

Relevance of Biostatistics in Medical Research

Baboo NS* , Sathian B⁺, Jayadevan S[#]

ABSTRACT

Medical scientists should be well informed of the importance biostatistics has in the realms of their research work. Many researchers are not giving due importance to optimum size calculation, confidence intervals and testing of hypothesis while undertaking their research. This negligence results in wrong conclusions and thus reducing the quality of their research. Absence of evidence is not evidence of absence. Biostatistics links the two subjects of mathematics and biology for the better exposition of the facts. Therefore, every life scientist must honestly and seriously follow the tenets of biostatistics and the suggestions of a qualified biostatistician from the stage of conceptualization to the finality and publication of the work. A few suggestions regarding the essential statistical methods to be followed in Medical research are briefly explained in this paper.

IJASM 2011; 55(1): 50-55

KEY WORDS: Sample size, Power, Confidence Interval, Medical Research.

Introduction

The major applications of biostatistics started in the middle of the 17th century in the analysis of vital statistics. After the early developments in vital statistics, the field of genetics was the next area that benefited most from the new statistical ideas emerging in the works of Charles Darwin (1809-1882), Francis Galton (1822-1910), Karl Pearson (1857-1936), and Ronald A. Fisher (1890-1962). Now, the fields of application and areas of concern of biostatistics include, among others, bioassay, demography, epidemiology, clinical trials, surveys of human populations, community diagnosis, bio-mathematical modelling, etc. Findings of good research deserve to be presented well, and a good presentation is as much a part of the research as the painstaking collection and analysis of the data. Critical reviewers of the biomedical literature have consistently found that more than half of the published articles (including scientific articles, published even in the best journals) that used statistical methods contained unacceptable errors [1-8]. Evaluation of statistical methods used in articles published in 3 major Indian journals; Indian Journal of Medical Research (IJMR), Indian

Journal of Medical Science (IJMS) & Indian Journal of Preventive and Social Medicine (IJPS) had found errors of omission and commission in 35%-95% of the articles published in them. Out of these, 78% of the articles had atleast one serious error of methodology, 41% had errors in planning and 49% had errors in data collection procedure [7]. To screen out such errors the respective editors (full/part time) must have sufficient training in research and statistical methodologies and the workers must ensure active involvement of medical statisticians in research planning.

The term "statistics" in this context, has a wider meaning and includes the methodology of research, study design, or epidemiological methodology etc [9-15]. A recent study on the published literature of biomedical journals has shown

* Ex-Sr. Advisor & Professor of Physiology, IAM IAF, Bangalore & Ex-Professor of Physiology, Manipal College of Medical Sciences, Pokhara, Nepal.

+ Lecturer (Biostatistics), Dept. of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal.

Assistant Director (Research), Gulf Medical University, Ajman, United Arab Emirates

that these errors mainly concern the sample size, statistical power, agreement between aim and conclusion, distribution of data, as well as description of location and variability of data [1]. A brief glance through almost any recently published medical journal will show that statistical methods are playing an increasingly visible role in modern medical research. At the very least, most research papers quote at least one 'p-value' to communicate. At the same time, a growing number of papers are now presenting the results of relatively sophisticated, statistical analyses of complex sets of medical data [8]. Referring to the available literature and from the personal experience in this important topic, the authors would like to suggest a few methods for the improvement of various situations in medical research. The authors believe that this brief discourse will be of help to all personnel involved in medical research.

Basic biostatistical concepts applicable to medical research [1-15]

A researcher should be well aware of the concepts of different types of data and variables, two types of errors (type I and type II errors), calculation of sample size, significance level, & confidence interval, and testing of hypothesis & power of this testing.

Data and Variables [9-15]

Researchers aim to find out the relationship between one or more events or characteristics (persons, places or things). The information about the various characteristics collected during a study is known as data. The characteristics which take different values for different situations are known as variables. Variables are characterized as quantitative and qualitative. Quantitative data is classified as discrete and continuous whereas qualitative data are classified as nominal and ordinal. The variables may be either dependent or

independent. Independent variables are presumed to be the causes of changes of other variables those are called dependent variables. Research tries to find out the relation between these two also. The selection of an appropriate statistical treatment depends on the type of the data.

Qualitative Variables

These assume non-numerical values. It is classified as Nominal, Dichotomous and Ordinal. Nominal Variables describe characteristics of people, objects or events into categories. Eg. Colour of the eyes. Dichotomous (binary) variable is a type of nominal variable in which a certain characteristics fall into two groups as either *present / absent, male / female* or *yes / no* responses. Ordinal Variables are characteristics which can be put into ordered categories. Eg. Socio-economic status (*low/ medium/high*).

Quantitative Variables

These are variables which assume numerical values. It is classified as discrete and continuous. Discrete Variables are variables which assume a finite or countable number of possible values. It is usually obtained by counting. E.g. Number of patients. Continuous Variables are variables which assume an infinite number of possible values. It is obtained by measurement. E.g. height, weight, etc.

Testing of hypothesis [9-15]

The hypotheses (two opposing statements about a population) are often conjectures or beliefs (even suspicions) concerning the value of a property of the study population. As a rule of thumb the null hypothesis will usually contain an = sign (same group). The alternate hypothesis will contain <, > or "" signs (different group). These hypotheses are to be tested to find out which of the statements is true or not. Statistical tests that are used for these

purposes are called 'tests of hypothesis'. These tests deal with the occurrence of an event under study is actual (true) or by chance (false). There are two types of hypothesis known as 'Null hypothesis' and 'Alternative hypothesis'

Null hypothesis (H_0): Assuming no difference, no change or no effect between the two opposing statements about a population parameter. (e.g. The standard and the new treatment for a disease have the same effect)

Alternative hypothesis (H_a): Assuming there is difference, change or effect between the two opposing statements about a population parameter. (e.g. The standard and the new treatment for a disease have different effect).

The researcher cannot decrease Type I and Type II errors simultaneously. So, one should fix the serious error and try to reduce the other. Type I error is the most serious error because in reality there is no relationship or significant difference but the data is causing the significant relationship or difference. This is just like inference relating coffee drinking to the risk of getting cancer. Really there is no relationship but the data procured in the concerned study (sample size, inclusion and exclusion criteria, sampling, choice of the tests etc.) resulted in this virtual relationship. So this very important error must be fixed to avoid erroneous conclusions. Probability of type I error is known as level of significance (α). For example, if it is fixed at a $p = 0.05$ then only 5% chance of this error is allowed in the study. Once α is fixed, then the 'power of the test' can be estimated. For a statistical test applied in a given situation, the power of the test is the probability of rejecting a false H_0 and will indicate the possibility of the alternative hypothesis (vide supra) to be correct in that situation (population). This enables a worker to detect and test true difference or relationship between a set

of sample data. It is equal to $1-\beta$ where β is the probability of type II error. Power of the test is usually fixed at 0.8 (80%). The researcher should fix these two errors at the formulation stage itself because all subsequent statistical treatments (calculation of sample size, confidence interval and testing of hypothesis etc.) are closely related with it. For example, to calculate the sample size when the significance level is fixed as 0.05 then the confidence interval must be 95% with cut off p value of 0.05 in testing of the hypothesis lest the results and conclusion will be wrong. Purpose of hypothesis testing is to aid the researcher to reach a correct opinion concerning a population by examining a sample from that population. In medical research, statistical inferences are obtained frequently by hypothesis testing (significance tests).

Steps in hypothesis testing [9-15]

1. State the null and alternate hypothesis.
2. Select the appropriate test statistic and calculate the value of the test statistic. (This step is the most important step, here researcher should decide which test formulae is appropriate for the particular hypothesis and calculate the result by substituting values in the appropriate formula using the data obtained from the sample.)
3. Compute the p value (usually from the statistical table, using the value of the test statistic).
4. If the p-value < 0.05 , reject the null hypothesis and accept the alternate hypothesis. If the p-value > 0.05 , accept the null hypothesis and reject the alternate one (Table-1 depicts the course of action after testing the hypothesis).
5. Reach the conclusion (inference) according to the correct hypothesis.

Table 1. Possible course of action after testing the hypothesis

Null hypothesis (H_0)	Course of action	
	Do not reject H_0	Reject H_0
True	Right	Wrong Type I error ($p = \alpha$)
False	Wrong Type II error ($p = \beta$)	Right Power = $1 - \beta$

Sample characteristics

The sample chosen for the study has to be truly representative of the population whose estimates are to be determined, significance tested and conclusions drawn. It should not be biased by being selective, suited to availability or convenience of handling. Its composition should be age, gender, disease specific etc as deemed by the scope of the study. A randomized selection will answer the conditions.

Sample size and significance [9-15]

A suitable sample should be of optimum size. If the sample size is inadequate, then the study will fail to detect a real difference between the effects of two approaches. On the contrary, if, the sample size is too large than what is needed, then the study will become cumbersome, ethically prohibitive, expensive, prolonged and without any added advantage. All the hypothesis testing hinges on the standard error (of the mean) SE (M) which is a measure of the extent to which the sample size deviate from the true population mean. This, in turn, is inversely related to the sample size. So the bigger the sample size, the smaller the P or SE? value, and the greater the likelihood that a small effect will be significant.

Confidence Interval (CI) [9-15]

Generally, confidence limits are equal to the sample mean plus or minus z score obtained from the table (for appropriate level of confidence level) multiplied by the SE. Now, 95% confidence limits (which are the ones conventionally used in medical

research) are approximately equal to sample mean plus or minus two standard errors.

The CI gives a measure of the precision (or uncertainty) of the results for making inferences about the population. CI is the range of values within which we can be 95% sure that the population value (population mean or proportion) lies. In other words, if the study is repeated 100 times in the same population which was sampled; 95 times the means or proportions will lie in the 95% confidence interval calculated from the initial sample. It is a very important concept. If researcher will use confidence interval in his research then the institutional or sample based study will become more applicable to the population. The CI approach places a clear emphasis on quantification, in direct contrast to the p values (vide infra) which arise from the significance testing approach. The p value is not an estimate of any quantity but rather a measure of the strength of evidence against the null hypothesis. The p value by itself tells us neither about the size nor the direction of that difference of the parameters under test. The p values on their own are thus not informative in research conclusions. Now a days CI is getting more superiority comparing with p value.

p value [9-15]

It is the probability of chance modifying the outcome of the research. Chance means the small difference that is not due to any real effect (reason) but because of some inaccurate measurements or inadvertent inclusion of errors in the study. In other words, it is the probability of deciding the null hypothesis as true.

Table-2. The choice of statistical tests according to type of data variables

Explanatory(Independent) variables	Response (Dependent) variables		
	Dichotomous (Binary)	Nominal with more than two categories	Continuous
Dichotomous (Binary)	Chi-square test, Logistic regression and log linear models	Chi-square test, and log linear models	t-tests and Mann Whitney U Test
Nominal with more than two categories	Chi-square test, Polytomous logistic regression and log linear models	Chi-square test, and log linear models	Analysis of variance (ANOVA) and Kruskal Wallis Test
Continuous	Dose response models including logistic regression	*	Correlation and Multiple regression
Mix of continuous and categorical	Logistic regression model	*	Analysis of co-variance and multiple regression

Testing of Hypothesis [9-15]

Testing of hypothesis is divided into two broad types known as parametric and non parametric tests. If the data is following normal distribution [The mean of the study variable is more than double the variance and histogram with frequency curve is a bell shaped curve (normal curve)] then parametric test and otherwise non parametric test is to be used to test the hypothesis. Power of parametric tests is more than that of non parametric tests. Parametric methods tests hypothesis related to population parameters like mean (*t* and *z* tests) or variance (*F* test) whereas non parametric techniques test nominal or ordinal data (*chi-square* test). They do not assume the normal distribution of data. If researcher doesn't understand the distribution pattern of the data, a wrong test is liable to be used for testing the hypothesis and lead to false conclusion of the study. If parametric test is used for non-parametric data then statistical inferences become meaningless. In such situations a significant difference can be identified as non significant and vice versa. Now a days almost all the tests can be performed using

powerful statistical softwares and p-values can be obtained. But one should be careful to choose the appropriate test. Table-2 depicts the choice of statistical tests according to type of data variables.

Conclusion

The importance of biostatistics in this era of Evidenced Based Medicine is emphasized in this article. The common errors can be avoided by understanding the basic statistical concepts by authors, peer reviewers and editors. A medical scientist need to have only a sound basic knowledge of statistics and profound knowledge is needed for the statistician only. The real solution to poor statistical reporting will come when authors learn more about the statistical methods in research design and when statisticians improve their ability to communicate statistics to authors, editors, and readers; when researchers begin to involve statisticians from the beginning of research, not at its end; when manuscript editors begin to understand to apply statistical reporting and editing guidelines; when the journals are able to screen the articles containing statistical analyses more carefully; and

when readers learn more about how to interpret statistics and begin to expect adequate statistical reporting [7-8]. Tables are part of the results intended to convey the magnitude of changes and its statistical significance. These will be complemented by graphs and diagrams. While showing such figures it is essential to choose an appropriate scale so as not to inflate or suppress facts. Researcher should never hesitate for professional assistance of a biostatistician to plan the study or experiment. If the results do not conform to established pattern or trend the researcher should locate the flaw/lacuna in the design or conduct of the study and recast the trials. Always remember the dictum "well-planned is almost half-done".

Acknowledgement

The authors are deeply indebted to Sri. PLN Rao (Ex-Head of Statistics, IAM IAF, Bangalore) for his valuable guidance in the preparation of this article.

Conflict of interest: None

References

1. Zaugg CE. Common biostatistical errors in clinical studies; *Schweiz Rundsch Med Prax* 2003; 92: 218-24
2. Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. *Circulation* 1980; 61: 1-7
3. Mainland D. Chance and the blood count. *CMAJ* 1993; 148: 225-27.
4. Schor S, Karten I. Statistical evaluation of medical journal manuscripts. *JAMA* 1966; 195: 1123-28
5. White SJ. Statistical errors in papers in the *British Journal of Psychiatry*. *Br J Psychiatry* 1979; 135: 336-42
6. *Statistics in Biomedical Journals* [Editorial]. *Ind J Med Res* 1977; 66: 696-703.
7. Hemminki E. Quality of reports of clinical trials submitted by the drug industry to the Finnish and Swedish control authorities. *Eur J Clin Pharmacol* 1981; 19: 157-65.
8. Gore SM, Jones G, Thompson SG. The Lancet's statistical review process: areas for improvement by authors. *Lancet* 1992; 340: 100-2
9. D.E. Matthews and V.T. Farewell. *Using and Understanding Medical Statistics*. 3rd ed. Basel, Switzerland: Karger, 1996.
10. Herbst AL, Ulfelder H, Poskanzer DC. Adenocarcinoma of the vagina: Association of maternal stilbestrol therapy with tumor appearance in young women. *N Engl J Med* 1971; 284: 878-81
11. Hill OW Jr. Rethinking the "significance" of the rejected null hypothesis. *Am Psychol* 1990; 45: 667-8
12. Lipid Research Clinics Program. The Coronary Primary Prevention Trial: Design and implementation. *J Chronic Dis* 1979; 32: 609-31
13. Rubin DH, Krasnikoff PA, Leventhal JM, et al. Effect of passive smoking on birth-weight. *Lancet* 1986; 2: 415-7
14. Shepherd J, Cobbe SM, Ford I, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Engl J Med* 1995; 333: 1301-7
15. Zito RA, Reid PR. Lidocaine kinetics predicted by indocyanine green clearance. *N Engl J Med* 1978; 298: 1160-3.