RADIATION RESEARCH 123, 138–152 (1990)

# Poisson Regression Analysis of the Mortality among a Cohort of World War II Nuclear Industry Workers

EDWARD L. FROME,* DONNA L. CRAGLE,† AND RICHARD W. MCLAIN†

*Mathematical Sciences Section, Oak Ridge National Laboratory, Oak Ridge, Tennessee, and †Center for Epidemiologic Studies, Oak Ridge Associated Universities, Oak Ridge, Tennessee

FROME, E. L., CRAGLE, D. L., AND MCLAIN, R. W. Poisson Regression Analysis of the Mortality among a Cohort of World War II Nuclear Industry Workers. *Radiat. Res.* 123, 138–152 (1990).

A historical cohort mortality study was conducted among 28,008 white male employees who had worked for at least 1 month in Oak Ridge, Tennessee, during World War II. The workers were employed at two plants that were producing enriched uranium and a research and development laboratory. Vital status was ascertained through 1980 for 98.1% of the cohort members and death certificates were obtained for 96.8% of the 11,671 decedents. A modified version of the traditional standardized mortality ratio (SMR) analysis was used to compare the cause-specific mortality experience of the World War II workers with the U.S. white male population. An SMR and a trend statistic were computed for each cause-of-death category for the 30-year interval from 1950 to 1980. The SMR for all causes was 1.11, and there was a significant upward trend of 0.74% per year. The excess mortality was primarily due to lung cancer and diseases of the respiratory system. Poisson regression methods were used to evaluate the influence of duration of employment, facility of employment, socioeconomic status, birth year, period of follow-up, and radiation exposure on cause-specific mortality. Maximum likelihood estimates of the parameters in a main-effects model were obtained to describe the joint effects of these six factors on cause-specific mortality of the World War II workers. We show that these multivariate regression techniques provide a useful extension of conventional SMR analysis and illustrate their effective use in a large occupational cohort study. © 1990 Academic Press, Inc.

## INTRODUCTION

This report is the first phase of a comprehensive study of the mortality of all workers employed at federal nuclear plants in Oak Ridge, Tennessee, and we limit our analysis to individuals who were employed only during the World War II era (1943–1947). During this period, radiation monitoring programs were in developmental stages; only persons considered at risk or representative samples of those at risk were monitored. Thus it is possible to designate workers only as "probably exposed" or "probably nonexposed" to ionizing radiation. Additional factors available for analyses are socioeconomic status, place of employment, length of employment, birth year, and period of follow-up. An important question considered is the comparison of age-cause-specific death rates for different birth cohorts. For this phase of the analyses, birth year and age at risk of exposure are synonymous since only short-term workers employed during the World War II era are included. Workers were assigned a relative radiation exposure index based on job and department codes, and a variable related to socioeconomic status based on the likelihood that their jobs were professional, skilled, or unskilled. The place of employment is an important exposure-related variable since the type of radiation exposure was different at each facility.

Subsequent phases of the study will include all workers who were employed in later years and were monitored for radiation exposure. Studies of two of the Oak Ridge plants have been published separately (*1–3*), and subcohort analyses have been published for special interest groups (*4–6*). Three studies (*3, 7, 8*) have included individual radiation dose estimates, and each group was studied separately primarily because dosimetry efforts were at different levels of completion. As additional information becomes available, we will pool the data from all three facilities to obtain a larger study population. This large population of workers with personal exposure data will be used to evaluate the health effects of protracted exposure to low levels of ionizing radiation and to conduct methodologic research in occupational epidemiology.

Studies of occupational cohorts frequently span several decades of plant operation. Wen *et al.* (*9*) have pointed out that selecting subgroups by different dates of hire may serve to isolate potentially interesting groups for study and also to attempt to control for time-related exposure risk. In this study, isolation of subcohorts is designed to group workers with common periods of work and common methods of exposure assessment. Each phase of the study will address a different methodologic problem in occupational epidemiology. In this first phase, we will demonstrate that Poisson regression methods can be used effectively to describe the effect of multiple risk factors on cause-specific mortality.

0033-7587/90 $3.00
Copyright © 1990 by Academic Press, Inc.
All rights of reproduction in any form reserved.

## MATERIALS AND METHODS

### *Description of Facilities*

In 1938 U.S. physicists confirmed an earlier report in Germany by Hahn and Strassman: a new type of nuclear reaction—nuclear fission—had taken place. By the end of 1939 the potential for a chain reaction as the result of the release of neutrons during nuclear fission was established. In 1940 the rarer isotope $^{235}U$ was identified as the one that fissioned with slow neutrons. It was determined that for a chain reaction with explosive potential to be possible, uranium containing higher concentrations of $^{235}U$ than found in nature would be required. This meant that an isotope separation process was needed. It was also clear that if a chain reaction were possible, the result would be a new power source and possibly an explosive with tremendous destructive force. The potential of powerful explosives in the hands of Nazi Germany had already established a need for federal support in the uranium research program. This led to the creation of a new relationship between science and government in the interest of national defense. By the end of 1941 U.S. scientists had confirmed the conclusions of scientists in England that $^{235}U$ could be separated and made into an effective bomb. An extensive historical account of the scientific and political events that surrounded the development of atomic fission has been given by Hewlett and Anderson (*10*).

On December 18, 1941, 11 days after Pearl Harbor, the decision was made to focus atomic energy research on the development of a weapon for the war. To pursue this goal, it was necessary to develop full-scale production facilities to produce enriched uranium with absolute secrecy and as quickly as possible. The Army was assigned the responsibility for this design and construction project of unprecedented size. The headquarters for the project were 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 pile that would be used to produce plutonium for further research in the war effort.

The major focus of the war effort in the newly created community of Oak Ridge was producing enriched uranium in large quantities. At the end of 1942 there were four feasible approaches to separating $^{235}U$—electromagnetic, centrifuge, gaseous diffusion, and thermal diffusion. The centrifuge method was the weakest and was eliminated from further consideration. The diffusion method appeared to have the best overall chance of success but would require completion of the entire plant before useful amounts of material could be produced. In December of 1942, President Roosevelt authorized construction of a full-scale gaseous-diffusion plant, a smaller electromagnetic separation plant, and a full-scale experimental pile for plutonium production. It was considered essential to build these facilities with utmost speed and without the benefit of an intermediate step that would normally occur in the transition from laboratory research to full-scale production.

An air-cooled experimental pile, a chemical separations plant, and supporting laboratories were constructed by the DuPont Co. at the X-10 site in Oak Ridge. This facility was officially named Clinton Laboratories (later renamed Oak Ridge National Laboratory). Most of the research staff came from the Metallurgical Laboratory at the University of Chicago, and industrial staff and managerial staff were transferred to Oak Ridge by DuPont. The implications of a chain-reacting pile as a radiation source caused Compton, director of the Metallurgical Laboratory, to establish a Health Division. The importance of radiation as an industrial health hazard was clearly recognized, and personnel working in high radiation areas were issued pocket ionization chambers and later film badges to monitor external whole-body exposure. The maximum permissible radiation exposure levels were those recommended by the Advisory Committee on X-ray and Radium Protection in the 1930s (i.e., 0.1 R per day for X rays and 0.01 R equivalent physical per day for fast neutrons). A detailed account of the Health Division practices is given by Hacker (*11*).

The primary purpose of the natural-uranium and graphite pile at the Clinton Laboratory was to produce plutonium for further research in the war effort. This goal was reached by the summer of 1944 and the laboratory quickly transformed itself into the first well-rounded institution for nuclear research. The X-10 pile provided an abundant supply of neutrons for research in physics, produced radioactive isotopes with potential for use in science and industry, and offered unprecedented opportunities for research into the biological effects of radiation.

The major portion of the war effort in Oak Ridge was devoted to obtaining enriched uranium. The Y-12 Plant, operated by the Tennessee Eastman Corp. (TEC) from June 1943 to May 1947, was engaged in the enrichment of uranium by the electromagnetic separation process. This approach to separating uranium isotopes required a mass-spectrograph unit called a "Calutron" (named after the University of California where it was developed). The workers at Y-12/TEC handled uranium compounds which had a low radium content and were weak $\gamma$-ray sources. The most important potential hazard at TEC was the inhalation of uranium compounds. The uranium dust levels were highest in the "alpha" stage of the process, which was discontinued in September 1945. The radiation hazard continued, however, due to the higher $^{234}U$ and $^{235}U$ content of the product (the feed material for the second stage was enriched uranium chloride obtained from the first stage). The type and solubility of the uranium compounds also changed over time. After late 1945, uranium hexafluoride ($UF_6$) was received directly from the K-25 site in Oak Ridge (later named the Oak Ridge Gaseous Diffusion Plant).

The main purpose of the facility at the K-25 site was to produce enriched uranium via the gaseous diffusion process. In this process, $UF_6$ gas containing 0.7% $^{235}U$ flows through a barrier which separates the lighter $^{235}U$ from the heavier $^{238}U$. The enrichment of the final $^{235}U$ was low initially (less than 5% $^{235}U$) but increased to over 90% by 1947. In support of the enrichment process, the facility also operated a plant that produced barrier material, ran a feed materials operation, and conducted laboratory research. Occupational exposures at the plant included uranium as soluble $UF_6$ and relatively insoluble uranium oxides, metallic nickel, hydrofluoric acid, and a variety of chemicals used in laboratory quantities, especially fluorocarbons.

### *Definition of Cohort and Study Factors*

The total number of workers ever employed in the Oak Ridge facilities between 1943 and 1985 was 118,588. This total includes all race/sex groups. The cohort designated as World War II workers was defined to be all white males who were employed at least 30 days between the start of facility operations and December 31, 1947, and at no other time after that date. The date of December 31, 1947, was chosen because it differentiated between the large population of workers who came to work during the war and left shortly after that large-scale production effort was completed, and those who were more long-term, stable workers. Also, before 1948 there were very few radiation monitoring data available. The total white male worker population who met the length of employment criterion for the three facilities numbered 34,726 of whom 28,008 had sufficient data to be included in this study. Over 80% ($N = 5491$) of the workers excluded from the analyses were lost to follow-up before the beginning of the study period, which was January 1, 1950. Additional reasons for omission from analyses included missing values for socioeconomic status, exposure code, and date of birth. Because of the criteria for data analyses, the workers who were lost to follow-up would not have contributed any person-years to the study. Missing socioeconomic status or exposure variables were most likely a result of missing job title, which was the variable used to assign values to these analytic variables. Since the majority of the exclusions involved workers who had no potential for contribution of person-years, it was felt that these exclusions would not bias the study results.

*Radiation.* Assignment of indices for radiation exposure for specific job title/department combinations was accomplished using a two-tiered scheme. Each job/department combination was assigned a code of "Y" or

"N." An N assignment was given to job title/department combinations with no rational expectation of continuing or even irregular contact with radioactive materials and sources (e.g. financial accountants), and where there was no reasonable expectation that radioactive materials were used, transported, or stored. A 'Y' assignment was given to job title/department combinations where workers had continuing contact with radioactive materials (e.g., recovery operator), infrequent but regular contact (e.g., inspector), or irregular contact, either frequent or infrequent (e.g., maintenance personnel), or job titles requiring presence in hazardous areas for regular but short periods (e.g., superintendent). The assignments were based on information obtained from several sources including plant records, procedure manuals, and job descriptions. The total days worked in a job title/department combination that had a 'Y' assignment were computed using work history records for each individual. For the purpose of analysis, a person was considered "exposed" if the total days on a job designated 'Y' were at least 30. Note that since working at least 30 days was the minimum length of employment criterion, all workers can potentially be classified as exposed. Further subclassification of this variable is problematic since the intensity of exposure within job title/department combinations is not known.

*Facility.* Each worker in the study was categorized by the facility in which he worked. If a worker was an employee of more than one facility during the study period, he was placed in a fourth category which was designated as "multiple." Therefore, all cohort members fell into one of the following categories: TEC, K-25, X-10, or multiple.

*Socioeconomic status.* A listing of all job titles for all cohort members was generated. This listing was reviewed and assignments of professional, skilled, or unskilled were made for each unique job title. A job that required a high school education or less was assigned to the unskilled category (e.g., chauffeur, laborer, helper). A skilled job was one in which at least a high school education would be required or in some cases vocational training (e.g., carpenter, mechanic, welder). Professional job titles were those that would require college degrees (e.g., biologist, chemist, physicist). When the assignments were completed, each cohort member had a socioeconomic status category for each job he had during all years of employment. The socioeconomic status of the first job of each cohort member was chosen for the analysis because it was judged to be a good indicator of level of education.

*Length of employment.* Length of employment was calculated as the length of time between first hire and last termination if employment was continuous. If there were breaks in employment, the period of time that the worker was not employed at one of the three study facilities was not counted. For the purpose of analysis, length of employment was categorized in two levels: <6 months or 6 months or longer (6+).

*Period of follow-up.* Each person entered the study on January 1, 1950, and was followed until they were lost to follow-up, died, or reached the end of the study on December 31, 1979. Observed and expected deaths were computed for three decades of follow-up: the 1950s, the 1960s, and the 1970s.

*Birth year.* The workers in the cohort were categorized into two groups according to year of birth: before 1910 and 1910 or later. This division was chosen in order to separate the cohort by draft eligibility based on age when hired. It was theorized that stratification might help when interpreting the results if the group which was draft age when hired had a significant proportion of workers who were medically unfit for combat.

### *Vital Status*

Vital status of the cohort was ascertained primarily through the Social Security Administration. Additional sources of information included Civil Service active and retired rolls, states' driver's license bureaus, and the Health Care Finance Agency. Death certificates were retrieved for the cohort and were coded to the eighth revision of the International Classification of Diseases adapted for use in the United States. All cohort members who were of unknown vital status were allowed to accumulate person-years in the study only until the date that they became unknown. Since follow-up for the whole cohort commenced on January 1, 1950, persons who became unknown or who died prior to this date were not included.

### *Analysis Methods*

Analyses of the mortality experience of the World War II cohort involved two approaches. The first was a modified traditional standardized mortality ratio (SMR) analysis using U.S. white males as the external control group. The modification consisted of an evaluation of the relative change in the SMR over six 5-year intervals. A trend statistic and an adjusted SMR were computed for each selected cause of death of interest. The trend statistic describes the rate of change (percentage per year) in the SMR over the 30-year interval from 1950 to 1979. The purpose of this analysis was to evaluate whether or not cause-specific mortality for the 57 cause-of-death categories used by Monson (*12*) was excessive or showed a significant rate of change over the period of follow-up.

The second approach was undertaken to evaluate the joint influence of several variables on the mortality risk for the cohort. Recently developed multivariate statistical methods were used to describe the effect of six factors on the cause-specific mortality—see Breslow *et al.* (*13*), Breslow (*14, 15*), and Whittemore (*16*). Observed and expected deaths were tabulated for each of the subgroups that are defined by the study factors of interest using a computer program that we have developed (*17*). Poisson regression analysis (*18, 19*) was then used to describe the joint effects (using a multiplicative main effects model) of the study factors on the cause-specific mortality. The expected deaths were computed using the cause-age-calendar year-specific rates for U.S. white males (*12*). Person-years for each subgroup were accumulated across 5-year age-specific and calendar-year-specific categories from the beginning of follow-up (January 1, 1950) until the date of death or December 31, 1979, whichever occurred first. Subjects whose vital status could not be ascertained were withdrawn on the date of their last known contact.

The goal of the multivariate analysis was to evaluate the combined effect of duration of employment, SES, radiation exposure, facility, birth year, and period of follow-up on cause-specific mortality of interest (all causes, diseases of the circulatory system, all malignant neoplasms, lung cancer, other cancer, and external causes).

The strategy that we have followed in the multivariate analysis is as follows (see the Appendix for a more detailed discussion):

(i) Generate a six-dimensional table of observed and expected deaths for the cause-of-death category of interest.

(ii) Fit the main-effects model and record the parameter estimates and their standard deviations.

(iii) Use the screening procedure described under Appendix to obtain the abridged table of $\chi^2$ statistics.

(iv) Review and interpret the results from steps ii and iii.

## RESULTS

### *Vital Status*

The vital status statistics are shown by facility in Table I. Approximately 42% of the World War II workers have died and death certificate retrieval is 96.8% complete. The percentage of workers with unknown vital status is approximately 2% for all facilities. The percentage of deaths varies by facility with a high of 44% for the K-25 population to a low of 30% for the X-10 population. The majority of the workers and deaths in the study come from the K-25 and TEC populations.

TABLE I
Vital Status Ascertainment and Death Certificate Retrieval Ratios by Facility—1950 to 1980

| Status | X-10 | TEC | K-25 | Multiple | Combined |
|---|---|---|---|---|---|
| Alive | 784 (67.7)[a] | 8,651 (57.0) | 5271 (54.4) | 1095 (55.1) | 15,801 (56.4) |
| Dead | 353 (30.5) | 6,202 (40.9) | 4278 (44.1) | 838 (42.2) | 11,671 (41.7) |
| With certificate | 342 (96.9)[b] | 6,007 (96.9) | 4139 (96.8) | 812 (96.9) | 11,300 (96.8) |
| Unknown | 21 (1.8) | 314 (2.1) | 148 (1.5) | 53 (2.7) | 536 (1.9) |
| Total | 1158 | 15,167 | 9697 | 1986 | 28,008 |

[a] Percentage of total.
[b] Percentage of deaths.

*Population Distribution by Study Factors*

The univariate distribution of the worker population by factor levels is shown in Table II. The length of employment distribution indicates that there was a fairly high rate of employee turnover. The K-25 plant experienced the greatest percentage of short-term (<6 months) employees (52.6% of the worker population) followed by X-10 (40.2), TEC (36.1), and multiple (33.2). For multiple facility workers, the length of employment was based on the total time of employment at all facilities.

The distribution of workers by SES reveals less than 6% professional workers at TEC, K-25, and multiple, while the X-10 worker population has 31.3% in the professional category. The highest percentage of skilled workers is found in the TEC population (79.8%) followed by the multiple workers (65.0%), the K-25 workers (59.9%), and finally the X-10 workers (44.8%). Unskilled workers comprised slightly more than 30% of the K-25 and multiple workers, nearly 24% of the X-10 workers, and approximately 14% of the TEC workers. The distribution of exposure shows that each facility has a different percentage of the population rated in the "yes" category. The population with the highest percentage of exposed workers is TEC (75.0%) and K-25 has the lowest percentage of exposed workers (41.5%). The birth year distribution is similar for all populations; between 73 and 80% of each population were born after 1910. It can be seen that there are substantial differences in the distribution of the study factors (with the possible exception of birth year) between the four population groups included in the analyses. Each of the study factors is a potential predictor of disease outcome either alone or in conjunction with the other variables in the data set. Table III shows the joint distribution of World War II workers cross-classified according to the five study factors of interest.

TABLE II
Marginal Distribution of Study Factors by Facility

| Factor | X-10 | TEC | K-25 | Multiple |
|---|---|---|---|---|
| Duration | | | | |
| <6 months | 465 (40.2)[a] | 5,468 (36.1) | 5099 (52.6) | 660 (33.2) |
| 6+ months | 693 (59.8) | 9,699 (63.9) | 4598 (47.4) | 1326 (66.8) |
| Socioeconomic status | | | | |
| Professional | 362 (31.3) | 894 (5.9) | 550 (5.7) | 64 (3.2) |
| Skilled | 519 (44.8) | 12,108 (79.8) | 5807 (59.9) | 1291 (65.0) |
| Unskilled | 277 (23.9) | 2,165 (14.3) | 3340 (34.4) | 631 (31.8) |
| Radiation | | | | |
| No | 352 (30.4) | 3,792 (25.0) | 5670 (58.5) | 617 (31.1) |
| Yes | 806 (69.6) | 11,375 (75.0) | 4027 (41.5) | 1369 (68.9) |
| Birth year | | | | |
| <1910 | 224 (19.3) | 3,817 (25.2) | 2577 (26.6) | 429 (21.6) |
| 1910+ | 934 (80.7) | 11,350 (74.8) | 7120 (73.4) | 1557 (78.4) |

[a] Percentage.

*SMR Trend Analysis*

Table IV presents the results of a traditional SMR analysis in the first four columns, where the observed number of deaths, the expected number of deaths, and the SMR are displayed for each cause of death category. Because of the multiple comparisons being made, a significance level of 0.01 was used to identify statistically significant effects in Tables IV and V. The healthy worker effect is not observed in this population, and in fact, there is a statistically significant 10% increase in all causes of death combined. Other causes of death showing statistically significant increases include tuberculosis (37%), mental, psychoneurotic and personality disorders (60%), cerebrovascular disease (10%), diseases of the respiratory system (24%), emphysema (24%), symptoms, senility, and ill-defined conditions (204%), all external causes (32%), all accidents (27%), and motor vehicle accidents (43%). A statistically significant deficit is found only in the diseases of the blood and blood-forming organs category of deaths (61% deficit).

Examination of the data by 5-year intervals of death over the 30-year period of the study reveals a statistically significant trend of increasing SMRs for all causes of death. This trend can be seen by comparing the numbers in Columns 5 through 9 of Table IV and is shown graphically in Fig. 1. Column 5 presents the unadjusted SMR in log percentage units ($L\%$) and is a transformation of the number in Column 4 (i.e., it is 100 times the natural logarithm of the

TABLE III
Number of Oak Ridge World War II Workers by Factor Levels ($N$ = 28,008)

| Socioeconomic Status: | Skilled | | | | Unskilled | | | | Professional | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Duration: | <6 mo | | 6+ mo | | <6 mo | | 6+ mo | | <6 mo | | 6+ mo | |
| Radiation: | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Facility: X-10 | | | | | | | | | | | | |
| Birth year | | | | | | | | | | | | |
| <1910 | 19 | 63 | 42 | 74 | 13 | 22 | 38 | 17 | 4 | 6 | 13 | 39 |
| 1910+ | 44 | 113 | 57 | 107 | 52 | 55 | 49 | 31 | 8 | 66 | 13 | 213 |
| Facility: TEC | | | | | | | | | | | | |
| Birth year | | | | | | | | | | | | |
| <1910 | 436 | 1058 | 636 | 2669 | 351 | 98 | 304 | 286 | 25 | 23 | 87 | 163 |
| 1910+ | 629 | 2116 | 632 | 3932 | 408 | 248 | 193 | 277 | 24 | 52 | 67 | 453 |
| Facility: K-25 | | | | | | | | | | | | |
| Birth year | | | | | | | | | | | | |
| <1910 | 592 | 411 | 579 | 672 | 738 | 119 | 493 | 139 | 31 | 20 | 57 | 35 |
| 1910+ | 966 | 819 | 689 | 1079 | 992 | 289 | 405 | 165 | 52 | 70 | 76 | 209 |
| Facility: Multiple | | | | | | | | | | | | |
| Birth year | | | | | | | | | | | | |
| <1910 | 58 | 47 | 73 | 257 | 47 | 25 | 66 | 76 | 0 | 1 | 4 | 11 |
| 1910+ | 104 | 174 | 87 | 491 | 110 | 91 | 59 | 157 | 2 | 1 | 7 | 38 |

SMR). Column 6 presents the adjusted log SMR for the midpoint of the study (which is 1965), and Column 7 is the estimated standard deviation. The adjusted log SMR is calculated by fitting the model $y_j = \mu_j^0 \exp(\alpha_a + \theta t_j)$ to the observed and expected deaths for each 5-year interval from 1950 to 1980, where $t_j$ represents the midpoint of the $j$th 5-year interval and $\alpha_a$ represents the logarithm of the adjusted SMR (see Appendix). Column 8 presents the trend statistic, which describes the annual rate of change of the SMR over the 30-year period, and the estimated standard deviation of the trend statistic is given in Column 9. Therefore, it is easily seen, for example, that the 0.74% per year increase in the SMR for all causes of death is statistically significant, since it is six times larger than its estimated standard deviation.

Of particular interest in this study are the deaths from all malignant neoplasms. The unadjusted "crude" SMR indicates that there is a 4.6% nonstatistically significant increase in deaths from malignant neoplasms in this worker population. There is, however, a statistically significant trend over the period of the study that shows a 1.05% per year increase in the SMR (see Fig. 1). The results for circulatory system deaths appear to be similar to those for all malignant neoplasms. Again, the unadjusted SMR is not statistically significant, but there is a significant increasing trend in the SMRs over time at a rate of 1.05% per year (see Fig. 1). The fact that the trend is identical for all cancer deaths indicates that these two lines are parallel (see Fig. 1). However, at the midpoint of the interval of follow-up, the SMR for circulatory disease was still less than one, whereas the SMR for all cancer was essentially unity. These trends would not have been noted if the unadjusted SMR had been the only statistic computed.

Occupational epidemiologists generally examine only elevated SMRs with regard to time trends. However, using this system of analysis, the statistically significant trend for all malignant neoplasms would probably have been missed. Similarly, the significant trend for diseases of the circulatory system and arteriosclerotic heart disease also would have been missed. Arteriosclerotic heart disease, a subset of diseases of the circulatory system, represents slightly more than 70% of the deaths in that category. So similar SMRs and trends for these two groupings should be expected.

The SMR for diseases of the respiratory system is significantly elevated and exhibits a significant trend of 1.53% increase per year. It can be seen from Fig. 1 that the SMR for respiratory diseases was greater than one during all but the first 5-year interval of follow-up. Suicide also shows a significant trend over time with a 2.45% increase per year. However, the overall SMR for suicide is not significantly elevated.

Table V displays the summary statistics for mortality from selected cancer causes of death. Significant deficits are noted for cancers of the digestive organs, and peritoneum, where it can be seen that the adjusted log SMR is −23.6 and there is no significant trend with time. This indicates that deaths classified in this category of disease are below the national level throughout the period of the study and, if the SMRs were plotted (as in Fig. 1), they would be less than one and would show a slight upward trend over the 30-year

**TABLE IV**
**Summary Statistics for Mortality from Selected Causes of Death among White Males Who Worked in Oak Ridge between 1943 and 1947 ($N$ = 28,008)**

| Cause of death (ICDA-8) | Observed number | Expected number[a] | SMR | SMR ($L$%) $\alpha_u$[b] | SMR ($L$%) $\alpha_a$[c] | SMR ($L$%) SD | Trend $\theta$[c] | Trend SD | Deviance |
|---|---|---|---|---|---|---|---|---|---|
| All causes (001–999) | 11,671 | 10,537.60 | 1.11 | 10.2† | 8.0 | 1.0 | 0.74* | 0.12 | 12.99 |
| All infective and parasitic (000–139) | 150 | 136.36 | 1.10 | 9.5 | 12.4 | 8.4 | 1.23 | 0.93 | 6.53 |
| Tuberculosis (010–019) | 108 | 78.73 | 1.37 | 31.6† | 44.9 | 10.2 | 3.44* | 1.19 | 3.77 |
| All malignant neoplasms (140–209) | 2,207 | 2,108.29 | 1.05 | 4.6 | 0.4 | 2.5 | 1.05* | 0.28 | 0.79 |
| Benign neoplasms (210–239) | 32 | 27.61 | 1.16 | 14.8 | −5.6 | 21.6 | 6.05* | 2.33 | 5.17 |
| Allergic, endocrine, etc. (240–279) | 152 | 176.07 | 0.86 | −14.7 | −17.4 | 9.0 | 0.77 | 1.06 | 1.78 |
| Diabetes mellitus (250) | 136 | 149.33 | 0.91 | −9.4 | −12.6 | 9.6 | 0.95 | 1.13 | 0.88 |
| Disease of blood, bloodforming organs (280–289) | 9 | 23.12 | 0.39 | −94.3† | −92.4 | 34.3 | −0.87 | 4.08 | 3.58 |
| Mental, psychoneurotic, personality disorders (290–317) | 81 | 50.52 | 1.60 | 47.2† | 48.3 | 11.4 | −0.50 | 1.30 | 6.27 |
| Nervous system (320–389) | 76 | 81.76 | 0.93 | −7.3 | −10.5 | 12.3 | 1.21 | 1.40 | 3.08 |
| Circulatory system (390–458) | 5,750 | 5,702.19 | 1.01 | 0.8 | −2.6 | 1.5 | 1.05* | 0.17 | 24.02 |
| Chronic rheumatic (393–398) | 108 | 110.70 | 0.98 | −2.5 | −3.2 | 9.8 | −0.45 | 1.18 | 2.72 |
| Arteriosclerotic (410–413) | 4,090 | 4,002.49 | 1.02 | 2.2 | −1.6 | 1.8 | 1.07* | 0.21 | 11.12 |
| Cerebrovascular disease (430–438) | 833 | 753.54 | 1.11 | 10.0† | 6.7 | 3.9 | 0.99 | 0.46 | 10.94 |
| Respiratory system (460–519) | 792 | 634.11 | 1.25 | 22.2† | 15.0 | 4.4 | 1.53* | 0.50 | 5.89 |
| Pneumonia (480–486) | 250 | 222.38 | 1.12 | 11.7 | 9.4 | 7.1 | 0.61 | 0.84 | 5.77 |
| Emphysema (492) | 209 | 168.06 | 1.24 | 21.8† | 25.8 | 8.4 | −0.89 | 1.10 | 6.84 |
| Asthma (493) | 26 | 25.95 | 1.00 | 0.2 | 4.7 | 20.0 | 2.10 | 2.54 | 4.92 |
| Digestive system (520–577) | 470 | 506.15 | 0.93 | −7.4 | −10.2 | 4.9 | 1.23 | 0.59 | 11.17 |
| Gastric and duodenal ulcer (531–533) | 88 | 81.58 | 1.08 | 7.6 | 7.7 | 10.7 | 0.42 | 1.35 | 4.56 |
| Cirrhosis of liver (571) | 218 | 251.01 | 0.87 | −14.1 | −20.5 | 7.5 | 2.13 | 0.90 | 8.68 |
| Diseases of genitourinary system (580–629) | 148 | 145.30 | 1.02 | 1.8 | 2.1 | 8.3 | −0.27 | 0.97 | 1.72 |
| Chronic nephritis (582) | 52 | 52.65 | 0.99 | −1.2 | 1.0 | 14.1 | 1.11 | 1.71 | 4.41 |
| Diseases of skin, cellular tissue (680–709) | 6 | 7.44 | 0.81 | −21.5 | −26.0 | 44.3 | 1.59 | 5.20 | 3.74 |
| Diseases of bone, organs of movement (710–738) | 25 | 17.52 | 1.43 | 35.6 | 37.1 | 21.1 | −0.54 | 2.50 | 1.45 |
| Symptoms, senility, ill-defined conditions (790–799) | 338 | 110.92 | 3.05 | 111.4† | 115.7 | 5.6 | −1.59 | 0.68 | 11.36 |
| All external causes (800–998) | 1,051 | 792.29 | 1.33 | 28.3† | 28.5 | 3.1 | 0.39 | 0.37 | 5.44 |
| All accidents (800–949) | 694 | 542.45 | 1.28 | 24.6† | 24.8 | 3.8 | 0.11 | 0.46 | 7.98 |
| Motor vehicle (810–949) | 339 | 235.60 | 1.44 | 36.4† | 37.6 | 5.5 | 0.78 | 0.66 | 10.50 |
| Suicide (950–959) | 221 | 198.21 | 1.12 | 10.9 | 9.7 | 6.8 | 2.45* | 0.83 | 0.95 |

[a] Number of expected deaths is based on U.S. white male mortality rates (*12*).

[b] The unadjusted SMR in log percentage ($L$%) units is $\alpha_u = 100 \log_e(\text{observed/expected})$.

[c] The adjusted log SMR, $\alpha_a$, and the trend statistics $\theta$ are obtained by fitting a model $y_t = \mu_t^o \exp(\alpha_a + \theta t)$ to the observed ($y_t$) and the expected ($\mu_t^o$) deaths for each 5-year interval from 1950 to 1980 (see the Appendix). The trend statistic describes the rate of change of the SMR over this 30-year period. For example, the SMR for diseases of the respiratory system increased at a rate of 1.53% per year. The parameter $\alpha_a$ is used to estimate the SMR at the middle of the interval (1965), e.g., for the respiratory system exp(0.15) = 1.16. Each of these statistics can be compared with the null value of zero by dividing by the estimated standard deviation (SD). The deviance is approximately distributed as a $\chi^2$ statistic with 4 *df* and provides a measure of lack-of-fit of the trend equation.

† SMR significant at 0.01 level.

* Trend significant at 0.01 level.

period of follow-up. A similar pattern is observed for cancers of the large intestine and rectum, which are subsets of the larger category of cancers of the digestive organs and peritoneum. Note, however, that the trend in the SMRs for cancer of the rectum is downward.

Cancer of the respiratory system and, more specifically, cancer of the lung are statistically significantly elevated. These categories represent logical end points to be examined in relation to exposure to uranium dust. The log SMR for cancer of the lung is 16.7, which indicates that the worker population has lung cancer death rates higher than those of the national average at the midpoint of the study.

TABLE V
Summary Statistics for Mortality from Selected Cancer Causes of Death among White Males Who Worked in Oak Ridge between 1943 and 1947 ($N$ = 28,008)

| Cause of death (ICDA-8) | Observed number | Expected number[a] | SMR | SMR ($L$%) $\alpha_u$[b] | SMR ($L$%) $\alpha_a$[c] | SMR ($L$%) SD | Trend $\theta$[c] | Trend SD | Deviance |
|---|---|---|---|---|---|---|---|---|---|
| All cancer | 2207 | 2108.29 | 1.05 | 4.6 | 0.4 | 2.5 | 1.05* | 0.28 | 0.79 |
| Buccal cavity and pharynx (140–149) | 75 | 69.01 | 1.09 | 8.3 | 4.0 | 12.9 | 1.27 | 1.53 | 3.50 |
| Digestive organs, peritoneum (150–159) | 490 | 611.73 | 0.80 | −22.2† | −23.6 | 4.9 | 0.46 | 0.57 | 7.78 |
| Eosphagus (150) | 49 | 51.00 | 0.96 | −4.0 | −2.3 | 15.5 | −0.49 | 1.84 | 3.84 |
| Stomach (151) | 93 | 119.92 | 0.78 | −25.4 | −25.1 | 10.4 | −0.27 | 1.26 | 1.31 |
| Large intestine (153) | 149 | 191.78 | 0.78 | −25.2† | −26.5 | 9.3 | 0.31 | 1.07 | 4.38 |
| Rectum (154) | 38 | 71.16 | 0.53 | −62.7† | −60.9 | 16.4 | −1.11 | 2.01 | 5.04 |
| Liver (155–156) | 34 | 45.27 | 0.75 | −28.6 | −40.9 | 19.6 | 4.15 | 2.23 | 3.41 |
| Pancreas (157) | 110 | 118.40 | 0.93 | −7.4 | −8.7 | 10.7 | 0.36 | 1.26 | 15.44 |
| Respiratory system (160–163) | 886 | 708.84 | 1.25 | 22.3† | 15.5 | 4.2 | 1.36* | 0.48 | 1.58 |
| Larynx (161) | 33 | 33.03 | 1.00 | −0.1 | 5.9 | 17.9 | −2.31 | 2.19 | 6.43 |
| Lung (162–163) | 850 | 667.99 | 1.27 | 24.1† | 16.7 | 4.4 | 1.44* | 0.49 | 0.94 |
| Bone (170) | 11 | 10.35 | 1.06 | 6.1 | −0.3 | 32.1 | −4.95 | 3.76 | 7.75 |
| Skin (172–173) | 34 | 34.26 | 0.99 | −0.8 | −15.9 | 20.7 | 3.93 | 2.33 | 10.72 |
| Prostate (185) | 150 | 141.96 | 1.06 | 5.5 | 3.2 | 10.4 | 0.42 | 1.16 | 2.00 |
| Testes (186–187) | 7 | 9.61 | 0.73 | −31.7 | −40.4 | 42.1 | 8.43 | 4.76 | 4.58 |
| Bladder (188) | 54 | 66.22 | 0.82 | −20.4 | −25.2 | 15.9 | 1.14 | 1.82 | 2.93 |
| Kidney (189) | 44 | 52.63 | 0.84 | −17.9 | −27.6 | 18.0 | 2.35 | 2.05 | 3.24 |
| Eye (190) | 4 | 2.07 | 1.93 | 65.9 | 66.9 | 50.6 | −4.55 | 6.26 | 3.62 |
| Brain, CNS (191–192) | 69 | 59.57 | 1.16 | 14.7 | 16.9 | 12.1 | −2.10 | 1.45 | 3.43 |
| Thyroid (193) | 2 | 4.50 | 0.44 | −81.2 | −136.8 | 113.3 | 11.66 | 11.81 | 2.58 |
| All lymphatic (200–209) | 195 | 200.02 | 0.98 | −2.5 | −7.2 | 7.9 | 1.47 | 0.92 | 1.44 |
| Lymphosarcoma, reticulosarcoma (200) | 39 | 45.80 | 0.85 | −16.1 | −20.7 | 17.1 | 2.02 | 2.13 | 4.07 |
| Hodgkin's disease (201) | 18 | 23.00 | 0.78 | −24.5 | −23.8 | 23.6 | 1.98 | 2.93 | 3.65 |
| Leukemia, aleukemia (204–207) | 92 | 81.17 | 1.13 | 12.5 | 8.3 | 11.4 | 1.36 | 1.33 | 2.93 |
| Other lymphatic tissue (202–203, 208) | 40 | 48.23 | 0.83 | −18.7 | −22.0 | 19.8 | 0.62 | 2.17 | 8.53 |

[a] Number of expected deaths is based on U.S. white male mortality rates (*12*).
[b] The unadjusted SMR in log percentage ($L$%) units is $\alpha_u$ = 100 $\log_e$(observed/expected).
[c] The adjusted log SMR, $\alpha_a$, and the trend statistics $\theta$ are obtained by fitting a model $y_t = \mu_t^o \exp(\alpha_a + \theta t)$ to the observed ($y_t$) and the expected ($\mu_t^o$) deaths for each 5-year interval from 1950 to 1980 (see the Appendix). The trend statistic describes the rate of change of the SMR over this 30-year period. For example, the lung cancer SMR increased at a rate of 1.44% per year. The parameter $\alpha_a$ is used to estimate the SMR at the middle of the interval (1965), e.g., for lung cancer exp(0.167) = 1.18. Each of these statistics can be compared with the null value of zero by dividing by the estimated standard deviation (SD). The deviance is approximately distributed as a $\chi^2$ statistic with 4 DF and provides a measure of lack-of-fit of the trend equation.
† SMR significant at 0.01 level.
* Trend significant at 0.01 level.

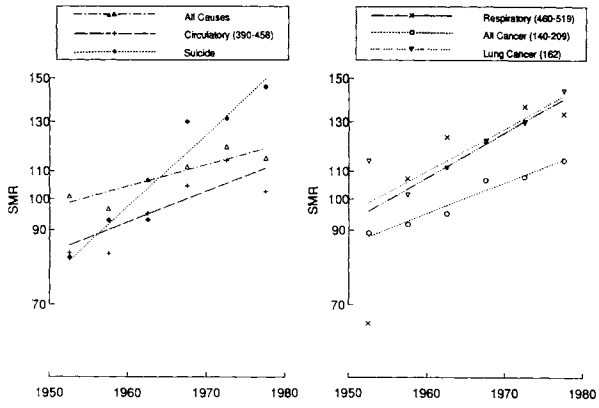


FIG. 1. Oak Ridge World War II workers. Observed and estimated SMRs for selected cause of death categories (see text and Appendix). Note that the vertical scale is in natural logarimithic units. The estimated SMR values for each category were obtained from the equation $\exp(\hat{\alpha}_a + \hat{\theta} t_j)$ using the values of $\hat{\alpha}_a$ and $\hat{\theta}$ in Tables IV and V.

In fact, a plot of the data (see Fig. 1) indicates that the worker population had elevated lung cancer rates in the first 5 years of follow-up and that the trend increased throughout the period of follow-up at a rate of 1.44% per year. The trend is statistically significant. Examination of the relationship between study factors and causes of death showing either significantly elevated SMRs or significant positive trends over time could help to generate some interesting hypotheses related to this population.

### *Multivariate Analysis*

The initial SMR trend analysis identified diseases or categories of diseases with interesting distributions in this population. We chose several for further study on the basis of either a statistically significant elevation of the SMR and/or a significantly increasing trend over the course of the study.

The reference group for all analyses is the group of workers who had no occupational radiation exposure, were born before 1910, were in follow-up period 1960s, worked less than 6 months, and were skilled workers. In the next stage of the analysis Poisson regression methods are used to examine the relationship between some diseases and the six factors described earlier. A detailed discussion of the methods and interpretation of the results are provided under Appendix.

Looking at the all causes of death column in Table VI, the same temporal trend that was observed in the overall SMR analysis is evident from the period effects. From Table IV, the SMR at the midpoint of the study is estimated to be 1.08 [exp (0.08)] and it is statistically significantly elevated. Table VI shows significantly elevated midpoint SMRs for three of the four facilities and a decreased SMR for one facility. Radiation exposure does not appear to alter this observation. Other significant factors in determining major overall increased death rates in this population of workers appear to be birth year before 1910, short length of employment, and unskilled socioeconomic status. The group experiencing the lowest death rate can be defined by choosing the lowest effect estimates for each factor. This group is

| | |
|---|---|
| X-10 | −6.6 |
| Radiation | −2.2 |
| Born after 1910 | 0 |
| Follow-up period 1950s | −10.2 |
| Worked ≥6 months | −11.4 |
| Professional | −39.1 |

The estimated effect of the combination of these factors is −69.5, from which we obtain 0.50 as an estimate of the SMR for all causes.

The group that had the worst overall mortality during this period of time was

| | |
|---|---|
| Multiple facility | 28.8 |
| No radiation | 0 |
| Born before 1910 | 12.2 |
| Followed in the 1970s | 6.2 |
| Worked <6 months | 0 |
| Unskilled | 7.1 |

The estimated combined effect is 54.3, from which we obtain 1.72 as an estimate of the SMR. By changing only the unskilled category to professional and keeping all the other factors constant in the example above, the SMR becomes 1.08. In fact, for each cause of death the socioeconomic status is the strongest predictor of outcome and strongly influences the calculation of the SMR.

The same strategy for examination of the estimated SMRs may be used for each cause of death in Tables VI and VII. Table VI displays multivariate analyses for all causes, diseases of the circulatory system, and external causes of

**TABLE VI**

**Oak Ridge World War II Workers Parameter Estimates for Main-Effects Model for Selected Cause of Death Categories**

| | All causes | | Circulatory system | | External causes | |
|---|---|---|---|---|---|---|
| Parameter | Estimate[a] | SD | Estimate | SD | Estimate | SD |
| Facility | | | | | | |
| X-10 vs US | −6.6 | 5.9 | −11.9 | 8.4 | 29.9 | 18.6 |
| TEC vs US | 9.9 | 2.7 | 4.4 | 3.9 | 22.7 | 9.7 |
| K-25 vs US | 17.2 | 2.6 | 9.9 | 3.7 | 39.2 | 9.3 |
| Multiple vs US | 28.8 | 4.3 | 16.0 | 6.3 | 51.7 | 14.1 |
| Radiation | | | | | | |
| Yes vs no | −2.2 | 2.1 | −5.6 | 3.0 | 3.6 | 7.1 |
| Birth year | | | | | | |
| <1910 vs 1910+ | 12.2 | 2.0 | 6.6 | 2.9 | 16.1 | 6.4 |
| Period | | | | | | |
| 1950s vs 1960s | −10.2 | 2.7 | −18.8 | 4.1 | −9.4 | 7.4 |
| 1970s vs 1960s | 6.2 | 2.1 | 7.6 | 2.9 | 2.9 | 7.6 |
| Duration | | | | | | |
| 6+ months vs <6 months | −11.4 | 1.9 | −8.0 | 2.8 | −18.2 | 6.5 |
| Socioeconomic status | | | | | | |
| Unskilled vs skilled | 7.1 | 2.3 | 5.3 | 3.2 | 7.2 | 7.6 |
| Professional vs skilled | −39.1 | 4.8 | −29.2 | 6.6 | −88.9 | 18.9 |

[a] The "facility" factor is used as the reference category so that the first four parameter estimates represent the natural logarithm of the SMR at the reference level of each of the remaining factors. The reference levels for internal comparisons are radiation, no; birth year, after 1910; period, 1960s; duration, less than 6 months; and socioeconomic status, skilled. The parameter estimates and standard deviation are in $L\%$ units (see the Appendix).

TABLE VII
Oak Ridge World War II Workers Parameter Estimates for Main-Effects Model for All Cancer, Lung Cancer, and All Other Cancer

| | All cancer | | Lung cancer | | Other cancer | |
|---|---|---|---|---|---|---|
| Parameter | Estimate[a] | SD | Estimate | SD | Estimate | SD |
| Facility | | | | | | |
| X-10 vs US | −1.7 | 12.9 | 5.7 | 21.4 | −3.9 | 16.2 |
| TEC vs US | −0.6 | 6.4 | 15.9 | 10.5 | −8.6 | 8.1 |
| K-25 vs US | 8.3 | 6.2 | 15.8 | 10.3 | 6.0 | 7.8 |
| Multiple vs US | 18.0 | 10.1 | 40.4 | 15.9 | 5.7 | 13.1 |
| Radiation | | | | | | |
| Yes vs no | 8.5 | 4.9 | 7.6 | 8.0 | 9.0 | 6.2 |
| Birth year | | | | | | |
| <1910 vs 1910+ | 3.3 | 4.6 | 6.1 | 7.3 | −1.3 | 5.9 |
| Period | | | | | | |
| 1950s vs 1960s | −11.4 | 6.8 | −9.9 | 12.5 | −10.0 | 8.1 |
| 1970s vs 1960s | 8.7 | 4.8 | 14.9 | 7.9 | 2.8 | 6.1 |
| Duration | | | | | | |
| 6+ months vs <6 months | −10.5 | 4.5 | −10.2 | 7.2 | −10.4 | 5.7 |
| Socioeconomic status | | | | | | |
| Unskilled vs skilled | 1.1 | 5.4 | 5.8 | 8.8 | −1.4 | 6.9 |
| Professional vs skilled | −43.0 | 11.0 | −66.9 | 19.5 | −29.8 | 13.3 |

[a] The facility factor is used as the reference category so that the first four parameter estimates represent the natural logarithm of the SMR at the reference level of each of the remaining factors. The reference levels for internal comparisons are radiation, no; birth year, after 1910; period, 1960s, duration, less than 6 months; and socioeconomic status, skilled. The parameter estimates and standard deviation are in $L\%$ units (see the Appendix).

death. For each cause of death category, the multiple population workers were significantly above the U.S. standard. This finding is consistent with the assumption that workers who worked at more than one facility in a short period of time may represent a group with unstable workers who may not be able to hold a job due to lack of education or because of poor work habits or poor health. As might be expected, radiation exposure does not have a significant relationship to any of the three disease categories, and in fact shows a slight negative effect for two of the disease groups. Those born before 1910 had significantly higher death rates in the three disease categories compared to those born after 1910. This is most likely reflective of a "cohort" effect where each successive birth cohort in the United States exhibits generally longer life expectancy than earlier cohorts. The observed period effects confirm a significant positive trend for all causes and circulatory system deaths that was noted in Table IV. Although there appears to be a slight positive trend for external causes, examination of the standard deviations for the coefficients reveals that the trend is not statistically significant. The period effect is stronger for cardiovascular diseases than for all causes. This is probably due to the strong healthy worker effect that is usually associated with cardiovascular diseases in an occupational cohort. In all cases, those who worked more than 6 months had reduced mortality when compared to short-term workers (<6 months). This finding reinforces the unstable worker hypothesis that was proposed for the multiple plant population. For each disease category in Table VI, socioeconomic status is a strong predictor of outcome. The differences between skilled and unskilled are small, whereas the comparison of professional to skilled reveals a strong reduction in death rates for professional workers.

Table VII displays multivariate analyses for all cancer, lung cancer, and other cancer deaths in the study population. Although there is some variation in the parameter estimates for the different facilities, only the estimate for lung cancer in multiple facilities is statistically significantly above the U.S. standard. Radiation exposure has a positive contribution in the main effects model for all cancer, but it does not change the magnitude of the SMR as much as other factors in the model. Birth year does not show as strong an effect for the cancer causes as was seen for the set of categories in Table VI. This would indicate that the 'cohort' effect does not operate strongly for cancers in this group. A positive trend is observed for each of the cancer categories when period of death is considered. Employment of ⩾6 months is associated with reduction of death from all cancer, lung cancer, and other cancer when compared to the <6-month employment group. No differences were noted between skilled and unskilled workers for the cancer categories. However, strong significant parameter estimates are obtained when professional workers are compared to skilled workers. The contribution of the socioeconomic status is the most significant in these models as was seen in the models presented in Table VI. Being a professional worker

is generally protective for all causes of death; however, it is seen to be particularly protective for external causes of death and lung cancer deaths in the examples above.

## DISCUSSION

The purpose of this study was to demonstrate the use of Poisson regression methods in an occupational cohort mortality analysis where there were multiple risk factors and multiple causes of death to be investigated. This type of descriptive hypothesis generating analysis is frequently employed as a first step in searching for any possible relationship between specific causes of death and employment in an occupational cohort or exposure to a certain substance in an occupational setting.

The cohort that was chosen for demonstration of these analytic techniques is unique because all members of the cohort terminated employment prior to the beginning of the follow-up period. This restriction allowed a data set where exposure was fixed and did not have to be treated as a time-dependent variable. Also, because of this approach, period of follow-up (as measured by calendar time) and duration of follow-up are synonymous. Every person in this data set who survived past the end of the study was followed for the maximum of 30 years. This long length of follow-up coupled with the selection bias introduced by including only terminated workers yields a cohort where there is no evidence of the "healthy worker effect."

This cohort also had a potential selection bias during the initial hiring period because of the effect of World War II. Most of the healthy young males in the U.S. population were already in the Armed Forces and any of the workers hired by the facilities in this study were either too old to enlist or not physicially fit enough to pass the Armed Forces entrance physical. Our analyses indicate that there was no healthy worker effect and that the younger draft-age workers did not have poorer mortality than the initially older workers. A similar result has been observed in other cohorts that were employed during the same era (*9, 20*).

Of particular interest in this study is the effect of radiation exposure on the risk of death from cancer. No effect was observed when all cancers were taken together or when lung cancer was specifically examined, although the parameter estimates were all positive for the three cancer categories displayed. Several reasons may be examined for this apparent lack of association. The radiation exposure variable was inferred after examination of job titles and department of work for each worker in the study. The procedure is subject to error that would result in misclassification of exposure status and lead to underestimation of risk in the exposed group. Also, since the radiation variable was dichotomous, and many workers were employed <6 months, a worker may be classified as exposed where the exposure is not significantly different than those in the unexposed category. Thus the exposure variable may not be a sensitive measure of true differences in dose. The population that was examined terminated employment before 1948 and the maximum length of employment was slightly less than 6 years. Thus the maximum length of exposure to radiation was less than 6 years. It is possible that this is not a sufficient length of time to cause increased risk for death from cancer.

Unfortunately, it has not been possible to quantify further the radiation exposures for the period of this study, since radiation monitoring programs were in developmental stages and personal monitoring data were not available for most workers. However, the importance of the study lies in clearly showing that future studies which will have more complete exposure data should not ignore the very strong socioeconomic status effects for most causes of death and possible birth cohort effects for noncancer deaths. The use of the Poisson regression and the multiplicative main effects models provides an effective data reduction method that is useful for looking at the multiple factors. This approach will be extended for use in the more rigorous studies on groups of workers who were monitored during their employment. These studies will necessarily incorporate time-dependent analyses to evaluate mortality patterns with respect to radiation exposure among all monitored workers.

## APPENDIX

### *Trend Analysis*

In occupational cohort studies, the SMR is often used to describe the mortality experience of the study group over an extended period of time—see, e.g., Whittemore (*16*) for a recent review. An implicit assumption in this summary is that the death rates differ from the reference population by a constant multiple that does not change over time. In this study, follow-up began on January 1, 1950, and since all workers were employed in Oak Ridge between 1942 and 1947 calendar time and duration of follow-up are synonymous. We divided the follow-up period into six 5-year intervals and determined the observed and expected deaths for each cause of death for each of the six intervals. The results for lung cancer are shown in Table A1. To determine if there is a systematic change in the SMR over this 30-year interval, we used a statistical model

$$E(y_j) = \mu_j^0 \exp(\alpha_a + \theta t_j), \qquad \text{(A1)}$$

where $j$ is the time interval, $y_j$ denotes the observed deaths, and $\mu_j^0$ is the "expected deaths" in the $j$th interval. The expected deaths, $\mu_j^0$, are computed using the age-specific cause-specific death rates for U.S. white males from Monson's program. The deaths are treated as observed values of Poisson variates, and the Poisson regression analysis is used to obtain maximum likelihood estimates of the unknown parameters $\alpha_a$ and $\theta$—see, e.g., Frome and Checkoway (*19*) and Whittemore (*16*). The parameter, $\theta$, represents the yearly rate of change in the SMR over the 30-year interval

TABLE A1
Summary Data for Lung Cancer Trend Analysis

| | | Deaths | | SMR | |
|---|---|---|---|---|---|
| $j$ | $t_j$ | Observed $y_j$ | Expected $\mu_j^0$ | Observed $y_j/\mu_j^0$ | Estimated $\exp(\hat{\alpha} + \hat{\theta}t_j)$ |
| 1 | −0.125 | 33 | 29.00 | 1.14 | 0.99 |
| 2 | −0.075 | 54 | 53.24 | 1.01 | 1.06 |
| 3 | −0.025 | 94 | 84.47 | 1.11 | 1.14 |
| 4 | 0.025 | 155 | 127.18 | 1.22 | 1.22 |
| 5 | 0.075 | 221 | 170.39 | 1.30 | 1.32 |
| 6 | 0.125 | 293 | 203.71 | 1.44 | 1.42 |

from 1950 to 1980. The predictor variable $t_j$ = (midpoint of $j$th interval minus 1965)/100 so that $\hat{\theta}$, the "trend statistic," is in percentage per year units. The estimated "intercept parameter," $\hat{\alpha}_a$, represents the natural logarithm of SMR at the midpoint of the 30-year interval. The data have been scaled so that $\hat{\alpha}_a$ is in log percentage units ($L\%$)—see Törnquvist *et al.* (*21*). The maximum likelihood estimates and their standard deviations are given in Columns 5 thru 9 of Tables IV and V. For the lung cancer data in Table A1 we obtained $\hat{\alpha}_a$ = 16.7 (SD = 4.36) and $\hat{\theta}$ = 1.44 (SD = 0.49). The estimated SMRs, exp (0.167 + 1.44$t_j$), at the midpoint of each 5-year interval are shown in Column 6 of Table A1. The SMR is increasing at a rate of 1.44% per year and the estimated value at the midpoint is exp(16.7/100) = 1.18. The observed and estimated SMR values are shown graphically in Fig. 1. If the trend statistic is significant, this indicates that the age-specific death rates of the cohort are changing relative to the U.S. population over the 30 years of follow-up. A test of the null hypothesis that the SMR is constant over the 30-year follow-up interval (i.e., $\theta = 0$) is obtained by dividing the ML estimate of the trend statistic by its standard deviation. For the lung cancer data $\hat{\theta}$/SD ($\hat{\theta}$) = 1.442/0.49 = 2.94, which is significant at the 0.01 level ($Z_{0.005}$ = 2.57). We have identified those cause-of-death categories that have trend statistics that are statistically significant at the 0.01 level in Tables IV and V. The last column of Tables IV and V gives the value of the deviance for each cause-of-death category. The deviance is approximately distributed as a $\chi^2$ statistic with 4 *df* and provides a measure of "lack-of-fit" of the trend equation—see the Appendix of Frome and Checkoway (*19*). For the lung cancer data, the value of the deviance is 0.94, which indicates that the observed and estimated SMR values are in close agreement.

TABLE A2
Oak Ridge World War II Workers Parameter Estimates for Main Effects Model for All Malignant Neoplasms

| Parameter | Estimate ($L\%$) | SD | SMR/RR | 95% CI |
|---|---|---|---|---|
| f—Facility | | | | |
| X-10 vs US | −1.7 | 12.9 | 0.98 | (0.76, 1.26) |
| TEC vs US | −0.6 | 6.4 | 0.99 | (0.88, 1.13) |
| K-25 vs US | 8.3 | 6.2 | 1.09 | (0.96, 1.23) |
| Multiple vs US | 18.0 | 10.1 | 1.20 | (0.98, 1.46) |
| x—Radiation (no) | | | | |
| Yes | 8.5 | 4.9 | 1.09 | (0.99, 1.20) |
| b—Birth year (before 1910) | | | | |
| After 1910 | 3.3 | 4.6 | 1.03 | (0.94, 1.14) |
| p—Period (1960s) | | | | |
| 1950s | −11.4 | 6.8 | 0.89 | (0.78, 1.02) |
| 1970s | 8.7 | 4.8 | 1.09 | (0.99, 1.20) |
| d—Duration (<6 months) | | | | |
| +6 months | −10.5 | 4.5 | 0.90 | (0.82, 0.98) |
| s—Socioeconomic status (skilled) | | | | |
| Unskilled | 1.1 | 5.4 | 1.01 | (0.91, 1.12) |
| Professional | −43.0 | 11.0 | 0.65 | (0.52, 0.81) |

TABLE A3
Oak Ridge World War II Workers Abridged $\chi^2$ Statistics for Marginal and Partial Association All Malignant Neoplasms (Total Deviance = 301 with 284 *df*)

| Term[a] | *df* | Marginal | Probability | Partial | Probability |
|---|---|---|---|---|---|
| x | 1 | 0.84 | 0.3594 | 3.03 | 0.0817 |
| b | 1 | 2.34 | 0.1261 | 0.51 | 0.4751 |
| p | 2 | 11.73 | 0.0028 | 10.61 | 0.0050 |
| d | 1 | 7.15 | 0.0075 | 5.47 | 0.0193 |
| s | 2 | 19.68 | 0.0001 | 17.63 | 0.0001 |
| f | 3 | 7.35 | 0.0615 | 6.70 | 0.0821 |
| xb | 1 | 0.01 | 0.9203 | 0.02 | 0.8875 |
| xp | 2 | 0.10 | 0.9512 | 1.51 | 0.4700 |
| xd | 1 | 0.28 | 0.5967 | 0.14 | 0.7083 |
| xs | 2 | 0.38 | 0.8270 | 0.12 | 0.9418 |
| xf* | 3 | 8.21 | 0.0419 | 7.08 | 0.0694 |
| bp | 2 | 2.28 | 0.3198 | 2.47 | 0.2908 |
| bd | 1 | 2.23 | 0.1354 | 1.71 | 0.1910 |
| bs | 2 | 2.14 | 0.3430 | 1.44 | 0.4868 |
| bf* | 3 | 6.56 | 0.0873 | 7.98 | 0.0464 |
| pd | 2 | 2.21 | 0.3312 | 1.74 | 0.4190 |
| ps | 4 | 4.68 | 0.3217 | 3.71 | 0.4467 |
| pf | 6 | 1.30 | 0.9717 | 0.72 | 0.9940 |
| ds | 2 | 0.15 | 0.9277 | 0.40 | 0.8187 |
| df | 3 | 3.04 | 0.3855 | 4.07 | 0.2540 |
| sf | 6 | 10.43 | 0.1077 | 9.05 | 0.1708 |

[a] Terms are defined in the text.
* Significant at the 0.05 level.

*Multivariate Analysis*

The analysis of mortality data from long-term follow-up studies has traditionally been based on SMR analysis (see Whittemore (*16*) for a recent review). The traditional approach is limited in the extent to which the combined effect of multiple risk factors can be disentangled. A second approach to cohort analysis that does not use an external standard requires stratification on age and calendar year as well

as the covariates of interest. This approach has been proposed for use in cohort studies of radiation-induced cancer (*22, 23*). A general discussion of recently developed biostatistical methods for the multivariate analysis of categorical and survival data has been given by Breslow (*14, 15*) and Breslow and Day (*24*). The approach that we have used is based on modeling the disease rates as a function of covariate levels using both external standard rates and internal comparisons—see (*23, 24*). This approach is simple and easy to interpret when the covariates are discrete with a small number of levels. In our situation, there are six factors that are of interest (see Materials and Methods):

| Term | Description | Level |
|---|---|---|
| $b$ | Birth year | 2 |
| $p$ | Period of follow-up | 3 |
| $f$ | Facility | 4 |
| $x$ | Radiation exposure | 2 |
| $d$ | Duration of employment | 2 |
| $s$ | Socioeconomic status | 3 |

This leads to a six-dimensional table ($2 \times 3 \times 4 \times 2 \times 2 \times 3$) that contains the observed and expected deaths for each combination of factor levels. Table B shows the complete six-dimensional table for all cancer (ICD 140–209) causes and will be used to explain the multivariate analysis procedure. Note that we have only partially collapsed over age and calendar year since two of the factors in our analysis are birth year and period of follow-up. This approach allows us to investigate temporal patterns in the relative risk as well as interactions between radiation exposure and time since exposure. It also allowed us to see that the SMRs for a number of disease categories were not constant over time. We also carried out a more detailed analysis using 5-year age and calendar year intervals with and without the external standard rates for all cancer causes. The resulting table contained 2790 cells with nonzero person-years, and results obtained for the main effects model were in close agreement with those given in Table VII.

The cause-specific hazard rate for the $i$th individual in category $j$ is assumed to satisfy the multiplicative relation

$$\lambda(t; \mathbf{z}_j, \boldsymbol{\beta}) = \exp(\mathbf{z}_j\boldsymbol{\beta})\lambda^0(t)$$

If the $i$th individual enters the study at age $a$ and is in view until age $t$, then $e_i = \int_a^t \lambda^0(u)du$ is the contribution this person makes to the expected deaths, where $\lambda^0(t)$ represents the age-cause-specific death rates for some external standard population (U.S. white males). The expected deaths for group $j$ ($\mu_j^0 = \Sigma_i e_{ij}$) are obtained by summing the contribution for each individual in the group. The observed deaths, $y_j$, are the number of individuals in group $j$ that died during the period of follow up from the cause of death of interest. We then treat the observed deaths, $y_j$, as Poisson variates with expectation $\mu_j^0\exp(\mathbf{z}_j\,\boldsymbol{\beta})$, where the $\mu_j^0$ are known constants, $\mathbf{z}_j$ is a covariate vector of indicator variables, and $\boldsymbol{\beta}$ is a vector of regression coefficients. The expected deaths were computed using the person-years approach and the observed and expected deaths were cross-tabulated using a system of SAS macros that we have developed (*17*). Poisson regression methods (*18*) were used to obtain maximum likelihood estimates of the parameters in a main-effects model for the six-dimensional table of observed and expected deaths. In our analysis, we have used the "facility" factor as the reference category so that the first four parameter estimates in Table A2 represent the log of the SMR for all cancer at the reference level of each of the remaining factors; e.g., the estimated SMR for K-25, at the reference level of each of the remaining five factors, is $\exp(8.3/100) = 1.09$. The remaining parameter estimates in Table A2 provide estimates of the relative risk associated with the levels of each of the factors. This log-linear main-effects model is

$$\begin{aligned}\log(\text{SMR}) = {} & \log(\text{facility SMR}) \\ & + (\text{effect due to radiation exposure}) \\ & + (\text{effect due to birth year}) \\ & + (\text{effect due to period of follow-up}) \\ & + (\text{effect due to duration of employment}) \\ & + (\text{effect due to socioeconomic status}).\end{aligned}$$

The effect estimates represent the relative risk (on a logarithmic scale in $L\%$ units) for each level of the five additional factors. For example, the estimated effect of radiation exposure on cancer risk is $8.5L\%$ with a standard error of $4.9L\%$. Thus, on taking exponents we see that the relative risk is $\exp(0.085) = 1.089$ with a multiplicative standard error of 1.050. The estimated effect of duration of employment is $-10.5L\%$ so that the relative risk (long-term vs short-term) is $\exp(-0.105) = 0.90$. To obtain an estimate of the SMR for any combination of factor levels, we add the effect estimates to the log(SMR) for the facility of interest. For example, consider K-25 workers with radiation exposure, born after 1910, during the 1970s, that worked more than 6 months and were classified as being skilled workers,

$$\begin{aligned}\log(\text{SMR}) = {} & 8.3 + 8.5 + 3.3 \\ & + 8.7 - 10.5 + 0.0 = 18.3L\%,\end{aligned}$$

and taking the exponent, the SMR = 1.20.

To evaluate the relative importance of each of the factors and to investigate the possibility of "interactions" we used the model screening procedure developed by Brown (*25*). In our situation the factors that define the multidimensional table are all viewed as "predictor variables" and the "response" variable is the SMR. This is analogous to a facto-

**TABLE B**
**ORFCOM—Oak Ridge World War II Workers Observed and Expected Deaths for All Cancer (ICD 140-209)**

| Socioeconomic Status: | | Skilled | | | | Unskilled | | | | Professional | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Birth year: | | <1910 | | 1910+ | | <1910 | | 1910+ | | <1910 | | 1910+ | |
| Radiation: | | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Facility: X-10 Duration: LT 6 | Birth year: | | | | | | | | | | | | |
| Period | | | | | | | | | | | | | |
| 1950s | obs | 0 | 2 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | exp | 0.47 | 1.83 | 0.12 | 0.50 | 0.26 | 0.59 | 0.14 | 0.16 | 0.15 | 0.12 | 0.03 | 0.21 |
| 1960s | obs | 1 | 2 | 2 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| | exp | 0.83 | 3.51 | 0.34 | 1.69 | 0.60 | 0.92 | 0.44 | 0.47 | 0.32 | 0.24 | 0.09 | 0.67 |
| 1970s | obs | 0 | 7 | 0 | 2 | 2 | 1 | 0 | 1 | 1 | 0 | 0 | 3 |
| | exp | 1.05 | 4.40 | 0.98 | 4.40 | 0.71 | 1.15 | 1.04 | 1.25 | 0.58 | 0.51 | 0.26 | 1.99 |
| Facility: X-10 Duration: GE 6 | | | | | | | | | | | | | |
| Period | | | | | | | | | | | | | |
| 1950s | obs | 1 | 1 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| | exp | 0.84 | 1.68 | 0.21 | 0.45 | 0.84 | 0.42 | 0.19 | 0.10 | 0.32 | 0.72 | 0.06 | 0.78 |
| 1960s | obs | 2 | 3 | 1 | 3 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 2 |
| | exp | 1.81 | 2.88 | 0.63 | 1.44 | 1.58 | 0.92 | 0.53 | 0.29 | 0.54 | 1.74 | 0.21 | 2.64 |
| 1970s | obs | 3 | 4 | 4 | 2 | 3 | 1 | 3 | 1 | 2 | 1 | 0 | 3 |
| | exp | 2.69 | 3.57 | 1.60 | 3.72 | 2.16 | 1.36 | 1.30 | 0.75 | 0.63 | 2.93 | 0.55 | 7.71 |
| Facility: TEC Duration: LT6 | | | | | | | | | | | | | |
| Period | | | | | | | | | | | | | |
| 1950s | obs | 11 | 18 | 2 | 7 | 6 | 3 | 1 | 0 | 0 | 0 | 0 | 1 |
| | exp | 10.88 | 24.01 | 2.31 | 7.34 | 10.39 | 2.79 | 1.26 | 0.72 | 0.76 | 0.64 | 0.12 | 0.17 |
| 1960s | obs | 17 | 42 | 8 | 29 | 10 | 5 | 9 | 4 | 0 | 0 | 0 | 0 |
| | exp | 18.69 | 45.93 | 7.38 | 22.88 | 17.88 | 4.94 | 3.84 | 2.25 | 1.27 | 1.23 | 0.38 | 0.56 |
| 1970s | obs | 23 | 62 | 25 | 74 | 17 | 5 | 13 | 14 | 1 | 0 | 0 | 3 |
| | exp | 23.12 | 59.09 | 18.88 | 56.79 | 17.86 | 5.48 | 9.86 | 6.04 | 0.99 | 1.87 | 0.96 | 1.57 |
| Facility:TEC Duration: GE 6 | | | | | | | | | | | | | |
| Period | | | | | | | | | | | | | |
| 1950s | obs | 19 | 53 | 6 | 15 | 6 | 8 | 1 | 1 | 3 | 2 | 0 | 1 |
| | exp | 17.59 | 61.92 | 2.59 | 16.09 | 10.15 | 9.36 | 0.71 | 1.02 | 2.36 | 3.56 | 0.29 | 1.48 |
| 1960s | obs | 25 | 115 | 9 | 50 | 18 | 13 | 1 | 6 | 2 | 7 | 2 | 3 |
| | exp | 29.91 | 117.21 | 8.24 | 52.14 | 15.23 | 15.12 | 2.26 | 3.07 | 3.60 | 7.21 | 0.92 | 4.87 |
| 1970s | obs | 33 | 165 | 19 | 132 | 10 | 16 | 10 | 8 | 3 | 9 | 0 | 14 |
| | exp | 33.03 | 146.27 | 20.72 | 131.22 | 14.21 | 14.42 | 5.55 | 7.60 | 4.21 | 11.28 | 2.41 | 14.02 |
| Facility: K-25 Duration: LT 6 | | | | | | | | | | | | | |
| Period | | | | | | | | | | | | | |
| 1950s | obs | 17 | 9 | 3 | 4 | 25 | 1 | 2 | 1 | 0 | 0 | 0 | 0 |
| | exp | 16.27 | 10.09 | 3.30 | 3.12 | 21.13 | 3.89 | 3.37 | 0.82 | 0.92 | 0.49 | 0.18 | 0.20 |
| 1960s | obs | 34 | 24 | 13 | 11 | 42 | 6 | 10 | 1 | 2 | 0 | 1 | 0 |
| | exp | 26.51 | 17.08 | 10.07 | 9.66 | 33.11 | 5.41 | 10.39 | 2.47 | 1.57 | 0.86 | 0.60 | 0.65 |
| 1970s | obs | 37 | 26 | 30 | 34 | 41 | 9 | 27 | 12 | 0 | 1 | 1 | 2 |
| | exp | 27.18 | 18.84 | 25.11 | 24.11 | 32.11 | 5.41 | 26.06 | 6.42 | 1.54 | 1.11 | 1.60 | 2.03 |
| Facility: K-25 Duration: GE 6 | | | | | | | | | | | | | |
| Period | | | | | | | | | | | | | |
| 1950s | obs | 11 | 19 | 3 | 11 | 10 | 4 | 2 | 0 | 1 | 0 | 0 | 0 |
| | exp | 16.55 | 16.25 | 2.65 | 4.09 | 15.98 | 5.19 | 1.43 | 0.55 | 1.79 | 0.69 | 0.31 | 0.66 |
| 1960s | obs | 25 | 32 | 6 | 11 | 23 | 9 | 4 | 1 | 2 | 2 | 0 | 0 |
| | exp | 25.79 | 28.47 | 8.30 | 12.75 | 24.14 | 6.92 | 4.54 | 1.64 | 3.08 | 1.37 | 1.00 | 2.16 |
| 1970s | obs | 26 | 43 | 16 | 33 | 29 | 7 | 11 | 7 | 2 | 0 | 0 | 2 |
| | exp | 28.51 | 33.86 | 20.61 | 31.71 | 23.10 | 6.02 | 11.70 | 3.96 | 3.10 | 1.61 | 2.76 | 6.31 |

TABLE B—*Continued*

| Socioeconomic Status: | | Skilled | | | | Unskilled | | | | Professional | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Birth year: | | <1910 | | 1910+ | | <1910 | | 1910+ | | <1910 | | 1910+ | |
| Radiation: | | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Facility: Multiple | | | | | | | | | | | | | |
| Duration: LT 6 | | | | | | | | | | | | | |
| Period | | | | | | | | | | | | | |
| 1950s | obs | 2 | 3 | 1 | 0 | 0 | 0 | 0 | 1 | NA | 0 | 0 | 0 |
| | exp | 1.50 | 1.07 | 0.38 | 0.55 | 1.15 | 0.65 | 0.35 | 0.28 | NA | 0.01 | 0.01 | 0.0 |
| 1960s | obs | 3 | 4 | 0 | 1 | 2 | 0 | 0 | 0 | NA | 0 | 0 | 0 |
| | exp | 2.37 | 1.52 | 1.16 | 1.69 | 2.14 | 1.20 | 1.05 | 0.82 | NA | 0.03 | 0.02 | 0.01 |
| 1970s | obs | 3 | 2 | 2 | 7 | 5 | 2 | 4 | 5 | NA | 1 | 0 | 0 |
| | exp | 2.55 | 1.52 | 2.77 | 4.07 | 1.91 | 1.14 | 2.45 | 2.08 | NA | 0.06 | 0.04 | 0.03 |
| Facility: Multiple | | | | | | | | | | | | | |
| Duration: GE 6 | | | | | | | | | | | | | |
| Period | | | | | | | | | | | | | |
| 1950s | obs | 1 | 5 | 0 | 2 | 0 | 2 | 0 | 1 | 0 | 1 | 0 | 1 |
| | exp | 2.09 | 5.59 | 0.32 | 1.69 | 1.98 | 2.28 | 0.18 | 0.55 | 0.10 | 0.30 | 0.02 | 0.11 |
| 1960s | obs | 1 | 14 | 2 | 10 | 4 | 2 | 0 | 2 | 0 | 2 | 0 | 1 |
| | exp | 3.49 | 10.08 | 1.01 | 5.28 | 3.13 | 3.58 | 0.51 | 1.75 | 0.24 | 0.34 | 0.08 | 0.35 |
| 1970s | obs | 5 | 16 | 3 | 16 | 5 | 4 | 2 | 5 | 0 | 0 | 0 | 1 |
| | exp | 3.39 | 12.49 | 2.39 | 13.27 | 3.49 | 3.57 | 1.16 | 4.59 | 0.39 | 0.51 | 0.23 | 1.01 |

rial experiment in which the relative importance of experimental factors is evaluated using an analysis of variance approach. The parallelism between log-linear models and the general linear hypothesis is used to compute test statistics for various model terms of interest. These test statistics are called marginal and partial associations and are used to evaluate the relative importance of terms in a log-linear model. The importance of a given term is measured by the change in the deviance that occurs when that term is added to or removed from the model. The resulting likelihood ratio test statistic will be approximately distributed as $\chi^2$ under the null hypothesis. Consider, for example, line three in Table A3 labeled "p" for period of follow-up. This term has 2 *df*, and the effect of including it in the model when no other terms are present is a decrease in the deviance of 11.73 (see Column 3 of line 3 in Table A3). This is a marginal test for the equality of the period effects and is equivalent to forming a marginal table (by collapsing over all factors except p) and then testing for the equality of the SMRs in this one-dimensional table. The value of the partial $\chi^2$ for the period effect when all other factors are present in the model is given in Column 5 of Table A3, and is obtained by fitting the main-effects model and then fitting the model with the p term removed. The value of the deviance (which is equivalent to the residual sum of squares) for the main effects model (x + b + p + d + s + f) is 253.97 with 274 *df*. The value of the deviance with this term deleted is 264.58 with 276 *df* and the difference, 10.61 with 2 *df*, is given in the fifth column of Table A3. If the change in the deviance is large then the term is important. The probability of obtaining a $\chi^2$ value of 10.61 (with 2 *df*) or larger is 0.005 (see the last column of Table A3). Consequently, we conclude that period of follow-up is an important variable for describing cancer mortality in the World War II cohort.

The next section of Table A3 (lines 7 through 21) is used to evaluate the relative importance of two-factor interaction terms. The $\chi^2$ statistic for marginal association can be obtained by fitting a main-effects model to the two-dimensional marginal table (that correspond to the term listed in Column 1). The value of the deviance for this marginal two-factor main-effect model is a measure of the interaction between the two factors when no other factors are present. The $\chi^2$ statistic for partial association for a two-factor interaction is obtained by first fitting the model with all main-effect and two-factor interactions in the model and then recording the change in the deviance that occurs when the two-factor interaction is removed from the model. This model screening procedure evaluates the relative importance of all possible higher-order interaction terms by following a procedure similar to that described for main effects and two-factor interactions (see Brown (*25*) for a more detailed discussion). The resulting table contains 62 lines with deviance values for marginal and partial association. For summary purposes, we have reviewed each of these tables and deleted all lines for three-factor or higher-order actions that were not significant at the 0.01 level. This resulted in an abridged table of $\chi^2$ statistics for each cause-of-death category that was studied. The purpose of this table is to evaluate the relative importance of different effects for each cause of death category. We emphasize that, due to

nonorthogonality, no single test can be performed to determine the importance of an effect. The two tests used here provide approximate bounds on the change in the deviance that occurs when that effect is either added to or deleted from the model.

## ACKNOWLEDGMENTS

The research was supported by the U.S. Department of Energy, Office of Energy Research, under Contract DE-AC05-76OR00033 with Oak Ridge Associated Universities. The authors thank Julia McClanahan, Martha Wray, Michael Zeigler, and Toni Newport of the Center for Epidemiologic Studies for their help in data collection, and Rhonda Harbison for assistance in manuscript preparation. We also thank the members of the Oak Ridge Associated Universities/University of North Carolina Research Planning Group and two referees for their helpful comments. 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), under the direction of Dr. T. F. Mancuso. 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.

RECEIVED: October 19, 1989; ACCEPTED: March 26, 1990

## REFERENCES

1. A. P. POLEDNAK and E. L. FROME, Mortality among men employed between 1943 and 1947 at a uranium-processing plant. *J. Occup. Med.* **23,** 169–178 (1981).
2. A. P. POLEDNAK, Mortality among men occupationally exposed to phosgene in 1943–1945. *Environ. Res.* **22,** 357–367 (1980).
3. H. CHECKOWAY, R. M. MATHEW, C. M. SHY, J. E. WATSON, JR., W. G. TANKERSLEY, S. H. WOLF, J. C. SMITH, and S. A. FRY, Radiation, work experience, and cause specific mortality among workers at an energy research laboratory. *Br. J. Indust. Med.* **42,** 525–533 (1985).
4. D. L. CRAGLE, D. R. HOLLIS, J. R. QUALTERS, W. G. TANKERSLEY, and S. A. FRY, A mortality study of men exposed to elemental mercury. *J. Occup. Med.* **26,** 817–821 (1984).
5. D. L. CRAGLE, D. R. HOLLIS, C. M. SHY, and T. H. NEWPORT, A retrospective cohort mortality study among workers occupationally exposed to metallic nickel powder at the Oak Ridge Gaseous Diffusion Plant. In *Nickel in the Human Environment* (F. W. Sunderman *et al.*, Eds.), pp. 57–63. Scientific Publication No. 53, International Agency for Research on Cancers, Lyon, 1984.
6. A. P. POLEDNAK and D. R. HOLLIS, Mortality and causes of death among workers exposed to phosgene in 1943–1945. *Toxicol. Indust. Health* **1,** 137–147 (1985).
7. A. V. CARPENTER, W. D. FLANDERS, E. L. FROME, D. J. CRAWFORD-BROWN, and S. A. FRY, Central nervous system cancers and radiation exposure: A case-control study among workers at two nuclear facilities. *J. Occup. Med.* **39,** pp. 601–604 (1987).
8. H. CHECKOWAY, R. M. MATHEW, C. M. SHY, J. E. WATSON, JR., W. G. TANKERSLEY, S. H. WOLF, J. C. SMITH, and S. A. FRY, Radiation doses and cause-specific mortality among workers at a nuclear materials fabrication plant. *Am. J. Epidemiol.* **127,** 255–266 (1988).
9. C. P. WEN, S. P. TASI, N. S. WEISS, and R. L. GIBSON, Long-term study of oil refinery workers: V. Comparison of workers hired before, during, and after World War II (1940–1945) with a discussion on the impact of study designs on cohort results. *Am. J. Indust. Med.* **9,** 171–180 (1986).
10. R. G. HEWLETT and O. E. ANDERSON, *The New World,* 1939/1946. WASH-1214 Vol. 1. National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia, 1972.
11. B. C. HACKER, *The Dragon's Tail.* Univ. of California Press, Berkeley, 1987.
12. R. R. MONSON, Analysis of relative survival and proportional mortality. *Comp. Biomed. Res.* **9,** 325–332 (1974).
13. N. E. BRESLOW, J. H. LUBIN, P. MAREK, and B. LANGHOLZ, Multiplicative models and cohort analysis. *J. Am. Statist. Assoc.* **78,** 1–12 (1983).
14. N. E. BRESLOW, Cohort analysis in epidemiology. In *A Celebration of Statistics* (A. C. Atkinson and D. E. Fienberg, Eds.), pp. 109–143. Springer-Verlag, New York, 1985.
15. N. E. BRESLOW, Multivariate models and cohort analysis. *NCI Monogr.* **67,** 149–156 (1985).
16. A. S. WHITTEMORE, Methods old and new for analyzing occupational cohort data. *Am. J. Indust. Med.* **12,** 233–248.
17. R. W. MCLAIN and E. L. FROME, A system for the analysis of cohort mortality data. In *SAS Users Group International Conference Proceedings: SUGI 11,* pp. 874–876. SAS Institute, Cary, NC, 1986.
18. E. L. FROME, The analysis of rates using Poisson regression models. *Biometrics* **39,** 665–674 (1983).
19. E. L. FROME and H. CHECKOWAY, Use of Poisson regression models in estimating incidence rates and ratios. *Am. J. Epidemiol.* **121,** 309–323 (1985).
20. E. A. DUPREE, D. L., CRAGLE, R. W. MCLAIN, D. J. CRAWFORD-BROWN, and M. J. TETA, Mortality among workers at a uranium processing facility, the Linde air products company ceramics plant, 1943–1949. *Scand. J. Work Environ. Health* **13,** 100–107 (1988).
21. L. TÖRNQUVIST, P. VARTIA, and Y. O. VARTIA, How should relative changes be measured? *Am. Statist.* **39,** 43–46 (1985).
22. E. L. FROME and D. R. HUDSON, A general statistical data structure for epidemiologic studies of DOE workers. In *Proceedings of the 1980 DOE Statistical Symposium* (T. Truett, D. Margolis, and R. W. Mensing, Eds.), Report Number CONF-801045, Oak Ridge National Laboratory, Oak Ridge, TN 1981.
23. D. A. PIERCE and D. L. PRESTON, Developments in cohort analysis with application to radiation induced cancer. In *Proceedings of the 46th Session of the International Statistical Institute,* Amsterdam, 1986.
24. N. E. BRESLOW and N. E. DAY, *Statistical Methods in Cancer Research, Volume II: The Design and Analysis of Cohort Studies.* Scientific Publication No. 82, International Agency for Research on Cancer, Lyon, 1987.
25. M. B. BROWN, Screening effects in multidimensional contingency tables. *Appl. Statist.* **23,** 37–46 (1976).