Why Biostatistics?

Abrol M*

Government P. G. College for Women, India

Submission: November 09, 2017; **Published:** February 09, 2018

***Corresponding author:** Meeru Abrol, Associate Professor, Government P. G. College for Women, Gandhi Nagar, Jammu. 180004, India, Tel: 09419122430; Email: meeru.abrol@yahoo.in

Opinion

The study of Statistics is important as the world is full of variations and sometimes it is hard to tell the real differences from natural variations. Statistics would not be needed if everybody in the world were exactly like everyone else, if there were no differences and we knew the life expectancy or whether or not a new drug was effective in eliminating cancer. But this is not the case. The downside of all the variability is that it makes it more difficult to determine how a person will respond to some new-fangled treatment regimen or react in a situation. There has been an explosive growth in the development of statistical methodology over the past several decades in all the fields whether it be agriculture, medicine or public health. This study is known as Biostatistics. Biostatistics, a branch of applied Statistics forms an integral part and has wide applications in the field of Agriculture, Veterinary Science, Fishery Science, Medicine, Forestry, Pharmacy, Botany, Zoology, Biochemistry, Microbiology, Biotechnology and Bioinformatics.

With or without biostatistical principles, it is human nature to try to convert observations into knowledge. It is also human nature to try to recognize patterns and make sense of observations. So, biostatistics could be defined initially as the descriptive enterprise of observing patterns in biomedical data. Precisely, biostatistics is the discipline concerned with how we ought to make decisions when analyzing biomedical data. It is the evolving discipline concerned with formulating explicit rules to compensate both for the fallibility of human intuition in general and for biases in study design in particular.

To see why we need such a discipline, consider the example of "What is the probability that both children [in a family] are boys given that we know at least one of them is a boy?" If the probabilities of a boy and of a girl are 50% each, and if the genders of the two children are independent of each other, then one could argue that the first is a boy and the second has a 50% chance either way, so the answer is 50%. This answer confuses the question asked with another similar question, specifically, "What is the probability that both children are boys given that we know that the first one is a boy?" The second question excludes two of the *a priori* possible arrangements, GB and GG, leaving only BB and BG, each equally likely. So, 50% is the answer. Our original question did not exclude GB, however, and rather excluded only GG. The remaining possibilities are BB, BG, and GG, all equally likely, and the answer is 33%. Even this answer is not necessarily correct, however, because it is still based on two assumptions, specifically, that boys and girls are equally likely and that the genders of consecutive children are independent. This classic example has analogue in biostatistics.

Many researchers view biostatistics as merely an annoyance, only for satisfying journal editors, regulatory authorities, or funding agencies. But the necessity of biostatistics is apparent when considering the relationship between our intuitions and formal analytical methods. The implication is that few of us have sufficiently "exact minds" to always intuit the correct answer-which is where analytical decision tools come into play. Given this, is there a sense in which intuition can combine with biostatistical methodology to improve understanding? The answer depends on the context in which the question is asked. As Reichenbach noted (1938), there is a "context of discovery" and a "context of justification." That is, when a researcher is formulating a hypothesis (the "context of discovery"), a researcher can draw on any number of insights-intuition, the opinion of esteemed colleagues, peer-reviewed journal articles etc.-in formulating the hypothesis to be tested. However, when the researcher must justify the conclusion reached while testing that hypothesis (the "context of justification"), the appropriate biostatistical methodology rather than intuition must be relied on. In such a situation, not using proper biostatistical principles is rushing to obtain answers to important questions without bothering to ensure the integrity or accuracy of these answers.

This would be akin to shunning a map on the basis of being a driver, and not a navigator. Yes, you will still arrive at a destination, but that destination may not be the one for which you set out. When driving to the wrong location, this fact becomes immediately apparent upon arriving there. In medical studies, the ability to check obtained results against the truth is never available, so the only recourse is to become the best navigator

possible. Only by navigating towards the truth can we hope to find the truth. Because our intuition often can be wrong even when steps are taken to address this very problem, this lesson has to be learned anew by each new generation of researchers. The process of continually trying to improve on past methods is what "biostatistics" means to us. Biostatistics mainly consists of various steps like generation of hypothesis, collection of data, and application of statistical analysis. Biostatistical methods are necessary to draw valid conclusion from the data. One should be aware of different types of data, measures of central tendencies, and different tests commonly used in biostatistics, so that they would be able to apply these tests and analyze the data. For example, in order to test the efficiency of a new drug or medicine, t - test is used or to compare the efficiency of two drugs or two medicines, t - test for the two samples is used. More and more applications of statistics are at present used in clinical investigation.

Recent developments in the fields of computer technology and information technology have enabled biostatistics to integrate their models and thus make statistics a part of decision making procedures. Bio statistical computations have been made very feasible owing to availability of computers and suitable software programs. Now a days, computers are mostly used for performing various statistical tests as it is very tedious to perform it manually. Commonly used software's are MS Office Excel, Graph Pad Prism, SPSS, NCSS, Instant, Dataplot, Sigmastat, Graph Pad Instat, Sysstat, Genstat, MINITAB, SAS, STATA, and Sigma Graph Pad. Free website for statistical software which is frequently used is: www.statistics.com, http://biostat. mc.vanderbilt.edu/ wiki/ Main/ Power Sample Size. There are also many software packages available for solving design of experiments, forecasting simulation problems etc. SYSTAT, a software package offers mere scientific and technical graphing options than any other desktop statistics package. SYSTAT supports all types of scientific and technical research in various diversified fields. For example, Archaeology (Evolution of skull dimensions); Epidemiology (Tuberculosis) and Medical research (Clinical investigations).

Research in these fields which has increased dramatically during this period, has been both a beneficiary of this methodology as well as a source of new problems. A lot of study has been done on the contributions of statistical methodology in designs, implementation and analysis as they relate to the role that biostatistics and biostatisticians play in the field of medical research. A lot has been said about the acceptance by the medical community that bio statistical concepts are an integral part of sound (medical) research; the sometimes unrealistic expectations placed on biostatistics and biostatisticians given limited resources and/or limited control; some controversies among biostatisticians and the need for emphasizing the design and implementation phases of medical investigations. Bio statistics with all its wide application has its own limitations. It is not suitable to the study of qualitative phenomenon: as it deals with a set of numerical data and is only applicable to the study of those subjects of enquiry, which can be expressed in terms of quantitative measurements. As a matter of fact, qualitative phenomenon like honesty, poverty, beauty, intelligence etc., cannot be expressed numerically and any statistical analysis cannot be directly applied on this qualitative phenomenon. Nevertheless, statistical techniques may be applied indirectly by first reducing the qualitative expressions to accurate quantitative terms. For example, the intelligence of a group of students can be studied on the basis of their marks in a particular examination.

Also it does not study individuals and does not give any specific importance to the individual items, in fact it deals with an aggregate of objects. Individual items, when they are taken individually do not constitute any data and do not serve any purpose for any enquiry. Also it is well known that mathematical and physical sciences are exact. But statistical laws are not exact and statistical laws are only approximations. Statistical conclusions are not universally true. They are true only on an average. Statistics table may be misused. Statistics must be used only by experts; otherwise, statistical methods are the most dangerous tools in the hands of the inexpert. The use of statistical tools by the inexperienced and untrained persons might lead to wrong conclusions. Statistics can be easily misused by quoting wrong figures of data. Statistics is only, one of the methods of studying a problem. Statistical method do not provide complete solution of the problems because problems are to be studied taking the background of the countries culture, philosophy or religion into consideration. Thus the statistical study should be supplemented by other evidences.

Collecting, analyzing, and interpreting data are essential components of biomedical research and require biostatistics. Doing various statistical tests has been made easy by sophisticated computer software. It is important for the investigator and the interpreting clinician to understand the basics of biostatistics for two reasons. The first is to choose the right statistical test for the computer to perform based on the nature of data derived from one's own research. The second is to understand if an analysis was performed appropriately during review and interpretation of others' research.