

Harry Spector

Reprinted from ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Vol. 52, Art. 6, Pages 889-902, March, 1950

STATISTICS IN MICROBIOLOGICAL ASSAY

By Lila F. Knudsen

Food and Drug Administration, Washington, D. C.

It is the purpose here to give a brief explanation of the logic and common sense that serve as a basis for the statistical methods applicable to results obtained from microbiological assays. In an attempt to present a simplified explanation, only a few formulas and illustrations are cited, and these to illustrate important classes of technics that are now in general use. Bacteriologists and biochemists are prone to avoid the use of computations involving complicated mathematical formulas. Fortunately, these are not necessary. Simpler formulas are often a satisfactory compromise and can be used to obtain both an estimate of potency from a set of assay data and some measure of the precision of that estimate. There are, of course, several different ways they may be expressed. The use of "short cut" statistical procedures often helps to reduce either the series of determinations required to give a desired precision or the number of assays necessary to give assurance as to acceptability of a given lot of material. Saving time by this means makes for greater efficiency and better utilization of an analyst's time in any analytical laboratory. Also, the use of shortened calculation procedures means that more time can be devoted to additional assays.

The applied statistician is continually beset with requests to simplify these calculations so that they may be easily applied by laboratory technicians. The original procedures rest on certain assumptions, and still others are required in their simplification. The latter assumptions are usually easier to test, and the validity of those underlying the simplifications presented here has been the subject of appropriate tests. Statistics can also be invaluable in giving an objective measure of the validity of an assay by means of testing the linearity of the dosage-response curve, where such linearity is assumed, or testing whether the slope of the dosage-response curve is significantly different from zero. However, no attempt will be made to discuss this phase of statistical methods. Further detailed procedures can be obtained from some of the references given at the end of this discussion.

In general, there are at least three statistical approaches commonly applied to the results of microbiological assay for calculating (1) an estimate of the potency, and (2) some measure of how much variation may be expected in a number of estimates of the potency of a given substance assayed in the same laboratory. These three approaches depend on the type of response in a particular assay: (1) assays having an undefined dosage-response relationship, such as those involving a daily standard curve relating dose and response; (2) assays, such as that for nicotinic acid, involving a linear relationship between dose and response and the straight lines for standard and unknown intersect at zero dose; and (3) assays, such as the penicillin plate assay, wherein there is a linear relationship between the response and the logarithm of the dose but with parallel straight lines for standard and

unknown. In all three types, estimates of the precision of the assay potencies may be determined, as well as estimates of the potency itself. The first type is to be avoided whenever possible, however, since it does not make efficient use of all the data in obtaining estimates of potency and precision.

In the illustrations given here, it is taken for granted that preliminary work has established that linear relationships exist where they are assumed to be, and that the unknown gives the same type of response as the standard.

An estimate of the precision of the method as applied to one laboratory's results—or how closely one laboratory can check its own results—is usually given by the standard error of the assay. This measure of precision cannot be said to hold from one laboratory to another, unless the method has been studied collaboratively, and it can be demonstrated that the one laboratory can check the other's results as closely as it can check its own. This is true of chemical and physical methods as well as those of microbiological and biological assays.

The first type of microbiological assay procedure, and one that is in common use, involves the standard curve. This type of assay gives an estimate of the potency and, as will be shown here, can be used to give a fair estimate of the standard error of the assay, even though it is a very inefficient use of the amount of information given by the assay. In some instances, however, such as the turbidimetric assay of streptomycin, the dosage-response curve of which is shown in FIGURE 1, it seems to be the only procedure that can be used, since no simple transformation has been found as yet that will make the dosage-response curve linear. The official method used by the Food and Drug Administration for turbidimetric assay of streptomycin will serve as an illustration. Since the official description can be found in the Federal Register for April 4, 1947, only the statistical part of the assay will be given here.

A solution containing a definite amount of the standard is prepared and labeled as containing "100 per cent of standard." Eight additional solutions are made to contain 60, 70, 80, 90, 110, 120, 130, and 140 per cent of standard. A solution of the unknown is prepared to contain an amount equivalent to 100 per cent of "standard," on the basis of its assumed potency. Six tubes are used for each level of the standard and for the one level of the unknown. After inoculation with the proper bacteria and incubation for the proper period, *etc.*, the light-transmission reading (labeled on FIGURE 1 as "turbidimetric response") is made on a photo-electric colorimeter. The colorimeter is adjusted so that the series of tubes containing 60 per cent of the standard will have a light transmission reading of about 10 and the series containing 140 per cent of the standard will have a reading of about 90. A record is made of the turbidimetric response for each tube, as shown in TABLE 1. These responses are then plotted on cross-section paper against the dilution as a percentage of "standard." Two calculations are made for the responses to each dilution: the average and the range (the latter is the difference between the highest and the lowest results on a single dilution). The standard curve, as shown in FIGURE 1, is drawn by connecting the averages with straight lines. The potency of the unknown is read from the "curve."

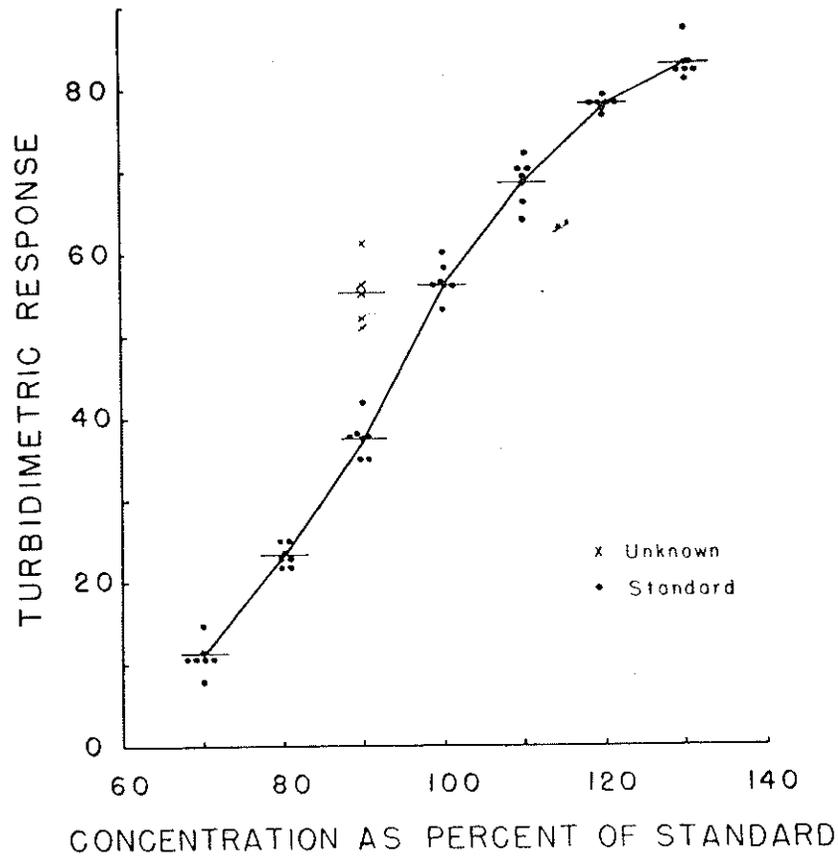


FIGURE 1. Standard curve for turbidimetric assay of streptomycin.

TABLE I

<i>% Standard</i>	<i>Turbidimetric response</i>						<i>Average response</i>	<i>Range</i>
140	90	90	90	90	90	90	90.	0
130	87	82	82	83	82	81	82.8	6
120	78	77	78	78	78	79	78.0	2
110	72	76	66	64	69	70	68.5	8
100	60	53	58	55	55	55	56.0	7
90	42	38	38	38	35	35	37.6	7
80	25	22	23	23	22	25	23.0	3
70	15	11	11	11	8	11	11.1	7
60	4	5	5	5	6	5	5.0	2
Unknown.....	61	51	55	52	56	56	55.1	10

For the unknown illustrated here (on the response scale only), the potency is 99.5 per cent of the standard. A notation is made of the two doses of

standard between which the response to the unknown falls. For purposes of calculation, these average responses are labeled S_H and S_L . The average response to the unknown is labeled U . Quite simple formulas may be developed for calculation of the potency and the standard error of the assay, for example,

$$\text{potency} = \bar{x} + \frac{5V}{W},$$

where \bar{x} is the dose halfway between the doses corresponding to S_H and S_L , $V = 2U - S_H - S_L$, and $W = S_H - S_L$. An approximation to the standard error of the assay can be calculated by the formula:

$$\text{standard error of the assay} = \frac{0.93R}{W} \sqrt{3 + \frac{V^2}{W^2}},$$

where R is the sum of the ranges ($R = R_{SH} + R_{SL} + R_U$). The maximum value of the quantity under the square root sign is 4, and the minimum is 3. Using the maximum value:

$$\text{standard error of the assay} = \frac{1.86R}{W}.$$

To illustrate how these formulas work, the data obtained in the previously cited assay may be substituted to obtain potency and standard error: $R = 7 + 7 + 10 = 24$; $W = 56.0 - 37.6 = 18.4$; and $V = 2(55.1) - 56.0 - 37.6 = 16.6$. Thus,

$$\text{potency} = 95.0 + \frac{5(16.6)}{18.4} = 99.5,$$

$$\text{and standard error of the assay} = \frac{1.86(24)}{18.4} = 2.4.$$

This formula has been found to give a very good estimate of how closely an estimated potency can be checked from one time to another (see Oswald and Knudsen).

The second type of microbiological assay yields a linear relationship between dose and response, and the dosage-response lines for standard and unknown intersect at zero dose. The potency is the ratio of the slopes of the two lines. This situation can be stated by the equation $Y = a + bX$, where the intercept a is the same for standard and unknown and the slope b differs. Finney and Wood have described this type of assay as applied to nicotinic acid. The simplest they describe is the 3-point assay with an equal number of observations on each point, as shown in FIGURE 2. Using the assumed potency of the unknown, its dose is adjusted so that it is equal to the dose of the standard. Since equal doses are assumed, the potency can be calculated by dividing the difference between the response to the unknown and the response at zero dose by the difference between the response

to the standard and the response to zero dose. The potency and standard error of the assay can be calculated as follows:

$$\text{potency} = B = \frac{U - S_0}{S_1 - S_0}$$

$$\text{and standard error of assay} = \frac{2R\sqrt{1 - B + B^2}}{D(S_1 - S_0)\sqrt{2n}}$$

where n = no. of observations for each average, \bar{R} = average of ranges of the three groups, and D = number of std. deviation units in the average

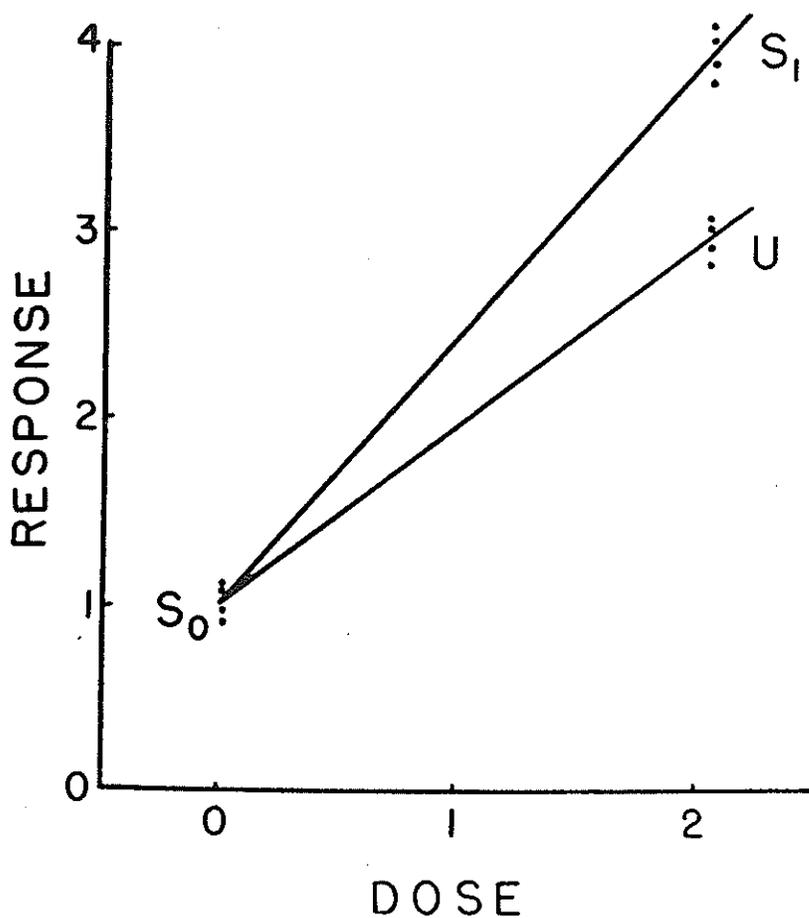


FIGURE 2. A three-point assay of nicotinic acid.

range ($D = 2.059$ for $n = 4$). These are slightly different from the equations given by Finney and Wood, in that the range divided by the appropriate figure D has been substituted for the standard deviation (this can be done with little loss of efficiency for values of n less than 10). For values of

n , other than $n = 4$, values for D can be found in several texts (such as Snedecor's *Statistical Methods*, Table 5.5) and in Karl Pearson's *Tables for Statisticians and Biometricians*, II: 165. The value of D can also be found from the order statistic-expectation table in Fisher and Yates's *Tables for Statisticians and Biometricians*, Table XX, where D equals twice the score for the first member of the sample.

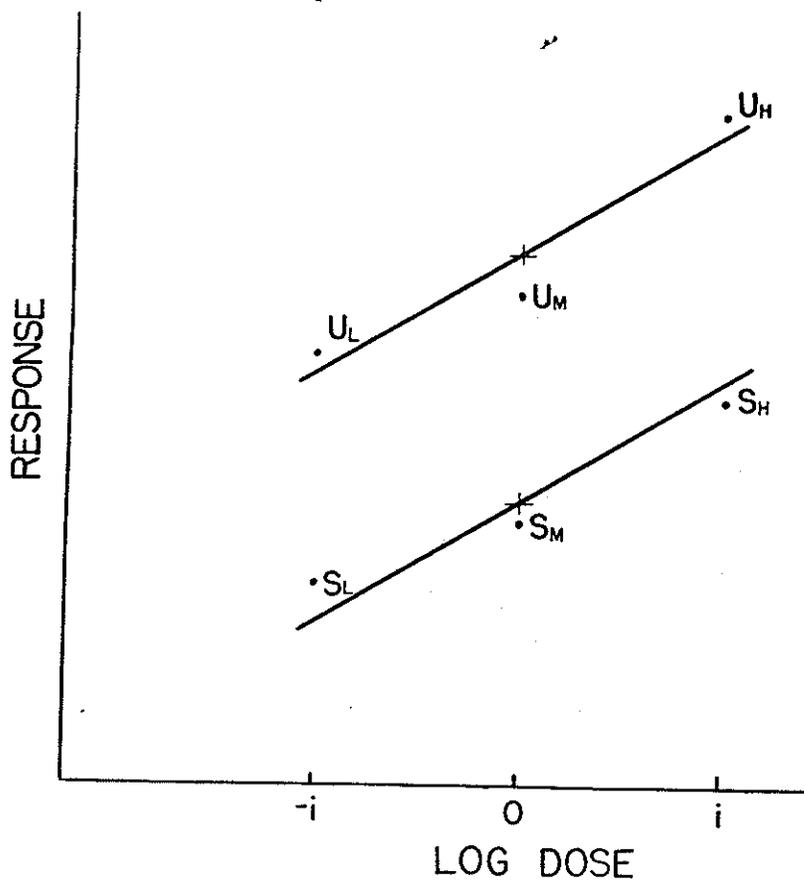


FIGURE 3. A three-dose penicillin assay

The third type of assay cited usually involves two or more doses each of the standard and the unknown, with equal logarithmic intervals between the doses. This can be illustrated by FIGURE 3. Here, the dosage is three of the unknown (low, medium, and high doses), whose responses are labeled U_L , U_M , and U_H , and three of the standard, whose responses are labeled S_L , S_M , and S_H . The difference between the logarithms of the doses is i . This approach can be illustrated by the equation: $Y = a + b \log X$, where the slope b is the same for standard and unknown, but a differs.

FIGURE 4 shows a two-dose assay (two doses on each of standard and unknown). It is based on the penicillin cup-plate assay, wherein the measure of

response is the diameter of the zone of inhibition of growth of the test organism. Relative potency is usually calculated as the ratio between the doses of standard and unknown that result in equal responses.

In terms of logarithms, this ratio is a difference—the horizontal distance M between the two parallel lines in FIGURE 4. The slope of a line is equivalent

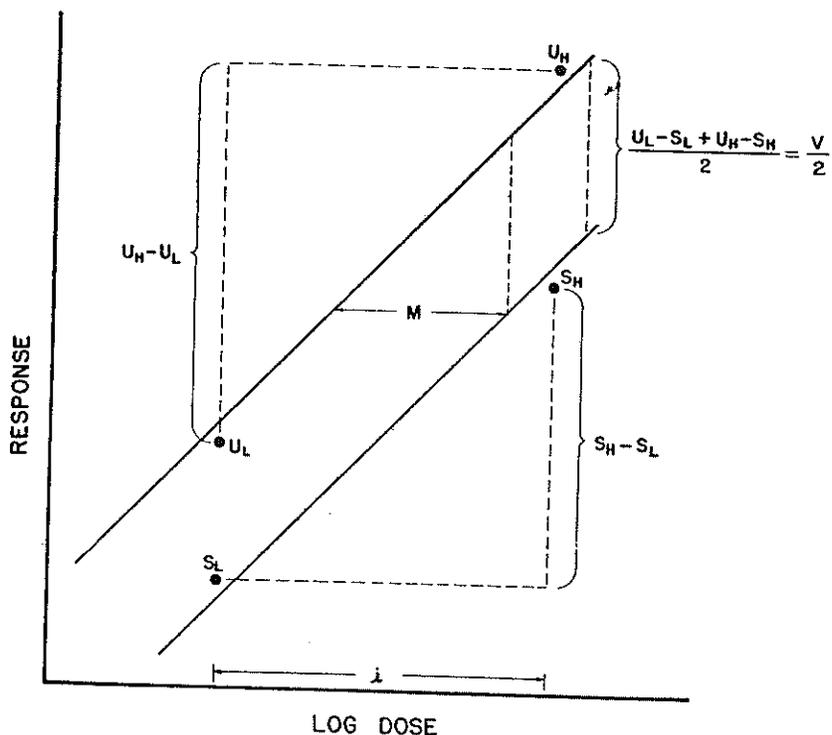


FIGURE 4. Two-dose penicillin cup-plate assay.

to the change in response for a unit change in log dose. In this case,

$$\text{slope} = \frac{(U_H - U_L) + (S_H - S_L)}{2i} = \frac{w}{2i}$$

As can be seen in FIGURE 4, however, the slope is also equal to the vertical distance between the lines divided by M . Therefore, since the vertical distance between the lines can be calculated as

$$\text{vertical distance} = \frac{U_L - S_L + U_H - S_H}{2} = \frac{v}{2},$$

then

$$M = \frac{\text{vertical distance}}{\text{slope}} = \frac{iv}{w},$$

and potency as per cent of standard = $\text{antilog}(2 + M)$.

There are many different mathematical ways of stating the formula for the potency in this type of assay, but all are based on similar reasoning. The slope involved in assays having 3 or more doses is usually calculated by least squares, and the formula for the vertical distance between the two parallel lines usually involves the slope.

FIGURE 4 represents one plate of a four-plate assay. The standard error of this type of assay can be calculated by considering the variation between plates. It can be calculated from the formula,

$$\text{standard error of assay} = \frac{k(\text{potency})}{W} \sqrt{R_v^2 + \frac{R_w^2 V^2}{W^2}},$$

where $k = 2.3026 i \sqrt{n/D}$, n = the number of plates, and, as before, D is the average number of standard deviations in the range. (For $n = 4$ and $i = 0.602$, $k = 1.3464$.)

TABLE 2
PENICILLIN PLATE ASSAY; RATIO OF DOSES = 4:1

Plate no.	s_L 0.25 u/ml	s_H 1.0 u/ml	u_L estimated 0.25 u/ml	u_H estimated 1.0 u/ml	v or $(u_L + u_H) -$ $(s_L + s_H)$	w or $(s_H + u_H) -$ $(s_L + u_L)$
	mm	mm	mm	mm		
1	16.0	22.5	15.0	20.0	-3.5	11.5
2	16.2	22.5	14.5	19.5	-4.7	11.3
3	16.0	22.5	15.0	22.0	-1.5	13.5
4	15.0	22.0	14.0	21.0	-2.0	14.0
Sum.....	63.2	89.5	58.5	82.5	-11.7 = V	50.3 = W
Range.....					3.2 = R_v	2.7 = R_w

The values of v and w are calculated as given in FIGURE 4. For a four-plate assay, one value of v and w is calculated for each plate. R_v is the highest value of v minus the lowest value and is called the range of the v 's. R_w is in the range of the values of w . The value of V is the sum of the values of v for individual plates and W is the sum of the values of w .

TABLE 2 may clarify the method of handling data obtained by this type of assay. Here $i = 0.602$, $V = -11.7$, $W = 50.3$, $R_v = 3.2$, and $R_w = 2.7$. Substituting the values in the equations for potency and standard error of the assay we obtain

$$M = \frac{.602 (-11.7)}{50.3} = -0.1400.$$

$$\text{Potency as \% of standard} = \text{antilog} (2 - 0.1400) = 72.4.$$

$$\text{Standard error of the assay} = \frac{1.3464 (72.4)}{50.3} \sqrt{(3.2)^2 + \frac{(2.7)^2 (11.7)^2}{(50.3)^2}}.$$

$$\text{Standard error of the assay} = 6.3.$$

These values can also be obtained by using a chart, such as shown in FIGURE 5, for obtaining potency from V and W and the nomograph for obtaining the ratio of the standard error of the assay to the potency, as shown in FIGURE 6.

If a routine assay procedure is conducted in the laboratory, it may be possible to have the assay results in statistical control, so that control charts can be kept on the values W and R_w , such as shown in FIGURE 7. Here the data obtained from each assay are plotted against time. The average values over this period of time are calculated and plotted as horizontal straight lines. "Control limits" are calculated from the variation indicated

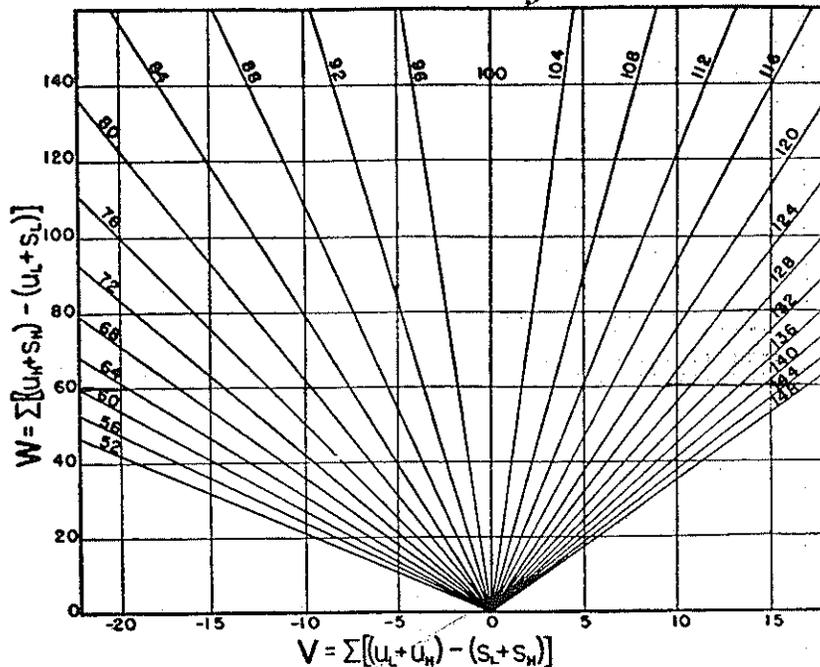


FIGURE 5. Chart for determining potency as a percentage of the standard from the two-dose plate method, where the ratio of high and low doses is 4:1.

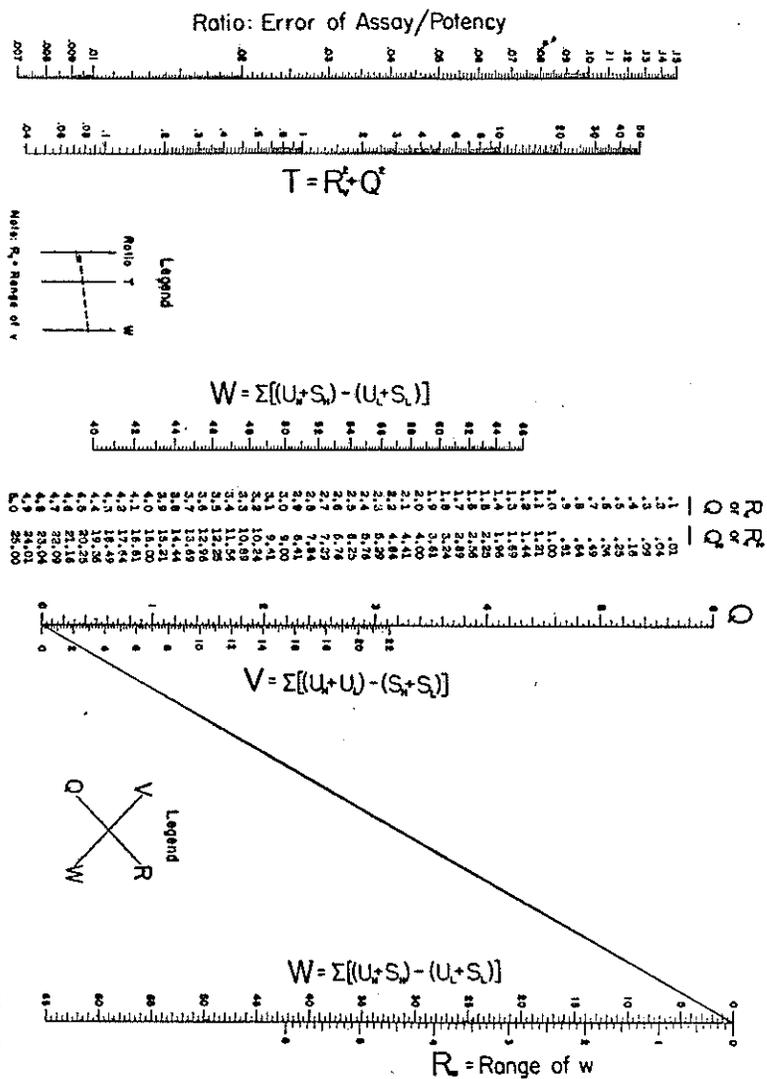
within an assay, and these control limits are plotted as horizontal dotted lines. If the assays are in control (*i.e.*, within the dotted lines), the average value of W can be used. Thus, the chance variation in \bar{W} , the average value of W, becomes very small and can be disregarded, and the formula for the standard error of the assay becomes

$$\text{standard error of assay} = \frac{kR_v (\text{potency})}{\bar{W}}$$

Thus, for a series of assays in statistical control over a period of time, the standard error of the assay is a more or less constant percentage of the potency.

It is possible to show the effect of the size of the standard error of the

FIGURE 6. Nomograph for estimating the error of the assay from the two-dose, four-plate assay, where the ratio of high and low doses is 4:1.



assay on the type of material determined to be acceptable by a laboratory's routine assay procedure. Every laboratory has a working limit for routine assays such that, if the potency result from an assay should fall above that limit, the lot or batch from which the sample was taken is regarded as passing and, if the potency result from the assay of a sample falls below that limit, the lot or batch is regarded as failing. In some cases, there are both lower and upper limits. For instance, lots may be rejected if the assayed potency is greater than ± 20 per cent from the standard, as in the case U. S. Pharmacopeia specification for digitalis."

In making a decision about a batch or lot of materials where that decision is on the basis of the results of an assay, two kinds of errors may be made: (1) one may decide to accept a lot when it should be rejected (*i.e.*, when an average of many assay results would result in its rejection); (2) one may decide to reject a lot when it should be accepted.

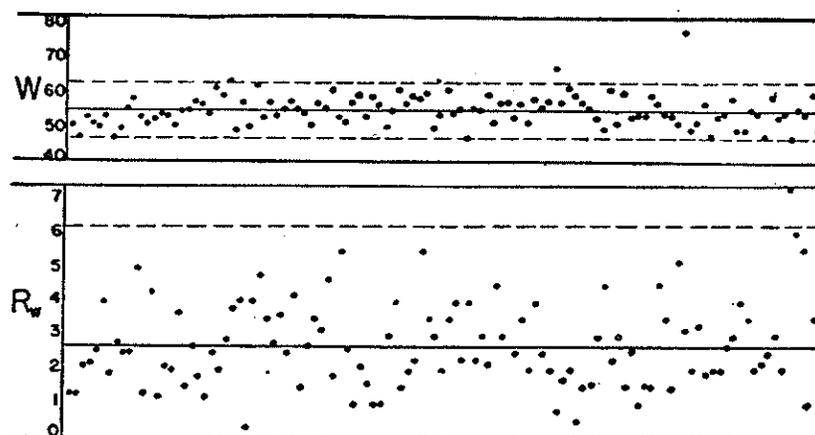


FIGURE 7. Control charts for two-dose, four-plate penicillin assay.

Suppose we consider only a lower limit and assume that the standard error of the assay is 10 per cent. We can then calculate the probabilities of making these wrong decisions. A product will be accepted half of the time ($P = .50$ on FIGURE 8), if its potency, as measured by an infinite number of determinations at that laboratory (here it is labeled "real potency," but it may not be true potency), is at the acceptance limit. Lots whose "real potency" at that laboratory is 10 per cent below the limit will be accepted 15 per cent of the time, and lots whose "real potency" at that laboratory is 20 per cent below the limit will have a probability of acceptance of 0.01. Likewise, lots whose "real potency," as measured at that laboratory, is 20 per cent above the limit will be accepted 97 per cent of the time and rejected 3 per cent of the time, and lots 10 per cent above the limit will be rejected 17 per cent of the time. This type of graph is called an "operating characteristic curve." It is one way of showing how closely a laboratory can expect to check its own results.

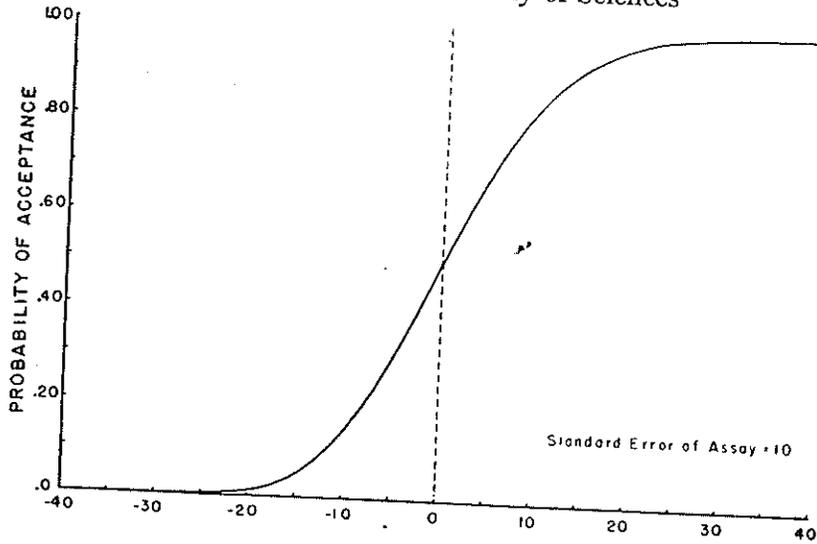


FIGURE 8. Operating characteristic curve for an assay procedure, where the standard error of the assay equals 10 per cent and acceptance or rejection of the lot is based on the results of one assay.

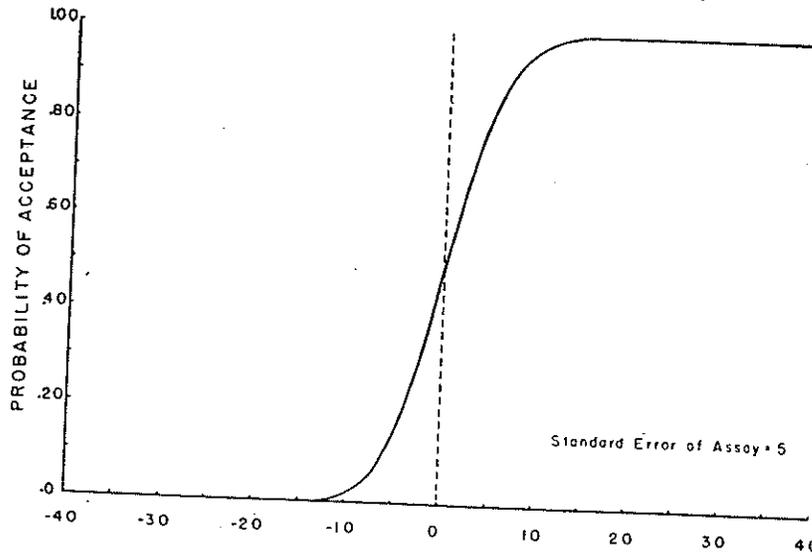


FIGURE 9. Operating characteristic curve for an assay procedure where the standard error of the assay equals 5 per cent and acceptance or rejection of the lot is based on the results of one assay.

Suppose, however, that the standard error of the assay was 5 instead of 10 per cent and the same lower limit was set for "passing" or "failing" material; then the operating characteristic curve will be shown as in FIGURE 9.

The amount of material accepted at the acceptance limit will still be 50 per cent, but 10 per cent below the limit, only 2 per cent of the lots will be accepted, instead of 15 per cent, as shown on FIGURE 8. At 5 per cent below the limit, 15 per cent of the lots will be accepted. At 5 per cent above the limit, 16 per cent of the lots will be rejected. At 10 per cent above the limit, only 3 per cent will be rejected, as contrasted to 17 per cent in FIGURE 8.

If the standard error of the assay is still 5, but a batch is accepted as up to standard or rejected as being below standard on the average of 3 assays instead of on the results of a single assay, the operating characteristic curve will be as shown in FIGURE 10. Here, only 3.8 per cent of the lots whose real potency (as measured by that laboratory) is 5 per cent below the limit

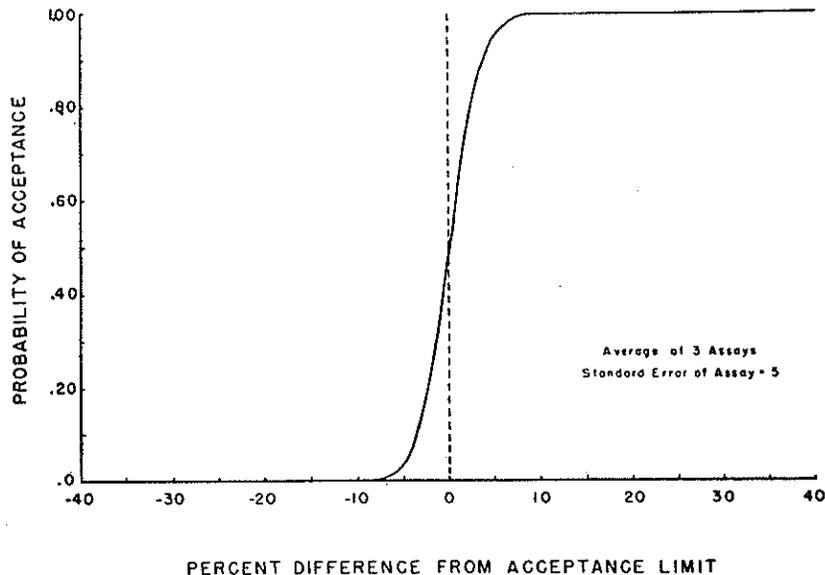


FIGURE 10. Operating characteristic curve for an assay procedure, where the standard error of the assay equals 5 per cent and acceptance or rejection of the lot is based on the average of three assay results.

will be accepted and only 4.6 per cent of the lots whose real potency is 5 per cent above the limit will be rejected.

To operate efficiently in an assay laboratory, one should know the chances of rejecting a "good" lot and passing a "bad" lot. In order to do this, one must have an estimate of the precision of a particular assay at one's laboratory. Statistical methods can play an important part in giving an objective estimate of precision rather than relying on very subjective "impressionistic statistics."

In summary, statistical methods can be applied to various types of microbiological assays. These methods can be greatly simplified, so that they can be easily utilized in the laboratory. In all instances, common sense and logic must be used in applying the statistical method.

An admonition from D. J. Finney will serve to emphasize the need for caution in selecting the proper statistical approach: "The unwary are fre-

quently entrapped by forgetting that the possibility of performing certain arithmetical operations provides no guarantee that the corresponding statistical technique is appropriate to the data."

References

- BLISS, C. I. 1944. A simplified calculation of the potency of penicillin and other drugs assayed biologically with a graded response. *J. Am. Stat. Assoc.* **39**: 479-487.
- BLISS, C. I. 1946. A revised cylinder-plate assay for penicillin. *J. Am. Pharm. Assoc.* **35**: 6-12.
- BLISS, C. I. 1946. An experimental design for slope-ratio assays. *Ann. of Math. Stat.* **17**: 232-237.
- DE BEER, E. J. & M. B. SHERWOOD. 1945. The paper-disc agar-plate method for the assay of antibiotic substances. *J. Bact.* **50**: 459-467.
- FIELLER, E. C. 1947. Some remarks on the statistical background in bio-assay. *J. Soc. of Public Analysts* **72**: 35-82.
- FINNEY, D. J. 1945. The microbiological assay of vitamins: the estimate and its precision. *Quart. J. Pharm. & Pharmacol.* **18**: 77-82.
- FINNEY, D. J. 1947. The principles of biological assay. *Suppl. J. Roy. Stat. Soc.* **9**: 46-91.
- FISHER, R. A. & F. YATES. 1948. *Statistical Table for Biological, Agricultural and Medical Research*. 3rd Ed. Hafner. New York.
- IRWIN, J. O. 1943. On the calculation of the error of biological assay. *J. of Hygiene* **43**: 121-128.
- KNUDSEN, L. F. & W. A. RANDALL. 1945. Penicillin assay and its control chart analysis. *J. Bact.* **50**: 187-200.
- OSWALD, B. J. & L. F. KNUDSEN. 1950. A turbidimetric method for the assay of streptomycin and its critical evaluation. *J. Amer. Pharm. Ass'n Sci. Ed.* **3**:(2). In press.
- SNEDECOR, G. W. 1946. *Statistical Methods*. 4th Ed. Iowa State College Press. Ames, Iowa.
- SHERWOOD, M. B., E. A. FALCO, & E. J. DE BEER. 1944. A rapid quantitative method for the determination of penicillin. *Science* **88**: 247-248.
- WOOD, E. C. & D. J. FINNEY. 1946. The design and statistical analysis of microbiological assays. *Quart. J. of Pharm. & Pharmacol.* **19**: 112-127.
- WOOD, E. C. 1946. The theory of certain analytical procedures with particular reference to microbiological assays. *The Analyst* **71**: 1-48.

References for Operating Characteristic Curves and Control Charts

- EISENHART, C., M. W. HASTAY, & W. A. WALLIS, (Editors). 1947. *Selected Techniques of Statistical Analysis for Scientific and Industrial Research and Production and Management Engineering*. McGraw-Hill. New York.
- GRANT, E. L. 1946. *Statistical Quality Control*. McGraw-Hill. New York.