12/11/26/91 Ans (1)

PNL-7867
UC-605

Some Computer Simulations Based on the Linear Relative Risk Model

E. S. Gilbert

October 1991

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

Battelle

PNL-7867

DISCLAIMER

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

Printed in the United States of America

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831; prices available from (615) 576-8401. FTS 626-8401.

Available to the public from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161.

PNL--7867

DE92 003355

SOME COMPUTER SIMULATIONS BASED ON THE LINEAR RELATIVE RISK MODEL

E. S. Gilbert

October 1991

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

Pacific Northwest Laboratory
Richland, Washington 99352

SUMMARY

This report presents the results of computer simulations designed to evaluate and compare the performance of the likelihood ratio statistic and the score statistic for making inferences about the linear relative risk model. The work was motivated by data on workers exposed to low doses of radiation, and the report includes illustration of several procedures for obtaining confidence limits for the excess relative risk coefficient based on data from three studies of nuclear workers.

The computer simulations indicate that with small sample sizes and highly skewed dose distributions, asymptotic approximations to the score statistic or to the likelihood ratio statistic may not be adequate. For testing the null hypothesis that the excess relative risk is equal to zero, the asymptotic approximation to the likelihood ratio statistic was adequate, but use of the asymptotic approximation to the score statistic rejected the null hypothesis too often. Frequently the likelihood was maximized at the lower constraint, and when this occurred, the asymptotic approximations for the likelihood ratio and score statistics did not perform well in obtaining upper confidence limits. The score statistic and likelihood ratio statistics were found to perform comparably in terms of power and width of the confidence limits. It is recommended that with modest sample sizes, confidence limits be obtained using computer simulations based on the score statistic.

Although nuclear worker studies are emphasized in this report, its results are relevant for any study investigating linear dose-response functions with highly skewed exposure distributions.

CONTENTS

TABLES

1.0 INTRODUCTION

Relative risk regression models have been investigated and applied extensively since Cox (1972) first proposed his lifetable regression model. In the form originally suggested by Cox, the relative risk was assumed to be an exponential function of the covariates, and this form has been shown to have many desirable statistical properties. More general forms of the relative risk function have subsequently been proposed (Barlow 1985, Breslow and Storer 1985, Moolgavkar and Venzon 1987, Thomas 1981).

The linear or additive relative risk model, in which the relative risk is a linear function of the covariates, has received special attention. Thomas (1981) and Moolgavkar and Venzon (1987) include discussion of this model in their consideration of more general models, and note that the distribution of the maximum likelihood estimate of the linear excess relative risk coefficient is likely to be highly skewed, and that confidence intervals based on inverse information can be seriously misleading. These investigators indicate that basing inferences on the asymptotic chi-square distribution for the likelihood ratio statistic is a preferable approach.

Prentice and Mason (1986) investigated the behavior of statistical procedures based on the linear relative risk model by conducting computer simulations. These simulations were based on a model with a single covariate with 24 sampling configurations and samples sizes (number of failures) of 50 and 100, and confirm the inadequacy of confidence intervals for the linear risk coefficient based on inverse information. The chi-square approximations to the score statistic with expected information or to the likelihood ratio statistic were found to be reasonably satisfactory, but the approximation to the score statistic with observed information was not satisfactory.

Prentice and Mason also note the need to constrain linear relative risk coefficients to avoid negative relative risks, and consider a reparametrization of the form $\alpha = \ln(\beta + \beta_0)$, where β

is the excess relative risk coefficient, and β_0 might be chosen as the reciprocal of the largest possible observation in the study. Simulations conducted to evaluate confidence intervals based on the inverse information for α indicated that these intervals are likely to be overly conservative, but the score statistic based on α with observed information was found to perform satisfactorily. The score statistic with expected information and the likelihood ratio statistic under this reparametrization are identical to those for the original parameter β.

Linear dose-response functions have proved especially useful for describing health risks resulting from exposure to radiation. Linear functions have provided an adequate fit to data from several epidemiologic studies, and have provided a practical expression of risk for radiation protection purposes. In addition, a linear model has been justified by certain biological and mechanistic arguments. Recent applications of the linear relative risk model in radiation studies include analyses of the Japanese A-bomb survivors (National Academy of Sciences 1990, Preston and Pierce 1987, Shimizu et al. 1990), and analyses of data on underground miners exposed to radon progeny (National Academy of Sciences 1988, Whittemore and McMillan 1983).

The problem that motivated this report was the need to provide a direct assessment of the effects of low-level radiation exposure based on data from epidemiologic studies of persons exposed at low-levels, particularly studies of nuclear workers. Several studies of workers exposed occupationally to radiation have been conducted (Beral et al. 1985, Beral et al. 1988, Checkoway et al. 1985, Gilbert et al. 1989, Howe et al. 1987, Smith and Douglas 1986, Wilkinson et al. 1987). Generally these populations have been exposed at levels where statistically detectable effects would not be expected if linear extrapolation from data on populations exposed at high levels (such as the Japanese A-bomb survivors) are reasonably correct. Although there are exceptions, most studies have provided little evidence

of positive correlations of radiation dose and cancer mortality, thus confirming that extrapolation from high level exposures has not seriously underestimated risks.

Studies of populations exposed at low levels pose special statistical problems. Because risks cannot be estimated precisely, interest centers on whether the data are consistent with risk estimates obtained from the A-bomb survivors and other populations exposed at high doses of radiation, and the extent to which low-level data confirm these estimates. Confidence limits are of greater interest than the estimates themselves, and accurate assessment of uncertainty is especially important. The number of cases from diseases of interest is often much smaller than the sample sizes of 50 and 100 considered by Prentice and Mason, while the dose distributions are usually extremely skewed. Under these conditions, asymptotic procedures, even those based on the likelihood ratio and score statistics, may not perform adequately. Constraint problems are a frequent occurrence. Similar problems can be expected to arise in studies of persons exposed to residential radon.

In this report, the special problems posed by data from populations exposed to low levels of radiation are illustrated using data from three studies of nuclear workers. Several procedures for obtaining confidence limits for the linear relative risk coefficient are illustrated using these data, including a procedure based on computer simulations. This report also presents the results of additional computer simulations designed to provide a general evaluation of various statistical procedures for making inferences regarding the linear relative risk model. Specifically, the performance of the likelihood ratio statistic and the score statistic with expected informatior are evaluated and compared.

2.0 THE STATISTICAL MODEL AND NOTATION

The linear relative risk model as it has been applied to nuclear worker data is one in which the hazard at time t is assumed to be of the form

$$\lambda_s(t)\ [1 + \beta\ z(t)], \qquad (1)$$

where $\lambda_s(t)$ is the baseline risk at time t for stratum s, and the single covariate z(t) is a measure of radiation exposure at time t. The exposure measure has usually been taken to be the dose accumulated by t - x, where x is a specified lag period introduced to account for a minimal latency period. The coefficient β is referred to as the excess relative risk, and is expressed as a proportional increase over the baseline risk per unit of dose.

Estimates of the excess relative risk, $\hat{\beta}$, are obtained by maximizing the partial log-likelihood function, given by

$$L(\beta) = \sum_i \left\{ \sum_\ell \log\ (1 + \beta\ Z_{i\ell}) - m_i\ \log(1 + \beta\ M_i) \right\} \qquad (2)$$

where i indexes strata defined by the indices t and s in equation (1), m_i is the number of failures (or deaths from the disease of interest) in stratum i, $Z_{i\ell}$ is the dose of the ℓ th death in stratum i and M_i is the mean of the doses of all workers in stratum i. It is useful to define

$$R(\beta) = \pm \sqrt{\{\ 2\ L(\hat{\beta}) - 2\ L(\beta)\ \}}$$

where the plus sign applies when $\hat{\beta} \geq \beta$, and the minus sign applies when $\hat{\beta} < \beta$. Approximate 90% confidence limits based on the likelihood ratio statistic may be obtained by determining β such that $R(\beta) = \pm 1.645$.

The score statistic for this model is given by

$$U(\beta) = \sum_i \left\{ \sum_\ell Z_{i\ell}/(1 + \beta\ Z_{i\ell}) - m_i\ M_i/(1 + \beta\ M_i) \right\}. \qquad (3)$$

The variance of U(ß) can be estimated by the expected information

$$V(\beta) = \sum_i \sum_k \frac{1 + \beta Z_{ik}}{N_i (1 + \beta M_i)} \left[\frac{Z_{ik}^2}{(1 + \beta Z_{ik})^2} - \frac{M_i^2}{(1 + \beta M_i)^2} \right]$$

where k indexes observations in stratum i, and N_i is the number of observations in stratum i. Note that V(ß) is constant for all samples. Approximate 90% confidence limits based on the score statistic may be obtained by determining ß such that $W(\beta) = U(\beta)/\sqrt{V(\beta)} = \pm 1.645$.

In general the likelihood ratio based approximate confidence limits are easier to calculate because only observed (Z_{il}) and expected (M_i) doses associated with failures are needed, and because the equations that must be solved are simpler. The use of the score statistic with expected information requires information on individual observations in the risk sets.

In calculating maximum likelihood estimates, it is necessary to constrain the parameter ß to avoid negative relative risks at high doses. This means that ß must be greater than $-1/Z_M$, where Z_M is the maximum dose in the study population. With a small number of failures, it may happen that the likelihood function is a decreasing function of ß for all $\beta > -1/Z_M$, and thus the maximum value is achieved at the lower constraint, $\beta = -1/Z_M$. The statistic R(ß) can still be calculated when this occurs, but the validity of statistical inferences based on the asymptotic approximation for R(ß) is questionable under these conditions. This was one reason for conducting the simulations presented in this report.

As noted above, Prentice and Mason suggest reparametrizing the linear relative risk model by taking $\alpha = \log(\beta + \beta_0)$, where $-\beta_0$ is the lower constraint, and β_0 might be chosen as $1/Z_M$. Alternatively, the parameter might be constrained to be non-negative by choosing $\beta_0 = 0$. In epidemiologic studies of low-dose radiation, this latter choice reflects an <u>a priori</u> view that negative estimates are not correct, but result from random

fluctuation, or from bias due to unidentified confounders. The inverse information for α may also be used to obtain approximate confidence limits. However, if the likelihood is maximized at the lower constraint, the maximum likelihood estimate for α will be undefined, and the method cannot be applied.

It can be shown that the comparable score test statistic to $W(\beta)$, calculated with expected information but based on α, is identical to $W(\beta)$. It can also be shown that if the upper confidence limit obtained from the asymptotic approximation for the score statistic with $\beta_0 = 0$ is positive, then this limit will be the same as the upper confidence limits based on negative constraint choices with $\beta_0 > 0$. By contrast, the upper confidence limit based on the likelihood ratio statistic approximation depends on the constraint chosen. Specifically, if the first derivative of $L(\beta)$ with respect to β is negative at $\beta = 0$, then the upper limit based on a constraint of zero is larger than the upper limit based on a negative constraint.

3.0 COMPUTER SIMULATIONS

If the number of failures is small, it may be desirable to obtain confidence limits using computer simulations (Gilbert 1989). Computer simulations are used in this report to evaluate the general behavior of several statistical procedures. Simulations may be implemented by first choosing a candidate upper limit B, and then randomly selecting S samples such that each sample consists of m_i observations from each risk set, where m_i is the number of failures in stratum i as noted under equation (2). This random selection is carried out so that the probability of selecting observation k is

$$\frac{1 + B\, Z_{ik}}{N_i(1 + B\, M_i)} \qquad (4)$$

where k indexes the N_i observations in the stratum i risk set. This expression gives the probability that observation k is the case, given that the risk set includes one case, and that B is the true value of ß.

For simulations based on the score statistic, U(ß) (see equation (3)) is calculated for each of the S samples, and the one-sided upper (lower) confidence level associated with B is estimated as the proportion of times that U(ß) is less (greater) than the comparable score statistic based on the actual failures in the data set. The process is repeated until a value of B is obtained with the desired confidence level.

Computer simulations based on the likelihood ratio statistic are computationally more difficult because the likelihood must be maximized for each of the S samples. However, once the maximum likelihood estimate is obtained, the procedure is similar to that described above, except that R(ß) is substituted for U(ß). In this report, where maximum likelihood estimates of ß were required, a Newton-Raphson iteration was used, taking the known true value of ß as the starting value. Iterations were based on

$\alpha = \log (\beta + \beta_0)$. Observed information was used unless it was negative or extremely small, in which case expected information was substituted. An estimate was judged to have converged when successive values of β had a relative difference of less than 0.01 or an absolute difference of less than 0.001. The iteration was discontinued after 20 iterations. Special procedures were used to check if the likelihood was maximized at the lower constraint, or at infinity (that is, the likelihood approached an asymptote as β approached infinity.) In most cases evaluated in this report, all samples had converged in 20 iterations, and in no instance did the percentage of samples failing to converge exceed 0.1%. For simulations comparing the performance of different statistics, e.g. $W(\beta)$ and $R(\beta)$, the same samples were used.

4.0 ILLUSTRATION BASED ON NUCLEAR WORKER DATA

Data on workers employed at the Hanford Site, at Oak Ridge National Laboratory (ORNL), or at the Rocky Flats Nuclear Weapons Plant are used to illustrate the application of the procedures described above. Detailed results of analyses of combined data from these studies, including additional description of the study populations and methodological details, have been reported (Gilbert et al. 1989b, 1990). Results of analyses of data from each of the three individual facilities have also been reported, and a description of these study populations and the methods used for determining vital status and cause of death, may be found in publications describing these results (Checkoway et al. 1985, Gilbert et al. 1989a, Wilkinson et al. 1987).[a]

For all three facilities, the effects of exposure to external radiation are of interest, and analyses have focused on such effects. Workers at all facilities were monitored for external radiation exposure through the use of dosimeters worn by the workers, and estimates of the whole body dose were available for each year of monitoring. Dose[b] is expressed in millisieverts (mSv). The study population for analyses in this report consisted of white male workers who were employed at one of the facilities for at least six months and who were monitored for external radiation exposure.

---

[a]Findings from the ORNL study have recently been updated to include seven additional years of follow-up (Wing et al 1991). Unlike earlier analyses (Checkoway 1985, Gilbert et al 1989b, 1990), the updated analyses showed a significant positive correlation with cumulative radiation dose. Analyses in this report are based on earlier data, which are adequate for illustrating methods, the objective of this report.

[b]Because most of the external exposure at the three facilities was to low linear-energy-transfer γ radiation (with a quality factor of one), the absorbed dose (in mGy) and the dose equivalent (in mSv) expressions were similar for most workers. Thus we refer to "dose" even though "dose equivalent" is more accurate terminology.

Illustrative analyses in this report are given for leukemia mortality and for all cancer mortality. For all analyses, age in years served as the time variable t with additional stratification on calendar year, time since initial employment, and number of years monitored (Hanford only). Combined analyses based on all three facilities were also stratified by facility. To facilitate computations based on the score statistic and computer simulations, analyses were based on grouped exposure data using the following categories: 0-, 5-, 10-, 20-, 50-, 100-, 150-, 200-, 300-, 400-, and 500+ mSv. Doses were lagged by two years for analyses of leukemia mortality, and by ten years for analyses of all cancer mortality. Additional detail on these choices and the way they were implemented is provided in Gilbert et al. (1989b, 1990).

Table 1 shows confidence limits for leukemia based on the asymptotic approximation to the score statistic, $W(\beta)$, and to the likelihood ratio statistic, $R(\beta)$; for the latter, results are shown for a lower constraint of zero, and for a lower constraint that is the negative of the inverse of the dose associated with the highest dose category ($-1/Z_M$). Table 2 shows similar limits for all cancers. The results labeled "estimated actual confidence level" were each based on 5000 samples generated under the assumption that the true excess relative risk β was the approximate upper limit. Simulations evaluating the likelihood ratio procedures for all cancer consumed considerable computer time, and thus were limited to a few cases.

Except for leukemia in Rocky Flats, all maximum likelihood estimates were negative, and thus the use of the two constraints yielded different approximate upper limits. In several cases, the likelihood was maximized at the lower constraint. The approximate limits obtained using the three methods did not agree even for analyses of all cancer based on the combined data, where the number of deaths exceeded 1000.

Most of the upper limits based on the score statistic and on the likelihood ratio statistic with a constraint of zero had

TABLE 1. Upper One-Sided 95% Confidence Limits for the Leukemia Excess Relative Risk (ß) Based on Asymptotic Approximations and Based on Simulations.

	Hanford	ORNL	Rocky Flats	Combined
Number of deaths	27	11	4	42
Maximum likelihood estimate	--(a)	--(a)	4.4%	--(a)
A. Based on the Score Statistic.				
Upper limits based on asymptotic approximation(b)	4.8%	14.0%	52.0%	3.4%
Estimated actual confidence level(c)	95.8%	95.5%	93.6%	96.0%
Upper limits based on computer simulations(d)	4.3%	12.0%	60.0%	2.8%
B. Based on the Likelihood Ratio Statistic with the Constraint ß ≥ 0.				
Upper limits based on asymptotic approximation(b)	5.0%	11.0%	72.0%	3.5%
Estimated actual confidence level(c)	96.2%	95.6%	94.2%	96.1%
Upper limits based on computer simulations(d)	4.4%	9.5%	85.0%	2.8%
C. Based on the Likelihood Ratio Statistic with the Constraint ß ≥ $-1/Z_M$.				
Upper limits based on asymptotic approximation(b)	3.3%	8.6%	72.0%	2.3%
Estimated actual confidence level(c)	93.2%	94.2%	93.7%	93.6%
Upper limits based on computer simulations(d)	4.4%	9.5%	85.0%	2.8%

(a) Likelihood maximized at the constraint $-1/Z_M$.
(b) Upper limits (based on two-sided 90% limits) obtained from asymptotic normal approximation.
(c) Percent of samples out of 5000 samples for which the calculated statistic was less than -1.645.
(d) The simulated limits have the property that for 5000 samples at this limit, the calculated statistic exceeded the comparable statistic based on the actual cases in 95% of the samples.

actual confidence levels that exceeded 0.95, the only exception being leukemia in Rocky Flats. However, the actual confidence levels for the likelihood ratio statistic with a negative constraint were less than 0.95. Overall, the approximations were reasonably good considering that some of the sample sizes were very small, and that the dose distributions were very skewed.

Tables 1 and 2 also show upper limits based on computer simulations. These were calculated as described above, and in each case the confidence level was consistent with a one-sided

TABLE 2. Upper One-Sided 95% Confidence Limits for the All Cancer Excess Relative Risk (ß) Based on Asymptotic Approximations and Based on Simulations.

	Hanford	ORNL	Rocky Flats	Combined
Number of deaths	833	140	63	1036
Maximum likelihood estimate	-0.80	-0.96%	--(a)	-0.99%
A. Based on the Score Statistic.				
Upper limits based on asymptotic approximation(b)	0.85%	3.2%	2.8%	0.38%
Estimated actual confidence level(c)	95.4%	96.1%	96.5%	95.5%
Upper limits based on computer simulations(d)	0.80%	2.9%	2.2%	0.30%
B. Based on the Likelihood Ratio Statistic with the Constraint ß ≥ 0.				
Upper limits based on asymptotic approximation(b)	0.93%	3.0%	3.2%	0.65%
Estimated actual confidence level(c)	96.0%	95.2%	97.5%	--
C. Based on the Likelihood Ratio Statistic with the Constraint ß ≥ $-1/Z_M$.				
Upper limits based on asymptotic approximation(b)	0.66%	2.8%	1.7%	0.22%
Estimated actual confidence level(c)	--	92.9%	94.6%	--

(a) Likelihood maximized at the constraint $-1/Z_M$.
(b) Upper limits (based on two-sided 90% limits) obtained from asymptotic normal approximation
(c) Percent of samples out of 5000 samples for which the calculated statistic was less than -1.645.
(d) The simulated limits have the property that for 5000 samples at this limit, the calculated statistic exceeded the comparable statistic based on the actual cases in 95% of the samples.

95% level (that is, based on 5000 samples, the estimated confidence level was not significantly different from 0.95 at the 0.10 rejection level). The likelihood ratio based simulated intervals, which were calculated only for leukemia, did not depend on whether the constraint was negative or zero. Apparently, the difference in the approximate intervals was due to inadequacies in the asymptotic approximation. In view of the fact that the actual confidence levels associated with the approximate limits were close to 95%, the differences in the simulated and approximate limits indicate a very flat likelihood function in the region of the upper limits. As would be

expected, agreement of simulated and approximate limits was generally better with the larger sample sizes.

5.0 RESULTS OF SIMULATIONS

To investigate the effect of sample size and of the size of the excess relative risk on the behavior of various procedures, simulations were conducted in which all risk sets had the exposure distribution obtained by averaging the exposure distributions from the 42 risk sets used in the combined nuclear worker leukemia results.[a] This distribution is referred to as the "worker-based distribution", and is shown in Table 3. These simulations were based on sample sizes (number of failures) of 10, 20, 50, 100 and 500, and with ß chosen to be 0.0, 0.025, 0.05, 0.10, and 0.20.

TABLE 3. Worker-based dose distribution.(a)

Dose category	Dose assigned to category (mSv)	Proportion of doses in category (mSv)
0-	1.4	0.448
5-	7.3	0.140
10-	14.3	0.160
20-	31.0	0.127
50-	70.5	0.057
100-	122.2	0.019
150-	173.0	0.014
200-	243.7	0.016
300-	343.9	0.011
400-	441.7	0.005
500+	590.7	0.003

(a)This distribution was obtained by averaging the exposure distributions from the 42 risk sets used in the combined nuclear worker leukemia results.

To evaluate the behavior of various procedures for different exposure distributions, simulations were conducted in

[a]With larger sample sizes (m), simulations with the same risk set for each failure used far less computer time than those based on the actual risk sets. This efficiency gain occurred because it was no longer necessary to include the m separate M_i (see equation 2) in the iteration process.

which all risk sets had one of six trinomial distributions where Z could take on one of three values 0.0, S2, or S3 with respective probabilities 1-P2-P3, P2, P3. The values of S2, S3, P2, and P3 are shown in Table 4. For odd numbered populations, S3 was chosen to be two times S2, whereas for even numbered populations, S3 was chosen to be ten times S3. In all cases, S2 and S3 were scaled so that the mean dose was equal to that for the worker-based distribution shown in Table 3. These latter simulations were conducted with sample sizes of 50, and with ß equal to 0.0 and 0.05. To facilitate comparisons, tables presenting results for the trinomial distributions also repeat the worker-based results for this sample size and these values of ß.

TABLE 4. Trinomial distributions

Dose distribution	S2[a]	S3[a]	P2[b]	P3[b]
1	28.3	56.6	0.333	0.333
2	7.7	77.2	0.333	0.333
3	43.5	87.1	0.350	0.150
4	15.3	153.0	0.350	0.150
5	28.6	257.3	0.180	0.020
6	74.5	744.7	0.180	0.020

(a) In these trinomial distributions, the dose Z could take on one of three values: S1 = 0, S2, and S3. These values were scaled so that the mean dose would be 28.3, the mean of the worker-based distribution shown in Table 3.

(b) P2 is the probability that Z = S2; P3 is the probability that Z = S3; P1 = 1 - P2 - P3 is the probability that Z = 0.

For each case evaluated, the likelihood ratio statistics were calculated with constraints set to 0.0 and to $-1/Z_M$, where Z_M was the maximum exposure value. Tables 5 and 6 show the percentage of samples for which the likelihood was maximized at the lower constraint. For the worker-based distribution (Table 5), this occurred frequently with smaller sample sizes and smaller ß. As would be expected, with ß = 0 and with ß

TABLE 5. Percentage of times (out of 5000) that likelihood was maximized at lower constraint. Based on the worker-based dose distribution.

Number of failures	Constaint of 0	Constaint of $-1/Z_M$
ß=0.0		
10	62	57
20	57	45
50	54	31
100	54	20
500	51	0.6
ß=0.025		
10	44	39
20	34	25
50	20	8.8
100	11	2.2
ß=0.05		
10	32	29
20	19	14
50	7.1	2.7
100	1.7	0.1
ß=0.10		
10	19	16
20	7.2	4.6
50	0.9	0.3
100	0	0
500	0	0
ß=0.20		
10	7.1	5.8
20	1.5	1.0
50	0.02	0.02
100	0	0

constrained to be greater than or equal to zero, this occurred about 50% of the time even with large sample sizes. In the remaining cases (ß > 0 or with the negative constraint), the percentage decreased with increasing sample size.

For the trinomial distributions (Table 6) and ß = 0, only for the highly skewed distribution #6 was the likelihood frequently maximized at the negative lower constraint; these simulations were all based on a sample size of 50. With

TABLE 6. Percentage of times (out of 5000) that likelihood was maximized at lower constraint. Based on trinomial dose distributions and a sample size of 50 failures.

Dose Distribution	Constaint of 0	$-1/Z_M$
ß=0.0		
#1	53	0
#2	50	0
#3	51	0
#4	51	0
#5	58	6.1
#6	58	35
Worker-based	54	31
ß=0.05		
#1	26	0
#2		
#3	17	0
#4	7.8	0
#5	7.2	0.1
#6	2.4	1.0
Worker-based	7.1	2.7

ß = 0.05, the likelihood was maximized at the zero constraint more frequently for the less skewed distributions (#1, #2, #3) than for the remaining distributions.

Tables 7 and 8 show the percentage of times the statistics W(ß) and R(ß) were greater than the 5% and 1% critical values from the standard normal distribution. These upper tails are relevant for determining lower confidence limits for ß. The results for ß = 0 are relevant for a one-sided test of the null hypothesis. The upper tails are unaffected by the choice of constraint for ß.

For the worker-based distribution, and for the highly skewed trinomial distribution #6, the score statistic W(ß) was skewed to the right resulting in too many samples with results in the upper tail. This would lead to lower confidence limits that are larger than they should be, and to rejecting the null hypothesis too often. As expected, the normal approximation improved as the

TABLE 7. Percentage of times (out of 5000) that simulated test statistics exceeded upper 5% (1%) nominal level obtained from the asymptotic approximation. Based on worker-based dose distribution.

Number of failures	Score statistic		Likelihood ratio statistic	
	5%	1%	5%	1%
ß=0.0				
10	8.7	3.8	5.0	1.2
20	7.4	2.9	4.4	0.8
50	6.7	2.2	4.5	0.8
100	6.2	2.0	4.7	1.0
500	5.7	1.7	4.9	1.1
ß=0.025				
10	7.2	2.6	4.7	0.9
20	6.2	1.8	4.3	0.7
50	6.7	1.9	5.1	1.1
100	5.3	1.4	4.5	1.0
ß=0.05				
10	6.9	2.1	5.1	1.2
20	6.2	1.6	4.8	0.7
50	5.9	1.5	4.8	0.8
100	5.1	1.2	4.5	0.9
ß=0.10				
10	6.4	1.9	5.4	1.4
20	5.2	1.4	4.4	1.0
50	5.8	1.3	5.1	0.8
100	5.9	1.5	5.4	1.2
500	5.0	1.2	4.9	1.1
ß=0.20				
10	5.4	1.0	5.3	1.2
20	5.5	1.2	5.2	1.1
50	5.4	1.1	5.0	1.0
100	4.6	1.1	4.5	0.9

sample size increased. For most cases examined, the approximation also improved as ß increased.

For the likelihood ratio statistic, the approximation provided by the normal distribution to the upper tails was quite good, even with sample sizes as small as 10. Tests of the null hypothesis and, more generally, determination of lower confidence limits for ß would probably be reasonably accurate if based on the likelihood ratio statistic.

TABLE 8. Percentage of times (out of 5000) that simulated test statistics exceeded upper 5% (1%) nominal level obtained from the asymptotic approximation. Based on trinomial dose distributions and a sample size of 50 failures.

Dose distribution	Score statistic		Likelihood ratio statistic	
	5%	1%	5%	1%
ß=0.0				
#1	5.2	1.0	5.2	0.9
#2	5.0	1.1	4.5	1.0
#3	5.9	1.2	4.9	1.1
#4	5.9	1.2	4.5	0.8
#5	4.6	1.2	3.9	0.8
#6	7.3	2.6	4.2	0.8
Worker-based	6.7	2.2	4.5	0.8
ß=0.05				
#1	5.1	0.9	5.5	1.3
#2				
#3	5.2	1.0	5.1	0.9
#4	5.3	1.0	4.5	0.7
#5	5.3	1.0	4.7	0.9
#6	5.5	1.5	4.7	0.8
Worker-based	5.9	1.5	4.8	0.8

Tables 9 and 10 show the lower tails of the distributions, or the percentage of times the statistics W(ß) and R(ß) were less than the 1% and 5% critical values from the standard normal distribution. These lower tails are relevant for determining upper confidence limits for ß. With ß = 0, and a constraint of zero, R(ß) can never be less than zero, and thus can never be less than the negative 1% and 5% critical values; however, one would usually not be interested in evaluating ß = 0 as an upper confidence limit. In other situations, with small sample sizes, the minimum possible value for R(ß) may also be larger than the lower-tail critical values.

For the worker-based distributions (Table 9) and for the more highly skewed trinomial distributions #4, #5, and #6, the score statistic yielded too few samples in the lower tails, and this would lead to upper confidence limits that were too large.

TABLE 9. Percentage of times (out of 5000) that simulated test statistics were less than lower 5% (1%) nominal level obtained from the asymptotic approximation. Based on worker-based dose distribution.

Number of failures	Score statistic		Likelihood ratio statistic			
			Constraint of 0		Constraint of $-1/z_M$	
	1%	5%	1%	5%	1%	5%
ß=0.0						
10	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0
50	0.0	2.2	0.0	0.0	0.0	0.5
100	0.1	3.4	0.0	0.0	0.0	5.0
500	0.4	3.8	0.0	0.0	0.9	5.1
ß=0.025						
10	0.0	0.1	0.0	0.0	0.0	0.0
20	0.0	2.5	0.0	0.0	0.0	3.9
50	0.3	4.0	0.0	3.1	0.8	7.1
100	0.5	4.4	0.3	5.3	1.6	6.0
ß=0.05						
10	0.0	1.9	0.0	0.0	0.0	1.9
20	0.1	3.8	0.0	3.9	0.1	7.4
50	0.5	4.2	0.7	6.1	1.7	6.2
100	0.7	4.5	1.3	5.8	1.3	5.8
ß=0.10						
10	0.1	3.9	0.0	5.6	0.0	8.0
20	0.5	4.1	0.8	6.4	1.5	6.5
50	0.8	4.8	1.4	5.7	1.4	5.7
100	0.9	4.2	1.2	5.2	1.2	5.2
500	0.9	5.2	1.0	5.5	1.0	5.5
ß=0.20						
10	0.6	4.1	0.9	6.5	1.3	6.6
20	0.7	4.4	1.3	5.9	1.4	5.9
50	0.8	4.7	1.0	5.1	1.0	5.1
100	0.9	4.9	1.2	5.2	1.2	5.2

As shown in Tables 5 and 6, the likelihood was frequently maximized at constraint values. For those instances in which the likelihood was maximized at a value between the negative constraint and zero, the two likelihood ratio statistics took on different values, which led to different behavior in the lower tail of the distribution. As sample size and ß increased, the percentage of samples for which the likelihood was maximized at a negative value approached zero.

TABLE 10. Percentage of times (out of 5000) that simulated test statistics were less than lower 5% (1%) nominal level obtained from the asymptotic approximation. Based on trinomial dose distributions and a sample size of 50 failures.

Dose distribution	Score statistic		Likelihood ratio statistic			
			Constraint of 0		Constraint of $-1/z_M$	
	1%	5%	1%	5%	1%	5%
ß=0.0						
#1	0.8	4.9	0.0	0.0	0.9	4.9
#2	0.5	4.6	0.0	0.0	1.0	5.6
#3	1.0	6.0	0.0	0.0	1.4	5.6
#4	0.3	4.1	0.0	0.0	1.4	5.0
#5	0.3	3.6	0.0	0.0	1.3	6.7
#6	0.0	0.1	0.0	0.0	0.0	2.7
Worker-based	0.0	2.2	0.0	0.0	0.0	0.5
ß=0.05						
#1	1.0	4.6	0.0	1.2	0.9	4.3
#2						
#3	1.1	4.6	0.06	2.8	1.2	4.6
#4	0.6	4.5	0.5	5.1	1.0	5.1
#5	0.6	4.8	0.6	4.8	1.1	4.8
#6	0.4	4.7	1.1	5.5	1.2	5.6
Worker-based	0.5	4.2	0.7	6.1	1.7	6.2

There was some tendency for the use of the negative constraint to lead to too many samples yielding results below the normal critical values, while the use of the zero constraint led to too few samples yielding results in this region. However, for very small samples and small ß, both procedures led to too few samples below the lower critical values. At certain combinations of sample size and ß, upper confidence limits would be too large when based on a zero constraint, but too small when based on a negative constraint; this situation was observed in several of the examples based on the nuclear worker data (Tables 1 and 2). For those instances for which the lower tail did not depend on the constraint, the normal distribution provided a good approximation.

The average value of the maximum likelihood estimate of ß across samples was also calculated, and the estimates were always found to be biased upward, a result that agrees with findings of

Prentice and Mason. The behavior of the score statistic based on $\alpha = \log(\beta + \beta_0)$ with observed information was investigated. This statistic behaved well only in situations where no constraint problems were encountered as in trinomial populations #1, #2, and #3. Even in those populations for which constraint problems were relatively infrequent, the statistic did not perform as well as $R(\beta)$ and $W(\beta)$, especially in the lower tail. Often there were no samples with observations below the 1% and 5% critical values. The asymptotic approximation for the test statistic using the maximum likelihood estimate of α with inverse expected information was also found to be inadequate and was usually worse than that for the score statistic based on α with observed information.

A comparison of power, or estimation precision, based on the statistics $W(\beta)$ and $R(\beta)$ is also of interest. For testing $\beta = 0$ when β is actually positive, the score test based on $W(\beta)$ usually had slightly greater power, but this is probably partly due to the fact that the Type I error based on the asymptotic approximation to $W(\beta)$ was often larger than 5%. Similarly, $R(\beta)$ usually had slightly greater power for rejecting positive β when β was actually zero, but again the Type I error for this test was sometimes too large.

A "fairer" comparison of power is to compare the probabilities of rejecting various hypothesized β using cutpoints that yielded Type I errors of 5%, rather than the cutpoint provided by the normal approximation. Such a comparison would be the appropriate one for tests and confidence limits based on simulations. To evaluate power for testing the null hypothesis $\beta = 0$, the 95 th percentile of $W(0)$ and $R(0)$ was estimated based on 5000 samples generated with $\beta = 0$. The number of times that $W(0)$ and $R(0)$ exceeded this percentile was then obtained based on 5000 samples generated with $\beta > 0$. Power for rejecting positive β when β was actually zero was estimated in an analogous manner. Results of these calculations are shown in Tables 11-14. The power of the different tests is very similar for all sample sizes and values

of ß evaluated. With respect to Tables 12 and 14, it is noted that for larger sample sizes and ß, all samples for which the likelihood is maximized at ß ≤ 0 fall in the rejection region, and thus identical results are obtained for the zero and negative constraints.

TABLE 11. Percentage of times (out of 5000) that simulated test statistics would reject the null hypothesis (ß = 0) based on the estimated 5% critical values of the distributions W(ß) and R(ß) when ß is actually equal to the values indicated. Based on worker-based dose distribution.

Number of failures	Score statistic	Likelihood ratio statistic
ß=0.01		
10	0.082	0.077
20	0.087	0.088
50	0.132	0.133
100	0.171	0.174
ß=0.025		
10	0.126	0.125
20	0.167	0.171
50	0.281	0.285
100	0.414	0.419
ß=0.05		
10	0.213	0.209
20	0.307	0.320
50	0.545	0.554
100	0.776	0.789
ß=0.10		
10	0.361	0.369
20	0.569	0.597
50	0.862	0.881
100	0.984	0.986
500	1.000	1.000
ß=0.20		
10	0.578	0.620
20	0.820	0.856
50	0.994	0.996
100	1.000	1.000

TABLE 12. Percentage of times (out of 5000) that simulated test statistics would reject the null hypothesis (ß = 0) based on the estimated 5% critical values of the distributions W(ß) and R(ß) when ß is actually equal to the values indicated. Based on trinomial dose distributions and a sample size of 50 failures.

Dose distribution	Score statistic	Likelihood ratio statistic
True ß=0.05		
#1	0.141	0.171
#2	0.289	0.276
#3	0.284	0.250
#4	0.472	0.478
#5	0.539	0.582
#6	0.775	0.747
Worker-based	0.545	0.554

TABLE 13. Percentage of times (out of 5000) that simulated test statistics would reject the hypothesis that ß = 0.05 or ß = 0.10 when ß is actually equal to 0.0. Based on the estimated 5% critical values of the distributions of W(0) and R(0) when ß = 0.05 or ß = 0.10. Based on worker-based dose distribution.

Number of failures	Score statistic	Likelihood ratio statistic	
		Constraint of 0	Constraint of $-1/Z_M$
	ß = 0.05		
10	0.126	0.129	0.128
20	0.217	0.225	0.228
50	0.474	0.478	0.467
100	0.757	0.758	0.758
	ß = 0.10		
10	0.260	0.261	0.261
20	0.470	0.485	0.487
50	0.839	0.849	0.849
100	0.988	0.991	0.991
500	1.000	1.000	1.000

TABLE 14. Percentage of times (out of 5000) that simulated test statistics would reject the hypothesis that $\beta = 0.05$ or $\beta = 0.10$ when β is actually equal to 0.0. Based on the estimated 5% critical values of the distributions of W(0) and R(0) when $\beta = 0.05$ or $\beta = 0.10$, on trinomial dose distributions, and a sample size of 50 failures.

	$\beta = 0.05$		
Dose distribution	Score statistic	Likelihood ratio statistic	
		Constraint of 0	Constraint of $-1/Z_M$
#1	0.188	0.188	0.194
#2	0.267	0.260	0.264
#3	0.261	0.259	0.259
#4	0.427	0.427	0.429
#5	0.494	0.539	0.539
#6	0.675	0.699	0.699
Worker-based	0.474	0.478	0.467

6.0 CONCLUSIONS AND RECOMMENDATIONS

The computer simulations described in this report indicate that with small sample sizes and highly skewed dose distributions, asymptotic approximations to the score statistic or to the likelihood ratio statistic may not be entirely adequate. For testing the null hypothesis that the excess relative risk is equal to zero, the likelihood ratio statistic performed better than the score statistic; the latter tended to reject too often, and thus to underestimate the significance level.

Frequently the likelihood was maximized at the lower constraint, and when this occurred, the asymptotic approximation to the likelihood ratio statistic did not perform well in obtaining upper confidence limits. The asymptotic approximation to the score test statistic was also inadequate in these situations leading to upper confidence limits that were too large.

In view of these problems, it is recommended that with modest sample sizes, confidence limits be obtained using computer simulations. If these simulations are based on the score statistic, the necessary computations can be performed fairly easily. The score statistic and likelihood ratio statistics were found to perform comparably in terms of power and width of the confidence limits.

7.0 REFERENCES

Beral, V., H. Inskip, P. Fraser, M. Booth, D. Coleman, and G. Rose. 1985. "Mortality of Employees of the United Kingdom Atomic Energy Authority, 1946-1979." Br. Med. J. 291:440-447.

Beral, V., P. Fraser, L. Carpenter, M. Booth, D. Coleman, and G. Rose. 1988. "Mortality of Employees of the Atomic Weapons Establishment, 1951-1982." Br. Med. J. 297:757-770.

Barlow, W. E. 1985. "General Relative Risk Models in Stratified Epidemiologic Studies." J. of the Royal Statistical Society Series C(Applied Statistics) 34:246-257.

Breslow, N. E., and B. E. Storer. 1985. "General Relative Risk Functions For Case-Control Studies." Am. J. Epidemiol. 122:149-162.

Checkoway, H., R. M. Mathew, C. M. Shy, J. E. Watson, W. G. Tankersley, S. H. Wolf, J. C. Smith, and S. A. Fry. 1985. "Radiation, Work Experience, and Cause Specific Mortality Among Workers at an Energy Research Laboratory." Br. J. Ind. Med. 42:525-533.

Cox, D. R. 1972. "Regression Models and Life Tables (with discussion)." J.R. Stat. Soc. B 34:187-220.

Gilbert, E. S. 1989. "Issues in Analysing the Effects of Occupational Exposure to Low Levels of Radiation." Statistics in Medicine 8:173-187.

Gilbert, E. S., G. R. Petersen, and J. A. Buchanan. 1989a. "Mortality of Workers at the Hanford Site: 1945-1981." Health Phys. 56:11-25.

Gilbert, E. S., S. A. Fry, L. D. Wiggs, G. L. Voelz, D. L. Cragle, and G. R. Petersen. 1989b. "Analyses of combined mortality data on workers at the Hanford Site, Oak Ridge National Laboratory, and Rocky Flats Nuclear Weapons Plant." Radiation Research 120:19-35.

Gilbert, E. S., S. A. Fry, L. A. Wiggs, G. L. Voelz, D. L. Cragle, and G. R. Petersen. 1990. "Methods for analyzing combined data from studies of workers exposed to low doses of radiation." Am. J. Epidemiol. 131:917-927.

Howe, G. R., J. L. Weeks, A. B. Miller, A. M Chiarelli, and J. Etezadi-Amoli. 1987. A Study of the Health of the Employees of Atomic Energy of Canada Limited. IV. Analysis of Mortality During the Period 1950-1981. Atomic Energy of Canada Limited AECL-9442, Pinawa, Manitoba ROE 1L0.

Moolgavkar, S. H., and D. J. Venzon. 1987. "General Relative Risk Regression Models for Epidemiologic Studies." Am. J. Epidemiol. 126:949-961.

National Academy of Sciences. 1988. Health Risks of Radon and Other Internally Deposited Alpha-Emitters. BEIR IV. Report of the Advisory Committee on the Biological Effects of Ionizing Radiation, Division of Medical Sciences, National Academy of Sciences, National Research Council. Washington, DC: National Academy of Sciences.

National Academy of Sciences. 1990. Health Effects of Exposure to Low Levels of Ionizing Radiation. BEIR V. Report of the Advisory Committee on the Biological Effects of Ionizing Radiation, Division of Medical Sciences, National Academy of Sciences, National Research Council. Washington, DC: National Academy of Sciences, 1980.

Prentice, R. L., and M. W. Mason. 1986. "On the Application of Linear Relative Risk Regression Models." Biometrics 42:109-120.

Preston, D. L., and D. A. Pierce. 1987. The Effect of Changes in Dosimetry on Cancer Mortality Risk Estimates in the Atomic Bomb Survivors. RERF Technical Report 9-87 (Hiroshima City 732, Japan: Radiation Effects Research Foundation).

Shimizu, Y., H. Kato, and W. J. Schull. 1990. "Studies of Mortality of A-Bomb Survivors. 9. Mortality, 1950-1985: Part 2. Cancer Mortality Based on the Recently Revised Doses (DS86)." Radiation Research 121:120-141.

Thomas, D. C. 1981. "General Relative-Risk Models for Survival Time and Matched Case-Control Analysis." Biometrics 37:673-686.

Smith, P. G., and A. J. Douglas. 1986. "Mortality of workers at the Sellafield plant of British Nuclear Fuels." Br. Med. J. 293:845-852.

Whittemore, A. S., and A. McMillan. 1983. "Lung cancer Mortality Among U.S. Uranium Miners: A Reappraisal." J. National Cancer Inst. 71:489-499.

Wilkinson, G. S., G. L. Tietjen, L. D. Wiggs, W. A. Galke, J. F. Acquavella, M. Reyes, G. L. Voelz, and R. J. Waxweiler. 1987. "Mortality Among Plutonium and Other Radiation Workers at a Plutonium Weapons Factory." Am. J. Epidemiol. 125:231-250.

Wing, S., C. M. Shy, J. L. Wood, S. Wolf, D. L. Cragle, and E. L. Frome. 1991. "Mortality Among Workers At Oak Ridge National Laboratory: Evidence of Radiation Effects in Follow-up Through 1984." JAMA 265:1397-1402.

PNL-7867
UC-605

DISTRIBUTION

No. of Copies

12 DOE/Office of Scientific and Technical Information

R. Hornung
DSHEFS, NIOSH
4676 Columbia Parkway
Cincinnati, OH 45226

10 Terry L. Thomas, Ph.D.
Office of Epidemiology and Health Surveillance
EH-42, GTN
U.S. Department of Energy
Washington, DC 20545

Robert Goldsmith, Ph.D.
Director
Office of Epidemiology and Health Surveillance
EH-42, GTN
U.S. Department of Energy
Washington, DC 20545

Donna Cragle, Ph.D.
Center for Epidemiologic Research
Oak Ridge Associated Universities
P.O. Box 117
Oak Ridge, TN 34831-0117

Janice Watkins
Center for Epidemiologic Research
Oak Ridge Associated Universities
P.O. Box 117
Oak Ridge, TN 34831-0117

No. of Copies

Laurie Wiggs, Ph.D.
Epidemiology Group
HSE-14 MS K404
P.O. Box 1663
Los Alamos, NM 87545

Edward Frome, Ph.D.
Center for Epidemiologic Research
Oak Ridge Associated Universities
P.O. Box 117
Oak Ridge, TN 34831-0117

Arthur G. Johnson, Ph.D.
Oregon State University
Director, Radiation Center
35th and Jefferson A-100
Corvallis, OR 97331-5903

Roy E. Shore, Ph.D.
Laboratory of Biostatistics and Epidemiology
Institute of Environmental Medicine
New York University Medical Center
341 East 25th Street
New York, NY 10010-2598

Norman E. Breslow, Ph.D.
Chairman, Department of Biostatistics, SC-32
University of Washington
Seattle, WA 98195

Robert W. Miller, M.D., Ph.D.
Clinical Epidemiology Branch
National Cancer Institute
EPN-400
Executive Plaza North #400
Bethesda, MD 20892

No. of Copies

Gary H. Spivey, M.D., MPH
Epidemiology & Environmental Medicine
Unocal Corporation
1201 W 5th
P.O. Box 7600
Los Angeles, CA 90051

Gilbert S. Omenn, M.D.
Dean, School of Public Health and Community Medicine
University of Washington
Seattle, WA 98195

FOREIGN

Elisabeth Cardis, Ph.D.
Unit of Biostatistics Research and Informatics
International Agency for Research on Cancer
150, Cour Albert-Thomas
69372 Lyon Cedex 08
France

Colin R. Muirhead, Ph.D.
National Radiological Protection Board
Chilton, Didcot, Oxon OX11 0RQ
Great Britain

Lucy Carpenter, Ph.D.
London School of Hygiene and Tropical Medicine
Keppel Street
London WC1E 7HT
United Kingdom

No. of Copies

ONSITE

U.S. Department of Energy Field Office, Richland

PW Kruger

Hanford Environmental Health Foundation

EL Omohundro

29 Pacific Northwest Laboratory

WJ Bair
ES Gilbert (20)
JA Mahaffey
RG Stevens
Publishing Coordination
Technical Report Files (5)

END

DATE
FILMED
12/24/91