EUROPEAN JOURNAL OF BIOMEDICAL AND PHARMACEUTICAL SCIENCES
http://www.ejbps.com

ISSN 2349-8870
Volume: 3
Issue: 1
400-401
Year: 2016

SIGNIFICANCE OF CORRECTION OF P VALUE IN STATISTICAL COMPARISONS AMONG MULTIPLE GROUPS IN BIOLOGICAL RESEARCH

Shashi Sharma[1], Atul Juneja[2], Tulsi Adhikari[2], Apoorva Anand[3]

[1]Institute of Cytology and Preventive Oncology (ICMR), Sector 39 NOIDA UP. 201301
[2]National Institute of Medical Statistics (ICMR) New Delhi-110029
[3]National Institute of Biologicals Sector 62 NOIDA UP-201309

***Author for Correspondence: Dr. Shashi Sharma**
Institute of Cytology and Preventive Oncology (ICMR), Sector 39 NOIDA UP. 201301

Article Received on 18/11/2015 **Article Revised on 10/12/2015** **Article Accepted on 01/01/2016**

ABSTRACT
The statistical comparisons between two groups among multiple groups are a common phenomenon in biomedical research including clinical trials, observational studies and more importantly genetic studies. It is important to have an idea about interpreting the p value or type 1 error in accepting or rejecting the null hypothesis when there are only two groups vis a vis when there are more than two groups and two groups of interest are being compared. Situation is most commonly encountered in genetic studies when multiple genes are being studied. Such situation of comparing sub sets from the multiple groups demands correcting the p value which is technically known as Family Wise Error rate. This communication presents the concept of this correction particularly to the medical or biological researchers when post hoc test are being employed. This would help in the better appreciation of the results through p values with due consideration of statistical and biological/medical and its importance when sub set from the multiple groups are being compared.

KEYWORDS: P value, multiple comparisons, Post Hoc tests family wise correction error rate.

INTRODUCTION
For any research problem including biomedical research, statistics forms an important component in substantiating the findings of the research. The researcher many times finds it an imposed activity. This happens because the genuine need of applying the statistical tools are not appreciated because of the lack of understanding of theoretical concept associated with that statistical tool being used there. Comparing different groups for one or more factors, as per the thrust area of research, forms the goal of the research activity. Obviously the comparability of the groups or controlling of factors is taken into consideration through the design employed in the study.[1] The magnitude of the difference between the groups (two or more) could be measured by applying statistical tests which depends upon the nature of the data. The statistical tests and the number of software are available including online analytical software for carrying out the computations which we do not intend to discuss in this paper.

Need of Correction of P value
The objective of this communication is to address the issue of interpreting the results of comparisons particularly when more than two groups are involved in term of accepting or rejecting the statistical hypothesis. This involves assumptions about type I and type II errors. The type 1 error refers to rejecting a null hypothesis when it is true in other words; it is the measure of chance that the two groups are same and popularly among medical researchers known as 'P' values.[2,3] Conventionally when test statistics value falls in less than 0.05 area of the distribution, the null hypothesis is rejected, which in other words means that the chance that two groups which have been compared by employing statistical tests are same i.e. less than 0.05. It may be reminded that P value will never be zero in view of the nature of the normal distribution.

All these issues of assessing and interpreting the P values are relatively simple when two groups are in consideration for comparison. When there are multiple groups then there is need to correct the P values for interpreting the decision about rejecting the null hypothesis.

In biological situation, the multiple comparison is a very common situation. For instance if there are multiple age groups then the factors under research need to be compared between the groups. Since the groups are more than two, the critical value of 0.05 cannot be considered as magnitude of wrongly rejecting the null hypothesis. If

there are three groups and three possible comparisons such as Av/sB, B v/s C and Av/s C, then a number of statistical tests such as-Tukey Kramer method,-Scheffé's method, Duncan's new multiple range test Dunnett's test Bonferroni–Dunn test etc. can be employed for carrying out multiple statistical comparisons.[4] The intention of this communication is not to discuss the details of the above referred statistical tests but the basic concept behind correcting the p value. The most commonly encountered situation among geneticists is comparing gene expression among many genes between the two groups such as normal and the diseased.[5,6] The comparison of two groups among available groups cannot be considered significant at P=0.05(type 1 error) as in multiple comparison the probability values exceeds the value of type I error which is conventionally fixed at 0.05. It is just like estimating a chance of getting a black ball from a box which has 100 balls with 1 black ball and 99 white balls, as 1/100 in a single attempt. If the numbers of draws are increased, the chance of getting the favourable event (black ball) also increases additively. Hence going by this logic, the probability of considering a difference to be statistically significant at the pre-defined Type I error,it has to be divided by the number of possible group comparisons. This corrected p value is referred as family wise error value where family is nothing but the groups in consideration for evaluation of factor (s). A situation which involves comparison of 100 gene expressions between normal and abnormal tissue, the magnitude of expression would be considered significant only at Type I error/ 100 (no. of genes). This will ensure the false positivity of null hypothesis at the desired fixed level. It may also be mentioned that if the gene expression is large enough to be detected significant at this low level of probability (α /100), which ideally should be $\alpha/{}^{100}C_2$ where ${}^{100}C_2$ is the possible number of groups but this will yield a very small corrected p value, this may also deprive a gene to be detected as significant even if it has biological relevance.[7,8] The results there have to be seen duly considering statistical significance along with clinical and biological relevance. This may not be the scenario in other situation such as clinical trials or epidemiological evaluations where the groups are likely to be around 5 to 6, making the p value for the individual comparison as p/5(0.05/5). This implies that to maintain the magnitude of Type I error at 0.05 in the group, the individual comparison has to be fixed at the p=0.01. Technically this is done by employing Bon Ferroni correction while attempting Post hoc tests.

CONCLUSION

The ultimate message of this communication is to apprise the biomedical researcher relatively new to the research about interpreting the p values or type 1 error (α) in individual groups where groups are more than two. With this knowledge and concept the researcher would be in a better position to appreciate the results in their research for clinical or biological significance. In the absence of the knowledge or facilities to carry out advanced statistical tests in multiple comparisons this basic knowledge would help in correcting the p value if basic chi square test is employed to evaluate one group of larger interest among multiple groups i.e. dividing the fixed type 1 error by the possible number of comparisons or number of groups that are too large and the reasons for doing the same.

REFERENCES

1. Fleiss JL. The design and analysis of clinical experiments. New York: John Wiley & Sons, 1986; 1–32.
2. http://www.statsdirect.com/help/default.htm#basics/p_values.htm
3. Panagiotakos DB. The Value of p-Value in Biomedical Research. The Open Cardiovascular Medicine Journal, 2008; 2: 97-99. doi:10.2174/1874192400802010097.
4. Isella C, Renzulli T, Corà D, Medico E. Mulcom: a multiple comparison statistical test for microarray data in Bioconductor. BMC Bioinformatics, 2011; 12: 382. doi:10.1186/1471-2105-12-382.
5. Kim H-Y. Statistical notes for clinical researchers: post-hoc multiple comparisons. Restorative Dentistry & Endodontics, 2015; 40(2): 172-176. doi:10.5395/rde.2015.40.2.172.
6. Hulshizer R, Blalock EM. Post hoc pattern matching: assigning significance to statistically defined expression patterns in single channel microarray data.BMC Bioinformatics, 2007; 8: 240. doi:10.1186/1471-2105-8-240.
7. Streiner DL, Norman GR.Correction For Multiple TestingChest, 2011; 140(1): 16-18. doi:10.1378/chest.11-0523.
8. Armstrong RA.Ophthalmic Physiol Opt, Sep, 2014; 34(5): 502-8. doi: 10.1111/opo.12131. Epub 2014 Apr 2. Review. PMID: 24697967