

최근 3년간(2010-2012) 한국한의학연구원논문집 게재 논문의 통계기법에 관한 연구

강경원, 이민희, 김정은, 이상훈, 최선미

한국한의학연구원

Analysis of Articles Published in the Journal of Korea Institute of Oriental Medicine - from 2010 to 2012

Kyungwon Kang, Minhee Lee, Jungeun Kim, Sunmi Choi

Korea Institute of Oriental Medicine

Background and Purpose : This study was to investigate statistical validities and trends of previously reported papers that used various statistical techniques such as t-test and analysis of variance.

Methods : To analyze the statistical procedures, 38 original articles using those statistical methods were selected from Journal of Korea Institute of Oriental Medicine(JKIOM) published from 2010 to 2012.

Results : Analysis of variance and t-test were used in 20 papers (38.5%), 16 papers (30.8%) of 52 papers. Four articles(10.5%) did not report α values and nineteen papers(50.0%) of 38 ones were not tested for normal distribution. Five papers (13.2%) misused t-test and 3 papers (7.9%) did not carry out the multiple comparison.

Conclusions : To improve the quality of JKIO, The participation of statisticians in research design will reduce the significant errors in statistical interpretation of the results.

Keywords : Statistical error, Statistical method, Journal of Korea Institute of Oriental Medicine

I. 서론

통계분석은 자료를 분석하고 객관적이고 타당한 결론을 도출하는데 중요한 역할은 한다¹⁾. 이런 역할을 달성하기 위해서는 임상계획서에 연구의 목적을 명확히 정의해야 하며, 연구의 목적을 바탕으로 타당한 연구방법 및 통계적 방법에 대한 내용을 기술해야 한다²⁾. 그러나, 대부분 연구자들은 통계에 대한 체계적인 교육을 받지 못한 상태에서 적절하지 못한 연구방법 및

통계기법 선택으로 인해 통계적 오류를 초래하고 있다. 이러한 그릇된 선택은 외국의 유명 의학학술지에도 예외가 아닌 실정인으로서, 게재된 논문의 70- 80% 이상에서 통계적 방법을 사용하고 있으며, 그 중 40-70%에서 통계적 오류가 발견된다고 한다³⁻⁵⁾. 국내에서도 부적절한 표본의 선택, 잘못된 통계적 기법의 적용, 결과해석에서의 오류 등 다양한 형태의 통계적 오류가 있었다고 보고하였다⁶⁻⁹⁾. 이와 같이 연구방법 및 통계적 방법에서 객관적이지 못할 경우에 논문의 그릇된 결론을 도출할 수 있기 때문에 올바른 통계 기법의 적절한 선택과 적용여부가 바로 과학적 연구의 성패를 가늠한다고 할 수 있다¹⁰⁾. 따라서 한국한의학연구원논문집에 게재된 최근 3년간(2010-2012)

접수 ▶ 2012년 10월 31일 수정 ▶ 2012년 10월 31일 채택 ▶ 2012년 11월 22일

교신저자 최선미, 대전광역시 유성구 유성대로 1672 한국한의학연구원

Tel 042-868-9485 Fax 042-863-9464 E-mail smchoi@kiom.re.kr

의 논문들을 중심으로 통계기법의 적용과 통계적 오류를 평가 분석하여 향후 보다 적절한 통계기법의 사용에 대한 개선책을 제시하고자 한다.

II. 연구방법

1. 연구기간 및 대상

연구대상은 2010년도 제16권 1호부터 2012년도 18권 2호까지 발간된 한국한의학연구원 논문집(The Journal of Korea Institute of Oriental Medicine: 이하 JKIO)의 총 125편 논문을 대상으로 통계기법을 사용한 논문 38편을 대상으로 하였다.

2. 방법

JKIO에 게재된 논문들을 통계기법의 종류와 통계기법의 오류로 나누어 조사가 되었다. 통계 기법의 종류는 그래프의 이용, 통계분석 프로그램의 이용, 통계분석 기법의 이용으로 세분화하여 조사가 되었으며, 통계기법 오류는 통계처리에 대한 서술의 오류, 유의수준 오류, 중심경향(대표치) 또는 산포도에 대한 기술의 오류, 정규분포성에 대한 기술의 오류, 다중비교 시행에 대한 오류, 반복 측정된 자료에 대한 분석의 오류 등으로 나누어 조사하였다.

III. 결과

1. 통계기법의 종류

1) 그래프의 이용

통계분석 또는 빈도표를 제시한 논문 59편 중 34편(57.6%)의 논문에서 그래프를 이용하였으며, 총 47회의 그래프가 이용되었다. 막대그래프가 19건(40.4%)으로 가장 많이 이용되었으며, 다음으로는 선형으로 이루어진 그래프가 16건(34.0%) 사용되었다. 그 외 원 그래프, 히스토그램, 점그래프, ROC curve 등이 사

용되었다 (Table 1).

〈Table 1〉 Number of graphs in JKIO

Vol.	Chart type						Total
	bar	histo-gram	line	pie	scatter	ROC curve	
16	8	1	5	2	0	0	16
17	6	2	6	2	1	1	18
18	5	0	5	3	0	0	13
Total	19	3	16	7	1	1	47

2) 통계분석 프로그램의 이용

통계분석을 사용한 38편의 논문에서 이용한 분석 프로그램 중 어떤 프로그램을 이용하였는지 밝힌 경우가 29건(76.3%)이었으며, 이 중 SPSS를 이용한 경우가 25편(65.8%)으로 가장 많았다. 그 외 GraphPad PRISM, Origin, sigma plot, SigmaStat이 각 1편에서 사용되었다. 어떤 프로그램을 이용하였는지 밝히지 않은 경우는 9편(23.7%)으로 나타났다 (Table 2).

〈Table 2〉 Number of articles classified by analysis program in JKIO

Vol.	Analysis Program						Total
	SPSS*	GraphPad PRISM	Origin	sigma plot	SigmaStat	Unknown	
16	11	1	1	1	0	4	18
17	9	0	0	0	1	4	14
18	5	0	0	0	0	1	6
Total	25	1	1	1	1	9	38

* SPSS: Statistical Package for Social Science

3) 통계분석 기법의 이용

통계분석을 이용한 38편의 논문에서 통계분석 기법은 총 52회 이용되었으며, 이 중 비모수적인 방법은 8회(15.4%) 이용되었다. 모수적인 방법 중에서는 단일 표본 검정방법이 1회(1.9%) 이용되었고, 두 그룹을 비교하기 위한 검정이 15회(28.8%), 세 그룹 이상을 비교하기 위한 검정이 20회(38.5%) 이용되었다. 그 외 카이제곱 검정, 회귀분석 등의 검정이 이용되었으며, 어떤 검정법을 사용하였는지 언급하지 않은 경우도 2회(3.8%) 있었다 (Table 3).

<Table 3> Number of methods of analysis in JKJOM

methods of analysis	Vol.			Total
	16	17	18	
single sample t-test	0	1	0	1
Student's t-test	2	5	4	11
paired t-test	2	2	0	4
ANOVA*	8	5	4	17
Repeated Measured ANOVA	0	2	1	3
Mann-Whitney U test	3	0	1	4
Wilcoxon signed rank test	1	1	0	2
Kruskal-Wallis	2	0	0	2
Chi-square	2	0	1	3
Regression	0	1	1	2
ROC curve	0	1	0	1
Unknown	2	0	0	2
Total	22	18	12	52

* ANOVA: analysis of variance

2. 통계기법의 오류

1) 통계처리에 대한 서술의 오류

사용한 통계처리에 대하여 서술하는 방법에 있어서, 38편의 논문 중에서 4편(10.6%)이 통계처리 방법을 밝히지 않았다. 16권에서는 통계처리 방법을 밝히지 않은 비율이 16.7%였던 것에 비하여 17권에서는 7.1%, 18권에서는 0%로 통계처리에 대하여 서술을 하는 비율이 높아졌다 (Table 4).

<Table 4> Number of Description about Statistics Methods in JKJOM

Vol.	Description of Statistical methods		
	None	Described	Total
16	3	15	18
17	1	13	14
18	0	6	6
Total	4	34	38

2) 유의수준(α)에 대한 오류

유의수준은 0.05를 사용한 논문이 32편 (84.2%)으로 가장 많았는데, 논문에서 유의수준을 0.05/0.01/0.001을 함께 사용한 논문 1편은 0.05로 포함해서 합계를 구했다. 유의수준을 제시하지 않고 p-value만 기록한 것이 4편 (Unknown, 10.5%) 있었으며, 유의수준과 p-value를 모두 제시하지 않은 논문이 1편

(None, 2.6%)이었다 (Table 5).

<Table 5> Number of articles classified by significant level in JKJOM

Vol	Significant Level				Total
	0.05	0.1	Unknown	None	
16	14	1	3	0	18
17	12	0	1	1	14
18	6	0	0	0	6
Total	32	1	4	1	38

3) 중심경향(대표치) 또는 산포도에 대한 기술의 오류

중심경향과 산포도에 대한 기술에 있어서 평균±표준편차 (Mean±SD)로 나타낸 경우가 24편 (63.2%)으로, 평균±표준오차 (Mean±SE)로 나타낸 6편 (15.8%)보다 많았다. 그리고 평균±표준편차인지 평균±표준오차인지 명시하지 않는 경우가 3편 (7.9%)이었으며, 평균만 제시한 경우가 1편 (2.6%)으로 나타났다. 범주형 변수로만 이루어져 평균이 아닌 빈도만 제시한 경우는 4편 (10.5%)이었다. 그리고 비모수 검정을 시행한 논문 중 중심경향과 산포도를 중앙값과 사분위범위 등으로 나타낸 경우는 없었다 (Table 6).

<Table 6> Number of articles classified by describing central tendency and dispersion in JKJOM

Vol.	Central tendency and Dispersion					Total
	Frequency	Mean	Mean±SD*	Mean±SE**	Unknown	
16	3	1	11	3	0	18
17	0	0	8	3	3	14
18	1	0	5	0	0	6
Total	4	1	24	6	3	38

* SD: Standard Deviation, ** SE: Standard Error

4) 정규분포성에 대한 기술의 오류

표본 집단의 정규분포성을 검정하여야 할 논문 38편 중 정규분포성에 대해 기술된 논문은 1편 (2.6%)이었지만, 정규분포 검정을 하였다고만 명시하였고 정규분포를 만족하는지에 대한 결과는 제시되어 있지 않았다. 또한 정규분포성을 검정하지 않은 논문 중 표본의 크기가 30례 이상이어서 정규분포를 가정할 수 있는 경우는 18편 (47.4%)이었으며, 표본의 크기를 언급하지 않은 경우가 13편 (34.2%)으로 나타났다 (Table 7).

<Table 7> Number of articles classified by describing normal distribution in JKIOm

Vol	Normal Distribution				Total
	Normal	Above 30cases	Unknown	None	
16	1	8	6	3	18
17	0	6	5	3	14
18	0	4	2	0	6
Total	1	18	13	6	38

5) 다중비교 시행에 대한 오류

세 그룹 이상을 비교하기 위한 분석 방법이 쓰인 논문은 총 22편이었는데, 이 중 분산분석을 시행한 경우가 17편이었고, t-test를 이용한 경우가 5편이었다. 분산분석을 시행한 논문 17편 중 다중비교를 시행한 경우가 14편 (82.4%)이었고, 다중비교를 시행하지 않은 경우가 3편 (17.6%)이었다. 또한 다중비교를 t-test로 수행한 후 유의수준을 보정해주지 않은 경우가 1편으로 나타났다 (Table 8).

<Table 8> Number of articles classified by post hoc in JKIOm

Vol	post hoc							Total
	Duncan	Dunnnett	LSD	Scheffe	t-test	Tukey	None	
16	2	2	2	0	0	1	1	8
17	2	1	0	1	1	0	0	5
18	1	0	0	1	0	0	2	4
Total	5	3	2	2	1	1	3	17

6) 반복 측정된 자료에 대한 분석의 오류

한 개체에 대하여 여러 번 반복적으로 측정되어 반복측정 분석법을 써야하는 논문 5편 중 반복측정 분산분석법을 이용하여 바르게 분석된 논문은 3편이었으며, 2편은 쌍체 t 검정 또는 분산분석법을 이용하여 결과를 나타내었다.

IV. 고찰 및 결론

연구자들은 적절한 통계기법을 선정하는데 있어서 많은 오류를 범하고 있는데, 이는 통계에 대한 개념을 명확히 이해하지 못하거나 또는 통계가 필요하다고 느

끼는 경우라도 이를 적절하게 사용하기가 어려워 연구 결과를 제시할 때에 통계 기법을 충분히 활용하지 못하는 경우가 있으며, 최근에는 사용자 중심 통계분석 프로그램의 보급으로 쉽게 통계결과를 얻고 있지만 정확하고 적절한 통계기법에 의한 것이었는지는 여전히 의문으로 남아있다¹¹⁾.

분석 결과를 제시하기 위해서는 표와 그래프 등을 이용할 수 있는데 통계분석을 하거나 빈도표를 제시한 총 59편의 논문 중 그래프를 이용한 논문은 34편 (57.6%)이었으며, 총 47회의 그래프가 이용되었다. 막대그래프와 선그래프가 각 19건 (40.4%), 16건 (34.0%)으로 가장 많이 이용되었다. 그래프를 이용하면 표로 제시되는 것보다 결과를 빠르게 한 눈에 파악할 수 있기 때문에 결과의 비교가 용이하므로 적절히 그래프를 이용할 필요가 있다.

통계분석을 위한 프로그램으로는 통계분석을 사용한 38편의 논문 중에서 SPSS를 이용한 논문이 25편 (65.8%)로 가장 많았으며, 이용한 프로그램을 밝히지 않은 경우가 9편 (23.7%)으로 나타났다. 분석 결과에 대한 신뢰도를 높이기 위하여 이용한 프로그램, 버전이 통계분석방법에서 기술되어야 한다.

적용된 통계방법 중에는 분산분석이 전체 52건 중 20건 (38.5%)으로 가장 많이 사용되었고, 그 다음으로 전체 52건 중 16건 (30.8%)으로 t-검정 순이었다. t-검정 및 분산분석으로 대표되는 모수적 방법으로 36건 (69.2%), Mann-Whitney U test 및 Wilcoxon signed rank test로 대표되는 비모수적 방법으로는 8건 (15.4%)이었다. 박용규⁷⁾에 의하면 카이제곱검정이 31.7%로 가장 많이 적용되었고, 그 다음으로 t-검정과 분산분석 그리고 상관분석 순으로 많이 적용되는 통계분석이라고 보고하고 있다. 이 연구에서는 연구자들이 사용한 통계분석은 비교적 쉬운 통계기법을 사용하고 있다고 볼 수 있는데, 고급 통계기법이 폭넓게 이용될 수 있도록 통계교육 및 통계전문가의 도움이 필요할 것으로 생각된다.

통계 검정법에는 모수적 검정과 비모수적 검정으로 나누어지는데, 모수적 검정이란 정규분포 (normal distribution)를 하는 자료의 평균을 비교하는 대부분의 검정, 즉 t-검정, 분산분석, 및 회귀분석을 일컫는다. 모수적 검정을 하기 위해서는 정규분포 외에도 각

집단의 분산이 서로 같아야 한다는 요구 조건이 만족되어야 한다. 반면 비모수적 검정은 자료가 정규분포를 하는 경우, 모수적 검정에 비해 낮은 검정력을 갖지만, 정규분포가 아닌 경우에는 모수적 검정보다 더 높은 검정력을 갖는다¹²⁾ (Table 9). 모수적 검정에서는 일반적으로 평균과 표준편차를 제시하고 비모수적 검정에서는 중앙값과 범위를 제시하는데, 검토된 논문 중 비모수 검정을 하였음에도 중앙값과 범위를 제시한 논문은 없었다. 이 연구결과에서는 정규성 검정을 하지 않는 경우가 19편 (50.0%)이었으며, 기술만 한 경우가 1편 (2.6%)이었고, 30례 이상이어서 정규분포를 가정할 수 있는 경우가 18편 (47.4%)이었다.

〈Table 9〉 모수적 검정과 비모수적 검정

모수적 검정		비모수적 검정
Pearson correlation	⇔	Spearman correlation
Paired t-test	⇔	Wilcoxon의 부호순위검정
Two-sample t-test	⇔	Wilcoxon의 순위합검정
일원 분산분석	⇔	Kruskal Wallis 검정
랜덤화블록 계획법	⇔	Friedman 검정

적절한 통계기법의 선정은 첫째 분석하고자하는 변수의 수, 둘째 표본 집단의 수, 셋째 측정된 자료의 형태에 따라서 결정된다¹³⁾ (Table 10). 집단이 3개 이상이며, 정규성을 만족하면 분산분석을 시행하여야 하는데, 본 연구에서 t 검정과 관련된 오류로는 보정 없이 t 검정을 반복 사용한 경우가 5건이 있었는데, 이는 3군 이상간의 비교에서 보정 없이 반복적으로 t 검정을 사용함으로써 α 가 증가되어 잘못된 결과를 초래할 수 있다¹¹⁾. 카이제곱 검정은 자료의 빈도 또는 비율을 비교하는 방법으로 전체 표본수가 20이하이거나, 어느 한 셀의 기대치가 5이하인 경우에는 카이제곱 검정을 그대로 적용하는 것은 옳지 않으며 연속성을 수정하거나 Fisher의 정확성 검정을 사용한 경우는 없었다. 분산분석을 한 결과 유의한 차이가 있을 때, 구체적으로 어떤 집단 간에 차이가 있는지를 추가로 검정할 필요가 있는데, 이 때 사용하는 검정방법이 다중비교(multiple comparison)인데, 모든 집단들간을 비교하는 경우에서 각 집단의 표본수가 같다면 Tukey, Duncan, Newman-Keuls 등의 검정방법을 사용하고, 각 집단의 표본수가 다르다면 Scheffe의 방법을 사용한다⁷⁾.

집단의 수	명 목 (범주형)	순 위	구간, 비(연속)
2			
Related(쌍체표본)	McNemar	Wilcoxon signed rank	Paired t
Independent	Chi-square Fisher exact	Wilcoxon rank sum (Mann-Whitney U)	Two-sample t
3 이상			
Related	Cochran Q	Friedman	Randomized block Multivariate ANOVA Repeated measures
Independent	Chi-square	Kruskal-Wallis	ANOVA

한 개체에 대하여 여러 번 반복적으로 측정된 자료를 분석하기 위한 분석법으로는 반복측정 분산분석법 등 반복 측정 자료를 위한 분석 방법을 이용해야 한다. 일원분산분석을 이용하는 경우 개체 내 차이에 대한 고려가 되지 않고 각 관측치가 서로 독립이 아니기 때문에 결과가 왜곡될 수 있다. 검토된 논문 중 반복적으로 측정된 자료를 이용한 논문은 전체 38편 중 5편 (13.2%)이었는데, 그 중 반복측정 분산분석법을 이용하여 바르게 분석된 논문은 3편이었다.

회귀분석은 독립변수와 종속변수의 결합 관계를 알아보기 위하여 수행되는데, 오차항의 등분산성과 독립성, 정규성을 만족하여야 수행할 수 있다. 검토된 논문에서 회귀분석이 이용된 논문은 2편이었는데, 2편 모두 회귀식은 제시하였지만 오차항에 대한 검정은 하지 않았으며, 한 편은 회귀식의 R²만 제시하고 그 외의 p-value는 제시하지 않았다.

연구 시작 전 통계방법론의 기술에서 반드시 포함해야 하는 항목은 양측/단측 검정, 사용한 통계프로그램, 유의수준(α), 검정력 등이다. 이 연구에서 유의수준을 나타내지 않고 연구결과와 함께 p-value를 제시한 논문이 4편(10.5%)이었다. 이러한 오류는 유의수준에 대한 이해가 부족한 것으로 생각되며, 연구방법에서 미리 설정을 해서 연구를 진행해야 할 것으로 생각된다.

본 연구의 제한 점으로는 통계기법의 오류에 대한 설명을 포괄적으로 기술함으로써 각각의 통계기법에 대한 전제조건, 장점, 및 단점에 대해서 독자들이 간과할 우려가 있기 때문에 연구 시작 전에 충분히 통계전문가와 논의를 할 필요가 있다고 생각한다.

이 결과를 통해 앞으로 한의학 연구를 수행하는데 있어 한의학에 적용될 수 있는 통계방법론의 개발의

필요해 보이며, 올바른 통계기법 사용과 통계프로그램의 이용에 대한 교육이 선행되어야 할 것이다. 또한 연구의 시작부터 통계전문가의 참여로 통계적인 오류를 사전에 방지하는 노력이 필요할 것으로 보인다.

V. 감사의 글

이 논문은 2012년도 한국한의학연구원의 「고혈압 위험군에 대한 침구경락요법 효과 근거확립」 과제의 연구비 지원으로 수행되었습니다. (연구과제번호: K12203)

VI. 참고문헌

1. 박희철, 최두호, 안성복, 강진오 등. 「대한방사선종양학회지 게재 논문의 통계적 오류 현황」. 『대한방사선종양학회지』. 2008;26(4):289-294.
2. 식품의약품안전청. 「임상시험계획서 및 결과보고서의 통계적 고려사항」. 2009.
3. 강현철. 「최근 3년간 (1999-2001) 대한간호학지의 통계기법 활용에 대한 분석」. 『대한간호학회지』. 2002;32(6):929-935.
4. 박동규, 최용근, 김영수, 신상완. 「대한치과보철학회지에 게재된 논문의 통계적 오류: 국문논문과 영문논문의 비교」. 『대한치과보철학회지』. 2009;47(3):273-285.
5. Altman DG. 「Statistical Reviewing for Medical Journals」. 『Stat Med』. 1998;17:2661-2674.
6. 박용규. 「1993년 3월 게재 논문의 통계적 오류」. 『가정의학회지』. 1999;20(4):412-413.
7. 박용규. 「가정의학회지 논문의 통계적 적용」. 『가정의학회지』. 1998;19(5):137-143.
8. 안윤옥, 고응린. 「자료처리 과정에 대한 통계적 검토- 일부 의학잡지에 게재된 논문 예를 중심으로」. 『예방의학학회지』. 1973;6(1):81-85.
9. 이형기, 허봉렬, 안윤옥. 「1980년대에 발표된 국내 의학 연구 논문의 방법론 및 통계처리기법의 타당성에 관한 평가연구」. 『가정의학회지』. 1991;12(6):46-67.
10. 이선미, 이승욱. 「국내 간호학 연구논문에 활용된 통계 기법의 타당성 평가 연구」. 『한국보건통계학회지』. 1998;3(1):42-64.
11. 유태영, 진영호, 정태오, 이재백. 「대한응급의학회지에 게재된 논문의 통계기법에 관한 고찰」. 『대한응급학회지』. 2003;14(5):475-480.
12. 박용규. 「가정의학회지 논문의 통계적 적용」. 『가정의학회지』. 1998;19(5):137-143.
13. 박용규. 「학회지 게재논문의 통계기법 적용: 논문심사의 관점에서」. 『가정의학회지』. 1998;19(11):1093-1098.