ejpmr, 2017,4(2), 394-398

SJIF Impact Factor 4.161

EUROPEAN JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.ejpmr.com

Review Article

ISSN 2394-3211

EJPMR

WHEN TO USE A PARTICULAR STATISTICAL TEST IN BIOMEDICAL RESEARCH?: A BRIEF REVIEW

[1]Dr. Abdalwahab MA Zwiri, [2]Dr. Muhammad Nadeem Baig, [3]Dr. Sudhakar Vundavalli, [4]Dr. Radhika Doppalapudi, [5]*Dr. Guntipalli M. Naidu

[1]Assistant Professor, Department of Preventive Dentistry, College of Dentistry, Al Jouf University, KSA.
[2]Assistant Professor, Department of Preventive Dentistry, College of Dentistry, Al Jouf University, KSA.
[3]Lecturer, Department of Preventive Dentistry, College of Dentistry, Al Jouf University, KSA.
[4]Assistant Professor, Dept of Oral Medicine & Radiology, Leonara Institute of Dental Sciences, Rajahmundry, India.
[5]Reader, Department of Public Health Dentistry, Drs. S & NR Siddhartha Institute of Dental Sciences, Gannavaram, Andhra Pradesh, India -521286.

***Corresponding Author: Dr. Guntipalli M. Naidu**
Reader, Department of Public Health Dentistry, Drs. S & NR Siddhartha Institute of Dental Sciences, Gannavaram, Andhra Pradesh, India - 521286.

Article Received on 30/11/2016 **Article Revised on 21/12/2016** **Article Accepted on 11/01/2017**

ABSTRACT
Statistics helps in interpreting results of study correctly and critically evaluating literature. As a researcher, knowing which test of significance is indicated in comparing different variables is sufficient and rest can be done by software. Test of significance is one of the common basic functions in statistics;this review presents various tests of significance in a simple form which should help researchers for quick reference in selecting an appropriate test for their research work.

KEY WORDS: statistics, test of significance, post hoc analysis.

INTRODUCTION
The findings of biomedical research are often met with considerable skepticism, even when they have apparently come from studies with sound methodologies that have been subjected to appropriate statistical analysis. This is may be implied to the finding of research which suggest few aspects of day to day activities are responsible for diseases. One of the popular skepticism about epidemiology made by James Le Fanu, a popular medical journalist in *The Rise and Fall of Modern Medicine*, went a step ahead in suggesting that the possible solution would be for rise in people's health problems would be shutting down of all epidemiology departments[1].

The confusion that exists in today's practice of hypothesis testing dates back to a controversy that raged between the founders of statistical inference more than 60 years ago[2]. The idea of significance testing was introduced by R A Fisher[2]. Researchers often seek Statistician's advice for applying various tests of significance. It is not necessary to memorize the formulas used for various tests of significance; now a day's hardly anyone do manual calculations and most of them use statistical software packages. So, most of the times, just knowing which test of significance is indicated in comparing different variables is sufficient and rest can be done by software.

It may be tedious and time taking for researchers to refer various text books and scientific articles for choosing appropriate test of significance for their research data. Hence the aim of this article is to review various tests of significance in a simple form which would help researchers as a quick reference in selecting an appropriate test for their research work.

UNDERSTANDING VARIOUS TERMS
Before discussing about tests of significance a brief description of commonly used statistical terms:

A] Variable:

A variable is any characteristic of an object that can be measured or categorized[3].

Types:

A. Depending on the data collection

1. Dependent variable: It is the outcome of interest, which would change in response to some intervention.

2. Independent: It is the intervention or possible casual factor or what is being manipulated.

Examples: A study is conducted to find out the relation between smoking and prevalence oral cancer. Here dependent variable (outcome) is oral cancer and independent variable is smoking. In another example, a survey is conducted to assess the prevalence of dental caries in school children, in the report comparisons were made between age Vs DMFT, sex Vs DMFT, etc- here age and sex are independent variables and DMFT is dependent variable.

B. Variables can be classified as either qualitative or quantitative[3].
i) Qualitative variable: It is a characteristic of people or objects that cannot be naturally expressed in a numeric value. Examples of a qualitative variable are sex, hair color, cause of tooth extraction, religion, caste, names of tooth brushes, level of oral hygiene, etc
ii) Quantitative variable: It is a characteristic of people or objects that can be naturally expressed in a numeric value. Examples of quantitative variables are Age, Height, Blood pressure, loss of attachment, serum cholesterol, Hb, etc.

C. Variables are classified into two categories according to the number of different values that they assume:
i). Discrete: A discrete variable is a random variable that can take on a finite number of values or a countably infinite number (as many as there are whole numbers) of values i.e. only whole numbers[3].
Ex: Size of a family, It can be 1,2,3,4, etc, Number of erupted teeth, Number of patients with cancer, etc
ii). Continuous: A continuous variable is a random variable that can take on a
range of values on a continuum; that is, its range is uncountable or infinite. It may take any value, within a defined range and keeps changing[3].

Example for Continuous variables are treatment time, temperature, periodontal pocket depth, amount of new bone growth, diastolic blood pressure, concentration level of anesthesia, etc.

The actual measurements of continuous variables are necessarily discrete due to the limitations in the measuring instrument. For example, the thermometer is calibrated in one degree Fahrenheit, speedometer in 1 mile per hour, and the pocket depth probe in 0.5 mm. As a result, our measurement of continuous variables is always approximate. On the other hand the discrete variables are always measured exactly. The number of amalgam fillings is 4, and the number of patients scheduled for surgery on Monday is 7, but the pocket depth 4.5 mm can be any length between 4.45 mm and 4.55 mm. Many discrete variables can be treated as continuous variables for all practical purposes. The number of colony-forming units (CFUs) in a dental waterline sample may be recorded as 260,000, 260,001, 260,002, where the discrete values approximate the continuous scale[3].

D. In statistics, it is convenient to arrange the data into four mutually exclusive categories according to the type of measurement scale: dichotomous variables, nominal variables, ordinal variables, and continuous variables[4].
1. Dichotomous variables: These are binary variables which have only two categories.
Eg: Normal/abnormal, male/female, treatment/placebo, male/female are few examples for dichotomous unpaired variables. Difference in pain before and after treatment, difference in satisfaction before and after treatment, etc are examples of Dichotomous paired variables.
2. Nominal variable: These are “naming” or categorical variables that have no measurement scales. E.g. blood groups, race, ethnicity, color of skin, tooth paste brands, etc.
Note: Dichotomous variables and nominal variables are sometimes called discrete variables because the different categories are completely separate from each other.
3. Ordinal: Same as nominal + ordered categories.
E.g.: Mild, moderate, severe, very severe, etc. Dean’s fluorosis index, level of satisfaction, etc. None of these variables is measured on a exact measurement scale (like height, weight, etc), but more information is contained in them than in nominal variables.
In ordinal scale the difference (interval) between each variable may or may not be equal i.e. in Dean’s fluorosis index the difference between questionable –very mild is not equal to difference between very mild to mild. If the differences (intervals) are equal such scale is known as interval variable, but it does not have true meaningful Zero[4]. For example IQ score representing the level of intelligence- IQ score 0 is not indicative of no intelligence.
4. Continuous and/ ratio variable: the ratio scale possesses the same properties of interval scale and there exists a true zero[4]. Many types of biological data are measured on continuous measurement scales. Patients’ heights, weights, blood pressure, serum glucose level, loss of periodontal attachment, DMFT score; number of teeth present, etc are the examples of data measured on continuous scale.

For most statistical analyses, including significance testing, the distinction between continuous and ratio variables is usually not important.

WHY TESTS OF SIGNIFICANCE?
As a rule of nature, no two objects or observations are similar; some amount of minute variation will be present[5]. If an investigator conducted a clinical trial about the efficacy of new mouth rinse in reducing the plaque, in which investigator tries to measure and compare before (using mouth rinse) and after plaque scores of study participants.

There might be difference in plaque scores which are measured before and after, here investigator cannot simply come to conclusion that this difference is due to mouth rinse, it can be due to rule of nature(by chance) as mentioned earlier, not a real one due to mouth rinse. The tests of significance answers this uncertainty, by testing what is the probability of this difference in plaque scores being natural phenomenon/by chance? For all tests of significance we get two values- test value and P-value. The test value implies the magnitude of the difference for that particular test (which was applied) and from this test value P-value will be derived. For example for the above mentioned experiment test value is 6.4 and P-value is 0.03, it means the probability of above found difference in plaque scores being a natural phenomenon/by chance is 3%(3 times out of 100 measurements) and 97% of the times it is real effect of mouth rinse. If P-value is 0.001, it means the probability of above found difference in plaque scores being a natural phenomenon/by chance is 0.1% (1 time out of 1000 measurements). So, the above found difference in plaque scores is statistically significant.

Statistical significance is different from clinical significance. For example, a new antihypertensive drug is tested in a group of patients, systolic blood pressure is measured before and after giving the drug and the results conclude that there is 3mm of Hg reduction in systolic blood pressure due to this new drug which is statistically significant(P<0.01). In this experiment, reduction in blood pressure of 3mm of Hg will have very little clinical significance. So, statistical tests help us in answering the variations in observations as they are real or by chance and it is the clinician or researcher who should judge for clinical significance[4].

Not all studies require tests of significance; whether to use statistical tests or not is based on objectives of the research and its methodology. In few studies, the researchers are interested only in descriptive information, such as the sensitivity or specificity of a screening test or laboratory tool, in which case there may be no reason to perform statistical test of significance and in some studies investigators would like to find out facts and figures i.e. Questionnaire based surveys which again does not require tests of significance. Tests of significance are mainly used for inferential statistics.

Table-1: Various Tests of Significance for comparing different variables

First variable	Second variable	Test of choice
Continuous E.g.: Age	Continuous E.g.: Blood pressure	Pearson correlation coefficient; Linear regression
Continuous E.g.: Age	Ordinal data E.g.: Level of satisfaction	Group the continuous variables and use Spearman correlation coefficient; One wayANOVA
Continuous E.g.: Blood sugar level	Dichotomous unpaired E.g: Gender(M/F)	Student's t-test
Continuous E.g.: Difference in body Weight	Dichotomous paired E.g.: Before Vs after treatment.	Paired t-test
Continuous E.g.: Blood sugar level	Nominal E.g: Caste	ANOVA
Ordinal *E.g. : Satisfaction.*	*Ordinal* *E.g: severity of pain*	*Spearman correlation coefficient; Kendall correlation coefficient*
Ordinal *E.g. : Satisfaction*	*Dichotomous unpaired* *E.g.: Gender*	*Mann-Whitney U test*
Ordinal *E.g. : Satisfaction.*	*Dichotomous paired* *E.g: before Vs after treatment*	*Wilcoxon matched-pairs signed-ranks test*
Ordinal *E.g. : Satisfaction*	*Nominal* *E.g.: Religion*	*Kruskal-Wallis test*
Dichotomous E.g: Success/failure	Dichotomous unpaired E.g: Treated/untreated groups	Chi-square test; Fisher exact probability test
Dichotomous E.g.: Change in success/ Failure	Dichotomous paired E.g.: Before Vs after treatment	McNemar Chi-square test
Dichotomous E.g: Success/Failure	Nominal E.g: Blood group	Chi-square test
Nominal E.g.: Religion	Nominal E.g.: Blood group	Chi-square test

Table-2: Multivariable analysis

Dependent variable	Independent variable	Approximate statistical procedure
Continuous	All are categorical	ANOVA
Continuous	Some are categorical, some are continuous	ANCOVA
Continuous	All are continuous	Multiple Linear regression
Ordinal	--	There is no formal multivariable procedure for ordinal dependent variables. Either treat the variables as if they were continuous(perform above

		procedures) or perform log linear analysis.
Dichotomous	All are categorical	Logistic regression, log- linear analysis
Dichotomous	Some are categorical, some are continuous	Logistic regression
Nominal	All are categorical	Log- linear analysis
Nominal	Some are categorical, some are continuous	Group the continuous variables and perform log linear analysis.
Nominal	All are continuous	Discriminant function analysis' group the continuous variables and perform log- linear analysis

Table-3: Post Hoc analysis

Statistic	Properties	Specific indications
Fisher's Least significant difference(LSD)	Too liberal i.e. it is too likely to find a difference (significance); problem of multiplicity; Less advised.	Less advised
Tukey's Honestly significant difference(HSD)	Medium stringent	Can be used for all types of data i.e. mean, range, etc
Tukey's wholly significant difference (WSD)	Medium stringent	Used mainly to compare large number of groups i.e. six or more groups.
Neumann's- Keuls test	Medium stringent	Mainly used for pair wise comparisons
Dunnet's test	Medium stringent	To compare multiple means Vs control mean; Eg; various drugs Vs placebo
Scheffe's method	Too stringent(less likely to get significance)	
All the above mentioned tests are used only when there is homogeneity among groups; if heterogeneity is present the below four tests are used.		
Tomahane's t2	Too stringent	
Games- Howell test	Too liberal	
Dunnet's T3	Medium stringent	
Dunnet"s C	Medium liberal	

Table-4: Efficiency comparison between parametric and non parametric tests

Application	Parametric test	Non parametric test	Efficiency rating of non parametric test for normal distribution
Matched pairs of sample data	t test for dependent samples	Sign test Wilcoxon's matched pairs test	0.63 0.95
Two independent sample	t test for independent samples	Mann-Whitney U test	0.95
Several independent samples	ANOVA	Kruskal- Wallis test	0.91
Correlation	Linear correlation	Rank correlation test	0.91

PARAMETRIC AND NON PARAMETRIC TESTS[4,5]

1. Parametric tests assume some characteristics of data i.e. data is collected from random sampling and normally distributed.

E.g.: Students 't' test, Anova, Pearson's correlation coefficient, Regression, Chi-square test, etc.

2. Non- parametric tests do not assume any characteristics i.e. can be used for nonrandom sampling and skewed data.

E.g.: Kruskal- Wallis test, Mann-Whitney U test, Chi-square test, etc.

BIVARIATE ANALYSIS[4]

Bivariate analysis is most commonly used analysis in which comparison was made between two variables i.e. one dependent variable (possible casual) and one dependent (outcome) variable. Various tests for bivariate analysis was mentioned in table-1.

MULTI VARIABLE ANALYSIS[3,4]

This analysis is required if more than one independent variable is required to compare with single dependent variable and various tests for multivariable analysis was mentioned in table-2.

POST HOC ANALYSIS[7,8]

Post Hoc analysis should be done only if statistically significance is found in pre hoc(e.g. ANOVA), otherwise it is not necessary. Various post hoc tests and their qualities and indications were mentioned in table-4.

IMPORTANCE OF CHOOSING A PROPER TEST OF SIGNIFICANCE

It is very important to choose an appropriate test for data or else it will reduce the efficacy of test. Table-4 depicts the efficiency rating of non parametric test with normal distribution, for example efficiency of Sign test in comparison to't' test will be reduced by nearly 40% if Sign test is used for normal distribution[7].

CONCLUSION

Nature of data mainly determines the choice of appropriate test. Statistical tests in clinical research are used to show that the findings are not likely due to chance. However, it is easy to misinterpret the results of statistical test. Investigators should consider both the clinical and statistical significance of the research findings.

REFERENCES

1. Le Fanu J. The clinical Scientist as an endangered species; Text book on "*The rise and fall of modern medicine"*. New York: Little, Brown,1999; Page: 230
2. Brian G. Leroux Statistical Methods for Studying Associations Between Variables; Textbook on Statistical and Methodological Aspects of Oral Health Research, Edited by E. Lesaffre, J. F eine, B. Leroux, D. Declerck © 2009 John Wiley & Sons, Ltd pg: 183-200.
3. Geoffrey R. Norman, David L. Streiner; Repeated measures Anova; Text book on Biostatistics: The Bare Essentials.3rd edition 2011, People's medical publishing house; page-107-118.
4. James F. Jekel, David L. Katz, Joann G. Elmore. Bivariate analysis; Text book on Epidemiology, Biostatistics and Preventive Medicine;1st edition 2007; page 175-200**.**
5. Jesper W. Schneider; Caveats for using statistical significance tests in research assessments. Available at: http://arxiv.org/ftp/arxiv/papers/1112/1112.2516.pdf [Accessed on: 24th August 2016]
6. When to use a particular statistic? SFU publications.Available at: http://www.sfu.ca/~ber1/iat802/pdfs/When%20to%20use%20what%20test.pdf. [Accessed on: 20th August 2016].
7. Bernard Rosner and Richard Rosner , Non parametric methods ; Text book on Fundamentals of Biostatistics , 6th edition 2005,Thamson & Brooks; page-296-305
8. Anthony Hilton and Richard Armstrong; Post hoc Anova tests; Available: http://eprints.aston.ac.uk/9317/1/Statnote_6.pdf. [accessed on: 6th August 2016]
9. Multiple Comparisons/Post-Hoc Tests. Available at: http://www.gigawiz.com/ simplecomp.html. [Accessed on : 7th August 2016].