

## Analysis of Statistical Methods and Errors in the Articles Published in the *Korean Journal of Pain*

Department of Anesthesiology and Pain Medicine, Seoul National University Bundang Hospital, Seongnam,  
\*Seoul National University Hospital, Seoul, Korea

Kyoung Hoon Yim, MD, Francis Sahngun Nahm, MD, Kyoung Ah Han, MD\*,  
and Soo Young Park, MD\*

### Background:

Statistical analysis is essential in regard to obtaining objective reliability for medical research. However, medical researchers do not have enough statistical knowledge to properly analyze their study data. To help understand and potentially alleviate this problem, we have analyzed the statistical methods and errors of articles published in the *Korean Journal of Pain (KJP)*, with the intention to improve the statistical quality of the journal.

### Methods:

All the articles, except case reports and editorials, published from 2004 to 2008 in the *KJP* were reviewed. The types of applied statistical methods and errors in the articles were evaluated.

### Results:

One hundred and thirty-nine original articles were reviewed. Inferential statistics and descriptive statistics were used in 119 papers and 20 papers, respectively. Only 20.9% of the papers were free from statistical errors. The most commonly adopted statistical method was the t-test (21.0%) followed by the chi-square test (15.9%). Errors of omission were encountered 101 times in 70 papers. Among the errors of omission, "no statistics used even though statistical methods were required" was the most common (40.6%). The errors of commission were encountered 165 times in 86 papers, among which "parametric inference for nonparametric data" was the most common (33.9%).

### Conclusions:

We found various types of statistical errors in the articles published in the *KJP*. This suggests that meticulous attention should be given not only in the applying statistical procedures but also in the reviewing process to improve the value of the article. (*Korean J Pain* 2010; 23: 35-41)

### Key Words:

data interpretation, statistical analysis, statistics.

Received October 16, 2009. Revised October 26, 2009. Accepted November 11, 2009.

Correspondence to: Francis Sahngun Nahm, MD

Department of Anesthesiology and Pain Medicine, Seoul National University Bundang Hospital, 166, Gumi-ro, Bundang-gu, Seongnam 463-707, Korea

Tel: +82-31-787-7499, Fax: +82-31-787-4063, E-mail: hiitsme@hanmail.net

© This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © The Korean Pain Society, 2010

## INTRODUCTION

The statistical method of analysis is to collect, arrange, and draw general regularity from data; this is recognized as the most fundamental and universal method to prove soundness of conclusions in all scientific research. Undoubtedly, the statistical method of analysis is important in that the final purpose of medical research is clinical application, since inappropriate statistical techniques may deteriorate the quality of research articles or cause decisive error, thus leading to wrong treatment.

Although rapid progress of the computer programs for statistical analysis in recent times has allowed for convenience in analyzing data, there has been an increased danger in obtaining the wrong results from statistical analysis or misinterpreting the analyzed results if correct understanding of fundamental statistical concepts is lacking [1].

Although many articles have been published since the first issue of the *Korean Journal of Pain (KJP)* was published in 1988, and the journal was registered with the National Research Foundation of Korea in 2004, a research paper with respect to the statistical method applied to each article and its statistical errors has been unavailable. Therefore, we have analyzed the statistical techniques and errors in all the articles published in the *KJP*, covering articles from the first issue of volume 17 in 2004, when the journal was registered with the National Research Foundation of Korea, to the second issue of volume 21 in 2008, with the objective of encouraging a more appropriate use of statistical techniques in order to contribute to quality improvement of the journal for the future.

## MATERIALS AND METHODS

Among the 296 articles published in the *KJP*, from the first issue of volume 17 in 2004 to the second issue of volume 21 in 2008, excluding 22 editorials and 131 case reports, the targets of this study were 139 articles, out of the 143 original articles, in which statistical analyses were used.

For the cases where only descriptive statistics was used, the number of those articles was counted, while the types and frequency of the used statistical methods were analyzed in the cases where inferential statistics was applied. The validity of the statistical method in each article was evaluated by using the revised Checklist for

Assessing the Methodological and Statistical Validity of Medical Articles (Table 1) [2]. On the checklist, items such as the type of study, type of applied statistical method, and validity of applying the statistical method were included. The item regarding the validity of applying the statistical method was divided into 2 categories; "errors of omission" and "errors of commission". "Errors of omission," were caused by an insufficient report of the analysis procedure and data by the researcher, and "errors of commission," were caused by statistical mistreatment. In the "errors of omission," items included the following: ① Incomplete description of basic data, ② Incomplete description of applied statistical methods, ③ No statistics were used even though statistical methods were required, and ④ No evidence that described statistical methods was used. In the "errors of commission," items included the following: ① Inadequate description of measures of central tendency or dispersion, ② Incorrect analysis, and ③ Unwarranted conclusion.

The statistics checklist for individual articles was filled by statistics professionals and pain medical specialists together. If more than one statistical method was used in one article, the number of times was added to the calculation individually. If there were different statistical errors in one article, each error count was added up, while only one time was added to the calculation if the same error was repeated more than once in one article.

The completed checklists were statistically analyzed with SPSS statistics version 17.0 (SPSS Inc., Chicago, USA) to derive the frequency and percentage of each item.

## RESULTS

A total of 20 (14.4%) articles out of the 139 articles employed only descriptive statistics; and inferential statistics was used in 119 (85.6%) articles (Table 2). The inferential statistics was used 252 times in the 119 articles, among which the t-test was the most frequently used at 53 times (21.0%), followed by the  $\chi^2$  test at 40 times (15.9%), the analysis of variance (ANOVA) at 25 times (9.9%), the Mann-Whitney U test at 23 times (9.1%), and the paired t-test at 22 times (8.7%). The distribution of each of the applied statistical methods is shown in Table 3.

Out of the 139 target articles, 29 (20.9%) articles were free from statistical errors. From the 110 (79.1%) articles where the statistical analysis was inappropriately applied,

**Table 1.** Checklist for Assessing the Statistical Errors of the Articles in the *Korean Journal of Pain*


---

1. Types of the article	
① Case report	② Review article
	③ Original article
2. Types of the statistics used in the article	
① Not used	② Descriptive statistics
	③ Inferential statistics
3. Types of the inferential statistics used in the article	
① $\chi^2$ test	② Fisher's exact test
③ T-test	④ Paired t-test
⑤ ANOVA (analysis of variance)	⑥ Repeated measure of ANOVA
⑦ Mann-Whitney U test (Wilcoxon rank sum test)	⑧ Kruskal-Wallis test
⑨ Wilcoxon signed rank test	⑩ Pearson's correlation analysis
⑪ Spearman's correlation analysis	⑫ Linear regression
⑬ Logistic regression	⑭ Survival analysis
⑮ Others ( )	
4. Validity of applying the statistical method	
1) Errors of omission	
① Incomplete description of basic data	
② Incomplete description of applied statistical methods	
③ No statistics were used even though statistical methods were required	
④ No evidence that described statistical methods were used	
2) Errors of commission	
① Inadequate description of measures of central tendency or dispersion	
② Incorrect analysis	
a. Repeatedly used t-test without correction	
b. T-test without considering variable independency: paired or unpaired	
c. Chi-square test on the data with inappropriate sample size	
d. Analysis of variance without appropriate multiple comparison	
e. Parametric inference for nonparametric data	
f. Failure to select proper statistical methods	
: Comparison of incidence was dealt with comparison of mean	
: Comparison of 2 independent means was dealt with correlation analysis	
③ Unwarranted conclusion	
a. Illogical conclusion	
b. Incorrect use of statistical terminology	
c. Extrapolation of the study results	

---

**Table 2.** Statistical Methods Used in the *Korean Journal of Pain*


---

Year	Interferential statistics	Descriptive statistics	No statistics	Total
2004	30	3	4	37
2005	22	7	0	29
2006	25	3	0	28
2007	21	5	0	26
2008	21	2	0	23
Total	119	20	4	143

---

Each number represents the number of articles.

the number of errors found according to the statistics checklist (Table 1) was 266 (2.4 time/article).

"Errors of omission" were found 101 times in 70 articles

(1.44 time/article). Among these, the most frequent error was "no statistics were used even though statistical methods were required" at 41 times (40.6%), followed by "incom-

**Table 3.** The Incidence of Inferential Statistics in the 119 Articles Published in the *Korean Journal of Pain*

Statistical method	2004	2005	2006	2007	2008	Total
T-test	9	13	14	9	8	53 (21.0%)
$\chi^2$ test	9	8	9	8	6	40 (15.9%)
One way ANOVA	5	7	4	5	4	25 (9.9%)
Mann-Whitney U test	3	4	7	3	6	23 (9.1%)
Paired t-test	7	4	3	3	5	22 (8.7%)
Repeated measures ANOVA	6	6	2	0	4	18 (7.1%)
Fisher's exact test	2	2	6	3	1	14 (5.6%)
Wilcoxon signed rank test	3	0	5	0	2	10 (4.0%)
Kruskal-Wallis test	2	1	1	2	3	9 (3.6%)
Pearson correlation analysis	2	0	0	5	0	7 (2.8%)
Survival analysis	0	0	5	0	0	5 (2.0%)
Spearman correlation analysis	0	0	1	2	0	3 (1.2%)
Linear regression	2	0	1	0	0	3 (1.2%)
Logistic regression	1	0	0	1	0	2 (0.8%)
Others	2	2	6	5	3	18 (7.1%)
Total	53	47	64	46	42	252 (100%)

ANOVA: analysis of variance. Each different kind of statistical technique in the same article was counted separately.

**Table 4.** The Incidence of Errors of Omission

	2004	2005	2006	2007	2008	Total
Incomplete description of basic data	7	4	4	3	4	22 (21.8%)
Incomplete description of applied statistical methods	7	3	4	3	7	24 (23.8%)
No statistics used even though statistical methods were required	5	10	7	9	10	41 (40.6%)
No evidence that described method were used	1	3	8	1	1	14 (13.9%)
Total	20	20	23	16	22	101 (100%)

Each different kind of error in the same article was counted separately. But two or more of the same kind of errors in an article were counted as one.

**Table 5.** The Incidence of Errors of Commission

	2004	2005	2006	2007	2008	Total
Inadequate description of measures of central tendency or dispersion	8	7	5	5	10	35 (21.2%)
Incorrect analysis						
Repeatedly used t-test without correction	1	2	2	1	1	7 (4.2%)
T-test without considering variable independency	0	4	3	2	0	9 (5.5%)
Chi-square test on the data with inappropriate sample size	5	5	3	6	5	24 (14.5%)
ANOVA without appropriate multiple comparison	2	4	1	6	2	15 (9.1%)
Parametric inference for nonparametric data	12	21	9	6	8	56 (33.9%)
Inappropriate method used with data characteristics	4	2	2	1	3	12 (7.3%)
Unwarranted conclusion						
Illogical conclusion	4	2	0	0	1	7 (4.2%)
Incorrect use of statistical terminology	0	0	0	0	0	0 (0.0%)
Extrapolation of study results	0	0	0	0	0	0 (0.0%)
Total	36	47	25	27	30	165 (100%)

ANOVA: analysis of variances. Each different kind of error in the same article was counted separately. But two or more of the same kind of errors in an article were counted as one.

plete description of applied statistical methods" at 24 times (23.8%) (Table 4).

"Errors of commission" were found 165 times in 86 articles (1.92 time/article). Among these, "inadequate description of measures of central tendency or dispersion" was registered at 35 times (21.2%), "incorrect analysis" at 123 times (74.5%), and "unwarranted conclusion" at 7 times (4.2%). Out of the 123 occasions where incorrect statistical analysis was used, "parametric inference for nonparametric data" was the most common error at 56 times (33.9%), followed by "chi-square test on the data with inappropriate sample size" at 24 times (14.5%) (Table 5).

## DISCUSSION

The importance of statistical analysis in medical research papers is ever increasing day by day, therefore, it can be said that evaluation of statistical validity in medical research articles is very important nowadays when evidence based medicine is highly valued.

Since 1990s, several academic societies in Korea have investigated the current status of statistics applied in the articles published in their journals [3–7], and the Checklist for Assessing the Methodological and Statistical Validity of Medical Articles [2], has been revised for individual academic societies with the intent of being used for the analysis of the articles [3,5,7]. Also in this study, we used the checklist [2], which has been used many times in previous studies, in analyzing the target articles as objectively as possible.

We could verify that many kinds of statistical methods have been used in the original papers published in the *KJP* from 2004 to 2008. The result that the t-test was the most frequently used statistical method in the *KJP* has no significant meaning other than the fact that it is due to the characteristic of the individual professional field, since alternatively, survival analysis is most frequently used in the *Journal of Korean Society for Therapeutic Radiology and Oncology* [3] and descriptive statistics are used in the *Korean Journal of Clinical Pathology* [4]. Thus, it is obvious that comparison of means and cross-tabulation analysis are most frequently used in the articles published in the *KJP*, in which many of the articles deal with therapeutic effects or complications.

The occasions where "no statistics were used even though statistical methods were required," was the most

frequent item in the "errors of omission," as 41 articles in this study were given these results. One representative example was where different concentrations of a drug were given to individual groups. Although only the difference in means was compared and analyzed for the difference between the effects on the individual groups, the conclusion that "the effect was proportional to the dose of the drug" was made and reported. In such a case, the analytical procedure which can prove the correlation between dose and effect should be carried out in order to make the correct conclusion. Another example is the case where the number of animals or the number of experimental targets was not clearly described; rather, it was described as "5–7 for each group" or "18–20 persons for each group," which is an incorrect description. In addition, cases of "no evidence that described methods were used" were found among the "errors of omission." If a statistical method was actually used, it should be explicitly mentioned. In addition, the statistical method used for each analysis should be precisely classified and described rather than listing the names of statistical methods.

A considerable number of occasions where "incorrect analysis" from the "errors of commission" were verified in this study, including following examples:

First, the errors found in high frequency were "parametric inference for nonparametric data" (33.9%) and "chi-square test on the data with inappropriate sample size" (14.5%). Furthermore, this result is particularly significant for the quality evaluation of not only the individual articles, but also the *KJP*, since the most frequently used statistical methods in the *KJP* are the t-test and chi-square test, as mentioned above. In order to reduce such errors, nonparametric statistics should be used or the data should be modified by means of data transformation for the cases where the number of observed data is small, or normal distribution cannot be assumed for the data [8]. Special caution should be taken lest the researcher form completely wrong conclusions by applying parametric inference for nonparametric data without such a process. For a chi-square test in cross-tabulation analysis, it is suggested that if the total sample size is not more than 20, or if more than 20% of the expected frequencies are less than 5, then it is not correct to apply the chi-square test directly; rather, Fisher's exact test must be applied [9].

Second, 32 examples where the experimental data were expressed as "mean  $\pm$  standard error" were found.

Standard error is used to estimate how much the mean value can be varied when repeated sampling of a different sample with an equal sample size was carried out from a population. Therefore, since standard error is to be used to estimate the distribution of a population mean, the data observed by the researcher must be expressed in the form of "mean  $\pm$  standard deviation" rather than "mean  $\pm$  standard error" [10].

Third, when comparing the means of three groups or more, although it is necessary to show that there is a group with a different mean by post hoc analysis, the error to conclude that a specific group had a different mean without this process was found in 14 cases. The parametric statistical method used to compare the means of three different groups or more is a one way analysis of variance (ANOVA), of which the null hypothesis is: "The means of all the groups are equal ( $H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$ )."

Although it can test whether the means of all the groups are equal or not, it cannot specifically tell which groups have a difference in means among them. Therefore, if there is a difference between groups, it is necessary to check through post hoc analysis which specific group has the difference with others [11].

Fourth, there were three cases where the error of comparing the means on the categorical variables was made. For example, when measuring patients' satisfaction in the three classes of high, moderate, and low, it is not right that the researcher renders arbitrary scores for each class and compares the means, since the variable of patients' satisfaction is one of the categorical variables. In this case, the statistical method for the analysis of categorical variables must be applied. The most fundamental factor in statistical analysis is to understand the types of variables to analyze, because the analytical method to be used is dependent upon the type of scale.

Fifth, the dependence/independence of the variables to analyze is also important. One representative example is the test for paired samples. The paired t-test, which is frequently used in comparing the degree of pain before and after a treatment, should have the same sample size for the two groups since the t-test is supposed to compare the difference of the two dependent groups.

The fact that only 20.9% of the articles published in the *KJP* from 2004 to 2008 were free of statistical errors does not mean that only 20.9% of the articles are reliable, since we have simply counted the number of errors re-

gardless of the statistical errors that could decisively affect the interpretation of the result in each study.

Considering the analysis of the statistical errors found in other Korean journals, the proportion of the articles where no statistical error was found was 19.0% for the *Journal of the Korean Society for Therapeutic Radiology and Oncology* [3], 42.3% for the *Journal of Korean Society of Emergency Medicine* [5], and 33% for the *Journal of Korean Society of Plastic and Reconstructive Surgeons* [12]. In the case of the *Korean Journal of Anesthesiology*, the ratio was 3% in 1981, but it was reported that the proportion was increased to 33% in 1990. For the international journals, analysis of the statistical errors of the submitted manuscripts were performed in 1970-80s [13-17], and it was reported that 52% of the target articles included errors in applying statistics [16]. Even a serious case, where only 15% of the articles had no errors in applying statistics, has been reported [16]. To reduce such errors, some of the international journals provide statistics guidelines [18] or checklists [19] for data analysis.

Differentiating the major fatal errors from the minor statistical errors is important, since the former may raise serious questions regarding the validity and reliability of the study. Caution should be paid to avoid making the mistake of devaluating significant academic achievements by taking all types of statistical errors overly seriously, and thus, exaggerating minor errors to an unnecessary extent. Hence, this should be remembered in the reviewing process.

Although statistics has a vast range, only several statistical methods are employed for medical articles. Based on the analysis of 1,828 medical articles, one can understand and interpret 70% of the medical articles if one precisely understands descriptive statistics, the t-test, the chi-square test, and Fisher's exact test [20]. As understood from the results of our study that the most frequently used analytical methods in the *KJP* were the t-test and the chi-square test, the common statistical errors found in medical articles were not because difficult and complicated analytical methods were mistakenly applied but because simple and easy methods were not correctly applied. Therefore, it is important to have a precise understanding of the statistical methods frequently used in each professional field, when having the full understanding of all the statistical methods is not possible.

In conclusion, we have found many statistical errors

in the articles published in the *KJP* and verified the fact that only 20.9% of the articles were without statistical errors. To elevate the quality of the *KJP*, as well as that of individual articles, the efforts of the researchers to make appropriate usage of statistical methods and to give appropriate attention during the reviewing process are required.

## REFERENCES

- Rosenbaum SH. Statistical methods in anesthesia. In: Miller's anesthesia. 7th ed. Edited by Miller RD. Philadelphia, Elsevier Inc. 2009, pp 3075–86.
- Ahn YO, Lee HK. Development of a checklist for assessing the methodological and statistical validity of medical articles. *Korean J Med Educ* 1991; 3: 19–35.
- Park HC, Choi DH, Ahn SV, Kang JO, Kim ES, Park W, et al. Statistical errors in papers published in the journal of the Korean society for therapeutic radiology and oncology. *J Korean Soc Ther Radiol Oncol* 2008; 26: 289–94.
- Park IK, Kang JO, Kim TY, Keum DG. An Evaluation of the statistical techniques used in the 1995–1996 editions of the Korean journal of clinical pathology. *Korean J Clin Pathol* 1999; 19: 460–4.
- Yu TY, Jin Y, Jeong TO, Lee JB. Statistical methods in the articles in the journal of the Korean society of emergency medicine published from 1998 to 2002. *J Korean Soc Emerg Med* 2003; 14: 475–80.
- Shin MH, Son YI, Cho YS, Kim HY, Ha BS, Kim SI. The use of statistical analysis–In the Korean journal of otolaryngology–Head and Neck Surgery–. *Korean J Otolaryngol–Head Neck Surg* 1999; 42: 1364–9.
- Ko H, Kwak IY, Choe IH, Ham BM, Kim KW. Statistical methods in the articles of the journal of the Korean society of anesthesiologists from 1981 to 1990. *Korean J Anesthesiol* 1993; 26: 22–7.
- Song K, Han M, Cheong MY, Lim KS, Kim DK. Clinical research design and biostatistical methods. *Korean J Urol* 2005; 46: 835–41.
- Kim BS, Lee YG, Lee TR. Categorical data analysis. Seoul, Korea National Open University Press. 2004, pp 32–5.
- Song KI, Choi JS. Clinical data analysis by SPSS 15–A practical guide for clinicians. Seoul, Hannarae publishing Co. 2008, pp 31–3, 55–61.
- Ahn JU, Yoo KY. Statistical analysis in medicine and public health. 2nd ed. Seoul, Hannarae publishing Co. 2006, pp 311–78.
- Hwang K, Lee HJ, Kim YJ, Lee SI. Statistical errors in papers in the journal of Korean society of plastic and reconstructive surgeons (JKPRS). *J Korean Soc Plast Reconstr Surg* 2001; 28: 302–9.
- White SJ. Statistical errors in papers in the British journal of psychiatry. *Br J Psychiatry* 1979; 135: 336–42.
- MacArthur RD, Jackson GG. An evaluation of the use of statistical methodology in the journal of infectious diseases. *J Infect Dis* 1984; 149: 349–54.
- Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med* 1987; 317: 426–32.
- Gore SM, Jones IG, Rytter EC. Misuse of statistical methods: critical assessment of articles in *BMJ* from January to March 1976. *Br Med J* 1977; 1: 85–7.
- Avram MJ, Shanks CA, Dykes MH, 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.
- Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. *Br Med J (Clin Res Ed)* 1983; 286: 1489–93.
- Gardner MJ, Machin D, Campbell MJ. Use of check lists in assessing the statistical content of medical studies. *Br Med J (Clin Res Ed)* 1986; 292: 810–2.
- Reed JF 3rd, Salen P, Bagher P. Methodological and statistical techniques: what do residents really need to know about statistics? *J Med Syst* 2003; 27: 233–8.