Original Article

Comparison of statistical methods, type of articles and study design used in selected Pakistani medical journals in 1998 and 2007

Masood Hussain Rao,[1] Nazeer Khan[2]

Pakistan Medical Research Council,[1] Dow University of Health Sciences,[2] Karachi.

Abstract

Objectives: To compare the statistical methods, types of article and design of studies used in 1998 and 2007 articles of leading indexed and non-indexed medical journals of Pakistan.

Methods: Six leading medical journals of Pakistan: (1) JCPSP, (2) JPMA, (3) JAMC, (4) PJMS, (5) PJMR, and (6) PAFMJ, were selected for this study. Articles reviewed were 1057 to achieve the above mentioned objective.

Results: The articles reviewed for 1998 and 2007 were 366 and 691, respectively. Original articles contributed the maximum percentage of 65.6%, followed by case reports with 24.8%. The contribution of case reports in 1998 was 20.5% which increased to 27.1% in 2007. There was no statistically significant difference between 'indexed' and 'non-indexed' journals for different type of statistical methods in 1998 or 2007. In total, 749 articles were categorized as 'original articles' or 'short communication'. Among them, 51% articles mentioned study design and 67.3% of them were correct for the respective methodology. In 1998, 202 (74%) articles did not use any statistics or indicated only descriptive statistics, while in 2007, 239 (50.2%) articles did the same. The reader who was familiar with t-test and contingency tables in 1998 could have understood 97.4% of the scientific articles. However, this percentage dropped to 83.0% in 2007.

Conclusions: Quality of elaborating methods and usage of biostatistics in 6 leading Pakistani medical journals improved from 1998 to 2007, but has still to come up as compared to other western medical journals (JPMA 60:745; 2010).

Introduction

Medical research depends on biostatistical methods to form meaningful inferences from uncertain and haphazard data.[1] Each article, includes different study designs whether it is interventional or non-interventional, and statistical methods to strengthen its results. Nevertheless, lack of knowledge of most of the physicians about statistical methods has been repeatedly documented.[2-4] On the other hand, due to recent development of systematic reviews in the form of meta analysis and evidence based medicine, it is important now that clinicians should understand statistical methods not only to incorporate it in their own research papers/work, but also to correctly translate the published literature into improved patient care.[5] It is due to doctors' belief that the interpretations made in published materials are appropriate.[6]

Many reviews have been conducted in different countries, especially in Western countries[2,5,7-17] to describe the use of statistical techniques in medical journals. Few reviews are also organized for the medical journals of developing countries.[18-21] Almost all the reviews showed that there was a positive growth for the use of statistical techniques in medical writings. Better and more complex statistical analyses are introduced to prove the hypotheses. However, there has been no statistical review of Pakistani medical journals in the literature, as far as the authors' knowledge is concerned. Therefore, this review was conducted to assess the level of improvement and changes that occurred in six leading Pakistani journals regarding the use of statistics, study design

and types of article in 1998 and 2007.

Objectives: To compare the statistical methods, types of article and design of studies used in 1998 and 2007 articles of leading medical journals of Pakistan and to compare the indexed and non-indexed journals.

Methodology

Six leading medical journals of Pakistan were selected for this study. They were: (1) Journal of College of Physicians and Surgeons of Pakistan (JCPSP), (2) Journal of Pakistan Medical Association (JPMA), (3) Journal of Ayub Medical College (JAMC), (4) Pakistan Journal of Medical Sciences (PJMS), (5) Pakistan Journal of Medical Research (PJMR), and (6) Pakistan Armed Forces Medical Journal (PAFMJ). The first three are indexed journals in Medline and last three are also the leading non-indexed journals of Pakistan. It should be noted that the above three medical journals are the only medical journals which are in Medline. All these journals are peer reviewed and have statistical consultants on their advisory board. All the issues of these journals of year 1998 and 2007 were reviewed for the study. One of the authors (MR) reviewed all these articles manually. JCPSP was bi-monthly in 1998 and became monthly in 2007. JPMA was monthly in both the years of 1998 and 2007. JAMC and PAFMJ were six monthly in 1998 and became quarterly in 2007. While remaining two journals, PJMR and PJMS, were quarterly in year 1998 and 2007. Number of issues reviewed for 1998 and 2007 of these six journals were 30 and 40, respectively. After combining the journals into 'indexed' and 'non-indexed' categories, descriptive statistics, type of the articles and different statistical methods were compared. All together 1057 articles were reviewed. The information recorded from these journals was: types of article; study design, appropriateness of study design, any statistical methods used, types of statistics and p-values with or without any statistical test mentioned. Wang and Zhang[18] classification was employed for study design grouping and Emerson and Colditz's[9] classification was used for statistical methods categorization. Editorials, commentaries and review of the books were not considered in this review. If more than one statistical technique were employed in one article, all of them were recorded; however, if the same statistical method was repeatedly used in the same article, the method was recorded only once. Chi-square test was employed to compare the differences of percentages. The significant level was fixed at 5%.

Results

Number of articles reviewed for 1998 and 2007 were 366 and 691, respectively. The maximum number of the articles reviewed was from JCPSP, which contributed to 28.1% of the total articles reviewed. The maximum number of issues reviewed were from JPMA (24 issues) which contributed 25.4% of all the articles reviewed.

Table-1: Type of articles published in the selected journals in 1998 and 2007.

Journal	Type of the article	1998	2007	Total	p-value
JAMC	Case Report	5 (16.7)	13(13.8)	18(14.5)	
	Original article	23(76.7)	73(77.7)	96(77.4)	
	Short Communication	0(0.0)	6(6.4)	6(4.8)	0.321
	Review article	2(6.7)	2(2.1)	4(3.2)	
	Total	30	94	124	
JCPSP	Case Report	22(25.6)	92(43.6)	114(38.4)	
	Original article	60(69.8)	108(51.2)	168(56.6)	
	Short Communication	1 (1.2)	6(2.8)	7(2.4)	0.019
	Review article	3(3.5)	5(2.4)	8(2.7)	
	Total	86	211	297	
JPMA	Case Report	25(21.6)	35(23.0)	60(22.4)	
	Original article	75(64.7)	91(59.9)	166(61.9)	
	Short Communication	13(11.2)	15(9.9)	28(10.4)	0.37
	Review article	3(2.6)	11(7.2)	14(5.2)	
	Total	116	152	268	
PAFMJ	Case Report	11(30.6)	24(38.1)	35(35.4)	
	Original article	22(61.1)	35(55.6)	57(57.6)	
	Short Communication	0(0.0)	0(0.0)	0(0.0)	0.734
	Review article	3(8.3)	4(6.3)	7(7.1)	
	Total	36	63	99	
PJMR	Case Report	1(2.5)	0(.0)	1(1.6)	
	Original article	34(85.0)	22(91.7)	56(87.5)	
	Short Communication	1 (2.5)	1(4.2)	2(3.1)	0.691
	Review article	4(10.0)	1 (4.2)	5(7.8)	
	Total	40	24	64	
PJMS	Case Report	11(19.0)	23(15.6)	34(16.6)	
	Original article	39(67.2)	111(75.5)	150(73.2)	
	Short Communication	5(8.6)	8(5.4)	13(6.3)	0.645
	Review article	3(5.2)	5(3.4)	8(3.9)	
	Total	58	147	205	
Total	Case Report	75(20.5)	187(27.1)	262(24.8)	0.128
	Original article	253(69.1)	440(63.7)	693(65.6)	
	Short Communication	20(5.5)	36(5.2)	56(5.3)	
	Review article	18(4.9)	28(4.1)	46(4.4)	
	Total	366	691	1057	

Original articles contributed the maximum with 65.6%, followed by case reports of 24.8%. The contribution of case reports to all 1998 publications was 20.5%, increased to 27.1% in 2007. On the other hand, the percentage of original articles was reduced from 69.1% to 63.7% during this period. The major contribution of increase of case reports was due to JCPSP, where it was increased from 25.6% in 1998 to 43.6% in 2007. On the other hand, the publication of original articles in JCPSP was reduced from 69.8% to 51.2% in these two years. These changes were statistically significant (p=0.019). The original articles were also reduced by about 5% in JPMA and 6% in PAFMJ. In contrast, an increase of 7% and 8% was noted for original articles in PJMR and PJMS,

Table-2: Comparisons of design of studies between indexed and non-indexed journal.

Design	1998			2007			1998	2007	p-value
	Indexed journals	Non-indexed journals	p-value	Indexed journals	Non-indexed journals	p-value			
Cross section	5(13.2)	5(12.8)		107(54.6)	50(48.5)		10(13.0)	157(52.5)	
Retrospective	14(36.8)	13(33.3)		20 (10.3)	19(18.4)		27(35.1)	39(13.1)	
Prospective	14(36.8)	15(39.5)	0.999	29(14.9)	14(13.6)	0.399	29(38.2)	43(14.4)	<0.0001
Clinical trial	3(7.9)	3(7.9)		37(18.9)	19(18.4)		6(7.9)	56(18.8)	
Basic sciences	2(5.3)	2(5.1)		2(1.0)	1(1.0)		4(5.2)	3(1.0)	
Total	38	38		195	103		76	298	
Study design mentioned	37(21.6)	43(42.6)	<0.0001	196(65.6)	103(58.2)	0.108	80(29.4)	299(62.8)	<0.0001
Correct study design	22(59.5)	14(32.6)	0.016	146(74.5)	73(70.9)	0.502	36(45.0)	219(73.2)	<0.0001

Table-3: Comparison of statistical methods applied in indexed and non-indexed journals.

Statistical method	1998			2007			Comparison of cumulative percentages of statistical methods	
	Indexed journals (n = 172)	Non-indexed journals (n = 101)	p-value	Indexed journals (n = 299)	Non-indexed journals (n = 177)	p-value	1998 (n = 273)	2007 (n = 476)
No statistical methods or descriptive statistics only	129 (75)	73 (72.3)	0.621	145 (48.5)	94 (53.1)	0.331	74.0	50.4
P-value without mentioning the test	11 (6.4)	16 (15.8)	0.012	18(6.0)	19(10.7)	0.063	83.9	58.2
t-test	16(9.3)	9 (8.9)	0.914	61 (20.4)	34 (19.2)	0.753	91.6	65.3
Contingency tables	15(8.7)	3 (3.0)	0.065	75 (25.1)	35 (19.8)	0.216	97.4	83.0
Analysis of Variance	1 (0.6)	0 (0.0)	1.000*	14 (4.7)	11 (6.2)	0.469	97.8	87.6
Pearson Correlation	0 (0.0)	1 (1.0)	0.370*	12(4.0)	0 (0.0)	0.005*	98.2	89.0
Epidemiological studies	1 (0.6)	0 (0.0)	1.000*	13 (4.3)	3 (1. 7)	0.121	98.5	92.6
Simple linear regression	1 (0.6)	0 (0.0)	1.000*	8 (2.7)	3 (1. 7)	0.754	98.9	94.3
Multiple regression	1 (0.6)	0 (0.0)	1.000*	7 (2.3)	1 (0.6)	0.268*	99.2	96.0
Nonparametric statistics	2 (1.2)	0 (0.0)	0.532*	13 (4.3)	5 (2.8)	0.465	100	99.6
Non-Para metric correlation	0 (0.0)	0 (0.0)	---	2 (0.7)	0 (0.0)	0.532*	100	100
Statistics used	43 (25.0)	28 (27.7)	0.621	154 (51.5)	83 (46.9)	0.331		
Statistics appropriate	34 (79.1)	26 (92.9)	0.117	97 (63.0)	68(81.9)	0.002		

Chi-square tests were employed in all places except at with * where *Fisher Exact test was used.

respectively. However, these changes were not statistically significant (p>0.05) (Table-1).

Seven hundred forty nine articles were categorized as 'original articles' or 'short communication'. Fifty one percent of these articles mentioned the study design and 67.3% of them were correct for the respective methodology. In 1998, the study design mentioned by the non-indexed journals was significantly higher as compared to 'indexed' journals (p<0.0001). However, identification of correct study design was significantly higher in 'indexed' journals. No statistical significance was observed between 'indexed' and 'non-indexed' journals for identifying the correct study design in 2007 publications. When comparing the articles of 1998 to 2007 for correct study design, the percentages became more than doubled during this period. This difference was highly statistically significant (p<0.0001). The cross-sectional studies were increased by four-folds from 1998 to 2007, but retrospective and prospective studies were decreased by less than half during this period (Table-2). If the clinical trial studies are also considered as prospective studies, the difference of 'prospective studies' between 1998 and 2007 is decreased.

Two hundred seventy three articles of 1998 were categorized as 'original articles' or 'short communications'. Out of these 172 were published in 'indexed' journals and 101 were published in 'non-indexed' journals. Only 25% of the 'indexed' journals and 27.7% of 'non-indexed' journals computed some inferential statistics for their results. Six percent of indexed journals and 15.8 percent of non-indexed journals stated the 'p' value, without mentioning the test statistics. The difference was statistically significant (p = 0.012). Four hundred seventy six articles of 2007 were categorized as 'original articles' or 'short communication'. Out of these, 299 were published in 'indexed' journals and 177 were published in 'non-indexed' journals. None of the statistical techniques, including the indication of 'p' values without mentioning the test statistic was statistically significant, except 'Pearson correlation' which was employed significantly more frequently among 'indexed' journals than 'non-indexed' journals. The comparison showed that 71 (26%) articles of 1998 and 237 (49.8%) articles of 2007 used some

statistical techniques to justify their findings ($p<0.0001$). In 1998, 202 (74%) articles did not use any statistics or indicated only descriptive statistics, while in 2007,239 (50.2%) articles did the same ($p<0.0001$). Nevertheless, 84.5% of the articles used the statistics appropriately in 1998 issues, while only 69.6% articles of 2007 used correct statistics. This difference was also statistically significant ($p = 0.013$). 'T' tests, contingency tables, ANOVA, Pearson correlation, epidemiological statistics and non-parametric statistics were used significantly more in 2007 articles as compared to articles published in 1998 ($p < 0.05$). Indicating the 'p' values without mentioning the statistical tests was less frequent in 2007, as compared to 1998, but the difference was not statistically significant ($p = 0.319$). The last two columns showed the cumulative percentages for minimum statistical methods used in the reviewed articles. The minimum statistical methods consider the fact that many articles used more than one statistical method. To understand this phenomenon, consider the first three items of 1998 data: no statistical methods or descriptive statistics only, and p-value without mentioning the test and t-test. Since the item 'p-value without mentioning the test' is disjoint with any other mentioned statistical test, therefore the cumulative percentage for item 1 and 2 will be 83.9% (74.0% + 9.9%). However, in the articles in which the third item i.e. t-test' has been used; other statistics has also been applied. Therefore, the cumulative percentage could be less than the sum of all the three items, i.e. 74.0% + 9.9% + 9.5% = 93.4%. The cumulative percentage is only 91.6%. Therefore, only 7.7% (91.6 - 83.9) articles have used only t- test, without using any other statistics. Similarly, 97.4% of the articles have used 'no-statistical methods or descriptive statistics, p-value without mentioning the test, t-test and contingency tables, without using any other tests. Here, we see that the reader who was familiar with t-test and contingency tables in 1998 could have understood fully 97.4% of the scientific articles of these 6 journals. However, this percentage was dropped to 83.0% in 2007.

Discussion

An exponential growth can be seen in the usage of statistical techniques in medical articles in last few years. Furthermore, new methods have also been introduced by the investigators to make proper inferences. Reasons for this remarkable increase could be; (i) researchers became conscious that analytical approaches to analyze their data have significant effect in publications;[13] (ii) the request of editors and referees of the journals to include statistical analysis in the results have been increased;[22,23] (iii) Availability of menu-driven and friendly statistical soft-wares to the authors, even not very trained and fragile statistical knowledge, for easy computations of complex statistical techniques for their data;[24] and (iv) unintentional publication bias of accepting articles with statistical significant results, forcing the authors to use statistical methods, whether right or wrong, and get 'p' value lower than significant level.[25]

Number of articles, published during this period became almost doubled in the studied six journals. The forces behind this encouraging development are: creation of many new medical universities, new and enhanced criterions for promotions of academic staff and encouraging steps taken by Higher Education Commission to broaden the research environment.

Surprisingly, in 1998, only one-fifth articles mentioned the correct study design in indexed journals. It was significantly lower than the non-indexed journals. However, this weak spot of indexed journals has been removed in 2007. Three-fourth of the articles mentioned the correct study designs in 2007. It shows that researchers are becoming more knowledgeable in epidemiology and research methodology. Many continuing education lectures and workshops that are being conducted by College of Physicians and Surgeons and other institutions could be the positive effects in this regards. Comparing with the medical journals of ten years ago of other countries,[19] this study showed that specification of study design was significantly lower than other medical journals. When comparing different types of design with other studies, the results of this study showed that cross sectional studies were much lower than with other medical journals[7,14,15,19] ten years ago. However, after ten years by 2007, the percentage of this type of studies is comparable with the above mentioned cited references.

The problem of mentioning p-value without mentioning the statistical test is more common in non-indexed journals. It shows that the authors, reviewers and editors of indexed journals are more conscious of committing this mistake. The articles of indexed journals used little more complex statistics than only 't' test or ANOVA. However, the appropriateness of statistical tests used for that particular type of data is more accurate in non-indexed journals. This shows that statistical consultants of indexed journals might not be consulted for all the articles which have some statistics.

In the last ten years, the uses of statistics other than basic descriptive statistics in the study journals have been increased by 32%, which is a significant jump. However, it is still very low as compared to many other foreign journals shown in other studies.[1,5,10,12,14,15,18,19] The reason could be that almost half of the studies are descriptive in nature. Nevertheless, few recent articles[9,10,16] are in agreement with this result of 50% articles with 'no statistics' or only 'descriptive statistics'. A study[21] considering the Polish journals also agreed with this result. Comparing other studies the use of statistics other than descriptive statistics in Pakistani medical journals can be considered 'average'. Majority of articles used't' test and the tests related to 'contingency tables'

(Chi-square, Fisher Exact test, McNemar etc.). Surprisingly, some other studies[15,19] also show 23% use of test of statistics related to contingency tables. Analysis of Variance, Non-parametric statistics and epidemiological analysis (odd ratio, relative risk etc) are also being used in recent medical articles. Most of other studies[5,12,15,18,20] showed that these are the common statistics being used in medical journals with higher percentage than Pakistani journals. Logistic regression, which is now commonly used in the journals of developed countries,[15] is still not being used by the Pakistani researchers. The problem of mentioning 'p' values without mentioning the statistical test is still a serious problem in Pakistani journals. Every 12th original article showed this deficiency. Three most commonly used statistics are 'contingency tables', 't-test' and Analysis of variance'. If a reader understands these three concepts of statistics, s/he could be able to read 87.6% of medical articles in 2007; however, this figure was 97.8% ten years ago. Emerson and Colditz[9] showed that the 73% of the readers who know 't' statistic and 'contingency tables' can understand 'The New England Journal of Medicine'. Two other articles[13,20] showed that familiarities of these two types of statistics will be enough to access about 60% of articles. However, the present study showed that 83% of the readers can understand the articles of 6 Pakistani leading journals by knowing these two types of statistics. Therefore, the knowledge of a few basic statistics techniques, t-test, chi-square and analysis of variance are still enough to understand about 88% of the articles of Pakistani medical journals. Nevertheless, with the introduction of systematic review, Cochrane methods and Meta analysis, advance techniques of biostatistics are also needed to understand these types of articles. Therefore, there is a need of teaching biostatistics in the education of doctors who wish to undertake research. The greater emphasis should be given to the statistics concerning dichotomous data (chi-square, logistic regression, odd ratio etc) along with methods concerned with continuous data, like multivariate regression, analysis of co-variance etc. Initial training in descriptive statistics should be replaced by the methods mentioned above along with non-parametric statistics. Since there is a demand of using some statistics by the editors and enhanced chances for acceptance of the articles in leading journals with some statistics has been increased in recent years, it forced the authors to do the same. However, due to shortage of medical statisticians, the appropriate statistics are not used in the articles.

This study covered the articles of 6 leading Pakistani medical journals and shows the improvement that took place in the use of statistics. However, the results should be read with caution due to these being the top most journals among 62 medical journals published in Pakistan. It is likely that these results are better than the average level of Pakistani medical journals.

Conclusions

Quality of elaborating methods and usage of biostatistics in 6 leading Pakistani medical journals has gradually improved from 1998 to 2007, but they are still quite behind other medical journals of the western countries. Inappropriate use of statistics remains a problem. The major problems of omitting description of statistical methods and giving 'p' values without mentioning the statistical procedure suggest that the authors of medical articles need to enhance their knowledge in biostatistics and epidemiology to improve the quality of their articles.

Acknowledgement

The authors are grateful to Ms. Salma Arzoo Khattak for data entry and statistical work. We are also thankful to Ms. Sumaira Naz for secretarial help.

References

1. Rosenfeld RM, Rockette HE. Biostatistics in otolaryngology Journals. Arch Otolaryngal Head Neck Surg 1991; 117: 1172-6.
2. Horton NJ, Switzer SS. Statistical methods in the journal. N Engl J Med 2005; 353: 1977-9.
3. West CP, Ficalora RD. Clinician attitude towards biostatistics. MayoClin Proc 2007; 82: 939-43.
4. Khan N, Mumtaz Y. Attitude of teaching faculty towards statistics at a medical university in Karachi, Pakistan. J Ayub Med Coll 2009; 21: 166-7.
5. Kurichi JE, Sonnad SS. Statistical methods in the surgical literature. J Am Coll Surg 2006; 202: 476-84.
6. Goodman SN, Altman DG, George SL. Statistical reviewing policies of medical Journals: caveat lector? J Gen Intern Med 1998; 13: 753-6.
7. Mora Ripoll R, Ascaso Terren C, Senttis Vilalta J. [Current use of statistics in biomedical research: a comparison of general medicine journals.] Med Clin (Barc) 1996; 106: 451-6.
8. Kirisci L, Tarter RE, Moss H. Statistical approaches in alcohol research: a comparative survey of two major alcohol journals with four major Psychiatric journals. Alcohol Clin Exp Res 1994; 18: 183-6.
9. Emerson JD, Colditz GA. Use of statistical analysis in the New England Journal of Medicine. N Engl J Med 1983; 309: 709-13
10. Hammar AS, Buffington CA. Survey of statistical methods used in the veterinary medical literature. J Am Vet Med Assoc 1994; 205: 344-5.
11. Becker PJ, Viljoen E, Wolmarans L, Ijsselmuden CB. An assessment of the statistical procedures used in original papers published in the SAMJ during 1992. S Afr Med J 1995; 85: 881-4.
12. Silva ALC, Perez NC, Cuellar WI. The use of statistical methods in 2 medical journals with high impact factors. Gac Sanit 1995; 48: 189-95.
13. Elster AD. Use of statistical analysis in the AJR and Radiology: frequency, methods, and subspecialty differences. Am J Roentgenal 1994; 163: 711-5.
14. Scales CD Jr, Norris RD, Peterson BL, Preminger GM, Dahm P. Clinical research and statistical methods in the urology literature. J Urol 2005; 174: 1374-9.
15. Rigby AS, Armstrong GK, Campbell MJ, Summerton N. A survey of statistics in three UK general practice journal. BMC Med Res Methodol 2004; 4: 28-34.
16. Goldin J, Zhu W, Sayre JW. A review of statistical analysis used in papers published in Clinical Radiology and British Journal of Radiology. Clin Radiol 1996; 51: 47-50.
17. Fuki T, Rahman M, Sekimoto M, Hira K, Maeda K, Morimoto T, et al. Study design, statistical method, and level of evidence in Japanese and American clinical journals. J Epidemiol 2002; 12: 266-70.
18. Wang Q, Zhang B. Research design and statistical methods in Chinese medical journals. JAMA 1998; 280: 283-5.
19. Taddese Z. A survey of statistical methodology used in Ethiopian health science research journals. Ethiop Med J 2001; 39: 229-39.
20. Wang Y, Yu JM, Geng JP, Zheng JW. [A review of the statistical methods used in

the journal of prosthetic dentistry.] Shanghai Kou Qiang Yi Xue 2004; 13: 201-2.

21. Sych Z. [Evaluation of using statistical methods in selected national medical journals.] Ann Acad Med Stetin 1996; 42: 67-84.

22. Squires BP. Statistics in biomedical manuscripts: what editors want from authors and peer reviewers. CMAJ 1990; 142: 213-4.

23. Gardner MJ, Machin D, Campbell MJ. Use of check lists in assessing the statistical content of medical studies. Br Med J 1986; 292: 810-2.

24. Rothman KJ. Significance questing. Ann Intern Med 1986; 105: 445-7.

25. Sterling TD. Publications decisions and their possible effects on inferences drawn for tests of significance-or vice versa. J Am Stat Assoc 1959; 54: 30-4.