

Time-related factors as potential confounders and effect modifiers in studies based on an occupational cohort

by Neil Pearce, PhD, Harvey Checkoway, PhD, Carl Shy, MD¹

PEARCE N, CHECKOWAY H, SHY C. Time-related factors as potential confounders and effect modifiers in studies based on an occupational cohort. *Scand J Work Environ Health* 12 (1986) 97–107. Time-related factors which are potential confounders and effect modifiers in studies based on an occupational cohort are reviewed. The most frequently considered ones include age at first exposure, duration of exposure, interval from exposure to disease recognition, and age at risk. These factors are related to the “healthy worker effect,” which appears to be more pronounced among workers with the longest durations of employment and older ages, at date of hire, but weaker with longer length of follow-up and older age at risk. Hence, use of an internal comparison group may not eliminate bias since confounding will occur if the exposed and unexposed groups differ in their distributions across these factors. It is also shown, using the multistage model of carcinogenesis, that these factors may be important effect modifiers. Fortunately, generally straightforward methods of control exist both for stratified analyses and for the commonly used mathematical modeling approaches. Although no firm recommendations can be made, it would appear to be important to control for length of follow-up in the design or analysis of most studies based on an occupational cohort, and controlling for age at first exposure may also be desirable under many circumstances.

Key terms: carcinogenesis, epidemiology, healthy worker effect, industrial exposures, methodology, temporal effects.

The term “time-related factors” refers to important determinants of disease risk, or modifiers of the effect of the main risk factor of interest, which vary as a person ages (58). The most important time-related factor is the cumulative dose from the main exposure of interest but, for simplicity, it is assumed in this paper that an individual has a fixed lifetime cumulative dose. The discussion concentrates on other time-related factors which should be considered in the design and analysis of occupational cohort studies, or in case-referent studies based on an occupational cohort. The factors which have been the most frequently considered include age at first exposure, duration of exposure, interval from exposure to disease recognition, and age at risk (ie, the age at which incidence or mortality rates are computed) (58).

Some redundancy exists in the set of time-related factors under consideration because age at risk is determined once age at date of hire and length of follow-up have been specified. Hence it will be necessary, in some areas of the discussion, to concentrate on only two of these factors, the third factor being regarded as a linear function of the other two. In studies in which the exposure of interest is simply employment in a particular industry, or in a specific job type with-

in the industry, it is convenient to use age at hire as an indicator of age at first exposure and duration of employment as an indicator of duration of exposure. These terms have been used interchangeably in the discussion, depending on the context. The time between exposure and disease recognition involves the sum of two terms, ie, the time between exposure onset and disease onset and the time during which the disease is “latent” but not clinically detected. The former period has been termed the “induction time” (54). It depends on the stage of the disease process at which the exposure acts and approaches zero for an exposure which affects the final stage. The overall time between exposure onset and the detection of disease has been termed the “empirical induction time” (54). It is also commonly, but incorrectly, referred to as the “latency period.” In occupational cohort studies this time period is often approximated by the length of follow-up from first employment to the recognition of disease.

The principal objective of this paper is to demonstrate that these time-related factors can be considered as potential confounders and effect modifiers in studies based on an occupational cohort. There are many contexts in which these phenomena can occur, but the discussion has concentrated on two. The potential for confounding by time-related factors has been explored in an examination of their relationship to the “healthy worker effect,” and the potential for effect modification by time-related factors has been explored in the context of the Armitage-Doll multistage model of carcinogenesis. In general, no distinction has been drawn between cohort and case-referent studies, as there is

¹ Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, North Carolina 27514, United States.

Reprint requests to: Dr N Pearce, Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, North Carolina 27514, United States.

a straightforward relationship between an occupational cohort study and a case-referent study involving a sample of the same study base (3, 40).

Confounding

Confounding may be conceived as a distortion in the estimate of effect of the exposure(s) of interest due to the effect of an extraneous factor(s). Several formal definitions have been presented (42, 43, 56), and it is not intended to debate their relative merits here. In most practical situations, a confounder is a factor which is predictive of disease occurrence or detection in persons who are not exposed to the main risk factor(s) of interest and is associated with exposure in the study base. It is implied that surrogates for confounding factors may also be regarded as confounders. For example, age at risk is strongly predictive of disease and hence will be a confounder when the age distributions of the exposed and unexposed populations differ, even though it is usually just a marker of more fundamental biological changes or of long-term exposure to a risk factor. The discussion has concentrated, however, on problems of confounding which are more specific to studies based on an occupational cohort, in particular the potential for confounding by the time-related factors under consideration. The "healthy worker effect" is one complex manifestation of these issues, and it is first reviewed briefly.

Healthy worker effect

The healthy worker effect (12, 22, 33, 37, 38, 39, 47) commonly occurs in studies which involve a comparison of the mortality experience of an occupational cohort with the pattern that would have been expected on the basis of national mortality rates, or the mortality rates for the relevant state or county. The typically lower relative mortality, from all causes combined, in an occupational cohort occurs because relatively healthy individuals are likely to gain employment, and to remain employed, ie, that "membership" of the employed population is dependent on health status (26). At least three factors are involved in the healthy worker effect (23): (i) the selection of healthy members from the source population, (ii) the survival in the industry of healthier men, and (iii) the length of time the population has been followed up. Hernberg (27) has commented that the healthy worker effect is essentially the result of a trivial "comparison bias" since the general population does not fulfill even the most elementary criteria for a reference group, especially if one keeps in mind that such a group should reflect "what would have happened in the exposed cohort had there been no exposure [p 124]." Furthermore, Tola & Hernberg (59) have commented that, because of the multifaceted nature of the healthy worker effect, it is doubtful whether such a crude summary term is useful at all and that one should perhaps try to make distinctions between the different underlying

factors. This statement is particularly relevant in the current context since the strength of the healthy worker effect varies across the different levels of the time-related factors under consideration (62).

Length of follow-up

The mortality of employed persons, relative to that of the general population, is lowest during the period immediately following the commencement of employment (38, 59). This phenomenon is illustrated by data from the two studies presented in table 1. Fox & Collier (23), in a cohort study of polyvinyl chloride production workers, found that the all-causes mortality of men within five years of entering the industry was as low as 37 % of that expected. For circulatory and respiratory disease it was as low as 21 %, whereas for cancer it was 45 %. The effect decreased with length of time since entry and had almost disappeared after 15 years of follow-up with standardized mortality ratios of 94, 112, 91, and 93 for all-causes, cancer, circulatory disease, and respiratory disease mortality, respectively. A study by Gilbert (25), which used an internal comparison, found a similar, but weaker, pattern (table 1). Delzell & Monson (18), McMichael (37), Ott et al (50), and Seltzer & Jablon (57) have also found relatively low standardized mortality ratios for the early years of follow-up, the ratios slowly approaching those of the general population as follow-up continued.

Duration of employment

Although the relative mortality advantage of employed persons diminishes with length of follow-up, it is the most pronounced among workers with the longest duration of employment (31, 63). This latter association is attributable to the survival of relatively healthier persons in the industry. For example, Gilbert's study (25) of workers at the Hanford nuclear facility demonstrated risks more elevated for all-causes mortality among terminated workers than among employed workers, and among short-term workers than among long-term workers (table 2).

In Fox & Collier's study (23) the survival effect was measured by separating men who survived 15 years according to whether or not they were still in the industry after this period. The mortality rate among those who left was approximately 50 % higher than those still in the industry for all causes, cancer, circulatory disease, and respiratory disease and was approximately three times higher for lung cancer. Vinni & Hakama's study (60) of a random sample of 20 000 persons from the total Finnish population considered associations between changes in occupations from 1960 to 1970 and mortality during 1971—1975. Those staying in the same occupational category from 1960 to 1970 had 20 % lower overall mortality than those who did not. However, the latter group was not homogeneous; the investigators reported standardized mortality ratios

Table 1. Relative risk for all causes of mortality by length of follow-up.

Study	Type of comparison	Follow-up (years)				Total
		0—4	5—9	10—14	≥ 15	
Fox & Collier (23)	External	0.37	0.63	0.75	0.94	0.75
Gilbert (25) ^a	Internal	0.63	0.93	1.01	1.03	1.00

^a Follow-up groups do not exactly correspond to those in heading.

Table 2. Relative risk for all causes of mortality by duration of employment.

Study	Type of comparison	Duration (years)					Total
		0—1	2—4	5—9	10—14	≥ 15	
Gilbert (25)	Internal	1.15	0.92	0.97	0.94	0.92	1.00

Table 3. Relative risk for all causes of mortality by age at risk.

Study	Type of comparison	Age at risk (years)				Total
		< 54	55—64	65—74	≥ 75	
Fox & Collier (23)	External	0.64	0.79	0.96	0.60	0.75
McMichael (37)	External	0.81	0.89	0.95	1.13	0.98
Delzell & Monson (18)	External	0.8	0.9	0.9	1.0	0.9

Table 4. Relative risk for all causes of mortality by age at hire.

Study	Type of comparison	Age at hire (years)				Total
		25—44	45—54	55—64	65—74	
Fox & Collier (23) ^a	External	0.45	0.37	0.32	0.23	0.37

^a This paper does not explicitly present relative risks by age at hire. The data represent relative risks by age at death during the first five years of employment.

of 70 for persons changing to another occupational category between 1960 and 1970, 100 for persons retiring at the usual age of 65 years during the same period, and 130 for persons retiring early. Similarly, a study in Denmark by Olsen & Sabroe (49) found mortality rates to be 65 % higher for persons leaving the Danish carpenters union than for persons entering the union, but mortality was much lower for persons leaving for a higher paid job than it was for persons leaving for other reasons. Collins & Redmond (13) and Delzell & Monson (18) also found excess risks among early retirees, the latter study identifying the greatest risks in the first year following early retirement.

Hence, the healthy worker effect is strongest during active employment and rapidly disappears following the cessation of employment, particularly if employment ends before the usual retirement age. Wen & Tsai (63) have thus commented that the healthy worker effect is more characteristic of “active” workers and would be more accurately addressed as the “active worker effect.”

Age at risk

Several investigations have examined the relative mortality advantage of employed persons at different levels

of age at risk (table 3). McMichael (37), in a study of male rubber workers, found a standardized mortality ratio of 81 for all causes at age 40—54 years, while that for the age group ≥ 75 years was 113. Similar patterns were also observed by Fox & Collier (23), Delzell & Monson (18), Musk et al (46), and Olsen & Sabroe (48).

Age at date of hire

Two studies have presented standardized mortality ratios according to the age of the worker on the date of hire. Musk et al (46), in a study of mortality among firefighters, found an all-causes standardized mortality ratio of 92 for persons hired before age 40 and that of 88 for persons aged ≥ 40 years at hire. Fox & Collier's study (23) presented data for the first five years after hire, with standardized mortality ratios of 45 and 23 for the 25-to-44-year and 65-to-74-year age groups, respectively (table 4). These results suggest that the relative mortality advantage of employed persons may be greater with increasing age at hire — the opposite pattern to that for age at risk. This hypothesis has plausibility since, if a certain level of health is required to gain employment, then the proportion of persons attaining the required level is likely to be smaller in the

older age groups. Therefore, gradients of health between workers and the general population are likely to increase with increasing age at hire.

Implications

From the foregoing it can be stated that all of the time-related factors under consideration appear to be important predictors of mortality in studies based on an occupational cohort. However, it is not possible to isolate their biasing effects because most published data are stratified on one or at most two variables simultaneously. For example, Fox & Collier (23) found that standardized mortality ratios increased with length of follow-up (table 1) and age at risk (table 3). It could be expected that, for a fixed length of follow-up, the pattern for age at hire would parallel that for age at risk, yet their data suggest the opposite relationship (table 4). One possible explanation is that the observed relationship for age at hire may be correct, whereas that for age at risk may be a confounded effect of length of follow-up. Further studies, or reanalyses of previous studies, are needed to examine separately the effects of these time-related factors.

It is generally recognized that an important method of minimizing bias in studies based on an occupational cohort is to make an internal comparison of the mortality experience of workers exposed to the risk factor of interest and of workers from the same cohort who have not been exposed or have experienced lower levels of exposure. It is not so widely recognized, however, that this procedure does not guarantee an elimination of bias because the exposure groups may differ in their distribution across the time-related factors under consideration. For example, if exposure occurs in "dirty" jobs primarily assigned to short-term, transient workers, then bias may occur if the crude mortality rates for exposed persons are compared with those for a unexposed workforce with a longer average duration of employment because the transient workers may have different base-line disease risks attributable to unfavorable life-style factors (20). This phenomenon may be reflected in the findings of a study by Peto et al (51), of lung cancer in asbestos textile factory workers; it found the highest relative risks for long-term workers and for persons who had worked for less than one year. It might be hypothesized that the former effect was due to asbestos exposure, whereas the latter might have been due to confounding related to life-style. In addition bias could occur if a factory had been "cleaned up." The more recently employed workers — who would have had a lower level of exposure — might have been followed for a shorter period of time and hence have had lower mortality rates than those employed in a previous period.

In other words, the time-related factors under consideration are predictive of disease in occupational cohorts and will thus be confounders if their distributions are different in the exposed and unexposed groups, even when there is no association between ex-

posure and disease. In the situation where exposure is associated with disease, the picture is further complicated if persons who develop disease as a result of exposure may retire early (13, 18, 23, 25, 31, 35, 60), whereas persons who remain healthy may continue to accumulate exposure until the normal age of retirement.

Effect modification

Effect modification occurs when the estimate of effect of the main exposure of interest depends on the level of another factor. All factors which are predictive of disease modify either the rate ratio or the rate difference, as uniformity over one measure implies nonuniformity over the other when the background disease rates vary across levels of the secondary factor. Hence, the decision as to whether or not a factor is an effect modifier depends on the measure of effect in use. The statistical analogue of this issue is frequently referred to as a choice of scales, where additive and multiplicative are the two most obvious options. The concept of effect modification, in the statistical sense, is thus distinct from concepts of biological interaction (55), although biological interrelationships could be expected to be reflected in the observed patterns for at least one effect measure.

The Armitage-Doll multistage model

An important example of interaction (effect modification) which illustrates the interrelationship of biological processes, as depicted mathematically, is the multistage model of carcinogenesis. According to the Armitage-Doll multistage model of carcinogenesis (2, 19), it is assumed that at least one cell must pass through k , distinct, heritable changes in a particular sequence, and that the instantaneous probabilities of transition may depend on the intensity of exposure to a particular risk factor. The simplest model assumes that the background event rate for a particular transition (j) is constant, although in more sophisticated models it may vary with factors such as age, the condition of the neighboring cells, or the time since that cell suffered its previous change (52).

The theory thus assumes that cellular change j occurs with background rate R_j (independent of age). It is easily shown (2) that at age t the background incidence rate $B(t)$ satisfies the following condition:

$$B(t) \sim R_1 R_2 \dots R_k (t - w)^{k-1}$$

or

$$B(t) \sim (t - w)^{k-1},$$

where w is the time for a fully transformed cell to develop into a clinically detectable cancer, ie, latency. This approximation may be inaccurate if the transition rates become sufficiently large (43), but it does not alter the general patterns being explored.

It is further assumed (17) that cellular change j occurs in excess with carcinogen-induced rate $R_j^*c(t)$ that is proportional to dose rate $c(t)$ at age t . Hence, the transition rate in the exposed population of cells is:

$$R_j^*(t) = R_j + R_j^*c(t).$$

The model can be simplified if one assumes that the time to tumor detection (w) is zero, the dose-rate at a particular age is either zero or a constant (c), and that only one of the cellular changes (j) is affected by the exposure of interest.

Lifetime exposure

One special case is when exposure occurs throughout life. Then it can be shown (17, 64) that the excess incidence rate $E(t)$ at age t is proportional to t^{k-1} for all $j < k$. Since the background rate $B(t)$ is also proportional to t^{k-1} , it follows that the excess rate is proportional to the background rate at all ages, and the rate ratio is constant throughout life.

Exposure commencing subsequent to birth

Most epidemiologic studies, including those based on occupational cohorts, are generally more concerned with exposures which begin at some age subsequent to birth. To derive the age-specific instantaneous risk for an exposed individual, we assume that the additional exposure begins at age t_0 , remains at constant level c for time-period d until age $t_1 (= t_0 + d)$, and increases the cellular event rates to $R_j + R_j^*c$. The individual is then observed for a further period. At age t the individual has been followed for period $f (= t - t_0)$.

The model is generally formulated in terms of age at first exposure (t_0) and age at risk (t). Age at first exposure is clearly important since it is associated with the probability of at least one cell having undergone the previous $j-1$ transitions and hence being eligible for transition j on which the exposure of interest acts. Age at risk is of indirect importance since it indicates the probability of a cell having experienced the final stage transition. However, since these subsequent stages can only occur following stage j , the relevant interval is the follow-up period from age at first exposure to age at risk ($f = t - t_0$). Hence, the multi-stage theory predicts the increase of risk with age as a reflection of increasing time since exposure, rather than as an ageing phenomenon (9). This prediction has been confirmed experimentally (53).

Hence two factors are of primary importance in the model: age at first exposure (t_0), which indicates the probability of a cell having undergone the previous $j-1$ transitions, and length of follow-up (f), which indicates the probability of a cell having undergone the subsequent $k-1-j$ transitions. It follows intuitively that age at first exposure will not affect the excess rate from a first-stage carcinogen, since no prior transitions

are necessary, and that length of follow-up will not affect the excess rate from a penultimate stage carcinogen since the final transition will not be affected by exposure and will occur at a constant rate. These observations are examined more formally later.

It should also be noted that duration of exposure (d) is important in at least two respects. At a constant dose rate (c) duration is a surrogate measure of the total dose. A longer duration also implies an increased likelihood that exposure will occur at some time subsequent to the completion of the $j-1$ necessary prior transitions. In this latter situation duration of exposure acts as a surrogate for effective age at exposure. For simplicity, duration of exposure will initially be regarded as fixed, and the situation of continuing exposure will be discussed later.

Fixed duration of exposure

When the stage affected is the last, then the excess rate is zero immediately following the cessation of exposure (assuming a zero latency period).

When the stage affected is not the last, and a cell receives an increment of exposure at age $t_0 + x$, then the likelihood that the cell is caused to become fully transformed at age t depends on the following four factors: (i) the likelihood that the cell has undergone the $j-1$ necessary prior transitions by age $t_0 + x$ [proportional to $(t_0 + x)^{j-1}$], (ii) the likelihood that the exposure will cause transition j to occur (proportional to c), (iii) the likelihood that the $k-1-j$ subsequent transitions will occur by age t [proportional to $[t - (t_0 + x)]^{k-1-j}$], (iv) and the likelihood that the final transition will occur at age t (a constant). Whittemore (64) has thus derived the following simplified expression for excess cancer incidence $[E(t)]$:

$$E(t) \sim c \int_0^d (t - t_0 - x)^{k-1-j} (t_0 + x)^{j-1} dx. \quad (\text{eq 1})$$

Substituting f for $t - t_0$ yields:

$$E(t) \sim c \int_0^d (f - x)^{k-1-j} (t_0 + x)^{j-1} dx. \quad (\text{eq 2})$$

The implications of this formulation are illustrated for the situations $j=1$ and $j=k-1$.

When $j=1$:

$$\begin{aligned} E(t) &\sim c \int_0^d (f - x)^{k-2} dx \\ &\sim c [f^{k-1} - (f-d)^{k-1}]. \end{aligned} \quad (\text{eq 3})$$

Hence, the excess rate increases markedly with length of follow-up (f) and is independent of age at first exposure (t_0). The excess rate ratio is:

$$E(t)/B(t) \sim \frac{c [f^{k-1} - (f-d)^{k-1}]}{(t_0 + f)^{k-1}}. \quad (\text{eq 4})$$

It decreases with age at first exposure (t_0) and can be shown to increase with length of follow-up (f) for a relatively long period before eventually beginning to decrease.

When $j = k - 1$:

$$\begin{aligned} E(t) &\sim c \int_0^d (t_0 + x)^{k-2} dx \\ &\sim c [(t_0 + d)^{k-1} - t_0^{k-1}]. \end{aligned} \quad (\text{eq } 5)$$

The excess increases with age at first exposure (t_0) and is independent of length of follow-up (f), for fixed duration. The excess rate ratio is:

$$E(t)/B(t) \sim \frac{c [(t_0 + d)^{k-1} - t_0^{k-1}]}{(t_0 + f)^{k-1}}. \quad (\text{eq } 6)$$

It decreases with length of follow-up (f). The relationship with age at first exposure (t_0) depends on the relative lengths of follow-up and exposure. Day & Brown (17) have given an example in which the length of follow-up is much greater than the duration of exposure and, in this situation, the excess rate ratio increased with age at first exposure, for a fixed length of follow-up. A similar example is illustrated in figure 1. However, in the situation where exposure is still continuing, or has just been terminated (ie, $f = d$), the excess rate ratio will decrease with age at first exposure.

Continuing exposure

For continuing exposures, the length of follow-up (f) is equal to the duration of exposure (d), and the latter term should be substituted for length of follow-up (f) in the preceding formulas. It can then be shown (8, 17, 64) that, for a fixed duration (follow-up), the general patterns for age at first exposure are the same as those already presented. As would be expected, when duration (follow-up) is permitted to vary, the excess rate and excess rate ratio increase with duration. For an early-stage carcinogen $E(t) \sim d^{k-1}$. This pattern was observed by Doll & Peto (21), who found the incidence of lung cancer in cigarette smokers to be proportional to the fourth or fifth power of duration of smoking.

Implications

Figures 1 and 2 illustrate the preceding findings for a hypothetical five-stage carcinogenic process. Figure 1 illustrates the relationship of the rate difference and excess rate ratio to age at first exposure, given a fixed duration of employment (five years) and a fixed length of follow-up (20 years). Figure 2 illustrates the relationship of the same two effect measures to length of follow-up, given a fixed duration of employment (five years) and a fixed age at first exposure (15 years). The patterns for carcinogens affecting intermediate stages (not shown) are intermediate to those for first-stage and penultimate-stage carcinogens.

Figure 1 demonstrates that age at first exposure is an important potential effect modifier. In particular, if a carcinogen acts at any stage other than the first, the rate difference will increase with increasing age at first exposure since the necessary prior transitions will be more likely to have occurred. This pattern has been

observed in studies of solid tumors in heavily irradiated sites (15) and in a study of bladder cancer in dye-stuff workers (10) although the latter increase was relatively moderate.

Figure 2 demonstrates that length of follow-up is also an important potential effect modifier. The effect of exposure to an early-stage carcinogen will be small until sufficient time has elapsed for the later-stage transitions to occur. For example, if the A-bomb survivor studies had been terminated after about 15 years, the leukemia effect would have been detected but not the solid tumor effect (5); also a study of cancer mortality in vinyl chloride workers found larger standardized mortality ratios when the analysis was restricted to the period 15 years or more after first exposure (28). The relative risk for lung cancer associated with asbestos exposure also remains high for many years but eventually begins to decrease (61). On the other hand, the rate difference associated with a carcinogen affecting the penultimate stage remains approximately constant after exposure stops. This pattern has been observed in studies of cigarette smoking (19, 52).

A feature of figures 1 and 2 is that, although predictable relationships exist for age at exposure and length of follow-up, it is not possible to derive a consistent relationship for age at risk, which is their sum. For example, for a carcinogen which acts at the first stage, the rate difference will be constant across age at risk when length of follow-up is fixed and age at first exposure varies (figure 1). However, it will increase with age at risk when age at first exposure is fixed and length of follow-up varies (figure 2). Hence, few general conclusions can be drawn. For an early-stage carcinogen the rate difference will generally be approximately constant, while the rate ratio will decrease with age at risk (16), in populations which have similar variability for age at exposure and length of follow-up. For a late-stage carcinogen the rate difference will generally increase, and the rate ratio will be approximately constant with age at risk (16). For example, most radiation-induced cancers, apart from leukemia, seem to develop proportionally to the cancer rate for the attained age (5, 16, 29).

It should also be noted that, although the factors of age at risk and duration of exposure do not enter directly into the simple formulation of the multistage model, they can become effect modifiers, independently of the effects of other time-related factors, if some of the simplifying assumptions are relaxed. For example, if a carcinogen acts at an early or intermediate stage and one of the later stage transitions generally occurs at a specific age, then exposed persons will only develop disease when they become old enough for some target cells to have undergone the relevant transition — irrespective of the age at which they were first exposed. In other words, the background transition rate for the later stage will not be a constant (R_j) but will be a function of age [$R_j(t)$]. This phenomenon appears to be rare since cancer does not appear to be a

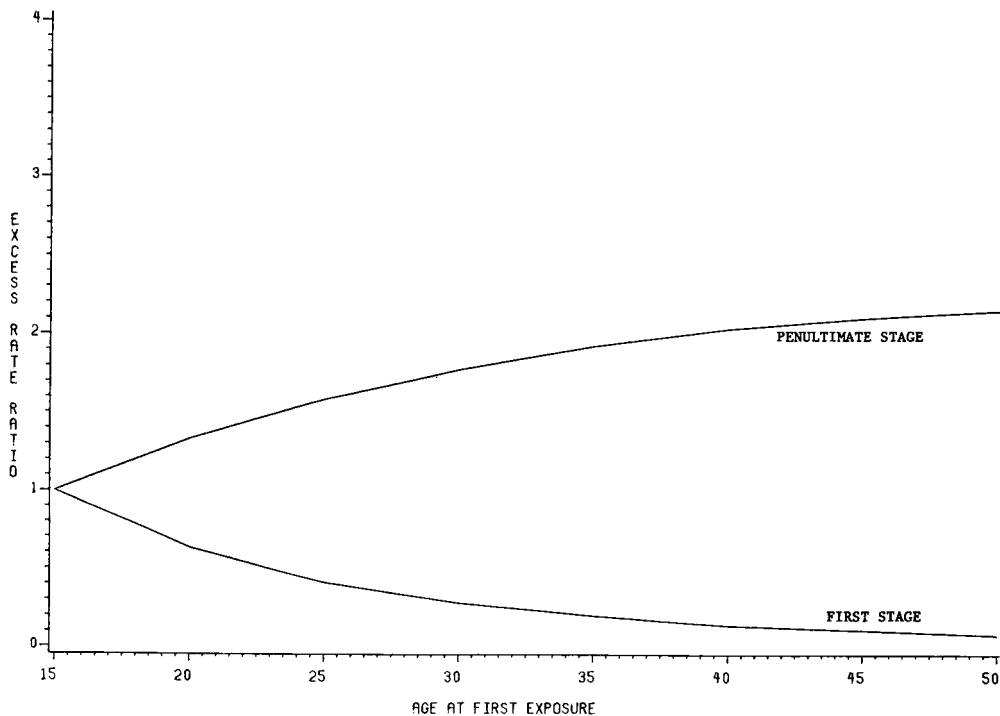
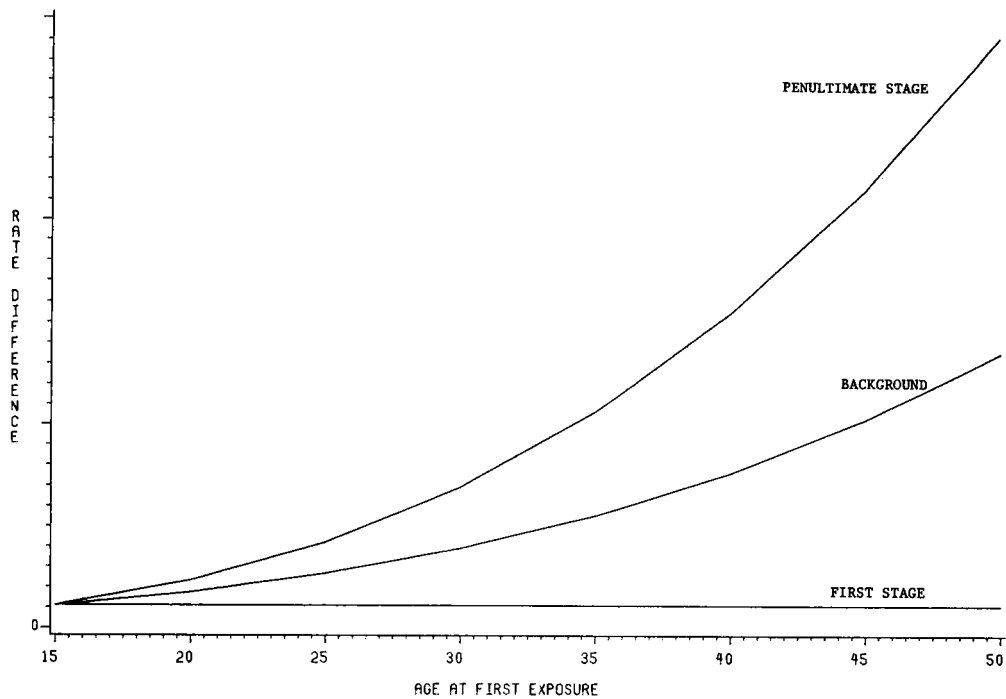


Figure 1. Rate difference and excess rate ratio by age (years) at first exposure for a fixed duration of exposure (5 years) and length of follow-up (20 years), for a hypothetical five-stage carcinogenic process.

natural consequence of ageing (52) but could occur for some specific hormonally related cancers, or if exposure to an exogenous late stage background factor generally occurred at a specific age. Such a phenomenon has been suggested for breast cancer, for which

radiation appears to have an early stage effect but operates in conjunction with a strong hormonally mediated promoting action (16, 45). In other words, radiation appears to initiate cellular changes that increase the likelihood that other, age-related events later in life

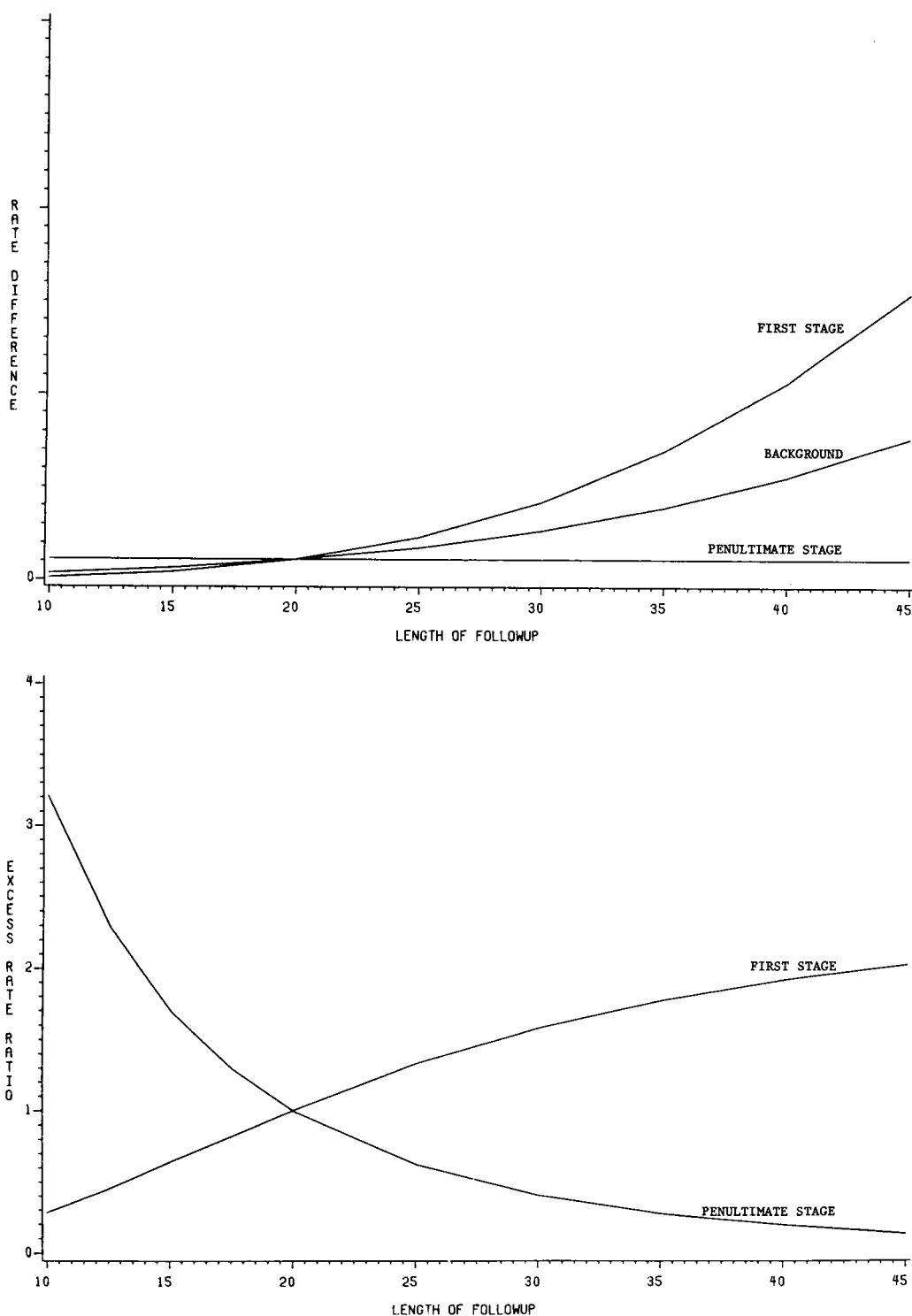


Figure 2. Rate difference and excess rate ratio by length of follow-up (years) for a fixed duration of exposure (5 years) and length of follow-up (20 years), for a hypothetical five-stage carcinogenic process.

result in breast cancer, with excess incidence rates proportional to the age-specific background incidence rates (32).

Duration of exposure will be an effect modifier if the effect of a given dose depends on the delivery rate, ie, if $R_j^*(t)$ is not a linear function of $c(t)$. For

example, a given nonlethal radiation dose may be less effective — due to the different cellular repair probabilities — if an equivalent dose is delivered in smaller fractions over a lengthy period than if it is delivered over a relatively short period (4). On the other hand, if the overall dose is very large, then it will be less effective if it is delivered over a relatively short period, as cell killing may occur (4).

The findings for duration of exposure in the simpler model are also of importance, as duration is predicted to be of more importance than exposure intensity for a carcinogen which affects only one or two stages (21, 52), suggesting that the characterization of dose as the simple product of intensity and duration may not be optimal (6). The choice of model for the expected long-term contribution of exposure is of considerable importance since, when the impact of an intervention is being predicted, it is common for the uncertainties in temporal extrapolation to be a greater source of imprecision than the uncertainties in low-dose extrapolation (58). Rothman has proposed taking the expected contribution of exposure to be constant over some “window” of time and zero outside of it, the optimal window being chosen by trial and error (54), and this is the approach which is usually followed in epidemiologic studies, eg, by ignoring exposure which occurred within a certain time period of the manifestation of disease. Other effective dose models have been developed wherein doses received during a worker’s period of exposure are assigned weights proportional to the time since occurrence (11). Alternatively, one can assume a log-normal distribution for the expected excess rate due to exposure (1, 34).

Discussion

Time-related factors are thus important potential confounders and effect modifiers in occupation cohort studies, or case-referent studies nested in an occupational cohort. Fortunately, generally straightforward methods of control exist both for simple stratified analyses and for the common multivariate approaches.

Control of time-related factors in the former situation involves stratifying the data according to specified levels of these factors (36, 41). The presence of effect modification can be assessed by a comparison of the effect estimates for each stratum, and statistical tests are available to help in such an assessment (6, 30). The assessment of a time-related factor as a potential confounder or effect modifier should ideally be carried out while other such factors are controlled, since one factor (eg, age at risk) may appear to be a confounder or effect modifier merely because of an association with other factors (eg, length of follow-up). This approach will usually encounter computational problems resulting from small numbers, and hence it may be necessary to use mathematical modeling to assess confounding and effect modification. The

major mathematical models used in studies of occupational cohorts include logistic regression (6), the Cox proportional hazards model (14), and Poisson regression (24). These models are all closely related (7), and the principles of control of potential confounders and the identification of effect modifiers are similar. The Poisson regression approach involves a grouped analytic strategy resembling that for a stratified analysis, whereas the other models, such as the Cox proportional hazards model, can incorporate either grouped (categorical) or ungrouped (continuous) data in the analysis.

It is difficult to make firm recommendations on the optimal analytic approach due to the variety of situations which can occur. However, some general guidelines can be drawn.

Studies involving an external comparison with general population mortality or incidence rates, with mortality ratios standardized for age at risk, will continue to be performed and should continue to be interpreted with caution. Apart from age at risk, the strongest predictor of disease in such studies appears to be length of follow-up, and it is thus important to examine the mortality patterns for each separate stratum of follow-up. Frequently, the overall standardized mortality ratio for an exposure-related disease will not be elevated, but the picture will be quite different if attention is focused on those persons who have been followed for the longest time periods. This phenomenon can occur both because of a decline in the healthy worker effect with follow-up and because of a long induction time for exposure-related disease. Duration of employment is also of importance because short-term workers may have greatly increased base-line disease risk, and they should be excluded from most analyses. Specification of the optimal value for minimum employment duration is seldom obvious, and trial and error may be necessitated in the absence of objective guidelines.

Length of follow-up will also be of prime importance as a potential confounder in studies involving an internal comparison. Some models, such as the proportional hazards model, implicitly control for length of follow-up because cases of disease are compared with the exposure experience over the same follow-up period of all the persons who were followed for at least as long but did not experience disease. The same principle ideally should be followed in the selection of referents for a case-referent study nested in an occupational cohort. It may also be necessary to control for age at first exposure or age at risk, in addition to length of follow-up — but it is usually not feasible to control for all three because of the problem of collinearity. Traditionally, most studies have controlled for age at risk since it is an important confounder in other contexts and relevant information is readily available, but the preceding discussion suggests that age at first exposure may be a more fundamental confounding factor in some situations. Furthermore, the discussion of the multistage model suggests that length of follow-

up and age at first exposure are likely to be effect modifiers in studies of occupational carcinogens and that age at risk should be regarded as merely the sum of these two factors. For example, these two factors are of prime importance in studies of radiation-induced solid tumors (15), and it will be necessary to consider these factors in the analysis in order to have sufficient power to detect increased risks for persons exposed to low doses of radiation. The same principle generally applies to other studies of occupational hazards. Hence, in many studies based on an occupational cohort, it may be preferable to assess the potential for effect modification or confounding by length of follow-up and age at first exposure before considering the inclusion of other time-related factors in the model.

Acknowledgments

This work was conducted during the tenure, by the first author, of an Overseas Research Fellowship of the Medical Research Council of New Zealand.

References

1. Armenian HK, Lilienfeld AM. The distribution of incubation periods of neoplastic diseases. *Am J Epidemiol* 99 (1974) 92—100.
2. Armitage P, Doll R. Stochastic models for carcinogenesis. In Neyman J, ed. *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*. Volume 4. University of California Press, Berkeley, CA 1961, pp 19—38.
3. Axelsson O. Elucidation of some epidemiologic principles. *Scand J Work Environ Health* 9 (1983) 231—240.
4. Beebe GW. Ionizing radiation and health. *Am Sci* 70 (1982) 35—44.
5. Boice JD, Land CE. Ionizing radiation. In: Schottenfeld D, Fraumeni JF, ed. *Cancer epidemiology and prevention*. WB Saunders, Philadelphia, PA 1982, pp 231—253.
6. Breslow NE, Day NE. *Statistical methods in cancer research*. Volume I (The analysis of case-control studies). International Agency for Research on Cancer, Lyon 1980.
7. Breslow NE, Lubin JH, Marek P, Langholz B. Multiplicative models and cohort analysis. *J Am Stat Assoc* 78 (1983) 1—12.
8. Brown CC, Chu KC. A new method for the analysis of cohort studies: Implications of the multistage theory of carcinogenesis applied to occupational arsenic exposure. *Environ Health Perspect* 50 (1983) 293—308.
9. Brown CC, Chu KC. Implications of the multistage theory of carcinogenesis applied to occupational arsenic exposure. *J Natl Cancer Inst* 70 (1983) 455—463.
10. Case RAM, Hosker ME, McDonald DB, Pearson JT. Tumors of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the British chemical industry. *Br J Ind Med* 11 (1954) 75—104.
11. Checkoway H. Methods of treatment of exposure data in occupational epidemiology. *Med Lav* (in press).
12. Ciocco A, Mancuso T, Thompson DJ. Four years mortality experience of a segment of the United States working population. *Am J Public Health* 55 (1965) 587—595.
13. Collins JF, Redmond CK. The use of retirees to evaluate occupational hazards. *J Occup Med* 18 (1976) 595—602.
14. Cox DR. Regression models and life tables. *J R Stat Soc B* 34 (1972) 187—220.
15. Darby SC, Nakashima E, Kato H. A parallel analysis of cancer mortality among atomic bomb survivors and patients with ankylosing spondylitis given X-ray therapy. *J Natl Cancer Inst* 75 (1985) 1—21.
16. Day NE. Radiation and multistage carcinogenesis. In: Boice JD, Fraumeni JF, ed. *Radiation carcinogenesis: Epidemiology and biological significance*. Raven Press, New York, NY 1980, pp 437—443.
17. Day NE, Brown CC. Multistage models and primary prevention of cancer. *J Natl Cancer Inst* 64 (1980) 977—989.
18. Delzell E, Monson RR. Mortality among rubber workers: IV General mortality patterns. *J Occup Med* 23 (1981) 850—856.
19. Doll R. The age distribution of cancer: Implications for models of carcinogenesis. *J R Stat Soc* 134 (1971) 133—166.
20. Doll R. Occupational cancer: A hazard for epidemiologists. *Int J Epidemiol* 14 (1985) 22—31.
21. Doll R, Peto R. Cigarette smoking and bronchial carcinoma: Dose and time relationships among regular smokers and lifelong non-smokers. *J Epidemiol Community Health* 32 (1978) 303—313.
22. Enterline PE. Not uniformly true for each cause of death. *J Occup Med* 17 (1975) 127—128.
23. Fox AJ, Collier PF. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. *Br J Prev Soc Med* 30 (1976) 225—230.
24. Frome EL, Checkoway H. Use of Poisson regression models in estimating incidence rates and ratios. *Am J Epidemiol* 121 (1985) 309—323.
25. Gilbert ES. Some confounding factors in the study of mortality and occupational exposures. *Am J Epidemiol* 116 (1982) 177—188.
26. Graham S, Graham-Tomasi R. Achieved status as a risk factor in epidemiology. *Am J Epidemiol* 122 (1985) 553—558.
27. Hernberg S. "Negative" results in cohort studies — How to recognize fallacies. *Scand J Work Environ Health* 7 (1981): suppl 4, 121—126.
28. Infante PF, Wagoner JK, Waxweiler RJ. Carcinogenic, mutagenic and teratogenic risks associated with vinyl chloride. *Mutat Res* 41 (1976) 131—142.
29. Kato H, Schull WJ. Studies of the mortality of A-bomb survivors. 7. Mortality, 1950—1978: Part I. Cancer mortality. *Radiat Res* 90 (1982) 395—432.
30. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic research: Principles and quantitative methods*. Lifetime Learning Publications, Belmont, CA 1982.
31. Koskela R-S. Occupational mortality and mortality in relation to selective turnover. *Scand J Work Environ Health* 8 (1982): suppl 1, 34—39.
32. Land CE, Tokunaga M. Induction period. In: Boice JD, Fraumeni JF, ed. *Radiation carcinogenesis: Epidemiology and biological significance*. Raven Press, New York, NY 1980, pp 421—436.
33. Lloyd JW, Ciocco A. Long-term mortality study of steelworkers. *J Occup Med* 2 (1969) 299—310.
34. Lundin FE, Archer VE, Wagoner JK. An exposure-time-response model for lung cancer mortality in uranium miners: Effects of radiation exposure, age and cigarette smoking. In: Breslow NE, Whittemore AS, ed. *Energy and health*. Society for Industrial and Applied Mathematics, Philadelphia, PA 1979, pp 243—264.
35. Mancuso TF, El-Attar AA. Dynamic changes in industrial cohort studies. *Ind Med Surg* 35 (1966) 1059—1067.
36. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 22 (1959) 719—748.
37. McMichael AJ. Standardised mortality ratios and the "healthy worker effect": Scratching below the surface. *J Occup Med* 18 (1976) 165—168.

38. McMichael AJ, Haynes SG, Tyroler HA. Observations on the evaluation of occupational mortality data. *J Occup Med* 17 (1975) 128—131.
39. McMichael AJ, Spirtas R, Kupper LL. An epidemiologic study of mortality within a cohort of rubber workers, 1964—72. *J Occup Med* 16 (1974) 458—464.
40. Miettinen O. Design options in epidemiologic research: An update. *Scand J Work Environ Health* 8 (1982): suppl 1, 7—14.
41. Miettinen OS. Standardisation of risk ratios. *Am J Epidemiol* 96 (1972) 383—388.
42. Miettinen OS. Confounding and effect modification. *Am J Epidemiol* 100 (1974) 350—353.
43. Miettinen OS, Cook EF. Confounding: Essence and detection. *Am J Epidemiol* 114 (1981) 593—603.
44. Moolgavkar SH. The multistage theory of carcinogenesis and the age distribution of cancer in man. *J Natl Cancer Inst* 61 (1978) 49—52.
45. Moolgavkar SH, Day NE, Stevens RG. Two-stage model for carcinogenesis: Epidemiology of breast cancer in females. *J Natl Cancer Inst* 65 (1980) 559—569.
46. Musk AW, Monson RR, Peters JM, Peters RK. Mortality among Boston firefighters, 1915—1975. *Br J Ind Med* 35 (1978) 104—108.
47. Olsen J. Some methodologic problems encountered in occupational health research. *Scand J Soc Med* 9 (1981) 19—24.
48. Olsen J, Sabroe S. Researching occupational mortality. *Scand J Soc Med* 7 (1979) 1—6.
49. Olsen J, Sabroe S. Health selection among members of a Danish trade union. *Int J Epidemiol* 8 (1979) 155—159.
50. Ott MG, Holder BB, Langner RR. Determinants of mortality in an industrial population. *J Occup Med* 18 (1976) 171—177.
51. Peto J, Doll R, Hermon C, Binns W, Clayton R, Goffe T. Relationships of mortality to measures of environmental asbestos pollution in an asbestos textile factory. *Ann Occup Hyg* 29 (1985) 305—355.
52. Peto R. Epidemiology, multistage models, and short-term mutagenicity tests. In: Hiatt HH, Watson JD, Winster JA, ed. *Origins of human cancer. Volume 4.* Cold Springs Harbor Laboratory, New York, NY 1977, pp 1403—1428.
53. Peto R, Roe FJC, Lee PN, Levy L, Clack J. Cancer and aging in mice and men. *Br J Cancer* 32 (1975) 411—426.
54. Rothman K. Induction and latent periods. *Am J Epidemiol* 114 (1981) 253—259.
55. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am J Epidemiol* 112 (1980) 467—470.
56. Schlesselman JJ. *Case-control studies: Design, conduct, analysis.* Oxford University Press, New York, NY 1982.
57. Seltzer CC, Jablon S. Effects of selection on mortality. *Am J Epidemiol* 100 (1974) 367—372.
58. Thomas DC. Statistical methods for analyzing effects of temporal patterns of exposure on cancer risks. *Scand J Work Environ Health* 9 (1983) 353—366.
59. Tola S, Hernberg S. Healthy worker effect. In: Chiazzze L, Lundin FE, Watkins D, ed. *Methods and issues in occupational and environmental epidemiology.* Ann Arbor Sciences, Ann Arbor, MI 1983, pp 85—92.
60. Vinni K, Hakama M. Healthy worker effect in the total Finnish population. *Br J Ind Med* 37 (1980) 180—184.
61. Walker AM. Declining relative risks for lung cancer after cessation of asbestos exposure. *J Occup Med* 26 (1984) 422—426.
62. Wang J-D, Miettinen OS. Occupational mortality studies: Principles of validity. *Scand J Work Environ Health* 8 (1982) 153—158.
63. Wen CP, Tsai SP. Anatomy of the healthy worker effect — A critique of summary statistics employed in occupational epidemiology. *Scand J Work Environ Health* 8 (1982): suppl 1, 48—52.
64. Whittemore AS. The age distribution of human cancer for carcinogenic exposure of varying intensity. *Am J Epidemiol* 106 (1977) 418—432.

Received for publication: 3 January 1986