IJE vol.33 no.5 © International Epidemiological Association 2004; all rights reserved.
Advance Access publication 20 May 2004

International Journal of Epidemiology 2004;**33**:1040–1045
doi:10.1093/ije/dyh151

METHODOLOGY

Quantitative evaluation of the effects of uncontrolled confounding by alcohol and tobacco in occupational cancer studies

David Kriebel,[1] Ariana Zeka,[1,2] Ellen A Eisen[1] and David H Wegman[1]

Accepted 30 January 2004

Background Uncontrolled confounding by personal exposures like smoking can limit the inferential power of occupational cohort studies. We developed and demonstrate a refinement of an existing type of sensitivity analysis, indirect adjustment, for evaluating the potential magnitude of confounding by alcohol and tobacco. Results of a large retrospective cohort study of laryngeal cancer and exposure to metalworking fluids (MWF) are used to illustrate the methods.

Methods Data on smoking and drinking habits representative of the study cohort were obtained from a sample of US manufacturing workers from the 1977 National Health Interview Survey (NHIS). Two different mechanisms were assumed to affect the distribution of confounding factors between MWF exposure groups: socially determined and chance differences. Chance variation was investigated with Monte Carlo sampling from the NHIS survey distribution of smoking and drinking. An upper bound on systematic differences in smoking and drinking was set by assuming that differences between exposure groups within the same unionized blue collar workforce were very unlikely to be larger than differences between blue and white collar manufacturing workers in the NHIS data.

Results Under plausibly large differences in smoking and drinking habits among MWF exposure groups occurring by either mechanism, the exposure–risk association was unlikely to have been over- or under- estimated by as much as 20%.

Conclusions When comparing exposure groups within the same working population, it is unlikely that either systematic or chance differences in smoking and drinking habits will cause as much as a 20% change in the relative risk in large studies. While this study focused on an occupational exposure and laryngeal cancer, there are many situations in which epidemiologists are concerned that unmeasured 'lifestyle factors' may differ among exposure groups, and it would appear that the likely confounding effect of such differences will often be modest.

Keywords Alcohol drinking, tobacco, confounding factors, laryngeal neoplasms, occupational exposure

Confounding is understood by epidemiologists to be perhaps the single most important limitation to causal inference in observational studies, and the statistical control of confounding to be among the most important tools of the epidemiologist's trade. When data are unavailable for a known potential confounder, however, our standard tools are not effective, and the researcher (or reader) often resorts to 'professional judgement' in evaluating how seriously compromised the study results may be. One common mistake that is made in these instances is the failure to appreciate that weak confounding cannot explain strong apparent exposure–disease associations.[1] But such a general principle, while correct, is not very useful without guidance on how to assess strength of potential confounding, in the absence of data.

[1] Department of Work Environment, University of Massachusetts Lowell, Lowell, MA, USA.

[2] Current affiliation: Environmental Epidemiology Program, Harvard School of Public Health, Boston, MA, USA.

Correspondence to: Dr D Kriebel, University of Massachusetts Lowell, 1 University Ave. Lowell, MA 01854, USA. E-mail: David_Kriebel@uml.edu.

This is one example of a kind of a 'what if' question that can often be usefully informed by sensitivity analyses conducted as a complement to standard epidemiological methods.[2,3] A particular type of sensitivity analysis, sometimes called indirect adjustment[4–6] or external adjustment,[3] has been used to assess potential confounding by smoking in the absence of individual confounder data, in occupational cohort studies. In this paper, we argue that the method of indirect adjustment could be used more frequently, and demonstrate extensions to it which may make it more widely accepted. An example is provided using results from a large retrospective cohort study of laryngeal cancer risk in automotive workers exposed to metalworking fluids (MWF), but the methods and findings are broadly applicable.

Occupational cohort studies are frequently conducted without individual data on potential confounding factors like smoking and alcohol consumption, because these studies typically rely on work records and disease registries and not on patient interviews. In a common scenario, researchers identifying an occupational exposure that appears to increase risk of a smoking-related cancer are criticized, and their findings challenged because of the possibility that differences in smoking habits among exposure groups may be the real cause of the association.[6–8] Epidemiologists have known for at least two decades that it is possible to estimate the magnitude of confounding that might occur in an occupational cohort study of a smoking-related cancer, given assumptions about the distribution of smoking habits.[4–8] Indirect adjustment involves calculating the confounding relative risk (RR_{co}), that would occur between different exposure groups in the study cohort, deriving solely from hypothetical but plausible differences in their smoking habits. An indirectly adjusted relative risk for the occupational exposure can then be calculated as: $RR_{adj} = RR_{crude}/RR_{co}$, following Miettinen's classic work on confounding.[9,10]

Several authors have used this method to evaluate the problem of potential confounding by smoking in occupational lung cancer studies[4–6,11,12] and the general conclusion has been that only modest confounding effects would be expected from differences in smoking habits among subcohorts with different levels of exposure to an occupational hazard. Using alternative approaches, including empirical evaluations, other authors have reached similar conclusions.[13–15] Despite this body of work, indirect adjustment is still not widely used, and occupational cancer studies continue to be criticized for failure to evaluate smoking confounding.[16] We think that one reason the method is not more widely applied is that it is difficult to decide upon a key assumption behind the sensitivity analysis: the degree to which smoking habits (or other 'lifestyle' factors) among different exposure groups might differ. We propose to at least partially resolve this difficulty by noting that there are two distinct ways that such differences might occur, and showing how each can be separately investigated. The two different explanations are: chance, due simply to the small numbers of subjects in different exposure groups; and systematic differences, due to unknown, probably socially determined, patterns of smoking within the cohort.

The methods are illustrated in an assessment of the magnitude of potential confounding by alcohol and tobacco in a study of laryngeal cancer and exposure to metalworking fluids (MWF). Elevated laryngeal cancer risk has been observed among workers with long-term exposure to MWF in a large cohort study of workers at three General Motors factories in Michigan.[17,18] Larynx cancer is strongly affected by alcohol and tobacco, with risks that are approximately multiplicative.[19] It seemed particularly important, therefore, to investigate the possibility that the observed associations with MWF might be explained by differences in smoking and drinking habits between heavily exposed and unexposed groups within the cohort.

Methods

To conduct indirect adjustment, two types of data are needed. First, a dose–response curve for the joint effects of alcohol and tobacco on cancer risk was developed from the published literature, using meta-regression methods. This work is described in a separate paper.[19] Tobacco and alcohol exposures were found to contribute additively to laryngeal cancer risk, in a log-linear model of the form:

$$\ln(OR) = \beta_{alcohol} * (grams\ of\ ethanol/week) + \beta_{tobacco} * (grams\ of\ tobacco/week) \quad \text{(equation 1)}$$

where: ln(OR) is the log of the odds ratio comparing the risk for those with a given level of alcohol and tobacco consumption with the risk among non-drinkers and non-smokers, $\beta_{alcohol}$ and $\beta_{tobacco}$ are the pooled measures of association. For laryngeal cancer, the coefficients of the model were: $\beta_{alcohol}$ = 0.19 per 100 g/week (robust standard error = 0.04), and $\beta_{tobacco}$ = 0.53 per 100 g/week (SE = 0.05). Note that an additive model of the logarithm of risk represents a multiplicative relation between the two risk factors. The studies which were used in the meta-regression analysis were all case-control studies, and because they all appeared to have used incidence density sampling, we assumed that the published odds ratios were valid estimates of incidence density ratios, or relative risks.

The second requirement was for data on smoking and drinking habits representative of those in the study cohort. Data on these behaviours among a stratified random sample of US workers were obtained from the National Health Interview Survey (NHIS).[20,21] A sample representative of the autoworkers study cohort was selected by using NHIS data for blue collar manufacturing workers between 20 and 65 years of age. For the present study only males were used for the analyses, due to very small numbers of females in the MWF cohort. We chose to use the data from 1977 because it was the survey with alcohol and tobacco consumption data closest in time to the middle of the period covered by our epidemiological study of MWF and cancer risk (1940–2000). The NHIS data provided the frequency and quantity of alcoholic drinks and tobacco consumed by the individuals studied in the NHIS. We converted the consumption data into consistent units of grams per day, based on published conversion factors: 1 drink = 30 ml hard liquor = 150 ml wine = 330 ml beer = 11.87 g of pure ethanol[22]; and 1 cigarette = 0.2 cigars = 0.4 pipes of tobacco = 1 g tobacco.[23]

Confounding by systematic differences

To quantify how much confounding could occur due to systematic differences in smoking and drinking habits, it was assumed that differences in smoking and drinking between groups within a factory who have different levels of exposure to MWF, or some other toxicant, could not plausibly be larger than the differences reported between blue collar and white collar manufacturing workers in the NHIS data. The differences in smoking and drinking habits between these two classes are well-known, and result from a complex set of strong social

determinants, probably linked to education, social supports, and perceptions of risk. It therefore seems very unlikely that different groups of blue collar workers within the same region, corporation and union, could have systematic differences in smoking and drinking habits as large as the observed blue collar–white collar differences. We therefore calculated the average tobacco and alcohol consumption for blue and white collar workers, and then used these figures and equation 1 to estimate the relative risks of cancer comparing the average blue collar or white collar worker with a non-drinking, non-smoking reference group. The ratio of the resulting relative risks (RR) for blue and white collar workers provided an estimate of RR_{CO}, the confounding relative risk due to the 'larger than plausible' difference in smoking and drinking habits between blue and white collar workers.

To illustrate the magnitude of the class differences, the prevalences of smoking and drinking habits for blue and white collar workers were also calculated for joint categories of tobacco and alcohol consumption. The following categories were used for alcohol: 0, 0–4, >4 drinks/day, and tobacco: 0, 0–30, >30 cigarettes/day. The strength of the potential confounding from smoking and drinking differences among groups with different consumption patterns were assessed by combining the alcohol and tobacco consumption prevalence data from the NHIS survey with effect estimates from equation 1 to estimate relative risks for different categories of alcohol and tobacco consumption among white male blue-collar workers.

Confounding by chance differences

If groups of workers who experience different exposures (to MWF or some other toxicant) are compared in an epidemiological study, then by chance, these groups may have different average smoking and drinking behaviours. The magnitude of confounding that might occur from chance variation was investigated by Monte Carlo simulations, in which samples were drawn repeatedly from the distribution of smoking and drinking in the blue-collar manufacturing worker NHIS sample, representative of the MWF worker cohort. These simulations were based on random samples or subcohorts of 10, 25, and 100, selected with replacement, from the blue-collar white male NHIS sample. These three sample sizes were chosen to represent common study sizes: exposure groups of 10 are found in small epidemiological studies, 25 might be expected in each group in a medium-sized study, and larger studies might have 100 subjects in each of the exposure groups. Two samples of a given size (10, 25, or 100) were selected simultaneously. For each pair of samples their survey-weighted average alcohol and tobacco consumption figures were calculated and then used in equation 1 to predict relative risks. The ratio of the estimated effects for each pair of samples was then obtained; and it represents an estimate of RR_{CO} due to chance variations in smoking and drinking habits among exposure groups. This procedure was carried out for a thousand pairs of trials of each sample size.

Results

From the NHIS survey of 1977 we selected all 1335 white male blue-collar workers and all 1433 white male white-collar workers between ages 20–65. The prevalences of smoking and drinking are shown in Table 1 (note that the actual calculations of RR_{CO} described below used the mean consumption from continuous data—the categories in Table 1 were used for presentation only). A majority of both classes of workers fell into just two cells of the matrix: those who drank >0 and <4 drinks/day, and were either non-smokers or smokers of up to 1.5 packs/day (30 cigarettes per day). Unsurprisingly, non-smokers who drank heavily, and non-drinkers who smoked heavily were uncommon—the prevalence of either of these combinations was less than 2%. When comparing the two classes of workers, the most striking difference was in the non-smoking, moderate drinking category, where there was 40.5% of the white collar workers, and only 26.7% of the blue collar workers. Blue collar workers were nearly twice as prevalent as white collar workers in the heavy alcohol consumption categories (except for the rare combination of non-smoking heavy drinkers).

Table 1 Prevalence of smoking and drinking habits. 1977 US National Health Interview Survey: Per cent of respondents by smoking and drinking habits. 'Blue collar' (top row; N = 1335) and 'white collar' (2nd row; N = 1433) white male manufacturing workers, ages 20–65 years

Tobacco consumption (cig/day)	Alcohol consumption (drinks/day)		
	0	>0–4	4+
0	7.6	26.7	1.7
	7.5	40.5	1.4
>0–30	6.7	37.6	5.8
	2.7	32.9	3.2
30+	1.8	9.1	3.1
	1.1	9.6	1.1

Table 2 Laryngeal cancer risks comparing categories of smoking and drinking. Rate ratios estimated relative to the intermediate category, which was also the most prevalent among blue collar workers (N = 1335)

Tobacco consumption (cig/day)	Alcohol consumption (drinks/day)		
	0	>0–4	4+
0	0.42	0.57	2.04
>0–30	0.73	1	3.55
30+	4.49	6.16	21.87

Laryngeal cancer risk was estimated to rise rapidly with either alcohol or tobacco consumption (Table 2). Compared with the 'typical' blue collar worker who was a drinker of <4 drinks/day and a smoker of <1.5 packs/day, those in the heavy smoking/drinking category were estimated to have more than a 20-fold increased risk for laryngeal cancer. Non-drinkers and non-smokers had about a 60% reduction in risk compared with the typical blue collar worker.

The average alcohol and tobacco consumption figures from the NHIS survey of blue and white collar workers were then used in equation 1 to estimate the expected RR comparing each of these two groups with non-drinkers and non-smokers. For blue collar workers, the RR was 2.22, while for white collar workers it was 1.89. These figures mean that, taken as a group, the smoking and drinking habits in 1977 of blue collar manufacturing workers conferred upon them slightly more than a twofold increased risk of laryngeal cancer, compared with those who neither drank nor smoked. For white collar workers, the RR was somewhat less. The ratio of these two figures, 1.17, is an estimate of RR_{CO}. If two subcohorts in a study were compared, and the only difference between them was that their respective smoking and drinking habits were typical of blue and

white collar workers respectively, then one would expect to see a RR of 1.17 for laryngeal cancer due solely to alcohol and tobacco. We consider it highly improbable that two subcohorts of blue collar workers in a typical occupational cohort study would differ in their smoking and drinking habits as much as do blue and white collar workers, and so we conclude that confounding by systematic differences in these habits would most likely cause less than a 17% increase in the RR.

The potential for chance differences in smoking and drinking to confound is illustrated by the empirical 90% and 95% CI for RR_{co} shown in Table 3, as well as in Figure 1. These show, for example, that when comparing small exposure groups or subcohorts (n = 10), RR_{co} might range from 0.51 to 1.94, with 95% confidence. As expected, tighter CI were obtained as the sample size increased. In a large study (n = 100), chance differences would have a much more limited effect on RR_{co} (95% CI: 0.82, 1.22).

The case of MWF and laryngeal cancer

These sensitivity analyses can be applied to the results of a recent cohort study of MWF and laryngeal cancer risk.[18] A case-cohort analysis was conducted in a cohort of 46 400 automotive workers who had worked at least 3 years at one of three automotive plants prior to 1 January 1985. All cohort members were blue collar workers, and members of the same union. Follow-up started in 1940 and ended in 2000, and included over 1.5 million person-years. Seventy-eight incident cases of laryngeal cancer were identified by the Michigan Cancer Registry from 1 January 1985 through 1 January 2000. A 10% random sampling of the whole MWF male cohort was selected as the referent group. Work histories of study participants and historical air monitoring data were used to generate lifetime cumulative MWF exposure estimates.[24] Cox proportional hazards models were fit, comparing subgroups with varying levels of cumulative exposure. All models were controlled for potential confounding by race, plant, birth cohort, year started work, and year entered the cohort. None of these factors appeared to change the effect estimates, however.

Laryngeal cancer incidence was found to be associated with exposure to straight MWF, with a linear relationship between cumulative exposure and log of relative risk.[18] Using a continuous measure of cumulative exposure, the relative risk was 1.08 for each 5 mg/m^3-years increment in straight MWF exposure, more than 10 years before diagnosis (95% CI: 1.03, 1.14).[18] In this cohort, the top 1% of the distribution of cumulative exposure to straight MWF included those with levels >45 mg/m^3-years, at least 10 years before diagnosis. At this exposure level, the predicted relative risk of laryngeal cancer was 1.9 (95% CI: 1.1, 3.4), compared with unexposed workers. Data on alcohol and tobacco consumption within this cohort were not available.

The assessments of the potential impacts of both systematic and chance differences in smoking and drinking habits presented above can be applied to these findings. The confounding effect of systematic, socially determined, differences in habits among exposure groups were evaluated using the 'larger than plausible' confounding effect which would result if all heavily exposed workers had the smoking and drinking habits of blue collar workers, and all unexposed workers the habits of white collar workers. The reverse situation is equally likely of course, and so our estimate of RR_{co} from systematic differences is not more than 1.17, and not less than 1/1.17 = 0.86.

Chance differences would yield roughly similar bounds on the confounding relative risk in this study because the MWF exposure strata contained hundreds of workers (Table 3). An extreme upper bound estimate for the combined effects of simultaneous chance and systematic differences in drinking and smoking habits among MWF exposure groups might be approximately $RR_{co} = 1.17 \times 1.22 = 1.43$, while at the other extreme, $RR_{co} = 0.86 \times 0.82 = 0.71$. An 'extreme case' estimate of the unconfounded RR for MWF exposure and laryngeal cancer risk therefore might be approximately $RR_{adj} = 1.9/1.4 = 1.3$, although it is just as likely that chance and systematic differences in smoking and drinking have biased the unadjusted RR *towards* the null, so that the less confounded relative risk would actually be larger than the observed RR of 1.9: $RR_{adj} = 1.9/0.7 = 2.7$. At neither of these extreme bounds of the range of adjusted relative risks was the MWF–laryngeal cancer association reduced to RR = 1.0, and so we conclude that uncontrolled confounding is very unlikely to fully explain the observed MWF–larynx cancer association in this cohort.

Table 3 90% and 95% empirical CI for confounding relative risk (RR_{co}) from Monte Carlo simulations

Sample size	90% CI		95% CI	
	Lower limit	Upper limit	Lower limit	Upper limit
10	0.63	1.68	0.51	1.94
25	0.73	1.37	0.68	1.56
100	0.86	1.17	0.82	1.22

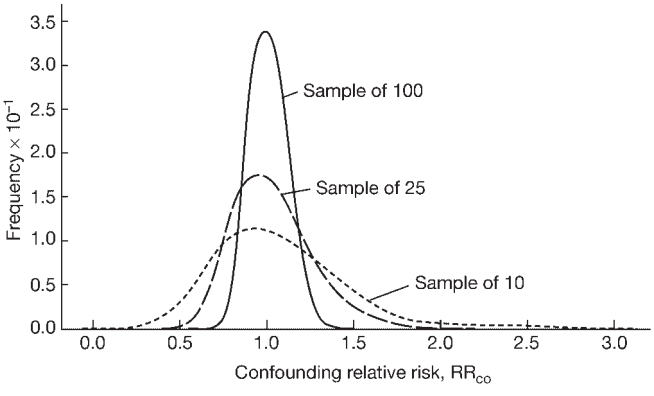

Figure 1 Monte-Carlo simulations (N = 1000) for subcohorts of three different sample sizes: 10, 25, and 100. The graph shows the frequency of alternative values of the confounding relative risk for laryngeal cancer, (RRco), comparing two groups of equal size and resulting from chance differences in alcohol and tobacco consumption

Discussion

Occupational cohort studies are frequently conducted without individual data on potential confounding factors like smoking and alcohol consumption.[25] Although unmeasured confounding may either increase or decrease the risk estimates, the concern is reasonably greatest in the presence of a modestly elevated relative risk. In such studies, concerns are often raised that the occupational exposure–disease associations may be explained by differences in smoking, drinking, or other lifestyle habits among exposure groups.[6–8]

A recent paper by Cole and colleagues[16] is a useful example. The paper consists of a critique of findings in several epidemiological

studies by Hansen *et al.* concerning cancer risk among Danish asphalt workers.[26–28] In Hansen's cohort mortality study, asphalt workers were found to have elevated risks of dying of cancers of the mouth, rectum, liver, larynx, and lung, all with rate ratios greater than 3.0, compared with a comparison group of unskilled service workers. No control for alcohol or smoking was performed in Hansen's study, which used census and mortality register data. Cole and colleagues critiqued these findings, arguing that uncontrolled confounding by tobacco and alcohol was a likely explanation for the findings. They wrote:

> We suggest that alcohol is a major confounder causing much of the excess disease that occurs among [asphalt] workers. Smoking, risk-taking behavior, and lesser factors are other confounders. Data of high quality are required to remove the effect of a confounder from the findings on any particular disease. Such data . . . should be subject-specific, valid, precise, and . . . highly similar for the two groups.

Further on, they state: '. . . even with ideal data, a strong confounding effect, such as that of smoking with lung cancer, is almost impossible to eliminate.' Our point in citing this example is not to argue that Hansen's findings are, or are not, confounded, but to point out that a fairly simple sensitivity analysis, in the form of indirect adjustment, would raise the level of the debate over these and similar findings. We argue, in contrast to Cole, that 'subject-specific' data on confounders are not necessarily needed to evaluate potential confounding.

Others have made a similar argument. Axelson, Steenland, and colleagues[4–6] are most frequently cited as the originators of the method of indirect adjustment, although earlier work making a similar point can be cited.[29–32] Axelson argued that only substantially different distributions of the confounding factor among exposed groups would cause a considerable confounding effect, even for strong confounder–disease associations, as in the case of smoking and lung cancer risk.[6]

We approached estimation of the magnitude of uncontrolled confounding in a new way. We examined quantitatively the boundaries within which hypothetical, but plausible confounding effects might occur. We chose distributions of smoking and drinking in the US population[20] stratified to resemble the MWF cohort. We assumed two distinct mechanisms that would affect the confounding factor distribution among exposure groups: systematic differences (socially determined), and chance differences (sample size effects). Considering a more complicated scenario than other investigators, we examined the joint distribution of two confounding factors: alcohol and tobacco consumption.

In this large MWF-exposed cohort, an approximately twofold increased risk of laryngeal cancer was found when comparing heavily exposed with unexposed workers.[18] We estimated an extreme, or worst case confounding scenario as follows:

> the heavily exposed MWF subcohort had the smoking and drinking habits of blue collar workers, while the unexposed had the habits of white collar workers; and
>
> chance differences in these habits between exposure groups were greater than would be expected in all but 2.5% of hypothetical repetitions of the study.

Under these conditions, the adjusted RR for the MWF– laryngeal cancer association was reduced from 1.9 to 1.3, comparing the upper 1% of the exposure distribution with the unexposed. An equally likely extreme was that the unadjusted RR was biased *towards* the null by confounding, and so the adjusted RR might be as high as 2.7. It is highly unlikely that this degree of confounding—in either direction—has occurred, and so we conclude that confounding by alcohol and tobacco is very unlikely to fully explain the observed MWF larynx cancer association in this cohort.

In this paper, we have focused on laryngeal cancer and uncontrolled confounding by alcohol and tobacco, but we believe that the broad conclusions are more generally applicable. There are many situations in which epidemiologists are concerned that unmeasured 'lifestyle factors' may be confounding a study. Following our findings, it would appear that the likely effects will often be modest. We note that the 20-fold increase in laryngeal cancer risk for heavy smokers/drinkers in Table 2 will not often be seen for the lifestyle factor(s) that one worries may be associated with the exposure under study. And so it would seem even less likely that serious confounding would occur. The method of indirect adjustment should be used more frequently in occupational studies, and can be applied more broadly, whenever risk estimates and exposure distributions for a known potential confounder can be obtained.

Acknowledgements

Supported by a grant from the US National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, R01-OH03575.

KEY MESSAGES

- Occupational cohort studies often are unable to gather data on individual habits like smoking and drinking, and so may be biased by uncontrolled confounding.
- Indirect adjustment, a type of sensitivity analysis, can be used to evaluate the potential magnitude of confounding by unmeasured covariates like alcohol and tobacco.
- In an occupational cohort study of a tobacco/alcohol-related cancer, it is unlikely that either systematic or chance differences in smoking and drinking habits will confound by more than about 20%.
- The specific example presented was an occupational cohort study of laryngeal cancer, but the methods can be used to assess potential confounding in other situations in which unmeasured 'lifestyle factors' may differ among exposure groups.

References

[1] Kleinbaum D, Kupper L, Morgenstern H. *Epidemiologic Research*. Belmont, CA: Lifetime Learning Publications, 1982.

[2] Greenland S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol* 1996;**25:**1107–16.

[3] Rothman K, Greenland S. *Modern Epidemiology*. Boston: Little Brown, 1998.

[4] Steenland K, Beaumont J, Halperin W. Methods of control for smoking in occupational cohort mortality studies. *Scand J Work Environ Health* 1984;**10:**143–49.

[5] Axelson O, Steenland K. Indirect methods of assessing the effects of tobacco use in occupational studies. *Am J Ind Med* 1988;**13:**105–18.

[6] Axelson O. Confounding from smoking in occupational epidemiology. *Br J Ind Med* 1989;**46:**505–07.

[7] Blair A, Steenland K, Shy C, O'Berg M, Halperin W, Thomas T. Control of smoking in occupational epidemiologic studies: methods and needs. *Am J Ind Med* 1988;**13:**3–4.

[8] Blair A, Stewart WF, Stewart PA *et al*. A philosophy for dealing with hypothesized uncontrolled confounding in epidemiological investigations. *Med Lav* 1995;**86:**106–10.

[9] Miettinen O. Components of the crude risk ratio. *Am J Epidemiol* 1972;**96:**168–72.

[10] Axelson O. Aspects on confounding in occupational health epidemiology. *Scand J Work Environ Health* 1978;**4:**85–89.

[11] 't Mannetje A, Kogevinas M, Chang-Claude J *et al*. Smoking as a confounder in case-control studies of occupational bladder cancer in women. *Am J Ind Med* 1999;**36:**75–82.

[12] Hooiveld M, Spee T, Burstyn I, Kromhout H, Heederik D. Lung cancer mortality in a Dutch cohort of asphalt workers: Evaluation of possible confounding by smoking. *Am J Ind Med* 2003;**43:**79–87.

[13] Siemiatycki J, Wacholder S, Dewar R, Cardis E, Greenwood C, Richardson L. Degree of confounding bias related to smoking, ethnic group, and socioeconomic status in estimates of the associations between occupation and cancer. *J Occup Med* 1988;**30:**617–25.

[14] Flanders WD, Khoury MJ. Indirect assessment of confounding: graphic description and limits on effect of adjusting for covariates. *Epidemiology* 1990;**1:**239–46.

[15] Levin LI, Silverman DT, Hartge P, Fears TR, Hoover RN. Smoking patterns by occupation and duration of employment. *Am J Ind Med* 1990;**17:**711–25.

[16] Cole P, Green L, Lash T. Lifestyle determinants of cancer among Danish mastic asphalt workers. *Reg Tox Pharm* 1999;**30:**1–8.

[17] Eisen EA, Tolbert PE, Hallock MF *et al*. Mortality studies of machining fluid exposure in the automobile industry. III: A case-control study of larynx cancer. *Am J Ind Med* 1994;**26:**185–202.

[18] Zeka A, Eisen EA, Kriebel D *et al*. Risk of upper aerodigestive tract cancers in a case-cohort study of autoworkers exposed to metal-working fluids. *Occup Environ Med* (In press).

[19] Zeka A, Gore R, Kriebel D. Effects of alcohol and tobacco on aero-digestive cancer risk: a meta-regression analysis. *Cancer Causes Control* 2003;**14:**897–906.

[20] Anonymous. *The National Health Interview Survey*. US National Center for Health Statistics. Washington DC: Department of Health and Human Services, 1977.

[21] National Center for Health Statistics, MG Kovar and GS Poe. The National Health Interview Survey design, 1973–84, and procedures, 1975–83. Vital and Health Statistics. Series 1, No. 18. DHHS Pub. No. (PHS) 85-1320. Public Health Service. Washington DC: US Government Printing Office, 1985.

[22] Anonymous. *Alcohol Drinking*. Volume 44 in International Agency for Research on Cancer Monographs on the Evaluation of Carcinogenic Risks to Humans. Lyons: IARC, 1988.

[23] Flanders WD, Rothman KJ. Occupational risk for laryngeal cancer. *Am J Public Health* 1982;**72:**369–72.

[24] Hallock MF, Smith TJ, Woskie SR, Hammond SK. Estimation of historical exposures to machining fluids in the automotive industry. *Am J Ind Med* 1994;**26:**621–34.

[25] Checkoway H, Pearce NE, Kriebel D. *Research Methods in Occupational Epidemiology, 2nd Edn.* Oxford: Oxford University Press, 2004.

[26] Hansen ES. Cancer incidence in an occupational cohort exposed to bitumen fumes. *Scand J Work Environ Health* 1989;**15:**101–05.

[27] Hansen ES. Cancer mortality in the asphalt industry: a ten year follow up of an occupational cohort. *Br J Ind Med* 1989;**46:**582–85.

[28] Hansen ES. Mortality of mastic asphalt workers. *Scand J Work Environ Health* 1991;**17:**20–24.

[29] Cornfield J, Haenszel WH, Hammond EC, *et al*. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst* 1959;**22:**173–203.

[30] Bross IDJ. Pertinency of an extraneous variable. *J Chron Dis* 1967;**20:**487–95.

[31] Schlesselman JJ. Assessing effects of confounding variables. *Am J Epidemiol* 1978;**108:**3–8.

[32] Gail MH, Wacholder S, Lubin JH. Indirect corrections for confounding under multiplicative and additive risk models. *Am J Ind Med* 1988;**13:**119–30.