



임상예방의료에 필요한 과학적 증거의 평가방법

이 중 엽¹ · 박 병 주^{1,2*} | 서울대학교¹ 의과대학/서울대학교병원 의학연구협력센터, ²의과대학 예방의학교실

A method to evaluate scientific evidence in clinical preventive medicine

Joongyub Lee, MD¹ · Byung-Joo Park, MD^{1,2*}

¹Medical Research Collaborating Center, Seoul National University College of Medicine & Seoul National University Hospital,

²Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea

*Corresponding author: Byung-Joo Park, E-mail: bjpark@snu.ac.kr

Received September 8, 2011 · Accepted September 22, 2011

Evaluation of scientific evidence in the medical literature is based on the research methodology, which can be the fundamental tool for medical professionals to improve their practice. Since analytic studies usually provide evidence for generating hypotheses and selecting appropriate research designs, this article aims to review the methods for evaluating the study quality in randomized controlled trials (RCTs), cohort studies, and case-control studies. Critical appraisal of systematic errors including selection bias, performance bias, attrition bias, detection bias, and publication bias are the essential elements of the evaluation. Clinical trials need to have an adequate description of sequence generation and allocation concealment, blinding for exposure measurement, completeness of follow-up and intention to treat analysis, and blinding for outcome evaluation. For cohort studies, like RCTs, appropriate control of confounding variables is needed to prevent selection bias, and blinding and completeness of follow-up are also very important. Matching of the case group to the control group, blinding of interviewers, and proper definition of the cases and controls are important to prevent bias in a case-control study. Since the process of critical appraisal depends on the quality of reporting, there have been efforts to improve the reporting quality of the medical literature. However, reporting all of the elements necessary to avoid bias does not automatically guarantee an exclusion of bias or the quality of a study; instead, what is important is the logical connection of the elements of a study and the eventual lucid expression of the elements.

Keywords: Randomized controlled trial; Cohort studies; Case-control studies; Bias; Critical appraisal

서 론

임상의사가 의학문헌을 읽을 때는 환자진료에 도움이 되는 새로운 정보를 얻음과 동시에 그 타당성에 대한 평가를 하게 된다. 즉, 어떤 필요성에 입각하여 연구를 한 것인지,

어떻게 연구를 수행하였는지, 연구결과 생산된 과학적 증거는 무엇인지를 살펴보고, 마지막으로 자신의 임상경험에 비추어 그 문헌에서 제시한 과학적 증거의 타당성과 함께 향후 환자진료에 어떻게 적용할 수 있을 것인지 헤아리게 된다. 마찬가지로 임상예방의학에서도 만성질환을 예방하기 위하

© Korean Medical Association

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.

Table 1. Evidence hierarchy of the research methodology

Characteristics	Study design	Evidence hierarchy
Descriptive study	Case report	Low
	Case series study	
	Cross-sectional study	
Analytic study	Case-control study	High
	Cohort study	
	Randomized controlled trial	

여 제공하는 임상예방의료서비스의 과학적 증거를 확인하기 위하여 기존에 발표된 의학문헌을 검색하고 평가하게 된다. 이러한 과학적 증거의 평가과정에서 독자는 연구의 타당성 및 적절성과 관련된 연구의 요소를 체계적으로 찾아서 검토하는 과정을 거친다. 타당성과 적절성을 평가할 때 검토의 대상이 되는 요소는 연구과정에서 갖추어지게 되는 것이므로 어떤 의미에서 과학적 증거의 평가방법을 안다는 것은 과학적 증거 생성의 원리, 즉 연구방법론을 안다는 것이 된다. 따라서, 연구방법론에 대한 지식은 의학연구자들의 전유물이라기 보다는 주어진 임상상황에서 환자 또는 건강인을 위하여 최선의 서비스를 제공하고자 애쓰는 모든 의료인이 가져야 할 필수적인 도구라 할 수 있다[1].

임상연구는 크게 기술적인 연구와 분석적인 연구로 나눌 수 있다(Table 1). 기술적인 연구에는 단면조사연구, 환자군연구, 환자사례보고 등이 있다. 환자사례보고는 기존에 알려져 있지 않은 특이한 질병양상을 보인 환자의 임상적 특성과 임상경과 및 의심되는 원인에 관하여 보고하거나, 특정한 약물을 복용한 후에 특이한 유해사례를 나타낸 환자에 대한 경과를 보고하는 것이다. 환자군연구는 동일한 질병을 가진 환자들을 모아서 공통점을 요약함으로써 질병의 원인에 관한 가설을 제기하는데 유용한 방법이다. 임상에서는 흔히 특정 질병을 진단받고 이미 치료를 받았던 환자들을 병원의무기록을 이용하여 후향적으로 찾아낸 후 환자들의 공통점을 파악하거나, 과거에 치료받았던 내용들을 비교하여 치료방법들 간 치료효과의 차이에 관한 가설을 제기하는데 주로 활용되고 있다. 단면조사연구는 질병과 특정 노출요인에 대한 정보를 같은 시점, 또는 짧은 시간 내에 얻는 역학적

연구 형태로 확률표본의 추출 여부, 그리고 연구참여율에 따라 비뚤림이 발생할 수 있다. 이들 기술적 연구는 비뚤림이 개입될 가능성이 높아서 연구결과는 낮은 수준의 근거로 평가받으며 분석적 연구의 전단계로 주로 가설을 생성하는 기능을 수행하게 된다[2].

한편, 분석적 연구는 주요 임상현상을 노출변수와 결과변수로 설정하고 이들 사이의 인과적 연관성을 입증하는 연구로서 환자-대조군연구, 코호트연구 및 임상시험 등이 있다. 분석적 연구는 대상자의 선정방법, 연구정보의 수집방법, 피험자 추적관찰의 완전성, 분석의 방향 등에 따라 내적 타당도와 외적 타당도가 달라진다. 본 고에서는 진료현장에서 실질적으로 임상적 판단의 근거로 많이 활용되며, 과학적 근거를 이차적으로 생산하는 방법인 체계적 문헌고찰과 메타분석, 경제성 평가, 성과연구 등에서 재료가 가장 많이 활용되는 환자-대조군연구, 코호트연구 및 임상시험에 대하여 살펴보고, 각각의 연구설계에 대한 평가방법을 설명한다.

무작위배정 비교임상시험

1. 무작위배정 비교임상시험의 특징

임상시험은 안전성이 완전히 파악되지 않은 새로운 치료제의 효능과 안전성을 평가하기 위하여 직접 사람을 대상으로 수행하는 실험적 연구이다. 무작위배정 비교임상시험은 비뚤림들을 최소화시킬 수 있는 연구설계로서 주된 가설이나 이차적인 가설들을 정량적으로 수립하고 이를 확인하는 것을 연구목적으로 한다. 임상시험의 형태는 설계에 따라 연구대상으로 선정된 환자를 연구시작 시점에 치료군과 비교군 가운데 한 군으로 무작위배정한 후 전체 연구기간 동안 해당 치료법만을 적용받는 평행설계, 연구대상 환자들에게 무작위로 치료법을 배정하여 일정한 기간 동안 치료를 시행한 후 일정한 휴약기간을 경과한 다음 교차된 약물을 투여하는 교차설계, 독립적인 두 가지 치료법의 효과를 동시에 평가하는 요인설계 등이 있다[3].

시험군과 비교군 간에는 치료내용 외에는 결과에 영향을 미칠 수 있는 모든 요인이 동등하게 분포하여야 한다. 즉, 비교군에 배정된 피험자들의 성, 연령 및 병기 등 연구대상 질

병의 경과에 영향을 미칠 수 있는 기초상태의 특성을 최대한 시험군과 유사하게 만들어 비교성을 확보하여야 한다. 비교군은 치료내용에 따라서 무치료군, 위약비교군, 표준치료제 비교군, 저용량비교군 등으로 구분이 가능하며, 진단기준이나 보조치료의 내용이 다를 가능성이 있는 과거비교군 보다는 치료군과 같은 시기에 선정되는 동시비교군을 설정하는 것이 바람직하다.

구체적 가설을 설정한 후 그러한 연구목적에 달성하는데 필요한 최소한의 연구대상수를 연구계획 수립 시에 결정하는 것이 윤리적이다. 하지만, 연구대상수가 너무 적은 경우에는 통계적 검정력이 낮아져 연구가설을 타당하게 검정할 수 없다. 따라서, 연구목적에 구체적으로 설정한 후 그러한 목적을 달성하는데 필요한 최소한의 연구대상수를 통계적으로 산출하여 연구계획서에 명시하여야 한다. 합리적인 연구대상수 산출을 위하여 과거에 수행되었던 유사한 연구결과가 있는지를 문헌검색을 통하여 파악하고, 파악된 연구결과를 종합하여 연구대상수 산출에 필요한 통계량을 확보하여야 한다.

무작위배정법은 치료군과 비교군의 환자에게 치료방법을 배정할 때 연구자의 주관적인 의지가 개입되지 않도록 무작위로 배정하는 방법으로서, 새로운 치료에서 비롯하는 예기치 않았던 중대한 유해반응으로 인한 피해나 예상하지 못하였던 탁월한 치료효과로 인한 혜택이 연구참여자들에게 공정하게 돌아가도록 만들므로 윤리적이며, 비교군간 예후요인의 분포를 같게 하여 비교성을 극대화함으로써 과학적 연구결론을 도출할 수 있도록 해주기 때문에 임상시험에서 가장 중요한 요소이다.

평가항목이 통증과 같은 임상증상의 완화나 인지기능의 개선 및 삶의 질 개선 등 주관적인 요소가 강할 때, 혹은 연구대상자나 치료효과 평가자가 치료내용을 알게 되면 치료효과 판정에 영향을 미쳐 정보비뚤림을 유발할 수 있는 경우에는 눈가림법을 적용하여야 한다. 연구대상이 되는 피험자만 치료내용을 모르게 하는 것을 단일눈가림법, 피험자를 비롯하여 환자와 접촉할 수 있는 모든 의료인력들인 의사, 약사, 간호사 및 의료기사들이 치료내용을 모르게 연구를 수행하는 것을 양측눈가림법, 여기에서 자료를 분석하고 결과를 해석하게 되는 임상역학자나 의학통계학자들까지 비밀로

한 다음 최종 결론이 내려진 후에 치료내용을 공개하는 경우를 삼측눈가림법으로 정의하고 있다.

2. 무작위배정 비교임상시험에서 비뚤림과 평가

선택비뚤림은 비교하고자 하는 치료집단 사이에 존재하는 계통적 차이 때문에 발생한다. 치료배정에 대한 정보가 연구참여자는 물론 연구자에게도 알려지지 않도록 적절한 방법을 사용하는 것은 대단히 중요하다. 선택비뚤림은 순서생성(sequence generation)과 배정은폐(allocation concealment)로 평가한다. 실행비뚤림은 임상시험이 진행되는 동안 시험군과 비교군에 제공되는 중재에 계통적 차이가 있음을 말한다. 의도하지 않은 치료 혹은 위약효과를 배제하기 위하여 제공하는 치료에 대해 ‘눈가림법’이 이루어져서 배정된 치료를 받는 사람들이 자신에게 배정된 치료의 내용을 몰라야 한다. 실행비뚤림은 연구참여자, 연구자, 결과평가자에 대한 눈가림법과 타당도에 대한 잠재적 위협이 존재하는지로 판정한다. 탈락비뚤림은 연구수행과정에 비교하고자 하는 집단에 속한 대상자 가운데 탈락하는 사람들간 계통적 차이 때문에 발생한다. 실제 연구수행에서 참여자 탈락(포기, 탈락, 프로토콜 위반)을 다루는 방법이 부적절한 경우가 많으며, 논문에서 추적관찰에 대한 설명 자체가 명확하지 않은 경우도 있다. 탈락비뚤림은 불완전한 결과자료의 처리와 눈가림법으로 평가한다. 결과확인비뚤림이란 결과를 평가하는 과정과 방법이 두 군 사이에 다르게 적용되므로써 발생할 수 있는 계통적 차이를 말한다. 치료법 배정에 대한 눈가림이 적절하게 이루어지면 이런 비뚤림이 발생할 가능성이 작아진다. 결과확인비뚤림은 눈가림법과 다른 형태의 비뚤림으로 판정한다[4].

코호트연구

1. 코호트연구의 개념과 특징

코호트연구는 연구대상이 되는 질환을 경험하지 않은 사람을 연구대상으로 선정한다. 연구대상을 특정 요인에 노출된 사람들과 노출되지 않은 사람들로 나눈 후 추적관찰하여 각 군의 질환발생률을 비교함으로써 특정 노출요인과 질

환발생간의 인과적 관련성을 평가하는 연구이다. 국내에서 시행된 코호트연구의 예로 1993년 이후 부산광역시에 거주하는 65세 이상 노인인구를 대상으로 구축된 한국노인약물 역학코호트가 있다. 한국노인약물역학코호트 연구는 1993년부터 1998년 사이의 부산광역시 공무원 및 사립학교교직원을 대상으로 하는 의료보험관리공단의 의료보험 피보험자 및 피부양자 46,113명을 대상으로 수행되었다. 연구방법은 의료보험관리공단에 청구된 약물처방자료를 통해 약물 노출 여부를 확인하고, 음주, 흡연, 식생활습관, 출산력, 수면 및 활동상황 등의 변수에 대한 정보를 얻기 위하여 자기 기입식 설문지를 이용한 설문조사를 수행하였다. 또한 병원 방문조사를 통하여 연구대상자들이 병의원을 방문하여 진료받았던 질병들에 대한 최종 진단명을 확인하였고, 암발생 여부는 국립암센터에서 운영하고 있는 한국중앙암등록본부의 1993년부터 1998년까지 암등록자료를 이용하여 확인하였다. 그리고 1992년부터 2004년까지의 통계청 사망자료를 이용하여, 사망연도, 월, 일, 사망원인 등을 확인하였다[5]. 이렇게 구축된 코호트 내에서 근위대퇴골 골절을 경험하지 않은 건강한 노인 가운데 항정신성 약물을 복용하고 있는 사람들과 항정신성 약물을 복용하지 않는 사람들을 선정한 뒤, 일정한 기간 동안 추적 관찰하여 두 군에서의 근위대퇴골 골절 발생률을 산출하여 비교함으로써 항정신성 약물이 노인에서 근위대퇴골 골절 발생간의 인과적 관련성이 있음을 확인하였다[6].

코호트연구의 장점은 한 가지 위험요인에 대한 노출로 인하여 발생할 수 있는 여러 건강결과를 함께 파악할 수 있고, 일반 인구집단에서처럼 특정 위험요인에 대한 노출빈도가 높지 않은 경우에도 적용할 수 있으며, 위험요인에 관한 정보를 연구대상(질병)이 발생하기 전에 파악하기 때문에 환자-대조군연구처럼 환자의 기억에 의존하지 않으므로 더욱 정확하게 파악할 수 있으며, 동시에 위험요인과 질병 발생간의 시간적인 선후관계가 명확하기 때문에 인과관계의 평가에 더욱 유리하다는 점 등을 들 수 있다. 하지만, 코호트연구는 발생률이 낮은 질병과 특정 위험요인간의 관련성을 밝히고자 하는 경우에는 대규모의 연구대상을 필요로 하여 연구수행 가능성이 떨어진다. 특정 위험요인에 노출된 사람들을

대상으로 발생률이 낮은 장기간에 걸친 추적관찰을 통하여 관찰하려면, 연구대상자뿐 아니라 연구자의 중도탈락이 문제가 되며, 연구비도 많이 들게 된다. 또한 추적 관찰하는 기간 동안에 연구대상 위험요인에 대한 노출량과 강도가 변하는 것을 정확하게 측정하기 힘들며, 건강결과의 발생 여부를 파악하기도 어려워지며, 발생한 건강결과의 정확한 진단명을 확인하기도 어려운 점들도 코호트연구의 단점이다.

2. 코호트연구의 비뚤림과 평가

코호트연구에서는 연구자가 연구대상 코호트 구성원으로부터 노출요인을 관찰하여 자료를 수집하는 것이지 노출요인을 대상자에게 직접 배정할 수 없다. 즉 노출요인에 대한 무작위배정이 불가능하다. 선택비뚤림을 통제하는 강력한 수단인 무작위배정을 활용할 수 없는 코호트연구에서는 선택비뚤림이 항상 문제가 된다. 선택비뚤림이란 연구대상자 모집시점에 이미 비교군간 체계적인 차이가 존재함을 의미한다. 관찰연구에서 연구대상이 되는 중재의 적용여부는 임상 의사의 판단뿐만 아니라 의사와 환자의 선호도, 병력 등에 영향을 받는다. 즉, 치료법 이외의 요인들에서 불균형이 존재하고 이 불균형이 치료의 선택에 영향을 미쳐 관찰하고자 하는 치료결과를 전적으로 치료의 선택에 따른 것으로 판단할 때, 치료효과를 과대 혹은 과소평가하게 된다. 대표적인 것으로 '적응증에 의한 교란'이 있으며, 이는 치료결과에 치료의 적응증이 선택비뚤림을 통하여 영향을 미치는 것을 말한다. 코호트연구에서 선택비뚤림은 교란변수의 통제를 중심으로 평가되었으며, 전통적으로 짝짓기, 층화, 보정 등의 방법이 활용되었다. 최근에는 측정되지 않은 교란요인의 영향까지 고려한 도구변수, 주변구조모형 등이 각광을 받고 있으며 코호트연구에서 적극 활용되고 있다[7,8].

한편, 코호트연구에서는 많은 수의 연구대상자에 대하여 장기간에 걸쳐 자료를 수집하는 것이 보편적이기 때문에 노출요인에 대한 측정방법을 표준화하여 실행비뚤림의 발생을 최소화하여야 한다. 추적관찰의 완전성은 코호트연구에서도 임상시험에서와 같이 중요하다. 노출군 또는 비노출군 중 어느 한 쪽에서 중도탈락이 상대적으로 많이 발생하면 연구결과는 비뚤어지게 된다. 노출군에서 비노출군에 비하여 건강결과

Table 2. Methodology to avoid the bias in analytic study designs

Bias	Case-control study	Cohort study	Clinical trial
Selection	Matching	Control for confounders	Random allocation and concealment of assignment
Performance	Measurement of exposure	Measurement of exposure	Blinding
Attrition	Completeness of follow up	Completeness of follow up	Completeness of follow up and ITT analysis
Detection	Case definition	Blinding	Blinding

ITT, intention to treat.

(질병)의 발생을 더 면밀히 찾음으로써 노출요인이 연구대상 질병의 원인으로 잘못 판정되는 결과확인비뚤림이 발생할 수 있다. 이를 방지하기 위한 눈가림법 적용 여부를 확인하는 것은 비뚤림의 개입 가능성을 평가하는데 반드시 필요하다[9].

환자-대조군연구

1. 환자-대조군연구의 개념과 특징

환자-대조군연구는 특정한 질병을 가진 환자군과 그 질병을 가지지 않으면서 이들과 유사한 특성을 가진 대조군을 선정하여 두 집단에 속한 사람들의 과거경험을 비교분석함으로써 질병발생에 유의하게 영향을 미치는 위험요인을 밝히고자 하는 연구설계이다. 국내에서 시행된 환자-대조군연구의 예로 2002년 출혈성 뇌졸중을 경험한 사람들을 환자군으로 선정하고, 출혈성 뇌졸중을 경험하지 않은 동일한 성과 연령을 가진 사람을 대조군으로 선정하여, 페닐프로파놀아민을 함유한 감기약 노출여부를 평가하여 페닐프로파놀아민과 출혈성 뇌졸중간의 인과적인 연관성을 연구한 급성출혈성뇌졸중연구(acute brain bleeding analysis)가 있다. 급성출혈성뇌졸중연구에서는 우리나라 11개 지역 33개 병원에서 출혈성 뇌졸중으로 확진된 30-84세 환자 940명과 동일한 연령과 성으로 짝지은 대조군 1,880명에 대하여 인구사회학적 변수, 과거병력, 사회력, 약물노출력을 의무기록 자료 및 면접조사를 통하여 수집하였다. 특히 약물력 노출 정보를 정확히 파악하기 위하여 당시 시판중인 소염진통제들의 실물과 포장지 실물 등을 수집하여 면접조사에 활용하였고, 모든 면접조사원들은 연구가설에 대하여 눈가림된 상

대로 설문조사를 수행하였다. 이렇게 수집된 자료로 조건부 로짓회귀분석을 시행한 결과 페닐프로파놀아민을 함유한 감기약은 출혈성뇌졸중과 연관성을 가진 것이 확인되었고 이를 근거로 식품의약품안전청에서는 페닐프로파놀아민을 시장에서 퇴출시켰다[10].

환자-대조군연구는 비교적 드물게 발생하는 건강결과와 관련된 위험요인을 파악하는데 유리하고, 단일한 건강결과를 발생시키는 여러 요인들을 동시에 파악할 수 있으며, 비교적 단시간에 연구결과를 얻을 수 있다는 장점이 있다. 하지만, 환자-대조군연구는 연구대상 위험요인에 노출된 사람이 적은 경우에는 적용하기 어려우며, 과거에 노출된 위험요인에 관한 정확한 정보를 파악하기 어려운 경우에는 거의 이용할 수 없다는 단점이 있다. 또한 특정한 위험요인을 대상으로 하여 위험요인에 노출된 사람에서 특정 건강결과와 발생을 관찰하는 연구가 아니고, 이미 건강결과가 발생된 환자로부터 과거에 경험한 위험요인에 관한 정보를 파악하는 연구이기 때문에 특정 요인에 따른 건강결과와 발생률을 직접 산출할 수는 없다.

2. 환자-대조군연구의 비뚤림과 평가

환자-대조군연구는 환자군과 비교할 대조군을 임의로 선발하게 되므로 대조군을 선정하는 과정에 비뚤림이 발생할 가능성이 크다. 선택비뚤림을 통제하며 대조군을 선정하기 위하여 주요 변수에 대한 짝짓기를 수행한다. 하지만, 짝짓기를 수행하더라도 통제되지 않는 비뚤림들이 발생할 수 있어 이에 대한 평가가 필요하다. 또한, 과거 위험요인에 노출된 정보를 수집하는 과정에서 정확성이 떨어지거나, 환자군과 대조군에 대하여 위험요인에 관한 정보를 수집하는 과정이 같지 않으면 두 군에서 위험요인에 대한 노출정도가 실제와 다르게 파악됨으로써 연구결과가 비뚤어지는 정보비뚤림도 발생할 수 있다. 따라서 환자-대조군연구에서도 조사도구의 표준화는 필수적인 과정이며, 도구의 개발과 타당도 평가가 반드시 선행되어야 한다. 환자군에서 대조군 보다

위험요인 노출여부를 더 면밀히 조사하는 수행비뚤림을 방지하기 위하여 조사자를 연구가설에 대해 눈가림하여 노출 요인에 관한 정보를 수집하도록 하여야 한다. 탈락비뚤림의 가능성은 추적관찰의 완결성으로 평가되며, 결과확인비뚤림은 환자군의 정의가 어떻게 되었는지, 그리고 환자여부를 어떻게 확인하였는지를 검토함으로써 평가할 수 있다.

평가방법의 한계점

이상으로 대상 선정 및 정보수집방법, 추적관찰의 완전성에 따라 발생할 수 있는 여러가지 비뚤림들을 통제하는 방법론을 연구설계별로 확인해 보았다(Table 2). 그런데, 이러한 비뚤림을 조절하며 적절하게 연구를 수행하였다 하여도, 연구를 수행한 결과를 정리하다 보면 표본오차나 측정오류 등으로 인하여 어떤 결과는 유의하게, 어떤 결과는 유의하지 않게 나오기 마련이다. 이 때 연구자 입장에서 유리하고 바람직한 결과는 보고하고, 그렇지 않은 결과는 보고하지 않게 되는 경향이 있는데, 이러한 선택적 보고로 인하여 발생하는 비뚤림을 보고비뚤림이라 한다. 이 비뚤림은 모든 연구설계에 있어서 발생할 수 있으며, 연구계획단계에 어떤 결과를 보고하고자 계획하였는지 알지 못하면 그 발생여부를 확인하기는 대단히 어렵다. 또한 거시적으로 보았을 때, 이러한 경향을 가진 연구자들이 많게 되면 후발 연구자들은 이전 연구자들이 부분적으로 보고한 결과를 근거로 자신은 다른 내용을 부분적으로 보고하는 것을 합리화할 수 있어서, 결과적으로 보고비뚤림은 보고비뚤림을 낳는 악순환에 빠져들게 된다. 또한 연구를 후원한 제약회사, 의학잡지의 편집인도 유의하지 않은 결과보다는 유의한 결과를 출판하고자 하는 경향을 가지며, 이러한 경향에 의하여 출판된 문헌을 거시적으로 평가할 때는 유의한 쪽으로 비뚤리는 출판비뚤림이 발생하게 된다. 이러한 경향은 유의한 결과를 보고하는 연구가 유의하지 않은 연구를 보고하는 연구보다 더 양질의 연구라는 보장이 없음에도 불구하고 발생한다[11]. 공개된 임상연구등록시스템에 연구를 등록할 때 연구에서 밝히고자 하는 주요 연구가설을 미리 등록하도록 하여 이러한 비뚤림을 방지할 수 있다. 임상시험에서는 연구의 사전등록이 이제는

국제적으로 보편화되어 있으나, 관찰연구에서는 아직 연구등록시스템 구축이 진행 중에 있어서 보고비뚤림과 출판비뚤림에 대한 평가는 정보의 부족으로 인하여 쉽지 않다.

이렇듯, 연구에 대한 평가는 보고된 내용에 기반하게 되므로 분석적 임상연구의 평가에 필수적인 요소를 반드시 보고하도록 권장하는 지침들이 있는데, 임상시험에서 consolidated standards of reporting trials (CONSORT) 지침이, 그리고 관찰적 연구에서는 strengthening the reporting of observational studies in epidemiology (STROBE) 지침이 활용되고 있다[12,13]. 많은 의학잡지들이 이 지침들을 활용하여 논문을 작성하는 것을 권장하고 있어서 독자들이 연구의 질을 직접 평가하는 것을 돕고 있다. 이러한 지침의 활용과 임상연구등록시스템 덕분에 무작위배정 비교임상시험의 경우 비교적 쉽게 문헌의 질을 평가할 수 있을 것이라 생각할 수 있다. 하지만 비뚤림을 줄이기 위한 조치를 단순히 기술하였다 하여 비뚤림 방지가 저절로 보장되는 것은 아니다. 연구전체를 조망하여 이러한 조치가 얼마나 기능한 지에 대한 평가가 이루어져야 하며, 이러한 평가에 근거하여 연구문헌을 평가하였을 때 연구결과가 얼마나 믿을만한가에 대한 결정을 내릴 수 있다. 관찰적 연구는 이러한 측면에서 그 평가가 간단하지 않다. 임상시험과 달리 근본적으로 제거할 수 없는 비뚤림을 가지고 있는 경우가 있으며, 모든 의학적 지식이 실험적 연구를 통하여 생산될 수 없는 까닭에 전체적인 근거의 위계에서 보았을 때 최고수준의 근거가 관찰적 연구에서 나올 수밖에 없는 경우들도 많기 때문에 관찰적 연구가 가지는 필연적인 비뚤림의 가능성만을 근거로 질이 낮은 연구로 평가하는 것도 부적절하다. 즉, 근거에 대한 평가는 개별적인 비뚤림의 요인을 어떻게 통제하였는가에 대한 평가, 그리고 그 통제가 전체 연구에서 어떻게 작동하여 내적 타당도와 외적 타당도를 높이고 있는 지에 대한 평가, 마지막으로 전체 근거체계의 맥락에서 해당 연구가 가지는 가치에 대한 평가까지 이루어질 때 완성된다.

결론

지금까지 과학적 근거생성의 주요 연구설계라 할 수 있는

환자-대조군연구, 코호트연구, 임상시험의 특징 및 연구설계별로 비뚤음을 줄이기 위한 조치를 살펴보았다. 이들 연구 방법론은 임상과학과 예방의학의 과학적 근거를 생산하고, 생산된 근거를 평가하는데 활용되고 있다. 이러한 연구 방법론의 장단점과 평가방법을 숙지하고 적절하게 활용하는 것은 변화하는 질병패턴에 맞추어 생성될 새로운 근거들을 과학적으로 평가하여 진료에 도입하거나 보건정책수립에 반영함으로써 국민의 건강수준을 향상시키는데 기여할 수 있다. 특히 질병이 발생하기 전이지만 만성질환에 걸릴 위험이 높은 고위험군에 대하여 제공하는 임상예방의료서비스가 효과적으로 질병을 예방하는 지를 과학적으로 평가할 수 있는 능력을 갖추어 실제 연구수행을 통하여 과학적 증거를 생산하는 것은 대단히 중요한 의미를 가진다. 그렇게 함으로써 외국에서 수행된 연구결과를 직수입하여 우리나라 사람에게 맞는지 여부를 확인하지도 않은 채 단순적용하는 단계에서 탈피하여 직접 우리나라 사람을 대상으로 수행한 연구결과로 과학적 증거를 삼아 예방의료서비스를 발전시켜 나가야 할 것이다.

핵심용어: 무작위배정 비교임상시험; 코호트연구; 환자-대조군연구; 비뚤림; 비판적 평가

REFERENCES

1. Crombie IK. The pocket guide to critical appraisal: a handbook of healthcare professionals. London: BMJ Publishing Group; 1996. 66 p.
2. Leon G. Epidemiology. 3rd ed. Pennsylvania: Saunders; 2004. 335 p.
3. Ahn YO. Epidemiology: the principles and applications. 1st ed. Seoul: SNU Press; 2005. 347 p.
4. Park BJ, editor. evidence-based healthcare. 1st ed. Seoul: Korea Medical Book; 2009. 364 p.
5. Park BJ, Cho YK, Kim SA. Construction of the Korea elderly pharmacoepidemiologic cohort: drug utilization review of cephalosporins in geriatric inpatients. *Pharmacoepidemiol Drug Saf* 2001;10:487-492.
6. Bae JM, Koo HW, Jung KO, Park BJ. A cohort study on the association between psychotropics and hip fracture in Korean elderly women. *J Korean Med Sci* 2002;17:65-70.
7. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology* 2006;17:268-275.
8. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550-560.
9. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. Tokyo: ITP; 1982. 529 p.
10. Yoon BW, Bae HJ, Hong KS, Lee SM, Park BJ, Yu KH, Han MK, Lee YS, Chung DK, Park JM, Jeong SW, Lee BC, Cho KH, Kim JS, Lee SH, Yoo KM; Acute Brain Bleeding Analysis (ABBA) Study Investigators. Phenylpropanolamine contained in cold remedies and risk of hemorrhagic stroke. *Neurology* 2007;68:146-149.
11. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *J Pharmacol Pharmacother* 2010;1:100-107.
12. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; Iniciativa STROBE. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Rev Esp Salud Publica* 2008;82:251-259.
13. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;337:867-872.



Peer Reviewers' Commentary

본 논문은 국내에 생소한 임상예방의학에 이해시킬 목적으로 임상연구방법론을 전반적으로 다루고 있다. 이는 최선의 예방의료 서비스가 이루어지기 위해서는 필요한 근거가 무엇인가를 알아내고, 임상연구를 통해 그 근거를 확보하는 것이 필수적이기 때문이다. 즉, 다양한 의학연구 방법론을 올바르게 이해하고 연구에 실제 적용하며, 그 결과들을 객관적으로 해석하는 능력을 길러서, 과학적이고 타당한 근거를 확보하는 것이다. 이렇게 얻어진 근거만이 의료 서비스를 제공하는 의료진뿐만 아니라 소비자인 환자들이 의사 결정 과정에 도움이 될 수 있기 때문이다. 관련된 기존의 문헌들을 체계적으로 고찰하는 연구방법론과 임상진료지침 갱신 등의 연구방법론들이 소개됨으로써 최신의 근거들이 예방의료서비스에 빠르게 도입될 수 있길 바란다.

[정리:편집위원회]