

근거중심 소아치과학 연구의 체계적 고찰과 메타분석

이 광 희

원광대학교 치과대학 소아치과학교실

국문초록

연구 목적은 근거중심 소아치과학 연구의 기본이 되는 체계적 고찰의 방법과, 체계적 고찰의 통계적 방법이 되는 메타 분석에 대하여 소개하는 것이었다. 연구 자료로서 Cochrane Handbook for Systematic Reviews of Interventions 등을 참고하였다. 체계적 고찰의 첫 단계이자 가장 중요한 것은 잘 규정된, 초점이 또렷한 질문을 만드는 것이고, 다음에 자료의 검색을 하고 선정된 자료에 대한 타당도 평가를 한 후, 변수의 종류에 따라 자료를 분석하고 결과를 해석하여 결론을 도출한다. 자료 분석의 통계학적 방법인 메타분석의 필요성과 원리 및 변수의 종류에 따른 자료별 메타분석 방법에 대하여 요약 기술하였다.

주요어 : 근거중심, 소아치과학, 체계적 고찰, 메타분석

I. 서 론

소아치과학은 치의학의 한 분야이고, 치의학은 넓은 의미에서 의학의 한 분야이며, 의학은 과학의 한 분야이다. 우리는 과학적 근거에 의해 환자를 진료한다고 생각하고 있으나, 실제로는 임상 진료의 많은 부분이 과학적 근거가 없이 행해지고 있다. 의료인들이 과거에 배운 지식이 과학적 근거가 결여된 경우가 많고 또 새로운 과학적 근거들이 통합되고 요약되어 의료인들에게 전달되기 어렵기 때문이다. 임상 진료가 과학적 근거에 근거해야 한다는 필요성에 따라 근거중심의학(evidence-based medicine)이 탄생하였다. 근거중심의학은 최고의 연구 근거를 임상 기술 및 환자의 가치와 통합하는 것이라고 정의되었다¹⁾.

근거중심의학은 1992년에 캐나다에서 처음으로 소개된 이후로 선진국을 중심으로 급속히 성장하고 있으며, 근거중심치의

학(evidence-based dentistry)도 발전하고 있다. 국내에서는 아직까지 근거중심치의학에 관한 관심과 연구가 부족한 실정이다. 이²⁾는 근거중심 소아치과학의 개념을 소개하고, 근거중심 소아치과학을 ‘어린이와 청소년의 치과 진료에서 최고의 과학적 연구 근거를 소아치과의사의 임상 기술, 그리고 어린이 환자 및 그 보육자의 가치와 통합하는 것’으로 정의할 것을 제안하였다.

근거중심의학의 연구는 문헌의 체계적 고찰(systematic reviews)과 임상진료지침의 수립을 기본으로 한다. 근거중심의학에서의 체계적 고찰은 일반적인 논문이나 증례의 문헌 고찰에서 흔히 볼 수 있는 서술적 고찰과 다르다. 체계적 고찰에서 다수의 연구결과들을 하나로 통합하는 데 사용되는 통계적 방법을 메타분석(meta analysis)라고 한다. 따라서 체계적 고찰과 메타분석은 근거중심의학의 연구에서 매우 중요한 방법이라고 할 수 있다.

이에 저자는 근거중심 소아치과학의 발전을 위한 연구의 일환으로서, 근거중심 소아치과학 연구의 기본이 되는 체계적 고찰의 방법과, 체계적 고찰의 통계적 방법이 되는 메타분석에 대하여 소개하고자 한다.

교신저자 : 이 광 희

전북 익산시 신용동 344-2

원광대학교 치과대학 소아치과학교실

Tel: 063-850-1955

E-mail: kwhlee@wonkwang.ac.kr

*이 논문은 2006년도 원광대학교의 교비지원에 의해서 수행됨.

II. 연구 자료 및 방법

근거중심치의학센터(Centre for Evidence Based Dentistry, <http://www.cebd.org>)에서 제공하는 자료들과 The British Medical Journal에 실린 체계적 고찰 및 메타분석에 관한 일련의 기사들을 참고하여, 체계적 고찰의 순서와 방법 및 메타분석의 기본 개념과 방법에 대하여 연구하였다 (Table 1~3). 이 중에서 특히 Cochrane Handbook for Systematic Reviews of Interventions은 Cochrane Collaboration에서 4분기마다 발행하는 Cochrane Database of Systematic Reviews의 방법론을 제시하는 문헌으로서, 체계적 고찰의 원리와 방법에 대해 잘 요약되어 있다. Cochrane Collaboration은 영국의 역학자(epidemiologist) Archie

Cochrane의 이름을 따서 1993년에 설립된 비영리 국제기관으로서, 체계적 고찰을 통해 최신 근거중심의학 정보를 제공하고 있다.

III. 체계적 고찰의 순서와 방법

체계적 고찰은 기본적으로 어떤 치료(intervention)의 효과에 관한 임상 시험의 결과들을 종합하여 임상진료지침을 만드는 것을 목적으로 한다. 지금 의학계에서는 많은 양의 정보가 계속 생성되고 있기 때문에 임상진료에서 올바른 판단을 내리기 위한 근거를 얻기 위하여는 정보들을 효율적으로 통합하는 과정이 필요하다.

Table 1. Resources for systematic reviews

QUOROM checklist ³⁾
MOOSE checklist ⁴⁾
QUADAS ⁶⁾
Cochrane Handbook for Systematic Reviews of Interventions ⁶⁾
Cochrane review production resources ⁷⁾
Undertaking Systematic Reviews of Research on Effectiveness ⁸⁾

Table 2. Systematic reviews in health care from The British Medical Journal

Assessing the quality of controlled clinical trials ⁹⁾
Investigating and dealing with publication and other biases in meta-analysis ¹⁰⁾
Systematic reviews of evaluations of diagnostic and screening tests ¹¹⁾
Systematic reviews of evaluations of prognostic variables ¹²⁾

Table 3. Meta-analysis series from The British Medical Journal

Meta-analysis: Potentials and Promise ¹³⁾
Meta-analysis: Principles and Procedures ¹⁴⁾
Meta-analysis: Beyond the Grand Mean ¹⁵⁾
Bias in Location and Selection of Studies ¹⁶⁾
Spurious Precision: Meta-analysis of Observational Studies ¹⁷⁾
Meta-analysis: Unresolved Issues and Future Developments ¹⁸⁾

Table 4. Key components of a question

What types of people (participants)?
1. Define the diseases of interest
2. Identify the population of interest
What types of comparisons (interventions)?
What types of outcomes?
What types of study designs?
Randomised controlled trials (RCTs) - therapeutic efficacy
Case-control and cohort studies - etiology or risk factors

1. 문제의 규정

체계적 고찰의 첫 단계이자 가장 중요한 것은 잘 규정된, 초점이 또렷한 질문을 만드는 것이다. 질문에 첫째로 포함되어야 하는 것은, 관심의 대상이 되는 질병을 정의하고, 관심의 대상이 되는 인구집단을 확인하는 것이다(Table 4). 또한 연구 디자인에서는 치료효과를 보려면 무작위 대조 시험(randomised controlled trial, RCT)을, 원인이나 위험요인을 보려면 환자-대조군 연구(case-control study)나 코호트 연구를 선택하여야 한다.

2. 자료의 검색

MEDLINE은 미국 국립의학도서관에서 발행한 Index Medicus에 따른 데이터베이스로서, PubMed를 통해 온라인에서 무료로 제공된다. EMBASE는 Elsevier가 발행하는 Excerpta Medica에 근거한 것이며 유럽의 MEDLINE이라고 할 수 있다. 포괄적인 검색을 위해서는 두 데이터베이스를 모두 이용하는 것이 권장된다. 그밖에 Cochrane Central Register of Controlled Trials (CENTRAL)은 대조 시험과 관련된 가장 포괄적인 자료이다. 또한 SciSearch는 과학기술문헌에 대한 데이터베이스이다. 전자 데이터베이스를 검색할 때에는 데이터베이스 이름, 호스트 이름, 날짜, 검색에 포함된 연도, 검색전략, 언어 제한 여부 등을 정보에 포함시켜야 한다.

그러나 모든 자료가 전자화되어 있지 않고 또 모든 단어가 전자 데이터베이스의 색인에 들어있지 않기 때문에 손으로 찾는 검색이 필요하다. 대한소아치과학회지를 비롯한 많은 국내 자료들이 데이터베이스에 포함되어 있지 않으므로 직접 검색해야 한다.

가장 편리한 방법은 기존 고찰들의 참고문헌을 보고 원 기사들을 찾는 것이다. Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects (DARE) 등에서 고찰들을 찾을 수 있다. 진행 중인 연구에 대한 정보는 TrialsCentral™ (www.trialscentral.org)과 Current Controlled Trials (www.controlled-trials.com)에서 검색할 수 있다.

일정한 형식에 따라서 자료를 모으게 되는데, 체계적 고찰을 위해 자료를 수집하는 컴퓨터 소프트웨어들이 개발되어 있다.

컴퓨터를 사용할 때에는 독립된 두 개의 평가 방법을 사용하여 이중으로 추출하는 과정을 거친 후 결과를 비교할 수 있다. QuatroPro, Excel, Lotus 같은 Spreadsheet 소프트웨어나 FoxPro, DataEasesuch 같은 데이터베이스 프로그램을 사용할 수 있다. 자료수집 형식에 포함되어야 하는 것은 문헌, 저자, 연구 대상 및 방법, 연구 성적과 결과 등이다.

3. 타당도의 평가

타당도(validity)란 연구의 설계와 수행이 치우침(bias)을 방지하는 정도를 가리킨다. 타당도가 높은 연구일수록, 보다 진리에 근접한 결과를 낳는다. 치료 효과를 과대 평가하는 쪽으로 치우친 연구에서는 가양성(假陽性) 결론이, 과소 평가하는 쪽으로 치우친 연구에서는 가음성(假陰性) 결론이 나올 것이다. 모든 연구들의 결론이 일관성이 있더라도 그 연구들이 모두 타당도가 낮을 가능성을 배제할 수 없다.

연구에서 치우침이 생기는 원인은 크게 네 가지로 구분된다(Table 5). 먼저 연구를 시작할 때 실험군과 대조군을 무작위로 배정하는 과정에서 표본의 편중(selection bias)이 생긴다. 다음으로, 효과를 연구하고자 하는 치료(intervention) 외에 다른 치료가 실험군과 대조군 간에 동일하지 않기 때문에 치우침(performance bias)이 생긴다. 또한 연구 기간 중에 연구대상자가 탈락함으로써 발생하는 치우침(attrition bias)이 있다. 끝으로 연구의 결과를 평가할 때 생기는 치우침(detection bias)이 있다.

연구의 타당도를 평가하는 방법에는 개별적 기준에 부합되는지 여부를 보는 방법과 여러 문항으로 된 척도나 checklist를 사용하여 점수를 부여하는 방법 등이 있다. 복잡한 방법이 더 나은 것은 아니며 단순한 방법이 바람직하다.

타당도 평가는 그 자료를 연구에 포함시킬 것인가 여부를 판정하는 기준으로, 연구들 간에 결과가 다른 것을 설명하는 근거로, 민감도 검사를 위해, 연구결과의 통계학적 분석을 위한 가중치로 사용될 수 있다. 타당도로 연구 결과의 차이를 설명하는 방법에는 평가된 타당도의 순위에 따라 연구 결과를 결합하는 것(cumulative meta analysis), 평가된 타당도에 따라 연구들에 가중치를 부여하는 통계학적 방법을 사용하는 것 또는 타당도와 효과 간의 관계를 탐색하기 위해 메타 회귀(meta regression)를 사용하는 것 등이 있다.

Table 5. Sources of bias

Study procedure	Bias
Allocation of control group	Selection bias
Intervention	Performance bias
Follow-up	Attrition bias
Outcomes	Detection bias

4. 자료의 분석

자료의 분석은 서술적일 수도 있고 통계 분석을 포함하여 정량적일 수도 있다. 두 개 이상의 분리된 연구들의 결과를 통계학적으로 결합하는 메타분석은 가장 많이 사용되는 통계 기법이다. Cochrane review writing software (RevMan)으로 다양한 메타분석을 할 수 있지만 모든 체계적 분석에서 메타분석이 적합한 것은 아니다. 자료의 종합을 위해서는 (치료나 약물의) 효과의 방향, 효과의 크기, 연구들 간의 효과의 일치 여부, 근거의 강도(strength of evidence)가 분석되어야 한다. 메타분석의 방법에 관하여는 아래의 IV. 메타분석에서 별도로 다루었다.

가. 명목변수 자료

명목변수 자료의 효과는 risk ratio, odds ratio, risk difference 등으로 측정된다(Table 6). Risk는 좋지 않은 결과가 나올 확률로서, 0에서 1 사이의 소수로 표현되며 백분율로 전환될 수 있다. Odds는 사건이 일어나지 않을 확률에 대한 사건이 일어날 확률로서, 0에서 무한대까지 어떤 수도 될 수 있다. 0.1의 odds는 1:10으로 쓸 수 있다.

Odds의 해석을 정확하게 하기 위해서는 odds를 risk로 전환하는 것이 간단한 방법이다. Odds가 0.1일 때, 사건이 일어나지 않는 사람 열 명에 대하여 사건이 일어나는 사람 한 명이 있으며, risk는 $0.1/(1+0.1) = 0.091$ 이다. Odds가 1일 때, 사건이 일어나지 않는 사람 한 명에 대하여 사건이 일어나는 사람 한 명이 있으며, risk는 $1/(1+1) = 0.5$ 이다. Risk와 odds의 상호 전환 식은 다음과 같다:

$$\begin{aligned} \text{risk} &= \text{odds} / 1 + \text{odds} \\ \text{odds} &= \text{risk} / 1 - \text{risk} \end{aligned}$$

Risk ratio 3은 처치했을 때 사건이 일어날 확률이 처치하지 않았을 때에 비해 세 배임을 뜻하며, 처치가 사건의 위험을 200% 증가시킨다고 말할 수 있다. Risk ratio 0.25는 처치했을 때 사건이 일어날 확률이 처치하지 않았을 때의 1/4임을 뜻하며, 처치가 사건의 위험을 75% 감소시켰다고 말할 수 있다. 임상적으로 risk ratio 0.75는 어떤 사건의 발생률이 80%에서 60%로 감소하는 것, 또는 4%에서 3%로 감소하는 것을 뜻한다. Odds ratio 또한 risk ratio로 전환해서 해석할 수 있다. Risk ratio(RR)와 odds ratio(OR)의 상호 전환 식은 다음과 같다(BR은 치료를 하지 않았을 경우에 사건이 일어날 전형적인 risk):

$$RR = \frac{OR}{1 - (BR \times (1 - OR))} \quad OR = \frac{RR(1 - BR)}{1 - (BR \times RR)}$$

Risk difference 0.05(5%)는 발생률이 90%로 높은 질병에서는 발생률이 85%로 감소하는 것을 의미하여 임상적 중요성이 적으나, 발생률이 5%로 낮은 질병에서는 발생률이 10%로 두 배로 증가하는 것이 된다.

나. 연속변수 자료

연속변수 자료의 효과는 mean difference와 standardised mean difference로 측정된다. Mean difference는 두 집단의 평균 간의 절대적 차이를 뜻하는 것이며, standardised mean difference는 연구들이 서로 다른 척도로 동일한 내용을 측정하는 경우에 메타분석에서 사용된다.

다. 서열변수 자료

서열변수 자료가 명목변수를 위한 방법으로 요약될 때에는 risk ratios, odds ratios, risk differences 등으로 효과를 분석할 수 있고, 연속변수 자료를 위한 방법으로 요약될 때에는

Table 6. Calculation of RR, OR, RD
The results of a clinical trial :

	Dental caries	No dental caries	Total
Fluoride application	a	b	a+b
Control group	c	d	c+d

risk ratio = risk of event in intervention group / risk of event in control group =

$$\frac{a/(a+b)}{c/(c+d)}$$

odds ratio = odds of event in intervention group / odds of event in control group =

$$\frac{a/b}{c/d} = \frac{ad}{bc}$$

risk difference = risk of event in intervention group - risk of event in control group =

$$\frac{a}{a+b} - \frac{c}{c+d}$$

mean difference와 standardised mean difference로 효과를 분석할 수 있다.

라. Count와 rate

희소한 사건의 수는 Poisson 자료라고 하며, rate는 그 사건이 일어날 수 있는 기간 중에 실제로 일어난 수를 가리킨다. 메타분석에서는 한 집단의 rate를 다른 집단의 rate로 나눈 rate ratio를 사용한다. 우식경험치 수(DMFT)와 같이 흔한 사건의 수는 연속변수 자료와 동일한 방법으로 다룰 수 있다.

5. 결과의 해석

체계적 고찰의 결론에는 근거의 강도, 응용가능성, 예상되는 효과와 치료 비용 간의 균형 등을 포함시켜야 한다.

가. 근거의 강도(strength of evidence)

시험들과 고찰의 방법론적인 한계들을 언급하여야 한다. 치료와 결과 사이의 인과관계를 규명하는 데 도움이 되는 한 가지 근거는 상관성의 간접적 근거이다. 여기에는 risk의 지표가 되는 생리적 또는 생화학적 측정치와 같은 중간 결과들, 동물실험을 포함한 상이한 집단을 대상으로 한 연구들로부터의 근거, 유사한 연구들로부터의 근거 등이 포함된다.

치료 효과에 관한 인과관계 추론에서 고려해야 할 사항은 시험의 질, 관찰된 효과의 크기와 유의성, 효과의 시험들간 일관성, 명백한 용량-반응 관계의 유무, 추론을 지지하는 간접적 근거의 유무, 관찰된 효과에 대한 가능한 다른 설명들이 배제되었는가 여부 등이다.

나. 적용가능성(applicability)

특정 치료의 잠재적인 이익과 해로움에 대하여 고찰이 타당성이 있는 정보를 제공하는지 여부를 먼저 결정하여야 한다. 다음으로 고찰에 포함된 연구들의 대상자들과 조건들이 우리의 임상 상황과 유사한지를 결정해야 한다. 부가적으로, 고찰에 포함된 연구들에서 제공된 치료들의 특성 또는 부가적인 치료를

고려하여야 한다. 체계적 고찰의 저자는 근거가 적용될 수 있는 상황들의 범주 및 상이한 상황에 따른 효과의 변이를 독자들에게 알려야 한다.

결과의 적용을 제한하는 것은 생물학적 및 문화적 변이, 순응도의 변이, 기초 risk의 변이 등이다. 결과를 고찰하고 결론을 이끌어내는 과정에서 저자들은 다른 사람들의 상황이 자기들의 상황과 같다고 전제하는 것을 피해야 한다. 생물학적 변이는 치료에 대한 반응에 영향을 끼칠 수 있는 남녀간의 차이와 같은 병리생리학적 차이 및 전염성 질병 등에서의 병원체의 차이를 포함한다. 시술자와 피시술자의 순응도 변이는 경제적 조건이나 태도의 차이 등으로 어떤 형태의 치료를 받는 것이 가능하지 않은 경우에 일어날 수 있다.

치료의 이익은 그 치료를 하지 않았을 때 나쁜 결과가 있을 risk 및 치료의 유효성에 달려 있다. 기초 risk와 관련하여 고려해야 할 두 가지 사항 중 첫 번째는 상대적인 이익과 해로움이 적용가능한지 여부이다. 예를 들어, risk가 높은 환자들에서 얻은 결과를 risk가 낮은 환자들에게 적용할 수 있는가는 의심할 만하다. 두 번째는 치료의 충격에서 예상할 수 있는 변이의 정도이다. 예를 들어, 연구들의 대조군들에서 관찰된 기초 risk의 범위를 고려할 수 있다.

결과가 적용될 수 있는 상황 내에서도 결과의 중요한 변이에 관해 토론하고 결론을 끌어내야 한다. 치료의 상대적 효과에서 예측가능한 변이, 그리고 반응이나 효과가 변하게 만드는 요인들을 확인해야 한다. 여기에 포함될 수 있는 것은 환자의 나이, 성별, 생화학적 지표, 치료의 타이밍 또는 강도, 호르몬 수용기 상태와 같은 질병 특성 등이다.

다. 다른 관련 정보

체계적 고찰의 결과를 다른 관련 정보의 맥락에서 토의해야 하는데, 여기에는 문제의 크기와 분포에 관한 역학적 자료, 원무자료나 진료조사에서 얻은 현행 임상진료정보, 치료비에 대한 정보 등이 있다. 이것은 종종 체계적 고찰의 범위를 넘어가기 때문에 임상진료지침을 개발하거나 기술평가를 하는 이들에 의해서 국가적 또는 지역적 수준에서 하는 것이 낫다.

Table 7. Categories of interventions

Evidence to provide clear guidelines for practice	Categories
Sufficient	1. that improve outcome 2. that should be abandoned 3. that involve important trade-offs between known benefits and known adverse effects
Insufficient	1. that appear promising 2. that have not been shown to have the effects expected from them 3. that are not effective for the purpose for which they have been used

라. 위해효과

위해효과와 예상하지 않았던 효과에 대한 확인과 보고는 일반적으로 잘 하지 않으나, 특정 치료에 따른 위해효과의 인과관계를 비평적으로 평가해야 한다.

마. Trade-offs

도출된 결론을 지지하는 근거의 강도를 고려함에 더해서, 서로 다른 결과들에 부여된 가치들(선호도)에 관해 명백한 설명이 있어야 한다. 치료는 일반적으로 이익뿐 아니라 비용과 위험을 수반한다. 치료의 실용성에 대한 결론을 도출하는 것은 추정된 이익과 추정된 비용 및 위험 간의 trade-off(타협을 위한 거래, 교환)를 하는 것을 수반한다.

바. 분류

상호 배타적인 여섯 개 범주로 분류하는 것이 유용할 수 있다 (Table 7). 첫 세 범주들은 충분한 근거가 있어서 상대적으로 확고한 임상적 결론에 도달한 것이고, 다음 세 범주들은 확고한 임상적 결론을 끌어내기 위하여는 더 많은 연구가 필요한 경우이다.

마. 결론 도출에서의 일반적 실수

근거가 결정적이지 않을 때 흔히 하는 실수는 '효과 없음의 근거'를 '효과의 근거가 없음'과 혼동하는 것이다. 근거가 결정적이지 않을 때 치료의 효과가 없다는 대조군의 경우와 다르다고 주장하는 것은 잘못이다. 결과의 감소나 증가 어느 쪽과도 양립할 수 있는 자료를 신뢰구간과 함께 보고하는 것이 안전하다. 효과가 긍정적이지만 통계학적으로 유의하지 않을 때 저자들은 이것을 흔히 유망한 결과로 묘사하지만, 같은 정도로 효과가 부정적일 때는 이것을 경고로 묘사하지 않는다.

결론을 희망사항의 틀에 맞추는 것도 실수이다. 연구결과에서 유병률이 통계학적으로 유의하지 않은 증가를 보였을 때에도 저자들은 '포함된 연구 수가 너무 적어서 유병률의 감소가 탐지되지 않았다'는 식으로 기술할 수 있다. 이러한 실수를 피하는 한 가지 방법은 결과의 방향을 반대로 하였을 때 결론이 어떻게 제시될 것인가를 고려하는 것이다. 치료 효과 차이의 신뢰구간이 효과 없음과 접한다면, 결론은 긍정적일 수도 있고 부정적일 수도 있다. 이 때 한 가지 가능성이 결론에 언급된다면 다른 한 가지 가능성도 언급되어야 한다.

또다른 흔한 실수는 근거를 넘어서는 결론을 도출하는 것이다. 부가적인 정보에 대한 언급이 없이 암묵적으로 이루어지는 경우가 가끔 있다. 부가적인 정보가 언급되더라도 부가적인 정보가 체계적으로 고찰되는 경우는 드물며 진료에의 적용은 특정 상황에 종종 의존한다. 저자들은 진료와 관련된 결론을 도출할 때 주의해야 하며 권고사항을 만드는 것을 피해야 한다.

또한 '더 많은 연구가 필요함' 같은 진부한 표현을 피하고, 어떤 연구가 왜 필요한지 정확히 말해야 한다. 부가적인 자료가 있으면 고찰이 어떻게 개선될 수 있는지에 대한 의견도 언급될 수 있다.

IV. 메타분석

1. 메타분석의 필요성

메타분석은 power를 증가시킨다. Power란 실제 효과가 존재할 경우 통계학적으로 유의하게 그 효과를 발견할 가능성을 말한다. 많은 개별 연구들은 규모가 작아서 작은 효과를 발견하기 어려우나, 여러 연구들을 합하면 효과를 발견할 가능성이 높아진다. 메타분석은 또한 precision을 개선한다. Precision은 무작위 오차로 이끄는 우연 효과의 가능성을 측정하는 것이며 효과 추정치의 신뢰구간에 반영된다. 보다 많은 정보에 기초할 때 치료 효과의 추정이 개선될 수 있다.

메타분석은 개별 연구들에 의해 제기되지 않은 질문들에 대한 답을 제공한다. 1차 연구들은 특정 유형의 환자와 명백히 규정된 치료들을 포함하는데, 특성들이 다른 연구들을 선택함으로써 효과의 일관성 및 차이의 이유를 조사할 수 있다. 또한 외관상 상충하는 연구들로부터 일어나는 논쟁을 가라앉히기 위해서 또는 새로운 가설을 세우기 위해서 메타분석이 필요하다.

메타분석에 대한 흔한 비판은 '사과를 오렌지와 결합시킨다'는 것이다. 다른 성질의 연구들을 함께 분석하려고 하는 것은 무의미하며 효과의 차이를 모호하게 할 수 있다. 질이 낮은 연구들을 메타분석하는 것은 심각하게 오도할 수 있다. 개별 연구들에 존재하는 오차들이 단순히 조합되어 생성된 그릇된 결과가, 메타분석을 했기 때문에 신뢰도가 높은 것으로 해석될 수 있다.

2. 메타분석의 원리

메타분석은 전형적으로 2단계 과정이다. 첫 단계에서, 각 연구의 요약 통계가 계산된다. 대조 시험의 경우, 이 수치들은 각 시험에서 관찰된 치료 효과를 나타낸다. 명목변수 자료인 경우에는 risk ratio가, 연속변수 자료인 경우에는 평균간 차이가 될 것이다.

두 번째 단계에서, 개별 연구들에서 추정된 치료 효과들의 가중 평균으로서, 요약된(종합된) 치료 효과 추정치가 계산된다. 가중 평균(weighted average)는 다음과 같이 정의된다:

$$\text{weighted average} = \frac{\sum TiWi}{\sum Wi}$$

여기서 Ti는 연구에서 추정된 치료효과, Wi는 연구에 주어진 가중치이다. 만일 모든 가중치가 동일하다면 가중 평균은 평균 치료효과와 같다. 연구에 주어진 가중치가 클수록 가중 평균치에 대한 기여도가 높아진다. 비율 측정에서는 Ti가 로그이다.

치료효과 추정치들의 결합은 연구들이 동일한 치료효과를 추정하고 있지 않지만 연구들의 전체 분포를 따르는 치료효과를 추정한다는 가정을 선택적으로 포함할 수도 있다. 이것이 무작위 효과 메타분석의 근거가 된다. 반대로, 각 연구가 정확히 동

일한 양을 추정하고 있다고 가정할 경우에는 고정 효과 메타분석이 수행된다.

요약된 치료효과와 표준오차는 요약 추정치의 precision을 나타내는 신뢰구간을 이끌어내는 데, 그리고 치료효과가 없다는 귀무가설에 대항하는 근거의 강도를 나타내는 유의수준 P값을 이끌어내는 데 사용될 수 있다.

중합된 효과를 정량적으로 요약하는 것과 더불어, 메타분석의 모든 방법들은 개별 연구들의 결과들 간의 변이를 무작위 변이와 비교하는 평가를 포함할 수 있다. 변이가 충분히 크다면 연구들의 치료효과가 불일치함을 가리킬 것이다.

3. 역분산을 통한 메타분석

매우 일반적이고 간단한 메타분석 방법은 역분산법(inverse variance method)으로 알려져 있다. 이 방법은 효과 추정치의 분산의 역(표준오차 제곱의 역)을 각 연구의 가중치로 선택하는 것이다. 연구 규모가 클수록 표준오차가 작아지고 가중치가 커지게 된다.

역분산법을 사용하는 고정효과 메타분석은 가중치를 다음과 같이 계산한다:

generic inverse variance weighted average =

$$\frac{\sum(T_i/S_i^2)}{\sum(1/S_i^2)}$$

여기서 T_i 는 연구에서 추정된 치료효과, S_i 는 그 추정치의 표준오차이다. 분석을 위해 필요한 기초 자료는 각 연구의 치료효과 추정치와 그 표준오차이다.

역분산법의 한 변형은 서로 다른 연구들이 서로 다르지만 연관된 치료효과들을 추정하고 있다는 가정을 포함시키는 것이다. 이것이 무작위 효과를 통한 메타분석이며, 가장 간단한 버전이 DerSimonian & Laird method이다. 위 공식에서 S_i 는 변이의 정도 또는 이질성의 측정치를 포함하도록 조정된다.

4. 명목변수 자료의 메타분석

가. Mantel-Haenszel method

Cochrane review writing software인 RevMan에 고정효과 메타분석 방법으로 프로그램되어 있다. 발생율이 낮거나 표본크기가 작아서 자료가 적을 경우에는 역분산법에 사용되는 효과추정치의 표준오차 추정이 불량할 수 있다. Mantel-Haenszel법에서는 사용되는 효과 추정치의 종류(risk ratio, odds ratio, risk difference)에 따라서 서로 다른 가중치를 부여한다.

나. Peto odds ratio method

이 방법은 odds ratio들을 종합하는 데만 사용할 수 있다.

역분산법을 사용하나 로그 odds ratio를 추정하는 데 근사법을 활용한다. 다른 식으로 본다면 Peto법은 'O ? E' 통계의 합계이다. O는 각 시험의 실험군에서 관찰된 사례 수이고 E는 기대된 사례 수이다. 치료효과가 작을 때, 즉 odds ratio가 1에 가까울 때, 실험군과 대조군의 사례 수가 비슷할 때, 로그 odds ratio의 근사치 계산이 잘 이루어진다. 사례 수가 0인 경우에 보정이 필요하지 않기 때문에 사례 수가 매우 희소한 경우에 적합하다.

다. DerSimonian & Laird 무작위효과법¹⁹⁾

이 방법은 서로 다른 연구들이 서로 다르지만 연관된 치료효과를 추정하고 있다는 가정에 근거해 있다. 역분산법에 기초해서, 다양한 치료효과간의 변이 또는 이질성의 정도에 따라 가중치를 조정한다. 연구들 사이에 이질성이 없을 경우에는 역분산법과 동일한 결과가 나올 것이다. 이질성이 존재하는 경우에는 고정효과법에 비해 평균 치료효과와 신뢰구간이 더 넓어질 것이며 그에 따라 통계학적 유의성의 주장이 보다 신중해질 것이다. 표본 크기와 관찰된 치료효과 사이에 상관성이 있다면 치료효과와 중앙 추정치가 변할 가능성도 있다.

5. 연속변수 자료의 메타분석

역분산법을 사용하는 고정효과법과 DerSimonian & Laird 무작위효과법으로 할 수 있다. 이질성이 없을 경우에 두 방법의 결과는 동일할 것이고, 이질성이 있을 경우에는 고정효과법에 비해 무작위효과법에서 평균 치료효과와 신뢰구간이 더 넓어지고 P값의 유의수준도 낮아질 것이다. 연속자료의 메타분석에 전제된 가정은 각 연구의 각 치료도구의 결과들이 정규분포를 하고 있다는 것이며, 매우 규모가 큰 연구에서는 문제가 되지 않지만 이 가정이 언제나 맞는 것은 아니다.

6. 명목변수 자료와 연속변수 자료의 결합

동일한 자료가 한 연구에서는 명목변수 자료로 제시되고 다른 연구에서는 연속변수 자료로 제시될 수 있다. 예를 들어, 어떤 진정요법의 효과를 Frankl의 4등급 척도로 조사한 경우에 결과를 평균으로 발표할 수도 있고 비협조군과 협조군으로 둘로 나누어 협조군의 백분율로 발표할 수도 있다. 명목변수화한 경우에 이해하기가 더 쉽다. 그러나 다른 척도의 경우에는 둘로 나누는 기준점이 임의적일 수 있고, 연속변수를 명목변수로 전환하는 과정에서 정보가 상실된다.

각 실험군에서의 연속 측정의 분포가 logistic distribution을 하고, 실험군과 대조군에서 결과의 변이가 동일하다는 것을 전제로, 다음 공식을 사용하여 odds ratio를 standardised mean difference로 바꿀 수 있다:

$$SMD = \frac{\sqrt{3}}{\pi} \log OR$$

Log odds ratio의 표준오차는 동일한 상수(0.5513)를 곱함으로써 standardised mean difference의 표준오차로 전환될 수 있다. 반대로 standardised mean difference는 $\pi/\sqrt{}$ = 1.8140을 곱함으로써 log odds ratio로 전환될 수 있다. 메타분석에서 모든 시험들의 standardised mean difference와 표준오차가 일단 계산되면 역분산법을 사용하여 결합시킬 수 있다.

7. 서열변수 및 측정척도 자료의 메타분석

시험자가 원래의 분석을 한 방식에 따라 명목변수 자료 또는 연속변수 자료로 메타분석할 수 있다. 서열 척도가 적은 수의 범주를 가지며, 각 범주에 들어가는 수를 알 수 있고, 동일한 서열 척도가 모든 시험에 사용된 경우에는, proportional odds model을 사용하여 자료를 분석하는 것이 가능하다. 이 방법은 명목변수화하는 것보다 모든 자료를 더 효율적으로 사용할 수 있으나, 고급 통계 소프트웨어를 필요로 하고 요약 통계의 결과에서 임상적 의미를 발견하는 일이 쉽지 않다.

Proportional odds model은 치료 차이로서 proportional odds ratio를 사용한다. 세 개의 범주가 있고 1이 가장 바람직하고 3이 가장 바람직하지 않은 순서로 되어 있다고 가정할 때, 자료는 두 가지 방법으로 명목변수화될 수 있다. 범주 1을 성공으로, 범주 2, 3을 실패로 할 수도 있고, 범주 1, 2를 성공으로, 범주 3을 실패로 할 수도 있다. Proportional odds model은 이 두 경우에 동일한 odds ratio가 있다고 가정한다. 따라서, 서열 범주를 어떻게 성공 또는 실패로 나누는가와 관계 없이, proportional odds model로부터 계산된 odds ratio를 치료 성공의 odds로 해석할 수 있다.

8. Count와 rate의 메타분석

연구대상자 각자가 사건을 경험할 수 있고 또 한 번 이상 경험할 수 있는 경우에는 결과를 count 자료로 표현할 수 있다. 예를 들면, 치과병원 방문 횟수가 있다. 이 사건은 일어나지 않을 수도 있지만, 사건이 일어나더라도 개인에서 이론적 최대치가 존재하지 않는다. Count 자료와 rate 자료는 명목변수 및 연속변수 자료와 같은 방법으로 분석될 수 있다.

관찰되는 기간에 걸쳐 각 참가자마다 count가 측정된다면 rate 자료가 생성된다. Count되는 사건이 희소할 때 특히 적합하다. 예를 들어, 2년의 추적기간 동안 2회의 구내염이 발생하였다면 구내염의 rate는 연당 1이다. Rate는 전통적으로 집단 수준으로 요약되는데, 여기에는 사건이 일어날 확률이 참가자 들마다 그리고 관찰기간 중에 일정하다는 전제가 있다. 상황에 따라 이 전제가 맞는지 검토해야 한다. Count 자료를 rate로 분석하는 것이 항상 적절한 것은 아니다. Risk가 일정하다는 가정이 맞지 않을 수 있고, 다른 종류의 자료들처럼 통계법이 잘 개발되어 있지 않기 때문이다.

시험 결과를 rate ratio로 나타낼 수도 있는데, 이것은 대조군의 rate에 대한 실험군의 rate의 ratio이다. 실험군에서 X 참가자-년(participant-years)당 A회의 사건이, 대조군에서 Y 대조군-년당 C회의 사건이 일어났다고 할 때 rate ratio는 $(A/X)/(C/Y) = AY/CX$ 가 된다. 시험들의 rate ratio 로그를 역분산법으로 결합시킬 수 있다.

V. 결 론

연구 목적은 근거중심 소아치과학 연구의 기본이 되는 체계적 고찰의 방법과, 체계적 고찰의 통계적 방법이 되는 메타분석에 대하여 소개하는 것이었다. 연구 자료로서 Cochrane Handbook for Systematic Reviews of Interventions 등을 참고하였다. 체계적 고찰의 첫 단계이자 가장 중요한 것은 잘 규정된, 초점이 또렷한 질문을 만드는 것이고, 다음에 자료의 검색을 하고 선정된 자료에 대한 타당도 평가를 한 후, 변수의 종류에 따라 자료를 분석하고 결과를 해석하여 결론을 도출한다. 자료 분석의 통계학적 방법인 메타분석의 필요성과 원리, 변수의 종류에 따른 자료별 메타분석 방법에 대하여 요약 기술하였다.

참고문헌

1. Sackett DL, Strauss SE, Richardson WS, et al. : Evidence-Based Medicine: How to Practice and Teach EBM. Chirchill Livingstone, 2nd edition, 2000.
2. 이광희 : 근거중심 소아치과학의 개념과 응용. 대한소아치과학회지, 33:269-280, 2006.
3. <http://www.consort-statement.org/QUOROM.pdf>
4. <http://www.consort-statement.org/Initiatives/MOOSE/moose.pdf>
5. <http://www.biomedcentral.com/1471-2288/3/25>
6. <http://www.cochrane.org/resources/handbook/index.htm>
7. <http://www.cochrane.org/resources/revpro.htm>
8. <http://www.york.ac.uk/inst/crd/report4.htm>
9. Juni P, Altman DG, Egger M : Assessing the quality of controlled clinical trials. BMJ, 323:42-46, 2001.
10. Sterne JAC, Egger M, Smith GD : Investigating and dealing with publication and other biases in meta-analysis. BMJ, 323:101-105, 2001.
11. Deeks JJ : Systematic reviews of evaluations of diagnostic and screening tests. BMJ, 323:157-162, 2001.
12. Altman DG : Systematic reviews of evaluations of prognostic variables. BMJ, 323:224-228, 2001.

13. Egger M, Smith GD : Meta-analysis: Potentials and promise. *BMJ*, 315:1371-1374, 1997.
14. Egger M, Smith GD, Phillips AN : Meta-analysis: Principles and procedures. *BMJ*, 315:1533-1537, 1997.
15. Smith GD, Egger M, Phillips AN : Meta-analysis: Beyond the grand mean? *BMJ*, 315:1610-1614, 1997.
16. Egger M, Smith GD : Meta-analysis bias in location and selection of studies. *BMJ*, 316:61-66, 1998.
17. Egger M, Schneider M, Smith GD : Meta-analysis Spurious precision? Meta-analysis of observational studies. *BMJ*, 316:140-144, 1998.
18. Smith GD, Egger M : Meta-analysis: Unresolved issues and future developments. *BMJ*, 316:221-225, 1998.
19. DerSimonian R, Laird N : Meta-analysis in clinical trials. *Control Clin Trials*, 7:177-188, 1986.

Abstract

SYSTEMIC REVIEWS AND META-ANANLYSIS
IN RESEARCH OF THE EVIDENCE-BASED PEDIATRIC DENTISTRY

Kwang-Hee Lee

Department of Pediatric Dentistry, College of Dentistry, Wonkwang University

The purpose of study was to introduce the method of systemic reviews and the meta-analysis as the statistical method of the systemic reviews. The resource of study was mainly the Cochrane Handbook for Systematic Reviews of Interventions. The first and most important decision in preparing a review is to determine its focus by asking clearly framed questions. Then searching for databases, quality assessment of the studies, collecting data, analysing and presenting results, and interpreting results follow. The necessity and principle of meta-analysis were described and the methods of meta-analysis by the data types were summarised.

Key words : Evidence-based, Pediatric Dentistry, Systemic Reviews, Meta-analysis