Chapter 04 / Capítulo 04

Evidence-based biomedicine: methodology for research, standardization, and scientific procedural aspects (English Edition)
ISBN: 978-9915-704-43-2
DOI: 10.62486/978-9915-704-43-2.Ch04
Pages: 20-25

©2025 The authors. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY) 4.0 License.

Observational Studies, Descriptive Studies, Cohorts, and Case-Control Studies / Estudios Observacionales, Estudios Descriptivos, Cohortes y Casos de Control

Once the working hypothesis and research objectives have been established, the next step is to select the most appropriate study design to test the hypothesis, ensuring a valid and useful conclusion.

This section explains Observational Studies, in which the researcher does not intervene in assigning the study factor or intervention to the groups, acting solely as an observer of reality. In contrast, in experimental studies, the researcher assigns the study factor. Observational studies are useful for describing a situation (as in descriptive cross-sectional studies) or for identifying cause-and-effect relationships (in longitudinal or analytical studies). However, they are often less conclusive than experimental studies.

Below is a diagram (figure 4.1) showing the different types of observational studies, followed by a brief description of the most common ones.

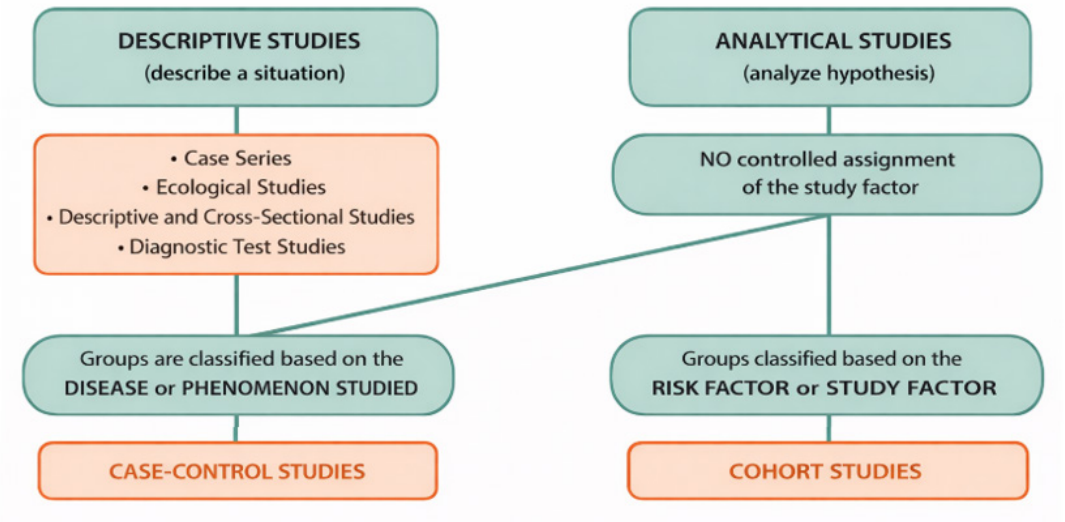

Figure 4.1. Types of observational studies

4.1. Descriptive studies

These designs focus on describing diseases or health phenomena, as well as the factors associated with them. Among the most common are Cross-sectional or Prevalence Descriptive Studies, which are widely used, in addition to other types mentioned in the corresponding figure.

4.1.1. Descriptive cross-sectional/prevalence studies

These studies examine exposure and disease in a specific population at a given time. They are mainly used to determine the prevalence of a disease or risk factor. An example cited in

the literature aims to estimate the prevalence of self-medication among older people and to describe the characteristics of this practice and the condition of home medicine cabinets.

They are useful for calculating the prevalence (at a specific time or period) of the phenomenon under investigation.

They require a clear definition of the reference population for extrapolating results; the population eligible for the study sample; the selection and definition of variables to characterize the process; the measurement scales to be used; and the definition of "case" for calculating prevalence.

Advantages: they are useful for generating etiological hypotheses, are easy to conduct, are relatively inexpensive, allow the study of multiple diseases and/or risk factors simultaneously, characterize the distribution of a disease in relation to various variables, require little time to execute, and are valuable for health planning and management.

Disadvantages: they are not effective for rare or short-term diseases. There are risks of information and selection bias. The main limitation is that they do not allow definitive conclusions about cause-and-effect relationships.

4.1.2. Validity studies of diagnostic tests

This is a type of descriptive study that deserves particular attention. It focuses on evaluating a test's diagnostic capacity, that is, its ability to correctly classify a person based on the presence or absence of a disease. An example article of this type of study will be analyzed later.

The study design involves comparing the test under evaluation with a gold standard.

The results are evaluated in terms of sensitivity, specificity, predictive values, and likelihood ratios, taking into account true positives, true negatives, false positives, and false negatives (figure 4.2).

The study should include patients with as wide a spectrum of disease severity as possible, from the least to the most severe.

The gold standard should be applied to both subjects who test positive and those who test negative to avoid bias.

	DISEASED	HEALTHY	TOTAL
TEST +	a (TP)	b (FP)	a+b
TEST −	c (FN)	d (TN)	c+d
TOTAL	a + c	b + d	a+b+c+d
TOTAL	a + c	b + d	a+b+c+d

Figure 4.2. Distribution of results from a diagnostic test study

4.2. Observational cohort studies

These are non-experimental analytical designs that involve observing and following two groups of subjects over time: one with the risk factor (exposed cohort) and one without (unexposed cohort). The objective is to monitor the occurrence of the phenomenon under study (such as a disease). A classic example of a cohort study is the Framingham Study.

4.2.1. Prospective/ retrospective

Depending on when the study begins and when the disease occurs, there are two approaches: retrospective, in which both exposure and disease have already occurred at the start of the study; and prospective, in which exposure may or may not have occurred but the disease has not yet manifested. It is also possible to have a design that integrates both approaches.

4.2.2. Methodological aspects

Identification of Cohorts: It is necessary to define the subjects who will comprise the cohorts clearly. Ideally, these subjects should come from the same population.

Measurement of Exposure: It is crucial to specify what is considered "EXPOSURE" and how it will be measured.

Outcome Measurement: The outcome variable must be precisely defined. For example, different studies may estimate cardiovascular risk with varying percentages compared to others that use similar tables (such as SCORE, REGICOR, etc.), due to differences in outcome measurement, which may or may not include fatal cardiovascular events.

Follow-up: A time period for follow-up and its specific characteristics must be determined. Measurements should be taken in both groups during the same time period and at the same intervals throughout the study.

4.2.3. Analysis

In cohort studies, the analysis is performed by calculating the incidence in each group, the relative risk (RR), the attributable risk, and the attributable risk fraction. The RR is used to identify the relationship between the risk factor analyzed and the phenomenon or disease (as illustrated in figure 4.3 and explained in a later section). The RR indicates how much more likely a person is to develop the disease. RR values can be >1 (indicating an association between the risk factor and the phenomenon studied), =1 (no differences between groups, implying no association), or <1 (where the risk factor acts more as a protective factor). These studies are also known as prognostic studies.

In this analysis, it is important to consider explanatory variables that may confound the interpretation of the results. An example would be a cohort with a higher mean age than another cohort.

4.2.4. Advantages

They calculate incidence rates based on the principle that exposure precedes disease, following a causal logic. They are particularly effective for studying rare exposures. They also prevent bias that could arise from knowing the effect in advance.

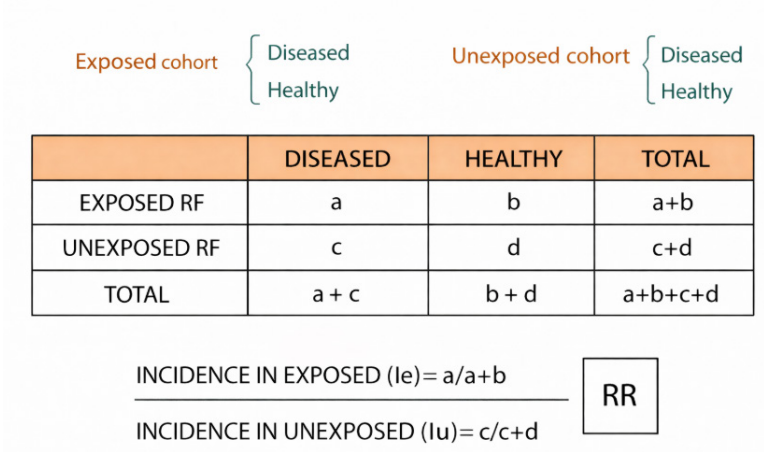

Figure 4.3. Calculation of incidence in each cohort and relative risk

4.2.5. Limitations

They are expensive and can be complicated to perform. They are not suitable for rare diseases or those with a long latency period. They generally require a large sample size. Over time, diagnostic methods and criteria may change, and there is a risk of losing track of participants.

4.3. Observational case-control studies

These are non-experimental analytical designs in which two groups are selected from the population: one composed of subjects with the disease (cases) and another composed of those without it (controls). Then, a retrospective examination is conducted to determine whether the subjects in each group were exposed to the study factor in question.

4.3.1. Methodological aspects

It is essential to clearly and precisely define what constitutes a "CASE," and to establish specific criteria for the "Cases" to be included in the study (e.g., mild disease, severe disease, etc.), depending on the objectives of the study.

Controls should be selected from the same population as the cases. It is important that they have the possibility of having been exposed to the study factor and that both groups, cases and controls, have a similar structure to ensure their comparability.

It is crucial to determine the source and reliability of the information, both for disease and exposure records. CASES and CONTROLS must obtain their information from the same sources, whether through questionnaires or surveys, from family members or friends, or from medical records.

4.3.2. Analysis

In case-control studies, the analysis involves calculating the frequencies in each group, the odds of the cases, and the odds of the controls, which facilitates the calculation of the odds ratio. The latter has a similar interpretation to that of the relative risk (RR), as shown in figure 4.4. It is also important in this analysis to take into account the explanatory variables that could act as confounding factors when interpreting the results.|

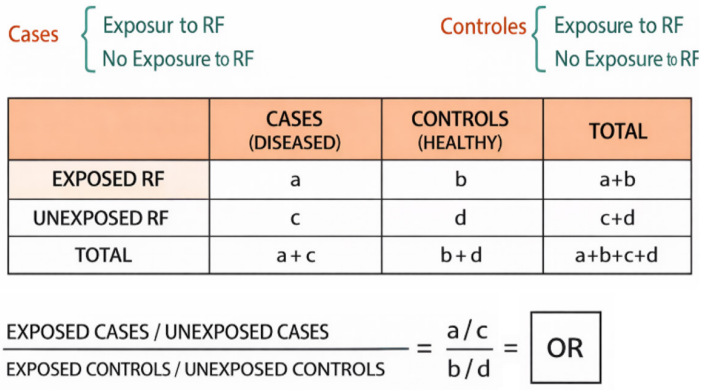

Figure 4.4. Odd calculation and control study outline

4.3.3. Limitations

Since case-control studies start with existing cases, they do not directly estimate incidence. Furthermore, it is difficult to establish the temporal sequence between the risk factor and the disease. They are not effective for analyzing infrequent exposures.