PB92-184282

A Dose-Response Analysis and Quantitative Assessment of Lung Cancer Risk and Occupational Cadmium Exposure

Leslie Stayner, Ph.D.[a]
Randall Smith, M.A.[a]
Michael Thun, M.D. M.Sc.[b]
Teresa Schnorr, Ph.D.[c]
Richard Lemen, M.S.[d]

September 4, 1990

a. Division of Standards Development and Technology Transfer, National Institute for Occupational Safety and Health.

b. Formerly with the Industrywide Studies Branch, Division of Surveillance, Hazard Evaluations and Field Studies, National Institute for Occupational Safety and Health. Currently with the American Cancer Society.

c. Industrywide Studies Branch, Division of Surveillance, Hazard Evaluations and Field Studies, National Institute for Occupational Safety and Health.

d. Washington Office, National Institute for Occupational Safety and Health.

Abstract

We have performed a quantitative assessment of the risk of lung cancer from exposure to cadmium based upon a retrospective cohort mortality study of workers exposed to cadmium. The study facility has processed cadmium metals and compounds since 1925, and was previously used as an arsenic and lead smelter. The study population consisted of white male workers who were employed for at least six months at this facility between January 1, 1940 and December 31, 1969 and who were first employed at the facility on or after January 1, 1926.

A life-table analysis was used to estimate standardized mortality ratios (SMRs), and to generate a matrix of person-years and lung cancer deaths for Poisson regression analyses. Several functional forms (i.e., exponential, linear, additive relative rate, and power) for modeling cumulative cadmium exposure were evaluated in the Poisson regression analyses. Analyses in which cadmium exposure was lagged 5, 10, 15, or 20 years were also performed. Similar models were also fitted using the Cox proportionate hazards model for comparison purposes. Estimates of working lifetime risk (45 years) were developed using an approach which corrects for competing causes of death.

A total of 606 white male workers, contributing 16,898 person-years and 162 deaths met our selection criteria and were included in the analysis. Vital status was successfully determined for over 98% of this cohort. A nearly significant ($p=0.076$, 2 tails) excess in mortality from lung cancer was observed for the entire cohort (SMR=149). Mortality from lung cancer was significantly elevated among non-Hispanic workers (SMR=211), workers in the highest cadmium exposure group (SMR=272), and among workers with 20 or more years since first exposure (SMR=161).

A statistically significant dose-response was evident in nearly all of the Poisson regression models evaluated. Lagging the exposures for five years appeared to maximize the dose-response relationship in the Poisson regression analyses, and this lag period was used for the risk assessment. The additive relative rate form of the Poisson model was judged to provide the best representation of the dose-response for the risk assessment. Similar results were obtained from modeling using the Cox model, although the coefficients representing exposure were somewhat lower.

Based upon our analyses the lifetime excess lung cancer risk at the current OSHA standard for cadmium fumes of 100 $\mu g/M^3$ is estimated to be approximately 50 to 111 lung cancer deaths per 1000 workers exposed to cadmium for 45 years. The proposed OSHA standard would significantly reduce this risk; however, efforts should be made to further reduce the risk by lowering occupational exposures to cadmium to the lowest level feasible.

INTRODUCTION

Occupational exposures to cadmium occur in the production of batteries, pigments, and stabilizers, electroplating, in the refinement of cadmium, lead, copper, and zinc sulphide and in a wide range of other U.S. industries. The Occupational Safety and Health Administration (OSHA) has recently estimated that approximately a half million (512,125) U.S. workers are currently exposed to cadmium [OSHA 1990].

A dose-related increase in carcinoma of the lung has been induced in rats following long-term inhalation of aerosols of cadmium chloride, oxide (dust and fume), sulfide, and sulfate [Takenaka 1983, Oldiges 1989]. Epidemiologic studies of workers exposed to cadmium have been conducted in the U.S., England, and Sweden. A statistically significant increase in mortality from lung cancer has been reported among U.S. cadmium recovery workers with two or more years employment [Lemen 1976, Thun 1985], in nickel-cadmium battery workers in the U.K. [Sorahan 1987] and Sweden [Elinder 1985, Jarup 1990], and in a British study of workers at 17 plants including a large zinc smelter [Armstrong and Kazantzis 1983, Kazantzis et al., 1988]. Evidence of a dose-response relationship between lung cancer and cadmium exposure exists in four of these populations [Thun 1985, Sorahan 1987, Kazantzis 1988, Jarup 1990]. However, controversy exists about the extent to which the excess lung cancer risk observed among cadmium workers could be explained by concomitant workplace exposures (i.e., arsenic or nickel) or cigarette smoking.

In 1976 investigators [Lemen et al., 1976] at the National Institute for Occupational Safety and Health (NIOSH) reported on the findings from a retrospective cohort mortality study of workers exposed to cadmium at a U.S. cadmium production facility. The study by Lemen et al., was the first to report a statistically significant excess of lung cancer mortality (SMR=235) among cadmium-exposed workers. Lemen et al. [1976] also reported a statistically nonsignificant increase in mortality from nonmalignant respiratory disease (SMR=159), and a statistically significant excess of mortality from prostatic cancer among workers with more than 20 years since first exposure (latency) to cadmium (SMR=452). Based on an update of the NIOSH cadmium cohort, Thun et al. [1985] reported a statistically significant excess in mortality from lung cancer (SMR=229) among workers who were hired after 1925 and had been employed for more than two years, and a statistically significant dose-response relationship between cumulative exposure to cadmium and lung cancer mortality. However, Thun et al., failed to detect a statistically significant increase in mortality from prostatic cancer or nonmalignant respiratory diseases.

Data from the NIOSH cadmium study [Thun et al. 1985] have been used by the U.S. Environmental Protection Agency [EPA 1984] and by the Occupational Safety and Health Administration [OSHA 1989] for quantitative risk assessment, because they provide the only human epidemiologic data relating quantitative estimates of cumulative exposure to cadmium to lung cancer mortality. Both of these risk assessments relied on modeling of the SMRs reported in the paper by Thun et al.

[1985]. Reliance on SMRs may have biased these assessments towards underestimation of the risk, since workers generally experience lower mortality than the U.S. population ("healthy worker effect"), the referent group in this analysis.

In this report we present the results from modeling of the dose-response relationship between cadmium exposure and lung cancer mortality, and projections from these models of the risks associated with varying levels of cadmium exposure. The modeling techniques used in this assessment are based on comparison of rates within the cohort, and are thus not subject to the bias related to the use of U.S. population rates. This analysis is based upon the most recent follow-up of the NIOSH cadmium cohort through 1984, which was previously analyzed and presented at a workshop by Thun et al. [1986].

MATERIALS AND METHODS

Background

A detailed description of the production process and the criteria used to define the NIOSH cadmium cohort has been previously reported [Thun et al., 1985]. Briefly, the study facility has processed cadmium metals and compounds since 1925. The facility had operated as an arsenic smelter from 1918 to 1925, and as a lead smelter from 1886 to 1918. Small quantities of high-purity arsenic continue to be produced at this facility by a few individuals in separate buildings, and some arsenic is evolved during the (early stages of the) cadmium process.

The study population was identified from plant personnel records and consisted of all hourly employees and foremen who had worked for at least six months in a production area of the facility between January 1, 1940, and December 31, 1969. NIOSH identified this cohort jointly with a company representative and senior union officials reviewed the list of study subjects. Females and nonwhites were excluded from the analysis because of small numbers (total=13). In order to minimize potential confounding by arsenic exposure, the analysis was restricted to only include workers who were first hired on or after January 1, 1926.

Air-monitoring data for cadmium has been collected by the company since the 1940s, which demonstrate that cadmium exposures vary appreciably by time period and among departments within the facility. The company's air-monitoring data was used by Smith et al. [1980a, 1980b] to construct

a cadmium exposure matrix by time period and department, that was adjusted for respirator usage. This exposure matrix provided the basis for the cumulative cadmium exposure estimates used in this analysis .

Each period of a worker's employment history was grouped into one of seven broad job categories, because many of the personnel records only specified general work categories rather than specific departments. Each worker's cumulative exposure over time was then estimated as the sum of the number of days worked (including weekends and holidays) in a given job category multiplied by the average inhalation exposure of that category for the relevant time period. For a complete description of how the exposure categories were constructed the reader is referred to the previous report by Thun et al. [1985].

In the previous report [Thun et al., 1985], vital status was ascertained as of December 31, 1978, using records from the Social Security Administration, state vital statistics offices, company and union, and direct telephoning. Records from the National Death Index have been used to update the vital status from the previous report to December 31, 1984. An analysis based on this updated cohort was presented by Thun et al. [1986] at a workshop on cadmium and cancer in Oxford, England. Death certificates for deceased individuals were obtained and coded by a trained nosologist using the revision in effect at the time of death. Deceased workers for whom no death certificate could be located were assumed dead on the date specified by the reporting agency, with cause of death unknown.

Life-Table Analysis

In this study, the NIOSH life-table program [Waxweiler et al. 1983] was used to compare the mortality experience of the cohort to that of the U.S. population and to create a data file for modeling the relationship between lung cancer mortality and cumulative cadmium exposure. Using this program, expected numbers of death were computed by multiplying cause, 5-year-age, and 5-year-calendar-time specific mortality rates from white males in the U.S. population by the corresponding person-years distribution of the study population. Person-years for the cadmium cohort were calculated from the time an individual had been employed for at least six months at the facility or after January 1, 1940, whichever occurred later; until the end of the study (December 31, 1984), or until the individual was lost to follow-up, or until the date of death. Standardized mortality ratios (SMRs) were calculated by dividing the observed number of deaths by the number expected and multiplying by 100. Statistical tests (2-sided) and associated 95% confidence intervals were estimated based upon the Poisson distribution.

The life-table analysis was stratified into four cumulative dose categories (i.e., $\leq$584, 585-1,460, 1,461-2,920, and $\geq$ 2,921 mg/M^3-days[1]) and three times since first exposure categories (i.e., <10, 10-19, 20+ years). The cumulative exposure categories were chosen previously by Thun et al. [1986], in the analysis presented at the Oxford workshop. Separate life-table analyses were performed for hispanics and non-

[1]milligrams of cadmium per cubic meter of air-days

Hispanics, because hispanics have been reported to experience lower lung cancer rates than non-Hispanics [Savitz 1986, Key 1981, Samet 1980].

Separate analyses were also performed in which cadmium cumulative exposures were "lagged" 5, 10, 15 and 20 years. In these analyses, each person-year of observation and death was classified according to the cumulative exposure level that was achieved 5, 10, 15 or 20 years previously. This technique has been used as a method for discounting exposures that may be etiologically irrelevant to the development of cancer [Rohman 1981, Checkoway et al. 1990].

Dose-response Modeling

The observed deaths and person-years from the life-table analyses were entered into a computer file for the regression analysis. The number of age- and calendar-year categories in this computer file were reduced to avoid producing marginal hazard rates with no deaths. Five-year-age groups were used except for the youngest group which was less than 50, and the oldest age-group which was greater than 75. Calendar time was stratified into four intervals that were approximately 10-years-wide except for the first and last categories (i.e., 1940-1960, 1960-1970, 1970-1980, 1980-1984). The output from the life-table analyses for hispanics and non-Hispanics were combined into one computer file with separate strata for each of these groups.

Regression models based upon the assumption of a Poisson distribution were used to model the hazard rate (incidence density) as a function of

cumulative cadmium exposure and the other covariates (i.e., age, calendar year and hispanic ethnicity). These models were fitted to the data using GLIM [Baker and Nelder, 1978], except for the additive relative rate model (described below), which was fitted using a customized program developed with the NLIN procedure of SAS [SAS 1987]. Both of these programs use iteratively reweighted least squares to derive the parameter and variance estimates. These procedures have been shown to provide equivalent results to maximum likelihood estimation when the data are from a regular exponential family model such as the Poisson model [Jennrich and Ralston 1978].

Several models based upon different functional relationships between the dependent (hazard rate) and independent variables were evaluated in this analysis. The reader is referred to Breslow and Day [1987] for a full description of the Poisson assumption, and of the techniques used to fit these models. Copies of the GLIM and NLIN programs used to fit these models will be provided upon request.

The following functional forms were fitted to the data:

Categorical-

$$\lambda=\exp(\alpha + \sum_j(\theta_j W_j)+ \Gamma X + \delta Y + \sum_k(Z_k\beta_k))$$

Exponential (log-linear) -

$$\lambda=\exp(\alpha + \sum_j(\theta_j W_j)+ \Gamma X + \delta Y + \beta Z)$$

Linear -

$$\lambda=\alpha + \sum_j(\theta_j W_j)+ \Gamma X + \delta Y + \beta Z$$

Power -

$$\lambda=\exp(\alpha + \sum_j(\theta_j W_j)+ \Gamma X + \delta Y)*([Z+1]^{\beta})$$

Additive Relative Rate -

$$\lambda=\exp(\alpha + \sum_j(\theta_j W_j)+ \Gamma X + \delta Y)*[1 + \beta Z]$$

Where: λ represents the hazard rate, W_j represents the calendar-year groups, X represents hispanic ethnicity, Y represents age, Z represents cumulative cadmium exposure[2], α represents the intercept, θ_j represents the coefficients for the calendar-year groups, Γ represents the regression coefficient for hispanic ethnicity, δ represents the regression coefficient for age, β represents the regression coefficient for cumulative cadmium exposure, j indexes the calendar-year groups, and k indexes the exposure groups.

The linear model and additive relative rate models are similar in functional form to the "absolute" and "relative risk" models fitted in the risk assessment performed by OSHA [1990]. In fitting the linear model, it was necessary to restrain the boundaries of the model such that it did not produce negative predicted values for the hazard rate. This was accomplished in GLIM by using an adaptation of a program described by Wacholder [1986] for binomial regression. The restriction on the boundary space for this model implies that the usual confidence intervals based upon the normal approximation may be misleading. The exponential model is the functional form that is most often used in epidemiologic analyses and the power function model has been commonly used in epidemiologic risk assessment.

---

[2]For the categorical model Z represents the cumulative dose categories, which are indexed by k.

Calendar-year and hispanic ethnicity were modeled as categorical variables; whereas, age was modeled as a continuous variable in these analyses by using the midpoint of each age category (i.e.,35, 52, 57, 62, 67, 72, and 80 years) to represent the rates for each data point in the model. Modeling the data with categorical variables for age did not appreciably improve the deviance of the model, or change the coefficient for cumulative cadmium exposure. Thus, models using a continuous variable for the age categories were judged to be appropriate for this assessment.

The four cumulative exposure groups were generally modeled as a continuous variable in this analysis, although exponential models using categorical variables to represent the exposure groups were also fitted for comparison purposes. The median dose for the person-years in each of the four cumulative exposure categories was approximated by running the lifetable analysis program with successively finer dose categories and identifying the cumulative exposure at which approximately 50% of the person-years in the category were above and below this value. The median mg/M^3-days exposure estimates derived from the life-table analyses were converted into units of $\mu g/M^3$-years by multiplying by 1,000 and dividing by 365. The median exposure estimates that were used in the models to represent the rates for the four exposure categories were 795, 2,466, 5,699 and 10,836 $\mu g/M^3$-years for the unlagged analysis, and 740, 2,948, 5,753, 10,575 $\mu g/M^3$-years for the 5-year lagged analysis.

Confounding and potential interactions between cadmium exposure and the other covariates were evaluated in all of the models. The significance of the addition of a quadratic term for cumulative cadmium exposure was also evaluated. The statistical significance of the parameters was tested by fitting the model with and without the inclusion of the parameters and computing the change in deviance which has an approximate chi-square distribution with degrees of freedom equal to the number of parameters. This test statistic is equivalent to a likelihood ratio test. Confounding was assessed by fitting the models with and without the parameters representing the covariates and monitoring the change in the parameter representing cadmium exposure. Parameters representing all of the potential confounders (age, calendar-year, and hispanic ethnicity) were included in the final models, because deletion of these parameters was found to alter the parameter estimates for cadmium exposure and the inclusion of these variables did not appear to destabilize the models in terms of increasing the variability of the exposure coefficients.

Asymptotically the deviance derived from these models has a chi-square distribution with degrees of freedom equal to the number of data points minus the number of parameters, and may be used to test the goodness of fit of the models. It should be noted, however, that given the small study size the assumption that the deviance has an approximate chi-square distribution may not be tenable. Thus, the deviance should probably not be viewed as a formal goodness of fit statistic in this analysis, but rather as a useful measure for comparing the models.

Finally, the Cox proportional hazards model [Cox 1972] was fitted in order to check the assumptions inherent in the Poisson regression models. Cumulative exposure to cadmium and calendar year were treated as time dependent covariates in these analyses, while hispanic ethnicity was treated as a fixed variable. Age was used as the time dimension in the Cox model, and thus age did not need to be included in the regression model. Functional forms equivalent to the Poisson regression models described above were fitted with the Cox model using the BMDP 2L program [BMDP 1985], except for the linear model (which cannot be fitted with the Cox model).

Excess Risk Estimation

Estimates of excess lifetime risk of dying from lung cancer for varying levels of cadmium exposure were developed based upon a method described by Gail [1975] that accounts for the influence of competing risks. For this analysis it was assumed that workers were exposed to a constant cadmium concentration for 45 years between the ages of 20 and 65, and that the life expectancy is 74 years. Death rates from 1984 for U.S. males (all races) for lung cancer and all causes were used as the background rates for this estimation procedure.

The formula used to estimate lifetime excess lung cancer risk may be expressed mathematically as:

$$\sum_{i=20}^{74} (RR_i - 1)\, q_l(i)\, \exp\left[-\sum_{j=20}^{i} \{ (RR_j - 1) q_l(j) + q_a(j) \} \right]$$

Where RR_i is the rate ratio estimate from the model, $q_1(i)$ represents the background age specific lung cancer rate, $q_a(i)$ represents the background age specific mortality for all causes, and i indexes age.[3]

For each year the cumulative cadmium exposure at the midpoint of that year was used for this algorithm. For example, if an individual was exposed to a constant 8-hour time-weighted average (TWA) exposure level Z beginning at age 20 then at age 24, the cumulative exposure would be 4.5*Z, and after age 65 the cumulative exposure would be 45*Z. When the risk estimates were derived from the lagged analysis, the cumulative exposure estimates used to estimate excess risk were also lagged.

Excess risk estimates were derived for 45 years of exposure to varying 8-hour TWA levels. "Continuous" exposure estimates were also presented which were estimated based on the assumption that workers in this study were exposed for approximately eight hours per day and 240 days per year (i.e., TWA*8/24*240/365). The continuous exposure may be thought of as the equivalent for exposures occurring in the general environment, which are generally constant during the day and year.

[3]A computer program to compute excess lung cancer risk was kindly provided to us by Elizabeth Grossman from OSHA.

RESULTS

Of the 625 white males in the cohort, a total of 606 workers contributing 16,898 person-years met our selection criteria and were thus included in this analysis. Vital status was successfully ascertained for approximately 98% of this cohort. One hundred sixty-two deaths were identified and the extended follow-up resulted in the identification of eight lung cancer deaths that were not included in the previous report by Thun et al. [1985].

Life-Table Analysis

The results from the life-table analysis for cancers of the trachea, bronchus, and lung (henceforth lung cancer) stratified by cumulative cadmium exposure, time since first exposure, calendar year, age, and hispanic ethnicity are presented in Table 1. Mortality from lung cancer was not quite significantly ($p=0.076$, 2 tails) elevated for the entire cohort (SMR=149, 95%CI=95,222). However, among non-Hispanics lung cancer mortality was significantly ($p<0.01$) elevated (SMR=211, 95%CI=131,323). Lung cancer mortality was less than expected among hispanics (SMR=49, 95%CI=10,143), as would be expected given that the referent rates used were for U.S. white males.

Lung cancer mortality appeared to increase with cumulative exposure to cadmium, and was significantly elevated ($p<0.05$) in the highest exposure group ($\geq$ 2,921) for the combined cohort (SMR=272) and for the three highest exposure groups among non-Hispanics. A significant ($p<0.05$)

excess of lung cancer mortality was also observed among workers in the longest time since first exposure category (≥20 years) for the combined cohort (SMR=161) and for non-Hispanics (SMR=233). Finally, the lung cancer excess among the combined cohort and non-Hispanics was largely evident among older workers (i.e., 70-74 and ≥ 75 years), and there did not appear to be a consistent trend in lung cancer mortality with calendar year.

Poisson Regression

The results from fitting the various functional forms of the Poisson regression model are presented in Table 2 for the unlagged analysis and in Table 3 for the 5-year lagged analysis. The regression coefficients representing cumulative exposure to cadmium were statistically significant ($p<0.05$, 2 tails) based on the log likelihood ratio statistic in all of the models examined, except for the linear models which fit the data poorly. Generally, the results from these models indicate a significant dose-response relationship between lung cancer risk and cumulative cadmium exposure. It is evident from Tables 1 and 2 that lagging the exposures by 5 years only slightly increased the magnitude of the cadmium exposure parameter estimates. Lagging the exposures for longer than 5 years (i.e., 10, 15 and 20 years) reduced the magnitude of the exposure parameters and thus decreased the likelihood of the model (i.e. worsened the fit). Thus, 5 years was chosen as the most appropriate lag period for this analysis.

The relationship between cumulative cadmium exposure and the lung cancer hazard rate predicted from these models are illustrated in Figure 1 for the lagged analysis. The data points from the categorical model were plotted on this graph to illustrate the degree of correspondence between these point estimates and the curves derived from the other models.

The linear models provided by far the worst fit to the data as judged by the model deviances for either the unlagged or lagged analyses. The reason for this lack of fit may be clearly seen in Figure 1, where the linear model misses the lowest dose point. The linear model failed to converge when a quadratic term for exposure was added; however, a significant interaction between age and exposure was observed in this model. Thus, the linear model was rejected for assessing risk based upon lack of fit to the data.

Although the power model yielded the lowest deviance, this model was also not chosen for this risk assessment. Reliance on the power model would have resulted in extremely large risk estimates. This is because, as can be seen in Figure 1, the baseline rates derived from this model are unreasonably low. For example, based on the power model, the lung cancer rate for white males, age 70 in 1970–1979, is estimated to be 0.0004, which is approximately two orders of magnitude lower than the corresponding lung cancer rate in the U.S. population. Since the background rates are the denominator for the rate ratios, the power model results in unreasonably high rate ratios (and predicted risks).

Of the remaining two models, the additive relative rate model yielded a slightly better fit (i.e., lower deviance) than the exponential model and this model was chosen as the best functional form for this assessment. The model failed to converge on a solution for the quadratic cumulative cadmium exposure term, or for all of the interaction terms except for the term representing the interaction with hispanic ethnicity, which was not statistically significant. The parameters representing hispanic ethnicity, age, and cumulative cadmium exposure were all found to be statistically significant based on the likelihood ratio test; whereas, the parameters representing calendar year were not statistically significant. Dropping age from the model had a large effect on the coefficient for cumulative cadmium exposure; whereas, dropping calendar year or hispanic ethnicity had only a slight effect on the exposure coefficient. Thus, age appeared to be a strong confounder, and hispanic ethnicity and calendar time appeared to be only weak confounders in this analysis.

Cox Proportional Hazards Model

As in the Poisson regression analysis described above, the additive relative risk form of the Cox model fit the data better than the exponential form, but not as well as the power function form. A comparison of the results from fitting the additive relative rate functional forms of the Cox proportional hazards and Poisson regression models to the 5-year lagged data is presented in Table 4. It is evident from this table that the Cox model coefficients for cumulative cadmium exposure are somewhat less than the corresponding coefficients from the

Poisson regression model. This difference may reflect errors introduced by the modeling of the exposure categories in the Poisson regression analysis. However, it is interesting to note that the calendar-year coefficients were also similarly reduced in the Cox analysis relative to the Poisson regression analysis, which may indicate that the difference in the exposure coefficients may be related to a change in the baseline rate used in these models.

Excess Risk Estimation

Based upon the analyses described above, the results from modeling using the additive relative rate functional form of the models was judged to be the most appropriate form for risk estimation purposes. Using this functional form estimates were derived from both the Poisson regression and Cox proportional hazards models. These excess risk estimates are presented in Table 5 as a function of the 8-hour TWA and corresponding "continuous" dose assuming 45 years of exposure to these concentrations between the ages of 20 to 65. The excess risk estimates derived from the Poisson regression model were approximately twice as high as the estimates derived from the Cox model.

DISCUSSION

The excess in mortality from lung cancer observed in this study is consistent with our previous reports on this cohort [Lemen et al. 1976, Thun et al. 1985], recent studies of Swedish [Elinder 1985, Jarup et al. 1990], and British cadmium workers [Kazanatsis et al. 1988]. The extended period of follow-up resulted in the identification of eight additional lung cancer cases and a slightly stronger overall estimate of lung cancer risk (SMR=149) for workers employed after January 1, 1926, than previously reported (SMR=147) by Thun et al. [1985]. Although overall lung cancer mortality was not quite significantly elevated (p=0.076, 2 tails) in our analysis, statistically significant elevations in mortality were observed for non-Hispanic workers and for workers in the highest cadmium exposure and longest time since first exposure groups.

The significant dose-response relationship between lung cancer mortality and cumulative cadmium exposure observed in the previous report by Thun et al. [1985] was also evident in the results presented in this paper. The availability of this information presented a unique opportunity for modeling the dose-response relationship and for producing quantitative estimates of risk for varying levels of exposure to cadmium. Based upon these analyses, the lifetime excess lung cancer risk at the current OSHA standard for cadmium fumes of 100 $\mu g/M^3$ is estimated to be approximately 50 to 111 lung cancer deaths per 1,000 workers exposed to cadmium for 45 years. OSHA [1990] has recently proposed revising its standard to 1 or 5 $\mu g/M^3$ for all cadmium compounds, which based on this

analysis corresponds to a risk of 0.5–1.2 and 2.6–6.0 lung cancer deaths per 1,000 workers respectively.

At least two important sources of bias should be considered when interpreting the results from this analysis. The first is the potential influence of cigarette smoking on our study findings. The potential for confounding by cigarette smoking was greatly reduced by the modeling procedures used in this paper. These procedures rely on internal comparisons within the cohort as opposed to the external comparisons made with the U.S. population in the SMR analysis. In order for smoking to confound this analysis one would have to propose that smoking habits vary between the exposure categories used in the analysis, which seems unlikely.

The potential influence of smoking was further reduced in this analysis by the inclusion of a parameter for hispanic ethnicity in the regression models. Hispanics in the Southwest experience lower death rates from lung cancer [Savitz 1986, Key 1981, Samet 1980], and smoke fewer cigarettes per day than non-Hispanics [Samet et al. 1982, Mitchell 1990]. Hispanics were observed to have a deficit in mortality from both lung cancer and cardiovascular diseases (SMR=41) in this analysis relative to the U.S. population. Hispanic ethnicity may therefore be thought of as a surrogate for cigarette smoking in this analysis. In actuality, the inclusion of hispanic ethnicity had little effect on the estimated cadmium exposure coefficients in our regression models suggesting that

hispanic ethnicity (and hence smoking) was not a strong confounder in this analysis. Thus, given the internal nature of this analysis and the control of ethnicity in the models, it seems unlikely that residual confounding by cigarette smoking would have a large influence on our findings.

The other major potential source of bias in this analysis is confounding by exposure to arsenic, a potent respiratory carcinogen [Pinto et al. 1977]. The potential for bias from arsenic exposure was reduced, but not eliminated, by the exclusion of workers first employed prior to January 1, 1926, when the plant was an arsenic smelter. A detailed analysis of the possible influence of arsenic exposure on respiratory cancer mortality at this facility was presented in the previous report on this cohort [Thun et al. 1985]. Based on this analysis, Thun et al. [1985] estimated that no more than 0.77 of the 16 lung cancer deaths observed among workers first employed after January 1, 1926 could be related to arsenic exposure. Thus, it would appear unlikely that arsenic exposure could fully explain the excess of lung cancer observed in this or the previous analysis of this cohort.

The influence of arsenic exposure on our findings may also have been limited by the nature of the study facility, and the analytic procedure employed in this investigation. Arsenic exposure occurs at this facility primarily in the departments that process incoming feed materials. Other stages of the process are housed in separate buildings where workers are exposed to cadmium but not arsenic. Entry-level workers were generally

assigned to the departments with potential for arsenic exposure; whereas, senior workers were able to bid out of these departments into areas with only cadmium exposures. Thus, workers with short tenure and low cumulative cadmium exposures might be expected to have approximately the same potential for cumulative arsenic exposure as workers with longer tenure and higher cumulative cadmium exposure. If this is the case, then the internal analytic procedures used in this investigation should have reduced any potential bias related to arsenic exposure, since all of the exposure groups modeled would have similar cumulative arsenic exposures.

Several other potential sources of uncertainty in this analysis should be considered. The first is related to the choice of the model for risk estimation. Although the power model provided the best fit to this data set, this model was rejected because it appeared to produce unreasonably high risk estimates due to an unusually low background rate. The exponential model evaluated in this paper also produced nearly as good a fit as the additive relative risk model which was used to estimate risk in this assessment. However, as can be seen from Figure 1, the exponential model predicted rates are very similar to the additive relative rate model, and the use of the exponential model would yield risk estimates that are just slightly lower than our reported estimates.

The cadmium exposure coefficients derived from the Cox proportional hazard model were approximately half as large as the coefficients derived from the Poisson regression models. One possible explanation for this discrepancy is that error was introduced in the modeling of the dose

categories in the Poisson regression analyses. However, the coefficients representing the calendar year categories were also reduced in the Cox model relative to the Poisson regression model (see Table 4). This fact suggests that the discrepancy may be related to the baseline hazard, which was paramaterized in the Poisson regression models and unspecified in the Cox model. Thus, it is unclear whether the difference in the exposure coefficients derived from these models may be attributed to bias in the modeling of dose in the Poisson regression model, or in the different assumptions made regarding the background hazard rate in the two models. We have therefore chosen to present risk estimates derived from both models in this paper.

Another important source of uncertainty in our analysis are the estimates of exposure used in this analysis. In addition, while cadmium exposure was treated as a known constant in this analysis, clearly there was some error in our estimation of exposure in this study. However, it is not possible to estimate the error in our exposure estimates or the degree to which these errors might influence our risk estimates.

Finally the choice of a 5-year "lag" period for this analysis was based upon our empirical determination that this assumption maximized the dose-response. This may be viewed as a somewhat arbitrary decision, and the assumption of other lags would alter our findings. The use of this lag only slightly increased the dose-response relationship and our estimates of risk in this analysis. Unfortunately, there is no biologic information available to aid us in selecting the appropriate lag period.

OSHA [1990] in a recently published proposal for a revised cadmium standard presented the results from two quantitative assessments of the risk of lung cancer associated with occupational exposure to cadmium. In the first analysis the quantal form of the multistage model was fitted to the rat inhalation bioassay data reported by Takenaka et al. [1983]. The second analysis fitted additive (form similar to our linear model) and relative risk (form equivalent to our additive relative rate model) models to the summary SMRs reported in the analysis by Thun et al. [1985]. It was suspected that the latter analysis might have underestimated risk, since it relied on modeling SMRs which are generally expected to underestimate risk due to the "healthy worker effect," and particularly in this case because of the large percentage of hispanics in this cohort. The analysis presented in our paper is not subject to the "healthy worker effect," since it was based on internal comparisons, and not with comparisons with the U.S. population. The excess risk estimates produced by our analysis are approximately two to four times greater than the estimates produced by OSHA's relative risk model of the Thun study, and 1/2 to 1/4 of the estimates produced by OSHA's multistage modeling of the animal bioassay data. Our risk estimates are actually remarkably close to OSHA's estimates given differences in the data sets, modeling techniques, and species used for these assessments.

CONCLUSIONS

In summary, we observed an excess in lung cancer mortality in this study of cadmium-exposed workers, which is consistent with previous reports in animals and man. The lung cancer excess was statistically significant among non-Hispanic workers, and among workers with high exposures and long time since first exposure.

A statistically significant dose-response relationship between cumulative cadmium exposure and lung cancer risk was evident in the models fitted to the cohort data. Although our risk estimates are somewhat higher than OSHA's assessment based on our previous report, our estimates are remarkably close considering differences in the methodology, and data used for these assessments.

Table 1

Lung Cancer Standardized Mortality Ratios (SMR), Observed (OBS), and Expected (EXP) Deaths Stratified by Cumulative Exposure to Cadmium (EXPOSURE), Time Since First Exposure (Latency), Age at Risk (AGE), Calendar Year at Risk (YEAR) and hispanic ethnicity.

	NON-HISPANIC			HISPANIC[a]			COMBINED		
CATEGORY	OBS	EXP	SMR	OBS	EXP	SMR	OBS	EXP	SMR
OVERALL	21	9.95	211**	3	6.12	49	24	16.07	149
EXPOSURE[b]									
≤ 584	1	3.35	29	1	2.38	42	2	5.73	34
585–1460	7	2.64	265*	0	1.64	0	7	4.28	163
1461–2920	6	1.55	386*	0	1.20	0	6	2.75	217
≥2921	7	2.41	290*	2	0.90	223	9	3.30	272*
LATENCY (YEARS)									
< 10	0	0.41	0	1	0.28	363	1	0.69	145
10–19	2	1.41	142	0	1.00	0	2	2.41	83
≥ 20	19	8.13	233**	2	4.84	41	21	12.97	161*
YEAR									
1940–1959	2	0.89	225	0	0.38	0	2	1.26	158
1960–1969	5	2.24	223	1	1.27	78	6	3.51	171
1970–1979	10	4.43	228*	2	2.88	69	12	7.30	164
≥ 1980	4	2.39	167	0	1.59	0	4	3.98	101
AGE (YEARS)									
< 50	0	0.78	0	1	0.50	201	1	1.28	78
50–54	2	1.01	198	0	0.67	0	2	1.68	118
55–59	1	1.61	62	2	1.00	200	3	2.60	115
60–64	5	2.20	227	0	1.20	0	5	3.40	146
65–69	4	2.12	188	0	1.13	0	4	3.25	123
70–74	5	1.37	366*	0	0.84	0	5	2.20	227
≥ 75	4	0.87	547*	0	0.79	0	4	1.66	241

a. U.S. rates for white males were used as the referent group for hispanic and non-Hispanic males in this analysis.
b. Milligrams cadmium per cubic meter of air - days

*p < 0.05 (two-tails)
**p < 0.01 (two-tails)

Table 2
Results from Poisson Regression Models Fitted to Unlagged Data File.[a]

EQUATION FOR RATE RATIO (RR) AS A FUNCTION OF CADMIUM EXPOSURE (Z)	DEGREES OF FREEDOM	DEVIANCE	EXPOSURE PARAMETER ESTIMATE(β)	STANDARD ERROR
1. CATEGORICAL[b] $RR=\exp(Z_k\beta_k)$	192	77.76	1.49 1.78* 1.82*	0.81 0.83 0.80
2. LINEAR[c] $RR=(\lambda_o + Z\beta)/\lambda_o$	194	96.84	4.90E-08	2.41E-08
3. EXPONENTIAL $RR=\exp(Z\beta)$	194	81.88	1.03E-04*	5.00E-05
4. POWER FUNCTION $RR=[Z+1]^{\beta}$	194	79.50	0.56*	0.23
5. ADDITIVE RR $RR=1+Z\beta$	194	80.69	4.30E-04*	4.45e-04

a. All models include categorical variables to control for calendar year and hispanic ethnicity, and a continuous variable to control for age.
b. The parameter estimates for the categorical model represent the effect of each category relative to the low dose category (<584 mg/M^3-days).
c. Rate ratios derived from the linear model are a function of the background rate (λ), which varies with age, year and hispanic ethnicity. Standard errors for the linear model may be unreliable due to the constraints imposed on this model.

Table 3

Results from Poisson Regression Models Fitted to Five-Year Lagged Data File.[a]

EQUATION FOR RATE RATIO (RR) AS A FUNCTION OF CADMIUM EXPOSURE (Z)	DEGREES OF FREEDOM	DEVIANCE	EXPOSURE PARAMETER ESTIMATE(β)	STANDARD ERROR
1. CATEGORICAL[b] $RR=\exp(Z_k\beta_k)$	193	77.81	1.49 1.80* 1.85*	0.81 0.83 0.80
2. LINEAR[c] $RR=(\lambda_o + Z\beta)/\lambda_o$	213	97.71	8.03E-08	3.95E-08
3. EXPONENTIAL $RR=\exp(Z\beta)$	213	82.29	1.15E-04*	5.14E-05
4. POWER FUNCTION $RR=[Z+1]^{\beta}$	213	79.28	0.58**	0.23
5. ADDITIVE RR $RR=1+Z\beta$	213	80.70	6.12E-04*	6.65E-04

a. All models include categorical variables to control for calendar year and hispanic ethnicity, and a continuous variable to control for age.
b. The parameter estimates for the categorical model represent the effect of each category relative to the low dose category (<584 mg/M^3-days). The person-years from the "0" dose group were dropped from this model, but were included in the other models. Inclusion of these person-years would cause this model to fail to converge, since there were no deaths in the "0" dose category.
c. Rate ratios derived from the linear model are a function of the background rate (λ), which varies with age, year and hispanic ethnicity. Standard errors for the linear model may be unreliable due to the constraints imposed on this model.

Table 4

Parameter Estimates, Standard Errors from the Poisson Regression and Cox Proportional Hazard Additive Relative Rate Models Based on the 5-Year Lagged Data Analyses.

	POISSON REGRESSION		COX PROPORTIONAL HAZARDS[c]	
PARAMETER	PARAMETER ESTIMATE	STANDARD ERROR	PARAMETER ESTIMATE	STANDARD ERROR
Intercept	-1.261E+01	1.376	-	-
Age	8.379E-02**	2.103E-02	-	-
Hispanic[a]	-1.356*	6.201E-01	-1.341*	6.238E-01
Calendar Year[b]				
1960-1969	4.792E-01	8.380E-01	2.615E-01	8.196E-01
1970-1979	7.201E-01	8.215E-01	4.751E-01	7.887E-01
≥ 1980	3.444E-01	9.539E-01	1.380E-01	9.154E-01
Cd Exposure	6.121E-04*	6.647E-04	2.632E-04*	2.379E-04

*$p < 0.05$ (two tails) based on the likelihood ratio test
**$p < 0.01$ (two tails) based on the likelihood ratio test

a. Represents the effects of hispanic ethnicity in reference to non-Hispanics.
b. The calendar year parameters presented represent the effect of the calendar periods in reference to the baseline period of 1940-1960.
c. The intercept for the Cox model is undefined, and age was controlled for by matching on survival to the same age.

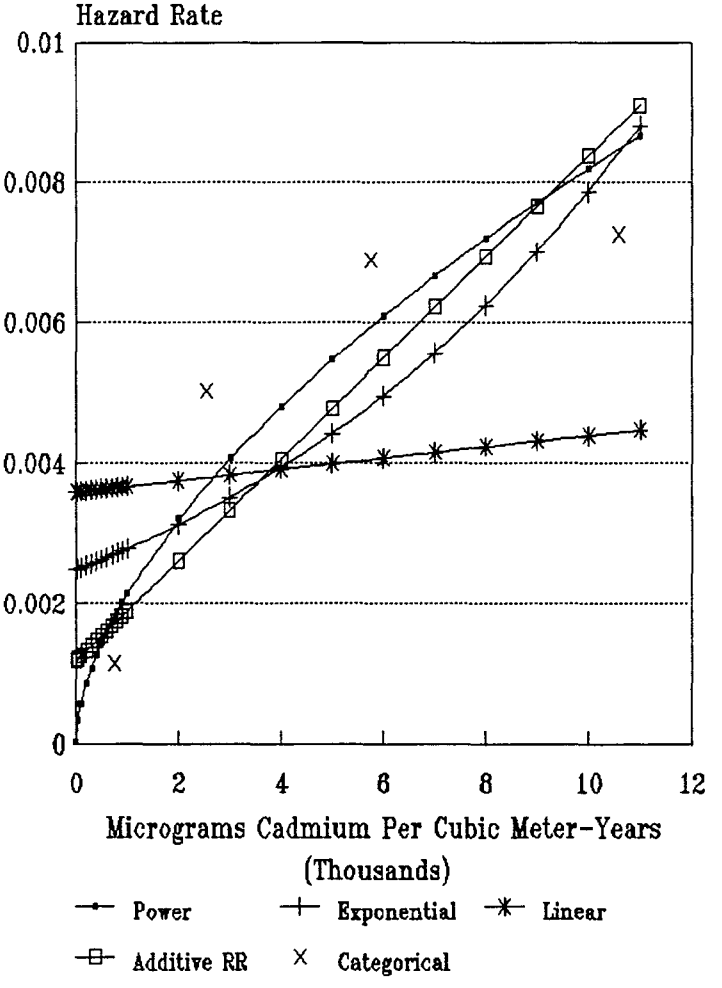
Table 5

Estimates of Excess Risk per 1,000 Workers Based on the Poisson Regression and Cox Proportional Hazards Additive Relative Rate Models (Lagged 5 years) Assuming 45 Years of Exposure to Cadmium and Varying the Time-Weighted Average (TWA) Exposures and Continuous (CONT) Exposures.[a]

EXPOSURE ($\mu g/M^3$)		EXCESS RISK ESTIMATES (PER THOUSAND WORKERS)	
TWA[b]	CONT[c]	POISSON MODEL	COX MODEL
1	0.22	1.2	0.5
3	0.66	3.6	1.5
5	1.10	6.0	2.6
7	1.53	8.4	3.6
10	2.19	11.9	5.2
20	4.38	23.7	10.3
40	8.77	46.5	20.4
50	10.96	57.7	25.4
100	21.92	110.9	49.9
200	43.84	205.2	96.4

a. Risk estimates are based on the results from the 5 year lagged analysis.
b. The TWA multiplied by 45 years is equivalent to the cumulative exposure estimates that were used for fitting the regression models.
c. Continuous exposure estimates were estimated based on the assumption that workers in this study were exposed for approximately 8 hours per day and 240 days per year (i.e. CONT=TWA*8/24*240/365). The continuous exposure may be thought of as the equivalent for exposures occurring in the general environment, which are generally constant during the day and year.

Figure 1
Model Estimated Hazard Rates as a
Function of Cumulative Cadmium Exposure



*Based on Lagged 5 Year Analysis
for White Males, Age 70, 1940-1960.

REFERENCES

Armstrong BG, Kazantzis G (1983). The mortality study of cadmium workers. Lancet; 1:1425-1427.

Baker RJ, Nelder JA (1978). Generalized linear interactive modeling (GLIM). Release 3, Numerical Algorithms Group, Oxford, England.

BMDP (1985). BMDP Statistical Software, 1985 Printing. University of California Press, Berkeley, California.

Breslow NE, Day NE (1987). Statistical Methods In Cancer Research. Volume II-The Design and Analysis of Cohort Studies. Oxford University Press, United Kingdom.

Checkoway H, Pearce N, Hickey JLS, Dement JM (1990). Latency analysis in occupational epidemiology. Arch Env Health; 45(2); pp. 95-100.

Cox DR (1972). Regression models and life tables. J Royal Stat Society; Series B 34; pp. 187-202.

Elinder CG, Kjellstrom T, Hogstedt C, Andersson K, Spang G (1985). Cancer mortality of cadmium workers. Brit J Ind Med; 42:651-655.

EPA, Office of Health and Environmental Assessment (1984). Updated assessment of mutagenicity and carcinogenicity assessment of cadmium. EPA 600/8-83-025B, Washington D.C., April 1984,

Gail M (1975). Measuring the benefits of reduced exposure to environmental carcinogens. J Chronic Dis: 28;135-147.

Jarup L, Elinder CG, Spang G (1990). Lung cancer mortality in cadmium exposed battery workers. Abstract accepted at Congress of Occupational Medicine, Montreal, September 1990.

Jennrich RI, Ralston ML (1978). Fitting nonlinear models to data. Technical Report No. 46. BMDP Statistical Software. Los Angeles, CA.

Kazantzis G, Lam TH, Sullivan KR (1988). Mortality of cadmium-exposed workers, a five-year update. Scand J Work Environ Health; 14:220-223.

Key CR (1981). Cancer incidence and mortality in New Mexico, 1973-77. In: Young JL, Percy CL, Asire AJ (eds). Surveillance, Epidemiology and End Results: Incidence and Mortality Data, 1967-77. pp. 489-595. Washington D.C. Govt Printing Office.

Lemen RA, Lee JS, Wagoner JK, Blejer HP (1976). Cancer mortality among cadmium production workers. Ann NY Acad Sci; 271:273-279.

Mitchell BD, Stern MP, Haffner SM, Hazuda HP, Patterson JK (1990). Risk factors for cardiovascular mortality in Mexican Americans and Non-Hispanic Whites. Am J Epid; 131: 423-433.

NIOSH (1984). Current Intelligence Bulletin #42: Cadmium. Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 84-116.

NIOSH (1990). NOSH testimony on the Occupational Safety and Health Administration proposed rule on occupational exposure to cadmium: 29 CFR Part 1910, Docket No. H-057a, July 17, 1990. NIOSH policy statements. Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health.

Oldiges H., Hochrainer D., Glaser U (1989). Long-term inhalation study with Wistar rats and four cadmium compounds. J Toxicol Environ Chem; 23:35-40.

OSHA (1990). Occupational exposure to cadmium; proposed rule. Federal Register 29 CFR Part 1910, pp. 4052-4147, Feb. 6, 1990.

Pinto SS, Enterline PE, Henderson V, Varner MO (1977). Mortality experience in relation to a measured arsenic trioxide exposure. Environ Health Perspect; 19:127-130.

Rothman KJ (1981). Induction and Latent Periods. Am J Epi; 114(2):253-259.

Samet JM, Schrag SD, Howard CA, Key CR, Pathak DR (1982). Respiratory disease in a New Mexico population sample of Hispanic and non-Hispanic whites. Am Rev Resp Dis; 125:152-157.

Samet JM, Key CR, Kutvirt DM, Wiggins CL (1980). Respiratory disease mortality in New Mexico Indians and Hispanics. AJPH; 70: 492-497.

Savitz D (1986). Changes in Spanish surname cancer rates relative to other whites, Denver areas, 1969-71 to 1979-81. AJPH; 76:1210-1215.

SAS Institute Inc (1987). SAS/STAT Guide for Personal Computers, Version 6 Edition. Cary, NC., pp. 675-712.

Smith TJ, Anderson RJ, Reading JC (1980a). Chronic cadmium exposures associated with kidney function efforts. Am J Ind Med; 1: 319-337.

Smith TJ, Ferrell WC, Varner MO, Putnam RD (1980b). Inhalation exposure of cadmium workers: Effects of respirator usage. Am Ind Hyg Assoc J; 41:624-629.

Sorahan T (1987). Mortality from lung cancer among a cohort of nickel cadmium battery workers: 1946-84. Br J Ind Med; 44:803-809.

Sorahan T, Waterhouse JA (1983). Mortality study of nickel-cadmium battery workers by the method of regression models in life tables. Br J Ind Med; 40:293-300.

Takenaka S, Oldiges M, Konig H, Hochrainer D, Oberdorster G (1983). Carcinogenicity of cadmium chloride aerosols in W rats. JNCI; 70:367-373.

Thun MT, Schnorr TM, Smith AB, Halperin WE, Lemen RA (1985). Mortality among a cohort of US cadmium production workers—an update. JNCI; 74:325-333.

Thun MJ, Schnorr TM, Halperin WE (1986). Mortality from lung and prostatic cancer in U.S. Cadmium Workers. Presented at the Workshop on Cadmium and Cancer, Oxford, England, September 29-October 1, 1986.

Wacholder S (1986). Binomial regression in GLIM: Estimating Risk Ratios and risk differences. Am J Epi; 123(1): 174-184.

Waxweiler RJ, Beaumont JJ, Henry JA, Brown DP, Robinson CF, Ness GO, Wagoner JK, Lemen RA (1983). A modified life-table analysis system for cohort studies. J Occup Med; 25:115-124.