ON AN APPROXIMATE METHOD OF DETERMINING THE MEDIAN EFFECTIVE DOSE AND ITS ERROR, IN THE CASE OF A QUANTAL RESPONSE

By J. O. IRWIN and E. A. CHEESEMAN

Of the Medical Research Council's Statistical Staff

From the Division of Epidemiology and Vital Statistics, London School of Hygiene and Tropical Medicine

The statistical method of treating biological assays based on a quantal (all or none) response and, in particular, of finding the median effective dose and its error has been dealt with by Gaddum (1933), Bliss (1935*a*, *b*) and one of the present writers (Irwin, 1937). The exact statistical treatment is, perhaps, rather laborious. It involves finding the dosage-response relation by an application of regression-technique which may involve several successive approximations. The methods of applying the technique which have been used differ in detail. There is no doubt that in exact work the maximum likelihood solution should be used. The method of obtaining this has been described by Bliss (1938) and by Fisher & Yates (1938, Introduction, Ex. 7). In cases where there are a relatively large number of doses with a small number of animals in each and in consequence responses of 0 and 100% are not infrequent, a large number of successive approximations may be necessary.

It seems possible that a simple approximate method of finding the median effective dose *and its error*, together with instructions for using it, would be welcome to bacteriologists. Kärber's method already fulfils the former requirement, while the approximation to the error, described here, may also be rapidly obtained. Kärber's method (Kärber, 1931; Gaddum, 1933) is appropriate to the case where we have a series of doses increasing in a constant ratio with the same number of animals on each. It consists in calculating

$$m_r = X_r - d\{\tfrac{1}{2}p_1' + p_2' + p_3' + \ldots + p'_{r-1} + \tfrac{1}{2}p_r'\},$$

where m = logarithm of the median effective dose, X_r = logarithm of a dose to which all the animals react, d = logarithm of the constant ratio between two consecutive doses and p_1', p_2', ..., p_r' are the observed proportional responses (mortalities if the effect observed is death). We always start with $p_1' = 0$. If the first dose given does not result in a zero response, we assume the next lower dose would have done so. In practice we calculate the sums $\frac{1}{2}(p_1' + p_2')$, $\frac{1}{2}(p_2' + p_3')$, etc., and add the results together.

If we knew the true proportional responses at each dose, the standard error of the logarithm of the median effective dose could be obtained from the expression $d\sqrt{\{S(pq/n)\}}$, where n is the number of animals on each dose, p the

true proportional response to a particular dose, $q = 1 - p$ and S denotes summation over all doses. However, p, q have to be estimated from the data. Actually, the observed proportional responses are apt to be very irregular if n is small, and it is advisable to smooth them by fitting a straight line. The arithmetical procedure is described below.

The validity of a standard error, calculated by any method based on probability, depends on the degree to which the experimentalist succeeds in randomizing his material in respect to all factors other than those with which his experiment is concerned. In the particular case with which we are dealing, the essential condition is the randomization of the experimental animals. An illustration of the relation between theory and practice in this particular field is afforded by a trial made by our colleagues Mrs Joyce Wilson and Prof. Topley, the results of which are set out in Tables I and II. In this test 250 male

Table I. *Deaths among ten differently labelled samples of mice injected with increasing doses of the same sample of a toxic fraction from* Bact. typhi murium

Dose (mg.)	A	B	C	D	E	F	G	H	J	K
0·0625	1	1	0	2	0	0	0	1	0	1
0·125	2	2	0	0	0	0	0	3	0	0
0·25	3	1	5	5	3	2	4	2	3	5
0·5	5	5	4	5	4	1	3	5	3	4
1·0	5	4	4	5	5	5	5	5	2	5
2·0	5	5	5	5	5	5	5	5	5	5
4·0	5	5	5	5	5	5	5	5	5	5

All bucks, 28–32 g. in weight. Picked out by weight from normal stock and placed five in a cage. These cages were taken at random—the first five mice (one cage) were given 0·0625 mg. and placed in five separate cages labelled A 0·0625, the next five mice were given the same dose and labelled B, after fifty mice had been given 0·0625 mg. (labelled A–K), the same process was repeated with a dose of 0·125 mg. and so on.

mice were picked from normal stock. They were taken at random, the only condition for selection being that any mouse picked should weigh between 28 and 32 g. These mice were placed five in a cage. A toxic fraction isolated from *Bact. typhi murium* was made up to a concentration of 8 mg./c.c. and from this six progressive twofold dilutions were prepared. Ten of the seventy cages in which the 250 mice had been disposed were selected at random, and each mouse was injected intraperitoneally with 0·5 c.c. of the highest dilution containing 0·0625 mg. of the toxic fraction. The ten cages in the order in which they were injected, were labelled A–K. The same procedure was carried out with each dilution, making seven increasing doses, the largest of which was 4 mg. The seven cages labelled A formed one group of thirty-five mice, in which five mice were injected with each of the seven increasing doses, the seven cages labelled B formed a second group, and so on; all mice were observed for four days, and the deaths recorded.

The data may therefore be used for two purposes. First they provide ten comparisons of the approximate with the exact method of calculating the L.D. 50 and its error. Secondly, they may be used to test the extent to which

Table II. *Comparison of exact and approximate methods of obtaining the* L.D. 50 *and its error*

	Exact method (maximum likelihood)				
				Limits of error %	
Sample of mice	log L.D. 50	S.E.	L.D. 50 mg.	$P=0{\cdot}99$	$P=0{\cdot}95$
A	$\bar{1}{\cdot}171$	0·114	0·148	51–197	60–167
B	$\bar{1}{\cdot}341$	0·144	0·219	43–234	52–191
C	$\bar{1}{\cdot}391$	0·113	0·246	51–195	60–166
D	$\bar{1}{\cdot}099$	0·108	0·126	53–189	62–163
E	$\bar{1}{\cdot}438$	0·085	0·274	60–165	68–147
F	$\bar{1}{\cdot}656$	0·097	0·453	56–177	65–155
G	$\bar{1}{\cdot}440$	0·097	0·276	56–178	65–155
H	$\bar{1}{\cdot}162$	0·130	0·145	46–216	56–180
J	$\bar{1}{\cdot}672$	0·127	0·470	47–212	56–177
K	$\bar{1}{\cdot}241$	0·110	0·174	52–192	61–165
Mean	$\bar{1}{\cdot}361$	0·114	0·230	51–197	60–167

	Approximate method (L.D. 50 by Kärber's method error by method described in text)				
				Limits of error %	
Sample of mice	log L.D. 50	S.E.	L.D. 50 mg.	$P=0{\cdot}99$	$P=0{\cdot}95$
A	$\bar{1}{\cdot}187$	0·113	0·154	51–195	60–167
B	$\bar{1}{\cdot}368$	0·132	0·233	46–219	55–181
C	$\bar{1}{\cdot}368$	0·129	0·233	47–215	56–179
D	$\bar{1}{\cdot}127$	0·105	0·134	54–186	62–161
E	$\bar{1}{\cdot}428$	0·098	0·268	56–179	64–156
F	$\bar{1}{\cdot}669$	0·103	0·467	54–184	63–159
G	$\bar{1}{\cdot}428$	0·103	0·268	54–184	63–159
H	$\bar{1}{\cdot}187$	0·100	0·154	55–181	64–157
J	$\bar{1}{\cdot}669$	0·129	0·467	47–215	56–179
K	$\bar{1}{\cdot}247$	0·100	0·177	55–181	64–157
Mean	$\bar{1}{\cdot}368$	0·112	0·233	51–194	60–166

the randomization has been successful, by examining whether the ten results differ significantly among themselves.

The degree of accuracy obtained may be illustrated from the data in Table I. The table gives the deaths among the ten differently labelled samples of mice injected with increasing doses of the same sample of the toxic fraction from *Bact. typhi murium*. There were, as stated above, seven doses increasing in a two-fold ratio and five mice on each dose. Table II shows, for the exact maximum likelihood solution and for the approximate method here described, the logarithm of the L.D. 50, its standard error, the actual L.D. 50 and the limits of error ($P=0{\cdot}99$ and $P=0{\cdot}95$). Limits of error ($P=0{\cdot}99$) of, for example, 50–200% mean that 99 times out of 100 the result would lie between 50 and 200% of its true value.

It is clear that the approximate values are remarkably close to the true ones. The average difference between the two estimates of the L.D. 50 and of its error, is less than 2%. The maximum difference between two individual estimates of the L.D. 50 is 8%. The individual estimates of error, as one might expect, differ rather more. The greatest difference, that for sample H, is about 20%. However, this is not of great importance as the experimenter would not

greatly care whether his limits of error ($P = 0{\cdot}99$) were 46–216% or 55–181%. He would be content to say they were about 50–200%.

We must now enquire how far our randomization has been successful. The standard deviation of the logarithms of the ten results is 0·199, while the average standard error is 0·114 (or 0·112 by the approximate method). The former is significantly greater than the latter; there is therefore some heterogeneity between the groups of animals used in the ten different tests. This heterogeneity may also be shown by comparing the forty-five differences between pairs of results with their own standard errors. Ten differences are found to be greater than twice their standard error and four greater than three times.

It must be remembered that our standard errors have been calculated on the assumption that the animals are homogeneous. Before we conclude that the randomization has been unsuccessful it is necessary to examine whether there is not just as much heterogeneity between the groups of animals used in the same tests as between those used in different tests. The calculation of the standard error by the exact statistical technique involves the fitting of a dosage response curve. When log-doses and the normal equivalent deviations corresponding to the mortalities are used, this curve, on the assumption of homogeneity, becomes a straight line. The goodness of fit of this straight line may be tested by appropriately performing a χ^2 test. If χ^2 significantly exceeds its expected value, it is necessary to multiply the standard error calculated on the basis of homogeneity by $\sqrt{(\chi^2/n)}$, where n is the number of degrees of freedom on which χ^2 is based (see Bliss, 1935*a*, *b*, and Irwin, 1937 for a discussion of these points). In the present instance the dosage-response curve was calculated for the ten tests individually and for the pooled results. Since the theoretical χ^2 distribution is not closely realized with small numbers the latter has been used for the present purpose. χ^2 was found to be 9·70 with 3 degrees of freedom. $\sqrt{(\chi^2/n)} = 1{\cdot}8$ and the correct average standard error for an individual test is therefore $(1{\cdot}8 \times 0{\cdot}114) = 0{\cdot}205$. This is almost identical with the standard deviation of the ten results (0·199). Of course when this value is used there are no significant differences between the pairs of tests.

Thus it seems that the animals are intrinsically heterogeneous, not that the randomization was unsuccessful. The practical conclusion is that the approximate standard error should be used with caution to detect differences between different toxic-fractions. For animals of similar heterogeneity to those used here, differences should only be regarded as significant if they are from 3 to 3·5 times their standard errors so calculated. Otherwise we may merely be detecting differences between the animals. The animals used were of the same sex and age, from a stock bred in the laboratory for many generations, but not by brother-sister mating.

We will illustrate the arithmetic of the approximate method by writing out the L.D. 50 and its error for a particular example—group B in Table I.

For working purposes we call the lowest dose unity and use logarithms of the doses to the base 2. In other words we call the doses 0, 1, 2, 3, 4, 5, 6 and we convert back to actual doses at the end.

(1) *To find the* L.D. 50

(*a*) Write down the deaths and the mortalities at each dose:

Dose				
Actual (mg.)	Working	Deaths	Mortality	
0·0625	0	1	0·2	0·1
0·125	1	2	0·4	0·3
0·25	2	1	0·2	0·3
0·5	3	5	1·0	0·6
1·0	4	4	0·8	0·9
2·0	5	5	1·0	0·9
4·0	6	5	1·0	3·1

(*b*) Assume the previous dose to the lowest would have given zero mortality, and take half the sum of each pair of mortalities, stopping at the first 100% beyond any reversals.

(*c*) Sum these quantities. Result, 3·1.

(*d*) Deduct 3·1 from 5, the last dose included. Result, 1·9.

Then working L.D. 50 = 1·9

$$\log \text{actual L.D. } 50 = \log(0{\cdot}0625) + 1{\cdot}9 \log 2$$
$$= 2{\cdot}7959 + 0{\cdot}5720 = 1{\cdot}3679,$$
$$\text{L.D. } 50 = 0{\cdot}233 \text{ mg.}$$

(2) *To find the error of the* L.D. 50

We start by smoothing the observed mortalities. It is good enough to smooth by fitting a straight line, although the true curve of mortality against log-dose is sigmoid. We again assume that the dose previous to the lowest gives zero mortality and have to fit a straight line to 0, 0·2, 0·4, 0·2, 1·0, 0·8, 1·0. (We again stop at the first 100% beyond any reversals.)

(*a*) Write down the mortalities in a column:

Observed (p)	$S(p)$	Smoothed (p)
0	0	0
0·2	0·2	6/35
0·4	0·6	12/35
0·2	0·8	18/35
1·0	1·8	24/35
0·8	2·6	30/35
1·0	3·6	36/35
3·6 = S_1	9·6 = S_2	

and sum them. Call the sum S_1.

(*b*) Take the successive sums of the items in the first column and sum again. Call this S_2.

(*c*) Calculate (where n is the number of group mortalities used):

$$a = \frac{1}{n} S_1, \quad b = \frac{2}{n(n+1)} S_2, \quad b' = a - b, \quad Y_1 = a + 3b', \quad \Delta Y_1 = \frac{-6}{(n-1)} b'.$$

Then Y_1 is the last smoothed value and ΔY_1 is the constant difference of the smoothed values. The others are then calculated by successively adding ΔY_1.

In the present case

$$a=\frac{3\cdot6}{7}=\frac{18}{35},\quad b=\frac{2}{56}(9\cdot6)=\frac{12}{35},\quad b'=\frac{6}{35},$$

$$Y_1=\frac{18}{35}+\frac{18}{35}=\frac{36}{35},\quad \Delta Y_1=\frac{-6}{6}\left(\frac{6}{35}\right)=\frac{-6}{35}.$$

The last smoothed value works out to be slightly greater than 1, illustrating the approximate nature of the method. It can be counted as 1 in the subsequent calculation. Similarly the first one will sometimes be slightly negative, it can be counted as 0.

(*d*) We then calculate $\sigma^2=S\,(pq/n)$ as follows:

$35p$	$35q$	$35^2\,pq$
6	29	174
12	23	276
18	17	306
24	11	264
30	5	150
		$1170=35^2S\,(pq)$

So

$$\sigma^2=\frac{1170}{5\times(35)^2}=0\cdot1910,$$

$$\sigma=0\cdot437.$$

Standard error of log L.D. 50 $=0\cdot437\log 2=0\cdot132$.

We thus have

log (actual L.D. 50) $=\bar{1}\cdot3679$, with S.E. 0·132.

It is known that 95% of results should be between $\pm1\cdot96$ S.E., and 99% between $\pm2\cdot576$ S.E.

Now antilog $(1\cdot96\times0\cdot132)=$ antilog $(0\cdot2587)=1\cdot814$,

antilog $(2\cdot576\times0\cdot132)=$ antilog $(0\cdot3400)=2\cdot188$.

Hence limits of error ($P=0\cdot95$) are 55–181%,

($P=0\cdot99$) are 46–219%.

In the above example it was easier to work with fractions. But this is not always so. As another example let us take case A, where the deaths are 1, 2, 3, 5, 5, 5, 5. Repeating the above steps we have

(1) L.D. 50

Dose			Mortality	
Actual (mg.)	Working	Deaths		
0·0625	0	1	0·2	0·1
0·125	1	2	0·4	0·3
0·25	2	3	0·6	0·5
0·5	3	5	1·0	0·8
1·0	4	5		1·7
2·0	5	5		
4·0	6	5		

Working L.D. 50 $=3-1\cdot7=1\cdot3$

log (actual L.D. 50) $=\bar{2}\cdot7959+1\cdot3\log 2=\bar{2}\cdot7959+3913=\bar{1}\cdot1872$.

Actual L.D. 50 $=0\cdot154$.

(2) *Error*

Observed		Smoothed		
(p)	$S(p)$	(p)	(q)	(pq)
0	0	−0·04 (taken as zero)	—	—
0·2	0·2	0·20	0·80	0·1600
0·4	0·6	0·44	0·56	0·2464
0·6	1·2	0·68	0·32	0·2176
1·0	2·2	0·92	0·08	0·0736
$2·2=S_1$	$4·2=S_2$			0·6976

$$a=\frac{2·2}{5}=0·44, \quad b=\frac{2}{30}(4·2)=0·28, \quad b'=0·16,$$

$$Y_1=0·44+0·48=0·92, \quad \Delta Y_1=\frac{-6}{4}(0·16)=-0·24.$$

$$\sigma^2=\frac{0·6976}{5}=0·13952,$$

$$\sigma=0·374.$$

Standard error of log L.D. 50 = 0·374 log 2 = 0·113,

log actual L.D. 50 = $\bar{1}$·1872, S.E. 0·113,

actual L.D. 50 = 0·154,

limits of error ($P=0·95$) = 60–167 %,

($P=0·99$) = 51–195 %.

Summary

Kärber's method is, for many purposes, a sufficiently accurate way of determining the median effective dose, when we have a series of doses increasing in a constant ratio and a small number of animals on each. The standard error of the determination may also be rapidly found by a simple approximate method. Illustrative examples are given.

The standard error calculated by the approximate method is based on the assumption that the animals are homogeneous. It must be used with caution when any heterogeneity is suspected. With animals of similar heterogeneity to those used here, the differences between two different toxic-fractions should only be regarded as significant if they are from 3 to 3·5 times their standard errors so calculated.

REFERENCES

Bliss, C. I. (1935*a*). *Ann. appl. Biol.* **22**, 134.

—— (1935*b*). *Ann. appl. Biol.* **22**, 307.

—— (1938). *Quart. J. Pharm. Pharmacol.* **11**, 192.

Fisher, R. A. & Yates, F. (1938). *Statistical Tables for Biological, Agricultural and Medical Research.* Edinburgh and London.

Gaddum, J. H. (1933). *Spec. Rep. Ser. med. Res. Coun., Lond.*, no. 183.

Irwin, J. O. (1937). *J. Roy. Statist. Soc. Suppl.* **4**, 1.

Kärber, G. (1931). *Arch. exp. Path. Pharmak.* **162**, 480.

(*MS. received for publication*, 27. VI. 39.—Ed.)