Getting Started: The Anatomy and Physiology of Clinical Research

Stephen B. Hulley, Thomas B. Newman, and Steven R. Cummings

This chapter introduces clinical research from two viewpoints, setting up themes that run together through the book. One theme is the anatomy of research-what it's made of. This includes the tangible elements of the study plan: the research question, design, subjects, measurements, sample size calculation, and so forth. An investigator's goal is to create these elements in a form that will make the project fast, inexpensive, and easy.

The other theme is the physiology of research-how it works. Studies are useful to the extent that they yield valid inferences, first about what happened in the study sample and then about generalizing these events to people outside. the study. The goal is to minimize the errors, random and systematic, that threaten conclusions based on these inferences.

Separating these two themes is artificial in the same way that the anatomy of the human body does not make much sense without some understanding of its physiology. But the separation has the same advantage: It clarifies our thinking about a complex topic.

● THE ANATOMY OF RESEARCH: WHAT IT'S MADE OF

The structure of a research project is set out in its protocol, the written plan of the study. Protocols are well kn~wn as devices for seeking grant funds, but they also have a vital scientific function: helping the investigator to organize her research in a logical, focused, and efficient way. Table 1.1 outlines the components of a protocol. We will introduce the whole set here, expand on each of them in the ensuing chapters of the book, and return to put the completed pieces together in Chapter 19.

The Research Question

The research question is the objective of the study, the uncertainty that the investigator wants to resolve. Research questions often begin with a general concern that must be narrowed down to a concrete, researchable issue. For example,

TABLE 1.1

Outline of the Study Protocol

Element	Purpose
Research questions	What questions will the study address?
Significance (background)	Why are these questions Important?
Design	How is the study structured?
Time frame	
Epidemiologic approach	
Subjects	Who are the subjects and how will they be selected?
Selection criteria	
Sampling design	
Variables	What measurements will be made?
Predictor variables	
Confounding variables	
Outcome variables	
Statistical issues	How large is the study and how will It be analyzed?
Hypotheses	
Sample size	
Analytic approach	

Initial research question: Should women take hormones after menopause?

This is a good place to start, very practical and important, but the question must be focused before planning efforts can begin. Often this involves breaking the whole question into its constituent parts and singling out one or two of these to build the protocol around.

More specific research questions:

How commonly women take estrogen after menopause?
Does taking estrogen after menopause lower the likelihood of developing coronary heart disease (CHD)?
Are there other benefits and harms of estrogen treatment?

A good research question should pass the "50 what?" test. Getting the answer should contribute usefully to our state of knowledge. The acronym *FINER* denotes five essential characteristics of a good research question: that it be *feasible, interesting, novel, ethical,* and *relevant* (Chapter 2).

Significance

The significance section of a protocol sets the proposed study in context and gives its rationale: What is known about the topic at hand? Why is the research question important? What kind of answers will the study provide? This section cites previous research that is relevant (including the investigator's own work) and indicates the problems with that research and what questions remain. It makes clear how the findings of the proposed study will help resolve these uncertainties, leading to new scientific understanding and influencing clinical and public health policy.

The Design

The design of a study is a complex issue. A fundamental decision is whether to take a passive role in observing the events taking place in the study subjects in

an observational study or to apply an intervention and examine its effects on these events in a clinical trial (Table 1.2). Among observational studies, two of the most common designs are cohort studies, in which a group of subjects is followed over time, and cross-sectional studies, in which the observations are made on a single occasion. A third common option is the case-control design, in which the investigator compares a group of subjects who have a disease or condition with another group of subjects who do not. Another design decision is whether to deal with past events in a retrospective study or to follow study subjects prospectively for events .that have not yet occurred. Among clinical trial options, the randomized blinded trial is often the best design but unblinded or time~series designs may be more suitable for some research questions.

No one approach is always better than the others, and each research question requires a judgment about which design is the most efficient way to get a satisfactory answer. The randomized blinded trial is often held up as the gold standard for establishing causality and the effectiveness of interventions, but there are many situations for which an observational study is a better choice or the only feasible option. The relatively low cost of case-control studies and their suitability for uncommon outcomes makes them attractive for many questions. Special considerations apply to choosing designs for studying diagnostic tests. These issues are discussed in Chapters 7 through 12.

A typical sequence for studying a topic begins with observational studies of a type that is often called descriptive. These studies explore the lay of the land, for example describing distributions of diseases and health-related characteristics in the population *(How common is estrogen treatment in women after menopause?)* or the sensitivity and specificity of a diagnostic test. Descriptive studies are usually

TABLE 1.2

Examples of Common Clinical Research Designs Used to Study Whether Hormone Therapy After Menopause Prevents Coronary Heart Disease

Study Design	Key Feature	Example
Observational Designs		
Cohort study	A group followed over time	The Investigator examines a cohort of women yearly for several years, observing the incidence of heart attacks in hormone users and non-users.
Cross-sectional study	A group examined at one point in time	She examines the group of women once, observing the prevalence of a history of heart attacks in hormone users and nonusers
Case-control study	Two groups, based on the outcome	She examines a group of women with heart attacks (the "cases") and compares them with a group of healthy women (the controls), asking about hormone use.
Experimental Design		
Randomized blinded trial	Two groups created by a random process, and a blinded intervention	She randomly assigns women to receive hormone or identical placebo, then follows both treatment groups for several years to observe the incidence of heart attacks.

followed or accompanied by analytic studies that evaluate associations to discover cause-and-effect relationships *(Is taking estrogen after menopause associated with lower risk of CHD?)*. The final step is often a clinical trial to establish the effects of an intervention *(Does hormone treatment alter the incidence of CHD?)*. Experiments usually occur later in the sequence of research studies, because they tend to be more difficult and expensive, and answer more narrowly focused questions that arise from the findings of observational studies.

It is useful to characterize the design in a single sentence that begins with its name. Some studies do not easily fit into these molds, however, and classifying them can be a surprisingly difficult exercise. It is worth the effort-a precise description of the type of study clarifies the investigator's thoughts and is useful for orienting colleagues and consultants. (This single sentence is the research analog to the opening sentence of a medical resident's report on a new hospital admission: "This 62-year-old white policewoman was well until 2 hours before admission, when she developed crushing chest pain radiating to the left shoulder.") If the study has two major phases, the design for each should be mentioned.

> *Research design:* This is a cross-sectional study of the prevalence of estrogen treatment among women aged 50 to 69 years, followed by a prospective cohort study of whether estrogen treatment is associated with low risk of subsequent heart attacks.

The Study Subjects

Two major decisions must be made in choosing the study subjects (Chapter 3). The first is to specify selection criteria that define the target population: the *kinds* of patients best suited to the research question. The second decision concerns how best to recruit enough women from an accessible aspect of this population who will be the actual subjects of the study. For example, the study of hormones and CHD in women might select women aged 50 to 69 years attending primary care clinic at the investigator's hospital, and the investigator might decide to invite the next 1,000 such patients. These design choices represent trade-offs; studying a random sample of all U.S. women of that age would enhance generalizability but be formidably difficult and costly.

The Variables

Another major set of decisions in designing any study concerns the choice of which variables (characteristics that vary from one study subject to another) to measure (Chapter 4). In a descriptive study the investigator looks at individual variables one at a time. A study of the prevalence of hormone treatment, for example, might record the presence or absence of the self-report of taking estrogen. .

In an analytic study the investigator studies the associations among two or more variables in order to predict outcomes and to draw inferences about cause and effect. In considering the association between two variables, the one that precedes (or is presumed on biologic grounds to be antecedent) is called the predictor variable; the other is called the outcome variable. * Most observational

‒Predictor variables are often termed independent and outcome variables are termed dependent, but we find this usage confusing, particularly since independent means something quite different in the context of multivariate analyses.

studies have many predictor variables (e.g., estrogen treatment, blood cholesterol, age, race), and several outcome variables (e.g., heart attacks, strokes).

Experiments study an intervention (a special kind of predictor variable that the investigator manipulates), such as treatment with estrogen or placebo. This design allows her to observe the effects on the outcome variable, often using randomization to control for the influence of confounding variables (other predictors such as age that could confuse the interpretation of the findings).

Statistical Issues

The investigator must develop plans for managing and analyzing the study data. For experiments this always involves specifying a hypothesis, a version of the research question that provides the basis for testing the statistical significance of the findings (Chapter 5).

> *Hypothesis:* Women who receive estrogen treatment after menopause will have fewer heart attacks than those who do not.

The hypothesis also allows the investigator to estimate the sample size, the number of subjects needed to observe the expected difference in outcome between study groups with a reasonable degree of probability, or power (Chapter 6).

Most observational studies have an analytic component that also benefits from having a prior hypothesis. For purely descriptive studies (e.g., the prevalence of CHD in women 50 to 69 years of age), an analogous approach estimates the number of subjects needed to produce an acceptable level of precision when confidence intervals are calculated for the means, proportions, or other descriptive statistics.

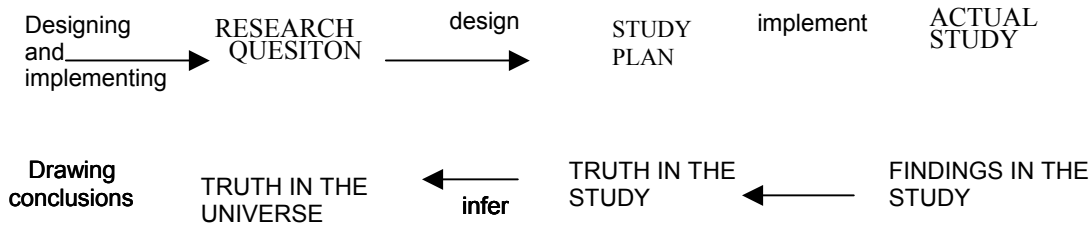## . THE PHYSIOLOGY OF RESEARCH: HOW IT WORKS'

The goal of clinical research is to draw *inferences* from the study results about the nature of truth in the universe (Fig. 1.1). With this in mind, the process of doing research reverses this sequence. Beginning with a decision about what health problem the investigator wishes to address, the investigator undertakes a study that will answer this research question.

This undertaking involves two distinct steps (Fig. 1.2). The first is to *design* a study plan with subjects and measurements chosen to enhance the process of appropriately answering the research question and generalizing these conclusions to the people and phenomena addressed by the research question. The second step is to *carry out* the study in a fashion that enhances the likelihood of getting
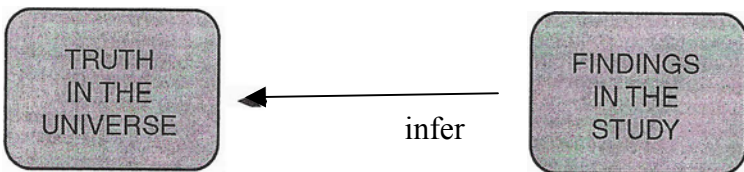
● FIGURE 1.1

The findings of a study lead to Inferences about the universe outside.

Designing and implementing: RESEARCH QUESITON → design → STUDY PLAN → implement → ACTUAL STUDY

Drawing conclusions: TRUTH IN THE UNIVERSE ← infer ← TRUTH IN THE STUDY ← FINDINGS IN THE STUDY

● **FIGURE** 1.2.
The process of designing and implementing a research project sets the stage for drawing conclusions from the findings.

the right answer; in other words, to draw the correct conclusions about what actually happened in the study.

In this section we will first address the design side of Fig. 1.2, then the implementation side. We will then turn to the errors that threaten the validity of clinical research inferences.

Designing the Study

The research question, as noted earlier, is what the investigator really wants to answer. For the purpose of illustration, we will use a simple descriptive question: *How common is it for women to take estrogen after menopause?*

This question cannot be answered with perfect accuracy because it would be impossible to study all postmenopausal women and because our instruments for discovering whether a woman is taking estrogen are imperfect. So the investigator must settle for a related question that can be answered by the study, such as the following: *"Among postmenopausal women seen for the first time in her primary care clinic, what proportion report on a questionnaire that they are taking estrogen?"* The transformation from research question to study plan is illustrated in Fig. 1.3.

One major component of this transformation is the choice of a sample of subjects that will represent the population. The group of subjects specified in the protocol can only be a sample of the population of interest because there are practical barriers to studying the entire population. The decision to study patients entering the UCSF primary care clinics is a compromise. This is a sample that is feasible to study but one that may produce a different prevalence of estrogen treatment than that found in all women.

The other major component of the transformation is the choice of variables that will represent the phenomena of interest. The variables specified in the study plan are usually proxies for these phenomena. The decision to use a self-report questionnaire to assess estrogen treatment is a fast and inexpensive way to collect information, but it will not be perfectly accurate. Some women will not give the right answer because they know they're taking hormones but don't recognize the phrase *estrogen treatment* that appears in the questionnaire, or because they may have forgotten what they are taking.

In short, each of the differences in Fig 1.3 between the research question and the study plan has the purpose of making the study more practical. The cost of this increase in practicality, however, is the risk that the study may produce a wrong answer to the research question. Figure 1.3 illustrates the important fact that errors in designing the study are a common reason for getting the wrong answer to the research question.

Implementing the Study

Returning to Fig. 1.2, the right-hand side is concerned with implementation, and the degree to which the actual study matches the study plan. At issue here is the problem of a wrong answer to the research question because the way the sample was actually drawn and the measurements made differed in important ways from the way they were designed (Fig. 1.4).

The actual sample of study subjects is almost always different from the intended sample. The plans to study all age-eligible women entering primary care clinics, for example, would probably be disrupted by incomplete attendance (say only 150 of the 278 patients who are scheduled for first visits ever show up during the year of the study) and by refusal to participate (say only 100 of these consent to be studied). The 100 patients who agree to be studied may have a different prevalence of estrogen treatment from those who do not show up or refuse. In addition to these problems with the subjects, the actual measurements can differ from the intended measurements. If the format of the questionnaire is unclear, the women may get confused and check the wrong box, for example, or they may simply omit the question by mistake.

These differences between the study plan and the actual study can alter the answer to the research question. Figure 1.4 illustrates the important fact that errors in implementing the study are the other common reasons (besides errors of design) for getting the wrong answer to the research question.

Drawing Causal Inference

A special kind of validity problem arises in studies that examine the association between a predictor and an outcome variable in order to draw causal inference. If the study finds an association between hormone therapy and heart attacks, does this represent cause and effect, or is there some other explanation? Reducing the likelihood of confounding and other rival explanations is one of the major challenges in designing an observational study (Chapter 9).

The Errors of Research

No study is free of errors, and the inferences that have been described are never perfectly valid. The goal is simply to maximize the validity of drawing inferences from what happened in the study sample to reach conclusions about the nature of things in the population. Erroneous inferences can be addressed in the analysis phase of research, but the best strategies are focused on design and implementation (Fig. 1.5): Preventing errors from occurring in the first place, to the extent that it is practical and economic to do so.:
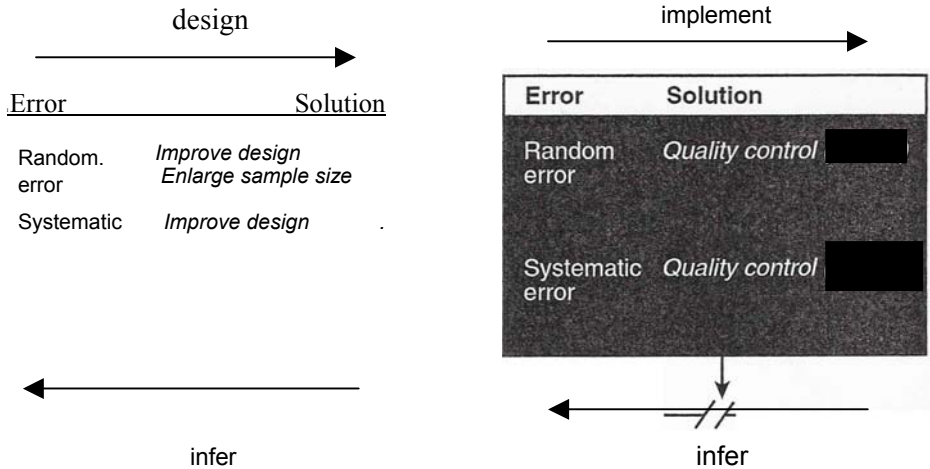
The two main kinds of error that interfere with research inferences are random error and systematic error. The distinction is important because the strategies for minimizing them are quite different.

Random error is a wrong result due to chance - unknown sources of variation that are equally likely to distort the sample in either direction. If the true prevalence of estrogen treatment in 50- to 69-year-old women is 20%, a well-designed sample of 100 patients from that population might contain exactly 20 patients with this disease. More likely, however, the sample would contain a nearby number such as 18, 19, 21, or 22. Occasionally, chance would produce a substantially different number, such as 12 or 28. Among several techniques for reducing the influence of random error, the simplest and best known is to increase the sample size. The use of a larger sample diminishes the likelihood of a wrong result by increasing the precision of the estimate-the degree to which the observed prevalence approximates 20% each time a sample is drawn.

Systematic error is a wrong result due to bias (sources of variation that distort the study findings in one direction). An illustration is the decision in Fig. 1.3 to use patients who come to the primary care clinic, who might be more likely than average to adopt medical treatments. Increasing the sample size has no effect on systematic error. The only way to improve the accuracy of the estimate (the degree to which it approximates the true value) is to design the study in a way that either reduces the size of the various biases or gives some information about them. An example would be to draw a second sample of women from a setting that may be less likely to bias the proportion of women treated with estrogen (e.g., employees in a corporation), and to compare the observed prevalence in the two samples.

The examples of random and systematic error in the preceding two paragraphs are components of sampling error, which threatens inferences from the study subjects to the population. Both random and systematic errors can also contribute to measurement error, threatening the inferences from the study measurements to the phenomena of interest. An illustration of random measurement error is the variation in the response when a questionnaire is administered on several occasions. An example of systematic measurement error is the underestimation of the prevalence of estrogen treatment due to lack of clarity in how the question is phrased. Additional strategies for controlling all these sources of error are presented in Chapters 3 and 4.

The concepts presented in the last several pages are summarized in Fig. 1.6. Here is an important bottom line: Getting the right answer to the research question is a matter of designing and implementing the study in a fashion that keeps the extent of inferential errors at an acceptable level.

design   implement

Error	Solution
Random. error	*Improve design* *Enlarge sample size*
Systematic	*Improve design* .

Error	Solution
Random error	*Quality control*
Systematic error	*Quality control*

infer   infer

. FIGURE 1.5

Research errors. These can have both random and systematic elements, as indicated in this blown-up version of the error boxes In Figs. 1.3 and 1.4. The boxes summarize the design and Implementation strategies for minimizing the effects of these errors.

DESIGNING THE STUDY

Developing the Study Protocol

The first step in designing a study is to establish the research question. This task is discussed at length in Chapter 2. Once the research question is in hand, the process of developing the study plan can begin. There are three versions of the study plan that are produced in sequence, each larger and more detailed than the preceding one.

- A one- to two-page outline of the elements of the study. We recommend the sequence in Table 1.1. It serves as a standardized checklist to remind the investigator to include all the components. Just as important, the sequence has an orderly logic that helps clarify the investigator's thinking on the topic.
- The study protocol, an expansion on the one- to two-page outline that can range from five to 25 or more pages. The full protocol is the main document used to plan the study and to apply for grant support; we discuss parts of it throughout this book and put them all together in Chapter 19.
- The operations manual, a collection of specific procedural instructions, questionnaires, and other materials designed to ensure a uniform and standardized approach to carrying out the study with good quality control

The study question and one- to two-page outline should be written out at an early stage. Putting thoughts down on paper leads the way from vague ideas to specific plans and provides a concrete basis for getting advice from colleagues and consultants. It is a challenge to do it (ideas are easier to talk about than to write down), but the rewards are a faster start and a better project.

Appendix 1.1 provides an example of a one-page study plan. As usual, this plan deals more with the anatomy of research (Table 1.1) than with its physiology (Fig. 1.6), so the investigator must remind herself to worry about the errors that may result when it comes time to draw inferences about what happened in the study sample and formulate conclusions for the population.

A study's virtues and problems can be revealed by explicitly considering how the question the study is likely to answer differs from the research question, given the plans for acquiring subjects and making measurements, and given the likely problems of implementation.

With the one- to two-page outline in hand and the intended inferences in mind, the investigator can proceed with the details of her protocol. This includes getting advice from colleagues, drafting specific recruitment and measurement methods, considering scientific and eth!cal appropriateness, changing the study question and outline, pretesting specific recruitment and measurement methods, making more changes, getting more advice, and so forth. This iterative process is the nature of research design and the topic of the rest of this book.

SUMMARY

1. The anatomy of research is the set of tangible elements that make up the study plan: the research question, design, study subjects, and measurement approaches. The challenge is to design a study plan with elements that are fast, inexpensive, and easy to implement.

2. The physiology of research is how the study works. The study findings are used to draw inferences about what actually happened in the study sample and about events in the universe outside. The challenge here is to design and implement a study plan with adequate control over two major threats to these inferences: random error (chance) and systematic error (bias).

3. A good way to develop the study plan is to write a one-sentence summary and to expand this into a one- to two-page outline that sets out the study elements in a standardized sequence. Later on the study plan will be expanded into the protocol and the operations manual.

4. The next step is to consider the main inferences that will be drawn from the study subjects to the population and from the study measurements to the phenomena of interest. At issue here are the relationships between the research question (what the investigator really wants to answer in the universe), the study plan (what the study is designed to answer), and the actual study (what the study will actually answer, given the errors of implementation that can be anticipated).

5. Good judgment by the investigator and advice from colleagues are needed for the many trade offs involved and for determining the overall viability of the project.

- ## APPENDIX 1.1
  ### Outline of a Study*

Element	Example
Title	Effects of Hormone Treatment after Menopause on lipoprotein (a) (Lp(a))
Research question	What are the effects of treatment with estrogen plus progestin (compared with placebo) on Lp(a) levels in postmenopausal women?
Significance	1. Epidemiologic studies suggest that hormone treatment after menopause may help prevent coronary heart disease, the largest cause of death in women.
	2. Lp(a) Is an understudied lipoprotein that has been found to be an Independent risk factor for coronary disease In several studies.
	3. Among conventional lipid-lowering drugs, only nicotinic acid In high doses lowers Lp(a) levels; however, previous studies have suggested that hormone treatment may have this effect.
	4. There is a need to confirm this finding for the estrogen plus progestin treatment that Is now commonly used after menopause, and to extend It to women with existing coronary disease.
Design	Randomized blinded trial with one year of follow-up.
Subjects	
Entry criteria	Postmenopausal women with documented coronary disease (evidence for prior myocardial Infarction or coronary artery surgery, or 50% obstruction on angiography).
Recruitment	Consecutive sample of all women who qualify In 20 clinical centers, recruited In cardiology clinics and by mailings and advertisements.
Variables	
Predictor	Randomization to a dally tablet containing conjugated equine estrogen (0.65 mg) and medroxy-progesterone acetate (2.5 mg), or to a placebo identical In appearance.
Outcome	Change in serum level of Lp(a) between baseline and 1 year after randomization, measured Immunochemlcally with a sandwich ELISA assay that uses a monoclonal antibody to apo(a) as the capture antibody (Strategic Diagnostics, Newark, DE).
Statistical Issues	
Hypothesis	There will be a greater decrease in Lp(a) levels In the hormone-treated group than In the placebo group.
Sample size and power	The number of women in the existing HERS trial available for this ancillary study was 2,763. This allows detection of a reduction in Lp(a) of 2 mg/dL with a power of 90%, using at-test and two-tailed alpha of 0.05.