

[결핵 및 호흡기질환] 게재 논문의 통계적 기법 활용에 대한 평가

¹성균관대학교 의과대학 내과학교실 삼성서울병원 호흡기내과, ²강원대학교 의과대학 내과학교실

³한림대학교 의과대학 내과학교실, ⁴한림대학교 의과대학 사회의학교실

고원중¹, 이승준², 강민중³, 이훈재⁴

An Evaluation of the Use of Statistical Methods in the Journal of Tuberculosis and Respiratory Diseases

Won-Jung Koh, M.D.¹, Seung-Joon Lee, M.D.², Min Jong Kang, M.D.³, Hun Jae Lee, M.D.⁴

¹Division of Pulmonary and Critical Care Medicine, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul

²Department of Internal Medicine, College of Medicine, Kangwon National University, Chuncheon, ³Department of Internal Medicine, College of Medicine, Hallym University, Seoul ⁴Department of Social and Preventive Medicine, College of Medicine, Hallym University, Chuncheon, Republic of Korea

Background : The statistical analysis is an essential procedure ensuring that the results of researches are based on evidences rather than opinion. The purpose of this study is to evaluate which statistical techniques are used and whether these statistical methods are used appropriately or not in the journal of Tuberculosis and Respiratory Diseases.

Materials and Methods : We reviewed 185 articles reported in the journal of Tuberculosis and Respiratory Diseases in 1999. We evaluated the validity of used statistical methods based upon the checklist that was developed on the basis of the guideline for statistical reporting in articles for medical journals by International Committee of Medical Journal Editors.

Results : Among 185 articles, original articles and case reports were 110 (59.5%) and 61 (33.0%) respectively. In 112 articles excluding case reports and reviews, statistical techniques were used in 107 articles (95.5%). In 94 articles (83.9%) descriptive and inferential methods were used, while in 13 (11.6%) articles only descriptive methods were used. With the types of inferential statistical techniques, comparison of means was most commonly used (64/94, 68.1%), followed by contingency table (43/94, 45.7%) and correlation or regression (18/94, 19.1%). Among the articles in which descriptive methods were used, 83.2% (89/107) showed inappropriate central tendency and dispersion. In the articles in which inferential methods were used, improper methods were applied in 88.8% (79/89) and the most frequent misuse of statistical methods was inappropriate use of parametric methods (35/89, 39.3%). Only 14 articles (13.1%) were satisfactory in utilization of statistical methodology.

Conclusion : Most of the statistical errors found in the journal were misuses of statistical methods related to basic statistics. This study suggests that researchers should be more careful when they describe and apply statistical methods and more extensive statistical refereeing system would be needed. (*Tuberc Respir Dis* 2004; 57:168-179)

Key words : Statistics, Periodicals, Medical Journalism, Pulmonary Disease, Korea.

서 론

환자를 직접 연구대상으로 하는 임상의학연구에서 연구의 목적을 효과적으로 달성하기 위하여는 연구설계를 타당하게 구상하여야 할 뿐 아니라, 연구 수행으

로 얻어진 자료로부터 합리적이고 객관적인 결론을 끌어낼 수 있도록 적절한 통계적 분석기법을 적용하여야 한다¹⁻³. 또한 최근에는 의학연구결과를 발표하고 보고하기 위해서뿐만 아니라 발표된 연구결과를 의학 학술지를 통해 접하는 독자 입장에서도 보고자가 사용한 연구방법 및 통계기법에 대한 지식이 없이는 보고자의 연구결과와 여기서 얻어진 결론을 효과적으로 평가하기 어렵게 되었고, 통계학은 의사들이 익혀야 할 필수적인 분야가 되었다고 할 수 있다¹⁻³.

따라서 통계적 기법의 잘못된 활용은 연구 논문에서 심각한 문제를 일으키게 된다. 이러한 문제는 외국 유명 의학학술지도 예외가 아닌 실정이며, 게재

Address for correspondence : **Seung-Joon Lee, M.D.**
Department of Internal Medicine
Kangwon National University Hospital
17-1 Hyoja-3-dong, Chuncheon-si
Kangwon-do 200-947, Republic of Korea.
Phone : (82) 33-258-2377 Fax : (82) 33-258-2455
E-mail: medfman@knuh.or.kr
Received : Nov. 10. 2003.
Accepted : Jun. 29. 2004.

Table 1. Uniform requirements for manuscripts submitted to biomedical journals: Statistics²³⁻²⁵

1. 연구 내용을 이해할 만한 독자가 연구 데이터를 확보했을 때 논문에 나타난 결과를 확인할 수 있을 정도로 통계학적 방법을 자세하게 기술하여야 한다.
2. 가능하면 연구 소견을 정량화하고 측정 오차 또는 불확실성의 정도(예를 들면 신뢰 구간 등)를 표시하는 지표를 같이 제시한다.
3. P-값 (P value)만 제시하는 경우와 같이 통계학적 가설 검정 결과에만 의존하여 중요한 정량적 정보를 전달하지 못하는 일은 피해야 한다.
4. 연구 대상의 적합성을 따져야 한다.
5. 무작위 추출이나 무작위 배정에 대한 세부 사항을 기록한다.
6. 결과 관찰 때의 맹검 방법을 기록하고 그 성공여부를 기록한다.
7. 치료의 합병증을 기록한다.
8. 관찰 회수를 기록한다.
9. 추적 관찰에 실패한 예의 수를 보고한다(임상 실험의 탈락자 수 등).
10. 연구 설계나 통계학적 방법에 대한 참고문헌은 통계 이론의 원전보다는 잘 설계되고 표준적으로 실행한 연구 보고를 (페이지와 함께) 인용하는 것이 좋다.
11. 일반적으로 사용하는 컴퓨터 프로그램도 명시한다.
12. 일반적으로 기술한 연구 방법은 “방법” 항목에서 기술한다. 자료를 “결과” 항목에서 요약한 경우 자료 분석에 사용한 통계방법을 같이 기록한다.
13. 표와 도표는 논문이 주장하는 논점을 설명하고 뒷받침하는데 필요한 것에 한정시킨다. 항목(entry)이 많은 경우에는 표보다 도표로 작성하도록 한다; 같은 자료를 표와 도표로 중복 제시하지 않는다.
14. “무작위”(random) (무작위화를 의미하는 말), “정상” (normal), “유의한”(significant), “상관성”(correlations), “표본”(sample)과 같은 기술적 전문 통계용어를 비전문적으로 사용하지 않는다.
15. 통계용어, 약어, 기호의 의미를 정의한다.

된 논문의 70-80% 이상에서 통계적 기법을 사용하고 있으며, 그 중 40-70%에서 통계적 오류가 발견된다고 한다⁴⁻⁸. 또한 국내 의학학술지를 대상으로 한 여러 연구에서도 게재된 논문의 30-90%가 통계적 기법을 사용하고 있으며⁹⁻²², 그 중 통계적 오류가 적게는 60-80%^{10,11,13,15,17,20}, 많게는 95-100%^{9,12,21}에 이르는 것으로 보고되고 있다.

대한결핵 및 호흡기학회 회원들이 가장 흔히 접하고 정보를 얻을 수 있는 출판물인 [결핵 및 호흡기질환]은 창간 후 현재까지 연구자와 독자간의 지식 교류의 장으로 큰 역할을 해 오고 있다. 본 연구는 [결핵 및 호흡기질환]에 게재된 논문에서 사용된 통계적 기법 활용 현황과 오류의 실태를 알아보고자 하였다.

대상 및 방법

1. 연구대상

본 연구는 1999년에 발간된 총 12권의 [결핵 및 호흡기질환]에 게재된 논문 185편을 대상으로 하였다. 주요 평가영역은 연구유형, 통계적 기법의 활용현황 그리고 통계학적 방법론의 타당성 평가 등이었다.

2. 점검표의 활용

국제의학학술지 편집위원회에서 1997년 제시한 “생의학 학술지에 투고하는 원고의 통일 양식(Uniform Requirements for Manuscripts Submitted to Biomedical Journals)” 개정 5판²³⁻²⁵의 통계적 서술원칙 15개항(Table 1) 중 점검항목으로 정형화하기가 곤란하거나 자료의 분석차원을 넘어 연구방법론 전반에 관련된 사항이라 할 수 있는 일부 내용(제4항-제9항)을 제외한 나머지 항목을 기본으로 반영하고, 국내외 관련연구에서 제시하고 있는 점검표 내용을 참고하여 이 등이 개발한 점검표²²¹를 사용하였다(Table 2).

3. 연구방법론 유형과 통계적 기법의 활용 현황

의학논문의 연구방법론에 따른 분류방법은 학자에 따라 다양하게 제시되고 있으나, 본 연구에서는 학회지 게재논문을 분류하는데 일반적인 기준이 되는 방법²⁶에 따라 종설, 원저, 증례보고, 편집인에게의 통신과 같은 기타로 구분하여 현황을 파악하였다.

통계적 기법의 활용현황은 기술통계학적 기법과 추측통계학적 기법의 적용 유무로 구분하였고, 추측통계학적 기법을 적용한 경우 그 종류를 분석하였다^{1,2}.

Table 2. Checklist for evaluation of statistical validity for the articles in the journal of Tuberculosis and Respiratory Diseases.

1. 연구종류는?
 ① 종설 () ② 원저 () ③ 증례 () ④ 기타 ()

2. 통계기법의 사용범위는?
 ① 사용안함 () ② 기술통계 () ③ 기술통계 + 추측통계 ()
 ④ 기타 (내용 :)

3. 적용한 추측(분석)통계기법의 종류는? (여러가지 통계적 기법을 사용한 경우는 복수표시)
 ① T-test, z-test () ② Wilcoxon rank sum test, Mann-Whitney U test ()
 ③ ANOVA test with multiple comparison () ④ Kruskal-Wallis test ()
 ⑤ paired t-test () ⑥ Wilcoxon signed rank test ()
 ⑦ repeated measures ANOVA test ()
 ⑧ Pearson chi-squared test () ⑨ Fisher's exact test ()
 ⑩ Mantel-Haenszel trend test ()
 ⑪ simple linear regression analysis () ⑫ Pearson correlation analysis ()
 ⑬ Spearman correlation analysis () ⑭ Sensitivity, Specificity, Predictive value ()
 ⑮ 다변량 분석 (종류 :)
 ○ Not defined () * 기타 (내용 :)

4. 통계적 기법 적용의 타당성
 4-1. 기술통계학적 기법 적용은 적절한가? ① 적절하다 () ② 부적절하다 ()
 4-2. 자료의 특성에 맞는 추측(분석)통계학적 기법을 선택하였나?
 ① 그렇다 () ② 아니다 ()
 * 잘못 적용된 기법은? (① , ② , ③) → 권고되는 기법의 종류는? (① , ② , ③)
 4-3. 추측통계학적 기법을 적용한 경우 통계기법의 적용과정에서 요구되는 통계적 절차 또는 가정을 고려하였는가?
 ① 그렇다 () ② 아니다 (내용 :)

5. 통계학적 방법론 및 결과 서술의 타당성
 5-1. 통계학적 방법론과 분석결과를 상세히 기술하였는가?
 ① 그렇다 () ② 아니다 ()
 5-2. 추측통계학적 기법을 사용한 경우 유의수준 또는 신뢰구간 구역을 제시하였는가?
 ① 그렇다 () ② 아니다 ()
 5-3. 검정을 시행한 경우 정확한 p-값을 제시하고 있는가? ① 그렇다 () ② 아니다 ()
 5-4. 적용했다고 기술한 각 통계적 기법의 결과가 모두 제시되었는가?
 ① 그렇다 () ② 아니다 ()
 5-5. 사용한 컴퓨터 프로그램을 명시하였는가?
 ① 그렇다 () ② 아니다 ()
 5-6. 표와 도표 등 보조적 표현수단을 적절히 사용되었는가?
 ① 그렇다 () ② 아니다 (내용 :)

6. 통계적 분석결과 해석의 타당성
 6-1. 검정을 시행한 경우 유의수준과 p-값 등에 따라 검정결과를 바르게 해석하고 있는가?
 ① 그렇다 () ② 아니다 (내용 :)
 6-2. 통계적 분석결과와 그 한계를 바르게 이해하여 결론을 도출하였는가?
 ① 그렇다 () ② 아니다 (내용 :)

7. 기타
 7-1. 통계용어를 적절하게 사용하였는가?
 ① 그렇다 () ② 아니다 (잘못 사용한 용어:)
 7-2. 전문적인 연구방법론 용어, 약어, 기호를 명확히 사용하였는가?
 ① 그렇다 () ② 아니다 (잘못 사용한 예:)
 7-3. 통계학적 방법론에 관한 참고문헌을 기재하였는가?
 ① 그렇다 () ② 아니다 ()

4. 통계학적 방법론의 타당성 평가

자료의 성격상 통계적 기법의 적용이 가능한 것으로 판단되었더라도 해당 논문의 저자가 통계적 분석을 시행하지 않았던 경우는 통계학적 방법론의 타당

성 평가대상에 포함하지 않았다.

통계학적 방법론의 타당성에 대한 평가영역은 표 2의 점검표에서와 같이 통계적 기법 적용의 타당성, 통계학적 방법론 및 결과 서술의 타당성, 통계적 분석결과 해석의 타당성 그리고 기타의 크게 4부분으로 나누

Table 3. Selection criteria for the articles which were satisfactory in utilization of statistical methodology

세부항목에 대한 평가결과
1. 기술통계학적 기법 적용이 적절
2. 자료의 특성에 맞는 추측통계학적 기법 선택
3. 통계학적 방법론과 분석결과를 상세히 기술
4. 추측통계학적 기법을 사용한 경우 유의수준 또는 신뢰구간을 제시
5. 적용했다고 주장한 각 통계적 기법의 분석결과를 모두 서술
6. P-값을 유의수준과 비교하여 바르게 해석
7. 통계적 분석결과와 그 한계를 바르게 이해하고 결론을 도출

었으며 총 14개항의 각 영역별 세부항목에 대해 평가하였다. 각 세부항목에 대한 구체적인 평가지침은 이등이 제시한 기준²을 따랐다. 또한 14개의 세부항목 중 통계학적 방법론의 타당성 여부에 큰 영향을 미치게 되는 주요 7개 세부항목을 선별하여 이에 대한 평가결과가 모두 양호한 경우를 통계학적으로 수용 가능하다고 평가하였다^{2,21}(Table 3).

최종적인 평가방법은 기본적인 통계지식을 갖고 있는 3인의 내과 전문의가 연구대상 논문 모두를 점검표에 따라 개별적으로 평가하고, 토론을 통하여 평가가 일치되는 항목과 일치되지 않는 항목을 정리한 후, 통계전문가 1인이 포함된 4인 평가자의 종합토의를 통해 최종합의를 도출하였다.

5. 통계분석

통계분석은 Windows용 SPSS 프로그램(SPSS Release 10.0, SPSS Inc, Chicago, Illinois, USA)을 이용하였고, 빈도분석을 시행하여 각 항목의 빈도와 백분율을 산출하였다.

결 과

1. 학회지 게재논문의 연구방법론 유형 및 통계적 기법 활용현황

평가대상 논문 185편은 원저가 110편(59.5%), 증례 보고가 61편(33.0%), 종설이 12편(6.5%) 그리고 학술 위원회 보고와 연구보고가 각각 1편이었다.

증례보고와 종설을 제외하고 원저 110편과 학술위

원회 보고, 연구보고 등 112편의 논문을 대상으로 통계적 기법의 활용현황을 살펴보면, 기술통계학적 기법과 추측통계학적 기법을 함께 활용한 논문이 94편(83.9%)으로 가장 많았으며, 기술통계학적 기법만을 적용한 경우가 13편(11.6%) 그리고 통계적 기법을 활용하지 않은 논문이 5편(4.5%)이었다. 추측통계학적 기법이 활용된 94편의 논문 중 9편(9.6%)은 적용한 통계적 기법이 무엇인지를 명시하지 않고 p-값만을 기술하였다. 85편(90.4%)의 논문은 적용한 추측통계학적 기법이 무엇인지 명시하였으며, 평균 2.0가지(범위 1-5가지)의 추측통계학적 기법을 사용하였다.

사용된 추측통계학적 분석기법 및 활용빈도를 분석하면 t-test, Pearson chi-square test가 각각 34편(36.2%), 32편(34.0%)으로 가장 많았고 이어서 민감도와 특이도, 예측도 등이 13편(13.8%), Mann-Whitney U test가 12편(12.8%), Fisher's exact test가 11편(11.7%) 등의 순서로 나타났다(Fig. 1). 이를 다시 분류하면 평균값의 비교를 위한 t-test, Mann-Whitney U test, paired t-test, 분산분석법(analysis of variance, ANOVA) 등이 64편(68.1%), 비연속 변수의 분석을 위한 Pearson chi-square test, Fisher's exact test가 43편(45.7%) 그리고 상관 및 회귀분석이 18편(19.1%)에서 사용되어 기본적인 추측통계학적 분석방법이 가장 빈번히 활용되고 있음을 알 수 있었다. 상대적으로 고급 분석방법이라고 할 수 있는 다변량 분석과 생존분석은 각각 9편(9.6%)과 8편(8.5%)에서만 활용되고 있었다(Fig. 1).

2. 통계학적 방법론의 타당성 평가

1) 통계적 기법 적용의 타당성

기술통계학적 기법이 적용된 논문에 대해서는 변수의 척도에 맞는 대표값과 산포도를 제시하고 있는지 여부를 평가하였으며, 기술통계학적 기법을 사용한 107편의 논문 중 18편(16.8%)이 적절한 기법을 적용하였고, 89편(83.2%)이 부적절하였다.

명목척도나 순위척도 등의 질적 변수에 대한 기술에서, 각 구간에 대한 횡수(frequency) 또는 비율(percent) 중 한가지만 제시한 경우가 52편(48.6%)에

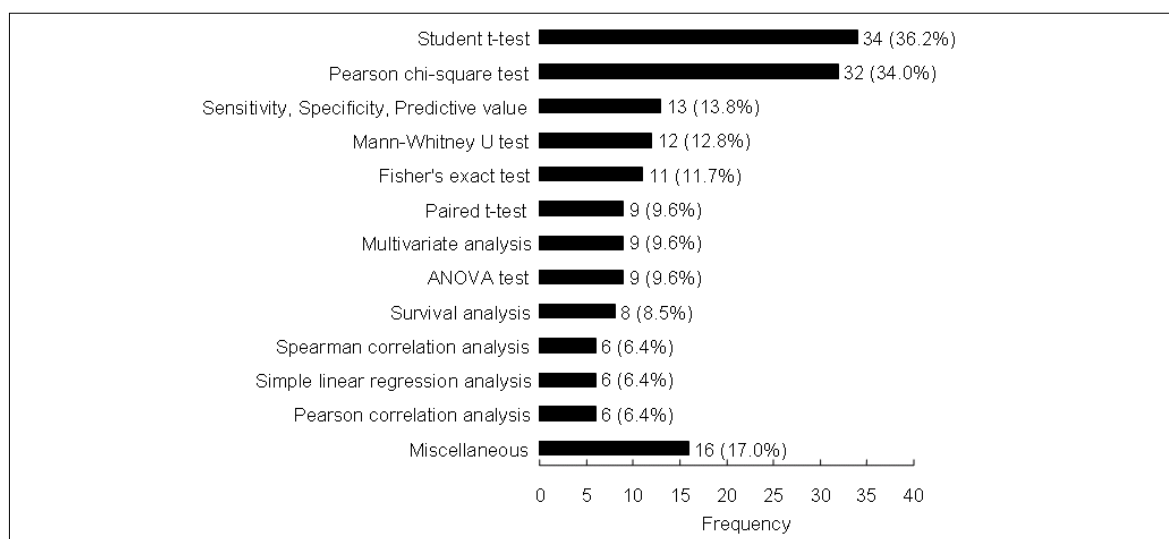


Figure. 1. Frequency of each inferential statistical method used in the 94 articles in the journal of Tuberculosis and Respiratory Diseases published in 1999.

서 관찰되어 가장 흔한 오류였으며, 조직학적 분화 정도나 환자의 만족도와 같은 순위척도를 평균과 표준편차로 제시한 경우가 3편(2.8%)이었다.

구간척도나 비척도 등 양적 변수에 대한 기술에서는, $a \pm b$ 로 제시한 결과에서 a 와 b 의 의미를 기술하지 않은 경우(39편, 36.4%)가 가장 흔하였으며, 평균보다 큰 표준편차를 제시하면서 범위를 부기하지 않은 경우(26편, 24.3%) 비모수 검정기법으로 분석한 변수를 평균과 표준편차로만 요약하는 경우(16편, 15.0%), 그리고 평균 등 대표값만 제시하고 산포도는 생략한 경우(12편, 11.2%), 표준편차 대신 표준오차를 제시한 경우(8편, 7.5%) 등의 순이었다.

추측통계학적 기법 적용의 타당성은 추측통계학적 기법을 활용한 94편의 논문 중 민감도, 특이도, 예측도 등 진단 통계만을 사용한 5편을 제외하고 89편의 논문을 대상으로 분석하였으며, 이 중 10편(11.2%)만이 자료의 특성에 맞는 통계적 분석방법을 적용하고 있었다.

자료의 특성에 맞지 않는 통계적 기법을 적용한 예를 살펴보면, 표본수가 적고 정규성이 의심되는 변수에 모수적 검정기법을 적용한 경우가 35편(39.3%)으로 가장 많았다. 그 외 분할표 분석시 표본수가 적거나 기대빈도가 작은 칸이 있음에도 불구하고 Pearson chi-square test를 적용한 경우(18편, 20.2%), 세 군 이상 또는 3회 이상 반복 측정된 자료분석에 t-test와

paired t-test 등으로 중복검정을 시행한 경우(17편, 19.1%), 연구의 목적상 경향분석이 필요함에도 불구하고 Pearson chi-square test를 적용한 경우(11편, 12.4%), 예후 요인 또는 만족도에 미치는 요인 분석 등과 같은 연구제목을 표방하였음에도 불구하고 다양한 요인들간의 상호관계를 고려한 다변량 분석을 시행하지 않고 t-test나 Pearson chi-square test와 같은 기본적인 기법만을 적용한 경우(11편, 12.4%), 그리고 두 변수간의 종속관계 또는 선후관계가 불명확함에도 불구하고 특별한 설명 없이 상관분석이 아닌 회귀분석을 적용한 경우(8편, 9.0%) 등의 순이었다.

평가대상논문 89편 중 61편(68.5%)에서는 추측통계학적 기법의 적용과정에서 요구되는 통계적 절차 또는 가정이 고려되지 않고 있는 것으로 나타났다.

2) 통계학적 방법론 및 결과 서술의 타당성

평가대상논문 89편 중 6편(6.7%)만이 통계학적 방법론과 분석결과를 비교적 상세히 기술하고 있었다. 83편(93.3%)의 논문은 “연구 대상 및 방법” 항목에서 적용한 통계적 기법을 분석대상에 따라 명확히 서술하지 않고 나열식으로만 제시하거나, “결과” 항목에서 p-값만 기술하고 이것이 어떤 분석기법의 결과인지 구체적으로 밝히지 않았다. 유의수준 또는 신뢰구간 구역을 제시한 경우는 54편(60.7%), 검정을 시행한 경

우 정확한 p-값을 정량적으로 제시한 경우는 19편(21.3%)이었다. 31편(34.8%)의 논문은 “연구 대상 및 방법”에서는 여러 가지 통계적 기법을 사용한 것으로 기술하였으나, 실제 “결과” 항목에서는 일부 결과가 누락되어 있었다. 활용한 통계 프로그램이 무엇인지 명시한 경우는 43편(48.3%)이었고, 표와 도표 등 보조적 표현수단은 46편(51.7%)에서 적절히 사용되고 있었다.

3) 통계적 분석결과 해석의 타당성

평가대상논문 89편 중 86편(96.7%)은 p-값과 유의수준의 개념을 구분하여 바르게 해석하고 있었다. 그럼에도 불구하고 통계적 분석결과와 그 한계를 바르게 이해하고 결론을 도출한 경우는 58편(65.2%)에 불과하였다. 31편(34.8%)의 논문에서는 p-값 또는 유의수준에 대한 언급이 없이 단정적인 결론을 내리거나, 상관분석 또는 회귀분석에서 나타난 두 변수간의 관련성을 의학적인 인과성으로 확대 해석하거나 또는 p-값이 0.05의 유의수준보다는 크지만 0.1보다는 작은 경우 즉, p-값이 경계역 상의 유의성을 보임에도 불구하고 이에 대한 합리적인 의미부여 없이 “차이가 없었다” 등과 같이 이분법적으로 결론을 단정하였다.

4) 기타 영역에 대한 평가

평가대상논문 89편 중 43편(48.3%)에서만 의학용어집²⁷ 또는 통계용어사전²⁸의 공인된 통계용어를 적절히 활용하고 있었다. “유의수준이 0.05 이하인 경우 통계적으로 유의한 것으로 간주하였다”와 같이 유의수준을 p-값과 혼동하거나 질적변수에 대한 분할표 분석시 Pearson chi-square test, Fischer's exact test 등 통계학적 방법을 구체적으로 기술하지 않고 “빈도수분석”과 같이 명확하지 않게 표현한 경우 등이 흔하였다.

기타 전문적인 용어, 약어, 기호 등의 경우는 69편(77.5%)의 논문에서 비교적 명확한 사용을 하고 있었다. 20편(22.5%)의 논문에서는 수신기작동특성곡선(receiver operating characteristics curve) 등의 전문적인 연구방법론 용어 또는 약어를 구체적인 해설이 없이 사용하거나 또는 전문통계용어인 상관성(correlation)을 인과성의 개념으로 사용하거나 또는 무작위

(random)를 정확한 방법에 대한 기술이 없이 편의적 개념으로 사용하였다.

통계학적 방법론에 관한 참고문헌을 기재한 논문은 2편(2.2%)에 불과하였다.

5) 통계학적 방법론이 적절히 활용된 논문

통계학적 방법론에 대한 세부 평가항목 중 상대적으로 연구논문의 타당성에 큰 영향을 미치는 7가지 항목(Table 3)에 대한 평가결과가 전부 양호한 경우를 통계학적으로 수용 가능하다고 가정하였을 때, 통계학적 기법이 활용된 107편의 논문 중 14편(13.1%)만이 여기에 해당하였고, 93편(86.9%)에서는 중요한 통계적 오류가 발견되었다.

고 찰

본 연구의 분석대상이 된 1999년도 [결핵 및 호흡기 질환] 게재 논문 중 증례와 종설을 제외한 112편 중 통계적 기법이 활용된 논문이 전체의 95.5%로 매우 높게 나타났다. 이는 국내 의학학술지를 대상으로 보고된 30-90%⁹⁻²²보다 상대적으로 통계적 기법의 활용빈도가 높은 수준임을 보여주었다. 또한 단순히 자료를 요약 제시하는 기술통계학적 기법의 적용수준을 넘어 연구자가 증명 또는 추론하고자 하는 목적으로 추측통계학적 기법을 활용한 논문이 83.9%에 달하고 있어 통계적 기법 활용의 내용적 측면에서도 상당히 발전된 모습을 보여주었다.

추측통계학적 기법의 개별적인 활용빈도를 보면 집단간 평균값의 비교가 64편(68.1%), 비연속 변수의 분할표 분석이 43편(45.7%) 그리고 상관 및 회귀분석이 18편(19.1%)에서 사용되어 기본적인 통계적 기법이 빈번하게 사용되고 있음을 알 수 있다. 이는 국내외 학술지를 대상으로 조사한 선행연구결과^{4-22,29-32}와 유사한 것이다. 하지만 최근 게재되는 해외학술지 논문에 비해서는³³⁻³⁵ 아직까지 다변량 분석, 로지스틱 회귀분석, 생존 분석 등의 고급 통계 기법은 아직까지는 그 활용빈도가 상대적으로 적음을 알 수 있다.

미국의학협회지(Journal of the American Medical Association)에서는 의학연구에서의 통계학적 방법론

이 차지하는 중요성을 인식하여 지난 1965년부터 게재신청논문에 대한 통계학적 심사과정을 최초로 의무화하였으며³⁹, 이후 외국의 많은 주요 의학학술지에서도 해당 학술지에 게재된 논문들의 통계적 오류를 사후 평가하거나 그 개선방안을 제시하는 연구들이 1970년대 이후 잇따르고 있다. 이러한 추세에 따라 1988년 국제의학학술지 편집인위원회에서는 의학학술지에 투고하는 원고의 통일양식에 통계학적 방법론에 관한 서술지침을 포함하게 되었다²³⁻²⁵. 본 연구에서는 이러한 서술지침을 기본으로 반영하고 실제 국내의학학술지에서 자주 관찰되는 통계적 오류 유형을 고려하여 이 등이 개발한 점검표²²¹를 사용하였다. 본 점검표는 대한성형외과학회지 게재논문에 대한 통계적 오류의 평가²¹에 이미 사용된 적이 있으며, 임상연구자들에게도 도움이 될 수 있도록 구체적이고 평이하게 작성되었다는 점에서 실용적인 가치가 있다고 사료된다.

본 연구에서 크게 4부분으로 구분한 통계학적 방법론에 대한 평가영역별 분석결과 중 먼저 통계적 기법 적용의 타당성에 대한 평가결과를 보면 기술통계학적 기법의 적용은 전체의 16.8%에서만 적절한 것으로 나타났다, 추측통계학적 기법의 경우는 11.2%만이 자료의 특성에 맞는 적합한 분석방법을 논문 전체에서 일관성 있게 적용하고 있었다. 추측통계학적 기법을 적용하는 과정에서는 합당한 통계적인 가정 또는 절차를 따라야 하지만 이를 준수한 경우는 전체의 32.5%에 불과하였고, 이는 국내 다른 의학학술지를 대상으로 한 결과와도 유사하다^{9,10,12-15,17,19-21}.

기술통계학의 목적은 관찰에 의하여 모아진 자료를 정리하거나 간결하게 요약함으로써 독자들이 연구결과의 윤곽을 쉽게 파악하게 하는 것이다². 사용하는 기술통계학적 기법은 변수의 척도에 따라 달라지는데 명칭척도나 순위척도 등의 질적 변수는 각 범주에서의 횟수와 분율을 제시하여 자료를 요약하는 것이 일반적이며, 간격척도나 비척도 등의 양적 변수는 대개 수학적인 조작과정을 거쳐 관찰값의 대표값과 산포도를 제시한다. 대표값의 종류에는 평균, 중앙값, 최빈수, 등분위수 등이 있으며, 산포도의 종류에는 최소값과 최대값 또는 그 차이로 제시되는 범위, 제3사분위수와 제1사분위수 값의 차이인 사분위수의 범위(inter-

quartile range), 표준편차, 분산 등이 있다. 연속변수의 요약에 위해 적용하는 기술통계학적 기법은 표본수, 산포도, 정규성 등 자료의 특성과 추후 적용하고자 하는 추측통계학적 기법을 고려하여 선택하여야 한다. 표본수가 적고 산포도도 커서 비모수적 검정 기법을 적용하여야 할 자료를 확일적으로 평균 \pm 표준편차 형식으로 요약하는 것은 적절하지 못한 기술통계학적 기법이다^{26,9,40}. 본 연구에서 평균보다 큰 표준편차를 제시하면서 범위를 부기하지 않은 경우가 24.3%, 비모수 검정기법으로 분석한 변수를 평균과 표준편차로만 요약하는 경우가 15.0%의 논문에서 관찰된 것이 그 예라 할 수 있다.

양적 변수에 대한 기술에서 가장 흔한 오류는 $a \pm b$ 로 제시한 결과에서 a 와 b 의 의미를 기술하지 않은 경우로 36.4%의 논문에서 이러한 문제점이 나타났다. 연구자의 의도가 a 는 평균, b 는 표준편차를 의도하는 것으로 해석할 수도 있으나, $a \pm b$ 로 결과를 제시하면서 b 가 표준편차가 아닌 표준오차라고 밝힌 논문이 7.5%인 것을 볼 때 정확한 의미를 기술하는 것이 중요하다고 사료된다. 또한 평균 등 대표값만 제시하고 산포도는 생략한 경우도 11.2%에 달해 자료의 특성을 요약 제시하는 기술통계학적 기법의 중요성에 대한 연구자들의 주의가 요망된다 할 것이다.

추측통계학이란 시간적 경제적 제약으로 인해 자료의 일부인 표본을 관찰하여 모집단의 특성을 추정하거나 연구가설을 검정하는 통계학을 말한다². 추측통계학에서는 연구목적에 따른 가설이 포함된 독립변수 및 종속변수의 척도가 정해지면 연구가설을 증명할 통계분석법이 결정된다. 추측통계를 적용할 때는 변수의 척도와 표본자료의 수, 자료의 정규성, 분석의 목적 등을 고려하여 적절한 통계적 기법을 적용하는 것이 중요하다^{1,2}.

[결핵 및 호흡기질환] 게재 논문 중 가장 흔히 관찰되는 오류는 비모수적 기법을 적용해야 하는데 모수적 기법을 적용한 경우(39.3%)였다. 평균치 분석에서 t-test나 ANOVA test와 같은 모수적 방법을 적용하기 위해서는 엄격한 기준이 필요하다. 각각의 표본은 독립적이어야 하며, 모집단에서 무작위적으로 추출되어야 하고, 모집단에 관해 특정 분포를 상정할 수 있어

야 한다. 또한 자료들은 순위척도나 명목척도가 아닌 간격척도나 비척도이어야 한다^{1,2}. 이러한 기준들이 충족되지 않으면 모수적 기법은 유의성 결정에 신뢰성을 잃게 되며 모집단을 적절히 설명할 수 없게 되는데, 표본이 큰 경우는 모수적 기법을 적용하여도 큰 무리가 없다고 할 수도 있으나, 정규성 검정 결과에 따라 통계적 기법을 선택하는 것이 올바르다 할 것이다^{1,2}. 표본수가 적고 정규성이 의심되는 변수에 Wilcoxon rank sum test 대신 t-test를 적용하거나, Kruskal-Wallis test 대신 ANOVA test를 적용한 경우 그리고 Spearman correlation analysis 대신 Pearson correlation analysis를 적용한 경우 등은 이러한 원칙을 지키지 않은 예라 할 수 있다.

그리고 세 집단 이상 또는 3회 이상 반복 측정된 자료에서 ANOVA test나 repeated measures ANOVA test를 사용하지 않고 t-test와 paired t-test 등으로 중복검정을 시행하는 경우도 19.1%에서 관찰되는 오류로, 이렇게 각각 두 개의 군끼리 t-test나 paired t-test를 반복하면 제1종 오류(type I error)의 증가를 가져와 유의하지 않은 결과를 유의한 것으로 잘못 결론짓을 가능성이 발생하게 된다^{1,2}.

질적변수에 대한 분할표 분석시 2×2 표에서 기대치가 5보다 작은 칸이 있음에도 불구하고 Fisher's exact test를 사용하지 않았거나 각 범주를 통합하여 각 칸의 기대빈도를 크게 하는 등의 시도를 하지 않고 Pearson chi-square test를 그대로 적용하는 경우도 20.2%에서 관찰되었다. 또한 암의 병기와 같이 순위척도가 높아질수록 사망률이 높아지는 지를 평가하는 연구와 같이 연구의 목적상 경향분석이 필요함에도 불구하고 Pearson chi-square test를 적용한 경우도 12.4%에서 관찰되었다.

이상과 같이 [결핵 및 호흡기질환] 게재 논문에서 적절하지 못한 통계적 기법을 사용한 오류의 대부분은 두 군 또는 세 군의 평균값 비교, 상관 및 회귀분석, 질적변수에 대한 분할표 분석과 같은 비교적 흔히 사용되는 기본적인 통계적 기법과 관련되어 있었다. 이는 의학연구에서 흔히 발견되는 통계적 오류가 결코 난해하고 복잡한 분석방법을 잘못 적용했기 때문이 아니라 오히려 쉽고 평이한 방법들을 올바르게 사용하

지 못함에 기인한 것이라는 주장을 뒷받침해준다⁹. Emerson 등은 1978년과 1979년 2년간 New England Journal of Medicine에 게재된 760편의 논문 중 통계학적 기법을 사용하지 않았거나 기술통계학적 기법만을 사용한 논문이 58%이며, t-test를 사용한 논문이 24%, 질적 변수에 대한 분할표 분석이 15%에서 사용되고 있어, 기술통계학적 기법과 함께 추측통계학적 기법 중 t-test와 분할표 분석에 대한 통계학적 이해만으로도 게재된 논문의 73%에 대해 통계적 이해가 가능하다고 하였다²⁹. 기본적인 통계적 기법의 기술과 적용에 대한 연구자들의 좀 더 깊은 이해와 노력이 필요할 것으로 사료된다.

두 번째로 통계학적 방법론 및 결과 서술의 타당성에서는 통계학적 방법론과 분석결과를 상세히 기술하였는가의 항목에서 6.7%에서만 양호한 것으로 나타나 이 영역에 대한 평가결과가 타 항목에 비해 상대적으로 낮았으며, 이는 국내학술지^{10,12-15,17,19-21} 뿐만 아니라 외국학술지³¹에서도 가장 흔히 관찰되는 오류 중 하나로 보고되고 있다. 국제의학학술지 편집인위원회에서 권고하고 있는 통계학적 방법론에 대한 서술지침 중 가장 첫 항목은 연구 내용을 이해할 만한 독자가 연구 데이터를 확보했을 때 논문에 나타난 결과를 확인할 수 있을 정도로 통계학적 방법을 자세하게 기술하여야 한다는 것으로²³⁻²⁵ 이에 대한 중요성을 강조하고 있다.

세 번째로 통계적 분석결과 해석의 타당성 평가에서는 p-값을 유의수준과 비교하여 유의성 여부를 바르게 판정한 경우가 96.7%로 다른 항목보다 상대적으로 매우 높았다. 하지만 통계적 분석결과와 그 한계를 바르게 이해하여 결론을 도출한 경우는 65.2%로 낮았다. 이는 미국의 소아과 전공의들을 대상으로 통계에 대한 기본개념의 이해 정도를 조사한 연구결과에서 전체 응답자의 절반 이하만이 p-값에 대해 올바르게 이해하고 있는 것으로 드러나기도 했듯이 기본적인 통계용어의 개념에 대한 이해 부족현상은 외국도 예외는 아닌 실정이다³⁷.

마지막으로 기타 영역에 대한 평가에서 통계용어나 의학연구방법론 용어 등의 정확한 사용여부는 각각 48.3%와 77.5%의 논문이 양호한 평가를 받았다. 용어 등의 부정확한 활용이 연구논문의 질 저하에 미치는

영향이 크지는 않더라도 연구논문의 형식적인 완성도와 직결되는 문제이고, 국제의학학술지 편집인위원회에서도 이를 강조하고 있음을 상기할 때 의학논문을 작성하는 연구자로서 간과해서는 안될 부분이다. 한편 통계학적 방법론에 대한 참고문헌의 기재도 국제의학학술지 편집인위원회에서 제시한 통계적 서술 통일지침에서 규정하고 있으나 단 2편(2.2%)의 논문만 이를 기재하였다. 어느 정도의 통계적인 지식이 있는 연구자라고 하더라도 연구논문 작성을 위해 통계분석을 시행할 경우는 대부분 통계전문서적 또는 유사한 분석방법을 활용한 선행 연구논문을 참고할 것이다. 또한 독자들에게도 때로는 연구논문에서 제시하는 의학적인 결론보다도 연구방법론에 대한 참고문헌이 더욱 큰 관심의 대상이 된다는 점에서 통계학적 방법론에 대한 참고문헌을 가능한 구체적으로 기술하는 것이 바람직하다 할 것이다².

본 연구에서는 점검표의 평가항목 중 부록 3에서 제시한 7가지 주요항목들에 대해서 모두 만족스러운 평가가 내려진 경우를 비교적 통계학적 방법론이 올바르게 사용된 논문이라고 가정한 결과 통계학적 기법이 활용된 논문 중 13.1%만이 이에 해당하였다. 이는 반대로 [결핵 및 호흡기질환] 게재 논문 중 나머지 86.9%에서는 통계학적 방법론이 타당하지 못하게 적용되었다는 것이다.

학술지 게재논문들을 대상으로 통계학적 방법론의 타당성을 평가한 주요 연구들은 통계적 오류의 평가 범위와 기준 등이 서로 다르기 때문에 그 결과를 직접적으로 비교하는 것은 의미가 없지만, 이를 통해 학술 논문에서의 통계적 오류 현황을 개략적으로 파악할 수는 있다. British Medical Journal, Circulation, Journal of Infectious Diseases 등 몇몇 외국 의학학술지를 대상으로 한 기존의 연구에서는 게재되는 논문 중 40-70%에서 통계적 오류가 발견된다고 보고되었다⁴⁻⁸. 또한 대한의사협회지 등 여러 국내 의학학술지를 대상으로 한 연구에서는 통계적 오류가 적게는 60-80%^{10,11,13,15,17,20}, 많게는 95-100%^{9,12,21}에 이르는 나 타나고 있어, 이러한 문제가 국내 의학학술지만의 특별한 문제이거나 또는 [결핵 및 호흡기질환]에 국한된 문제가 아니라는 것을 알 수 있다.

본 연구는 [결핵 및 호흡기질환] 게재 논문의 질을 평가하고자 시도된 것이 아니며 본 연구에서 사용한 점검표가 완벽한 평가서라고 장담할 수는 없다. 다만 [결핵 및 호흡기질환] 게재 논문에서 사용된 통계분석 방법의 종류와 활용 빈도 그리고 이의 사용이 적절한지의 현황을 파악하여 향후 이러한 오류를 최소화하기 위한 노력의 필요성을 제시하고자 하였고, 이의 기본적인 자료를 제공하고자 하였다.

통계적 기법 적용의 오류를 개선하기 위해서는 무엇보다도 연구자 스스로의 노력이 가장 중요하다고 할 수 있다. 본 연구에서 나타난 것과 같이 흔히 사용되는 기본적인 통계적 기법을 연구자 스스로 숙지하는 것은 연구논문의 통계적 측면에서의 질을 향상시키는 가장 빠른 방법일 것이다. 이를 위해 필요하다면 결과를 분석하는 단계가 아닌 연구를 계획하는 단계에서부터 연구방법론 및 통계기법 적용에 대해 전문가의 조언을 얻는 방법도 적극 활용하여야 할 것이다.

투고된 논문의 심사와 게재 그리고 게재 이후 평가과정에서의 개선 노력이 또한 필요할 것이다. 이는 통계학적 방법론에 대한 서술지침의 제정, 투고된 논문에 대해 연구방법론과 통계적 기법에 대한 타당성 평가 체계의 확립, 학술지에 독자의 반론을 실을 수 있는 지면의 마련, 학회 회원들의 연구능력 향상을 위한 지상 또는 특별 강좌의 활성화 등이 포함될 수 있을 것이다^{6,41}. 최근 국내 몇몇 학술지에서 논문 작성과정에서부터 통계 자문 및 분석 의뢰의 제공⁴², 학회지에 게재된 논문에 대해 사후 연구방법론 및 통계적 기법에 대한 평가를 담은 논평의 수록^{43,44}, 통계적 기법에 대한 지상 강좌의 마련^{45,46} 등이 시도되고 있는 점은 이러한 면에서 긍정적으로 평가할 수 있을 것이다.

요 약

배 경 :

의학연구에서는 표본자료로부터 의학적 의미를 찾아내고 그 결과를 일반화시키기 위하여 통계적 기법과 원리를 활용하게 된다. 현재까지 여러 국내 의학학술지를 대상으로 한 연구에서는 60-100%의 논문에서 통계적 오류가 있는 것으로 보고되었다. 본 연구는 [결

핵 및 호흡기질환]에 게재된 논문에서의 통계적 기법의 활용 및 오류 실태를 파악하고자 하였다.

대상 및 방법 :

1999년에 발간된 총 12권의 [결핵 및 호흡기질환]에 게재된 논문 185편을 대상으로 하였다. 주요 평가영역은 연구유형, 통계적 기법의 활용현황 그리고 통계적 기법 적용의 타당성 등이었으며, 국제의학학술지편집위원회에서 제시한 통계적 서술 통일양식을 기초로 개발된 점검표를 활용하여 평가하였다. 평가는 연구자들간의 종합토의를 통해 합의를 도출하였다.

결 과 :

1) 논문을 연구유형에 따라 분류하면 원저 110편(59.5%), 증례보고 61편(33.0%), 종설 12편(6.5%), 기타 2편(1.1%) 순이었다. 증례보고와 종설을 제외한 112편의 논문 중 통계기법은 107편(95.5%)에서 활용되었으며 기술통계학적 기법만을 적용한 경우가 13편(11.6%), 그리고 추측통계학적 분석을 함께 활용한 경우가 94편(83.9%)이었다.

2) 추측통계학적 분석을 활용한 논문에서는 평균 2.0가지(범위 1-5개)의 통계기법을 적용하였으며, 집단간 평균값의 비교가 64편(68.1%), 비연속 변수의 분할표 분석이 43편(45.7%) 그리고 상관 및 회귀분석이 18편(19.1%)에서 사용되었다. 9편(9.6%)의 논문에서는 p-값을 제시했음에도 불구하고 어떤 분석기법을 이용했는지 제시하지 않았다.

3) 기술통계학적 기법을 적용하여 자료를 요약 제시한 107편의 논문 중 변수의 특성을 고려하지 않고 적절하지 못한 대표값 또는 산포도를 제시한 경우가 89편(83.2%)이었다.

4) 추측통계학적 기법을 활용한 94편의 논문 중 민감도와 특이도, 예측도 등 진단 통계만을 사용한 5편의 논문을 제외한 89편의 논문 중 79편(88.8%)에서 자료의 특성에 맞지 않는 분석기법을 적용하고 있었으며, 가장 흔한 오류는 정규성이 의심되는 변수에 모수적 검정기법을 적용한 경우였다(35편, 39.3%).

5) 전체적으로 통계적 기법이 활용된 107편의 논문 중 93편(86.9%)에서 중요한 통계적 오류가 관찰되었다.

결 론 :

[결핵 및 호흡기질환] 게재논문 중 통계적 기법을 활용한 많은 논문에서 통계적 오류가 발견되고 있으며, 이는 대부분 기본적인 통계적 기법의 적용과 관련되어 있었다. 통계적 기법을 적용할 때 연구자의 보다 세심한 주의와 함께, 논문의 심사와 게재 과정에서 적절한 지침과 평가가 필요하리라 사료된다.

감사의 글

본 연구결과를 세심히 검토해 주시고 많은 조언을 해 주신 서울아산병원 호흡기내과 오연목 선생님께 깊은 감사를 드립니다.

참 고 문 헌

1. Ahn YO, Yoo KY, Park BJ. Manual for medical statistics. Seoul: Seoul National University Press; 1990.
2. Lee HJ. Medical research and use of statistics. Seoul: Korea Medical Book Publisher; 2000.
3. Bailar JC III, Mosteller F. Medical uses of statistics. 2nd ed. Boston: NEJM Books; 1992.
4. 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.
5. White SJ. Statistical errors in papers in the British Journal of Psychiatry. Br J Psychiatry 1979;135:336-42.
6. Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. Circulation 1980;61:1-7.
7. Felson DT, Cupples LA, Meenan RF. Misuse of statistical methods in Arthritis and Rheumatism. 1982 versus 1967-68. Arthritis Rheum 1984;27:1018-22.
8. 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.
9. Lee HK, Huh BY, Ahn YO. An assessment of methodological and statistical validity of medical articles published in Korea, from 1980 to 1989. J Korean Acad Fam Med 1991;12:46-67.
10. Cho SC, Ahn YO. Present status of description and application of statistics in papers in the Journal of Neuropsychiatric Association. J Korean Neuropsychiatric Assoc 1991;30:462-76.

11. 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.
12. Kim YY, Lee EI, Jung HR. Research designs, research methods, and statistical techniques in Korean ophthalmic articles. *J Korean Ophthalmol Soc* 1994;35:460-73.
13. Song H, Park KH, Kim WH, Jun TG. Present statistical status in papers in the Korean Journal of Thoracic and Cardiovascular Surgery. *Korean J Thorac Cardiovasc Surg* 1994;27:732-7.
14. Park IK, Kang JO, Kim TY, Keum DG. Statistical techniques used in 1981~1995 edition of Korean Journal of Clinical Pathology. *Korean J Clin Pathol* 1997;17:878-84.
15. Lee SK, Kim JH, Hwang HY, Lee HK, Chung HS, Kim HS. The investigation of statistical descriptions in the original articles of Journal of the Korean Radiological Society. *J Korean Radiol Soc* 1997;37: 953-8.
16. Kim JH, Kim JS. Statistical methods used in the Journal of Korean Academy of Rehabilitation Medicine. *J Korean Acad Rehabil Med* 1998;22:46-55.
17. 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.
18. Kim YJ, Choi YJ. Statistical analysis used for the studies in the Journal of the Korean Orthopaedic Association. *J Korean Orthop Assoc* 1999;34:625-9.
19. Park IK, Kim TY. An evaluation of the statistical techniques used in the 1994~1998 editions of The Journal of the Korean Rheumatism Association The Journal of the Korean Rheumatism Association. 2000;7:120-6.
20. Ahn WS. Statistical methods in the articles in the Korean Journal of Anesthesiology Published from 1994 to 1998. *Korean J Anesthesiol* 2000;39:706-1.
21. 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.
22. Kim SN, Huh S. Can statistics used in the medical journals from Korea be covered by Korean statistical program, dBSTAT? *Korean J Med Educ* 2002;14:111-7.
23. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Intern Med* 1997;126:36-47.
24. Bailar JC 3rd, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. *Ann Intern Med* 1988;108: 266-73.
25. Korean Association of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. 1997.
26. Yoon BJ, Chi JG. How to write medical papers and use literature retrieval system. Seoul: Seoul National University Press; 1992.
27. Korean Medical Association. English-Korean Medical Terminology. 3rd ed. Seoul: Academy Publisher; 1992.
28. Korean Statistical Society. Statistical term glossary. Seoul: Freedom Academy; 1992.
29. Emerson JD, Colditz GA. Use of statistical analysis in the New England Journal of Medicine. *N Engl J Med* 1983;309:709-13.
30. Colditz GA, Emerson JD. The statistical content of published medical research: some implications for biomedical education. *Med Educ* 1985;19:248-55.
31. Hokanson JA, Luttmann DJ, Weiss GB. Frequency and diversity of use of statistical techniques in oncology journals. *Cancer Treat Rep* 1986;70:589-94.
32. 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.
33. Altman DG. Statistics in medical journals: some recent trends. *Stat Med* 2000;19:3275-89.
34. Chinn S. Statistics for the European Respiratory Journal. *Eur Respir J* 2001;18:393-401.
35. Moss M, Wellman DA, Cotsonis GA. An appraisal of multivariable logistic models in the pulmonary and critical care literature. *Chest* 2003;123:923-8.
36. Feinstein AR. Clinical biostatistics. XLIV. A survey of the research architecture used for publications in general medical journals. *Clin Pharmacol Ther* 1978; 24:117-25.
37. Friedman SB, Phillips S. What's the difference? Pediatric residents and their inaccurate concepts regarding statistics. *Pediatrics* 1981;68:644-6.
38. Berwick DM, Fineberg HV, Weinstein MC. When doctors meet numbers. *Am J Med* 1981;71:991-8.
39. Schor S, Karten I. Statistical evaluation of medical journal manuscripts. *JAMA* 1966;195:1123-8.
40. Wulff HR, Andersen B, Brandenhoff P, Guttler F. What do doctors know about statistics? *Stat Med* 1987;6:3-10.
41. Appleton DR. What statistics should we teach medical undergraduates and graduates? *Stat Med* 1990;9: 1013-21.
42. The Korean Academy of Family Medicine. Request for statistical consultation and analysis. *J Korean Acad*

- Fam Med 2001;22:114.
43. Park YG. Statistical errors in the published papers in the Journal (2001. 1). J Korean Acad Fam Med 2001;22:241-3.
44. Lee MS. Methodological and statistical consideration to the original articles published in the Journal (Volume 1, No. 2). Korean J Health Promot Dis Prev 2002;2:82-6.
45. Lee HK. Comparison of two population means. Korean J Pediatr Infect Dis 1998;5:152-7.
46. Park YG. Logistic regression analysis. J Korean Acad Fam Med 2001; 22:1007-20.
-