evaluating the results of research. No doubt many will find that the increased objectivity and efficiency attending these modes of analysis will more than repay the time invested in designing their experiments and in computing their results. In no sense can the criteria be substituted for accurate and well-planned experimentation, without which they are of little use. Nor can they be expected to resolve unaided all difficulties in selecting the best method for the bioassay of digitalis with frogs, if only because of two important elements which are not weighed by the proposed criteria, namely, practicability and experimental objectivity. It goes without saying that these factors will be conclusive in choosing between two procedures which appear equally advantageous in the light of the tests proposed above.

COMPARISONS OF PROCEDURE BASED UPON THE PROPOSED CRITERIA.

Of the several modifications of procedure which are being investigated and weighed in the light of the proposed criteria, two have reached a stage where a report may be justified.

*Experimental Procedure.*—All of the tests followed the same general plan. As a rule, two factors involving four dosage-effect curves were compared in each experiment, all other factors being kept as uniform as possible. Three dosage levels were used in each experiment with fifteen to twenty frogs on each dose.

The frogs came from two sources. In tests conducted prior to May 1939, they were from Vermont, having been stored there on a lake bottom since November 1938, and shipped to the laboratory as needed in four-gross lots. Later tests were made on frogs shipped in similar lots from a dealer in Wisconsin. Upon arrival, the frogs were segregated as to sex and in some cases the males were divided as to weight into light and heavy groups, prior to storage in running water at a temperature below 15° C. It was not possible, because of limited facilities, to store the frogs for a uniform period prior to transfer to the assay tank on the afternoon of the day before the test. The frogs were kept singly as far as possible, especially during over night assays.

The preparation of tinctures from standard digitalis powders followed the U. S. P. XI method uniformly using 1 Gm. of powder for each 10 cc. of menstruum in cases where direct comparisons of two or more powders were involved. The mixture of powder and menstruum was shaken mechanically and intermittently for twenty minutes in every hour during the 24-hour maceration period.

In the preparation and injection of the test dilutions of the tinctures, every effort was made to obtain uniformity with respect to the relative volume of the fluid injected and its alcoholic content, unless these factors were varied experimentally. Except in special cases, the test dilutions of digitalis contained about 23 per cent alcohol. This represents about the maximum content that allows a margin of safety below the U. S. P. XI limit of 25 per cent. Furthermore, it can be obtained with a minimum of manipulation, since it is the average alcoholic content resulting from a 1:3 dilution with water of U. S. P. XI Tincture of Digitalis (67 to 72 per cent by volume of $C_2H_5OH$).[1]

Not infrequently it was necessary to dilute less than 1:3 since the volume of the injected fluid per Gm. of frog was kept the same for all three doses of each test. In such cases some of the alcohol was removed by evaporation in order to keep the final alcoholic content at 23 per cent. Data have been taken in this laboratory which make it possible to predict the amount of evaporation under an air stream needed to reduce the amount of alcohol sufficiently to yield a test dilution containing almost any desired alcoholic content. It is probably not generally recognized how difficult is the matter of removing all or even a large part of the alcohol from a tincture residue. For example, in the test of June 19 (Table III) it was desired to use a test dilution containing 5 per cent alcohol in a volume of 0.01 cc. per Gm. of frog. To fulfil these requirements for the highest dose needed in the one-hour method (0.0063 cc. per Gm. of U. S. P. XI Standard Preparation of Digitalis) it was necessary to evaporate off 70 per cent of the original weight of a measured sample of the tincture. The alcoholic content actually attained, as determined on a subdivision, was 4.7 per cent.

In each experiment a separate dilution was prepared for each dosage level, using an appropriate amount of tincture and adding or removing alcohol as required so that the final test dilutions contained the desired doses in either 0.01 or 0.02 cc. of fluids with the desired alcoholic contents. With a view toward convenience in the computations, the doses selected were such that their logarithms were round numbers, uniformly spaced. The interval between doses for the one-hour method was usually 0.10 log and for the over night method, 0.05 log. All the doses used in any experiment were assigned at random to the frogs (in groups of five) by shaking numbered tags, so that all groups had an equal chance of receiving each dose of each test preparation.

In reading the results of the one- and four-hour test, a rigid interpretation of the U. S. P. XI definition was adopted, namely, ventricle in systolic standstill upon exposure, or going into standstill upon gentle mechanical stimulation; auricles widely dilated and *in standstill.* In reading the results of the over night or lethal dose method, there was virtually no room for the exercise of personal judgment. Rarely does one encounter a frog which is not unquestionably dead or alive 16 to 24 hours after injection.

[1] In this connection it should be noted that the official Standard Preparation of Digitalis (U. S. P. XI, p. 397) when prepared as directed using a menstruum of 4 parts by volume of alcohol and 1 part of distilled water, contains about 74 per cent alcohol and in this respect differs from U. S. P. XI Tincture of Digitalis.

*The Influence of Alcohol on the Apparent Potency.*—In view of the results reported by Rowe (12) considerable attention was devoted to the possible influence of the alcoholic content of the injection fluid upon the apparent potency of the same sample of digitalis. Toward this end, eight comparisons were made between test dilutions of digitalis containing approximately 5 and 23 per cent alcohol, respectively, each comparison involving the use of ninety to one hundred twenty frogs. As may be noted from the summary of the data in Table III, three standard digitalis powders were used as test material in these comparisons under the wide variety of conditions indicated in columns 2, 3 and 5.

TABLE III.—INFLUENCE OF ALCOHOL CONTENT OF TEST DILUTION ON POTENCY.

| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|
| Date 1939. | Route. | Time. | Material. | Volume. | $M \pm s_M$. | $(b_5 - b_{23}) \pm$ s. e. diff. | $b_c \pm$ s. e. $b_c$. | $\chi^2$ 23 Per Cent. | $\chi^2$ 5 Per Cent. |
| 4/18 | *L* | 1h | U. S. P. | .02 | −.0151 ± .0668 | 1.739 ± 2.941 | 3.996 ± 1.468 | .692 | 1.225 |
| 4/18 | *L* | 1h | 1936 | .02 | −.0081 ± .0471 | 4.264 ± 3.076 | 5.514 ± 1.534 | .069 | .000 |
| 4/20 | *L* | 1h | U. S. P. | .02 | .0108 ± .0447 | 2.131 ± 3.504 | 6.236 ± 1.750 | 1.650 | .124 |
| 4/20 | *L* | 1h | 1926 | .02 | −.0350 ± .0536 | 3.951 ± 3.449 | 5.218 ± 1.719 | .727 | .120 |
| 6/19 | *M* | ON | U. S. P. | .01 | .0327 ± .0170 | −9.039 ± 8.097 | 18.073 ± 3.946 | 1.672 | .220 |
| 6/19 | *M* | 1h | U. S. P. | .01 | −.0174 ± .0423 | −3.912 ± 3.184 | 6.217 ± 1.573 | 1.279 | .589 |
| 6/24 | *M* | ON | U. S. P. | .01 | .0244 ± .0254 | − .050 ± 6.654 | 10.824 ± 3.321 | 1.811 | .000 |
| 6/24 | *M*[1] | ON | U. S. P. | .01 | −.0085 ± .0230 | 1.618 ± 6.281 | 10.966 ± 3.140 | 1.309 | 1.994 |
| Weighted Mean | $\bar{M}$ = | | | | +.0131 ± .0107 | 1.227 ± 1.350 | | | |

Conclude: Potency with 5 per cent alcohol = 103.1 ± 2.5 per cent of that with 23 per cent alcohol.
In determining $M$ and the difference in respective slopes, the preparation of 23 per cent alcohol content was taken as the standard. Thus, in column 7, $(b_5 - b_{23})$ indicates the difference in the slopes obtained, respectively, by the use of 5 and 23 per cent alcohol contents. Also tabulated is the standard of errors of the difference in slopes.

[1] Female frogs used; all others were male frogs.

The value of $M$ and $s_M$ for each comparison is listed in column 6. In only one case (6/19, over night method) did $M$ exceed its standard error, $s_M$, and then not significantly, so that it was unnecessary to compute $\chi_M{}^2$ to test the agreement of values of $M$ prior to combining them. The weighted mean of the series, $\bar{M}$, and its standard error were 0.0131 ± 0.0107. In terms of percentages, this result indicates that the same digitalis sample when injected with 5 per cent alcohol exhibited a potency 103.0 ± 2.54 per cent of that exhibited when injected with 23 per cent alcohol. Thus under the conditions of our experiments, a fourfold change in the alcoholic concentration did not alter the outcome of the assay significantly. The experimental error of this series was sufficiently low, as indicated by its standard error, that an $\bar{M}$ corresponding to less than 93.8 or more than 106.6 per cent would have pointed to a highly significant influence ($P = 0.99$) of the alcoholic content of the test dilutions upon the estimated relative potency. No one test had a precision within these limits but because the individual values from assays on more than 720 frogs were all in agreement, they could be combined so that the error of the final estimate was reduced to one-third of that observed in some of the individual assays.

The influence of variation in alcoholic content upon the slope of the dosage-effect curve may be judged from the data listed in column 7, which show the difference in slope and the standard error of the difference for each comparison. In no experiment was the difference significantly greater than its standard error. The trend of the data is in the direction of a steeper dosage-effect curve with the use of 5 per cent alcohol, particularly for lymph sac injections, and with sufficient replication of suitably designed experiments, this trend might prove to be significant. Any advantage found, however, would have to be considerable to outweigh the inconvenience of evaporating off the alcohol to such a low concentration.

Column 8 lists for each comparison the value of $b_c$ and its standard error, the significance of which will be discussed in the following section.

In columns 9 and 10 are listed the respective values of $\chi^2$ for the two percentages of alcohol tested. Although none of the values indicates any noteworthy degree of heterogeneity, so that both techniques would be considered equally valid, it is evident that the $\chi^2$ is generally lower in the series with the 5 per cent alcohol. In two experiments of this series, the $\chi^2$ was zero since all three of the observed points fell exactly on a straight line. Any further significance of these comparative values of $\chi^2$ cannot be evaluated at present but remains a subject for further investigation.

*The Influence of Route of Administration upon Absorption.*—As indicated in the numerical example given above, two routes of administration of digitalis have been tested in our survey of the

TABLE IV.—COMPARISON OF INTRAMUSCULAR AND LYMPH SAC INJECTION.

| (1) Date 1939. | (2) Time. | (3) Volume Cc./Gm. | (4) Material. | (5) $M \pm sM$[4]. | (6) $(b_m - b_l) \pm$ s. e. diff. | (7) $b_c \pm$ s. e. $b_c$. | (8) $\chi^2$ Muscle | (9) $\chi^2$ Lymph |
|---|---|---|---|---|---|---|---|---|
| 5/24 | 1h | [1] | U. S. P. | .3194 ± .0425 | 1.862 ± 3.671 | 6.806 ± 1.835 | .045 | .038 |
| 6/30 | 1h | .02 | U. S. P. | .2421 ± .0482 | .694 ± 3.523 | 5.975 ± 1.761 | .087 | .597 |
| 7/11 | 1h | .02 | U. S. P. | .3207 ± .0463 | − .226 ± 3.526 | 6.431 ± 1.763 | .089 | .676 |
| Weighted Mean | 1h | | U. S. P. | .2969 ± .0263 | .748 ± 2.062 | | | |
| 5/25 | 1h | [1] | 1936 | .0775 ± .0533 | −1.729 ± 3.165 | 5.188 ± 1.549 | .112 | .043 |
| 7/11 | 1h | .02 | 1936 | .2387 ± .0478 | 1.048 ± 3.455 | 5.779 ± 1.727 | .087 | .001 |
| Weighted Mean | 1h | | 1936 | .1668 ± .0356 | − .462 ± 2.333 | | | |
| 6/12 | ON[2] | .01 | U. S. P. | .0486 ± .0214 | 3.761 ± 6.848 | 12.977 ± 3.422 | .393 | .009 |
| 6/12 | ON[3] | .01 | U. S. P. | .0108 ± .0157 | −3.880 ± 6.411 | 16.870 ± 3.199 | 1.538 | .897 |
| 6/22 | ON | .01 | U. S. P. | .0261 ± .0160 | 1.029 ± 6.835 | 16.663 ± 3.412 | 2.936 | .465 |
| 6/22 | ON | .02 | U. S. P. | .0725 ± .0356 | 5.007 ± 6.389 | 9.366 ± 3.194 | .680 | .008 |
| 6/30 | ON | .02 | U. S. P. | .0107 ± .0424 | −6.522 ± 7.172 | 6.963 ± 3.574 | .239 | .262 |
| Weighted Mean | ON | | U. S. P. | .0274 ± .0093 | .034 ± 3.000 | | | |

Conclude: U. S. P. (one-hour): Potency of Intramuscular = 198.1 ± 12.0 per cent of lymph sac.
International 1936 (one-hour): " " " = 146.8 ± 12.0 " " " " "
U. S. P. (over night): " " " = 106.5 ± 2.3 " " " " "

Male frogs were used on all experiments except on 6/30/39 when all females were used. Alcohol content of injected material approximately 23% in all experiments.

---

[1] Volume of injected material was confounded with route of injection, *i.e.*, 0.01 cc./Gm. given intramuscularly and 0.02 cc./Gm. given by lymph sac.

[2] Light-weight frogs used (average 19.4 Gm.).

[3] Heavy frogs used (average 27.2 Gm.).

[4] In determining $M$ and the difference in respective slopes, results of lymph sac injection were taken as standard. Thus, in column 6 ($b_m - b_l$) indicates the difference in the slopes obtained, respectively, by the intramuscular and lymph sac injections; also tabulated is the standard error of the difference in slopes.

factors which may influence assay results with frogs. The report of Dooley and Higley (6) suggested that the intramuscular route was worthy of an early trial. In Table IV the data are summarized from ten comparisons of intramuscular with lymph sac injection. These were made with both the one-hour and the over night methods (column 2) and two comparisons were made with the four-hour method, these data not being included because of their incompleteness. Both the U. S. P. XI Reference Digitalis Powder and the 1936 International Standard Powder were used (column 4). The volume of the injected fluid is recorded in column 3 and, as noted, volume was confounded with route on May 24 and 25. It is clear from the data in column 5 that the influence of route on the potency was determined largely by the length of time the digitalis was allowed to act. The two powders also seemed to differ significantly in this respect, so that the data obtained by the one-hour method have been combined separately for each powder.

From the combined results of three assays by the one-hour method, it is seen that the U. S. P. XI Reference Digitalis Powder had twice as great a potency when injected intramuscularly as when injected into the lymph sac. It may be inferred, therefore, that in one hour not more than one-half of the active glucosides of the powder were absorbed effectively from the lymph sac as compared with the amount absorbed following intramuscular injection. Similarly it may be inferred that only about two-thirds of the activity of the 1936 International Standard Powder was absorbed effectively from the lymph sac in one hour. The limited data available indicate that the observed difference in the two standards is probably significant. Further work is now in progress to establish this point.

The results of the five assays on the U. S. P. XI Reference Digitalis Powder by the over night method agreed among themselves within the sampling error ($\chi_M^2 = 3.87$ with 4 degrees of freedom). Their combined value, $\overline{M} = 0.0274 \pm 0.0093$, demonstrates that more complete absorption was apparently obtained following intramuscular injection even with the longer period of observation. Although the difference in potency was small, 106.5 ± 2.3 per cent, the ratio of $\overline{M}$ to its standard error was 0.0274/0.0093 = 2.95. Thus these tests on about 500 frogs show that by the over night method the potency of this digitalis powder was slightly but significantly ($P = 0.99$) greater by the intramuscular route than *via* the lymph sac.

The differences in slope and their standard errors are listed in column 6. From this point of view it is obvious that both routes of administration were equal. The standard errors of the differences in slope were much larger with the over night technique than with the one-hour method, due to the use of smaller interval between doses for the over night method necessitated by its steeper slope (0.05 log as against 0.1 log for the one-hour). Column 7 lists the combined slope, $b_c$, and its standard error for each assay. Examination of these data and those of column 8, Table III, reveals that the dosage-effect curve was much steeper for the over night method than for the one-hour assay. This finding is not in agreement with the result reported by Edmunds, Moyer and Shaw (13), who concluded from experiments with the lymph sac method of injection "that regardless of the period of observation (one-hour or lethal dose) . . . the characteristic curve for digitalis has the same slope." Without the data underlying the above quotation, this discrepancy cannot be examined further, but in the experiments at this laboratory the steeper slope of the over night curve is the outstanding feature of data on comparisons of the methods embodying the two periods of observation. It is this steeper slope that accounts for the greater precision of the estimate of potency obtained by the over night method, as shown by $s_M$ in Table IV. Thus application of the second criterion proposed above reveals a decided advantage in favor of the over night method.

Scrutiny of the respective chi-square values for the two routes of injection reveals no significant difference. With one exception (over night assay of 6/22 with 0.01 cc. volume) the chi-squares did not approach a value which would indicate a significant degree of heterogeneity.

## SUMMARY.

1. Further improvement in the digitalis assay on frogs can be facilitated by objective criteria for testing the effectiveness of any proposed modification of procedure. Due to the nature of the assay, such criteria are necessarily statistical in character and in a well-designed experiment can be computed from the data of a self-contained assay. By transformation of dosages to logarithms and of percentage

effect to probits, the sigmoid dosage-response curves for the two materials or procedures involved in the test are fitted by parallel straight lines. Then the horizontal distance between them measures with a calculable precision the log-ratio of their potencies, $M \pm s_M$. The computation of these and related parameters is illustrated with a numerical example.

2. Methods are described for increasing the precision of $M$ by an efficient design of the individual assay, by utilizing past experience relative to the slope of the standard curve where applicable and by combining the results of replicated assays to obtain a more precise weighted mean log-ratio of potencies.

3. On the above statistical basis, four criteria are proposed for evaluating an assay procedure: (*a*) a value of chi-square which indicates that the several groups of frogs used in the test were homogeneous and comparable, (*b*) a consistent and relatively steep slope of the parallel dosage-effect curves, (*c*) a consistently low value of the standard error of the log-ratio of potencies, $s_M$, and (*d*) agreement of replicated determinations of $M$ within the sampling error. In conjunction with a knowledge of the practicability and experimental objectivity of a procedure, these criteria should be conclusive in the development of an improved technique.

4. These criteria have been applied in a comparison of the effects produced by injection of the same sample of digitalis in test dilutions containing 5 and 23 per cent of alcohol. The apparent potencies of the U. S. P. XI Reference Digitalis Powder and of the International 1926 and 1936 Standard Digitalis Powders were not influenced by this fourfold change in the alcoholic content of the injected test dilutions, the same digitalis powder when injected with 5 per cent alcohol exhibiting a potency 103.0 ± 2.54 per cent of that exhibited when injected with 23 per cent alcohol.

5. The relative effectiveness of digitalis injected intramuscularly has been compared with that injected in the lymph sac. One hour after injection the effective absorption of U. S. P. XI Reference Digitalis Powder from the lymph sac was about one-half that from the thigh muscles and even over night the potency of digitalis was slightly but significantly greater (106.5 ± 2.3 per cent) by the intramuscular route than *via* the lymph sac. The difference was less pronounced with the 1936 International Standard Powder. The over night assays showed a smaller standard error than the one-hour tests because of a consistently steeper dosage-effect curve.

REFERENCES.

(1) Trevan, J. W., *Proc. Roy. Soc., B*, 101, 483 (1927).
(2) Gaddum, J. H., *Med. Res. Council, Special Report Series*, No. 183 (1933).
(3) Bliss, C. I., *Science*, LXXIX, 409 (1934).
(4) Bliss, C. I., *Ann. Appl. Biol.*, 22, 307 (1935).
(5) Bliss, C. I., *Quart. J. Pharm. Pharmacol.*, 11, 192 (1938).
(6) Dooley, M. S., and Higley, C. D., JOUR. A. PH. A., 11, 911 (1922).
(7) Fisher, R. A., and Yates, M. A., "Statistical Tables for Biological, Agricultural & Medical Research," Oliver & Boyd, London (1938), Table IX.
(8) Fisher, R. A., and Yates, M. A., *Ibid.*, Table XI.
(9) Fisher, R. A., and Yates, M. A., *Ibid.*, Table IV.
(10) Cochran, W. G., *Ann. Appl. Biol.*, 25, 426 (1938).
(11) Burn, J. H., "Biological Standardization," Oxford Univ. Press (1937).
(12) Rowe, L. W., JOUR. A. PH. A., 27, 844 (1938).
(13) Edmunds, C. W., Moyer, C. A., and Shaw, J. R., *Ibid.*, 26, 290 (1937).