University of Nebraska Medical Center
BREAKTHROUGHS FOR LIFE®

University of Nebraska Medical Center
DigitalCommons@UNMC

Journal Articles: College of Nursing

College of Nursing

7-1995

Basics of research (part 3): Research study design

Edward A. Panacek
University of California, Davis

Cheryl Thompson
University of Nebraska Medical Center, cbthompson@unmc.edu

Follow this and additional works at: https://digitalcommons.unmc.edu/con_articles

Part of the Nursing Commons

Recommended Citation

Basics of Research (Part 3): Research Study Design

Edward A. Panacek, MD,[1] Cheryl B. Thompson, PhD, RN, CS[2]

1. Division of Emergency Medicine and
   Clinical Toxicology University of
   California, Davis
2. University of Utah, College of Nursing

Key Words: clinical research, research,
research design, research methodology

Address for correspondence: Edward A.
Panacek, MD, University of California, Davis,
Medical Center, Department of Emergency
Medicine, 2315 Stockton Blvd., Sacramento,
CA 95817

The first article in this series discussed developing an area of general interest, and generating a proposed research question or hypothesis. The second article discussed reviewing the relevant body of literature on the subject and confirming that the research question is an appropriate one. The next step is planning the research project. Translating a research idea into an actual project requires an understanding of research study designs. This article (Part 3) discusses the basic research design types that are most appropriate for clinical investigations. Without having an understanding of the full spectrum of research designs, it is difficult to select the design that would be most appropriate to answer an individual research question.

The study design is the "general plan" for setting up and testing a specific hypothesis. In a sense, the design directs the who, what, how and when of the research project. Think of the design as the basic foundation or infrastructure for the project. Layered on top of the design are the specific elements of the study protocol itself, which will be discussed in detail in a future part of this series.

A common misconception is that for every research project there is one single "best design" to answer that research question. In reality, there are usually many different research designs that can be used to approach a given research question, and each one of the designs has advantages and disadvantages. Decisions regarding which research design to use generally represent a compromise between the goal of rigorous scientific integrity versus limited resources and clinical reality. It is easy to assume that we always should strive to achieve the "gold standard" of a "prospective, randomized, blinded, controlled clinical trial." However, such trials are the most expensive to perform and often are not the most appropriate to answer certain types of research questions. Another common misunderstanding is the belief that retrospective studies are worthless from a scientific standpoint. However, in reality many research questions only can be answered through the use of retrospective study designs. Every research design has a potential application in a given setting, and each design has its limitations. The large prospective, rigorous, clinical trials, which we now take for granted, are a relatively recent development. It was not until 1970 when the first large, exemplary, multicenter clinical trial was published. During the past 20–30 years, the important elements of research study design and clinical-trial development have been refined further and are now well-established.

There are many different research study designs in existence. To best understand how the various designs interrelate, a classification system is needed. There are a number of different classification systems in existence, and individual textbooks use different systems, sometimes resulting in confusion. Examples of such systems include classification by "time frame of data collection," by "assignment of study groups" (i.e., randomized vs nonrandomized), by "degree of masking" (i.e., blinding), and by "degree of overall scientific rigor" or "scientific validity." One of the oldest systems simply divides the research designs into retrospective or prospective categories. Retrospective studies are those in which

Table 1

Classification of Research Designs by Degree of Scientific Rigor

True experimental

- Have all three design elements
- Are always prospective
- Have high scientific validity

Quasi-experimental

- Have manipulation
- Lack control or randomization (or both)
- Are prospective in nature
- Are moderate in scientific validity

Nonexperimental

- Lack manipulation
- May lack other design elements
- Are generally retrospective
- Have low scientific validity

the events of interest occurred before the onset of the study. Even when the research question and design are generated prospectively, if the events that the investigation will be studying have already occurred, then it is a retrospective study. In prospective studies, the events of interest have not yet occurred when the study begins, providing the ideal opportunity to maximize the accurate collection of relevant data. This will be discussed in greater detail as we go through individual study designs.

The research design classification system that is in most common use is "classification by scientific rigor"(Table 1). This system organizes the research designs based on lines of overall scientific integrity. In other words, it asks the important question, "If this research design is used, how scientifically valid are the study results?" *True experimental* designs are those that have a structure that generally result in highly valid results. *Quasi-experimental* are one step down and have a moderate level of scientific validity. *Nonexperimental* are those research designs that, by virtue of their overall structure, give results that do not have strong scientific validity. The ability to draw firm conclusions from the study results is directly proportional to the level of scientific validity of the design. This is the classification system that will be used throughout the remainder of this article.

First, some basic definitions must be

established (Table 2). There are three important variables that apply to all designs. The *independent variable* is the specific study "intervention." For example, treatment A versus treatment B or diagnostic technique A versus diagnostic technique B. The *dependent variable* is the "outcome" that is being measured, which presumably is being influenced by the intervention (e.g., mortality rates or complication rates). Either the independent variable or the dependent variable could be the parameter of greatest interest in a given study, depending on the research question being asked. The most common study model is to manipulate the independent variable (e.g., to give drug A or drug B) and then measure resultant outcomes (the dependent variable). The *extraneous variables* are all the outside influences that are not being introduced or controlled directly as part of the research design, but might have an important impact on the results. These are factors that may distort the relationship between the independent and dependent variables, and may be unequally distributed among the study groups. For example, if the study involves treatment of asthma, then smoking history might be an important potential extraneous variable. Extraneous variables must be considered carefully because they directly can alter the study results, and sometimes even invalidate the entire study. Unfortunately, it is difficult to predict all of the important extraneous variables in advance. Nonetheless, some research designs are inherently better at controlling for extraneous variables.

Research designs also have three primary elements: manipulation, control and randomization. Manipulation is the ability of the design to interact with the study subjects and direct the "independent variable." For example, prospective interventional studies clearly have this element. However, observational studies do not interact with the study subjects and simply record data as they are observed, without any "manipulation." "Control" refers to whether the design has influence over the study environment itself, and, thus, the ability to limit any confounding variables. For example, does the design include formal inclusion and exclusion criteria to

define a specific target population? Prospective studies almost always have some degree of control, whereas retrospective studies rarely do. The third element in research design is randomization. "Randomization" refers to how subjects are assigned to study groups. Designs where randomization is used provide each subject with a known probability of being assigned to each of the study groups (e.g., experimental or control). In most studies the probability of assignment to the two groups is equal. However, if a larger experimental group is needed for a specific reason, the randomization process might assign two subjects to the experimental group for every one subject assigned to the control group. Several methods can be used for ensuring random assignment to groups. Drawing numbers out of a hat (1 = control, 2 = experimental), use of a random number table, or flipping a coin are common methods for assuring random assignment.

Note should be taken that random assignment is not the same as random selection. Random selection is the process of choosing subjects to enroll in the study and is not a component of study design, per se. Subject selection is important to the validity of the study, but will be discussed in a subsequent article on sampling techniques. Only random assignment impacts on the classification of the study as experimental, quasi-experimental or nonexperimental.

Using the classification system by "degree of scientific rigor," true experimental designs have all three of the primary elements (i.e., manipulation, randomization and control). As such, these designs have the most safeguards against sources of bias and, therefore, the greatest degree of overall scientific validity. Quasi-experimental designs have the element of manipulation, but do not have all three elements. Usually the element missing is randomization. As such, these designs have the next highest level of scientific validity, but have scientific limitations. Nonexperimental designs lack manipulation, and usually one or more of the other primary elements also. These are generally "ex post facto," i.e., retrospective-type designs. As a result, they have the lowest level of scientific validity, and, to variable degrees, the findings always are open to

Table 2

Research Design Structure

Three research variables:

- Independent variable (the intervention)
- Dependent variable (the outcome)
- Extraneous variables (confounding factors)

Three design elements:

- Manipulation (the ability to influence the independent variable)
- Control (the ability to minimize potential extraneous variables)
- Randomization (unbiased [random] study subject assignments)

Figure 1

True Experimental Designs*

Two Arm

R O X O
R O   O

Three Arm

R O X_1 O
R O X_2 O
R O   O

Follow-up

R O X O_1 O_2
R O O_1 O_2

Factorial

R O X_1 O
R O X_2 O
R O X_1 X_2 O
R O X_2 X_1 O
R O   O

Crossover

R O X_1 O X_2 O
R O X_2 O X_1 O

* R = Randomization, O = Observation, X = Intervention

question. It is important not to overstate the conclusions from studies using a research design from this category.

True Experimental Research Designs

True experimental research designs are always prospective in nature. A true experiment can support effectively a "cause-and-effect" relationship. They are the most effective at demonstrating efficacy of a new intervention or treatment. To bring a new pharmaceutical product to market, the Food and Drug Administration (FDA) will require the compelling evidence of efficacy shown in a true experimental research design study (i.e., a prospective, randomized, controlled, blinded clinical trial). On the downside, these types of studies are the most demanding in terms of time, cost

and other resources. Also, these studies are focused and only can look at a narrow research question. Therefore, if that research question is not relevant, a great deal of resources can be spent "barking up the wrong tree." Before using a true experimental design, there should be preliminary work already performed that supports the focused research question.

There are many "true experimental" research designs in existence. The most common ones are outlined in Figure 1. In those descriptions, observation (O) can be any measurement or other data collection. Intervention (X) is the manipulation of study drugs, new diagnostic studies, etc. The *classic two-arm* design is that which would be used to study an intervention, such as a new drug. The design uses a placebo control and examines the outcome in two different groups of patients, one that receives the new intervention and one that receives the placebo. A *three-arm* study is very similar but compares two different drugs to placebo. The *extended follow-up* design takes the principles of the two-arm design and makes multiple subsequent observations over a longer period of time (e.g., hospital admission rates, length of stay, relapse rates). This design takes longer to perform but can provide important outcome data. The *factorial* design looks at the effect of multiple interventions, both individually and in various combinations. For example, two different drugs can be evaluated for individual effects, as well as their cumulative effects, depending on the order in which they are given. This design is attractive to examine each intervention, as well as sequences, but can be very resource expensive because the additional study arms require more total study patients. The *crossover* design has the advantage that each patient becomes its own control, thereby directly controlling for most extraneous variables. The disadvantage is that it requires a lengthy period of study and, therefore, is generally not appropriate for the acute-care setting. Crossover designs are used more commonly in a clinic setting with long follow-ups, and require an adequate "wash-out" period between each of the successive interventions. Otherwise, there can be unintended interactions between each intervention.

Separate from their cost, there are

other disadvantages to true experimental designs. It is not always possible to randomly assign patients to the respective study groups. For example, if a research goal is to study the effects of cocaine use, it obviously would be impossible to assign randomly patients to cocaine use, or not. Such a study cannot be done in humans using a "true experimental" design. True experimental designs also are impractical for clinical events or conditions that are very uncommon and in settings where the clinical environment cannot be controlled. As we will discuss, other study designs are more appropriate for those circumstances.

Quasi-experimental Designs

Quasi-experimental designs have manipulation of the independent variable but lack either control or randomization, usually the latter. Though the degree of scientific validity is not as high as in true experimental designs, for some research questions these are the best designs available. Quasi-experimental designs can help to validate treatment methods or establish potential associations. However, because they usually lack random patient selection, and random patient assignment to study groups, there is an increased potential for bias and their validity is limited. As such, they can sometimes be used as a stepping stone to establish the rationale for subsequent focused true experimental designs in the same field. Quasi-experimental designs are generally less expensive than true experimental designs and are sometimes the best or only realistic option for ethical reasons.

The most common quasi-experimental designs are listed and outlined in Table 3. The *group sequential* design is sometimes known as a "single group times series." A single population of subjects is selected and used as its own controls as it goes through a series of observations and interventions. The advantages of the design are two. First, the design controls for potential extraneous variables by using each patient as his/her own control, much as the cross-over design did. Second, the design requires fewer subjects and, therefore, has an application in settings where the number of potential study candidates is limited. The trade-off is that scientific va-

lidity is lower because randomization is absent. In addition, particularly in the acute-care setting, it can be very difficult to track study subjects for lengthy periods of time and put them through a series of sequential interventions.

Cohort studies are among the most popular of the quasi-experimental type. These studies are sometimes called follow-up or longitudinal studies. The term cohort comes from the old Roman armies, where it was used to describe a large circumscribed group of relatively identical soldiers (e.g., all foot soldiers, calvary or archers). In a cohort study, two similar populations are selected. One group has the independent variable of interest at the time of study entry and the other group does not. The groups can be selected concurrently or sequentially, but in either case, the patients are followed for development of the dependent variable (outcome) of interest. For example, a cohort study would be ideal for examining the effects of cocaine use. The control group should consist of closely matched subjects who are clearly not using cocaine. Each group would then be followed and data would be collected. This is why cohort studies are sometimes called follow-up or longitudinal studies. Sometimes the term "cohort analytic study" is used to describe prospective studies meant to establish prognosis. A cohort with a se-

lected characteristic is compared to a concurrent matched control group without that characteristic. Both groups are then followed for development of the dependent variable (e.g., mortality rates). Cohort studies are ideal for studying the effects of relatively rare exposures on outcomes and measuring incidence rates. Sometimes they are also the best option for ethical or other reasons. Because they are a prospective study design, the data collection is generally very complete and, therefore, the scientific validity is much higher than for retrospective studies. However, in addition to some limitations in scientific validity, cohort studies also have other disadvantages. The prolonged follow-up period can be relatively costly and difficult to perform. When the outcomes of interest are rare, there are better study designs (e.g., case control). The results of cohort studies also are better for establishing association between variables than supporting true causality.

Cross-sectional studies are a one-time survey or observation of a population, similar to taking a "snap shot" at a single point in time. The design does not look backwards at antecedent events and, therefore, does not have a retrospective component. The design also does not look forward at outcome or subsequent events. Ideally, cross-sectional studies should be performed in a prospective fashion with decisions made in advance about what data will be collected, thereby increasing the likelihood of a complete data set. These studies are best for establishing prevalence rates (i.e., how common is a given condition?). They also are useful for describing the frequency with which two variables co-exist (e.g., chest pain and dyspnea). These studies are easy and quick to perform while also providing better data than retrospective studies. They are very useful for collecting preliminary data to support subsequent more extensive studies. However, cross-sectional studies have important limitations. First, co-existence of two variables does not prove that they are truly associated. Common variables always will be found to co-exist at a high rate. Second, even if there is strong evidence to establish an association between variables, it does not prove causality. Third, the temporal relationships are often un-

clear (i.e., which was first?). Even though a cross-sectional study might show a high rate of co-existence 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 these 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.

Nonexperimental Designs

Nonexperimental designs are generally retrospective in nature and are sometimes called "ex post facto" research. There is not any manipulation of independent variables possible, because those events have already transpired. In addition, the dependent variable (i.e., the outcome) already has occurred prior to study initiation. These designs also lack the element of "control," making it very difficult to account for potential extraneous variables. There are steps that can be taken to increase the validity of these studies, which are discussed in greater detail in the recommended texts at the end of this article. However, regardless of those precautions, there is tremendous potential for bias in these studies, and they have the lowest level of scientific validity.

The second type of nonexperimental studies often are referred to as "pre-experimental" designs. Pre-experimental designs may be prospective but examine a situation without providing controls to the environment in which the study is conducted or examine a single group of subjects.

Despite their notable limitations, a large number of research publications in clinically oriented journals and presentations at national meetings use nonexperimental study designs. The reason is that these studies are by far the least time consuming, least expensive and easiest to perform of all design types. In addition, a nonexperimental design is a perfectly appropriate stepping stone in the early stages of investigation. Also, there are some research questions that only can be answered through the use of nonexperimental designs. As was true with the

Table 4

Nonexperimental Designs

Case-control

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

Before-and-after

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

Historical controls

- 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.

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.

quasi-experimental designs, it would be unethical or impossible to answer some types of research questions with anything other than a nonexperimental design. These designs are ideal for sorting through large amounts of data in an effort to identify possible factors that then can be studied formally and prospectively. Retrospective derivation of criteria using nonexperimental designs followed by prospective validation using true experimental designs, is a common research process. Development of the Ottawa ankle rules is a good example.[1]

The category of nonexperimental designs is the most heterogeneous of the three general classification categories. Examples of the most common nonexperimental designs are listed in Table 4. Although, in general, this category has the lowest level of scientific validity, each design within this category varies as to its own individual level of scientific validity. In Table 4, the design types are listed in general order of decreasing validity.

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 backward, which is a good description of the relationship between the two study designs. A case-control de-

sign selects two similar populations of patients. One has the dependent variable of interest (i.e., the outcome), and the other does not. The investigator then looks backward, retrospectively, for the independent variables of interest, i.e., looking for the causes for that outcome. This is the opposite of a cohort design where you identify patient populations based on the independent variables being present, and then look prospectively for the subsequent outcomes. Because case-control studies select 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 be studied in a prospective fashion. For such situations, a case-control study is often the best design option. In a case-control study, the selection of the control group is of critical importance. If the control group is not well-matched, there is a potential for important confounding variables, and the results of the entire study can be invalidated. Even though case-control studies are retrospective in nature and are within the nonexperimental 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 inherently was 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 FDA on the basis of a single, nonexperimental design study.

A *before-and-after* design takes advantage of a change being implemented within the environment to look at the effects before and after that change. The investigator is rarely responsible for the change, but simply selects that change as the study intervention. The investigator is an observer to the process, simply taking advantage of the opportunity and collecting data. This unique design can provide intriguing information and sometimes is the only option available for either logistic or ethical reasons. Common examples include studies that have looked at changes in vehicular trauma rates before and after implementation of motorcycle-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 one after the intervention, and the difference between those two measurements is significant, the logical conclusion is that the intervention resulted in the change. However, without multiple data points, it is impossible to tell if the change is both real and persistent. It is possible that separate from the "intervention" the outcome variable was already changing, perhaps due to extraneous variables. It is also possible that initial changes, immediately after the intervention, are only temporary (e.g., after a 55-mph speed limit, everyone eventually returns to faster highway speeds). Multiple pre- and post-measurements would demonstrate these phenomena. Before-and-after designs also can be strengthened through the use of stratification or the use of extensive base-line demographic information for the study populations, both before and after the intervention, to demonstrate comparability. Even with these steps, the potential for substantial confounding variables always exists in this type of design. It is almost impossible to account for all the potential confounding variables. As a result, because of the significant limitations of this

study design, any resultant conclusion should be supported by strong and compelling study results.

When attempting to evaluate the efficacy of a new therapeutic intervention, sometimes the underlying disease process is so severe that it is untenable to have a true control or placebo study group. Sometimes it is the patient or the patient's family that has difficulty accepting random assignment that may include a placebo control. Sometimes it is the clinical investigators who ethically have difficulty with such a study protocol. In these situations, one approach is to use a *historical control* design. Because all three design elements (randomization, manipulation and control) are absent in the control comparison group, and because randomization is generally also absent in the study group, this constitutes a nonexperimental design. In many ways, this is very similar to a before-and-after design, where a change in a therapeutic protocol is instituted and the patient outcomes before and after those changes are compared. Even though this is a nonexperimental design, the results have a higher level of scientific validity if it can be shown that the two study groups are indeed very similar. It is helpful if the historical controls are very close in their temporal relationship to the new study group. Nonetheless, this design has clear scientific limitations with a high potential for investigator bias and should be used only as a last resort. The investigator should consider carefully whether this design approach is absolutely necessary and whether randomization of patients to a control group is truly unethical.

The use of *surveys* or *questionnaires* is a commonly employed nonexperimental research design. It seems straightforward, i.e., ask questions, get answers and tabulate the results. Everyone is used to filling out questionnaires and, therefore, generally feels comfortable with their use. It seems 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. Most journals and national meetings have an inherent bias against survey-type research, often for good reason. Unfortunately, most surveys are done poorly, and the data are not very scientific. Often, there is a strong bias on the part of the investigator behind the survey. Nonetheless, surveys do have a role and are an important study-design option. There are a number of steps that can be taken to improve their quality.

First, keep the survey as simple, short and focused as possible. It is attractive to try to accumulate additional data and answer corollary questions at the same time. However, the longer and more complex the survey becomes, the lower the response rate. One of the important goals of any survey is to maximize the response rate, thereby increasing the potential validity of the findings. As a general rule, any survey with a response rate less than 75% is highly suspect. The goal should be a response rate of more than 85%, if at all possible. There is a logical reason for this. Individuals with particularly strong opinions are those who are most motivated to complete a questionnaire on a given topic. Therefore, the initial survey results may represent only the most "radicalized" segment of the survey population. When considering the results of a survey study, ask 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. Other steps to increase the validity of survey-based studies are to use existing validated-measurement instruments whenever possible. Rather than developing a new scoring system or a new measurement scale for a parameter (e.g., pain or happiness), it is much better to use something that already has been validated through other studies.

No matter how well you design a survey or questionnaire, there are always potential misunderstandings that will occur. It is best to pilot test the questions and revise, revise, revise before general circulation of the final survey forms. Despite multiple limitations, there are certain research questions that only can be addressed using surveys. Unfortunately, there are many examples of poorly done surveys and, as a result, a general bias against them in the research community.

Case series can range from a comprehensive retrospective review to an expansion of the individual case report. Usually a case series involves 3–10 patients, all with a defined condition. To be worthy of publication, that condition must either be newly recognized, quite unique or highly educational. Case series are generally more compelling than a single case report because they demonstrate that the condition has existed more than once. However, case series can be strengthened by making them a comprehensive and consecutive collection of patients that includes all relevant cases within a set time period and a given institution. There is no concurrent control group. Case series often are useful to bring attention to a new area of concern that can then be investigated in a more scientific and prospective fashion. However, for very uncommon disorders, a "case series" may be the only way to study the subject, and prospective investigations may be impossible. For example, historically it was recommended that all patients with spontaneous pneumomediastinum be admitted to the hospital and receive serial chest radiographs. However, a single large (retrospective) case series[2] showed that neither of those steps were necessary, and such patients could be discharged home safely, thus, changing national practices. This is a disorder that could never be studied prospectively because of its rarity.

The *case report* continues to hold a surprisingly strong position of respect within the medical literature. The case report has its origins in the way medicine was originally taught, i.e., largely through apprenticeship with an emphasis on powers of careful observation. As such, education largely 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 direct hands-on experience rather than simply hearing about events second hand. The case report is meant to replicate this sort of experience. Currently, there are three kinds of case reports that merit publication. First, the highly unique case that may represent a previously undescribed syndrome or disease. Second, the case

that demonstrates an unexpected associ-ation between two or more diseases or disease manifestations and may repre-sent an unsuspected causal relationship. Last, the case with an unexpected evolu-tion suggesting a surprising new thera-peutic effect or adverse drug effect. Case reports are intended to be somewhat en-tertaining, as well as educational. The goal is for the reader response to be "that's interesting" instead of "so what?". When the findings in a single case report are not always fully adequate or entirely representative of the disease process, it is common to include a discussion of previously described cases that have some features in common. This kind of paper that collates and interprets previ-ous reports, as well as the reported case, is known as "a case report with a review of the literature." Even in such situations though, the reported case must be suffi-cient to stand largely on its own. If not, then it is best to drop the case entirely and consider writing an article that re-views the literature on the subject. In other words, a weak case report cannot be rescued simply by adding in a "review of the literature." There is an art to writ-ing a case report. Some additional sug-gested readings on this subject are listed at the end of this article.

Miscellaneous Designs

There are a number of research method-ologies that do not necessarily fit the "sci-entific rigor" classification system. Some of these are relatively recent develop-ments that reflect newer statistical tech-niques, or the impact of personal computer capabilities to handle and ana-lyze large data bases. These research de-signs do not involve collection of new data themselves, but rather the reanalysis of data previously collected and reported by others. In some cases, modeling tech-niques are used to generate theoretical data, where none actually exists. The de-gree of scientific validity of these designs is highly variable and often open to subjec-tive interpretation. The use of these de-signs can be an area of tremendous controversy, however. With growing inter-est in both evidence-based medicine and cost-effectiveness, it is likely that the num-ber of publications using these method-ologies will increase in the future.

META-Analysis is the statistical analysis

of results from two or more independent studies, for the purpose of integrating their findings and developing overall con-clusions. 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 impor-tant objective is to suggest directions for future research based on questions that remain unanswered by the literature. The statistical methods used in performing a META-Analysis varies between studies, and their appropriateness often is de-bated. With newer and easier personal computer-based statistics programs, no one need be a sophisticated statistician to perform this type of analysis. More impor-tantly, the criteria used to include a previ-ous study in the META-Analysis is 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, blind-ing or at least a control group? The over-all 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.

Economic evaluation studies involve a comparative analysis of alternative courses of action in terms of costs, conse-quences and relative benefits. These analyses involve two important elements. One is a medical "modeling" system that maps out decision alternatives and courses of action in a theoretical fashion. In addition, they involve a large num-ber of "assumptions" regarding the fre-quency of clinical events and the costs of therapeutic decisions. There are gener-ally two categories of economic evalua-tion. A *cost-benefit analysis* is a form of economic assessment in which the costs of medical care are compared with the economic benefits of that care. The bene-fits generally include a calculation for in-creased earnings due to improved health,

as well as potential reductions in future health-care costs. Generally, these calcu-lations are done from a societal perspec-tive. A *cost-effectiveness analysis* compares alternative programs, therapies or other interventions, in terms of their overall costs per degree of clinical effect. For ex-ample, it could be cost-per-life saved, per additional year of life gained or per in-crease of 1% in the hematocrit. Though the scientific approaches used in eco-nomic evaluation analyses are becoming more sophisticated, this methodology is not yet standardized and has tremendous variability in the methods used within dif-ferent studies. This type of research gen-erally requires a team effort, including panels of clinicians, biostatisticians and, sometimes, economists. The work itself is often quite tedious. This is not a form of research for the novice.

Retrospective reviews involve a review of a body of literature on a given subject, usually with resultant conclusions and recommendations. It can be the work of 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 compulsive and comprehensive review of all literature on the subject in a highly ob-jective fashion. Other times, such articles are little more than the biased opinion of the author(s), and the resultant scientific validity is nearly nonexistent. Now that sophisticated computer literature search capabilities are readily available, the abil-ity to comprehensively identify the body of literature on a given subject is much easier. It is no longer appropriate to rely on one's personal filing system to identify the relevant literature. Retrospective re-views are strengthened by a discussion of the methodology used both in identify-ing the relevant body of literature, as well as analyzing and weighting individual ar-ticles. Such reviews can make significant contributions to the scientific literature if performed properly and are being used increasingly to support "evidence based medicine" recommendations.

Summary

There are a large number of clinical re-search designs that have stood the test of time. There is no single research design that is best to answer all research ques-tions, and every research design has ap-

propriate applications. This article describes the most common designs and a recommended classification system. However, in the process of performing research, there are always other options and "hybrid" studies combining elements from different designs, are not uncommon. Though 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, as well as disadvantages. In addition, there are usually many different ways to answer the same research question. Which design is most appropriate is dependent largely on a stage of evolution of the investigative process.

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, it is best to use the most scientifically valid design that the circumstances will allow. However, the actual decision regarding the research 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 for the research, including the time frame that is available. Realize that research is done in incremental steps, and it is unusual to be able to answer the entire research question in a single study. It must be emphasized that the process of planning and revising the protocol, prior to 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 actual data also is important.

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. Stiell IG, McKnight RD, Greenlier GH, et al: Implementation of the Ottawa ankle rules. *JAMA* 1994;271:827–831.
2. Panacek EA, Singer AJ, Sherman BW, Rutherford WF: Spontaneous pneumomediastinum: Clinical and natural history. *Ann Emerg Med* 1992;21:1222–1227.

Suggested Readings

1. Nakam AM: The clinical case report: "Pot boiler" or scientific literature? *Head and Neck Surgery* 1979;3:291–292.
2. DeBakey L, DeBakey S: The case report: I. Guidelines for preparation. *Int J Cardiol.* 1983;4:357-364.
3. Okolo EN: *Health Research Design and Methodology*, Boca Raton, Fla., CRC Press Inc., 1990.
4. Hulley SB, Cummings SR: *Designing Clinical Research*, Philadelphia, Pa., Williams and Wilkins, 1988.
5. Task Force on Principles: Economic analysis of health-care technology: A report on principles. *Ann Intern Med* 1995;122:61–70.
6. Chalmers I: The Cochrane collaboration: Preparing, maintaining and disseminating systematic reviews of the effects of health care. *Ann NY Acad Sci* 1993;703:156–163.
7. Mulrow CD: Rationale for systematic reviews. *Brit Med J* 1994;309:597–599.
8. Oxman AD: Checklists for review articles. *Brit Med J* 1994;309:648–651.
9. Eysenck HJ: META-Analysis and its problems. *Brit Med J* 1994;309:789–792.
10. Campbell DT, Stanley JC: *Experimental and Quasi-Experimental Designs for Research*, Boston, Mass., Houghton Mifflin Company, 1963.
11. Pollt DF, Hungler BP: *Nursing Research: Principles and Methods* (5th ed.), Philadelphia, Pa., Lippincott Company, 1995.