

원저

《대한침구학회지》에 실린 중재 관련 비무작위 비교연구논문의 질적 평가

선종인 · 백용현 · 이상훈 · 이재동

경희대학교 한의과대학 침구학교실

Abstract

Quality Assessment of Non-randomised Comparative Clinical Studies in *The Journal of Korean Acupuncture and Moxibustion Society*

Seon Jong-in, Baek Yong-hyeon, Lee Sang-hoon and Lee Jae-dong

Dept. of Acupuncture & Moxibustion, College of Oriental medicine, Kyung Hee University

Objectives : Since *The Journal of Korean Acupuncture and Moxibustion Society*(JKAMS) started publishing in 1984, randomised controlled trials(RCTs) and non-randomised clinical studies(NRSs) consistently have increased in quality and quantity. There has been no study on quality assessment of NRSs. Thus, this study evaluated non-randomised comparative clinical trials in the JKAMS to assess the quantity and quality.

Methods : Upon extracting NRSs for assessing the intervention effects from all the articles published in the JKAMS from 1984 to 2009, assessments were made on methodological index for non-randomised studies(MINORS). Also, the analysis were made upon the proportion of non-randomised comparative clinical trials within original articles. The mean scores by research methods and years, and total scores and mean scores of yearly research methods were analyzed.

Results : A total of 44 trials on non-randomised comparative clinical trials were selected. In 1980s, 4.3% of the total selected original articles, in 1990s, 4.1%, and in 2000s 14.2% were NRSs. According to the research designs, the mean scores of MINORS were Case-control study 18.3, Controlled before and after 18.5, Quasi randomised trial 18.5, and non randomised trial 17.9(out of 24), respectively. Cross-sectional study was the most frequently used(41.9%).

· 접수 : 2010. 1. 11. · 수정 : 2010. 1. 25. · 채택 : 2010. 1. 28.
· 교신저자 : 이재동, 서울특별시 동대문구 회기동 1번지 경희의료원 한방병원 침구과
Tel. 02-958-9208 E-mail : ljdacu@khmc.or.kr

Conclusion : Although NRSs consistently increased in quantity, the assessed mean scores were low and most articles used Cross-sectional study. Thus, there should be studied using appropriate research methods in the future.

Key words : The Journal of Korean Acupuncture and Moxibustion Society ; Non-randomised comparative clinical studies ; Quality Assessment ; Methodological index for non-randomised studies(MINORS)

I. 서론

1984년 《대한침구학회지》가 발간된 이후 초기의 학회지는 문헌연구가 대부분을 차지했으며, 2000년대에 접어들면서 다양한 연구설계 방법을 통한 논문의 많은 수적 증가를 보였다.

중재 효과를 보는 연구설계방법에는 크게 실험연구(Experimental study)와 관찰연구(Observational study)로 나눌 수 있다. 실험연구에는 무작위대조연구(Randomised controlled trial, RCT), 준무작위연구(Quasi-randomised trial, QRT), 비무작위연구(Non-randomised trial, NRT)가 있고, 관찰연구에는 대조전후연구(Controlled before and after study), 단면연구(Cross-sectional study), 증례연구(Case-series) 등이 있다¹⁾.

서 등²⁾은 1980년대 《대한침구학회지》에 게재된 논문을 분석한 결과, 문헌논문과 실험논문이 대부분을 차지하고 임상논문은 상대적으로 찾아보기 어려웠다고 하였다. 하지만 2000년대에 들어서면서 학회지에 게재된 임상연구논문은 관찰연구뿐만 아니라 실험연구의 수도 증가하였고 중재의 효과를 본 연구도 증가하였다.

실험연구 중 무작위대조연구는 치료적 중재효과를 평가하는 데 가장 믿을 만한 연구방법으로 비뚤림(bias)을 최소화시키는 연구설계방법이다. 무작위대조연구의 경우 Jadad 질 평가기준(Jadad quality assessment scale) 등을³⁾ 이용하여 질 평가를 수행한다. 하지만 연구의 윤리적인 문제, 비용적인 문제, 연구 여건 등의 문제로 인해 무작위대조연구를 시행할 수 없는 연구주제는 비무작위연구(Non-randomised study), 즉 준무작위연구, 비무작위연구/준실험연구, 대조전후연구, 전후연구 등을 통해 연구를 수행하게 된다. 연구주제에 따라서는 환자-대조군 연구나 코호트 연구 등의 디자인을 통하여 중재 효과를 보는 경우도 있다.

비무작위연구는 치료 효과를 평가하는 데 있어 무작위대조연구에 비해 비뚤림 위험은 더 크지만 연구의 시행가능성을 높이는 연구설계방법이다.

연구의 질은 연구의 비뚤림 정도를 말하는 것으로 외적 타당성, 정밀도 등과는 구별된다. 연구 질 평가의 목적은 연구의 비뚤림 정도를 정확하게 파악하여 연구 결과를 적용할 때 이를 반영하기 위함이다.

기존 연구 중 《대한침구학회지》에 실린 논문의 현황과 분석에 대한 논문이 있었고²⁾, 통계방법에 대한 질 평가가 시행된 경우도 있었지만⁴⁾ 꾸준한 양적 성장을 보인 비무작위연구에 관한 질 평가는 아직까지 시행된 적이 없었다.

이번 연구에서 중재 관련 비무작위 비교연구, 즉 준무작위연구, 비무작위연구/준실험연구, 대조전후연구, 환자대조군연구 등의 연구에 대한 질 평가를 시행하고자 한다.

II. 대상 및 방법

1. 대상

1984년도 창간호부터 2009년 26권 4호까지 26년 동안 발간된 《대한침구학회지》에 실린 논문 1,737편을 분석하였다. 논문 중 중재효과를 본 논문을 뽑았으며 이중 무작위대조연구를 제외한 비무작위연구를 평가 대상논문에 포함시켰다. MINORS 평가는 비무작위연구 중 비교연구에 국한하였다(Fig. 1).

2. 비무작위연구의 분류

비무작위연구의 분류는 2003년 National Health Service(NHS)에서 사용한 방법과 정의를 따르되, 부족한

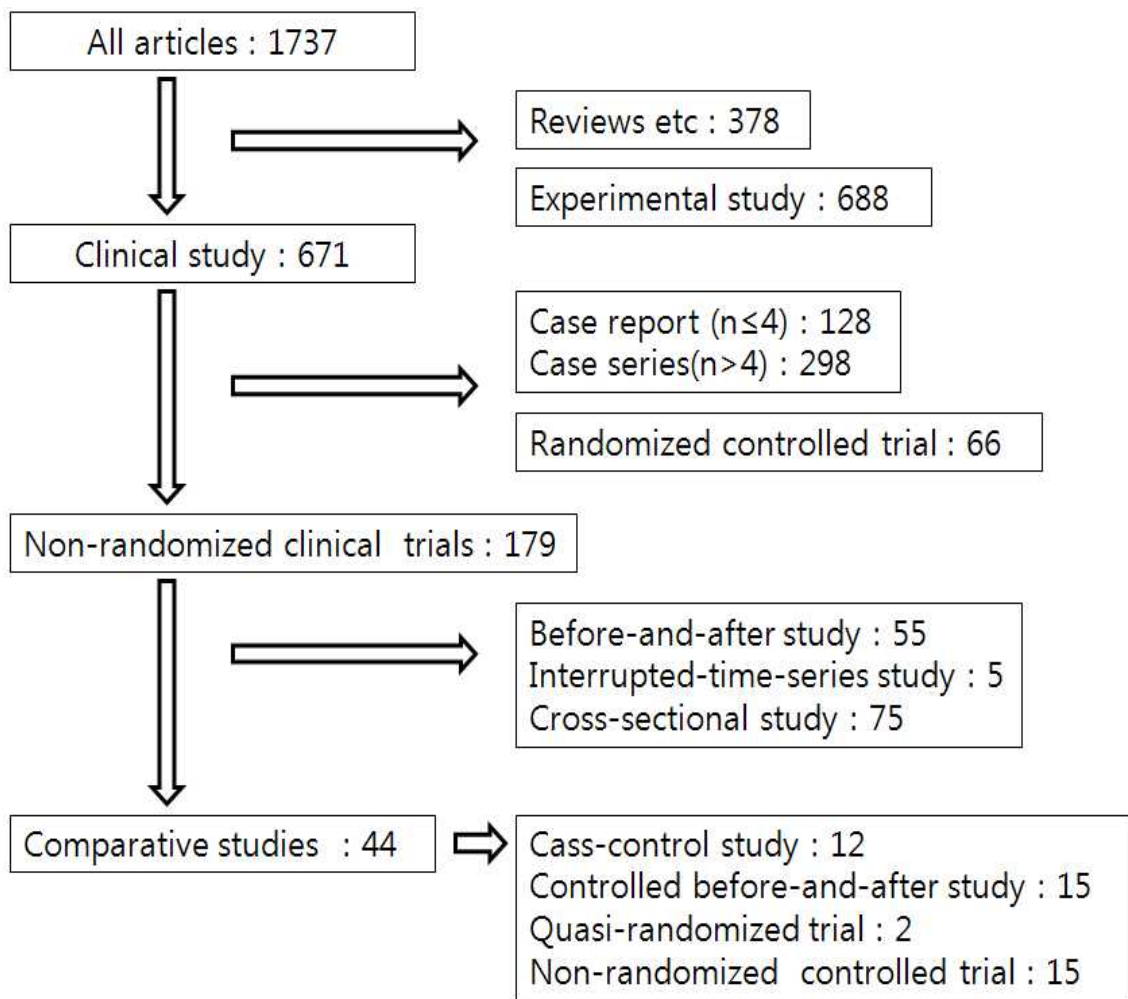


Fig. 1. Flow chart of the study - selection of the materials by research method

부분은 다른 문헌에서 보충하여 분류하였다(Table 1)¹⁵⁾.

비무작위연구를 비교대상이 없는 연구, 즉 전후연구(Before and after study, BAS), 단속적시계열 연구(Interrupted time series, ITS)와 비교대상이 있는 연구, 즉 대조전후연구(Controlled before and after study, CBA), 준무작위연구, 비무작위시험으로 나누어 분석하였다.

각각의 정의에 따라 침구과 전공자 2인이 개별적으로 평가를 하였고 평가 차이는 제3자가 조정을 하여 비무작위연구논문을 최종 분류하였다.

3. 분류된 비무작위 비교연구논문의 질 평가 도구

비무작위 비교연구의 질 평가는 MINORS로 하였다. MINORS는 비무작위연구를 평가하기 위해 개발

되었으며 내적 일관성, 외적 타당도, 평가자간의 동의, 검사-재검사법으로 검증된 평가도구이다.

평가는 각각 ① 분명한 목적이 있는지, ② 환자가 연속적으로 포함되었는지, ③ 데이터가 전향적으로 수집되었는지, ④ 연구목적에 적절한 결과인지, ⑤ 연구결과가 비뉘림없이 평가되었는지, ⑥ 추적기간은 적절했는지, ⑦ 탈락이 5% 미만인지, ⑧ 연구크기가 전향적으로 계산되었는지 등의 공통항목이 있고, 대조군이 있는 경우, ⑨ 적절한 대조군이 있었는지, ⑩ 연구군 모집이 동시적이었는지, ⑪ 연구군이 기저상태에서 유사한지, ⑫ 적절한 통계분석이 이루어졌는지에 대한 항목을 평가하도록 되어 있다. 비교가 없는 비무작위연구는 ①~⑧번까지 공통항목 8개만을 평가하며 비교대상이 있는 연구에 대한 항목이 ⑨~⑫번까지 4개를 추가하여 평가한다. 개별 문항에 대해서는 0점(not reported), 1점(reported but inadequate), 2점

Table 1. Taxonomy of Study Designs to Assess the Effectiveness of an Intervention

Experimental designs

A study in which the investigator has control over at least some study conditions, particularly decisions concerning the allocation of participants to different to different intervention groups.

① Randomised controlled trial

Participants are randomly allocated to intervention or control groups and followed up over time to assess any differences in outcome rates. randomisation with allocation concealment ensure that on average known and unknown determinants of outcome are evenly distributed between groups

② Quasi-randomised trial

Participants are allocated to intervention or control groups by the investigator, but the method of allocation falls short of genuine randomisation and allocation concealment(e.g. allocated by date of birth, hospital record number, etc.)

③ Non-randomised trial / Quasi-experimental study

The investigator has control over the allocation of participants to groups, but does not attempt randomisation(e.g. patient or physician preference). Differs from a “cohort study” in that the intervention is experimental rather than observational.

Observational designs

A study in which natural variation in interventions(or exposure) among study participants is investigated to explore the effect of the interventions(or exposure) on health outcomes.

④ Controlled before-and-after study

A follow-up study of participants who have received an intervention and those who have not, measuring the outcome variable both at baseline and after the intervention period, comparing either final values if the groups are comparable at baseline, or change scores. It can also be considered an experimental design if the investigator has control over, or can deliberately manipulate, the introduction of the intervention.

⑤ Concurrent cohort study

A follow-up study that compares outcomes between participants who have received an intervention and those who have not. Participants are studied during the same(concurrent) period either prospectively or, more commonly, retrospectively.

⑥ Historical cohort study

A variation on the traditional cohort study where the outcome from a new intervention is established for participants studied in one period and compared with those who did not receive the intervention in a previous period, i.e. participants are not studied concurrently.

⑦ Case-control study

Participants with and without a given outcome are identified(cases and controls respectively) and exposure to a given intervention(s) between the two groups compared.

⑧ Before-and-after study

Comparison of outcomes from study participants before and after an intervention is introduced. The before and after measurements may be made in the same participants, or in different samples. It can also be considered an experimental design if the investigator has control over, or can deliberately manipulate, the introduction of the intervention.

⑨ Cross-sectional study

Examination of the relationship between disease and other variables of interest as they exist in a defined population at one particular time point.

⑩ Case series

Description of a number of cases of an intervention and outcome(no comparison with a control group)

⑪ Interrupted-time-series studyz

A study that uses observations at multiple time points before and after an intervention(the ‘interruption’). The design attempts to detect whether the intervention has had an effect significantly greater than any underlying trend over time.

Table 2. Methodological Items for Non-randomised Studies

Methodological Items for Non-randomised Studies(MINORS)
① A clearly stated aim : the question addressed should be precise and relevant in the light of available literature
② Inclusion of consecutive patients : all patients potentially fit for inclusion(satisfying the criteria for inclusion) have been included in the study during the study period(no exclusion or details about the reasons for exclusion)
③ Prospective collection of data : data were collected according to a protocol established before the beginning of the study
④ Endpoints appropriate to the aim of the study : unambiguous explanation of the criteria used to evaluate the main outcome which should be in accordance with the question addressed by the study. Also, the endpoints should be assessed on an intention-to-treat basis
⑤ Unbiased assessment of the study endpoint : blind evaluation of objective endpoints and double-blind evaluation of subjective endpoints. Otherwise the reasons for not blinding should be stated
⑥ Follow-up period appropriate to the aim of the study : the follow-up should be sufficiently long to allow the assessment of the main endpoint and possible adverse events
⑦ Loss to follow up less than 5% : all patients should be included in the follow up. Otherwise, the proportion lost to follow up should not exceed the proportion experiencing the major endpoint
⑧ Prospective calculation of the study size : information of the size of detectable difference of interest with a calculation of 95% confidence interval, according to the expected incidence of the outcome event, and information about the level for statistical significance and estimates of power when comparing the outcomes
Additional criteria in the case of comparative study
⑨ An adequate control group : having a gold standard diagnostic test or therapeutic intervention recognized as the optimal intervention according to the available published data
⑩ Contemporary groups : control and studied group should be managed during the same time period(no historical comparison)
⑪ Baseline equivalence of groups : the groups should be similar regarding the criteria other than the studied endpoints. Absence of confounding factors that could bias the interpretation of the results
⑫ Adequate statistical analysis : whether the statistics were in accordance with the type of study with calculation of confidence intervals or relative risk

(reported and adequate)으로 나누어 점수를 측정해 총점을 대조군이 없는 경우 16점, 대조군이 있는 경우 24점 만점으로 평가하였다(Table 2)⁶⁾.

점수 측정은 침구과 전공자 2인이 개별적으로 측정하였고 평가에 대한 차이는 제3자가 조정하여 점수를 부여하였다.

4. 자료의 분석

1) 비무작위연구논문의 양적 변화

1983~1990년, 1991~2000년, 2001~2009년으로 크

게 3군으로 나누어 각 군의 총 연구논문에서 비무작위연구논문이 차지하는 비중과 연도별 연구설계 방법에 따른 분류를 시행하였다. 또한 연구설계방법에 따른 비율로 분석하였다.

2) 비무작위 비교연구논문의 질적 평가

MINORS를 이용해 연도별 연구설계방법에 따른 총점과 연도별 평균점수를 분석하였고, 각 평가 항목별 점수와 평균 점수를 분석하였다.

Ⅲ. 결 과

1. 비무작위연구의 양적 변화

포함된 중재관련 비무작위연구논문은 총 179편이었다. 연도별로 1984~1990년까지는 4편으로 원저의 4.3%에 불과하였지만, 1991~2000년까지는 24편(4.1%),

Table 3. The Quantative Change of Non-randomised Clinical Trials According to the Years

Years	Articles(no)	Non-randomized trials(%)
1984~1990	93	4(4.3%)
1991~2000	577	24(4.1%)
2001~2009	1067	151(14.2%)
	1737	179(10.3%)

Table 4. Classification & Rate of the Research Method According to the Years

Year	BAS	ITS	CSS	CCS	CBA	QRT	NRT	Total
1986	1	0	0	0	0	0	0	1
1987	0	0	0	0	0	0	1	1
1988	0	0	1	0	0	0	0	1
1990	0	0	0	0	0	0	1	1
1992	0	0	1	0	0	0	1	2
1994	1	0	0	0	0	0	0	1
1995	1	1	0	0	0	0	1	3
1997	1	0	0	0	0	0	0	1
1998	1	0	1	0	0	0	1	3
1999	2	0	2	0	2	0	0	6
2000	3	0	4	0	1	0	0	8
2001	7	0	3	2	3	0	0	15
2002	4	0	3	3	3	0	0	13
2003	10	1	5	0	0	1	1	18
2004	6	1	8	3	2	1	2	23
2005	4	1	9	1	1	0	3	19
2006	11	0	9	1	2	0	1	24
2007	1	1	11	1	0	0	2	16
2008	1	0	12	1	1	0	1	16
2009	1	0	6	0	0	0	0	7
	55	5	75	12	15	2	15	179
Rate (%)	30.7	2.8	41.9	6.7	8.4	1.1	8.4	

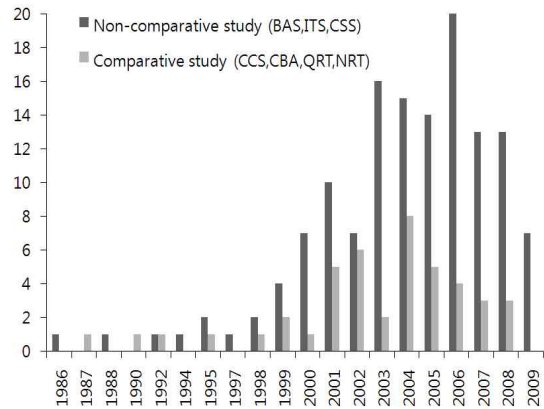


Fig. 2. Classification of the non-comparative studies vs comparative studies according to the years

2001~2009년까지는 151편(14.2%)으로 양적 증가를 보였다(Table 3).

연도에 따른 연구방법의 분류와 비율을 살펴보면 단면연구가 총 75편(41.9%)로 가장 많이 시행되는 연구 방법이었고, 전후연구 55편(30.7%), 비무작위시험과 대조전후연구 8.4%, 환자대조군연구 6.7% 단속적시계열연구 5편(2.8%) 등의 순서였다. 연구방법론 중 코호트연구는 없었다(Table 4).

비교군이 없는 연구 비율은 1990년대 65%(17편 중 11편), 2000년대 77%(159편 중 122편)로 급증하는 양상을 보였다(Fig. 2).

2. 비무작위 비교연구의 질적 평가

MINORS를 이용한 연도별 비무작위 비교연구논문의 평균점수는 16~19점의 분포를 보였다. 연도가 올라가면서 점수가 나아지는 경향을 보이지는 않았다. 1987년에 점수가 가장 낮았고(16점), 1998, 1999, 2000, 2003년에 가장 높았으며(19점), 2009년에는 18.7점이었다(Table 5).

설계방법에 따른 평균점수는 환자대조군연구는 18.3점, 대조전후연구 18.5점, 준무작위시험 18.5점, 비무작위시험 17.9점이었다. MINORS를 이용한 연구방법에 따른 항목별 점수를 보았을 때 2, 8, 10번 항목은 모든 연구방법에서 2점이었고, 5, 6항목은 0~0.2점, 0~0.4점이었다. 낮은 점수를 받은 문항은 ⑤ '연구결과가 비뮴립없이 평가되었는지', ⑥ '추적기간은 적절했는지'였다(Table 6).

Table 5. Relative & Total Mean Scores of the Research Method According to the Years Using MINORS

Year	CSS	CBA	QRT	NRT	Mean score(point)
1987	0	0	0	16	16
1990	0	0	0	17	17
1992	0	0	0	18	18
1995	0	0	0	17	17
1998	0	0	0	19	19
1999	0	19	0	0	19
2000	0	19	0	0	19
2001	17.5	19.6	0	0	18.6
2002	18	18	0	0	18
2003	0	0	19	19	19
2004	18.3	17	18	18	17.8
2005	17	15	0	19	17
2006	19	19.5	0	16	18.2
2007	18	0	0	18.5	18.3
2008	20	19	0	17	18.7

Table 6. The item Scores and Mean Scores According to Research Methods Using MINORS

Item	CSS	CBA	QRT	NRT
1	1.7	1.8	2	1.8
2	2	2	2	2
3	2	1.8	2	2
4	1.7	2	2	1.7
5	0	0.1	0	0.2
6	0	0.1	0	0.4
7	1.9	1.8	2	1.9
8	2	2	2	2
9	1.8	1.9	1.5	1.1
10	2	2	2	2
11	1.2	1	1	1.1
12	2	2	2	1.8
Mean score(point)	18.3	18.5	18.5	17.9

IV. 고 찰

1984년 《대한침구학회지》 창간호부터 2009년 26권 4호까지 비무작위연구논문은 많은 수는 아니지만

꾸준한 성장을 보였다. 1984~2009년 26권 4호까지 1,737편의 연구논문 중에서 1984~1990년까지 93편 중 4편으로 4.3%, 1991~2000년까지 577편 중 24편으로 4.1%, 2001~2009년까지 1,067편 중 151편으로 14.2%를 차지해 초기 20년에 비해 2000년대에 들어서면서 6배 이상의 양적 성장을 보였다. 비교가 있는 연구인 환자-대조군연구, 비교전후연구, 준무작위시험, 비무작위시험의 비율이 각각 6.7%, 8.4%, 1.1%, 8.4%로 낮게 나타났으며, 비교대상이 없는 연구는 약 74.6%를 차지했다. 비교대상이 없는 연구논문과 비교대상이 있는 연구논문을 비교했을 때 2000년대에 들면서 비교군이 없는 논문의 비율이 큰 폭으로 높아졌음을 알 수 있었다.

MINORS를 이용해 비무작위 비교연구의 질을 평가한 결과 연도별 평균점수가 적게는 16점에서 많게는 19점으로 전반적으로 중등도의 점수를 보였다. 연도별 평균점수는 예상과는 달리 연도별로 높아지는 경향을 보이지는 않았다. 침구학회 회원, 연구자, 간행위원회, 심사자 등을 대상으로 한 연구설계방법 교육이 필요하다고 할 수 있다.

《대한침구학회지》에 실린 논문들의 질이 다른 학회지나 다른 나라와 비교하여 어떠한지에 대해서는 직접적인 비교대상이 없어서 말하기 힘들다.

낮게 평가된 문항은 ‘비뿔림’과 관련된 문항에서 0~0.2점, ‘추적기간과 관련된 문항’에서 0~0.4점을 나타내 거의 0점이었다. 이 항목은 이중맹검을 실시하지 않은 것이나, 단회성 상태 파악으로 인해 낮은 점수를 받게 된 것이다. 향후 연구 수행과 보고에서 이 부분에 대한 집중적인 관심이 필요하다고 할 수 있다.

근거중심의학이 도입되면서 Jadad scale³⁾ 같은 질 평가 도구가 개발되었으며 《대한침구학회지》에서도 연구 논문에 관한 분석을 수행한 연구논문이 발표되었으나²⁾, 비무작위연구의 질 평가 도구는 그리 많지 않다. Deeks 등¹⁾은 현재까지 나온 비무작위연구의 질 평가 도구를 검토하였고, 6개의 내적 타당도 항목 중 5개 이상, 4개의 중심 항목 중 3개 이상이 포함되어야 좋은 도구로 인정되는데 이를 만족하는 도구는 6개였다⁷⁻¹²⁾. 하지만 이들 도구들은 측정해야 할 항목이 많고, 측정하는데 많은 시간이 필요해 논문의 질을 평가하는데 실제적으로 적용하기는 어려운 단점이 있었다. MINORS는 12개의 항목과 0~2점으로 평가할 수 있고 신뢰도와 타당도가 입증되어있는 평가도구로 이러한 제한점을 극복할 수 있는 장점이 있어 본 연구의 질 평가 도구로 이용했다.

이번 연구는 《대한침구학회지》에 실린 비무작위 연구논문을 처음으로 양적 및 질적 평가를 한 의의가 있다. 현재까지 《대한침구학회지》에 게재된 비무작위 연구의 경우 양적 성장은 있었으나 논문 대부분이 비교대상이 없는 단면연구였고, 질도 낮게 평가되었다.

결론적으로 비무작위연구의 양적 증가에 따르는 질적 향상을 위해서는 점수가 낮은 항목을 잘 고려해 연구를 설계하고, 질 평가에 대한 교육을 실시하고, 체크리스트 개발을 통해 향상시키며, 연구설계방법의 다양화를 통해 객관적인 결과를 얻어 그 결과를 임상에 적용할 수 있도록 해야겠다.

V. 결 론

포함된 증재 관련 비무작위연구논문은 총 179편이었다. 연도별로 1984~1990년까지는 4편으로 원저의 4.3%에 불과하였지만, 1991~2000년까지는 24편(4.1%), 2001~2009년까지는 151편(14.2%)로 양적증가를 보였다. 연구설계방법의 종류에 따른 평균점수는 환자대조군연구 18.3점, 비교전후연구 18.5점, 준무작위시험 18.5점, 비무작위시험 17.9점으로 측정되었다. 가장 흔한 연구설계방법은 단면연구로 41.9%를 차지했고 QRT, NRT는 각각 1.1%, 8.4%로 많지 않았다.

VI. 참고문헌

- Deeks JJ, Dinnes J, D'Amico R, Sakarovitch C, Song F et al. Evaluating non-randomised intervention studies. *Health Technol Assess.* 2003 ; 7 : 1-173.
- 서정철, 남상수, 이재동, 최도영, 안병철, 박동석, 이운호, 최용태. 대한침구학회지에 수재된 논문에 대한 분석고찰. *대한침구학회지.* 1999 ; 16(1) : 125-46
- Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ et al. Assessing the quality of reports of randomised clinical trials : Is blinding necessary? *Control Clin trials.* 1996 ; 17 : 1-12.
- 이승덕. 대한침구학회지 논문의 통계적 오류에 관한 연구. *대한침구학회지.* 2004 ; 21(1) : 176-88
- National Institute for Health and Clinical Excellence. *Methods for development of NICE public health guidance.* London : National Institute for Health and Clinical Excellence. 2006.
- Slim K, Nini E, Forestier D, Kqiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomised studies(minors-MINORS) : development and validation of a new instrument. *ANZJ Surg.* 2003 ; 73 : 712-6.
- Cowley DE. Prosthesis for primary total hip replacement : a critical appraisal of the literature. *Int J Technol Assess Health Care.* 1995 ; 11 : 770-8.
- Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health.* 1998 ; 52 : 377-84.
- Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Tugwell P. *The-Newcastle-Ottawa Scale(NOS) for assessing the quality of non-randomised studies in meta-analyses.* Ottawa : Clinical Epidemiology Unit. 1999.
- Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. *Pediatrics.* 1989 ; 84 : 815-27
- Thomas H. *Quality assessment tool for quantitative studies.* Effective public health practice project. Toronto : McMaster University. 2003.
- Zaza S, Wright-De Aguero LK, Briss PA, Truman BI, Hopkins DP, Hennessy MH et al. Data collection instrument and procedure for systematic reviews in the Guide to community Prevention Services : Task Force on Community Preventive Services. *Am J Prev Med.* 2000 ; 18(1) : 44-74.