

Selecting the appropriate study design for your research: Descriptive study designs

Aamir Omair

Department of Medical Education, Research Unit, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia

ABSTRACT

This article describes the importance of selecting the appropriate epidemiological study design for a given study question. It provides an explanation to the different terms used in describing study designs with regards to observational versus interventional and descriptive versus analytical types of study designs. This article focuses on the description of the different types of descriptive study designs, that is, case report, case series, correlational, and cross-sectional study designs. The requirements for selecting these study designs are discussed along with the advantages and disadvantages of each study design. The descriptive studies are similar in the context that they are based on a single sample with no comparative group within the study design. Their basic purpose is to describe the characteristics of the sample with regards to the characteristics that are present and so are useful in generating a hypothesis. The absence of a comparative group is the main limitation of the descriptive studies, and this is the reason they cannot be used to determine an association by testing a hypothesis showing a relationship between a risk factor and disease. The analytical study designs will be discussed in the next article in this series.

Keywords: Case report, case series, correlational study, cross-sectional study

INTRODUCTION

One of the basic issues in deciding how to conduct a study is to first determine the appropriate epidemiological study design for achieving the stated aim and objectives of the proposed research question.^[1] Choosing the right study design is the most important decision to make in determining the methodology of any research study. This is important for the way in which the study will be conducted, especially the sampling and data analysis.^[2,3] Study designs are different for qualitative and quantitative research. The qualitative study designs will be discussed separately in an article on qualitative research. This article will describe the types of descriptive study designs used in quantitative

research and the next article in this series will focus on analytical study designs.

The quantitative research study designs are broadly classified either as descriptive versus analytical study designs or as observational versus interventional. Study designs are arranged in a hierarchy starting from the basic 'Case Report' to the highly valued 'Randomised Clinical Trial'^[4] as shown in Table 1. Before discussing the advantages and disadvantages of each study design, it is important to understand what is meant by the terms 'descriptive', 'analytical', 'observational', and 'interventional', as used to classify the different types of study.

CLASSIFICATION OF STUDY DESIGNS

Descriptive study designs are useful for simply describing the desired characteristics of the sample that is being studied, e.g., an abnormal presentation of a disease in a case report or a case series which includes a collection of cases with the same disease/condition. A descriptive study may also try to generalise the findings from a representative sample to a larger target population as in a cross-sectional survey.^[5] The common aspect between the descriptive study designs is that there is only one single sample without any comparison group.^[6] Analytical study designs, on the other hand, start with the presumption of comparing two or more

Access this article online

Quick Response Code:



Website:

www.thejhs.org

DOI:

10.4103/1658-600X.159892

Address for correspondence:

Dr. Aamir Omair, Department of Medical Education, Research Unit, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia.
E-mail: amir.omair@gmail.com

Table 1: The hierarchy of epidemiological study designs

Observational studies	Strength of evidence	
Descriptive study designs		
Case report		Single case
Case series		Collection of similar cases
Correlational		Population based study - using secondary data
Cross-sectional (descriptive)		Single sample from larger population- no comparison
Analytical study designs		
Cross-sectional (analytical)		Single sample from larger population-compares two or more groups in the sample
Case-control		Compares risk factors between diseased (cases) and non-diseased (controls) groups
Cohort		Compares outcomes between groups exposed and non-exposed to a risk factor for a disease
Interventional study		
Clinical trial	Investigator allots the subjects to different groups - intervention versus non-intervention	

groups and follows them over a period of time. The term retrospective refers to a collection of data pertaining to events which have occurred in the past, irrespective of how/when the subjects are enrolled in the study. Hence, it is more appropriate to reserve the terms prospective and retrospective only for cohort studies.^[9]

The choice of the appropriate study design depends upon the way the research question is stated.^[1] It is important to note here that different study designs may be applicable for the same research problem, but it is the way that the research question is framed that determines which study design is most appropriate. It should also be added that there are many grey areas in which researchers may differ in opinion about the type of study design, but the important factors to consider include the research question as well as the way the subjects were selected for the study.^[3,6] We will consider ‘shisha smoking and heart disease’ as an example for identifying the application of the different types of study design in this article. The research problem is that the practice of shisha (hubble-bubble) smoking is becoming very common and is not considered by the general public as a serious health problem as compared to cigarette smoking.

groups and the samples are selected accordingly from the different groups. These groups may be based on people who are diseased/not diseased for a ‘case-control study’, or they may be based on people who are exposed/not exposed to a particular risk factor in a ‘cohort study’.^[3] A clinical trial is the third type of analytical study design in which one group is considered as the intervention (or experimental) group, which is compared to the non-intervention group (comparison) group.^[7]

A clinical trial is classified as an ‘interventional’ study because the investigator determines who is to be placed in the experimental or the comparison group. All the other types of study designs are classified as ‘observational’ since the investigator only labels a subject as being diseased/non-diseased or exposed/non-exposed based on his/her previous status.^[8] The terms prospective and retrospective, as applicable to epidemiological study designs, refer to whether a subject is being followed up in the future or are being asked/investigated about events or exposure in the past.^[8] These terms are now used only for the cohort studies among the different observational study designs. The collection of subjects or data in the future does not classify a study as ‘prospective’ so it is not appropriate to classify a cross-sectional study as a prospective study even though the subjects may be enrolled over a period of time. An interventional study design such as a clinical trial is generally prospective since the investigator assigns the subjects to different

CASE REPORT/CASE SERIES

A case report in this example may just present a study of a young adult aged less than 30 years who has been smoking shisha since the age of 8 years with his elder brothers. The case report is a presentation of an abnormal finding or outcome which, otherwise, would not be present. It describes the case in detail presenting the abnormal findings but is generally limited in demonstrating an association between the risk factor and the disease.^[10] The argument is that this finding may be due to chance or coincidence, without there being a real association between the risk factor and the outcome. Another example may be of an adverse (or beneficial) side-effect of a new drug, which may not have been documented before. These case reports are useful in identifying a new phenomenon and the reporting of similar cases from other observers, thereby helping generate a hypothesis for the association between exposure and outcome.^[10]

A “Case Series” is a collection of cases with the same outcome/finding. The general number of cases reported in a case series range from 20 to 50, but may vary from as few as 5 to as many as more than 100. In the above example, a study may show a number of young adult ‘shisha smokers’ presenting with heart disease. These kind of studies are more useful than case reports in generating a viable hypothesis for an association between exposure and outcome.^[11] However, there is

still a chance of this being only an incidental finding and the real reason may be due to some 'other' common factor. For example in the above study, there may be other risk factors such as cigarette smoking, lack of exercise, unhealthy diet and family history, etc. The main limitation of a case series study is the absence of a comparative group. So, it cannot be stated whether the outcome is really associated with the exposure or not, unless it can be shown that the group that is not diseased has a different exposure rate from the cases being studied. It is generally easy to select a matching group of subjects who are not diseased and determine the exposure to the risk factor in that group as well. This will convert the case series into an analytical case-control study design (discussed in the next article in the series), which is more useful in showing an association between the exposure and outcome.^[12]

CORRELATIONAL STUDIES

These are also called ecological surveys that are generally based on secondary data. They should not be confused with the term 'correlation' as used in statistical analysis. The term 'correlation' needs to be used with caution as it is misused/misinterpreted very commonly by researchers. The term 'association' is more appropriately used instead of correlation when establishing the relationship between different variables. Furthermore, it is important to remember that the correlation coefficient is used in statistical analysis to determine the association between two numerical (quantitative) variables.^[13] The use of the test for correlation can be applied in any type of analytical epidemiological study design, but this does not classify that study as a correlational study.

Correlational studies as described use secondary data for two or more variables from different sources and try to determine an association between these variables.^[14] The most commonly used sources of secondary data are from governmental databases or reports of international agencies, e.g., cigarette sales per capita (in dollars) versus incidence of lung cancer or number of cars registered in a city versus number of deaths due to road traffic accidents. This can be also done on an international level for different countries, e.g., literacy level or income per capita versus infant mortality rates.^[15] As discussed above, correlational studies are more commonly conducted at a national or international level. However, it is possible to conduct a correlational study on hospital-based data, e.g., number of patients being admitted in a day (from the hospital records) versus number of dosing errors (from the Pharmacy/Quality Assurance Department) or nosocomial infections (from the Infection Control Department).

Correlational studies have the advantage of being quick and inexpensive since they are based on secondary data, which is already available from different sources. However, it should be kept in mind that any probable association shown in the correlational studies might be due to some other underlying factor and not due to the variables being considered.^[14] This is known as ecological fallacy and can be best exemplified by the example that a correlational study could show a positive association between the number of television sets per household (in different countries) and the death rate due to myocardial infarction. While there may be a link between a more sedentary lifestyle leading to increasing heart disease, the association may also be due to the difference in the gross income per capita between these countries leading to difference in food consumption or higher cigarette consumption, etc.

CROSS-SECTIONAL STUDIES

These are also called cross-sectional surveys or prevalence studies. They are easy to conduct and they are the most common study designs being reported in most medical journals. The survey can be completed in a relatively short time depending upon the sample size required and access to the study population. The main aspect of this study design is that it takes a representative sample (cross-section) from the population to generalise the findings for the study population.^[16] It is important to ensure that the sample is randomly selected using the appropriate probability sampling technique as discussed in the previous article on sampling. The unique advantage of cross-sectional surveys is that it is possible to determine the prevalence of an outcome or risk factor from this type of survey. For example, if a sample is selected from the study population of young adults living in Riyadh, it will be possible to identify what proportion of them are shisha smokers. This information can be used to determine the 95% confidence interval, which gives a range for estimating the prevalence of shisha smoking among young adults in Riyadh with a specific degree of confidence.^[9] The presence of the outcome such as heart disease or hypertension as well as other risk factors (cigarette smoking, obesity, sedentary lifestyle, family history, etc.) can also be determined at the same time in cross-sectional surveys.

The cross-sectional studies can also be classified as analytical study designs if the relationship is determined between the exposure and the outcome.^[17] These are considered as point-in-time surveys where both the risk factor and the outcome are determined at the same time, that is, whether the subject is a shisha smoker or not and also whether he has heart disease/hypertension or not. The relationship between

the risk factor(s) and outcome can be determined using the appropriate epidemiological measure with odds ratio (to be described with the case-control study design) or other suitable statistical measures.^[9] It is important to note that this analysis will only show an association and not causation since both the risk factor(s) and the outcome are being measured at the same time point. So, it is not possible to determine in many situations if the risk factor(s) actually preceded the outcome or not.^[16]

CONCLUSION

In summary, descriptive study designs are useful for identifying the risk factors that may be associated with a disease condition. They can generate hypothesis about the probable risk factors for a disease but cannot be used to test the hypothesis that the disease is actually associated with the risk factor. The case report and case series are useful for identifying new diseases or different presentations for a particular disease while the correlational studies are more applicable on general data from secondary sources. The cross-sectional study design is the most commonly used design and generally has an analytical component to test the association between the risk factor and the disease. The analytical study designs of case-control, cohort and clinical trial will be discussed in detail in the next article in this series.

REFERENCES

1. Checkoway H, Pearce N, Kriebel D. Selecting appropriate study designs to address specific research questions in occupational epidemiology. *Occup Environ Med* 2007;64:633-8.
2. Silva IS. Overview of study designs. In: *Cancer Epidemiology: Principles and Methods*. Lyons: International Agency for Research on Cancer; 1999. p. 83. Available from: <http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/>. [Last accessed on 2015 Jun 11].
3. Williams CF, Nelson KE, editors. *Study design*. In: *Infectious Disease epidemiology*. 2nd ed. Sudbury, MA: Jones and Bartlett Publishers; 2007. p. 63. Available from: http://www.jblearning.com/samples/0763728799/28799_CH03_061_116.pdf. [Last accessed on 2015 Jun 11].
4. McKeon PO, Medina JM, Hertel J. Hierarchy of research design in evidence-based sports medicine. *Athl Ther Today* 2006;11:42-5. Available from: <http://www.humankinetics.com/acucustom/sitename/Documents/DocumentItem/6836.pdf>. [Last accessed on 2015 Jun 11].
5. Grimes DA, Schulz KF. Descriptive studies: What they can and cannot do. *Lancet* 2002;359:145-9.
6. Bickman L, Rog DJ. *Applied research design*. In: Bickman L, Rog DJ, editors. *Applied Social Research*. 2nd ed. Thousand Oaks, CA: Sage Publications; 2009. p. 15-6. Available from: http://www.sagepub.com/upm-data/23770_Ch1.pdf. [Last accessed on 2015 Jun 11].
7. Association of Faculties of Medicine of Canada. *Assessing Evidence and information*. Available from: <http://www.phprimer.afmc.ca/Part2-MethodsStudyingHealth/Chapter5AssessingEvidenceAndInformation/ExperimentalorInterventionalStudies>. [Last accessed on 2015 Jun 06].
8. Thiese MS. Observational and interventional study design types; an overview. *Biochem Med (Zagreb)* 2014;24:199-210.
9. Mann CJ. Observational research methods. *Research design II: Cohort, cross sectional, and case-control studies*. *Emerg Med J* 2003;20:54-60.
10. Nissen T, Wynn R. The clinical case report: A review of its merits and limitations. *BMC Res Notes* 2014;7:264.
11. Dalziel K, Round A, Stein K, Garside R, Castelnuovo E, Payne L. Do the findings of case series studies vary significantly according to methodological characteristics? *Health Technol Assess* 2005;9:iii-iv, 1-146.
12. Cole P. Introduction. In: Breslow NE, Day NE, editors. *Statistical Methods in Cancer Research: The Analysis of Case-control Studies*. IARC Publication No. 32. Vol. 1. Lyon: International Agency for Research on Cancer; 1980. Available from: http://www.iarc.fr/en/publications/pdfs-online/stat/sp32/SP32_vol1-1.pdf. [Last accessed on 2015 Jun 11].
13. Omar A. Understanding the process of statistical methods for effective data analysis. *J Health Spec* 2014;2:100-4.
14. *Survey and Correlational Research Designs*. Available from: http://www.sagepub.com/upm-data/57732_Chapter_8.pdf. [Last accessed on 2015 Jun 10].
15. Zakir M, Wunnava PV. Factors affecting infant mortality rates: Evidence from cross-sectional data. *Appl Econ Lett* 1999;6:271-3. Available from: http://www.community.middlebury.edu/~wunnava/Recent_Papers/IMRAEL1999.pdf. [Last accessed on 2015 Jun 10].
16. Levin KA. Study design III: Cross-sectional studies. *Evid Based Dent* 2006;7:24-5.
17. *Health Knowledge. Epidemiology*. Available from: <http://www.healthknowledge.org.uk/e-learning/epidemiology>. [Last accessed on 2015 Jun 10].

How to cite this article: Omar A. Selecting the appropriate study design for your research: Descriptive study designs. *J Health Spec* 2015;3:153-6.
Source of Support: Nil. **Conflict of Interest:** None declared.