ELSEVIER

Confounding after Risk-Set Sampling in the Beryllium Study of Sanderson et al.

KENNETH J. ROTHMAN, DrPH, AND PAUL L. MOSQUIN, PhD

PURPOSE: Beryllium is classified as carcinogenic on the basis largely of limited human data showing a modest increase in lung cancer from occupational exposure. With occupational exposure now curtailed, earlier results merit more scrutiny. We simulated data to understand the design implications of a landmark case-control study.
METHODS: We generated datasets from the original occupational cohort by randomly assigning lung cancer events to workers independently of their exposure. We analyzed the simulated data on the basis of different modes of risk-set sampling, with risk sets defined by calendar time, age, or both, to assess how much bias existed using several exposure metrics. We controlled for several time related variables to assess confounding. Finally, we re-analyzed the data from the original study, controlling for time-related covariates.
RESULTS: No bias occurred using any type of risk-set sampling with unlagged exposures. When exposure was lagged 10 or 20 years, however, there was considerable confounding by year of birth and year of hire, which remained uncontrolled in the original study.
CONCLUSIONS: Simulations and reanalysis show that much of the reported association with lagged exposure is attributable to confounding by year of birth and year of hire. Lagging changes the exposure variable and can thus lead to changes in the amount of confounding.
Ann Epidemiol 2011;■:■–■.

KEY WORDS: Beryllium, Case-Control Study, Risk-Set Sampling, Confounding, Simulation, Lagged Exposure.

INTRODUCTION

Beryllium was first classified as carcinogenic to humans by the International Agency for Research Against Cancer in 1993 (1). Much of the human evidence came from a series of occupational studies among U.S. workers involved with beryllium processing, culminating in a publication in 1992 (2) that combined data from 7 U.S. processing facilities. This report found an overall standardized mortality ratio for lung cancer of 1.26, comparing workers in the plants with population data. Because the formerly high occupational exposures to beryllium no longer exist, no new sources of data relating to substantial beryllium exposure are likely. Thus, considerable attention has been focused on the existing data. Critics have argued that the association is weak and confounded by smoking and other occupational exposures (3, 4). Another concern related to external comparisons of workers with general population data. These concerns were among the motivations for a nested case-control study by Sanderson et al. (5), conducted within one of the seven plants included in the earlier research.

Sanderson et al. (5) used risk-set sampling in their study, a version of density-based sampling that samples controls from the risk set for each case (6). The risk set for each case includes those who were at risk to become a case at the time of each case's occurrence; the time scale may be calendar time, age, or another time index, such as time since start of treatment in clinical studies or time since start of employment in occupational studies (7–9). Whichever variable is chosen as the time index, if associated with exposure it must then be taken into account as a matching factor and controlled in the analysis. Other time-related variables that are not matched may be confounders, as in any other study.

Considerable debate has arisen about the validity of the study design used by Sanderson et al. (10–16). Some critics have implied that risk-set sampling could have introduced a bias, and others have rejected that argument. Like Wacholder (16), we expect that risk-set sampling should not in itself introduce any bias, provided that the time-matching variable is conditioned in the analysis. Nevertheless, as outlined previously, time-related

From RTI Health Solutions, Research Triangle Institute, Research Triangle Park, NC.
Address correspondence to: Kenneth J. Rothman, DrPH, RTI Health Solutions, 200 Park Offices Drive, Research Triangle Park, NC 27709. Tel.: 919-541-6000. E-mail: KRothman@rti.org.
This study was funded by a contract between Brush Wellman Corporation and RTI Health Solutions. RTI Health Solutions is a business unit of the Research Triangle Institute, which is an independent nonprofit research organization that does work for government agencies and private companies. RTI had full control over the content of this work.
Received January 3, 2011; accepted March 6, 2011.

360 Park Avenue South, New York, NY 10010
1047-2797/$ - see front matter
doi:10.1016/j.annepidem.2011.03.008

confounding factors can still bias a study with risk-set sampling. Our purpose here is to investigate the extent to which the risk-set sampling and analysis as conducted by Sanderson et al. was subject to time-related confounding.

METHODS

Sanderson et al. (5) conducted a nested case-control study within an occupational cohort of 3569 male workers employed at a beryllium processing plant in Reading, Pennsylvania, between January 1, 1940, and December 1, 1969, as described by Ward et al. (2). They sampled controls from risk sets by using the computer program of Beaumont et al. (9), with age as the time-matching variable for defining the risk sets. Controls were sampled randomly without replacement from workers whose age at death or age at loss to follow-up was greater than the case's age at death, and whose age at hire was less than the case's age at death, a version of risk-set sampling that defines the risk sets by age rather than calendar time. Exposures were estimated for the final case-control sample according to the methods outlined in Sanderson et al. (17). Calendar time was not controlled. The matching was taken into account in their analysis through the use of a conditional logistic regression model that conditioned on the risk sets.

The dataset for the studies of Ward et al. and of Sanderson et al. were made available by National Institute for Occupational Safety and Health to the sponsor, Brush-Wellman Corporation. Three subjects had unknown date of birth in the data of Ward et al., and we excluded them, leaving 3566 subjects for our analysis. Date of birth was specified only by month and year. Our copy of the dataset did not contain data on race, which was controlled by Sanderson et al. through matching. Otherwise, we used the same exposure and covariate data as Sanderson et al., generating 10,000 datasets that contained lung cancer cases that were simulated on the basis of population life tables but without regard to their actual beryllium exposure (Appendix).

We obtained cause-specific, calendar-year-, and age-specific population mortality rates for white males from National Institute for Occupational Safety and Health LTAS.NET (18). For each of the 3566 Reading cohort subjects, date and cause of mortality was determined randomly by the use of application of life table rates appropriate to their date of birth, starting from their hire date and continuing until death or exit from the cohort (Appendix). Exposures were assigned by simple random sampling (with replacement) from the 852 historical exposures in the data from Sanderson et al. The simulation process is described in greater detail in the Appendix. The simulation guaranteed a null association between beryllium exposure and lung cancer.

We then used several different approaches for sampling of controls and data analysis to estimate how much bias affected the original study design and analysis. For the first set of analyses, we used risk-set sampling with calendar time matching of controls to cases, and no other matching variable. We included all eligible controls in each risk set. The exposure of controls was calculated from exposure information before the occurrence of lung cancer in the matched case. We analyzed the data by using a conditional logistic model that conditioned on the matched sets. To evaluate confounding after using this sampling scheme, we used four different conditional logistic regression models. The first model included only terms for the three highest exposure quartiles. The second model added a single continuous term for year of birth. The third substituted year of hire for year of birth. The fourth included both year of hire and year of birth as two continuous terms.

In our second approach to sampling, we replicated the methodology of Sanderson et al., who used age rather than calendar time as the risk-set variable. This approach requires controls to have survived to at least the same age as the matched case, with exposures truncated for controls at the age at which the matched case died. The analysis of Sanderson et al. did not include any control for calendar time. In our third approach to sampling, we reverted to calendar time as the risk-set variable, but we also restricted control selection to controls born in the same year as the case, thus matching simultaneously for both calendar time and age.

Sanderson et al. used several measures to assess amount of beryllium exposure: (i) tenure, defined as total days worked; (ii) cumulative exposure, expressed as $\mu g/m^3$ days; (iii) average daily exposure, expressed as $\mu g/m^3$; and (iv) maximum exposure, expressed as $\mu g/m^3$. For zero lag time, the strongest association with lung cancer was seen with average daily exposure, so that is the metric that we used in our analyses. Following Sanderson et al., we divided the scale of average exposure into four categories as defined by the quartile cutpoints of the unlagged exposure from the study of Sanderson et al. and chose the lowest of these four categories as the referent. Sanderson et al. recalculated their quartile cutpoints for each of their analyses as they changed the exposure with lagging. We applied the quartile cutpoints from the unlagged analysis to all the lagged analyses to have consistent demarcations for the exposure categories for all the analyses. We did, however, run additional analyses by using the changing boundaries recalculated as in Sanderson et al. to examine what effect, if any, the changing category boundaries had on the results.

Finally, we used the actual data from Sanderson et al. (apart from race) to attempt to replicate their design and analysis. We then added control of year of birth and year of hire by including indicator terms for quinquennia of these

variables and compared our findings with and without controlling those variables. We assumed that the change in estimate as a result of controlling for these variables reflected uncontrolled confounding by these factors.

RESULTS

Our first set of analyses used risk-set sampling on the basis of risk sets defined as those who were being followed in the study cohort at the same calendar time that the case died ("traditional" risk-set sampling), and with no other matching variable. In the first ("crude") analysis, we used only exposure to predict lung cancer. Note that calendar time is controlled by matching and the use of a conditional logistic model. Because cases were generated independently of exposure, we would hope to see that all rate ratio estimates are unrelated to exposure. Thus, in the absence of confounding by other variables, the rate ratio estimates should center around 1.0. The results of this first analysis are shown in the first row of Table 1, which shows effect estimates across the four exposure quartiles (using the lowest quartile as the referent) for each of the three lag times.

For zero lag time, there was in fact no relation with exposure and the median rate ratio estimates were all near 1.0. With exposure calculated after lagging of 10 or 20 years, however, we note increasing bias with increasing lag time. The effect estimates are similar for all three upper quartiles, indicating that it is the lowest quartile exposure group that appears to be different from the others. For a 10-year lag, the median rate ratio was increased approximately 29% in each of the upper three quartiles, and for a 20-year lag, the median rate ratio estimate was increased approximately 54%, indicating substantial bias for these lagged data, especially for the 20 year lag.

In our second analysis, we used the same sampling approach as the first analysis, but we controlled for birth year by including a single continuous term for birth year in the conditional logistic regression model. (Throughout, we used birth year instead of age, because with the sampling scheme used by Sanderson et al., some controls were no longer alive in the calendar year of their matched case, leaving age undefined for them at that time. For someone alive, birth year is tantamount to age at a specified time, and for someone who has already died, birth year is well defined, unlike age.) These results are in row 2 of Table 1. As before, there was essentially no bias for the unlagged exposure, but there was bias for the 10-year lagged exposure and more bias for the 20-year lagged exposure. With analytic control of birth year, however, the bias was considerably smaller, approximately 6% for exposure lagged 10 years and between 7% and 8% for exposure lagged 20 years. It appears that the bias seen in the crude analysis is substantially mitigated by controlling for birth year.

Row 3 of Table 1 shows results when the analysis was repeated using year of hire instead of birth year as a single continuous term in the conditional logistic model. The findings were similar to those in which birth year was controlled as a covariate, with slightly less bias for the exposure lagged 20 years, but slightly more with it lagged 10 years. In row 4 of Table 1 we present the results obtained when controlling for both year of birth and year of hire in the logistic regression model. There is still evidence of some bias for the lagged exposure, but it is smaller than the bias with control of only one of these two variables, and, interestingly, it is slightly smaller for the exposure lagged 20 years than it is for the exposure lagged 10 years.

The next analysis of the simulated data mirrored the sampling scheme of Sanderson et al., with the use risk sets defined by age, without regard to calendar time, and with no control of birth year, calendar year, or year of hire in the analysis. The results are given in the first row of Table 2. There is no discernible bias for the unlagged exposure, but considerable bias for exposure lagged by 10 or 20 years, with approximately a 16% inflation of the rate ratio for the exposure lagged 10 years and approximately 28% inflation of the rate ratio (or 25% on the log rate ratio scale) for the exposure lagged 20 years. As seen in our other simulations, the bias is nearly constant across the second through fourth quartiles of exposure, indicating that the problem may be confined to the referent category of exposure. When we

TABLE 1. Median rate ratio estimates using calendar time matched risk-set sampling, by exposure quartile and lag time, for 10,000 simulated analyses with lung cancer cases generated randomly, independently of beryllium exposure

	Lag Time (yrs)								
	Zero			10			20		
	Exposure Quartile			Exposure Quartile			Exposure Quartile		
Covariates in Analytic Model	2	3	4	2	3	4	2	3	4
None	1.003	0.998	1.005	1.291	1.284	1.290	1.539	1.533	1.547
Birth year	1.004	0.998	1.004	1.061	1.059	1.062	1.074	1.071	1.079
Year of hire	1.003	0.999	1.003	1.072	1.067	1.072	1.056	1.051	1.062
Birth Year, Year of Hire	1.003	0.998	1.003	1.045	1.042	1.046	1.034	1.028	1.037

TABLE 2. Median rate ratio estimates using age-matched risk-set sampling, and time and age-matched risk set sampling, by exposure quartile and lag time, for 10,000 simulated analyses with lung cancer cases generated randomly, independently of beryllium exposure

	Lag Time (yrs)								
	Zero			10			20		
	Exposure Quartile			Exposure Quartile			Exposure Quartile		
Time Variables for Risk-Set Sampling	2	3	4	2	3	4	2	3	4
Age	1.005	0.999	1.006	1.156	1.153	1.159	1.278	1.272	1.280
Time and Age	1.001	0.996	1.005	1.008	1.006	1.010	1.002	0.997	1.007

added calendar time as a matching variable, in addition to age, there was no discernible bias for any of the exposure metrics, even for the exposure lagged by 20 years (row 2 of Table 2).

We repeated this analysis matching by year of hire instead of birth year, and obtained essentially identical results (results not shown). We noted that matching by birth year or year of hire, coupled with a conditional logistic regression conditioning on the matching factors, was more effective than simply controlling either as a covariate in the logistic model. The latter approach depends on the regression model capturing the correct functional relation between the covariate and the study outcome, whereas the matching guarantees the absence of confounding by the matched variables apart from bias that stems from too loose a match.

To understand the bias better, we selected a typical simulation generated using the control sampling method used by Sanderson et al. We selected the simulation run for which the effect estimate in the greatest exposure category with 20-year lagging was closest to the median of that measure over all simulations. We then stratified the data from that simulation by year of birth to examine age confounding in the stratified simulated data. In Table 3, we show the year-of-birth-stratified data from a typical simulation using the control sampling method employed by Sanderson et al. There is considerable confounding by year of birth, as a comparison of the crude odds ratio of 1.27 and the Mantel-Haenszel adjusted odds ratio of 1.05 indicates. (Although stratification is no longer the usual approach to control residual confounding in matched studies, it has the advantage of giving a clearer picture of the origin of the confounding.) Most of the confounding is removed by a simple Mantel-Haenszel adjustment on the basis of strata of year of birth. The table gives a clue to the mechanism of the confounding. The quartile boundaries were established on the basis of the zero lag data, but in these data with a 20-year lag, the exposure of many subjects is shifted toward lower categories. This shift occurs for a greater overall proportion of controls than cases. In Table 3, there are more than twice the number of controls in the first quartile than in the fourth quartile of exposure, but for cases there are only 64% more cases in the lowest versus the highest quartile.

In Table 4, we examine the actual data from the study by Sanderson et al. for the greatest level of exposure (quartile 4) with a 20-year lag. They reported 1.76 for their odds ratio,

TABLE 3. Data comparing quartile 4 with 1 for a typical simulated dataset with a 20-year lag time, using the sampling of Sanderson et al. and stratified by year of birth

	Quartile			
Year of birth	Cases quart. 4	Contr. quart. 4	Cases quart. 1	Contr. quart. 1
[1865, 1885)	1	0	1	38
[1885, 1890)	0	1	0	37
[1890, 1895)	0	2	9	30
[1895, 1900)	4	11	2	23
[1900, 1905)	2	12	7	26
[1905, 1910)	3	21	6	26
[1910, 1915)	6	25	9	24
[1915, 1920)	4	10	4	20
[1920, 1925)	4	25	4	29
[1925, 1930)	3	18	3	18
[1930, 1940)	1	8	1	7
Total	28	133	46	278

$OR_{crude} = 1.27$; $OR_{MH} = 1.05$.

TABLE 4. Data from the study of Sanderson et al., stratified by year of birth, comparing quartile 4 with 1 with a 20-year lag time

	Quartile			
Year of birth	Cases Quart. 4	Contr. Quart. 4	Cases Quart. 1	Contr. Quart. 1
[1865, 1885)	0	0	2	49
[1885, 1890)	0	4	2	28
[1890, 1895)	2	7	3	28
[1895, 1900)	1	4	2	19
[1900, 1905)	2	21	4	16
[1905, 1910)	5	19	1	18
[1910, 1915)	5	21	2	12
[1915, 1920)	8	31	2	9
[1920, 1925)	5	43	4	8
[1925, 1930)	4	14	0	4
[1930, 1940)	0	5	0	7
Total	32	169	22	198

$OR_{crude} = 1.70$, $OR_{MH} = 1.08$.

which corresponds closely to the crude value of 1.70 that we show in the table. When these data are stratified by year of birth, however, the Mantel-Haenszel summary odds ratio over the age strata is 1.08, indicating striking confounding by this variable in their data and their reported results, confirming what we found in our simulations.

A potential difference between our simulation results and the analysis of Sanderson et al. is their recalculation of exposure scale cutpoints for each analysis. We used consistent exposure boundaries that were determined by the quartile cutpoints for the zero-lag exposure data. We therefore reanalyzed the data from Sanderson et al. using our method of consistent exposure cutpoints across various analyses, to compare with their results using recalculated cutpoints. We found some small to moderate differences in the effect estimates between the two approaches. In nearly all cases the effect estimates were larger when using the varying cutpoints employed by Sanderson et al. than using consistent cutpoints.

As a final step, we repeated the full analysis of Sanderson et al. with their actual data, rather than simulating the data. First we used the same analysis that they used, which did not control for year of birth or year of hire in the analysis. Then we controlled these two variables in the analysis by including them in the conditional logistic regression model. Without controlling for year of birth and year of hire, and using a 20-year lag, we found that the estimates of odds ratio for quartiles 2, 3, and 4 were 2.06, 3.03, and 1.80 relative to the lowest quartile. These values are nearly equal to those reported by Sanderson et al., but differ slightly, presumably because we had to surmise precisely what the quartile boundaries were that were used by Sanderson et al., and possibly because we were unable to control for race.

With control of year of birth and year of hire, the effect estimates from the study of Sanderson et al. were greatly reduced, the extent depending on whether we included these covariates as continuous terms in the logistic model or as "factored variables," a term used to describe converting a variable into a set of categorized indicator terms in a model. By using indicator variables for 11 categories of year of birth (10 terms) and 6 for year of hire (5 terms), we obtained odds ratios for exposure quartiles 2, 3, and 4 of 1.55, 1.68, and 0.99, with similar results when including these two variables as single continuous terms. Thus, most of the reported effect in the study of Sanderson et al. appears to be attributable to confounding by these two variables. From analyses that controlled for one but not both of the two variables, we found that the confounding seemed to be contributed approximately equally by both factors.

DISCUSSION

These simulations demonstrate confounding by time-related variables in a case-control study with risk-set sampling and lagged exposures. Such confounding is hardly remarkable, inasmuch as risk-set sampling does not preclude confounding by time-related variables. In our simulations, data were generated so that there was no overall association between exposure and disease, so that any systematic departures from null effect estimates must reflect bias. The simulations indicate that regardless of whether the risk-set sampling is done using age or calendar time as the time variable, with the lagged analyses there appears to be residual confounding by age that needs to be controlled.

The validity of the beryllium studies of Sanderson et al. and others has been debated in discussion that has called into question the reliability of risk-set sampling as a method of control sampling (14–16, 19). We agree with Langholz and Richardson (14), who argued that there is no fundamental bias in risk-set sampling, and with Wacholder, who argued that nested case-control studies are inherently valid. We also agree with Wacholder that "Lagging is simply one way to measure exposure, and does not differ fundamentally from choosing other metrics such as average exposure, peak exposure, or cumulative exposure without lagging. As long as exposure is measured only up to the time of the event, the particular choice of exposure summary cannot introduce bias in comparing cases and controls" (16). Nevertheless, our simulations show bias that increased with greater lagging of exposures in the study done by Sanderson et al. and in various alternative designs that we simulated. Confounding by year of birth in the actual data of Sanderson et al. is evident in our Table 4; note that of those controls who had exposure in the fourth quartile or the first quartile, the proportion that was in the fourth quartile was only 11% for those with birth year before 1900, whereas this proportion was 77% for those born in 1920 or later (Table 4). In contrast, with the zero-lag exposure, the corresponding proportions were 58% and 42% (data not shown), a reversed and smaller difference. The difference in associations for the lagged and unlagged exposures provides some insight into why the degree of confounding with year of birth increased with increasing lag time.

These biases, however, are conventional biases that stem from confounding by uncontrolled, primarily age-related variables. We have demonstrated that no bias remains, for example, when year of birth is controlled using conventional methods such as stratification or inclusion of year of birth in a conditional logistic regression model. Why should bias exist for one definition of exposure but not another? Exposures with different lag times are distinct variables, and the magnitude of their associations with potential confounders such as age and calendar time will differ as the definition of exposure changes. Thus, although lagging of exposures does not in principle create any validity problem itself, different lag times can result in a different amount of confounding, which will need to be controlled in the analysis.

Another investigation involving simulations and matching on attained age has been reported by Hein et al. (15). Their work differs from ours in several important respects. Their interest was in comparing matching on attained age versus matching on attained age and age at death. Their findings are difficult to interpret, because it is unclear what age at death adds beyond age itself, and because they report both a bias toward the null and an increased type I error, which appear incompatible. In any event, they did not start with the actual data of Sanderson et al. (5), but generated entirely new simulated data, for which they modeled specific exposure-disease relations, focusing on mean cumulative exposure as their primary exposure metric. In contrast, our approach retains the dates of birth and hire of the original Reading cohort, assigning exposures at random from the Sanderson et al. data exposures. This approach guarantees a null relation between exposure and outcome. We focused on average daily exposure, for which Sanderson et al. reported the strongest effects. Nothing we report is discrepant from the results of Hein et al., but our intent was to address residual confounding after control of the one time-related variable used by Sanderson et al. in their matching. Our results show that matching on one time-related variable and appropriate adjustment for that matching in the analysis is not sufficient to control confounding by other time-related variables.

Part of the problem highlighted here may stem from the nature of the exposure metric that showed the strongest associations in the study by Sanderson et al. Average daily exposure is not a cumulative exposure measure. A worker who worked for one day could have a much higher average exposure than a 30-year worker if that one day was a day of high exposure. This measure reflects age and time-specific exposure conditions in the workplace and may be more susceptible to confounding with lagging of exposure than alternatives, although we have not explored those alternatives in this analysis.

REFERENCES

1. IARC. 1993. Beryllium, Cadmium, Mercury and Exposures in the Glass Manufacturing Industry. IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, vol. 58. Lyon, France: International Agency for Research on Cancer.
2. Ward E, Okun A, Ruder A, Fingerhut M, Steenland K. A mortality study of workers at seven beryllium processing plants. Am J Ind Med. 1992;22:885–904.
3. Beryllium Industry Scientific Advisory Committee. Is beryllium carcinogenic in humans? J Occupation Environ Med. 1997;39:205–208.
4. Trichopoulos D. On The Alleged Human Carcinogenicity of Beryllium. Safety Facts SF002-Version 1.3, December 2008. Available at: http://www.brushwellman.com/EHS/SafetyFacts/SF002.pdf. Accessed March 28, 2011.
5. Sanderson WT, Ward EM, Steenland K, Petersen MR. Lung cancer case-control study of Beryllium workers. Am J Ind Med. 2001;39:133–144.
6. Rothman KJ, Greenland S, Lash TL. Case-control studies. In: Rothman KJ, Greenland S, Lash TL, eds. Modern Epidemiology. 3rd ed. New York: Lippincott, Williams and Wilkins; 2008, pp. 124–125.
7. Langholz B, Goldstein L. Risk Set Sampling in epidemiologic cohort studies. Stat Sci. 1996;11:35–53.
8. Clayton D, Hills M. Statistical Models in Epidemiology. New York: Oxford University Press; 1993.
9. Beaumont JJ, Steenland K, Minton A, Mayer S. A computer program for incidence density sampling of controls in case-control studies nested within occupational cohort studies. Am J Epidemiol. 1981;129:212–219.
10. Deubner DC, Lockey JL, Kotin P, Powers MB, Miller F, Rogers AE, et al. "Re: Lung cancer case-control study of beryllium workers. Sanderson WT, Ward EM, Steenland K, Petersen MR. Am J Ind Med 2001;39:133–144." Am J Ind Med. 2001;40:284–285.
11. Sanderson WT, Ward EM, Steenland K, et al. "Response to criticisms of Lung cancer case-control study of beryllium workers. Am J Ind Med. 2001;39:133–144." Am J Ind Med. 2001;40:286–288.
12. Deubner DC, Roth HD, Levy PS. Empirical evaluation of complex epidemiologic study designs: workplace exposure and cancer. J Occup Environ Med. 2007;49:953–959.
13. Levy PS, Roth HD, Deubner DC. Exposure to beryllium and occurrence of lung cancer: A reexamination of findings from a nested case-control study. J Occup Environ Med. 2007;49:96–101.
14. Langholz B, Richardson D. Are nested case-control studies biased? Epidemiology. 2009;20:321–329.
15. Hein M, Deddens J, Schubauer-Berigan M. Bias from matching on age at death or censor in nested case-control studies. Epidemiology. 2009;20:330–338.
16. Wacholder S. Bias in full cohort and nested case-control studies? Epidemiology. 2009;20:339–340.
17. Sanderson WT, Petersen MR, Ward EM. Estimating historical exposures of workers in a beryllium manufacturing plant. Am J Ind Med. 2001;39:145–157.
18. Schubauer-Berigan MK, Raudabaugh WR, Ruder AM, Hein MJ, Silver SR, Chen B, et al. LTAS.NET: A NIOSH life table analysis system for the windows environment. Ann Epidemiol. 2005;15:656.
19. Deubner DC, Roth HD, Levy PS. Empirical evaluation of complex epidemiologic study designs: Workplace exposure and cancer. J Occup Environ Med. 2007;49:953–959.

APPENDIX.

Historical datasets for the Reading, Pennsylvania cohort and the case-control study of Sanderson et al. (5) were obtained from NIOSH. The Reading file contains complete date of birth and date of hire information for 3566 of the 3569 subjects in the cohort. Date of birth was known only to the month, and so all subjects were assumed to have been born on the 15th of the month, and date of hire and age at hire were recalculated as decimal values. The Sanderson file provides date of birth, date of hire, date of death, cause of death (either lung cancer or other) and estimated exposure for the 142 individuals of the Reading cohort who died of lung cancer and 710 matched controls who were selected by risk-set sampling for the study of Sanderson et al.

Each of our simulated cohorts was generated starting with the date of birth and date of hire information for each of the 3566 subjects with complete data in the original Reading

cohort. For each of these subjects, date of death and cause of death (either lung cancer or other cause) was simulated following the life table approach that follows. Average exposure was then assigned to subjects by a simple random sample with replacement from the 852 historical exposures of the Sanderson dataset. There was therefore no dependence on case-control status in the assignment of exposure in the simulated data. A total of 10,000 simulated cohorts were created by this approach.

To assign date and cause of death to each cohort subject, we used mortality rates from the National Institute for Occupational Safety and Health (NIOSH)'s LTAS.NET (18). Mortality rates for white males from lung cancer were taken directly from the LTAS.NET tables, whereas the rates for other causes were set equal to the sum of all rates for non-lung-cancer causes. The NIOSH rates are yearly rates for 5-year periods and 5-year age ranges. We assigned these yearly rates to each of the 25 age-year combinations represented by a single NIOSH risk, here represented as r^L_{jk} for lung cancer and r^O_{jk} for other causes for $j = 1865,\ldots,1993$, $k = 14,\ldots,120$. Because several of the subjects were hired at age 14, we assigned the age 15 values to age 14 as well. We assigned the age 85 value to all ages above 85 in the same birth year. Subject risks r^L_{jk} , r^O_{jk}, for calendar year j were set equal to the risk that applied for the age that the person was on first day of that year.

For each subject, starting with their first year of employment and continuing for each year thereafter up to cohort censoring at end of follow-up in December 31, 1992, we simulated whether death or survival occurred, where the probability death (by lung cancer or other causes) depends only on the LTAS risks and the duration of employment in the year. More specifically, for a subject alive and employed at the start of a given year, as a randomizing device, two independent uniform random variables $0 < u_1, < 1, 0 < u_2 < 1$, were generated, with death occurring by lung cancer if $u_1 < r^L_{jk} f$, and by other causes if $u_2 < r^O_{jk} f$ where $f = 1$, except for the year in which the subject was hired, when f is the fraction of the year spent in employment. If the subject was found to have died from both other causes and lung cancer in a given year, then the cause of death is lung cancer if $u_1/r^L_{jk} \leqslant u_2/r^O_{jk}$, and other causes if not. Age at death was set equal to age at the start of the year plus $(1-f) + f * u_1/r^L_{jk}$ if the cause of death was lung cancer, or age at the start of the year plus $(1-f) + f * u_2/r^O_{jk}$ if the cause of death was other causes. Follow-up ended at the end of 1992, at which time all subjects still alive were considered withdrawn alive.