

Research Study Designs: Non-experimental

Editors' note: This article is the fourth in a multipart series designed to improve the knowledge base of readers, particularly novices, in the area of clinical research. A better understanding of these principles should help in reading and understanding the application of published studies. It should also help those involved in beginning their own research projects.

The category of non-experimental designs is the most heterogeneous of the three classification categories (experimental, quasi-experimental, and non-experimental). Examples of the most common non-experimental designs are listed in [Table 1](#). Although, in general, this category has the lowest level of scientific rigor, each design within this category varies as to its own individual level of scientific validity.

Commonly, non-experimental studies are purely observational and the results intended to be purely descriptive. For example, an investigator may be interested in the average age, sex, most common diagnoses, and other characteristics of pediatric patients being transported by air. They may be interested in the prevalence of a clinical presentation pattern or a specific symptom for a given disease. In such studies, the research question would be focused on prevalence rates, or such, rather than causality. They may propose some associations but cannot effectively prove them.

Most non-experimental designs are retrospective in nature and are sometimes called “ex post facto” (after the fact) research. Because a retrospective study is examining activities that have already occurred, manipulation of independent variables and randomization is not possible. In addition, the dependent variable (ie, the outcome) has occurred before study initiation. Therefore, retrospective designs also generally lack the element of “control” over the study setting, making it very difficult to restrict potential extraneous variables. For this reason, non-experimental designs are the most prone to bias (ie, to have invalid results). Steps that can be taken to increase the validity of these studies are discussed in greater detail in the recommended texts at the end of this article. In addition, a future article in this series will focus specifically on performing retrospective (chart review) type research. However, although precautions can be taken to limit potential for bias in these studies, significant potential for bias exists whenever retrospective study designs are used.

Table 1. Non-Experimental Designs

Cross-sectional study

- Records observations in a selected group at a single point in time. Useful for calculating prevalence rates and collecting much preliminary data quickly.

Case-control study

- Generally used to test possible causes of a disorder. Looks backward from effect to cause. Has a control group.

Before-and-after study (retrospective)

- Takes advantage of a change in therapy or a change in the environment to compare outcomes between the two times.

Historical controls (retrospective)

- Similar to the above but does not require such an absolute period of “change.” Used when a concurrent control group is not possible for ethical or other reasons. Uses prior (historical) patients as the comparison control group.

Surveys/questionnaires

- As the name implies, queries the research question(s) directly and collates the answers. Potential for tremendous bias if poorly designed or executed.

Case series

- Description of a series of patients with a defined characteristic. Can be done as a consecutive group or selectively. Does not have a control group.

Case report

- Description of a single case that reports a new finding or is uniquely educational.

Cross-sectional studies are a one-time survey or observation of one or more groups of subjects. If multiple groups are used, they usually are selected to vary on an important characteristic. This is similar to taking a “snap shot” of these subjects at a single point in time.¹ Cross-sectional studies are ideal for calculating simple prevalence rates. For example, to calculate the prevalence of lung injury in all patients with blunt chest trauma. Investigators also can use this design to attempt to examine the “natural history” of an injury, illness, or other phenomenon by performing multiple cross-sectional observations. For example, rather than follow a single group of subjects across time to observe changes, multiple cross-sectional observations could identify these changes by looking at *different* samples at the different points in time of interest. For example, a cross-sectional study examining the recovery of patients from traumatic brain injury may examine subjects at 1, 6, 12, and 18 months after injury to determine the trajectory of symptoms during recovery. A cross-sectional design would allow researchers to identify symptoms at 18 months after injury without waiting 18 months from time of injury to collect data.

Cross-sectional studies do not look backward at antecedent events and therefore do not have a retrospective component. The design also does not look forward at outcome or subsequent events so there is no prospective or longitudinal component. Nonetheless, individual cross-sectional studies could be performed prospectively or retrospectively. Whenever possible, cross-sectional studies should be performed in a prospective fashion. This allows decisions, made in advance, about when and how the study data will be collected. Doing so increases the likelihood of a more complete and more accurate data set. The greatest advantage of cross-sectional studies is the simplicity of performing them, while also providing better quality data than retrospective studies. They are very useful for collecting preliminary data to support subsequent, more extensive studies.

Case-control studies are perhaps the most respected design within this category. The studies are sometimes referred to as “case-referent,” “case-comparison,” or “trohoc” studies. The latter term is simply cohort spelled backwards, which is a good description of the main difference between the two study designs. A case-control design selects two similar populations of patients based on their outcome. One has the dependent variable of interest (ie, the outcome), and the other does not. The investigator then looks backward, retrospectively, for the independent variables of interest, that is, looking for the causes for that outcome. This is the opposite of a cohort design in which you identify patient populations based on the independent variables being present (or not) and then generally look prospectively for the subsequent outcomes.

Because case-control studies select subjects based on the dependent variable (outcome), the design is ideal for situations in which the outcomes are relatively rare. Some medical conditions are simply so uncommon that they can never realistically be studied in a prospective fashion. For such situations, a case-control study is often the best design option available. In a case-control study, the selection process for the control group is of critical importance. If the control group is not an appropriate

match for the study group, there is a potential for important confounding variables that could invalidate the study results or require complex adjusted statistical analyses.

Even though case-control studies are retrospective in nature and are within the non-experimental design category, sometimes their results can be so compelling as to demonstrate causality. As an example, the association between the use of thalidomide and birth defects was established through a relatively small case-control study. Although the scientific validity of that study design was inherently open to some question, the results were so compelling, and the consequences so severe, that the use of thalidomide in the United States was disallowed by the Food and Drug Administration (FDA) on the basis of a single study, using a non-experimental design.

A *before-and-after* design takes advantage of a change being implemented within the environment to look for changes before and after that change. The investigator is rarely responsible for the change but simply takes advantage of that change as the main study intervention. The investigator is an observer to the process, simply recognizing the opportunity and collecting data. These are sometimes called “natural experiments.” This unique design can provide intriguing information and sometimes is the only option available for either logistic or ethical reasons. Common examples of before-and-after studies are those that have looked at changes in vehicular trauma rates before and after implementation of motor cycle-helmet laws or seat-belt laws.

The results of a before-and-after design can be strengthened through the use of multiple pre- and post-observations. If you simply have a single measurement before and a single measure after the intervention, and the difference between those two measurements is significant, the logical conclusion may be that the intervention resulted in the change. However, without multiple data points, telling whether the change is both real and persistent is impossible. The outcome variable may have been changing already, regardless of the presence or absence of the “intervention,” perhaps because of extraneous variables. Initial changes, immediately after the intervention, also may be just temporary (eg, after a 55-mph speed limit, initially drivers slow down but eventually return to faster highway speeds). Multiple pre- and post measurements would help demonstrate these phenomena.

Before-and-after designs also can be strengthened through stratification of the study population or the use of extensive baseline demographic information. This way, the study populations, both before and after the intervention, are examined in detail to demonstrate comparability. *Historical controls* that resemble the subjects in the “after” group can be used as the “before” comparison group if sufficient demographic information is available to demonstrate group equivalence for potential extraneous variables. It is also helpful if the historical controls are as recent as possible, so as to be close in their temporal relationship to the new study group.

Even with these steps, the potential for substantial confounding variables (or bias) always exists in this type of design. Randomization is not an option in these studies. Usually at least the “before” group is studied retrospectively. As a result, accounting for all the potential confounding

variables is very difficult, if not impossible. Therefore, the investigator should consider carefully whether this design approach is the only option and whether randomization of patients to a control group is not possible. Because of the significant limitations of this study design, any resultant conclusions should be conservative and supported by strong and compelling study results.

The use of *surveys or questionnaires* is a commonly employed research design. This design appears to be a simple and easy form of research to perform, but the reality is that it is not simple or easy to do correctly. Although there are highly scientific techniques available for performing surveys, they are inadequately used in medical studies. As a result, surveys are usually classified in the non-experimental category. In addition, most journals and national meetings have an inherent bias against survey-type research. This is often for good reason, because much survey research is very poorly done.

A primary problem with surveys is “response bias.” In many cases the individuals who return a survey vary greatly from those who do not; thus, the researcher cannot state with certainty that the results obtained represent the beliefs of the entire sample. The lower the rate of survey return encountered, the higher the chance that the results are biased.² Because only the most “radicalized” segment of the survey population may have responded to the survey, ask yourself the question, “If everyone who did not respond to the survey had responded with answers that were the opposite of the study findings, would it substantially change the study conclusions?” If the answer to that is yes, then the results of the survey are at least suspect, if not invalid. A target response rate of 85% is a desirable goal because even if the remaining 15% had entirely different responses, they would be unlikely to change the overall conclusions substantially. However, response rates in most studies are generally much lower than 85%. Thus, increasing response rates is perhaps the most important method for increasing the validity of a survey research study.

There are multiple ways to try to maximize response. The first is to keep the survey as simple, short, and focused as possible. Do not try to accumulate additional data and answer multiple corollary questions with a single survey. The longer and more complex the survey becomes, the lower the response rate. Second, do not distribute a survey at a time anticipated to be inconvenient for the selected sample. For example, most surveys should not be distributed in December because of the holidays. Do not distribute a survey to farmers during harvest, to faculty or students during finals week, or to transport team members a week before the national or regional conference. Third, some investigators have found it helpful to include an incentive with the survey. At a minimum include a self-addressed stamped envelope. Inclusion of a small monetary incentive (usual \$1) or a small item such as mailing labels or a bookmark may also encourage subjects to return the survey.^{3,4}

Another important principle is, whenever possible, to use existing validated instruments rather than developing a new measurement scale in your survey. No matter how well you design a survey or questionnaire, potential misunderstandings

Case reports are often intended to be entertaining and educational; the goal is to elicit a reader's response of “That's interesting” instead of “So what?”

always occur. When available, adapt prior surveys that have already dealt with this. Researchers should always pilot test any new questions or surveys in a group of people resembling the desired sample. Revise the survey with feedback from the pilot test. If a large number of changes were made, the survey should be pilot tested a second or even third time to eliminate

as much potential for confusion as possible. A future article in this review will focus on performing surveys research.

Case series are sometimes not even thought of as a study design. They certainly have a very low level of scientific validity. However, they can range from a comprehensive retrospective review to a simple expansion of the individual case report. Usually a case series involves 3 to 30 patients, all with a defined condition. To be worthy of publication, that condition must be either newly recognized, rarely seen, or highly educational. Case series are more compelling than a single case report because they demonstrate that the condition or some new treatment has occurred more than once. Case series can be strengthened by making them a comprehensive and consecutive collection of patients that includes all relevant cases within a set period and a given institution. Such steps limit concerns about selection bias.

Case series do not have a concurrent control group; therefore, making comparisons or measuring associations is not possible. However, for very uncommon disorders, a retrospective “case series” may be the only way to study the subject. In some cases, prospective investigations may be impossible because of the inability to collect a sample of sufficient size in a practical timeframe. They often are useful to bring attention to a new area of concern that can then be investigated in a more scientific and prospective fashion. In fact, Albrecht, Meves, and Bigby⁵ have identified 23 published controlled clinical trials that resulted from case reports or case series.

Case reports continue to hold a relatively strong position within the medical literature. The case report has its origins in the way medicine was originally taught: largely through apprenticeship with an emphasis on powers of careful observation and learning directly from patient experiences. Before clinical studies, knowledge advancement occurred through the experience and discussion of individual cases. We all know from personal experience that the educational process is much more memorable when there is a personal hands-on experience or direct clinical relevance. The case report is meant to replicate this type of experience.

Currently, three basic types of case reports merit publication. The first is the “highly unique case” that may represent a previously undescribed syndrome or disease. The second type is the case that demonstrates an unexpected association between two or more diseases or disease manifestations and may represent “an unsuspected causal relationship.” The third type is the case with an unexpected outcome, suggesting a surprising “new beneficial therapeutic effect or adverse effect.” Case reports are often intended to be entertaining as well as educational. The goal is to elicit a reader's response of “That's interesting” instead of “So what?”

When the findings in a single case report are not fully adequate or appropriately representative of the disease process, including a discussion of previously described cases that have some features in common is typical. This kind of paper that collates and interprets previous reports, as well as the reported case, is often referred to as “a case report with a review of the literature.” Even in such situations, the reported case must be sufficient to stand largely on its own. If not, then it is best to drop the case entirely and consider writing an article that reviews the literature on the subject. In other words, a weak case report cannot be rescued simply by adding in a review of the literature.

Despite their notable limitations, a large number of research publications in clinically oriented journals and presentations at national meetings use non-experimental study designs. There are multiple reasons for this. First, these studies are by far the least time consuming, least expensive, and easiest to perform. In addition, a non-experimental design is a perfectly appropriate stepping stone in the early stages of investigation. Last, some research questions can only be addressed through the use of non-experimental designs. As was true with the quasi-experimental designs, it would be unethical or impossible to answer some types of research questions with anything other than a non-experimental design. These designs are ideal for sorting through large amounts of data in an effort to identify possible factors that then can be studied more formally and prospectively. Retrospective derivation of diagnostic criteria or clinical rules, using nonexperimental designs, followed by prospective validation using true experimental designs, is a common research process.

Miscellaneous Designs

A number of research methods do not necessarily fit the “scientific rigor” classification system. The degree of scientific validity of other designs is variable and open to subjective interpretation.

Correlational studies are similar to cross-sectional studies in that they attempt to describe sample characteristics. Correlational studies can be classified in a variety of ways. Polit and Beck⁶ equate them to ex post facto designs but discuss the fact that the literature is inconsistent on this classification. Burns and Grove¹ categorize correlational studies as a fourth type of design, not experimental, quasi-experimental, or descriptive (non-experimental). Their main purpose is to look for relationships among variables.

However, correlational studies have important limitations. First, co-occurrence of two variables does not prove that they are truly related. The association may only be by chance because of a high rate of occurrence in the population of interest. Second, even strong evidence to establish an association between variables does not prove causality, that one causes the other. Third, temporal relationships are often unclear (ie, which was first?) in correlational studies. Even though a correlational study might show a high rate of coexistence of two variables and intuitive logic may indicate that one variable could be the cause of the other, this design cannot prove causality. This often results in a “chicken versus the egg” dilemma. Nonetheless, because correlational studies are performed easily and can accumulate a large amount of data relatively quickly, they have an important role within the overall spectrum of research design options.

A *meta-analysis* is an approach that does not involve collection of new data, but rather the reanalysis of data previously collected and reported by others. With meta-analysis the statistical analysis of results from two or more independent studies are analyzed for the purpose of integrating their findings and developing overall conclusions. One objective of meta-analysis is to accumulate evidence about a given treatment or other procedure to provide guidance to clinicians in treating future patients. Another less common but important objective is to suggest directions for future research based on questions that remain unanswered by the literature. The term meta-analysis was first used by Glass in 1976.^{7,8} The first research study classified as a meta-analysis within the health care literature was not until 1982⁹ (www.pubmed.com). However, with the growing interest in both evidence-based medicine and cost-effectiveness, the number of publications using these methodologies has increased, as has the respect for the techniques.

The statistical methods used in performing a meta-analysis vary among studies, and their appropriateness is often debated. One of the most important features of such studies is the criteria used to include a study in the final meta-analysis. If not selected carefully, the inclusion criteria can be an important source of bias in these studies. For example, if performing a meta-analysis on benefits of neuromuscular blockade for emergency intubation, do you include all studies on this subject, or only those that used randomization, blinding, or at least a control group? The overall quality of the meta-analysis can be no better than the general quality of the studies that are included in that analysis. In other words, this process is no stronger than its weakest link; when the articles analyzed represent “garbage in,” the final results often are no more than “garbage out.”

Nonetheless, rigorously performed meta-analyses have been an important addition to medical literature, as well as an important new option among research study designs. They are particularly helpful when a number of small but well-done studies exist in a specific field. Each study may individually be too small to find statistical significance for observed clinical differences. A meta-analysis can pool the results of the smaller studies and provide a better overall conclusion.

A related activity is a *retrospective systematic review*, which involves a review of a body of literature on a given subject, usually with resultant conclusions and recommendations. A retrospective review differs from a meta-analysis in that meta-analytical statistical techniques are not used to summarize the findings. The work can be from a single investigator or a group effort. The degree of scientific validity can vary tremendously depending on the process used. Sometimes these works involve a very comprehensive review of all literature on the subject in a highly objective fashion, generally called a “systematic review.” Other times, such articles are little more than the biased opinion of the author(s), and the resultant scientific validity is questionable.

Now that sophisticated computer literature search capabilities are readily available, the ability to comprehensively identify the body of literature on a given subject is much easier. If problems are encountered, medical librarians can be helpful to improve the quality of such searches.

Table 2. Other Design Types

- Correlational studies
- Meta-analyses
- Systematic reviews
- Cost-benefit studies
- Cost-effectiveness studies

Retrospective reviews are strengthened by a discussion of the methodology used both in identifying the relevant body of literature, as well as analyzing and weighting individual articles. Such reviews can make significant contributions to the scientific literature if performed properly and are being used increasingly to support evidence-based medicine recommendations.

A *cost-benefit* analysis is an additional type of research study that does not fit the usual classification systems. These studies are a form of economic assessment in which the costs of medical care are compared with the economic benefits of that care. The benefits generally include a calculation for increased earnings due to improved health, as well as potential reductions in future health care costs. Generally, these calculations are done from a societal perspective. A number of assumptions are always built into these analysis models. For example, what is the value of an additional year of life? The authors should clearly identify the nature, range, and reasons for any assumptions.

A *cost-effectiveness analysis* is similar but compares alternative programs, therapies, or other interventions in terms of their overall costs per degree of clinical effect. For example, it could be cost-per-life saved, per additional year of life gained or per increase of 1% in the hematocrit.

Although the scientific approaches used in economic evaluation analyses are becoming more sophisticated, this methodology is not yet standardized and has tremendous variability in the methods used within different studies. This type of research generally requires a team effort, including panels of clinicians, biostatisticians, and, sometimes, economists. The work itself is often quite tedious and is not a form of research for the novice.

Summary

A core set of clinical research designs have stood the test of time and are repeatedly used in most studies. No single research design is best to answer all research questions, and every research design has appropriate applications. This and the previous articles have described the most common designs. However, in the process of performing research, there are always other options (Table 2), and “hybrid” studies combining elements from different designs are not uncommon. This article does not have sufficient space to discuss subtypes and hybrid designs.

Although it is a good practice to think of research designs in terms of degree of scientific integrity or rigor, it must be recognized that every design type has both advantages and disadvantages. In addition, usually many different ways are available to answer the same research question. Which design is most appropriate depends largely on a stage of evolution of

the investigative process and the resources available. An understanding of the full breadth and spectrum of research study designs is necessary to select the model that is most appropriate for a given investigation.

In general, the best approach is to use the most scientifically valid design that the circumstances will allow. However, the actual decision regarding the design usually represents a compromise between lofty scientific goals and the clinical or resource limitations of the research setting. Therefore, be realistic about the resources available, including the time frame and finances. Realize that research is done in incremental steps, and it is unusual to be able to answer an entire important research question in a single study. The process of planning and revising the protocol, before starting the actual data collection, is critically important. The extra time spent planning will pay off in time savings during the actual study itself. Involvement of a statistician during the planning process, before collecting any data, also can be quite helpful.

Once you have a sense of which research study design is most appropriate to answer your research question, the next step is to flesh out the actual research protocol itself. That process will be addressed in the next part of this series.

References

1. Burns N, Grove SK. The practice of nursing research: conduct, critique, & utilization. 5th ed. Philadelphia: W.B. Saunders Company; 2005.
2. Johnson LC, Beaton R, Murphy S. First record: a methodological approach to counter sampling bias. *Psychol Rep* 2004;95(2):391-392.
3. King KA, Vaughan JL. Influence of paper color and a monetary incentive on response rate. *Psychol Rep* 2004;95:432-434.
4. Everett SA, Price JH, Bedell AW, Telljohann SK. The effect of a monetary incentive in increasing the return rate of a survey to family physicians. *Eval Health Prof* 1997;20:207-214.
5. Albrecht J, Meves A, Bigby M. Case reports and case series from *Lancet* had significant impact on medical literature. *J Clin Epidemiol* 2005;58:1227-32.
6. Polit DF, Beck CT. Nursing research: principles and methods. 7th ed. Philadelphia: Lippincott Williams & Wilkins; 2004.
7. Mullen PD, Ramirez G. The promise and pitfalls of systematic reviews. *Ann Rev Pub Health* 2006;27:81-102.
8. Glass V. Primary, secondary, and meta-analysis of research. *Educ Res* 1976;5:3-8.
9. Gardner MJ, Jusko WJ. Effect of age and sex on theophylline clearance in young subjects. *Pediatr Pharmacol (New York)* 1982;2:157-169.

Other Suggested Readings

- Cook TD, Campbell DT, Shadish WR, Shadish W. Experimental and quasi experimental designs for generalized causal inference. Boston: Houghton Mifflin; 2001.
- Friedman LM, Furberg CD, DeMets DL. Fundamentals of clinical trials. St. Louis, MO: Mosby-Year Book Inc; 1996.
- Hulley SB, Cummins SR, Browner WS, et al. Designing clinical research. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 2001.
- Stiell IG, McKnight RD, Greenlier GH, McDowell I, Nair RC, Wells GA, et al. Implementation of the Ottawa ankle rules. *JAMA* 1994;271:827-831.

Cheryl Bagley Thompson, RN, PhD, is an associate professor, Assistant Dean Informatics and Learning Technologies, and Director of the UNMC Health Informatics Program at the University of Nebraska Medical Center, College of Nursing, in Omaha, Nebraska. She can be reached at cbthompson@unmc.edu. Edward A. Panacek, MD, MPH, is a professor of emergency medicine and director of clinical trials at the UC Davis Medical Center in Sacramento, California.

1067-991X/\$32.00
Copyright 2007 Air Medical Journal Associates
doi:10.1016/j.amj.2006.10.003