ANL-4486
Health and Biology

ARGONNE NATIONAL LABORATORY
P. O. Box 5207
Chicago 80, Illinois

# DETERMINATION OF THE $LD_{50}$ BY USE OF PROBIT, ANGULAR, AND LOGIT TRANSFORMATIONS

by

Sylvanus A. Tyler and Joan Gurian

Biology Division

March 1950

Operated by the University of Chicago
under
Contract W-31-109-eng-38

## DISCLAIMER

| Distribution | No. of Copies |
|---|---|
| Argonne National Laboratory | 64 |
| Armed Forces Special Weapons Project | 1 |
| Atomic Energy Commission - Washington | 2 |
| Battelle Memorial Institute | 1 |
| Brush Beryllium Company | 1 |
| Brookhaven National Laboratory | 4 |
| Bureau of Medicine and Surgery | 1 |
| Bureau of Ships | 1 |
| Carbide and Carbon Chemicals Division (K-25 Plant) | 4 |
| Carbide and Carbon Chemicals Division (Y-12 Plant) | 4 |
| Chicago Operations Office | 1 |
| Columbia University (J. R. Dunning) | 1 |
| Columbia University (G. Failla) | 1 |
| Dow Chemical Company | 1 |
| H. K. Ferguson Company | 1 |
| General Electric, Richland | 3 |
| Harshaw Chemical Corporation | 1 |
| Idaho Operations Office | 1 |
| Iowa State College | 2 |
| Kansas City Operations Branch | 1 |
| Kellex Corporation | 2 |
| Knolls Atomic Power Laboratory | 4 |
| Los Alamos Scientific Laboratory | 3 |
| Mallinckrodt Chemical Works | 1 |
| Massachusetts Institute of Technology (A. Gaudin) | 1 |
| Massachusetts Institute of Technology (A. R. Kaufmann) | 1 |
| Mound Laboratory | 3 |
| National Advisory Committee for Aeronautics | 1 |
| National Bureau of Standards | 3 |
| Naval Radiological Defense Laboratory | 2 |
| New Brunswick Laboratory | 1 |
| New York Operations Office | 3 |
| North American Aviation, Inc. | 1 |
| Oak Ridge National Laboratory | 8 |
| Patent Branch (Washington) | 1 |
| RAND Corporation | 1 |
| Sandia Corporation | 1 |
| Santa Fe Operations Office | 2 |
| Sylvania Electric Products, Inc. | 1 |
| Technical Information Division (Oak Ridge) | 15 |
| USAF, Air Surgeon (Lt. Col. R. H. Blount) | 1 |
| USAF, Director of Armament (Captain C. I. Browne) | 1 |
| USAF, Director of Research and Development (Col. R. J. Mason, Fred W. Bruner) | 2 |

| Distribution | No. of Copies |
|---|---|
| USAF, Eglin Air Force Base (Major A. C. Field) | 1 |
| USAF, Kirkland Air Force Base (Col. Marcus F. Cooper) | 1 |
| USAF, Maxwell Air Force Base (Col. F. N. Moyers) | 1 |
| USAF, NEPA Office | 2 |
| USAF, Office of Atomic Energy (Col. H. C. Donelly, A. A. Fickel) | 2 |
| USAF, Offutt Air Force Base (Col. H. R. Sullivan, Jr.) | 1 |
| USAF, Wright-Patterson Air Force Base (Rodney Nudenberg) | 1 |
| U. S. Army, Atomic Energy Branch (Lt. Col. A. W. Betts) | 1 |
| U. S. Army, Army Field Forces (Captain James Kerr) | 1 |
| U. S. Army, Commanding General, Chemical Corps Technical Command (Col. John A. MacLaughlin through Mrs. Georgia S. Benjamin) | 1 |
| U. S. Army, Chief of Ordnance (Lt. Col. A. R. Del Campo) | 1 |
| U. S. Army, Commanding Officer, Watertown Arsenal (Col. Carroll H. Dietrick) | 1 |
| U. S. Army, Director of Operations Research (Dr. Ellis Johnson) | 1 |
| U. S. Army, Office of Engineers (Allen O'Leary) | 1 |
| U. S. Army, Office of the Chief Signal Officer (Curtis T. Clayton through Maj. George C. Hunt) | 1 |
| U. S. Army, Office of the Surgeon General (Col. W. S. Stone) | 1 |
| U. S. Geological Survey (T. B. Nolan) | 2 |
| USAF, Director of Plans and Operations (Col. R. L. Applegate) | 1 |
| U. S. Public Health Service | 1 |
| University of California at Los Angeles | 1 |
| University of California Radiation Laboratory | 5 |
| University of Rochester | 2 |
| University of Washington | 1 |
| Western Reserve University | 2 |
| Westinghouse Electric Company | 4 |
| Naval Medical Research Institute | 1 |
| Total | 194 |

# DETERMINATION OF THE $LD_{50}$ BY USE OF PROBIT, ANGULAR, AND LOGIT TRANSFORMATIONS*

Sylvanus A. Tyler and Joan Gurian

## I. INTRODUCTION

In statistics, extensive use is made of transformations on variates that simplify the extraction of essential parameters and inferences about a population from its sub sets.(1) Especially in the evaluation of biological assays that admit quantal responses have transformations proved indispensible to statistical analysis. In many cases the relation between two concomitant variates can well be represented by considering the normal deviate or probit of one variate as a linear function of an appropriate measure of the other variable. Particularly useful is the probit transformation when a linear relationship exists between these transformed quantities.(2,3,4,5) However, because of the labor involved in response analysis by the probit method, many investigators have sought other transformations that would 1) provide a simple method of computation for assays that admit analysis by probits and 2) determine transformed variates that can be made to satisfy the requirements for statistical inference. Two other transformations meet these requirements with precision when the number of subjects used is large--the angular and the logit transformations. A short discussion of each transformation is given and an example, typical of those encountered in bioassay, is used to illustrate the computational procedure peculiar to each transformation.

## II. PROBIT ANALYSIS

Let us investigate the probit transformation by considering a typical problem in pharmacology. A number of animals, n, is exposed to a known degree of physical or chemical stimulation and a number, s, survives while the remainder dies. The graph of the proportion of dead animals as ordinate and some measure of the corresponding dosage received as abscissa will be sigmoid in shape. If we assume that such a graph is descriptive of a distribution of susceptibilities between individuals of our population, then the ordinate at any dosage level will divide the area of the dosage-mortality curve into two parts that give the proportions of animals possessing individual susceptibilities below and above the dosage indicated. If we now plot the normal probability deviate that corresponds to the per cent deaths against an appropriate function of dosage, a straight-line relation between these quantities will result. To avoid the inconvenience of the negative sign, five is added to the normal deviates and the resulting values

---

*A compilation of techniques for computational purposes.

are called probits.[2] The regression line can now be found and the relationship between mortality and dosage is thus known.

Since the sampling variance of the probit is dependent upon the probit value, the problem of weights in the regression analysis must be considered. Also, special consideration must be given to the case of 0 and 100 per cent survival since great departure from normality is expected.*

We recall that at a given dosage level, s survivors were observed out of n animals tested. If P and Q represent the probabilities of death and survival, respectively, and are estimated by

$$P = 1 - \frac{s}{q} \qquad Q = 1 - P = \frac{s}{q} \tag{1}$$

from the binomial distribution, s survivors are expected in $\psi$ of the cases, where

$$\psi = \frac{n!}{s!\,(n-s)!} P^{n-s} Q^{s} \tag{2}$$

To fit the corrected regression line of the probit y on dose x by the method of maximum likelihood we equate to 0 the sum of the differentials with respect to y of the logarithms of quantities (2) for all dosage levels. The expressions for maximum likelihood are then of the form

$$\frac{d(\log \psi)}{dy} = \frac{d(\log \psi)}{dP}\,\frac{dP}{dy} = \frac{(Qn - s)}{PQ}\,\frac{dP}{dy} = \frac{(Qn - s)\,Z}{PQ} \tag{3}$$

where $\frac{dP}{dy} = Z$ is the value of the ordinate to the normal curve at y. If nP and nQ are both large, the binomial distribution closely approximates the normal, and the factor (Qn - s) is proportional to the differences between the expected and observed probits. Then

$$Qn - s \approx n(Y - y)\frac{dQ}{dy} = n(y - Y)\,Z$$

and equation (3) becomes

$$\frac{d(\log \psi)}{dy} = (y - Y)\frac{nZ^2}{PQ} \tag{4}$$

---

*The binomial distribution is used as the probability model.

where Y is the probit expected. To determine this expression, the values of P and Q are taken from the uncorrected regression line and the corrected line computed by the usual regression procedures.

The binomial distribution when n or s are small numbers, will depart appreciably from normality and the quantities (Qn - s) and (Y - y) will not, in general, be proportional.(6) The case of 0 survivors is an example.(7) Here, y is infinite. Use is made of a working probit that will satisfy the equation

$$Qn - s = nZ\,(y_w - Y) \text{ or } y_w = Y + \frac{1}{Z}\,(Q - \frac{s}{n}) \tag{5}$$

where $y_w$ is the working probit and P, Q, Y are estimated from the uncorrected regression line. With this substitution, the calculation of the corrected regression line follows as before. Let us enumerate a few facts that are embodied in expressions (4) and (5).

1. For 0 survivors (s = 0)

$$y_w = Y + \frac{Q}{Z} \tag{6}$$

2. For a Y less than five, the expression

$$y_w = Y - \frac{1}{Z}\,(P - \frac{n - s}{n}) \tag{7}$$

is convenient to use.

3. The weight $(\frac{nZ^2}{PQ})$ is symmetrical about Y = 5 at which point its value is a maximum.

Tables of the above quantities can be found in several books.(4,8)

a. Application

The computational procedure for determining the dosage necessary to affect 50 per cent of the subjects tested by the probit method is outlined in detail. Illustrative material was taken from the 20-day, X-ray mortality data for mice obtained by G. Sacher.(9) The corrected regression line is calculated and the $LD_{50}$ with associated standard error computed.

Table 1 contains the raw data used and the quantities required for determining the provisional regression line.

Table 1

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| x | n | d | p | q | y | $Y_c$ |
| 467 | 110 | 20 | .182 | .818 | 4.0922 | 3.792 |
| 500 | 112 | 17 | .152 | .848 | 3.9721 | 4.227 |
| 533 | 127 | 38 | .299 | .701 | 4.4727 | 4.662 |
| 567 | 112 | 56 | .500 | .500 | 5.0000 | 5.111 |
| 600 | 124 | 102 | .822 | .178 | 5.9230 | 5.546 |
| 633 | 110 | 86 | .782 | .218 | 5.7790 | 5.981 |
| 667 | 126 | 118 | .936 | .064 | 6.5220 | 6.430 |
| 700* | 94 | 94 | 1.000 | .000 | -- | -- |
| 733 | 94 | 93 | .989 | .011 | 7.2904 | 7.300 |

*Dose group omitted from the analysis.

$N = 8$ $\qquad \bar{x} = 587.5$

$\Sigma x = 4{,}700$ $\qquad \Sigma xy = 26{,}022.0007$

$\Sigma x^2 = 2{,}816{,}534$ $\qquad \Sigma y = 43.0514$

$$b = \frac{N\Sigma xy - \Sigma x \Sigma y}{N\Sigma x^2 - (\Sigma x)^2} = \frac{8(26022.0007) - (4700)(43.0514)}{8(2{,}816{,}534) - (4{,}700)^2} = 0.01319$$

$$a = \bar{y} = \frac{\Sigma y}{N} = 5.381$$

$$Y_c = a + b(x - \bar{x}) = 5.381 + 0.01319\,(x - 587.5) = 0.01319x - 2.368 \qquad (8)$$

The quantities entered in Table 1 are, by columns, the following:

Column 1 - dosage (in r) administered to group
Column 2 - total number of animals in group
Column 3 - number of animals dead within 20 days after exposure
Column 4 - proportion dead within 20 days (Col. 3 + Col. 2)
Column 5 - proportion alive at the end of the 20-day period (1 - Col. 4)
Column 6 - empirical probits of corresponding values in Col. 4 (Table IX, ref. 8)
Column 7 - the expected probits, based on the provisional regression line of empirical probits (y) versus dosage (x)

The expected probits may be obtained from a visually fitted line with, of course, a sacrifice in the accuracy of the resulting determinations. To

establish the dosage-mortality relationship in reference 10, the expected probits were estimated from a visually-drawn trend line and therefore differ slightly from the trend established herein. The provisional regression equation and necessary computations are given immediately below Table 1.

The quantities in Table 2, columns 5 and 6 correspond, respectively, to the maximum working probit and the range associated with each expected probit, obtained by linear interpolation of values in Table II, reference 8. The working probits, $y_w = Y_c + \frac{Q}{Z} - \frac{q}{z}$ are tabulated in column 8. The weighting coefficients, (Table II, reference 8) and the weight assigned each $Y_c$ are recorded in columns 9 and 10. Since the weighting coefficient $\frac{Z^2}{PQ}$ is symmetrical about $Y_c = 5$, its value for an expected probit less than five ($Y_c < 5$) is the same as that for an expected probit equal to $(10 - Y_c)$. Thus to calculate the working probit from an expected probit value greater than five, the quantity $Y' = 10 - Y_c$ is used in place of $Y_c$, the value p is substituted for q and $y_w$ found from $y_w = 10 - y'_w$ (Table 3).

Table 2

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Dose | Number Exposed | % Surv. | Table I | Table II, ref. 4 | | Cols. 3x6 | Cols. 5-7 | Table II ref. 8 | Cols. 2x9 |
| x | n | q | $Y_c$ | $Y_c + \frac{Q}{Z}$ | $\frac{1}{Z}$ | $\frac{q}{Z}$ | $y_w$ | $\frac{Z^2}{PQ}$ | w |
| 467 | 110 | .818 | 3.792 | See Table 3 | | | 4.1500 | .36756 | 40.432 |
| 500 | 112 | .848 | 4.227 | | | | 4.0085 | .51043 | 57.168 |
| 533 | 127 | .701 | 4.662 | | | | 4.4800 | .62201 | 78.995 |
| 567 | 112 | .500 | 5.111 | 6.2611 | 2.5234 | 1.2617 | 4.9994 | .63355 | 70.958 |
| 600 | 124 | .178 | 5.546 | 6.3978 | 2.9143 | .5187 | 5.8791 | .57036 | 70.725 |
| 633 | 110 | .218 | 5.981 | 6.6435 | 4.0615 | .8854 | 5.7581 | .44486 | 48.935 |
| 667 | 126 | .064 | 6.430 | 6.9623 | 6.9915 | .4475 | 6.5148 | .29212 | 36.807 |
| 733 | 94 | .011 | 7.300 | 7.6786 | 35.3020 | .3883 | 7.2903 | .07563 | 7.109 |

Table 3

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Dose | Number Exposed | % Killed | Table I | 10 - $Y_c$ | Table II ref. 8 | | Cols. 3x7 | Cols. 6-8 | 10 - Col. 9 |
| x | n | p | $Y_c$ | Y' | $Y' - \frac{Q}{Z}$ | $\frac{1}{Z}$ | $\frac{p}{Z}$ | $y'_w$ | $y_w$ |
| 467 | 110 | .182 | 3.792 | 6.208 | 6.7983 | 5.2107 | .9483 | 5.8500 | 4.1500 |
| 500 | 112 | .152 | 4.227 | 5.773 | 6.5160 | 3.3846 | .5145 | 5.9915 | 4.0085 |
| 533 | 127 | .299 | 4.662 | 5.338 | 6.3146 | 2.6575 | .7946 | 5.5200 | 4.4800 |

From the entries in Table 2, cols. 1, 8 and 10, the quantities below were found:

$$\Sigma w = 411.129 \qquad \Sigma wy_w = 2{,}094.784712$$

$$\Sigma wx = 232{,}975.286 \qquad \Sigma wy_w^2 = 10{,}980.924078$$

$$\Sigma wx^2 = 133{,}626{,}934.604 \qquad \Sigma wxy_w = 1{,}208{,}478.186544$$

$$\bar{x} = \frac{\Sigma wx}{\Sigma w} = 566.672 \qquad a = \bar{y}_w = \frac{\Sigma wy}{\Sigma w} = 5.095$$

$$b = \frac{\Sigma wy_w x - \frac{(\Sigma wy_w)(\Sigma wx)}{\Sigma w}}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}} = 0.0133$$

The weighted regression line,

$$Y = 5.095 + 0.0133\,(x - 566.672) = 0.0133x - 2.4417 \qquad (9)$$

where Y denotes the proportion of deaths (in probits) and x is the dosage (in r), describes the dosage-mortality relationship characterizing this experimental sample.

The question now arises as to the "goodness of fit" with which the computed relation (9) describes the empirical trend. The $X^2$ test gives the answer.

Table 4

THE $X^2$ TEST FOR GOODNESS OF FIT

| | Item | $X^2$ | Degrees of Freedom* | P |
|---|---|---|---|---|
| a | Total | $\Sigma(wy_w^2) - \frac{(\Sigma wy_w)^2}{\Sigma w}$ | N-1 | |
| b | Regression | $\frac{(\Sigma(wxy_w) - \Sigma wy_w \Sigma wX)^2}{\Sigma wX^2 - \frac{(\Sigma wX)^2}{\Sigma w}}$ | 1 | Any $X^2$ table |
| c | Residual (error) | Total - Regression | N-2 | Any $X^2$ table |
| a′ | Total | 307.995970 | 7 | |
| b′ | Regression | 278.490953 | 1 | $<<.001$ |
| c′ | Residual (error) | 29.505017 | 6 | $<.001$ |

*N = number of dose groups

The $X^2$-test indicates the agreement between expected and observed probits. For homogeneous data (i.e., data for which regression accounts for a significant portion of the total variance and the residual or error variance is nonsignificant) the residual variance ($V_y$) about the regression line is taken equal to one. In this example, however, a significant error variance is found and the heterogeneity that is not accounted for by any systematic nonlinear departure of the empirical data from the linear least-square fit (Figure 1) proportionally contributes to the error inherent in the regression parameters. The error variance is equal to the error-sum-of-squares (total minus regression-sum-of-squares) divided by the corresponding degrees of freedom (number of dosage groups minus two). Thus the variances of the slope (b) and the mean working probit (a) are respectively $V_b = \frac{V'_{y_w}}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}}$, and $V'_a = \frac{V'_{y_w}}{\Sigma w}$. The square root of these variances gives the corresponding standard errors. The variances associated with the regression line of our example are

$$V'_{y_w} = \frac{\left\{\Sigma wy_w^2 - \frac{(\Sigma wy_w)^2}{\Sigma w}\right\} - \left\{\frac{(\Sigma wy_w x - \frac{(\Sigma wy_w \Sigma wx)^2}{\Sigma w}}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}}\right\}}{N-2} = \frac{29.505017}{6} = 4.917502 \qquad (10)$$

//

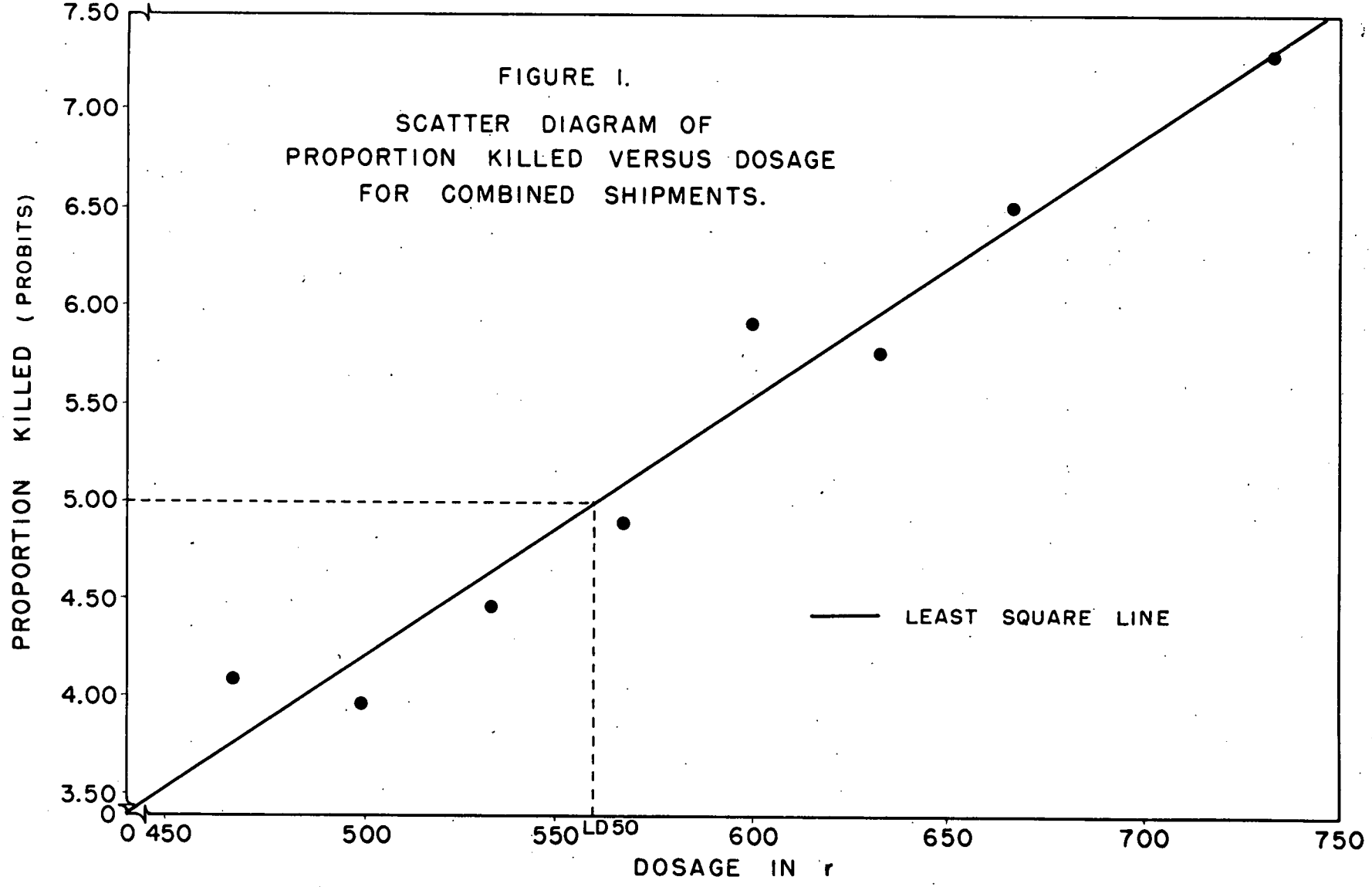


FIGURE I.
SCATTER DIAGRAM OF
PROPORTION KILLED VERSUS DOSAGE
FOR COMBINED SHIPMENTS.
PROPORTION KILLED (PROBITS)
7.50
7.00
6.50
6.00
5.50
5.00
4.50
4.00
3.50
0
0 450
500
550
LD50
600
650
700
750
DOSAGE IN r
LEAST SQUARE LINE

$$V_b = \frac{V_{y'_w}}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}} = \frac{4.917503}{1{,}606{,}363.336} = .306126 \times 10^{-5}; \quad \sigma_b = \sqrt{V_b} = .00175$$

$$V_a = \frac{V_{y'_w}}{\Sigma w} = \frac{4.917503}{411.129} = .0119610; \qquad \sigma_a = \sqrt{V_a} = .109$$

A quantity usually desired in toxicity studies is the amount of an agent necessary to produce an effect upon p per cent of the subjects tested. When the agent is administered in dosages and death of the subject is the response criterium, the dosage, lethal to p per cent of the subjects is notationally written $LD_p$. Mathematically, determining the $LD_p$ is equivalent to finding a dosage (x), given a measure of mortality (Y). For if $Y = a + b(x - \bar{x})$, then $x = \bar{x} + \left(\frac{Y - a}{b}\right)$, and in particular,

$$LD_p = \bar{x} + \frac{(p' - a)}{b} \tag{11}$$

where $p'$ is the probit transpose of p. The variance of the $LD_p$ is found to be (ref. 5, pp. 250-1)

$$V_{LD_p} = \frac{1}{b^2} V_a - (LD_p - x)^2 V_b \tag{12}$$

The computation of the $LD_{50}$ and associated variance and standard error for the data of our example is given below.

$$LD_{50} = \bar{x} + \frac{(5.0 - a)}{b} = 566.672 + \frac{(5.0 - 5.095)}{0.0133} = 559.529$$

$$V_{LD_{50}} = \frac{1}{b^2}\left(V_a + (LD_{50} - \bar{x})^2 V_b\right) = \frac{1}{(.0133)^2}\left\{.0119610 + (7.143)^2(.306126 \times 10^{-5})\right\}$$

$$= 68.501289$$

$$\sigma_{LD_{50}} = \sqrt{V_{LD_{50}}} = 8.276$$

b. Comparison of several probit regression lines.

The animals from which the data of Table 1 were collected consisted of 8 shipments,[9] each of which was subdivided into experimental groups and exposed to varying numbers of the 9 dosages. A probit regression line was established for each shipment and for the combined shipments (Figure 2). Section IIa gives the detailed calculation for determining the regression line based on the combined data. In table 5 are given the statistics resulting from

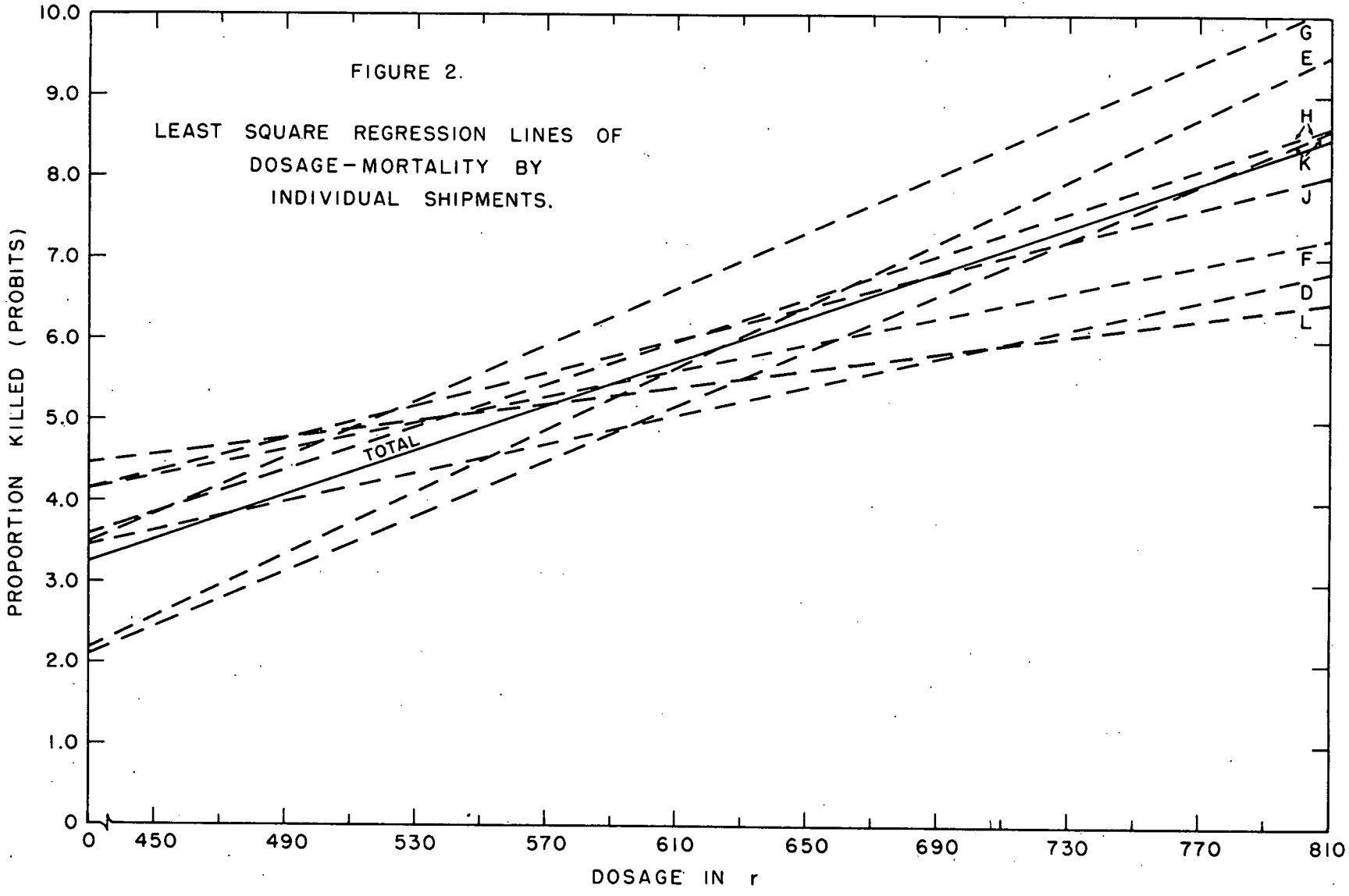


FIGURE 2.
LEAST SQUARE REGRESSION LINES OF DOSAGE-MORTALITY BY INDIVIDUAL SHIPMENTS.
PROPORTION KILLED (PROBITS)
DOSAGE IN r
TOTAL
G
E
H
K
J
F
D
L
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0
0
450
490
530
570
610
650
690
730
770
810

a probit analysis of this data by shipment.[10] The statistics included in Table 4 were based on least-square fitted provisional lines, and therefore differ slightly from the results tabulated in Table 8, ref. 10, which were determined from visually fitted lines.

Table 5

| Shipment* s | $N_s$ | $b \pm \sigma_b$ | $a \pm \sigma_a$ | $LD_{50} \pm \sigma_{LD_{50}}$ | $\overline{V}_{y_{w_s}}$ |
|---|---|---|---|---|---|
| D | 5 | .0088 ± .0055 | 4.72 ± .25 | 603.74 ± 34.58 | 2.8066 |
| E | 4 | .0189 ± .0013 | 5.17 ± .06 | 577.62 ± 3.22 | 0.1008 |
| F | 5 | .0079 ± .0082 | 5.20 ± .54 | 536.72 ± 73.23 | 11.9057 |
| G | 3 | .0176 ± .0007 | 5.13 ± .03 | 516.28 ± 2.01 | 0.0236 |
| H | 5 | .0131 ± .0065 | 4.96 ± .29 | 537.60 ± 22.25 | 3.7252 |
| J | 7 | .0101 ± .0003 | 5.56 ± .18 | 515.64 ± 22.60 | 1.6090 |
| K | 5 | .0153 ± .0038 | 4.98 ± .16 | 599.79 ± 10.79 | 1.1171 |
| L | 6 | .0055 ± .0022 | 5.23 ± .12 | 536.78 ± 28.15 | 0.8638 |
| Combined | 8 | .0133 ± .0018 | 5.095 ± .109 | 559.53 ± 8.28 | |

*Shipments for which either 100 per cent survival or 100 per cent mortality resulted were omitted.

To test whether the response pattern of individual shipments represents populations for which the dosage-mortality dependence is the same, analysis of variance is used.*(11) The pooled variance ($V_{y_w}$) from individual shipments is found by use of the expression

$$V_{y_w} = \frac{\sum_{s=D}^{L} (N_s - 2)\overline{V}_{y_{w_s}}}{\sum_{s=D}^{L} (N_s - 2)} \qquad (13)$$

*The validity of this procedure can be questioned since the variances $\overline{V}_{y_{w_s}}$ (Table 4) are not constant. A variance ratio close to the criterion accepted would contain no statistical information.

$$\overline{V}_{y_{w_s}} = \frac{\left\{\Sigma w_s Y_{w_s} - \frac{\left(\Sigma w_s Y_w\right)^2}{\Sigma w_s}\right\} - \left\{\frac{\left(\Sigma w_s Y_{w_s} - \Sigma w_s Y_{w_s} \Sigma w_s X_s\right)^2}{\Sigma w_s X_s^2 - \frac{\left(\Sigma w_s X_s\right)^2}{\Sigma w_s}}\right\}}{(N_s - 2)} \qquad (14)$$

The symbols used in equations (13) and (14) are defined as follows:

$y_{w_s}$ = the working probits for shipment s (s runs from D through L)

$N_s$ = the number of dose groups in shipment s

$x_s$ = the dosages from shipment s

$w_s$ = corresponding weights attached to probits in shipment s

$\overline{V}_{y_{w_s}}$ = variance from regression of weighted probits in shipment s

The difference between the sums of squares associated with $V_{y_w}$ and $V_{\acute{y}_w}$, the variance of the group working probits from the combined regression line (10), divided by the difference between the corresponding degrees of freedom gives an independent estimate ($V_{\acute{\acute{y}}_w}$ of the variance residual which is then compared with $V_{y_w}$.

$$V_{\acute{\acute{y}}_w} = \frac{V_{y_w}\Sigma(N_s-2) - V_{\acute{y}_w}(N-2)}{\Sigma(N_s-2) - (N-2)} \qquad (15)$$

(N = dosage groups in combined data)

The variance ratio $\frac{V_{y_w}}{V_{\acute{\acute{y}}_w}}$ is computed and the probability of its occurrence determined.

Table 6

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | F | P |
|---|---|---|---|---|---|
| From individual shipment lines | $V_{y_w}\Sigma(N_s-2)$ | $\Sigma(N_s-2)$ | $V_{y_w}$ | $\frac{V}{V''}$ | Table V ref. 8 |
| From combined data line | $V_{\acute{y}_w}(N-2)$ | $(N-2)$ | $V_{\acute{y}_w}$ | | |
| Residual | $V_{y_w}\Sigma(N_s-2) - V_{\acute{y}_w}(N-2)$ | $\Sigma(N_s-2) - (N-2)$ | $V_{\acute{\acute{y}}_w}$ | | |

Values of the quantities in Table 6 for this data are entered in Table 7.

Table 7

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | F | P |
|---|---|---|---|---|---|
| From individual shipment lines | 70.3894 | 24 | 2.9329 | 1.291 | P>.20* |
| From combined data line | 29.5050 | 6 | 4.9175 | | |
| Residual | 40.8844 | 18 | 2.2714 | | |

*From Table V, ref. 8, for $n_1 = 24$, $n_2 = 18$.

The probability P of the chance occurrence of such a difference between the shipment trends and the total group trend is greater than 20 per cent. Thus, with the 5 per cent mark as the significance criterion, the combined regression adequately describes the response behavior of this set of shipment groups.

## III. THE ANGULAR TRANSFORMATION

A population, each of whose elements possess one of two mutually exclusive and exhaustive attributes can be described by the binomial distribution. If p represents the proportion of the population possessing one of two attributes, then $q = 1 - p$ is the proportion possessing the other attribute. The probability of finding any combination of the two attributes in random samples of n elements is given by (2). The variances of p ($V_p = \frac{pq}{n}$), is a function of both p and n. To eliminate the dependence of the variance on p, a transformation which substitutes an angle $\phi$ for the proportion p is performed. If

$$p = \sin^2\phi \qquad 0 \le p \le 1 \qquad 0 \le \phi \le \frac{\pi}{2}, \tag{16}$$

the variance of the transform $\phi$ is approximately*

$$V\phi = \begin{cases} \frac{1}{4n} & (\phi \text{ measured in radians}) \\ \frac{820.7}{4n} & (\phi \text{ measured in degrees}) \end{cases} \tag{17}$$

*The proportion p is generally determined empirically and approximates the corresponding population proportion to a goodness that depends on the size of sample from which the estimate is made.

The variance (17) is still dependent on the number (n) of elements in the sample used, but by designing an experiment so that the resulting sample numbers are equal, homoscedacity is achieved. The dosage-response relationship is essentially linear over most of the dosage range.

The angular or arc-sine transformation is applied to the 20-day-survival data used in section II and the weighted* regression equation is calculated in the usual manner for dosage (x) versus the angular transform of the proportion killed ($y = \phi$). The $LD_{50}$ and associated parameters are computed. A table of values and the necessary computations follow.

Table 8

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| x | n | d | p | $y = \phi$* | $w = \frac{n}{820.7}$ |
| 467 | 110 | 20 | .182 | 25.2 | .1340 |
| 500 | 112 | 17 | .152 | 23.0 | .1365 |
| 533 | 127 | 38 | .299 | 33.1 | .1547 |
| 567 | 112 | 56 | .500 | 45.0 | .1365 |
| 600 | 124 | 102 | .822 | 65.0 | .1511 |
| 633 | 110 | 86 | .782 | 62.1 | .1340 |
| 667 | 126 | 118 | .936 | 75.4 | .1535 |
| 733 | 94 | 93 | .989 | 84.1 | .1145 |

*Table 12, reference 4

$N = 8$

$\Sigma w = 1.1148$ $\quad\quad$ $\Sigma wxy = 35,297.3013$

$\Sigma wx = 652.4736$ $\quad\quad$ $\Sigma wy = 57.1256$

$\Sigma wx^2 = 389,079.1208$ $\quad\quad$ $\Sigma wy^2 = 3,440.8725$

$\bar{x} = \frac{\Sigma wx}{\Sigma w} = 585.28$

For the weighted least square line $y = a + b(x - \bar{x})$

*Weights (reciprocal of the corresponding variances) were assigned each dose group since the number of animals between experimental groups differed.

$$b = \frac{\Sigma wxy - \frac{\Sigma wx \Sigma wy}{\Sigma w}}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}} = .25871$$

$$a = \bar{y} = \frac{\Sigma wy}{\Sigma w} = 51.243$$

$$y = 51.243 + .25871\ (x - 585.28) = .25871x - 100.17$$

$$V_y = \frac{\Sigma wy^2 - \frac{(\Sigma wy)^2}{\Sigma w} - \left\{ \frac{\left( \Sigma wxy - \frac{\Sigma wx \Sigma wy}{\Sigma w} \right)^2}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}} \right\}}{N - 2} = 5.2886 \qquad (18)$$

$$V_b = \frac{V_y}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}} = 7.346 \times 10^{-4}; \qquad \sigma_b = \sqrt{V_b} = .02710$$

$$V_a = \frac{V_y}{\Sigma w} = 4.7440; \qquad \sigma_a = \sqrt{V_a} = 2.1781$$

$$LD_{50} = \bar{x} + \frac{(45.0 - a)}{b} = 561.15 \text{ r} \qquad (\text{For } p = .50, \phi = 45.0°)$$

$$LD_{50} = \frac{1}{b^2} \left[ V_a - V_b (LD_{50} - \bar{x})^2 \right] = 77.2754 \text{ r}; \quad \sigma LD_{50} = \sqrt{V_{LD_{50}}} = 8.7906$$

## IV. THE LOGIT TRANSFORMATION

The logistic function,

$$p = \frac{1}{1 + e^{a - bx}} \qquad a,b = \text{constants} \qquad (19)$$

which is frequently called the "growth function" because of its association with the description of population growth, has achieved wide use in investigations of physiochemical processes. Because of the closeness with which this function approximates the integrated normal curve, on which the probit transform is based, the logistic function as a theoretical model might well serve as a substitute for the longer probit method in the characterization of "all or none" responses.(12)

From (19) a relationship linear in x is easily deduced.

$$\ell = \ln \frac{(1 - p)}{p} = a - bx \qquad (20)$$

The quantity $\ell = \ln \frac{(1 - p)}{p}$ is called a logit and transforms mortality data into a linear function of the dosages when p is the observed proportion dead at each dosage (x). With weight, w = npq, given to each observed mortality (p = 1 - q) a least square solution of the dosage-mortality dependence can be obtained. Details of the application of this transformation to our example are given in Table 9.

Table 9

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| x | n | d | p | q | y | w | $y_1$ | $P_c$ |
| 467 | 110 | 20 | .182 | .818 | 1.503 | 16.38 | 2.054 | .114 |
| 500 | 112 | 17 | .152 | .848 | 1.718 | 14.44 | 1.324 | .210 |
| 533 | 127 | 38 | .299 | .701 | .852 | 26.62 | .593 | .356 |
| 567 | 112 | 56 | .500 | .500 | .000 | 28.00 | - .160 | .540 |
| 600 | 124 | 102 | .822 | .178 | -1.530 | 18.14 | - .890 | .708 |
| 633 | 110 | 86 | .782 | .218 | -1.277 | 18.75 | -1.621 | .835 |
| 667 | 126 | 118 | .936 | .064 | -2.683 | 7.07 | -2.374 | .915 |
| 733 | 94 | 93 | .989 | .011 | -4.500 | 1.02 | -3.835 | .979 |

$N = 8$

$\Sigma w = 130.42$ $\qquad$ $\Sigma wxy_1 = 11{,}836.232$

$\Sigma wx = 73{,}150.02$ $\qquad$ $\Sigma wy_1 = -3.150$

$\Sigma wx^2 = 41{,}483{,}157.76$ $\qquad$ $\Sigma wy_1^2 = 243.535$

$\bar{x} = 560.88$ $\qquad$ $\bar{y} = a = -.02415$

$$b = \frac{\Sigma wxy_1 - \frac{\Sigma wx \Sigma wy_1}{\Sigma w}}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}} = -.022142$$

$$y_1 = -.02415 - .022142\ (x - 560.88) = 12.3949 - .022142\ x \qquad (21)$$

$$V_y = \frac{\Sigma wy_1^2 - \frac{(\Sigma wy_1)^2}{\Sigma w} - \left\{ \frac{\Sigma wy_1 x - \frac{\Sigma wy_1 \Sigma wx}{\Sigma w}}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}} \right\}}{N - 2} = 3.4415$$

$$V_b = \frac{V_y}{\Sigma wx^2 - \frac{(\Sigma wx)^2}{\Sigma w}} = 7.567 \times 10^{-6}; \qquad \sigma_b = .002751$$

$$V_a = \frac{V}{\Sigma w} = .02639; \qquad \sigma_a = .162$$

$$LD_{50} = \bar{x} - \frac{a}{b} = 559.79$$

$$V_{LD_{50}} = \frac{1}{b^2}\left[V_a + (LD_{50} - \bar{x})^2 V_b\right] = 5.3839; \qquad \sigma_{LD_{50}} = 2.32$$

## V. COMPARISON OF THE THREE TRANSFORMATIONS

The probit, angular and logit transformations were applied to the same data and the $LD_{50}$ and associated standard error computed. These values are listed in Table 10.

Table 10

| Transformation | $LD_{50}$ (in r) | $\sigma LD_{50}$ |
|---|---|---|
| Probit | 559.53 | 8.28 |
| Angular | 561.15 | 8.79 |
| Logit | 559.79 | 2.32 |

Close agreement between computed values of the $LD_{50}$ is evident. The logit transformation gives an appreciably lower standard error for the $LD_{50}$ than the probit or angular transformations. This result is in accord with the findings of other investigators. However, the probit method seems to be the best and most complete way to analyze response data. This method is based on the assumption that the susceptibilities of a population to some stimulant are normally distributed with respect to a measure of the stimulant given. Impetus is given to the acceptance of this model by empirical verification. Although dosage groups for which a 0 or 100 per cent mortality resulted were not considered in the analysis, no difficulty is encountered when such groups are included.

If, however, an experimental investigation of response is restricted to a set of dosages close to the median lethal dosage, the angular and logit transforms closely approximate the probit transform in describing the mortality-dosage relationship (Figure 3) and provide a simple means of analysis.

## LITERATURE CITED


(1) Bartlett, M. S.,"The Use of Transformations," Biometrics, Vol. 3, No. 1 (March 1947), 47-49

(2) Bliss, C. I.,"The Calculation of the Dosage-Mortality Curve," Annals of Applied Biology, 22 (1935) 134-167

(3) Bliss, C. I.,"The Comparison of Dosage-Mortality Data," Annals of Applied Biology, 22 (1935) 307-333

(4) Finney, D. J., Probit Analysis, Cambridge University Press 1947

(5) Mather, K., Statistical Analysis in Biology, Interscience Publisher Inc. N. Y. 1947

(6) Bliss, C. I., "The Determination of Dosage-Mortality Curves from Small Numbers," Quart. J. Pharm. 11 (1938) 192-216

(7) Fisher, R. A., Appendix to Bliss, C. I., "The Case of Zero Survivors," Annals of Applied Biology, 22 (1935) 164-5

(8) Fisher, R. A., and Yates, F., Statistical Tables for Biological, Agricultural and Medical Research, Oliver and Boyd, Edinburgh, 1938

(9) Tyler, S., and G. Sacher, Quarterly Report, Biological and Medical Divisions, August to November 1948 (ANL-4227) 44-45

(10) Tyler, S., Quarterly Report, Biological and Medical Divisions, November 1948 to February 1949 (ANL-4253) 61-64

(11) Tippett, L. H. C., The Methods of Statistics, Williams and Norgate Ltd. London, 1948, Chapter VIII

(12) Berkson, J., "Application of the Logistic Function to Bioassay," Journal of the American Statistical Association, 39 (1944) 357

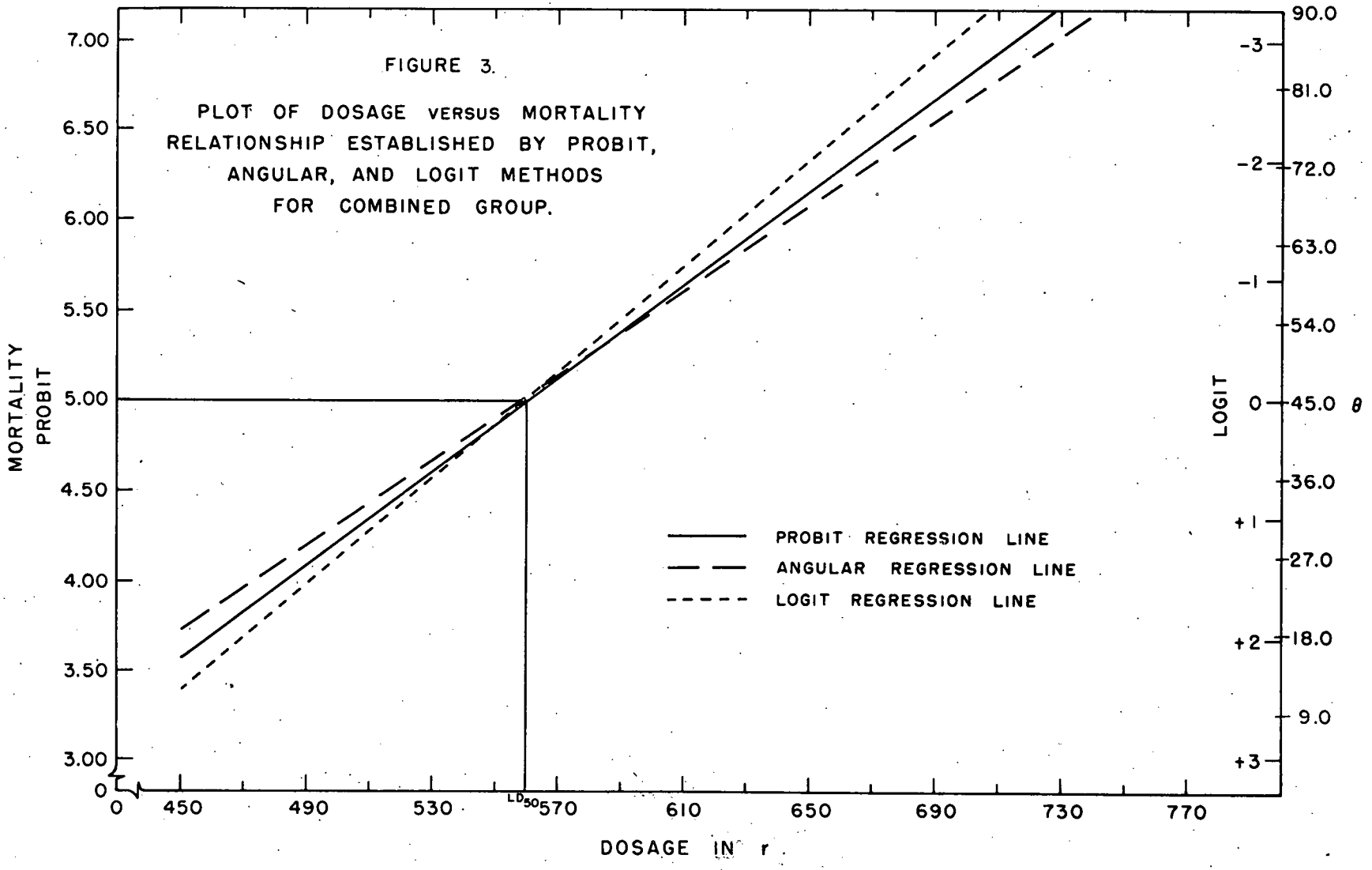
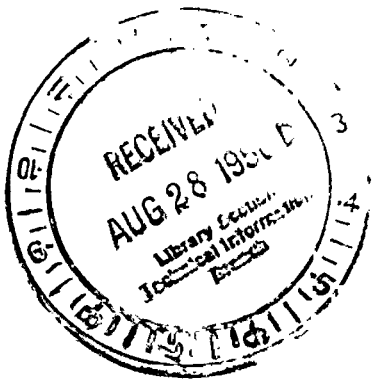



FIGURE 3.

PLOT OF DOSAGE VERSUS MORTALITY RELATIONSHIP ESTABLISHED BY PROBIT, ANGULAR, AND LOGIT METHODS FOR COMBINED GROUP.

1-22


RECEIVED
AUG 28 19
Library Section
Technical Information
Branch

OR Form 315 (Revised Feb. 8, 1950)

DOCUMENT COVER SHEET - MASTER COPY

LIBRARY BRANCH
TECHNICAL INFORMATION DIVISION
OAK RIDGE EXTENSION

REPORT NUMBER ANL-4486
CLASSIFICATION U
SUBJECT CATEGORY C-19

Requested By _____

INITIAL REVIEW BY MTLS. & INFO. BR.

| Received From | No. Cys. | Date Received | Loan | Retention |
|---|---|---|---|---|
| ANL | 15 | 8/28/50 | ☐ | ☒ |
| _____ | ___ | _____ | | |
| _____ | ___ | _____ | | |

Distribution Authorized By JaR 9/19/50

Category _____ As Indicated _____

List ☐ Abstract ☒ Do Not List Or Abstract ☐

Distribution Made By _____ Date _____

Comments: _____

| TID Dist. Auth'd On | | Installation | Subsequent Distribution | | | |
|---|---|---|---|---|---|---|
| Init. Review | Subs. Review | | Auth'd By | Date | Dist. Made By | Date |
| | | Argonne National Laboratory | | | | |
| | | Armed Forces Special Weapons Project | | | | |
| | | ATOMIC ENERGY COMMISSION, WASHINGTON | | | | |
| | | Atomic Energy Commission, Brookhaven Area | | | | |
| | | Battelle Memorial Institute | | | | |
| | | Brookhaven National Laboratory | | | | |
| | | Brush Beryllium Company | | | | |
| | | Bureau of Medicine & Surgery | | | | |
| | | Bureau of Ships | | | | |
| | | Bureau of Yards & Docks | | | | |
| | | Carbide & Carbon Chemicals Division (K-25 Plant) | | | | |
| | | Carbide & Carbon Chemicals Division (Y-12 Plant) | | | | |
| | | Carnegie Institute of Technology (Creutz) | | | | |
| | | CHICAGO OPERATIONS OFFICE | | | | |
| | | Cleveland Area Office | | | | |
| | | Columbia University (Dunning) | | | | |
| | | Columbia University (Failla) | | | | |
| | | Dow Chemical Company | | | | |
| | | H. K. Ferguson Company | | | | |
| | | General Electric Company, Richland | | | | |
| | | HANFORD OPERATIONS OFFICE | | | | |
| | | Idaho Operations Office | | | | |
| | | Iowa State College | | | | |
| | | Kirtland Air Force Base | | | | |
| | | Knolls Atomic Power Laboratory | | | | |
| | | Los Alamos | | | | |
| | | Mallinckrodt Chemical Works | | | | |
| | | Mass. Inst. of Tech. (Eliassen) | | | | |
| | | Mass. Inst. of Tech. (Gaudin) | | | | |
| | | Mass. Inst. of Tech. (Kaufmann) | | | | |

| TID Dist. Auth'd On | | Installation | Subsequent Distribution | | | |
|---|---|---|---|---|---|---|
| Init. Review | Subs. Review | | Auth'd By | Date | Dist. Made By | Date |
| | | Mound Laboratory | | | | |
| | | National Advisory Committee for Aeronautics | | | | |
| | | National Bureau of Standards | | | | |
| | | Naval Radiological Defense Laboratory | | | | |
| | | Naval Research Laboratory, Washington | | | | |
| | | NEPA Project | | | | |
| | | New Brunswick Laboratory | | | | |
| | | NEW YORK OPERATIONS OFFICE | | | | |
| | | North American Aviation, Inc. | | | | |
| | | Oak Ridge National Laboratory | | | | |
| | | PATENT BRANCH, WASHINGTON | | | | |
| | | Rand Corporation | | | | |
| | | Public Health Service | | | | |
| | | Sandia Corporation | | | | |
| | | SANTA FE OPERATIONS OFFICE | | | | |
| | | Sylvania Electric Company | | | | |
| | | TECHNICAL INFORMATION DIVISION, ORE | | | | |
| | | UCLA Medical Research Laboratory (Warren) | | | | |
| | | U. S. Air Force | | | | |
| | | U. S. Geological Survey | | | | |
| | | University of California Radiation Laboratory | | | | |
| | | University of Rochester | | | | |
| | | University of Washington | | | | |
| | | Western Reserve University (Friedell) | | | | |
| | | Westinghouse Electric Corporation | | | | |
| ✓ | 2 | 2LT AEC | 2mw | 4-19-51 | Cmb | 4-24-51 |

AEC, Oak Ridge, Tenn.-2-28-50--1500-A19904