HOMOGENEITY OF RESPONSE OF MOUSE AND GUINEA PIG STRAINS TO VIRULENCE TESTS WITH *BACILLUS ANTHRACIS* AND *PASTEURELLA TULARENSIS*

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

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

Received for publication April 20, 1959

At best, the estimation of virulence is a variable with wide confidence limits because the characteristic being measured represents a complex interaction between host and pathogen. This interaction is influenced by an infinite number of variables in the heredity of the living organisms and their cells and by the environment in which these systems interact. Roth *et al.* (1956) and Fernelius *et al.* (1958) reported that a graded response virulence test based on time to death of all animals challenged appeared favorable for increased test uniformity. This paper reports the homogeneity of various animal strains and makes comparisons between the reliability of the graded and quantal method of virulence testing. Our results show a feasible way to reduce the cost of virulence tests by reducing the number of animals, time, and physical facilities required to conduct a virulence test at any given level of precision.

METHODS

Animal strains. Five inbred or F_1 hybrid strains of mice and two inbred strains of guinea pigs, furnished by the National Institute of Health, Bethesda, Maryland, and noninbred strains of mice or guinea pigs, produced at Fort Detrick, were used in this study. Designation of the strains used was:

Mouse: A/HEN, A/LN, BALB/c AnN, CAF_1, and FD (Swiss).

Guinea pig: no. 2, no. 13, and FD (Hartley). Descriptions of the National Institute of Health (NIH) strains are given by the Committee on Standardized Nomenclature (1952). The Fort Detrick (FD) strains of both mice and guinea pigs are albino, have a high reproduction rate, and have relatively little inbreeding, being maintained by a closed colony breeding system. The hybrid between two inbred lines, CAF_1, is, in this report, referred to as an inbred.

Roth *et al.* (1956) showed the sex of mice had no significant effect on mean time to death. For this work, both male and female mice of the FD strain were pooled in all tests.

Pathogens. Spores of 8 strains of *Bacillus anthracis* were produced by methods given by Roth *et al.* (1955). Those lots designated as nos. 3, 4, 17, 18, and 27 were spores produced at different times, all being derived from the Vollum strain. Three other strains, 30R, 32R, and 298R were stocks characterized by large rough colonies and of relatively low virulence.

The SCHU strain of *Pasteurella tularensis* was used for all the work. Cultures were produced by methods given by Boyles and Lincoln (1958).

Inoculation and testing. Spores of *B. anthracis* and cells of *P. tularensis* were serially diluted in gelatin-phosphate diluent. Animals were inoculated intraperitoneally (ip) with 0.5 ml of the indicated dilution of culture. Observation of animals began 8 hr after inoculation and continued at ½-hr intervals through 48 hr then at 4-hr intervals. Challenge doses, 1 log apart, were administered. Actual doses are indicated in the text and tables.

Statistical Methods. Statistical methods were used to test for uniformity of animal strains, and for contrasting animal requirements by strains for bio-assay.

In the test of uniformity of an animal strain, the death time distribution or variability of individual animals associated with a killing dose of microorganisms can be defined by the variance of the distribution, σ^2. We have used the log of death time as the variable since it is assumed to be more normally distributed about the mean than death time itself. Variances associated with different strains of animals are compared in the usual variance ratio tests, and those strains with the smallest variances are considered the most uniform hosts for study of a particular pathogen. Although we are not aware of reported work where variability in animal death times was used to contrast inbred strains, Jay (1955) and

TABLE 1

Equations in N, the approximate animal requirements, for both graded and quantal bio-assays

Parameter to be Estimated by Assay	Type Assay	Animal Requirements for Each of "Standard" and "Unknown" Materials	Equation No.
Potency ratio of doses to yield equal responses	Graded	$N = \frac{32\lambda^2}{L^2}$	(2)
LD_{50}	Quantal	$N = \frac{38}{L^2b^2}$	(3)
Potency ratio of LD_{50}'s	Quantal	$N = \frac{77}{L^2b^2}$	(4)

L = length of 95 per cent confidence interval of log potency ratio or length of 95 per cent confidence interval of log LD_{50} for appropriate formula; b = probit slope of dose-response curve; and $\lambda = \sigma/b'$.

McLaren and Michie (1954) use the variability in sleeping time associated with barbiturates to contrast uniformity of inbred, hybrid, and non-inbred strains of mice.

For contrasting the animal requirements by strain for use in graded bio-assay, consideration must be given to both the slope of the dose response curve and the variance of the animal response. In the graded response bio-assay, regression analyses are used and the statistic of interest as shown by Finney (1952*a*) is λ, which is the ratio of the standard deviation, σ, to the slope of the dose response curve, b'.

For contrasting the animal requirements by strain for use in the quantal response bio-assay, the most important statistic is the probit slope of the dose-response curve (Finney, 1952*b*).

In this paper the comparison of animal strains for use in either the graded or quantal bio-assays has been based on the contrast of animal requirements of any two given strains to yield equal precision in the results. The proportion of animals required from one strain (N_1) to those required from another strain (N_2) to yield equal precision in the bio-assay can be approximated by the ratio, R

$$R = N_1/N_2 \qquad (1)$$

Equations in N, the approximate animal requirements, for both graded and quantal bio-assays have been given by DeArmon and Lincoln (1959) and are shown in table 1.

RESULTS

Graded response using spores of B. anthracis. Replicate tests were made on various strains of mice, the survival time of each mouse being observed. Using log death time, the variances and dose-response slopes for each strain of mouse was computed along with the median death time of the group. These values are given in table 2.

(1) Death time distribution:—Since the per animal variance represents the death time distribution for that particular strain (or sex) variances may be compared by the variance ratio. Before performing statistical tests between variances, the comparable variances were pooled within each of the two replications. Variances associated with male and female mice of the CAF_1 and A/LN strains were not statistically significant and were pooled with the strain. However, the variances, each with approximately 400 degrees of freedom, associated with the highly virulent strains (0.0021 for lot nos. 17, 18, and 14 of replication 2) of *B. anthracis*, were smaller than those associated with the less virulent variant strains (0.0049 for lots nos. 32R, 30R, and 298R). This difference between these variances is partly or totally attributed to the fact that mean time of death between the two types of *B. anthracis* was widely different. Doses should be adjusted so that time of response is approximately equal, otherwise bio-assays comparing spores of high and low virulence must be evaluated carefully as to the homogeneity of variances when making tests of significance.

Before making tests of significance between strains of mice, the variances were pooled over sex and over replication into single estimates as shown in table 3, column 3. Statistical tests of variance ratios show that the variances associated with the FD or C_3HP and the CAF_1 or A/LN mice, respectively, are not significantly different. However, the pooled variance of CAF_1 and A/LN mice (0.0018) is significantly smaller than the pooled variances of FD and C_3HP mice (0.00453) ($P < 0.01$, two tail F).

(2) Slopes of the dose-response curves:—All except 2 of the 30 dose-response curves computed were linear over the three doses used. Although

TABLE 2

Summary of mean, slope, and variance based on time to death of mice injected with various lots of spores of Bacillus anthracis

Mean No. Spores injected $\times 10^7$	Spore Lot No.	Mouse Strain	No. Mice Challenged	Median Time to Death (MTD) Adjusted*	Observed Mean Log Slope	Per Animal Variance	Degrees of Freedom
$\bar{x}$				*hr*	*log units*		
Replication 1†							
	3 + 27	FD	144	17	−0.092	0.0040	131
	3 + 27	CAF_1 (F)	56	16	−0.063	0.0003	50
11.0	3 + 27	A/LN (F)	56	15	−0.049	0.0007	49
	3 + 27	A/LN (M)	72	16	−0.048	0.0013	66
	3 + 27	C_3HP (F)	72	18	−0.075	0.0042	64
Replication 2							
	17	FD	72	18	−0.07	0.0031	66
	17	CAF_1 (F)	18	18	−0.09	0.0014	12
11.5	17	CAF_1 (M)	24	18	−0.10	0.0010	18
	17	A/LN (F)	24	17	−0.07	0.0024	18
	17	A/LN (M)	24	17	−0.09	0.0010	18
	18	FD	72	20	−0.12	0.0018	66
	18	CAF_1 (F)	18	20	−0.10	0.0008	12
8.5	18	CAF_1 (M)	24	20	−0.11	0.0014	18
	18	A/LN (F)	24	19	−0.10	0.0010	18
	18	A/LN (M)	24	20	−0.09	0.0014	18
	4	FD	72	12	−0.09	0.0038	66
	4	CAF_1 (F)	18	15	−0.08	0.0026	12
10.5	4	CAF_1 (M)	24	16	−0.03	0.0010	18
	4	A/LN (F)	24	14	−0.06	0.0003	17
	4	A/LN (M)	24	15	−0.05	0.0004	18
	32R	FD	71	22	−0.10	0.0028	65
	32R	CAF_1 (F)	18	22	−0.13	0.0004	12
5.0	32R	CAF_1 (M)	24	23	−0.13	0.0009	18
	32R	A/LN (F)	24	20	−0.13	0.0016	18
	32R	A/LN (M)	24	20	−0.16	0.0020	18
	30R	FD	71	31	−0.07	0.0057	65
	30R	CAF_1 (F)	18	33	−0.12	0.0040	12
3.0	30R	CAF_1 (M)	24	32	−0.06	0.0032	18
	30R	A/LN (F)	24	22	−0.11	0.0034	17
	30R	A/LN (M)	24	26	−0.08	0.0031	18
	298R	FD	67	31	−0.10	0.0119	61
	298R	CAF_1 (F)	18	36	−0.09	0.0047	12
5.5	298R	CAF_1 (M)	23	31	−0.03	0.0048	17
	298R	A/LN (F)	24	29	−0.05	0.0043	18
	298R	A/LN (M)	24	29	−0.06	0.0063	18

(F) = female; (M) = male.

* Adjusted by $MTD = MTD_{obs.}(10^7/\bar{x})^{-0.08}$].

† The test was of two virulent lots, nos. 3 and 27. Since the results in no case differed significantly, the data are combined.

TABLE 3

Pooled statistical constants for mouse strain challenged with spores of Bacillus anthracis

Mouse Strain	Degrees of Freedom	Per Animal Variance	Slope of Dose-Response Curves, $-b \pm s_b$	$\lambda \pm s^*_\lambda$
FD	520	0.00457	0.089 ± 0.006	0.76 ± 0.06
CAF_1	229	0.00173	0.085 ± 0.006	0.49 ± 0.04
A/LN	329	0.00185	0.077 ± 0.005	0.56 ± 0.05
C_3HP	64	0.00420	0.075 ± 0.012	0.86 ± 0.15

* Approximate $V(\lambda) = \lambda^2\left(\frac{V(s)}{s^2} + \frac{V(b)}{b^2}\right)$.

considerable variability was encountered between slopes of the dose-response curves, no explanation could be attributed to any specific deviation, therefore, weighted slopes were pooled as to mouse strain as shown in column 4, table 3. Test of significance relative to the pooled slopes indicated that the A/LN and C_3HP mouse strains are significantly smaller in slope than the FD and CAF_1 strain of mouse.

(3) Contrast of animal strains for bio-assay of *B. anthracis*:—In using the graded response for bio-assay, the precision of the assay is a function of the per animal variance and the slope of the dose-response curve. The square root of the per animal variance, s, divided by the slope of the dose-response curve, b, is defined as λ. Since the standard deviation and the slope of the dose-response curve have been estimated for each strain of mouse under consideration, the λ values associated with each mouse strain can be calculated and are given in table 3, column 5.

Tests of significance indicate that the values for λ associated with the FD and C_3HP mice are significantly larger than the values for λ associated with the CAF_1 and A/LN mice. By utilizing equation (1) from the Statistical Methods section, the proportion of animals from each strain required for equal precision of results has been calculated for strain comparisons having significantly different λ values and is shown in table 4.

For example, it is estimated that in this graded bio-assay it would require 240 FD mice to obtain the same precision as with 100 CAF_1 mice; therefore, a 58 per cent savings in animals would result if CAF_1 mice were used instead of FD mice. There would be no significant savings if CAF_1 mice were used rather than A/LN since there was no significant difference between the observed values for λ associated with these strains of mice.

To obtain some idea about the size of a bio-assay between two *B. anthracis* materials under study, the approximate animal requirements for a given degree of precision has been estimated from equation (2) (table 1) according to strain of mouse and shown in table 5.

TABLE 4

Proportion of animals from each strain required for equal precision of results

Strains Compared	Relative Animal Requirement for Equal Precision	Savings in Animals if More Uniform Strains Were Used
		%
FD to CAF_1	2.4 to 1	58*
FD to A/LN	1.9 to 1	47
C_3HP to CAF_1	3.1 to 1	68
C_3HP to A/LN	2.4 to 1	58

* Per cent savings $= 100 - \left(\frac{1}{2.4}\right)100$.

TABLE 5

Animal requirements for a graded bio-assay with spores of Bacillus anthracis

Desired Length of 95% Confidence Interval of Log Potency Ratio	Approx No. Animals Required per Treatment Using Indicated Strain			
	FD	CAF_1	A/LN	C_3HP
0.2	443	184	240	574
0.4	110	46	60	144
0.5	71	30	39	92
0.6	50	21	27	64
0.7	36	15	20	47

Homogeneity of animals challenged by cells of P. tularensis. (1) Graded ip response in mice:—The graded response by ip challenge of mouse strains A/HEN, BALB, CAF_1, A/LN, C_3HP, and FD to cells of *P. tularensis* was determined with

TABLE 6

Summary of median time to death, slope, and variance of strains of mice injected intraperitoneally (ip) with cells of Pasteurella tularensis (lot no. 3)

Mouse Strain	No. Mice	Median Time to Death*	Slope and standard deviation of Dose-Response Curve, $-b \pm s_b$	Per Animal Variance	Degrees of Freedom
		hr			
FD	72	20.9†	0.187 ± 0.009	0.00197	61
A/HEN (F)	14	21.4	0.083 ± 0.010	0.00033	12
BALB/c (M)	32	19.5	0.154 ± 0.008	0.00058	30
CAF_1 (M)	26	20.4	0.153 ± 0.009	0.00050	24
CAF_1 (F)	40	21.9	0.165 ± 0.011	0.00116	38
A/LN (M)	40	22.9	0.170 ± 0.016	0.00233	36
A/LN (F)	40	23.4	0.184 ± 0.013	0.00213	31
C_3HP (M)	34	20.4	0.140 ± 0.011	0.00108	32
C_3HP (F)	40	21.9	0.178 ± 0.010	0.00270	36

(F) = female; (M) = male.

* Mean dose injected ip was 32×10^8 organisms except in FD mice, mean dose was 100×10^8.

† Survival time adjusted to 32×10^8 organisms injected ip using slope −0.187.

regard to time to death response curves. Three strains, CAF_1, A/LN, and C_3HP, were divided into groups by sex. Cells of *P. tularensis* from one lot were used for all injections. The inbred strains were challenged by doses of 100 and 10×10^8 cells, whereas the FD strain was challenged by doses of 1000, 100, and 10×10^8 cells. A summary of the observed data, median time to death, slope of the dose-response curve, and per animal variance is shown for each strain of mouse in table 6.

a. Death time distribution. The death time distribution associated with strains of mice is indicated by the per animal variance. Variance attributable to sex where tested did not differ significantly and was pooled. Tests of significance indicate that the per animal variance for these six strains could be consolidated into two general groupings as follows: namely, a first group composed of the FD, A/LN, and C_3HP strains having a mean variance of 0.00205 for 196 degrees of freedom and a second group composed of the A/HEN, BALB, and CAF_1 strains having a variance of 0.00074 for 104 degrees of freedom.

The ratio of the two variances is 2.77, with P less than 0.01 for a two tail distribution. This indicates that in three strains of mice, A/HEN, BALB, and CAF_1, the death time distribution and the standard deviation of these distributions is much smaller than in the FD, A/LN, and C_3HP strains when injected with cells of *P. tularensis*.

b. Slope of the dose-response curve. In estimating the slope of the dose-survival time response curve the slope for the FD mice using three dose levels has a much better slope estimation than does any other strain as these were based on a two-dose unreplicated test. The CAF_1, A/LN, and C_3HP slopes were pooled for the male and female mice.

TABLE 7

Strains grouped and pooled with regard to like dose-response slopes

Mouse Strain	Pooled Slopes, $-b \pm s_b$
FD, CAF_1, A/LN, C_3HP	0.171 ± 0.009
BALB/c	0.154 ± 0.008
A/HEN	0.093 ± 0.010

The strains have been grouped and pooled with regard to like dose-response slopes and are shown in table 7.

c. Contrast of animal strains for bio-assay of *P. tularensis*. The observed values of λ (the square root of the per animal variance divided by the slope of the dose response curve) for the six strains of mice are shown in table 8. The values for λ fall into three general groups, and within each group they do not differ significantly from each other. The A/HEN line differs from other lines by the large standard deviation of λ and indicates that our knowledge of λ is not very precise. The mean λ of the second grouping is significantly larger than the mean λ associated with the third group. Using equation (1)

TABLE 8

Observed λ (square root of the per animal variance divided by the slope of the dose response curve) for 6 strains of mice

Group	Mouse Strain	$s/b = \lambda \pm s_\lambda$*
1	A/HEN	0.195 ± 0.045
2	FD	0.237 ± 0.024
	A/LN	0.267 ± 0.028
	C_3HP	0.275 ± 0.028
3	BALB	0.154 ± 0.026
	CAF_1	0.187 ± 0.019

* Approximate $V(\lambda) = \lambda^2\left(\frac{V(s)}{s^2} + \frac{V(b)}{b^2}\right)$.

TABLE 9

Animal requirements for a graded bio-assay with cells of Pasteurella tularensis

Desired Length of 95% Confidence Interval of Log Potency Ratio	Approx No. Animals Required per Treatment Using Indicated Strain					
	FD	A/HEN	BALB	CAF_1	A/LN	C_3HP
0.1	173	117	73	108	219	233
0.2	44	30	19	27	55	59
0.3	20	13	9	12	25	26
0.4	11	8	5	7	14	15

(Methods section) it can be shown that a savings of 60 per cent would be made if animals of group 3 were used rather than animals in group 2.

The approximate animal requirement for a graded bio-assay between two lots of *P. tularensis* has been estimated from equation (2) (table 1) according to the six mouse strains for several intervals of L as shown in table 9.

(2) Graded ip response in guinea pigs:—Three strains of guinea pigs were used in an experiment to determine the distribution of death times and slope of the dose-response curve. All three strains were characterized by weight and sex and challenged with 12.5 and 1.25 × 10^9 cells of *P. tularensis*. The survival time of each animal was recorded. As there was no correlation between animal weight and death time it was concluded that a covariance analysis would not change the death time distributions pattern, therefore, analyses were performed on the survival time observations using log death time as the variable. A summary of the analysis showing strain, sex, number of animals, weight range, median death time, slope of dose-response curves, and per animal variance is given in table 10.

a. Death time distribution. The death time distribution associated with strains of guinea pigs is indicated by the per animal variance. The per animal variances, pooled over sex, are given in table 11, column 3.

Tests of significance indicate that Hartley and NIH no. 2, the NIH no. 2 and NIH no. 13, and the Hartley and NIH no. 13, differ significantly in variances. It is expected, therefore, that the FD Hartley strain would yield the smallest distribution of death times.

b. Slopes of the dose-response curves. Tests of significance indicated no significant differences between slopes associated with sex of the same strain of guinea pig, therefore, the slope estimates by strain were pooled over sex and are given in table 11, column 4.

Tests of significance among the weighted mean slopes for the three strains indicate borderline significance for one pair of slopes, 0.140 and 0.298. Since these experiments were not replicated it is believed that the slope values should be considered homogeneous, having a mean value of 0.193 ± 0.020.

The calculated values for λ for the three strains of guinea pigs using the mean slope of 0.193 are given in table 11, column 5.

Tests of significance among the values for λ indicate that only the Hartley and the NIH no. 13 strains are significantly different. This indicates that in estimating animal requirements it would take on the average approximately 4.3 NIH no. 13 guinea pigs for every FD Hartley guinea pig to gain equal precision in the graded assay. Although the difference between the λ values of the FD Hartley strain and NIH no. 2 does not test to be significant, the difference is great enough to be considered in planning experiments with animals as expensive as guinea pigs. The approximate guinea pig requirements by strain for a graded bio-assay between two lots of *P. tularensis* as estimated from equation (2) (table 1) are shown in table 12.

(3) IP quantal response in mice:—By the graded response of mice to ip challenge to *B. anthracis* or *P. tularensis*, it has been demonstrated that the per animal variance associated with some inbred strains of mice was less than that associated with the FD noninbred mice. The probit slope of the quantal dose-mortality

TABLE 10

Summary of mean, slope, and variance based on survival time of guinea pigs injected with cells of Pasteurella tularensis

No. Animals	Guinea Pig Strain	Wt Range	Mean Time to Death*	Slope, $-b \pm s_b$	Per Animal Variance	Degrees of Freedom
		g	*hr*			
24	Hartley (M)	240–300	49.8	0.17 ± 0.03	0.0065	22
24	Hartley (F)	260–313	49.9	0.22 ± 0.03	0.0058	22
24	NIH no. 2 (M)	218–414	48.3	0.09 ± 0.05	0.0131	22
36	NIH no. 2 (F)	213–356	54.0	0.17 ± 0.04	0.0128	34
14	NIH no. 13 (M)	179–314	60.8	0.19 ± 0.06	0.0290	12
18	NIH no. 13 (F)	160–314	78.3	0.38 ± 0.08	0.0250	16

(F) = female; (M) = male.

* Injected with geometric mean dose 3.95×10^9 organisms.

TABLE 11

Pooled statistical constants for guinea pig strains challenged with cells of Pasteurella tularensis

Guinea Pig Strain	Degrees of Freedom	Per Animal Variance	Slope of Dose-Response Curves, $-b \pm s_b$	$\lambda \pm s_\lambda$
FD Hartley	44	0.0062	0.190 ± 0.020	0.408 ± 0.061
NIH no. 2	56	0.0129	0.140 ± 0.030	0.588 ± 0.083
NIH no. 13	28	0.0267	0.298 ± 0.058	0.846 ± 0.143
Mean			0.193 ± 0.020	

response, theoretically, should increase as the uniformity of the animal strain increases. This theory was tested when two inbred strains of mice, CAF_1 and C_3HP, were compared with the FD strain for probit slope of the LD_{50} response to cells of *P. tularensis*.

The C_3HP and FD mice were each divided into three equal size groups for injection of one group on successive days. Due to a shortage of CAF_1 mice, only two groups for each sex could be obtained and these were challenged on the first and second days of the experiment. Six dose levels obtained by a serial dilution process were made each day. Response was taken as death of an animal within a 10-day observation period. The LD_{50}'s and probit slopes were calculated using the maximum likelihood method and are shown in table 13. The LD_{50} estimates are not significantly different for the strains of mice. Tests of significance showed that the several slopes associated with each mouse strain were homogeneous. The weighted pooled probit slopes are as shown in table 14.

Tests of significance indicated no real difference between the pooled probit slopes associated with strain of mouse, although both the inbred CAF_1 and C_3HP mice have an observed slope slightly larger than the FD noninbred mice. The animals available were used in these tests; however, larger numbers of animals and better selection of inbreds as judged by the graded response should be used before the results are considered conclusive.

TABLE 12

Animal requirements for a graded bio-assay with cells of Pasteurella tularensis

Desired Length of 95% Confidence Interval of Log Potency Ratio (L)	Approx No. Guinea Pigs per Treatment with Indicated Strain		
	FD Hartley	NIH no. 2	NIH no. 13
0.2	128	270	550
0.3	57	120	244
0.4	32	68	138
0.5	21	43	88
0.6	15	30	62
0.7	11	22	45

TABLE 13

Summary of observed LD_{50} and probit slope by strain of mouse and day of injection

Mouse Strain	Date of Injection	Approximate LD_{50} (cells)	Probit Slope
CAF_1 (F)	6 June*	0.50	1.14
	7 June*	1.04	0.72
CAF_1 (M)	6 June	0.40	0.83
	7 June	6.00	1.12
FD	6 June	0.80	1.79
	7 June	3.00	0.74
	8 June†	0.30	0.83
C_3HP	6 June	0.80	1.83
	7 June	2.10	0.77
	8 June	0.40	1.74
Over-all weighted mean		0.76 (0.55 to 1.14)‡	0.94 (0.76 to 1.12)‡
In logs		−0.1192 (−0.2596 to 0.0569)‡	

(F) = female; (M) = male.

* Challenge doses were 105, 21, 10.5, 2.1, 1.05 and 0.21 cells.

† Challenge doses were 32.5, 6.26, 3.25, 0.63, 0.33 and 0.063 cells.

‡ Ninety-five per cent confidence interval of mean.

TABLE 14

Weighted, pooled probit slopes for LD_{50} (quantal) test of mice challenged with cells of Pasturella tularensis

Mouse Strain	Probit Slope
CAF_1(F)	0.86
CAF_1(M)	0.90 } 0.87 pooled
FD	0.80
C_3HP	1.08
Over-all mean (95% CI)	0.94(0.76 to 1.12)

(F) = female; (M) = male; CI = confidence interval.

Contrast of animal strains for quantal bio-assays.

In most of the quantal response bio-assays conducted in our laboratory, the average probit slope equals approximately 1.0. The mean probit slope in this experiment was 0.94. Meynell (1957) claims that with an entirely uniform strain of animal the probit slope should approach 2.0. This might be accomplished with an inbred strain of mouse as yet untested.

With the equations developed by DeArmon and Lincoln (1959) and shown as equations (3) and (4) in table 1 of this report, a contrast in animal requirements for bio-assay can be made between animals producing a probit slope of 1.0 and those contemplated to produce a probit slope of 2.0. The approximate animal requirements for

TABLE 15

Animal requirements for quantal response bio-assay

Length of 95% Confidence Interval of Log LD_{50} or Potency Ratio	Probit Slopes	Animal Requirements for Each Standard or "Standard" and "Unknown" Material	
		LD_{50}*	Potency ratio based on 2 LD_{50}'s†
0.21	1.0	860	—
	2.0	215	—
0.34	1.0	328	660
	2.0	80	160
0.50	1.0	150	300
	2.0	37	75
0.77	1.0	63	126
	2.0	16	32
1.26	1.0	—	46
	2.0	—	12

* $N = \frac{38}{L^2b^2}$.

† $N = \frac{77}{L^2b^2}$.

a LD_{50} determination (equation 3) and the LD_{50} for the "unknown" and "standard" materials to produce the potency ratio between the two LD_{50}'s (equation 4) are given in table 15.

These estimates indicate that with our present methods and animals, the animal requirement for reasonable precision of an LD_{50} estimate is large. When a small number of animals is used in the determination of LD_{50}'s, the difference in LD_{50}'s must be large in order for the potency ratio between the LD_{50}'s to be significant. A considerable savings in animals for bio-assay would result if animals could be found uniform enough with a given pathogen to produce a probit slope significantly greater than the present slopes of 1.0. These estimates should be used only as a guide to the experiment size. Should a large number of 100 or 0 per cent responses occur, the equations will underestimate the animal requirements.

DISCUSSION

Experiments should be designed to furnish the maximum amount of information for a given expenditure of material or energy with bio-assays. This condition will be obtained by using animal strains giving least variance with the desired test, and it appears that no one strain will be optimum for all kinds of tests that may be performed. This concept is illustrated by arraying all lines used in the three tests in order of increasing per animal variance:

B. anthracis, graded response, CAF_1, A/LN, FD, C_3HP.

P. tularensis, graded response, A/HEN, BALB, CAF_1, C_3HP, FD, A/LN.

P. tularensis, quantal response, C_3HP, CAF_1, FD.

It seems probable that only if inbred lines are selected for uniformity to test conditions while inbreeding is in progress will maximum uniformity be obtained. A laboratory not now using inbred animals for bio-assays could make great improvement in bio-assay uniformity by testing presently developed inbred strains and using those strains that prove least variable. It would seem that for any long term program, a number of inbred strains or hybrids would be evaluated, and the one of maximum uniformity used for later tests. This preliminary survey would assure that for any given expenditure of animals, (a) each conclusion made would be made with maximum reliability, (b) more conclusions could be made, or (c) both more conclusions and greater reliability could be achieved. Reduction in the number of animals used results in savings other than the cost of animals, particularly manpower of handling and observing, holding space, and reduced risk to personnel when pathogens are assayed.

The increased variance of inbred guinea pigs used in these tests compared with the noninbred FD Hartley strain guinea pig was unexpected but not completely in contrast to the reaction with mouse inbreds. Actually, for each of the graded response tests using mice, one or more inbred lines had a variance not significantly different from the FD noninbred animal. More guinea pig lines should be tested to obtain more uniform lines as the per animal variance is extremely high as compared to the variance of mouse strains. This observation furnishes proof to the belief that no one inbred or strain of animals is suitable for all uses and parallels the observations of McLaren and Michie (1954) and Emmons (1939) that inbred lines may give a more variable response than noninbred animals.

These data indicate that a mouse furnishes as much information as a guinea pig in regard to the virulence characteristic of *P. tularensis*. For tests conducted where this relationship exists, the cheapness and size of the mouse allows the opportunity of obtaining virulence data based on an adequate number of animals, for as noted in the text, the LD_{50} test requires the use of a substantially larger number of animals than frequently is now used to establish virulence information.

It is interesting to note that the bio-assay test for *P. tularensis* is more precise than the test for *B. anthracis*. This is indicated by a comparison of the λ values associated with each microorganism. The range of λ values for *P. tularensis* was 0.15 to 0.28, whereas the range of λ values for *B. anthracis* was 0.49 to 0.86. The same information has been presented more practically by indicating the number of animals required for a comparable level of precision using either of the pathogens.

In conducting routine virulence tests using the graded response, the median death time of the censored group of animals has been found to be as reliable an estimate as the mean time to death based on observing all animals on test. Methods of calculating means and variances of censored

groups have been given by Sahran and Greenburg (1956). Censoring the latter portion of the distribution shortens the time during which observations need to be made. Animals challenged with large numbers of cells give 100 per cent response, and response time is remarkably uniform. In all of the individual tests made in this work using *B. anthracis*, for all tests in which the challenge dose was about 1 billion spores, the minimum time of response for 10 to 12 animals was 11 to 12 hr and the most variable time was 11 to 18 hr. Comparable uniformity of response to *P. tularensis* was observed. It is suggested that this test concept may prove very useful in obtaining diseased animals in a uniform condition or state of response.

Results of this work show that with the graded response measurement the time to perform an assay is reduced to one day or less as compared to about 10 days for the quantal response test. Results further indicate that for equal precision in the bio-assay between two materials the graded response test can be conducted with about ¼ the animals necessary with the quantal response test. These results combine into a savings of animals, time, and physical facilities required to conduct a virulence test. The results of our virulence tests using several strains of *P. tularensis* agree in general with those of Moody and Downs (1955); therefore the application of the graded response test with this pathogen is limited to comparisons within the strain of pathogen.

ACKNOWLEDGMENT

It is a pleasure to acknowledge the encouragement of Dr. George Jay of the National Institutes of Health, Bethesda, Maryland, and to thank him for furnishing the inbred animals used in these experiments.

SUMMARY

Five inbred lines of mice and two inbred lines of guinea pigs were compared with the Fort Detrick (FD) noninbred mouse and guinea pig strain with regard to the uniformity of their response to virulence tests. The uniformity of response was evaluated using both a graded and quantal response test in mice and a graded response test in guinea pigs. Formulas are used for estimating the minimum number of animals needed to determine the virulence of a culture depending upon the desired length of statistical confidence interval.

When challenged with spores of *Bacillus anthracis* certain inbred lines of mice were significantly more uniform than other inbred lines and the FD strain. The dose-response curves were linear over the range of doses used and varied according to strain of animal used. Savings in animals up to 58 per cent would occur if the more uniform inbred lines, such as CAF_1, were used rather than the noninbred FD mouse. When challenged by *Pasteurella tularensis*, using a graded response test, three of the inbred lines were significantly more uniform than the remaining two inbred lines and the FD noninbred strain. The dose-response slope of one strain was significantly less than the slopes obtained with all other strains. If mouse inbred strain BALB were used instead of the FD noninbred strain a savings up to 58 per cent could be made in animal use for equal precision.

Virulence tests using guinea pigs and a graded response procedure showed the inbred lines to be significantly more variable than the FD noninbred guinea pig strain. The slope of the dose-response curves was linear over the range of doses used and did not vary with strain of guinea pigs; however, this conclusion was reached using a minimum number of guinea pigs, a single route of challenge, and only one pathogenic organism.

REFERENCES

BOYLES, W. A. AND LINCOLN, R. E. 1958 Separation and concentration of bacterial spores and vegetative cells by foam flotation. Appl. Microbiol., **6**, 327–334.

Committee on Standardized Nomenclature for Inbred Strains of Mice. 1952 Standardized nomenclature for inbred strains of mice. Cancer Research, **12**, 602–613.

DEARMON, I. A., JR., AND LINCOLN, R. E. 1959 Number of animals required in the bio-assay of pathogens. J. Bacteriol., **78**, 651–657.

EMMONS, C. W. 1939 Response of inbred mice to oestrone. J. Endocrinol., **1**, 373–377.

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.

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

FINNEY, D. J. 1952*b* *Probit analysis*, 2d ed. Cambridge University Press, London.

JAY, G. E, JR. 1955 Variation in response of various mouse strains to hexabarbitol (Evipal). Proc. Soc. Exptl. Biol. Med., **90**, 378–390.

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

MCLAREN, A. AND MICHIE, D. 1954 Are inbred strains suitable for bio-assay? Nature, **173**, 686–689.

MOODY, M. D. AND DOWNS, C. M. 1955 Studies on tularemia. I. The relation between certain pathogenic and immunogenic properties of variants of *Pasteurella tularensis*. J. Bacteriol., **70**, 297–304.

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.

ROTH, N. G., LIVELY, D. H., AND HODGE, H. M. 1955 Influence of oxygen uptake and age of culture on sporulation of *Bacillus anthracis* and *Bacillus globigii*. J. Bacteriol., **69**, 455–459.

SARHAN, A. E. AND GREENBURG, B. G. 1956 Estimation of location and scale parameters by order statistics from singly and doubly censored samples. I. The normal distribution up to sample size 10. Ann. Math. Stat., **27**, 427–451.