ORNL-6785R

Computer Science and Mathematics Division
Mathematical Sciences Section

A MORTALITY STUDY OF EMPLOYEES OF THE NUCLEAR INDUSTRY IN OAK RIDGE, TENNESSEE

E. L. Frome [*]
D. L. Cragle [†]
J. P. Watkins [†]
S. B. Wing [‡]
C.M. Shy [‡]
W. G. Tankersley [†]
C. M. West [†]

[*] Mathematical Sciences Section
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831-0117
[†] Center for Epidemiologic Research
Oak Ridge Institute for Science and Education
Oak Ridge, Tennessee 37831-0117
[‡] Department of Epidemiology
University of North Carolina
Chapel Hill, North Carolina 27599-7400

Date Published: July 1997

The research was conducted by Oak Ridge Institute for Science and Education (ORISE) under Contract DE-AC05-76OR00033 with the U. S. Department of Energy, Office of Energy Research and Office of Environment, Safety and Health. Additional funding and oversight have been provided by the National Institute for Occupational Safety and Health. An abridged version of this report is published in *RADIATION RESEARCH*,148, July 1997.

Prepared by the
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
managed by
Lockheed Martin Energy Research Corporation
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-96OR22464

Contents

List of Tables

List of Figures

A MORTALITY STUDY OF EMPLOYEES OF THE NUCLEAR INDUSTRY IN OAK RIDGE, TENNESSEE

E. L. Frome
D. L. Cragle
J. P. Watkins
S. B. Wing
C.M. Shy
W. G. Tankersley
C. M. West

Abstract

An analysis was conducted of 27,982 deaths among 106,020 persons employed at four Federal nuclear plants in Oak Ridge, Tennessee between 1943 and 1984. The main objectives were to extend the evaluation of the health effects of employment in the nuclear industry in Oak Ridge to include most workers who were omitted from earlier studies; to compare the mortality experience among the facilities; to address methodological problems that occur when individuals employed at more than one facility are included in the analysis; and to conduct dose-response analyses for those individuals with potential exposure to external radiation.

All cause mortality and all cancer mortality were in close agreement with national rates. The only notable excesses occurred for white males for lung cancer—standardized mortality ratio (SMR) = 1.18, 1849 deaths— and nonmalignant respiratory disease (SMR = 1.12, 1568 deaths). A more detailed analysis revealed substantial differences in death rates among workers at the Oak Ridge plants. Evaluation of internally adjusted log SMRs using Poisson regression showed that workers employed only at TEC or K-25, and at multiple facilities had higher death rates than similar workers employed only at X-10 or Y-12, and that the differences were primarily due to noncancer causes. Analysis of selected cancer causes for white males indicated large differences among the facilities for lung cancer, leukemia, and other lymphatic cancer.

Dose-response analyses for external penetrating radiation were limited to a subcohort of 28,347 white males employed at X-10 or Y-12. Their collective recorded dose equivalent was 376 Sv. There was a strong "healthy worker effect" in this subcohort—all cause SMR = 0.80 (4786 deaths) and all cancer SMR = 0.87 (1134 deaths). Variables included in the analyses were age, birth cohort, a measure of socioeconomic status, length of employment, internal radiation exposure potential, and facility. For external radiation dose with a ten year lag the excess relative risk was 0.31 per Sv (95% CI = -0.16,1.01) for all causes and 1.45 per Sv (95% CI = .15, 3.48) for all cancer. The estimated excess relative risk for

leukemia was negative but imprecisely determined.

A preliminary dose adjustment procedure was developed to compensate for missing dose but not other dosimetry errors. Results of the analyses using the adjusted doses suggest that the effect of missing dose is an upward bias in dose-response coefficients and likelihood ratio test statistics.

This report, related data sets, and computer programs are available via the Internet at `http://www.epm.ornl.gov/~frome/`—see **Oak Ridge Mortality Study**.

1. Introduction.

This report is the second phase of a study of the mortality of most workers employed at Federal nuclear plants in Oak Ridge, Tennessee between 1943 and 1984. In the first phase the analysis was limited to white males who were employed only during the World War II era when radiation monitoring programs were in the developmental stages [16]. Workers were assigned a relative radiation exposure index based on job and department codes and were designated as "probably exposed" or "probably non-exposed". Additional factors considered in the analyses were socioeconomic status, place of employment, length of employment, birth year, and period of follow-up. The place of employment is an important exposure-related variable since the type of radiation exposure was different at each facility. Previous dose-response studies have been limited to subgroups of white males and have excluded individuals that were employed at more than one facility after 1947 [38, 39]. Mortality study results for white males employed only at X-10 with follow-up through 1984 have also been reported in combination with populations of workers employed at other Department of Energy (DOE)[1] facilities [20].

The main objectives of this second phase are to extend the evaluation of the health effects of employment in the nuclear industry in Oak Ridge to include Oak Ridge workers who were omitted from earlier studies; to compare the mortality experience among the facilities; to carry out dose-response analyses for those individuals who were potentially exposed to external radiation; and to address methodological problems that occur when individuals employed at more than one facility are included in the analyses. These difficulties are due to the fact that the nature and extent of the radiation hazards and monitoring programs were different at each facility and varied over time.

Three types of analyses were done for this report. **First**, cause-specific mortality for the Oak Ridge cohort was compared with the U. S. population by race and gender groups. In the **second** type of analyses, Poisson regression methods were used to evaluate the joint influence of several risk factors on cause-specific mortality. Observed and expected deaths were cross-classified by levels of facility of employment, duration of employment, socioeconomic status, and period of follow-up and a main- effects model was used to describe the joint effects of the

[1]Abbreviations used: ADS,Analytic data structure;AIC,Akaike information criteria; AG, age group factor in main effects model; B,birth cohort factor in main effects model;CI,confidence interval; CEDR,Comprehensive Epidemiologic Data Resource; df,degrees of freedom; D, external radiation dose; DOE, Department of Energy;ERR,excess relative risk; F,facility factor in main effects model; FTR, Freeman-Tukey residual; GLIM, generalized linear interactive modeling; IG, internal radiation exposure group factor in main effects model; K-25, Oak Ridge facility (also known as Oak Ridge Gaseous Plant); L, length of employment factor in main effects model; L%, logarithmic percent; LRT, likelihood ratio test; MULT, Oak Ridge workers employed at more than one facility; S, factor that is surrogate for SES in main effects model; SAS, Statistical Analysis System; SE, standard error; SES, socioeconomic status;SMR, standardized mortality ratio; TEC, Oak Ridge facility consisting of workers employed at the Y-12 site when the plant was operated by Tennessee Eastman Corporation; X-10, Oak Ridge facility (also known as Oak Ridge National Laboratory); X-10/Y-12, subcohort of white males employed at X-10 or Y-12; Y-12, Oak Ridge facility at Y-12 site after May, 1947

four risk factors. In the **third** type of analysis the goal was to evaluate the potential adverse health effects of low level occupational exposure to ionizing radiation. These analyses were limited to white males who were ever employed at X-10 or Y-12. A uniform approach to radiation exposure assessment (based on ordered exposure categories for internal and external radiation for each facility) was used, and cause-specific mortality patterns were evaluated for potential association with radiation exposure. Length of employment, socioeconomic status, birth year, and age are considered as covariates.

During the course of this study it was determined that external radiation doses at X-10 and Y-12 were underestimated for some individuals—before 1957 at X-10 and before 1961 at Y-12 when not all workers were monitored. A preliminary procedure for computing adjusted doses was developed by Watkins *et al* [38, 37, 39] and used to evaluate the bias that may result in dose-response coefficients as the result of this dose underestimation.

2. Materials and Methods.

2.1. Description of Cohort.

On December 18, 1941, 11 days after the Japanese attack on Pearl Harbor, the decision was made to focus atomic energy research on the development of a weapon for the war. The Army was assigned the responsibility for this design and construction project of unprecedented size. The headquarters for the project was established in New York and named the Manhattan Engineer District. In September 1942 an isolated area in east Tennessee was selected as the site for the development of the full scale production facilities for uranium separation and for the construction of an experimental nuclear pile that would be used to produce plutonium for research in the war effort. An air-cooled experimental pile, a chemical separation plant, and supporting laboratories were constructed by the Du Pont Co. at the X-10 site. This facility was officially named the Clinton Laboratories—later renamed Oak Ridge National Laboratory (ORNL)—and will be referred to in this report as the **X-10 facility**. By the summer of 1944 the primary goal of the war effort was achieved and the laboratory was transformed into the first well rounded institution for nuclear research. A short history of ORNL is available via the World Wide Web [2].

The major portion of the war effort at the Oak Ridge site [3] was devoted to obtaining enriched uranium. Part of this work was done at the Y-12 site using an electromagnetic separation process. The Tennessee Eastman Corporation operated the facilities at the Y-12 site from June 1943 to May 1947 and is referred to as the **TEC facility**. In May of 1947 work at the Y-12 site changed from uranium enrichment to nuclear materials fabrication and Union Carbide Corpo-

[2] See *The History of ORNL* at URL http://www.ornl.gov/swords/swords.html.
[3] See URL http://epawww.ciesin.org:7777/atsdr/oakridge.html

ration became the operating contractor. There was also a significant change in the workforce with only 6.7% of the workers staying on to work in the new operations. In this report **Y-12 facility** refers to the plant at the Y-12 site that began operation in May 1947 and continued until the end of this study. The main purpose of the facility at the K-25 site was to produce enriched uranium via the gaseous diffusion process. In support of the enrichment process the **K-25 facility** (later named the Oak Ridge Gaseous Diffusion Plant) also operated a plant that produced barrier material, ran a feed mill operation, and conducted laboratory research.

A more detailed description of these facilities and the potential for occupational exposure to hazardous material is given by Watkins *et al* [38, 39]. A total of 118,588 workers are known to have been employed at one of the four nuclear facilities in Oak Ridge between 1943 and 1984. The study cohort consists of 106,020 workers that were employed for at least thirty days and whose records did not have any critical errors (e.g. unknown sex, race, date of birth, or employment dates).

2.2. Definition of Study Factors

2.2.1. *External Radiation Exposure*

The nature and extent of the radiation hazard differed considerably among facilities and over time. A detailed description of the monitoring programs for external and internal radiation at each facility, and results summarizing the exposure characteristics by race and gender are given in [38, 39]. Annual external dose estimates were obtained for each monitored worker at each Oak Ridge facility. Evaluation of these results [38, 39] showed that over 93 percent of the total recorded external dose was received by the 28,770 white males who had ever been employed at the X-10 or Y-12 sites, and that about 30 percent of these workers were employed at more than one Oak Ridge facility. Results from the facility comparison analyses (see Results) show that most of the workers at K-25 and TEC had higher death rates than those at X-10 and Y-12, so inclusion of K-25 and TEC workers would potentially lead to overestimation of baseline mortality rates and underestimation of radiation effects. For these reasons the dose-response analyses for external radiation were restricted to white males who were ever employed at X-10 or Y-12. The other race/gender groups at these facilities were not included since they would contribute little to the assessment of effects of low-level radiation and would unduly complicate the dose-response analyses.

To provide a preliminary evaluation of the potential effect of “missing dose” on radiation dose-response estimates the doses for X-10 workers prior to 1957 and for Y-12 workers prior to 1961 were adjusted. The adjustment procedure increased dose estimates for some person-years under assumptions about the minimum detectable dose and the monitoring policies that were followed at X-10 and Y-12. Other sources of systematic and random measurement error are not addressed by this procedure. A detailed discussion of how these crude adjustments were made and summary results are provided in [38, 37, 39].

2.2.2. *Internal Radiation Exposure*

There were considerable differences in the potential for internal radiation exposure and the monitoring policies that were followed at each facility. For this reason internal radiation exposure was represented as a factor (that is referred to as IG) with three levels:

level 1- *eligible* for monitoring but *not monitored* (EN);

level 2- *eligible* for monitoring and *monitored* (EM); and

level 3- *not eligible* for monitoring (NE).

The precise definition of this variable was facility dependent and changed over time–see [38, 39]. Exposure analysis files that contained yearly values for external dose and the internal exposure indicator variable for each individual at each Oak Ridge facility were created (see Appendix). In the dose-response analyses radiation exposure is a time dependent variable, and cumulative dose (with a lag) was used to form dose groups for external radiation. For internal exposure an individual was assigned to the appropriate category based on the sequential yearly values for the internal exposure indicator variable.

2.2.3. *Additional Study Factors*

Facility. Workers were initially categorized according to their facility of first employment—X-10, TEC, Y-12, or K-25. An individual who worked at more than one Oak Ridge facility was assigned to a fifth category "MULT", at the time of first entry into the second facility. In the dose-response analysis of white males ever employed at X-10 or Y-12 (X-10/Y-12), an individual could also enter follow-up as a multiple facility worker if his entry into X-10 or Y-12 was preceded by employment at K-25 or TEC. Entry dates were adjusted to the appropriate lag (two, ten, or twenty years).

Socioeconomic Status (SES). Each individual was categorized as either a nonmonthly paid worker or a monthly paid worker based on the variable *Paycode* in the demographic analysis file [38, 8]. This variable is used as a surrogate for SES in the regression analyses with non-monthly workers as the referent category so that estimated coefficients represent relative risks for monthly workers.

Length of Employment. Length of employment was calculated as the length of time between first hire and last termination at one of the Oak Ridge facilities. In all analyses, length of employment was categorized into two levels: level 1- one year or greater; or level 2- less than one year. The reference level used is one year or greater, so that estimated coefficients represent relative risks for "short term" workers relative to "long term workers".

Birth year. Workers were divided into five categories according to year of birth: before 1900, 1900-1909, 1910-1919, 1920-1929, and 1930 or after.

Age. Age at risk is divided into 15 groups (15-, 20-, ..., 80-, 85+) starting at 15-19 and continuing through 85+. Age was alternately treated as a factor (AG) and as a continuous variate A = (age -52.5)/100, where age is the interval mid-point.

2.2.4. *Vital Status*

Vital status was ascertained primarily through the Social Security Administration. All cohort members with unknown vital status were allowed to contribute person-years up to the date at which they became “unknown”. Underlying cause of death and non-underlying cancer causes were coded to the International Classification of Disease, Adapted, Eight Revision (ICDA-8)—see [38, 39] for further details.

2.3. Data Reduction

The six analysis files described in [38, 37, 39] contain demographic, vital status, and annual exposure history data for each cohort member. Regression analyses are based on an analytic data structure (ADS) that consists of a table of observed deaths, “expected” deaths (based on U. S. rates), and person-years at risk for each combination of levels of three or more risk factors for each cause of death of interest that are derived from these files. In the simplest analysis the factors are age at risk (five year intervals), birth cohort, and facility as defined above. The next level of analysis focuses on the comparison of cause specific mortality at the Oak Ridge facilities that include length of employment, SES, and calendar year. The resulting ADS is a four dimensional table of observed and expected deaths for each cause of death. Results are presented in tabular form for race gender groups. The most complex ADS is required for the dose-response analyses of the X-10/Y-12 white males. Each cause of death requires a seven dimensional table of observed deaths, expected deaths, and person-years at risk. The factors that define the table are socioeconomic status (denoted by S with 2 levels), length of employment (denoted by L with2 levels), birth cohort (denoted by B with 5 levels), age at risk (denoted by AG with 15 levels), facility (denoted by F with 3 levels), internal exposure (denoted by IG with 3 levels), and external exposure (denoted by XG with 10 levels). The ten external dose categories were defined using cut points at 5, 10, 20, 40, 80, 160, 320, and 640 mSv. In the case of the last four (time dependent) covariates, workers contribute person time to all levels that they attain. An entry criterion of at least 365 days of follow-up was used for all analyses. Each observation in the ADS consists of the index value of each of the stratifying factors, the observed deaths, the expected deaths, the person-years, and the lagged average cumulative dose (see the Appendix).

2.4. Statistical Methods

2.4.1. *Traditional SMR Analyses*.

In these "external" analyses the mortality of Oak Ridge workers was compared with that of the general population of the United States using standardized mortality ratios (SMRs). SMRs were computed for race and gender groups for each cause of death category for which rates are provided by Monson [30]. The statistical strength of the deviation of observed deaths from those expected based on U.S. rates is indicated using the Freeman-Tukey residual (FTR) [13]. Under the null hypothesis the FTRs are approximately normally distributed with mean zero and variance one.

2.4.2. *Grouped Data Poisson Regression*.

The second and third approaches use Poisson regression methods [18, 17, 15, 5] to describe the joint effect of several risk factors on cause- specific mortality. The general form of the regression function is

$$E(y_{ijk}) = n_{ijk}\lambda_{ijk} = n_{ijk}\lambda^0_{ijk}R(Z_j, X_k), \tag{1}$$

where λ_{ijk} represents the unknown mortality rate, y_{ijk} is the number of deaths, and n_{ijk} denotes the person-years at risk in the ijk^{th} cell of an ADS. The i subscript refers to age, the j subscript indicates the dimensions of the ADS that correspond to factors of secondary interest (e.g., birth-cohort, SES) and the k subscript indicates exposure related covariates that are of primary interest, i.e. facility, internal exposure, and external exposure. The "external/internal" model is used—[5], chapter 4—in which the baseline rates (λ^0_{ijk}) are proportional to known external standard rates (λ^*_i), and the y_{ijk} are treated as observed values of Poisson variates with expectation given by Eq. (1). For a given cell in the ADS, the value of λ^*_i will also depend on birth cohort. **This approach has two advantages. First, it is not necessary to provide a parametric description of the age component of the baseline rates, and second the "intercept" terms (when all explanatory variables are at their reference level) provide estimates of the SMRs for the internal control group.**

2.4.3. *Facility Comparison Analyses*.

Facility is the risk factor of primary interest. Mortality differences between work-forces at these facilities reflect a combination of occupational and other factors. A detailed discussion of the potential for exposure to internal and external radiation and other related hazards at each facility is given by Watkins *et al* [38, 39]. In these analyses an evaluation of differences of cause specific mortality among the Oak Ridge plants is emphasized. A preliminary analysis of death rates due to all causes with three explanatory variables (birth cohort, age at risk, and facility) is presented to demonstrate the relationship between the two approaches that are used to "adjust" for age at risk. The rest of the results are based on Eq. (1) with a multiplicative main effects model using external age-cause specific death rates (from U.S. vital statistics) to compute expected deaths. The resulting ADS contains y_{jk} the observed deaths, μ^*_{jk} the "expected" deaths, and the level of each of the factors (which is equivalent to a covariate vector) SES, calendar period, length of employment, and facility. The multiplicative main effects model is expressed on a logarithmic scale as

$$\log\,(y_{jk}/\mu^*_{jk}) = F + S + L + t. \tag{2}$$

For convenience in describing results the convention—[12] Chapter 22—of dropping Greek letters (that represent the unknown parameters) and listing the explanatory variables that define the relative risk function is used. This corresponds to standard GLIM notation [14] for a log-linear model in which F is the facility factor (5 levels), S is the socioeconomic status proxy paycode(2 levels), L is the length of employment factor (2 levels), and t = (calendar year - 1965)/100, which represents a calendar period trend in the log of the SMR. The indicator variables for the factors F, S, and L are coded so that the coefficients for the facilities represent the log of the SMR for each facility at the reference level of S and L when $t = 0$ (i.e. in 1965 which is the midpoint of follow-up). The coefficient for t describes the change in the log SMR over follow- up. The results of fitting the main effects model for selected causes of death by race gender groups are presented in tabular form.

2.4.4. *Dose-Response Analyses***.**

These analyses are based on Eq. (1) with a multiplicative main effects model for the variable A, B, S, L, IG, and F. The effect of dose is represented with an exponential relative risk function or an additive excess relative risk (ERR) function. The main effects model with exponential relative risk is expressed as

$$\log\,(r_{ijk}) = A + B + S + L + IG + F + D, \tag{3}$$

where $r_{ijk} = y_{ijk}/n_{ijk}\lambda^*_{ijk}$, A = (age -52.5)/100, and D is external dose in Sv. In Eq (3) B, S, L, IG, and F are factors and A and D are continuous variates. Score test statistics [35] for external dose are presented for selected cause of death categories using all ten dose groups. Additional summary results—parameter estimates, standard errors, and likelihood ratio test (LRT) statistics—are given for each cause of death category with the highest dose group deleted (low dose analysis). The score test for D in Eq. (3) is identical [10] to that obtained for D variable in Eq. (4) below, and can be compared to the standard normal distribution to evaluate the strength of the dose-response relation. These analyses are then repeated using adjusted doses. The score test and low dose exponential relative risk are used as a screening procedure to identify cause-of-death categories that may show a strong association with dose.

Most summary statistics (estimates and SEs) for relative risk parameters are expressed in log percent (L%) units, i.e. they are given in logarithmic units multiplied by 100—see [12] Chapter 22,[36] and the Appendix. For the ERR estimates likelihood based intervals are given—see [1, 34].

More detailed results are presented for several cause of death categories using a main effects model with the additive excess relative risk (ERR) function to describe the dose-response relation for external radiation, e.g.

$$r_{ijk} = exp\,(A+B+S+L+IG+F)\,(1+D). \tag{4}$$

The main effects model provides an overall descriptive summary of the the effects of each stratification variables on cause specific mortality. Thus inclusion of these potential confounding variables (A,B,S,L) and exposure variables (F and IG) provides a broader context in which to evaluate the relative importance of the estimated effect of external radiation. A saturated model for the confounding variables AG,B,S, and L was also considered and was found to have little effect on the dose parameter estimate. Detailed results are given for all cancer in the Appendix (see Table AVII). The score statistic for a linear dose term for the main effects model and the saturated model (i.e. stratified analysis) were routinely calculated and no important difference for any cause of death categories were found.

A detailed analysis for all cancer mortality that uses the Akaike Information Criteria (AIC) to contrast the effectiveness of several exponential and ERR models is presented [2, 19]. For Poisson data the AIC = Deviance + α(number of parameters), combines a measure of the discrepancy between the fitted values and the data (as measured by the deviance) and the simplicity of the model as reflected in the number of parameters. McCullagh and Nelder—see Sec 3.9— [27] discuss a general approach to model selection which includes the AIC and note situations when α= 2 (used here) provides a reasonable choice. Ostrouchov and Frome [32] discuss various approaches to model selection based on AIC and related criteria where the objective is to find a subset of models that adequately describe a large ADS.

3. Results.

3.1. Traditional SMR Analyses.

Results of the SMR analyses by race and gender groups for selected cause of death categories are given in Table I. For white males the all cause SMR is 1.00 and the all cancer SMR is .98. Death rates are elevated in the Oak Ridge cohort for lung cancer, diseases of the respiratory system, and all external causes (note that large positive/negative values of the FTR indicate that observed deaths are higher/lower than expected after accounting for the number of deaths involved). Death rates are lower than the U.S. white male rates for cancer of the buccal cavity, digestive system, bladder, and thyroid. Death rates are also lower than expected for white males in Oak Ridge for diseases of the blood, nervous system, circulatory system, digestive system, and genitourinary system. For nonwhite males the all cause SMR is .96. The only elevated cancer SMR with a relatively large FTR is cancer of the large intestine. Death rates for diseases of the circulatory system, digestive system, and liver cancer are lower that U.S. nonwhite male rates.

3.2. Facility Comparison Analyses.

3.2.1. *Preliminary Results: White Males All Cause Mortality*

Facility is the risk factor of primary interest and our evaluation is limited to white males for all causes of death. A detailed discussion of the potential for exposure to internal and external radiation and other related hazards at each facility is given in [38, 39].

As a preliminary step in the comparison of the death rates of workers at the Oak Ridge plants an ADS was constructed based on birth cohort, age at risk, and facility. Each worker in a specified birth-cohort enters follow-up at the age of first hire, and subsequently contributes person-time at risk to appropriate five year age intervals until he dies or reaches the last day of follow-up. Poisson regression is used to estimate the age-specific death rates for each facility adjusted for birth cohort, and the 1915 birth cohort is used as the referent group. These estimated rates are shown in Fig. 1 for each of the Oak Ridge facility groups and the age-specific death rates for U.S. white males are also shown as the dotted lines. Fig. 1 shows that the log of the age-specific death rates increase linearly with age for all cause mortality, and that the rates for TEC, Y-12, K-25, and multiple facility workers are close to or above the U.S. rates, while X-10 rates are generally lower. To further evaluate the differences among the facilities a model is used to “adjust” for age, using two different approaches. In the *first approach* (internal analysis) the age-specific death rates are assumed to increase linearly with age at risk on a logarithmic scale (see Fig. 1). The regression function is

$$\lambda_{ijk} = exp\left[\alpha_{jk} + \theta(age - 50)\right], \tag{5}$$

where λ_{ijk} is the rate for the ith age group for birth cohort j and facility k, and α_{jk} represents the log of the age-specific death rate for the jth birth cohort in facility k at age 50 years. The maximum likelihood estimates of these “intercept terms” for this internal analysis are shown in panel A of Table II, and the corresponding estimated death rates (see the vertical axis in Fig. 1 corresponding to age 50) are given in panel B. This approach uses the well known Gompertz law of mortality to “adjust” for age. The maximum likelihood estimate of the “slope” parameter is $\hat{\theta} = 0.0794$ (SE = 0.00069) and this, together with the $\hat{\alpha}_{jk}$ can be used to estimate the median survival time, conditional on having reached age 50–see panel C of Table II.

Panels A-C of Table II are different ways to describe the all cause mortality experience of these facility/birth cohort groups. For statistical purposes the $\alpha' s$ are the parameters of direct interest, whereas the estimated death rates $exp(\alpha)$ are more convenient for interpretation. For example, for the 1915 birth cohort at age 50 the estimated annual death rates for K-25 workers is 9.48/1000, which is 1.62 times higher than the X-10 rate of 5.85/1000. Another way to state this result is that the estimated *relative risk for K-25 workers, with X-10 as the internal referent group* is 1.62. The estimated median survival time at age 50 for the 1915 birth cohort is 24.15 years for K-25 workers which is 5.34 years less than that for X-10 workers (see panel C of Table II) in the same birth cohort. The corresponding estimated median survival time for the 1915 birth cohort for U.S. white males at age 50 is 25.74 years.

In the *second approach* the external age-specific death rates λ^*_{ij} are used for each birthcohort and it is assumed that the rates for each facility/birth cohort group are proportional to the external rates. The regression function is

$$\lambda_{ijk} = \lambda^*_{ij} exp(\alpha^*_{jk}), \tag{6}$$

where the α^*_{jk} correspond to the log of the SMR for each of the birth-cohort/facility groups. These estimates are given in panel D of Table II in logarithmic percent units (L%) [36], and the corresponding estimated SMRs—exp $(\hat{\alpha}^*/100)$—are given in panel E. For example, for K-25 workers in the 1915 birth cohort $\hat{\alpha}^*$ = 10.8L% and SMR = 1.11, i.e. their rates are 10.8L% units higher than the U.S. white male rates which corresponds to an SMR that is elevated by 11%. The SMR can be interpreted as a relative risk with U.S. white males as the referent group. The estimates obtained using the external death rates can be used to make internal comparisons. For example, the relative risk for the 1915 birth cohort of K-25 workers with X-10 as the referent group in L% units is 10.8 - (-36.7) = 47.5L%, and the relative risk is exp(.475)= 1.61 which is in close agreement with the result obtained using the internal comparison approach. Table II shows that workers born after 1930 have much lower age-adjusted death rates than other Oak Ridge workers (see panel B) and than U. S. white males (see panels D and E). Further, X-10 workers have lower death rates than other Oak Ridge workers (panel B) and the U.S. white male comparison group (panels D and E). White males employed only at the K-25 site have the least favorable mortality experience with death rates that are higher than both other Oak Ridge workers and the U.S. except for those born after 1930.

This preliminary analysis illustrates two approaches to using Poisson regression methods to adjust for the effect of age at risk when comparing the mortality rates at the Oak Ridge facilities. To further describe the cause-specific mortality of Oak Ridge workers the analysis is extended to include the additional variables SES, length of employment, and calendar period

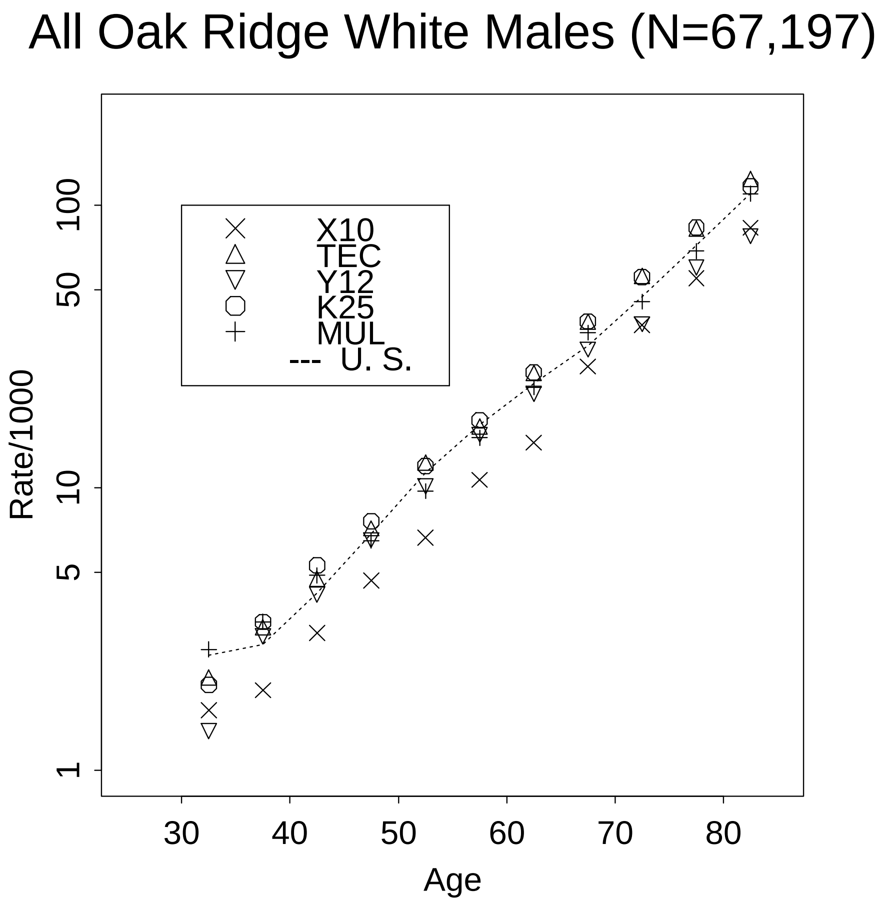


Figure 1: Estimated age-specific death rates for Oak Ridge workers for the 1915 cohort for all causes of death.

and have adjusted for age by using the external rates. This leads to a four dimensional ADS for each race/gender group for each cause-of-death of interest. Results for white males and females are given in Table III, and results for nonwhite males and females are given in Table IV for selected disease categories. Table V presents result for white males for selected cancer causes. A detailed description of the data reduction procedure and of how to read these tables is given in the Appendix. The observed deaths (column 2 of Tables III-V) are slightly less than those given in Table I since only individuals with at least one year of follow-up were included in these analyses.

3.2.2. *Detailed Results By Race/Gender Groups*

From column 10 of Table III it is clear that monthly workers have much lower death rates than nonmonthly workers. The all cause death rates are lower by 41.5 L% for both white males and females—this corresponds to a relative risk of exp(-.415) = 0.66 for monthly versus nonmonthly workers. The log relative risk estimates for length of employment are all positive—see column 11 of Table III. This may reflect a more hazardous environment and/or lower SES of short term workers. The facility effects in columns 4 through 8 of Table III are estimates of the SMRs (in L% units) for each facility at the middle of follow-up (1965-69) for nonmonthly long term workers, and provide a direct comparison with the external referent population. These estimates are internally "adjusted" for SES, length of employment, and calendar time; and for attained age through the use of the external rates. For example, for all cause mortality for white males employed only at TEC the facility effect is -5.9L%, and the estimated SMR is exp (-.059) = .94. An estimate of the log relative risk with one facility as the referent facility is obtained as the difference of the two facility effect estimates. Using the facility effect estimates for all cause mortality for white males (see line 1 columns 4-8 of Table III) with X-10 as the referent group leads to:

	X-10	TEC	Y-12	K-25	MULT
Effect Difference (L%)	0.0	20.9	9.7	25.8	20.8
Relative Risk	1.0	1.23	1.10	1.29	1.23 .

This *internal comparison* shows that white males employed only at TEC or K-25 and multiple facility workers have higher death rates than similar workers employed only at X-10 or Y-12. The large value of the LRT statistic in column 9 of the Table III—LRT = 101.0, 4 degrees of freedom (df)—indicates that the facility differences are not due to small numbers of deaths, i.e. unstable rates. Evaluation of the corresponding estimates for white females (see row 1 of the bottom panel of Table III) reveals a similar pattern. The facility effect differences for all cause mortality for white females (using X-10 as the referent group) are

	X-10	TEC	Y-12	K-25	MULT
Effect Difference (L%)	0.0	24.6	-23.0	41.0	40.3
Relative Risk	1.0	1.28	0.79	1.51	1.50 .

As was noted earlier, the facility effect estimates for females are possibly biased downward as the result of under ascertainment of vital status, but the internal comparisons are valid if failure to determine vital status is the same across facility groups.

Due to the potential for vital status ascertainment problems and the small size of the cohort of nonwhite Oak Ridge workers (3,909 male and 2,826 females), the results in Table IV have been limited to a few selected causes. The results for nonwhite Oak Ridge workers in Table IV follow the same pattern that was observed for white Oak Ridge workers. Monthly workers have much lower death rates than nonmonthly workers, and short term workers have higher rates than long term workers. For nonwhite females there were so few monthly workers, that estimates of the relative risk for this factor could not be calculated.

Table V shows the results of the main effects analyses for selected cancer causes for white males. Death rate differences among the facilities are large enough and based on a sufficient number of deaths to yield large LRT statistics for lung cancer (LRT = 21.6, 4 df), leukemia (LRT = 10.7, 4 df) and other lymphatic tissues (LRT = 12.3,4 df). The facility effect estimates indicate that lung cancer is low at X-10 (-34.8L%), and that other lymphatic tissue rates are low for X-10 (-92.8L%) and highest for Y-12 (71.9L%). The leukemia estimate is lowest at Y-12 (-64.5L%) and is high for X-10 (61.7L%).

3.3. Dose-Response Analyses.

3.3.1. *Selected Results for Main Effects Model Using ERR*.

The marginal distribution (dose group by facility) of person years, observed and expected cancer death, and average dose is given in Appendix Table AIV. The parameter estimates for the ERR main effects model (see Eq. 4) with external rates and using unadjusted doses with a ten year lag are given in Tables VI and VII. A detailed description of the data analysis procedure and example of how to interpret the parameter estimates is given in the Appendix. The *referent group is nonmonthly workers employed at least one year at the X-10 facility who were eligible but not monitored for internal radiation exposure and had zero external occupational dose*. The parameter estimates for the factor B (rows 2-6 of Tables VI and VII) are estimates of the log SMR for the referent group for each birth cohort for each cause of death. The large negative values for all causes of death (column 2 of Table VI) indicate a strong “healthy worker” selection effect for the X-10 facility, and that the effect is strongest among workers born after 1930 (SMR = exp [-0.63] = 0.53). Review of the birth cohort effect estimates shows a similar pattern for diseases of the circulatory system and external causes of death. For all cancer causes lower rates were observed primarily for those born after 1920 (see column 4 of Table VI). By contrast, the birth cohort effect estimates for selected cancer causes in Table VII show a much different pattern. The birth cohort effect estimates for those born before 1930 are large and positive for cancer of the prostate and leukemia, indicating that for the internal referent group these death rates are higher that those for U. S. white males.

The SES effect estimates (row 8 of Table VI) are large and negative indicating that death rates are substantially lower for monthly workers than for nonmonthly workers. For example, for all cancer the relative risk for monthly versus nonmonthly workers is exp(-.419) = 0.66 with (95% CI: 0.56, 0.78). The SES effect is most pronounced for lung cancer where the relative risk for monthly workers is 0.43 with (95% CI: 0.31, 0.59).

The log relative risk for all cause mortality for short term workers is 9.9L%, mostly due to higher rates for external causes. Differences in mortality rates among the three levels of the internal radiation risk group (IG) were unstable and in varying directions (see lines 11-13 of Tables VI and VII). LRT statistics (not shown in the table) for the null hypothesis of no difference among the levels of this factor were all below the 95 percentile of the chi-square distribution with 2 df.

There are differences in the death rates in the facility groups in the X-10/Y-12 subcohort. All cause death rates for those employed only at Y-12 are 12.4L% higher than X-10 only workers, and multiple facility workers are higher by 8.4L% (see lines 14-16 of Tables VI and VII). These differences appear to be primarily due to higher lung cancer rates for the Y-12 workers (log relative risk = 46L%) and multiple facility workers (log relative risk = 33L%). A notable exception occurs for leukemia where the rates are substantially lower for the Y-12 (-121.4L%) and multiple facility (-64.0L%) workers than for X-10 workers.

The last lines in Tables VI and VII provide an estimate of and 95% CI for the ERR per Sv. The estimated dose-response relations are represented graphically in Figs. 2 and 3 (see Section 3.3.2 for additional discussion). The average lifetime dose for X-10/Y-12 subcohort is about 10 mSv (see Fig.4). This coefficient can be used to compute the relative risk of the "average" worker in the subcohort of Oak Ridge workers with potential for exposure to external penetrating radiation. The result can be compared with the other estimates of the log relative risk parameters for each cause of death. Using all cancer as an example, the log relative risk for the average worker is about 100*log(1+ 0.0145)= 1.44L%. By contrast, the effect of being a nonmonthly worker (see Table VI line 8) is 41.4L %, suggesting that SES factors are far more influential than external radiation exposure in determining cancer risk in this group of workers.

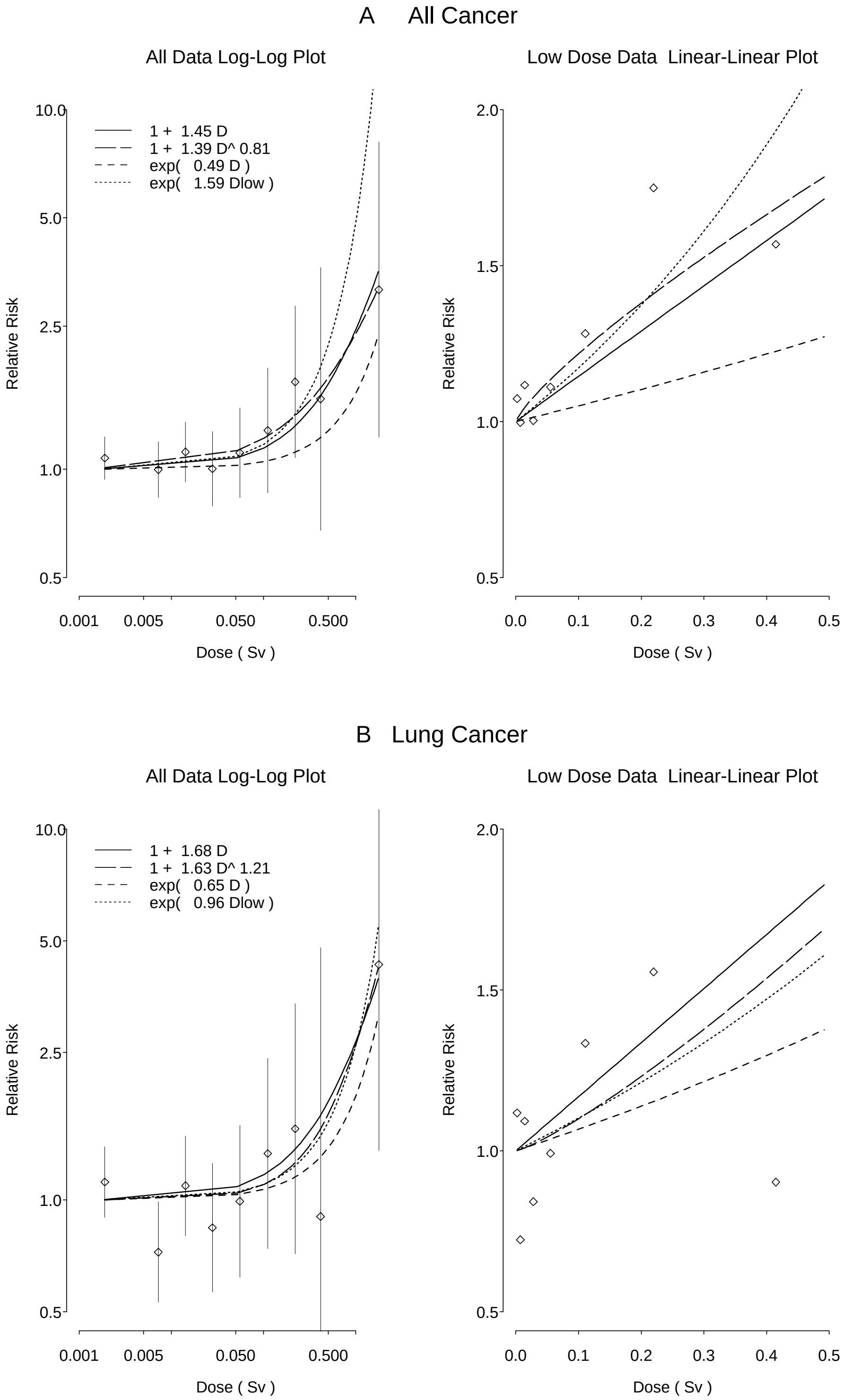
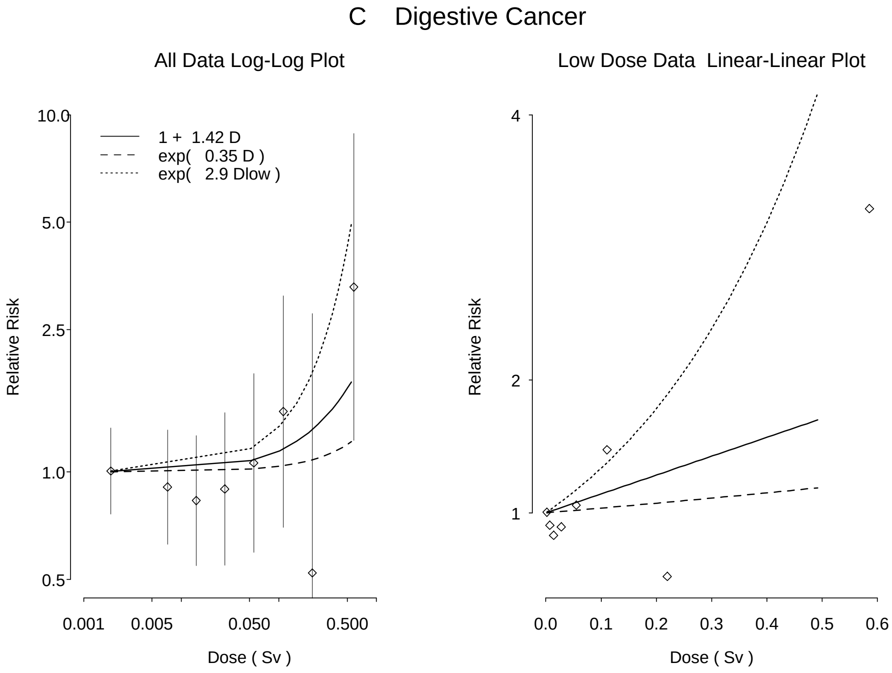


Figure 2: Fitted dose-response functions: all cancer and lung cancer. Point estimates and 90% CIs (vertical lines) are shown for each dose group. Dlow indicates that high dose group was not used in the fit.



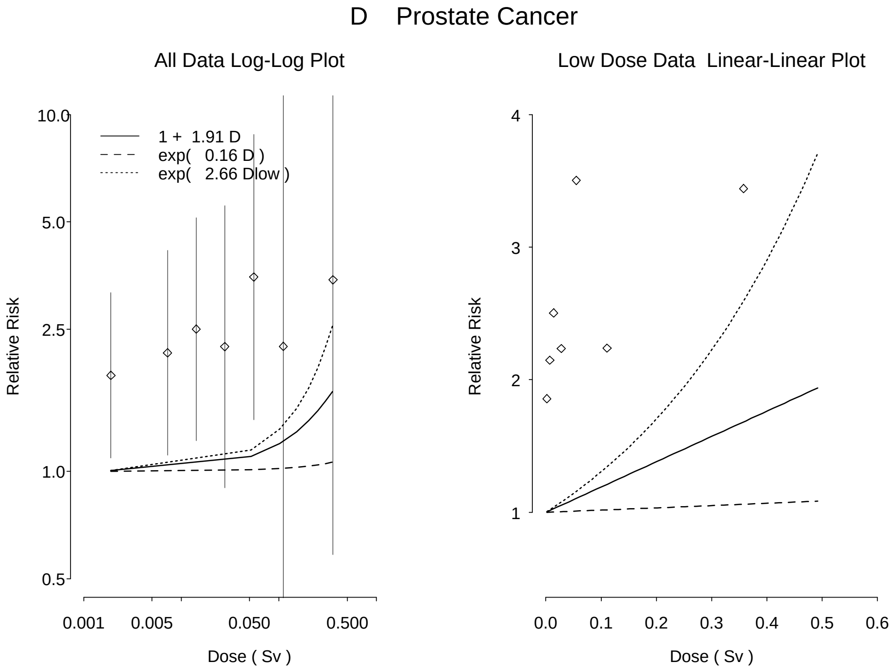


Figure 2: (Cont.) Fitted dose-response functions digestive and prostate causes. Point estimates and 90% CIs (vertical lines) are shown for each dose group.Dlow indicates that high dose group was not used in the fit.

3.3.2. *Results of the Screening Procedure for Selected Causes of Death***.**

Table VIII gives the unadjusted SMR (in % and L%), the value of the score test statistic based on all of the data, and the estimated dose-response coefficient based on the main effects model with exponential relative risk using the "low dose" data. In this low dose analysis all cells in the ADS in the highest (640mSv+) dose group were omitted so that once an individual's cumulative dose exceeded 640 mSv he was censored. This choice was based on the results of the regression diagnostics —see Fig. AII in the Appendix—and the fact that a radiation worker would be very unlikely to receive a cumulative lifetime dose above 640mSv. There are 41 cells with 220 persons years (total person years is 603,365) and 3 cancer deaths in the highest dose group (see Appendix Table AIV). The regression diagnostics in the Appendix (which are based on Eq. 3 using all data) suggest these few cells with cumulative dose values that exceed 640 mSv have a large influence on the estimate of the dose-response coefficient for the exponential relative risk model. For the X10/Y12 subcohort about 0.1 percent of the workers exceeded this value. (It may be more appropriate to think of this as a low to medium dose analysis.)

The effect of the high dose group is to decrease the trend estimate when the exponential model is used, since most of the 41 cells have no events and high leverage values. For all cancer causes the estimate obtained using all the data is lower than that obtained when cells for the highest dose group are omitted. The low dose estimate (1.59 per Sv) is three times larger than the estimate using all the data (0.49 per Sv). Fig. 2A shows the relative risk estimates for all cancer for each dose group (ten year lag). These estimates were obtained using the ten level factor XG in the main effects model and are therefore "adjusted" for factors B, S, L, IG, F, and age through use of the external rates (see Appendix). Fig. 2A shows that the exponential dose-response is reasonable in the low dose region, but that it does not provide a good description of the relation between risk and dose over the range of doses in this study. Note that while the relative risk for the highest dose group is lower than that predicted from the exponential model fitted over the lower doses, it is roughly in line with the linear ERR model.

The score test values in column 6 of Table VIII are based on all the data and the LRT values in column 9 were obtained with the cells in the highest does category excluded. When all the data are used the results for all cancer and lung cancer show the strongest association between external dose and deaths certified to these causes (see column 6 of Table VIII). When the high dose data are excluded there is a decrease in strength of the association for all cancer, and a large decrease occurs for lung cancer (see LRT statistics column 9 of Table VIII). The opposite pattern is observed for digestive system cancer and emphysema. Estimates of the relative risk by dose group and those derived from the exponential relative risk and ERR models are shown graphically for lung cancer in Fig. 2B and for digestive cancer in Fig. 2C. For digestive cancer there were no deaths in the highest dose group so the relative risk was estimated for the last two groups combined. The dose values used for the graphical displays in Figs. 2 and 3 are the

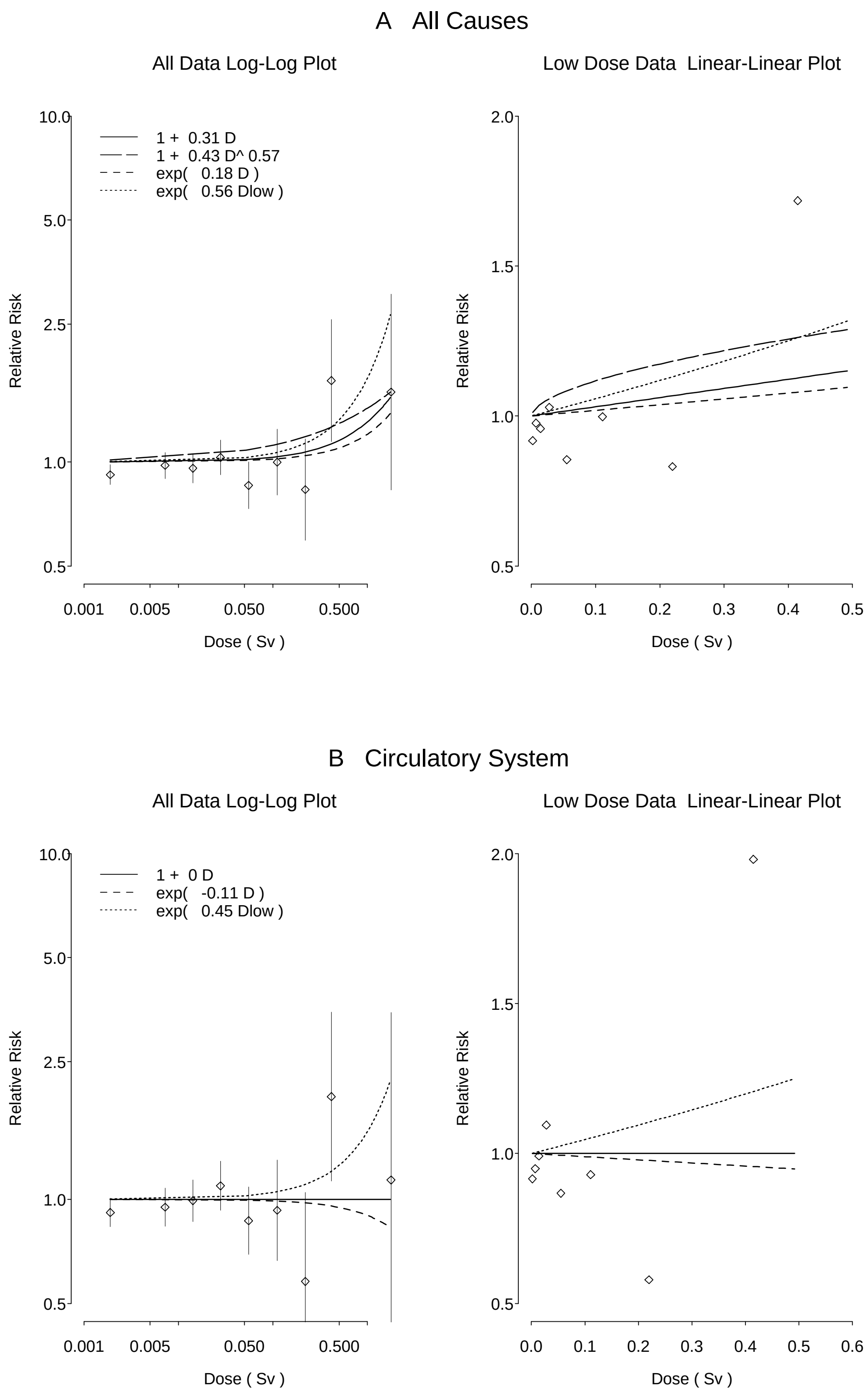


Figure 3: Fitted dose-response functions for all causes and circulatory system. Point estimates and 90% CIs (vertical lines) are shown for each dose group. Dlow indicates that high dose group was not used in the fit.

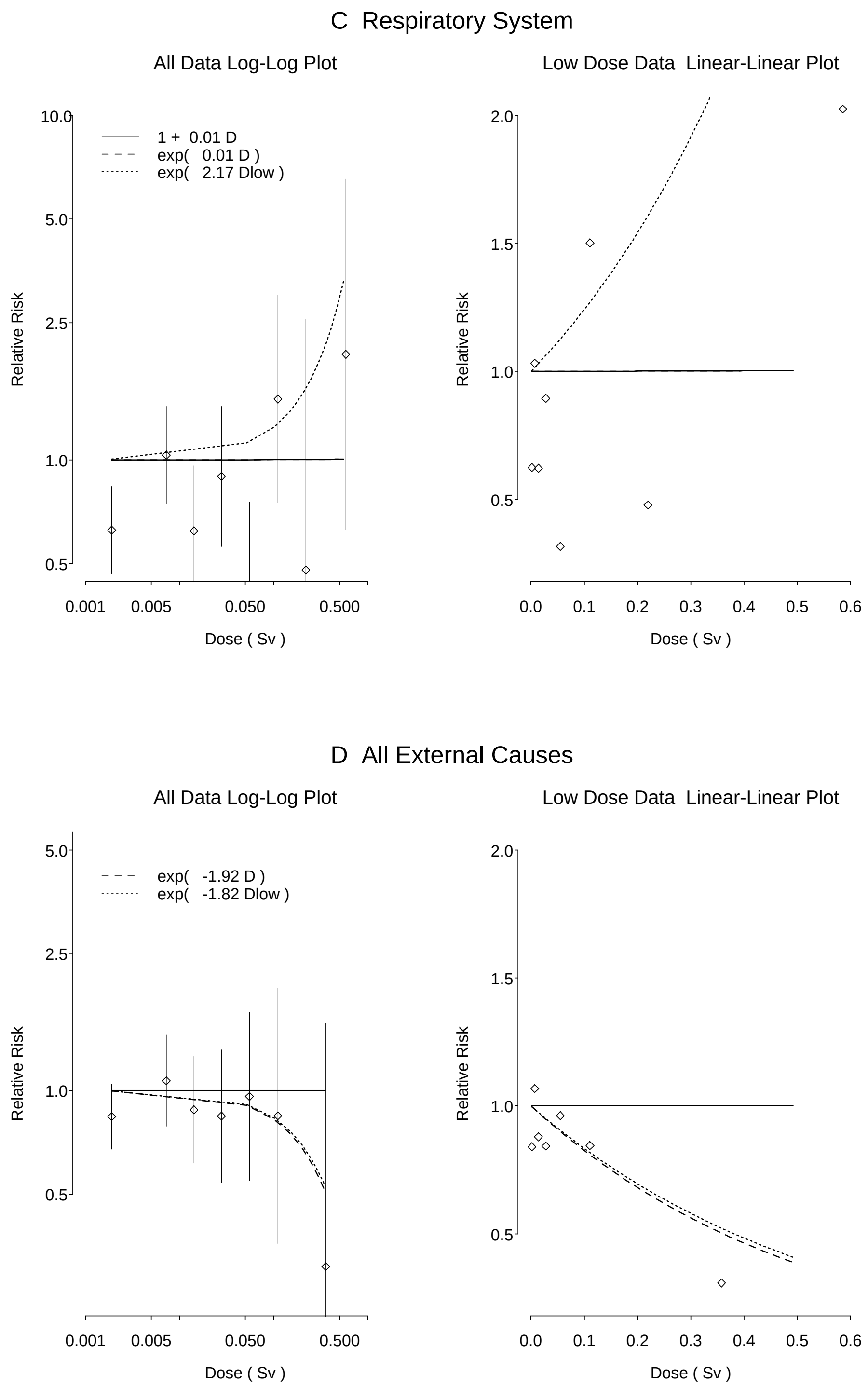


Figure 3: (Cont.) Fitted dose-response functions for respiratory system and external causes. Point estimates and 90% CIs (vertical lines)are shown for each dose group. Dlow indicates that high dose group was not used in the fit.

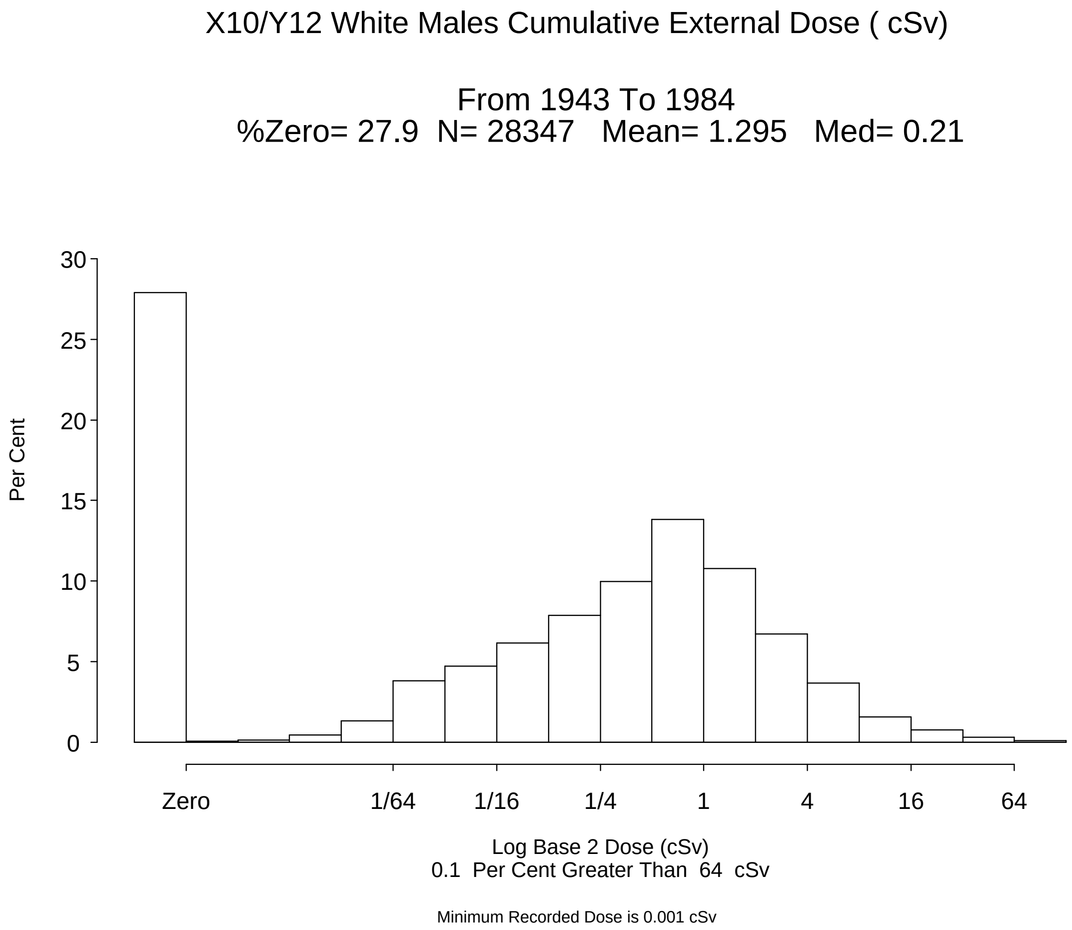


Figure 4: Cumulative External Dose for X10/Y12 White Males

person-year weighted average of the doses associated with each stratum in the ADS (see the Appendix) for a given dose category. For prostate cancer there were no deaths in the highest two dose categories so the relative risk at the highest dose value in Fig. 2D is a combined estimate for the three highest dose groups. The dose-response coefficient for prostate cancer is positive but the score test and LRT do not show a strong association with dose (see row 6 of Table VIII). Fig. 2D shows that the risk is increased for workers with external doses greater than zero (recall that the internal referent group is based on cumulative lifetime occupational dose with a ten year lag). An *ad hoc* analysis yields a relative risk estimate of 70.4L% (SE = 29.8) for those with any external radiation exposure versus those with zero dose. This *ad hoc* result was obtained using the main effects model with an indicator variable for external radiation exposure. The unadjusted SMR for the ten year lag zero dose group is -19.8L% (SE = 21.8), and for those with dose greater than zero the unadjusted SMR is 16.4L% (SE = 13.4).

Fig. 3A shows relative risk versus doses for all cause mortality. Figs. 3B-3D show similar results for diseases of the circulatory system, nonmalignant respiratory disease, and all external causes of death. Dose-response coefficients, score test statistics, and LRT statistics are given in Table VIII.

3.3.3. *Dose-Response Using Adjusted Doses.*

Table IX shows the effect of using the adjusted doses on the score test values and the dose-response coefficients and LRT statistics (for the low dose data) for each cause of death category that was presented in Table VIII. The results suggest that the effect of missing dose is an overestimation of the strength of the dose-response association, and an upward bias in the dose-response coefficients.

3.3.4. *ERR Estimates by Facility and Lag for All Cancer.*

To further evaluate the effect of the dose adjustment procedure at each facility on the ERR dose-response coefficients, a more detailed analysis was done using two year, ten year, and twenty year lag intervals for all cancer causes. Facility differences in the impact of the dose adjustment procedures are of interest because external radiation monitoring differed in coverage and other respects at X-10 and Y-12 before 1961. Most dose adjustment at Y-12 resulted from the replacement of missing values for unmonitored workers with one of three assumed values, whereas at X-10 most workers were monitored for external radiation after 1947 and individual yearly doses for each worker were adjusted using the preliminary *ad hoc* procedure [38, 39]. In the analysis with follow-up through 1984, the cumulative adjusted dose assignments for the twenty year lag are based almost entirely on annual doses obtained during the period when the adjustments were made.

The first row of panel A in Table X shows the estimated ERR per Sv for each lag using the unadjusted doses. The second line gives the value of the LRT statistic comparing the estimated ERR to the assumption of zero slope. Lines three, four and five show the estimated ERRs for each facility separately, and the final line is the LRT for facility differences in slope compared to the assumption of a common slope. Panel B gives the same results for adjusted doses. The estimated ERRs for all X-10 and Y-12 workers combined based on unadjusted doses are approximately 50% larger than those based on adjusted doses, and the largest ERR and LRT values are observed for the ten year lag.

The lower lines of each panel show that the change in ERR with increasing lag differs between X-10 workers and the other groups. Estimated ERRs at X-10 increase with increasing lag for both unadjusted and adjusted doses, while estimates for Y-12 and multiple facility workers are positive for two and ten year lags but negative for the twenty year lag, the period when the majority of Y-12 workers were not monitored for external radiation and adjusted annual doses for Y-12 workers took on one of three assigned values. The LRT for facility differences in ERR increases with increasing lag. The difference between ERRs based on adjusted and unadjusted doses is proportionately larger for Y-12 than for X-10 workers. Unadjusted values for multiple facility workers are closest to zero and change the least with adjustment assumptions.

3.3.5. *Alternative Representations of Dose-Response Relation***.**

The results in Tables VIII and IX and Fig. 2 show that the restriction of results to the low dose region has a stronger effect on β (using the exponential relative risk model) than does dose adjustment. The graphical results in Fig. 2A and regression diagnostics (Appendix Fig. A2) further indicate that the main effects model with cumulative dose represented as a linear term in the exponent (see Eq. 3) does not provide an adequate description of the dose-response relation over the entire range of doses. This is the dose-response model that was used in previous studies of the X-10 only subcohort of Oak Ridge workers [40, 41]. The results presented in Tables VI and VII are based on the ERR (see Eq. 4) which has been widely used in radiation epidemiology [20, 31, 6]. These two regression functions cannot be compared directly using likelihood ratio tests but can be compared indirectly using the values of the deviance and df for each model. A more extensive analysis of all cancer mortality was conducted to further explain and clarify the differences between the ERR and exponential relative risk functions. Table XI contains the resulting summary statistics for seven relative risk functions for the all cancer mortality using unadjusted doses with a ten year lag for the X-10/Y-12 subcohort. This analysis is based on a main effects model with external rates for the baseline risk (see Appendix). The deviance for a Poisson regression model is an overall summary of the discrepancy between the fitted values and the data for each cell in the ADS and provides a measure of "unexplained variation" similar to the residual sum of squares in least squares regression, i.e. smaller values indicate a "better fit". The deviance for each relative risk function in column 1 is given in column 3, and these statistics can be used to calculate LRTs for nested models. The deviance for the null model of no dose-response relation (i.e. a constant relative risk of one) is 2020.05 (see line 1 of Table XI). The value of the LRT statistic (3.55) for the null hypothesis of zero slope in the exponential model is obtained by subtracting the deviance on line 2 from that on line 1. The last line in Table XI is an unconstrained model, i.e. a relative risk parameter is estimated for each dose group, and can be used to construct a "lack-of-fit" test for any of the other models in Table XI that impose constraints on the relative risk- dose relation [19]. For example, the lack-of-fit test for the linear excess relative risk function (line 4 of Table XI) with the unconstrained model as the alternative yields a LRT statistic of 2014.75 - 2012.08 = 2.67 with 8 df, which does not indicate lack-of-fit. The LRT for zero slope for the linear ERR model is 5.3 with 1 df. **These results suggest that the ERR model provides a better description of the relative risk dose relation over the entire dose range than the linear exponential relative risk function.**

Another less formal approach that can be used to identify "good" models is the AIC (small values indicate a "better fit"). The AIC values in column 4 of Table XI suggest that the best model among those considered is the linear ERR (see line 4) since it has the smallest AIC value. The fitted values for this linear ERR function are shown in Fig. 2A, where it appears to describe the relative risk over the entire dose range. The exponential relative risk estimates based on the low dose data are close to the linear ERR estimates in the low dose range. Results for a power law model for the ERR are given on line 6 of Table XI and in Fig. 2A. Both the power law model and the linear-quadratic model (line 5 Table XI) can be used to construct likelihood ratio lack-of-fit tests for the linear ERR dose-response. These results further support the conclusion that the linear ERR model is a reasonable representation of the relative risk dose relation over the entire dose range. Fig. 2 also presents the results of fitting the exponential and linear ERR function to lung, digestive, and prostate cancer. Fig. 3 presents similar results for all cause mortality, diseases of the circulatory system, nonmalignant respiratory disease, and all external causes.

4. Discussion.

This study presents results of the combined analysis of all workers employed for at least thirty days at one or more of the four nuclear industry facilities in Oak Ridge, Tennessee. Numerical and graphical summaries based on Poisson regression methods that effectively describe the joint influence of several risk factors on cause-specific mortality were presented. The primary focus of this report was on the statistical aspects of this multiple facility analysis. A detailed account of the nature and extent of the radiation hazards and monitoring programs is presented in a companion report [38, 39].

4.1. Results For SMR Analyses.

All cause mortality rates for male Oak Ridge workers over the entire follow-up period were similar to national rates (Table I; SMR = 1.00 and 0.96 for white and nonwhite males,respectively). This is unusual in occupational studies where workers in large industries typically show a "healthy worker effect"— see Checkoway *et al*[9] chapter 4. One possible explanation is the large proportion of male workers who were hired at young ages during the war years – all cause SMR = 1.11 [16] – and who worked only for a short amount of time. These may have been transient workers not eligible for the draft due to poor health, and/or they may have been subjected to more hazardous working conditions because of the war effort. Notable excesses in mortality occurred only for white males for lung cancer and nonmalignant respiratory disease, and notable deficits occurred for diseases of the digestive system (both malignant and nonmalignant) and diseases of the circulatory system. For both white and nonwhite females the SMRs based on U.S. rates are generally lower than one. This apparent negative bias may be due to underascertainment of vital status. It is known that females are more difficult to follow through some national data sources. If, however, failure to determine vital status is not differentially related to one or more risk factors of interest, then internal comparisons of disease rates are appropriate. This assumption is implicit in our presentation of results for the facility comparison analyses.

4.2. Results For Facility Comparison Analyses.

A more detailed analysis based on facility, SES, length of employment, and period trend revealed substantial differences in the death rates among workers at the four Oak Ridge sites for white males and females (see Table III). Estimates of the internally adjusted log SMRs (facility effect estimates) show that the TEC, K-25, and multiple facility workers have higher death rates than similar workers employed only at X-10 or Y-12. The LRT statistics indicate that the differences among the facilities are primarily due to noncancer causes of death. However, further analyses of selected cancer causes (see Table IV) shows that there are large differences among the four facilities for lung cancer, leukemia, and other lymphatic cancer. Monthly paid workers had substantially lower mortality than nonmonthly workers; this is in agreement with other studies showing socioeconomic differentials in mortality. For white males, paycode differentials in mortality were greatest for respiratory diseases in general and emphysema in particular. These observations are consistent with studies showing socioeconomic differentials in smoking. Excess mortality was also observed for workers employed for less than one year, and among white males the relative risk was greatest for external causes of death. Over calendar time SMRs tended to increase. For white males, all cause SMRs increased at an average rate of 0.49L% per year, with values of 0.58L% for all cancer, 0.59L% for diseases of the circulatory system, and 0.97L% for respiratory diseases. Among the causes of death listed in Table III for white males, only emphysema showed decreasing SMRs over time. The magnitude of changes in SMRs with calendar time and differences in trends for causes of death of interest, including cancer and smoking-related diseases, suggests the importance of continued follow-up of these populations.

4.3. Dose-response Analyses For X-10/Y-12 Subcohort.

All dose-response analyses were restricted to the subcohort of white males employed at X-10 or Y-12 (see Materials and Methods). In previous studies with follow-up through 1984 the dose-response analyses were based on white males (N = 8318, with 346 cancer deaths) employed only at X-10 [40, 41]. The addition of Y-12 and multiple facility workers increased the cohort size to 28,347 with 1,038 cancer deaths. The methods used in this study differ in several ways from those in our previous analysis of the X-10 only group. **First**, there were ten dose groups equally spaced on a logarithmic scale (instead of eight equally spaced on a linear scale with a width of 20mSv). **Second**, the person-year weighted average dose was used in each cell of the ADS. In the previous analysis the median value of all of the person-years in all of the cells in the highest (greater than 120 mSv) dose category was used as the value of D for all cells in the highest dose group, and the interval mid-point was used for all of the other cells. **Third**, external rates and a multiplicative main effects model were used to describe the baseline rates, and internal radiation monitoring status and facility factors were included in the model. In the previous dose-response analyses external rates were not used and the baseline rates were described with a parametric model – see the Appendix for further discussion. **Fourth**, in previous dose-response analyses deaths where cancer was a "contributing cause" were included. Only underlying causes of death were used in the current study. **Fifth**, all of the previous dose-response analyses were limited to the exponential relative risk model.

In all dose-response analyses the potential biases associated with time-related factors and time dependent exposures were dealt with through the use of an internal control group based on birth cohort, age at risk, and length of employment. Different approaches to dealing with time dependent variables are possible—see e.g.[5, 9] and alternative approaches to dealing with these and other variables in mortality studies of nuclear industry worker in the U.S. [20] and the United Kingdom [3, 22, 24, 7] have been presented.

4.3.1. *All Cancer Mortality With a Ten Year Lag.*

For all cancer mortality with a ten year lag the estimated ERR for the X-10/Y-12 subcohort was 1.45 per Sv with a likelihood ratio based 95% CI =(0.15, 3.48) - see Table XI and Fig. 2A. Results were also presented for all cancers for two and twenty year lags by facility group. The results based on unadjusted doses for the X-10 only subcohort (see line 3 of Table XI) are consistent with results in earlier studies which showed a positive dose-response relation for all cancers [40, 41]. The dissimilarity of dose response associations between X-10, Y-12 and multiple facility workers increases with the lag assumption, as indicated by both the slope estimate and the LRT. A similar pattern is observed for adjusted doses. Results for the twenty year lag are based on doses received before 1965, the time period when radiation monitoring was substantially more complete at X-10 than at Y-12—see [38, 39].

4.3.2. *Results For Leukemia and Prostate Cancer.*

Most dose-response analyses were based on unadjusted doses with a ten year lag. There was no evidence for an association between leukemia deaths (based on a two year lag) and external radiation dose. Except for those born after 1930, leukemia mortality rates at X-10 were higher than those for U. S. white males and higher than those for similar Y-12 and multiple facility workers (see Table VII). See the next to last paragraph of this section for further discussion of leukemia rates. Prostate cancer rates were elevated relative to the U. S. rates and are about two times higher in workers with any amount of recorded external doses relative to those who received no occupational dose (see Fig. 2D). There was, however, no evidence for a smoothly increasing dose-response for prostate cancer. Two studies of UK workers have reported statistical associations between radiation dose and mortality from cancer of the prostate [4, 3]. Cardis and colleagues [6], however, note that more recent studies indicate that the association with external dose was largely a result of correlation between external radiation and radionuclide contamination.

4.3.3. *Results For Lung Cancer.*

The only specific cancer for which there was evidence for a positive association with external radiation was lung cancer (see Tables VII and VIII and Fig. 2B). There were two lung cancers in the highest dose group and exclusion of high dose data results in a LRT statistic of 0.43 for the multiplicative dose-response model. The estimate of the lung cancer ERR per Sv is 1.68 with a likelihood ratio based 95% CI of (0.03, 4.94). Results in Table VII show that there was a strong SES effect, that baseline rates were higher for Y-12 and multiple facility workers, and that the relative risk for workers monitored for internal exposure was slightly elevated. Information on cigarette smoking for this cohort is not available for analysis and residual confounding by cigarette smoking cannot be ruled out—see Wing *al*[41] for further discussion and references. Such confounding could bias dose response estimates in either direction. There was no evidence for an association between diseases of the circulatory system or nonmalignant respiratory disease and external radiation (see Tables VII and VIII and Fig.s. 3B and 3C). Note, however, that the largest LRT statistic in Table VIII occurred for emphysema for the low dose data (there were two deaths in the 320 to 640 mSv dose group with 0.16 expected based on U. S. rates).

4.3.4. *Effect of "Missing Dose" on Analyses.*

Another important issue that has not been considered previously for this cohort is the "missing dose" that may have occurred as the result of recording a zero for below-detectable doses, especially for frequent (weekly) film badge readings at X-10. The results presented here are based on a preliminary dose adjustment procedure that was derived from a small sample of hard copy X-10 personnel dose records [38, 39, 37]. The sample was needed because daily and weekly radiation monitoring data prior to 1957 are not available in electronic form. Future studies that include X-10 workers should use all available dosimetry data to deal with the uncertainty (systematic bias and random measurement error) in these data [23, 29, 42]. Our preliminary dose adjustment method does not consider other sources of systematic (e.g. neutron exposures) and random measurement error, nor did we consider the effect of measurement error on other covariates (internal exposures, SES) used in the analyses. Results based on adjusted doses are consistent with earlier studies, but suggest that the effect of missing dose was to introduce an upward bias in the dose-response coefficient and the score test statistics. Differences in measurement of external radiation at X-10 and Y-12 are relevant to the results reported in Table X, which show consistently increasing dose coefficients for longer lags only at X-10. The lag-20 coefficients are based entirely on exposure measurements made before 1965. For most of those years, less than one quarter of Y-12 workers routinely wore external dosimeters, while the proportion for white male X-10 workers after 1946 was much higher [42]. If there was an association between radiation exposures and cancer mortality during this period, it would be difficult to detect if few workers were monitored. In addition, recorded external exposures were higher at X-10 than at Y-12 during this period, resulting in a less skewed dose distribution upon which to base the dose- response estimates.

4.3.5. *Results For Internal Radiation.*

The factor IG was included as a crude indicator of the occupational hazard associated with internal radiation monitoring status in all dose-response analyses. The results for this factor are never impressive and no doubt mix selection factors (including facility selection) with any effect of heterogeneous exposures. It is likely that these estimates are affected by measurement error since the accuracy and completeness of internal exposure data varies considerably by facility and over time–see [38, 39]. The EM coefficient for lung cancer is positive and may reflect the contribution of Y-12 workers to this category [25]. The coefficient for circulatory diseases is negative, suggesting that the positive EM coefficient for lung cancer is not merely a smoking-related phenomenon. The NE group has a large coefficient for external causes, which has been interpreted in other places as an SES effect. This is noteworthy because it is adjusted for paycode, facility and cohort.

4.3.6. *Comparison With Results From Recent Studies.*

In the most recent combined analyses of the data for workers in several countries exposed to low-level external radiation, Cardis *et al.* [21, 6] present an excellent discussion of the importance of and difficulties encountered in studies of occupational effects of external radiation. The subgroup of workers from Oak Ridge that are in these combined studies are referred to as the ORNL subpopulation and correspond to X-10 only workers in this report. Their estimate of ERR per Sv for the ORNL subpopulation for all cancer excluding leukemia is 1.66 90% CI = (.04, 4.4) and for leukemia excluding CLL it is -1.06 with 90% CI = ($<$ 0, 4.8)—see [6] Table V). These results are consistent with the results presented here for the larger X-10/Y-12 subcohort of Oak Ridge workers that were potentially exposed to external ionizing radiation. Cardis *et al* —see [6] Table IV—give estimates of ERR per Sv for all cancer for their combined data of -0.02 90% CI: (-.34, .35); for leukemia the estimate is 1.55 with 90% CI: (-0.21, 4.7).

A recent high dose study by Pierce *et al.*[33] presented risk estimates based on mortality in the Atomic Bomb Survivors through 1990 (see Table AI). Their excess relative risk estimates for all cancers [0.37 per Sv 90% CI= (0.31,0.44)] and lung cancer [0.42 per Sv 90% CI= (0.24,0.63)] are compatible with those derived here. Failure of this study to detect a significant dose-response for leukemia may be due to a lack of power. Another possibility is the presence of an unidentified toxic chemical that is only present at X-10 (e.g. some type of solvent) that is independent of, or negatively correlated with, external dose, and causes leukemia. This would also be a possible explanation of the increase in leukemia risk at X-10 relative to U.S rates and other Oak Ridge facilities.

The combined risk estimates derived from the present study are generally comparable to the positive dose response estimates in Pierce *et al.* for those categories where there were larger numbers of deaths. However, this study also suggests that there are differences in dose response between worker cohorts in the same location followed over about the same time period, and that these differences are associated with measurement issues. Specifically, dose response associations are stronger for X-10 workers, the population with more complete external radiation monitoring. Differences in dose response between X-10 and other workers are larger under longer latency assumptions that depend more on exposure classification based on radiation data from that historical period when monitoring differences between populations were greatest. This observation suggests that investigations of external radiation effects in combined facility studies should be sensitive to differences in other exposures and in radiation measurement characteristics between the cohorts.

5. Acknowledgments.

The research was conducted by Oak Ridge Institute for Science and Education (ORISE) with the collaboration of the School of Public Health, University of North Carolina at Chapel Hill, and the Computer Science and Mathematics Division of ORNL under Contract DE-AC05-76OR00033 with the U. S. Department of Energy, Office of Energy Research, and Office of Environment, Safety and Health. Additional funding and oversight have been provided by the National Institute for Occupational Safety and Health. The authors thank the staff of ORISE's Center for Epidemiologic Research and Lockheed-Martin Energy Systems for their help in data collection, and Marjorie Harris, Richard L. Neubert and Stelmo Poteet for assistance in manuscript preparation. Part of the data used was collected under a previous contract from the U. S. Atomic Energy Commission and later from the U. S. Energy Research and Development Administration (E11-1-3428). The authors acknowledge the vital statistics offices of the individual states as the source of death record data and appreciate the offices' technical support of this research. The authors are solely responsible for the data analysis and the interpretation of the results.

6. Tables.

3. References

[1] M. Aitkin, D. Anderson, B. Francis, and J. Hinde. *Statistical Modelling in GLIM*. Oxford Statistical Science Series. Clarendon Press, Oxford, 1989.

[2] A. C. Atkinson. Likelihood ratios, posterior odds and information criteria. *J Econometrics*, 15:15–20, 1981.

[3] V. Beral, P. Fraser, L. Carpenter, M. Booth, A. Brown, and G. Rose. Mortality of employees of the Atomic Weapons Establishment, 1951–1982. *Br. Med. J.*, (297):757–770, 1988.

[4] V. Beral, H. Inskip, L. Fraser, M. Booth, D. Coleman, and G. Rose. Mortality of employees of the United Kingdom Atomic Energy Authority, 1946–1979. *Br. Med. J.*, (291):440–447, 1985.

[5] N. E. Breslow and N. E. Day. *Statistical Methods in Cancer Research, Volume II: The Design and Analysis of Cohort Studies*. Number Scientific Publication 82. International Agency for Research on Cancer, Lyon, 1987.

[6] E. Cardis, E.S. Gilbert, L. Carpenter, G. Howe, I. Kato, B.K. Armstrong, V. Beral, G. Cowper, A. Douglas, J. Fix, S.A. Fry, J. Kaldor, C. Lave, L. Salmon, P.G. Smith, G.L. Voelz, and L.D. Wiggs. Effects of low doses and low dose rates of external ionizing radiation: cancer mortality among nuclear industry workers in three countries. *Radiat. Res.*, 142:117–132, 1995.

[7] L. Carpenter, C. Higgins, A. Douglas, P. Fraser, V. Beral, and P. Smith. Combined analysis of mortality in three United Kingdom nuclear industry workforces, 1946-1988. *Radiat. Res.*, 138:224–238, 1994.

[8] CEDR. *Comprehensive Epidemiologic Data Resource*. U. S. Department of Energy, Office of Environment Safety and Health, Washington, D. C., May 1995. URL http://cedr.lbl.gov/CEDRhomepage.html.

[9] H. Checkoway, N. E. Pearce, and D. Crawford-Brown. *Research Methods in Occupational Epidemiology*. Oxford University Press, 1989.

[10] Chan-Fu Chen. Robustness aspects of score tests for generalized linear and partially linear regression models. *Technometrics*, 27(3):277–283, 1985.

[11] D. Clayton. The analysis of event history data: a review of progress and outstanding problems. *Stat. in Med.*, 7:819–841, 1988.

[12] D. Clayton and M. Hills. *Statistical Models in Epidemiology*. Oxford University Press, 1993.

[13] E. L. Crow. Freeman - Tukey test. In S. Kotz and N. L. Johnson, editors, *Encyclopedia of Statistical Sciences*, volume 3. John Wiley and Sons, 1983.

[14] B. Francis, M. Green, and C. Payne. Glim 4 the statistical system for generalized linear interactive modelling. Release 4, Royal Statistical Society, Oxford University Press, 1993.

[15] E. L. Frome. The analysis of rates using Poisson regression models. *Biometrics*, (39):665–674, 1983.

[16] E. L. Frome, D. L. Cragle, and R. W. McLain. Poisson regression analysis of the mortality amomg a cohort of World War II nuclear industry workers. *Radait. Res.*, (123):138–152, 1990.

[17] E. L. Frome and D. R. Hudson. A general statistical data structure for epidemiologic studies of DOE workers. In T. Truet, D. Margolis, and R. W. Mensing, editors, *Proceedings of the 1980 DOE Statistical Symposium*, Report Number CONF801045, Oak Ridge, TN, 1981. Oak Ridge National Laboratory.

[18] E. L. Frome, M. K. Kutner, and J.J. Beauchamp. Regression analysis of Poisson distributed data. *J. Am. Statist. Assoc.*, 68:935–940, 1972.

[19] E. L. Frome and M. D. Morris. Evaluating goodness of fit of Poisson regression models in cohort studies. *The Am Statist*, 43(3):144–147, 1989.

[20] E.S. Gilbert, D.L. Cragle, and L.D. Wiggs. Updated analyses of combined mortality data for workers at the Hanford Site, Oak Ridge National Laboratory, and Rocky Flats weapons plant. *Radiat. Res.*, 136:408–421, 1993.

[21] IARC. Study Group on Cancer Risk among Nuclear Industry Workers. Direct estimates of cancer mortality due to low doses of ionizing radiation: an international study. *Lancet*, 344:1039–1043, 1994.

[22] G. M. Kendall, C. R. Muirhead, B. H. MacGibbon, J. A. O'Hagan, A. J. Conquest, A. A. Goodill, B. K. Butland, T. P. Fell, D. A. Jackson, M. A. Webb, R. G. E. Haylock, J. M. Thomas, and T. J. Silk. Mortality and occupational exposure to radiation: First analysis of the National Registry for Radiation Workers. *Br. Med. J.*, 304:220–225, 1992.

[23] G. D. Kerr. Missing dose from mortality studies of radiation effects among workers at Oak Ridge National Laboratory. *Health Phys.*, 66:206–208, 1994.

[24] M. P. Little, G. M. Kendall, C.R. Muirhead, B.H. MacGibbon, R.G.E. Haylock, J.M. Thomas, and A. A. Goodill. Further analysis, incorporating assessment of the robustness of risks of cancer mortality in the national registry for radiation workers. *J. Radiat. Prot.*, 113:95–108, 1993.

[25] D. Loomis and S. Wolf. Mortality of workers at a nuclear materials production plant at Oak Ridge,Tennessee, 1947-1990. *Am. J. Indust. Med.*, 29:131–141, 1996.

[26] M. Macaluso. Exact stratification of person-years. *Epidemiology*, 3:441–448, 1992.

[27] P. McCullagh and J. A. Nelder. *Generalized Linear Models*. Chapman and Hall, 2 edition, 1989.

[28] R. W. McLain and E. L. Frome. A system for the analysis of cohort mortality data. In *SAS Users Group International Conference Proceedings*, SUGI II, pages 874–876, Cary, NC, 1986. SAS Institute.

[29] T. J. Mitchell, G. Ostrouchov, E.L. Frome, and G. D. Kerr. A method for estimating occupational dose to individuals, using weekly dosimetry data. *Radiat. Res.*, 147:195–207, 1997.

[30] R. R. Monson. Analysis of relative survival and proportional mortality. *Comp. Biomed. Res.*, (9):325–332, 1974.

[31] Committee on the Biological Effects of Ionizing Radiation (BEIR V) National Research Council. *Health Effects of Exposure to Low Levels of Ionizing Radiation*. National Academy of Sciences, Washington, D.C., 1990.

[32] G. Ostrouchov and E. Frome. A model search procedure for hierarchical models. *Computational Statistics and Data Analysis*, 15:285–296, 1993.

[33] D. A. Pierce, Y. Shimizu, D. L. Preston, M. Vaeth, and K. Mabuchi. Studies of the mortality of atomic bomb survivors. report 12, part i. cancer:1950-1990. *Radiat. Res.*, 146:1–27, 1996.

[34] D. L. Preston, J. H. Lubin, D. A. Pierce, and M.E. McConney. **epicure** Users's Guide. Technical report, Hirosoft International Corporation, Seattle, WA, 1993.

[35] C. R. Rao. *Linear Statistical Inference and Its Application*. John Wiley & Sons, 1975.

[36] L. Törnquvist, P. Vartia, and Y. O. Vartia. How should relative changes be measured? *Am. Statist.*, (39):43–46, 1985.

[37] J. P. Watkins, D.L. Cragle, E.L. Frome, C.M. West, D.J. Crawford-Brown, and W.G. Tankersley. Adjusting external doses from the ORNL and Y-12 facilities for the Oak Ridge nuclear facilities mortality study. Technical Report G-34, ORISE, 1994.

[38] J. P. Watkins, J. L. Reagan, D.L. Cragle, E.L. Frome, C.M. West, D. Crawford-Brown, and W.G. Tankersley. Collection, validation, and description of data for the Oak Ridge nuclear facilities mortality study. Technical Report J42, ORISE, 1993.

[39] J. P. Watkins, J. L. Reagan, D.L. Cragle, E.L. Frome, C.M. West, D. Crawford-Brown, and W.G. Tankersley. Collection, validation, and treatment of data for a mortality study of nuclear industry workers. *Appl. Occup. Environ. Hyg.*, (12):195–205, 1997.

[40] S. Wing, C. M. Shy, J. L. Wood, S. Wolf, D. L. Cragle, and E. L. Frome. Mortality among workers at Oak Ridge National Laboratory: Evidence of radiation effects in follow-up through 1984. *J. Am. Med. Assoc.*, (265):397–402, 1991.

[41] S. Wing, C. M. Shy, J. L. Wood, S. Wolf, D. L. Cragle, W. Tankersley, and E. L. Frome. Job factors, radiation, and cancer mortality at Oak Ridge National Laboratory: Follow-up through 1994. *Am. J. Indust. Med.*, (23):265–279, 1993.

[42] S. Wing, C. M. West, J. L. Wood, and W. Tankersley. Recording of external radiation exposures at Oak Ridge National Laboratory: Implications for epidemiologic studies. *J. Exposure Analysis and Environmental Epidemiology*, 4(1):83–93, 1994.