



침 임상연구방법론 연구: 요통 임상연구 사례를 중심으로

박유리

대한한의사협회 글로벌센터, 한의학교육평가원

Study on Clinical Research Methodologies for Acupuncture: Based on Review of Clinical Studies on Back Pain

Yu Lee Park

Global Center, The Association of Korean Medicine, Institute of Korean Medicine Education and Evaluation

Objectives : The results from randomized controlled trials on traditional medicine have been negative or ambiguous. The appropriateness of the research methodology started to be questioned. However, the research methodology appropriate for studying traditional medicine has not been well-established. This study aims to explore the appropriate research methodologies for clinical studies on acupuncture. **Methods :** Clinical researches evaluating the efficacy and effectiveness of acupuncture for back pain were systematically searched and evaluated focusing on research methodology. The acupuncture treatment protocol is evaluated by 9 items which is individual treatment based on diagnosis, combining proximal and distal acupoints, exact location of acupoint, De Qi, adequate stimulation, duration of acupuncture, treatment sessions, treatment interval, and establishing relationship between patients and doctors. **Results :** Systematic review of acupuncture researches on low back pain suggested that observational studies and pragmatic trials rather than explanatory studies could incorporate these elements. Qualitative studies also could help understanding complex characteristics of acupuncture treatment which could not be evaluated by quantitative studies. **Conclusions :** In designing clinical researches on acupuncture, various study designs, especially observational study and pragmatic study, should be considered according to research objectives and subjects. The efforts to incorporate the core characteristics of acupuncture practiced in everyday clinical setting are also essential.

Key words : acupuncture, clinical research, research design, systemic review

서 론

“근거중심의학(Evidence-Based Medicine, EBM)”이라는 용어는 1991년 캐나다 McMaster 대학의 임상역학자들에 의해 처음으로 사용하기 시작한 것으로 그들은 근거중심의학이 의학의 “새로운 패러다임 전환(Paradigm Shift)”이라고 주장하였다¹⁾. 즉, “근거중심의학”이란, 환자를 치료하는데 있어 현존하는 최선의 근거를 신중하고 세심하게 활용하는 의료행위로, 근거중심의학 이전의 의료

는 임상경험과 병리적 이론, 전문가의 견해 등에 근거하여 이뤄졌지만, 근거중심의학이라는 새로운 패러다임에서의 의료는 환자에게 최적의 의료서비스를 제공하기 위해서 체계적 관찰에서 나온 임상연구의 근거들을 활용해야 한다는 것이다²⁾. 이러한 “근거중심의학”의 실질적인 정의는 의료행위에 있어서 다양한 의학적 근거 가운데 무작위배정 대조 연구(Randomized Controlled Trial, RCT)의 결과를 가장 중시하는 것이라고 할 수 있다³⁾.

이러한 분위기 속에서 한의학계에서도 “근거중심의학”의 개념

Received December 17, 2014, Revised December 18, 2014, Accepted December 18, 2014

Corresponding author: **Yu Lee Park**

Global Center, The Association of Korean Medicine, Institute of Korean Medicine Education and Evaluation, Heojun-ro 91, Gangseo-gu, Seoul 157-200, Korea
Tel: +82-2-2659-5064, Fax: +82-2-2657-5096, E-mail: kiki2877@naver.com

© This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

이 강조되기 시작하였고, 한의학 치료의 과학적 근거에 대한 요구에 의해 임상연구가 활발하게 이뤄지기 시작하였다. 따라서 한의학 임상연구 역시 한의학 치료가 과학적으로 '정말 효과가 있는가'라는 질문을 중심으로 이뤄졌고, 이를 위해 근거중심의학에서 가장 과학적인 연구방법론으로 인정받고 있는 RCT를 활용한 연구가 급증하였다⁴⁾. 하지만, 이러한 임상연구가 증가하면서 실제 임상에서 환자와 의료인에 의해 보고되고 있는 긍정적인 치료효과와는 달리 부정적이거나 애매한 연구 결과들의 축적이 증가하게 되었고, 이에 많은 연구자들은 연구방법론의 적절성에 대한 문제들을 제기하기 시작하였다. 무엇보다 서양의학의 치료법, 특히 신약의 효과를 증명하기 위해 개발된 RCT가 전혀 다른 철학적 배경을 갖고 있는 한의학 연구방법론으로서 적절한가에 대한 논의가 활발하게 이뤄지기 시작하였다⁵⁻⁷⁾.

기준에 이뤄진 한의학 임상연구방법론에 대한 논의들을 살펴보면, RCT가 연구방법론 가운데 가장 과학적인 연구방법이기 때문에 RCT의 틀은 유지하면서 한의학 진료의 특징을 반영할 수 있도록 설계해야 한다는 주장들이 있었고⁸⁾, 한편으로는 현행의 실험적 연구가 한의학 치료의 효과를 평가하는데 부적절함을 비판하며 관찰 연구나 질적 연구 등의 다른 대안을 제시하는 논의들이 있었다^{10,11)}. 특히, 한의학을 환자 개인에 맞는 개별화된 진단 및 치료를 통해 환자 스스로 정신적, 육체적 건강을 유지하고 회복할 수 있는 능력을 배양시키기 위한 의학(Whole Systems)이라고 규정하고, 이러한 특징을 잘 반영할 수 있는 연구방법(Whole Systems Research)을 개발하기 위해 활발하게 연구를 진행하고 있는 연구자들도 있었다¹²⁾.

또한, 일부의 학자들은 한의학의 발달과정과 임상특징을 고려할 때 현재 근거중심의학에서 강조하고 있는 근거피라미드의 순위가 한의학에는 적절치 못하기 때문에 한의학의 근거피라미드 순위는 서양의학에서 중시되는 순위와 반대로 적용되어야 함을 주장하기도 하였으며¹³⁾, 혹은 근거의 수직적 상하관계가 아닌 수평적 관계를 강조하는 학자들도 있었다¹⁴⁾.

국내에서는 엄석기 등이 "역사적 근거중심의학"을 주장하며, 동아시아 전통의학의 전통적 근거는 문헌자료와 문헌에 근거한 임상가의 경험이고, 현대의학에서 전문가의 견해나 증례(군)보고, 기술적 연구 등이 근거로 인정되고 있는 것처럼 동아시아 전통의학의 전통적 근거 또한 '근거중심의학'의 근거의 범주에서 재해석되어야 함을 강조하며 6단계의 구체적인 연구단계를 제안하였다¹⁵⁾.

한편, 이처럼 다양한 논의가 있었음에도 불구하고, 지금까지 한의학적 특징에 근거하여 기존의 임상연구들을 연구방법론 중심으로 체계적으로 분석하거나 고찰하여 문제를 도출하고 이를 바탕으로 구체적인 연구방법론의 대안을 제시한 연구는 없었다. 이에 본

연구에서는 한의학의 특징적인 임상과정을 잘 반영하고 있으며, 환자에게 직접적으로 물리적, 심리적 영향을 끼치는 시술행위이기 때문에 무작위배정대조연구를 적용하는데 여러 가지 문제점이 제기되어 임상연구방법론에 대한 연구가 비교적 활발하게 이뤄진 침 임상연구에 국한하여 한의학적 특징을 가장 잘 반영할 수 있는 임상연구방법론을 모색하고자 한다. 기존의 많은 연구들에서 일반적인 임상연구방법론에 대한 고찰은 이뤄졌기 때문에 본 연구에서는 '요통'이라는 특정 질환에 대한 기존의 침 임상연구들을 조사하여 실질적인 사례들을 통해 침 임상연구에 활용된 다양한 연구방법론들을 정리하고, 연구설계를 비교 고찰함으로써 각 연구방법론의 장단점을 도출하고자 한다.

대상 및 방법

요통의 침 임상연구의 체계적 문헌고찰을 위해 다음과 같은 검색기준과 전략을 통해 논문을 검색하고 선택하였다.

1. 연구 선택 및 배제기준

연구 디자인은 요통 치료와 관계된 임상평가를 수반한 모든 연구디자인을 포함하였으며, 실험연구, 유사 실험적 연구, 모든 관찰 연구(코호트, 환자대조군연구, 단면연구, 증례군보고) 등이 포함되었다. 단, 기존의 연구 결과를 체계적으로 고찰한 연구나 연구대상이 한 명인 증례보고는 배제하였다.

연구 참여자의 특징은 이환기간이 4주 이상인 요통을 앓고 있는 성인으로, 요통은 비특이성 요통과 특이성 요통을 모두 포괄하였다. 한편, 4주 이하의 요통이나 임신성 요통처럼 일반적인 요통이 아니라 특수한 상황에서 발생하는 요통은 배제하였다.

침 시술법은 연구의 목적에 따라 한의학 이론에 근거 혹은 전통적 穴位에 자침한 연구, 자극에 있어서는 수기자극과 전침자극을 사용한 연구, 도구에 있어서는 호침을 사용한 침 임상연구를 포함하였다. 호침 외에 피내침 등 다른 도구를 활용한 연구들은 배제하였다.

대조군이 있는 연구의 경우, 플라시보 대조군, 현행 양방 치료 대조군, 보완대체의학 치료 대조군, 침 치료군을 모두 포함시켰다.

결과 평가에 있어서는 임상적 결과평가가 이뤄진 연구를 모두 포함시켰다. 따라서 요통 임상연구에서 가장 많이 활용되고 있는 통증 강도, 기능적 상태, 전반적 평가(전반적 호전, 호전된 환자의 비율, 주관적 통증 호전 평가), 직장 복귀 등의 기본적인 평가 중 한 가지만 만족시켜도 본 연구에 포함되었다.

2. 연구의 검색과 선택

위의 선정 기준에 근거하여 PUBMED(<http://www.pubmed.com>), EMBASE(<http://www.embase.com>), Cochrane Library (<http://www.thecochranelibrary.com>)의 검색엔진을 활용하여 논문을 검색하였으며, 기간은 2000년에서 2010년으로, 언어는 한국어, 영어, 중국어로 제한하였다. 검색어는 “acupuncture”, “back pain”, “low back pain”을 사용하였고, 연구 선택은 혼자서 시간 간격을 두고 반복적으로 체크하였다.

3. 결과 평가

각 연구디자인별로 연구설계의 전반적인 특징에 대하여 평가하고 비교하였다. 평가의 항목은 연구 환경, 연구대상자의 특징, 침 시술의 특징, 결과평가 항목 및 평가시기 4가지 항목으로 이뤄졌다. 이 중 침 시술의 특징은 임상 실제와의 근접성을 평가하기 위해 ‘침 시술은 단순히 침의 삽입을 뛰어넘어 다양한 요소들이 긴밀하게 연결되어 있어 정확하게 어떤 요소들이 특이적 효과를 내는 것인지 규정짓기 어려운 “복합적 치료”¹⁶⁾라는 전제하에서 선혈의 과정에 있어서 한의학적 진단과정(개별치료)과 配穴, 得氣, 手技, 留鍼시간, 치료횟수, 치료기간, 다른 치료의 병행 여부, 의료인과 환자의 긴밀한 관계형성의 9가지 항목^{17,18)}으로 평가하였다.

4. 통계적 분석

각 연구디자인별로 특징을 비교하기 위해 변수의 특징과 분포의 성격에 따라 Mann-Whitney test, independent t-test, chi-square test를 활용하였으며, 통계프로그램은 SPSS ver 18.0 for Windows(IBM, USA)를 사용하였다.

결 과

1. 체계적 고찰에 포함된 연구의 분류와 특징

총 987편의 논문이 검색되어 논문의 제목, 초록, 혹은 논문 전체의 내용에 근거하여 미리 계획된 포함기준과 배제기준에 따라 33편의 논문이 체계적 고찰에 포함되었다(Fig. 1).

최종적으로 선택된 33편의 논문 가운데 23편은 실험연구, 8편은 관찰연구, 2편은 질적 연구였으며, 실험연구는 Gartlehner 등¹⁹⁾의 연구에서 제안한 7개의 항목을 기준으로 실용적 연구와 설명적 연구로 분류한 결과, 17편은 설명적 연구, 6편은 실용적 연구로 분류되었다. 유사 실험연구와 관찰연구는 Linde 등²⁰⁾의 기준을 참고하여 분류한 결과, 5편의 코호트연구와 3편의 증례군보고가 있었다.

각 연구의 일반적인 특징을 연구환경, 연구참여자, 침 시술, 결과 평가를 기준으로 비교한 결과는 다음과 같았다(Table 1).

2. 실험연구와 관찰연구 비교

관찰연구와 실험적 연구의 특징을 비교한 결과, 가장 현저한 차이를 보이는 것은 관찰연구에서는 실험적 연구의 장기추적조사연구 외에는 모두 대조군이 없었다는 것과 실험적 연구는 모두 비특이적 요통을 대상으로 한 반면 관찰연구는 3편의 연구가 특이적 요통을 대상으로 연구하였다는 점이었다.

연구 설계의 각 항목에 대해 구체적으로 살펴보면, 연구 장소에 있어서 관찰연구는 실험연구에 비해 1차 의료기관에서 이뤄진 비율이 높았고(67% vs. 30%, $p=0.061$) 다기관 연구도 더 많았으나(44% vs. 26%, $p=0.314$), 그 차이가 통계적으로 유의미하지는 않았다. 또한, 연구 참여자의 나이나 성별, 요통의 이환기간 등에 있어

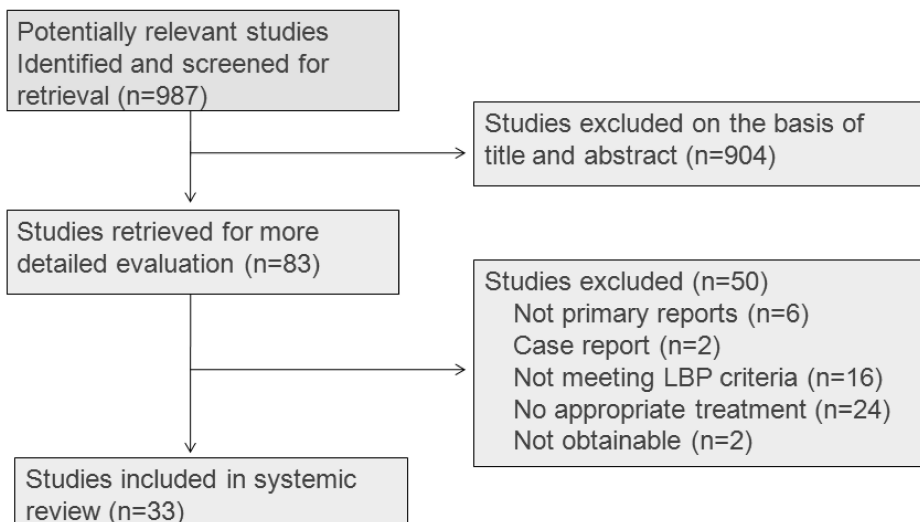


Fig. 1. Study flow diagram.

Table 1. Comparison of Experimental Studies and Observational Studies Included in Systematic Review

	Experimental Studies(n=23)		Explanatory vs. Pragmatic	Observational Studies (n=9)	Experimental vs. Observational
	Explanatory RCT(n=17)	Pragmatic RCT(n=6)	p value		p value
Research Site					
Primary health care center	1(5.9%)	6(100%)	<0.001	6(67%)	0.061
Multicenter(sites ≥2)	2(11.8%)	4(66.7%)	0.008	4(44%)	0.314
Characteristics of the participants					
Number median(range)	50(26~298)	450(75~3093)	<0.001	150(48~204,775)	0.046
Age(y) median(range)	53.8(27~76)	46.8(42.7~52.9)	0.150	52.1(32.5~58.2)	0.199
Proportion of female(%) median(range)	60(0~82.7)	59(50~62)	0.779	59(33~85)	0.801
Non-specific LBP	17(100%)	6(100%)	1	6(67%)	
Duration of LBP (y) median(range)	7(0.9~14.7)	7.2(0.3~8)	0.459	6.1(0.4~16)	0.985
Characteristics of Acupuncture Treatment					
Diagnosis(individual treatment)	9(52.9%)	5(83.3%)	0.190	9(100%)	0.027
Acupointscombination	13(76%)*	n.a		n.a	
Sensation(De Qi)	14(82%)	n.a		n.a	
Manipulation	9(53%)**	n.a		n.a	
Needle retention time(m) median(range)	20(0~30)	20(18~30)	0.883	30(20~30)	0.199
Number of treatment median(range)	10(1~20)	10(8.1~24)	0.433	8.7(7~24)	0.949
Period of treatment(w) median(range)	5(4~12)	11(5~12)	0.036	n.a	
Co-interventions	1(5.9%)	2(33.3%)	0.086	7(78%)	<0.001
Practitioner-patient relationship	0(0%)	5(83.3%)	<0.001	8(89%)	0.001
Evaluation of outcomes					
Quality of life and functional change	14(82.4%)	6(100%)	0.270	6(67%)	0.186
Long-term follow up ≥6m	5(29.4%)	5(83.3%)	0.022	5(56%)	0.538

LBP: Low Back Pain. *One study excluded, †Two studies excluded.

서도 큰 차이를 보이지 않았으나 관찰연구가 실험적 연구에 비해 더 많은 환자들을 대상으로 이뤄졌다는 점에서는 차이를 보였다(150 vs. 60, $p=0.046$).

관찰연구일수록 침 치료에 있어서 환자에 따른 개별진단과 치료를 허용하였고(100% vs. 61%, $p=0.027$), 침과 함께 뜸이나 부항과 같은 병행치료를 허용하였으며(78% vs. 13%, $p<0.001$), 의료인과 환자와의 치료적 관계형성도 자율적으로 허용하였다(89% vs. 21.7%, $p<0.001$). 한편, 한의학적 침 치료의 특징 가운데 配穴, 得氣, 手技 등의 항목에 대해서는 관찰연구와 실용적 임상연구 모두에서 명확하게 기술된 바가 없어 평가가 불가능하였다. 결과평가에 있어서는 실험적 연구에서 삶의 질이나 요통과 관련된 기능적 장애를 평가하는 도구를 사용한 연구가 더 많았지만(87% vs. 67%) 통계적으로 의미 있는 차이는 아니었으며($p=0.186$), 6개월 이상의 장기추적조사는 관찰연구에서 더 많이 이뤄졌으나 실험적 연구와 비교하여 큰 차이를 보인 것은 아니었다($p=0.538$).

3. 실험연구 중 설명적 연구와 실용적 연구의 특징 비교

설명적 연구와 실용적 연구는 여러 가지 설계상의 차이를 보였는데, 먼저 연구 장소에 있어서 실용적 연구일수록 일차의료기관에서 이뤄졌으며(100% vs. 6%, $p<0.001$), 여러 의료기관이 참여하는 경향성을 보였다(67% vs. 12%, $p=0.008$). 또한, 연구 참여자수에 있어서도 실용적 연구가 훨씬 많은 수의 인구집단을 대상으로 연구가 이뤄져서 통계적으로 유의미한 차이를 보였다(450 vs. 50, $p<0.001$).

침 치료의 특징에 있어서는 53%의 설명적 연구와 83%의 실용적 연구가 환자에 따른 개별 치료를 허용하여 절대적 수치로는 큰 차이를 보였으나 통계적으로 유의미한 차이는 아니었다($p=0.190$). 본 연구에서는 표준화된 穴位와 선택 가능한 穴位를 함께 제시한 반 표준화된 침 프로토콜과 아시혈 침법을 모두 개별 치료의 범주에 포함시켰다. 이에 반 표준화된 침 프로토콜과 아시혈 침법을 사용한 연구를 배제한 후 비교해본 결과, 실험연구에서는 17개 중 3개

Table 2. Acupoints Used in the Experimental Studies

Studies	Acupoints used in each study
Zaringhalam, 2010	BL: BL23, BL25, BL28, BL32, BL60/ GB: GB30, GB34
Cherkin, 2009	GV: GV4/ BL: BL23, BL40/ KI: KI3/ Ashi points
Itoh, 2009	BL: BL23, BL25, BL32, BL40, BL60/ GB: GB30, GB34
Szczurko, 2007	GV: GV3, GV4/ BL: BL23, BL25, BL40
Brinkhaus, 2006b*	(Standardized) GV: GV3, GV4, GV5, GV6/ BL: BL20-BL34, BL50-BL54/ GB: GB30 (Individual) GV: GV14, GV20/ BL: BL40, BL60, BL62/ GB: GB31, GB34, GB41/ LR: LR3/ KI: KI3, KI7/ SI: SI3
Itoh, 2004	BL: BL23, BL25, BL40, BL60/ GB: GB30, GB34/ Ashi points
Tsui, 2004	BL: BL25, BL26, BL40/ GB: GB30/ ST: ST36
Kerr, 2003	GV: GV4/ BL: BL23, BL25, BL40/ GB: GB30/ KI: KI3
Meng, 2003*	(Standardized) GV: GV3, GV4/ BL: BL23, BL24, BL25, BL28 (Individual) BL: BL36, BL37, BL40, BL54/ GB: GB30, GB31
Yeung, 2003	BL: BL23, BL25, BL40/ SP: SP6
Leibing, 2002	GV: GV3, GV4/ BL: BL23, BL25, BL31, BL32, BL40, BL60/ GB: GB34/ SP: SP6/ Extra points
Molsberger, 2002	BL: BL23, BL25, BL40, BL60/ GB: GB30, GB34
Carlsson and Sjölund, 2001*	BL: BL24, BL25, BL26, BL40, BL57, BL60/ LI: LI4, LI11

BL: Bladder meridian, GB: Gallbladder meridian, GV: Governor Vessel, KI: Kidney meridian, LR: Liver meridian, SI: Small intestine meridian, ST: Stomach meridian, SP: Spleen meridian, LI: Large intestine meridian. *Studies using semi-standardized acupuncture protocol.

의 연구만이 진정한 의미의 개별 치료를 허용하였고, 실용적 연구에서는 6개 중 4개의 연구가 개별 치료를 허용하여 통계적으로 유의미한 차이를 보였다($p=0.025$).

실제로 요통 치료에 활용한 穴位들을 제시한 설명적 연구들과 실용적 연구²¹⁻³³들을 토대로 임상연구에 활용된 穴位에 대해 정리한 결과는 다음과 같다(Table 2).

한편, 配穴과 得氣, 手技에 대해서는 실용적 연구에서 구체적으로 기술하지 않아 비교가 불가능하였으나, 아시혈을 사용한 연구를 제외한 대부분의 설명적 연구에서 近位穴과 遠位穴을 配穴하였으며, 得氣와 手技을 시도하였고, 手技은 得氣를 유발시키거나 유지하기 위한 자극법으로 시도되었다. 총 침 치료횟수는 두 연구 모두 중간값이 10회로 차이가 없었으나 설명적 연구에서는 1회의 치료만 이뤄진 연구가 있어 1회에서 24회의 넓은 범위를 보였고, 치료기간에 있어서는 설명적 연구의 중간값은 5주, 실용적 연구의 중간값은 11주로 유의미한 차이를 보였다($p=0.036$). 각 연구에서 사용된 총 침의 개수는 실용적 연구에서는 10개의 중간값을, 설명적 연구에서는 12개의 중간값을 보여 큰 차이가 없었으나, 실용적 연구에서 사용된 침의 개수의 범위는 10개에서 17개의 범위인 반면, 설명적 연구에서는 1개에서 20개로 치료횟수와 마찬가지로 극단적인 최소값과 최고값을 보였다. 환자와의 치료적 관계 형성 허용에 있어서 대부분의 설명적 연구에서는 통제하였으나 많은 실용적 연구에서 의사와 환자와의 관계 형성을 자유롭게 허용하여 큰 차이를 보였다($p<0.001$).

결과평가에 있어서는 두 연구 모두 비슷한 비율로 삶의 질, 요통

과 관련된 기능적 변화를 평가하는 도구를 활용하였고, 특히 모든 실용적 연구에서는 SF-12, SF-36과 같은 삶의 질을 평가하는 도구를 이용하여 침의 효과를 평가하였다. 한편, 6개월 이상의 장기추적 조사는 실용적 연구에서 월등하게 많았고(83% vs. 29%, $p=0.022$), 6개의 실용적 연구 가운데 5개의 연구에서 임상적 결과 평가와 함께 경제성 평가를 병행하였다^{22,24,34-36}.

4. 질적 연구

요통의 침 치료에 대한 질적 연구에 대해서는 각 연구의 연구방법과 연구주제, 핵심적인 결론에 대해 살펴보면, 다음과 같다.

먼저, MacPherson 등에 의한 연구³⁷는 침의 만성 요통 치료 효과를 평가하기 위한 실용적 연구³⁰에 병행되어 이뤄진 연구였다. MacPherson 등은 임상연구에 참여한 6명의 침시술자를 대상으로 치료를 마친 후 침 치료를 위한 진단과 침 치료과정의 핵심적인 요소들, 환자와 의료인의 치료관계, 침시술자들이 임상연구에 참여하면서 느낀 한계점 등에 대해 심층 면접을 하였다. 이 연구를 통해 그들은 다음과 같은 결론을 내렸다³⁷. 첫째, 침시술자들이 중시하는 침 치료과정의 핵심은 환자에 따른 개별 치료였다. 둘째, 침 치료는 치료 직후의 통증 완화 효과보다는 장기적인 치료효과를 목표로 하고 있었다. 셋째, 이러한 장기적 치료효과를 목표로 환자와의 신뢰관계 형성, 치료과정에 있어서 환자의 적극적 참여, 환자의 증상과 전신적인 상태에 따른 개인별 치료가 치료과정에서 중시되었다. 넷째, 이를 위해 침시술자들이 구체적으로 선택한 전략은 환자와의 적극적이고 개방적인 대화시도, 진단과정 중 환자와의 상호교

류, 한의학적 발병원인, 치료과정, 목표에 대한 설명을 통해 환자의 질병에 대한 이해 도모와 생활 변화의 동기유발 등이었다. 다섯째, 침시술자들이 치료 중에 환자들에게 제안하는 운동, 식생활 변화, 감정조절, 생활환경 변화 등에 대한 조언은 치료의 부수적인 부분이 아닌 한의학적 진단에 근거한 치료과정의 일부였다.

다음으로 Hsu 등에 의한 질적 연구³⁸⁾는 침, 마사지, 요가, 카이로프랙틱, 태극권과 심리요법 등의 요통 치료 효과를 평가하기 위한 5개의 무작위배정 대조 연구에서 환자들에게 전화 인터뷰를 통해 환자들이 주관적으로 느낀 치료효과에 대해 조사한 결과를 바탕으로 이뤄진 연구였다. 총 844명의 연구 참여자들 가운데 483명(57%)의 환자들이 침 치료를 받았는데, 침 치료를 받은 환자들의 응답만을 살펴보면 다음과 같다. 빈도순으로 새로운 치료법에 대한 가능성 인식과 그로 인한 희망이 생김이 가장 많았고, 다음으로 인체와 발병원인, 치료에 대한 인식의 전환, 인체에 대한 관심의 증가, 요통과 관련된 근육 및 전신적, 정신적 이완, 감정적 안정, 전반적 건강 증진, 마음과 영혼, 육체의 밀접한 관계에 대한 인식 등이었다. Hsu 등은 이 결과를 바탕으로 기존의 효과 평가도구들은 환자들이 주관적으로 느끼는 다양한 효과들을 평가에는 부족하다는 결론을 내렸다.

고찰

지금까지 요통의 침 치료 임상 연구에 대한 고찰을 통해 침 치료의 핵심적 요소 반영 여부를 중심으로 각 연구들의 연구 설계 특징들에 대해 살펴보았다. 그 결과, 각 연구마다 연구 디자인의 종류에 따라 서로 다른 특징들을 보였고, 침 치료의 핵심적 요소의 반영 여부에서도 큰 차이를 나타냈다. 이에 각 연구 디자인별로 고찰을 통해 도출된 각 연구의 장점과 한계점들에 대해 정리해보고자 한다.

실험적 연구의 가장 큰 장점은 무작위배정을 통해 비교집단 간의 동질성을 확보할 수 있다는 것이었다. 즉, 연구 결과의 내적 타당도가 높다는 점이다. 특히 실험적 연구 중 설명적 연구는 표준화된 혹은 부분적으로 선택 가능한 경혈들을 사용함으로써 연구의 재현성이 높았다는 점과 환자의 맹검을 통해 침의 비특이적 효과를 통제함으로써 '특정 穴位에 자침하는 것이 요통 치료에 효과가 있는가'라는 침의 특이적 효과 평가를 목적으로 하는 연구의 내적 타당도를 향상시킬 수 있다는 점이 장점이다. 한편, 17개의 설명적 연구 가운데 16개의 연구가 전문화된 3차 의료기관에서 이뤄졌고, 환자군을 모집할 때 협소한 포함기준과 많은 배제기준을 설정하여 지나치게 선택적인 환자집단을 대상으로 연구가 이뤄진 것 등으로

인해 연구결과의 일반화 가능성이 낮다는 문제점이 있다. 또한, 표준화된 침 치료 프로토콜이 과연 침 치료의 실재를 반영할 수 있을 것인가라는 문제, 설명적 연구에서 사용된 혈위들이 모든 환자에게 동일하게 적용될 수 있을 것인가와 선정 근거의 타당성 문제와 맹검을 통해 침의 비특이적 효과를 배제하는 것이 복합적 치료과정인 침 치료의 핵심을 배제하는 것이라는 문제점으로 인식되기도 한다. 즉, 설명적 연구를 침 시술의 핵심적인 요소들을 중심으로 분석해 본 결과, 많은 설명적 연구는 침을 다양한 치료 요소가 복합적으로 작용하는 치료과정으로 인식하기보다는 "침의 삽입(needleling)"이라는 행위로 단순화하여 '경혈에 침을 삽입하는 것이 과연 효과가 있는 것인가'라는 침의 특이적 효과만을 분리하여 평가하려는 연구였다. 이는 무작위배정 대조 연구가 신약의 효과 평가를 위해 개발된 연구디자인이라는 점을 고려해볼 때, 무작위배정 대조 연구의 전형적인 형태인 설명적 연구에서 침을 양약과 같이 단순하게 재구성하려는 시도가 이뤄지고 있다고 할 수 있으며, 이러한 시도는 서양의학의 환원주의적 관점에서 침의 정체성을 왜곡시킬 가능성이 있다고 할 수 있다.

실용적 연구의 장점은 모두 일차 의료기관에서 이뤄졌고, 환자의 모집도 실제 임상에서 이뤄지고 있는 양상대로 환자의 선택 혹은 의료인의 의뢰를 중심으로 이뤄졌다. 또한 환자의 포함기준 및 배제기준이 덜 엄격하여 다양한 환자를 포괄하였기 때문에 외적 타당도는 높았다. 또한, 침 시술의 양상도 실제 임상을 그대로 반영하려고 하였고, 평가 역시 본 연구에 포함된 모든 실용적 연구에서 삶의 질, 요통과 관련된 기능적 변화 등을 다각도에서 평가하였고, 6개 중 5개의 연구에서 6개월 이상의 장기적 추적조사를 통해 침 치료의 장기적 효과를 평가하였다. 한편, 실용적 연구에서는 개별적인 침 치료 시술을 허용함으로써 연구의 재현성이 낮다는 점과 플라시보 대조군 없이 대부분 현행의 양방 치료와 효과를 비교하였기 때문에 환자들의 맹검이 불가능했고, 침 치료군에 참여한 환자들이 갖고 있던 기대심리, 믿음 등 침의 비특이적 효과가 강하게 작용했을 가능성이 높음으로 인해 침의 특이적 효과를 평가하기 어렵다는 문제점이 있었다. 이와 함께 침 외에 뜸이나 부항 등 부수적인 치료법들을 병행하여 침 치료 패키지에서 어떤 치료가 효과를 나타냈는지 평가가 불가능하였으며, 연구에 참여한 의료인이 2명 이상인 경우가 많아 의료인의 편차가 결과에 끼칠 가능성도 있었다.

관찰연구는 실용적 연구와 비슷한 측면이 많았으며, 가장 큰 장점은 실제 임상을 그대로 반영할 수 있다는 점이었다. 또한, 많은 수의 환자를 확보할 수 있었다는 점, 본 연구에 포함된 관찰연구에서 알 수 있듯 디스크 탈출을 확진 받은 요통, 혹은 반척추분리증, 고질적인 요부 근막염과 같은 특이성 요통을 대상으로 연구할 수

있다는 점이 장점이었다. 한편, 실용적 연구와 마찬가지로 침 시술의 특이적 효과를 평가할 수 없다는 문제점과 대조군 부재로 인한 내적 타당성 저하, 평가시점에 참여자들의 낮은 응답률 등으로 인한 타당성 저하 등의 문제가 있었다.

마지막으로 질적 연구는 일반적으로 체계적 고찰 연구의 대상은 아니지만, 다양한 임상 관련 연구 디자인을 살펴보기 위해 본 연구에 포함시켰으며, 가장 큰 장점은 양적 연구로 반영할 수 없었던 침 치료의 과정, 치료의 목표와 환자들이 느끼는 침 치료의 효과에 대해 깊게 탐색함으로써 수치로서 평가될 수 없었던 침 치료의 핵심적인 부분을 보여줄 수 있었다. 또한, 이러한 연구들의 결과는 향후 침 시술 프로토콜의 설계나 결과 평가도구를 개발하거나 선택함에 있어서 방향제시를 할 수 있다는 장점이 있었다. 한편, 질적 연구는 침 치료의 효과를 직접적으로 평가할 수 없기 때문에 연구의 결과가 끼칠 수 있는 영향력이 낮고, 연구자의 선입견이나 편견이 개입될 가능성이 높다는 한계가 있었다.

각 연구디자인의 장점과 한계점을 정리하면 다음의 표와 같다 (Table 3).

지금까지 요통의 체계적 고찰의 결과에 근거하여 각 연구디자인 별로 장점과 단점에 대해 살펴보았다. 각 연구디자인마다 장점이 있었고, 내재적 한계와 각 연구의 설계가 미흡하여 발생한 극복 가능한 한계점들이 있었다. 따라서 어떤 특정 디자인만이 우월한 연구가 될 수 있는 것이 아니라 연구의 목적과 연구 대상의 특징에 따라 각 연구디자인의 장점과 한계점들을 고려하여 가장 적절한 연구디자인을 선택해야 하는 것이다. 이에 본 연구에서 고찰한 각 연구디자인의 장점과 한계점들을 바탕으로 연구대상과 목적에 따라 각 연구디자인이 어떻게 활용되어야 할 것인가에 대해 살펴보도록

하겠다.

침 치료가 일반화된 질환에 대해서는 코호트 연구처럼 많은 수의 환자를 장기적으로 추적할 수 있는 연구디자인이 적절할 것이다. 특히, 코호트 연구는 치료의 장기적 효과를 평가할 수 있기 때문에 만성적 질환에 적절할 것이다.

침 치료가 일반화되어 있지만 서양의학에 있어 효과가 있는 치료법이 확립되지 않은 질환의 경우에는 침 치료와 서양의학의 치료법의 효과를 비교하는 실용적 임상연구가 가능할 것이다³⁹⁾. 또한, 실용적 임상연구는 실제 임상에서 이뤄지고 있는 침 시술의 양상을 그대로 반영하고 있기 때문에 표준화된 침 시술의 적용이 어렵고 환자와 침 시술자 간의 치료적 관계 형성이 치료에 있어서 핵심적 역할을 하는 내상질환일 수록 더욱 적절하다고 할 수 있다. 만약 현실적, 윤리적 이유로 무작위배정이 불가능하여 실용적 연구가 이뤄질 수 없다면, 유사 실험적 연구 중 대조군이 있는 전후 평가 연구가 활용될 수 있을 것이다. 특히 임상적 효과 평가뿐만 아니라 경제성 평가의 병행을 고려한다면, 실용적 임상연구가 우선적으로 검토되어야 할 것이다.

한편, 穴位나 침 시술의 방법에 있어서 정형화된 시술이 이뤄지고 있는 일부 질환들에 대해서는 설명적 연구가 가능할 것이다. 무엇보다 설명적 연구는 침의 효과가 관찰되거나 실용적 연구를 통해 어느 정도 밝혀진 상태에서 특정 穴位에 자침하는 것이 정말 효과가 있는가라는 구체적인 질문에 대한 연구를 할 때 적절하다고 할 수 있다. 이와 함께 체계적 고찰에 포함된 2편의 설명적 연구^{26,40)}에서 알 수 있듯 침 시술의 빈도, 자침 깊이 혹은 서로 다른 종류의 침 시술법의 효과를 비교하는데도 유용한 연구디자인이라고 할 수 있다. 설명적 연구는 ‘침이 효과가 있는가’라는 광범위한 질문보다

Table 3. Strengths and Limitations of Study Designs

		Strengths	Limitations
Experiment al study	Explanatory	<ul style="list-style-type: none"> • High internal validity • High Reproducibility • High quality of report on acupuncture treatment 	<ul style="list-style-type: none"> • Low external validity • Difficulty to reflect acupuncture treatment in practice • Short-term follow up
	Pragmatic	<ul style="list-style-type: none"> • Reflection of acupuncture treatment in real practice • High external validity • Long-term follow up • Evaluation of various outcomes • Evaluation of cost-effectiveness 	<ul style="list-style-type: none"> • Impossibility to evaluate specific acupuncture treatment • Difficulty to apply the result to specific patients in practice • Low reproducibility
Observational study		<ul style="list-style-type: none"> • Reflection of acupuncture treatment in real practice • Inclusion of many participants • Possibility to study rare diseases 	<ul style="list-style-type: none"> • Impossibility to evaluate specific acupuncture treatment • Low internal validity, especially in studies without controls • Low response rate
Qualitative study		<ul style="list-style-type: none"> • Suggestion of various aspects of acupuncture treatment that cannot be evaluated in quantitative studies 	<ul style="list-style-type: none"> • Impossibility to control researcher's bias • Low impact compared to quantitative studies

는 침의 특이적 효과, 침 치료법의 비교 등 구체적인 질문에 대해 연구하고자 할 때 적절한 연구디자인이라고 할 수 있다.

마지막으로 침 치료가 일반화되지 않은 질환, 혹은 희귀해서 서양의학에서도 치료법이 확립되지 않은 질환들에 대해서는 증례군 연구나 증례보고 등이 이러한 질환들에 대한 침 치료 가능성을 보여줄 수 있을 것이다.

본 연구는 하나의 질환에 대한 임상연구들을 연구방법론 중심으로 체계적 문헌고찰을 시도한 연구로서 실험적 연구뿐만 아니라 관찰연구, 질적 연구까지 포함함으로써 다양한 연구방법론을 포괄적으로 고찰하고자 하였다. 한편, 본 연구의 가장 큰 한계점은 체계적 고찰을 혼자서 수행하였기 때문에 각 단계에서 오류 및 편견이 발생했을 가능성이 높다. 둘째, 체계적 고찰에서 언어와 시간, 검색 엔진을 제한함으로써 영어, 중국어, 한국어가 아닌 다른 언어로 쓰인 임상연구, 2000년 이전과 2010년 이후의 연구와 Pubmed, EMBASE, Cochrane Library에서 검색되지 않았던 연구들은 배제되었을 것이다. 셋째, 침 임상연구의 연구방법론을 고찰함에 있어 요통이라는 특정 질환에 국한하였기 때문에 본 연구의 결과를 다른 질환에도 적용할 수 있을 것인가의 일반화 가능성에 대한 문제가 있을 것이다. 넷째, 본 연구에 포함된 관찰연구와 질적 연구의 수가 각 연구 디자인의 일반적인 특징 및 장점과 단점을 평가하기에는 적절하지 못했다는 한계가 있다. 다섯째, 체계적 고찰에 포함된 연구들의 연구보고의 질에 있어 큰 편차가 있기 때문에 본 연구에서 발견된 연구 디자인에 따른 차이가 실제 이뤄진 연구의 차이가 아닌 보고의 차이로 발생했을 가능성이 있다.

결론

본 연구에서는 한의학 이론에 근거한 침 시술의 특징을 잘 반영할 수 있는 임상연구방법론 개발을 위해 요통의 침 임상연구들을 연구방법론 중심으로 체계적으로 고찰하여 다음과 같은 결론을 내렸다.

침 치료는 단순히 경혈에 침을 삽입하는 것 이상의 다양한 요소들이 결합하여 이뤄지는 복합적 치료과정으로 한의학적 진단에 근거한 개별 치료와 配穴, 정확한 穴位, 得氣, 증상에 따른 手技, 적절한 留鍼시간과 치료횟수, 치료간격, 의료인과 환자의 긴밀한 관계 형성 등이 침 치료의 핵심적인 요소라는 전제 하에 요통의 침 임상연구에 대해 체계적으로 고찰한 결과, 설명적 연구에 비해 관찰연구나 실용적 연구가 침 치료의 핵심적 요소를 잘 반영하였다. 질적 연구는 침 치료의 과정 및 치료효과에 대해 깊게 탐색한 연구들로

양적 연구로 평가되기 어려웠던 개별 치료의 중요성, 환자와 의사의 관계의 중요성 등이 밝혀졌다.

근거중심의학에서는 일반적으로 무작위배정 대조 연구(RCT)로 대표되는 설명적 연구를 가장 과학적인 연구디자인으로 중시하고 있으나 본 연구의 결과 전체적으로 한의학적 침 시술의 특징을 잘 반영할 수 있는 연구는 관찰연구와 실용적 연구였으며, 질적 연구는 양적 연구에서 평가될 수 없었던 침 치료과정과 치료목표, 치료 결과에 대해 깊이 있는 이해를 가능하게 하였다. 따라서 향후 침 임상연구를 설계함에 있어서 '과학적'이라는 이유로 설명적 연구를 우선적으로 선택하기보다는 연구의 목적과 대상에 따라 각 연구디자인의 장점과 한계를 고려하여 적절한 연구디자인을 선택해야 할 것이다. 또한, 각 연구디자인에는 내재적 한계가 있기 때문에 이를 극복하여 내적 타당도를 높이기 위해 노력해야 할 것이며, 어떤 연구디자인을 선택하든 침 치료의 핵심적 요소를 잘 반영할 수 있도록 설계해야 할 것이다.

References

1. Haynes RB. What kind of evidence is it that Evidence-Based Medicine advocates want health care providers and consumers to pay attention to? BMC Health Services Research 2002 ; 2 : 3.
2. Evidence-Based Medicine Working Group. Evidence-Based Medicine: a new approach to teaching the practice of medicine. JAMA 1992 ; 268(17) : 2420-5.
3. Sehon SR, Stanley DE. A philosophical analysis of the evidence-based medicine debate. BMC Health Services Research 2003 ; 3 : 14.
4. Jackson S, Scambler G. Perceptions of evidence-based medicine: traditional acupuncturists in the UK and resistance to biomedical modes of evaluation. Sociology of Health & Illness 2007 ; 29(3) : 412-29.
5. Richardson J. The use of randomised control trials in complementary therapies: exploring the issues. J Adv Nurs 2000 ; 32(2) : 398-406.
6. Tonelli MR, Callahan TC. Why alternative medicine cannot be evidence-based. Academic Medicine 2001 ; 76(12) : 1213-20.
7. Nahin RL, Straus SE. Research into complementary and alternative medicine: problems and potential. BMJ 2001 ; 322 : 161-4.

8. Mason S, Tovey P, Long AF. Evaluating complementary medicine: methodological challenges of randomised controlled trials. *BMJ* 2002 ; 325 : 832-4.
9. Walji R, Boon H. Redefining the randomized controlled trial in the context of acupuncture research. *Complement Ther Clin Pract* 2006 ; 12(2) : 91-6.
10. Coulter ID, Khorsan R. Is health service research the holy grail of complementary and alternative medicine research? *Alternative Therapies in Medicine* 2008 ; 14(4) : 40-5.
11. Herman PM, D'Huyvetter K, Mohler MJ. Are health services research methods a match for CAM? *Altern Ther Health Med* 2006 ; 12(3) : 78-83.
12. Ritenbaugh C, Verhoef M, Fleishman S, Boon H, Leis A. Whole Systems Research: a discipline for studying Complementary and Alternative Medicines. *Altern Ther Health Med* 2003 ; 9(4) : 32-6.
13. Schwartz L. Evidence-Based Medicine and Traditional Chinese Medicine: Not Mutually Exclusive. *Medical Acupuncture* 2000 ; 12(1).
14. Walach H, Falkenberg T, Fonnebo V, Lewith G, Jonas WB. Circular instead of hierarchical : methodological principles for the evaluation of complex interventions. *BMC Medical Research Methodology* 2006 ; 6 : 29.
15. Eom SK, Kim SH, Choi WC. Study on modernization of traditional research methodologies of traditional Korean medicine. *Journal of Medical History* 2010 ; 23(2) : 89-105.
16. MacPherson H, Thomas K. Self-help advice as a process integral to traditional acupuncture care: implications for trial design. *Complement Ther Med* 2008 ; 16(2) : 101-6.
17. MacPherson H, Schroer S. Acupuncture as a complex intervention for depression: A consensus method for develop a standardised treatment protocol for a randomised controlled trial. *Complement Ther Med* 2007 ; 15 : 92-100.
18. Yuan J, Kerr D, Park J, Liu XH, McDonough S. Treatment regimens of acupuncture for low back pain-A systemic review. *Complement Ther Med* 2008 ; 16 : 295-304.
19. Gartlehner G, Hansen RA, Nissman D, Lohr KN, Carey TS. A simple and valid tool distinguished efficacy from effectiveness studies. *J Clin Epidemiol* 2006 ; 59(10) : 1040-8.
20. Linde K, Scholz M, Melchart D, Willich SN. Should systematic reviews include non-randomized and uncontrolled studies? The case of acupuncture for chronic headache. *J Clin Epidemiol* 2002 ; 55(1) : 77-85.
21. Zaringhalam J, Manaheji H, Rastqar A, Zaringhalam M. Reduction of chronic non-specific low back pain : a randomised controlled clinical trial on acupuncture and baclofen. *Chin Med* 2010 ; 5 : 15.
22. Cherkin DC, Eisenberg D, Sherman KJ, Barlow W, Kaptchuk TJ, Street J, et al. Randomized trial comparing traditional chinese medical acupuncture, therapeutic massage, and self-care education for chronic low back pain. *Arch Intern Med* 2001 ; 161(8) : 1081-8.
23. Itoh K, Itoh S, Katsumi Y, Kitakoji H. A pilot study on using acupuncture and transcutaneous electrical nerve stimulation to treat chronic non-specific low back pain. *Complement Ther Clin Pract* 2009 ; 15(1) : 22-5.
24. Szczerko O, Cooley K, Busse JW, Seely D, Bernhardt B, Guyatt GH, et al. Naturopathic care for chronic low back pain : a randomized trial. *PLoS One* 2007 ; 2(9) : e919.
25. Brinkhaus B, Witt CM, Jena S, Linde K, Streng A, Wagenpfeil S, et al. Acupuncture in patients with chronic low back pain. *Arch Intern Med* 2006b ; 166(4) : 450-7.
26. Itoh K, Katsumi Y, Kitakoji H. Trigger point acupuncture treatment of chronic low back pain in elderly patients-a blinded RCT. *Acupunct Med* 2004 ; 22(4) : 170-7.
27. Tsui ML, Cheing GL. The effectiveness of electroacupuncture versus electrical heat acupuncture in the management of chronic low-back pain. *J Altern Complement Med* 2004 ; 10(5) : 803-9.
28. Kerr DP, Walsh DM, Baxter D. Acupuncture in the management of chronic low back pain: a blinded randomized controlled trial. *Clin J Pain* 2003 ; 19(6) : 364-70.
29. Meng CF, Wang D, Ngeow J, Lao L, Peterson M, Paget S. Acupuncture for chronic low back pain in older patients: a randomized, controlled trial. *Rheumatology* 2003 ; 42 : 1508-17.
30. Yeung CK, Leung MC, Chow DH. The use of electro-acupuncture in conjunction with exercise for the treatment of chronic low-back pain. *J Altern Complement Med* 2003 ; 9(4) : 479-90.
31. Leibing E, Leonhardt U, Köster G, Goerlitz A, Rosenfeldt JA,

- Hilgers R, Ramadori G. Acupuncture treatment of chronic low-back pain- a randomized, blinded, placebo-controlled trial with 9-month follow-up. *Pain* 2002 ; 96(1-2) : 189-96.
32. Molsberger AF, Mau J, Pawelec DB, Winkler J. Does acupuncture improve the orthopedic management of chronic low back pain- a randomized, blinded, controlled trial with 3 months follow up. *Pain* 2002 ; 99(3) : 579-87.
33. Carlsson CP, Sjolund BH. Acupuncture for chronic low back pain: a randomized placebo-controlled study with long-term follow-up. *Clin J Pain* 2001 ; 17(4) : 296-305.
34. Witt CM, Jena S, Selim D, Brinkhaus B, Reinhold T, Wruck K, et al. Pragmatic randomized trial evaluating the clinical and economic effectiveness of acupuncture for chronic low back pain. *Am J Epidemiol* 2006 ; 164 : 487-96.
35. Cherkin DC, Sherman KJ, Avins AL, Erro JH, Ichikawa L, Barlow WE, et al. A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. *Arch Intern Med* 2009 ; 169(9) : 858-66.
36. Thomas KJ, MacPherson H, Thorpe L, Brazier J, Fitter M, Campbell MJ, et al. Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain. *BMJ* 2006 ; 333(7569) : 623.
37. MacPherson H, Thorpe L, Thomas K. Beyond needling-therapeutic processes in acupuncture care: a qualitative study nested within a low-back pain trial. *J Altern Complement Med* 2006 ; 12(9) : 873-80.
38. Hsu C, BlueSpruce J, Sherman K, Cherkin D. Unanticipated benefits fo CAM therpaies for back pain: an exploration of patient experiences. *J Altern Complement Med* 2010 ; 16(2) : 157-63.
39. MacPherson H. Pragmatic clinical trials. *Complement Ther Med* 2004 ; 12 : 136-40.
40. Yuan J, Purepong N, Kerr DP, Park J, Bradbury I, McDonough S. Effectiveness of acupuncture for low back pain. *Spine* 2008b ; 33(23) : E887-E900.