

The Concept of Causation and some common Epidemiological study Designs

Manvi Gupta,

Student of online Advanced Post Graduate Diploma in Clinical Research & Medical Writing (APGDCRP-MR), James Lind Institute, India.

Corresponding Author: Manvi Gupta, Student of online Advanced Post Graduate Diploma in Clinical Research & Medical Writing (APGDCRP-MR), James Lind Institute, India

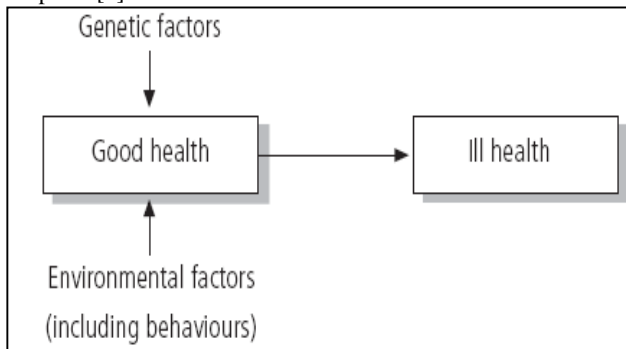
Received date: September 21 2020; **Accepted date:** November 07, 2020; **Published date:** November 15, 2020

Citation: Manvi Gupta (2020) The Concept of Causation and some common Epidemiological study Designs. J. Clinical Research and Clinical Trials. 2(4); DOI: [10.31579/2693-4779/019](https://doi.org/10.31579/2693-4779/019)

Copyright: Copyright © 2020, Manvi Gupta. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Epidemiology in its modern form is a relatively new discipline and uses quantitative methods to study diseases in human populations to inform prevention and control efforts [7]. Causation is an essential concept in epidemiology (Figure 1) [7] yet there is no single, clearly articulated definition for the discipline [1]. Causal inference may be viewed as a special case of the more general process of scientific reasoning, about which there is substantial scholarly debate among scientists and philosophers [2].



Causation (Figure 1)

Concept of Causation:

Definitions of causation from the epidemiological literature:-

1. Production: Causes are conditions that play essential parts in *producing* the occurrence of disease.
2. Necessary causes: A necessary cause is a condition without which the effect cannot occur.
3. Sufficient-component causes: A sufficient cause guarantees that its effect will occur; when the cause is present, the effect *must* occur.
4. Probabilistic cause: A probabilistic cause increases the probability of its effect occurring.
5. Counterfactual causes: A counterfactual cause makes a difference in the outcome (or the probability of the outcome) when it is present, compared with when it is absent, while all else is held constant [1].

Recently however, the term ‘causal inference’ has come to designate a specific set of tools and attitudes [3].

Epidemiological studies

The tools or research study designs used by epidemiologists are varied [6]. (Figure 2) [4]

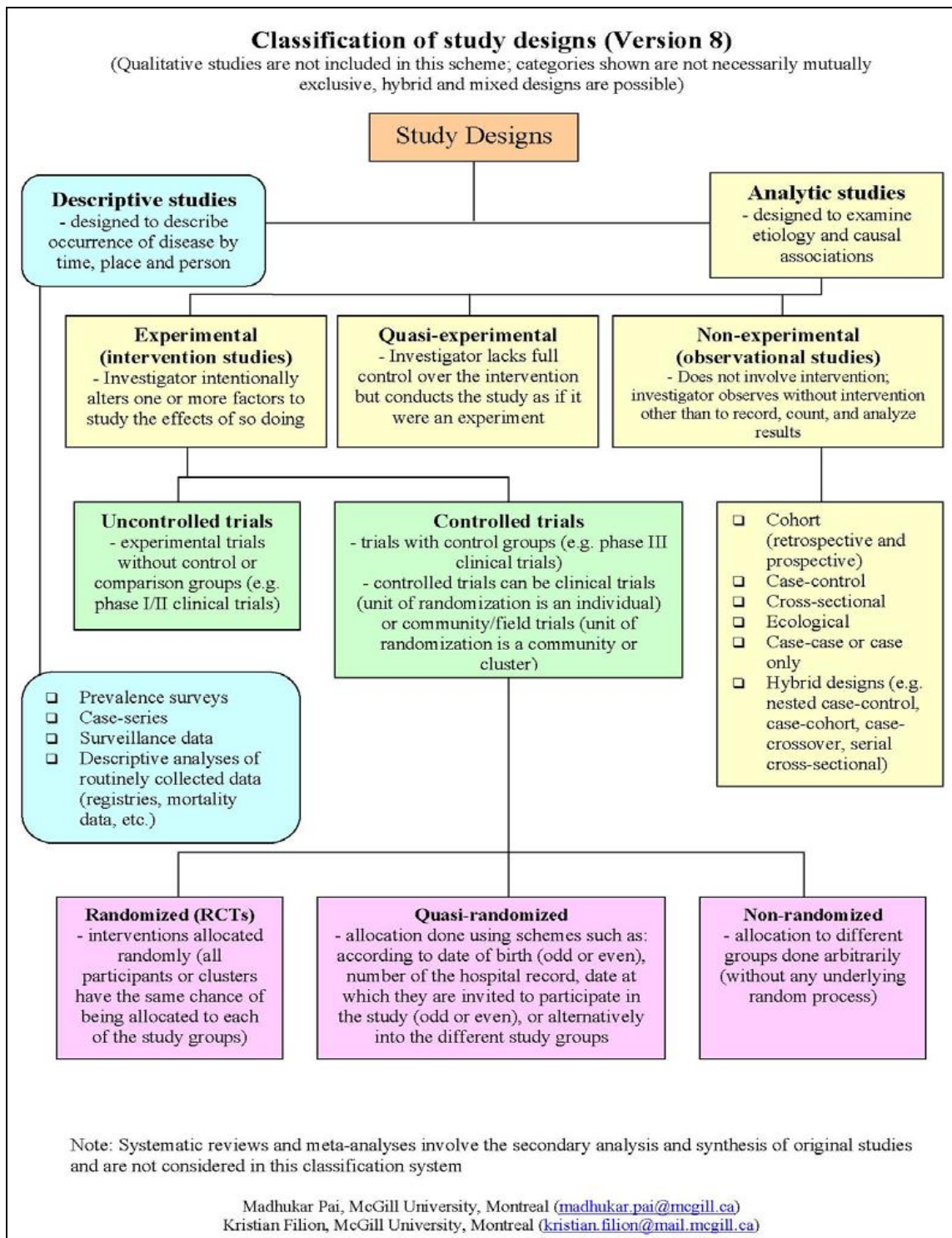


Figure 2

In all epidemiological studies it is essential to have a clear definition of a case of the disease being investigated by delineating the symptoms, signs or other characteristics indicating that a person has the disease. A clear definition of an exposed person is also necessary. In the absence of clear

definitions of disease and exposure, it is very difficult to interpret the data from an epidemiological study [7].

The hallmark of good research is the rigor with which it is conducted. A checklist of the key points in any study irrespective of the basic design is given in box 1 [8].

Box 1**Study purpose**

The aim of the study should be clearly stated.

Sample

The sample should accurately reflect the population from which it is drawn.

The source of the sample should be stated.

The sampling method should be described and the sample size should be justified.

Entry criteria and exclusions should be stated and justified.

The number of patients lost to follow up should be stated and explanations given.

Control group

The control group should be easily identifiable.

The source of the controls should be explained—are they from the same population as the sample?

Are the controls matched or randomised—to minimise bias and confounding.

Quality of measurements and outcomes

Validity—are the measurements used regarded as valid by other investigators?

Reproducibility—can the results be repeated or is there a reason to suspect they may be a “one off”?

Blinded—were the investigators or subjects aware of their subject/control allocation?

Quality control—has the methodology been rigorously adhered to?

Completeness

Compliance—did all patients comply with the study?

Drop outs—how many failed to complete the study?

Deaths

Missing data—how much are unavailable and why?

Distorting influences

Extraneous treatments—other interventions that may have affected some but not all of the subjects.

Confounding factors—Are there other variables that might influence the results?

Appropriate analysis—Have appropriate statistical tests been used?

Validity

All studies should be internally valid. That is, the conclusions can be logically drawn from the results produced by an appropriate methodology. For a study to be regarded as valid it must be shown that it has indeed demonstrated what it says it has. A study that is not internally valid should not be published because the findings cannot be accepted.

The question of external validity relates to the value of the results of the study to other populations—that is, the generalisability of the results. For example, a study showing that 80% of the Swedish population has blond hair, might be used to make a sensible prediction of the incidence of blond hair in other Scandinavian countries, but would be invalid if applied to most other populations.

Cohort, cross sectional, and case-control studies are collectively referred to as observational studies. Often these studies are the only practicable method of studying various problems [8].

Cohort study:

These are the best method for determining the incidence and natural history of a condition [8]. The cohort or follow-up study can be either retrospective or prospective, and sometimes a combination retrospective–prospective approach is used [9].

Prospective cohort studies: A group of people is chosen who do not have the outcome of interest (for example, myocardial infarction). Over a

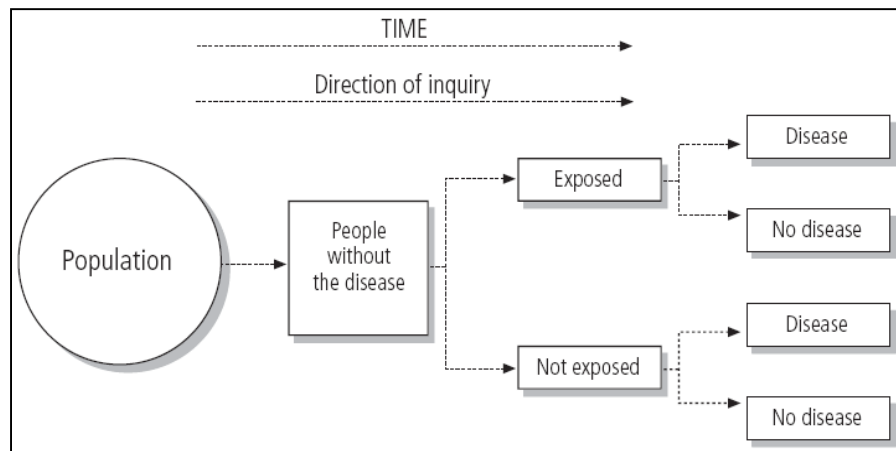
period of time the people in the sample are observed to see whether they develop the outcome of interest [8].

Retrospective cohort studies: The methodology is the same but the study is performed posthoc. The cohort is “followed up” retrospectively [8].

When the objective of a cohort study is to investigate the causal effect of an exposure on a health outcome, that effect is usually quantified in terms of a relative risk (also called a risk ratio) [13].

$$\text{Relative risk} = \frac{\text{risk in exposed people}}{\text{risk in unexposed people}}$$

Design of a Cohort study: [7]



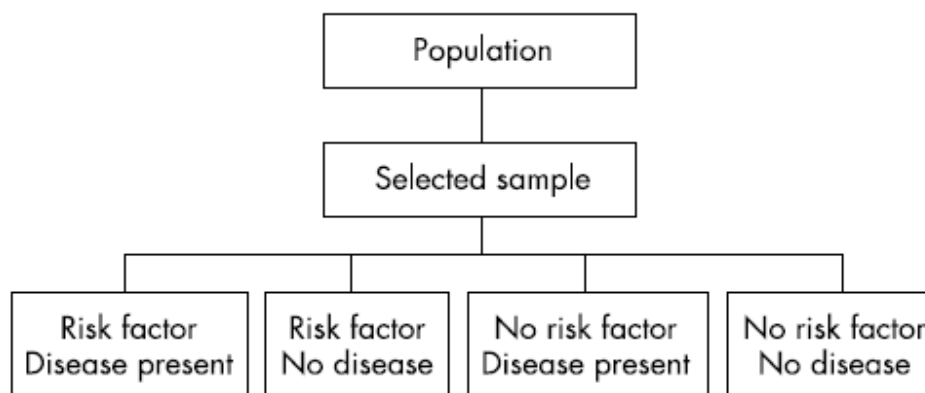
Because the data on exposure and disease refer to different points in time, cohort studies are longitudinal. Although conceptually simple, cohort studies are major undertakings and may require long periods of follow-up since disease may occur a long time after exposure [7].

Cross-sectional studies

Like cohort studies, cross-sectional studies conceptually begin with a

population base. But unlike cohort studies, in cross-sectional studies we do not follow individuals over time. Instead, we only look at the prevalence of disease and/or exposure at one moment in time. Cross-sectional studies begin by selecting a sample population and then obtaining data to classify all individuals in the sample as either having or not having the health outcome [10].

Study design for cross sectional study: [8]



A cross sectional study is particularly suitable for estimating the prevalence of a behavior or disease in a population [11].

$$\text{Prevalence}^{10} = \text{cases} / \text{total population}$$

Cross-sectional studies are used both descriptively and analytically.

Descriptive cross-sectional studies simply characterize the prevalence of a health outcome in a specified population.

Analytical cross-sectional studies, data on the prevalence of both exposure and a health outcome are obtained for the purpose of comparing health outcome differences between exposed and unexposed [10].

Cross sectional studies are generally quick, easy, and cheap to perform. They are often based on a questionnaire survey. There will be no loss to follow-up because participants are interviewed only once. However, a

cross sectional study may be prone to non-response bias if participants who consent to take part in the study differ from those who do not, resulting in a sample that is not representative of the population [11].

Case control studies

Case-control studies provide a relatively simple way to investigate causes of diseases, especially rare diseases. They include people with a disease (or other outcome variable) of interest and a suitable control (comparison or reference) group of people unaffected by the disease or outcome variable. The study compares the occurrence of the possible cause in cases and in controls. The investigators collect data on disease occurrence at one point in time and exposures at a previous point in time. Case-control studies are longitudinal, in contrast to cross-sectional studies (Figure 3) [7].

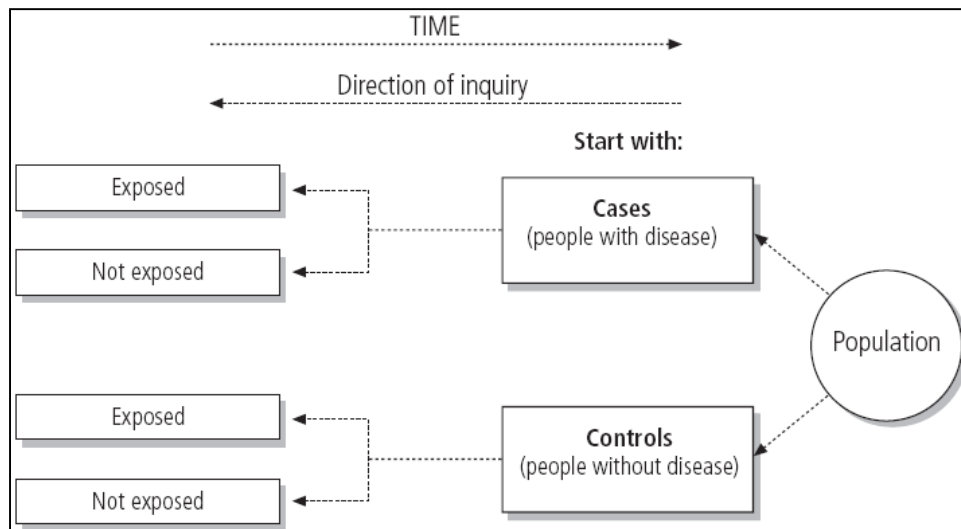


Figure 3

Exposure

An important aspect of case-control studies is the determination of the start and duration of exposure for cases and controls [7].

Odds ratio

The association of an exposure and a disease (relative risk) in a case-control study is measured by calculating the odds ratio (OR) [7].

$$\text{Odds of exposure in the cases}$$

$$\text{Odds (of exposure) ratio}^{12} = \frac{\text{Odds of exposure in the cases}}{\text{Odds of exposure in the controls}}$$

The hard part of designing case-control studies is deciding how to find controls that satisfy the requirement of having been sampled from the whole population without regard to exposure. Most (not all) case-control studies are conducted retrospectively. A particular problem arises when participants are asked to recall past exposures because cases may be more likely than non-cases to have reflected on, and therefore recall, past exposures. This can be a serious source of bias [13].

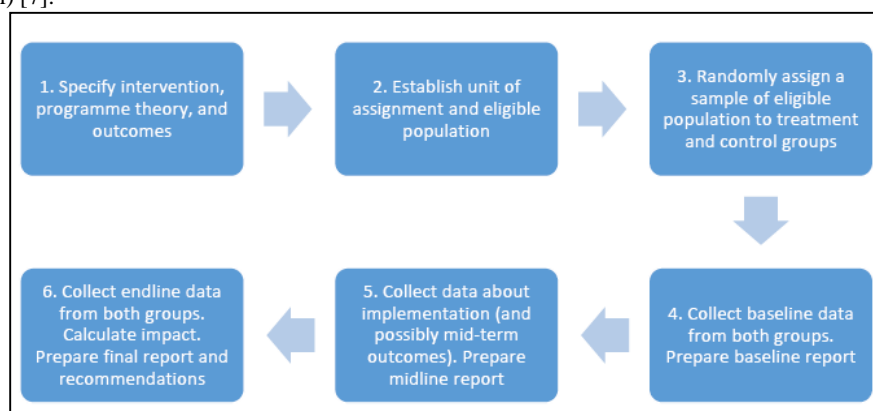
Randomized control trial

A Randomized controlled trial is an epidemiological experiment designed to study the effects of a particular intervention, usually a treatment for a specific disease (clinical trial) [7].

Randomized controlled trials (RCTs) are the gold standard for evaluating interventions to treat or prevent adverse health events, when it is feasible and ethical to use this study design. In a RCT, the investigator randomly allocates study units from a study population to two or more intervention groups. Randomized controlled trials are hypothesis-testing studies, and therefore, it is essential that a concurrent comparison group be used. An outcome may improve, stay the same or decline over time, with or without an intervention. Therefore, without a concurrent comparison group, it is not possible to determine the extent to which any changes in the outcome are the result of the intervention.

An important consideration with randomization is the ethics of the intervention(s) administered to the comparison group(s). Control or comparison groups fall into two broad categories. *Positive controls*: commonly used in RCTs where there is an existing intervention and failure to provide that intervention is not ethically acceptable. Positive controls allow the investigator to comment on the efficacy of the intervention of interest relative to a current intervention in the form of ‘as good as’ or ‘better than’. *Negative controls*: untreated although they should receive a placebo or sham intervention to when it is possible to do so [14].

Overview of conducting RCT (Figure 4) [15]



Potential errors in epidemiological studies

Random error

Random error is when a value of the sample measurement diverges – due to chance alone – from that of the true population value

Sample size

The sample size must be large enough for the study to have sufficient statistical power to detect the differences deemed important.

Systematic error

Systematic error (or bias) occurs in epidemiology when results differ in a systematic manner from the true values. A study with a small systematic error is said to have a high accuracy. Accuracy is not affected by sample size.

Selection bias

Selection bias occurs when there is a systematic difference between the characteristics of the people selected for a study and the characteristics of those who are not.

Measurement bias

Measurement bias occurs when the individual measurements or classifications of disease or exposure are inaccurate – that is, they do not measure correctly what they are supposed to measure.

Ethical issues

Ethical issues are those involving actions and policies that are right or wrong, fair or unfair, just or unjust [7].

Conclusion:

All study designs inherently acknowledge time and represent alternative approaches for sampling populations as exposed and non-exposed persons develop disease over time. Epidemiological associations may be causal; however, before causality can be assessed, each study must be evaluated to determine whether its design is appropriate, the study size is adequate and systematic bias has not influenced the observed association. In addition, the association should be consistent with prior hypotheses and previous study results, and its magnitude should be moderately large [9].

References

1. M Parascandola and D L Weed. (2001). Causation in epidemiology. *J Epidemiol Community Health* 55:905-912.
2. Kenneth J. Rothman, Sander Greenland. (2005). Causation and Causal Inference in Epidemiology. *American Journal of Public Health*, 95(1):144-150.
3. Jan P Vandembroucke, (2016). Alex Broadbent and Neil Pearce. Causality and causal inference in epidemiology: the need for a pluralistic approach. *International Journal of Epidemiology*, 45 (6):1776–1786.
4. Madhukar Pai. Overview of Epidemiological Study Designs. madhukar.pai@mcgill.ca
5. Jonathan M. (1998). Samet and Alvaro Munoz. Evolution of the Cohort Study. *Epidemiol Rev*, 20(1):1-14.
6. Chris Olsen and Diane Marie M. St. George. (2004) Cross-Sectional Study Design and Data Analysis. Copyright © 2004. All rights reserved.
7. R Bonita, R Beaglehole and T Kjellström. (2006). Basic epidemiology, 2nd ed. © World Health Organization.
8. C J Mann. (2003) Observational research methods. Research design II: cohort, cross sectional, and case-control studies. *Emerg Med J*, 20:54–60.
9. Chapter 5. Epidemiological study.
10. Lorraine K. Alexander, Brettania Lopes, Kristen Ricchetti-Masterson, Karin B. Yeatts. Cross-sectional Studies. ERIC Notebook 2nd ed. UNC CH Department of Epidemiology.
11. Philip Sedgwick. (2014) Cross sectional studies: advantages and disadvantages. *BMJ*, MAR, 348:1-2.
12. Chapter 9. Case control study.
13. Rob Herbert. (2017). Case-control studies. *Journal of Physiotherapy*, 63: 264–266.
14. J. M. Sargeant, D. F. Kelton and A. M. O'Connor. (2014). Randomized Controlled Trials and Challenge Trials: Design and Criterion for Validity. *Zoonoses and Public Health*, 61(1):18–27.
15. White, H., Sabarwal S. & T. de Hoop, (2014). Randomized Controlled Trials (RCTs), Methodological Briefs: Impact Evaluation 7, UNICEF Office of Research, Florence.