NUMBER OF ANIMALS REQUIRED IN THE BIO-ASSAY OF PATHOGENS

IRA A. DEARMON, JR., AND RALPH E. LINCOLN

U. S. Army Chemical Corps, Fort Detrick, Frederick, Maryland

Received for publication April 20, 1959

Experimental designs which prescribe either too few or too many animals per assay are wasteful, and it is important that experimenters in planning bio-assays use the number of animals consistent with the statistical precision actually desired in the results. The general philosophical treatment of experimental design, as noted by Fisher (1949), places controls on treatments, treatment combinations, randomization, and replication. Wilson and Miles (1955), in their application of statistics to bacteriology, accept these principles, as do numerous other writers. Experimental designs in the field of bio-assay, as treated by Finney (1952*a*, *b*), Bliss (1952), and Perry (1950), are somewhat specialized due to the use of two or more dose levels and of a standard preparation. A satisfactory experimental design for bio-assays estimates a mean effect, a slope, the degree of slope linearity, and some measure of error. To meet this bio-assay requirement at least 3 doses of "standard" and 3 doses of "unknown" culture must be used, as well as a number of animals for each separate dose. Therefore, treatment arrangements including dose, randomization, and total number of animals used are all essential parts of the bio-assay.

The results of a bio-assay express the "unknown" in terms of the "standard" by a ratio (called potency ratio) of dosages that manifest equal response from the host. The reliability of this ratio depends upon the length of its confidence interval (Fieller, 1940), a statistic the size of which is dependent upon assay variance, and the total number of animals per assay. The length of the confidence interval of the estimated parameter, therefore, is an objective criterion of the success of the bio-assay, and usually is proportional to the parameter value itself. The number of animals required in a bio-assay to obtain the desired or essential precision must be determined if efficient experimentation is to be conducted.

In a collaborative paper evaluating uniformity of response in different strains of animals to pathogenic organisms, we (Lincoln and DeArmon, 1959) determined the number of animals of each strain required to obtain a selected bio-assay precision. This paper derives simple equations useful in estimating the approximate animal requirements for bio-assays generally used in the study of organism pathogenicity. The animal requirements depend primarily upon the desired confidence interval the experimenter wishes to place on a LD_{50}, or the potency ratio as determined either by the quantal or graded response assays. The equations are derived independent of number of doses used and will estimate the approximate total animal requirements per assay; therefore, this number should be proportioned equally among the doses used.

ESTIMATION OF ANIMAL REQUIREMENTS IN BIO-ASSAY

For determining LD_{50} from the quantal response. According to Finney (1952*a*) the $\log_{10} LD_{50}$, m, is defined as:

$$m = \bar{x} + \frac{5.0 - \bar{y}}{B} \qquad (1)$$

with an approximate variance defined as:

$$V(m) = \frac{1}{B^2}\left(\frac{1}{\sum nw}\right) \qquad (2)$$

Where $\bar{x}$ is mean $\log_{10}$ weighted dose, $\bar{y}$ is mean weighted probit response and B is slope of the dose response curve. $\sum nw$ is the sum of the number of animals, n, tested at each dose multiplied by a weighting coefficient, w, which can be taken as a constant equal to approximately 0.40 (*unpublished data*).

Since the variance of m is only an approximation, the confidence interval (CI) of m is also an approximation but can be estimated as:

$$m \pm \frac{z}{B}\left(\frac{1}{0.40 \sum n}\right)^{1/2}$$

or for total length, L, as:

$$L = \frac{2z}{B}\left(\frac{2.5}{\sum n}\right)^{1/2} \tag{3}$$

Since $\sum n$ is N, the total animals in the experiment, then equation (3) can be expressed in terms of N as:

$$N = \frac{10z^2}{L^2 \cdot B^2} \quad \text{or} \quad \frac{38}{L^2 \cdot B^2} \tag{4}$$

where L is the length of the 95 per cent confidence interval of m, and z is taken as the normal deviate equal to 1.96.

With any given L, as chosen by the experimenter, equation (4) is useful in planning for the total number of animals, N, required for any LD_{50} determination. The calculation may be reversed, when an animal pool of a definite size is available, in order to determine whether the probable results would justify use of the animals. At best, estimations of N are only approximations. The accuracy of an estimate depends on the assumptions that (a) a homogeneous assay can be conducted, (b) the approximate slope of the dose-response curve can be estimated, and (c) the responses of animals will be somewhat symmetrical around the LD_{50}.

In our experience, when mice and guinea pigs are challenged by the intraperitoneal route by *Bacillus anthracis* or *Pasteurella tularensis*, the probit slope, B, of the dose-response curves has an average value of approximately 1.0.

Table 1 indicates the approximate number of animals required for the LD_{50} determination as computed from equation (4) when L and B have varying values. The dose response slope, B, for most bacterial infections is less than 2.0, (Meynell 1957), and therefore the computations shown in table 1 are generally useful. Column 1 of table 1 presents the CI of the LD_{50} as a proportion of the LD_{50} which might be used in the planning in place of the comparable logarithm.

TABLE 1

Number of animals required for LD_{50} determinations

Desired Length of 95% Confidence Interval as Proportion (p)* of LD_{50}	Length of 95% Confidence Interval of $\log_{10}$ LD_{50} (L)*	Approx No. Animals per Probit Slopes, B:		
		1.0	1.5	2.0
0.5 LD_{50}	0.21	860	380	220
0.8 LD_{50}	0.34	330	150	82
1.0 LD_{50}	0.42	220	96	54
2.0 LD_{50}	0.77	64	29	16

* L can be computed from p by: $L = 2 \log_{10} (p/2 + \sqrt{p^2/4 + 1})$.

For determining potency ratio between two LD_{50}'s. According to Finney (1952*a*), the $\log_{10}$ potency ratio (PR), M, is the difference between m_1 and m_2 where m was defined in equation (1); therefore,

$$M = \bar{x}_s - \bar{x}_u - \frac{(\bar{y}_s - \bar{y}_u)}{B_c} \tag{5}$$

where $\bar{x}_u$ and $\bar{x}_s$ refers to mean weighted $\log_{10}$ dose respectively of the unknown and standard, the $\bar{y}_u$ and $\bar{y}_s$ refers to respective mean weighted responses and B_c is the common slope of the dose-response curves.

From equation (3) it can be shown that the approximate length of the confidence interval of M can be expressed as L, where

$$L = \frac{2z}{B_c}\left(\frac{2.5}{\sum n_s} + \frac{2.5}{\sum n_u}\right)^{1/2} \tag{6}$$

where $\sum n_s$ or N_s and $\sum n_u$ or N_u are the total animals used with standard and unknown respectively.

Assuming N_u equals N_s equals N, then the approximate number of animals to be used with each "unknown" or "standard" can be expressed as:

$$N = \frac{20z^2}{B_c^2 L^2} \quad \text{or} \quad \frac{77}{B_c^2 L^2} \tag{7}$$

where L is the length of the 95 per cent confidence interval of M and z is taken as the normal deviate equal to 1.96.

It can be noted that in equation (7) N is just twice that of the N as determined by equation (4).

For equation (7), N is a function of the slope of the dose-response curve, B_c, and the length of the confidence interval, L. When the slope is known, the confidence interval of the PR is dependent on the number of animals used for each LD_{50} value of the unknown and of the standard; therefore the animal requirements can be estimated for any desired confidence interval.

A summary of animal requirements is given in table 2 for slopes of 1.0, 1.5, and 2.0, and for various lengths of the confidence interval of the PR. The accuracy of the approximation of

TABLE 2

Number of animals required for each standard and unknown material in determining potency ratio by quantal response assay method

Desired Length of 95% Confidence Interval as Proportion (p)* of the Potency Ratio (PR)	Length of the 95% Confidence Interval of Log Potency Ratio (L)*	Approx No. Animals per Probit Slope B_0:		
		1.0	1.5	2.0
0.5 PR	0.21	1700	760	440
0.8 PR	0.34	660	300	160
1.0 PR	0.42	440	200	110
2.0 PR	0.77	130	58	32

* $L = 2 \log_{10} (p/2 + \sqrt{p^2/4 + 1})$.

animal requirement as given by equation (7) depends on how well the following assumptions are met: (a) that a homogeneous assay can be conducted; (b) that both unknown and standard have parallel dose-response slopes; (c) that a common slope can be approximately estimated before the assay is conducted; and (d) that responses of animals will be reasonably symmetrical about their respective LD_{50}'s.

For estimating potency ratio of unknown and standard by the graded response. In the past, the graded response has been adapted to various assay uses with infectious microorganisms (Ipsen (1944), Pike and Mackenzie (1940), Gard (1940), and Brownlee and Hamre (1951)) and recently there appears to be a renewed interest in the method (Roth *et al.* (1956), Fernelius *et al.* (1958), and L'Heritier (1958)). The assay envisaged here is a bio-assay contrasting an "unknown" material to a "standard" by use of a graded response variable.

The graded response assay method relies on the classical regression theory and in deriving an equation to estimate the number of animals required in the assay, the estimate of the $\log_{10}$ PR or M as given by Finney (1952*b*) may be written as:

$$M = \bar{x}_s - \bar{x}_u - \frac{(\bar{y}_s - \bar{y}_u)}{b} \qquad (8)$$

with approximate variance of M as:

$$V(M) = s^2/b^2 \left(\frac{1}{N_s} + \frac{1}{N_u}\right) \qquad (9)$$

where $\bar{x}_s$ and $\bar{x}_u$ are, respectively, the mean $\log_{10}$ doses of standard and unknown and $\bar{y}_s$ and $\bar{y}_u$ are the corresponding respective responses, b is the pooled linear slope of the dose-response curve, and s is the pooled standard deviation per response unit. N_s and N_u are the respective number of animals used with "standard" and "unknown."

The third term in the variance estimate as given by Finney (1952*b*) is without practical value and was omitted above; therefore, the confidence interval of M can be written as:

$$M \pm ts/b \left(\frac{1}{N_s} + \frac{1}{N_u}\right)^{1/2} \qquad (10)$$

where t is the "t" statistics and varies with respect to the degrees of freedom and the given probability entering with the estimates of s^2.

If L is the total length of the confidence interval, then

$$L = \frac{2ts}{b} \left(\frac{1}{N_s} + \frac{1}{N_u}\right)^{1/2} \qquad (11)$$

In performing an assay, the number of animals assigned to the standard material is usually the same as that assigned to the unknown material; therefore, assume N_s equals N_u. Also, after the terminology of Finney (1952*b*), the ratio of s/b can be replaced by λ, and since $t = 2.0$ for a 95 per cent confidence interval, providing the total animals in both assays exceeds 30, equation (11) can be written as:

$$N = \frac{8\lambda^2 t^2}{L^2} \quad \text{or} \quad N = \frac{32\lambda^2}{L^2} \qquad (12)$$

N is therefore defined as the approximate number of animals to be used with each "standard" and "unknown" to obtain a 95 per cent confidence interval of length L.

The only assumption involved in the accuracy of determining N is that λ is approximately known; therefore, the standard deviation per animal, s, and the slope of the dose-response curve, b, must be approximately known. Approximations of values for λ have appeared in various reports in the literature and a number of reference values for λ, as well as s and b, have been listed in table 3. From these references, the λ values of infectious materials range from 0.15 to 0.76, whereas the λ values for toxic materials range from 0.07 to 0.20. Thus, there is an indication that toxic materials have smaller values for λ than do the infectious materials.

The approximate number of animals required

TABLE 3

Various values for λ observed in animal assays for infectious and toxic materials

Type Assay	Experimental Animal	Time to Death Response	Route of Challenge*	s	$-b$	λ	Reference
Bacillus anthracis	Noninbred mice	Log-hr	IP	0.07	0.09	0.76	Lincoln and DeArmon (1959)
	CAF_1 inbred mice	Log-hr	IP	0.04	0.09	0.49	Lincoln and DeArmon (1959)
Pasteurella tularensis	Noninbred mice	Log-hr	IP	0.05	0.17	0.24	Lincoln and DeArmon (1959)
	BALB inbred mice	Log-hr	IP	0.03	0.15	0.15	Lincoln and DeArmon (1959)
Neoarsphenamine	Mice	Log-min (coded)	IV	218	3168	0.07	Perry (1950)
Phenol, cresols	Goldfish	Log-min	—	0.30	1.91	0.16	Bliss and Cattell (1943)
Sodium arsenate	Silkworm larva	Log-min	—	0.10	0.49	0.20	Bliss and Cattell (1943)

* IP = intraperitoneally; IV = intravenously; — = unknown.

TABLE 4

Number of animals required for each standard and unknown material in determining the potency ratio by graded response assay method

Desired Length of the 95% Confidence Interval as Proportion (p)* of the Potency Ratio (PR)	Length of the 95% Confidence Interval of the $\log_{10}$ PR (L)*	No. Animals Required per λ:		
		0.75	0.50	0.25
0.5 PR	0.21	409	182	46
0.8 PR	0.34	156	70	17
1.0 PR	0.42	103	46	11
2.0 PR	0.77	31	14	—†

*$L = 2 \log_{10} (p/2 + \sqrt{p^2/4 + 1})$.

† Animal requirements less than 10 not listed.

for bio-assay using various values for λ and various lengths of the confidence interval of the PR have been calculated and presented in table 4. Once the approximate animal requirement for the assay has been determined, it must be decided by the experimenter whether to use a (1 + 1),[1] (2 + 2), (3 + 3), or a more complex dose assay (Perry (1950)). It is of interest to contrast columns 3 of tables 2 and 4 which indicate that for equal precision the quantal response assay requires about 4-fold more animals than does the graded response assay.

[1] One dose of unknown + one dose of standard.

EXAMPLES

Roth *et al.* (1956) presented quantal and graded (survival time) response data for mice challenged by two strains of *Bacillus anthracis*. Typical bio-assays between Pasteur no. 2 strain and a Vollum strain can be illustrated with their data in conjunction with all three prediction equations that have been derived.

To illustrate the two prediction equations (4) and (7), the quantal response data in replicate no. 1, as given by Roth *et al.* (1956) is used; calculations based on their data are shown in table 5. The LD_{50} of the Vollum strain of *B. anthracis* was 24 spores with a 95 per cent confidence interval of 13 to 46 spores or 33 organisms. The confidence interval as a proportion of the LD_{50} was 1.35 LD_{50}'s; therefore, the length of the confidence interval in log units was 0.55.[2] By using equation (4) the approximate number of animals necessary to obtain this confidence interval when the slope of the dose-response curve is about 1.05, is computed to be 114. This corresponds approximately to 116 animals actually used in determining the LD_{50}.

For the PR between strains of *B. anthracis*, the Pasteur no. 2 strain (LD_{50} of 15.5×10^3

$$^2L = 2 \log \left(\frac{1.35}{2} + \sqrt{\frac{(1.35)^2}{4} + 1.0} \right) = 0.55.$$

TABLE 5

LD_{50}'s and potency ratio in mice of two strains of Bacillus anthracis from replication 1 experiments of Roth et al. (1956)

Pasteur No. 2 Strain		Vollum Strain	
Spores per mouse	Dead/challenged	Spores per mouse	Dead/challenged
95,000	26/30	860	28/30
9,500	9/30	86	21/30
950	3/30	8.6	9/26
95	1/30	0.86	2/30
Slope, B.....	1.05	0.99	
LD_{50} (95% CL).	15500 (8320 to 29200)	24 (13 to 46)	
Length of 95% CI.....	20880	33	
Ratio of CI and LD_{50}.....	1.35	1.38	
Slope, B_c.....	1.01		
Potency ratio (95% CL).....	646 (260 to 1550)		
Length of 95% CI.....	1290		
Ratio of CI and potency ratio.	2.00		

CL = confidence limits; CI = confidence interval.

spores) needed 646 times more spores in mice to produce the same effect as the Vollum strain (LD_{50} of 24 spores). The 95 per cent confidence interval of 260 to 1950 was 1290 which is equivalent to 2.00 times the PR; therefore, the length of the confidence interval (L) in log units is 0.77.[3] By using equation (7), the approximate number of animals necessary to obtain the observed confidence interval when the combined slope of the dose-response curves was 1.01 is computed to be 128 which is a close approximation to the number actually used in the test.

Both examples show by inference that when an approximate slope estimate is available and a certain confidence interval of the LD_{50} or potency ratio of two LD_{50}'s is desired, then the approximate number of animals necessary to gain this confidence interval can be estimated by use of equation (4) or equation (7).

To illustrate the use of equation (12) for estimating animal requirements in a graded bio-assay, data extracted from table 7 of Roth *et al.* (1956) for a (2 + 2) bio-assay are shown in table 6 of this report. For this example, the PR between the strains of *B. anthracis* was estimated to be 143 with a 95 per cent confidence

TABLE 6

Survival time responses of mice for given doses of spores of two strains of Bacillus anthracis from experiments of Roth et al. (1956)

	Pasteur No. 2 Strain		Vollum Strain	
	Spores per mouse*		Spores per mouse*	
	52.6×10^6	0.526×10^6	40.7×10^6	0.407×10^6
	Survival time, hr	Survival time, hr	Survival time, hr	Survival time, hr
	18.0	25.5	10.5	16.5
	19.0	26.0	11.0	17.5
	19.0	26.5	12.5	18.5
	19.5	27.0	12.5	19.0
	19.5	28.5	12.5	20.0
	19.5	30.0	13.0	20.0
	20.5	30.5	13.5	20.5
	21.0	35.0	13.5	21.5
	21.0	38.0	14.5	22.0
	21.5	40.0	15.0	22.0
Geometric mean.....	20.0	30.2	12.9	19.5
Slope b, $\log_{10}$ units.....	0.092		0.093	
s, $\log_{10}$ units..	0.053		0.045	
λ.....	0.53			
Potency ratio (95% CL)...	143 (67 to 305)†			
Length of 95% CI.....	238			
Ratio of CI and potency ratio.....	1.66			

CL = confidence limits; CI = confidence interval.

* Mice were male, starved, weighing 20 to 22 g.

† From equations (8) and (9).

$$^3L = 2 \log \left(\frac{2.04}{2} + \sqrt{\frac{(2.04)^2}{4} + 1.0} \right) = 0.77.$$

interval of 67 to 305. Thus the Pasteur no. 2 strain required 143 (67 to 305) times more spores than the Vollum strain to yield equal response in mice. The 95 per cent confidence interval of the PR was 238 (305 minus 67) which is 1.66 times the PR. Therefore, the length of the confidence interval (L) in log units is 0.66.[4]

Now substituting λ as estimated from the data in table 6 and L equal to 0.66 into equation (12) computed N is equal to 20, the required number of animals per treatment for the bio-assay. In this example N was actually 20; therefore, by inference, the estimates of N would be sufficiently accurate provided λ is approximately known prior to the conduct of the bio-assay. The value of L can be set for any degree of precision at the pleasure of the experimenter.

DISCUSSION

The estimation of animal requirements for a given degree of precision of the LD_{50} is based on knowledge of the approximate LD_{50} and of the slope of the expected dose-response curve. For estimating animal requirements when pathogen or host with unknown response is involved, an approximate working estimate of the LD_{50} and slope could be obtained by conducting a preliminary virulence test, challenging a few animals at each of a wide range of doses spaced at 10-fold dose intervals and observing the responses. In practice, when pathogenic organisms are used, there is evidence that the probit slope will usually fall within the range of about 1.0 to 2.0, although with toxigenic materials the slope may be 5.0 or greater.

Once a preliminary LD_{50} and probit slope estimate have been obtained, the equation (4) in this paper becomes useful in estimating animal requirements for precision estimates of the LD_{50}, whereas equations (7) and (12) are useful without preliminary estimates of parameters since it is assumed that adequate information for planning is available from the "standards."

SUMMARY

The planning of a bio-assay involves knowledge of experimental designs and statistical analyses, and in this report an attempt has been made to extend the usefulness of statistical planning by estimating the approximate animal requirements for any given design. The criterion used for estimating the animal requirements has been taken as the confidence interval of the estimated parameter. Emphasis has been placed on designs in which bio-assay of infectious material is contemplated.

In a quantal response assay, once the desired confidence interval of the estimated LD_{50} or potency ratio is set, the animal requirement is dependent upon the slope of the dose-response curve. Equations have been derived that permit reasonable approximations of animal requirements when three assumptions are met that (a) a homogeneous assay can be conducted, (b) the approximate slope of the dose-response assay is known, and (c) response of animals will be approximately symmetrical about the LD_{50}.

In the graded response assay, once the confidence interval of the potency ratio has been set, the animal requirement is dependent upon λ, the ratio of the per animal standard deviation to the slope of the dose-response curve. An equation has been derived that permits reasonable approximations of animal requirements provided an approximate value of λ is known.

[4] $L = 2\log_{10}\left(\frac{1.66}{2} + \sqrt{\frac{(1.66)^2}{4} + 1.0}\right) = 0.66.$

REFERENCES

BLISS, C. I. 1952 *The statistics of bio-assay.* Academic Press, Inc., New York.

BLISS, C. I. AND CATTELL, M. 1943 Biological assay. Ann. Rev. Physiol., **5,** 479–539.

BROWNLEE, K. A. AND HAMRE, D. 1951 Studies on chemotherapy of vaccinia virus. J. Bacteriol., **61,** 127–134.

FERNELIUS, A. L., DEARMON, I. A., JR., KLEIN, F., AND LINCOLN, R. E. 1958 A comparison of virulence tests for *Bacillus anthracis* spores. Bacteriol. Proc., **1958,** 68.

FIELLER, E. C. 1940 The biological standardization of insulin. J. Roy. Statistical Soc., Suppl. **7,** 1–64.

FINNEY, D. J. 1952*a* *Probit analysis.* Cambridge University Press, London.

FINNEY, D. J. 1952*b* *Statistical methods in biological assay.* Hafner Publishing Co., New York.

FISHER, R. A. 1949 *Design of experiments,* 5th ed. Hafner Publishing Co., New York.

GARD, S. 1940 Encephalomyelitis of mice. II. A method for the measurement of virus activity. J. Exptl. Med., **72,** 69–77.

IPSEN, J. 1944 Quantitative Studien über

Mäuse Pathogene Fleckfieberrickettsien. Schweiz. Z. Pathol. u. Bakteriol., **7,** 129–151.

L'HERITIER, P. 1958 The hereditary virus of drosophilia. Advances in Virus Research, **5,** 195–245.

LINCOLN, R. E. AND DEARMON, I. A., JR. 1959 Homogeneity of response of mouse and guinea pig strains to virulence tests with *Bacillus anthracis* and *Pasteurella tularensis*. J. Bacteriol., **78,** 640–650.

MEYNELL, G. G. 1957 Inherently low precision of infectivity titrations using a quantal response. Biometrics, **13,** 149–163.

PERRY, W. L. M. 1950 Reports on biological standards. VI. The design of toxicity tests. Med. Research Council (Brit.). Spec. Rep. Ser. No. 270.

PIKE, R. M. AND MACKENZIE, G. M. 1940 Virulence of *Salmonella typhimurium*. I. Analysis of experimental infection in mice with strains of high and low virulence. J. Bacteriol., **40,** 171–195.

ROTH, N. G., DEARMON, I. A., JR., AND LIVELY, D. H. 1956 Survival time as a rapid method of determining virulence with *Bacillus anthracis*. J. Bacteriol., **72,** 666–672.

WILSON, G. S. AND MILES, A. A. 1955 *Topley and Wilson's principles of bacteriology and immunity*, 4th ed. Vol II, Ch. 43. The William & Wilkins Co., Baltimore.