THE CALCULATION OF MICROBIAL ASSAYS

C. I. BLISS

The Connecticut Agricultural Experiment Station and Yale University, New Haven, Connecticut

Microbial assays of the vitamins, and of other nutrients such as amino acids, are still a research necessity, despite the encroachment of physicochemical techniques for standardization. Although their role in standardization is shrinking, the need of industry for quantitative bioassays with a measurable precision has produced some useful designs and calculations that the research worker could adopt with profit. The quantitative biological assays in U.S.P. XV, for example, have been reduced to a few basic patterns. Included in this program are the two remaining U.S.P. microbial assays for the vitamins, those for vitamin B_{12} and for calcium pantothenate. The methods developed there are by no means restricted to these two vitamins but are easily adapted to the bioassay of all vitamins and amino acids in which a growth response is measured in a limiting nutrient. By investing a little more effort, the microbiologist can gain appreciable dividends in the form of more reliable and precise results. The purpose of this paper is to show how this may be done.

BASIC EXPERIMENTAL DESIGN

A standard pattern for a microbial assay is to pipette, into a replicated series of sterilized test tubes, doses varying from 1.0 to 5.0 ml of test solution. The experimenter adds to each tube 5.0 ml of basal medium stock solution and sufficient water to make 10 ml. The tubes are covered, sterilized in an autoclave, cooled, inoculated, and incubated for 16 to 24 hr at a constant temperature. The transmittance of the contents of each tube, well mixed, is then measured in an electrophotometer at a specific wave length and recorded as a percentage of the inoculated blank.

With the aid of rough, preliminary assays or past experience, test solutions are prepared of each sample or Unknown that have approximately the same concentration of vitamin, or other nutrient, as the test solution of Standard. The assay determines how much test solution of each Unknown will produce the same growth response, measured electrophotometrically, as the test solution of the Standard. This relative potency is adjusted for the known difference in concentration of the two test solutions to obtain the biologically effective content of the vitamin in each Unknown. Four conditions, developed primarily in studies of the vitamin B_{12} assay, are of general applicability and should increase the reliability of microbial assays. They may be considered in order.

A straight log-dose response curve. Apart from technical details, which vary with the material under test, much of the success of an assay depends upon the concentration of the test solutions. By covering a wide enough dosage range, the growth response, in terms of $y = 100 -$ (per cent transmittance), can be plotted against the log-dose as a sigmoid curve, the response then increasing (instead of decreasing) with the dose. The curve for the pantothenic acid Standard in figure 1 is an example. Its fiftyfold range (obtained by using two or more concentrations of test solution) is far wider than the threefold to fivefold range in a customary assay.

An effective assay requires only the central linear portion of the complete curve. There may exist some transformation of the response, based upon a sigmoid function, which would plot as a straight line against the log-dose over the entire range, but we have yet to find such a "cure-all." The angular or arc sine transformation of y extends the linear zone in figure 1 from a fivefold to a tenfold range of dosages but not much more. Given an appropriate concentration of test solution, and a suitable organism and medium, enough of the complete curve will be linear for assay purposes.

With some nutrients, the lower part of the curve yields a straight line when the growth response is plotted against arithmetic dosage units, leading to the so-called slope-ratio assays. These are covered by a well-developed series of statistical techniques which need not be considered here (Bliss, 1952; Finney, 1952). Sometimes the smaller growth responses can be converted to $z = \log y$ to obtain a straight line relating z to x, where $x =$ log-ml ($\log_{10}$ ml of solution taken). Conversely, at dosage levels giving comparatively

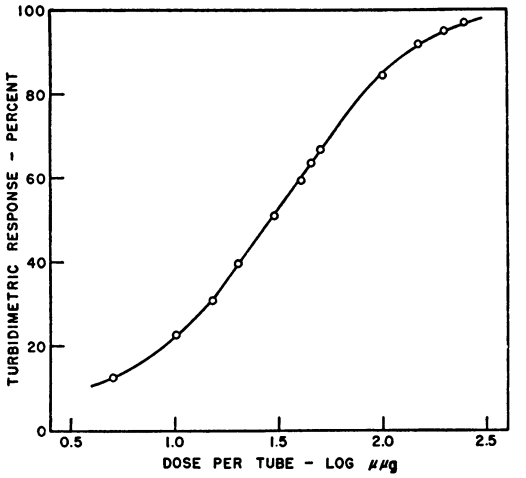

Figure 1. Log-dose response curve for pantothenic acid in an overnight turbidimetric assay with *Lactobacillus plantarum*, showing the range over which the response may be considered a linear function of the log-dose.

large responses, we may obtain a straight line with $z = 2 - \log$ (per cent transmittance). "Logging" the response in this way has proved effective in rectifying log-dose response curves for vitamin B_{12} (Bliss, 1957). The relatively long central portion of the curve, however, is experimentally indistinguishable from a straight line, and the simplest procedure is to select dosages that would be expected to fall within this range. The pantothenic acid curve in figure 1 provides a fivefold to sixfold range of working dosages, a range which can be matched, at least, with vitamin B_{12} and probably with other nutrients. Since a threefold range is usually ample, the arithmetic is minimized by selecting a concentration of test solution for which the simplest growth response, $y = 100 -$ (per cent transmittance), plots as a straight line against the log-ml per tube.

Randomization. The laboratory technician usually identifies a tube by its position in the rack, arranging the tubes of each preparation in a sequence of increasing doses and maintaining the same order in all stages of the assay. It has been shown, however, directly in at least two laboratories and indirectly in others, that the tube positions in a rack influence the observed response. When a number of racks were run with all tubes containing the same amount of vitamin B_{12}, Brownlee and Lapedes (1951) found in every instance significant drifts from one part of the rack to another. In a typical rack, the average transmission in four adjacent tubes varied from 73.9 per cent on one side to 82.5 per cent on the other. A substantial part of this drift, averaging 3.2 per cent in 10 racks, was caused by variation in the medium during autoclaving, as shown by experiments in which the vitamin was added aseptically after autoclaving. The effects of tube position in the autoclave and in the incubator were evaluated separately by Campbell *et al.* (1953), who showed in four blocks of 25 tubes each that the variance between rows and columns within blocks exceeded the error variance by 2.6 times in the autoclave and by 4.3 times in the incubator. In an analysis by Bliss (1956*b*) of the interpolated assayed potencies of the same four Unknowns in three independent systematic assays, the variance between dosage levels exceeded that between triplicates within levels by 3.4, 6.9 and 2.4 times in three different laboratories. Other examples could be cited.

There is no sure way of correcting this bias in an assay where the tubes are arranged systematically. Biologists faced with larger inherent errors have had to abandon systematic designs in order to obtain quantitatively valid results. The basic requirement is the assignment of treatments to experimental units at random, usually within restrictions which segregate the potentially larger sources of variation. The substitution of turbidimetry for titration in measuring the growth response has removed the main barrier to similar randomization in microbial assays. Derandomization just before reading the tubes, as practiced by Campbell *et al.* (1953), would avoid this technical difficulty in assays with a titrated end point.

Since randomization is new to many microbiological laboratories, it is suggested, but is not obligatory, in the U.S.P. XV assay for vitamin B_{12}. To compensate in part for positional effects, there is a new requirement that replicate tubes be placed in different racks or in different parts of the same rack.

Apparently, the main limitation to randomizing the tubes in a microbial assay is that they must be marked individually. This need not be a problem. A supply of tubes for each of the relatively few dosage levels in ml could be marked permanently and stored separately after each assay when the tubes are washed. Each preparation could be identified when pipetting the test

solutions by a mark added with a wax pencil or by a distinctive metal clip. On the assumption that bias during pipetting is negligible, randomization could be started after the tubes for each preparation are made to 10-ml volume. Replicates would then be placed at random in different racks, within each rack intermingling all dosage levels and preparations. After the subsequent stages of autoclaving, inoculation and incubation, the tubes could be read in order within each rack and each response recorded directly in its proper place in the laboratory notebook. A procedure as simple as this would soon be mastered by the laboratory technician and the gain in reliability should far exceed any additional time that it may cost.

When tubes are randomized, the efficient unit is the response of each tube and the variation among replicate tubes is an index to assay precision. But when tubes are arranged systematically, the variation among individual readings does not take account of any potential bias associated with their position. Under these conditions, the total response of the f replicates for each treatment (T_t) is the more logical unit. If preferred, this can be computed directly as $T_t = 100f - \sum$(per cent transmittance) and designated as y, following the U.S.P. XV assay for vitamin B_{12}.

Aberrant and missing values. When replicates are compared, grossly discrepant values may be discarded at once, but an occasional reading may seem suspiciously aberrant. An objective rule for identifying such aberrant tubes is given in U.S.P. XV from tables prepared by Dixon (1951). We assume that each reading is equally subject to experimental error, whether the response is large or small, and that these errors follow a normal distribution. It can then be shown that the differences between corresponding tubes in replicate racks will also be distributed normally.

In an assay with duplicate tubes, for example, each reading in rack 1 is subtracted from its mate in rack 2 to obtain a difference which may be plus or minus. When these differences are arranged in order of size, does an outlying end difference belong in the series? If the odds are less than 1 in 25 that it does belong, one of the two tubes in the difference is assumed to be aberrant, is identified from its relation to the rest of the assay, and rejected. In an assay with triplicate tubes, one reading in a set of three may seem out of line. To apply the same test, the suspected reading is multiplied by 2 and subtracted from the sum of its mates in the other two racks. A similar difference is computed for every other tube in the rack containing the suspected outlier, and the series is then handled as before.

To apply this criterion, the differences are numbered in order of size, beginning with the supposedly erratic difference or outlier, from y_1 to y_N where N is the number of differences. A relative gap G_1 is computed which depends upon N. When

$N = 3$ to 7, $G_1 = (y_2 - y_1)/(y_N - y_1)$;
$N = 8$ to 13, $G_2 = (y_3 - y_1)/(y_{N-1} - y_1)$;
$N = 14$ to 30, $G_3 = (y_3 - y_1)/(y_{N-2} - y_1)$.

If an observed G_i exceeds the critical level in table 1 and there are relatively few, if any, ties, one of the tubes in the largest difference is presumably aberrant; otherwise its value is accepted. The process may be repeated with the remaining differences if a second outlier is suspected.

The gap test may be applied to a randomized vitamin B_{12} assay in three replicated racks (table 2). The reading of 26.5 in rack II for 1.5 ml of Standard was marked "omit" by the experimenter. Was this justified? The $N = 13$ differences, I + III − 2(II), were formed for each treatment, and then rearranged in order of decreasing size. Since the relative gap, $G_2 = 0.469$, was less than 0.578, the critical level for $N = 13$ in table 1, the observed difference did not exceed the experimental error. All observations should be retained in computing the assay.

TABLE 1

Critical levels of G_1 to G_3

N	G_1	N	G_2	N	G_3	N	G_3	N	G_3
3	0.976	8	0.780	14	0.602	20	0.502	26	0.450
4	0.846	9	0.725	15	0.579	21	0.491	27	0.443
5	0.729	10	0.678	16	0.559	22	0.481	28	0.437
6	0.644	11	0.638	17	0.542	23	0.472	29	0.431
7	0.586	12	0.605	18	0.527	24	0.464	30	0.425
		13	0.578	19	0.514	25	0.457		

$P = 0.04$ in the gap test for outliers at either end of a series of N ordered differences (Dixon, 1951).

$G_1 = (y_2 - y_1)/(y_N - y_1)$.
$G_2 = (y_3 - y_1)/(y_{N-1} - y_1)$.
$G_3 = (y_3 - y_1)/(y_{N-2} - y_1)$.

TABLE 2

Gap test for outliers applied to a vitamin B_{12} assay in three randomized sets

Preparation	Dose	Growth Response in Rack			Difference I + III − 2 (II)	In Rank Order
		I	II	III		
	ml					
S	1.5	20.0	26.5	20.0	−13.0	$-13.0 = y_1$
	2	29.5	30.5	29.5	−2.0	−6.5
	3	45.5	44.5	45.0	1.5	$-5.5 = y_3$
	4	56.0	55.0	57.0	3.0	−3.0
	5	63.5	63.0	63.0	0.5	−2.0
U_1	1.5	37.0	41.0	38.5	−6.5	−2.0
	2	49.0	49.5	49.5	−0.5	−1.0
	3	63.0	66.0	63.5	−5.5	−0.5
	4	72.0	72.0	74.0	2.0	+0.5
U_2	1.5	20.5	20.5	19.5	−1.0	+1.5
	2	29.0	30.5	30.0	−2.0	+2.0
	3	45.0	45.5	43.0	−3.0	$+3.0 = y_{N-1}$
	4	58.5	54.5	54.0	3.5	+3.5

$$N = 13, G_2 = \frac{-5.5 - (-13.0)}{3.0 - (-13.0)} = \frac{7.5}{16.0} = 0.469.$$

A subsequent analysis of variance demonstrated good assay validity, with less variation about parallel straight lines for the three preparations than in the interaction of sets by treatments ($F = 0.79$, with $n_1 = 9$, $n_2 = 24$).

With tubes arranged in separate racks or blocks, a rejected or missing value is replaced for computational purposes by a number which depends not only upon the remaining replicate or replicates, but also upon how the rack with the missing value compares with the other racks in the assay. This technique has long been used by biologists in other areas; it is summarized on page 868 of U.S.P. XV and in the last section of this paper. More than one aberrant value can be replaced similarly by statistically valid substitutes, but if many tubes are aberrant or much outside the response range for the Standard, the assay would ordinarily be repeated.

Independent replicates. Another feature in the U.S.P. XV assay for vitamin B_{12} is the requirement that the potency of a given Unknown must be based upon two or more runs on different days with independently prepared test solutions. An informal poll of a vitamin B_{12} panel showed that some members repeated an assay on three or four different days before committing themselves as to the vitamin B_{12} content of a given Unknown. This tallied with the statistical evidence from two collaborative studies (Brownlee, 1952; Bliss, 1956*b*), that the variation among independent assays in the same and in different laboratories could be two to three times as large as that estimated from the variation within a single run. At least part, and perhaps all, of this discrepancy is a byproduct of the systematic arrangement of tubes. Thus, in randomized assays with 95 per cent confidence intervals of ±5 to ±10 per cent, Campbell *et al.* (1953) could check their assayed potencies upon repetition within limits based upon these internal errors. Until randomization has been tried more widely and been shown to be a complete cure, independently replicated runs are a necessary safeguard in the assay of vitamin B_{12}.

A similar but less marked discrepancy has been observed with calcium pantothenate and it is a potential factor in all microbial assays. Before it can be disregarded, the agreement of independent runs of the same preparation should be compared with the precision expected from the variability within each run.

Evaluating an assay. To take advantage of the greater inherent reliability stemming from these improvements, the assayed potency needs to be computed rather than interpolated graphically. Its reliability can then be estimated in terms of a confidence interval which would be expected to include the true potency of the Unknown in a given proportion of assays, usually in 19 out of 20. The calculation is not difficult and may take one of two forms; both will be described. The first is the numerical equivalent (with extensions) of the traditional interpolation in a plotted Standard curve. While more objective and precise than a graphic estimate, it uses only part of the latent information in an assay. Factorial analysis, the second procedure, makes the most of the data and, in consequence, gives more reliable estimates of potency and its precision. Factorial analysis has been used effectively for both animal and microbial assays.

These basic improvements in the design of microbial assays do not exhaust the possibilities for increasing their efficiency. Restricting the randomization of tubes within a rack by means of a balanced or confounded design may segregate additional patterns of variation, so that they will neither bias the estimated potency nor con-

tribute to the error. Brownlee and Lapedes (1951) have described several alternatives.

Another advance would increase the number of replicates of the Standard in multiple assays. Several Unknowns may be assayed against the same tubes of Standard, with one complete set of dosage levels for each Unknown in each randomized rack. For three to six Unknowns, maximum efficiency would call for allotting two tubes per rack to each dose of Standard; for seven to 12 Unknowns per rack, the number of tubes at each dose of Standard should be increased to three (Finney, 1952). Doubling or trebling the tubes of Standard in these two cases could increase the potential precision over equal numbers for each preparation by a maximum of 11 and 25 per cent, respectively (for four and nine Unknowns). One form that the calculation might then take is described elsewhere (Bliss, 1956c). These and other refinements are beyond the scope of the present paper.

CALCULATION FROM THE STANDARD CURVE

The range of log-doses against which the percentage growth response y, or its logarithm z, can be plotted linearly varies with the vitamin, the laboratory, and the assay within a laboratory. If the linear range were reasonably stable, the individual tubes were randomized, and the number of replicates increased, the most efficient design statistically would be a three-dose assay. However, unpredictable changes in the sensitivity of the test organism may shift the log-dose response curve sufficiently, within as well as between laboratories, that four dosage levels are usually preferred for each Unknown. The curve for the Standard often includes two additional doses, one below and one above the range adopted for the Unknowns. These are a check on the validity of linear extrapolation beyond range of the usual four doses and allow for inaccuracies in the assumed potency of the Unknowns.

The equation for the Standard curve. The first step is to plot the total response T_t of the f replicates at each dosage level against x, the log-ml of test solution. Inspection of the points will indicate any obvious curvature which could be corrected by transforming y to $z = \log y$, or to $z = 2 - \log$ (per cent transmittance). The plotted responses y or z may be approximately linear except at one end; without this end dose the curve may still cover a wide enough range for an acceptable assay. Convenient dosage series for various combinations of three to six levels in the range from 1 to 5 ml are given in table 3.

When the relation is apparently linear, the best-fitting straight line can be computed in units of T_t as

$$Y = \bar{T}_t + b(x - \bar{x}) = a' + bx \qquad (1)$$

where Y is the estimated total response, $\bar{T}_t = \sum T_t/k$ is the average total response over the k dosage levels of the Standard, b is the slope,

TABLE 3

Orthogonal polynomials x_1 and x_2 based upon the log-ml per tube for different dosage sequences

Dose per Tube		Polynomials for Indicated Sequence of k Doses																	
ml.	x	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2
1	0	−1	1					−13	1	−47	19					−34	11	−15	52
1.5	0.1761					−29	1	−6	−1					−20	20			−8	−21
2	0.3010	0	−2	−28	22	−12	−1			−6	−33	−61	28	−11	−12	−9	−14	−3	−44
3	0.4771			3	−53	12	−1	6	−1	18	−12	−12	−41	2	−23	5	−10	4	−35
4	0.6020	1	1	25	31	29	1			35	26	23	−21	11	−6	15	0	9	1
4.5	0.6532							13	1										
5	0.6990											50	34	18	21	23	13	13	47
$\sum x_1^2 = \frac{1}{2}c_b$		2		1418		1970		410		3794		6894		970		2016		564	
$\sum x_2^2 = \frac{1}{2}c_q$		6		4254		4		4		2270		4062		1550		586		8516	
$i' = i'$		0.30103		0.0056798		0.0073433		0.025117		0.0073416		0.0035839		0.013737		0.012293		0.025003	
$\bar{x} = \sum x/k$		0.30103		0.46007		0.38908		0.32661		0.34505		0.51980		0.45105		0.41584		0.37588	
$\frac{1}{2}ci$		0.20069		2.6847		3.6166		2.5745		6.9635		6.1769		2.6650		4.9565		2.3503	
$\frac{1}{2}c'i^2$		0.12083		0.030497		0.053115		0.12933		0.10225		0.044274		0.073218		0.12186		0.11753	

$i' = \sum\{x_1(x - \bar{x})\}/\sum x_1^2$.
$\frac{1}{2}ci = i'\sum x_1^2/k$.
$\frac{1}{2}c'i^2 = 2i'^2\sum x_1^2/k$.

TABLE 4

A turbidimetric vitamin B_{12} assay

Preparation	Dose	log-ml x	Response y in Set			T_t	Polynomials	
			I	II	III		x_1	x_2
	ml							
S	1	0	33	32.5	32.5	98.0	−15	52
	1.5	0.1761	41	39.5	41.5	122.0	−8	−21
	2	0.3010	46	47	48	141.0	−3	−44
	3	0.4771	55	53	53	161.0	4	−35
	4	0.6021	59.5	58.5	59.5	177.5	9	1
	5	0.6990	64	62.5	64.5	191.0	13	47
Total						890.5	0	0
							x'	M'
U_1	1.5	0.1761	48.5	46	45.5	140.0	0.1358	
	2	0.3010	53	52.5	52	157.5	0.1439	0.1470
	3	0.4771	62	58.5	62.5	183.0	0.1616	
	4	0.6021	66	66	65.5	197.5	0.1468	
Total							0.5881	
U_2	1.5	0.1761	41.5	42	43.5	127.0	0.0370	
	2	0.3010	47.5	48	49.5	145.0	0.0489	0.0264
	3	0.4771	55.5	52.5	55.5	163.5	0.0134	
	4	0.6021	58	61.5	59.5	179.0	0.0062	
Total							0.1055	
U_3	1.5	0.1761	42.5	43	42	127.5	0.0408	
	2	0.3010	48.5	49.5	46	144.0	0.0413	0.0454
	3	0.4771	57	54.5	53.5	165.0	0.0248	
	4	0.6021	61.5	62.5	64	188.0	0.0746	
Total							0.1815	
Total, T_r, 1.5–4 ml			843.0	834.5	841.0	2518.5		

Design: each set of tubes placed at random in a different rack; assumed potencies of Unknowns, B_{12} in each ml: $U_1 = 2.4\,\mu g$, $U_2 = 56\,\mu g$, and $U_3 = 28\,\mu g$; $f = 3$, $h = 3$.

Standard curve: $b = 1855.5/14.1017 = 131.58$ (equation 2), where $\Sigma(x_1T_t) = -15 \times 98.0 - 8 \times 122.0 - \cdots + 13 \times 191.0 = 1855.5$ and $i'\Sigma x_1^2 = 0.025003 \times 564 = 14.1017$ from table 3, last column; $a' = 148.42 - 131.58 \times 0.37588 = 98.956$ (equation 1), where $\overline{T}_t = (98.0 + 122.0 + \cdots + 191.0)/6 = 148.42$ for $k = 6$, and $\bar{x} = 0.37588$ from table 3, last column; $B^2 = 1855.5^2/3 \times 564 = 2034.80$ and $Q^2 = -150.5^2/3 \times 8516 = 0.887$ (equation 5), where $\Sigma x_2T_t = 52 \times 98.0 - 21 \times 122.0 - \cdots + 47 \times 191.0 = -150.5$.

Log-relative potency from Standard curve: at each dosage level, from equation 6, $X' = 0.007600(T_t - 98.956)$, where $b^{-1} = 1/b = 1/131.58 = 0.007600$; for U_1 at 1.5 ml, $X' = 0.0076(140.0 - 98.956) - 0.1761 = 0.1358$, at 2 ml, $X' = 0.0076(157.5 - 98.956) - 0.3010 = 0.1439$; for U_1, $M' = \overline{X}' = 0.5881/4 = 0.1470$.

Precision of M' from Standard curve: $s^2 = 2(0.100342 - 0.097484)/9 = 0.000635$ (equation 8), where $\Sigma X'^2 = 0.1358^2 + 0.1439^2 + \cdots + 0.0746^2 = 0.100342$, $\Sigma(\Sigma^2X'/k) = 0.5881^2/4 + 0.1055^2/4 + 0.1815^2/4 = 0.097484$, and $n = 3 \times 4 - 3 = 9$; $\frac{1}{2}L = 0.0252 \times 2.262/\sqrt{4} = 0.0285$ (equation 9), where $s = \sqrt{0.000635} = 0.0252$, $t = 2.262$ from table of Student's t at $n = 9$ and $P = 0.05$, and $k = 4$; percentage limits = 100 antilog(0.0285) − 100 = 100 × 1.068 − 100 = 6.8.

$\bar{x}$ is the mean log-dose in ml of test solution (table 3), and $a' = \bar{T}_t - b\bar{x}$. The slope b may be computed conveniently with the coded log-doses x_1 in table 3. These are small whole numbers, with nearly the same relative spacing as the x's, which total 0 in each series. The slope of T_t upon x is computed with the x_1's as

$$b = \sum(x_1T_t)/i'\sum x_1^2 \qquad (2)$$

where $\sum$ stands for the sum of the terms which follow it, and i' is the interval in logarithms per unit of x_1. Both i' and $\sum x_1^2$ are given in table 3 for each dosage sequence. Since the spacing of the x_1's differs slightly from that of the log-ml per tube, each i' is a weighted average, computed by weighting the ratio $(x - \bar{x})/x_1$ for each dosage level by x_1^2.

The numerical example in table 4 is a vitamin B_{12} assay based upon the procedure in U.S.P. XV but with three instead of two replicates. The tubes of each replicate or set were placed in a separate rack at random. The reading for each tube is in terms of $y = 100 -$ (per cent transmittance); these have been totaled in each row to obtain the T_t's. The T_t's for the Standard have been plotted against their log-doses x in figure 2 and fitted with the straight line: $y = 98.956 + 131.58\,x$. As shown beneath table 4, its slope by equation 2 is $b = 131.58$, its mean response is $\bar{T} = 148.42$, and its mean log-dose $\bar{x}$ from table 3 (last column) is $\bar{x} = 0.3759$. The relation is obviously linear.

A test for curvature. Occasionally, the plotted points may form a slightly convex or concave curve about the fitted straight line. A small curvature can be disregarded and potencies computed as if the log-dose response line were straight. To test whether this approximation is acceptable in a given case, we may fit a parabola and compare the amounts of variation attributable to its linear and quadratic terms.

The parabola is defined as

$$Y = a + b_1x + b_2x^2 \qquad (3)$$

Just as the log-doses x have been coded to x_1, the x^2's can be replaced similarly by orthogonal polynomials x_2, which measure the quadratic curvature, total 0 ($\sum x_2 = 0$), and have the property that $\sum(x_1x_2) \doteq 0$ (Bliss and Calhoun, 1954). The corresponding x_2's parallel the x_1's for each dosage sequence in table 3. When determined from the T_t's, the predicted parabolic

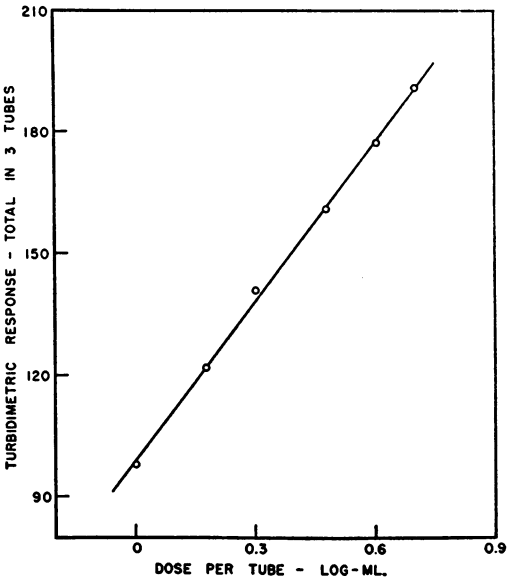

Figure 2. The log-dose response curve for the Standard in the assay of vitamin B_{12} in table 4, from a test solution containing 0.02 μg of cyanocobalamin per ml in an overnight turbidimetric assay with *Lactobacillus leichmanii*. The straight line has the equation $Y = 98.956 + 131.58x$, where Y is the predicted total response in three replicate tubes and x is the log-ml of Standard test solution per tube.

response at each dosage level may be computed as

$$Y = \bar{T}_t + B_1x_1 + B_2x_2 \qquad (4)$$

where

$$B_1 = \sum(x_1T_t)/\sum x_1^2$$

and

$$B_2 = \sum(x_2T_t)/\sum x_2^2.$$

The variation in units of y^2 attributable to B_1 and to B_2, respectively, is computed as

$$B^2 = \sum{}^2(x_1T_t)/f\sum x_1^2$$

and (5)

$$Q^2 = \sum{}^2(x_2T_t)/f\sum x_2^2$$

where f is the number of replicates or racks and $\sum^2$ indicates the square of the sum of the succeeding terms in parentheses. If $B^2/Q^2 > 100$, the error in the estimated potencies due to ignoring the curvature is usually negligible (Finney, 1952; Bliss, 1957).

From figure 2, the log-dose response curve for

the Standard in table 4 is obviously linear, but it may be used to illustrate the numerical test. The effect of the linear term or slope is $B^2 = 1855.5^2/(3 \times 564) = 2034.80$. From the coefficients x_2 for quadratic curvature, $\sum(x_2T_t) = -150.5$, from which $Q^2 = 150.5^2/(3 \times 8516) = 0.887$. The ratio $B^2/Q^2 = 2295$ is far above the proposed minimal level for disregarding the curvature.

Log potencies from the equation for the Standard. A separate estimate of the log-relative potency of each Unknown is computed from the T_t at each dosage level, substituting each T_t in turn for Y in equation 1. Each treatment total may be calculated alternatively as $T_t = 100f - \sum$ (per cent transmittance), or, if logarithms are required, as $T_t = \log\{100f - \sum(\text{per cent transmittance})\}$ or $T_t = 2 + \log f - \log \sum(\text{per cent transmittance})$. By rearranging equation 1, the log-relative potency X' from each dosage level of each Unknown is determined as

$$X' = b^{-1}(T_t - a') - x \qquad (6)$$

where x = log-ml of test solution. The k values of X' are then averaged to obtain for a given test solution the log-relative potency $M' = \overline{X}' = \sum X'/k$. A strongly marked increase (or decrease) in X' at successive dosage levels of the same test solution would throw doubt on the parallelism of the dosage-response curves of the Unknown and the Standard, and, in consequence, upon their qualitative equivalence—*if* the assay has been randomized. If the assay has not been randomized, the "drift" may be due instead to the relative positions of the tubes during the assay.

An assayed log-relative potency can be converted to vitamin content as

$$P_* = \text{antilog } M = \text{antilog } (M' + \log R) \qquad (7)$$

R is the μg of the vitamin assumed to be present in each mg (or ml) of the original Unknown when preparing its test solution.

For computing the log-relative potencies of the Unknowns in table 4, the reciprocal of the slope of the Standard curve is determined as $b^{-1} = 1/131.58 = 0.007600$. Substitution in equation 6 gives the four estimates for each Unknown as $X' = 0.0076(T_t - 98.956) - x$. Their averages ($\overline{X}'$) are the assayed log-relative potencies M' in the last column of table 4. Based on preliminary assays, the three Unknowns in table 4, all solutions, were assigned assumed potencies per ml of 2.4 μg, 56 μg and 28 μg, respectively. For U_1, for example, log R = log (2.4) = 0.3802, from which the log potency $M = 0.1470 + 0.3802 = 0.5272$ by equation 7. From the antilogarithm of M, the three Unknowns had assayed contents of vitamin B_{12} per ml of 3.367, 59.51, and 33.31 μg, respectively.

Precision of an assayed potency. The error variance of each log-relative potency $M' = \overline{X}'$ depends upon terms involving both the Standard curve and the Unknown. Although an exact calculation would be impracticable, it may be approximated from the k values of X' and their totals $\sum X'$ for each of the h Unknowns as

$$s^2 = 2\{\sum X'^2 - \sum(\textstyle\sum^2 X'/k)\}/n \qquad (8)$$

with $n = \sum k - h$ degrees of freedom. The confidence interval for each M' is then

$$L = 2st/\sqrt{k} \qquad (9)$$

where t is Student's t at $P = 0.05$ (tabled in U.S.P. XV, p. 875, and elsewhere). The approximate confidence limits in logs, $M' \times \frac{1}{2}L$, would enclose the true log-relative potency of the Unknown at odds of about 19 in 20 if there were no interassay error. To express these limits as a percentage of the assayed potency, compute 100 antilog($\frac{1}{2}L$) − 100. Approximate upper and lower confidence limits of the potency in terms of vitamin content are given by the antilogarithms of $X_M = M + \frac{1}{2}L$ and $M - \frac{1}{2}L$.

For the numerical example in table 4, the values of X' and their sums in the next to last column of the table lead by equation 8 to $s^2 = 0.000635$ with 9 degrees of freedom, from which $s = 0.0252$. With $t = 2.262$, the half-confidence interval in equation 9 is $\frac{1}{2}L = 0.0285$. From the antilog of $\frac{1}{2}L$, the approximate confidence limits on each side of an assayed potency are 6.8 per cent. The confidence limits for U_1 are 3.15 and 3.60 μg per ml; the others could be obtained similarly. Since this was a randomized assay, these are more likely to approximate the true limits than if a systematic design had been followed.

FACTORIAL ANALYSIS

Factorial analysis takes advantage of two assumptions basic to these and all other parallel-line assays: (1) that operationally the response is a linear function of the log-dose, and (2) that

all preparations have the same assay slope. The calculations are made with a relatively stable combined slope, which measures the average relation between the response and the log-dose over all preparations. The error variance and confidence interval are free of the approximations inherent in equations 8 and 9. If the tubes have been intermingled at random in the tube racks, the response in a single tube forms the preferred unit.

The analysis falls into several stages. Calculating the potency of each Unknown is the first and easiest. In many assays, nothing more may be required. Some assays are more precise than others, so that a statistic C is defined in the next section, which measures assay precision. If an assay does not exceed a predetermined level of C, the calculation may not be carried further. Confidence limits, measuring the precision of each assayed potency, follow readily from C. The statistical tests for curvature and nonparallelism in the next two sections are available for critical or doubtful assays; they need not be computed routinely. Since missing values are a normal experimental hazard, the calculation of incomplete assays is treated separately.

Calculation of the log-relative potency. Potencies are determined from the k dosage levels, in ml per tube, which are common to the Standard and the h Unknowns, a total of $h' = h + 1$ preparations. The totals T_t for the response to each treatment, as given in table 4, may be rearranged conveniently by preparations and dosage levels in the form shown in table 5. The columns and rows are totaled to obtain $T_P = \sum T_t$ for each preparation and $T_d = \sum T_t$ for each dosage level. For each preparation, the T_t's are multiplied in turn by the coefficients x_1 and x_2, and the products totaled to obtain $T_b' = \sum(x_1 T_t)$ and $T_q' = \sum(x_2 T_t)$. The same procedure, applied to the T_d's in the last column, checks the totals $T_b = \sum T_b'$ and $T_q = \sum T_q'$. If the row for T_q' includes both + and − values, the relation can ordinarily be considered as linear. If the range in T_b' does not exceed 15 or 20 per cent of its smallest value, the chances are good that the log-dose response lines for the h' preparations will prove to be parallel with a common assay slope based upon T_b. In most assays, a further test for curvature and for divergence in slope is not needed, and after scanning the results one can proceed directly to calculating the potency of each Unknown, initially in logarithms.

The log-relative potency M' is essentially a ratio of two statistics, the difference between two mean responses $(\bar{y}_U - \bar{y}_S)$ divided by the assay slope b of the log-dose response curve, or $M' = (\bar{y}_U - \bar{y}_S)/b$. In factorial assays, this ratio can be reduced to a very simple pattern. For an assay with one Standard and several Unknowns, each log-relative potency is

$$M' = \tfrac{1}{2}cih'T_a/T_b \qquad (10)$$

where the constant term,

$$\tfrac{1}{2}ci = i'\textstyle\sum x_1^2/k,$$

is given in table 1 for selected dosage sequences, and $T_a = T_U - T_S$ is the difference between the totals T_P for the given Unknown and for the Standard. M' is converted to the assayed potency P_* by equation 7.

In the numerical example of table 5, + and − values of T_q' tend to balance one another, so that there is clearly no curvature. The percentage range in T_b' for slope, $100(2006.5 - 1730.0)/1730.0 = 16$ per cent, is not so large as to compromise the assumption of parallel log-dose response lines. T_a is then computed from the T_P's. From table 3, $\frac{1}{2}ci = 3.6166$ and for the $h' = 4$ preparations, $\frac{1}{2}cih'/T_b = 3.6166 \times 4/7559.5 = 0.0019137$. By equation 10, $M' = 0.0019137\ T_a$ for each Unknown. After adjustment for their respective assumed potencies by equation 7, the three Unknowns had assayed potencies of $P_* = 3.362$, 59.30 and 30.99 μg of vitamin B_{12} per ml. The potencies obtained by interpolation in the equation for the Standard curve and by factorial analysis are in substantial agreement.

Assay precision. The precision of an assay is epitomized most conveniently by the statistic C, which measures the reliability of the assay slope. C depends upon the variation attributable to the slope (B^2), the error variance (s^2) for the assay, and t^2 from a table of Student's t, such as that in U.S.P. XV (p. 875), for the degrees of freedom in s^2.

The variation in y attributable to the assay slope (B^2) is essentially the same statistic as that computed for the Standard alone with equation 5 and $h' = 1$. Here, however, it is restricted to the k dosage levels that are common to both the Standard and the $h' - 1$ Unknowns, and it gains in reliability by including all h' prepara-

TABLE 5

Factorial analysis of the vitamin B_{12} assay of table 4

Dose (Ml per Tube)	Coefficients		Total Response T_t For:				Total
	x_1	x_2	S	U_1	U_2	U_3	T_d
1.5	−29	1	122.0	140.0	127.0	127.5	516.5
2	−12	−1	141.0	157.5	145.0	144.0	587.5
3	12	−1	161.0	183.0	163.5	165.0	672.5
4	29	1	177.5	197.5	179.0	188.0	742.0
$T_P = \Sigma T_t$			601.5	678.0	614.5	624.5	2518.5 = T
$T_b' = \Sigma x_1 T_t$			1849.5	1973.5	1730.0	2006.5	7559.5 = T_b
$T_q' = \Sigma x_2 T_t$			−2.5	−3.0	−2.5	6.5	−1.5 = T_q
$T_a = T_U' - T_S'$				76.5	13.0	23.0	
$M' = 0.0019137 T_a$				0.14640	0.02488	0.04402	
$\frac{1}{2}L$ ($C = 1.00309$)				0.02825	0.02696	0.02704	
For 2 to 4 ml only M'				0.15257	0.02086	0.04564	
$\frac{1}{2}L$				0.03582	0.03286	0.03308	

Assay precision ($k = 4$): $B^2 = 7559.5^2/(3 \times 4 \times 1970) = 2417.35$ and $Q^2 = (-1.5)^2/(3 \times 4 \times 4) = 0.047$ (equation 11); $s^2 = \{134916.75 - 132145.02 - 134860.75 + 2518.5^2/48\}/30 = 53.53/30 = 1.784$ (equation 12), where $\Sigma y^2 = 41^2 + 46^2 + \cdots + 53.5^2 + 64^2 = 134916.75$ and $\Sigma T_r^2/kh' = (843.0^2 + 834.5^2 + 841.0^2)/4 \times 4 = 132145.02$ from table 4 (omitting doses of 1 and 5 ml), $\Sigma T_t^2/f = (122.0^2 + 141.0^2 + \cdots + 188.0^2)/3 = 134860.75$, and $n = (4 \times 4 - 1)(3 - 1) = 30$; $t^2 = 4.171$ for $n = 30$ and $P = 0.05$ from a table of Student's t; and $C = 2417.35/(2417.35 - 1.784 \times 4.171) = 1.00309$ (equation 15). Alternative if not randomized: $[T_t^2] = 134860.75 - 132424.73 - 2417.35 - 0.05 = 18.62$ (equation 13), where $\Sigma T_p^2/kf = (601.5^2 + 678.0^2 + 614.5^2 + 624.5^2)/4 \times 3 = 132424.73$ and $n = 4(4 - 1) - 2 = 10$; $s_t^2 = 18.62/10 = 1.862$ (equation 14).

Test for nonparallelism ($k = 4$): $[B^2] = 2425.43 - 2417.35 = 8.08$ (equation 18), where $\Sigma T_b'^2/f\Sigma x_1^2 = (1849.5^2 + 1973.5^2 + 1730.0^2 + 2006.5^2)/3 \times 1970 = 2425.43$; $F = 8.08/3 \times 1.784 = 1.51$ with $n_1 = 3$ and $n_2 = 30$. Alternative if not randomized: $F = \{4(4 - 2) - 1\}8.0/3(18.62 - 8.08) = 1.79$ with $n_1 = 3$ and $n_2 = 7$.

Calculation in last 2 rows for 2, 3, and 4 ml only ($k = 3$): $T_b = 4117.5$; $T_a = 538.0 - 479.5 = 58.5$ for U_1, 8.0 for U_2, 17.5 for U_3; $M' = 2.6847 \times 4T_a/4117.5 = 0.002608T_a$ (equation 10); $s^2 = 44.56/22 = 2.025$ (equation 12); $t^2 = 4.301$ from table of Student's t; $B^2 = 4117.5^2/3 \times 4 \times 1418 = 996.34$ (equation 11); $C = 1.00882$ (equation 15); $L^2/4 = 0.00882(1.00882M'^2 + 0.12199)$ (equation 16).

tions. There is a similar gain in measuring the effect of single or quadratic curvature (Q^2), if it is required. They are computed as

$$B^2 = T_b^2/fh'\sum x_1^2 \tag{11}$$

and

$$Q^2 = T_q^2/fh'\sum x_2^2$$

where f is the number of replicate responses in each T_t.

In a randomized assay, the error variance s^2 is computed from the variation among the N individual responses y (or z), segregating the differences due to the totals for sets or racks T_r and for treatments T_t to obtain

$$s^2 = \{\sum y^2 - \sum T_r^2/kh' - \sum T_t^2/f + T^2/N\}/n \tag{12}$$

where $T = \sum T_r = \sum T_t$ and the degrees of freedom in s^2, $n = (kh' - 1)(f - 1)$, is diminished by 1 for every gap in the original table which has been filled by computation.

In all systematic assays, and in some that are randomized, the error variance is based upon the scatter of the treatment means rather than of the individual responses. The variation of the T_t's for each treatment around a series of parallel parabolas, in the same units of y^2 as equation 12, is computed as

$$[T_t^2] = \sum T_t^2/f - \sum T_P^2/kf - B^2 - Q^2 \tag{13}$$

leading to the alternate error variance of

$$s_t^2 = [T_t^2]/n \tag{14}$$

where $n = h'(k - 1) - 2$ degrees of freedom. If the ratio $F = s_t^2/s^2$ is considerably larger than

1, s_t^2 is a safer estimate than s^2 for the assay variance of a randomized assay. The error variance of a systematic assay is necessarily s_t^2.

Given B^2 from equation 11 and s^2 (or s_t^2), the slope factor C is computed as

$$C = B^2/(B^2 - s^2t^2) \qquad (15)$$

where t^2 depends upon the degrees of freedom in s^2 (or in s_t^2). The more nearly C approaches unity, the more precise is the assay. A microbial assay with $C > 2$ or C negative is of little or no value.

For the assay in table 5, the variation in y attributable to the assay slope is $B^2 = 7559.5^2/3 \times 4 \times 1970 = 2417.35$ and that due to curvature is $Q^2 = 1.5^2/3 \times 4 \times 4 = 0.047$ (equation 11). Since the tubes were randomized in the $f = 3$ racks, the error variance is computed from the y's in table 4, omitting the responses at doses of 1 and 5 ml. Substituting in equation 12, $s^2 = 53.53/30 = 1.784$ with 30 degrees of freedom. With $t^2 = 4.171$ for 30 degrees of freedom, $C = 2417.35/(2417.35 - 1.784 \times 4.171) = 1.00309$.

If the assay had not been randomized, its error variance would be based instead upon the variation in the T_i's about four parallel parabolas, giving by equation 13, $[T_i^2] = 18.62$, and by equation 14, $s_t^2 = 18.62/10 = 1.862$. The ratio, $F = 1.862/1.784 = 1.04$ is here so near unity that C would be computed with s^2 as above. If the assay had not been randomized, it would be computed instead with $s_t^2 = 1.862$ and $t^2 = 4.965$ for 10 degrees of freedom to obtain $C = 1.00384$.

The precision of M'. The confidence limits for factorial M' avoid the approximations inherent in those for an interpolated estimate of the log-potency. In 19 assays out of 20 they are expected to bracket the true potency of the Unknown, provided there are no interassay variance components. The confidence interval between these limits is a function of three terms, the log-relative potency M', the statistic C defined in the preceding section, and a constant determined by the design of the experiment and not by its outcome (Bliss, 1956*a*). This constant is the ratio of two variances, that for the numerator of M' divided by that for its denominator. It may be designated as $\frac{1}{2}c'i'^2h'$. For each sequence of k doses in table 3, $\frac{1}{2}c'i'^2 = 2i'^2\sum x_1^2/k$ as listed in the last row of the table.

Given the above terms, the confidence interval for each Unknown depends upon how closely its assayed potency confirms its assumed potency, *i.e.*, upon how nearly M' approaches 0. Each interval is computed as

$$L = 2\sqrt{(C - 1)(CM'^2 + \tfrac{1}{2}c'i'^2h')} \qquad (16)$$

In order to retain three-figure accuracy in the confidence interval, C and M' should be computed to five decimal places. M' may later be rounded to three or four decimals. The log-relative potency for each Unknown then has confidence limits of

$$X_{M'} = CM' \pm \tfrac{1}{2}L \qquad (17)$$

In terms of vitamin content, the upper and lower confidence limits are $R(\text{antilog } X_{M'})$, where R is defined as in equation 7.

In addition to $C = 1.00309$ and M' from table 5, the other term required for the confidence interval in our numerical example is $\frac{1}{2}c'i'^2h'$. For the present dosage sequence, $\frac{1}{2}c'i'^2 = 0.053115$ from table 3, and with $h' = 4$ preparations, the constant terms 0.21246. Substituting the M' for U_1 in equation 16,

$$\tfrac{1}{2}L = \sqrt{0.00309(1.00309 \times 0.14640^2 + 0.21246)} = 0.0282.$$

The half-intervals for the three Unknowns in table 5 are all somewhat smaller than the single half-interval of 0.0285 obtained with equation 9. If the assay had not been randomized, the half-intervals would have been determined instead with $C = 1.00384$ from the variation in the T_i's, giving larger values of $\frac{1}{2}L = 0.0300$, 0.0286, and 0.0287 for U_1 to U_3, respectively. To transform the confidence limits in logarithms to content of vitamin B_{12}, $X_{M'}$ is determined with equation 17; for U_1, $X_{M'} = 0.1464 \times 1.00309 \pm 0.0282 = 0.1751$ and 0.1187. As noted earlier, the assumed potency of U_1 was $R = 2.4$ μg per ml, from which the two limits are 2.4(antilog 0.1751 and 0.1187) = 3.59 and 3.15 μg per ml, bracketing the most probable value of 3.362 μg per ml.

Statistical test for curvature. The factorial calculation of potency is based upon an assumed linear relation between the response and the log-dose within the observed range of doses. A simple comparison of the signs of T_q', as described above, is usually adequate. If all values of T_q' are of the same sign, however, the variation attributable to combined curvature, as measured by Q^2 (equation 11), can be tested for signifi-

cance by computing $t^2 = Q^2/s^2$ or Q^2/s_t^2, with n equal to the degrees of freedom in s^2 or in s_t^2. In a randomized assay, the test is definitive, but if the tubes have been placed systematically, an apparently significant curvature, tested necessarily with s_t^2 rather than s^2, may be due to their arrangement in the racks rather than to a nonlinear relation of the response to the log-dose. In either case, a clearly significant Q^2 may be small enough, relative to B^2, that it can be disregarded and the assay analyzed as if the lines were straight. Empirical studies of the vitamin B_{12} assay (Bliss, 1956*b*, 1957) suggest that if $B^2/Q^2 > 100$, we can proceed as if the relation were linear.

When the assumption of linearity is adopted, should significant average curvature in the log-dose response curve be included or omitted in estimating the assay variance? If the true relation is parabolic, the variability measured by Q^2 is presumably common to all curves and would have little influence upon the internal error of the assay. If the tubes are arranged systematically, however, part of the assay error may appear as combined curvature, and the omission of Q^2 could lead to an underestimate of the variance. In the absence of randomization, a conservative rule is to modify equation 13 by not subtracting Q^2 in computing $[T_t^2]$ and s_t^2.

In the randomized assay in table 3, the combined curvature $Q^2 = 0.047$ is far less than the error variance $s^2 = 1.784$ and still smaller than $B^2 = 2417.35$. The relation is clearly linear.

Statistical test for nonparallelism. The agreement of preparations in slope can be tested critically, if the tubes have been placed in each rack at random. If an Unknown and the Standard were to differ significantly and materially in slope, the same relative potency would not hold at different levels of response. However, if the tubes have been arranged systematically, an apparent discrepancy in slope may be due to differences in the positions of the tubes, and the test for parallelism is no longer clean-cut.

The total variation due to differences in slope between the $h' = h + 1$ preparations is measured by

$$[B^2] = \sum T_b'^2/f \sum x_1^2 - B^2 \qquad (18)$$

with h degrees of freedom. In a randomized assay, any divergence in slope is tested by referring the variance ratio, $F = [B^2]/hs^2$, with $n_1 = h$ and $n_2 = n$ for s^2, to any standard table of F (Fisher and Yates, 1953). If it exceeds the tabular value materially at the 5 per cent point ($P = 0.05$), the parallelism of the slopes is open to question. Sometimes a single Unknown with a divergent slope can be spotted at a glance and the test repeated with the remaining Unknowns and the Standard to determine whether they are satisfactorily parallel. In doubtful cases, the test of parallelism may be merged in a more general test of assay validity by computing $F = s_t^2/s^2$, with n_1 and n_2 equal to the degrees of freedom (n) in s_t^2 and s^2, respectively. If F is not significantly larger than the tabular value at the 5 per cent level, the assay can be accepted as valid, even though an F that is considerably larger than 1 may lead to adopting s_t^2 as the assay variance.

In a systematic assay, the test for lack of parallelism includes possible discrepancies due to tube position. Within this limitation, we may compute

$$F = \{h'(k-2) - 1\}[B^2]/h([T_t^2] - [B^2]) \qquad (19)$$

with $n_1 = h$ and $n_2 = h'(k-2) - 1$ degrees of freedom. If F exceeds the tabular value for $P = 0.05$, the parallelism of the log-dose response lines for the h' preparations is in doubt.

The range in T_b' for the example in table 3 suggests satisfactory parallelism, but since the assay was randomized, this can be tested critically. In units of y^2, the sum of squares for differences in slope is $[B^2] = 2425.43 - B^2 = 8.08$ (equation 18), leading to $F = 8.08/3 \times 1.784 = 1.51$ with $n_1 = 3$ and $n_2 = 30$ degrees of freedom. This does not approach the 5 per cent level (2.92) for significant nonparallelism. If the assay had not been randomized, its error variance would have been based instead upon the variation in the T_t's (equation 13). For the corresponding test of parallelism by equation 19, $F = 7 \times 8.08/3(18.62 - 8.08) = 1.79$, with $n_1 = 3$ and $n_2 = 7$ degrees of freedom, the critical level at $P = 0.05$ being 4.35. Because of the randomization, the test here is not confounded with tube position but is less sensitive than in the preceding case.

CALCULATING INCOMPLETE ASSAYS

The initial balance of an assay may be destroyed by losses during the test or through the later rejection of an aberrant response. In other

assays, test solutions that are too concentrated or too dilute, because of a poor guess as to the potency of one or more Unknowns, may lead to an end response outside the range of the straight log-dose response line. The balance of an incomplete assay can usually be restored so that its calculation is changed but little. An end dosage level may be dropped or replacements may be computed to fill the one or more gaps.

Dropping an end dosage level. An assay initially with four or more dosage levels may still give quite respectable estimates of log-potency even though an end dose is eliminated. The assay is then computed as described above, but with coefficients and constants for the curtailed sequence of doses. This may be the simplest expedient for obtaining a linear log-dose relation, or unavoidable if there are many gaps at an end dosage level. Although the computed log potencies are unbiased, their precision is less than if the original complement of doses could have been retained.

To show the effect of basing an assay upon three instead of four dosage levels, the example in table 5 has been recomputed without the responses at 1.5 ml per tube. The coefficients x_1 and x_2, and other constants for the curtailed sequence of 2, 3, and 4 ml per tube, are given in table 3. From $T_b = 4117.5$ and new totals T_P and differences T_a, the recomputed values of $M' = 0.002608T_a$ are listed in the next-to-last row of the table. Their reduced precision is reflected in a larger $C = 1.00882$ and in half-confidence intervals ($\frac{1}{2}L$) in the last row that are 22 to 27 per cent wider than when computed from the original four-dosage levels.

Log potencies from assays with replacements. In order to retain valid observations and to simplify the analysis, each gap may be filled with a number which will minimize the error variance (Fisher, 1951). The replacement for a single lost tube is computed by the well-known equation

$$y' = \frac{fT_r' + kT_t' - T'}{(f-1)(k-1)} \qquad (20)$$

where f is the number of sets or tube racks, k is the number of treatments or doses, and T_r', T_t' and T' are the incomplete totals for the randomized set, treatment and assay from which an observation is missing. If more than one y is missing, the treatment mean is substituted temporarily in all empty places but one, for which y' is computed by equation 20. Each initial substitution is replaced in turn by a computed y', and the process repeated in successive approximations until a stable y' is obtained for each missing observation.

Occasionally an entire treatment total (T_t) may be replaced by the same equation. The remaining T_t's are arranged by dosage levels and preparations as in table 5 or 6. The missing total is then computed as

$$T_t' = \frac{h'T_P' + kT_d' - T'}{(h'-1)(k-1)} \qquad (21)$$

where T_P', T_d' and T' are the incomplete totals for the preparation, dosage level and assay from which the treatment total is missing, and there are h' preparations and k dosage levels.

An unbiased log-relative potency M' is computed from the completed set of readings for each Unknown, exactly as if there had been no replacement. Its precision, however, is less than if no observations had been lost, as shown in the next section.

The calculation of the log-relative potency (M') with a substitute value may be illustrated by a turbidimetric assay of calcium pantothenate with duplicate tubes ($f = 2$). Since the tubes were not randomized, only the T_t's are given in table 6. The assumed potency of one Unknown (U_2) was underestimated, so that at a dose of 5 ml per tube its observed response of $T_t = 164$ fell outside the linear range of the log-dose response curve. If the log-relative potencies were estimated from equation 6 for the Standard curve, there would be only three values of X' for U_2 but four for each of the remaining Unknowns.

For a factorial estimate, the original T_t for U_2 at 5 ml per tube is replaced by equation 21 with

$$T_t' = (5 \times 428 + 4 \times 571 - 2361)/(5-1)\ (4-1) = 172.$$

With this one replacement, the calculation is completed as shown in table 6. Since the range in T_b' is only 10 per cent of its smallest value, all preparations could be assumed to have parallel log-dose response lines. Although all values of T_q' are negative, the ratio of $B^2/Q^2 = 3584.09/10.95 = 327$ (equation 11) does not suggest excessive curvature. To determine the log-

TABLE 6

Factorial analysis with a replacement

Dose (Ml per Tube)	Coefficients		Total Response T_t For:					Total T_d
	x_1	x_2	S	U_1	U_2	U_3	U_4	
2	−61	28	103	56	123	97	110	489
3	−12	−41	126	86	145	124	134	615
4	23	−21	143	96	160	139	148	686
5	50	34	154	109	172*	150	158	743
$T_P = \Sigma T_t$			526	347	600	510	550	2533
$T_b' = \Sigma x_1 T_t$			3194	3210	3037	3292	2986	15719
$T_q' = \Sigma x_2 T_t$			−49	−268	−13	−187	−150	−667
$T_a = T_U' - T_S'$				−179	74	−16	24	
$M' = 0.0019648 T_a$				−0.35170	0.14540	−0.03144	0.04716	
$\frac{1}{2}L$ ($C = 1.00316$)				0.03399	0.03094	0.02768	0.02775	

Turbidimetric calcium pantothenate assay from $T_t = \Sigma y$ for $f = 2$ replicates, $h' = 5$, $k = 4$; tubes not randomized.

* Replaced end value: 5 ml of U_2 with observed $T_t = 164$ fell outside linear range of log-doses reponse curve and response was omitted. Replacement $T_t' = (5 \times 428 + 4 \times 571 - 2361)/(5 - 1)(4 - 1) = 172$, where $T_P' = 123 + 145 + 160 = 428$, $T_d' = 154 + 109 + 150 + 158 = 571$ and $T'' = 489 + 615 + 686 + 571 = 2361$. (equation 21).

Assay precision: $B^2 = 15719^2/2 \times 5 \times 6894 = 3584.09$ and $Q^2 = (-667)^2/2 \times 5 \times 4062 = 10.95$ (equation 11); $[T_t^2] = 168573.5 - 164960.62 - 3584.09 = 28.79$ (equation 13 + Q^2), $s_t^2 = 28.79/13 = 2.215$, where $n = 5(4 - 1) - 1 - 1 = 13$; $C = 3584.09/(3584.09 - 1.09066 \times 2.215 \times 4.667) = 1.00316$ (equation 23), where $t^2 = 4.667$ from table of Student's t, $\Sigma(x_1^2/f_b) = (61^2 + 12^2 + 23^2)/10 + 50^2/8 = 751.9$, and $j = 2 \times 5 \times 751.9/6894 = 1.09066$ (equation 22).

Confidence interval: $j' = (2 \times 4 \times 2 - 2)j/2(4 \times 2 - 2) = 14 \times 1.09066/12 = 1.27244$ (equation 24) with $r = 2$; $(\frac{1}{2}c'i^2)h'j = 0.044274 \times 5 \times 1.09066 = 0.24144$ (table 3), $L^2/4 = 0.00316$ $(1.00316M'^2 + 0.24144)$ for U_1, U_3 and U_4 (adjusted equation 16); $(\frac{1}{2}c'i^2)h'j' = 0.044274 \times 5 \times 1.27244 = 0.28168$, $L^2/4 = 0.00316(1.00316M'^2 + 0.28168)$ for U_2 (adjusted equation 16).

relative potencies, we read $\frac{1}{2}ci = 6.1769$ from table 3, and, with $h' = 5$ and $T_b = 15719$, substitute in equation 10 to obtain $M' = 0.0019648$ T_a for each Unknown, just as in a complete assay with no replacements.

The precision of assays with replacements. If each log-relative potency is computed from the equation for the Standard, its approximate error variance (s^2) and confidence interval (L) are given by equations 8 and 9. In equation 9, however, the number (k) of individual estimates of X' will be less for the Unknown with a missing T_t, and, with a smaller k, it will have a longer confidence interval than the other Unknowns.

When M' is computed factorially, a replacement requires several adjustments in computing the confidence intervals. The degrees of freedom (n) in the denominator of s^2 is reduced by one for each y that is replaced. If a T_t is replaced with equation 21, s_t^2 has similarly one less degree of freedom in the denominator, and s^2, if computed, loses f degrees of freedom. Because of the reduction in n, t^2 will be larger. A further adjusting term j is computed as

$$j = fh'\sum(x_1^2/f_b)/\sum x_1^2 \qquad (22)$$

where f_b is the number of observed responses y in the T_d for each x_1. To allow for the replacement, the slope factor in equation 15 is computed as

$$C = B^2/(B^2 - js^2t^2) \qquad (23)$$

If there is no replacement in T_a, the last term in equation 16 for the confidence interval, $\frac{1}{2}c'i^2h'$, is also multiplied by j. However, if T_a is computed with a replacement, $\frac{1}{2}c'i^2h'$ is multiplied instead by

$$j' = \frac{(2kf - r)j}{2(kf - r)} \qquad (24)$$

where r is the number of replaced y's.

Since the assay in table 6 was not randomized,

the more conservative course is not to subtract Q^2 (= 10.95) when computing $[T_t^2]$ = 28.79 by equation 13. The error variance is then determined by equation 14 as $s_t^2 = 28.79/13 = 2.215$ with $n = 13$ degrees of freedom and $t^2 = 4.667$. The adjusting term from equation 22 is $j = 2 \times 5 \times 751.9/6894 = 1.09066$, leading to $C = 3584.09/(3584.09 - 1.09066 \times 2.215 \times 4.667) = 1.00316$ by equation 23. For Unknowns U_1, U_3 and U_4, none involving the replacement, the constant term in computing the confidence interval L by equation 16 is $\frac{1}{2}c'i^2h'j = .044274 \times 5 \times 1.09066 = 0.24144$. For U_2, however, where T_t' replaces $r = f = 2$ values of y, $j' = 14j/12 = 7j/6$ from equation 24, and we have instead $\frac{1}{2}c'i^2h'j' = 0.28168$. The half-confidence intervals in the last row of table 6 have been determined with these values.

SUMMARY

Four practices which increase the precision and reliability of many microbial assays are: (1) selecting dosage levels and a function of the growth response that lead to a straight log-dose response line; (2) randomizing the location of culture tubes within racks during sterilization, inoculation, and incubation; (3) following an objective probability rule in identifying and rejecting supposedly aberrant observations; and (4) basing assayed potencies upon the means of two or more independent assays until experience has shown that duplicate runs agree routinely within the limits determined by their internal errors. To take advantage of the greater inherent reliability of assays based upon these designs, two calculations of potency and its precision are described. Each is illustrated numerically.

The first calculation is the arithmetic equivalent (with extensions) of the traditional interpolation in a plotted Standard curve. Methods are described for computing the equation of the Standard curve with coded log-doses, and for testing the importance of any apparent curvature with the aid of orthogonal polynomials. Solving this equation in reverse converts the response at each dosage level of each Unknown to its indicated log potency. Approximate confidence limits are defined for the mean log potency of each Unknown.

The second calculation by factorial analysis utilizes more fully and efficiently the information inherent in an assay. With the aid of factorial coefficients for the selected dosage sequence, little computation is required to determine the log-relative potency for each Unknown by methods that have long been used in animal assays. The inherent precision of an assay depends upon the reliability of the assay slope, which is measured by the statistic C, a number slightly larger than 1.0. Exact confidence limits defining the precision of each assayed potency follow directly from C. The validity of critical and borderline assays can be checked by statistical tests for curvature and for nonparallelism to complete the factorial analysis.

Since discrepancies may occur, or occasional tubes give aberrant readings or be lost, the last section concerns the calculation of incomplete assays. In some cases an end dose may be omitted. More often, replacements are determined with which the log potencies may be computed factorially. These are not as precise as in complete assays, so that adjustments are described for measuring their precision.

ACKNOWLEDGMENTS

I am indebted to Dr. J. A. Campbell of the Food and Drug Directorate in the Department of National Health and Welfare of Canada, and to Dr. O. L. Kline of the Food and Drug Administration in the U. S. Department of Health, Education and Welfare, for the numerical examples used in the present paper. These data were submitted originally as part of a U.S.P. collaborative study of microbial assay procedure.

REFERENCES

BLISS, C. I. 1952 *The statistics of bioassay with special reference to the vitamins.* Academic Press, Inc., New York, N. Y.

BLISS, C. I. 1956*a* Confidence limits for measuring the precision of bioassays. Biometrics, **12**, *in press*.

BLISS, C. I. 1956*b* The precision of microbial assays with special reference to vitamin B_{12}. J. Assoc. Office Agr. Chemists, **39**, 816–834.

BLISS, C. I. 1956*c* Analysis of the biological assays in U.S.P. XV. Drug Standards, **24**, 33–68.

BLISS, C. I. 1957 Bioassay from a parabola. Biometrics, **13**, *in press*.

BLISS, C. I., AND CALHOUN, D. W. 1954 *An outline of biometry.* Yale Co-Operative Corp., New Haven, Conn.

BROWNLEE, K. A. 1952 A statistical analysis of the 1950 collaborative study on the U.S.P. method for vitamin B_{12} assay. Drug Standards, **20**, 48–56.

BROWNLEE, K. A., AND LAPEDES, D. N. 1951 The effects of design upon the error of a microbiological assay for vitamin B_{12}. J. Bacteriol., **62**, 433–444.

CAMPBELL, J. A., MCLAUGHLAN, J. M., CLARK, J. A., AND DUNNETT, C. W. 1953 The six-point design in the U.S.P. microbiological assay of vitamin B_{12}. J. Am. Pharmac. Assoc., Sci. Ed., **42**, 276–283.

DIXON, W. J. 1951 Ratios involving extreme values. Ann. Math. Stat., **22**, 68–78.

FINNEY, D. J. 1952 *Statistical method in biological assay*. Chas. Griffin and Co., London, England.

FISHER, R. A. 1951 *The design of experiments*. 6th Ed. Oliver and Boyd, Edinburgh, Scotland.

FISHER, R. A., AND YATES, F. 1953 *Statistical tables for biological, agricultural and medical research*. 4th Ed. Hafner Publishing Co., Inc., New York, N. Y.

U. S. Pharmacopoeia 1955 15th Revision. Mack Publishing Co., Easton, Pa.