Editorial

Confounding from smoking in occupational epidemiology

The recognition of smoking as a cause of lung cancer and other disorders was a great victory for epidemiology, although somewhat Pyrrhic perhaps, when other epidemiological relations are questioned in favour of smoking, thought by many to explain almost any disease. So not only King Pyrrhus of Epirus comes to mind but also the philosopher Pyrrho of Elis and his questioning of infallible knowledge. Still, there are at least some guidelines that may help our thinking about disease, smoking, and other factors affecting human health.

When smoking is suggested as an alternative or contributory explanation to some epidemiological relation of exposure and disease, the first requirement has to be some evidence that smoking can cause the disorder. There must also be more or heavier smokers among the exposed than the unexposed reference population. Confounding, however, may not only be positive and an exaggeration but also negative and masking, as when the occupational group under study has smoked less than the reference population.

Another aspect needing consideration, as this is often confused with confounding but is actually fundamentally different, is the interaction between smoking and occupational exposure. The result may be stronger than the simple additive combination of effects of smoking on the one hand and industrial exposure on the other. Such combined effects may even have contributed to the discovery of some industrial health hazards. For example, lung cancer due to exposure to asbestos would probably have been difficult to recognise in the absence of smoking. But there is also the reverse aspect—namely, that if a widespread risk factor such as smoking raises the baseline occurrence of a disease—an additive contribution from other factors becomes relatively more marginal and therefore difficult to detect. Again, this phenomenon has nothing to do with confounding, although modifying the magnitude of the effect of the exposure in terms of risk ratio.

It follows that the potential for confounding by smoking may be recognised simply by determining the pattern of smoking in the exposed group as well as in the reference population. This applies irrespective of whether a direct comparison with a particular reference group is made or the exposed numbers are provided by disease or death rates in the general population, as in the traditional cohort study. If the proportion of smokers in the compared populations proves to be much the same and similarly distributed in relation to other determinants of the disease, such as age and sex, there is no need for concern about confounding from smoking, not even in a study of lung cancer. Usually, however, there are some differences in smoking between the exposed and the reference population, unless the study is of cohort type and smoking has been carefully matched. But as will be seen below, quite substantial differences are needed to cause more than a moderate influence from confounding on the risk ratio, even in a study of lung cancer.

Whenever smoking data can be obtained it is certainly wise to consider this factor in an epidemiological analysis, but commonly the information is lacking. This should not prevent sound conclusions being drawn from a study, since approximations based on assumptions or partial information about smoking are usually sufficient for the purpose.

For example, using a model, and assuming that a population under study has 30% non-smokers, 50% moderate smokers (with a risk ratio of 10 in relation to non-smokers), and 20% heavy smokers (with a risk ratio of 20), the overall effect from confounding is not impressive when the reference population has 50% non-smokers, 40% moderate smokers, and 10% heavy smokers; the risk ratio from confounding will be only 1.43. On the other hand, should the study population have 70% non-smokers and only 30% moderate smokers by comparison with the same reference population, negative confounding would result and reduce the risk ratio to 0.57. Thus confounding from smoking would rarely explain risk ratios of more than about 1.5 for lung cancer and even
smaller excesses for other disorders much less strongly related to smoking. This point may be generalised to suggest that uncontrolled confounding from other general life style factors would usually be rather weak in occupational health studies.

In view of the interest in smoking as a confounder in occupational epidemiology, several studies have focused on the extent to which smoking might vary between occupational groups. Usually, smoking habits are fairly similar among various groups of blue collar workers but, over a broad range of occupational groups, Asp reported from 7% to 49% non-smokers, with an average of 24%. Accounting for the risk of both smokers and ex-smokers, the confounding risk ratios calculated for various groups with a range of non-smokers from 10% to 41%, were found to vary from 0.67 to 1.31 for lung cancer, compared with the average population. Figures of this sort are obviously specific for regions and countries and might change considerably over time, especially as more and more people now realise the health hazards and are giving up smoking. Nevertheless, such studies indicate the potential magnitude of the effect that may be expected as a result of confounding from smoking.

Other approaches to the evaluation of epidemiological data have been proposed to obtain some idea of the possible influence of confounding from smoking, when specific information is missing. One possibility is to examine the range of smoking related disorders, since a general increase in risk ratios for all these conditions would suggest an effect attributable to smoking rather than some other exposure. A counter argument would be that the occupational exposure might also have multiple and somewhat similar effects to smoking on cancer and other diseases.

The enrolment of an unexposed comparison cohort from the same company to evaluate possible confounding from smoking, and to check general comparability with national or regional rates, may be helpful if there is good agreement with expectation. A lack of serious confounding could also be indicated by a clear dose response between the effect and exposure as a dose response relation is unlikely to depend entirely on uncontrolled confounding. Some correlation might occur, however, between cigarette pack-years and exposure-years to a particular agent, implying confounding. There is also the possibility that a weak dose response phenomenon could be obscured by negative confounding or modified as a result of the influence of smoking on baseline morbidity or mortality in the various exposure categories.

The control of confounding from smoking (or any other factor), creates a special problem in case-referent (case-control) studies. Random variation in the process of selecting referents from the background or "base" population, from which the cases were drawn, may sometimes lead to a different distribution of the confounding factor among the exposed and non-exposed referents than in the corresponding sectors of the base population—assuming a state or simple, dichotomous exposure. Thus positive, negative, or lack of confounding may obtain in case-referent data irrespective of what is true for the base population. A reasonable number of referents may offer some remedy by more properly reflecting the distribution of various factors in the base population; no better control is obtained from matching, although this is commonly carried out. In principle, adjustments in the analysis of case-referent data, aimed at the control of confounding, may even take an adjusted risk ratio further away from the truth than the crude risk ratio if, for example, positive confounding in the base population happens to come up negative in the case-referent data, or vice versa.

This effect of random variation in the identification of referents is worth considering, especially when planning a case-referent study within a cohort with the object of obtaining risk estimates adjusted for smoking or other factors. The utility of such an approach may even be questionable, owing to the influence of random variation on the sample of referents if this involves only some 50–100 subjects. Indirect adjustment of the risk ratio for confounding may therefore be as accurate and possibly more cost efficient.

It should perhaps be emphasised that poor assessment of exposure and of a confounding factor have different consequences. Thus crude information leading to poor discrimination between those considered exposed and non-exposed would tend to drive the risk ratio towards unity—that is, mask an effect and remove evidence of a dose response relation. A poor assessment of smoking or other confounding factors leads to remaining, uncontrolled confounding, with risk estimates exaggerated by positive and decreased by negative confounding. This would occur if only broad categories of cigarette consumption are used, or a simple dichotomy into smokers and non-smokers. This aspect is obviously most relevant when a relatively low risk ratio is obtained in a study and when the effect of smoking is strong, as in lung cancer.

Although the confounding effect from smoking tends to be limited in occupational epidemiology, it often causes great concern. This problem may become more relevant in the future as the current (and highly justified) anti-smoking propaganda may result in people declaring their smoking habits less openly, causing additional problems for the epidemiologist who wishes to control correctly for smoking.

Finally, a somewhat Pythian comment: occupational and especially environmental studies in the future may more often show low risks and so be increasingly vulnerable to the problem of both positive
and negative confounding whether from active or passive smoking, as well as from other factors, and their control will be correspondingly more difficult.

O AXELSON

Department of Occupational Medicine, University Hospital, 581 85 Linköping, Sweden.

References

Confounding from smoking in occupational epidemiology.

O Axelson

doi: 10.1136/oem.46.8.505

Updated information and services can be found at:

[http://oem.bmj.com/content/46/8/505.citation](http://oem.bmj.com/content/46/8/505.citation)

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
[http://group.bmj.com/group/rights-licensing/permissions](http://group.bmj.com/group/rights-licensing/permissions)

To order reprints go to:
[http://journals.bmj.com/cgi/reprintform](http://journals.bmj.com/cgi/reprintform)

To subscribe to BMJ go to:
[http://group.bmj.com/subscribe/](http://group.bmj.com/subscribe/)