

한의 임상 연구에 대한 국내외 제언 고찰

- 미래 임상 연구 전략 II -

정기용 · 이민혜 · 최유경 · 이충열¹ · 박종형 · 전찬용*

가천대학교 한의과대학 내과학교실, 1:가천대학교 한의과대학 생리학교실

Study on the Proposals for Clinical Research in Korean Medicine Worldwide

- Future Clinical Research Strategy II -

Ki Yong Jung, Min Hye Lee, You Kyung Choi, Choong Yeol Lee¹, Jong Hyeong Park, Chan Yong Jeon*

Department of Internal Medicine, 1: Department of Physiology, College of Korean Medicine, Gachon University

The purpose of this paper is to explore the strategy of future Korean medicine(KM) clinical research through the study on the proposals for KM clinical research worldwide. In this study, the papers published in English through Pubmed were investigated mainly. Among them, we analyzed the methodological proposals from the clinical research papers that were published in the KM related fields. Various proposals for improving the problems in KM clinical studies are as follows. First, KM clinical research should be designed based on understanding for the theory, backgrounds, paradigms and worldviews of KM. In addition, considering the model validity, KM clinical research model should include the diagnosis, interventions and outcomes measurement methods reflecting the characteristics and real practice in KM. The internal validity and external validity should be also taken into account. One of the most important thing is to identify the contents about various and complex 'real practice' in KM. A prospective observational study was suitable for the purpose of this study. Finally, we suggested a few improvement directions for RCTs studies in KM. First, we would be able to improve the quality and the internal validity in KM clinical research using the checklists of CONSORT(Consolidated Standards for Reporting Trials) Statement and STRICTA(Standards for Reporting Interventions in Clinical Trials of Acupuncture). Second, we could use various clinical research methods and the modified research of RCTs such as PCT(pragmatic clinical trial) to reflect the characteristics of actual KM practice. Consequently, we would be able to improve the external validity. Third, the KM diagnosis and outcomes measurement methods should be developed based on an actual KM practice and it should reflect a real practice. The 'pattern identification(辨證)' is the core to KM diagnosis. But in order to be applied to the clinical research, the pattern identification(辨證) should be objectified and standardized. Future KM clinical research model should reflect the characteristics and a real practice in KM. In addition, it should include the advantage of rigorous RCTs research. Specially, the diagnosis, interventions and outcomes measurement methods in KM clinical research should reflect this view.

keywords : Clinical Research, Korean Medicine(KM), Methodology, Proposals

서론

전 세계적으로 의학계에는 근거중심의학(Evidence-Based Medicine, 이하 EBM) 및 이를 뒷받침하는 임상 연구(clinical research)에 대한 중요성이 강조되고 있고 한의 분야 또한 예외는 아니라고 할 수 있다¹⁻³⁾. 이러한 상황 속에서 2000년대 이후 한의 임상 연구 분야에서 EBM을 위한 임상 연구의 핵심인 무작위화

(randomization), 이중 맹검(double blinding), 위약(placebo) 등으로 이루어진 무작위대조군시험(Randomized Controlled Trials, 이하 RCTs) 연구⁴⁾가 지속적으로 증가하고 있다⁵⁻⁷⁾.

문제는 기존 의학 연구에 일반적으로 적용되는 RCTs 연구를 비롯한 임상 연구 방법론(clinical research methodology)이 한의 분야에 그대로 적용되는 것이 과연 적절한가라는 점이다. 이에 대해 몇몇 연구들에서 기존 한의 분야 임상 연구에서 이루어진 RCTs 연

* Corresponding author

Chan Yong Jeon, Department of Internal medicine, Gachon University Gil Korean Medical Hospital, 1200-1, Guwoldong, Namdong-gu, Incheon, 405-760, Korea

·E-mail : joncy@gachon.ac.kr ·Tel : +82-70-7120-5007

·Received : 2015/02/05 ·Revised : 2015/03/03 ·Accepted : 2015/03/04

© The Korean Society of Oriental Pathology, The Physiological Society of Korean Medicine

pISSN 1738-7698 eISSN 2288-2529 <http://dx.doi.org/10.15188/kjopp.2015.04.29.2.115>

Available online at http://society.kisti.re.kr/sv/SV_svjsj03L.do?method=list&poid=ksomp&kojic=DRSRDH&sVnc=v28n5&menuid=1&subid=13

구 결과가 실제 한의 임상 치료 효과를 적절히 보여주고 있지 않음을 지적하였다⁵⁻¹⁰. 아울러 그 이유에 대한 다양한 분석과 함께 한의 임상 분야에 적절한 임상 연구 방법이 무엇인지에 대해 다양한 논의와 제안이 이루어졌다^{4-6,8,11-14}.

우리는 선행 연구를 통해 한의 임상 특성과 실제 임상 현상이 반영되지 않은 한의 임상 연구의 문제, 한의 임상 특성을 반영하기 어려운 기존 임상 연구 방법의 무비판적 적용과 이로 인한 문제, 질적으로 낮은 수준의 한의 RCTs 연구, 한의 진단(diagnosis) 및 중재(interventions)의 표준화 문제, 부적절한 결과 측정(outcomes measurement) 방법의 문제 등을 확인할 수 있었다⁷.

본 논문에서는 앞서 지적되었던 한의 임상 연구의 문제점들을 개선하기 위한 국내외의 다양한 제언들을 고찰하고자 한다. 이를 통해 실제 임상 영역에서 다양하게 이루어지고 있는 한의학의 치료 효과를 적절하게 검증할 수 있는 미래 한의 임상 연구 방법론이 무엇인지 모색하고자 한다.

연구대상 및 방법

본 연구에서는 2000년 이후 출판된 국내외 저널에서 한의 임상 연구 방법론에 대해 다룬 논문들을 대상으로 하였으며, 논문은 주로 Pubmed(www.ncbi.nlm.nih.gov/pubmed)에서 검색하였다. 그리고 필요할 경우 국내 논문 검색은 전통의학정보포털(OASIS, <http://oasis.kiom.re.kr>), 중국 논문 검색은 중국학술정보원에서 제공하는 의학 분야 중국 저널 검색 엔진인 China National Knowledge Infrastructure(CNKI, <http://search.cnki.net>)를 사용하였고, 일본 논문 검색은 Japan Science and Technology information Aggregator, Electronic(J-STAGE, <http://www.jstage.jst.go.jp>)을 이용하였다. 검색 범위는 한국의 한의학(韓醫學, Korean Medicine), 일본의 캠포의학(漢方醫學, Kampo medicine), 중국의 중의학(中醫學, Traditional Chinese Medicine, 이하 TCM), 서양에서 한의학이 포함되거나 관련되어 논의될 수 있는 분야라 할 수 있는 보완대체 의학(Complementary and Alternative Medicine, 이하 CAM), 전통 의학(Traditional Medicine), 통합 의학(Integrative medicine) 등에서 임상 연구에 대한 논의로 한정하였다. 임상 연구와 관련된 내용으로는 이들 한의 분야의 임상 연구 방법론(methodology)이나 구체적 임상 연구 방법들에 대한 제언 및 개선점을 논하고 있는 논문들을 수집하여 분석하였다. 그러나 한의 임상 연구 수행에 대한 논의와 제안은 많았지만 이 주제에 대한 체계적 접근이 담긴 출판 논문은 소수였다¹⁰. 따라서 본 논문에서는 임상 연구 방법론(methodology), 구체적 임상 연구 방법(예를 들어, RCTs 등)에 대한 검색뿐만 아니라 한의 임상 연구와 관련된 세부적인 영역(예를 들면, 결과 평가 방법, 진단(변증) 등)에 대한 검색을 함께 실시하였고, 각 논문들을 개별적으로 검토하여 본 논문의 주제와 연관 있는 내용을 포함시켰다.

실제로 완전히 일치한다고 할 수 없기 때문에 논란이 있을 수 있겠지만, 본 연구에서는 한의학, 중의학, 캠포의학, 보완대체의학, 전통의학, 통합 의학 등의 용어를 문맥 상 특별히 구별이 필요한 경우를 제외하고, '한의학' 및 '한'으로 통칭하여 기술하였다.

본 론

[문제 해결을 위한 다양한 제언 및 미래 임상 연구 전략]

1. 한의학의 이론, 배경, 패러다임, 세계관적 특성에 대한 이해를 바탕으로 임상 연구가 설계되어야 함.

한의학은 이론과 임상 완성 체계를 가진 의료 체계이다. 한의학은 환자의 전반적인 상태 및 체질을 중요시 하는 개별적(individualized) 특성을 반영한 치료 시스템으로 몸과 마음을 하나로 보는 전일론적인 치료적 접근(holistic therapeutic approach)을 중시하고 증상을 완화하고 몸의 기능의 균형을 회복시키는 것을 치료의 목적으로 한다^{6,11}. 또한 한의학은 서양의학과는 개념적으로 다른 중후군(證) 위주의 전일론적인 의학 체계(syndrome-oriented holistic medical system)이다. 여기서 중후군(證)은 한의학 이론에 따라 증상과 징후로 이루어진 그룹이다. 따라서 한의학의 증(證)은 서양의학의 질환 및 중후군 분류와 항상 일치하지 않는다¹⁴. 그리고 한의학 세계관에서 개인은 주변 환경, 타인, 사회 등과 모두 연결되어 있으며, 그 결과 환자와 의사 사이에 치료적 관계로 연결되어 있다¹¹.

이렇듯 한의 치료의 맥락(context), 패러다임(paradigms), 철학적 바탕 및 임상적 사용은 기존 서양의학과는 다르다. 따라서 연구자들은 한의 임상 연구에 있어 치료 과정이 무엇인지, 얼마나 많은 변수들이 있는지, 치료의 기저에 깔려 있는 철학적 기초들, 건강과 질병에 대한 개념, 그러한 치료의 맥락과 관련된 체계(contextual framework) 및 핵심 치료 요소들이 무엇인지 이해할 필요가 있다⁸. 즉 한의학 이론에 대한 이해가 필수적이라 할 수 있다^{14,15}. 이러한 이해를 바탕으로 다양한 연구 질문들이 제기될 수 있다. 아울러 구체적 임상 연구의 접근에 있어, 치료 환자군의 선택, 진단 및 치료 방법에 대한 이해, 나아가 효과 판정에 대한 기준 등을 세울 수 있으므로, 이를 통해 실제 한의 임상 현상이 반영된 임상 연구를 계획할 수 있게 된다.

그 구체적 예로 인삼과 관련된 RCTs를 분석한 Yan 등¹⁴의 연구는 참고할 만하다. Yan 등은 기존 임상 연구가 제대로 이루어지지 않은 원인으로 패러다임(paradigm)의 충돌을 들면서 한의학 이론이 반영된 연구 디자인을 통해 이 문제가 해결될 수 있다고 주장하였다. 그리고 그 예로 인삼 연구에 대한 기존 RCTs 연구들을 분석한 결과를 그 논증 근거로 들고 있다. 우선은 처방이 아닌 인삼 단일 약물을 이용한 연구, 기허(氣虛)로 진단된 환자들이 아닌 건강한 사람들을 대상으로 한 연구, 그리고 인삼의 종류(Panax ginseng, Panax quinquefolius)에 대한 구별을 하지 않은 연구, 인삼의 오용으로 인한 증상들(주로 氣鬱이나 氣滯에서 나타나는 증상들)이 언급되지 않은 연구 등이 한의학 이론과 맞지 않는 연구 디자인임을 지적하고 있다. 아울러 인삼에 대한 19개의 RCTs 논문을 분석하였는데, 인삼 효과가 긍정적으로 평가된 논문이 10편, 부정적으로 평가된 논문이 9편이었다. 분석 내용 중 주목할 만한 부분은 효과가 긍정적인 논문 10편 중 8편이 한의학적 진단 기준으로 氣虛라는 병리적 상태에 있는 대상자들에 대해 이루어진 연구였고, 이에 반해 부정적인 효과가 나온 연구 9편 중 7편이 건강한 대상자들에

대해 이루어진 연구였다는 점이다. 또한 긍정적 연구 결과가 나온 논문 10편 중 나머지 2편과 부정적 연구 결과가 나온 논문 9편 중 나머지 2편의 결정적 차이는 한의학 이론에 따른 적절한 품종의 인삼 사용 유무였다. 긍정적 연구 결과의 논문은 인삼 종류에 따른 적절한 연구 디자인이 이루어졌고, 부정적 연구 결과의 논문은 인삼 종류가 부적절하게 선택되었다. 그 결과 이 연구는 한의학 이론에 따른 인삼의 치료 적응증을 확인해 주었다. 아울러 한의학의 이론을 바탕으로 한 RCTs 연구 디자인이 미래의 한의학 분야의 임상 연구를 발전시키고 인정받도록 하는데 필수적 요소임을 시사해주고 있다¹⁴⁾. 즉, 한의학 임상 실체를 뒷받침하고 있는 고유 이론들과 원칙들이 임상 연구에 있어 반드시 고려되어야 한다는 것이다.

의학 이론의 우선적인 가치는 임상 진료에 대해 성공적인 가이드라인 역할을 하여 효과적인 치료가 가능하게 하는데 있다. 따라서 임상 연구를 통해 효능이나 효과를 입증하는 것은 한의학 이론에 대한 가장 좋은 실증적인 시험이 될 수 있을 것이다. 확인된 효능이나 효과에 대한 이론에 근거해서, 새롭게 발전되고 향상된 한의학 이론은 더욱더 견고한 근거가 될 수 있을 것이다^{15,16)}.

2. 모델 타당도(Model validity)에 대한 개념 가질 필요

앞서 언급한 바와 같이 한의 임상에 기반이 되는 한의 이론 및 배경을 이해한 뒤, 이러한 이해가 실제 임상 연구로 이어지기 위해서는 실제적 접근 및 평가가 필요하다. 이러한 점에서 모델 타당도(Model validity, Model fit validity)라는 개념이 도움이 될 수 있다.

모델 타당도를 CAM 분야에 적용할 것을 주장한 사람은 Cassidy인데, 그는 모델 타당도를 패러다임 적합(paradigm fit)으로 설명하였고, 조사(investigation) 및 연구 설계(research design)가 시스템의 설명 모델과 일치하는 경우 패러다임 적합이 있다고 주장하였다^{13,17)}. 아울러 모델 타당도를 평가하는 기준을 다음의 네 가지로 제시하였다.

첫째, 대표성(representativeness)으로 중재(intervention)가 현재 실제 임상 진료와 일치하는지, 효과를 나타낼 가능성이 있는지, 분명하고도 적절하게 묘사되고 있는지를 파악한다. 둘째, 균형(Equipoise)과 신뢰성(Credibility)에 대한 평가로 중재, 중재의 신뢰성, 기대되는 치료 효과에 대해 치료자와 환자 모두에 대한 균형 잡힌 평가가 이루어지고 있는지를 파악한다. 셋째, 모델 적합성(Model congruity)으로 진단, 중재(intervention), 결과 평가가 조사되는 의학 시스템에 적합한지를 평가하는 것이다. 넷째, 맥락(context)으로 중재에 대한 환자와 치료자의 믿음, 환자가 중재에 속한 것으로 생각하는 문화, 가족 및 의미에 얼마나 민감한지에 대해 평가하는 것이다.

이 가운데 네 번째를 제외한 모든 부분이 실제 임상 연구의 설계와 밀접한 관련을 가지고 있다고 할 수 있다. 즉, 한의 임상 연구 모델이 한의학의 학문적 특징 및 실제 임상 현장(real practice)을 반영하여 진단, 중재, 결과 평가 방법을 사용하였는지 검토함으로써(모델 타당도를 검토), 실제 한의 임상 치료를 검증할 수 있도록 설계되었는지를 알 수 있을 것이다. 특히, 모델 타당도와 가장 관련이 깊은 것은 진단 및 결과 평가 방법이라 할 수 있다¹³⁾.

이와 함께 임상 시험(clinical trials)의 두 가지 형태에 대한 고

려가 반드시 필요한데, 하나는 설명적 임상 연구(explanatory trials)이고 다른 하나는 실용적 임상 연구(pragmatical trials)라 할 수 있다. 전자의 대표적인 연구가 RCTs 연구로 가장 이상적인(optimal and ideal) 상황에서 중재의 효능(efficacy)을 검증하는 것으로 중재 외에 다른 다양한 요소들을 엄격하게 통제하여 내적 타당도(internal validity)를 극대화한다. 후자는 매일의 임상 치료 현장에서 실제로 시행되는 중재들의 전체적인 효과(effectiveness)를 검증하는 것으로 연구 결과들이 일반화될 수 있음을 보장함으로써 외적 타당도(external validity)를 극대화한다^{10,13,18-21)}. 주의할 점은 실용적 임상 연구가 설명적 임상 연구(RCTs)의 일종으로 완전 별개의 연구가 아니라는 점이다. 실제 임상 연구에서는 내적 타당도와 외적 타당도 사이의 적절한 긴장 관계가 중요하다¹⁹⁾. 한의 임상 연구의 경우 서양의학과 한의 임상 체계가 다르기 때문에 내적 타당도, 외적 타당도뿐만 아니라 모델 타당도에 대한 관심이 중요하다¹³⁾. 아울러 내적 타당도, 외적 타당도 및 모델 타당도는 서로 완전히 독립적인 것이 아니며, 한의 임상 연구에 있어 이 세 요소에 대한 균형 잡힌 접근이 필요하다 할 수 있다(Fig. 1).

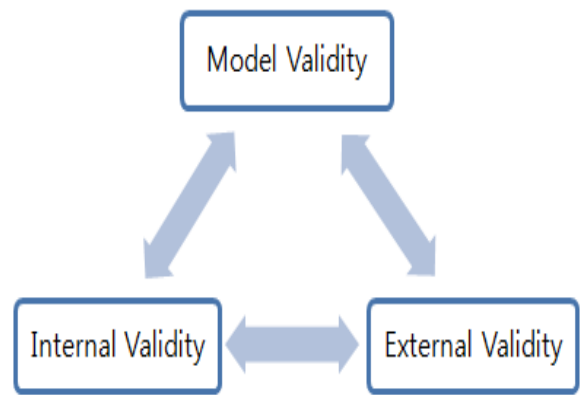


Fig. 1. The Balance of three Dimensions for Good Clinical Research in Korean Medicine.

3. 전향적 관찰 연구(prospective observational study)의 필요성

관찰 연구(observational study)는 RCTs가 적절하지 않은 실용적인 연구 분야에 적합하다^{12,22)}.

전향적 관찰 연구는 두 가지 점에서 한의 임상 연구 분야에서 우선적으로 필요한 연구라고 할 수 있다.

먼저, 전향적 관찰 연구는 한의 임상 시험 설계에 앞서 구체적인 실제 임상 진료 내용들을 파악할 수 있게 해준다는 점에서 중요하다. 선행 연구에서 언급한 바와 같이 한의 임상 연구는 한의 치료가 이미 폭넓게 임상에서 사용되고 있으므로 서양의학과 달리 연구의 출발점이 실제 임상 진료가 될 것이다^{7,8,15,16)}. 따라서 임상 연구에 앞서 실제 임상 진료의 내용들을 파악하는 일이 무엇보다 중요한 일이라 할 수 있다. 즉, 한의 치료로부터 얻을 수 있는 이점, 치료하고자 하는 질병 또는 환자의 상태, 구체적인 치료 방법(투약법, 침치료 방법 등), 치료 기간 등에 대한 지식이 요구되고, 이러한 지식은 필수적인 안전성 평가는 말할 것도 없이 치료받고 있는 환자의 호전에 대한 지식을 얻는데 도움을 줄 것이다. 이러한 모든 질문은 일상적인 치료(usual care)에서 이러한 측면들을 평가할 수 있

는 전향적 관찰 연구(prospective observational study)를 통해 그 답을 얻을 수 있다⁶⁾. 즉, 전향적 관찰 연구를 통해 실제 임상 현장에서 어떤 질환이 한의 치료의 대상이 되고, 어떤 구체적인 치료 방법으로, 얼마의 기간 동안 치료가 이루어지고 있는지, 어떤 방식으로 치료 효과를 평가하고 있는지 등을 파악할 수 있게 되고 이러한 지식을 기반으로 구체적인 임상 시험을 설계할 수 있을 것이다. 만약 실제 한의 임상 치료 현장에 대한 구체적인 자료를 근거로 임상 시험이 이루어지지 않는다면, 임상 시험의 결과가 한의 치료의 효과를 제대로 평가한 결과라고 볼 수 없을 것이다.

그 다음으로 전향적 관찰 연구는 그 자체로 치료에 대한 임상적 근거를 제공해주므로 가치가 있다고 할 수 있다. 최근 몇몇 연구들은 잘 고안된 관찰 연구가 RCTs에 맞먹는 강한 근거를 제공해줄 수 있다는 결과를 제시하고 있다. 관찰 연구는 개별화된 치료가 적용됨에 따라 강력한 외적 타당도를 가지고 있는 것으로 나타난다. 만약 적절하게 선택되고, 치료 결과가 폭넓게 평가된다면 모델 적합도를 향상시킬 수 있는 잠재성을 가지고 있다. 한의 분야에서 관찰 연구가 적용될 때는 장기적인 관점으로 시행되어야 한다. 그러나 관찰 연구는 비교군 또는 대조군이 없고 긍정적인 결과는 단순히 평균으로의 회귀를 반영할 수 있다. 이것은 독립적인 차트 리뷰, 각각의 환자에 대한 예측 생존 등에 대한 엄밀한 진료 기록, 기준 측정 방법 또는 비교군 등을 사용하여 내적 타당도를 높이려는 다양한 전략을 통해 개선될 수 있을 것이다¹³⁾.

한의 임상 연구에 있어 한약 처방과 침치료 방법(여러 침법, 침 처방) 등의 다양성 때문에 가장 적절한 한약 처방이나 침치료 방법 등을 선택하는 것이 RCTs를 포함한 한의 임상 연구에 있어 가장 중요한 부분이라 할 수 있다. 성공적인 임상 연구를 위해서 임상적으로 가장 효과적인 한약 처방이나 침치료 방법을 선택해야 할 것이다⁵⁾. 바로 이 부분에서 전향적 관찰 연구는 많은 정보를 제공해 줄 수 있을 것이다.

이렇듯, 전향적 관찰 연구가 가장 좋은 방법이지만, 한의 임상 분야에서 가장 빈번하게 보고되는 것은 후향적 관찰 연구라 할 수 있다. 후향적 관찰 연구는 잘 설계된 연구가 아니므로 그 한계가 분명하지만, 추후 더 나은 임상 연구를 위한 사전 자료 및 권고를 제공할 수도 있다는 점에서 의미를 찾을 수 있을 것이다⁵⁾.

4. 한의 분야에 적용되는 RCTs 연구 방법의 보완

앞선 연구에서 지적한 바와 같이 기존 한의 분야에 적용되었던 RCTs 연구는 연구의 질적 저하 문제, RCTs 연구 자체의 방법론적인 문제, 한의 임상 특성이 충분히 반영되지 못한 한계 등으로 개선이 필요하다 할 수 있다⁷⁾. 따라서 다음의 세 방향으로 개선이 이루어져야 할 것이다. 첫째 RCTs 연구의 질을 개선하여 내적 타당도를 증가시켜야 한다. 둘째, 기존 RCTs의 한계를 보완할 수 있는 다양한 임상 연구 방법들을 도입하여 외적 타당도를 증가시켜야 한다. 마지막으로 한의 임상 특성에 근거한 진단 및 결과 평가 방법을 도입하여 모델 타당도를 증가시켜야 한다. 구체적 내용은 다음과 같다.

1) RCTs 연구의 질 개선 - CONSORT와 STRICTA 적용

RCTs 연구의 질 개선은 세 가지 방면에서 이루어져야 한다.

먼저 적절한 규모의 연구가 진행되어야 한다. 표본 크기가 통계적 분석이 가능한 정도로 충분히 큰 규모여야 하며, 충분한 임상 연구 기간이 설정되어야 한다. 다음으로 중재가 표준화(standardization)되어야 하고 임상 연구 계획서(protocol)에 따라 임상 연구가 진행되어야 한다. 예를 들어, 침 치료의 경우 같은 경혈 위치, 같은 자극 강도 및 깊이, 같은 수기법 등을 사용해야 하고, 한약 치료의 경우 한약 처방의 구성 및 제조 방법이 같아야 한다¹⁾. 마지막으로 할당(allocation), 은닉(concealment) 등의 무작위화(randomization), 맹검(blinding), 대상자 선정, 대조군 설정, 결과 평가 방법 등 RCTs 연구에 맞게 연구가 설계되어 방법론적인 문제가 없도록 해야 한다.

이러한 부분들을 만족시킬만한 기준으로 고려할 수 있는 부분이 바로 CONSORT(Consolidated Standards for Reporting Trials, 이하 CONSORT) 성명서(statement)라 할 수 있다. 본래 CONSORT 성명서는 RCTs 연구의 보고(reporting) 질을 높이기 위한 목적으로 1996년 CONSORT 성명서 첫 번째 버전이 출판되었다. 이후 몇 차례 개정이 되었고 개정 과정에서 연구 디자인이 RCTs 뿐만 아니라 cluster trials(2004), 비열등성 임상 시험 및 동등성 임상 시험(non-inferiority and equivalence trials)(2006), 실용적 임상 시험(pragmatic trials)(2008) 등으로 확대되었고, 중재 또한 한약물 치료(2006), 비약물 치료(non-pharmacological treatment, 2008), 침 치료(2010)로 확대되었다²³⁾. 2010년에 개정된 CONSORT 성명서에서는 기존의 항목에 임상 시험 등록(registration), 연구계획서(protocol), 연구비용(funding)의 세 항목이 추가되어 25개 점검 항목이 제시되었다²³⁾. 25개의 항목 중 임상 시험 설계에 대한 기술, 연구 대상자, 중재, 결과 평가, 표본 크기, 무작위화(할당 방법, 은닉 방법, 누가 어떻게 시행하는지에 대한 과정), 맹검, 통계 방법, 연구 대상자 흐름(각각의 그룹의 인원수, 시행된 치료들, 일차 결과에 대한 분석들, 무작위화 이후의 배제 및 그 이유 등등), 모집(모집 기간 및 추적 관찰 기간, 종료 이유 등등), 결과 평가 방법 등²³⁾은 RCTs 연구의 핵심적인 내용으로 이에 대한 엄밀한 보고가 이루어진 RCTs 연구의 경우 질이 높을 수밖에 없을 것이다. CONSORT 점검 항목에서 요구하는 각 항목의 세부 사항들을 기준으로 RCTs 연구를 시행하기 전에 점검한다면 연구의 질이 높아질 수 있다. 따라서 CONSORT 성명서는 연구자들이 미래에 임상 연구를 설계하는데 도움을 줄 수 있고²³⁾ 임상 시험 설계에 있어 방법론적인 고려 사항들에 대한 개념적 지침으로 사용되어, 한의 임상 중재의 근거 기반을 증가시키는데 기여할 수 있을 것이다²⁴⁾.

한약물 치료에 대한 CONSORT 성명서의 적용은 2006년 발표되었는데, 가장 큰 특징은 기존 CONSORT의 중재(Interventions) 항목에 한약물 치료에 대한 특성이 반영되었다는 점이다. 추가된 중재 항목의 내용은 한약물 제품명, 한약물 특징, 복용법(용량 및 기간), 한약 처방 구성 성분 용량, 질적 검사(화학적 분석, 어떤 성분들이 제거되었는지, 제품의 화학적 구성 성분 등등), 위약/대조군 그룹(위약 제조 방법), 임상의(교육 및 임상 경력)에 대한 부분이었고, 나머지 점검 항목에서는 약간의 보완 정도만 추가되었다²⁵⁾.

침 치료의 경우 STRICTA(Standards for Reporting

Interventions in Clinical Trials of Acupuncture, 이하 STRICTA)가 있다. 2001년 처음 출판된 STRICTA 보고 지침은 침의 비교 임상 시험에서 중재 보고의 완벽성과 투명성을 향상시켜 이러한 임상 시험들을 보다 정확하게 해석하고 손쉽게 반복할 수 있도록 하기 위해 만들어졌다. 이후 2010년에 CONSORT와 비약물 치료 및 실용 임상 시험을 위한 CONSORT 확충 안에 맞추어 STRICTA가 개정되었다. STRICTA는 CONSORT 성명서에서 중재(interventions) 보고에 관련된 5번째 중재 항목의 일반적 내용을 확충한 점검표로 구성된다. 여기에서 침 임상 시험의 중재를 보고할 때 포함해야 할 정보로 침치료를 대한 논거, 자침에 대한 상세한 내용(자침 수, 경혈 이름, 자침 깊이, 유발한 반응, 침 자극의 형태, 유침 시간, 침의 형태), 처치 내용(치료 횟수, 빈도와 시간), 치료의 다른 구성 요소들(뜸, 부항, 한약, 생활 습관 조언 등, 치료 환경과 상황), 시술자의 배경(교육 및 임상 경력 등), 대조군 및 비교군 중재(대조군이나 비교군에 대한 논거 및 정확한 기술, 거짓침 등의 세부 사항)를 제시하였다²⁶⁾.

이들 CONSORT와 STRICTA에서 제시하고 있는 점검 항목들이 임상 연구 과정에서 잘 이행된다면 연구의 질 또한 높게 유지될 수 있을 것이다.

2) 기존 RCTs 연구 방법에 대한 보완 및 새로운 연구 방법들

RCTs 연구가 한의 분야에 많은 것들을 제공해줄 수 있지만, 한계를 가지고 있다. RCTs에 의해서 해결될 수 없는 간격을 메우기 위해 또 다른 방법론적인 접근이 필요하다¹²⁾. 여기에는 기존 RCTs 연구를 변형한 형태의 연구 방법들과 새로운 접근의 연구 방법들이 함께 포함된다. 본 연구에서는 몇 가지 대표적인 연구 방법들을 검토하고자 한다.

(1) 실용적 임상 연구(Pragmatic Clinical Trials)

한의 분야 RCTs 연구의 한계를 보완하기 위해서 자주 제시되는 연구 방법이 바로 실용적 임상 연구(Pragmatic Clinical Trials, 이하 PCT)다^{6,8,10,13,20,21)}. 앞서 언급한 바와 같이 PCT는 RCT의 일종으로서 실험군과 대조군을 무작위 배정하는 것은 동일하다. 실제로 CONSORT에 PCT가 추가되었고²⁷⁾, 연구 방법에서 일부 항목은 PCT와 RCT가 거의 같다고 할 수 있다¹⁰⁾. 연구 가설의 인과 관계를 검증하기 위한 RCT를 설명적 임상연구(Explanatory Clinical Trials, ECT)라고 하며, 실용적 임상 연구는 임상 진료에서의 결정에 관한 정보를 주기 위한 RCT를 가리킨다. PCT는 치료의 '전체적인 효과(overall effectiveness)'를 평가할 수 있다. ECT는 특정 가설에 따른 '효능(efficacy)'을 검증하기 위한 것으로 PCT는 매일의 실제 임상 치료에 대한 내용과 효과 및 환자와 의사 사이의 치료를 잘 반영해준다^{18,20)}. 즉, 개별화된 치료 접근에 대한 평가가 가능한데, 환자는 무작위 배정되지만, 중재(interventions)는 실제 임상 현상을 반영하여, 외적 타당도를 높이게 된다¹³⁾.

PCT는 맹검을 하지 않으므로 환자-의사 관계, 치료 효과에 대한 믿음을 반영하고 있으며, 치료 또한 표준화된 단일 치료가 아닌 실제 일상적인 치료(개별화된 치료 및 복합적인 치료)를 허용하고 있다. 아울러 비특이적이고 복잡한 만성적인 질환에 주로 적용되고 치료의 이득을 확인하기 위해 장기간 추적 관찰이 이루어지고, 모집 환자군 또한 이질적인 환자군을 용인함으로써 환자 선정 기준이

넓다. 이와 함께 임상 결과의 평가 또한 복합적 치료가 이루어지므로 다차원적인 결과 평가 방법이 선호된다. 따라서 연구 결과가 실제 임상 현상을 반영하므로 임상에 곧바로 적용할 수 있다. 이러한 PCT 연구의 특징들이 한의 임상 의 특성과 어느 정도 일치한다고 할 수 있고, RCT의 부족한 부분을 메워준다고 할 수 있다^{13,18,21)}. 그러나 PCT는 몇 가지 한계를 가지고 있다. 먼저, 외적 타당도가 높으나 맹검이 되지 않으므로 내적 타당도가 떨어진다^{10,18)}. 아울러 복합적 치료를 구성하는 각 요소들이 치료 효과에 기여하는 정도를 정확히 결정할 수 없고, 개별 임상 의들의 능력에 따라 다른 결과가 나올 수 있고, 일반적으로 연구 대상수가 많기 때문에 실제 연구 수행에 연구비와 시간이 많이 든다¹⁸⁾.

(2) 선호 임상 시험(preference trials)

한의 임상에서 치료에 대한 환자들의 믿음과 선호는 중요한 부분 중 하나이다. 따라서 한의 임상 연구에서 임상 시험 참여자의 믿음, 실행 과정, 선호도 때문에 참여자의 모집과 무작위화(randomization)가 문제될 수 있다⁴⁾.

환자 선호도는 치료 효과와 관련 있기 때문에 무작위화는 적절치 않다. 이 연구의 경우 치료 선호가 없는 환자의 경우 일반적인 경우와 같이 무작위 배정을 하고 치료 선호가 있는 환자의 경우 선호하는 치료를 받게 한다. 이를 통해 치료 선호도와 효과 간의 상관 관계를 평가할 수 있다²⁸⁾.

(3) 요인 설계(Factorial designs)

여러 치료법들 사이의 다중적인 상호 작용 효과를 평가하기 위하여 하나의 치료법과 여러 치료법들의 조합을 비교하는 연구 방법이다. 이 방법은 두 개 이상의 치료 방법의 시너지 효과 뿐만 아니라 중재의 다양한 구성 요소들의 상대적인 효과에 대한 통찰 또한 제공해준다. PCT와 요인 설계를 결합할 경우, 외적 타당도를 증가시킬 수 있고, 잠재적으로 내적 타당도와 모델 타당도를 증가시킬 수 있을 것이다¹³⁾.

(4) 단일 증례 연구(Single-case research)

① N-of-1 임상 시험(N-of-1 trials)

RCTs 연구는 개개인에 대한 정보를 제공할 수 없고, 단지 개인이 속한 그룹의 평균 치료 효과만을 제공할 수 있다¹³⁾. N-of-1 임상 시험은 개개인의 치료에 대한 개별화된 관점을 보장할 수 있고²⁾, 이러한 부분이 한의 임상 의 특성을 반영할 수 있어 한의 임상 연구의 방법으로 몇몇 연구자들이 지속적으로 제안하였다^{6,29)}. N-of-1 임상 시험은 무작위화 단일 증례 실험 연구(randomized controlled single-patient trial)로 치료 기간과 위약 또는 표준 치료 기간 사이에 다중 교차(multiple crossovers)로 시행된다. 즉, 먼저 검증하고자 하는 중재를 시행하는 기간(Period 1)이 있고, 중간에 적절한 휴지기(Washout period)가 지난 뒤, 표준 치료 또는 위약 치료를 시행하는 기간(Period 2)이 있게 된다. 각 기간에 대한 효과를 판정함으로써 치료 효과를 검증하게 된다. 모든 기간에서 치료 방법은 무작위 배정되고 이상적으로는 의사와 환자의 이중 맹검이 적용될 수도 있다. 따라서 N-of-1 임상 시험은 엄격한 임상 시험으로 객관성을 가질 수 있는 장점이 있다. 즉, 한의 임상 의 특성과 EBM 모두에 어느 정도 부합한다고 할 수 있다³⁰⁾. 그러나 비용이 많이 들고³⁰⁾, 개인 수준의 자료만을 제공하여 그룹이나 인구 집단에

적용하기는 어렵다고 할 수 있다¹³⁾. 여러 개의 N-of-1 임상 시험의 결과가 합쳐질 경우 이러한 문제는 해결될 수 있을 것이다¹³⁾.

② 단일 증례 관찰 연구

임상 연구에 대한 잦은 비판은 임상 시험이 치료 증례 후에 개별 환자의 예후에 대해 거의 알려주지 못한다는 것이다. 한의 임상 치료는 매우 개별화된 치료로 증례 관찰 연구의 중요성이 폭넓게 받아들여지고 있다. 문제는 한의 임상 증례 보고의 질이 여전히 낮다는 것이다. 한의 임상에서 증례 연구를 보고하는 절차가 표준화되어야 한다. 현재 가장 우선시 되는 것은 임상 용어 및 기록 양식(record format)을 진단과 치료 효과 평가에 대한 명확하고 적절한 정보를 가지고 표준화하는 것이다⁵⁾. 덧붙여 미국 국립 암연구소(National Cancer Institute, NCI)의 암 CAM 사무국(NCI's Office of Cancer Complementary and Alternative Medicine, OCCAM)의 우수 증례 시리즈 프로그램(NCI Best Case Series(BCS) Program)의 기준(www.cancer.gov/cam/bestcase_intro.html)을 참고하여, 한의 임상 증례 관찰 연구의 질을 향상시킬 수 있을 것이다⁵⁾. 한 가지 추가하자면 증례 관찰 연구가 전향적으로 이루어지는 것이 가장 좋고, 후향적으로 이루어지더라도 충분한 임상 자료들이 갖추어진 경우에 이루어져야 증례 관찰 보고의 질이 높아질 수 있다.

(5) 질적 연구(qualitative research)

RCTs로부터 만들어진 대부분의 지식들은 일반적이고 전체적인 형태로, 연관성의 빈도 및 강도에 대한 질문에 답을 하고 있다. 반면에 질적 연구(qualitative research)는 개개인의 경험의 '왜'와 '어떻게'에 관하여 구체적 정보를 만들어내기 위해 설계된다⁴⁾. 한의 임상상의 경우 환자와 의사간의 주관적, 전일론적인(holistic), 상호적인(reciprocal) 관계라는 특성을 가지고 있어 질적 연구 방법이 필요하다 할 수 있다¹²⁾. 아울러 한의 임상 분야의 다양성(다양한 학술 유파 및 이론들, 다양한 침 치료 이론 및 방법 등)과 복잡성(여러 치료 방법들의 복합, 문화적 맥락 등) 등이 반영된 실제 임상 현장을 파악하는데 질적 연구가 도움이 될 수 있다. 따라서 몇몇 연구자들은 질적 연구와 양적 연구(RCTs)가 결합된 연구의 필요성을 주장하였다^{12,13)}.

질적 연구의 목적은 다음과 같다. 첫째, 연구에 참여하는 참여자 및 그들이 참여하는 사건(events), 정황(situations), 행동(actions)에 대한 의미를 확인하여 증례(interventions)에 대한 이해를 얻을 수 있다. 둘째, 참여자들이 행동하는 특정 맥락(context) 및 그러한 맥락이 그들의 행동에 미치는 효과를 이해할 수 있다. 나아가 가장 중요한 부분은 증례 시 환자-치료자 간의 상호 작용의 역할을 확인하는 것이다. 셋째, 사건과 행동이 일어나는 과정을 이해할 수 있다. 넷째, 증례에 대해 다양한 관계자(임상의, 환자, 연구자)들의 실제 관점이 어떻게 다른지 평가할 수 있다⁴⁾.

RCTs와 관련하여 질적 연구는 RCTs 연구 전에 적절한 결과 평가 방법의 개발을 위해 사용될 수 있고, 측정되는 상황 및 치료의 과정을 이해하기 위해 사용될 수 있다. 아울러 임상 시험의 결과를 설명하기 위해 RCTs 연구 종료 뒤에도 사용될 수 있다⁴⁾. 물론 결과 평가 방법 개발뿐만 아니라 한의 진단 관련 분야에서도 사용될 수 있다.

질적 연구는 주로 사회 과학(Social Science Research) 분야에

서 자주 사용되는 연구 방법이다. 이와 관련하여 Complementary Therapies in Medicine 저널에 발표된 내용들은 참고할만하다. 이 논문에서 2004, 2005년 연구 방법 시리즈(research methods series)를 기획하였고, 여기에 RCT, SR(Systematic Review), PCT 등의 양적 연구(quantitative research)뿐만 아니라 심층 면접(qualitative interviews), 대화 분석(conversation analysis), 민족지학(ethnography), 문서 분석(document analysis) 등의 질적 연구(qualitative research)가 언급되었다. 이어서 2007년 보고에서는 추가로 일기문 분석(diary analysis), 실천 연구(action research), 초점 집단(focus group), 델파이법(Delphi method) 등의 질적 연구 방법에 대한 제안이 있었다³¹⁾. 그러나 사회 과학의 질적 연구 방법론을 한의 임상 연구 분야에 어떻게 접목시킬지에 대한 실제적 가이드라인은 없었다. 질적 연구 방법들은 실제 한의 임상상의 효과 및 효능을 판정하는데 한계를 가지지만 한의 임상 현장을 보다 입체적으로 파악할 수 있게 해주는 도구가 될 수 있다는 점에서 의미가 있다고 할 수 있다.

이들 가운데 실제 임상 연구와 관련되어 접목될 수 있는 연구가 하나 있는데 바로 델파이법이라 할 수 있다. 델파이법은 전문가들 사이에 불일치되는 견해에 대한 문제를 해결하기 위한 전문가 합의법이다. 먼저 중요 문제에 대하여 설문지를 발송하여 개인적으로 전문가들에게 일련의 집중적인 질문을 한다. 첫 번째 설문지의 결과가 요약된 뒤, 그 결과가 수정을 위해서 전문가들에게 다시 보내진다. 매회 설문에 대한 반응을 수집·요약하며, 그것을 다시 전문가들에게 송환해 주게 되는데, 이에 따라 개별 전문가들은 자신의 견해나 평정을 수정해 간다. 전문가들이 더 이상 그들의 견해를 수정할 필요가 없는 의견 일치에 다다를 때까지 이 과정이 지속된다. 이를 통해 특정 주제에 대해 전문가 합의에 이르게 된다³¹⁾. 한의 임상 분야에 델파이법이 적용된 예로 중의 진단 과정의 평가자간 신뢰도(inter-rater reliability)를 향상시키고, 타당도를 측정하기 위한 연구에서 델파이법을 사용하여 연구에서 사용된 평가 도구가 한의 진단 범주와 일치하는지를 규명한 연구가 보고되었다³²⁾. 현재 한의 진단 및 결과 평가 방법이 충분히 표준화, 객관화 되어 있지 못한 것이 현실이다. 델파이법은 한의 임상 연구에 앞서 한의 전문가들의 의견을 모아 실제 임상 현장에 대한 현실을 파악할 수 있게 해주고, 전문가들의 의견을 공유하여 합의안을 만드는데 중요 도구로 사용될 수 있을 것이다.

3) 한의 임상 실제 및 특성에 기반을 둔 진단 및 결과 측정 방법

(1) 한의 진단(변증)

① 한의 진단(변증)의 표준화 및 객관화

한의 임상 연구에서 가장 중요한 부분이 바로 진단이라 할 수 있다. 진단은 환자군의 설정 및 모집(Inclusion and Exclusion criteria)의 기준이 되기 때문에 매우 중요하다. 선행 연구에서 지적한 바와 같이 기존 한의 임상 연구에서 바로 이 부분에 많은 문제가 있었다⁷⁾. 한의 임상 연구지만 한의 진단이 사용되지 않은 경우가 많았고, 사용된 경우라도 명확한 진단 기준이 제시되지 않았다^{5,6)}.

한의 진단은 곧 변증(辨證, pattern identification)이라 할 수 있다. 물론 한약 처방의 적응증을 진단 기준으로 삼는 경우도 있지만, 이 경우도 변증 진단 체계로 포함시킬 수 있다. 가장 중요한 핵

심 문제는 표준화된 변증 진단 기준이 부재하다는 것이다. 한의 변증을 정의하는 것이 한의 임상 연구의 가장 근본적인 문제가 될 것이다. 임상 연구가 수행되기 전에 목표가 되는 질환에 대해 받아들일 만한 데이터와 근거로 뒷받침된 표준 변증 진단 기준이 있어야만 한다⁵⁾.

실제 한의 임상 연구에 도움이 되는 표준 변증 진단을 개발하기 위해서 다음의 몇 가지 고려 사항을 제안하는 바이다. 첫째, 변증 전체에 대한 표준 변증 진단을 개발하는 방법과 임상 시험의 대상이 되는 특정 질환 또는 증후에 해당하는 표준 변증 진단을 개발하는 방법 중 어떤 방법을 선택할지 결정해야 한다. 전자의 경우 오랜 시간과 노력이 필요하지만 모든 질환에 적용할 수 있어 편리하고 후자의 경우 변증의 수가 많지 않아 비교적 쉽게 만들 수 있지만 매번 같은 작업을 해야 하는 불편함이 있다. 이러한 점을 고려하여 임상 시험을 위한 표준 변증 진단을 개발해야 하고 여러 변증 체계들 중에 실제 임상에서 사용되는 변증을 위주로 표준 변증 진단이 정립되어야 한다. 둘째, 실제 임상에서 중요시되는 부분들을 반영하여 각 변증을 구성하는 증상(symptoms) 및 징후(signs), 설진, 맥진 등이 확정되어야 한다. 아울러 변증을 구성하고 있는 각각의 증상들과 징후들이 명확하게 정의되어야 하고(모든 임상들이 동일하게 이해 및 기술 가능하도록) 각 요소의 중요도를 어떤 방식으로 반영할 것인지에 대해서도 고려해야 할 것이다. 셋째, 변증을 구성하는 증상 및 징후 항목들의 개수를 가지고 변증을 확정할 것인지 아니면 각각의 증상 및 징후들의 중요도를 정량화(수치화)할 것인지를 결정해야 한다. 만약 정량화 한다면 어떤 방식으로 정량화할 것인지를 결정해야 한다. 예를 들어, 변증을 구성하는 특정 증상에 대해 가중치를 줄 것인지, 가중치를 준다면 어느 정도 줄 것인지 등을 결정해야 한다. 마지막으로 확정된 변증에 대한 내적 타당도 및 외적 타당도에 대한 평가 및 수정이 이루어져야 한다. 즉, 만들어진 표준 변증 진단은 다시 임상 현장에 적용되어 임상 의사의 진단 일치도, 타당도 연구 등을 통해 지속적으로 수정되고 보완되어야 할 것이다.

앞서 제시한 고려 사항들을 실제로 만족시키기는 쉽지 않은 문제이다. 기존의 문헌 분석을 통한 연구와 함께 실제 한의 임상 현장 가운데 이루어지고 있는 변증 진단을 파악하기 위한 대규모의 전향적 관찰 연구, 대규모 설문 조사, 전문가 합의를 위한 델파이법 등의 다양한 연구 방법이 동원되어야 할 것이다. 아울러 체계적인 장기간의 연구가 일관성 있게 지속적으로 이루어져야 할 것이다.

② 한의 진단(변증)의 임상 연구 적용

한의학적인 치료는 한의 진단을 통해 이루어져야 하는데⁶⁾, 한의 진단이 충분히 객관화, 표준화되지 않고 서양 의학이 전 세계적으로 통용되는 현실을 감안할 때, 한의 임상 연구는 서양 의학적 진단과 한의 진단(double classification)을 함께 사용해야 할 것이다^{5,6)}. 여기서 서양 의학적 진단은 국제질병분류(international classification of diseases, 이하 ICD), 한의 진단은 변증이라 할 수 있다. 몇몇 연구에서 연구 설계를 실제 임상 치료에 맞추기 위한 시도로, 임상 시험의 통계 범위 안에서 한의 변증을 사용하였다. 이러한 방법은 1998년 Schnyer에 의해 만들어지고 최초로 출판되었다. 이 연구에서 환자들은 서양 의학적 진단에 따라 먼저 모집되었고 그리고 나

서 변증에 따라 진단되었다. 이들 연구에서 사용된 한의 변증은 주로 문헌 고찰을 통해 수집된 내용의 변증이었다³³⁾.

한의 진단인 변증을 임상 연구에 적용하고자 할 때, 몇 가지 가능한 연구 설계 방법들이 제시되었다. 첫째로 먼저 서양의학의 특정 질환을 가진 환자들을 모집한 후 한의학적 진단 방법으로 하위 집단(subgroups)으로 나눈 뒤 그들 모두에 한의 변증 진단에 따라 각각 다른 한의 치료를 시행한다^{1,6,15,34)}(Fig. 2). 여기서 두 가지 방법으로 연구를 진행할 수 있는데, 각각의 한의 변증에 대해서 무작위 배정으로 한의 치료와 위약 치료를 함께 하는 방법(Fig. 2)과 대조군 없이 한의 변증에 한의 치료만 시행하는 경우가 있을 수 있다. 전자의 경우 개인적 특성이 반영된 한의 치료 체계의 효과(efficacy)를 입증하기 위한 예비 연구나 소규모 연구에 적합하다⁶⁾. 후자의 경우 비용도 적게 들고 전자보다 연구가 간단하고, 복잡한 치료료가 가능한 장점이 있으나, 대조군이 없으므로 명확한 치료 효과를 입증하기 어려운 단점이 있다.

두 번째로, 서양 의학적 진단 후 무작위 배정을 통해 한의 치료군과 위약 치료군으로 나누고, 한의 치료군 내에서만 한의 진단을 통한 분류 및 개별 치료를 하는 것이다⁶⁾(Fig. 3). 한의 진단과 이에 따른 개별적인 치료를 통해 전체적인 임상 효과(effectiveness)를 평가할 수 있고, 각 변증에 따른 분류에서 복합적인 치료가 가능하고, 위약 또는 서양의학 치료 효과와 비교가 가능한 장점이 있다. 아울러 이 모델에서는 실용적 임상 연구(PCT)의 방법을 도입하는 것 또한 가능하다⁶⁾.

앞서 언급한 두 개의 연구 설계 모델에서 서양의학 진단에 대해 한의 변증 진단은 통상 두 개 이상인 경우가 많으므로, 많은 수의 하위 집단 분류가 있게 되어 많은 수의 증례가 필요하게 된다. 그러나 각 하위 집단에서 많은 수의 증례수를 수집하기 어려워 결국 하위 집단의 분석에서 통계적 파워가 떨어지고, 그 결과 각각의 한의 치료에 대한 일반화가 어렵게 된다^{6,15)}. 이러한 문제를 해결하기 위해 서양의학의 질환 환자들로부터 특정 변증과 일치하는 상태를 가진 환자만을 모집해 치료 효과를 검증하여 한의 치료 효과를 평가하는 방법이 있다^{1,5,6,15)}. 아울러 먼저 사전 분석을 통해 모아진 변증 진단(pooled patterns)을 이용하여, 매우 빈번한 변증에 대해서 하위 집단 분석을 미리 구체화하는 것이다⁶⁾. 즉, 모든 변증에 대해 연구를 진행하는 것이 어렵고 비효율적이므로 사전 분석을 통해 가능한 변증 전체를 모으고, 여기서 실제 임상 현장에서 다빈도로 이루어지는 변증 위주로 연구를 설계하는 것이다. 물론, 실제 임상에서 어떤 변증이 다빈도로 나타나는지에 대한 선행 연구가 이루어져야 할 것이다.

한의 변증 진단은 질병의 층화(stratification) 기준으로서 사용될 수 있는데, 바로 이 부분이 미래 한의 임상 연구의 핵심적인 이슈가 될 수 있을 것이다^{5,6)}. 그에 앞서 실제 임상 현장에서 이루어지는 변증 진단에 대한 지속적인 연구 및 이를 바탕으로 한 변증 진단의 객관화 및 표준화가 선행되어야 할 것이다. 이를 통해 임상 연구에 맞는 변증 체계 및 개별 변증의 빈도수 등을 알 수 있을 것이다. 아울러 실제 임상에서 가장 효과가 좋은 약물 처방이나 치료법을 찾는 노력이 함께 이루어질 때, 성공적인 한의 임상 연구가 될 수 있을 것이다.

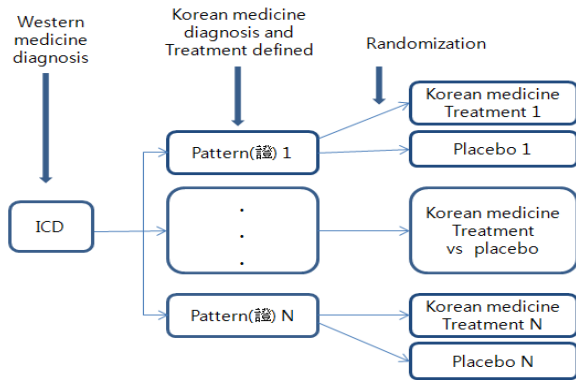


Fig. 2. Combining the Western and Korean medicine diagnosis(pattern identification) for a placebo-controlled trial: to evaluate the efficacy of Korean medicine treatment(an extract from Watanabe et al.⁶⁾)

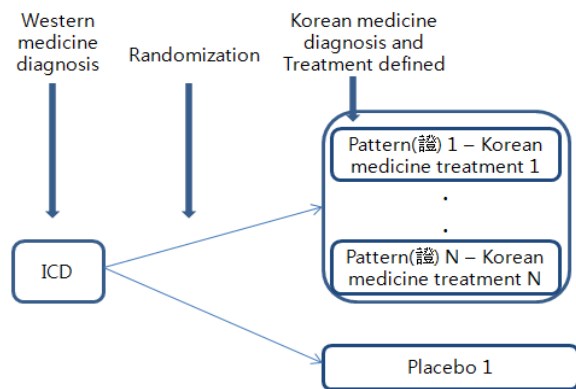


Fig. 3. Combining the Western and Korean medicine diagnosis(pattern identification) for a standard care controlled trial: to evaluate the effectiveness of Korean medicine treatment(an extract from Watanabe et al.⁶⁾)

(2) 결과 측정(outcomes measurement) 방법의 보완

한의 진단과 함께 한의 임상 연구의 성패를 결정하는 중요한 요인 중 하나가 바로 적절한 결과 측정 방법이다. Long 등은 한의 임상 연구에 사용되는 결과 측정 방법이 개념적으로 다음의 세 가지 조건을 만족시켜야 한다고 하였다. 첫째, 전일론적이며 개별화된 치료, 몸의 불균형의 회복 등과 같은 건강과 치유에 대한 한의학적 세계관이 반영되어야 한다. 둘째, 한의 임상에서 중요시되는 의사-환자 간의 관계 및 상호 작용에서 나타나는 효과를 평가할 수 있어야 한다. 셋째, 의사가 치료를 위해 사용하는 치료 기술들의 집합적 효과를 평가할 수 있어야 한다³⁵⁾.

실제적으로는 환자와 의사 모두에게 임상적으로 중요하고, 효과를 볼 수 있는 적절한 시점에서, 중재의 의도와 배경에 맞는 기대되는 효과를 평가하는 것이 중요하다^{13,35)}. 다양한 결과 평가 방법(주관적인 부분과 객관적인 부분을 모두 포함)의 사용은 내적, 외적, 모델 타당도를 완성하는데 도움을 준다¹³⁾. 그러나 한의 치료의 개념적 특성 및 임상 현장을 반영하면서도 객관성과 타당성이 함께 확보된 결과 평가 방법을 만드는 일은 쉬운 일이 아니다. 이러한 상황 속에서 몇 가지 고려해볼 수 있는 결과 평가 방법은 다음과 같다.

① 객관적 결과 평가 방법

앞서 설명한 바와 같이 한의 임상 연구에서 서양 의학적 진단

이 사용되므로 당연히 결과 평가 방법으로 서양 의학의 대상 질환의 진단 기준이나 결과 평가 방법이 동일하게 사용되어야 할 것이다. 이는 서양 의학 진단 및 결과 평가 방법인 여러 가지 영상의학 검사, 혈액 검사, 조직 검사 등의 객관적 검사 및 특정 질병의 진단 가이드라인 항목 등을 말한다고 할 수 있다. 물론 서양 의학적 결과 평가 방법만을 말하는 것은 아니고 여러 방법 중의 하나로 이러한 객관적 방법이 사용될 수 있다는 것이다. 무엇보다도 한의학 특성에 맞는 평가 방법과 함께 사용되어야 한다는 것이다^{1,15)}.

② 주관적 결과 평가 방법

한의 임상 진단 및 치료에서 중시되는 것은 바로 환자들이 느끼는 주관적 증상들과 삶의 질과 관련된 여러 호소들이다^{5,24)}. 따라서 주관적 결과 평가 방법은 일반적인 임상 연구에서는 이차적 평가 방법일 경우가 많지만 한의 임상 연구에서는 일차적 결과 평가 방법으로 많이 쓰인다^{5,24,37)}. 환자가 보고하는 형식의 주관적 평가 방법은 지난 20년간 꾸준히 증가하고 있다. 환자 보고 결과(patient-reported outcomes, 이하 PROs) 방법은 오늘날 한의 임상 연구를 포함한 임상 연구에서 종료 시점(endpoints)으로 자주 사용되고 있다³⁷⁾. 최근에는 PROs가 결과 평가 방법으로 사용된 임상 연구의 질을 높이기 위해 CONSORT 확장으로 CONSORT Patient-Reported Outcome도 발표되었다³⁸⁾.

현재 한의 임상 연구 분야 및 서양의학의 임상 연구에서 사용될 수 있는 주관적 결과 평가 방법이 다수 존재한다. 여기에는 삶의 질(quality of life) 또는 일반적인 건강에 대한 표준화된 평가 도구(예 ; 36 item short form health survey, SF-36), 증상 또는 장애 등의 특정 상태 평가 도구(예 ; the Arthritis Impact Measurement Scales), 정신적인 평가 도구(예 ; the profile of Mood States), 통증 측정 도구(예 ; visual analogue scales(VAS), numerical rating scales(NRS)), 환자-생산 도구(patient-generated instrument)(예 ; Measure Yourself Medical Outcome Profile, MYMOP) 등이 포함된다³⁷⁾. 한의학의 특성을 기반으로 한의 임상 영역 내에서 만들어진 평가 도구는 중국에서 개발된 ChQOL(Chinese medicine-based quality of life)이 대표적이다^{5,37,39)}.

PROs 방법과 관련되어 몇 가지 문제가 제기되었다. 먼저 여러 평가 방법들 중 적정 평가 방법을 선택하는 가이드라인이 없다는 것이다. 다른 하나는 한의 임상 연구에 사용된 PROs 방법의 결과가 임상적 의의와 항상 일치하지는 않는다는 것이다. 즉, PROs 결과의 통계적 유의성이 반드시 임상적 중요성을 의미하지 않을 수 있다는 것이다³⁷⁾. 한의 임상 분야 암 관련 연구에서 단지 약 20% 정도의 연구에서만 PROs 점수가 임상적으로 의미가 있다고 보고되었다⁴⁰⁾. 이는 평가 방법에서 나타나는 작은 차이(small difference)가 환자 대상자수가 많아질 경우 통계적으로는 의미 있는 결과가 나올 수 있기 때문이다. 그러한 작은 차이는 실제 임상적으로는 의미가 없을 수 있다. 전자는 임상 연구와 관련된 적절한 결과 평가 방법을 선택하기 위해, 관련 임상 연구에서 사용된 결과 측정 방법들을 직접적으로 비교하는 방법을 통해 해결될 수 있다. 문헌 분석, Canadian Interdisciplinary Network for CAM Research에서 만든 결과 평가 도구 데이터베이스(www.outcomesdatabase.org) 사용 등이 그

방법이라 할 수 있다³⁷⁾. 후자의 경우 한의 임상 연구에서 사용된 PROs 방법의 임상적 의의(clinical significance)를 확인하기 위해 ‘minimal (clinically) important difference(MID or MCID)’ 개념을 사용할 수 있다. MCID는 환자가 중요한 것으로 인식하는 결과의 가장 최소의 변화이다. 따라서 MCID는 환자에게 유의미한 의의가 있는 것으로 경험될 수 있는 결과 바로 위의 한계점을 제공해주어, 단순한 통계적 유의성의 문제를 피할 수 있게 해준다. 이러한 MCID를 계산하는 방법으로 distribution-based methods, anchor-based methods 등이 사용된다³⁷⁾.

아울러 2004년 미국국립보건원(National Institute of Health, NIH)이 PROMIS 협력 그룹(Patient-Reported Outcomes Measurement Information System, 이하 PROMIS)을 설립하였는데, 미래 한의 임상 연구 결과 평가 방법을 개선하는데 도움을 줄 것이다. PROMIS(www.nihpromis.org)는 건강 관련 다양한 결과 평가 방법들에서 많은 설문 항목들을 모아 항목들을 개발하고, 교정하고, 타당성을 평가하여 항목들의 풀(pool) 또는 은행(bank)을 만들어, 더 짧고 개별화된 방법을 개발할 수 있는 플랫폼(platform)을 제공해 준다. 여기에 포함되는 항목들은 엄격한 질적, 양적 검토가 이루어져 왔다³⁷⁾.

또한 ChQOL처럼 기존에 검증된 결과 평가 방법들을 한의 임상 특성을 반영하여 수정 보완하거나³⁹⁾ 한의 임상 특성을 반영할 수 있는 새로운 결과 평가 방법을 개발해야 할 것이⁴¹⁾, 이러한 결과 평가 방법들은 실제 한의 임상 속에 적용되어 지속적으로 수정되어야 할 것이다.

③ 결과 평가 방법으로서의 변증

한의 임상은 한의 변증을 통해 환자들이 진단되고 이를 바탕으로 한의학적 치료가 이루어진다. 이러한 변증을 진단뿐만 아니라 효과 판정에 사용하는 것 또한 완전히 불가능한 일은 아닐 것이다. 이에 대해 Berle 등³³⁾은 C형 간염 환자들을 대상으로 하여 변증을 통해 치료 결과를 평가하였다. 이 연구에서 먼저 문헌 고찰을 통해 간염과 관련된 문헌상의 모든 변증을 모집하였고, 그 다음으로 각 변증을 구성하는 증상(symptoms) 및 징후(signs)를 모아서 각 변증 진단명에 따라 그 증상 수를 확인하고 증상으로 이루어진 변증 설문표를 작성하였다. 환자들마다 모든 변증에 해당되는 증상을 체크한 뒤, 각 변증마다 체크된 증상 수를 각 변증에 해당되는 총 증상 수로 나누어 100을 곱해서 퍼센트로 환산하여 수치화하였다. 이 연구에서 세 차례 변증 진단을 하였고, 세 차례 변증 진단의 결과를 수치화하여 대조군과 비교하여 통계 분석을 하였다³³⁾. 이 연구는 변증을 결과 평가 도구로 사용한 최초의 연구라는 점에서는 의미가 있지만, 몇 가지 문제점을 가지고 있다. 첫째, 변증 진단의 객관성이 떨어지고, 타당도 평가가 이루어지지 않았다. 둘째, 변증을 구성하는 증상들 모두를 같은 비중으로 다루었다는 점이다. 변증을 구성하는 여러 증상들 중 일부는 그 비중이 더 클 수 있다. 이에 대한 가중치를 두는 방안이 실제 변증 진단에 가깝다고 할 수 있다. 셋째, 변증을 계량화하는 과정에 대한 객관성 및 타당성이 결여되어 있다. 이러한 수치화가 과연 실제 임상적인 상황을 반영하는 것인지 알 수 없다. 추후에 이러한 지적 사항들이 보완되고, 실제 한의 임상 현상이 반영된 변증 진단이 이루어지고, 변증의 계량화 및 이에

대한 임상적 타당성에 대한 엄밀한 연구가 이루어진다면, 결과 평가 방법으로 변증이 사용될 수 있을 것이다.

④ 적절한 연구 기간

마지막으로 고려할 사항은 한의 임상 연구에서 결과를 평가하는 시점이다. 치료 기간이 너무 짧게 설정될 경우, 치료 효과가 미처 나타나기 전에 연구가 종료될 수 있다. 임상 연구에서 증개의 효과를 검증하기 위해서 먼저 증개가 치료 효과를 나타내는 시점에 대한 기준이 필요하다. 먼저, 같은 질환을 대상으로 한 선행 임상 연구들이 그 기준이 될 수 있을 것이다. 한 가지 고려할 사항은 이러한 선행 연구들 가운데 서양 의학 분야의 임상 연구들의 치료 기간은 참고로 해야 한다는 점이다. 선행 연구에서 지적한 바와 같이 한의 임상 연구의 대상되는 질환들이 만성 질환인 경우가 많고, 한의 치료의 효과가 완만하다는 것을 함께 고려할 때 치료 기간이 서양 의학의 임상 연구에 비해 길어질 수 있는 가능성이 있다^{2,7,11)}. 또한 가지는 한의 임상 연구에 있어 한의 임상 실체를 반영하여 임상 연구의 치료 기간을 설정해야 한다는 것이다. 이를 위해 실제 임상 현장에서 이루어지는 연구 대상 질환의 치료 기간을 파악하기 위한 예비 연구를 시행하는 것이 좋을 것이다. 문헌 고찰(관련 질환 한의 임상 분야의 국내외 논문, 그 수가 충분할 때), 전향적 관찰 연구, 관련 질환의 임상 전문의를 대상으로 한 델파이법, 많은 수의 임상 의들을 대상으로 한 설문 조사, 특정 질환에 대한 한의 분야의 건강 보험 청구 현황 조사 등이 예비 연구의 방법으로 가능할 것이다.

결 론

미래 한의 임상 연구는 RCTs 연구를 포함한 기존 임상 연구 방법의 장점을 반영할 뿐만 아니라 그 한계를 극복하는 가운데 확립될 것이다.

이번 연구를 통해 볼 때, 이상적인 미래 한의 임상 연구는 다음의 두 가지 측면에 대한 충분한 고려가 있을 때 가능하지 않을까 생각된다.

첫째, 한의 임상 연구는 실제 한의 임상 현장에서 이루어지고 있는 한의 임상의 특성을 반영해야 한다. 이와 함께 한의학적 진단 및 치료의 기저에 있는 한의학적 이론, 배경, 패러다임, 세계관적 특성 등에 대한 이해가 선행되어야 한다. 그리고 임상 연구가 실제 한의 임상 현장을 제대로 반영하고 있는지에 대한 평가는 모델 타당도의 개념을 통해 검토할 수 있을 것이다. 덧붙여 실제 한의 임상 현장을 파악하기 위한 다양한 선행 연구가 이루어져야 할 것이다. 예를 들어, 다양한 시공간적 차이를 고려한 고전 의서의 해석 및 분석, 국내외 임상 논문 검색을 통한 체계적 문헌 고찰(Systematic Reviews) 등의 다양한 문헌 연구, 전향적 또는 후향적 관찰 연구, 대규모 설문 조사 등의 방법이 이용될 수 있을 것이다. 이를 통해 실제 한의 임상 현상이 파악된 후 이를 기반으로 다음 단계의 대규모 임상 연구가 이루어질 수 있을 것이다.

둘째, 한의학적 특성을 반영하기 위해 기존 RCTs 연구 방법을 수정 보완하여 한의 임상 연구 방법을 확립할 수 있을 것이다. 먼저 기존 한의 임상 연구에서 문제점으로 지적되었던 한의 분야 RCTs 연구의 질적 보안을 위해 CONSORT 성명서, STRICTA 보고 지침

등을 적용하여 임상 연구의 내적 타당도를 높일 수 있을 것이다. 다음으로 RCTs 연구 자체의 한계를 보완하기 위한 PCT 연구, 질적 연구 등의 다양한 연구 방법을 추가하여 임상 연구의 외적 타당도를 높일 수 있을 것이다. 마지막으로 실제 한의 임상 연구의 특성이 반영된 한의 진단(변증)의 표준화 및 객관화, 결과 평가 도구 개발, 적절한 임상 연구 기간 설정 등에 대한 연구가 이루어져야 한다. 이를 통해 한의 임상 연구의 모델 타당도를 높일 수 있을 것이다.

본 연구에서는 임상 연구 방법론에 대한 논의가 주로 이루어졌다. 추후 연구에서는 한의 임상 연구의 확대를 위한 다양한 방안이 논의될 수 있을 것이다. 여기에는 한의 임상 연구의 주체가 대학 한방 병원으로 한정되는 것이 아니라, 대부분을 차지하고 있는 개원의를 임상 연구에 참여시키는 방안이 포함될 수 있을 것이다. 그 방법으로 표준화된 진료지침의 개발, 중례기록지(clinical research form)의 개발, 중례기록지 수집(전자중례기록지 등)과 관리 및 임상 연구 교육을 담당하는 국립 한의 임상 연구 기관 등의 설립 등이 논의될 수 있을 것이다. 아울러 충분한 인력과 시설이 갖추어져서 임상 연구와 다양한 질환들을 함께 진료할 수 있는 대규모의 국립 임상 연구 선도 한방병원 설립에 대한 연구도 논의에 포함시킬 수 있을 것이다.

한의 임상 연구 분야는 정체되어 있지 않으며, 수행되는 연구의 질이 점차적으로 개선되고 있으며, 적용되는 연구의 양도 빠르게 증가하고 있다^{5-7,12}). 그러나 한의 임상 연구를 위한 충분한 재정 지원의 부족과 한의 임상 연구 관련 전문가의 부족은 큰 문제라 할 수 있다^{2,42}). 이러한 상황 속에서 앞서 제시한 한의 임상 연구 방법이 미래에 바람직한 방향으로 발전하기 위해 보다 많은 한의사들의 참여와 관심이 필요할 것이다. 아울러 통계학, 보건학, 생물학, 의학, 철학 등 다양한 관련 학문 분야와의 공동 연구가 필요하며, 무엇보다도 한의 임상 연구를 발전시키려는 정부의 정책적 배려와 지원이 필요하지 않을까 생각한다. 추후 미래 한의 임상 연구 전략 수립에 대한 보다 나은 연구가 지속적으로 발표되어 양질의 한의 임상 연구가 활발히 이루어지길 기대해본다.

Acknowledgements

본 연구는 보건복지부 한의약선도기술개발사업의 지원에 의하여 이루어진 것임. (과제고유번호 HI13C0530)

References

1. Shea, J.L. Applying evidence-based medicine to traditional chinese medicine: debate and strategy. *J Altern Complement Med.* 12(3):255-263, 2006.
2. Ernst, E. Methodological aspects of Traditional Chinese Medicine(TCM). *Ann Acad Med Singapore.* 35(11):773-774, 2006.
3. Leung, P.C. A practical way of research in Chinese medicine. *Ann Acad Med Singapore.* 35(11):770-772, 2006.
4. Verhoef, M.J., Casebeer, A.L., Hilsden, R.J. Assessing efficacy of complementary medicine: adding qualitative research methods to the "Gold Standard". *J Altern Complement Med.* 8(3):275-281, 2002.
5. Jiang, M., Yang, J., Zhang, C., Liu, B., Chan, K., Cao, H., Lu, A. Clinical studies with traditional Chinese medicine in the past decade and future research and development. *Planta Med.* 76(17):2048-2064, 2010.
6. Watanabe, K., Matsuura, K., Gao, P., Hottenbacher, L., Tokunaga, H., Nishimura, K., Imazu, Y., Reissenweber, H., Witt, C.M. Traditional Japanese Kampo Medicine: Clinical Research between Modernity and Traditional Medicine-The State of Research and Methodological Suggestions for the Future. *Evid Based Complement Alternat Med.* 2011: 513842, 2011.
7. Jung, K.Y., Go, H.Y., Sun, S.H., Jeong, J.J., Park, J.S., Song, Y.K., Kim, T.H., Hong, S.I., Choi, Y.K., Go, S.G., Lim, E.M., Lee, C.Y., Park, J.H., Jeon, C.Y. Study on the Current Situation and Issues for Clinical Research in Korean Medicine worldwide - Future Clinical Research Strategy I -. *Korean J Oriental Physiology & Pathology.* 28(2):137-145, 2014.
8. Fønnebo, V., Grimsgaard, S., Walach, H., Ritenbaugh, C., Norheim, A.J., MacPherson, H., Lewith, G., Launsø, L., Koithan, M., Falkenberg, T., Boon, H., Aickin, M. Researching complementary and alternative treatments-the gatekeepers are not at home. *BMC Med Res Methodol.* 11:7:7, 2007.
9. Cochrane, S., Possamai-Inesedy, A. Looking outside the square: the use of qualitative methods within complementary and alternative medicine: the movement towards rigour. *Complement Ther Med.* 21(1):73-76, 2013.
10. Porzolt Franz, Eisemann. M., Habs, M. Complementary alternative medicine and conventional medicine should use identical rules to complete clinical trials. *Eur J Integr Med.* 2(1):3-7, 2010.
11. Mason, S., Tovey, P., Long, A.F. Evaluating complementary medicine: methodological challenges of randomised controlled trials. *BMJ.* 325(7368):832-834, 2002.
12. Carter, B. Methodological issues and complementary therapies: researching intangibles?. *Complement Ther Nurs Midwifery.* 9(3):133-139, 2003.
13. Verhoef, M.J., Lewith, G., Ritenbaugh, C., Boon, H., Fleishman, S., Leis, A. Complementary and alternative medicine whole systems research: beyond identification of inadequacies of the RCT. *Complement Ther Med.* 13(3):206-212, 2005.
14. Yan, J., Engle, V.F., He, Y., Jiao, Y., Gu, W. Study

- designs of randomized controlled trials not based on Chinese medicine theory are improper. *Chin Med*. 2009;4:3, 2009.
15. Tang, J.L. Research priorities in traditional Chinese medicine. *BMJ*. 333(7564):391-394, 2006.
 16. Tang, J.L., Leung, P.C. An efficacy-driven approach to the research and development of traditional Chinese medicine. *Hong Kong Med J*. 7(4):375-380, 2001.
 17. Cassidy, C.M. methodological issues in investigations of massage/bodywork therapy. AMTA Foundation, Evanston. pp 10-15, 2002. Available from: URL: http://www.massagetherapyfoundation.org/clientuploads/Resources/Cassidy_paper.pdf. Accessed Feb 4, 2015.
 18. MacPherson, H. Pragmatic clinical trials. *Complement Ther Med*. 12(2-3):136-140, 2004.
 19. Godwin, M., Ruhland, L., Casson, I., MacDonald, S., Delva, D., Birtwhistle, R., Lam, M., Seguin, R. Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. *BMC Med Res Methodol*. 22;3:28, 2003.
 20. Firenzuoli, F., Gori, L. Herbal medicine today: clinical and research issues. *Evid Based Complement Alternat Med*. 4(Suppl 1):37-40, 2007.
 21. Cardini, F., Wade, C., Regalia, A.L., Gui, S., Li, W., Raschetti, R., Kronenberg, F. Clinical research in traditional medicine: priorities and methods. *Complement Ther Med*. 14(4):282-287, 2006.
 22. Nahin, R.L., Straus, S.E. Research into complementary and alternative medicine: problems and potential. *BMJ*. 322(7279):161-164, 2001.
 23. Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., Elbourne, D., Egger, M., Altman, D.G. Consolidated Standards of Reporting Trials Group. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol*. 63(8):e1-37, 2010.
 24. Sikorskii, A., Wyatt, G., Victorson, D., Faulkner, G., Rahbar, M.H. Methodological issues in trials of complementary and alternative medicine interventions. *Nurs Res*. 58(6):444-451, 2009.
 25. Gagnier, J.J., Boon, H., Rochon, P., Moher, D., Barnes, J., Bombardier, C. CONSORT Group. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med*. 144(5):364-367, 2006.
 26. MacPherson, H., Altman, D.G., Hammerschlag, R., Li, Y., Wu, T., White, A., Moher, D. STRICTA Revision Group. Revised STAndards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement. *Acupunct Med*. 28(2):83-93, 2010.
 27. Zwarenstein, M., Treweek, S., Gagnier, J.J., Altman, D.G., Tunis, S., Haynes, B., Oxman AD, Moher D; CONSORT group; Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*. 337:a2390, 2008.
 28. Bower, P., King, M., Nazareth, I., Lampe, F., Sibbald, B. Patient preferences in randomised controlled trials: conceptual framework and implications for research. *Soc Sci Med*. 61(3):685-695, 2005.
 29. Li, J., Tian, J., Ma, B., Yang, K. N-of-1 trials in China. *Complement Ther Med*. 21(3):190-194, 2013.
 30. Johnston, B.C., Mills, E. n-of-1 randomized controlled trials: an opportunity for complementary and alternative medicine evaluation. *J Altern Complement Med*. 10(6):979-984, 2004.
 31. Broom, A., Adams, J. Current issues and future directions in complementary and alternative medicine (CAM) research. *Complement Ther Med*. 15(3):217-220, 2007.
 32. Schnyer, R.N., Conboy, L.A., Jacobson, E., McKnight, P., Goddard, T., Moscatelli, F., Legedza, A.T., Kerr, C., Kaptchuk, T.J., Wayne, P.M. Development of a Chinese medicine assessment measure: an interdisciplinary approach using the delphi method. *J Altern Complement Med*. 11(6):1005-1013, 2005.
 33. Berle, C.A., Cobbin, D., Smith, N., Zaslowski, C. A novel approach to evaluate Traditional Chinese Medicine treatment outcomes using pattern identification. *J Altern Complement Med*. 16(4):357-367, 2010.
 34. Critchley, J.A., Zhang, Y., Suthisisang, C.C., Chan, T.Y., Tomlinson, B. Alternative therapies and medical science: designing clinical trials of alternative/complementary medicines is evidence-based traditional Chinese medicine attainable?. *J Clin Pharmacol*. 40(5):462-467, 2000.
 35. Long, A.F. Outcome measurement in complementary and alternative medicine: unpicking the effects. *J Altern Complement Med*. 8(6):777-786, 2002.
 36. Sikorskii, A., Wyatt, G., Victorson, D., Faulkner, G., Rahbar, M.H. Methodological issues in trials of complementary and alternative medicine interventions. *Nurs Res*. 58(6):444-451, 2009.
 37. Eton, D.T., Bauer, B.A., Sood, A., Yost, K.J., Sloan, J.A. Patient-reported outcomes in studies of complementary and alternative medicine: problems, solutions, and future directions. *Explore (NY)*. 7(5):314-319, 2011.

38. Calvert, M., Brundage, M., Jacobsen, P.B., Schönemann, H.J., Efficace, F. The CONSORT Patient-Reported Outcome (PRO) extension: implications for clinical trials and practice. *Health Qual Life Outcomes*. 11: 184, 2013.
39. Leung, K.F., Liu, F.B., Zhao, L., Fang, J.Q., Chan, K., Lin, L.Z. Development and validation of the Chinese Quality of Life Instrument. *Health Qual Life Outcomes*. 3: 26, 2005.
40. Efficace, F., Horneber, M., Lejeune, S., Van Dam, F., Leering, S., Rottmann, M., Aaronson, N.K. Methodological quality of patient-reported outcome research was low in complementary and alternative medicine in oncology. *J Clin Epidemiol*. 59(12):1257-1265, 2006.
41. Zhang, L., Zhang, J., Chen, J., Xing, D., Mu, W., Wang, J., Shang, H. Clinical Research of Traditional Chinese Medicine Needs to Develop Its Own System of Core Outcome Sets. *Evid Based Complement Alternat Med*. 2013:202703, 2013.
42. Normile, D. Asian medicine. The new face of traditional Chinese medicine. *Science*. 299(5604):188-190, 2003.