The Importance of Employment Status in Occupational Cohort Mortality Studies

Kyle Steenland and Leslie Stayner

Person-years at risk in occupational cohort mortality studies may be defined as "active" (when a person is working) or "inactive" (after a person has left employment at the plant under study). To investigate the effects of employment status (active/inactive) both across studies and within them, we have analyzed ten large cohort studies conducted by the National Institute for Occupational Safety and Health in which no occupational risk had been observed. These ten data sets included 89,376 workers, 1,984,505 person-years, and 18,840 deaths. In these ten studies, the SMR for all causes was positively correlated with the percentage of inactive person-years in the study ($r = 0.57, p = 0.08$). Considering only inactive person-years, the all-causes SMR was 1.12 (approximately 1.25 before age 65, dropping to 1.00 after age 65). Stratification of inactive person-years by time-since-last-employment showed markedly increased mortality during the first year following employment. The all-causes SMR during active person-years was 0.40 and was fairly constant across age categories. With active and inactive person-years combined, a strong negative trend in SMRs with duration of employment was observed for all causes and for heart disease. These trends were not apparent when person-years were stratified by employment status. These results indicate that investigators should evaluate the effects of employment status when comparing SMRs between multiple cohorts or when interpreting trends in rate ratios within cohorts. (Epidemiology 1991;2:418–423)

Keywords: epidemiology, occupation, mortality, employment status, neoplasms, coronary disease.

Occupational cohort mortality studies compare the mortality rates of an exposed group (the working population) to an unexposed group (often the general population). The resulting measure is typically a standardized mortality ratio (SMR). Because a working population is generally healthier than one that is not employed, when the general population (which includes many people not employed) is used as a referent group, SMRs for overall mortality are usually lower than 1.0. This "healthy worker" effect has been discussed extensively in the literature.[1-8]

Mortality rates are composed of observed deaths divided by person-years at risk. The person-years in occupational studies can be divided into two types: person-years while working at the study plant ("active" person-years) and person-years subsequent to working at that plant ("inactive" person-years, although since workers who leave a plant may work elsewhere, "inactive" is somewhat of a misnomer). Relatively few people die while working, hence the mortality rate for active person-years would be expected to be low compared with the general population, absent any occupational risk. The mortality rates during inactive person-years, conversely, might be expected to be similar to or even higher than that of the general population (especially before the age when most people in the general population retire), depending on what percentage of those who leave work find work elsewhere and on the percentage of individuals who are not employed in the general population.

The percentage of active versus inactive person-years could therefore be expected to influence the findings of different cohort mortality studies. Similar considerations come into play when considering analyses of trends with duration within one cohort. Investigators frequently calculate SMRs by duration of employment or exposure (for example, 1–10, 10–20, 20+ years). A positive trend is taken as strengthening the case for an occupational risk.

Regarding duration, workers who work for only a short time (for example, less than a year) at the plant in question contribute few active person-years to the study but may contribute many inactive person-years (if there are many short-term workers). Their cumulative duration is short, and all of their person-years may be contributed to the short duration category (for example, a category of less than one year). Longer-duration workers will contribute some active person-years to the short duration category while passing through, but these active person-years may be relatively few. Hence, the category of shortest duration will often be composed of primarily

inactive person-years and will tend to have higher mortality rates.

Note that it is not that short-term workers per se have particularly high mortality, but rather that their contributed person-years are primarily inactive. The inactive person-years contributed by short-term workers should not necessarily show higher mortality rates than the inactive person-years contributed by long-term workers. Short-term workers are unlikely to leave work because they are ill. If their mortality after leaving employment is higher than the corresponding mortality for long-term workers, such higher mortality might be a result of life-style factors (for example, smoking, drinking), rather than preexisting illness.

In light of all this, the employment status (active/inactive) of the person-years might be expected to influence observed trends with duration and can be considered a potential confounder or effect modifier of such trends. Furthermore, if exposure increases the probability of leaving work due to exposure-induced chronic disease, then employment status itself will be a risk factor for death and will be an intermediate variable on the pathway from exposure to disease. Even if exposure itself does not cause increased mortality, exposure may increase the probability that individuals already ill for other reasons will leave work (for example, via irritation), and these individuals will again have higher mortality owing to their preexisting illness. In this situation, employment status will play the role of an intermediate variable, even though there is no direct causal link between exposure and mortality.

Consideration of employment status is generally absent in occupational cohort mortality studies. There has been some discussion of the issue in methodological articles. Gilbert,[9] in a discussion of confounding factors in mortality studies, noted that mortality rates were higher for short-term workers (less than two years) compared with longer-term workers and noted that this phenomenon might result in an inverse trend in mortality with duration of employment. She also stratified person-years by employment status and noted that inactive person-years showed a higher mortality than active person-years and that short-term workers had more inactive person-years. Wen and coauthors[4] divided a large cohort into active and inactive person-years, showed that inactive person-years had higher mortality rates than the United States population while the opposite was true of active person-years, and observed that the healthy worker effect might in fact be called the "active" worker effect. Howe and coauthors,[8] in an analysis of factors influencing the healthy worker effect, stratified the person-years at risk in their cohorts between active and inactive person-years.

They noted that the active person-years had a low mortality rate that did not increase with increasing time since first employment, in contrast to the situation for inactive years at risk.

Robins[10] has provided the most thorough theoretical discussion of the role of employment status in evaluating the effect of exposure. He pointed out that if workers tend to leave work owing to their exposures (which may be irritating or cause work-prohibiting illness), and if unemployment is a risk factor for the ultimate health outcome of interest (independently of exposure), then employment status cannot be treated simply as a confounder to be controlled via conventional methods. He also suggested a model for analysis in which the current exposure history of cases and controls is compared after matching on prior history of exposure and employment status.

To investigate further the role of employment status, we have examined ten large occupational cohorts previously studied by the National Institute for Occupational Safety and Health.

Methods

We reviewed National Institute for Occupational Safety and Health cohorts that had existing computer files in standard format and for which a written report (usually a publication) was available ($n = 27$). From these, we chose all those in which there were at least 500 deaths, in which there were no major categories of death for which an excess risk (compared with the United States population) was observed, and in which there were no pronounced positive trends with duration of exposure for any major category of deaths. Thus, these were generally "negative" studies. Analyses were restricted to hourly male workers, as the patterns of reemployment after leaving a work site may differ for females and salaried individuals. Furthermore, we restricted attention to those cohorts in which any minimum length-of-employment criterion used to define the cohort was less than or equal to one year, so as not to eliminate all short-term workers, who were partly the focus of our study. We considered all-causes mortality, mortality from all cancers (ICD 140–208, 9th revision), and mortality from ischemic heart disease (ICD 410–414, 9th revision).

We conducted standard life-table analyses using the National Institute for Occupational Health and Safety life table system.[11] In these analyses, person-years at risk were distributed into categories by duration of employment, age, and calendar time. As an individual was observed through time, his person-years were categorized into higher categories for these time-related variables. Cutpoints for duration (chosen a priori) were less than 1

TABLE 1. Description of the Study Population

Number of workers	89,376
Deceased	18,840 (21%)
Vital status unknown at study end	3,284 (4%)
Total person-years	1,984,505
Active	909,055 (46%)
Inactive	1,075,450 (54%)
Average year of first employment	1953
Average year of last employment	1965
Average age at first employment	29
Average length of follow-up (years)	24
White (%)	96
Cumulative employment (%)	
< 1 year	24
1–2 years	10
2–5 years	15
5–10 years	11
10–20 years	16
20+ years	24

TABLE 2. Person-Years, Percent Inactive Person-Years, and All-Causes SMRs for Each Study*

Study	Person-Years	All-Causes SMR	Percent Inactive Person-Years	Percent Employed < 2 yrs
1	140,147	0.95	83	73
2	471,304	0.85	37	25
3	125,393	0.82	49	36
4	656,648	0.96	52	29
5	155,070	1.02	77	47
6	58,960	0.76	45	13
7	54,343	0.97	80	58
8	174,924	0.82	52	9
9	62,639	0.93	68	20
10	85,077	0.83	66	16

*The referent group for all SMR analyses is the United States population.

year, 1–2 years, 2–5 years, 5–10 years, 10–20 years, and 20+ years for duration of employment. Age and calendar time were categorized into 5-year intervals, and person-years were stratified by race (white versus nonwhite).

Only general work histories (first and last employment dates) were considered, and gaps in employment were ignored (several of the larger cohorts lacked detailed work histories). Person-years began at date first employed and continued until the study end date or the date last observed, if the date last observed preceded the study end date.

Person-years were also categorized as either active or inactive. Active person-years began when an individual started employment and ended one week after the end of employment. One week was allowed to account for inaccuracies in personnel records regarding actual last day worked, and to allow for deaths that may have followed an acute illness at work by only a few days. Inactive person-years began seven days after an individual left employment and continued until the end of follow-up. This definition of active and inactive is arbitrary, since workers may become ill at work, leave work, and die at some point thereafter.

All SMR analyses used the United States population as the comparison population. Trend tests for SMRs were

conducted using the method suggested by Breslow and coauthors.[12] Internal comparisons were also conducted using Poisson regression.[13] Results from these analyses were consistent with SMR results and are not presented here.

Results

Ten studies fulfilled our criteria for inclusion in the analysis. The ten studies involve the following industries: leather (1), pesticide production (1), paper and pulp production (1), mining (1), rubber (1), chemical (2), shipbuilding (1), automobile (1), and munitions (not exposed to nitroglycerine) (1). Six studies had no minimum employment entry criterion, two had a one-month minimum, one had a five-month minimum, and one had a one-year minimum.

Table 1 describes the study population. Table 2 presents the percentage of inactive person-years and short-term workers (less than two years) in each of the ten studies. This table indicates that SMRs generally increase as inactive person-years increase. A regression weighted by the person-years showed that there is a linear relationship between the percentage of inactive person-years and the SMR ($r = 0.57$, $p = 0.08$), which explains 33% of the variance of the SMRs. The increase of

TABLE 3. Percent Active Person-Years by Duration Category

	Duration (years)					
	< 1	1–2	2–5	5–10	10–20	20+
Person-years	418,114	208,422	319,731	316,913	402,182	319,147
Active person-years	56,568	55,386	134,717	186,582	273,705	202,097
Active person-years (%)	14	27	42	59	68	63

TABLE 4. SMRs for All Causes, by Active and Inactive Person-Years

Study	Person-Years	SMR Combined	SMR Active (Deaths)	SMR Inactive (Deaths)	Ratio SMR Inactive/Active
1	140,147	0.95	0.31 (63)	1.06 (1,285)	3.42
2	471,304	0.85	0.45 (1,008)	1.12 (3,676)	2.48
3	125,393	0.82	0.29 (119)	1.19 (699)	4.10
4	656,648	0.96	0.30 (551)	1.26 (5,138)	4.20
5	155,070	1.02	0.61 (142)	1.09 (1,402)	1.79
6	58,960	0.76	0.27 (75)	0.96 (642)	3.56
7	54,343	0.97	0.37 (271)	1.00 (1,781)	2.70
8	174,924	0.82	0.47 (36)	1.04 (547)	2.21
9	62,639	0.93	0.46 (63)	1.04 (603)	2.26
10	85,077	0.83	0.50 (87)	0.91 (652)	1.92
Total	1,984,505	0.90	0.39 (2,415)	1.12 (16,425)	2.87

inactive person-years is also, as one would expect, correlated with the percentage of short-term workers ($r = 0.69, p = 0.03$).

Table 3 shows that the percentage of active person-years, for all studies combined, varies substantially by duration of employment category. Hence, employment status can potentially be a confounder in assessing trends with duration. Decreased percentages of active person-years occur in categories of short employment.

Table 4 shows the overall mortality rate, compared with the United States, for all person-years combined as well as for person-years stratified by employment status, for each of the ten studies. It is clear from Table 4 that relatively few people die while "at work" as determined by employment records. Those that do may actually have been on sick leave but are indicated as still active in their work history. Table 4 also indicates that the healthy worker effect exists only during active person-years.

Table 5 shows how mortality rates of inactive and

TABLE 5. SMRs for All Ten Studies Combined, by Age Group

Age Group	SMR Inactive (Deaths)	SMR Active (Deaths)	Ratio SMR Inactive/Active
All causes			
<45	1.24 (1,717)	0.40 (586)	3.10
45–54	1.28 (2,800)	0.40 (792)	3.20
55–64	1.22 (4,453)	0.40 (917)	3.05
65+	1.00 (7,455)	0.42 (120)	2.38
All cancers			
<45	1.26 (206)	0.26 (44)	4.85
45–54	1.24 (537)	0.27 (102)	4.59
55–64	1.20 (1,028)	0.29 (141)	4.14
65+	1.07 (1,589)	0.35 (19)	3.05
Heart disease			
<45	1.23 (280)	0.58 (138)	2.12
45–54	1.19 (912)	0.55 (406)	2.16
55–64	1.20 (1,683)	0.53 (485)	2.18
65+	0.97 (2,825)	0.55 (60)	1.76

active person-years vary by age. This table shows that overall mortality of inactive person-years is remarkably constant (SMRs 1.22 to 1.28) across age until age 65, at which point the SMR drops to 1.00. After age 65, the United States referent group is presumably composed also of primarily inactive persons; hence the drop at that age point would be expected. All-causes mortality during active person-years is also remarkably constant (SMRs 0.38 to 0.42) across all age categories, including after age 65. The patterns for cancer and heart disease are quite similar to that of all causes.

Table 6 shows that a clear inverse trend of mortality with duration of employment is occurring for all causes and for heart disease. No such trend is seen for all cancers. These trends were generally consistent across all plants (see below).

Table 7 shows trends of mortality with duration of employment after dividing the data between active and inactive person-years. The downward trends seen in Table 6 for all causes and heart disease are no longer apparent for the stratified data in Table 7. There is, in fact, an upward trend for cancer among inactive person-years. Plant-specific analyses indicated that this overall increase was due to a significant positive trend for cancer among inactive person-years at two plants.

Table 8 considers inactive person-years by time since last employment. This table shows that the high SMRs for inactive person-years are concentrated in the first one to two years after employment, which is probably a reflection of the fact that some workers end employment because they are ill.

Discussion

We have limited our presentation to empirical descriptions of differing mortality by employment status. In our analyses of studies in which no effect of exposure had been observed by conventional analyses, we observed an inverse relation between duration of exposure and mortal-

TABLE 6. SMR Results by Duration of Employment, All Studies Combined

Cause of Death	Duration Category (years)						All Durations	Total Deaths
	< 1	1–2	2–5	5–10	10–20	20+		
All causes	1.12	0.98	0.90	0.87	0.83	0.83	0.90	18,840
Ischemic heart disease	1.10	0.99	0.93	0.89	0.91	0.88	0.93	6,790
All cancers	0.99	0.81	0.88	0.91	0.85	0.95	0.91	3,662

ity. This inverse relation disappeared when the analysis was stratified by employment status. Furthermore, our analysis showed that mortality was quite high in the first year or two after employment, supporting the thesis that workers who are ill leave employment and suffer high mortality shortly after leaving employment.

Robins[10] has pointed out a possible artificial reason contributing to high mortality during inactive person-years as defined in this study, when general instead of detailed work histories are used. General work histories do not take into account intermittent absences from employment over the course of total employment at a plant. When such intermittent employment is due to a relatively short layoff, some percentage of those who do not return to work may be those who die during the layoff period (although many intermittent absences are not due to layoffs but to voluntary change of jobs). In addition to causing an increase in mortality immediately after date of final employment, this phenomenon contributes to the overall elevation of mortality during inactive person-years (as defined here). Use of detailed work histories would enable the investigators to correctly classify person-

years of intermittent unemployment as inactive. However, proportionally few inactive person-years would be expected to be added to the denominator of the inactive mortality rate, so that the lowering of that overall rate is unlikely to be substantial. Furthermore, an initial increase in mortality upon the beginning of inactive status might still be expected even with the use of detailed work histories. Any correction for the "layoff phenomenon" remains limited by one of the major general limitations of any analysis of "inactive" person-years, namely the lack of a complete employment history outside the plants under study to determine when someone was actually not working.

The interpretation of the mortality trends we have observed here is complex; at least two alternative hypotheses might explain the observed relation between duration of employment (or exposure), employment status, and mortality. First, under the assumption that duration of exposure (or cumulative exposure) is unrelated to employment status, workers leaving their jobs (that is, inactive workers) may be ill (not due to exposure). Due to this illness, or for other reasons related to the status of

TABLE 7. SMR Results and Number of Deaths by Duration of Employment, by Employment Status

Cause of Death	Duration Category (years)						All Durations
	< 1	1–2	2–5	5–10	10–20	20+	
Active person-years							
All causes	0.45	0.43	0.38	0.34	0.41	0.38	0.39
Deaths	80	81	191	297	791	975	2,415
Ischemic heart disease	0.54	0.60	0.52	0.44	0.58	0.55	0.55
Deaths	16	21	55	99	371	527	1,089
All cancers	0.26	0.27	0.38	0.27	0.30	0.26	0.28
Deaths	5	6	24	34	102	135	306
Inactive person-years							
All causes	1.16	1.05	1.05	1.14	1.14	1.11	1.12
Deaths	1,454	701	1,785	3,092	4,934	4,459	16,425
Ischemic heart disease	1.12	1.02	1.00	1.05	1.10	1.08	1.14
Deaths	402	228	641	1,055	1,747	1,628	5,701
All cancers	1.01	0.86	1.47	1.15	1.16	1.36	1.14
Deaths	489	145	320	613	886	907	3,360

TABLE 8. SMR Results and Number of Deaths by Time Since Last Employment

Cause of Death	Years Since Last Employment					
	< 1	1–2	2–5	5–10	10–20	20+
All causes	2.40	1.16	0.96	1.04	1.05	1.14
Deaths	1,454	701	1,785	3,092	4,934	4,459
Ischemic heart disease	1.96	1.10	1.00	1.00	1.01	1.12
Deaths	489	145	320	613	886	907
All cancers	4.20	1.23	0.87	1.04	0.99	1.04
Deaths	402	228	641	1,055	1,747	1,628

not being employed (for example, lack of health insurance), they may suffer increased mortality. In this case, employment status would be a confounder, and conventional methods (for example, stratification) for controlling confounding could be used. Second, workers may terminate employment early for reasons related to the exposure itself (for example, noxious fumes, irritation, or some exposure-induced chronic debilitating condition). In this latter case, exposure is a determinant of employment status which will then be a risk factor for mortality. In this situation, Robins[10] has shown that standard methods of control (stratification on employment status and calculation of a summary measure across strata) may lead to biased estimates of the effect of exposure on mortality, whether or not exposure is causally related to disease. Robins has proposed a nonparametric test (the G-null test) for the overall association between exposure and mortality, controlling for employment status, as well as some parametric models to accomplish the same goal. The essence of the G-null test is a comparison of cases and matched controls for exposure assignment at serial points of time across work history, matching for previous exposure and employment status history prior to each particular time point evaluated.

Investigators who are evaluating trends by duration of exposure in cohort mortality studies may find it worthwhile to stratify their data on employment status to evaluate this potential source of bias, although such stratification must be viewed with caution in light of the possibility that employment status is an intermediate variable. Control over the potential bias introduced by employment status may require more sophisticated approaches such as those proposed by Robins.[10] For exposures with biological latent periods greater than two to three years, our analyses also confirm the suggestions by

Gilbert[9] and Robins[10] that the potentially confounding effects of employment status may be controlled partially by lagging exposures by three years (or more).

Acknowledgments

References

1. Fox A, Collier P. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. Brit J Prev Soc Med 1976;30:225–230.
2. McMichael A. Standardized mortality ratios and the "healthy worker effect": Scratching beneath the surface. JOM 1976;18:165–168.
3. Wen C, Tsai S, Gibson R. Anatomy of the healthy worker effect: A critical review. JOM 1983;25,4:283–289.
4. Wen C, Tsai P, Gibson R, McClellan WA. Long-term mortality of oil refinery workers. II. Comparison of the experience of active, terminated, and retired workers. JOM 1984;26,2:118–127.
5. Monson R. Observations on the healthy worker effect. JOM 1989;28,6:425–433.
6. Sterling T, Weinkham J. Extent, persistence, and constancy of the healthy worker or healthy person effect by all and selected causes of death. JOM 1986;28,5:348–353.
7. Weed D, Tyroler H, Shy C. The healthy worker effect in actively working communications workers. JOM 1987;29,4:335–339.
8. Howe G, Chiarelli A, Lindsay J. Components and modifiers of the healthy worker effect. Am J Epi 1988;128,6:1,364–1,375.
9. Gilbert E. Some confounding factors in the study of mortality and occupational exposures. Am J Epidem 1982;116,1:177–188.
10. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. J Chron Dis 1987;40(suppl 2):139S–161S.
11. Steenland K, Beaumont J, Spaeth S, Brown D, Okum A, Jurcenko L, Ryan B, Phillips S, Roscoe R, Stayner L, Morris J. New development in the life table analysis system of NIOSH. JOM 1990;32,11:1091–1098.
12. Breslow N, Lubin J, Marek P, Langholz B. Multiplicative models and cohort analysis. JASA 1983;78,381:1–12.
13. Frome E, Checkoway H. Use of Poisson regression in estimating incidence rates and ratios. Am J Epidem 1985;121,2:309–323.