

## 대한방사선종양학회지 게재 논문의 통계적 오류 현황

성균관대학교 의과대학 삼성서울병원 방사선종양학교실\*, 연세대학교 의과대학 예방의학교실<sup>†</sup>, 경희대학교 의과대학 방사선종양학교실<sup>†</sup>, 순천향대학교 방사선종양학교실<sup>§</sup>, 울산대학교 의과대학 방사선종양학교실<sup>||</sup>, 고려대학교 의과대학 방사선종양학교실<sup>¶</sup>, 동국대학교 의과대학 방사선종양학교실<sup>#</sup>, 국민건강보험공단 일산병원 방사선종양학과<sup>\*\*</sup>, 서울대학교 의과대학 방사선종양학교실<sup>††</sup>, 국립암센터 양성자치료센터<sup>††</sup>, 건국대학교 의과대학 방사선종양학과<sup>§§</sup>

박희철\* · 최두호\* · 안성복<sup>†</sup> · 강진오<sup>†</sup> · 김은석<sup>§</sup> · 박 원\* · 안승도<sup>||</sup>  
양대식<sup>¶</sup> · 윤형근<sup>#</sup> · 정은지<sup>\*\*</sup> · 지의규<sup>††</sup> · 표홍렬<sup>††</sup> · 홍세미<sup>§§</sup>

**목적:** 본 연구는 대한방사선종양학회지 게재 논문의 통계 오류 현황을 파악하고 이에 근거한 문제 제기를 통해 학회지의 학술적 발전에 기여하고자 하였다.

**대상 및 방법:** 2006~2007년 사이에 대한방사선종양학회지에 게재된 총 77편의 논문을 연구 대상으로 하였다. 각각의 논문에 적용된 통계 방법론의 적정성 평가는 통계점검표를 활용하였다. 통계점검표에는 연구의 종류, 통계 기법의 사용 범위, 각 논문에 적용된 통계 기법의 종류, 통계적 기법 적용의 타당성 항목이 포함되었다. 통계 오류는 '생략의 잘못'과 '시행의 잘못' 항목으로 나누었다. 한 논문에서 서로 다른 항목이 여러 가지 관찰된 경우 각각 횟수로 측정하였다. 같은 항목이 2회 이상 측정된 경우 1회로 횟수를 측정하였다. 통계 전문가가 개별 논문을 대상으로 통계점검표를 작성하였다. 일차 평가자가 방사선종양학 전문가가 아닌 것에서 올 수 있는 평가 오류가 있을 수 있으므로 개별 통계점검표는 간행위원회에서 한 차례 더 세부 점검을 하였다. 작성된 통계점검표의 통계 분석은 SAS (version 9.0, SAS Institute, NC, USA) 소프트웨어를 이용하였고 빈도분석을 시행하여 각 항목의 빈도와 백분율을 산출하였다.

**결과:** 총 77편 중 원문을 다운로드 할 수 없었던 4편을 제외한 73편의 게재 논문을 대상으로 평가하였다. 증례보고는 5편, 원저 논문은 68편이었다. 대상 논문 중 46편의 논문에서 통계적 추론을 사용하였고, 16편의 논문에서 단순한 기술통계를 사용했으며, 11편의 논문에서는 통계적 기법을 사용하지 않았다. 추론통계를 사용한 46편의 논문에서 사용된 추론통계의 횟수는 각각 분할표분석 17회(37.0%), 비교통계분석 23회(50.0%), 회귀분석 7회(15.2%), 상관분석 5회(10.9%), 생존분석이 27회(58.7%)였다. 통계 기법을 활용하여 연구 결과를 분석할 때 통계 적용의 오류가 없는 논문은 19%였다. '생략의 잘못'은 34편(50.0%)의 논문에서 총 50회 관찰되었다. '시행의 잘못'은 35편(51.5%)의 논문에서 총 47회 관찰되었다. '생략의 잘못'과 '시행의 잘못'이 모두 발견된 논문은 21편(30.9%)이었다.

**결론:** 대한방사선종양학회지에 게재된 논문에서 통계 분석 과정의 다양한 영역에 걸쳐 크고 작은 통계적 오류가 있음을 확인하였다. 향후 대한방사선종양학회지 투고 논문의 심사 과정에서 통계 오류에 관한 적절한 심사를 추가하는 것이 필요하다고 판단된다.

**핵심용어:** 통계, 방사선종양학

## 서론

연구자들은 자신들의 임상 경험과 연구 결과를 바르고 정확하게 분석하여 논문의 형태로 발표함으로써 동료 연구자들과 교류하고 이런 상호 교류를 통해 관련 학문을 발전시킨다. 통계 분석은 논문을 쓰기 위하여 연구 자료를 분석하고 객관적이고 타당한 결론을 도출하는데 필수적인

이 논문은 2008년 10월 23일 접수하여 2008년 11월 3일 채택되었음.  
책임저자: 최두호, 삼성서울병원 방사선종양학과  
Tel: 02)3410-2436, Fax: 02)3410-2619  
E-mail: dohochoi@skku.edu



대상 및 방법

결 과

2006~2007년 사이에 대한방사선종양학회지에 게재된 총 77편의 논문을 연구 대상으로 하였다. 각각의 논문에 적용된 통계 방법론의 적정성 평가는 통계점검표를 활용하였다. 통계점검표는 통계 전문가와 대한방사선종양학회 간행위원회 회원의 종합 토의 및 문헌 고찰을 통하여 결정하였으며, 통계처리기법의 타당성을 평가하기 위한 기존 연구에서 안과 이가 제시한 통계점검표를 위주로 수정, 보완하였다(Fig. 1).<sup>8)</sup> 통계점검표에 포함된 평가 항목은 연구의 종류, 통계 기법의 사용 범위, 각 논문에 적용된 통계 기법의 종류, 통계적 기법 적용의 타당성 항목이었다.

연구 논문의 종류는 대한방사선종양학회지에서 사용하고 있는 분류에 따라 종설, 증례보고, 원저 논문으로 구분하였다. 평가 대상 논문이 통계 방법론을 사용하고 있는지의 여부를 조사하고, 통계 방법론을 사용한 경우 기술통계(descriptive statistics)와 추론통계(inferential statistics)로 나누었다. 이 중 추론통계를 사용한 논문을 대상으로 해당 통계 방법론의 빈도를 측정하였다. 만약 한 논문에서 여러 가지 통계 방법론이 사용된 경우 이를 각각 횟수로 측정하였다. 증례보고를 제외한 원저 논문을 대상으로 통계적 기법 적용의 타당성(통계 오류)을 조사하였다. 통계 오류는 같은 논문에서 서로 다른 오류가 두 가지 이상 있는 경우 이를 각각 횟수로 측정하였고 같은 오류가 2회 이상 반복된 경우 1회로 횟수를 측정하였다.

통계적 기법 적용의 타당성 항목에 분류한 통계 오류는 ‘생략의 잘못’과 ‘시행의 잘못’으로 나누었다. ‘생략의 잘못’은 기초자료에 대하여 충분히 기술되지 않았거나 사용된 통계처리기법 및 이의 검정력과 신뢰구간에 대하여 충분히 기술되지 않은 경우를 포함하였다. ‘시행의 잘못’은 대표치 및 산포도에 대한 부적절한 기술, 자료의 성격을 고려할 때 잘못된 분석 방법을 사용한 경우, 통계 분석의 결과와 다른 결론을 유도한 경우 등을 포함하였다.

통계 전문가가 개별 논문을 대상으로 통계점검표를 작성하였다. 원문은 대한방사선종양학회 홈페이지에서 다운로드 하여 사용하였다. 일차 평가자가 방사선종양학 전문가가 아닌 것에서 올 수 있는 평가 오류가 있을 수 있으므로 개별 통계점검표는 대한방사선종양학회 간행위원회에서 한 차례 더 세부 점검을 하였다.

작성된 통계점검표의 통계 분석은 SAS (version 9.0, SAS Institute, NC, USA) 소프트웨어를 이용하였고, 빈도분석을 시행하여 각 항목의 빈도와 백분율을 산출하였다.

2006년 1월부터 2007년 12월까지 대한방사선종양학회지에 게재된 논문은 총 77편이었으며, 이 중 원문을 다운로드 할 수 없었던 4편을 제외한 73편의 게재 논문을 대상으로 평가를 진행하였다. 통계 전문가에 의하여 분석된 73편의 논문 중 종설은 없었고, 증례보고는 5편, 원저 논문은 68편이었다(Table 1). 이 중 46편의 논문에서 통계적 추론을 하였고, 16편의 논문에서 단순한 기술통계를 사용했으며, 11편의 논문에서는 통계적 기법을 사용하지 않았다.

추론통계를 사용한 46편의 논문에서 사용된 추론통계의 횟수는 각각 분할표분석 17회(37.0%), 비교통계분석 23회(50.0%), 회귀분석 7회(15.2%), 상관분석 5회(10.9%), 생존분석 27회(58.7%)였다. 그 외 Hierarchical Clustering, Mantel-Haenzel Chi-square test, Analysis of Covariance (ANCOVA), Receiver Operating Curve (ROC) Analysis, Duncan test, Scheffee’s test가 각각 1회씩 사용되었다. 사용된 각각의 통계 기법 별 분포는 Table 2와 같다.

대상 논문 중 통계 적용의 오류가 없는 논문은 19%였다(Table 3). 기초자료에 대하여 불충분하게 기술하거나 적용된 통계 기법의 종류 및 처리 결과의 검정력과 신뢰구간에 대하여 불충분하게 기술한 ‘생략의 잘못’은 34편(50.0%)의 논문에서 총 50회 관찰되었다. ‘시행의 잘못’은 35편(51.5%)의 논문에서 총 47회 관찰되었다. 자료의 척도나 특성을 고려하지 않고 무분별하게 평균이나 표준편차, 표준오차 등을 사용한 경우가 4회 관찰되었다. 온당하지 않은 분석 방법을 사용한 경우가 10회 관찰되었다. 표본에 대한 통계 처리를 전혀 하지 않은 상태에서 모집단에 대한 결론을 유

Table 1. The Type of Research and the Statistical Method used in the Journal of the Korean Society for Therapeutic Radiology and Oncology from 2006 to 2007

	Year		Total (n=73)
	2006 (n=39)	2007 (n=34)	
Type of research			
Original article	35	33	68 (93.2%)
Case report	4	1	5 (6.8%)
Use of statistical method			
Descriptive plus inferential statistics	20	26	46 (63.0%)
Descriptive statistics	11	5	16 (21.9%)
No statistics	8	3	11 (15.1%)

**Table 2. The Incidence of Inferential Statistics used in the Journal of the Korean Society for Therapeutic Radiology and Oncology from 2006 to 2007**

Statistical method	Year		Total (n=46)
	2006 (n=20)	2007 (n=26)	
Chi-square test	5	5	10 (21.7%)
Fisher's exact test	2	5	7 (15.2%)
ttest	6	6	12 (26.1%)
Paired ttest	1	1	2 (4.3%)
ANOVA test	3	3	6 (13.0%)
Repeated measures ANOVA test	0	1	1 (2.2%)
Mann-Whitney U test (Wilcoxon rank sum test)	1	0	1 (2.2%)
Kruskal-Wallis test	0	1	1 (2.2%)
Wilcoxon signed rank test	0	0	0 (0.0%)
Pearson correlation analysis	3	1	4 (8.7%)
Spearman correlation analysis	0	1	1 (2.2%)
Linear regression	0	1	1 (2.2%)
Logistic regression	2	4	6 (13.0%)
Survival analysis	13	14	27 (58.7%)
Others	1	5	6 (13.0%)

**Table 3. Number of Articles with Statistical Errors in the Journal of the Korean Society for Therapeutic Radiology and Oncology from 2006 to 2007**

Type of statistical error	Year		Total (n=68)
	2006 (n=35)	2007 (n=33)	
1. Errors of omission			
① Incomplete description of basic data	5	7	12 (17.6%)
② Incomplete description of statistical test performed	12	18	30 (46.9%)
③ Incomplete description of statistical power	3	5	8 (11.8%)
2. Errors of commissioning			
① Inadequate description of measures of central tendency or dispersion	3	1	4 (5.9%)
② Incorrect analysis	4	6	10 (14.7%)
③ Multiplicity on hypothesis testing	0	0	0 (0.0%)
④ Unwarranted conclusion			
- with no statistical test	5	2	7 (10.3%)
- illogical conclusion	3	9	12 (17.6%)
- inadequate statistical terminology	3	6	9 (13.2%)
- extrapolation of study results	3	2	5 (7.4%)
3. No statistical error	9	4	13 (19.1%)

추한 경우가 7회, 통계 분석 결과를 확장하여 비논리적 결론을 유도한 경우가 12회, 연구 결과를 일반 집단에 외삽(extrapolation)하면서 그 한계를 언급하지 않은 경우가 5회 관찰되었다. 통계 용어를 잘 못 사용한 경우도 9회 관찰되었다. ‘생략의 잘못’과 ‘시행의 잘못’이 모두 발견된 논문은 21 (30.9%)편이었다.

연도별로 2006년과 2007년 사이에 통계점검표를 이용한 평가 결과의 차이는 없었다.

### 고안 및 결론

대한방사선종양학회지에 2년간 게재된 논문에서 통계적 기법의 적용, 통계 방법론, 결과의 서술과 해석 등 다양한 영역에 걸쳐 크고 작은 통계적 오류가 있음을 확인하였다. 통계 적용의 오류가 없는 논문은 19%였다. 외국의 연구에 의하면 전문학회지에서 논문 투고 시 적절한 통계적 처리를 강조하고 이를 요구한 경우 투고 논문에 사용된 통계 분석의 질이 발전하였다.<sup>9,10)</sup> 이 결과를 볼 때 대한방사선

종양학회지의 투고 규정과 저자점검표에 통계적 기법의 활용에 관한 점검 항목 추가하고, 통계 오류에 관한 적절한 심사를 하는 것이 필요하다고 판단된다.

연구자가 통계에 관한 기본 지식을 가지지 않고서는 연구 자료로부터 통계적 추론을 통해 올바른 결론에 도달할 수 없다. 물론 통계전문가에게 의뢰하여 통계 분석의 신뢰성을 향상시키기 위한 노력을 할 수 있다. 그러나 통계전문가도 연구 논문의 요점을 이해하지 못하는 경우 연구 데이터의 성격 자체만으로 분석을 시행하는 오류를 일으킬 수 있다. 그러므로 대한방사선종양학회지의 질적 측면을 향상시키기 위하여 통계 심사를 강조하는 것 이외에도 회원을 대상으로 통계방법론에 대한 꾸준한 교육을 병행하는 것이 필요하다고 생각된다.

대한방사선종양학회지 게재 논문 중 추론통계의 방법론을 적용한 연구 논문은 63%로 대한이비인후과학회지의 39.2%보다는 많았고 대한응급의학회지의 69.1%와는 비슷한 분포를 보였다.<sup>5,7)</sup> 이러한 정도의 차이는 해당 전문학술지에 게재되는 논문의 특성을 반영하는 것으로 생각된다. 추론통계 중 높은 빈도로 사용된 통계 방법론은 생존분석(58.7%), 비교통계분석(50%), 분할표분석(36.9%)의 순서였다. 회귀분석(15.2%)과 상관분석(10.9%)은 상대적으로 적은 빈도로 사용되었다. 대한방사선종양학회지 게재 논문에는 생존분석이 58.7%로 가장 많이 사용된 반면 대한임상병리학회지와 대한마취과학회지의 게재 논문의 경우 생존분석이 사용된 경우가 없었고, 대한성형외과학회지, 대한응급의학회지, 대한이비인후과학회지의 경우 극히 일부의 논문에서만 생존분석이 사용되었다.<sup>3~7)</sup> 이는 대한방사선종양학회지의 게재 논문이 방사선치료 후 연구 대상 환자들의 생존을 보고와 생존율에 영향을 미치는 요인에 대한 추론을 시행한 연구가 많았기 때문으로 생각된다.

대한응급의학회지 게재 논문의 경우 ‘생략의 잘못’ 항목이 70.6%, ‘시행의 잘못’ 항목이 29.6%로 ‘생략의 잘못’이 더 많았던 것에 비하여 대한방사선종양학회지 게재 논문의 경우 두 가지 잘못이 각각 50% 가량의 논문에서 관찰되어 비슷한 정도였다.<sup>7)</sup> 이는 대한방사선종양학회지 게재 논문의 통계 적용 오류가 통계 분석 방법론의 전체 항목에 걸쳐 골고루 발생하고 있음을 의미한다.

본 연구에서는 연구 결론의 신뢰성에 영향을 미칠 수 있는 통계 오류인 ‘시행의 잘못’이 적지 않은 빈도로 일어남을 확인하였다. ‘시행의 잘못’ 항목 중 높은 빈도로 발생한 오류는 통계 분석 결과를 확장하여 비논리적 결론을 유도한 경우(17.6%), 온당하지 않은 분석 방법을 적용한 경우(14.7%), 통계학적 연관(association)과 인과 관계(causal relation)

등을 혼동하여 사용하는 등 통계용어를 잘못 사용한 경우(13.2%)의 순서였다. 아무런 통계 분석이 없이 결론을 유추하는 경우도 10%의 게재 논문에서 관찰되었다. 온당하지 않은 분석 방법을 적용한 경우는 자료의 성격과 맞지 않는 통계방법을 사용한 경우, 불충분한 표본 수에서 Chi-square test를 시행한 경우가 대부분이었다. 이와 같은 ‘시행의 잘못’ 항목의 오류는 개별 논문뿐만 아니라 대한방사선종양학회지의 질적 평가에 영향을 미칠 수 있는 부분이라고 판단된다.

그러나 본 연구에서 19%의 게재 논문만이 통계 적용의 오류가 없었다는 사실을 19%의 논문만이 신뢰할만하다는 의미로 판단해서는 안 될 것이다. 본 연구에서는 통계 오류를 통계점검표 항목별 빈도만으로 측정하였고 심각하고 중대한 통계 오류와 사소한 잘못을 구별하지 않았다. 한치의 오류도 없이 연구가 이루어지기는 어려우므로 중요한 것은 어떠한 통계 오류가 연구의 타당성과 신뢰성에 치명적인 것인지 결정될 필요가 있다.<sup>8)</sup> 모든 유형의 통계 오류를 심각하게 받아들임으로써 세세한 잘못들이 필요 이상으로 확대 해석되고 중요한 연구 업적들이 평가절하될 수 있다. 따라서, 향후 대한방사선종양학회지의 투고 과정에서 통계 오류에 관한 심사를 강화하더라도 이 점을 간과해서는 안 될 것이다. 점검표의 활용에 의해 통계 오류가 지적될 경우라도 지적된 오류가 전체 연구를 무효화할 정도로 부적절한 내용인지 반드시 고려해야 한다.

결론적으로 본 연구에서는 대한방사선종양학회지에 게재된 논문에서 다양한 영역에 걸쳐 크고 작은 통계적 오류가 존재함을 확인하였다. 향후 대한방사선종양학회지 투고 논문의 심사 과정에서 통계 오류에 관한 적절한 심사를 추가하는 것이 필요하다고 판단된다. 또한, 대한방사선종양학회지의 질 관리 측면에서 게재 논문의 통계 방법론의 적정성에 대한 주기적인 모니터링이 필요하리라 판단된다.

## 참 고 문 헌

1. 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;1:85-87
2. 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;64:607-611
3. 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;26:22-27

4. 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;19:460–464
5. 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;42:1364–1369
6. 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;28:302–309
7. 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;14:475–480
8. 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
9. Schor S, Karten I. Statistical evaluation of medical journal manuscripts. JAMA 1966;195:1123–1128
10. Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. Circulation 1980;61:1–7

---

*Abstract*

## Statistical Errors in Papers Published in the Journal of the Korean Society for Therapeutic Radiology and Oncology

Hee Chul Park, M.D.\*, Doo Ho Choi, M.D.\*, Song-Vogue Ahn, M.D.<sup>†</sup>, Jin Oh Kang, M.D.<sup>†</sup>, Eun-Seog Kim, M.D.<sup>§</sup>, Won Park, M.D.\*, Seung Do Ahn, M.D.<sup>||</sup>, Dae Sik Yang, M.D.<sup>¶</sup>, Hyong Geun Yun, M.D.<sup>#</sup>, Eun Ji Chung, M.D.\*\*, Eui Kyu Chie, M.D.<sup>††</sup>, Hongryull Pyo, M.D.<sup>††</sup> and Semie Hong, M.D.<sup>§§</sup>

Department of Radiation Oncology, \*Sungkyunkwan University School of Medicine, Seoul <sup>†</sup>Yonsei University College of Medicine, Seoul, <sup>††</sup>Kyung Hee University School of Medicine, Seoul, <sup>§</sup>Soonchunhyang University College of Medicine, Asan, <sup>||</sup>University of Ulsan College of Medicine, Ulsan <sup>¶</sup>Korea University College of Medicine, Seoul, <sup>#</sup>Dongguk University College of Medicine, Seoul, <sup>\*\*</sup>NHIC Ilsan Hospital, Goyang, <sup>††</sup>Seoul National University College of Medicine, Seoul, <sup>††</sup>Proton Therapy Center, National Cancer Center, Goyang, <sup>§§</sup>Konkuk University School of Medicine, Seoul, Korea

**Purpose:** To improve the quality of the statistical analysis of papers published in the Journal of the Korean Society for Therapeutic Radiology and Oncology (JKOSTRO) by evaluating commonly encountered errors.

**Materials and Methods:** Papers published in the JKOSTRO from January 2006 to December 2007 were reviewed for methodological and statistical validity using a modified version of Ahn's checklist. A statistician reviewed individual papers and evaluated the list items in the checklist for each paper. To avoid the potential assessment error by the statistician who lacks expertise in the field of radiation oncology; the editorial board of the JKOSTRO reviewed each checklist for individual articles. A frequency analysis of the list items was performed using SAS (version 9.0, SAS Institute, NC, USA) software.

**Results:** A total of 73 papers including 5 case reports and 68 original articles were reviewed. Inferential statistics was used in 46 papers. The most commonly adopted statistical methodology was a survival analysis (58.7%). Only 19% of papers were free of statistical errors. Errors of omission were encountered in 34 (50.0%) papers. Errors of commission were encountered in 35 (51.5%) papers. Twenty-one papers (30.9%) had both errors of omission and commission.

**Conclusion:** A variety of statistical errors were encountered in papers published in the JKOSTRO. The current study suggests that a more thorough review of the statistical analysis is needed for manuscripts submitted in the JKOSTRO.

---

**Key Words:** Statistics, Radiation Oncology