

Research design and statistical methods in Pakistan Journal of Medical Sciences (PJMS)

Sohail Akhtar¹, Syed Wadood Ali Shah², M. Rafiq³, Ajmal Khan⁴

ABSTRACT

Objective: This article compares the study design and statistical methods used in 2005, 2010 and 2015 of Pakistan Journal of Medical Sciences (PJMS).

Methods: Only original articles of PJMS were considered for the analysis. The articles were carefully reviewed for statistical methods and designs, and then recorded accordingly. The frequency of each statistical method and research design was estimated and compared with previous years.

Results: A total of 429 articles were evaluated (n=74 in 2005, n=179 in 2010, n=176 in 2015) in which 171 (40%) were cross-sectional and 116 (27%) were prospective study designs. A variety of statistical methods were found in the analysis. The most frequent methods include: descriptive statistics (n=315, 73.4%), chi-square/Fisher's exact tests (n=205, 47.8%) and student *t*-test (n=186, 43.4%). There was a significant increase in the use of statistical methods over time period: *t*-test, chi-square/Fisher's exact test, logistic regression, epidemiological statistics, and non-parametric tests.

Conclusion: This study shows that a diverse variety of statistical methods have been used in the research articles of PJMS and frequency improved from 2005 to 2015. However, descriptive statistics was the most frequent method of statistical analysis in the published articles while cross-sectional study design was common study design.

KEY WORDS: Biostatistics, Statistical analysis, Study design, Statistical methods, Bibliometric analysis.

doi: <http://dx.doi.org/10.12669/pjms.321.9033>

How to cite this:

Akhtar S, Shah SWA, Rafiq M, Khan A. Research design and statistical methods in Pakistan Journal of Medical Sciences (PJMS). Pak J Med Sci. 2016;32(1):151-154. doi: <http://dx.doi.org/10.12669/pjms.321.9033>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Nowadays, biostatistics play an extremely active role in the field of medical research to provide good services to public health. New analytical techniques

and tools for biostatistics are consistently being found. There has been a tremendous growth in the development of statistical techniques and methods over the past few decades. These promising areas of application consist of medical trials, investigative genomics, observational type study, imaging and structural study. Simply, statistics has become indispensable to evaluate treatment effects in medical research.¹⁻³ Therefore, the appropriate use of bio-statistical techniques and methods are becoming more and more important in biomedical research. Many biomedical journals have hired statisticians to review the appropriateness of statistical methods and approaches.

During the last few years, a number of papers have been published to illustrate the use of statistical techniques and study design in biomedical journals.⁴⁻¹⁷ Most of these review papers showed that there has been a positive improvement in the use of statistical methods in medical research. Two

1. Dr. Sohail Akhtar, PhD (UK).
Department of Statistics,
2. Syed Wadood Ali Shah (PhD Candidate).
Department of Pharmacy,
3. M. Rafiq (M.Phil)
Department of Statistics,
4. Ajmal Khan,
Department of Zoology,
- 1-4: University of Malakand, Lower Dir,
KPK, Pakistan.

Correspondence:
Dr. Sohail Akhtar,
Assistant Professor, Department of Statistics,
University of Malakand, Lower Dir, KP, Pakistan.
E-mail: s.akhtar@uom.edu.pk

- * Received for Publication: September 21, 2015
- * 1st Revision Received: September 28, 2015
- * 2nd Revision Received: November 24, 2015
- * Final Revision Accepted: November 28, 2015

similar studies have been published for Pakistani journals which mainly focused on the comparison of statistical methods and research designs of indexed and non-indexed journals.^{18,19} However, in this article, we compared the level of improvement of statistical content and research design used during three different years (2005, 2010, 2015) of Pakistan journal of Medical Sciences. Pakistan journal of Medical Sciences is a bi-monthly and ISI indexed journal. The journal has been successfully continuing to hold a broad range of articles on general medicine and health, and has been provided a significant source of medical research in Pakistan. So this study will help the biomedical readers and researchers to understand the level of statistics that have been published in Pakistan Journal of Medical Sciences.

METHODS

To perform the analysis, all the original papers were downloaded from the PJMS web-site for these each years from those published in the year 2005, 2010 and 2015 (Jan to August). A total of 429 articles were studied. Only original research articles were included in the study and case reports, reviews, letters to the editor, editorials, systematic reviews, conference proceedings, meta-analysis, publication audit were excluded from the analysis.

Each article was carefully reviewed to draw out the research design and type of statistical methods used in their analysis. An excel spreadsheet was used to record the required information from the

articles. Through the investigation, we found all the major research designs (cross-sectional survey, prospective study, retrospective study, randomized clinical trial (RCT)). On the other hand, an extensive variety of statistical methods were found in the review of PJMS articles. The statistical methods were categorized in the almost similar manner as published in earlier literature with some modification (Table-I).^{4,20} Initially, the statistical methods were categorized in parametric and non-parametric tests. These tests were categorized into basic and advance. In the investigation, a wide number of basic tests were found, for example, chi-square/Fisher's exact test, student *t*-test, analysis of variance (ANOVA), regression analysis, etc. On the other hand several advanced statistical methods were found, such as, logistic regression, receiver operating characteristic (ROC), sensitivity, specificity, odd ratio, etc. A comparison of these three years (2005, 2010, 2015) was made by calculating the frequency of each research design and statistical content used. Chi-square test was then used to check the statistical difference between these frequencies of three time periods. Statistical software SPSS version 20 was used for all calculations and value with $p < 0.05$ was considered significant.

RESULTS

A total of 429 articles were studied for this investigation: 74 (17.2%) from 2005, 179 (41.7%) from 2010, and 176 (41.0%) from 2015. Majority of the articles used a cross-sectional study design

Table-I: Classification of statistical methods as used in PJMS in 2005, 2010 and 2015.

| <i>Statistical Method</i> | <i>Description</i> |
|----------------------------|--|
| Descriptive statistics | Ratios and percentages means, standard deviations, standard errors. |
| Student t-tests | One-sample, two-samples t-tests and paired samples |
| Contingency tables | Chi-square (χ^2) tests, Fisher's exact test, McNemar's test |
| Parametric Correlation | Pearson correlation coefficient, |
| linear regression | Least squares regression with one outcome variable and one or more predictor variables |
| Epidemiological statistics | Statistics relative risk, odds ratio, log odds, measures of association, sensitivity, specificity, receiver operating characteristic (ROC) |
| Multi-way tables | log-linear models |
| Non-parametric test | Sign test, Wilcoxon signed-rank test, Mann-Whitney test, Kruskal-Wallis Test, Friedman Test, Shapiro-Wilk test |
| Non-parametric Correlation | Kendall rank, Spearman |
| Analysis of variance | Analysis of variance (ANOVA), analysis of co-variance (ANCOVA) |
| Multiple comparisons | To find out the source of the significant differences among the group means: Dunnett's procedure, Tukey's Method, Fisher's LSD procedure. |
| Regression for survival | Includes logistic regression, Poisson regression |
| Sensitivity analysis | Kaplan-Meier method, log-rank test |
| Other | Anything not fitting above headings, Factor analysis, Bartlett's Test, Shapiro-Wilk's. |

Table-II: Summary of research design used in the samples.

| | 2005 | | 2010 | | 2015 (Jan. to Aug.) | | P-Value |
|------------------------------|----------|----|----------|----|---------------------|----|---------|
| | Articles | % | Articles | % | Articles | % | |
| Cross-sectional survey | 28 | 38 | 75 | 42 | 68 | 39 | 0.055 |
| Retrospective study | 10 | 14 | 30 | 17 | 38 | 22 | |
| Prospective study | 18 | 24 | 47 | 26 | 51 | 29 | |
| Randomized controlled trials | 6 | 8 | 17 | 9 | 11 | 6 | |
| Not clear from study | 12 | 16 | 10 | 6 | 8 | 5 | |
| | 74 | | 179 | | 176 | | |

*chi-square statistic is 15.227. The p-value is 0.055. The result is not significant at $p < 0.05$.

(n=171, 40%), prospective study design (n=116, 27%), retrospective study (n=78, 18%), randomized clinical trials (n=34, 8%). There were no significant differences (with $p=0.055$) in study designs used over the three periods (2005, 2010, 2015). We also noted that (n=30, 7%) manuscripts did not mention a very clear research design.

A number of statistical methods were found during the review of published articles (Table-III). Majority of the articles used descriptive statistics: 50 (67.6%) from 2005, 135 (75.4%) from 2010 and 130 (74.0) from 2015. The p-value (0.430) shows that the trend of descriptive statistics was insignificant over the three periods at 5% level of significance. The trend of the use of correlation, regression and analysis of variance were slightly greater in percentage over time periods but not statistically significant. On the other hand there was a significant increase in the use of statistical methods: student *t*-test (48.9% in 2015 vs. 27.0% in 2005), logistic regression (9.7% in 2015 vs 5.4% in 2005), chi-square/Fisher's exact test (51.1% in 2015 vs. 20.3% in 2005), epidemiological statistics (17.6% in 2015 vs. 4.1 in 2005), non parametric tests (24.4% in 2015 vs 5.4% in 2005).

A number of other different statistical methods were reported only once or twice in the PJMS articles, for example, factor analysis, component

analysis, and Poisson regression, etc (see Table-II). Therefore, we ignored these statistical methods in the Table-III.

DISCUSSION

Our findings show that cross-sectional research design was the most common study in the published articles of PJMS during each of the three years. This result is consistent with other published studies.^{21,22} Prospective study designs is used in almost one third of the published articles and is the second most frequent type of research design. On the other hand, there is a low frequency of retrospective study and randomized clinical trials in the published articles. Furthermore, there was no statistical significance difference between research designs used over three times periods.

Previous reports shows that descriptive statistics is the most frequent method of analysis in publications of PJMS.²³ Simple statistical methods like student *t* - tests, correlation, chi-square and Fisher's exact tests, Mann-Whitney, Kruskal-Wallis and Wilcoxon signed rank test are frequently used. There is a significant increase in the number of articles using chi-square and *t* - tests between the three periods. This is consistent to previously published studies.²³ We also found that there was a significant increase in the reports using

Table-III: Outline of statistical methods used in PJMS.

| Statistical methods | 2005 | | 2010 | | 2015 (Jan. to Aug.) | | p-value |
|--|------|------|-------|------|---------------------|------|---------|
| | n=74 | % | n=179 | % | n=176 | % | |
| Descriptive statistics | 50 | 67.6 | 135 | 75.4 | 130 | 74.0 | 0.430 |
| t- test (independent/paired) | 20 | 27.0 | 80 | 44.7 | 86 | 48.9 | 0.005 |
| Logistic regression | 4 | 5.4 | 10 | 5.6 | 17 | 9.7 | 0.030 |
| Regression | 2 | 2.7 | 4 | 2.2 | 10 | 5.7 | 0.202 |
| ANOVA/ANCOVA | 3 | 4.1 | 16 | 8.9 | 24 | 13.6 | 0.058 |
| Correlation | 10 | 13.5 | 20 | 11.2 | 21 | 11.9 | 0.872 |
| Contingency table(Chi-square/Fisher's exact) | 15 | 20.3 | 100 | 55.9 | 90 | 51.1 | 0.000 |
| Epidemiological statistics | 3 | 4.1 | 14 | 7.8 | 31 | 17.6 | 0.001 |
| Nonparametric | 4 | 5.4 | 17 | 9.5 | 43 | 24.4 | 0.000 |

chi-square, *t*-test, epidemiological statistics and logistic regression between the three time periods. However, the frequency of complex statistical methods like logistic regression, Poisson regression and factor analysis is quite low as compared to simple statistical methods

CONCLUSION

In conclusion, the current analysis shows that Pakistan Journal of Medical Sciences (PJMS) used all research study designs in published articles but there is still much room for progress concerning advance statistical methods. Most of the published articles continue to use a cross-sectional design while retrospective study and randomized clinical trials still require more attention as compared to journals of advanced countries. It is noted that a variety of statistical methods were found in the published articles but the frequency of advance statistical methods was quite low as compared to advanced countries journals. Furthermore, it is concluded that the medical researchers and readers should have at least the basic knowledge of biostatistics to understand the biomedical research and its interpretation.

Grant Support & Financial Disclosures: None.

REFERENCES

- Cadarso-Suárez C, González-Manteiga W. Statistics in Biomedical Research. *Arbor Ciencia Pensamiento y Cultura*. 2007;(725):353-361. doi:10.3989/arbor.2007.i725.108
- Armitage P, Berry G, Matthews JNS. Statistical methods in medical research. John Wiley & Sons: 2008. Edition 4th.
- Forthofer RN, Lee ES, Hernandez M. Biostatistics: A Guide to Design, Analysis and Discovery: 2006; Academic Press; 2 edition.
- Emerson JD, Colditz GA. Use of statistical analysis in the New England Journal of Medicine. *N Engl J Med*. 1983;309:709-713. doi: 10.1056/NEJM198309223091206
- Schwartz SJ, Sturr M, Golberg G. Statistical methods in rehabilitation literature: a survey of recent publications. *Arch Phys Med Rehabil*. 1996;77:497-500. doi: 10.1016/S0003-9993(96)90040-4
- Morris RW. A statistical study of papers in the journal of bone and joint surgery (BR) 1984. *J Bone Joint Surg Br*. 1988;70:242-246.
- Juzych MS, Shin DH, Seyedsadr M, Siegner SW, Juzych LA. Statistical techniques in ophthalmic journals. *Arch Ophthalmol*. 1992;110:1225-1229.
- Bailar JC, Mosteller F. Guidelines for statistical reporting in articles for medical journal: amplifications and explanations. *Ann Intern Med*. 1988;108:266-273. doi:10.7326/0003-4819-108-2-266.
- Avram MJ, Shanks CA, Dykes MHM, Ronai AK, Stiers WM. Statistical methods in anesthesia articles: an evaluation of two American journals during two six-month periods. *Anesth Analg*. 1985;64:607-611. doi: 10.1213/00000539-198506000-00009.
- 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. doi:10.1111/j.1365-3148.1995.tb00194.x.
- Horton NJ, Switzer SS. Statistical methods in the journal. *N Engl J Med*. 2005;353:1977-1979. doi:10.1080/03610926.2013.858166.
- Jaykaran, Kantharia ND, Preeti Y, Bharddwaj P, Goyal J. Reporting statistics in clinical trials published in Indian journals: a survey. *African Health Sci*. 2010;10(2):204-207. doi: 10.1371/journal.pone.0121268.
- Solanki P, Saxena D, Kantharia ND. Study design and statistical methods in Indian medical journals. *Journal of Pharmaceutical Negative Results*, 2011;2(1):35-38. doi: 10.4103/0976-9234.82983
- Jaykaran, Preeti Y. Quality of reporting statistics in two Indian pharmacology journals. *J Pharmacol Pharmacother*. 2011;2(2):85-89. doi: 10.4103/0022-3859.97174.
- Yergens, DW, Dutton DJ, Scott BP. An overview of the statistical methods reported by studies using the Canadian community health survey. *BMC Med Res Methodol*. 2014;14(15). doi:10.1186/1471-2288-14-15
- Vähänikkilä H, Leo T, Pentti N. The statistical reporting quality of articles published in 2010 in five dental journals. *Acta Odontologica Scandinavica*. 2015;73(1):76-80. doi:10.3109/00016357.2014.954612
- Vähänikkilä H, Käkilehto T, Pihlaja J, Pääkilä J, Tjäderhane L, Suni J, et al. A data-based study on survival of permanent molar restorations in adolescents. *Acta Odontol Scand*. 2014;72:380-385.
- Rao MH, Khan N. Comparison of statistical methods, type of articles and study design used in selected Pakistani medical journals in 1998 and 2007. *J Pak Med Assoc*. 2010;60(9):745-749. 13. doi: 10.1016/j.jclinepi.2007.05.003
- Khan N, Rao MH. Survey of statistical methods and type of articles published in the selected Pakistani medical journals in 1999 to 2007. *J Dow Univ Health Sci*. 2011;5(1):6-11.
- Colditz GA, Emerson JD. The statistical content of published medical research: some implications for biomedical education. *Med Educ*. 1985;19:248-255. doi: 10.1111/j.1365-2923.1985.tb01315.x.
- Wang Q, Zhang B. Research design and statistical methods in Chinese medical journals. *JAMA*. 1998;280:283-285. doi:10.1001/jama.280.3.283.
- Rigby AS, Armstrong GK, Campbell MJ, Summerton N. A survey of statistics in three UK general practice journal. *BMC Med Res Methodol*. 2004;4:28. doi: 10.1186/1471-2288-4-28
- Otwombe KN, Petzold M, Martinson N, Chirwa T. A review of the study designs and statistical methods used in the determination of predictors of all-cause mortality in HIV-infected cohorts: 2002-2011. *PLoS ONE*. 2014;9(2):e87356. doi:10.1371/journal.pone.0087356.

Authors' Contribution:

Sohail Akhtar and **Syed Wadood Ali Shah** conceived, designed and did statistical analysis & editing of manuscript.
M Rafiq and **Ajmal Khan** did data collection and manuscript writing.