Research Designs

Maj Gen (R) Suhaib Ahmed, HI (M)
MBBS; MCPS; FCPS; PhD (London)
Genetics Resource Centre (GRC)

Introduction

Designing scientific research is an art that can be learnt only by practicing. We all are involved in doing research in one form or the other for example writing a dissertation, thesis, or a scientific paper. Probably no aspect of clinical research is as neglected as study design. Eager young investigators attend classes on medical statistics, find dozens of ways to compute "P" values, but rarely learn how to organize a clinical research project. Yet careful study design is the foundation of quality clinical research.

Laying down the objectives

Following the conception of a scientific idea, the starting point should be to phrase the study objective (study question). The objectives should be as clearly and crisply stated as is possible. Usually only one or at the most two objectives should be tackled in one study. If there are more than two objectives then it may be appropriate to address the additional objectives through a separate study. The objectives should be:

- Usually one or two
- Clearly spelled out
- Realistic and measurable
- Achievable in a reasonable frame of time
- Tailor the study design to achieve the objective(s)

Study Designs

The research designs may be roughly categorized as "Observational" or "Experimental". The experimental studies are done on variables whose control lies in the researcher's hand. In an observational study the researcher observes things as they happen.

- Observational Studies
  - For factors that can not be controlled by investigators like age race etc.
  - Essentially descriptive
  - Most epidemiological surveys
  - May be comparative
  - May be prospective, retrospective or cross sectional
- Experimental Studies
  - For factors that may be controlled by investigators
  - Always prospective
  - Usually comparative
  - Most clinical trials
  - Inferences are stronger than observational studies

Observational Studies

In an observational study the researcher only observers and documents the things as they happen. Observational studies are done on variables that can not be controlled by the researcher. Such studies are mostly descriptive. However, comparative observational studies may also be done. Observational studies may be of the following three types (Fig: 1):

- Cohort Study
- Case Control Study
- Cross sectional Study

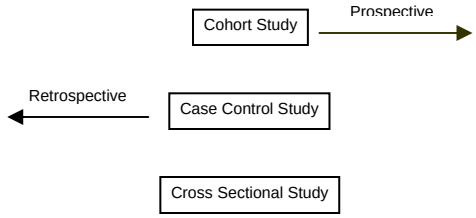

Fig: 1. Different types of observational studies.

Cohort Study

In a cohort study a group of individuals is prospectively followed forward in time. The essential features of a cohort study are:

- A cohort (group) of subjects is identified
- Followed prospectively (longitudinally)
- Subsequent findings are recorded
- More than one cohorts may be studied and compared
- The best choice out of the observational studies
- Problems:
  - May be very long and expensive
  - Unsuitable for rare outcomes
  - Loss to follow-up
  - Careful selection of subjects and controls

Some examples of cohort study may be:

- Patients with Cirrhosis who develop Hepatocellular carcinoma
- Smokers and non smokers who develop Carcinoma lung
- Patients of Myelodysplastic syndrome who develop leukaemia

Case Control Study
In a case control study a group of subjects with a disease (cases) are compared with an unaffected group (controls) for the past history of exposure to a factor of interest. Its essential features are:

- Always retrospective
- Simple and inexpensive
- Suitable for rare outcomes
- Problems:
    - Selection of controls
    - Selection of cases
    - Recall bias

Special care is required in the selection of cases and the controls. The two groups should be as identical as is possible except for the disease under investigation. The recall bias is a phenomenon in which the cases with disease may give an exaggerated (biased) response to the question regarding exposure to a certain factor while the controls who did not develop the disease may not give the same importance to a similar factor as they did not develop the disease.

Case Control studies are always retrospective by design. However, these must be differentiated from retrospective data collection in which the archival hospital/institution records are used to record information of epidemiological nature. On the contrary a case control study usually moves forward in time and only the previous history of exposure to a factor of interest is recorded.

Cross Sectional Study
The subjects or the controls are examined and the findings are recorded only at one point in time. The essential features of a cross sectional study includes:

- Information is collected at only one point in time
- Easy and cheap
- May be used as an alternative to cohort or case-control studies
- Most epidemiological surveys are cross-sectional in design
- Problems:
    - Sample selection

Since most epidemiological studies fall within this category of research it is extremely important that the sample analysed is representative of the population at large.

Experimental Study

In an experimental study the researcher deliberately influences events and investigates the effects of intervention. All experimental studies also include a control group against which the effects of the intervention are noted. The requirements of an experimental study may include:

- Randomization
- Blinding
- Cross over

Randomization

An essential requirement of an experimental study is the order in which the experimental conditions are allocated to the patients in each group. Randomization does not mean haphazard distribution.

- Each patient should have an equal chance of receiving the experimental conditions
- Simple randomization
  - Tossing a coin
  - Table of random numbers
  - Computer
- Block randomization
- Stratified randomization

In a study with small number of patients randomization may result in unequal numbers in each group. This may be avoided by "block randomization" in which blocks of 2-4 patients are made and are then randomized. Stratified randomization may be required when the response to the treatment is conditional for example in patients of different ages or sexes etc. Randomization is then done in each stratum (age groups or sexes etc.).

Blinding

To avoid any bias the patient and/or the investigator may be blinded (no knowledge as to which patient is getting what treatment). In a single blind study the investigator knows but the patients are unaware of the type of treatment that they are getting. In a double blind study the investigator as well as the patient is blinded.

Cross over

In some experimental studies it may be possible to administer each patient both forms of treatment at different occasions. Generally the cross-over design is possible when the effect of treatment is only short lived. A typical example may be the use of two different forms of treatment for migraine.

Experimental Laboratory Studies

The experimental study design is mostly suitable for clinical trials. However, this design may be used in some laboratory studies. An experimental study may be designed to

"study the effect of an anticoagulant-X on the blood cell morphology". The blood samples, divided into two groups, could be randomly allocated to each group. The test group could be collected in the anticoagulant-X while the "controls" could be collected in a standard anticoagulant like EDTA. The changes on blood cell morphology, as seen under the microscope, could be recorded at regular intervals like, 1 hour, 6 hours, 12 hours etc. At the end of the experiment the findings in each group could be compared.

Many lab studies are designed to compare the sensitivities of two diagnostic methods with a predefined sensitivity or specificity. Such comparisons can not form an experiment because by definition the researcher in an experimental study should be able to control the variable of interest. Take the example of comparison of sensitivity of Test-A and Test-B in detection of antibodies in the serum. Here a comparison is being made between two techniques with predefined (well known) sensitivities. The researcher would not have any control over the amount of antibody in the test sample. This study design can be best described as observational. However, if the researcher had a control over changing the sensitivities of the two tests or a control over the amount of antibody in the sample then this could form an experimental study.

Choosing a study design

The first step in any study is to define the objective. Next the study design is tailored to achieve the desired objectives. The choice between an observational and an experimental study is generally straightforward. The experimental studies, when possible, are always a better choice. However, most studies are not experiments. Out of the observational studies cohort studies are the best choice. Regarding observational studies one needs to keep in mind the large number of possible biases that may easily lead to false results.

Sampling

All scientific studies are done on a limited sample size with a view to extrapolate the results to the population at large. Therefore, both the choice as well as the size of the sample is important.

- The sample should be representative of the population
- Care in formation of "Groups" with respect to known sources of variation
- Confounding Variables
- Adequate sample size

The importance of confounding variables can be understood in the study of "increased infant mortality due to genetic causes amongst the offspring of consanguineous couples". The cousin marriage poses a higher genetic risk for the offspring. Such effects may be exaggerated if one fails to take into account the confounding variables like poverty and illiteracy. The latter two factors are an important cause of increased infant mortality and at the same time consanguineous marriage is also more common amongst the illiterate and the poor section of the population. Failure to control for the confounding variables can lead to biased results.

Sample size

Adequate sample size is another important factor in a scientific study. Too small or at times too large a sample size is undesirable. As a general rule the size of the sample is inversely proportional to the expected frequency of the variable under study. Higher the expected frequency smaller would be the size of the sample required and vice versa. Additional information regarding the level of confidence and accuracy required are also used in calculation of the sample size. Good computer packages are available that can be used to calculate the adequate sample size.

Is it possible to rectify a bad study design?

The data from a good study can be analysed in a number of different ways. On the contrary it is not possible to compensate for the incorrect study design by any trick in data analysis.