Occup. Med. Vol. 49, No. 4, pp. 225–229, 1999

0962-7480/99

A review of the healthy worker effect in occupational epidemiology

C.-Y. Li* and F.-C. Sung†
**Department of Public Health, College of Medicine, Fu-Jen Catholic University, Hsinchuang, Taipei Hsien 24205, Taiwan, ROC; †Graduate Institute of Environmental Health, College of Public Health, National Taiwan University, Taipei 100, Taiwan, ROC*

This review article aims to anatomize sources of the healthy worker effect (HWE) and to summarize advantages and limitations of several approaches frequently proposed to eliminate the HWE. Although the HWE is frequently addressed in the context of selection bias, our review suggests that the selection of occupational cohorts with advantageous health status would preferably be addressed as a source of confounding biases. The authors also conclude that the exclusion of unhealthy workers at employment and the study of active workers are the two main sources of HWE, and that the use of the general population as a comparison group in occupational epidemiology should be avoided if possible. The authors encourage investigators to make distinctions between the underlying factors related to the use of the general population as the comparison group in occupational epidemiology.

Key words: Confounding; healthy worker effect; occupational epidemiology; selection bias; validity.

Occup. Med. Vol. 49, 225–229, 1999

Received 23 June 1998; accepted in final form 4 December 1998

INTRODUCTION

The healthy worker effect (HWE) is a term applied to the deficit of both morbidity and mortality ascribed to various employment-associated factors when workers and the general population are compared. First used by McMichael *et al.*,[1] the HWE reflects that an individual must be relatively healthy in order to be employable in a workforce, and both morbidity and mortality rates within the workforce are usually lower than in the general population. As a result, real excesses in both morbidity and mortality due to harmful exposures at work might be wholly or partially masked.[1] Although well-recognized, the HWE has been considered to be a poorly defined phenomenon and a popular but vague concept.[2–4] To a certain extent, the above-stated criticisms are justifiable as neither rigorous definition of the HWE nor a consensus upon how to deal with the HWE have been available. Additionally, although years of effort by occupational epidemiologists have been devoted to reducing or even eliminating the HWE, it remains one of the most annoying methodological difficulties in the study of occupational hazards and human health. This paper is a review of the publications dealing with the HWE in occupational studies, and systematically presents the sources, components, magnitude, effect modifiers and strategies for reducing the HWE.

Correspondence and reprint requests to: C.-Y. Li, Department of Public Health, College of Medicine, Fu Jen Catholic University, 510, Chung Cheng Rd, Hsinchuang, Taipei Hsien 24205, Taiwan. Tel: (+886) 2-29031111 ext. 3432; Fax: (+886) 2-29086229; email: chungyi@mails.fju.edu.tw

SOURCES AND COMPONENTS OF THE HEALTHY WORKER EFFECT

The HWE has long been considered as a source of selection bias.[5–8] It is true that there is a selection process of excluding unhealthy individuals from the workforce, and this selection process leads to a difference in health status between workers and the general population. From this perspective, in an industry free of significant life-shortening hazards, both morbidity and mortality rates within the workforce of interest are likely to be lower than that in the general population. In addition to bias due to the selection process at employment, occupational studies of both morbidity and mortality, which compare workers and the general population, appear to be influenced by additional sources of biases. For example, healthier workers are more likely to stay in the workforce than those who are sick, which may also give rise to a healthier occupational cohort. From this perspective, the HWE can be viewed as a consequence of selecting an occupational cohort with a process based on health and/

or survival effects. This review classifies the HWE and the other biases related to the comparison of workers and general population into election bias, information bias, and confounding bias as addressed below.

Selection bias

Many investigators have suggested that incomplete follow-up of the section of workers who leave employment could be a source of the HWE.[8–11] Such incomplete follow-up can be attributable to (1) good health required of workers for continued employment and (2) the tendency for those who develop diseases to leave their employment. Thus, if comparisons are made between workers who remain in employment during the time period of observation (*i.e.*, active workers) and the general population, the HWE may arise.

In addition to incomplete follow-up, lower morbidity and mortality rates of workers could be simply a consequence of an improper local *vs.* national comparison. If the worker population belongs to a region with better health conditions than large geographic areas or a nation as a whole, then regional differences in the occurrence of a particular disease may contribute to observed deficits of morbidity and mortality among workers.[12] Such regional differences may result from dissimilar qualities of health and clinical care and/or local peculiarities in diagnostic criteria, and have little to do with the good health of workers.[12]

Information bias

The comparison of both morbidity and mortality between workers and the general population might leave room for information bias. For example, the differences in both morbidity and mortality between an occupational cohort and the general population may arise from different criteria in the diagnosis of diseases, or from differences in the methods and quality of recording health outcomes between the two populations being compared.[13,14] The difference in the mortality ascertainment may entail different degrees of misclassification of diseases between populations. If the number of deaths among workers is under-ascertained, a study may report deficits in both morbidity and mortality among workers, and such deficits are unrelated to the selection process of an individual into the workforce.

Confounding bias

As mentioned earlier, many researchers consider the HWE to be a source of selection bias because it is a result of the selection process of relatively healthy individuals into industries. Such selection processes may have the following consequences. First, because people who are diagnosed with illnesses with a symptomatic pre-diagnostic phase are less likely to obtain employment, the selected workers may have a better-than-average health status. Second, employment regulations set by industries may restrict certain risk factors for diseases and causes of death. For example, some health-related behaviours, such as smoking, are not allowed during the work hours, and some personal traits, such as obesity, may be thought unfit for particular labour forces by industry.[15] A study exemplified such phenomena by demonstrating a significant deficit in lung cancer mortality among petrochemical workers in which the authors hypothesized that since smoking was prohibited in the workplace, the proportion of smokers in that group of workers was smaller than in the general population.[16] In addition to such selection processes, the differences in both morbidity and mortality between specific groups of workers and the general population may be attributed to non-comparability of socio-economic status. It has been suggested that once hired, workers in large industries may have greater access to medical services that protect them from diseases.[15] Thus, the HWE may reflect the selection of workforces for study rather than selection of individuals into workforce.

BEHAVIOUR OF THE HEALTHY WORKER EFFECT

In addition to the sources and components of the HWE, determining size and effect-modifiers of the HWE has been another challenge to researchers. Specifically, the HWE can be very serious in some studies, but may be moderate in others. We summarized, in the following section, factors affecting the HWE.

Causes of death

Many investigators have argued that the HWE is of little or no consequence in interpreting data on cancer mortality.[5,14,17,18] The reason for this is that it is unlikely that factors predicting eventual cancer deaths would be presented at 20 years of age, when many people become employed, which may not be true for factors that predict other causes of death. In other words, most cancers are not associated with a prolonged period of ill-health that would affect employability for a long time before death occurred. Although the verification of this argument is almost unfeasible empirically, it is quite reasonable to conceptually be convinced that the influence of the HWE should be relatively moderate for mortality studies of diseases with an absence of a prolonged disabling illness preceding death such as cancer.

Demographic factors

Fox and Collier analysed data collected in a cohort study of all men ever exposed to vinyl chloride monomer in manufacturing processes in Great Britain. The results showed that the standardized mortality ratio (SMR) for all causes was lower for younger workers than that for older workers, even after adjustment for length of employment.[5] This finding is contradictory to the common belief that older individuals seeking employment are healthier than individuals of the same age in the general population while it is not so obvious for younger people looking for employment. Thus, if age at start of employment modified the size of the HWE, it would favour the

older workers more than the younger workers. However, it is again not possible to empirically verify the above-stated argument. Even if the SMRs are found to be equal across the different age ranges, it provides no convincing evidence of equal operation of the HWE at different ages, and may only reflect the mix of possible different HWEs and different age-specific effects from occupational hazards. Additionally, Hernberg[19] argued that the HWE would have greater influence on male workers than on female workers, since women are less likely to be rejected from the workforce for poor health status than men are.

Types of occupational cohort

Different workforces usually have different hiring policies with respect to physical fitness and/or certain health-related behaviours, such as smoking. As a result, the HWE is likely to be different across industries. The HWE in the study of active employees can be even more serious since workers who remain in the workforce are generally healthier than those who are retired or disassociated from the workplace.[20] Thus, the HWE would tend to be more observable for physically demanding occupations than for those with little need of physical labour.

The time elapse

Some studies have noted that SMRs were approaching one with increased time before follow-up and concluded that the advantage of a health selection process at the initial stage disappeared gradually. Thus, for older age groups, the proportion of healthy persons in the general population and that in the occupational cohort would become more and more alike.[5,21] The decline of the HWE with time since first employment may be because the effect of selective exclusion from entry into work only operates during the period when an illness impairs employability. For example, a man who dies from chronic obstructive lung disease may have been too ill to obtain a job for 10 years before his death, but it is less likely to have restricted him from employment 40 years before his death. Breslow[23] demonstrated this in a cohort of smelter workers by examining the joint effects of time of hire, birth place, years since employment and levels of arsenic exposure on the SMRs for respiratory cancer mortality. He noted that the change of SMR with time was largely determined by time of hiring. Additionally, one could also speculate that the advantageous health status of workers at the start of employment would decline with the passage of time because the advantages resulting from the selection process would gradually disappear as a result of physical and psychological work pressure.

With the exception of the above-stated argument for the increase of SMR with time, researchers have different viewpoints on the change of SMR. Firstly, it has been suggested that the comparison of two age-adjusted SMRs should not be allowed unless there is homogeneity across strata of ratios of mortality rates in cohort 1 and cohort 2 and in the general population.[22] As time goes on, the age-specific mortality rate might change and the condition for comparing SMRs derived at two different points in time could be violated. Thus, the customary observation of the increase of SMR over time may not necessarily be due to the disappearance of HWE (or the increase in mortality risk), but is instead an artefact of SMR methodology.[9] Secondly, researchers have argued that the increase of SMR for certain diseases with time may simply be a consequence of accumulated hazardous exposures rather than the disappearance of the advantages of the selection process at employment.[9]

STRATEGIES FOR REDUCTION OF HEALTHY WORKER EFFECT

Among a variety of biases arising from the comparison between workers and the general population, the selection bias can be effectively minimized if studies include not only active workers but also pensioners and those who left work before retirement. Additionally, the information bias would be much less serious than the selection bias if an appropriate general population with comparably accurate information on both morbidity and mortality was identified. However, once the general population is used as a comparison population and the industry's recruitment of workers is based on health status or/and certain health-related behaviours, the confounding bias would invariably occur. The only way of adjusting for confounding bias is to conceive the baseline health status and risk factor distributions of the occupational cohort and of the general population, which is, unfortunately, unrealistic.

A number of strategies for minimizing the HWE have been frequently proposed. Nearly every strategy has its strengths and limitations and these are comprehensively summarized in the literature.[13] Among the strategies, 'use external work comparison groups' and 'use internal comparison groups', in our view, are the most methodologically plausible. Ideally, we should identify a theoretically correct external comparison population, or a representative sample from it, which comprises other employed persons who have entered and remained in the workforce through an equivalent selection process. Moreover, the correct external population should consist of workers from certain occupations who are comparable with the index occupation in terms of extraneous effects on the outcomes of interest.

In addition to 'use external work comparison groups', researchers may consider another strategy that examines variation in the health outcome rate across a gradient of increasing exposure within the workforce, *i.e.*, 'use internal comparison groups'. This strategy is justifiable in that employees from the same industry tend to experience a similar selection process, and they are likely to share a similar potential confounding effect. As a result, the presence of the HWE can be effectively controlled by comparing rates of the health outcomes of interest

between individuals with high exposure and those with low or no exposure. The nested case–control study is equivalent, in the sense of selecting controls, to the use of internal comparison groups.

DISCUSSION AND CONCLUSION

This review recognizes that the HWE is one of the most annoying methodological aspects in occupational studies which use the general population as the control group. The HWE has long been addressed in the context of selection bias since unhealthy people are more likely to be removed from employment and healthier individuals are more likely to remain in the workforce. An alternative viewpoint was offered by Monson[17] who argued that 'the HWE occurs mainly through confounding by the factor of "good health status" that is associated with both the outcome (morbidity or mortality) and the exposure (employment in the industry)'. Examining the sources of the HWE in this review further suggests that the loosely defined HWE should be considered a mixture of selection bias and confounding bias in as much as no single bias source is sufficient to characterize the HWE.

It is likely that healthier workers tend to remain in the workforce (the survival effect). Several studies have demonstrated such selection bias due to incomplete follow-up. Gilbert[24] studied nuclear facility workers and found an elevated mortality from all causes in short-term workers and terminated workers. Similarly, Fox and Collier reported that the mortality rate was approximately 50% higher among those who had left work than those who had remained working.[5] Several other studies have also showed that workers who left the workforce early,[25] changed occupation category[26] or retired early[27] experienced higher mortality rates. Wen and Tsai thus concluded that 'The HWE mostly characterizes actively employed workers and would be more accurately termed as the "active worker effect" '.[3] Apparently, the study of active workers may entail serious problems in occupational studies. Thus, an attempt to trace pensioners and persons disassociated with the industry should be made in occupational studies.

Among HWE-related biases, the disparity of health status between study cohorts and the general population (confounding bias) and the study of active workers (selection bias) are of great importance. Information on good health status in workers is somewhat subtle and is rarely obtainable. Thus, adjustment for such potential confounders is not feasible empirically. Additionally, data from previous studies have demonstrated that types of occupational cohort, causes of death and age at hire may modify the effect related to the selection process of health. These facts pose difficulties in creating confounder-adjustment methods suitable at all times for the elimination of the HWE.

Although sources of HWE can be well-conceived, the methods to minimize its effect appear to be unknown. The most straightforward way of avoiding or minimizing the HWE is not to use the general population as the reference group. However, many researchers still consider using the general population for comparison to be a very useful strategy not only for economic reasons but also because of difficulties in knowing or identifying proper reference groups. Additionally, the use of the general population may yield adequate statistical power, which is a common shortage encountered in other alternatives. We may conclude that it may be more acceptable to use the general population for comparison when the end point of interest is cancer mortality or morbidity, and the HWE is expected to be lower.

Among the alternatives for reducing the HWE, we consider 'use internal comparison groups' or 'use external work comparison groups' to be methodologically superior to the others. Individuals working in the same workforce are likely to be employed through the same selection process, which makes internal comparison one of the most effective ways of reducing the HWE. However, it is not likely that all occupational hazards pose gradient effects on human health. Additionally, such internal comparisons provide little information on the risk of diseases for workers as compared to the general population that might be of great interest in public health practices. Concerning 'use external work comparison groups', two potential disadvantages must be addressed.[8] First, the comparison work cohort might not be large enough to produce stable morbidity and mortality rates. Second, despite the fact that another occupational cohort is superior to the general population with respect to reducing biases from differential workers' health status and health selection process, differences in other extraneous variables such as demographic characteristics and an array of occupational hazards might produce confounding.

Feasibility is another issue in using an external work comparison group. Research has suggested that although an alternative occupational population could be more 'comparable', it is not necessarily 'better' since a different comparison population may yield different interpretation.[9] Additionally, almost all industries of which the health selection process is required are subject to an adverse health effect due to exposure to certain occupational hazards. Thus, it would be very difficult to find an occupational population without work-related hazards other than the one of interest among index occupational cohorts.

In summary, the HWE has not been sufficiently and systematically defined. Considering its multifaceted nature, it is very doubtful that such a crude term is useful at all.[3] This review also concludes that there is no particular comparison population preferable in all aspects and careful interpretation of the results from the studies using different comparison populations are strongly recommended. We suggest that investigators should not focus on the term 'healthy worker effect' and should perhaps try to make comprehensible distinctions between the different underlying factors related to the comparison between workers and the general population. This approach would be, we believe, a better way of gaining

insight into and handling HWE-related biases in occupational epidemiology.

REFERENCES

1. McMichael AJ, Spirats R, Kupper LL. An epidemiological study of mortality within a cohort of rubber workers, 1964–72. *J Occup Med* 1986; 18: 165–168.
2. Hernberg S. Evaluation of epidemiological studies in assessing the long-term effects of occupational noxious agents. *Scand J Work Environ Health* 1980; 6: 163–169.
3. Tola S, Hernberg S. Healthy worker effect. In: Chiazze L Jr, Lundin F, Watkins D, ed. *Methods and Issues in Occupational and Environmental Epidemiology.* Ann Arbor, MI (USA): Ann Arbor Science Publishers, 1983: 85–92.
4. Arrighi HM, Hertz-Picciotto I. Definitions, sources, magnitude, effect modifiers, and strategies of reduction of the healthy worker effect. *J Occup Med* 1993; 35: 890–892.
5. Fox AJ, Collier PF. Low mortality rates in industrial cohort studies due to selection for worker and survival in the industry. *Br J Prev Soc Med* 1976; 30: 225–230.
6. Wang JD, Miettinen OS. Occupational mortality studies, principles of validity. *Scand J Work Environ Health* 1982; 8: 153–158.
7. Wen CP, Tsai SP, Gibson RL. Anatomy of the healthy worker effect: A critical review. *J Occup Med* 1983; 25: 283–289.
8. Monson RR. Observations on the healthy worker effect. *J Occup Med* 1987; 28: 425–433.
9. Sterling TD, Weinkam JJ. Extent, persistence, and constancy of the health worker or healthy person effect by all and selected causes of death. *J Occup Med* 1986; 28: 348–353.
10. Bell CM, Coleman DA. Models for the healthy worker effect in industrial cohort. *Stat Med* 1987; 6: 901–909.
11. Ostlin P. The 'healthy-related selection effect' on occupational morbidity rates. *Scand J Work Environ Health* 1989; 17: 265–270.
12. Meijers JM, Swaen GM, Volovics A, Lucas LJ, Van Vliet K. Occupational cohort studies: the influence of design characteristics on the healthy worker effect. *Int J Epidemiol* 1989; 18: 901–909.
13. Industrial Disease Standards Panel (IDSP). *Reports to the Workers' Compensation Board on the Healthy Worker Effect.* Toronto, Canada: IDSP Report No. 3, 1988.
14. McMichael AJ. Standardized mortality ratio and the healthy worker effect: scratching beneath the surface. *J Occup Med* 1976; 81: 165–168.
15. Wilcosky T, Wing S. The healthy worker effect: selection of workers and workforces. *Scand J Work Environ Health* 1987; 13: 70–72.
16. Thériault G, Prvencher S. Mortality study of oil refinery workers: five-year follow-up. *J Occup Med* 1987; 29: 357–360.
17. Monson RR. *Occupational Epidemiology.* Boca Raton, FL (USA): CRC Press, 1990.
18. Enterline PE. Not uniformly true for each of cause. *J Occup Med* 1975; 17: 127–128.
19. Hernberg S. *Introduction to Occupational Epidemiology.* Chelsea, MI (USA): Lewis Publishers, Inc., 1991: 129–134.
20. O'Berg MT. Study of active workers. In: Chiazze L Jr, Lundin F, Watkins D, ed. *Methods and Issues in Occupational and Environmental Epidemiology.* Ann Arbor, MI (USA): Ann Arbor Science Publishers, 1983: 93–96.
21. Howe GR, Chiarelli AM, Lindsay JP. Components and modifiers of the healthy worker effect: evidence from the three occupational cohorts and implications for industrial compensation. *Am J Epidemiol* 1988; 128: 1364–1375.
22. Armstrong BG. Comparing standardized mortality ratios. *Ann Epidemiol* 1995; 5: 60–64.
23. Breslow NE, Day NE. *Statistical Methods in Cancer Research, Vol II.* Lyon, France: IARC Sci. Publ. No. 82, 1987.
24. Vinni K, Hakrama M. Healthy worker effect in the total Finnish population. *Br J Ind Med* 1980; 37: 180–184.
25. Olsen J, Sabroe S. Health selection among members of a Danish trade union. *Int J Epidemiol* 1979; 8: 155–159.
26. Delzell E, Monson RR. Mortality among rubber workers; IV. General mortality pattern. *J Occup Med* 1981; 23: 850–856.
27. Wen CP, Tsai SP. Anatomy of the healthy worker effect. A critique of summary statistics employed in occupational epidemiology. *Scand J Work Environ Health* 1982; 8(**Suppl 1**): 48–52.