Statistical Analysis of Litter Experiments in Teratology

**R. Williams
R. L. Buschbom**

November 1982

**Prepared for the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830**

**Pacific Northwest Laboratory
Operated for the U.S. Department of Energy
by Battelle Memorial Institute**

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC06-76RLO 1830

Printed in the United States of America
Available from
National Technical Information Service
United States Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22151

NTIS Price Codes
Microfiche A01

Printed Copy

Pages	Price Codes
001-025	A02
026-050	A03
051-075	A04
076-100	A05
101-125	A06
126-150	A07
151-175	A08
176-200	A09
201-225	A010
226-250	A011
251-275	A012
276-300	A013

STATISTICAL ANALYSIS OF
LITTER EXPERIMENTS IN TERATOLOGY

R. Williams
R. L. Buschbom

November 1982

Prepared for
the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

Pacific Northwest Laboratory
Richland, Washington 99352

TABLE OF CONTENTS

TERATOLOGICAL EXPERIMENTS	1
METHODS OF ANALYZING TERATOLOGICAL DATA	3
OTHER CONSIDERATIONS.	9
SUMMARY	11
REFERENCES.	13
SUPPLEMENTAL BIBLIOGRAPHY	15

STATISTICAL ANALYSIS OF LITTER EXPERIMENTS IN TERATOLOGY

TERATOLOGICAL EXPERIMENTS

Teratology is the study of congenital defects in newborns caused by the mother's exposure to a chemical. The word derives from the Greek "teratos," meaning "monster." A teratogen is an agent which can cause birth defects. Known teratogens for humans include thalidomide, androgenic steroids and alcohol.

Suspected teratogenic agents are tested on laboratory animals. It is hoped that any substance which can have toxic effects on a human fetus will also have such effects on the experimental animals. (The case of thalidomide, which caused no birth defects when tested on laboratory mice, provides a tragic example of the dangers of such extra-species extrapolation.) The animals commonly employed in this type of testing are rats, mice and rabbits. The experiment usually proceeds by randomly assigning a number of pregnant females to the "control" and "exposed" groups. The exposure can be administered in several ways, including injection, inhalation or ingestion. These experiments can be more complicated with numerous exposure groups receiving different dosages. The animals are all sacrificed as they near term and the fetuses examined. At this time, experimenters may record many variables such as maternal weight gain, number of eggs implanted, number of dead fetuses, number of live fetuses that are malformed. The objective is to determine statistically if these measures differ significantly between the control and exposed groups, or exhibit a trend which corresponds to dosage.

For most of this discussion, we will restrict our attention to the simple case of a single exposure group versus a control group. Moreover, we will consider only two variables to be observed on each mother: the litter size (total number of fetuses) and the number of fetuses which are affected in a specific way (e.g., dead or malformed). Let k_0 be the number of females assigned to the control group and k_1 be the number in the exposed group. Let

n_{ij} = the total number of fetuses in the j^{th} litter of the i^{th} group ($i = 0$ is the control group, $i = 1$ denotes exposed group), $j = 1, 2, \dots, k_i$; while x_{ij} = the number of affected fetuses in this same litter. Let $p_{ij} = x_{ij}/n_{ij}$ be the proportion of affected fetuses in the $(i,j)^{\text{th}}$ litter. Let

$$x_{i.} = \sum_{j=1}^{k_i} x_{ij} \quad \text{and} \quad n_{i.} = \sum_{j=1}^{k_i} n_{ij} ; \quad i = 0, 1$$

be the total number affected and total number of fetuses, respectively, within each group.

A naive method of comparison of the groups is to pool the litter counts within each group, i.e., form $x_{i.}$ and $n_{i.}$, $i = 0, 1$, and then perform a contingency table analysis. In this simple case, the table is the following 2x2:

	Control	Treated
Affected	x_0	x_1
Not Affected	$n_0 - x_0$	$n_1 - x_1$
	n_0	n_1

A test of equality of treatment group proportions can be carried out by a chi-square test, a t-test, or the Fisher-Irwin Exact test. This type of analysis has been done by Lyon (1970) and Suter (1975). Such a procedure takes the individual fetuses as the experimental units and disregards their natural "nesting" in litters. It assumes a binomial distribution for the occurrence of affected fetuses within each group:

$$x_{i.} | n_{i.} \approx b(n_{i.}, p_i) ,$$

where p_i is the population proportion of affected fetuses in group i , and

$$\hat{p}_{i.} = \frac{x_{i.}}{n_{i.}} = \hat{p}_i .$$

These procedures are called "per fetus" tests.

Over the last several years, however, opinion has mounted that such a simple binomial model is inadequate to explain the extreme variation which can occur in the litter proportions p_{ij} within a fixed group. That is, the proportion of affected fetuses in a litter commonly varies considerable within a treatment group, sometimes even ranging from 0 to 1. This phenomena is taken as evidence of a "litter effect," that is, a tendency for litter mates to develop more like one another than like fetuses in other litters within the same treatment groups. To the extent that a litter effect exists, the fetuses of a single litter are not independent of each other and, as a consequence, individual fetuses are inappropriate as the experimental unit. Several studies (Kruger, 1970; Stapes and Haseman, 1974; Haseman and Soares, 1976) have concluded that the litter effect is real and that the litter proportion of affected fetuses often varies more than it should if the treatment group proportion were constant.

In light of this development, a variety of different approaches are being used or have been proposed for analysis of teratological data. The purpose of this report is to survey these statistical methods.

METHODS OF ANALYZING TERATOLOGICAL DATA

Table 1 contains data originally reported by Weil (1970) and later reexamined by Williams (1975). It shows the results of an experiment of the type described in Section 1. One group of 16 pregnant female rats was fed a control diet during pregnancy, while a second group of 16 pregnant females received a diet containing a chemical. For each litter, the values of n_{ij} , x_{ij} and p_{ij} are given. Note that the treated group exhibits the maximum possible amount of variation in its litter proportions.

For purposes of illustration, let us use two different methods to examine these data for a possible difference between the groups as regards the overall proportion of affected fetuses. First, we assume a simple binomial model for each group, pool the litters and create the contingency Table 2. The chi-squared test statistic which results is 8.25, which corresponds to a p-value < 0.005 . Thus, this approach strongly indicates a difference between

TABLE 1. Teratological data set.

CONTROL GROUP			TREATED GROUP		
Litter Size (n_{ij})	Number Affected (x_{ij})	Proportion Affected (p_{ij})	Litter Size (n_{ij})	Number Affected (x_{ij})	Proportion Affected (p_{ij})
13	0	.000	12	0	.000
12	0	.000	11	0	.000
9	0	.000	10	0	.000
9	0	.000	9	0	.000
8	0	.000	11	1	.091
13	1	.077	10	1	.100
12	1	.083	9	1	.111
10	1	.100	9	1	.111
10	1	.100	5	1	.200
9	1	.111	9	2	.222
13	2	.154	7	3	.429
5	1	.200	10	5	.500
7	2	.286	6	3	.500
10	3	.300	10	7	.700
10	3	.300	7	7	1.000
$n_0 = 158 \quad x_0 = 16 \quad p_0 = .101$			$n_1 = 144 \quad x_1 = 32 \quad p_1 = .222$		

TABLE 2. Contingency table of teratological data.

	Control	Treated
Affected	16	32
Not Affected	142	112

the groups. If, on the other hand, the litters are kept separate and the litter proportions p_{ij} are taken as the data for analysis, the Mann-Whitney test results in a test statistic which is not significant at the 0.100 level. This disparity clearly indicates that the method of analysis that is used with teratological data can make a great difference in the resulting conclusion. In particular, the issue of per fetus versus per litter tests is seen to be crucial.

As was stated in Section 1, it is now widely held that it is inappropriate to pool the litters within a group and perform a per fetus test. This assumes no litter effect at all, that is, independence of the fetuses. The result is a test based upon an overly simple model, and ultimately, a level of significance different from the nominal one. Haseman and Soares (1976) and Gladen (1979) each report simulation studies in which per fetus tests such as chi-squared and Fisher's Exact have actual α -levels 35-150% above the intended level. Hence, an experimenter using one of these procedures would be running an increased risk of erroneously finding a difference between the treatment groups when in fact there is none. In evaluating the two different results we obtained with the data in Table 1, it therefore appears that some, if not all, of the discrepancy is due to the chi-squared test attaining an excessive level of significance.

Of course, level of significance and power are opposite sides of the same coin - gains are made on one at the expense of the other. When two test procedures give different results, it is not always clear whether one is more powerful than the other or if it is operating at an inflated level of significance. The objective is to find a procedure, from among those which adhere reasonably well to the nominal α -level, which gives the greatest power.

With these thoughts in mind, we now shall consider the various alternatives which have been proposed for this type of data. Only per litter tests will be considered in light of what has been learned about the defects of per fetus tests. That is, we shall take the p_{ij} as the data to be compared.

There are two general types of tests: parametric and nonparametric. Parametric methods involve hypothesizing an underlying model for the

experimental data and the subsequent analysis entails estimating the parameters of the model. The quality of the resulting test is often directly related to the fit of the data to the assumed model. On the other hand, nonparametric procedures are model-free. As a general rule, they provide less power than a well suited parametric approach but a more stable level of significance when the data fits the model poorly.

The Mann-Whitney U Test is a commonly used nonparametric procedure (Yasuda and Maeda, 1972; Gaylor, 1978; Soares and Sheridan, 1975) for comparing the litter proportions of affected fetuses from the two treatment groups. It is simple and well known. It avoids the problem of fitting the data to a model, either simple or complex. Its power appears to be about the same as the simpler parametric approaches we will consider but lags well behind some of the more sophisticated ones. A minor drawback of the Mann-Whitney test is the occasional occurrence of data containing many ties at an extreme value such as $p_{ij} = 0$. The test assumes the random variables of interest are continuous and thus should have few if any ties. This problem has been addressed, however, by Haseman and Hoel (1974), who have proposed a version of the Mann-Whitney test that is conditional on the total number of observed ties.

The parametric approaches to the analysis of teratological data fall into two categories themselves. The first category consists of the various "transform-then-ANOVA" techniques. These methods transform the litter proportions in some way,

$$z_{ij} = f(p_{ij}) \quad ,$$

in the hope of creating samples which satisfy the assumptions of the t-test for independent samples. Specifically, the samples must be approximately normally distributed and possess approximately equal variances. Commonly used transformations include the square root,

$$z_{ij} = p_{ij}^{1/2} \quad ,$$

the arcsine (Chrisman, 1974)

$$z_{ij} = \begin{cases} \sin^{-1}(p_{ij})^{1/2} & , \quad \text{if } 0 < p_{ij} < 1 \\ \sin^{-1}\left(\frac{1}{4n_{ij}}\right)^{1/2} & , \quad \text{if } p_{ij} = x_{ij} = 0 \\ \sin^{-1}(1) - \sin^{-1}\left(\frac{1}{4n_{ij}}\right) & , \quad \text{if } p_{ij} = 1 \end{cases}$$

and the Freeman-Tukey binomial arcsine (Biddle, 1975),

$$z_{ij} = \frac{1}{2} \left(\sin^{-1}\left(\frac{x_{ij}}{n_{ij} + 1}\right)^{1/2} + \sin^{-1}\left(\frac{x_{ij} + 1}{n_{ij} + 1}\right)^{1/2} \right) .$$

The validity of this approach rests, of course, on how well the transformed data satisfy the assumptions of the t-test. Haseman and Soares (1976) and Gladen (1979) report that their large scale simulation studies indicate that the arcsine transform, followed by the t-test, gives a level of significance which is reliable and power which is comparable to that of Mann-Whitney. However, when the data of Table 1 is transformed by the arcsine, the result is samples with variances of 0.0256 (control) and 0.1296 (treated). These values refute the hypothesis of equal population variances at the 0.005 level of significance, so that a t-test using these z_{ij} appears ill-founded. One must bear in mind that any analysis based upon transformed data is "ad hoc"; the transformed data must always be scrutinized to determine if it has the desired properties.

The second type of parametric approach is to model the p_{ij} themselves. As has already been seen, an adequate model will need to be, in general, more complicated than a simple binomial distribution. Williams (1975) proposed using the beta-binomial distribution. In this procedure, it is assumed that within each litter the binary responses form a set of Bernoulli trials whose probability parameter varies between litters in the same treatment group rather than remains constant. In this case, the probability parameter is assumed to vary according to a two-parameter beta distribution. That is

$$x_{ij}|n_{ij} \approx \beta(n_{ij}, p_i) \quad ,$$

where p_i is the unknown binomial parameter for treatment i , and

$$p_i \approx \text{beta}(\alpha, \beta) \quad .$$

In several cases the beta-binomial model has provided a much better fit to observed sets of data compared to the binomial model (Aeschbacher et al., 1977; Kupper and Haseman, 1978). However, in some cases its lack of fit remained statistically significant (Haseman and Jupper, 1979). Use of a mixing distribution with the binomial is appealing in that it directly addresses the problem of excessive variation in the litter proportions of affected fetuses. On the other hand, the beta distribution appears to have been chosen purely for mathematical convenience. The resulting procedure is more complicated computationally than the others we have considered, involving as it does maximizations with boundary conditions. It is also far less familiar to non-statisticians. The question is whether these negative aspects are outweighed by the gain in power which the test procedure promises. There is evidence that this increased power is realized. For example, when beta-binomial distributions are fit to each of the treatment groups in Table 1, and the hypothesis of identical distributions tested, the resulting test statistic exceeds the upper 2.5 percent point of the chi-square distribution with two degrees of freedom (see Williams, 1975, for computational details. Recall that the Mann-Whitney test finds the difference not significant at the 10 percent level.

Kupper and Haseman (1978) have proposed an alternative to the beta-binomial model, the so-called "correlated-binomial." This is similar to the beta-binomial in that it allows non-zero correlations between the response of the fetuses but it extends the concept by permitting these correlations to be negative as well as positive. Kupper and Haseman performed a study in which both the beta-binomial and the correlated-binomial model were fit to several real-life data sets. In each case there was very little difference in the fits provided by the two models and both were a marked improvement over the ordinary binomial distribution.

OTHER CONSIDERATIONS

The jackknife procedure for estimating a treatment group proportion p_i from the litter proportions p_{ij} , $j = 1, 2, \dots, k_i$, has been proposed by Gladen (1979). This estimator is given by

$$\tilde{p}_i = \hat{p}_i + \frac{k_i - 1}{k_i} \left[\sum_{j=1}^{k_i} \frac{x_{ij} - n_{ij}\hat{p}_i}{n_{i.} - n_{ij}} \right],$$

where $\hat{p}_i$ is any initial estimate of p_i (say, $\hat{p}_i = x_{i.}/n_{i.}$). The jackknife has two desirable properties. First, if the initial estimator $\hat{p}_i$ is a biased estimator of p_i , then the jackknifed estimator $\tilde{p}_i$ has bias of a lesser order. The second property is that the jackknife procedure also produces an easily-computed estimate of the standard error of $\tilde{p}_i$, and it is this fact that leads one to hope for a more powerful t-test. In essence, this approach transforms only the mean of each sample rather than transforming all of the p_{ij} and then calculating the sample means. In either case a t-test (or ANOVA) is performed to test for differences. Gladen's own simulation studies, however, indicate that the jackknife method is comparable in power to the various standard transformation.

An important matter that is little discussed in the literature is that of sample size: how many litters are needed in each treatment group to be able to have a certain probability of detecting a desired difference in effect rates? Of course, in practice such decisions are often dictated by considerations such as cost or the size of the animal care facility. It would be helpful nonetheless to have some idea of the optimal size in various situations. A table (reprinted from Busch and Leidel, 1979) containing some of this information appears on page 10 of this report. This table gives the sample size required to detect a difference $\Delta = p_2 - p_1$, with power 0.95 when using a one-tailed binomial test at $\alpha = 0.05$. The test upon which the table is based is a per fetus test so that in many instances, as we have already seen, its actual level of significance is higher than the stated one. Hence, for teratological data, we should expect the values of k_i given in this table to be understated. The matter of sample size requirements for the Mann-Whitney

Sample size (per group) to detect a difference
between binomial incidences θ_1 and θ_2 .*

θ_1	Δ = Increase in Incidence								
	.01	.015	.02	.05	.10	.20	.25	.30	.50
0	299	199	149	59	29	14	11	9	5
.00001	576	379	282	110	54	26	20	17	9
.00005	623	403	297	114	55	26	21	17	9
.0001	658	423	310	117	56	27	21	17	9
.0005	840	517	368	130	61	28	22	18	10
.001	1030	605	423	142	65	29	23	18	10
.005	2002	1071	699	197	81	34	26	21	11
.01	3113	1577	992	250	96	39	31	23	12
.02	5220	2525	1529	342	121	46	34	26	13
.05	11182	5156	3025	585	183	61	44	29	15
.10	20123	9166	5286	937	269	82	56	41	17
.15	28012	12754	7262	1243	342	98	66	48	19
.20	35199	15986	9091	1509	403	111	74	53	20
.25	49023	18294	10411	1726	479	121	79	56	20
.30	45550	20370	11494	1898	490	128	83	58	20
.40	55121	23425	13134	2114	535	134	85	58	17
.50	55220	24378	13531	2165	534	128	79	53	5

*Symbols as follows:

θ_1 = lower incidence

θ_2 = higher incidence

Δ = $\theta_1 - \theta_2$

size (significance probability) = .05 (1-tail test)

power = .95 (probability of detecting a
significant effect)

Ref: Chapter 3: "Statistical Design and Data Analysis Requirements,"
by Kenneth A. Busch and Nelson A. Leidel, from Patty's Industrial
Hygiene and Toxicology, Volume 3, Wiley, 1979.

test or the more sophisticated parametric approaches is either terribly complicated or impossible except in the presence of powerfully simplifying assumptions. Let us also remark, regarding the use of the table, that teratological experimenters commonly assume p_0 , the incidence of malformations in control groups of rats and rabbits, to be on the order of 0.001 (1 per 1000).

Finally, we briefly consider the situation in which there are more than two treatment groups, e.g., a control and several exposure levels, or several different methods of exposure. Many of the procedures that have been discussed, such as the t-test and the Mann-Whitney, have natural extensions (the ANOVA and the Kruskal-Wallis test, respectively) to the multiple sample experiment. Crowder (1978) has discussed a beta-binomial ANOVA for proportions. In the particular case in which the treatment groups are increasing exposure levels of a substance, the alternative hypothesis of interest is apt to be the "one-sided" statement that there is an increasing trend in the incidence of the effect. Johckheere's test (Lin and Haseman, 1976) offers a straightforward nonparametric procedure having this alternative.

The problem of the level of significance must be kept in mind when one wishes to make several treatment versus control pairwise comparisons: if each of several statements has probability p of being true, then the probability is not p that all of them are true. An adjustment can be made in the α -level of the individual pairwise comparisons in order that the overall level is roughly the desired one. The Bonferroni method, in which each of the N tests are conducted at the α/N level of significance, does this. Otherwise, a test such as Dunnett's, which automatically maintains the overall level of significance, should be used.

SUMMARY

Teratological data is binary response data (each fetus is either affected or not) in which the responses within a litter are usually not independent. As a result, the litter should be taken as the experimental unit. For each litter, its size, n , and the number of fetuses, x , possessing the effect of

interest are recorded. The ratio $p = x/n$ is then the basic data generated by the experiment.

There are currently three general approaches to the analysis of teratological data: nonparametric, transformation followed by t-test or ANOVA, and parametric. The first two are currently in wide use by practitioners while the third is relatively new to the field. These first two also appear to possess comparable power levels while maintaining the nominal level of significance. When transformations are employed, care must be exercised to check that the transformed data has the required properties. Since the data is often highly asymmetric, there may be no transformation which renders the data "nearly normal." The parametric procedures, including the beta-binomial model, offer the possibility of increased power.

Thus the teratological experimenter today has a choice among several acceptable approaches. Which of them will eventually become preeminent remains to be seen. Perhaps an entirely different approach will be developed. Whatever happens, the situation is in flux at present.

REFERENCES

- Aeschbacher, H. W., Vuataz, L., Sotek, J., and Stalder, R. 1977. Use of the beta-binomial distribution in dominant-lethal testing for "weak" mutagenic activity. Mutation Research 44, 369-390.
- Biddle, F. G. 1975. Teratogenesis of acetazolamide in the CBA/J and SWV strains of mice. II. Genetic control of the teratogenic response. Teratology 11, 37-46.
- Busch, K. A. and Leidel, N. A. 1979. "Statistical Design and Data Analysis Requirements." Chapter 3 of Patty's Industrial Hygiene and Toxicology, Vol. 3, John Wiley.
- Chrisman, C. L. 1974. Aneuploidy in mouse embryos induced by diethylstilbestrol diophosphate. Teratology 9, 229-232.
- Crowder, M.J. 1978. Beta-binomial ANOVA for proportions. Applied Statistics 27, 34-37.
- Gaylor, D. W. 1978. "Methods and Concepts of Biometrics Applied to Teratology." Chapter 14 of Handbook of Teratology, Vol. 4: Research Procedures and Data Analysis. Wilson and Fraser (eds.), Plenum Press.
- Gladden, B. 1979. The use of jackknife to estimate proportions from toxicological data in the presence of litter effects. Journal of the American Statistical Association 74, 278-283.
- Haseman, J. K. and Hoel, D. G. 1974. Tables of Gehan's generalized Wilcoxon test with fixed point censoring. Journal of Statistical Computation and Simulation 3, 117-135.
- Haseman, J. K. and Kupper, L. L. 1979. Analysis of dichotomous response data from certain toxicological experiments. Biometrics 35, 281-293.
- Haseman, J. K. and Soares, E. R. 1976. The distribution of fetal death in control mice and its implications on statistical tests for dominant lethal effects. Mutation Research 41, 277-288.
- Kruger, J. 1970. Statistical methods in mutation research. In Chemical Mutagenesis in Mammals and Man. Vogel and Rohrborn (eds.), Springer-Verlag.
- Kupper, L. L. and Haseman, J. K. 1978. The use of a correlated binomial model for the analysis of certain toxicological experiments. Biometrics 34, 69-76.
- Lin, F. O. and Haseman, J. K. 1976. A modified Jonckheere test against order alternatives when ties are present at a single extreme value. Biometrische Zietschrift 18, 623-631.

- Lyon, M. F. 1970. X-ray induced dominant lethal mutations in male guinea pigs, hamsters and rabbits. Mutation Research 10, 133-140.
- Soares, E. R. and Sheridan, W. 1975. Lack of induction of dominant lethals in mice by orally administered AF-2. Mutation Research 31, 235-240.
- Staples, R. E. and Haseman, J. K. 1974. Selection of the appropriate experimental units in teratology. Teratology 9, 259-260.
- Suter, K. E. 1975. Studies on the dominant-lethal and fertility effects of the heavy metal compounds methylmercuric hydroxide, mercuric chloride and cadmium chloride in male and female mice. Mutation Research 30, 365-374.
- Weil, C. S. 1970. Selection of the valid number of sampling units and a consideration of their combinations in toxicological studies involving reproduction, teratogenesis or carcinogenesis. Food and Cosmetics Toxicology 8, 177-182.
- Williams, D. A. 1975. The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. Biometrics 31, 949-952.
- Yasuda, M. and Maeda, H. 1975. Teratogenic effects of 4-chloro-2-methylphenoxyacetic acid ethylester (MCPEE) in rats. Toxicology and Applied Pharmacology 23, 326-333.

SUPPLEMENTAL BIBLIOGRAPHY

Haseman, J. K. and Hogan, M. D. 1975. Selection of the experimental unit in teratology studies. Teratology 12, 165-172.

Haseman, J. K. 1975. "Statistical Aspects of Teratology Studies." Presented at the Fourth Conference of the European Teratology Society, Freudenstadt, Federal Republic German, September 1975.

Jonckheere, A. R. 1954. A distribution-free k-sample test against ordered alternatives. Biometrika 41, 133-145.

DISTRIBUTION

No. of
Copies

OFFSITE

Dr. George E. Stapleton
Health Effects Research Division
U.S. Department of Energy
Washington, DC 20545

27 DOE Technical Information Center

ONSITE

DOE Richland Operations Office

H. E. Ransom

46 Pacific Northwest Laboratory

W. J. Bair
R. L. Buschbom (20)
D. B. Carr
H. Drucker
M. F. Gillis
P. L. Hackett
B. J. Kelman
J. A. Mahaffey
D. D. Mahlum
S. Marks
A. R. Olsen
J. F. Park
R. A. Pelroy
R. D. Phillips
D. N. Rommereim
L. B. Sasser
M. R. Sikov
D. L. Springer
D. L. Stevens
H. D. Tolley
Publishing Coordination (2)
Technical Information (5)