Introduction to Biostatistics: Part 1, Basic Concepts

Statistical methods commonly used to analyze data presented in journal articles should be understood by both medical scientists and practicing clinicians. Inappropriate data analysis methods have been reported in 42% to 78% of original publications in critical reviews of selected medical journals. The only way to halt researchers' misuse of statistics and improve the clinician's knowledge of statistics is through education. This is the first of a six-part series of articles intended to provide the reader with a basic, yet fundamental knowledge of common biomedical statistical methods. The series will cover basic concepts of statistical analysis, descriptive statistics, statistical inference theory, comparison of means, χ^2, and correlational and regression techniques. A conceptual explanation will accompany discussion of the appropriate use of these techniques. [Gaddis ML, Gaddis GM: Introduction to biostatistics: Part 1, basic concepts. Ann Emerg Med January 1990;19:86-89.]

Monica L Gaddis, PhD
Gary M Gaddis, MD, PhD
Kansas City, Missouri

From the Departments of Surgery and Emergency Health Services, Truman Medical Center, University of Missouri, Kansas City.

Received for publication September 1, 1989. Accepted for publication October 4, 1989.

Address for reprints: Monica L Gaddis, PhD, Department of Surgery, Truman Medical Center, 2301 Holmes, Kansas City, Missouri 64108.

INTRODUCTION

The practicing physician must remain abreast of new information and techniques and meet standards for continuing medical education. A major source of new information is the medical journal. Given the competition involved in achieving publication of research, as well as the review and editing process, the reader might justifiably assume that published studies are correct in their methodology, regardless of the conclusions made. Although this assumption seems reasonable, it is not correct. Between 1979 and 1984, 42% to 78% of original publications from selected medical journals used inappropriate statistical analysis methods.[1-5]

Statistical analysis involves organization and mathematical manipulation of data. This process can describe characteristics studied and help to infer conclusions from the data, thus guiding the acceptance or rejection of a given treatment or theory. Incorrect data analysis is a grave error in the research process, often leading to inappropriate conclusions, continued study of erroneous hypotheses, and curtailed study of viable therapies and therapeutic adjuncts. Additional dangers ensue when a physician uses nonefficacious treatment on a patient. From this, it becomes apparent that a basic knowledge of statistics can be an important tool for any clinician, whether in performing research or simply reading about it.

However, statistical analysis of data is a task commonly delegated to statistical consultants. This is often justified by citing that there are those more qualified to perform this function than the principal investigator of a study. Though this seems logical, the principal investigator remains the individual ultimately responsible for the content and conclusions of a research project. The principal investigator cannot effectively meet this obligation fully if he does not have an adequate working knowledge of biomedical statistics. The investigator cannot plead innocence through ignorance when serious errors are made.

Thus, there is an obvious need for statistical education among clinicians, not only to provide for a better understanding when reading the biomedical literature but also to aid medical researchers in communication with consulting statisticians and for selection of appropriate data analysis techniques.

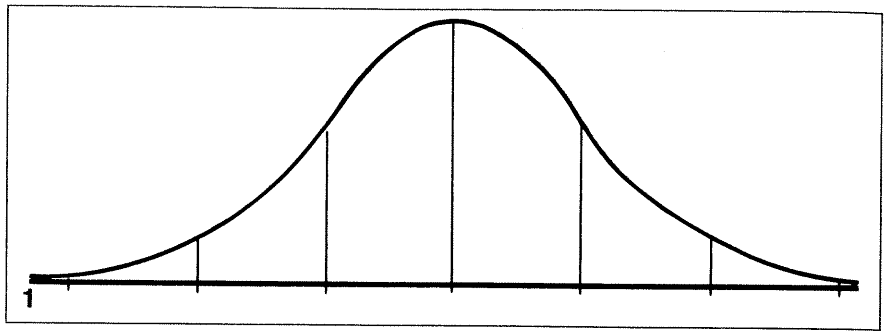

FIGURE 1. *A bell-shaped or normal distribution curve.*

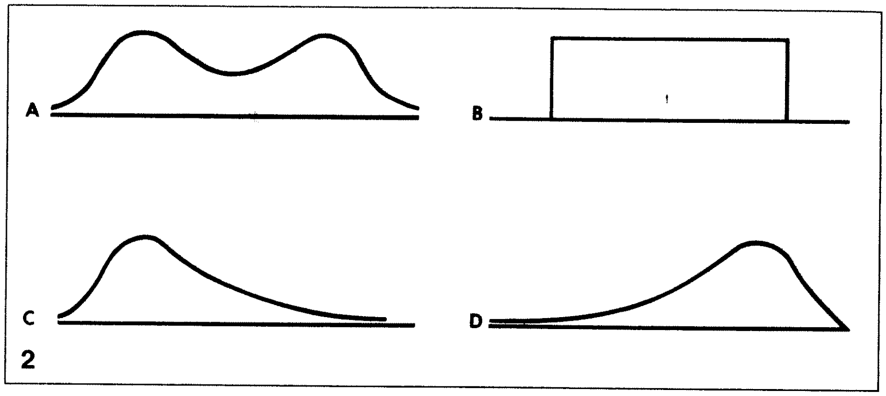

FIGURE 2. *Frequency distributions:* A, *bimodal;* B, *rectangular;* C, *positively skewed;* D, *negatively skewed. (Hopkins KD, Glass GV:* Basic Statistics for the Behavioral Sciences. *Englewood Cliffs, New Jersey, Prentice-Hall Inc, 1978, p 36.)*

STATEMENT OF PURPOSE

It is our purpose to present a six-part series discussing the basics of biomedical statistics, with the intent to familiarize the reader with the terminology and appropriate use of data analysis techniques commonly used in original papers published in *Annals of Emergency Medicine* and other clinical medical journals. Learning will be directed toward a conceptual understanding of statistical analysis methods rather than computational exercises.

Part 1, in this issue of *Annals*, presents some basic concepts of statistical analysis. Knowledge of these building blocks is necessary for a complete understanding of biomedical statistics and the topics of future articles in this series.

Part 2 will address descriptive statistics. These include measures of central tendency (mean, median, and mode), measures of variability (standard deviation and standard error of the mean), and confidence intervals. Appropriate uses of these will be discussed.

Part 3 will introduce statistical inference theory. An explanation of the concept of hypothesis testing, definition of the probability value *P*, a discussion of the terms alpha, beta, and power, and a discussion of clinical versus statistical significance will be included. Sensitivity, specificity, and predictive value will also be addressed.

Part 4 will present parametric and nonparametric methods used for the comparison of means. Included will be analysis of variance (ANOVA), the Student's *t* test, the Mann-Whitney *U* test, and methods of multiple comparisons.

Part 5 will present a discussion of χ^2 and Fisher's exact tests. Both statistical methods should be understood by readers of biomedical research involving a study of the efficacy of experimental medical treatments.

Part 6 will give a basic discussion of correlation and regression, along with the series' concluding remarks.

SAMPLE VERSUS POPULATION

It is clearly impossible for most scientific studies to survey all individuals of a group about which conclusions are to be drawn. Cost and time considerations force the researcher to settle for studying a subset of a group in order to form conclusions about the entire group. For example, if one wished to know the systolic, diastolic, and mean blood pressures of the entire population of the United States, this testing could be done as part of the next census, in which all members of the population would be surveyed. However, the cost of training census takers for this time-consuming task and the need to hire additional census takers because of the increased time required to check each individual would make sampling the entire population an impractical task. Similarly, in biomedical research, time and financial costs preclude study of every member of the population.

However, if a representative sample of an appropriate population can be obtained and studied, conclusions about the sample can be properly extrapolated to the defined population. The key to obtaining a representative sample of that population lies in random selection of a study sample from the applicable population. Randomization can be accomplished by sample subject selection using a random number table, drawing numbers or names out of a hat, or the like. Random sampling implies that all individuals in a population have an equal chance to be included in the sample.

It is when random allocation of study subjects to treatment groups is violated that bias is introduced. Bias can easily lead to erroneous conclusions because control and treatment groups may inherently differ in relevant characteristics before the study is initiated. Therefore, post-treatment differences between groups can erroneously be ascribed to an effect of the experimental treatment! Improper allocation of subjects to control and treatment groups remains a significant problem confounding current biomedical research.

DATA SCALES

Before an analysis method can be selected, the type of data that will be generated by the research process must be defined. Data will fit a nom-

Statistical analysis	The organization and mathematical manipulation of data, used to describe characteristics studied and/or to help infer conclusions from the data
Population	A large group possessing a given characteristic or set of characteristics. A population may be finite (the states of the United States) or infinite (blood pressure measurements of all infants born in New York)[6]
Sample	The studied subset of members of a defined population[6]
Random sample	The process of selection of a sample from a population whereby each member of the population has an equal and independent chance of being chosen[6]
Nominal scale	Numbers are arbitrarily assigned to characteristics for data classification
Ordinal scale	Numbers are used to denote rank-order, without defining a magnitude of difference between numbers
Interval scale	Numbers denote units of equal magnitude as well as rank order on a scale *without* an absolute zero
Ratio scale	Numbers denote units of equal magnitude as well as rank order on a scale *with* an absolute zero
Distribution	The systematic organization of a collection of data
Normal distribution	A grouping of data that is graphically symmetrical and bell shaped. Many human anatomic and physiologic characteristics are normally distributed
Parametric methods	Used when the data studied are from a sample or population that is normally distributed. Data must be of an interval or ratio scale
Nonparametric methods	Used when the data studied are from a sample or population that deviates from a normal distribution. Ordinal data are analyzed using nonparametric methods of analysis

3

FIGURE 3. *Summary of biostatistical terms.*

inal, ordinal, interval, or ratio scale.

Nominal Scale

Nominal is the most primitive of the data scales. Information classified by an assigned number or code to make the data numeric fits a nominal scale. For instance, gender may be defined as 1 for "female" and 2 for "male." Clinical diagnoses may also be assigned representative numbers, such as 1 for "renal failure," 2 for "congestive heart failure," 3 for "diabetes," and so forth. It must be emphasized that the numbers selected are purely arbitrary and are chosen at the discretion of the researcher without regard to any order of ranking of severity.[6,7]

Ordinal Scale

Ordinal scale data can be ranked in a specific order, be it low to high or high to low. An example would be data from a questionnaire in which a response of "strongly agree" is scored as 5, "agree" is scored 4, "no opinion" is scored 3, "disagree" is scored 2, and "strongly disagree" is scored 1. In this example, the responses are scored on a continuum, without a consistent level of magnitude of differences between ranks. However, unlike the nominal scale, numbers of an ordinal scale are not arbitrary because the order of numbers is meaningful.[6] For instance, in the above example, progressively larger numbers indicate progressively greater agreement with the question presented. The Glasgow Coma Score[8] allots a progressively decreased classification number that implies progressively worsened obtundation. Other examples of ordinal scales familiar to many emergency physicians would include the Trauma Score[9] and the Injury Severity Score.[10]

Average values of ordinal scale data are often calculated but are usually misleading because of the lack of any consistent magnitude of difference between units of the scale.[11] Also, it is not uncommon in trauma outcome studies to see such data reported with standard deviation values. Calculation of such statistics from nonparametric, and thus nonnormally distributed data, is highly questionable because use of the standard deviation assumes that the data are normally distributed.[11]

Interval Scale

Interval scale data are a step more sophisticated than ordinal scale data. Not only is there a predetermined order to the numbering of the scale, but also there is a consistent level of magnitude of difference described between the observed data units.[6] Interval data also have a clearly defined unit of measure. However, "the zero point on the scale is arbitrary, and does not correspond to a total absence of the characteristic measured . . . ."[6] The Farenheit scale for temperature is interval in nature, as the numbering of the scale is consistent, yet the zero value is arbitrary. Because of the consistent numbering of the scale and equal magnitude between measurement units, average values and measures of variability of interval scale data are meaningful. An interval scale can be converted to an ordinal scale to show ranking, but an ordinal scale ordinarily cannot be converted to an interval scale.[6]

Ratio Scale

The ratio scale is simply an interval scale with an absolute zero.[6] A predetermined order to the numbering of the scale is present, as is a consistent level of magnitude between each unit of measure. Ratio data can be converted to ordinal data. Heart rate, blood pressure, distance, time, and degrees Kelvin represent examples of ratio scale data.

DISTRIBUTIONS

Once data are collected, they can be organized into a distribution, or graph of frequency of occurrence. This is a visually descriptive tool that allows the researcher to begin to define and analyze data.

Figure 1 is a theoretical frequency distribution of resting heart rate of emergency physicians. Its shape is symmetrical and bell-shaped, and is defined as the "normal distribution." Many human anatomic and physiologic characteristics approach the normal distribution.[6] Other terminology used synonymously with "normal distribution" includes Gaussian curve, curve of error, and normal probability curve. An understanding of the characteristics of the normal distribution is fundamental in the development of even a basic knowledge of biostatistics. This topic will be discussed in greater detail in Part 2 of this series.

Other distributions are depicted in Figure 2.[6] Bimodal distributions have two peaks of cluster, or areas with a high frequency level. For example, if the weights of American adults were plotted, there would be two definitive points of cluster, one for female weight and one for male weight.

Data that are rectangularly distributed show equal frequency of occurrence for all levels of a characteristic. The date of birth (month and day) of a sample of all patients seen in an emergency department approaches this distribution pattern.

Skewed data are those that tail off to either the high or low end of measurement units. The annual income of patients seen in the ED of an inner-city community hospital will be defined by a distribution that is positively skewed, showing a high frequency of lower annual incomes. A negatively skewed distribution has a cluster of data on the high end of the unit scale and tails off toward the low end.[6]

Given the nature of data collected in medical science research, the normal distribution is the one with which the clinician will be most familiar. Most statistical methods applied to interval or ratio data assume that the data are normally distributed. It is when this assumption is violated that significant controversy exists in the statistical community regarding the proper application of statistical tests.

PARAMETRIC VERSUS NONPARAMETRIC METHODS

Statistical methods are defined as being parametric or nonparametric. The type of analysis method that should be selected is dependent on the nature of the data to be analyzed. Parametric methods use data extrapolated from a sample of the population studied to numerically describe some characteristic of a population. Parametric methods are valid only when that characteristic follows or nearly follows the normal distribution in the population studied.[12] Thus, parametric methods can be applied properly to most interval and ratio scale data when those data come from a sample of a normally distributed population. Parametric methods can be used to derive measures of central tendency and variability, which will be discussed in Part 2 of the series.

Nonparametric methods are applied to non-normally distributed data and/or data that do not meet the criteria for the use of parametric methods. Data that fit the ordinal scale definition should be analyzed by nonparametric methods. Examples include Glasgow Coma Scale,[8] Trauma Score,[9] the Injury Severity Score,[10] and other similar ordinal scales data.

SUMMARY

This has been the introductory installment of a series of articles outlining the proper use of biostatistics. We have introduced some basic concepts regarding data and its classification, which will be needed for understanding of topics to be presented subsequently (Figure 3). Measures of central tendency, measures of variability, confidence intervals, and the appropriate use of these statistical concepts will be discussed in part 2 of this series.

REFERENCES

1. Glantz SA: Biostatistics: How to detect, correct, and prevent errors in the medical literature. *Circulation* 1980;61:1-7.

2. Felson DT, Cupples LA, Meenan RF: Misuse of statistical methods in arthritis and rheumatism: 1982 versus 1967-68. *Arthritis Rheum* 1984;27:1018-1022.

3. Thorn MD, Pulliam CC, Symons MJ, et al: Statistical and research quality of the medical and pharmacy literature. *Am J Hosp Pharm* 1985;42:1077-1082.

4. Avram MJ, Shanks CA, Dykes MHM, et al: Statistical methods in anesthesia articles: An evaluation of two American journals during two six-month periods. *Anesth Analg* 1985;64:607-611.

5. MacArthur RD, Jackson GG: An evaluation of the use of statistical methodology in the *Journal of Infectious Diseases*. *J Infect Dis* 1984;149:349-354.

6. Hopkins KD, Glass GV: *Basic Statistics for the Behavioral Sciences.* Englewood Cliffs, New Jersey, Prentice-Hall Inc, 1978.

7. Elenbaas RM, Elenbaas JK, Cuddy PG: Evaluating the medical literature. Part II: Statistical analysis. *Ann Emerg Med* 1983;12:610-620.

8. Teasdale G, Jennett B: Assessment of coma and impaired consciousness: A practical scale. *Lancet* 1974;2:81-84.

9. Champion HR, Sacco WJ, Carnazzo AJ, et al: Trauma score. *Crit Care Med* 1981;9:672-676.

10. Baker SP, O'Neill B, Haddon W Jr, et al: The Injury Severity Score: A method for describing patients with multiple injuries and evaluating emergency care. *J Trauma* 1974;14:187-196.

11. Sokal RR, Rohlf FJ: *Biometry*, ed 2. New York, WH Freeman and Co, 1981, pp 39-52.

12. Glantz SA: *Primer of Biostatistics*, ed 2. New York, McGraw-Hill Book Co, 1987.