

# *Collection of Biostatistics Research Archive*

## COBRA Preprint Series

---

*Year* 2015

*Paper* 117

---

# STATISTICAL HANDLING OF MEDICAL DATA - AN ETHICAL PERSPECTIVE

Ajay Kumar Bansal Dr\*

\*University College of Medical Sciences, University of Delhi, [akchbansal@yahoo.com](mailto:akchbansal@yahoo.com)

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

<http://biostats.bepress.com/cobra/art117>

Copyright ©2015 by the author.

# STATISTICAL HANDLING OF MEDICAL DATA - AN ETHICAL PERSPECTIVE

Ajay Kumar Bansal Dr

## Abstract

Medical Science is a delicate subject and the clinical data generated from the medical trials must be reliable and of good quality. Not only the quality of generated data is important, but the management is also crucial and is to be handled very carefully. In this paper, the ethical aspect of statistical handling of such data is discussed.

Every profession has some set of norms to follow to achieve its objectives. These norms are called professional ethics which shows the essence of human behaviour. Same way, the field of medical research is expected to follow ethical norms, to obtain reliable and true results. Being a biostatistician, I would like to share my experiences to highlight that how the ethical norms are violated especially during data collection and data analysis, while conducting a trial in a health setup. In a follow-up study, correlated data recorded by the postgraduate students, health workers and field workers is highly biased in the desire of establishing the favourable effectiveness of their intervention. It is also seen that they have the tendency of manipulating the data at the time of analysis to have desirable outcomes. They fix in their mind to prove a drug effective, even when it is ineffective. Students are more prone to this practice. Sometimes, even the statistician apart from the students and health professionals, does not have adequate knowledge to apply the right test, which leads to wrong decisions. All these practices are un-ethical and all the members involved at any level in a trial must be made well aware, and must understand that these results are going to be applied on human beings, where one do not have a second chance to try. So, they should stick to the principles of honesty and truthfulness. In a clinical trial, reliability of data and results is very important. Biostatisticians must also understand that instead of applying wrong test, it is better to consult their seniors and should not indulge themselves

in suppressing the true facts. It is their responsibility to translate the statistical interpretation of the results into medical interpretation.

# STATISTICAL HANDLING OF MEDICAL DATA - AN ETHICAL PERSPECTIVE

## Abstract

Medical Science is a delicate subject and the clinical data generated from the medical trials must be reliable and of good quality. Not only the quality of generated data is important, but the management is also crucial and is to be handled very carefully. In this paper, the ethical aspect of statistical handling of such data is discussed.

Every profession has some set of norms to follow to achieve its objectives. These norms are called professional ethics which shows the essence of human behaviour. Same way, the field of medical research is expected to follow ethical norms, to obtain reliable and true results. Being a biostatistician, I would like to share my experiences to highlight that how the ethical norms are violated especially during data collection and data analysis, while conducting a trial in a health setup. In a follow-up study, correlated data recorded by the postgraduate students, health workers and field workers is highly biased in the desire of establishing the favourable effectiveness of their intervention. It is also seen that they have the tendency of manipulating the data at the time of analysis to have desirable outcomes. They fix in their mind to prove a drug effective, even when it is ineffective. Students are more prone to this practice. Sometimes, even the statistician apart from the students and health professionals, does not have adequate knowledge to apply the right test, which leads to wrong decisions. All these practices are unethical and all the members involved at any level in a trial must be made well aware, and must understand that these results are going to be applied on human beings, where one do not have a second chance to try. So, they should stick to the principles of honesty and truthfulness. In a clinical trial, reliability of data and results is very important. Biostatisticians must also understand that instead of applying wrong test, it is better to consult their seniors and should not indulge themselves in suppressing the true facts. It is their responsibility to translate the statistical interpretation of the results into medical interpretation.

**Key Words:** Ethical Handling, Honesty, Truthfulness, Sense of integrity, Medical data

## Introduction

Ethics is neither an oath nor a set of rules to follow but it is the essence of human behaviour which is favourable or not favourable. It gives sense to differentiate between right and wrong.

Statistical Handling of Medical Data.....

It depends on someone's degree of moral development and consciousness. But, many different disciplines, institutions, and professions have norms for behaviour that suit their particular aims and goals. These norms also help members of the discipline to coordinate their actions or activities and to establish the public's trust of the discipline [1]. Same way, the field of medical research is also expected to have a set of norms to obtain the reliable and true results for the betterment of health care. To optimize these results, Ellwood [2] suggested to use the outcome management as a technology of experiences, designed to help patients, payers, and providers.

In medical sciences, the purpose of a clinical trial is to obtain facts that affect the decision of a doctor to select an effective treatment for improving its outcome. According to NIH [3] Clinical trials produce the best data available for health care decision making. The true efficacy of any treatment depends on the battery of determinants but the main are 1) ability of a team to design & administer the trial, 2) proper facilities to conduct the trial, 3) patient's availability as per inclusion criteria & their compliance, 4) fair data collection and 5) proper data handling, analysis & interpretation of results.

Being a biostatistician, I take up the last 2 determinants to share my experiences where the ethical norms are violated while collecting, handling and analysing the data. Collection of data is the responsibility of the team who is administering the trial but the handling, analysis and statistical interpretation translated into medical interpretation is the responsibility of a biostatistician.

### **Data Collection**

I have one very simple example, when a medical researcher is to verify or establish the efficacy of a pain killer on surgically treated patients in an ICU, where he has to take his observations repeatedly after a gap of time on visual analogue scale (VAS) [4]. VAS is a psychometric response scale which can be used in questionnaires to record the pain score. VAS is a 10 point scale scoring system where the observer asks the patient to rate his pain from 1 to 10. 1 is no pain and 10 means severe pain.

After the surgery, when a patient is asked to rate his pain, suppose he says 8, a pain killer whose efficacy is to be tested is administered to relieve the pain. The observer is supposed to ask the patient after every 30 minutes to check whether the pain is relieved or not? After first 30 minutes, the patient says a pain score of say 7, after second 30 minutes, he says it is 5 and after

Statistical Handling of Medical Data.....

third 30 minutes, he says it is 6. Then, it is seen that the observer records either 5 or less than 5 for the 3<sup>rd</sup> reading as he assumes that the drug is to relieve the pain or to decrease the pain score with the increase of time. Now, this reading is biased and affected by his **objective** of proving that the pain killer is really effective, which is unethical. Same thing happens, when the studies are done on lowering down of blood pressure, respiratory rate, arterial pressure, etc. It is also observed that students and other untrained staff are more prone to this practice and fix in their mind to prove their objective as positive & are not willing to take much of the headache, thereby collecting biased data irrespective of obtaining the true facts.

### **Data Analysis**

Recently I have come acrossed one PG Thesis, where they were studying the effect of an Ayurvedic Medicine in comparison to an Allopathic Drug to reduce the size of a kidney stone over a period of 1 year. At the time of analysis, the student told that Ayurvedic Medicine is an established drug and it is expected that this drug cause a substantial decrease in size. After the analysis, we found that there was no statistical significant difference in the size reduction between the two drugs. The reduction in Allopathic Drug was of 1.0 mm and in Ayurvedic Medicine was of 1.8 mm. The student said, sir, this Ayurvedic Medicine is an established drug and the reduction cannot be so low. So, you please change few readings to make it 2.5 to 3.0 mm so that it comes out significant. Here the student tried to prove an ineffective medicine as effective, which is highly unethical.

In a second example, a new faculty member was studying the growth of fungal infection in Psoriasis patients with medication and without medication. In without medication group, the rate was 30% and in with medication, it was 60%, which was not significant ( $p=0.057$ ) [5]. Since the results were not significant, he asked me to reduce the rate a little bit in without drug group to obtain the significance. When I made it 25%, the p-value comes 0.026, which was significant. It was just a borderline case, where due to inadequate sample size, this problem arise.

The idea is only to explain that students, researchers and sometimes, some of the young faculty members support to adopt such unethical practice to obtain the significance, which leads to false results.

Sometimes lack of proper knowledge about the statistical technique is the reason to adopt wrong practice. In one of the example, where an article was referred to me for reviewing, unpaired t-test was applied on qualitative data. When authors were asked to consult a biostatistician, they replied, it is done by a statistician.

In another example, Pearson's Chi square p-value instead of Fishers Exact p-value, was used even when the expected frequencies were very low. Pearson's Chi-sq p-value was 0.041 and Fisher's exact p-value was 0.078. Just for the sake of reporting significant results, the authors used  $p=0.041$  irrespective of whether it is convincing or not. When the authors were asked about who has performed this analysis, they replied, a senior resident of PSM department. So, lack of knowledge also leads to wrong results.

Recently a sample size was to be calculated for the PG thesis protocol on effect of low PAPP levels on adverse pregnancy outcome. I referred few of the articles brought by the student. In one of the articles, they obtain the association of PAPP levels (Low or High including normal) with pregnancy outcome (normal or abnormal) using Chi-Sq test. In that article, they described one big paragraph on linear by linear association which was significant although there is no role of linear association with 2 categories.

These are the few examples, which I could narrate, but otherwise there are many cases where Bonferroni correction is not applied, Linear regression is done on qualitative dependent variable, instead of logistic regression, etc.etc.

## Conclusion

Whenever a research is planned, data should be collected by some competent and trained staff. Before the collection process starts, they must be made well aware of, the pros & cons of recording the false data and the basic ethical principles of honesty, integrity and truthfulness. Because, the final result will depend upon the reliability of the data input only and thus, in-turn affects the welfare of the patient. ***They must be made clear in their mind that the results are not going to be applied on non living things.*** We cannot afford to take a second chance. The next subject for treatment can be their own near & dear. So, they must be very careful & honest. Stated simply, it is unethical to carry out a bad scientific experiment [6].

Collection of Biostatistics  
Research Archive

The biostatistician or whosoever is conducting the analysis work, must also keep the same in his mind that to the best of his knowledge & belief, he must not suppress the true results and must have a strong desire to report the unfavourable facts also. It is always better to consult a senior biostatistician, if one is not very much sure about his knowledge.

## References

1. National Institute of Environmental Health Sciences,  
<http://www.niehs.nih.gov/research/resources/bioethics/whatis/> accessed on 07-05-14.
2. Ellwood P. Shattuck Lecture--Outcomes Management: A technology of patient experience. *N Engl J Med* 1988;318:1549-1556.
3. National Heart, Lung, and Blood Institute,  
<http://www.nhlbi.nih.gov/studies/clinicaltrials/> accessed on 21-08-2014.
4. Visual Analgue scale, [http://en.wikipedia.org/wiki/Visual\\_analogue\\_scale](http://en.wikipedia.org/wiki/Visual_analogue_scale) accessed on 10-05-2014.
5. Statistical Package for Social Sciences, IBM SPSS Statistics, IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.
6. Denham MJ, Foster A, Tyrell DAJ. Work of a district ethical committee. *Br Med J* 1979.II::1042-45

