

Basic study designs in health research

Fyezah Jehan, Abdul Momin Kazi

Department of Pediatrics and Child Health, Aga Khan University, Karachi, Pakistan.

Most studies related to clinical health research have one or more of the following aims:

- 1) To define and quantify occurrence of disease (incidence or prevalence) OR
- 2) To identify associated factors such as an environmental exposure OR
- 3) To study or compare efficacy of a medical intervention to the disease.

To address these questions various types of study designs may be employed. Each design not only represents a

different way of answering a question but is also different in the types of questions it can answer. The type of study design selected depends on many factors for example the particular research question, validity, efficiency, practicality and ethical considerations. Study designs may broadly be classified into descriptive and analytical. The simplest are the descriptive (non analytical) studies. These studies are primarily "hypothesis generating", have no comparison group and are a precursor to analytical studies where hypotheses are tested. No description of studies to assess causal relationships would be complete without mentioning experimental studies. In the

hierarchy of research study designs, experimental study designs are the most superior in showing disease etiology relationships and for evaluating the effect of a drug or treatment on disease. Figure represents the traditional hierarchy of the basic study designs in use to assess causal relationships. Table provides their basic characteristics, strengths and weaknesses. We aim to give an overview of the basic architecture of various study designs in health research, considerations for their use as well as their strengths and weaknesses. Adherence to a particular design will determine the way results are analyzed and conclusions are presented; eventually contributing to scientific quality and clinical relevance.

Case Reports and Series

Among them case report and case series are done to give detailed description of the occurrence of a disease and are employed mostly in the clinical setting. Case report for single

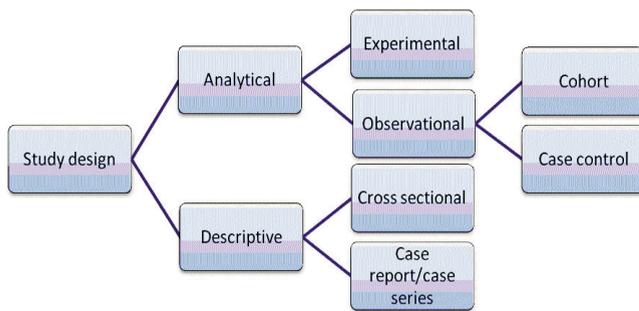


Figure: Traditional hierarchy of basic study designs in health research.

case or case series for multiple cases provide context and detail of a new disease or problem which occurs out of routine. For example to describe an uncommon presentation of a paediatric renal mass or to explain an unconventional or novel method of evaluating laryngeal function in patients after intubation.^{1,2} However case reports suffer from subjectivity and qualitative data, hence can be generalized in a particular context only. Case studies are not designed for doing a detailed statistical analysis. For the purpose of description of a rare disease or to show the unusual manifestation of a known disease it is justified to perform case studies.³

Cross Sectional Studies:

Cross sectional studies are the most popular type of descriptive studies. They are conducted at one point in time over a short period providing the disease and to identify the characteristics associated with it in a snapshot, at a particular time.⁴ This kind of study is commonly employed when the aim is to describe a certain disease with respect to a set of risk factors or to learn the prevalence of a disease or risk factor. For example, to assess the sero positive prevalence of hospitalized, pregnant women infected with Hepatitis B.⁵ These studies have a major benefit that they are inexpensive and can be conducted over a short period of time. Their utility is however limited by the fact that they do not show causality between risk factor and disease. One important problem is that of "survival bias" that is seen in cases of diseases that have long term survivors where a risk factor associated with survivorship will be over represented and appear to be associated with disease.⁶ In spite

Table: Study designs in health research, their basic characteristics and strengths and weaknesses.

Name of design	Type of research question	Characteristics	Strengths	Weaknesses
Case report or case series	Descriptive	Detailed description of the occurrence of one or more cases of a rare disease or rare manifestation of common disease; done by clinicians	Fast, cheap, exploratory	Context is narrow, hence poor generalizability. Not designed for establishing causality
Cross sectional study	Descriptive	Snapshot of a disease and its risk factors measured at a point in time	Fast;cheap; generates hypothesis for future work. Measures prevalence	Limited potential for establishing causality as risk factor and disease are measured at same point in time. Survival bias
Case control study	Analytical observational	Cases(diseased)compared with controls(non diseased) with regards to risk factor information obtained retrospectively	Efficient for rare diseases and more than one exposure. Fast, relatively cheap. Usually smaller numbers required	Can only study one outcome. Some potential to establish causality. Problems with recall. Selection bias
Cohort study	Analytical observational	Cohort of subjects with and without risk factors and free of disease are followed for development of disease	Measures incidence. Can study multiple outcomes. Good for studying rare exposures	Expensive. Long and resource exhaustive. Selection bias
Experimental study	Analytical, experimental	Randomized controlled trials are gold standard. Experimenting a new treatment in controlled environment	Establishes causality and effect of intervention	Expensive and resource exhaustive. Poor generalizability

of these problems, cross sectional studies are highly beneficial in determining prevalence and enumerating risk factors for understanding disease etiology and generating hypotheses.

Cohort Study:

A cohort study is performed by examining one or more risk factors at two or more different levels and observing for factor related disease. In a classical cohort study varying levels of risk factors are assessed in a non-diseased group. These groups are then followed for development of disease and then the two groups are compared with regards to the risk factors and disease. A famous example is the Framingham heart study which looked into the risk factors of different patterns of coronary heart diseases.⁷ Cohort studies provide incidence and natural history of disease. Since the temporal sequence of risk factor to outcome is clearly evident such studies are appealing to clinicians doing health research. These studies are most suitable where the factor is rare. However, these studies may require large numbers of subjects especially for uncommon disease outcomes and may be extremely time and resource intensive. If the period of follow-up is long there can be problems of drop outs (lost to follow up). Given the above, cohort designs are most powerful in showing causal associations between risk factor and disease if appropriate extraneous factors are taken care of.

Case Control Studies:

Case control studies are also done to identify risk factors that may contribute to an outcome. Subjects are selected on the basis of those who have disease (cases) and those who do not have the disease (controls); and risk factors are identified looking back in time. The numbers of subjects that need to be studied are smaller than those in cohort studies. The association of lung cancer with smoking was demonstrated mostly from case control studies.⁸ Thus this design is more efficient especially for studying rare outcomes.⁹ The benefits are somewhat offset by the fact that like cross sectional studies, case control studies also lack temporal flow. Since risk factor information is obtained in the past it may be incomplete and not in the context of the study. A researcher might not get all the information that is needed. They also have the problem of survivor bias that was described with cross sectional studies. Special care needs to be taken to obtain controls from the same population as the cases to avoid "selection bias". Having said this case control studies continue to be a design favored by clinicians who want to show etiological relationships.

Randomized Control Trial (RCT):

The gold standard of experimental studies is the RCT. They are used to assess the efficacy of therapies and interventions. In a RCT every study subject is randomly allocated to either an intervention group or a control. Through

this randomization the assignment to an intervention group is made purely on the basis of chance. Every subject is then followed for effect or disease. This achieves groups that are fairly balanced with regards to known and unknown confounders. If executed well, RCTs can deal with the common problem of extraneous factors (confounders) and selection bias seen in observational studies. However, RCTs need strict monitoring with investigator accountability for adverse events related to the intervention. It results in resource intensive studies. Also, because of ethical issues related to performing experiment on humans RCTs can be difficult to design.

While experimental research is the a superior method for establishing risk factor to disease associations, their design and implementation is hindered by important practical and ethical issues. If implemented properly, observation study designs are a good alternative. Cohort designs are most efficient to establish causal associations of risk factor with disease outcome particularly if the risk factor is rare and several outcomes need to be studied. However careful consideration needs to be made if the period of follow up is long to minimize "loss to follow up". Also added information is required for confounder control. Case control design is preferred to study risk factors if the disease has already occurred. This is in particular applicable to the disease with a long latent time and for studying many risk factors. However as risk is assessed after disease has occurred there may be inherent biases such as recall bias, misclassification of risk factor or incomplete information (information bias). Also if non-diseased (controls) are not selected from a population that is similar to cases invalid association can be produced (selection bias). Cross sectional studies are efficient to quantify disease burden and exploration of new risk factors and generating avenues for further analytical studies. Case reports are reserved for reporting new or new manifestations of old disease or innovative therapies. As medical research advances there is an increasing recognition of the individual strengths of each design and their uniqueness in answering a particular research question.

References

1. Smith EA, Styn N, Wan J, McHugh J, Dillman JR. Xanthogranulomatous pyelonephritis: an uncommon pediatric renal mass. *Pediatr Radiol* Aug; 40: 1421-5.
2. Van der Meer G, Ferreira Y, Look JW. The S/Z ratio: a simple and reliable clinical method of evaluating laryngeal function in patients after intubation. *J Crit Care* Sep; 25: 489-92.
3. Vandenbroucke JP. Case reports in an evidence-based world. *J R Soc Med* 1999; 92: 159-63.
4. Levin KA. Study design III: Cross-sectional studies. *Evid-based Dent* 7: 24-5.
5. Maclean B, Hess RF, Bonvillain E, Kamate J, Dao D, Cosimano A, et al. Seroprevalence of hepatitis B surface antigen among pregnant women attending the Hospital for Women & Children in Koutiala, Mali. *S Afr Med J* Jan; 102: 47.
6. Neyman J. Statistics; servant of all sciences. *Science* 1955; 122: 401-6.
7. Lerner DJ, Kannel WB. Patterns of coronary heart disease morbidity and mortality

in the sexes: a 26-year follow-up of the Framingham population. *Am Heart J* 1986; 111: 383-90.

8. Wynder EL, Graham EA. Landmark article May 27, 1950: Tobacco Smoking as a possible etiologic factor in bronchiogenic carcinoma. A study of six hundred and

eighty-four proved cases. By Ernest L. Wynder and Everts A. Graham. *JAMA* 1985; 253: 2986-94.

9. Rothman KJ, Greenland S, Lash TL. Case-Control Studies. *Encyclopedia of Quantitative Risk Analysis and Assessment*: John Wiley & Sons, Ltd; 2008.
-