0363-5465/96/2424-0393$02.00/0
THE AMERICAN JOURNAL OF SPORTS MEDICINE, Vol. 24, No. 3


Current Concepts

# A Statistics Primer

Mary Lou V. H. Greenfield,* MPH, John E. Kuhn, MS, MD, and Edward M. Wojtys, MD

*From MedSport and the Section of Orthopaedic Surgery, University of Michigan, Ann Arbor, Michigan*

Modern societies survive on data. Businesses collect data on potential clients, consumers need data to compare products, and voters require information on economic trends to help with decision-making at election time. Sports fans everywhere are masters of data consumption: batting averages, passing efficiency, shooting percentages, and other facts. Not surprisingly, the scientific community requires even more data. Unfortunately, data are only useful when the masses of numbers that we collect can be put into usable order. Fortunately, this ordering of masses of numbers can be accomplished by an understanding of statistics.

Medical practitioners require a basic understanding of statistical principles (but not necessarily statistical formulas) because we are consumers of medical research. To make sound clinical decisions, we must be able to discern good studies from bad, to verify whether conclusions of a particular study are valid, and to understand the limitations of a study.[1] Although we need to know the fundamentals of statistics to understand medical research and to draw intelligent conclusions from what we read, we must beware, however, of Disraeli's dictum: "There are three kinds of lies: lies, damned lies, and statistics." We must be ever vigilant in our review of studies that employ statistics.

What is statistics? Mendenhall and Beaver,[7] in the classic text *Introduction to Probability and Statistics*, defines statistics as an area of science concerned with extraction of information from numerical data and its use in making inferences about a population from which data are obtained. Because the word "statistics" may evoke images of equations and formulas that can be intimidating to the uninitiated, many of us shy away from statistics. Its study is seen as too mathematical, mechanical, and beyond comprehension. So why must we use statistics? Because in health research, in particular, we need to make measurements and observations on subsets of the individuals for whom we would like to draw some conclusions or make estimations.[4]

Specifically, an understanding of statistics allows us to do several important things: 1) to generalize about a population from information about a sample of that population knowing the certainty of such judgments; 2) to reduce research data to meaningful indices for quick reference and comparison (for example, normal laboratory test values); 3) to judge the relationship between two or more variables or how well one can predict the value of one variable by knowing the value of another (for example, osteoporosis and hip fracture); and 4) to determine whether a set of research findings may be due to chance.[8]

The purpose of this series of articles is to provide a basic understanding of statistical terminology and principles so that the reader may become a better consumer of medical research. The goal is not to turn clinicians into statisticians. We are not statisticians; we are physicians and professional researchers whose work has forced us to become modestly proficient at understanding the underlying principles of the conduct of clinical trials and basic science studies. Our statistical knowledge should allow us to read the literature and critique presentations comfortably and fairly. We should also know enough to seek statistical assistance when study review or design requires more sophistication than we possess in this area.

## STATISTICAL TERMINOLOGY AND FACTORS THAT INFLUENCE STUDY CONCLUSIONS

### Part 1

Consider the following anecdote taken from the book *How to Lie with Statistics* by Huff.[5]

> Each of a dozen different independent studies showed that a considerable percentage of colds cleared up after treatment with antihistamines. A great fuss ensued, at least in the advertisements, and a medical product boom was on. It was based on an eternally springing hope and also on a curious refusal to look at a fact that

*Address correspondence and reprint requests to Mary Lou V. H. Greenfield, MPH, University of Michigan, Orthopaedic Surgery, TC2914G-0328, 1500 East Medical Center Drive, Ann Arbor, MI 48109.

No author or related institution has received any financial benefit from research in this study.

has been known for a long time. Henry G. Felsen, a humorist and no medical authority, pointed out some time ago, proper treatment will cure a cold in 7 days, but left to itself a cold will hang on for a week.

This is true for much of what we read and hear. Averages and relationships and trends and graphs are not always what they seem. There may be more in them than meets the eye, and there may be a good deal less.

To evaluate a study, we must understand statistical terminology. This will help in the recognition of factors that may influence study conclusions.

The term *population* has a broader meaning than a geographical area, although a geographical region is a perfectly reasonable statistical population. In statistics, however, it refers to a collection of all possible measurements (or data) that could be used to address a study question.[4] This collection of measurements may be from groups of people, institutions, events, or other subjects in a study.[9] There is the assumption that all members of the population have at least one characteristic in common.[2] For example, all women with tennis elbow or all men with beards or all motorcycle owners or all men with femur fractures are populations. Note that none of these have characteristics relating to residence or geographic location.

Populations may be large (all persons within the contiguous U.S.) or they may be small (all persons in Michigan over the age of 100). However, statistical procedures typically assume very large populations, which may present some problems. In scientific research, it is almost impossible to study entire populations. For example, it would be cost-prohibitive and logistically unfeasible (not to say impossible) to contact all Michigan children under the age of 17 at the time of their upper or lower extremity fractures to study psychosocial characteristics of these children. A subset of this population can be selected to study this variable.[5] This is called a *sample*. A sample is a group of subjects selected from the larger group in hopes that studying this smaller group (the sample) will reveal important things about the larger group (the population).[9]

It is important that all members in the sample drawn from the population have an equal chance of being selected for the sample.[2] This is the definition of a *random sample*. Random sampling is very important in statistics. The underlying principle is that only chance determines whether particular data in the population will be part of the sample.[4] Without this element of equal chance in the selection process for the sample, bias is introduced into the study.

*Bias* is defined as a "preconceived personal preference or inclination that influences the way in which a measurement, analysis, assessment, or procedure is reported or performed."[6] Two types of bias to be considered are selection bias and observation or information bias.

*Selection bias* refers to errors that occur when study populations are identified.[3] For example, patients hospitalized for hip fractures may be tested for shoulder strength. This study group would be biased because as a group it probably represents an inherently weaker group of study subjects than that of the population at large because this sample group has been hospitalized. Another type of selection bias occurs when physicians recruit patients themselves, and consciously (or unconsciously) exclude patients who might not be ideal candidates for followup. In either case, the results from these study samples may not reflect the population at large because the study sample is not representative.

*Observation or information bias* refers to bias that occurs when the outcome of interest is affected by characteristics of the study group itself. *Recall bias* is an example of observation bias in which the study participant who has experienced a shoulder injury, for example, may tend to think a lot about possible events leading up to this injury. This study participant may remember past events differently from study participants who have not had a shoulder injury.

Patients who are *lost to followup* also represent a potential for bias of study outcome; unfortunately, patients who are lost to followup are frequently not reported in some studies. When patients lost to followup are different in some characteristic than patients who remain in the study, there should always be concerns about study conclusions. For example, did patients who dropped out of the study do so because they were healthier (or sicker)?

*Confounding* is another concept that readers should be aware of when evaluating study results. For example, when reviewing a study of ankle injuries and playing tennis, possible confounders to a relationship between time and injury might be playing surface, type of footwear, previous ankle injuries, age, and sex, to name only a few. All of these factors can interfere with (confound) the investigator's ability to draw valid conclusions regarding ankle injuries and time spent playing tennis; that is, playing time may have nothing to do with ankle injuries whatsoever. Consider potential confounders when examining study results.

There are ways in which studies can minimize the effects of bias and confounding to produce more generalizable results. This information should be found in the "Materials and Methods" section of a paper. *Randomization* virtually assures that all confounding factors will be evenly distributed among the study groups. Randomization simply means that treatments are selected by a random process "and neither the patients nor the persons responsible for selection or treatment can influence the assignment. Assignment to a treatment remains unknown to both the patient and the staff until the patients have been determined to be eligible for enrollment into the study."[6]

Bias can be minimized in a study by *masking*. This is a condition whereby an individual *or* group of individuals is kept from learning or knowing some fact or observation regarding the study, for example, treatment assignment. (The term "blinding" is no longer used so as to avoid confusion in clinical trials where loss of vision is an outcome measure.)

Another means for reducing study bias is rigorous standardized training of study personnel, reproducibility testing, and clearly written protocols. Look for training protocols in the "Materials and Methods" section of research studies. Have well-standardized test instruments been used? Have interrater reliability tests—those that check the tendency for raters to agree among themselves when using measurement tools—been included? Have intraobserver reliability tests—those that demonstrate that one rater agrees with himself or herself in repeating the same measurements—been conducted?

*Matching* is a strategy that may control confounding variables and should be mentioned in the "Materials and Methods" section of a paper. For example, in a study of the relationship between ankle injuries and playing time, tennis players who have had no ankle injuries could be matched with patients who have ankle injuries. Both participants could be matched for age and sex, thereby controlling for the effects of these two confounders.

Strategies to minimize bias and confounding in studies should be addressed in the "Materials and Methods" sections of papers and presentations. If they are not there, check the "Discussion" section for a review or acknowledgment of potential limitations in study design and, hence, study conclusions. After your own review of the study, draw your own conclusions; then try implementing a study conclusion into your methods of clinical practice strategies, keeping in mind the limitations of the study. Remember that the sample must reflect the clinical group in question to warrant changes in your clinical practice based on those particular outcomes.

## ACKNOWLEDGMENT

The authors thank Dr. M. A. Schork, from the University of Michigan, School of Public Health, Department of Biostatistics, for his review of this manuscript.

## REFERENCES

1. Campbell M, Machin D: *Medical Statistics.* Chichester, England, John Wiley & Sons, 1990, p 1
2. Chase CI: *Elementary Statitistical Procedures.* New York, McGraw-Hill Book Company, 1984, pp 2–3
3. Hennekens CH, Buring JE (Mayrene SL, ed): *Epidemiology in Medicine.* Boston, Little, Brown, 1987, p 273
4. Hirsch RP, Riegelman RK: *Statistical First Aid: Interpretation of Health Care Data.* Boston, Blackwell Scientific Publications, 1992, pp 21–22
5. Huff D: *How To Lie With Statistics.* New York, Norton & Co, 1954, p 8
6. Meinert CL: *Clinical Trials.* New York, Oxford University Press, 1986, pp 283–300
7. Mendenhall W, Beaver RJ: *Introduction to Probability and Statistics.* Boston, PWS-Kent Publishing Company, 1991, pp 5–7
8. Pilcher DM: *Data Analysis for the Helping Professions.* Newbury Park, CA, Sage Publications, 1990, p 15
9. Vogt WP: *Dictionary of Statistical Methodology.* Newbury Park, CA, Sage Publications, 1993, pp 9, 174