Phyllis G. Supino • Jeffrey S. Borer
Editors

Principles of Research Methodology

A Guide for Clinical Investigators

Foreword by Stephen E. Epstein

Springer

Contents

Contributors

Jeffrey S. Borer, MD Department of Medicine, Division of Cardiovascular Diseases, Howard Gilman Institute for Valvular Heart Diseases, and Cardiovascular Translational Research Institute, State University of New York (SUNY) Downstate Medical Center, Brooklyn, NY, USA

Todd A. Durham, MS Axio Research, LLC, Seattle, WA, USA

Helen-Ann Brown Epstein, MLS, MS, AHIP Clinical Librarian, Samuel J. Wood Library and C.V. Starr Biomedical Information Center, Weill Cornell Medical College, New York, NY, USA

Peter L. Flom, PhD Peter Flom Consulting, LLC, New York, NY, USA

Joseph A. Franciosa, MD Department of Medicine, SUNY, Downstate Medical Center, Brooklyn, NY, USA

Eli A. Friedman, MD Department of Medicine, SUNY, Downstate Medical Center, Brooklyn, NY, USA

Mario Guralnik, PhD Synergy Research Inc, Irvine, CA, USA

Gary G. Koch, PhD Department of Biostatistics, University of North Carolina at Chapel Hill Gillings School of Global Public Health, Chapel Hill, NC, USA

Lisa M. LaVange, PhD Department of Biostatistics, University of North Carolina at Chapel Hill Gillings School of Global Public Health, Chapel Hill, NC, USA

Martin L. Lesser, PhD Biostatistics Unit, Departments of Molecular Medicine and Population Health, Feinstein Institute for Medical Research, Hofstra North Shore-LIJ School of Medicine, Manhasset, NY, USA

Lorenzo Paladino, MD Department of Emergency Medicine, SUNY Downstate Medical Center, Brooklyn, NY, USA

N. Philip Ross, BS, MS, PhD SUNY Downstate Medical Center, Bethesda, MD, USA

conduct a research project" and to expect to design, execute, and complete it in that time frame.

There is general consensus that information gathering, including reviewing and synthesizing the literature, is a critically important activity to be undertaken by an investigator. However, in and of itself, it is not research. The same can be said for data gathering activities aimed at personal edification or those undertaken to resolve organization-specific issues. So what, then, characterizes research?

Tuckman [3] has argued that in order for an activity to qualify as research, it should possess a minimum of five characteristics:

1. *It should be systematic.*
   While some important research findings have occurred serendipitously (e.g., Fleming's accidental discovery of penicillin, Pasteur's chance finding of microbial antibiosis), most arise out of purposeful, structured activity. Structure is engendered by a series of the rules for defining variables, constructing hypotheses, and developing research designs. Rules also exist for collecting, recording, and analyzing data, as well as for relating results to the problem statement or hypotheses. These rules are used to generate formal plans (or protocols) which guide the research effort, thereby optimizing the likelihood of achieving valid results.
2. *It should be logical.*
   Research employs logic that may be inductive, deductive, or abductive in nature. Inductive logic is employed to develop generalizations from repeated observations, abductive logic is used to form generalizations that serve as explanations for anomalous events, and deductive logic is used to generate specific assertions from known scientific principles or generalizations. Further elaboration of these distinctions is covered in Chap. 3. Logic is used both in the development of the research design and selection of statistics to ensure that valid inferences may be drawn from data (internal validity). Logic also is used to generalize from the results of the particular study to a broader context (external validity or extrapolability).
3. *It should be empirical.*
   Despite the deductive processes that may precede data collection, the findings of research must always be based on observation or experience and, thus, must relate to reality. It is the empirical quality of research that sets it apart from other logical disciplines, such as philosophy, which also attempts to explain reality. Recognition of this fact may pose a problem for physicians who, according to some researchers [4, 5], have a cognitive style that tends to be more deterministic than probabilistic, causing personal experience to be valued more than data. Under these circumstances, the importance of subordinating the hypothesis to data may not be fully appreciated. As part of the education of the physician scientist, he or she must learn that when confronted with data that do not support the study hypothesis, it is the hypothesis and *not* the data that must be discarded, unless it is abundantly clear that something untoward occurred during the performance of the study.
4. *It should be reductive.*
   As Tuckman [3] has noted, a fundamental purpose of research is to reduce "the confusion of individual events and objects to more understandable categories of concepts" (p. 11). One heuristic tool used by scientists for this purpose is the creation of abstractive constructs such as "intervening variables" (e.g., *resistance* and *solubility* in the physical sciences, *conditioning or reflex reserve* in the behavioral sciences) to explain how phenomena cause or otherwise interact with each other [6]. Another powerful tool available to the researcher for this purpose is a constellation of techniques for numerical and graphical data analysis (the specific methodology employed depending on the objectives and design of the study as well as the number of observations generated by the study). As Tuckman observes, whenever data are subjected to analysis, some information is lost, specifically the uniqueness of the individual observation. However, such losses are offset by gains in the capacity to

conceptualize general relationships based on the data. As a result, the investigator can explain and predict, rather than merely describe.

5. *It should be replicable and transmittable.*
   The fact that research procedures are documented makes it possible for others to conduct and attempt to replicate the investigation. The ability to replicate research results in the confirmation (or, in some unhappy cases, refutation) of conclusions. Confirmation of conclusions, in turn, results in the validation of research and confers upon research a respectability that generally is absent in other problem-solving processes. In addition, the fact that research is transmittable also enables the general body of knowledge to be extended by subsequent investigations based on the research. For this reason, researchers are encouraged to present their findings as soon as possible at local, national, and international scientific sessions and to publish them expeditiously as letters (communications) or full-length articles in peer-reviewed journals (to ensure their quality and validity).
6. *It should contribute to generalizable knowledge.*
   The Tuckman criteria speak to the structure and process of research, but not to its intended objectives. The Belmont Report [7], which codified the definition of human subjects research for the US Department of Health and Human Services, argues additionally that for an activity to be considered research, it must contribute to generalizable knowledge (the latter expressed in theories, principles, and statements of relationships). For knowledge to be generalizable, the *intent* of the activity must be to extrapolate findings from a sample (e.g., the study subjects) to a larger (reference) population to define some universal "truth," and be conducted by individuals with the requisite knowledge to draw such inferences [8]. Because research seeks generalizable knowledge, it differs from information gathering for diagnosis and management of individual patients. It also differs from formal evaluation procedures (e.g., review of data performed for clinical quality improvement [CQI] or formative and summative appraisals of educational programs) which, while employing many of the same rigorous and systematic methodologies as scientific research, principally aim to inform decision making about particular activities or policies rather than to advance more wide-ranging knowledge or theory. As Smith and Brandon [9] have noted, research "generalizes" whereas evaluation "particularizes."

Types of Research

There are multiple ways of classifying research, and the categorizations noted below are by no means exhaustive. Research can be classified according to its theoretical versus practical emphasis, the type of inferential processes used, its orientation with respect to data collection and analysis, its temporal characteristics, its analytic objective, the degree of control exercised by the investigator, or the characteristics of the measurements made during the investigation. These yield the following categorizations: basic versus applied versus translational, hypothesis testing versus hypothesis generating, retrospective versus prospective, longitudinal versus cross-sectional, descriptive versus analytic, experimental versus observational, and quantitative versus qualitative research.

Basic Versus Applied Versus Translational Research

Traditionally, research in medicine, as in other disciplines, has been classified as basic or applied, though the lines between the two can, and do, intersect. In basic research (alternatively termed "fundamental" or "pure" research), the investigation often is driven by scientific curiosity or interest in a conceptual problem; its objective is to expand knowledge by exploring ideas and questions and developing models and theories to explain phenomena. Basic research typically does not seek to provide immediate solutions to

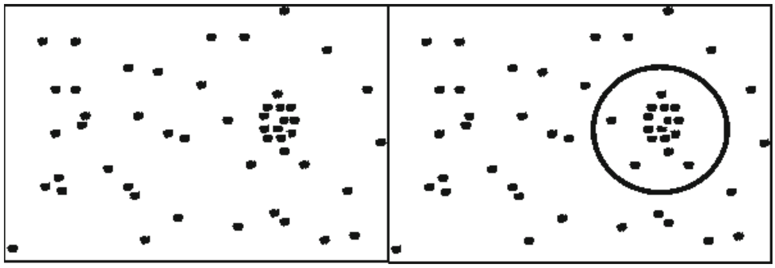

Fig. 1.1 The Texas sharpshooter fallacy

Fig. 1.2 Variables included in an exploratory dataset based on 95 patients with chronic coronary artery disease

- Age at study entry
- Incidence of sudden death
- Hair color
- Angina severity
- MI history
- Number of coronary vessels diseased
- Systolic blood pressure
- Income
- Sudden death
- Marital status
- Preop Ischemia severity
- CABG benefit
- Body mass index
- Family history of MI
- History of hypertension
- Height
- History of PTCA
- Beta blocker use
- Gender
- College education
- Zip code
- Number of children
- Exercise capacity
- Heart rate
- Eye color
- Smoking history
- Ethnicity
- IQ score
- ETOH consumption
- LV mass

target practice, he would again be able to get that many bullets in the circle? Note: the Texan defined his target only after he saw his results. He also ignored the bullets that were not in the cluster! This parable illustrates what epidemiologists call the "Texas Sharpshooter Fallacy" [16] to underscore the dangers of forming causal conclusions about cases of disease that happen to cluster in a population due to chance alone or to reasons other than the chosen cause. As per Atul Gawande, in his classic article in *The New Yorker*, of the myriad of cancer "clusters" studied by scientists in the United States, "not one has convincingly identified an underlying environmental cause" [17]. In a more general sense (and particularly germane to the activities of some biomedical researchers), the Texas Sharpshooter Fallacy is related to the "clustering illusion," which refers to the tendency of individuals to interpret patterns in randomness when none actually exists, often due to an underlying cognitive bias.

Consider a more clinical example: A resident inherits a dataset that contains information about 95 patients with chronic coronary artery disease. Figure 1.2 depicts the variables in that dataset.

He believes that he could satisfy his research elective if he could draw inferences about this study group, though he has no a priori idea about what relationships would be most reasonable to explore. He recruits a friend who happens to have a statistical package installed on his computer, enters all of the variables in the dataset into a multiple regression model, and comes up with some statistically significant findings, as noted below:

- Ischemia severity and benefit of coronary artery bypass grafting (CABG): $p<0.001$
- Hair color and severity of myocardial infarction (MI): $p<0.03$
- Zip code and height: $p<0.04$

He concludes that he has *confirmed the hypothesis* that there is a strong association between preoperative ischemia severity and benefit of coronary artery bypass grafting because not only was the obtained probability (p) value low, his hypothesis also makes clinical sense. He also decides that he would not report the other findings because, while also statistically significant, he cannot explain them. What methodological error has the resident made in drawing his conclusion?

The answer is that, analogous to the rifleman who defined his target only after the fact, the resident "confirmed" a hypothesis that did not exist before he examined patterns in his data. The fallacy would not have occurred if the resident had, in mind, a prior expectation of a particular association. It also would not have occurred had the resident used the data to generate a hypothesis and validated it, as he should have, with an independent group of observations if he wanted to draw such a definitive conclusion. This is an important distinction because the identification of an association between two or more variables may be the result of a chance difference in the distribution of these variables—and hypotheses identified this way are suggestive at best, not proven. What one cannot do is to use the same data to generate and test a hypothesis.

Moreover, the resident compounded his error by capitalizing on only one association that he found, ignoring all of the others. Working with hypotheses is like playing a game of cards. You cannot make up rules after seeing your hand, or change the rules midstream if you do not like the hand that you have been dealt. Similarly, if you gather your data first and draw conclusions based only on those you believe to be true, you have, in the words of the famed behavioral scientist, Fred Kerlinger, violated the rules of the "scientific game" [18]. The most important take-home point is *if you wish to test it*, a hypothesis always should be generated before data collection begins.

Hypothesis-testing studies (especially randomized clinical trials [RCTs]) are highly regarded in medicine because, when based on correct premises, properly designed, and adequately powered, they are likely to yield accurate conclusions [19]; in contrast, conclusions drawn from hypothesis-generating studies, even when well designed, are more tentative than those of hypothesis-testing studies due to the myriad of explanations (hypotheses) one can infer from the observation of a phenomenon.

For these reasons, hypothesis-generating studies are appropriately regarded as exploratory in nature. These differences notwithstanding, there is general consensus that hypothesis-testing and hypothesis-generating activities *both* are vital aspects of the research process. Indeed, the latter are the crucial initial steps for making discoveries in medicine. As Andersen [20] has correctly argued, without hypothesis-generating activities, there would be no hypotheses to test and the body of theory and knowledge would stagnate. The critical role of the hypothesis in the research process and the logical issues entailed in formulating and testing them are further discussed in Chap. 3.

Retrospective Versus Prospective Research

Research often is classified as retrospective or prospective. However, as pointed out by Catherine DeAngelis, former editor-in-chief of the Journal of the American Medical Association (JAMA), these terms "are among the most frequently misunderstood in research" [21] in part because they are used in different ways by different workers in the field and because some forms of research do not neatly fall within this dichotomy. Many methodologists [22, 23] consider research to be retrospective when data (typically recorded for purposes other than research) are generated prior to initiation of the study and to be prospective when data are collected starting with or subsequent to initiation of the study. Others, including