Exposure to Beryllium and Occurrence of Lung Cancer: A Reexamination of Findings From a Nested Case–Control Study

Paul S. Levy, ScD
H. Daniel Roth, PhD
David C. Deubner, MD, MPH

Objective: *Our aim was to reanalyze a nested case–control study of beryllium and lung cancer because we identified analysis and study design issues that could have led to the elevated odds ratios obtained in the study.* ***Methods:*** *We reanalyzed the data using nontransformed exposure metrics instead of log-transformed metrics used in the publication. We identified and examined effects on estimated odds ratios of imbalances between cases and controls caused by the control selection method.* ***Results:*** *This reanalysis found no elevated odds ratios for any exposure variable.* ***Conclusions:*** *Our conclusions differ from the authors' interpretation that the findings are due to a causal relationship between beryllium exposure and lung cancer. Our alternative explanation is that they may be due to methodological problems that could have been controlled by closer matching of controls to cases.* ***Clinical Significance:*** *This study challenges conclusions made from a large case–control study concerning beryllium–lung cancer associations. Occupational medicine practitioners may want to integrate findings from this study into advice they give beryllium-exposed workers concerned about lung cancer.* (J Occup Environ Med. *2007;49: 96–101*)

Beryllium has been labeled a carcinogen in humans by both the International Agency for Research in Cancer and the National Toxicology Program of the National Institute of Environmental Health Sciences. The major epidemiologic evidence for this characterization comes primarily from two articles emanating from a single retrospective cohort mortality study conducted by the National Institute on Occupational Safety and Health (NIOSH) in 9225 workers employed in seven U.S plants between 1940 and 1969 with mortality follow up through 1988. The first article used employment duration as the exposure metric and found an overall low level but statistically significant standard mortality ratio (SMR) for lung cancer equal to 1.26.[1] Like with prior studies, there was an inverse exposure response effect, as the SMR decreased with increasing duration of employment,[2] and a large proportion of the lung cancers were in workers employed for less than 1 year. In our subsequent reexamination of the data from that study using what we felt were more appropriate population rates and adjustment techniques, we obtained an SMR equal to 1.04 with 95% confidence intervals between 0.92 and 1.17.[3]

The second article, also by NIOSH investigators, addresses more fully the issue of exposure response.[4] It presents the findings of a nested case–control study, constructed from the same NIOSH cohort study and conducted in a Reading, Pennsylvania, beryllium manufacturing plant, which was the

From RTI International (Dr Levy), Research Triangle Park, North Carolina; Roth Associates, Inc. (Dr Roth), Rockville, Maryland; and Brush Wellman Inc. (Dr Deubner), Elmore, Ohio.

Brush Wellman Inc is a producer of beryllium materials.

Dr Deubner is an employee of Brush Wellman Inc, support for the other two authors was provided by Brush Wellman Inc. through contract with Roth Associates and also to RTI through contracts between Roth Associates and RTI.

Address correspondence to: Paul S. Levy, ScD, RTI International, 3040 Cornwallis Rd., P.O. Box 12194, Research Triangle Park, NC 27709-2194; E-mail: levy@rti.org

DOI: 10.1097/JOM.0b013e31802db595

largest of the seven plants in the original cohort study. In this spinoff study, for 142 lung cancer cases, 710 controls were selected individually matched 5:1 to each case by means of a procedure known as risk set sampling,[5] density sampling,[6] and referred to as incidence density sampling in the article reporting the results of this study.[4]

Because much of the discussion in this report focuses on risk set sampling, we briefly discuss below how risk set sampling is performed to select controls in nested case–control studies, particularly those constructed from occupational cohort mortality studies. A strong motivation and justification for risk set sampling is that, when used along with conditional logistic regression, it is consistent with the way data would be analyzed if the entire cohort was available and proportional hazard regression (a frequently used method for survival data analysis) were used.[5] Risk set sampling, as generally used in occupational nested case–control mortality studies, has the following elements:

1. For each case, its risk set consists of all cohort members whose age at death or when last observed, if alive (called attained age in the NIOSH nested case–control studies) is greater than or equal to that of the case and who are either alive at that age or, if dead, do not have the target disease (eg, lung cancer) as the cause of death. Also, the age when first employed must be less than the age at death of the case.
2. Controls are selected by random sampling from the risk set of the corresponding case. Generally, five controls (if available) are taken as matches for each case.
3. Measurement of exposure in the case is from the age first employed to the age at termination of employment. Measurement of exposure in the matching controls would be from the age at which employment began to the age of employment termination or to the age of death from the target disease of the case, whichever comes first. In other words, exposure measurement in the controls is "benchmarked" to the matching case in the sense that it is measured only to the age of death of the case.
4. When disease latency is considered, the same control group is used and exposure in the case is measured from the age at first employment to the age at which employment ends minus the assumed latency period. For example, if the age at death of the case is 65 years and disease latency is considered to be 10 years, any exposure that occurs after age 55 is not measured. Exposure in the control is measured analogously from the age of first employment to the age at which employment terminates or to the age at death of the case minus the assumed disease latency time, whichever comes first. For example, for a cohort member who is alive at age 70 when last observed and in the risk set of the case described previously, and who began employment at age 45 and terminated it at age 60, exposure would be measured from age 45 to age 55 (the age at death of the case minus the 10-year latency period). Exposure from age 55 to age 60 would be excluded because it falls outside of the latency period of the case.

Using work histories and available industrial hygiene data, the authors of this nested case–control study constructed four exposure metrics (cumulative, maximum, and average beryllium exposures as well as employment tenure). Two of the three additional exposure metrics (maximum and average exposure) furnish information on intensity of exposure; the third provides information on cumulative exposure as does employment tenure, which was the only exposure metric used in the original cohort study. They also extended follow up of the workers through 1992. In their major analyses, they examined relationships between disease status (case vs control) and the natural logarithm of each of the exposure metrics (rather than the original untransformed variable). Their main findings are shown descriptively in Figure 1, which shows for 0-, 10-, and 20-year disease latency the geometric means of each of the four log-transformed metrics for cases and controls. Using conditional

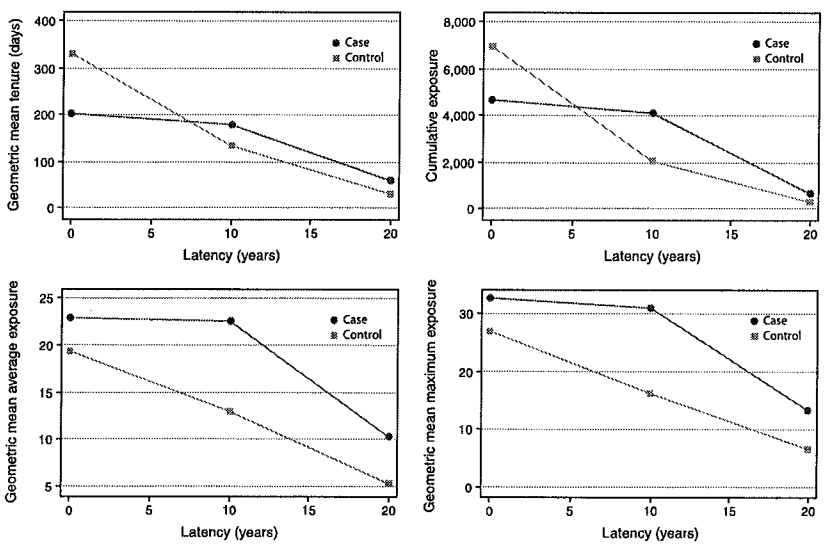

Fig. 1. Geometric means of exposure metrics for each of four exposure metrics used in the National Institute on Occupational Safety and Health nested case–control study by case–control status and latency status.

TABLE 1
Major Findings of Conditional Logistic Regression From NIOSH Nested Case Control Study as Reported in the Original Article[4]‡

Exposure Variable	Exposure Odds Ratios		
	No Latency	10-Yr Latency	20-Yr Latency
Log employment duration (days)	0.91*	1.05	1.05
Log cumulative exposure ($\mu g/m^3$ days)	0.94	1.06*	1.05*
Log average exposure ($\mu g/m^3$)	1.12	1.20†	1.09
Log maximum exposure ($\mu g/m^3$)	1.10	1.20†	1.11*

*$P \leq 0.05$.
†$P \leq 0.01$.
‡The authors' original findings were expressed as regression coefficients; the figures in this table use the exponential transform to express them as odds ratios. This was done for descriptive purposes and does not affect the P values.

logistic regression with the natural logarithm of exposure as the independent variable, they estimated exposure odds ratios (ORs) for each metric both with and without consideration of disease latency, latency was examined by means of exposure lagging,[7] and they did not control for other covariates in their analysis.

From this analysis they found, as shown in Table 1, no significantly elevated exposure odds ratios when disease latency was not considered, but some significantly elevated ORs when 10- and 20-year disease latency were considered. They found no evidence of increase in OR elevation with increase in latency for any of the exposure metrics.

We wanted to explore further the findings in their latency-based analysis of significantly elevated odds ratios without a concomitant increase in odds ratios with increase in latency and requested data from NIOSH of both the nested case–control study and the parent cohort study (which they generously provided). We present here the findings of our investigation.

Materials and Methods

Our reexamination focuses on the methods used in the NIOSH nested case–control study for analysis of exposure–disease associations under considerations of disease latency. Most importantly, we were concerned about the use in that study of the logarithm of the beryllium exposure metrics to represent exposure in their conditional logistic regression models rather than the original variables, because logistic regression methodology permits use of independent variables that can be categorical as well as quantitative and are not confined to normally distributed variables. Their use of this logarithmic transformation required assignment of a value (they chose 0.1 probably because it was close to but still below the minimum value for each of the metrics used) to subjects who had no exposure within the disease latency period. They were required to do this because the logarithm of 0 is not defined and this permitted such subjects to remain in the analysis set. The logarithmic transformation itself alters the shape of any distribution, and the assignment of a particular number (eg, 0.1) to those measurements that are not defined in the transformed distribution may alter it further. Concerned with this possibility, we examined the effect of this transformation on the findings of the NIOSH case–control study.

With these issues in mind and using the data file provided us by NIOSH, we reanalyzed their data using conditional logistic regression[8] (as was used and presented in Table 4 of the NIOSH nested case–control study) for each of the four exposure metrics discussed previously and under each of the two models described subsequently.

Model 1. Case–control status as the dependent variable; logarithm of the exposure metric as the independent variable; subjects having no exposure during the latency period given a value of 0.1 for exposure and included in the analysis set. This was the method used in the NIOSH case–control study.

Model 2. Case–control status as the dependent variable; the untransformed value of the exposure metric as the independent variable; subjects having no exposure during the latency period given a value of 0 for exposure and included in the analysis set.

In the following sections, we show the findings of our reanalysis, confining our presentation to the scenario in which the latency for lung cancer is assumed to be 10 years. This is the scenario that showed the highest odds ratio elevations in the NIOSH case–control study.

Results

Although our analysis was predicated on our assumption that we were using the same data that was used in the NIOSH case–control publication, in a few instances involving conditional logistic regression analysis, we obtained regression coefficients that differed slightly from those published in the original article. In all of these, the means we obtained for the independent and dependent variables were equal to those in the article (when given), so we conjecture that the differences in the regression coefficients are likely due to the fact that we used a different software package for conditional logistic regression than was used by the authors (We used the *clogit* module in Stata Release 9.1[9]; the particular software module they used for their conditional logistic regression was not specified exactly, although they stated they used SAS for their analyses.). These slight discrepancies would have no effect on the interpretations of our findings.

TABLE 2
Regression Coefficients From Conditional Logistic Regression of Case–Control Status Against Beryllium Exposure Metric Lagged 10 Yr; Cells Contain Logistic Regression Coefficients, 95% Confidence Intervals, and *P* Values

Exposure Metric	Natural Logarithm of Exposure Metric Used as Independent Variable; All Cases and Controls Used in Analysis With Those Having 10-Yr Lagged Exposures Equal to Zero Given a Value for the Exposure Variable Equal to 0.1	Original Untransformed Exposure Metric Used as Independent Variable; All Members of Original Control Group Used in Analysis
Employment duration (days)	−0.058 (−0.141 to 0.026) $P = 0.177$	-0.0567×10^{-3} $(-0.013 \times 10^{-2}$ to 0.0152×10^{-3} $P = 0.122$
Log cumulative exposure ($\mu g/m^3$ days)	0.0575 (0.0026 to 0.1124) $P = 0.042$	-8.77×10^{-8} $(-1.17 \times 10^{-6}$ to $9.94 \times 10^{-7})$ $P = 0.874$
Average exposure ($\mu g/m^3$)	0.1977 (0.0847 to 0.3107) $P = 0.001$	-0.0337×10^{-3} $(-0.0413 \times 10^{-2}$ to $0.0347 \times 10^{-2})$ $P = 0.864$
Maximum exposure ($\mu g/m^3$)	0.1811 (0.0780 to 0.2842) $P = 0.001$	0.0100×10^{-2} (−0.0001 to 0.0003) $P = 0.363$

The conditional logistic regression coefficients that we obtained from this 10-year disease latency-based analysis are shown in Table 2. In column 2 of this table, we show the findings using, as was done in the NIOSH nested case–control study, the logarithmic transformation of the exposure metrics and also using a value of 0.1 for those subjects that had no exposure during the latency period (model 1). Although employment duration showed no significant relationship with lung cancer, all three of the other exposure metrics (cumulative exposure, average exposure, and maximum exposure showed significant associations with lung cancer as was stated in the NIOSH case–control study).

The regression coefficients shown in column 3 of this table are based on model 2 and use the untransformed exposure metrics as the covariates with subjects having no latency exposure still included as in model 1 and given a value equal to 0. These regression coefficients all had values very close to 0, which is the value indicating no relationship between exposure to beryllium products and occurrence of lung cancer.

Discussion

The NIOSH occupational retrospective cohort study and its subsequent nested case–control study provide a wealth of information concerning relationships between occupational exposure to beryllium products and mortality from lung cancer. The case–control study supplements the cohort study in that it uses an internal rather than an external comparison group and supplements employment tenure with more refined measures of beryllium exposure based on work histories and industrial hygiene data. Our own reanalysis of the NIOSH nested case–control study, however, has uncovered some issues that we feel pose a challenge to the authors' interpretation of their findings.

The most important issue involves the results of their analysis choice in using the logarithms of the four exposure metrics rather than the original values of the variable to express exposure to beryllium products. Although it is true that the log-transformed variable has a less skewed and somewhat more "bell-shaped" or Gaussian distribution than the original variable, the method of analysis, conditional logistic regression, does not make any distributional assumptions concerning the independent variables and hence does not require a transformation for its use.[8] The use of the log transform has entailed that they either: 1) exclude from their analysis all of the observations that had no exposure eligible to be measured before the cutoff defined by the exposure lagging, because the logarithm of the value, "0," is not definable; or else 2) give the subjects having "0" exposure some small value so that they can be included in the analysis. They chose the latter and gave them a value equal to "0.1." As shown in Table 2, the conditional logistic regression analysis has resulted in point estimates of regression coefficients that, for three of the four exposure metrics, were statistically significant in the direction of an exposure–disease relationship.

The main reasons for differences in the results when the log transform of the exposure metrics are used rather than the original exposures can be explained in large part by the following.

The original untransformed exposure variables have a highly skewed distribution with the mean considerably greater than the median and the minimum value, 0, relatively close to the mean. The logarithm of the exposure variables, on the other hand, with the value "0.1" substituted for "0" have a negatively skewed distribution with the median being greater than the mean and the value, −2.303 (the logarithm of 0), being considerably further away from the mean of

its distribution than is the value "0" from the mean of the original distribution. This is seen in Figure 2, which shows for the exposure metric, maximum exposure, the histogram of the untransformed variable (graph on the left), and that of the log transformed variable (graph on the right). The result is that the log transform gives the cohort members who have no 10-year lagged exposure a more extreme value (further from the mean of the log distribution than the value they had in the original distribution). This gives them greater impact in the analysis based on the log transform than in the original distribution.

Giving a value of 0.1 to subjects having no latency-based exposure would not have had such an impact if the proportion of cases and controls having this value were approximately equal. However, the control selection process resulted in construction of a control group that was not well balanced with the cases on variables reflecting age at death or when last observed, age when employment in the beryllium plant began, and age at termination of employment. These imbalances, as shown in row 5, columns 2 and 3 of Table 3, have led to a much greater percentage of controls (10.2%) than cases (1.4%) having no exposure during the latency period and thus being assigned low values of 0.1 in the NIOSH analysis. This, in turn, has contributed to 10-year latency-based logistic regression coefficients (and subsequent odds ratios) that are elevated when compared with those obtained using the untransformed variable.

A major reason for the imbalances between cases and controls observed in columns 2 and 3 of Table 3 is that the risk set sampling allows the age at death or last observation for a control to be much greater than that of the matching case so long as it is at least as great as that of the case and so long as the age of the control at first employment at the plant is not greater than the age at death from lung cancer of the matching case. In this case–control study, the controls were, on average, 9.7 years older than the matching cases at age of death or last observation approximately 4 years older when first employed at the plant and nearly 6 years older at termination of employment. These imbalances made the controls more likely to have all or part of their exposure outside of the latency period of the matching case because their age at first employment and at employment termination was on average several years greater than that of the matching case.

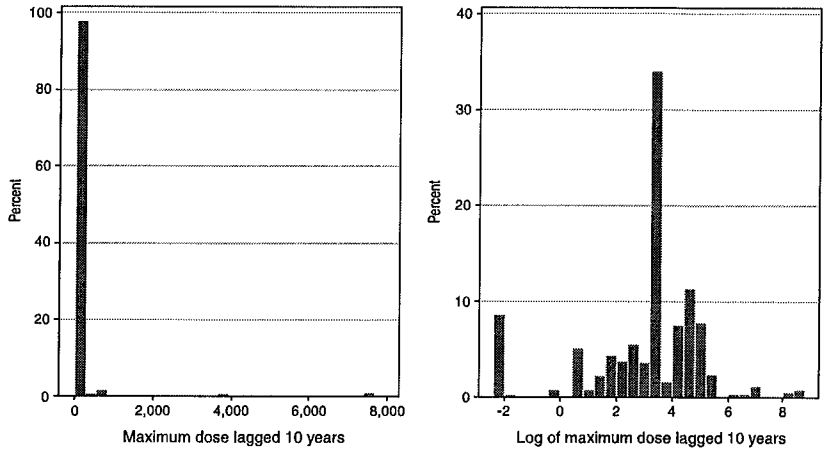

Fig. 2. Histogram of exposure metric, maximum exposure; untransformed exposure metric, on left; logarithmic transformation on right. (Figure was constructed from Table 3 of Sanderson et al.[4])

TABLE 3
Summary Statistics for Workers in Case–Control Study

Variable	Cases (n = 142)	All Controls (n = 710)	Subset of Controls Matched to Case Within 3 Yr of Age at Death or When Last Observed (n = 200)
Mean age at death or last observation (yrs)	65.8	74.5*	71.7
Mean age first employed (yrs)	33.2	37.1†	34.6
Mean age at termination of employment (yrs)	36.9	42.6†	40.0
Percentage of subjects having exposure truncated to 0 owing to 10-yr exposure lagging	1.4%	10.2%†	4.0%
Percentage of subjects having some truncation of exposure owing to 10-yr exposure lagging	10.6%	24.6%†	15.5%

*Three subjects not in analysis set because of missing values.
†Three subjects not in analysis set because of missing values.

To illustrate this effect, we performed the logistic regression using only those 200 controls that were matched to the cases within 3 years of age at death or last observation, and we show the result of this in Table 4 column 3. Closer matching of controls to cases resulted in fewer imbalances between cases and controls (comparing column 2 with column 4 in Table 3) and, as seen in Table 4, the odds ratios obtained using only those controls that were matched to the cases within 3 years of age at death or at last observation

TABLE 4
Effects of Matching Imbalances Between Cases and Controls on 10-yr Latency Analysis Odds Ratios

Exposure Metric Log-Transformed and Lagged 10 Yr; Subjects Having 0 Latency-Based Exposures Included in Analysis	Exposure Odds Ratios (95% CI)	
	All Cases and Controls Included in the Analysis (n = 142 cases; 710 matched controls)	All Cases Included in the Analysis; Only Controls Matched to Case Within 3 Yr of Age Included (n = 142 cases; 200 matched controls)
Employment duration	0.94 (0.87–1.03)	0.91 (0.82–1.02)
Cumulative exposure	1.06 (1.00–1.12)	0.99 (0.91–1.08)
Average exposure	1.22 (1.09–1.36)	1.11 (0.95–1.31)
Maximum exposure	1.20 (1.08–1.33)	1.06 (0.92–1.22)

CI indicates confidence interval.

were considerably lower than those obtained from the complete control group. Also, for all of the exposure metrics, the 95% confidence intervals for the estimated odds ratios overlapped unity (Table 4, column 3).

Our discussion regarding the matching issue applies only to the use of risk set sampling in combination with exposure lagging and, to our knowledge, has not been addressed in the literature. It would not be an issue when disease latency is not under consideration and hence there is not exposure lagging. Although we would recommend closer matching of cases to controls in studies such as the one considered here, which does consider latency, we realize that this may not be easy to do in nested case–control studies constructed from mortality studies when a high proportion of the cohort is still alive at the last follow-up date. Still, as was shown in Table 3, specifying that controls should be at least within 3 years of age at last ascertainment would have led to fewer imbalances.

The discrepancies that we found between findings using the untransformed and those using the transformed exposure metrics was a product not only of the lack of closeness in the matching, but also to the particular relationships in these data among the employment variables, namely age at first employment, age at termination of employment, and age at death or last follow up. We are presently in the process of exploring empirically how widespread are these particular patterns and to what extent the findings here would generalize to nested case–control studies that have different employment patterns. We are also involved in methodological analysis that would identify the conditions that would lead to greater exposure truncation in the controls as compared with the cases.

With respect to this specific application, however, we feel that the findings of our reanalysis of the NIOSH nested case–control study challenges the interpretation given by the authors that the best explanation for their findings is that of a causal association between exposure to beryllium products and lung cancer occurrence.

References

1. Ward E, Okun A, Ruder A, Fingerhut M, Steenland K. A mortality study of workers at 7 beryllium processing plants. *Am J Ind Med.* 1992;22:885–904.
2. Doll R. Occupational cancer: a hazard for epidemiologists. *Int J Epidemiol.* 1985;14: 22–31.
3. Levy PS, Roth, HD, Hwang P, Powers T. Beryllium and lung cancer: a review of a cohort mortality study. *Inhal Toxicol.* 1985;14:101–113.
4. Sanderson WT, Ward EM, Steenland K, Peterson MR. Lung cancer case-control study of beryllium workers. *Am J Ind Med.* 2001;39:133–144.
5. Langholz B, Goldstein L. Risk set sampling in epidemiologic studies. *Stat Sci.* 1996;11:35–53.
6. Gail M. Density sampling, In: Gail M, Benichou J, eds. *Encyclopedia of Epidemiologic Methods.* Chichester, UK: John Wiley and Sons; 2001:300.
7. Checkoway H, Pierce N, Kriebel D. *Research Methods in Occupational Epidemiology,* 2nd ed. Oxford, UK: Oxford University Press; 2004:170–173.
8. Hosmer DW, Lemeshow S. *Applied Logistic Regression,* 2nd ed. New York: Wiley and Sons; 2000:223–257.
9. StataCorp. *Stata Statistical Software; Release 9.* College Station, TX: StataCorp LP; 2005:216–232.