AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 121, No. 2
Printed in U.S.A.
Copyright © 1985 by The Johns Hopkins University School of Hygiene and Public Health
All rights reserved

AN INVESTIGATION OF BIAS IN A STUDY OF NUCLEAR SHIPYARD WORKERS

E. ROBERT GREENBERG,[1] BERNARD ROSNER,[2] CHARLES HENNEKENS,[2] ROBERT RINSKY[3] AND THEODORE COLTON[4]

E. R. Greenberg (Norris Cotton Cancer Center, Hanover, NH 03756), B. Rosner, C. Hennekens, R. Rinsky and T. Colton. An investigation of bias in a study of nuclear shipyard workers. *Am J Epidemiol* 1985;121:301–8.

The authors examined discrepant findings between a 1978 proportional mortality study and a 1981 cohort study of workers at the Portsmouth, New Hampshire, Naval Shipyard to determine whether the healthy worker effect, selection bias, or measurement bias could explain why only the proportional mortality study found excess cancer deaths amoung nuclear workers. Lower mortality from noncancer causes in nuclear workers (the healthy worker effect) partly accounted for the observed elevated cancer proportional mortality. More important, however, was measurement bias which occurred in the proportional mortality study when nuclear workers who had not died of cancer were misclassified as not being nuclear workers based on information from their next of kin, thereby creating a spurious association. Although the proportional mortality study was based on a small sample of all deaths occurring in the cohort, selection bias did not contribute materially to the discrepant results for total cancer deaths. With regard to leukemia, misclassification of occupation in the proportional mortality study and disagreement about cause of death accounted for some of the reported excess deaths.

epidemiologic methods; mortality; neoplasms; radiation effects

Selection of subjects in a 1978 proportional mortality study of workers at the Portsmouth, New Hampshire, Naval Shipyard involved identification of 1,722 death certificates in New Hampshire, Maine, and Massachusetts that indicated occupation at a shipyard. Next of kin, if contacted (the investigators contacted only 34 per cent), were asked whether the decedent had worn a radiation detector and had engaged in work involving radiation exposure ("nuclear work"). Using age-specific proportional mortality data for all US white males, the investigators determined expected numbers of deaths. Among nuclear workers, they observed statistically significant excesses of deaths from all cancers

Received for publication November 29, 1983, and in final form February 29, 1984.

Abbreviation: NIOSH, National Institute for Occupational Safety and Health.

[1] Norris Cotton Cancer Center and Departments of Community and Family Medicine and Medicine, Dartmouth-Hitchcock Medical Center, Hanover, NH.

[2] Channing Laboratory, Departments of Medicine and Preventive Medicine and Clinical Epidemiology, Harvard Medical School, Boston, MA.

[3] Division of Surveillance, Hazard Evaluations and Field Statistics, National Institute for Occupational Safety and Health, Center for Disease Control, Cincinnati, OH.

[4] Boston University School of Public Health, Boston, MA.

Reprint requests to Dr. E. Robert Greenberg, Norris Cotton Cancer Center, Hanover, NH 03756.

This work was partially funded by Grant CA23108, and Contracts HEW-PHS-CDC-NIOSH-80-1574 and HEW-PHS-CDC-NIOSH-79-2887.

The authors thank Dr. Tom Najarian, who kindly provided us with copies of the original data used in the PMR study; the staff of the National Institute for Occupational Safety and Health, who prepared a tape of their data; and Vera Bergen, who assisted in preparing the manuscript.

(56 observed, 31.5 expected) and leukemia (six observed, 1.1 expected). Among nonnuclear workers, they found no statistically significant excess deaths from cancers or leukemia. The authors acknowledged potential sources of bias in their study and suggested the need for "more careful and thorough cohort studies of workers in naval yards where nuclear powered vessels are serviced" (1).

Investigators at the National Institute for Occupational Safety and Health (NIOSH) later conducted a historical cohort study of mortality among Portsmouth Naval Shipyard workers (2). From shipyard files, they identified 24,545 individuals who had worked at least one day at the shipyard between 1952 and 1977, of whom 4,762 had died. They used nosologist-coded death certificates to calculate standardized mortality ratios (SMRs) for three groups according to shipyard records of radiation exposure: those engaged in nuclear work and occupationally exposed to radiation (subcohort I); those not involved in nuclear work (subcohort II); and those who had qualified for nuclear work and had been monitored but had no recorded occupational radiation exposure (subcohort III). Age-specific mortality rates for all US white males served as the standard for comparison. The NIOSH investigators found no significant increase in mortality from cancer or leukemia for any of the three subcohorts. The mortality from cancer was actually higher for nonnuclear workers (SMR = 100) than for radiation-exposed nuclear workers (SMR = 92) or for qualified but unexposed nuclear workers (SMR = 84). Shipyard workers who had qualified for nuclear work (whether or not they had been exposed to radiation) had a significantly lower all-cause mortality than US white males, a finding the authors attributed to the "healthy worker effect." No excess leukemia mortality was observed in the nuclear workers.

We evaluated several possible explanations for the discrepant results between the original proportional mortality study and the cohort study. First, the initial study involved proportional mortality ratios, while the cohort study employed standardized mortality ratios. The elevated cancer proportional mortality ratios in nuclear workers might have largely resulted from lower noncancer mortality (the healthy worker effect) (3). Second, the mortality experience of subjects identified by death certificate in the proportional mortality study might not have been representative of all shipyard workers (selection bias). Third, in the proportional mortality study, the mortality experience of the subgroup whose next of kin were successfully contacted might have been atypical (selection bias). Fourth, the next of kin might have erred in their recall and reporting of the subjects' occupational category (measurement bias). We restricted our statistical analyses to deaths from all cancers because of the smaller numbers of deaths from leukemia and other specific hematologic cancers. We did, however, examine the deaths reported as leukemia in the proportional mortality study to determine their occupational radiation exposure and nosologist-coded cause of death.

METHODS

Following the approach used in the proportional mortality study, we considered all individuals who had been monitored for radiation exposure (subcohorts I and III in the NIOSH study) as nuclear workers and the remainder (subcohort II) as nonnuclear workers. The published NIOSH report (2) contained the observed and expected numbers of deaths by cause in the cohort. The expected numbers were calculated from US white male death rates by five-year calendar time periods. From the tables in the NIOSH report, we calculated, for nuclear and nonnuclear workers, the cancer standardized mortality ratios.

The NIOSH investigators provided a computer tape listing the name, radiation exposure, date of birth, date of death, and cause of death for all Portsmouth, New Hampshire, Naval Shipyard workers employed after 1952 and known to have died. We eliminated from this list those workers for whom no cause of death information had been obtained by the Institute, leaving 4,416 decedents who constituted the "NIOSH list."

Dr. Thomas Najarian supplied a list of names, dates of death, and causes of death for subjects in the initial study. The list also indicated whether next of kin had been contacted, and, if so, the reported occupation as "nuclear worker" or "nonnuclear worker." Of the 1,750 names on this list, 116 were duplicates and 31 were women. We matched the remaining 1,603 names and dates of death with those on the NIOSH list to determine which of the individuals identified through death certificates had actually been employed at the shipyard during the period of interest. Using these procedures, we found a match for 1,174 individuals (73 per cent), who constituted the death certificate review list. There were 429 unmatched individuals, possibly including those who stopped work at the shipyard before 1952, those who had worked at the shipyard as nongovernment contract employees, those who worked at other naval shipyards, and those who never worked at a shipyard. Cancer deaths were those coded as ICD-8 (Eighth Revision of the *International Classification of Diseases*) 140–207; the remainder were noncancer deaths. Leukemia deaths were coded 204–207.

We defined the proportional mortality ratio as the proportion of all deaths that were caused by cancer. To adjust for possible effects of age differences between nuclear and nonnuclear workers, we also calculated an age-adjusted proportional mortality ratio using the total population of Portsmouth Naval Shipyard deaths as the standard. In particular, let p_i = the proportion of deaths due to cancer in age group i in the total population of shipyard deaths, C_{i1},D_{i1} = the number of cancer deaths and total number of deaths in age group i for nuclear workers, and C_{i2},D_{i2} = the number of cancer deaths and total number of deaths in age group i for nonnuclear workers. The age-adjusted proportional mortality ratio (PMR) for nuclear and nonnuclear workers, respectively, is given by $PMR_1 = \Sigma C_{i1}/\Sigma D_{i1}p_i$, $PMR_2 = \Sigma C_{i2}/\Sigma D_{i2}p_i$.

We assessed whether selection bias affected the initial proportional mortality study results by examining the cancer proportional mortality ratios for those nuclear workers and nonnuclear workers enrolled in that study and for those whose next of kin had been successfully contacted. We assessed the possible effects of measurement bias by examining the cancer proportional mortality ratios for nuclear and nonnuclear workers classified correctly and incorrectly by next of kin. Age-adjusted proportional mortality ratios were also computed for these various subgroups with the same methods as given above for the total group of nuclear and nonnuclear workers.

To determine the characteristics predictive of a worker being selected and correctly classified, we performed three multiple logistic regression analyses in which the respective outcomes were 1) whether or not he was enrolled in the death certificate review list, 2) once enrolled in the list, whether or not his next of kin was contacted, and 3) once next of kin was contacted, whether or not they correctly classified his exposure status. The predictor variables considered in these analyses were year of birth, year of death, cause of death (cancer or noncancer), years employed at the shipyard after 1952, type of worker (nuclear or nonnuclear), and the interaction between cause of death and type of worker.

RESULTS

Characteristics of the groups

Of the 4,762 deceased Portsmouth Naval Shipyard workers, 4,416 (93 per cent) had a known cause of death, 1,174 (25 per cent) were identified in the proportional mortal-

TABLE 1

Characteristics of the deceased worker population identified by NIOSH, of the subgroup identified through death certificate review (DCR) list, and of the subgroup whose next of kin were contacted among deceased workers at the Portsmouth, NH, Naval Shipyard, 1952-1977

	NIOSH	DCR list	Next of kin contacted
No.	4,416	1,174	470
Mean death age (years)	62.4	65.5	65.5
Mean year of death	1967.6	1970.2	1971.2
% nuclear workers	22.3	32.4	36.2
% cancer deaths	22.2	22.1	23.4

ity study, and relatives of 470 (10 per cent) were contacted. Table 1 presents the characteristics of these three groups with regard to age at death, year of death, proportion who were nuclear workers, and proportion who died of cancer. The group identified on the death certificate review list included relatively more nuclear workers than occurred in the NIOSH group. Among the 470 individuals whose next of kin were contacted, there were proportionally more nuclear workers and slightly more cancer deaths than were on the death certificate review list. Thus, preferential selection of nuclear workers occurred at each of the two steps in assembling the study group, but there was no major tendency to select cancer deaths.

The healthy worker effect

Tables 3-5 in the NIOSH report contained information on all observed deaths and expected deaths (calculated from US white male mortality rates) among 8,960 nuclear and 15,585 nonnuclear workers. From these data, we determined the all-cause and the cancer standardized mortality ratios (SMRs) (table 2). We also calculated the cancer proportional mortality ratios for the deaths with cause known to NIOSH and listed on the computer tape (table 2). Among nuclear workers, there was a small deficit in cancer mortality (SMR = 90) and a relatively large deficit in all-cause mortality (SMR = 75). Among nonnuclear workers, cancer mortality (SMR = 100) and all-cause mortality (SMR = 98) were virtually identical with that expected among US white males. Although the standardized mortality ratio for cancer in nuclear workers was 10 per cent less than that in nonnuclear workers (ratio of SMRs = 0.90), the cancer proportional mortality ratio (PMR) was about 20 per cent higher (ratio of PMRs = 1.20) because the nuclear workers had a lower all-cause mortality. This represents the healthy worker effect and accounts for part of the discrepancy in results between the proportional mortality study and the cohort study.

Selection bias

For all workers identified in the original proportional mortality study (the death certificate review list), the proportional mortality ratios for cancer were 0.268 in nuclear workers and 0.198 in nonnuclear workers (table 3), yielding a ratio of 1.36, higher than the ratio of 1.20 for the NIOSH list. The difference between these two figures is a consequence of selection bias occurring in the identification of deceased shipyard workers through death certificate review.

Among the 1,174 deceased workers on the death certificate review list, there were 470 whose next of kin had been contacted. Among these, the proportional mortality ratios for cancer were 0.207 in nonnuclear workers and 0.282 in nuclear workers, giving a ratio of 1.37 (ratio of age-adjusted PMRs = 1.53). Thus, we found only a modest selection bias resulting from contact of only about one third of the next of kin (table 3).

Measurement bias

Among the 470 deceased workers whose next of kin were contacted, when subdivided according to the report of their relatives, the proportional mortality ratios for cancer were 0.193 in those reported as nonnuclear workers and 0.336 in those reported as nuclear workers, yielding a ratio of 1.74 (ratio of age-adjusted PMRs = 1.96), substantially higher than that calculated for this group when actual occupational infor-

TABLE 2
The healthy worker effect among Portsmouth, NH, Naval Shipyard workers, 1952–1977

Occupation	Total workers	All deaths		Cancer deaths			
		No.*	SMR*	No.*	SMR*	PMR†	Age-adjusted PMR†
Nuclear	8,960	1,029	75	251	90	0.26	0.25
Nonnuclear	15,585	3,733	98	726	100	0.21	0.21

Ratio of SMRs = 0.76
Ratio of cancer SMRs = 0.90
Ratio of cancer PMRs = 1.20
Ratio of age-adjusted cancer PMRs = 1.20

* Number of deaths and standardized mortality ratios (SMR) taken from reference 2 and from NIOSH records.
† Calculation of the proportional mortality ratios (PMR) and the age-adjusted proportional mortality ratios based on the 986 nuclear worker deaths and 3,430 nonnuclear deaths with cause of death known by January 1982.

TABLE 3
Determination of proportional mortality ratios (PMRs) for deceased workers at the Portsmouth, NH, Naval Shipyard, 1952–1977

Group*	Occupation	Cause of death†		PMR	Ratio of PMRs	Age-adjusted PMR	Ratio of age-adjusted PMRs
		Cancer (*n*)	Noncancer (*n*)				
NIOSH list	Nuclear	252	734	0.26	1.20	0.25	1.20
	Nonnuclear	730	2,700	0.21		0.21	
Death certificate review list	Nuclear	102	278	0.27	1.36	0.26	1.37
	Nonnuclear	157	637	0.20		0.19	
Contacted group (occupation per navy)	Nuclear	48	122	0.28	1.37	0.28	1.53
	Nonnuclear	62	238	0.21		0.18	
Contacted group (occupation per next of kin)	Nuclear	45	89	0.34	1.74	0.34	1.96
	Nonnuclear	65	271	0.19		0.17	

* See text for definition of groups.
† Based on the death with cause known as of January 1982.

mation was used (table 2). We attribute the difference between these two figures to *measurement bias*, or misclassification of exposure status.

This bias was largely due to inaccurate information from the next of kin regarding the work exposure histories of those nuclear workers who died from noncancer causes. The effect in the 2 × 2 table was to diminish the numbers in the "nuclear worker/not cancer" cell and increase the numbers in the "nonnuclear worker/not cancer" cell. There was a smaller tendency to misclassify nuclear workers who died of noncancer causes: the net effect introduced a bias away from the null. The age-adjusted proportional mortality ratios confirmed these results.

The original report of the proportional mortality study included 55 subjects whose names were not found on the NIOSH list. The data from these subjects did not appreciably affect the results, since the ratio of proportional mortality ratios in that study was 1.65, while the ratio for the contacted group was 1.74 (not age-adjusted).

Results of multiple logistic analyses

The most important term in the logistic regression analyses was the cross-product term "cancer × nuclear worker," indicating whether the association between cancer and nuclear work differed in various subgroups. The subgroups considered included the following comparisons: 1) persons on the death certificate review list versus those on the NIOSH list but not on the death certificate list; 2) persons on the death certificate review list whose next of kin were contacted versus those on that list whose next of kin were not contacted; 3) persons whose next of kin provided correct assessment of occupation versus those whose next of kin provided incorrect assessment.

As regards whether a person on the NIOSH list was selected for the death certificate review list, after controlling for year of birth, year of death, and years employed since 1952 (all of which were significant predictors of selection for the death certificate review list), the variables cancer death, nuclear worker, and cancer death × nuclear worker were not statistically significant (table 4). Thus, there was no significant tendency for either cancer deaths or deaths of nuclear workers to appear on the death certificate review list. Similarly, the nonsignificant cross-product term means that the proportions of cancer deaths selected for the death certificate review list were not significantly different in the subgroups of nuclear and nonnuclear workers, respectively, and indicates that no statistically significant bias occurred at this stage of the selection process.

The second analysis concerned the possible bias introduced by successfully contacting only a small proportion of the next of kin. Only year of death and years employed since 1952 were significant predictors, while neither the main effects of cancer death or nuclear worker nor the cross-product term relating these two variables was statistically significant. This indicates that no statistically significant selection bias occurred as a result of contact with only one third of the next of kin.

The third analysis concerned the possible bias introduced by inaccurate information about occupational exposure. Year of death was the only significant confounding variable, and both nuclear worker and the cross-product term were significant predictors of correct assessment of occupational exposure. The occupational exposure information was incorrect significantly more often if the deceased was a nuclear worker. Of particular importance, the significant cross-product term indicates that this relative misclassification of nuclear versus nonnuclear workers occurred more often for noncancer than for cancer deaths, thus creating a bias.

TABLE 4

Results of logistic regression analyses for deceased workers at the Portsmouth, NH, Naval Shipyard, 1952–1977

Variable	Predictors of selection onto death certificate review list (n = 4,416)			Predictors of successful contact with next of kin (n = 1,159)*			Predictors of correct classification of occupation by next of kin (n = 467)†		
	R	SE	p value	R	SE	p value	R	SE	p value
Constant	−30.21			−117.26			−242.46		
Year of birth	−0.052	0.0043	<0.001	−0.0085	0.0079	NS‡	0.019	0.018	NS
Year of death	0.065	0.0072	<0.001	0.067	0.151	<0.001	0.106	0.034	0.002
Years employed since 1952	0.136	0.0086	<0.001	0.041	0.016	0.013	−0.0064	0.038	NS
Cancer death (1 = yes, 0 = no)	−0.208	0.108	NS	0.049	0.186	NS	−0.426	0.446	NS
Nuclear worker (1 = yes, 0 = no)	−0.071	0.106	NS	0.033	0.168	NS	−2.381	0.341	<0.001
Cancer death × nuclear worker§	0.190	0.191	NS	0.072	0.303	NS	1.504	0.599	0.012

* Fifteen persons excluded because of missing data.
† Three persons excluded because of missing data.
‡ NS, not significant.
§ Defined as 1 if both a cancer death and a nuclear worker and 0 otherwise.

Leukemia deaths reported in the proportional mortality study

The original report, based on 525 deaths, described eight deaths from leukemia, six in the nuclear workers (1.1 expected) and two in the nonnuclear workers (2.8 expected). For these eight decedents, NIOSH records, with cause of death coded by an independent nosologist, were as follows: Four were nuclear workers, of whom two were coded as having died of leukemia, one of "bone marrow failure," and one could not be assigned a cause of death because the death certificate was not available; four were nonnuclear workers, of whom two were coded as having died of leukemia and two of other causes, with leukemia only a contributing cause. The nuclear worker who had bone marrow failure listed as the cause of death on the death certificate, was said by his family to have died of complications following treatment for leukemia. His death certificate was coded as ICD 289.9, i.e., aplastic anemia. Therefore, of the eight deaths identified in the proportional mortality study, four had the principal cause of death assessed as nonleukemia by the nosologist in the NIOSH study.

Discussion

This evaluation indicates that the discrepancies with regard to all cancers between the initial proportional mortality study and the subsequent cohort study of Portsmouth, New Hampshire, Naval Shipyard workers resulted largely from 1) the healthy worker effect, i.e., the lower all-cause mortality among nuclear compared with nonnuclear workers, and 2) measurement bias, a more substantial effect, which resulted from misclassification of decedents' occupational exposure. Selection bias contributed less to the discrepant results, although the proportional mortality study included only a small proportion of all deaths that occurred in the cohort.

The healthy worker effect may have occurred because health requirements for entry into nuclear work were more stringent than those for nonnuclear work (2). Similar effects have been reported in several occupational cohort studies of mortality (4–6). The potential presence of the healthy worker effect, an inherent limitation in the interpretation of all proportional mortality studies (3, 6–8), partly explains why such studies are used to suggest associations rather than to test specific hypotheses.

Other investigators have explored the problem of inaccurate exposure data obtained from next of kin (9–14). Their reports, however, have dealt primarily with misclassifications occurring equally among study groups, which always bias observed results toward a null effect (15, 16). In the proportional mortality study, misclassification of occupation occurred principally in the nuclear worker/noncancer group, thereby introducing a substantial bias away from the null hypothesis of no association between nuclear work and cancer death. When use of next of kin information is unavoidable, this type of bias is an inherent possibility that may be difficult to overcome, particularly if the study hypothesis is generally known.

In our analysis, we used the occupational classification designated in the NIOSH records, and we did not take into account total radiation dosage when classifying subjects as nuclear or nonnuclear workers. In this regard, our definition did not permit exploration of possible relationships between radiation dose and cancer death.

Since we confined our analysis to the category of all cancer deaths, we cannot quantitatively address discrepancies concerning deaths from site-specific cancers. Nevertheless, for leukemia deaths, misclassification of both occupation and cause of death appeared important, although the numbers were too small to reach firm conclusions.

In summary, misclassification of workers' occupation and lower all-cause mortality among nuclear workers contributed to the initial finding of an elevated cancer proportional mortality ratio among deceased nuclear shipyard workers. Bias occurring in the selection of subjects for the initial study accounted for a smaller part of the elevation in cancer proportional mortality ratios.

REFERENCES

1. Najarian T, Colton T. Mortality from leukemia and cancer in shipyard nuclear workers. Lancet 1978;1:1018–20.
2. Rinsky RA, Zumwalde RD, Waxweiller RJ, et al. Cancer mortality at a naval nuclear shipyard. Lancet 1981;1:231–5.
3. Kupper LL, McMichael AJ, Symons MJ, et al. On the utility of proportional mortality analysis. J Chronic Dis 1978;31:15–22.
4. McMichael AJ, Spirtas R, Kupper LL. An epidemiologic study of mortality within a cohort of rubber workers, 1964–72. J Occup Med 1974; 16:458–64.
5. Fox AJ, Collier PF. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. Br J Soc Med 1976; 30:225–30.
6. McMichael AJ. Standardized mortality ratios and the "healthy worker effect": scratching beneath the surface. J Occup Med 1976;18:165–8.
7. Decouflé P, Thomas TL, Pickle LW. Comparison of the proportionate mortality ratio and standardized mortality ratio risk measures. Am J Epidemiol 1980;111:263–9.
8. Wong O, Decouflé P. Methodological issues involving the standardized mortality ratio and proportionate mortality ratio in occupational studies. J Occup Med 1982;24:299–304.
9. Enterline PE, Capt KG. A validation of information provided by household respondents in health surveys. Am J Public Health 1959;49:205–12.
10. Adelson SF. Some problems in collecting dietary data from individuals. J Am Diet Assoc 1960; 36:453–61.
11. Moore MC, Moore EM, Beasley CdeH, et al. Dietary-atherosclerosis study on deceased persons. J Am Diet Assoc 1970;56:13–22.
12. Rogot G, Reid DD. The validity of data from next-of-kin in studies of mortality among migrants. Int J Epidemiol 1975;4:51–4.
13. Kolonel LN, Hirohata T, Nomura AMY. Adequacy of survey data collected from substitute respondents. Am J Epidemiol 1977;106:476–84.
14. Jain M, Howe GR, Johnson KC, et al. Evaluation of a diet history questionnaire for epidemiologic studies. Am J Epidemiol 1980;111:212–19.
15. Bross I. Misclassification in 2 × 2 tables. Biometrics 1954;10:478–86.
16. Copeland KT, Checkoway H, McMichael AJ, et al. Bias due to misclassification in the estimation of relative risk. Am J Epidemiol 1977;105:488–95.