



Control of Confounding in Study Design

It is possible to minimize confounding by utilizing certain strategies in the design of a study:

- Restriction
- Matching
- Randomization (in intervention studies only)

There are also analytical techniques that provide a way of adjusting for confounding in the analysis, provided one has information on the status of the confounding factors in the study subjects. These techniques are:

- Stratification
- Multiple variable regression analysis

Restriction

One of the conditions necessary for confounding to occur is that the confounding factor must be distributed unequally among the groups being compared. Consequently, one of the strategies employed for avoiding confounding is to restrict admission into the study to a group of subjects who have the same levels of the confounding factors. For example, in the hypothetical study looking at the association between physical activity and heart disease, suppose that age and gender were the only two confounders of concern. If so, confounding by these factors could have been avoided by making sure that all subjects were males between the ages of 40-50. This will ensure that the age distributions are similar in the groups being compared, so that confounding will be minimized.

This approach to controlling confounding is simple and effective, but it has several limitations:

- It reduces the number of subjects who are eligible (may cause sample size problem).
- Residual confounding can occur if you don't restrict narrowly enough. For example, in the study on exercise and heart disease, the investigators might have restricted the study to men aged 40-65. However, the age-related risk of heart disease still varies widely within this range as do levels of physical activity.
- You can't evaluate the effects of factors that have been restricted for. For example, if the study is limited to men aged 45-50, you can't use this study to examine the effects of gender or age (because these factors don't vary within your sample).
- Restriction limits generalizability. For example, if you restrict the study to men, you may not be able to generalize the findings to women.

Matching

Instead of restriction, one could also ensure that the study groups do not differ with respect to possible confounders such as age and gender by matching the two comparison groups. For example, for every active male between the ages of 40-50, we could find and enroll an inactive male between the ages of 40-50. In this way, the groups we are comparing can artificially be made similar with respect to these factors, so they cannot confound the relationship. This method actually requires the investigators to control confounding in both the design and analysis phases of the study, because the analysis of matched study groups differs from that of unmatched studies. Like restriction, this approach is straightforward, and it can be effective. However, it has the following disadvantages:

- It can be time-consuming and expensive.
- It limits sample size.
- You can't evaluate the effect of the factors you that you matched for.

Nevertheless, matching is useful in the following circumstances:

- When one needs to control for complex, multifaceted variables (e.g., heredity, environmental factors)
- When doing a case-control study in which there are many possible controls, but a smaller number of cases (e.g., 4:1 matching in the study examining the association between DES and vaginal cancer)

Randomization in Clinical Trials

You previously studied randomization in the online module on Clinical Trials. Given the more detailed discussion in this current module of the conditions necessary for confounding to occur, it should be obvious why randomization is such a powerful method to control prevent confounding. If a large number of subjects are allocated to treatment groups by a random method that gives an equal chance of being in any treatment group, then it is likely that the groups will have similar distributions of age, gender, behaviors, and virtually all other known and as yet unknown possible confounding factors. Moreover, the investigators can get a sense of whether randomization has successfully created comparability among the groups by comparing their baseline characteristics.