ABSTRACT

Gulnar Azevedo Silva Mendonça

Selection of subjects in case-controls studies

Instituto Nacional do Câncer e Instituto de Medicina Social — UFRJ

The selection of cases and controls seems to be the most difficult question in planning and conducting a case-control study and remains essential regarding the practical requirements for all the steps of this kind of epidemiological design. In the limit, it will be the indicator that can measure validity. The presence of selection bias may influence the results of a given study, making the observed association between the factor under investigation and disease stronger or weaker than t6he actual association.

Selection bias must be considered for both the case and control series, since its related to their comparability. For Breslow and Day selection bias can arise when cases and controls are included in or excluded from a study because of some characteristics they exhibit are related to the exposure factor being studied. This kind of bias according to the type of case-control study, the disease, the risk factors and other factors. Frequently they occur when the exposition status is influencing the selection of individuals differently for cases and controls. The probability of being selected might be the same for both groups of comparison in the source population in order to prevent selection bias.

Some specific times of selection bias should be outlined here as the Bearkson's bias, which is caused by an increased tendency for persons with multiple diagnosis to seek or enter treatment and then fall into study populations drawn from treatment sources. In the specific situation there different probabilities for those with the index disease under study and for individuals with the exposure. The selective survival bias is another type of bias that can occur diferentially for cases and controls and is associated with the disproportionate loss of eligible subjects before selection.

Definition of cases

In the case-control design, the main aim is to ensure that all true cases of a given population have the same probability of taking part of the study. This pressupose a standard definition of the disease being studied and all logistic proceedings to find the cases. The source of cases, in general, are based on hospital or clinic lists, death certificates and special health records as cancer registries or disese surveillance systems. The criteria of elegibility should be applied equally to potential cases and controls. The definition of a "case" involves the establishment of objective criteria for diagnoses and a statement of elegibility criteria of selection fo individuals. The exclusion criteria depends on the definition of exposure and might be very well stablished at the exclusion criteria depends on the definition of exposure and might be very well stablished at the beginning of the study. There are some circumstances that can lead to a detection bias. It occurs when the disease in study is nott detected (even at death) or, in the other hand, when individuals of high risk have more intensive medical surveillance.

Control Selection

It seems to be a consensus among epidemiologists that the choice of controls is the most difficult judgement in case-control research. Rothman believes that the central aim of control recruiment raises on the assumption that participate as controls in a given study, should be potential cases in the same study. This condition, however, is sometimes hard to meet in a research protocol. The principle that the cotrols should be selected from the same risk population of the selected cases is intuitively acceptable. Wacholder et al. have

proposed three aspects that must be considered in the selection of controls: the study base, the confounding elimination and the comparable accuracy. If the researcher achieves these three aspects then the possibility of selection bias, cofounding and information bias would be reduced.

In order to avoid future problems with controls selection the researches might define which is the base population and target population for the study when the research is being planned.

Type of Controls

Usually, controls are selected among hospital or clinic patients or among the general population. The logistic difficulties to access randomly population controls points to the need of using other types of community controls as friends, neighbours or relatives.

Hospital or clinic controls

During a long time, and even nowadays, this type of control has been largely used. Some of the advantages of the hospital controls lay on the facts that they would come, hypothetically, from the same population base that generate the cases and that the quality of information would be equivalent to that of cases. It is not obviously, however, that the cases and controls from the same hospital or similar would have the same population base. The best way to check if the hospital controls are representative of the catchment population is validating their choice by comparing them with community controls. Unfortunatly, this is not possible in every case-control study.

Population controls

This would be the ideal choice of control if some of the principles in selecting controls can be assured as: the guaran-tee of the same study base in the two groups of comparison; the stablishment of precise criteria of exclusion and the possibility of extrapolation to the base the distribution of exposures in the controls for purposes such as calculations of absolute or attributable risk. There are some disadvantages of using population controls specially, if the case ascertainment is incomplete and when the sample of the study base is impossible to be obtained.

Neighborhoods controls

In that special type, some operational points are easily succeed as a cooperative participation of the controls and the ability to match on variables related to social class. Theyare usually chosen nonrandomly within a geographical set what may lead to an independence of the selection process and the exposure. The use of neighborhood controls may lead to "over-matching" on the study exposure if the exposure is related to residence.

Friend controls

This seems to be an attrractive alternative for case-control study in terms of resources and logistics aspects. They can be selected from a list of friends or associates obtained from the case. As the case and friend control will have similar socioeconomic status, the biases due to social classd are reduced.They can also lead to overmatching regarding factors that depends on life-style and occupational exposures. Friend controls may be useful in some circumstances, such as the study of genetical disorders.

Other types of control are also employed in epidemiological studies based on death certificate or disease registry lists. The dead controls might be used when the cases in the study but die before being interviewed[6]. The information might be obtained from a "proxy" informant.