

Study designs, use of statistical tests, and statistical analysis software choice in 2015: Results from two Pakistani monthly Medline indexed journals

Masood Ali Shaikh

Abstract

Assessment of research articles in terms of study designs used, statistical tests applied and the use of statistical analysis programmes help determine research activity profile and trends in the country. In this descriptive study, all original articles published by Journal of Pakistan Medical Association (JPMA) and Journal of the College of Physicians and Surgeons Pakistan (JCPSP), in the year 2015 were reviewed in terms of study designs used, application of statistical tests, and the use of statistical analysis programmes. JPMA and JCPSP published 192 and 128 original articles, respectively, in the year 2015. Results of this study indicate that cross-sectional study design, bivariate inferential statistical analysis entailing comparison between two variables/groups, and use of statistical software programme SPSS to be the most common study design, inferential statistical analysis, and statistical analysis software programmes, respectively. These results echo previously published assessment of these two journals for the year 2014.

Keywords: Statistics, Software, Programme, Pakistan Publication Research Journal.

Introduction

Assessment of research articles published in biomedical journals, in terms of study designs used, statistical tests applied and the use of statistical analysis programmes, helps elucidating research activity profile and trends in the biomedical specialties as well as countries where these journals are published.¹⁻⁴ Statistical tests are applied to make sense of data and glean information from it; "Statistics is the science of learning from data, and of measuring, controlling, and communicating uncertainty; and it thereby provides the navigation essential for controlling the course of scientific and societal advances".⁵

Two studies on the use of statistical tests, used in the articles published by Pakistani biomedical journals,

report that over the years, there has been an increase in the use and sophistication of statistical tests applied.^{6,7} A review of six biomedical journals of Pakistan (three each Medline indexed and non-indexed) for the years 1998 and 2007, reported that there were 299 'original articles' and 'short communications' published in the year 2007.⁶ However, there were only 7 (2.3%) articles in the Medline indexed journals reported use of multiple linear regression model, while only 1 (0.6%) out of 177 articles published in the non-indexed journals used this model. While in comparison to the year 1998, only 1 (0.6%) article out of 172 in the Medline indexed journals and none out of 101 articles in non-Medline indexed journals used multiple regression model. A review of all original articles published in the year 2014 by the three Medline indexed journals of Pakistan reported dramatic increase in the use of multivariate models.⁷

In Pakistan, two biomedical journals are published on a monthly basis i.e. Journal of Pakistan Medical Association (JPMA) and Journal of the College of Physicians and Surgeons of Pakistan (JCPSP). In this descriptive study, all original articles published by these two monthly Medline indexed journals in the year 2015 were reviewed in terms of study designs used, application of statistical tests, and the use of statistical analysis programmes.

Methods and Results

All 'original articles' published by the JPMA and JCPSP in the year 2015 were downloaded from the websites of JPMA and JCPSP, in November 2016. Both journals provide full text of all the articles free of cost. Analysis entailed categorization of each original article based on the study design used, type of statistical tests applied, and the statistical software used for analysis of data. In all instances, use of statistical tests, and study designs, as reported in the published article were taken at face value; in very few instances where such information was not reported, this information was gleaned from the description provided in the 'Methodology' section. Data were analyzed in terms of frequencies and percentages

.....
Independent Consultant, Karachi.

Correspondence: Email: masoodali1@yahoo.com

Table-1: Type of study designs used in the published original articles by two monthly Pakistani Medline indexed journals in 2015.

Study Design	JPMA N = 192 Frequency (%)	JCPSP N = 128 Frequency (%)
Descriptive	5 (2.6)	19 (14.8)
Observational	1 (0.5)	6 (4.7)
Case Series	1 (0.5)	14 (10.9)
Cross-sectional	74 (38.5)	41 (32.0)
Cohort	0	2 (1.6)
Case-Control	15 (7.8)	4 (3.1)
Randomized Clinical Trial/Clinical Trial/ Non-Randomized Clinical Trial/ Quasi-Experimental/Experimental/ Interventional/Pre-Post	42 (21.9)	30 (23.4)
Comparative	0	1 (0.8)
Analytical	1 (0.5)	5 (3.9)
Correlational	1 (0.5)	1 (0.8)
Validation	0	1 (0.8)
Mixed Methods	3 (1.6)	1 (0.8)
Prospective	17 (8.9)	0
Retrospective	21 (10.9)	2 (1.6)
Qualitative	2 (1.0)	0
Other	9 (4.8)	1 (0.8)

JPMA: Journal of Pakistan Medical Association

JCPSP: Journal of the College of Physicians and Surgeons of Pakistan.

using the freely available statistical analysis programme R version 3.3.2.

Cumulatively, 320 original articles were published by the two Pakistani monthly Medline indexed journals in the year 2015, with 192 (60.0%) by JPMA and 128 (40.0%) by JCPSP. On average JPMA published 16 (Range: 14 - 18) original articles per month. While JCPSP published 10.7 or 11 (Range: 10 - 12) original articles per month.

Study designs, as reported in the original articles, were taken at the face value. The structured abstract for original articles in JCPSP includes a sub-heading of "Study Design", which made it easier to categorize each original article. For original articles in JPMA, study design was primarily garnered from the methodology section. Table-1 presents the frequency and percentages of study designs used in the original articles published by the two journals in 2015. The most common study design in both journals was 'cross-sectional', cumulatively reported in 115 (35.9%) original articles.

Regarding use of statistical tests reported; in this study, statistical tests were combined in several groups, as most of the articles reported use of more than one

Table-2: Type of statistical tests reported by original articles published in two Pakistani Medline indexed monthly journals in 2015.

Statistical Tests	JPMA N = 192 Frequency (%)	JCPSP N = 128 Frequency (%)
Group 1 Frequencies, Percentages, Proportions Mean/SD, Median/IQR, Confidence Interval	43 (22.4)	35 (27.3)
Group 2 t-test, Chi Square/McNemar, Mann-Whitney U, Median test, Pearson Correlation, Spearman Correlation, Intra Class Correlation Coefficient, Kendall's tau, Relative Risk, Odds Ratio, Kappa, Wilcoxon Signed Rank	76 (39.6)	57 (44.5)
Group 3 Kaplan Meier Survival Curve, Log-Rank Test, Cox Proportional Hazards	3 (1.6)	2 (1.6)
Group 4 Linear Regression, Logistic Regression Multinomial Regression, Factor Analysis, ANOVA, MANOVA, ANCOVA, Kruskal-Wallis, Friedman	58 (30.2)	22 (17.2)
Group 5 Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, Accuracy, Area Under The Curve Receiver Operator Characteristic Curve	9 (4.7)	10 (7.8)
Group 6 Qualitative Analysis	2 (1.0)	1 (0.8)
Group 7 None Reported/Required/Other	1 (0.5%)	1 (0.8)

JPMA: Journal of Pakistan Medical Association

JCPSP: Journal of the College of Physicians and Surgeons of Pakistan.

statistical test. Table-2 lists the frequency and percentages of all the groups used, along with various statistical tests comprising each group. Univariate analysis entailing frequencies, percentages, mean, standard deviation, median, inter-quartile range etc. were grouped together. Statistical tests involving comparison between two groups/variables e.g. Chi-Square, t-tests, and their non-parametric counterparts, in addition to correlations, were all combined together in one group. While any analysis involving sensitivity, specificity, positive or negative predictive values were grouped separately. Statistical tests involving comparison between more than two groups/variables e.g. various types of regression models, Factor analysis, Analysis of Variance (ANOVA) and Kruskal-Wallis test etc. were combined together as a separate group. Statistical analysis involving time-to-event data were kept as a separate group. Using hierarchical approach to grouping of statistical tests entailing primary focus of study and non-overlapping analysis was espoused; where for instance original articles using Chi-square test in combination with ANOVA were counted as belonging to

Table-3: Statistical software programmes used in original articles published by two monthly Pakistani Medline indexed journals in 2015.

Statistical Software Programme	JPMA N = 192 Frequency (%)	JCPSP N = 128 Frequency (%)
SPSS	142 (74.0)	113 (88.2)
Excel	4 (2.1)	1 (0.8)
NCSS	2 (1.0)	0
STATA	2 (1.0)	1 (0.8)
Other	3 (1.6%)	0
Minitab	0	2 (1.6)
Epi Info	0	1 (0.8)
Use of two or more programmes	5 (2.6)	5 (3.9)
Not Specified/Not Needed	34 (17.7)	5 (3.9)

JPMA: Journal of Pakistan Medical Association

JCPSP: Journal of the College of Physicians and Surgeons of Pakistan.

ANOVA group.

Table 2 presents frequency and percentages of statistical tests reported by original articles in the two journals. Most frequently reported group of statistical tests in original articles by both journals was bivariate parametric and non-parametric tests i.e. involving comparisons between two groups/variables e.g. Chi-square test, t-tests, and various types of correlations; cumulatively, 133 (41.6%) original articles reported using these tests.

Table-3 presents frequency and percentages of the use of statistical analysis software programmes for data analysis in the original articles by both journals in year 2015. The most commonly used statistical analysis programme reported by both journals was SPSS. Cumulatively, 255 (79.7%) original articles exclusively used SPSS for data analysis. The SPSS versions used ranged from version 10 to 22 for both journals. The SPSS version 16 being most commonly reported in JPMA original articles i.e. 34 (23.9%), while SPSS version 20 most commonly reported in JCPSP i.e. 28 (24.8) in original articles reporting use of SPSS.

Discussion

In the year 2014, JPMA and JCPSP published 189 and 147 original articles, respectively.⁷ While 192 and 128 original articles were published in the year 2015 by JPMA and JCPSP, respectively. However, the study on original articles published in 2014 did not assess the study designs used, but in year 2015, the most common study design reported was cross-sectional i.e. 115 (35.9%). A previous study based on Journal of Family and Community Medicine looked at study designs of 229 articles published from 1994 to 2010,

and also reported that 175 (76.4%) used cross-sectional design.¹

The most common inferential statistical analysis involved comparison of two groups/variables in 2015, with t-tests, Chi-square test etc. were reported by 133 (41.6) original articles in both monthly journals. Those articles that used multivariate analysis i.e. involving inferential analysis of more than two groups/variables, in addition to bivariate analysis are not included here. These bivariate statistical tests were also most commonly reported by JPMA and JCPSP in 2104, as 150 (44.6%) original articles used them.⁷ SPSS (SPSS Inc., Chicago, IL, USA) was again the most commonly used statistical analysis software programme; 255 (79.7%) original articles published by the two monthly Medline indexed journals exclusively used this programme in 2015, compared to 280 (83.3%) in the year 2014. for data analysis.⁷ A study on the use of statistical analysis software programmes in the health services research studies published from 2007 to 2009, in the United States, reported most commonly used analysis programmes to be STATA (StataCorp LP, College Station, TX, USA) and SAS (SAS Institute Inc., Cary, NC, USA).

Results of this study indicate that cross-sectional study design, bivariate analysis, and use of SPSS to be the most common study design, inferential statistical analysis, and statistical analysis software programmes, respectively. The findings presented here are restricted to original articles published by the two monthly Medline indexed journals of Pakistan, for the year 2015, only. Albeit, these results are in agreement with 2014 results, reported previously. Nevertheless, these findings cannot be extrapolated to all biomedical journals or health/medical research conducted in the country. Several caveats are in order: analysis was restricted to review of original articles; inclusion of Medline indexed journals of Pakistan with monthly frequency of publication; and the fact that articles published by Pakistani researches in international journals were — by design — not included. Finally, no attempt was made to elucidate whether study design matched the description provided in the methodology section, or application of statistical tests was correct and conclusions drawn were warranted.

It could reasonably be argued that Pakistani Medline indexed journals would ostensibly have more stringent critical appraisal and peer review processes vis a vis non-indexed journals in the country. Hence, to better understand the study designs used and statistical analysis carried out by some of the best Pakistani medical and

health researchers, publishing in the country's Medline indexed journals; future studies need to assess all types of research articles published in all the Medline indexed journals of Pakistan.

Disclaimer: None to declare.

Conflict of Interest: None to declare.

Funding Disclosure: None to declare.

References

1. Aljoudi AS. Study designs and statistical methods in the Journal of Family and Community Medicine: 1994-2010. *J Family Community Med.* 2013; 20:8-11.
 2. Choi E, Lyu J, Park J, Kim HY. Statistical methods used in articles published by the Journal of Periodontal and Implant Science. *J Periodontal Implant Sci.* 2014; 44:288-92.
 3. Dembe AE, Partridge JS, Geist LC. Statistical software applications used in health services research: analysis of published studies in the U.S. *BMC Health Serv Res.* 2011; 11:252.
 4. Yim KH, Nahm FS, Han KA, Park SY. Analysis of statistical methods and errors in the articles published in the Korean journal of pain. *Korean J Pain.* 2010; 23:35-41.
 5. Davidian M, Louis TA. Why statistics? *Science.* 2012; 336: 12.
 6. 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; 60:745-50.
 7. Shaikh MA. Use of statistical tests and statistical software choice in 2014: tale from three Medline indexed Pakistani journals. *J Pak Med Assoc.* 2016; 66:464-6.
-