Statistical Techniques Reported in Korean Dermatology Journals During 1990-1994

Sung Yul Lee, M.D., Jong Suk Lee, M.D.,
Jae Young Hwang, M.D., Chong Hyeok Kim*, M.D.

*Department of Dermatology, Chunan Soonchunhyang General Hospital, Soonchunhyang University College of Medicine Chungnam, Korea, Medical Corps**

Background **: Effective evaluation of an article's scientific merit requires familiarity with the methodology described, especially when quantitative techniques, such as statistical procedures, are invoked to clarify research findings or to summarize data.**

Objective **: The purpose of this study was to describe the frequency with which various statistical concepts were reported in journals important to dermatology. From these results, dermatologists can identify the major statistical skills needed to critically evaluate their literature.**

Methods **: All 376 original articles of the Korean Journal of Dermatology(KJD) and the Annals of Dermatology(KAD) during 1990-1994 were chosen for review. Each article was reviewed to determine the statistical content.**

Results **: The two most commonly used statistical techniques in the two journals were mean and standard deviation. The t-test was the next most frequently used statistical technique in the reviewed journals, followed by non-parametric, chi-square test, orphan p, ANOVA, and correlation/regression.**

Conclusion **: These results indicate the need for wider education about the use of descriptive and basic comparative statistics. It is impossible to evaluate the dermatological literature critically without these skills. (Ann Dermatol 8:(4)243~246, 1996).**

Key Words : Statistical techniques, Annals of Dermatology, Korean Journal of Dermatology

All physicians face the challenge of keeping abreast of a body of biological knowledge that is expanding at an astonishing rate. The medical journal is the primary channel for disseminating medical information to the physician[1]. The practitioner who is determined to maintain cognitive skills is already forced to choose those items which he will take notice of and those he will ignore. Subsequently, effective evaluation of an article's scientific merit requires familiarity with the methodology described, especially when quantitative techniques, such as statistical procedures, are invoked to clarify research findings or to summarize data[2].

Academic dermatologists must decide which quantitative methods to teach their trainees and how to incorporate teaching of these concepts into crowded training programs[3]. Identification of the statistical techniques used in major dermatology journals may provide one basis for these judgements. The purpose of this study was to describe the frequency with which various statistical concepts were reported in journals important to dermatology. From these results, dermatologists can identify the major statistical skills needed to critically evaluate their literature.

Received July 11, 1995.
Accepted for publication September 3, 1996.
Reprint request to : Sung Yul Lee, M.D., Department of Dermatology, Chunan Soonchunhyang General Hospital, Soonchunhyang University College of Medicine Chungnam, Korea, Medical Corps

Table 1. The use of analytical statistics in two Korean dermatology journals

	Korean Journal of Dermatology	Annals of Dermatology
Number of articles reviewed	330(100%)	46(100%)
Number with no statistics	171(51.8%)	21(45.7%)
Number of statistical procedures used	159(48.2%)	25(54.3%)

MATERIALS AND METHODS

All 376 original articles of the Korean Journal of Dermatology(KJD) and the Annals of Dermatology(KAD) during 1990-1994 were chosen for review(Table 1.). Each article was reviewed to determine the statistical content. The categories used to assess statistical content(Table 2 - 4) were modified from an article by Oliver and Hall[3]. Each statistical procedure was scored only once per article regardless of the number of times reported. Techniques were also scored even if they were mentioned negatively. No attempts were made to evaluate the appropriateness of the usage of specific statistical tests. Since it was not the aim of this study to analyze the relative statistical sophistication of journals, only reported frequencies of statistical use were tabulated(Table 2 - 4).

RESULTS

About half the articles(184 of 376) mentioned the use of one or more statistical techniques(Table 1.). Table 2. - 4. show the relative frequency with which each statistical method was used in the two dermatology journals reviewed. The two most commonly used statistical techniques in two journals were mean and standard deviation. The t-test was the next most frequently used statistical technique in the reviewed journals, followed by non-parametric, chi-square test, orphan p, ANOVA, and correlation/regression. Table 2.- 4. suggest that even though the frequency of use varies from journal to journal, a reader familiar with 8 statistical procedure(mean, standard deviation, t-test, non-parametric, chi-square test, orphan p, ANOVA, and correlation/regression) will be familiar with over 95% of the quantitative techniques reported in the dermatology journals surveyed.

DISCUSSION

This study demonstrates that a knowledge of simple descriptive statistics and the basic comparative statistics(t-test, non-parametric, chi-square test, ANOVA, correlation/regression) provides access to 95% of the analytical articles that are published in the dermatological literature. It is impossible for dermatological trainees to evaluate critically the dermatological literature unless they are conversant with these basic statistics3. There is a limit to the extent that colleges, and journals, can take remedial action to overcome deficits in the basic medical education. There is a need for a course in statistical comprehension during dermatological residency training[3].

Table 2. Usage of statistics in 184 original articles (KJD and KAD)

Statistics	No.	%
Descriptive statistics		
Central tendency(location)		
mean	118	64.1
median	0	0
mode	0	0
Dispersion		
standard deviation	105	57.1
range	4	2.2
standard error of mean	9	4.9
interquartile percentile	0	0
Comparative statistics		
t-test or analogue	78	42.4
Non-parametric	42	22.8
Chi-square	31	16.8
Analysis of variance	16	8.7
'Orphan P'	28	15.2
Correlation/regression	17	9.2
Survival analysis	0	0
Others	7	4.3

Table 3. Usage of statistics in 159 original articles of Korean Journal of Dermatology

Statistics	No.	%
Descriptive statistics		
Central tendency(location)		
mean	104	65.4
median	0	0
mode	0	0
Dispersion		
standard deviation	93	58.5
range	4	2.5
standard error of mean	7	4.4
interquartile percentile	0	0
Comparative statistics		
t-test or analogue	69	43.4
Non-parametric	35	22.0
Chi-square	28	17.6
Analysis of variance	15	9.4
'Orphan P'	25	15.7
Correlation/regression	15	9.4
Survival analysis	0	0
Others	7	4.4

Table 4. Usage of statistics in 25 original articles of Annals of Dermatology

Statistics	No.	%
Descriptive statistics		
Central tendency(location)		
mean	14	56
median	0	0
mode	0	0
Dispersion		
standard deviation	12	48
range	0	0
standard error of mean	2	8
interquartile percentile	0	0
Comparative statistics		
t-test or analogue	9	36
Non-parametric	7	28
Chi-square	3	12
Analysis of variance	1	4
'Orphan P'	3	12
Correlation/regression	2	8
Survival analysis	0	0
Others	1	4

The primary purpose of our study was not to report statistical errors in the dermatology literature; however, we noted two questionable practices. First, measures of dispersion were often described ambiguously either with the standard deviation(SD) and the standard error of mean(SEM) used interchangeably or with the identification of the dispersion parameter. The data cannot usually be given in full, and so are described by summary statistics. Commonly two statistics are quoted: a measure of location or ' center' of distribution of sample values, and a measure of dispersion or ' spread '[4]. With ordinal and skewed data, the median is a more informative measure of location than is the mean. Therefore, it is not appropriate in such cases to use means and standard deviations, as they give equal weight to equally spaced intervals. Again, to measure dispersion, a statistic such as the semiinterquartile range is appropriate. The standard deviation has the useful property that roughly 68% of the observations will be within 1 standard deviation of the mean and roughly 95% of the observations will be within 2 standard deviations of the mean[5]. This property makes the standard deviation a good way to summarize the variability in data with a single number. Some authors, however, fail to summarize their data with the standard deviation; they use the standard error of the mean. Unlike the standard deviation, the standard error of the mean does not summarize the variability in the observation or give the reader insight into the range of the observations. Thus, the standard error of the mean does not quantify variability in the observations, as the standard deviation does, but rather the precision with which a sample mean estimates the true population mean. That is, the standard error of the mean quantifies the certainty with which one can estimate the true population mean from the sample. Thus, the standard deviation, not the standard error of the mean, should be used to summarize data[5]. Why do some authors use the standard error of the mean to summarize their data? First, tradition; second, the standard error of the mean is always smaller than the standard deviation.

Second, we also noted the common omission of the identification of the methods used(orphan P) when a " P value " resulting from a statistical test of a hypothesis was reported[3]. The 15.7% incidence of orphan statistics found in this study is in accor-

dance with other reports. A review of analytical papers published in the Australian and New Zealand Journal of Surgery revealed that 13% of publications that used statistical comparisons failed to state the name of the test[6]. A recent review of 200 surgical publications reported an incidence of 12% for the orphan P[7]. We suggest that every test of statistical significance should be accompanied by an unambiguous identification of the techniques used to calculate this value. If statistical software is used for other than simple counts, the program and procedure used should be identified. We contend that authors should describe statistical analysis in sufficient detail so that readers can reproduce the calculations if the data were available[8].

In conclusion, these results indicate the need for wider education about the use of descriptive and basic comparative statistics. It is impossible to evaluate the dermatological literature critically without these skills.

REFERENCES

1. Hokanson JA, Stiernberg CM, McCracken MS et al: The reporting of statistical techniques in otolaryngology journals. Arch Otolaryngol Head Neck Surg 1987, 113; 45-50
2. Hokanson JA, Ladoulis CT, Quinn FB et al: Statistical techniques reported in pathology journals during 1983-1985. Arch Pathol Lab Med 1987, 111; 202-207
3. Olive D, Hall JC: Usage of statistics in the surgical literature and the ' orphan P ' phenomenon. Aust N Z J Surg 1989, 59; 449-451
4. White SJ: Statistical errors in papers in the British Journal of Psychiatry. Brit J Psychiat 1979. 135; 336-342
5. Glantz SA: Biostatistics: How to detect, correct and prevent errors in the medical literature. Circulation 1980, 61; 1-7
6. Hall JC, Hill D, Watts JMcK: Misuse of statistical methods in the Australian surgical literature. Aust N Z J Surg 1982, 52; 541-543
7. Reznick RK, Dawson-Saunders E, Foise JR: A rationale for the teaching of statistics to surgical residents. Surgery 1987, 101; 611-617
8. Elenbaas RM, Elenbaas JK, Cuddy PG: Evaluating the medical literature part 2: Statistical analysis . Ann Emerg Med 1983, 12; 610-620