Chapter 9 Case-Control Study

Definition

A **retrospective**, analytical, observational study often based on secondary data in which the proportion of cases with a potential risk factor are compared to the proportion of controls (individuals without the disease). The common association measure for a case-control study is the **odds ratio**.

Examples

Chambers, C. D., Hernandez-Diaz, S., Van Marter, L. J., Werler, M. M., Louik, C., & Jones, K. L. et al. (2006). Selective serotonin-reuptake inhibitors and risk of persistent pulmonary hypertension of the newborn. New England Journal of Medicine, 354(6), 579-587.

This study used a matched design, matching infants who had persistent pulmonary hypertension with infants who did not have it and compared the rates of exposure to SSRIs.

Importance

Case-control studies are important to investigate the association between exposure to a certain risk factor and developing an outcome (or a disease).

Advantages

→ Useful for initial risk evaluation

For initial risk evaluation, investigators can test their hypothesis from hospital records and existing databases to validate whether there is a possible relationship between the exposure to a particular risk factor and the development of an outcome (or disease). Case-control studies are cheaper in time and costs, and therefore, they are a better method for initial evaluation of the hypothesis. If an association relationship was found significant through case-control studies, the next step is to validate this association in a cohort study.

→ inexpensive evaluation of risk factors

Case-control studies do not cost as much as cohort studies or clinical trials. Since case-control studies usually rely on existing patient data.

→ Useful for rare conditions

For risk assessment in rare diseases, it is incorrect to design a cohort study because the incidence of the disease will be too low that few or no participants will develop the disease. A better solution for studying rare conditions it to design a case-control study that starts with identifying a group of individuals who already have developed the disease "cases" and a control group of those who have not developed the disease "controls" thus overcoming the limitation of the rare condition.

→ Useful for risk factors with long induction periods

In order to assess risk factors with long induction periods through a cohort study, it will be expensive in time. Alternatively, a case-control study is a suitable design in this case in order to save time and provide initial risk assessment.

Disadvantages

→ Weak empirical evidence even when properly done

Case-control studies usually provide evidence that is weaker than those provided by the cohort studies and clinical trials. The reasons are the retrospective nature of the case-control study and the lack of investigator control over confounding variables that occur during the period from exposure to risk factor until the development of the outcome.

→ Lack of control over risk assignment

Same as discussed before in the chapter of cohort studies

→ Limited to the available patient data

A main limitation of the case-control studies is that the investigators can only utilize the data items registered in the existing database (i.e., hospital records). Many other variables that might be important for risk evaluation might be not available.

→ Confounding bias

See later (chapter of error and bias)

Cohort study versus Case-control study

The direction from past to present or from present to future is called prospective direction. While the direction from future to present or from now to past is called retrospective direction.

The major difference is the basis of classifying study groups. If patients were classified according to the exposure to the risk factor (exposed vs. non-exposed groups), this study is a cohort study. If the participants are classified according to the outcome of interest, the study is a case-control study.

All case-control studies are retrospective in terms of time, but not all retrospective studies are case-control. Hence, it is important to consider the subject classification rather than the direction of time.

Example of the analysis of the case-control study

A research team conducted a case-control study to assess the association between eating processed meat and the risk of colorectal cancer. Two groups of individuals were enrolled. The first group includes individuals with colorectal cancer, and the second group includes age- and sex-matched controls without a family history of colorectal cancer. Data of the study are summarized in the following table.

	Cases (CRC)	Controls (Non-CRC)
With diet depending on processed meat	110	75
Diet free from processed meat	90	125

Odds of eating processed meat in the CRC group = 110/90=1.2

Odds of eating processed meat in the non-CRC group = 75/125=0.6

Odds ratio= 1.2/0.6=2

Interpretation of the OR: Colorectal cancer patients were twice as likely to eat processed meat as healthy controls, or the odds of eating processed meat was 2 times higher among CRC patients compared to healthy controls.