Original Article

REPORTING STATISTICS IN CLINICAL TRIALS PUBLISHED IN INDIAN JOURNALS: A SURVEY

Jay Karan[1], ND Kantharia[2], Preeti Yadav[3], Pankaj Bhardwaj[4]

ABSTRACT

objective: Clinical trials are having very important place in the hierarchy of evidence based medicine. It has been observed that current methods of use and reporting of statistics of clinical trials are responsible for errors in the interpretation of results.We decided to evaluate clinical trials published during 2007 and 2008 in four Indian journals to analyses statistical issues which may affect the interpretation of results.

Methodology: We analyzed all the clinical trials (46) published in Indian Pediatrics, Indian Journal of Pharmacology, Journal of Postgraduate Medicine and Indian Journal of Dermatology, Vanereology and Leprology in 2007-2008.

Results: Median number of end points reported in clinical trials as well as median number of end points which were used for testing of significance was four. Twenty one (45%) of trials reported repeated measurement. Eighteen (39%) trials had three or more than three treatment groups. median number of test of significance was 15. post hoc subgroup analysis was done in 19% (9) of trials. P value was the sole criteria for interpretation of results in most of the trials; confidence interval was calculated in 11 (23%) trials. Baseline comparison between the study groups was done in 41 (89%) trials. In all cases comparison was done by statistical tests. Exact sample size was calculated in 18 (39%) trials.

***Conclusion*:** There are great chances of committing error during the interpretation of results of these trials because of multiple treatment groups, repeated measurements of endpoints, multiple tests of significance, over reliability on P value and less use of confidence interval. Statistical methods available to reduce chances of errors should be used and result should be interpreted accordingly.

KEY WORDS: Clinical trials, Statistics, Type-1 error.

Pak J Med Sci January - March 2010 Vol. 26 No. 1 212-216

How to cite this article:

Karan J, Kantharia ND, Yadav P, Bhardwaj P. Reporting statistics in clinical trials published in Indian journals: A survey. Pak J Med Sci 2010;26(1):212-216

Correspondence:

Dr. Jay Karan,
MD (Pharmacology), PG Diploma in Research Methodology
Assistant Professor,
Department of Pharmacology,
Govt. Medical College,
New Civil Hospital
Surat - India 395001.
E-mail: drjaykaran@yahoo.co.in
pali6gadh@yahoo.com

* Received for Publication: October 8, 2009
* Revision Received: November 18, 2009
* Revision Accepted: November 21st 2009

INTRODUCTION

The randomized controlled clinical trial is considered as the gold standard for an evaluation of effectiveness of any new drug or intervention. The methodologies of clinical trials have improved over the years. Statistics used in analysis of data of clinical trials has also improved because of the availability of the software for analysis of data. By using the software more complex statistics can be performed like analysis of several endpoints, repeat measure-

ment of endpoints over the course of treatment, subgroup analysis and comparison of more than two treatments, but this leads to various types of statistical errors like danger of excessive use of the statistical testing and inflation of type 1 error. Morever if all findings are not expressed in article or abstract then it may lead to exaggeration or underplaying of actual treatment difference.

Studies published in different journals have pointed out poor quality of statistical reporting of clinical trials.[1] Unfortunately even simple statistical methods like t test and chi-square tests are misused for data analysis because there test assumptions were not evaluated sufficiently before application of tests.[2-4] If a study contains multiple endpoints necessitating multiple statistical tests, it is important to check the rate of false positive results and the potential inflation of type I errors by applying adequate multiple-comparison corrections.[5,6] post hoc subgroup analysis, not pre specified at the start of study should be avoided as it lead to bias.[7] We found that all these studies were related to articles published in journals of western countries and there is need of conducting these studies in articles published in India. So our aim in this study was to discuss few issues associated with the reporting of statistics of clinical trials published in four Indian journals in year 2008-2009. Very few studies have bee carried out to address these issues and most of them are not for Indian journals.[8]

METHODOLOGY

We evaluated 46 clinical trials which were published in three Indian medical journals in 2007 and 2008. Out of these 15 from Indian Pediatrics, 6 were from Indian Journal of Pharmacology, 18 were from Indian Journal of Dermatology, Vanereology and Leprology and 7 from Journal of Post Graduate Medicine. All authors evaluated each clinical trial for various parameters related to design, analysis and reporting. Discrepancies between the authors were resolved by consensus. Ten clinical trials were reevaluated by first author to check the reliability and no substantial difference was found (kappa – 0.62). Our survey included only the comparative trials.

Statistical Analysis: Descriptive statistics was used for the measurement of frequency of events. Exact frequency was reported with proportion and 95% confidence interval around the proportion.

RESULTS

We observed that the median number of patients of the trial was 65. In 7 trials sample size was less than 30 and in 8 trials it was more than 100. Range of sample size was 16 to 206.

Table-I shows the number of endpoints mentioned in each clinical trial report, as well as the number for which statistical test used to see the difference between groups. We consider endpoints as those event or outcome that can be measured objectively to determine whether the intervention being studied is beneficial. We did not included adverse effects as endpoints of trial. The median number of end points was four. Twenty five (0.54, 95%CI 0.40 to 0.67) trials reported number of end points 5 or less than 5. 9 or more than 9 end points were observed in 8 (0.17, 95%CI 0.09 to 0.30) trials. Median end points reported in the clinical trials of four journals were similar (4 for Indian journal of pharmacology, 4 in Indian pediatrics, 4 in Indian journal of dermatology, vanereology and leprology and 5 in journal of post graduate medicine). Significant test was used for most of the end points. Median end points for which significant test used was 4. Most of the trials contain the end points of both qualitative and quantitative type; only 1 (0.021, 95%CI 0.003 to 0.113) trial has survival analysis as end point. Primary end points were mentioned in 23 (0.50, 95%CI 0.36 to 0.63) trials. Multiple end points were not adjusted in any trial. In the studies where primary end points were mentioned, in none of the study statistically significant secondary endpoints were emphasized over the primary endpoint.

In trials related to quantitative end points, end points are usually measured before the start of intervention and many times after the start of

Table-I: Distribution of number of end points in 46 clinical trials

No. of end point mentioned	*No. of trials*	*No. of end point tested for significance*	*No. of trials*
3 to 5	25	3 to 5	29
6 to 8	13	6 to 8	12
9 to 11	8	9 to 11	5

intervention. In our study 21 (0.45, 95%CI 0.32 to 0.59) of trials reported repeated measurements. Among these eighteen (0.85, 95%CI 0.65 to 0.95) trials reported significant test at every measurement. There was no technique used to decrease type I error generated because of repeated measurement over the time in any trial.

Twenty one (0.45, 95%CI 0.32 to 0.59) trials were having two treatment groups. 18 (0.85, 95%CI 0.65 to 0.95) trials were having three or more than three treatment groups. Seven (0.15, 95%CI 0.07 to 0.28) trials were having two period crossover designs. In the trials where multiple treatment groups were analyzed in paired fashion, repeat measure ANOVA or its non parametric equivalent (Friedman test) was used to analyze data (17 (0.94, 95%CI 0.74 to 0.99) trials).

We found that median number of significance test was 15 (Table-II). It includes all significance tests including the subgroup analysis, multiple end points and repeated measurement. More than 20 significant tests were used in five (0.10, 95%CI 0.04 to 0.23) trials. Range of number of significance tests was 2 to 53.

In this study we found that subgroup analysis was done in 9 (0.19, 95%CI 0.10 to 0.33) of studies. Three (0.33, 95%CI 0.12 to 0.64) of these has more than one prognostic factor in the subgroup analysis.

Confidence interval was calculated in 11 (0.23, 95%CI 0.13 to 0.37) trials. We observed that baseline comparison between study groups were done in 41(0.89, 95%CI 0.76 to 0.95) articles. In all trials the comparison was done by a statistical test. Exact sample size was calculated in 18 (0.85, 95%CI 0.65 to 0.95) trials.

Table-II: No. of test of significance used in clinical trial of three Indian journals

No. of tests	*No. of trials*
1 to 5	7
6 to 10	18
11 to 20	15
21 to 50	3
> 50	2

DISCUSSION

We observed that multiple end points were evaluated in all of the trials. This enhances the chance of type I error.[9] One method of addressing this problem is by considering most important end point as primary end point during the design of trial and validity of hypothesis checked by using statistical test on this end point only. Other end points should be considered as secondary end points and should be analyzed as exploratory.[8] In this study 23 (50%) trials used this method. In some condition it is very difficult to identify single end point in advance. In that condition some other statistical methods can be used to adjust the inflation of type 1 error like Bonferroni correction method and composite end point method.[8-10] In a study done by Ton J et al (2006) it was found that among the 16 randomized controlled trials with positive results published in British medical journal in 2004, only 8 trials remain positive after Bonferroni correction.[9]

We also observed repeated measurement of end points in many of the trials (45%). It may also leads to type 1 error. During the design of trial a strategy should be planned to deal with the repeated measurement. There are various strategies available for this like comparing average effect of treatment over the time or one or two time point can be fixed in advance for comparison between the treatment or time period to attend specific threshold value can be compared.[8,9,11] In our study we found that in no trial these adjustments were done so there is profound chance of having type 1 error in these trials.

Same problem was also seen in the case of more treatment groups that will also leads to type 1 error. ANOVA and Friedman one-way ANOVA are used to decrease the inflation of type I error. This method was used in 17 trials. This is an encouraging finding. There are some other statistical methods available which can be used for adjustment of P values.[9,12]

We found that median number of statistical test per trial was 15. As compared to other similar studies done on western medical journals, the number is more.[8] Actual number of statistical tests may be more as authors usually report the significant tests only. The best scenario of validity of statistical test is, when only one predefined end point is used in analysis. More statistical tests lead to type I error.[13] So a primary end point should be defined before starting the clinical trial and secondary end points should be used for exploration of hypothesis.[9.12]

In our study subgroup analysis was done in 9 (19%) of trials done which is less compared to similar studies published in western medical journals.[14] It is common practice to explore a subgroup of patients if overall result is not significant.[14] If author decides to do the subgroup analysis then it should be designed before the start of trial. In our study we found that planned subgroup analysis was not reported in any study. Subgroup analysis can increase type I error. However, a preferred alternative to subgroup analysis is to combine the factors into a single predictive model (an equation, such as a regression analysis), rather than to analyze each subgroup separately. The researchers thus test for interactions between the variable and the endpoint to avoid subgroup analysis. We found that this method was not used in any of the 9 trials in which subgroup analysis was reported. Subgroup analyses always include fewer patients than does the overall analysis, they carry a greater risk of making a type II error—falsely concluding that there is no difference. Guideline should be followed to report the subgroup analysis of clinical trials.[14]

In this study we observed that in most of the trial P value was the only method of reporting the results. P value is often misinterpreted,[15] and even if it is interpreted rightly it has some limitations.[16] Not writing the exact P value enhanced the problem further.[13] For the main results absolute difference between the groups with 95% confidence interval should be reported with or without P value.[13]

We found that in many of the trials baseline comparison was done with the help of statistical test. In a properly randomized trial each participant has equal chance of assigment to any of the study groups so any difference in the prognostic factor is because of the chance and not due to bias. It shows statistical imbalance. The result of the trials should be adjusted for this statistical imbalance by regression methods.[8,13]

It was observed that exact sample size was calculated in only 39%.In clinical trials sample should be big enough to have a high chance of detection, as statistically significant, a worthwhile effect if it exists, and thus to be reasonably sure that no benefit exists if it is not found in trial. For sample size calculation in hypothesis testing researcher must know the effect size, standard deviation, significant level and power of study.[17] Effect size and standard deviation of new agents can be calculated by pilot study. In a review by Alexander M et al in 2007 these findings are confirmed.[18]

***Limitations of the study*:** We evaluate clinical trials on the basis of few statistical problems reported in clinical trials of western medical journals; this is the limitation of our study as more extensive criteria could have been used. We believe that very few studies are published regarding reporting of the statistics in Indian journals and this study is one of them. Another limitation of our study is less number of journals under study and less number of trials. Study with more Indian journals and more clinical trials may give better validity to these results.

We believe that clinical trials published in four representative Indian journals are not devoid of statistical problems and most important problem is type I error. Though there are various methods available to decrease these errors, they are not used during the reporting of these tri-

als. It leads to exaggeration of results. Reader should be careful during the formation of opinion on the basis of these trials.

REFERENCES

1. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA 1994;272:122–4.
2. MacArthur RD, Jackson GG. An evaluation of the use of statistical methodology in the J Infectious Diseases. J Infect Dis 1984;149:349–54.
3. Kanter MH, Taylor JR. Accuracy of statistical methods in Transfusion: a review of articles from July/ August 1992 through June 1993. Transfusion. 1994;34:697–701.
4. Dar R, Serlin R, Omer H. Misuse of Statistical Tests in three Decades of Psychotherapy Research. J Consult Clin Psychol 1994;62:75–82.
5. Goodman NW, Hughes AO. Statistical Awareness of Research Workers in British Anaesthesia. Br J Anaesth 1992;68:321–4.
6. Avram MJ, Shanks CA, Dykes M, Ronai AK, Stiers WM. Statistical Methods in Anesthesia Articles: An Evaluation of Two American Journals during Two Six-Month Periods. Anesth Analg 1985;64:607–11.
7. Moreira ED, Stein Z, Susser E. Reporting on methods of subgroup analysis in clinical trials: a survey of four scientific journals. Brazilian J Med Biol Res 2001;34:1441–6.
8. Pocock S, Hughes M, Lee R. Statistical problems in the reporting of clinical trials: a survey of three medical journals. N Eng J Med 1987; 317:426-32.
9. Ton J, Aeilko H. Clinical trials are often false positive: A review of simple methods to control this problem. Current Clinical Pharmacology 2006;1:1-4.
10. Cleophas TJ, Zwinderman AH, Cleophas AF. Multiple statistical inferences. In: Statistics applied to clinical trials. Kluwer Academic Publishers, Boston, MA, 2002, pp. 73-81.
11. Cleophas GM, Cleophas TJ. Clinical trials in jeopardy. Int J Clin Pharmacol Ther 2003; 41:51-56.
12. Sarah J, Val J, Anthony C. Multiple analyses in clinical trials: sound science or data dredging? MJA 2004;181:452-454.
13. Tom L. Twenty statistical error even YOU can find in biomedical research articles. Croat Med J 2004;45:361-370.
14. Rui W, Stephen W, James H, David J, Jeffrey M. Statistics in Medicine — Reporting of Subgroup Analyses in Clinical Trials. N Engl J Med 357;21:2190-2195.
15. Bailar JC, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Ann Intern Med 1988;108:266-73.
16. Gardner MJ, Altman D. Confidence intervals rather than P values: estimation rather than hypothesis testing. BMJ 1986;292:746-50.
17. Altman D. Practical statistics for medical research. London: Chapman and Hall, 1991. 456.
18. Alexander M, Qamruz Z, Karl P, Georg G, Hanno U. Statistical errors in medical research –a review of common pitfalls. SWISS MED WKLY 2007;137:44–49.

Authors:

1. Dr. Jay Karan,
   MD (Pharmacology),
   PG Diploma in Research Methodology,
   Assistant Professor
2. Dr. N. D. Kantharia
   Professor and Head
3. Dr. Preeti Yadav,
   MD (Pharmacology),
   Associate Professor
4. Dr. Pankaj Bhardwaj,
   MD (Community Medicine),
   Assistant Professor,
   Dept. of Community Medicine,
   Era's Lucknow Medical College,
   Lucknow - India.

1-3. Dept of Pharmacology,
   Govt. Medical College,
   New Civil Hospital,
   Surat India- 395001.