
Research study design

Sami L. Bahna, MD, DrPH,* and Steven A. Conrad, MD, PhD†

INTRODUCTION

Study design is the most important first step in a research project. Comprehensively, a research study comprises several components, starting with stating clear objectives and ending with conclusions and recommendations (Table 1). In general, studies may be descriptive or analytical. *Descriptive studies* provide description of a condition in one or multiple persons. *Analytical studies* are based on certain hypotheses, follow a specific protocol, and involve measurement, classification, and statistical analysis of data. Research studies may be classified into 2 main types, namely, epidemiologic and interventional (Table 2).

EPIDEMIOLOGIC STUDIES

Descriptive Studies

In a descriptive study, the author describes the characteristics of affected persons in a sample of one, a few, or a series of subjects. Observations may include age, sex, occupation, geographic distribution, physical findings, comorbid conditions, specific associated factors, and others. This type of study is easy and quick, but the provided information is largely left up to the authors, and hence it can be incomplete or biased. Statistical analysis would depend on the data collected but typically includes descriptive and comparative statistics.

Cross-sectional Studies

A cross-sectional study is the most common type of epidemiologic study. It is usually intended to answer a question of interest through collecting data on a population sample (a cohort) and its exposure factors. Criteria are specified for the target and accessible populations, and then appropriate methods are used for drawing the sample. The data to be collected should be tailored toward answering the research questions. The measurements are made once without a follow-up. It is a snapshot of the condition as it exists in a certain sample at a particular point of time or within a short interval. The cohort sample size should depend on the frequency of the disease in that population and the precision of estimate required (ie, the lower the prevalence, the larger the sample needed; see the article entitled "Sampling and Statistical Inferences" in this issue).

Affiliations: * Department of Pediatrics, Allergy and Immunology Section, Louisiana State University Health Sciences Center, Shreveport, Louisiana; †Departments of Medicine, Emergency Medicine, and Pediatrics, Louisiana State University Health Sciences Center, Shreveport, Louisiana.

Disclosures: Authors have nothing to disclose.

Received for publication February 4, 2009; Received in revised form March 27, 2009; Accepted for publication March 30, 2009.

Table 1. Components of Research Studies

State clear objective(s)
Delineate a hypothesis
Define the study population
Plan a sound study design
Choose appropriate methodology
Decide on a minimal valid sample size
Data collection and organization
Data analysis and presentation
Interpretation and discussion of findings
Conclusion(s) and recommendation(s)

Table 2. Types of Research Studies

I. Epidemiologic studies (observational)
A. Descriptive
B. Cross-sectional (survey)
C. Prospective (longitudinal, cohort, or follow-up)
D. Retrospective (case-control)
II. Interventional studies (clinical or experimental)
A. Therapeutic trials
B. Secondary prevention trials
C. Primary prevention trials

Cross-sectional studies reveal prevalence, which is the number of cases at one point in time divided by the number of people at risk at that time. They have the advantages of being relatively fast and inexpensive. In addition to describing the demographic and clinical characteristics of the study group, they may reveal possible associations and generate hypotheses for further studies. However, causal relationships may be difficult to establish by cross-sectional studies. Another limitation is that they do not provide worthwhile information on the prognosis or the natural history of the disease. Serial cross-sectional studies on a particular population would be valuable in providing information about changing patterns of the disease over time.

Prospective Studies

Prospective studies are also called longitudinal, cohort, or follow-up studies. In this type of study, an attempt at answering the questions of interest is performed by following up a group of initially nonaffected subjects longitudinally to document the development of the condition under study and the possible causative factors. The subjects being followed up may be enrolled as a cohort rather than individually in a sequential manner. Follow-up can be periodic during a certain interval or at the end of the study duration. Prospective studies may or may not include intervention. In the latter case, a specific intervention (exposure or treatment) is ap-

plied to one cohort and comparison is made with another similar cohort without intervention. Prospective studies have several advantages and disadvantages (Table 3). The most important advantage is the collection of complete and standardized data. The main disadvantages are the cost and long duration of follow-up. Also, outflux or influx in the population or changes in exposure can impose problems in data analysis or introduce bias.

Retrospective Studies (Case-Control)

In retrospective studies, a group of cases of disease is identified and its exposure history is compared with that of a control group from the same population. Retrospective studies are most desirable for studying diseases that have low prevalence or that require a long induction or latent period to develop, when exposure data are difficult to obtain, or when the study population is dynamic. Case-control studies can reveal factors that lead to the occurrence of disease and estimate the risk and odds ratios. Their greatest advantages are saving time and cost. A great limitation of retrospective studies is recalling of past events or gathering data from different sources that can vary in completeness or accuracy (Table 4).

The definition of *disease* or *case* has a great impact on the study findings. Validity of the results would depend on the criteria being applied (ie, specific vs nonspecific and objective vs subjective). For example, a case of asthma in different studies may be based on any of the following: self-diagnosis, history of recurrent wheezing, diagnosis by any health care professional, diagnosis by a specialist, response to bronchodilators, or specific pulmonary function test results. Inclusion and exclusion criteria should be explicit and clear. Restrictive criteria, though enhancing the purity of data, would reduce the number of subjects. This issue is of great importance in comparing the results of different studies. Once a case definition is settled, subjects can be identified through various database sources known for their high accuracy and efficiency (eg, clinics, hospitals, special registries). In some instances, advertisements in media or in selected common places can identify persons for screening.

The control group should be sampled independently of the case group and must represent the source population. It can be difficult to enroll controls for retrospective studies because they do not have the same level of interest or the recalling of past events as the cases do (recall bias). They should be selected randomly from a well-defined population or sources

Table 3. Advantages and Disadvantages of Prospective Studies

Advantages	Disadvantages
Uniformity and completeness of data	Require large sample size
Calculation of absolute and relative risk	Long follow-up periods
Can have multiple end points	High costs
Can study multiple outcomes	Dropouts
Provide incidence	

Table 4. Advantages and Disadvantages of Retrospective Studies

Advantages	Disadvantages
Short duration	Inaccuracy or incompleteness of data because of the recall factor
Low cost	Heterogeneity in data sources
Calculates odds ratios	May not identify weak exposure factors
	Susceptibility to sampling bias of using 2 populations
	Limited to one outcome measure

that guarantee no or minimal bias. The controls should be matched with the cases regarding confounding factors that might affect the development of the disease. The size of the control group should preferably be at least equal to the case group. Increasing the number of controls has the advantage of slightly increasing the power of the study, but a ratio of up to 4 controls to 1 case would be sufficient to avoid increasing the cost or duration of the study. Sometimes more controls than cases are used to ensure generalizability of results or to reduce the likelihood that unknown hidden variables may be responsible for outcomes.

INTERVENTIONAL STUDIES

Interventional studies are typically either therapeutic or preventive in which a particular procedure or measure is being evaluated regarding its efficacy and safety. Clinical or therapeutic trials are intended to test the effect of a certain treatment on the disease outcome. *Secondary prevention trials* are intended to test the effect of a certain intervention in preventing the recurrence or complication of a disease. *Primary prevention trials* are performed on apparently healthy subjects with the objective of preventing the development of the disease. To increase yield and cost-effectiveness, this type of trial is usually performed on subjects at high risk of developing the disease.

Subjects to be included in interventional trials are enrolled according to clearly stated predetermined inclusion and exclusion criteria. In addition to voluntary withdrawal by the subject at any time during the study, there should be certain criteria for withdrawing the subject before completing the trial. It is important to follow up all patients who entered a trial, even if they withdraw from the study because exclusion of these subjects may introduce bias. Each trial may have one or more well-defined end points and outcome measures, preferably including objective ones. Limiting the number of end points makes follow-up easier and also reduces the chances of bias due to variations in the follow-up period among subjects.

Crossover Designs

In studying chronic diseases, a certain therapy may be compared with another while the subject serves as his/her own control. Patients are randomly assigned to receive 2 types of therapy in different sequences (ie, first treatment A then B or first treatment B then A). The duration between the 2 treat-

ments (washout period) would depend on the duration of effect of the drug therapy after its discontinuation (ie, carry-over effect). The period of treatment with each drug depends on the expected duration until maximal response. The major limitation of crossover studies is the uncertainty of whether there is a carryover effect or not, and if so how long the washout period should be.

Factorial Designs

Interventional studies may need to address the effect of more than one treatment regimen; each may include more than one therapeutic agent or procedure. For example, a trial on patients with uncontrolled asthma may be designed to see whether the addition of an oral antihistamine (anti-H₁) or of an oral leukotriene antagonist (LTA) to an inhaled corticosteroid (IC) would have a beneficial effect. The subjects should be randomly assigned to 1 of 4 groups:

- A. IC only
- B. IC plus anti-H₁
- C. IC plus LTA
- D. IC plus anti-H₁ plus LTA

To avoid bias on the subject's part, every patient in each group should receive 3 "medications." Therefore, group C should receive a placebo in lieu of anti-H₁, group B receives a placebo in lieu of LTA, and group A receives 2 placebos: one for anti-H₁ and another for LTA. In this example, the effect of anti-H₁ is addressed by comparing group B with group A, that of LTA by comparing group C with group A, and that of the combination of anti-H₁ and LTA by comparing group D with group A. These comparisons are not performed individually but as part of an overall factor analysis.

In designing clinical trials, provisions for intervention should be in place for subjects whose disease deteriorates during the trial. Subjects may be withdrawn from the trial or may crossover from a treatment arm to the control arm if the latter comprises standard of care. In an intent-to-treat analysis, subjects removed from their initial assigned group are analyzed as if they remained in the group.

Control and Placebo Groups

Comparative therapeutic trials test the efficacy of one or more treatments against a control group. The control group is the benchmark against which an improvement is sought. The preferable type of control group is the *placebo* or an inactive treatment. A finding of efficacy is then conclusive evidence that the treatment has an effect. Although placebo groups are preferred, there may be ethical reasons to use an *active control*, which is an existing, approved, or widely accepted treatment that, if withheld, would deprive the subject of standard care. Although the trial may be designed to demonstrate efficacy over that of the active control, a more common approach is a *noninferiority design*, with the goal of demonstrating efficacy equal to that of an accepted therapy. The major pitfall of noninferiority designs is that the accepted therapy used as an active control may not have achieved its status as standard of care through rigorous trial design, and

therefore demonstration of noninferiority would not necessarily equate to efficacy.

Intent to Treat

In any clinical trial, some subjects who have been enrolled in the study and have been randomized to receive the treatment may drop out of the study at any time. Or, some patients in the placebo arm of the study may start taking the active drug on their own; these are called *drop-ins*. In either case, it may be preferable to compare the intent to treat of one group with the intent to treat of another group (ie, proceed with the analysis as if no patient dropped in or out). Exclusion of such patients not only reduces the sample size but may introduce bias, particularly if the withdrawal was because of a negative outcome during the treatment. However, if some patients after enrollment were found to be ineligible for the trial, they should be excluded.

QUESTIONNAIRE DEVELOPMENT

Using questionnaires is a common way of gathering uniform data in clinical studies. They may be completed by the responder (self-administered), by an interviewer, or by a combination. Both methods are susceptible to errors imposed by the memory effect and the tendency of responders to provide socially acceptable answers. The advantages and disadvantages of each method are summarized in Table 5. Telephone interviews can reduce the cost but may be difficult to conduct at an optimally convenient time to the subject or to the staff.

General Principles of Questionnaire Design

Designing an attractive, easy to complete questionnaire is of utmost importance in increasing the rate of response and the completeness and accuracy of data.

- At the beginning, include a brief description of the purpose of the study and how the data will be used.
- Provide simple clear instructions on how the questionnaire be filled.
- Begin with identification data: name (unless anonymous), age or birth date, sex, address (or geographic location), method of contact (eg, telephone, e-mail, fax), and date of completing the questionnaire.
- The number and contents of questions depend on the data that need to be collected. Questions relevant to a certain issue should be grouped together under a heading or short statement.

The questions may be constructed in 2 ways. Open-ended questions have the advantage of obtaining answers in the responder's own words but the disadvantage of difficulty to code or group so that statistical analysis is feasible. Closed-ended questions ask the respondent to choose from preselected, mutually exclusive potential answers, which saves time and allows easy grouping and analysis. This method has the disadvantage of limiting the answers to the investigator's choices and may not reveal important unexpected responses. This can be overcome by adding a category of "other (specify)." Providing only 2 potential answers may be suitable for

Table 5. Comparison Between Self-Administered and Interviewer-Administered Questionnaires

Type	Advantages	Disadvantages
Self-administered	Low cost Can be anonymous to protect privacy and encourage honest answers Exclude potential bias by interviewers Mailed questionnaires allow convenience for responders and time to recall or gather information	Potential misunderstanding of questions Incomplete or inaccurate data Poor handwriting Low return rate Influenced by the subjects' degree of literacy
Interviewer-administered	Interviewer can clarify questions Low effect of subjects' literacy level High degree of accuracy and completeness Allows interviewer to collect observational information or administer certain tests	Does not protect privacy Responders may provide socially acceptable answers resulting in inaccurate data Very personal questions may not be answered Require more staff, time, and cost

certain qualitative variables but may not be optimal for quantitative variables. The responder may be asked to choose only 1 answer or all that apply or to give a rank. Certain questions may include the option of “nonapplicable” or “do not know.” Questionnaires should be designed such that all questions are answered or denoted as not applicable or not answered. It is essential to data analysis to know whether a data value is missing or not. In some situations, blanks may be erroneously interpreted by statistical software as having a value of 0.

- Questions should be constructed with utmost neutrality, clarity, and simplicity.
- Borrow from others (ie, take advantage of questionnaires used in previous similar studies with the privilege of modifications).
- Quantitative categories should encompass the whole anticipated range, meaningfully classified, and should be exclusive (ie, the beginning of a category should not be the end of the preceding category).
- The sequence of potential responses should be in a logical order to facilitate answering and later coding.
- The questionnaire should be reviewed by experienced colleagues, including the statistician who will be involved in the data coding and analysis.
- Pretesting should be performed first on a small number of potential responders and after revision on a larger number. This process may result in clarification, estimation, or addition of questions.
- Interviewers should be selected for high skill and should receive training on administering the questionnaire to maximize the accuracy and minimize introducing bias.
- The final findings' reliability largely depends on the accuracy and completeness of data collected. Methods for handling missing values should be explicit in advance of data collection.

CLINICAL TRIAL PHASES

Development of new drugs begins in the laboratory through biochemical procedures and animal experiments. Once those

preclinical tests show promise of efficacy and probable safety, human trials can be performed in phases, usually 3, before their approval for marketing. A postmarketing fourth phase may be performed for additional purposes. Study design parameters, such as the need for control groups, randomization, and sample sizes, may differ considerably for each of the 3 study phases.

Phase 1

This preliminary trial is performed on a small number of patients (usually 20 to 30) primarily to determine the drug's safety. It involves frequent visits for recording symptoms, physical findings, and laboratory tests that may include complete blood cell count, liver function, renal function, electrocardiogram, etc. Different dosage regimens may be tested to determine the dose that causes minimal adverse effects. This phase is most commonly performed on healthy male volunteers and in addition to safety testing includes pharmacokinetic and pharmacodynamic studies in humans. Additional phase 1 trials may be performed on the target population (those with the disease to be treated) because the drug's pharmacokinetics and pharmacodynamics and safety may be altered in this group. In the case of toxic treatments, for example, cancer chemotherapy or radiation, it is unethical to use healthy volunteers, so actual patients with the disease must be the study subjects.

Phase 2

Once phase 1 shows an acceptable safety profile, the drug can be tested in a larger number of patients (usually >100) primarily for efficacy but also for safety. In this phase, the drug may be compared with a placebo or a current standard treatment in a double-blind fashion (ie, neither the patient nor the assessing investigator is aware of the type of treatment). This phase also involves frequent visits and laboratory tests but often less than what was recorded in the phase 1 trial. The sample size is typically small, and phase 2 findings are used to support the decision to move to phase 3 trials for purposes

of approval. There may be more than one treatment arm to further evaluate optimal dosing if phase 1 data are not conclusive.

Phase 3

After documenting a statistically significant efficacy of the new drug (usually above placebo), the drug is tried in a larger number (hundreds or thousands) of patients, primarily focused on efficacy while continuing to monitor safety. A head-to-head comparison is performed against a placebo control or a currently accepted treatment (active control) if a placebo would be unethical, usually in equal proportions. Monitoring this phase may include fewer parameters than in phase 2, unless findings from phase 2 warrant closer monitoring. Patients are randomly assigned to either treatment or control in a double-blind fashion. One, or typically more, of these phase 3 (registration) trials is performed for drug approval. In the United States, the decision for approval is made by the Food and Drug Administration based on a sound study design, significant efficacy, and an acceptable safety profile.

Phase 4

After marketing of the drug for some years, the manufacturer may choose to perform postmarketing clinical trials. The objective of the trial may be to compare the drug with other available treatments, to investigate the efficacy and safety of new dosage or combination regimens, or to explore new potential indications for the drug.

BIAS AND ITS SOURCES

Bias has a systematic influence on the observations in a trial. Bias can occur at any stage of the study and can have various sources. Its direction may be toward or away from the null

hypothesis. Its impact would depend on its degree. Every effort should be made to prevent it through a sound study design and careful performance of the study, beginning with the selection of the subjects, collection of measurements, and data analysis, presentation, and interpretation. Bias is of 2 main types: selection and observational.

Selection bias can arise from differences in selecting or following up the study groups. Therefore, it can occur from inappropriate selection of cases or controls, self-selection of subjects, or differential loss to follow-up.

Observational bias arises from differences in the way of collection of data. It can lead to enrolling subjects whose disease or exposure is incorrect. Its sources can be poor recall of past information, suboptimal interview method, or interviewer's error in obtaining measurements or data collection. Observational bias can be prevented by ensuring impartiality of the interviewers and the study subjects, appropriately designing the questionnaire or data collection form, and ensuring high accuracy of data.

REFERENCES

Included in a comprehensive list at the end of the Supplement.

Requests for reprints should be addressed to:

Sami L. Bahna, MD, DrPH

Department of Pediatrics

Allergy and Immunology Section

Louisiana State University Health Sciences Center

1501 Kings Highway

Shreveport, LA 71103-3932

E-mail: SBAHNA@LSUHSC.EDU