Editorial

Some observations on the healthy worker effect

The "healthy worker effect" has, in recent years, been referred to as "an unfortunate term,"[1] "a poorly defined phenomenon,"[2] and "a popular, but conceptually vague, term."[3] To a certain extent these criticisms are justified in that no rigorous definition can be found. The term was first used by McMichael *et al* and their original observations perhaps best convey the general concepts that lie behind it: "There is a strong selection process at play, wherein, to be employable in an industrial workforce, an individual must be relatively healthy and active. This selection factor acts to produce a "healthy worker effect," such that, in an industry free of significant life-shortening hazards, death rates within the workforce in question will be less than in the general population."[4]

The healthy worker effect is a general term used to cover the biases that result when the mortality experience of occupational cohorts, which consist of individuals having achieved a certain standard of health, is compared with that expected on the basis of mortality rates in the general population. Since the general population includes the chronic sick, the unemployed, and early retirees, all of whom are known to have a worse than average mortality experience, national mortality rates will usually be higher than those of an occupational group.

These biases pose a serious problem when interpreting the findings of investigations into the effects of potentially adverse exposures on the mortality of a workforce. Real excesses in mortality which result from harmful exposures may be wholly or partially masked by an apparent reduction in mortality of the workforce when compared with the mortality experience of the general population. Against this shortcoming should be weighed the advantages of these analyses. National mortality rates are readily available and are based on large numbers, which is particularly important when investigating rare causes of death. In some cases this form of analysis may be the only option available. Even when alternative comparison groups exist however—for example, an unexposed group of employees in the same industry— analyses using national rates still provide an important source of supplementary information. In any case, the usefulness of analyses involving national rates would clearly be improved with a greater understanding as to how the biases that constitute the healthy worker effect operate.

How much is understood about the healthy worker effect? Various observations have been made regarding its behaviour. It has been found to depend on the age at risk of the employees, being more pronounced in younger employees.[5,6] It is more important for employees who remain in active employment than for those who leave or for retirees.[5–7] It also appears to exert its greatest effect during the period shortly after employment starts but diminishes as the duration of follow up increases.[5,6] This is often referred to as the "wearing off" of the healthy worker effect. A further important observation is that it is more important for some diseases than others. In particular it is thought to be more serious for diseases which have "a considerable prediagnosis symptomatology"—for example, respiratory diseases—than for diseases without such symptomatology, cancers for example.[8]

These observations are important because they suggest ways in which the biases introduced by the healthy worker effect might be minimised. For example, there may be advantages to be gained from ignoring a certain period of follow up of a cohort immediately after the start of employment, since this appears to be where the biases are greatest. Such strategies, however, are only as good as the observations on which they are based. For this reason, it is important to examine carefully the origins of the observations that have been made on the healthy worker effect.

A major problem that arises when attempting to verify any observation on the healthy worker effect is one of measurement. Most researchers investigating the healthy worker effect have examined the ratios of observed to expected mortality (O/E × 100, or standardised mortality ratios (SMRs)). These are obtained from person-years type analyses using follow up data on all employees entering an industry over a given period.[9] A recent review of the all cause SMRs obtained from 50 such occupational cohort studies resulted in a distribution with a median SMR of 89 with tenth and ninetieth percentiles of 69 and 101 respectively.[10] At first sight this might be taken to support the view that the healthy worker effect is more serious in some cohorts than others, reflecting the different health requirements obtaining in

different industries. To what extent, however, do such differences in SMRs obtained from different cohorts reflect socioeconomic differences in mortality[11] rather than the healthy worker effect per se? This is difficult to answer in practice because adjustment for social class in occupational cohort studies is rarely carried out, usually because the required national rates are not available. In the United Kingdom, for example, rates are not routinely available by age, calendar year, cause of death, and social class. Crude scaling of ratios of observed to expected mortality using social class specific SMRs obtained from the Decennial Supplement[12] is usually the only option available but is rarely attempted in practice. This illustrates, albeit at the crudest level, the problems in trying to separate the healthy worker effect from the confounding effects of socioeconomic factors.

The problem of confounding by social class also arises when attempts are made to examine data pertinent to the wearing off of the healthy worker effect. The reduction in deficits in mortality with increasing duration of follow up was clearly illustrated by Fox and Collier.[6] In their much cited analysis of a cohort of over 7000 men who were potentially exposed to vinyl chloride monomer they observed SMRs for all causes of 37·4, 62·9, 75·1, and 94·2 at 0–4, 5–9, 10–14, and $\geqslant 15$ years of follow up respectively. It is not difficult, however, to find instances where this does not appear to be the case. Wen, for example, cites an all cause SMR of 87 after a 41 year follow up of Texas refinery workers.[13] Whereas some such persistent deficits in mortality might result from losses to follow up, particularly when long periods of observation are concerned, some will reflect socioeconomic differences. Cohorts consisting of highly skilled employees, and therefore predominantly higher social class, will tend to experience below average mortality rates no matter how long they are followed up. How much of such persistent deficits should be attributed to a continuing healthy worker effect and how much to unadjusted socioeconomic confounding?

There is a further, perhaps more subtle, reason why SMRs approaching 100 with increased duration of follow up may not merely reflect a wearing off of the healthy worker effect. This arises from the fact that most occupational cohorts include individuals who have been exposed to some hazard or other. Some researchers have rightly questioned the degree to which such observed patterns in mortality are confounded by the effects of accumulated exposures rather than simply the diminution of the healthy worker effect with time.[6 14]

The problem of confounding is also an issue when trying to confirm the observation that the healthy worker effect is more serious for younger employees. This is again well illustrated by the data of Fox and Collier.[6] They observed SMRs of 62·2, 65·4, 79·3, and 95·7 for age at risk groups 25–44, 45–54, 55–64, and 65–74 respectively, suggesting that the healthy worker effect declines with increasing age at risk. When they stratified the SMRs by both age at risk and duration of follow up, however, the SMRs within quinquennia of follow up appeared, if anything, to *decrease* with increasing age at risk. Although based on only one (relatively small) cohort of employees, this finding does serve to suggest that simple univariate analyses may give misleading predictions for the behaviour of the healthy worker effect. Consideration of the effects of other relevant factors, such as age at hire and calendar year of hire, which appear to have been little studied with reference to the healthy worker effect, serve to complicate the picture further. For example, if recruitment practices change across time (calendar period of hire) this may confound any observed patterns seen with increasing duration of follow up, since those employed in earlier years would tend also to be those followed up for the longest periods.

An example of the sort of analysis that might help to disentangle these various effects was undertaken by Breslow using data on a cohort of smelter workers.[15] He examined the joint effects of period of hire, birthplace, years since employment, and level of arsenic exposure on SMRs for respiratory cancer. Using appropriate linear models, he found that the change in SMR with time since hire was largely explained by the effect of period of hire. Such multivariate analyses are required to examine properly the biases resulting from the healthy worker effect and the relative importance of the different selection effects in occupational cohorts.

A search for supportive empirical evidence for the observation that the healthy worker effect is more serious for some diseases than others is also hampered by the difficulties identified above. For example, in most occupational cohorts the SMR for all cancers will be closer to 100 than those for respiratory diseases after 10 or more years of follow up. Where persistent deficits in cancer mortality are observed, however, how much of these result from socioeconomic confounding? Some have argued all,[16] thus supporting the view that the healthy worker effect is not a problem for cancer mortality.[17] Despite the difficulties of trying to verify this in practice, there is an underlying rationale for the lack of a healthy worker effect for diseases such as cancer which have an asymptomatic prediagnosis phase. This rests on the premise that, at selection for employment, potential entrants to the workforce who are shortly to develop such diseases are not identifiable on the basis of their present state of health. Against this, however, it might be argued that selection for employment will also operate on *health related behaviours* such as

smoking and alcohol consumption. While not being selected directly on the basis of health, this may still (favourably) affect the subsequent mortality experience of the cohort, including mortality from cancer. Should such selection come under the umbrella of the healthy worker effect? The line between the effects of selection on such health related behaviours and selection on health status (respiratory function or blood pressure, for example) is unclear. Taking a "purist" definition of the healthy worker effect it might be argued that the relevant selection effect is health status alone. Since in practice, however, it is impossible to separate these two types of selection effects, this purist definition leads to a prediction regarding the behaviour of the healthy worker effect that is impossible to verify or refute.

In conclusion, the term healthy worker effect has proved useful, insofar as it has communicated the seriousness of the biases that result from selection effects arising in occupational cohorts. At the same time, however, its acceptance into reports of occupational epidemiology has perhaps given the impression that the biases relevant to comparisons of occupational cohort mortality with that of the general population are well understood.[18] This is far from true, especially given the difficulties in trying to verify its behaviour. More detailed examinations are required of the different components that contribute to and confound the selection effects encompassed by the healthy worker effect in addition to those that are not.

LUCY M CARPENTER
Epidemiological Monitoring Unit,
Department of Epidemiology,
London School of Hygiene and Tropical Medicine,
London WC1E 7HT.

References

1 McDonald JC, Liddell FDK, Gibbs GW, Eyssen GE, McDonald AD. Dust exposure and mortality in chrysotile mining, 1910–75. *Br J Ind Med* 1980;**37**:11–24.
2 Tola S, Hernberg S. Healthy worker effect. In: Chiazze L, Lundin FE, Watkins D, eds. *Methods and issues in occupational and environmental epidemiology.* Ann Arbor, Mich: Ann Arbor Science, 1983:85–92.
3 Hernberg S. Evaluation of epidemiologic studies in assessing the long-term effects of occupational noxious agents. *Scand J Work Environ Health* 1980;**6**:163–9.
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 McMichael AJ. Standardized mortality ratios and the "healthy worker effect": scratching beneath the surface. *J Occup Med* 1976;**18**:165–8.
6 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* 1976;**30**:225–30.
7 Gilbert ES. Some confounding factors in the study of mortality and occupational exposures. *Am J Epidemiol* 1982;**116**:177–88.
8 Weed DL. An epidemiological application of Popper's method. *J Epidemiol Community Health* 1985;**39**:277–85.
9 Berry G. The analysis of mortality by the subject-years method. *Biometrics* 1983;**39**:173–84.
10 Waxweiler RJ, Haring MK, Leffingwell SS, Halperin WH. Quantification of differences between proportionate mortality ratios and standardized mortality ratios. In: Peto R, Schneiderman M, eds. *Quantification of occupational cancer.* (Banbury report No 9) New York: Cold Spring Harbour Laboratory 1981:379–89.
11 Wilkinson RG, ed. *Class and health.* London: Tavistock Publications, 1986.
12 Office of Population Censuses and Surveys. *Occupational mortality decennial supplement 1979–80, 1982–3.* London: HMSO, 1986. (Series DS No 6.)
13 Wen CP, Tsai SP, Gibson RL. Anatomy of the healthy worker effect: a critical review. *J Occup Med* 1983;**25**:283–9.
14 Sterling TD, Weinkam JJ. Extent, persistence and constancy of the healthy worker or healthy person effect by all and selected causes of death. *J Occup Med* 1986;**28**:348–53.
15 Breslow N. Multivariate cohort analysis. *National Cancer Institute Monograph* 1985: 149–56.
16 Day NE. Statistical considerations. In: Wald NJ, Doll R, eds. *Interpretation of negative epidemiological evidence for carcinogenicity.* Lyon: International Agency for Research on Cancer. 1985:13–29.
17 Enterline PE. Not uniformly true for each cause of death. *J Occup Med* 1975;**17**:127–8.
18 Anonymous. Healthy worker effects. *Lancet* 1985;ii:1443.