THE CENTRAL AFRICAN JOURNAL OF MEDICINE

Vol. 43, No. 1 January, 1997

Confounding and effect modification: their significance in medical research

Introduction

Confounding and effect modification, two very important statistical concepts that may occur both in the statistical modelling of epidemiological data and other medically related data are seldom considered during the analysis of data. This is rather unfortunate as this may introduce some degree of bias in the statistical interpretation following the analysis.

Confounding and effect modification are a major cause for concern in medicine. Epidemiology is an example of that branch of medicine where confounding and effect modification are commonly applied.

Confounding.

Confounding in the estimation of the effect of a given factor (in producing disease) is the distortion in the estimate attributable to an extraneous variate.[1,2] A confounder is defined as a variable which, if not controlled, produces a distortion in the estimated effect of a study exposure in the absence of misclassification.[3] For a variable to be a confounder, it must be a determinant of the disease and not intermediate in the causal pathway from exposure of interest to disease.[4] A covariate can be said to confound if the measure of an association between an exposure and a disease differs according to whether or not the estimate is adjusted for the disease.[5]

Confounding can be divided into two broad categories: the 'comparability based' and the 'collapsibility based'.[6]

The 'comparability based' is defined as the bias in the estimation of the effects of an exposure on a disease risk due to inherent differences between exposed and unexposed individuals. A practical way by which this can occur is in a clinical trial where two treatments (say Treatment A and Treatment B) are to be assigned to a group of patients or vice versa. In order to reduce confounding in this case, patients assigned to both treatments should be comparable. The 'comparability based' confounding occurs as a result of differences between certain stratified (conditional) statistical measures of association and the corresponding crude (unconditional or collapsed) measure. Most definitions of confounding fall into these categories.

From a 'collapsibility based' point of view, confounding is the failure of a crude parameter to equal the value of the parameter that would have been obtained upon control of confounders.[7] A 'collapsibility based' approach entails that the confounder should be specified and control of the confounder should be defined before the problem of confounding can be tackled. Confounding in the study design can be explained by the covariate's relationship to the way the population was sampled and/or from errors of observation associated with the covariate.[8]

Positive and Negative Confounding.

Breslow and Day[1] define the concept of positive and negative confounding. Positive confounding is defined as a situation whereby ignoring the confounder will make the association between the exposure and the disease risk more positive than it should be. Negative confounding on the other hand is defined as that situation where on ignoring the confounder will make the association between an exposure and a disease less positive than it should be.

Control of Confounding.

There are two possible errors that can arise from confounding:[1] these are when no attempt is made to control for the confounder and when a non confounder is controlled for.

Misclassification of a confounder leads to loss of ability to control for confounding, although control may still be useful provided that the misclassification of the confounder is non-differential.[9]

Misclassification of exposure may bring about a greater problem if factors which influence misclassification occur as confounders, and control for these factors may increase net bias.

It is, therefore, apparent that confounding may be controlled for in the design or at the analysis stage or at both stages.[3]

There are three major methods of controlling for confounding at the design stage.[2] The first method is randomization, which is usually practised in potential cohort designs. Secondly is restriction of the studies to narrow ranges of values for the potential confounders.[3] The third method of controlling for confounding is matching on potential confounding factors. Matching is carried out in both case control and cohort studies. Matching as claimed does not reduce confounding, but controls for it in the analysis. Matching also usually increases the precision of effect estimates.

In certain cases it is not possible to control for confounding in the study design. The best estimate of an association is obtained when the true confounding covariates are known and have been adjusted for in the analysis.[5] At the analysis stage, control for confounding involves stratifying the data according to the levels of the confounder(s) and calculating the effect estimate that summarizes the information across the strata of the confounder(s).

The major disadvantage in stratification is that it is seldom possible to control for more than two or three confounders in the analysis.[3] There is no guarantee that there will ever be a point in stratification at which the stratum-specific parameters will change when one stratifies on a different risk factor.[7] Statistical modelling is thus recommended as a means for the simultaneous control of more confounders.[3]

The problems of confounding are due to the fact that most authors have not taken into account the fact that there exist two separate and distinct phenomena that can give rise to confounding.[10] These phenomena, are the 'non-collapsibility' and the 'non-comparability'(inadequacy of the control) of the disease exposure association. Since confounding due to the inadequacy of the control group is primarily due to the observational nature of epidemiologic studies, eliminating confounding due to this source should be the main focus of the design phase of the study. 'Non-collapsibility', being a property of the target population should be identified and controlled for in the analysis. This is because non-collapsibility is a function of the target population and cannot be eliminated.

Effect Modification.

Effect modification is an interaction between a confounder and an exposure of interest.[11] This signifies that effect modification can be considered to be a special case of confounding. Effect modification can also be looked upon as the joint effects of two or more factors that occur when the study factor depends on the level of another factor in the study base.[12] In epidemiology, effect modification is of equally significant importance as confounding. Effect modification and confounding are related, but different concepts. This is perhaps elucidated in the above definition of effect modification.

Discussion

The ideas and principles underlying confounding and effect modification have been elaborately discussed. This paper should enable many inexperienced medical researchers and epidemiologists to understand in more detail the concepts of confounding and effect modification.

Many studies have been carried out where the possibility of confounding (or effect modification) was not taken into account. The outcome of these studies could have been seriously misleading.

Great care should be taken at the design and analysis stages of the study. It is thus recommended that medical researchers should ensure that the effect of confounding and effect modification is minimized. This is accomplished by identifying every potential confounding variable (this is not very plausible), and also methods by which these variables can be controlled for in the analysis, at the design stage or at both stages.

References

1. Breslow NE, Day NE. Statistical methods in cancer research. Vol 1. The analysis of case control studies. Lyon, France.International Agency for Research on Cancer; 1980.
2. Rothman KJ. Modern epidemiology. Boston: Little, Brown and Co, 1986.
3. Checkoway H, Pearce N, Crawford-Brown DH. Research methods in occupational epidemiology. New York: Oxford University Press,1989.
4. Grayson DA. Confounding confounding. *Am J Epidemiol* 1987;126:546-53.
5. Siemiatycki J, Wacholder S, Dewar RM, *et al.* Degree of confounding bias related to smoking, ethnic group and socio-economic status in estimates of the association between occupation and cancer. *J Occ Medicine* 1988;30:617-25.
6. Greenland S, Robins JM. Identifiability, exchangeability and epidemiological confounding. *Int J Epidemiol* 1988;15:413-19.
7 Greenland S, Morgenstern H, Poole C, Robins JM. Letters to the Editor, re: confounding confounding. *Am J Epidemiol* 1987;126:1086-89.
8. Miettinen OS, Cook EF. Confounding: essence and detection. *Am J Epidemiol* 1981;114: 593-603.
9. Greenland S. The effect of misclassification in the presence of covariates. *Am J Epidemiol* 1980;112:564-9.
10. Wickramaratne PJ, Holford TR. Confounding in epidemiologic studies: The adequacy of the control group as a measure of confounding. *Biometrics* 1987;43:751-65.
11. Clayton D, Hills M. Statistical models in epidemiology. Oxford University Press, Oxford, 1993.
12. Miettinen OS. Confounding and effect modification. *Am J Epidemiol* 1974;114:593-603.

Mr BA Clayinka
Department of Obstetrics and Gynaecology
University of Zimbabwe
PO Box A178
Avondale, Harare
Zimbabwe