
임상적 유용성 평가를 위한 간접비교 지침개발

2010 12 !

서울대학교 산학협력단

제 출 문

건강보험심사평가원장 귀중

건강보험심사평가원으로부터 수탁 받은 '임상적 유용성 평가를 위한 간접비교 지침개발'에 대한 보고서를 제출합니다.

2010년 12월

연구책임자:	한서경(서울대학교 의과대학 교수)
연구원:	안정훈(한국보건의료연구원 연구위원)
	장은진(한국보건의료연구원 책임연구원)
	김윤희(한국보건의료연구원 책임연구원)
연구보조원:	김선희(서울대병원 의학연구협력센터 연구원)
	강수희(서울대병원 의학연구협력센터 연구원)

목 차

1. 서론	1
1.1. 연구 배경	4
1.2. 연구 필요성	2
1.3. 연구 목적	3
2. 연구 방법	5
2.1. 해외 지침서 검토	5
2.1.1. 영국	5
2.1.2. 호주	6
2.1.3. 캐나다	6
2.1.4. 해외학회	7
2.2. 방법론 세미나	7
2.2.1. 정의 및 용어	7
2.2.2. 간접비교 개념 및 방법론	8
2.2.3. 간접비교 전제조건으로의 교환가능성	8
2.2.4. 간접비교의 근거 수준 비교	8
2.2.5. 다양한 지표의 사용 및 장단점	8
2.2.6. 혼합비교	8
2.3. 전문가 자문회의	8
2.3.1. 전문가 자문회의 및 관련단체 간담회	9
2.3.2. 관련단체 의견수렴	10
3. 연구결과	11
3.1. 해외 주요 지침서 고찰	11
3.1.1. 영국	11
3.1.1.1. Report to Methods Review Working Party	11
3.1.1.2. Guide to the Methods of Technology Appraisal	13
3.1.2. 호주	16
3.1.2.1. Report of the Indirect Comparisons Working Group	16

3.1.2.2. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee	18
3.1.3. 캐나다	18
3.1.3.1. Indirect Evidence: Indirect Treatment Comparisons in Meta Analysis	18
3.1.3.2. Guidelines for the Economic Evaluation for the Health Technologies	19
3.1.4. ISPOR	20
3.1.5. 소결	20
3.2. 간접비교 수행에 대한 이론적 고찰	22
3.2.1. 정의 및 용어	22
3.2.1.1. 단순비교	22
3.2.1.2. 공통대조군 간접비교	23
3.2.1.3. 혼합비교	24
3.2.1.4. 공통대조군	25
3.2.1.5. 동질성, 교환가능성, 일관성	25
3.2.1.6. 상대치료효과	26
3.2.2. 이론적 배경 및 전제조건	26
3.2.2.1. 이론적 배경	26
3.2.2.2. 간접비교 전제조건으로의 교환가능성	27
3.2.2.3. 공통대조군 결과와 전제조건에 대한 확인	31
3.2.3. 교환가능성 가정에 영향을 미칠 수 있는 요인과 이질성 평가	36
3.2.3.1. 합성된 상대치료효과의 이질성을 일으킬 수 있는 요인 평가	36
3.2.3.2. 합성된 상대치료효과의 측정치와 통계적 이질성	38
3.2.4. 적절한 결과지표의 조건 및 측정지표 선택시 주의 사항	39
3.2.4.1. 이분형 자료	39
3.2.4.2. 연속형 자료	41
3.2.5. 간접비교에 의한 상대치료효과 추정	43
3.2.5.1. 단순비교	43
3.2.5.2. 공통대조군 간접비교	43
3.2.5.3. 혼합비교	45
3.2.6. 교환가능성 확인 및 간접비교 수행 절차	45

4. 임상적 유용성 평가를 위한 간접비교 수행단계별 자료제출 지침 iii 48

4.1. 서론	48
---------	----

4.2. 검색전략에 대한 기술	50
4.2.1. 요구되는 정보	50
4.2.2. 검색전략	51
4.3. 간접비교에 포함될 무작위배정임상시험의 선정	52
4.3.1. 선정과정 요약	52
4.3.2. 검색결과의 제시	53
4.3.3. 최종 확인된 무작위배정임상시험의 목록 제시	56
4.3.4. 간접비교를 위한 무작위배정임상시험의 선정/제외	57
4.4. 교환가능성 평가	60
4.4.1. 간접비교에 고려된 무작위배정임상시험에서 비뚤림 위험의 요약평가	60
4.4.2. 간접비교에 고려된 무작위배정임상시험의 요약	61
4.4.3. 간접비교에 고려된 무작위배정임상시험별 결과변수	62
4.5. 간접비교 결과 제시 및 해석	63
4.6. 단순비교	65
5. 간접비교 평가 사례 제시	67
5.1. 간접비교 사례선정	67
5.2. 임상시험의 요약 및 교환가능성 평가	68
5.3. 간접비교 결과 제시 및 해석	75
참고문헌	80
부록	83
[부록 1] 혼합비교방법	83
1.1. 이론적 배경	83
1.2. 혼합비교 사례	85
1.2.1. 고정효과 혼합비교 모형	86
1.2.2. 단순 랜덤효과 혼합비교 모형	90
[부록 2] PBAC과 교신 자료	93
[부록 3] 비뚤림 위험 평가 지침	94
[별첨자료] 간접비교 사례제시 문헌	98

표 목 차

표 3-1	합성된 상대치료효과의 이질성에 영향을 미칠 수 있는 요인	37
표 3-2	OR와 RR 비교	40
표 4-1	간접비교를 위해 검색된 후보 무작위배정임상시험 목록	57
표 4-2	간접비교에서 무작위배정임상시험의 제외 이유	59
표 4-3	간접비교에 포함된 약제의 무작위배정임상시험별 요약	59
표 4-4	간접비교를 위한 후보임상시험별 비뚤림 위험 평가결과의 요약	60
표 4-5	간접비교를 위한 후보임상시험별 자료의 완결성에 대한 평가	60
표 4-6	간접비교를 위한 후보임상시험별 선정 제외기준의 요약	61
표 4-7	간접비교를 위한 후보임상시험별 피험자의 특성	62
표 4-8	간접비교를 위한 후보임상시험별 비교되는 중재	62
표 4-9	간접비교를 위한 후보임상시험별 결과변수	63
표 4-10	간접비교 결과의 요약	64
표 4-11	단순비교 결과의 요약	66
표 5-1	간접비교를 위해 검색된 후보 무작위배정임상시험 목록	68
표 5-2	간접비교에서 무작위배정임상시험의 제외 이유	69
표 5-3	간접비교에 포함된 약제의 무작위배정임상시험별 요약	69
표 5-4	간접비교를 위한 후보임상시험별 비뚤림 위험 평가결과의 요약	69
표 5-5	간접비교를 위한 후보임상시험별 자료의 완결성에 대한 평가	70
표 5-6	간접비교를 위한 후보임상시험별 선정 제외기준의 요약	71
표 5-7	간접비교를 위한 후보임상시험별 피험자의 특성	72
표 5-8	간접비교를 위한 후보임상시험별 비교되는 중재	73
표 5-9	간접비교를 위한 후보임상시험별 결과변수	74
표 5-10	간접비교 결과의 요약(24주에서의 ACR20 반응률)	76
표 5-11	간접비교 결과의 요약(24주에서의 ACR50 반응률)	77
표 5-12	간접비교 결과의 요약(24주에서의 ACR20 반응률) - Keystone 포함	78

그 림 목 차

그림 1-1 심평원 약제평가 흐름도	4
그림 1-2 직접비교와 간접비교	3
그림 3-1 간접비교의 분류	22
그림 3-2 단순비교	22
그림 3-3 공통대조군 간접비교의 여러 형태	23
그림 3-4 혼합비교 예	24
그림 3-5 신청약제 A에 대하여 비교약제 B, C, D가 있는 경우	24
그림 3-6 간접비교와 혼합비교의 가정	26
그림 3-7 두 개의 임상시험으로부터 얻어진 간접비교 자료	28
그림 3-8 교환가능성 확인과 간접비교의 수행	46
그림 4-1 제출자료 작성 흐름도	50
그림 4-2 출판된 문헌 검색 및 선정 흐름도	55
그림 4-3 제약사의 미출판 자료원 선정 흐름도	56

1. 서론

1.1. 연구 배경

국내에서는 2007 년부터 의약품 경제성 평가, 약가협상제도를 포함한 선별 등재방식이 본격적으로 시행되고 있으며, 제약회사가 제출한 자료를 바탕으로 건강보험심사평가원이 비용 대비 효과를 과학적으로 평가하고 이를 기초로 제약회사와 건강보험공단이 가격을 협상하게 됨으로써 소비자는 품질 좋은 의약품을 보다 적절한 가격에 복용할 수 있을 것으로 기대하고 있다.

비교약제 대비 효과의 개선 정도를 비용의 증가분에 대해 검토하는 의약품의 경제성 평가 시에도 비용효과성 평가에 앞서 비교약제 대비 신청약제의 효과 개선에 대한 임상적 유용성에 대한 평가가 이루어지고 있으며(그림1-1), 이를 위해 직접비교 임상시험(Randomized Controlled Trial)에 의한 자료의 이용이 가장 적절하나, 비교약제와의 직접비교 임상시험 결과 자료가 부족하거나 부재한 신약 등의 경우 적절한 간접비교를 통한 대체방안의 적용 및 이의 타당성에 대한 검토가 필요하다.

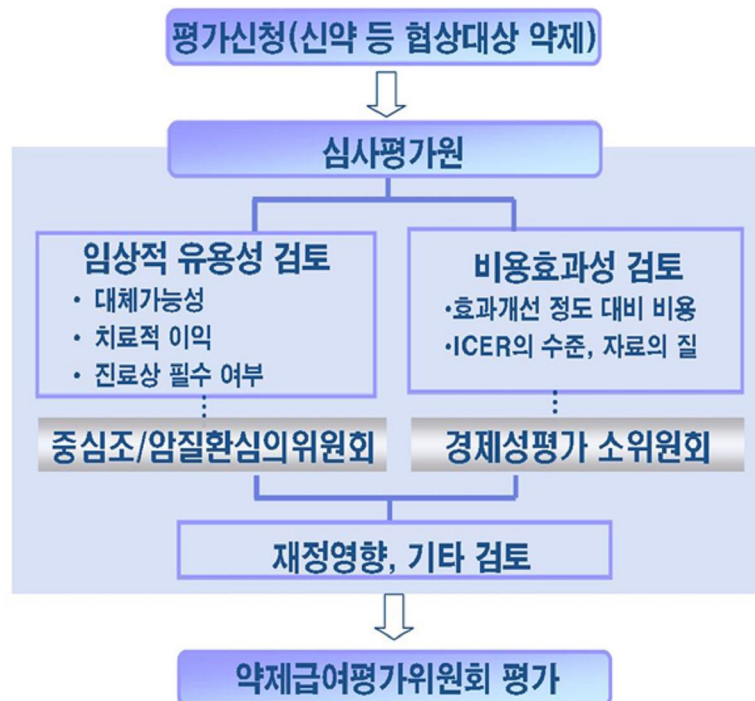


그림 1-1 심평원 약제평가 흐름도

관련 전문가의 층이 매우 얇고 현장의 실무자들이 해당 방법론에 대한 명확한 이해와 경험이 부족한 국내 현실에서, 급여비용 심사 및 평가 업무를 담당하고 있는 건강보험심사평가원의 해당 부서나 평가 자료를 제출해야 하는 제약회사 모두에게 과학적이고 효율적인 평가업무 수행을 도울 수 있는 표준지침서의 필요성이 대두된다.

이에 현장 실무자들로 하여금 기존의 방법론에 대한 이해를 돕고, 해외의 기존 지침서들을 검토하고 조정하여 국내 요양급여의 적정성평가 과정의 객관적이고 과학적인 수행과 표준적인 평가 업무를 도울 수 있는 지침서를 작성하고자 한다.

1.2. 연구 필요성

의약품 선별등재 제도 시행에 따라 신약의 임상적 및 경제적 가치 평가를 통해 우수약품 선별과정이 필요하다. 신약의 효과 등 신약에 대한 임상적 유용성 개선 평가를 위해서는 신약에 대한 비교약제를 선정하고 이와 직접비교 무작위배정임상시험(RCT)을 통하여 얻어진 결과에 근거한 평가가 가장 적절하나, 신약의 경우 비교약제와의 직접비교 임상시험 결과가 없는 경우가 빈번하다.

직접비교 임상시험이 부재하여 신청약제와 비교약제의 개별 임상시험 결과를 단순히 간접적으로 상호 비교할 경우, 개별 임상시험의 연구대상 및 시험방법 등의 상이함에서 유입될 수 있는 비뚤림(bias)이 개입될 가능성이 크므로 신약의 실제 효과 개선을 파악하기 어렵게 된다. 이런 경우, 공통대조군을 통한 간접비교, 혼합비교(mixed/multiple treatment comparison) 등 간접비교 수행을 위한 다양한 통계적 접근법들이 제시되어 온 바 있으나, 이는 여러 가정이 성립될 때 이용할 수 있는 분석기법으로 기존의 방법론들에 대한 명확한 이해를 바탕으로 타당한 접근을 통해 수행되어야 한다.

메타분석에 있어서 개별 임상시험자료를 이용한 간접비교에 대한 방법론이 90년대 말에 소개된 이래로, 네트워크를 통한 메타분석이나 베이지안 메타분석의 개념을 도입하여 보다 복잡한 가정과 방법론을 동원한 통계적 방법론에 대한 활발한 연구가 영국을 중심으로 2000년 초부터 진행되었다. 근래에는 이러한 접근의 타당성에

대한 검토가 이루어지고 있으며 영국 및 호주, 캐나다 등의 의료기술평가기관 및 관련 학회에서는 간접비교 평가지침 마련을 위한 검토와 연구들을 수년에 걸쳐 수행하여 이를 토대로 지침안을 제시하고 있다.

이에 국내에서도 간접비교 평가와 관련하여 그간 축적된 방법론적 성과의 토대 위에서 국내연구 여건 등을 감안한 지침안을 도출하여 간접비교를 수행하여 평가 자료를 작성하거나 제출된 자료에 대한 검토를 수행하는 데 있어 표준적으로 참고할 수 있는 자료의 제시가 필요하다.

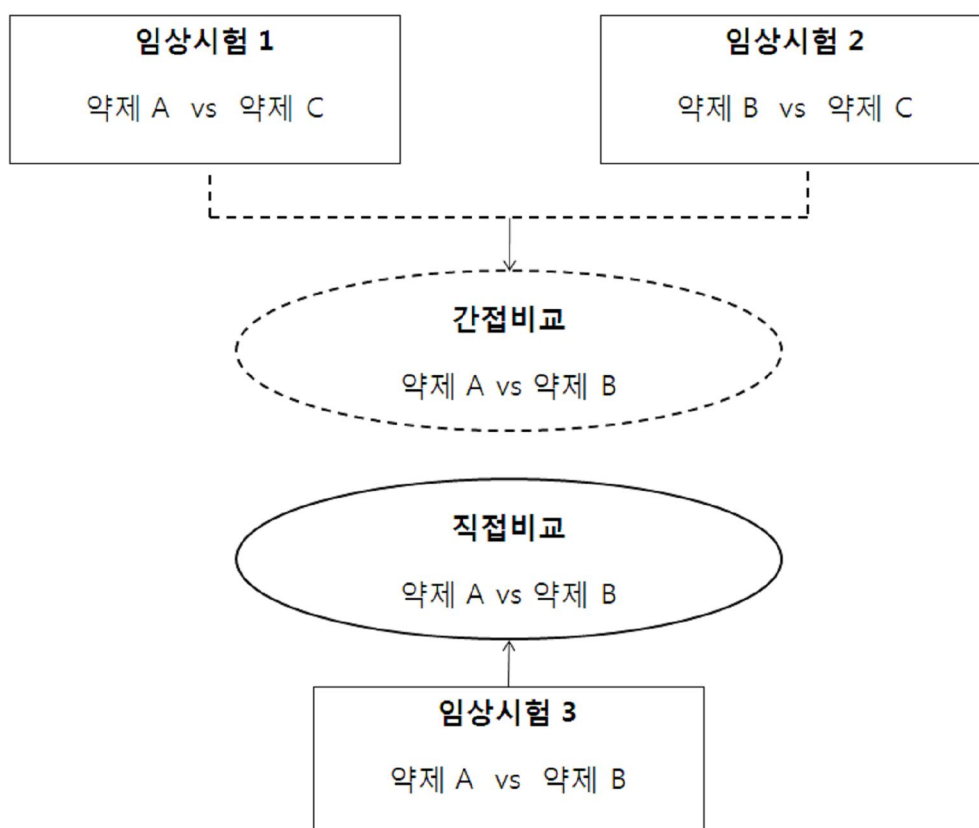


그림 1-2 직접비교와 간접비교

1.3. 연구 목적

본 연구는 비교약제와의 직접비교 자료가 부족한 신청약제에 대한 임상적 유용성 평가를 위한 간접비교지침을 개발하여, 객관적인 효과 비교자료를 생성할 수 있도록 하고 의약품의 요양급여의 적정성 및 경제성 평가수행의 근거로 활용할 수 있고

자 하였다.

개발된 간접비교 수행방법 및 결과 제시 지침을 활용하여 약제 결정신청자의 자료 작성 및 제출을 용이하게 하고 실무 검토자의 표준화된 검토기준으로 활용할 수 있도록 하고자 한다.

2. 연구 방법

국내 요양급여의 적정성평가 과정의 객관적이고 과학적인 수행과 표준적인 평가 업무를 도울 수 있는 지침서를 작성하기 위하여 해외 의료기술평가기관 및 관련 학회에서 발행한 지침 또는 각종 연구문헌들을 검토하고, 방법론 세미나를 통해 기존 간접비교 수행 방법론에 대한 문헌을 정리하고 간접 비교에 대한 개념 및 방법론을 제시하였다.

2.1. 해외 지침서 검토

2.1.1. 영국

□ 지침서: Guide to the methods of technology appraisal. UK National Institute for Health and Clinical Excellence; June 2008

□ 보고서: Report to Methods Review Working Party Key issues arising from workshop on evidence synthesis(indirect and mixed treatment comparison); June 2007

NICE(National Institute for Health and Clinical Excellence)를 중심으로 한 영국의 의료기술평가를 뒷받침해온 평가지침서인 Guide to the methods of technology appraisal과 2008년 6월 평가지침서를 위해 NICE에 제출된 2007년 6월 보고서인 Report to Methods Review Working Party Key issues arising from workshop on evidence synthesis를 검토하였다. 2008년 개정판에 수록되어 있는 간접비교수행에 관한 지침은, 그 이전 판까지의 지침을 재검토하고 혼합비교 등을 포함한 간접비교 관련 최신 방법론의 도입 여지와 사용의 타당성에 관한 제안을 마련하기 위하여, 관련 전문가 그룹으로 구성된 working group이 2년 이상 걸친 토론과 작업을 통해 NICE에 제출한 briefing paper와 공공간담회를 통해 모아진 의견에 기초하고 있다.

2.1.2. 호주

□ 지침서: Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Commonwealth of Australia Pharmaceutical Benefits Advisory Committee; December 2008

□ 보고서: Report of the Indirect Comparisons Working Group to the Pharmaceutical Benefits Advisory Committee: assessing indirect comparisons; 2007

호주에서도 영국과 유사한 간접비교 working group이 2007년 후반에 ESC (Economics Sub-Committee), PBAC (subcommittee of the Pharmaceutical Benefits Advisory Committee) 및 호주 의약품 규제기관인 TGA (Therapeutic Goods Administration) 후원과 참여 하에 추진되어 보고서가 작성되었다. 호주 PBAC을 중심으로 한 호주의 의료기술평가를 뒷받침해온 평가지침서인 Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (2008년 12월 개정판)와 2008년 12월 평가지침서를 위해 PBAC에 제출한 2007년 보고서인 Report of the Indirect Comparisons Working Group to the Pharmaceutical Benefits Advisory Committee를 검토하였다.

2.1.3. 캐나다

□ 지침서: Guidelines for the Economic Evaluation for the Health Technologies: Canada; March 2006

□ 보고서: Indirect Evidence: Indirect Treatment Comparisons in Meta - Analysis; March 2009

캐나다에서 의료기술평가는 CADTH (Canadian Agency for Drugs and Technology in Health)에서 담당하며, CADTH에서 2006년에 발표한 평가지침서인 Guidelines for the Economic Evaluation for the Health Technologies를 검토하였다.

캐나다의 경우 2006년 판 이후로 평가지침서에 대한 개정은 없었다. CADTH에서 2009년 3월에 발표한 보고서인 Indirect Evidence: Indirect Treatment Comparisons

in Meta - Analysis는 간접비교를 위한 구체적인 수행 지침을 담고 있으므로 검토에 포함되었다.

2.1.4. 해외학회

ISPOR(International Society for Pharmacoeconomics and Outcome Research)는 약물 경제학과 치료성과연구 분야에서 가장 권위 있는 학회로서 1995년에 설립돼 해당학문분야 발전에 리더역할을 해오고 있다.

2010년 5월 미국 애틀랜타에서 진행된 ISPOR-15th Annual International Meeting에 참석하여 간접비교와 관련된 이슈들에 대해 파악하였다. 현재 ISPOR에서는 간접비교를 위한 지침서는 개발 중에 있으며, 2010년 9월에 홈페이지에 공개한 초안을 검토하였다.

2.2. 방법론 세미나

간접비교 평가와 관련한 방법론은 아직 확립되지 않았다. 새로운 평가 기법들이 최근에도 계속 개발되고 있으며, 어느 정도 연구자들의 견해가 일치하는 분야도 있지만, 아직까지 몇 가지 방법에 대해서는 이론적 논쟁이 계속되고 있다.

연구기간 동안 정기적인 내부 세미나를 개최하여 다른 나라의 지침서와 주요 방법론적 이슈에 대한 검토 과정을 아래와 같이 수행하였다.

2.2.1. 정의 및 용어

- Network analyses와 Mixed network analyses의 차이 및 범위 확인
- Baseline risk, underlying baseline risk의 정의
- 기저상태의 차이가 보정된다는 것과 다중회귀분석모형에서 교란요인을 보정한다는 것과의 차이점 논의

2.2.2. 간접비교 개념 및 방법론

- 간접비교의 추정치가 비뚤림이 없다고 할 수 있는가?
- 간접비교 방법론이 복잡해질수록, 불확실성은 커지게 되는데, 어느 범위까지 허용할 것인가?
- Song et al.(2009)이 언급한 논문의 동질성과 유사성은 무엇을 의미하는가?

2.2.3. 간접비교 전제조건으로의 교환가능성

- 교환가능성과 유사성은 동일한 의미인가?
- 임상적 동질성이 보장되지 않아도 교환가능성은 성립할 수 있는가?

2.2.4. 간접비교의 근거 수준 비교

- 공통대조군 간접비교는 근거의 수준에서 비무작위 연구보다 높다고 할 수 있는가?
- 간접비교에서 무작위배정의 강점이 유지된다고 생각하는 이유는 무엇인가?

2.2.5. 다양한 지표의 사용 및 장단점

- 공통대조군의 기저위험이 다를 경우에 특정 지표의 사용을 추천할 것인가?
- 간접비교 방법론에서 연속형 자료의 사용을 제한해야 하는가?

2.2.6. 혼합비교

- 혼합비교 수행에 대하여 충분한 평가가 이루어져 있다고 할 수 있는가?

2.3. 전문가 자문회의

건강보험심사평가원 실무진과 연구 진행사항 및 자문가 위촉 등에 관하여 주기적으로 협의하였다. 결과의 객관적인 도출을 위하여 총 2회의 전문가 자문회의와 관련

단체와의 공감대 형성을 위하여 최종보고서 제출 이전에 1회의 이해관계자 간담회를 통해 의견을 수렴하였다.

2.3.1. 전문가 자문회의 및 관련단체 간담회

□ 자문위원

자문위원은 다음과 같다. (가나다 순)

- 김규흔(한국노바티스, 한국다국적의약산업협회)
- 배은영(상지대학교 보건과학대학, 한국보건의료기술평가학회)
- 손영택(덕성여자대학교 약학대학, 한국제약협회)
- 이상무(한국보건의료연구원, 신의료기술평가본부),
- 이상일(울산대학교 의과대학, 한국보건의료기술평가학회)
- 이재원(고려대학교 통계학과, 한국보건정보통계학회)

□ 자문회의 진행

i 임상적 유용성 평가를 위한 간접비교 지침개발i 과제 수행 중 관련전문가 및 건강보험심사평가원 내 해당 업무 관련 실무자 자문회의는 2010년 10월부터 2010년 12월까지 진행되었으며, 동 기간 중 전체 자문회의는 다음과 같이 개최되었다.

■ 1차 자문회의

- 일시: 2010년 10월 27일 18:00
- 장소: 건강보험심사평가원 9층 회의실
- 자문위원: 김규흔, 손영택, 이상무, 이상일, 이재원
- 건강보험심사평가원: 약제등재부 관련 실무자
- 회의안건: 간접비교 방법론에 관한 이론적 고찰 및 적용가능 범위 결정

■ 2차 자문회의

- 일시: 2010년 12월 3일 12:00
- 장소: 서울대학교 의과대학 의생명과학관 114호 회의실
- 전문가 자문위원: 김규훈, 손영택, 이상무, 이재원
- 건강보험심사평가원: 약제등재부 관련 실무자
- 회의안건: 중간보고서에 대한 1차 자문회의 논의결과에 따른 후속조치의 적절성 검토 및 수정 보완된 보고서 초안에 대한 자문단 검토의견 수렴

2.3.2. 관련단체 의견수렴

□ 간담회 진행

i 임상적 유용성 평가를 위한 간접비교 지침개발i 보고서 초안에 대한 발표 및 지침서 사용대상자들의 의견을 청취하고 수렴하기 위해 이해관계자 및 관련 학회를 대상으로 다음과 같이 간담회를 개최하였다.

■ 이해관계자 및 관련단체 간담회

- 일시: 2010년 12월 9일 13:00
- 장소: 서울대학교 의과대학 의생명과학관 115호 강의실
- 참석: 한국다국적의약산업협회, 한국제약협회, 한국보건의료기술평가학회
- 건강보험심사평가원: 약제등재부 실무자
- 내용: 보고서 초안에 대한 관련 협회 및 학회 의견 청취 및 수렴

3. 연구결과

3.1. 해외 주요 지침서 고찰

영국, 호주, 캐나다 등의 의료기술 평가기관에서 간접비교 평가지침 마련을 위한 검토와 연구들이 수년에 걸쳐 수행되었으며, 이를 토대로 지침안이 제시되고 있다. 이에 기존에 제시된 각국의 보고서와 지침서를 대한 검토하고 내용을 고찰하였다. 또한 최근 경제성 분석 및 성과 연구와 관련된 해외학회인 ISPOR를 통해 제시된 간접비교 평가지침에 대해서도 간략하게 소개하고자 한다.

3.1.1. 영국

3.1.1.1. Report to Methods Review Working Party

영국의 2007년 보고서(Report to Methods Review Working Party Key issues arising from workshop on evidence synthesis(indirect and mixed treatment comparison))는 간접비교와 혼합비교의 근거 합성에 대한 워크숍 논의에서 진행된 중요 이슈에 대해 소개하였다. 총 5가지 이슈가 제기되었고 내용은 다음과 같다.

- NICE의 의료기술 평가에서 간접비교와 혼합비교를 도입할 여지가 있는가?
- 간접비교와 혼합비교와 관련된 기술적인 이슈들은 무엇이 있는가?
- 간접비교에 포함될 연구와 결과변수를 어떻게 정의할 것인가?
- 간접비교와 혼합비교 분석 결과를 어떻게 제시할 것인가?
- 간접비교 또는 혼합비교 분석 방법에서 나온 결과는 얼마나 믿을 만한가?

워크숍 참석자들은 NICE의 의료기술 평가(technology appraisal)에서 간접비교와 혼합비교의 도입여지와 사용의 타당성에 대해 다양한 의견을 제시하였다. 이들 의견은 혼합비교를 모든 의료기술평가 대상에서 고려해야 한다는 주장부터 특정 상황에서만 사용하자는 주장, 절대 사용해서는 안 된다는 주장까지 다양했다.

근거의 네트워크를 통해 공통대조군이 연결이 안 되거나 무작위배정비교임상시험이 존재하지 않는 치료법에 대해서는 혼합비교는 도움이 되지 못한다는 것이 제한점으로 논의되었다. 또한 적절한 자료가 상세히 제공되지 않았거나 분석을 수행하는 연구자가 필요한 자료에 접근하지 못한다면 혼합비교를 수행할 수 있는 상황이 제한된다는 것이 언급되었다. 간접 비교를 하는 것이 의사 결정에 있어 유용한 측면이 있으나 혼합비교와 같이 추가적인 노력이 더 많이 필요한 분석을 하는 데 따른 비용-효과성도 고려해야 할 것으로도 보았다.

간접비교 및 혼합비교와 관련된 기술적인 문제에 대해서 참석자들은 간접비교나 혼합비교를 쓰는 분석인 경우 이질성, 일관성, 그리고 무작위배정의 보존에 대해 충분한 주의가 필요하다고 제안했다. 참석자들은 이질성을 임상시험 프로토콜과 인구집단 특성의 차이에 의한 치료효과의 차이로 정의하고 일관성은 간접비교와 직접비교 추정값들 간의 차이로 정의하였다. 많은 참석자들은 혼합비교를 위한 가정들이 직접비교 임상시험들을 이용하여 두 치료법을 비교하는 표준적인 메타분석을 위한 가정들과 차이가 없다고 강조하였다. 그러나 어떤 참석자들은 혼합비교에 필요한 가정들이 표준적인 메타분석의 가정들보다 더 강하다고 믿었다. 또한 참석자들은 서로 다른 시험들에서 한 가지 치료법의 치료효과만 모으는 방식에 대해 우려를 나타냈다. 참석자들은 무작위배정을 깨는 서로 다른 시험들에서 한 가지 치료법의 절대 치료 효과만 모으는 방법은 교란효과와 비뚤림을 피하기 위해서라도 사용되어서는 안 된다고 동의하였다.

간접비교와 혼합비교를 위한 네트워크를 정의하는 방법에 대한 포괄적인 합의는 없었다. 그러나 네트워크를 만들 때 거쳐야 하는 단계에 대한 평가와 방법론적인 지침의 범위에 대한 분명한 규칙을 설명하는 것이 필요하다고 생각하였다. 그러나 다른 연구자들은 네트워크를 너무 넓게 확장하는 것은 근거로부터 잘못된 결론을 유도할 수 있다고 우려 했다. 일부 연구자들은 네트워크에 포함될 치료법을 결정하는데 임상적인 적절성과 임상적인 판단의 중요성을 강조했다.

또한 Briefing paper for methods review workshop on evidence synthesis(indirect and mixed treatment comparison)에서 제시하고 있는 간접비교와 혼합비교 결과 제시 방법에 대한 논의도 있었다.

참석자들은 혼합비교는 많은 임상시험의 결과를 사용하기 때문에 분석결과는 더 강력하고, 치료 효과에 대한 불확실성을 줄일 수 있으며, 분석에 사용하기에 적절한 근거들을 분석에 포함시킨다고 생각한다. 그러나 다른 참석자들은 혼합비교 분석에 비일관성과 이질성이 존재하는 경우, 혼합비교 분석이 반드시 불확실성을 줄이기 위해 필요한 것은 아니라고 주장한다. 또한 일부 참석자들은 임상적으로 이치에 맞지 않을지라도 간접적인 혼합비교 분석은 치료효과의 일관된 추정치를 제시할 수 있다고 지적했다.

3.1.1.2. Guide to the Methods of Technology Appraisal

영국의 의료기술 평가 지침서(Guide to the Methods of Technology Appraisal)에서 간접비교와 혼합비교에 대해 언급하고 있는 부분은 5. Clinical and cost effectiveness and NHS impact에서 indirect and mixed treatment comparisons(5.3.13~5.3.22)이며 그 내용은 다음과 같다.

5.3.13

직접비교 무작위배정임상시험이 있으면, 이를 기본분석으로 제시해야 한다. 직접비교 무작위배정임상시험이 있을 경우 혼합비교 분석은 직접비교 임상시험에서 얻지 못한 추가 정보를 얻기 위해 수행할 수 있다. 이 경우 혼합비교 분석에 대해 충분히 기술해야 하며 기본 분석에 추가로 제시해야 한다. '혼합비교'라 함은 직접비교 무작위배정임상시험과 간접비교 무작위배정임상시험을 포함한 것을 말한다. 만일 평가하고자 하는 다수의 의약품이 하나의 무작위배정임상시험에서 비교되지 않았다면, 각각 쌍으로 비교한 직접비교 임상시험의 결과를 모두 제시해야 한다. 직접비교 임상시험이 충분하지 않은 경우 추가 정보로 혼합비교 분석 결과를 제시할 수 있다. 직접비교 임상시험이 없다면, 간접비교가 사용되어야 한다. '간접비교'는 임상시험의 네트워크(network of trials)에서 나온 자료를 통합하는 것이다. 일반적인 메타 분석의 원칙들을 혼합비교와 간접비교에서도 지켜야 한다.

5.3.14

NICE는 직접비교 임상시험으로부터 나온 결과를 선호하며 가능하다면 이를 기본

분석으로 제시해야 한다.

5.3.15

'간접비교'는 신청약제에 대한 직접비교 무작위배정임상시험이 없으나 간접적으로 다른 치료법과 비교한 무작위배정임상시험의 네트워크(network of trials)가 있을 경우 이들 임상시험 자료들을 통합하는 것이다. '혼합비교'는 직접비교한 임상시험과 이를 간접적으로 비교한 임상시험을 포함한 분석이다. 체계적 문헌 고찰과 메타 분석을 수행하는데 필요한 원칙들을 혼합비교와 간접비교를 수행할 때에도 지켜야 한다. 비교약제와 함께 임상시험 선정기준을 충분히 기술해야 한다. 통합 방법에 대해서도 자세하게 기술해야 한다. 개별 임상시험의 연구방법과 결과도 기재해야 한다. 만일 특정 임상시험이 적절한지 의심스럽다면 이 임상시험을 제외한 민감도 분석 결과를 제시해야 한다. 쌍으로 비교한 결과들의 이질성여부와 직접비교와 간접비교 결과간의 비일관성 여부를 제시해야 한다.

5.3.16

직접비교 임상시험에서 나온 자료가 충분하지 않은 경우 혼합비교에서 나온 추가적인 근거를 고려할 수 있다. 이 때 혼합비교는 기본 분석 결과와 분리하여 별도로 제시해야 하며 혼합비교를 수행한 이유를 기술해야 한다. 또한 앞서 언급한 필요한 원칙들을 따라야 한다.

5.3.17

다수의 치료법들을 평가할 때 각 치료법들을 각각 쌍으로 직접비교 임상시험에서 나온 자료를 제시해야 한다. 이 경우 혼합비교를 사용한 추가 분석을 수행할 수 있다.

Appraisal Committee는 통합 방법에 따른 분석 결과와 혼합비교 포함여부의 적절성을 확인할 것이다.

5.3.18

신청약제에 대한 직접비교 임상시험이 없는 경우가 있다. 이 경우 간접비교를 고

려해야 한다.

5.3.19

간접비교 또는 혼합비교를 사용하여 근거를 통합한 경우 무작위배정이 보존되어야 한다. 발생할 수 있는 비뚤림과 불확실성을 보정하기 위한 적절한 단계를 밟지 않았다면 서로 다른 임상시험에서 가져온 치료법의 결과를 단순히 비교하는 것은 받아들이지 않는다.

5.3.20

평가 과정에서 규정한 비교약제는 아니지만 간접비교 또는 혼합비교 분석의 네트워크상에 나타난 비교약제(예: 위약)들을 포함할 수 있다. 포함 여부에 대한 이유는 명확하게 기술해야 한다. 다시 말해 기본적인 원칙들이 여전히 적용된다.

5.3.21

혼합비교 결과는 다양한 방법으로 제시할 수 있다. 근거의 네트워크(network of evidence)는 표로 제시될 수 있다. 직접비교 결과와 간접비교 결과가 명확하게 구분되고 각 비교에서 사용한 임상시험의 수가 기재된다면 그림으로도 제시할 수 있다.

5.3.22

직접비교 임상시험에 대한 메타 분석 또는 간접비교, 혼합비교에서 충분하고 관련성 있는 타당한 자료가 없을 경우 분석 결과는 개별 연구와 그 결과를 평가하여 정성적인(qualitative overview)형태로 제시할 수 있다. 이 경우 Appraisal Committee는 분석 결과를 검토할 때 특별히 주의할 것이다.

3.1.2. 호주

3.1.2.1. Report of the Indirect Comparisons Working Group

호주의 2007년 보고서는 working group에 의해 2007년 지침서의 간접비교와 관련된 부분에 대한 검토와 고찰에 기반하여 작성되었다. 이 보고서는 신청약제와 비교약제에 대한 직접비교 임상시험 자료가 없는 경우 공통대조군을 이용한 교환가능성 가정과 같이 기본적으로 입증이 불가능한 가정에 의존하고 있는 간접비교의 평가와 그러한 평가결과를 얼마나 신뢰할 수 있을 것인지에 대한 판단을 돕고자 하고 있다.

우선 간접비교에 대한 정의, 보고서에 사용되고 있는 용어 및 가정들에 대한 정의에 대해 제시되어 있다.

교환가능성 가정은 간접비교가 이루어진 모든 임상시험에서 직접비교가 수행되었다 할지라도 간접비교를 통해 합성된 상대치료효과가 동일한 효과크기가 얻어질 수 있어야 함에 기반하고 있다. 이 교환가능성 가정은 입증하기 어렵다. 그러나 결과에 신뢰성을 갖추기 위해 다음과 같은 과정을 따라 평가를 수행한다.

1. 간접비교에 포함 가능한 모든 임상시험들에 대한 확인 및 기술
2. 분석에 선정된 임상시험들에 대한 타당성 제시(이질성 요인 측면, 공통대조군의 결과 측면, 통계적 이질성에 미치는 영향에 대한 측정치 선택 측면)
3. 간접비교 계산 방식에 대한 적절하고 명확한 기술과 신뢰구간 제시
4. 잠재적으로 논란의 여지가 있는 임상시험의 포함여부에 대한 민감도 분석 제시

위의 과정은 체계적 문헌고찰의 설계와 수행방식과 원칙적으로 매우 유사하며 이를 통해 비뚤림을 최소화할 수 있다.

또한 working group은 간접비교의 근거의 수준을 무작위배정 측면, 교란변수 측면, 근거의 질 측면에서 다른 연구설계의 근거 수준과 비교하고 정리하였다. working group은 단순간접비교에 대해 적절한 방법이 아니며 공통대조군 간접비교에 대해

비무작위배정 연구설계보다 선호하는 것은 방법론적인 측면이 아니라 실용적인 측면 때문이며, 관심 치료를 비교하기 위한 간접비교의 결과는 그 이후에 잘 수행된 무작위배정임상시험의 결과와 일치된다는 것을 보여야 한다고 결론을 내리고 있다.

상대치료효과를 계산하는데 어떤 추정치가 나은지에 대한 논의를 다루고 있다. 이분형 자료에 대한 직접비교에서 상대치료효과의 추정치 중 무엇이 좋은지에 대해서는 많은 연구들이 있으며, 모든 추정치는 각각 장점과 단점을 갖고 있다. 즉, odds ratio(OR)의 경우 임상적 평가에서 해석하기 어렵다는 단점이 있으나 Risk Difference(RD)에 대해서 일관된 추정치를 보여주고 이분형 사건발생율이 0과 1사이가 되도록 해준다는 장점이 있다. 따라서 working group은 OR, Relative risk of harm(RRH), Relative risk of benefit(RRB), RD, Number needed to treat(NNT)와 같은 모든 추정치를 사용하여 분석할 것을 제안하였다. 상대치료효과가 underlying 기저 위험과 상관없이 일정하다고 알려져 있는 경우 경제성 평가를 위한 RD를 추정할 때 OR을 사용하는 것이 선호된다. 즉, 간접비교에서 추정된 RD를 직접적으로 경제성 분석에 사용될 수 있다. 만일 상대치료효과의 추정치에 따라 결과가 달라질 경우 결과를 제출할 때 다른 추정치에 따른 결과의 범위를 제시해야 하며, 만일 특정 추정치를 사용한다면 이에 대한 근거를 명확히 제시해야 한다.

마지막으로 네트워크 분석과 혼합분석에 대한 working group의 입장이 정리되어 있다. working group은 네트워크 분석과 혼합분석은 보다 복잡한 통계적 방법들이 필요하고 따라서 단순한 간접비교와 관련된 기본적인 문제들의 영향을 독립적으로 평가하기가 더 어렵다는 점에서 두 방법론에 대해 부정적인 입장이다. 즉, 네트워크 분석과 혼합분석이 모든 적합한 근거들을 미리 결정된 분석들로 포함할 수 있어 분석 비뚤림을 줄일 가능성이 있다고 주장되지만 분석에 포함시키는데 영향을 줄 수 있는 출판 비뚤림이나 선택 비뚤림과 같은 비뚤림을 발견하고 설명하는데 어려움이 있고 이런 어려움은 간단한 간접비교와 전통적인 메타 분석을 비교할 때 네트워크 분석과 혼합분석에서 더 커질 것을 의미한다.

3.1.2.2. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee

2008년에 발표한 지침서 version 4.3은 1995년 이래 처음으로 전체 개정을 하였다. 이 지침서는 주 적응증에 대한 임상적 유용성 평가와 경제성 평가를 위해 요구되는 사항들에 대한 지침을 제시하고 있다.

간접비교와 관련된 부분은 Part3의 section B(i)Clinical evaluation for main indication: Presenting an indirect comparison of randomised trials이며, 이 부분에는 문헌 검색 전략, 검색된 문헌에 대한 선정/제외 기준, 질 평가, 간접비교에 사용된 개별 무작위 배정 임상시험들의 상대치료효과와 간접비교에 의한 상대치료효과의 추정치 제시, 약물의 부작용이나 의존성 제시, 간접비교 결과 해석에 대한 수행 방법이 제시되어 있으며, 각 과정에 대한 자료 제출 방식에 대해서도 언급하고 있다.

3.1.3. 캐나다

3.1.3.1. Indirect Evidence: Indirect Treatment Comparisons in Meta Analysis

2009년에 발표된 이 보고서는 다른 나라와는 달리 간접비교를 위한 소프트웨어 소개와 그 사용 및 예제를 포함하고 있는 기술 보고서의 성격을 띠고 있다.

우선 간접비교 방법론으로 Bucher-adjusted indirect comparison method, Lumley's network meta-analysis approach, and mixed treatment comparisons method에 대한 소개 및 비교를 하고 있다. 그 후 Bucher Method에 의한 간접비교의 확장으로 공통대조군이 여러 개 있는 간접비교에 대해 소개하고 측정치로 사용될 수 있는 Odds Ratio(OR), Relative Risk(RR), Hazard Ratio(HR), Risk Difference(RD)와 Mean Difference(MD)의 정의와 simulation을 통한 bias 및 MSE(mean square error)의 크기를 제시하였다. 간접비교를 위해 Visual Basic을 통해 개발된 ITC program에 대한 소개 및 사용방법과 예제가 있으며, 다양한 간접비교 방법(Bucher Method, Network Meta-analysis approach, 혼합비교)적용이 WinBUGS 코드와 함께 예제를

통해 설명되어있다.

3.1.3.2. Guidelines for the Economic Evaluation for the Health Technologies

고찰한 지침서는 CADTH(구 CCOHTA, Canadian Coordinating Office for Health Technology Assessment)에서 2006년에 발표한 1994년의 경제성 평가 지침서의 3 번째 개정판이다. 지침서는 경제성 평가에 대한 소개와 관련 용어에 대한 정의 및 경제성 평가 절차, 검토항목, 검토결과 보고 등에 대해 소개하고 있다. 간접비교와 관련된 부분은 비교대상 선정과 관련된 3.4 Comparators의 3.4.1 General consideration와 효과자료 선택과 관련된 3.6 Effectiveness의 3.6.2 Gathering best evidence의 일부분에 언급되고 있으며 구체적인 내용은 다음과 같다.

3.4 Comparators의 3.4.1 General considerations 중에서

만일 신청약제와 비교약제의 직접비교 무작위배정임상시험이 없는 경우 적절한 분석 방법에 근거를 둔 간접비교를 사용할 수 있다. 사용된 간접비교 방법론 및 사용된 간접비교 방법론의 한계, 잠재적인 비뚤림, 결과 해석에 있어서 문제점 등을 기술해야 한다. 비교대상에 대한 가정이 어떤 영향을 주는지 확인하기 위해 민감도 분석을 수행할 수 있다.

3.6 Effectiveness 중에서

3.6.2 Gathering best evidence

잘 수행된 임상적 검토가 평가의 기초가 된다. 여기에는 신청약제의 효능과 효과에 대한 가능한 근거들에 대한 체계적 문헌고찰도 포함된다. 신청약제의 효능과 효과에 대한 체계적 문헌고찰을 하지 못할 경우 이유를 기술해야 한다.

체계적 문헌고찰에 포함된 연구와 문헌고찰을 수행한 방법에 대하여 기술해야 한다. 문헌고찰은 다른 수준의 내적/외적 타당성을 반영하여 다양한 연구 설계를 포함할 수 있으며, 선정된 연구들에 대한 질을 파악할 수 있는 정보가 보고되어야 한다. 질이 낮은 연구(예를 들어 결측값이 많고 대상환자에 대한 자세한 설명이

없는 표본 크기가 작은 연구)는 높은 근거를 제공하지 않을 것이다. 각 형태의 근거에서 제공하는 보완적인 정보에 대해서는 민감도 분석을 통해 분석해 볼 수 있다. 자료를 분석하거나 통합하는 방법(메타 분석 또는 간접비교)을 기술하고 사용 근거를 제시해야 한다. 또한 이에 대한 분석 결과를 제시해야 한다.

3.1.4. ISPOR

ISPOR(International Society for Pharmacoeconomics and Outcomes Research)는 2009년 1월 보건의로 연구자들과 보건의로 관련 의사 결정자들에게 간접비교 지침을 제공하기 위한 Task Force를 구성하였다. 2010년 5월 미국 애틀랜타에서 진행된 15th Annual International Meeting에서 발표한 초안에 따르면 연구진들은 간접비교 자료를 이용하여 의사 결정을 하게 되는 사용자들을 위한 지침과 간접비교를 직접 수행하는 평가자들을 위한 지침을 별도로 작성하고 있다. 이들 지침에는 간접비교에서 필요한 가정과 통계적 기법들을 소개하고 이를 해석하는 방법, 그리고 수행된 간접비교의 질과 적절성을 확인할 수 있는 체크리스트를 담을 계획이라고 밝혔다. 최종 결과물은 2010년 11월에 발간될 Value in Health에 보고될 예정이다.

3.1.5. 소결

영국의 National Institute for Health and Clinical Excellence(NICE)는 1998년 공적 의료기술 평가와 의료보험 합리화를 목적으로 설립된 기관이다. 이를 중심으로 한 'Guide to the methods of technology appraisal'는 의료기술평가를 뒷받침해온 평가지침서로서 2001년 초에 초판이 나온 이후, 여러 차례의 개증보를 거쳐 2008년에 다시 개정을 이루었다.

2008년 개정판에 수록되어있는 간접비교 수행에 관한 지침은, 그 이전 판까지의 지침을 재검토하고 혼합비교 등을 포함한 간접비교 관련 최신 방법론의 도입 여지와 사용의 타당성에 관한 제안을 마련하기 위하여, 관련 전문가 그룹으로 구성된 working group이 2007년 초에 결성되고 2년 반이 넘는 시간에 걸친 토론과 작업을 통해 NICE에 제출한 briefing paper와 간담회를 통해 모아진 의견에 기초하고 있다.

복잡한 가정과 모델에 기반한 혼합비교 등과 같은 새로운 방법론에 대해서, 해당 방법론 개발자 및 연구자들이 대거 참여하여 조언하고 있는 만큼 영국의 지침서에서의 새로운 방법론의 차용에 대한 입장은 호주 등 다른 국가의 기관 지침서 보다 덜 보수적인 경향을 띤다.

호주에서도 영국과 유사한 간접비교 working group이 2007 년 후반에 Economics Sub-Committee(ESC), subcommittee of the Pharmaceutical Benefits Advisory Committee(PBAC) 및 호주 의약품 규제기관인 Therapeutic Goods Administration(TGA)후원 과 참여 하에 추진되어 보고서 작성 작업을 수행하였다.

최신 개발된 방법론 등을 포함하는 기존의 연구문헌들과 영국과 캐나다의 관련 기관 지침서에 대한 검토와 고찰에 기반 하여 작성되고 Pharmaceutical Benefits Advisory Committee에 제출된 제안서에 기초하여 호주의 Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee 2008년 12월 최신판의 간접비교 관련 부분이 개정되었다.

호주의 working group의 경우, 간접비교 방법론이 관찰적 연구에 우위하여 갖는 장점이나, 입증되기 어렵고 불명확한 많은 가정에 기반하는 최신 간접비교 관련 방법론을 통해 얻어진 결과가 과연 직접비교 임상시험을 통해 얻어져야 할 결과를 대체할 수 있을 것인가에 대한 이견이 많았으며, 이러한 보다 보수적인 입장과 견해가 지침서에 보다 구체적으로 반영되었다.

캐나다 Canadian Agency for Drugs and Technologies in Health(CADTH)의 Guidelines for the economic evaluation of health technologies 경우 아직 2006년 판 이후로 개정이 없었으며, 간접비교 관련 부분은 매우 간단히 언급되어 있으며 구체적인 수행 지침을 담고 있지는 않다.

본 보고서에서 제안하는 지침서는 호주(Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee)의 기본틀을 참고([부록2] PBAC과 교신 자료 참조)하여 구성하고 영국 및 그 외 지침서 및 보고서의 입장 및 내용도 반영하여 개발되었다.

3.2. 간접비교 수행에 대한 이론적 고찰

3.2.1. 정의 및 용어

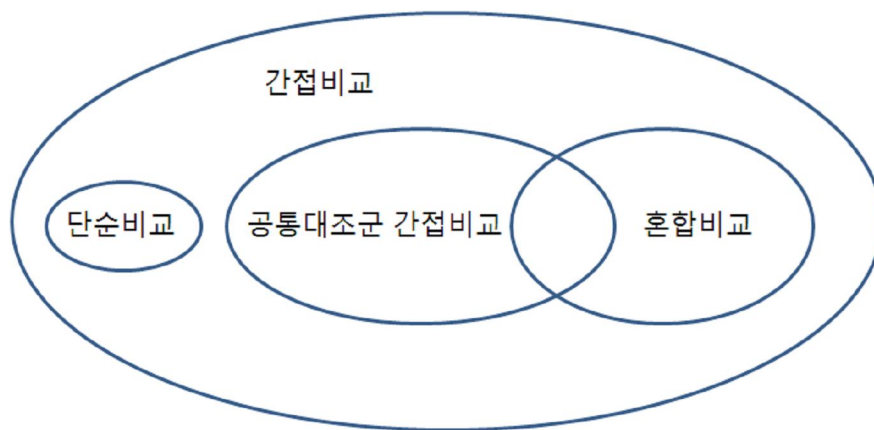


그림 3-1 간접비교의 분류

3.2.1.1. 단순비교

단순비교란 무작위배정임상시험들에서 보고자 하는 하나의 치료군의 결과만을 추출하여 메타분석 등으로 합쳐서 결과 값을 도출하고 같은 방법으로 추출한 다른 치료군의 결과 값과 단순히 서로 비교하는 방법이다 (그림3-2). 단순비교는 통상적으

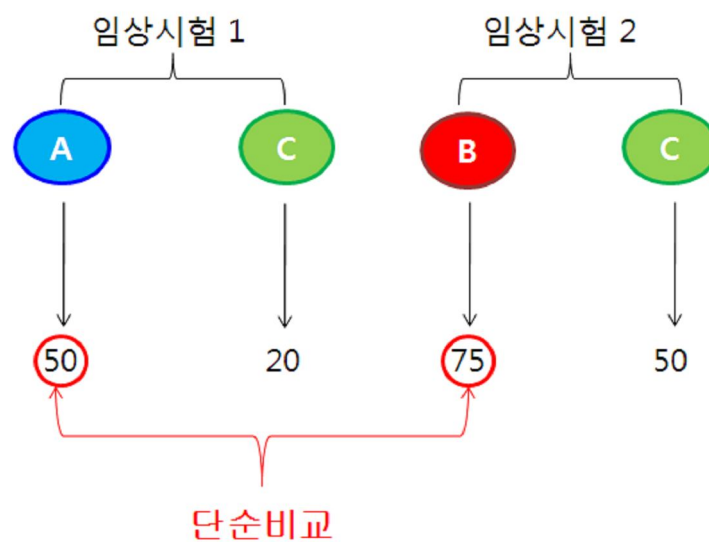


그림 3-2 단순비교

로 [naive indirect comparison]으로도 불리며, 간접적으로 군 간에 비교를 이루는 것이므로 이 또한 간접비교의 방법에 포함될 수는 있다. 그러나 본 지침서에서는 단순히 치료군 간 결과만을 비교하는 것 이상의 보정 또는 조정을 거치지 않은 단순비교는 수용 가능한 간접비교와 구별하여 사용할 것이다.

3.2.1.2. 공통대조군 간접비교

Bucher et al.(1997)에서 제안된 방법으로 약제 A와 약제 C를 직접비교한 무작위배정임상시험과 약제 B와 약제 C를 직접비교한 무작위배정임상시험이 있을 경우 약제 C를 공통대조군으로 하여 약제 A와 약제 B를 간접적으로 비교하는 방법이다. 해외에서 [adjusted indirect comparison]으로 불리기도 하나 [보정된 간접비교]라는 말은 보정될 것이 없다는 뜻으로 오인될 수 있으므로 [공통대조군 간접비교]라는 표현이 더 적절할 것으로 판단된다. 본 보고서에서는 공통대조군을 통한 간접비교를 통상적인 간접비교의 범주로 간주하며 이를 줄여 간접비교로 지칭하되 단순비교와 구분할 필요가 있을 때는 [공통대조군 간접비교]로 부르기로 한다.

공통대조군을 통한 간접비교에는 하나의 공통대조군(약제 C)을 통하여 약제 A와 B를 비교하는 가장 단순한 상황에서의 공통대조군 간접비교 (a), 하나의 공통대조군을 기준으로 비교대상 약제가 여러 개 존재하는 형태 (b), 비교하고자 하는 약물 간에 비교가 여러 단계의 공통 대조군을 거쳐서 이루어질 수 있는 형태 (c), 그리고 비교하고자 하는 약물 간에 공통대조군이 여러 개 존재 하는 형태 (d) 등으로 구분될 수 있다. (그림 3-3)

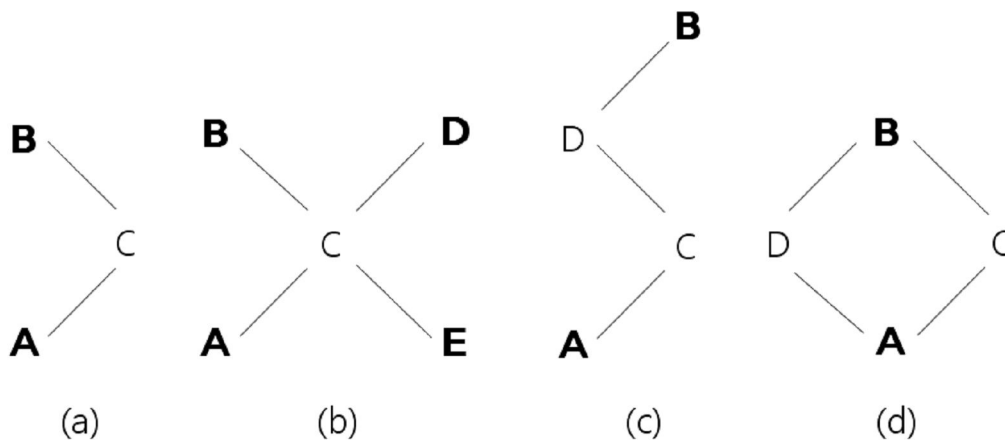


그림 3-3 공통대조군 간접비교의 여러 형태

3.2.1.3. 혼합비교

해외문헌들에서 'mixed treatment comparison'으로 불리는 혼합비교의 구체적인 정의는 문헌들 간에 사용자들 간에 다소 차이가 있으나 본 지침서에서는 다음의 경우를 모두 혼합비교로 분류하고자 한다.

1) 약제 B(비교약제)에 대한 약제 A(신청약제)의 효과를 알고자 할 때, 약제 A와 약제 B를 직접비교한 임상시험이 있고 약제 A와 약제 C, 약제 B와 약제 C를 비교한 임상시험들이 있어 약제 A와 약제 B의 직접비교와 약제 C를 통한 약제 A와 약제 B의 간접비교를 결합하여 두 약제의 치료효과를 비교하는 경우 (그림 3-4)

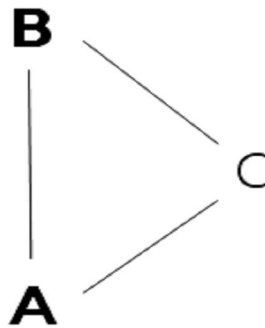


그림 3-4혼합비교의 예

2) 신청약제 A에 대하여 여러 개의 비교약제(B, C, D)가 있고 이들이 직접, 간접으로 연결되어 있는 경우 (그림3-5)

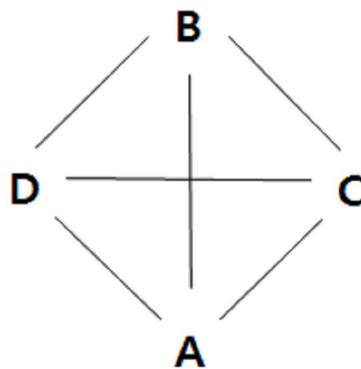


그림 3-5 신청약제 A에 대하여 비교약제 B, C, D가 있는 경우

3.2.1.4. 공통대조군

약제 A와 약제 B를 비교하고자 할 때 약제 A와 약제 C를 직접비교한 임상시험과 약제 B와 약제 C를 직접비교한 임상시험이 있을 경우 약제 C를 공통대조군이라 한다. 문헌에 따라 control group, placebo group(Bucher et al., 1997), common comparator, linking treatment(CADTH, 2009)로 사용되기도 한다.

3.2.1.5. 동질성, , 일관성

동질성은 임상적 동질성(clinical homogeneity)과 통계적 동질성(statistical homogeneity)으로 구분된다. 임상적 동질성이란 연구들 간 환자 선정/제외기준, 환자들의 기저위험, 연령 및 성별, 진단기준, 용량 또는 약제 복용 스케줄, 처리의 종류, 결과변수에 대한 정의, 추적기간, 연구 지역 및 연구 상황 등이 동질한 경우를 말한다. 통계적 동질성이란 각 연구결과에 관한 요약자료(처리효과에 대한 측정값 및 해당 신뢰구간)의 크기들이 통계적으로 차이가 없는 것을 의미한다.

간접비교에서도 전통적인 메타분석에서 요구하는 동질성(homogeneity)이 성립해야 한다. 약제 A와 약제 C를 비교한 임상시험들이 메타분석으로 합할 수 있을 정도로 동질하고 약제 B와 약제 C를 비교한 임상시험들도 동질하여 메타분석으로 합할 수 있어야 한다(Song et al., 2009).

약제 C를 대조군으로 하는 임상시험에서, 약제 A와 약제 C를 비교한 무작위배정임상시험과 약제 B와 약제 C를 비교한 무작위배정임상시험의 약제 A와 약제 B를 서로 바꾸어 임상시험이 이루어졌어도 기존의 A vs C 와 B vs C 임상시험에서 얻었던 것과 각각 동일한 결과가 나올 것으로 기대하는 경우 교환가능성(exchangeability)이 성립한다고 말한다. 이를 Song et al.(2009)은 similarity(유사성)라는 표현으로 사용하기도 하였다.

일관성이란, 혼합비교에서 직접비교에 의한 상대치료효과와 간접비교에 의한 상대치료효과가 동일한 경우를 의미한다.

보다 자세한 설명은 3.2.2장에서 다루었다.

3.2.1.6. 상대치료효과

약제 A와 약제 C를 비교한 무작위배정임상시험과 약제 B와 약제 C를 비교한 무작위배정임상시험에서 공통대조군 C의 결과를 보정하여 간접비교를 통해 산출된 약제 B(비교약제)에 대한 약제 A(신청약제)의 효과를 상대치료효과(comparative treatment effect)로 부르기로 한다.

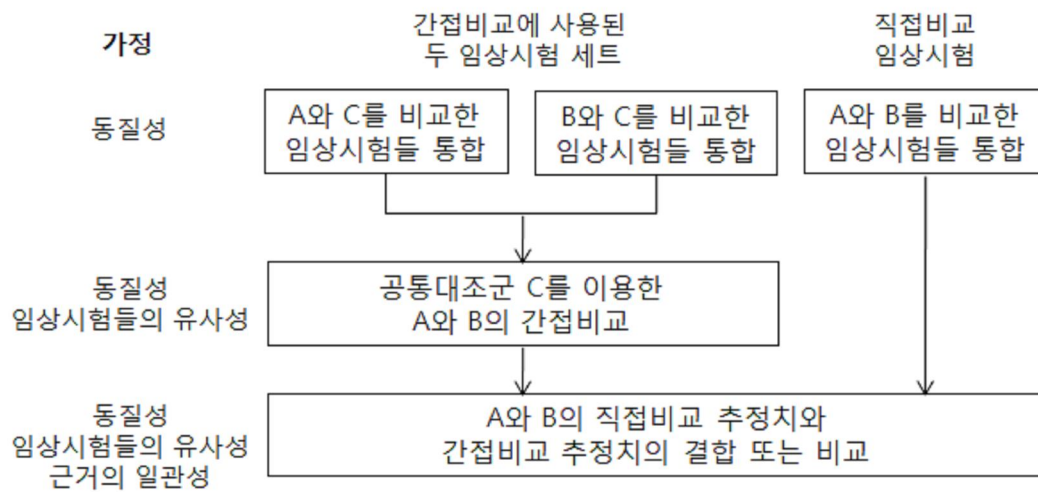


그림 3-6 간접비교와 혼합비교의 가정

자료원: Song et al.(2009)

3.2.2. 이론적 배경 및 전제조건

3.2.2.1. 이론적 배경

관심약제 A와 이에 대한 비교약제인 B를 직접 비교하여 수행된 무작위배정 비교임상시험이 아닌 제 3의 표준치료제 C와 각 치료약제 A, B가 비교 평가된 기존의 문헌결과를 통해서 약제 A와 약제 B간 치료효과차이의 추정치를 간접적으로 얻고자 간접비교를 아래와 같이 수행한다. 이때, 치료효과에 영향을 미칠 수 있는 다른 요인은 통제되었다고 가정한다.

관심약제로서 A를 표준치료제인 약제 C와 비교하여 약의 효과에 대해 평가하고자 수행된 대규모 무작위배정임상시험 연구와 약제 B를 표준치료제인 약제 C와 비교하여 약의 효과에 대해 평가하고자 수행된 대규모 무작위배정임상시험이 기존의 문

현으로 존재한다고 하자. 첫 번째 연구가 수행되었던 연구대상을 모집단1 이라고 하고 두 번째 연구가 수행되었던 연구대상은 모집단2 라고 하며, 이에 대한 효과측정치로서 부정적인 결과변수에 대한 위험률 R이 각 연구로부터 평가가 이루어진 상황이라고 할 때, 모집단1 로부터 약제 A와 B간의 치료효과차이로서 위험차(risk difference, RD) $\Delta_{AC}^1 = R_A^1 - R_C^1$ 가 추정될 수 있고, 모집단2 로부터 약제 B와 C간의 위험차가 $\Delta_{BC}^2 = R_B^2 - R_C^2$ 로 추정될 수 있다. 이는 상대위험도(relative risk, RR)이나 교차비(odds ratio, OR)에 대해서도 유사하게 표현될 수 있다.

위에서 언급한 바와 같이 각 모집단1, 2로부터 얻어진 각각의 효과크기 $\Delta_{AC}^1, \Delta_{BC}^2$ 를 비교하여 약제 A와 약제 B의 치료효과의 차이를 $\Delta_{BC}^2 - \Delta_{AC}^1$ 로서 측정한다. 이때 약제 A와 약제 B의 효과차이에 대한 타당한 측정치를 얻기 위해서는 연구1의 연구대상과 연구2의 연구대상에서 공통대조군 C에 대한 각 약제의 효과차이는 같아야 한다. 즉, $\Delta_{BC}^1 = \Delta_{BC}^2, \Delta_{AC}^1 = \Delta_{AC}^2$ 이 전제조건으로서 가정된다. 이를 교환가능성(exchangeability) 가정이라고 한다.

이러한 전제조건이 만족될 경우, $\Delta_{BC}^2 - \Delta_{AC}^1 = \Delta_{BA}^2 = \Delta_{BA}^1$ 이 될 수 있고, 이를 통해 약제 A와 약제 B의 치료효과 차이를 비교할 수 있다고 입증할 수 있게 된다.

직접비교에서 추정된 치료 효과(B vs C, A vs C)인 Δ_{BC}, Δ_{AC} 를 이용하여 간접적으로 비교하여 추정한 치료 효과(B vs A)로서 $\Delta_{BA} = \Delta_{BC} - \Delta_{AC}$ 를 추정하였을 때, 이의 분산은 $\text{Var}(\Delta_{BA}) = \text{Var}(\Delta_{BC}) + \text{Var}(\Delta_{AC})$ 와 같이 얻어지게 되고, 이를 통한 신뢰구간의 크기는 각각의 직접비교에서보다 넓어지는 결과가 된다. 이는 검정에 대한 보수적인 접근으로 이어질 것이나, 간접비교를 통한 불확실성의 증가를 감안한다면 이와 같은 보수적인 접근이 타당하다고 볼 수 있다.

3.2.2.2. 간접비교 전제조건으로의 교환가능성

간접비교에 대한 전제조건으로 앞서 제시한 교환가능성(exchangeability)은 간접비교에서 도출되는 상대치료효과 측정치가 타당성을 갖기 위해 만족되어야 하는 가장 핵심적인 가정이다. 문제는 각 약제 A와 약제 B에 대한 공통대조군과의 직접비교 임상시험에서의 각 대상 환자집단 1과 2에 약제 B와 약제 A를 교환하여 각각 적용한 후 결과를 얻어 확인해보지 않고서는 교환가능성은 입증하기 어려운 가정이라는

점이다.

간접비교는 기저 위험이 달라도 상대적인 치료효과는 동일하다는 가정하에 수행되는 경우가 많다. 즉, 공통대조군 기저위험에 대한 보정으로써 기저위험의 차이를 통제한 것으로 보는 것이다.

그러나, 이는 입증되기도 어렵고 교환가능성 가정에 영향을 줄 수 있는 요인들을 모두 적절히 감안했다고 보기도 어렵다. 기저 위험에 영향을 줄 수 있는 요인외의 다른 요인들에 대해서도 임상시험은 다르게 수행되었을 수 있고 이러한 요인들이 합성된 치료효과크기에 영향을 줄 수 있기 때문이다.

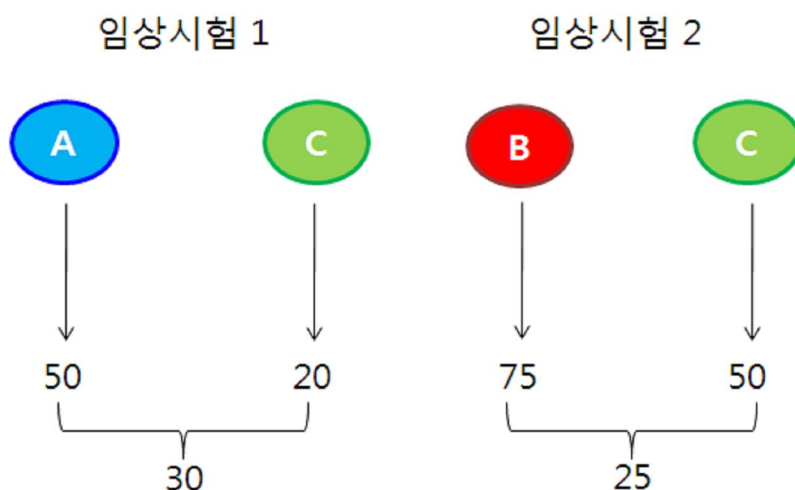


그림 3-7 두 개의 임상시험으로부터 얻어진 간접비교 자료

예를 들어, 약제 A와 B 그리고 공통대조군인 약제 C에 의한 질병의 호전정도를 측정하고 이에 대한 결과를 두 개의 임상시험으로부터 그림 3-7과 같이 얻었다고 할 때, 약제 A와 약제 B의 효과를 비교하기 위해 임상시험1에서 관찰된 A에 의한 호전정도 50과 임상시험2에서 얻어진 B에 의한 호전정도에 대한 결과 75를 바로 비교하여 약제 B의 효과가 A에 비해 25(75-50) 차이만큼 더 좋았다고 단순 결론을 내리는 것은 오류이다. 이것이 단순비교가 가진 문제이다.

위 그림에서 제시된 각 연구에서의 공통대조군 C의 결과를 보면, 연구 1에서는 20이었고 연구 2에서는 50임을 볼 때, 같은 약제 대해서 20과 50으로 호전정도에 대하여 다른 결과가 나타났다는 것은 두 연구에서 환자의 상태나 치료환경 등의 조건

이 달랐음을 의미할 수 있으므로, 이들로부터 얻어진 각각의 결과를 바로 비교하기는 어려움을 알 수 있다. 따라서, 이때 공통대조군 C에 대하여 얻어진 약제 A와 약제 B의 각 상대효과들의 비교를 통해 약제 A의 B에 대한 상대효과를 합성해내고자 하는 것이 간접비교의 개념이다.

다시 말하면, 공통대조군의 상태가 20과 50으로 달랐다는 것을 고려하여 감안할 때, 약제 A가 C에 비하여 30 만큼 개선되었고, 약제 B는 C에 비하여 25만큼 개선되었으므로, C에 대한 상대적인 개선정도로써 A와 B 두 약제를 비교해보면 A 약제가 B 약제보다 C에 대한 상대효과가 5만큼 더 좋았으므로 약제 A가 B보다 효과가 더 좋다고 보는 것이다. 즉, 공통대조군 C를 이용하여 연구 1에서 A의 치료효과가 추정되는 기저상태와 연구 2에서 B의 효과가 추정되는 기저상태가 다름에 대한 *i*일종의 보정을 행하는 것이다. 따라서 이와 같은 간접비교를 *i*보정된 간접비교(*adjusted indirect comparison*)*i* 라고도 부른다.

그러나, 여기서 주목할 것은 공통대조군에서의 결과가 50으로 나타나는 다른 환자군을(치료반응이 더 좋거나 기저상태가 더 좋았거나 등) 대상으로 약제 A의 상대적인 개선을 측정했을 때도 C에 대한 결과가 20이었던 환자를 대상으로 측정했을 때와 마찬가지로 C에 대해서 30만큼 더 좋은 것으로 나타날 것인지, C에 의한 질병의 호전을 20으로 보이는 환자군을 대상으로 약제 B를 평가했을 때도 C에 의한 결과에 대한 상대효과크기가 25차이만큼 더 좋은 것으로 나타날 것인지 확신하기 어렵다. 그렇기 때문에, 이와 같은 상황에서 각 공통대조군에 대한 상대치료효과크기인 30과 25을 비교하여 약제 A가 B에 비해 효과가 더 좋다고 판단하는 것에도 오류가 있을 수 있으며 사실상 보정이 이루어진 것이라고 보기도 어렵다는 것이 교환가능성이 보장되어야 하는 핵심적인 이유이다.

교환가능성이란 공통대조군 C에 대한 약제 A의 효과가 연구 1에서 30의 차이로 나타났다면 연구 2에서도 똑같이 30의 차이로 나타날 수 있음에 관한 것이다(약제 B의 공통대조군 C에 대한 효과에서도 마찬가지). 이것이 성립하기 위해서는 1차적으로 각 모집단이 매우 유사하고 모든 요소에서 동질해야 할 것이고, 이를 1차적으로 확인하기 위한 방법으로 공통대조군 C에 대한 결과가 연구 1, 2 모두에서 유사한지를 확인하는 것이 중요하다.

그러나, 위와 같이 공통대조군에서 다른 결과를 보인 두 개의 연구를 간접비교하는 것이 반드시 불가능한 것은 아니다. 공통대조군에 대한 호전을 20으로 보이는 환자를 대상으로 약제 A의 치료결과를 측정했을 때의 A에 의한 호전이 50으로서 대조군에 비해 상대치료효과가 30차이 만큼 더 좋은 것으로 나타났다면, 같은 치료법 C에 대한 호전 효과가 20이 아니라 50으로 나타나는 환자군에 대해서 약제 A를 평가했을 때는 호전 효과크기가 80일 것이라고 가정할 수 있다면, 공통대조군 C에 대한 효과차이를 이용한 상대효과를 도출하고 이를 합성하여 간접비교를 수행할 때의 i보정은 이론적으로는 유효하다고 볼 수 있을 것이다. 그러나 이러한 가정은 사실상 막연하며 입증하기 어렵거나 불가능한 가정이므로 가정에 대한 분명한 근거가 없다면 간접비교의 수행 또는 결과해석은 부적절할 수 있다.

만일, 공통대조군에 대한 치료결과가 50으로 크게 나타나는 치료반응이 좋거나 상태가 다른 환자집단에서는 약제 A에 의한 결과도 80보다는 90이나 100으로 훨씬 좋아질 수 있다면, 각 연구로부터 얻어진 상대효과크기 30과 25를 가지고 A가 B보다 좋다고 내린 결론은 오류일 수 있는 것이다.

따라서 교환가능성에 대한 전제조건으로서 1차적으로는 공통대조군의 결과를 연구간 비교하여 충분히 유사한지를 확인하고, 유사하지 않을 경우에는 1. 교환가능성이 성립하기 어려울 것으로 보고 간접비교를 수행하지 않거나, 2. 공통대조군 효과가 다른 경우에도 사용하고자 하는 평가지표를 이용한 상대치료효과크기는 동일하다고 할 수 있는 근거가 있다면 이를 제시하고 간접비교를 고려할 수 있을 것이다. 이에 대해서는 3.2.4 절을 참조할 수 있다.

결과값이 이분형으로 측정되는 자료의 경우, 앞서 설명된 공통대조군에서의 결과값은 공통대조군에서의 사건발생률을 의미하며, 이를 비교하는 것으로 각 연구에 포함된 환자집단간의 기저위험을 비교할 수 있고, 기저위험의 유사성은 곧 대상 환자집단의 유사성과 관련이 깊으므로 이를 통해 교환가능성을 평가하는 것과도 관련이 깊다. 그러나, 연속형으로 측정되는 결과변수의 경우에는 공통대조군의 결과 비교를 통해 대상환자군간 기저위험의 차이를 평가하고 이를 통한 연구의 유사성에 대하여 평가하는 것이 이분형 결과변수의 경우처럼 바로 연관되지 않으므로, 이때의 교환가능성에 대한 평가는 더욱 어려워진다. 따라서, 현재까지의 간접비교의 수행에 따

른 여러 문제점들에 대한 해결방안에 대한 방법론적인 개발 또는 간접비교를 통한 결과의 수용은 이분형으로 표현되는 결과변수에 한하여 제안되는 경우가 많다.

교환가능성에 대한 가정은 입증하기 어려우나, 결과에 확신을 갖기 위해서는 다음과 같은 과정을 따라 평가를 수행하기를 권장한다:

1. 수행하고자 하는 특정 간접비교에 잠재적으로 관련된 모든 무작위배정임상시험을 확인하여 적절히 기술하고,
2. 확인된 임상시험 중 간접비교에 포함될 임상시험의 선정이 합당한지 기술한다.

이러한 면에 있어 간접비교의 설계와 수행은 체계적 문헌고찰의 설계와 수행방식과 원칙적으로 매우 유사하며, 이를 통해 비뚤림을 최소화하고자 하여야 한다. 이에 관한 구체적인 세부절차는 4.2 및 4.3 장을 참조하여 수행하고 제시할 수 있다.

3.2.2.3. 공통대조군 결과와 전제조건에 대한 확인

이분형 자료로 표현되는 결과변수의 경우에는 공통대조군의 결과값이 기저위험에 대한 측정치로 고려 될 수 있으므로, 본 절에서는 공통대조군 결과값으로 서로 다른 연구의 연구대상 간 기저위험의 차이를 평가하고 이를 통해 간접비교의 전제조건을 확인하는 과정을 기술하였다.

● 2개의 무작위배정임상시험만 존재할 때

- 1) 공통대조군에서의 결과값이 임상시험 간 같을 때

이분형 자료의 경우, 공통대조군에서의 결과값, 즉 사건발생률이 유사하고, 임상시험 수행 방법이 유사(유사한 선정제외기준 사용, 추적관찰기간이 유사 등)할 때, 기저위험이 유사하다고 결론짓고 간접비교를 통해 상대치료효과의 합성을 시도할 수 있다.

이때 공통대조군을 통해 보정된 결과는 사실상 단순비교를 통해 얻어진 결과나 마찬가지이다. 공통대조군의 사건발생률은 단순히 상쇄되어져 없어지고 각 시험군에서의 사건발생률 간 비교를 하는 결과가 되기 때문이다. 그러나 이때 점추정치는

갈더라도 신뢰구간의 크기는 단순비교에서 더 좁게 산출되기 때문에 단순비교를 수행하는 것은 여전히 적절하지 않다.

두 개의 다른 임상시험의 환자군들간 기저 위험에 영향을 줄 수 있는 것 중에 하나로 표준 치료의 발전을 생각할 수 있다. 전통적으로 쓰이고 있는 약제 B가 개발되고 평가되었던 당시의 위약과 약제 B를 비교한 임상시험, 새로운 약제 A와 위약을 비교한 임상시험이 각각 존재하고, 이들을 통해 약제 A와 B를 간접비교하고자 한다면, 약제 B의 치료효과는 B에 대한 평가를 A vs C가 수행되었던 임상시험에서 수행했어도 동일하게 얻어지며 약제 A에 대해서도 마찬가지로 가정할 수 있어야 한다.

일반적으로는 두 임상시험의 위약군에서 같은 사건발생률이 보고되었을 경우 교환가능성 가정이 만족한다고 보게 될 것이나, 표준 치료의 발전이 환자의 기저위험에 영향을 미칠 수 있음에도 이것이 위약군에서의 사건발생률에는 반영되지 않아 두 위약군에서 사건발생률이 같게 나타날 수 있다. 또한 표준치료의 발전으로 인한 기저위험의 변화는 각 시험의 선정기준에도 반영되지 않을 수 있다.

예를 들면, 전통적인 치료 약제에 대한 오래된 임상시험보다 새로운 약제에 대한 임상시험에 정도가 더 심한 환자들이 모집되었을 수 있다. 왜냐하면 전통적인 치료에 반응하지 않은 환자들을 새로운 시험에서 모집하게 될 가능성이 높기 때문이다. 기저병력이나 연령 등과 같은 환자의 특성, 위약군에서의 사건발생률 등이 두 임상시험의 환자군에서 비슷하게 나타남으로써 질병의 정도가 비슷하게 보임에도 불구하고, 새로운 약물에 대한 임상시험에서의 환자는 기본적인 약물들을 더 많이 복용하고 있는 등 표준 치료를 달리 받고 있을 가능성이 크다. 따라서 새로운 약물의 치료효과가 전통적인 약제와 비슷하게 나타났다고 할지라도 이는 두 시험에 포함된 환자들의 질병의 중증도가 다른 것에 따른 교란효과의 결과일 가능성이 있다. 이는 공통대조군에서의 사건발생률이 같았을 때 오히려 기저위험이 다를 수 있음을 보여준다.

따라서 공통대조군의 결과가 두 임상시험 간에 비슷하다고 해도, 동반되는 치료법 등과 같이 상대치료효과가 달라질 수 있는 원인이 될 만한 요인들에 대해서 주의 깊게 비교해보는 것이 매우 중요하다. 두 개의 임상시험이 수행된 시기나 지리적

요건이 다를 때 표준 치료의 차이가 교란요인으로 환자군의 특성에 영향을 줄 수 있으므로 특히 주의가 필요하며, 이 경우 이에 대한 비교가능성에 대한 정당화가 제시되어야 한다.

2) 공통대조군에서의 결과값(사건발생률)이 임상시험 간 다를 때

공통대조군의 사건발생률이 다를 수 있는 경우는 다음과 같다. 그러나 이를 명확히 구분하기는 어렵고 모두 관련되어 있을 가능성 있다.

a) 우연히 발생

b) 중요한 교란요인에 대해 환자군이 다를 때(예. 기저 위험이 다른 경우)

c) 임상시험의 여타 측면에서의 차이가 있을 때(예. 임상시험 수행 방법이나 수행의 질 등이 다른 경우)

위약대조군 무작위배정임상시험이 서로 다른 시기에 수행된 경우 환자의 기저위험이 달라질 수 있다. 동반되는 치료가 임상시험이 수행된 시기에 따라 발전했을 경우, 실제로 발전된 동반치료법이 환자의 기저위험을 다르게 함으로써 위약군에서의 반응률이 변할 수가 있다. 앞의 예에서와는 반대로, 늦게 수행된 임상시험의 환자가 더 낮은 기저위험을 갖게 되고 이는 위약군의 반응률에서의 차이로 나타나게 될 수 있다. 이 경우 기저위험의 차이에 따른 교란효과의 결과로 새로운 약물의 효과가 전통적인 약물과 다른 것처럼 보일 수 있다.

따라서 간접비교를 수행할 때 상대치료효과의 이질성에 영향을 미칠 수 있는 요인들(표 3-1 참조)이 임상적으로 같은지 확인해야 하며 임상시험이 다른 시기에 수행된 경우 특히 주의가 필요하며, 수행시기가 다름이 치료결과에 영향이 없음에 대한 설명 등이 제시되어야 한다.

두 임상시험에서 다른 것이 기저위험 뿐이라면, 기저위험의 차이에 상관없이 상대 치료효과는 일정하다는 가정 하에 간접비교를 수행할 수 있다. 그러나 이는 매우 강력한 가정이며, 입증의 거의 불가능하고 교환가능성 가정에 영향을 미치는 요인들을 적절하게 고려하지 못할 수 있다. 따라서, 이러한 경우에도 간접비교를 수행하

기 위해서는 기저위험의 차이에 상관없이 상대치료효과가 일정하다는 가정을 뒷받침할 수 있는 근거가 함께 제시됨으로써 해당 상황에서의 간접비교 수행이 정당화되어야 하고, 근거를 제시할 수 없는 경우에는 간접비교 수행을 권장하지 않는다.

공통대조군의 사건발생률이 기저위험의 차이와 무관한 요인들에 의해서 달라진 것이라고 한다면(추적기간이나, 감시체계, 눈가림 등), 상대치료효과가 이러한 임상시험 수행 방법과 상관없이 일관될 수 있다는 가정 하에서 간접비교를 수행할 수 있다. 그러나 이러한 가정을 입증하는 것은 현실적으로 불가능하므로, 부득이 간접비교를 수행하고자 한다면, 공통대조군 결과의 차이가 기저위험의 차이와 무관하여 상대치료효과에 영향을 미치지 않을 것이라고 가정하게 되는 합당한 이유에 대하여 기술하고 및 이러한 입증이 불가능한 가정하에서의 수행된 결과이므로 결과에 불확실성이 크게 존재한다는 내용을 충분히 기술할 것을 권장한다.

● 메타 분석을 고려할 수 있을 때

임상시험이 2개 이상 있는 경우, 특정 임상시험을 분석에서 제외함으로써 분석에 포함된 임상시험들 간에 상대치료효과의 이질성 원인이 될 만한 요인에 대하여 유사성을 증대시킬 수 있다. 그러나 이러한 제외가 분석에 포함되는 임상시험을 선정하는데 비뚤림으로 작용하거나 검정력을 떨어뜨릴 수 있으므로 이는 논란의 여지가 있을 수 있다. 또한, 선정기준을 사전에 명확히 해놓기는 쉽지 않으므로 선정기준에 따른 결과에 대한 민감도 분석을 수행하는 것이 필요한지도 고려하고, 필요한 경우에는 민감도 분석을 수행해야 한다.

일반적으로 1) 기저위험과 관련 없는 요소들에서 상당한 이질성을 보이거나, 2) 공통대조군의 결과에 변동이 크거나, 3) 통계적 이질성을 증가시키는 연구는 교환가능성 가정에 대한 위배의 가능성을 최소화하기 위해 분석에서 제외하는 것을 고려할 수 있다.

간접비교에서 사용할 수 있는 여러 임상시험들에 대한 하나 또는 두 개의 메타분석이 존재할 때 전제조건 확인 및 해결을 위해 다음과 같은 시도가 활용될 수 있다. 메타분석 또는 메타회귀 분석의 수행에 대한 구체적인 수행방법론에 대해서는 참고문헌에 제시된 체계적문헌고찰에 관한 종합적인 안내서로서 코크란 핸드북을 참고하거나

Egger(2001) 또는 박병주 등(2010)의 서적 내에서 관련 장을 참고하면 도움이 될 것이다.

a) 약제 A와 C, B와 C 임상시험에서의 각 상대치료효과가 기저위험의 다름에 상관 없이 일정하다는 것을 보이기 위해 메타분석이 이용될 수 있다. 이는 B와 A간의 합성된 상대치료효과가 각 직접비교가 이루어진 서로 다른 상황에서 일정하다는 가정을 지지하는 것으로 활용할 수 있다.

b) 임상시험 간 환자군 또는 방법론 적인 차이를 통계적으로 보정하기 위해 메타회귀분석을 사용할 수 있다. 그러나 안정적인 결과를 얻기 위해서는 하나의 변수를 보정하기 위해 적어도 10개의 연구가 필요하다는 것을 고려하면 이는 통상적인 해결책이 되기는 어렵다.

c) 위에 대한 변형된 대안으로 네트워크 메타분석이 수행될 수도 있겠으나, 이 또한 많은 수의 임상시험 결과가 필요하다.

● 2개 이상의 공통대조군이 있는 경우 또는 다단계 간접비교

공통대조군을 통한 전제조건 확인에 대하여 앞서 설명된 원칙들은 2개 이상의 공통대조군이 있는 간접비교 또는 다단계 간접비교 (그림3-3 참고)에도 적용된다. 이러한 상황에서의 간접비교도 개념적으로는 수행 가능한 방법론이다. 그러나 분석의 단계 수가 늘어나거나 복잡해질수록 더 많은 불확실성이 개입되므로, 간접비교를 단순화하고 단계를 최소화하는 것이 필요하다.

그림 3-3의 마지막 경우와 같이 공통대조군의 하나 이상 있을 수 있는 경우에는 임상적인 근거 및 연구의 방법론적인 공통성 등의 근거를 기반으로, 가장 적절한 공통대조군을 선정하는 것을 권고한다.

● 소결

이분형 자료로서 사건발생률을 통해 기저위험을 평가할 수 있는 경우, 공통대조군의 사건발생률을 통한 간접비교의 평가방법을 요약하면 아래와 같다.

1) 공통대조군의 사건발생률이 임상시험 간에 유사

공통대조군의 사건발생률이 임상시험 간에 유사하고 사용된 임상시험 방법론이 유사하다면 임상시험에 포함된 환자군은 같은 기저위험을 가진 것으로 간주한다. 이 경우, 간접비교를 통해 절대차(RD, NNT) 또는 상대비(RR, OR)로 결과를 측정하여 통합한다.

2) 공통대조군의 사건발생률이 임상시험 간 다르나, 임상시험은 방법론적으로 유사 사건발생률이 상이한 것은 환자의 기저위험 차이 때문일 가능성 높다. 이 경우, 추정된 OR 또는 RR과 합성된 상대치료효과가 서로 다른 기저위험 간에 일정함을 추가로 가정하여 간접비교를 수행할 수 있다. 그러나 이러한 가정은 소수의 임상시험만의 간접비교에서 증명될 수 없다. 상대치료효과가 서로 다른 기저위험 간에 일정함을 정당화할 수 있는 근거를 제시할 수 없는 경우에는 결과에 불확실성이 존재함에 대해 충분히 기술해야한다

3) 공통대조군의 사건발생률과 임상시험 간 방법론이 모두 다름

이때 간접비교의 수행은 더욱 큰 문제를 가져올 수 있다. 임상시험수가 많다면, 메타회귀분석을 수행하여 방법론적인 차이나, 연구의 질, 환자특성 등을 설명하거나 보정할 수 있다. 또는 임상적인 판단에 크게 의존하는 수밖에 없다. 그러나 이 경우 OR 또는 RR이 여타 이질성에 대해서 일정하다는 가정을 자료로써는 입증할 길이 없다.

3.2.3. 교환가능성 가정에 영향을 미칠 수 있는 요인과 이질성 평가

3.2.3.1. 합성된 상대치료효과의 이질성을 일으킬 수 있는 요인 평가

무작위배정임상시험 수행 방법, 치료법, 결과변수 측정방식 등을 포함한 관련 요인들이 기저 위험에 영향을 미치지 않을 정도로 임상적으로 동질할 때, 합성된 상대치료효과가 포함된 전 임상시험에서 일정하다고 가정할 수 있게 된다. 이를 확인하기 위해 약제 A와 대조군 C를 비교한 무작위배정임상시험과, 약제 B와 대조군 C를 비교한 무작위배정임상시험 내에서와 두 임상시험들 간 모두에서 평가가 이루어져야 한다. 그러나 이러한 평가는 정량적인 통계적 추론보다는 정성적인 형태로 기술한다.

합성된 상대치료효과의 이질성에 영향을 미칠 수 있는 요인은 다양하나(표 3-1) 엄밀히 말하면, 치료효과를 교란시킬 수 있는 요인들에 대해서만 임상시험들이 유사하면 된다. 다시 말해 약제 A, B, C 모두에 전혀 영향을 미치지 않는다면 굳이 그 요인에 대해서 임상시험들이 유사할 필요는 없다. 문제는 해당 약물들이 그 요인에 전혀 영향을 받지 않는다는 것을 제시하는 것이 어렵다는 점이다. 임상적인 동질성을 평가하는데 가장 큰 제한점은 정보가 부족하다는 것이다. 뿐만 아니라 가능한 교란요인이 무엇일지 모두 확인하여 교란의 가능성을 모두 제외시킨다는 것은 사실상 불가능하다. 교란요인이 발견되었을 때는, 메타회귀분석 방법을 통해 보정할 수 있으나 요인 하나를 보정하기 위해서 필요한 무작위배정임상시험 수는 최소한 10개로서 이러한 보정이 안정적으로 이루어질 수 있는 상황은 드물다.

표 3-1 합성된 상대치료효과의 이질성에 영향을 미칠 수 있는 요인

A. 임상시험의 수행방법 또는 질	1) 무작위배정 2) 눈가림 3) 추적관찰기간 4) 추적손실(loss to follow up) 5) 교차치료(cross over)
B. 임상시험에 참여한 환자 특성	1) 연령 2) 성 3) 유전자 4) 진단방식 5) 감시체계 강도 6) 질병 중증도 7) 생리적 특성 8) 유병 기간 또는 진행단계 9) 선행 치료법 10) 동반질환 11) 표준치료의 발전
C. 임상시험 환경과 관련한 교란요인	1) 보건의료체계 2) 지리 3) 의료기관 4) 임상시험 시점
D. 치료약제의 차이 (공통대조군과 비교약물)	1) 용량 2) 복용기간 3) 복용시점
E. 결과변수 측정방식	1) 정의 2) 척도 3) 측정 주기 4) 측정시점

참고자료 원: PBAC, 2008

3.2.3.2. 합성된 상대치료효과의 측정치와 통계적 이질성

1) 측정지표

무작위배정임상시험간의 상대치료효과가 동질성은 상대치료효과의 측정지표에 따라 다를 수 있다. 이분형 자료 분석에서 상대적 측정치에 의한 치료효과가 절대적 측정치에 의한 치료효과보다 더 일관된 것으로 알려져 있다. 따라서 적절한 측정지표를 사용하면 교환가능성에 대한 가정이 더 잘 유지될 수 있다(각 지표에 대해서는 3.2.4장에서 자세히 다룸). 연속형 변수를 사용할 경우 변수값의 기저치와 환자 기저 특성을 비교하여 해당 결과변수의 최종 결과를 토대로 분석할 경우 이질성이 발생하지 않는지 확인해야 한다. 변화량을 토대로 한 지표를 사용할 경우 결과변수 기저치의 차이를 보정하는 측면이 있다. 그러나 이 경우에도 최종 결과에 영향을 미치는 요인(표 3-1 참조)들에 대하여 확인이 필요하다.

2) 통계적 동질성

약제 A와 약제 C를 비교한 무작위배정임상시험과 약제 B와 약제 C를 비교한 무작위배정임상시험이 여러 개 존재하여 이를 근거하여 간접비교를 수행할 때, 상대치료효과크기 값들의 이질성을 확인해야 한다. 약제 A와 약제 C의 상대치료효과가 모든 약제 A와 약제 C를 비교한 각각의 무작위배정임상시험들에 대해 일관되어야 하고, 약제 B와 약제 C를 비교한 무작위배정임상시험들에서도 마찬가지여야 한다. 이는 합성된 약제 A와 약제 B의 상대치료효과가 모든 약제 A와 약제 C를 비교한 무작위배정임상시험들, 약제 B와 약제 C를 비교한 무작위배정임상시험들에 대해 일관될 가능성을 보여주는 방법이다. 그러나 통계적 이질성 검사법은 검정력이 낮으므로 통계적 이질성이 나타나지 않더라도 이를 의심해 보아야 한다. 특히 약제 A와 약제 C를 비교한 무작위배정임상시험과 약제 B와 약제 C를 비교한 무작위배정임상시험간에 임상적인 이질성이 발견될 때는 통계적 동질성 확인만으로 교환가능성을 판단해서는 안 된다.

한편, 약제 A와 약제 B를 간접비교하는데 하나 이상의 공통대조군이 존재하는 경우라면 각 공통대조군에서 얻어진 합성된 약제 A와 약제 B의 상대치료효과 들 간에 동질성이 유지되어야 한다.

3.2.4. 적절한 결과지표의 조건 및 측정지표 선택시 주의 사항

3.2.4.1. 이분형 자료

이분형 자료 분석에서 사용하는 효과 측정지표는 크게 절대적 측정지표(absolute measure)와 상대적 측정지표(relative measure)로 구분된다. RD(Risk Difference), NNT(Number of patients Needed to be Treated)와 같은 절대적 측정지표는 두 그룹 간 평균값에서 절대적인 차이를 추정한 통계량으로 해석이 용이하고 개별 환자들에 대한 임상적 결정을 내릴 때 유용하다는 측면이 있다. 그러나 절대적 측정지표는 환자 집단의 기저위험에 따라 값이 달라질 가능성이 높아 기저위험이 다른 환자집단에 적용하기 어렵다는 문제가 있다.

반면 RR, OR과 같은 상대적 측정지표는 환자 집단의 기저위험과 상관없이 일정할 가능성이 높다. 따라서 특정 기저위험이 있는 하부집단들에 적용하는데 더 용이하다.

상대적 측정지표 중 RR은 OR보다 더 이해하기 쉬운 지표이다. 그러나 RR은 RD(0 기준으로 대칭)와 OR(1기준으로 대칭)과 달리 1 근처에서 대칭적이지 않다는 문제가 있다. 특히 위해에 대한 RR(RR for Harm: RRH)과 편익에 대한 RR(RR for Benefit: RRB)는 서로 다른 측정지표나 다름없다. 관심 있는 사건(편익 기준)을 반대 의 사건(위해 기준)으로 전환하면(예를 들어 생존 여부를 사망 대신 사용한 경우)이 때의 RRH는 RRB의 역수와 다르기 때문이다. 즉 1 근처에서 대칭적이지 않다. Montreuil(2005)은 RR을 측정지표로 사용했을 경우 수술 후 합병증이 발생한 것을 기준으로 했을 때와 수술 후 합병증이 발생하지 않은 사건을 기준으로 했을 때의 결과가 대칭되지 않음을 보여주었다. 즉, 수술 후 합병증이 발생한 것을 기준으로 했을 때 기존 약물에 비해 새로운 약물의 위험 감소 효과는 0.5로 기존 약물에 비해 수술 후 합병증 발생의 상대적 위험은 50%였던 반면 합병증이 발생하지 않을 상대적 위험은 1.02로 나타났다. 반면 OR의 경우 어떤 방향으로 결과를 보든지 일관된 것을 알 수 있다. 표 3-2는 사망여부를 효과지표로 사용했을 때와 사망여부 대신 생존여부를 사용한 두 무작위배정임상시험에서 OR와 RR를 사용했을 때의 결과를 비교하여 제시하고 있다. 임상시험1에서 사망여부에 기반을 둔 RR은 0.5, 생존

여부에 기반을 둔 RR은 1.02이고 임상시험2에서 사망여부에 기반을 둔 RR은 0.5였지만 생존여부에 기반을 둔 RR은 3이었다. 만일 사망을 효과 지표로 사용했다면 임상시험의 결과는 비슷한 상대치료효과를 보이지만, 생존여부를 효과 지표로 사용했다면 임상시험1의 결과는 효과가 없는 것으로 나타난 반면 임상시험2는 매우 큰 효과가 있는 것으로 나타난다. 반면, 이를 OR로 계산을 하면 사망여부를 기준으로 할 때 연구1에서는 0.49, 연구2에서는 0.167이며, 생존여부를 기준으로 하면 연구1에서는 2.04, 연구2에서는 6으로 결과가 일관된다. 사건이 드물게 발생할 때는 OR과 RR이 거의 같지만 위의 예는 사건 많이 발생하게 됨에 따라 이러한 관계가 왜곡됨을 보여준다.

표 3-2 OR와 RR 비교

구분		임상시험1		임상시험2	
		치료군	비치료군	치료군	비치료군
사망기준		2%	4%	40%	80%
생존기준		98%	96%	60%	20%
RR	사망기준	0.2/0.4=0.5		0.4/0.8=0.5	
	생존기준	0.98/0.96=1.02		0.6/0.2=3	
OR	사망기준	(0.2/0.98)/(0.4/0.96)=0.49		(0.4/0.6)/(0.8/0.2)=0.167	
	생존기준	(0.98/0.2)/(0.96/0.4)=2.04		(0.6/0.4)/(0.2/0.8)=6	

자료원 : Walter(2000), Montreuil et al.(2005)

공통대조군의 기저위험이 다를 경우 간접비교에서 RR을 사용하게 되면 합성한 상대치료효과의 수준뿐 아니라 방향까지도 일관되지 않을 수 있다. Eckermann et al.(2009)은 예를 통해 공통대조군 C를 사용한 약제 A와 약제 B의 간접비교에서 약제 A와 약제 C를 비교한 무작위배정임상시험에서의 약제 C의 기저위험과 약제 B와 약제 C를 비교한 무작위배정임상시험에서의 약제 C의 기저위험이 다를 경우 RR을 사용하면 생존여부에 기반을 둔 RRB에서는 A치료법이 B보다 더 효과적인 것으로 나타나고 사망여부에 기반할 경우(RRH) B가 A보다 더 효과적인 것으로 나타날 수 있음을 보여주었다. 반면 OR의 경우 간접비교에서 합성한 상대치료효과의 정도와 방향에서 일관된 경향을 보였다.

간접비교 수행시 측정지표의 선택은 간접비교에서 사용한 무작위배정임상시험의 특성에 달려있다. 만일 기저위험이 건강보험 적용 대상과 간접비교에 포함한 무작위배정임상시험에서의 환자군에서 동일하다면, 절대적 측정지표로서 RD를 상대치료효

과로 사용하는 것이 적절할 것이다.

A와 C를 비교한 임상시험 및 B와 C를 비교한 임상시험들에 포함된 중재군 중 어느 한 군에서라도 기저위험이 다를 경우 간접비교를 하려면 무작위배정임상시험들이 잠재적 교란요인을 고려했을 때 비슷한지 확인해야 한다(표 3-1). 만일 간접비교를하기로 결정했다면 상대적 측정지표(OR 또는 RR)가 RD보다 더 선호된다. 왜냐하면 위에서 언급했듯이 상대적 측정지표가 기저위험과 상관없이 일관될 가능성이 더 높기 때문이다.

드물지만 기저위험이 달라도 RD가 기저위험과 상관없이 일정한 경우가 있다. 이 경우 RD를 적용하는 것이 상대적 측정치에 의한 치료효과를 기준으로 점증적 효과를 잘못 계산하는 것보다 더 나을 수 있다. OR과 RR이 기저위험과 상관없이 일정하다면 OR를 더 선호한다.

간접 비교에서 사용한 어느 임상시험에서는 RRB를 사용하고 어느 임상시험에서는 RRH를 사용한 경우 간접비교를 하기에 앞서 RRH와 RRB를 전환하여 결과를 확인하는 것이 필요하다. 효과 측정지표에 따라 결과가 바뀌거나 관련된 상대치료효과의 수준을 바뀔 정도라면 결과를 해석하는데 주의가 필요하다.

결론적으로, 간접비교 수행시 기저위험과 관련된 문제가 없을 경우 이분형 자료의 합성을 위해 RD를 직접적으로 사용할 수 있고, 기저위험이 달라도 상대적 측정지표에 의한 치료효과가 기저위험과 상관없이 일정하다고 한다면, 간접비교에서 합성된 상대치료효과 추정은 상대적 측정치를 사용하여 계산하는 것을 권고한다. 이 경우 특정한 측정지표를 사용하는데 따른 근거를 명확히 제시해야 한다.

3.2.4.2. 연속형 자료

연속형 자료에서 사용하는 효과 측정치는 평균차(Mean difference, MD)와 표준화된 평균차(Standardized mean difference, SMD) 등 평균을 토대로 한 지표를 주로 사용하게 된다. 평균차는 임상시험에서 두 그룹간 평균값에서 절대적인 차이를 측정하는 것으로 대조군과 비교하여, 중재군이 평균적으로 의료결과를 변화시키는 양을 추정한다. 모든 연구들에서 의료결과의 측정법이 동일한 scale일 때, 이들 결과를 통합할 수 있다. 표준화된 평균차(SMD)는 모든 연구들이 모두 같은 의료결과를 평가

하지만, 다양한 방법으로 측정했을 때, 요약 통계량으로 사용된다. SMD는, 결과들을 통합하기 전에 단일 scale로 연구결과를 표준화하여 각 연구결과에서 관찰된 변이(표준편차)에 대한 그 연구에서 중재효과의 크기로 표현된다.

$$SMD = \frac{\text{그룹간평균차}}{\text{표준편차}}$$

SMD는 연구들의 표준편차의 차이는 연구대상자의 변이(variability)의 실제 차이가 아니라 measurement scales에서의 차이를 반영하는 것으로 보기 때문에 다른 연구들 간 변이의 실제 차이가 있을 것으로 판단되는 상황에서는 문제가 있을 수 있다.

Baseline으로부터 변화량을 토대로 한 분석은 최종 효과측정치를 비교하는 것보다 더 효율적이고 강력할 수 있다. 그러나, 변화량 계산은 결과(outcome)를 두 번 측정해야 하고 측정오차가 커져 덜 효율적일 수 있다. 변화량을 토대로 한 효과 측정치의 분포가 최종 결과를 토대로 한 효과 측정치들의 분포보다 덜 치우친 분포(skewed distribution)를 보인다면 변화량을 토대로 한 분석이 더 선호될 수 있다. 한편 최종 결과를 토대로 한 효과 측정치와 변화량을 토대로 한 효과 측정치를 표준화된 평균차(SMD)로 통합해서는 안 된다.

연속형 자료를 통합하는 기본적인 가정은 각 중재군의 결과가 각각의 연구에서 정규분포를 갖고 있는 것이다. 따라서 자료를 통합하는데 있어 정규성을 확인해야 하며 이를 보정하는 방법으로 log scale로 전환하여 사용할 수 있으나 이 경우 Log로 전환된 자료와 전환되지 않은 자료를 통합할 수 없으므로 주의해야 한다.

연속형 자료의 경우 기저위험과 관련된 문제가 없다면 자료의 정규성을 확인한 후 평균차나 표준화된 평균차이 중 선택하여 사용할 수 있다. 그러나 연속형 자료의 경우 이분형 자료에서와 같이 공통대조군 결과를 통해 모집단 간 기저위험률이 일정한지에 대한 평가 등이 불가능하므로 간접비교가 권장되지는 않는다.

3.2.5. 간접비교에 의한 상대치료효과 추정

3.2.5.1. 단순비교

무작위배정임상시험들에서 보고자 하는 하나의 치료군의 결과만을 추출하여 메타 분석 등으로 합쳐서 결과값을 도출하고 같은 방법으로 추출한 다른 치료군의 결과값과 비교하는 방법을 단순비교라고 한다(3.2.1.1 참조). 이러한 단순비교가 간접비교를 위해 권장되지 않는 가장 중요한 이유는 비교를 구성하는 각 무작위배정임상시험들에서 이루어진 무작위 배정을 무시하는 것이기 때문이다. 즉 공통대조군이 있을 경우에는, 공통대조군의 사건발생률이 비슷한 것을 이용하여 통합된 상대치료효과에 교란을 일으킬 수 있는 요인들이 각 시험에서 비슷한지 최소한의 평가가 가능한 반면 (사건발생률이 유사한 것이 교란요인들이 유사한 것의 필요조건) 단순비교는 그것을 불가능하게 만든다. 또한 A vs C 연구와 B vs C 연구간 변이 (variation)를 설명하지 못하며, 공통대조군 간접비교에 비해 분산이 작으므로 신뢰구간이 좁아지게 된다(ICWG Report, 2008).

3.2.5.2. 공통대조군 간접비교

Bucher et al.(1997)에서 제안된 방법으로 직접비교에서 추정된 두 치료 효과(A vs C, B vs C)를 각각 Δ_{AC} , Δ_{BC} 라고 하고 3.2.2의 전제조건들을 만족할 때, 약제 A와 약제 B간의 간접비교 치료 효과는 $\Delta_{BA} = \Delta_{BC} - \Delta_{AC}$ 로 나타낼 수 있고 이 때 분산은 $Var(\Delta_{BA}) = Var(\Delta_{BC}) + Var(\Delta_{AC})$ 와 같이 표현된다. Song et al.(2003)의 체계적 문헌고찰에 의하면 이 방법은 직접비교와 대개 일치하는 결과를 보여주나 항상 그렇지 않고 특히 선정된 시험에 비뚤림이 있을 경우 그대로 간접비교 결과에 영향을 미칠 수 있다. Glenny et al.(2005)에서도 같은 자료로 비슷한 결론을 내리고 있다.

직접비교와 비교해 간접비교는 상대적으로 추가적인 이질성이 존재할 가능성이 커서 랜덤효과모형(Random Effect Model)이 고정효과모형보다 선호된다(Glenny et al., 2005). 이 경우 신뢰구간이 더 커지게 되고 이것은 추가적인 불확실성을 반영하는 것이므로 타당한 것이라고 할 수 있다. 간접비교 방법의 대표적인 한계점은 이분형 자료가 아닌 연속형, 범주형, 생존자료 등의 경우 간접비교에 포함되기 힘들다는 점이다. 이분형이 아닌 자료의 경우 기저 위험률이 일정한지에 대한 평가를 하

기 어려우므로 ICWG Report(2008)의 경우에는 간접비교에서 배제하도록 권하고 있다.

Odds Ratio를 지표로 하여 공통대조군 간접비교를 통한 상대치료효과추정의 수식의 예는 다음과 같이 제시될 수 있다.

$$OR_{AB} = \frac{OR_{AC}}{OR_{BC}} = \frac{odds_A / odds_C}{odds_B / odds_C} = \frac{odds_A}{odds_B}$$

$$\ln OR_{AB} = \ln OR_{AC} - \ln OR_{BC}$$

$$SE(\ln OR_{AB}) = \sqrt{SE(\ln OR_{AC})^2 + SE(\ln OR_{BC})^2}$$

$$95\% \text{ CI} = [\text{Exp}\{(\ln OR_{AB}) - 1.96 \times SE(\ln OR_{AB})\}, \text{Exp}\{(\ln OR_{AB}) + 1.96 \times SE(\ln OR_{AB})\}]$$

공통대조군 간접비교를 수행하기 위해서는 위와 같은 공식을 이용할 수도 있지만 많은 시험들에서 추출해야 하는 경우는 CADTH에서 만든 ITC 소프트웨어 (<http://www.cadth.ca/index.php/en/itc-user-guide/download-software>)를 사용하면 보다 쉽게 수행할 수 있다. 단, 치료군 수 10개, 임상시험 수 20개의 제한이 있어 그 보다 더 많은 수가 필요한 간접비교의 경우 각각의 메타분석(랜덤효과모형)으로 합친 결과값들과 위의 수식을 이용할 수밖에 없다. 구체적인 사용법에 대한 상세한 안내는 해당 사이트에서 확인할 수 있다.

간접비교에 있어 포함될 무작위배정임상시험을 정하는 것이 중요한 역할을 하게 되는데 이때 적용할 수 있는 절대적인 원칙은 존재하지 않는다. 하지만 특정 시험의 제외 또는 선정의 결정이 논쟁의 여지가 있을 경우, 전체적인 임상적인 평가로부터의 결론에 특정시험의 제외여부가 결과에 차이를 만드는지 검토하는 민감도 분석을 제시하는 것을 추천한다. 즉, 그 시험을 제외했을 때와 선정했을 때의 각각의 간접비교 결과를 제시하는 것이 제외 또는 선정 결정의 정당성을 설명하는데 도움을 줄 수 있다.

3.2.5.3. 혼합비교

신청약제와 비교약제의 직접비교 임상시험이 존재할 경우 직접비교를 통해 두 약제의 상대치료효과를 추정하는 것이 원칙이나 작은 표본수, 비뚤림 가능성 등의 이유로 직접비교 임상시험을 신뢰할 수 없는 경우에는 직접비교와 간접비교를 결합하는 혼합비교방법을 통하여 두 약제의 치료효과를 비교할 수도 있다. 또한 비교약제가 여러 개인 경우에도 혼합비교방법을 통해서 두 약제의 상대치료효과를 추정할 수 있다. Lu and Ades(2004)는 그림 3-6에서 제시되는 가정이 모두 만족될 경우 베이지안 계층적 모형을 이용해 관심있는 상대치료효과를 추정하는 방법을 제시하였다. 그러나 혼합비교는 간접비교에 비해 더욱 복잡한 통계적 방법론을 사용하게 되므로, 이질성과 교환가능성 가정이 복잡한 형태의 분석 안에서 유지될 수 있겠는가를 독립적으로 평가하는 것이 더욱 어렵다. 따라서 혼합비교에 의한 결과는 보조적이고 추가적인 정보로 고려해야 한다. 혼합비교방법에 대한 이론적 배경 및 사례에 대해서는 부록에 기술되어 있다.

3.2.6. 교환가능성 확인 및 간접비교 수행 절차

교환가능성 가정에 대한 확인을 통한 간접비교 수행을 위해 다음의 단계를 따라 평가를 진행할 것을 권고한다. 구체적인 평가를 위한 사항은 표시된 해당 절의 설명을 참고한다.

1) 각 직접비교 임상시험에 대한 평가

먼저 약제 A와 공통대조군 C를 비교한 직접비교 임상시험에 대하여 다음의 사항을 평가한다.

- 이용 가능한 임상시험들에 대해 표 3-1을 참고하여 각 임상시험 간 여러 요인들의 차이를 기술하여 비교하고, 각 요인이 이질성에 미치는 영향을 확인하여 이질성 요인을 평가한다. (3.2.3.1장)
- 공통대조군에서의 결과값의 유사성 및 기저위험과의 관계를 확인한다. (3.2.2.3장)

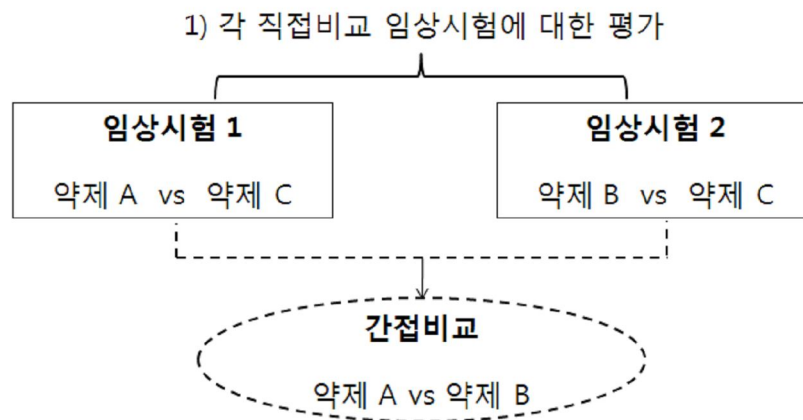
- 상대치료효과 측정치의 선택의 적절성을 평가한다 (3.2.4장)
- 임상시험들 간 상대치료효과의 통계적 동질성을 평가한다 (3.2.3.2장)

이와 같은 과정을 약제 B와 공통대조군 C를 비교한 임상시험들에 대해서도 동일하게 수행한다.

2) 간접비교 추정치에 대한 합성

약제 A와 약제 B의 상대치료효과에 영향을 미칠만한 요인들을 두 세트의 직접비교 임상시험들(A와 C, B와 C)간에 비교하여 앞서와 같은 방식으로 평가 한다

- 공통대조군에서의 결과값의 유사성을 평가한다
- 약제 B에 대한 약제 A의 상대치료효과 측정치 선택의 적절성을 평가한다



2) 간접비교 추정치에 대한 합성

그림 3-8 교환가능성 확인과 간접비교의 수행

이와 같은 평가과정이 명확하고 투명하게 제시될 수 있도록 간접비교 수행 단계별로 결과 제시방법을 아래와 같이 권고한다.

- 1) 간접비교에 포함 가능한 모든 임상시험들을 확인하고 이에 대해 기술한다. (4.2와 4.3장 참고)

2) 다음의 요인들을 고려하여 분석에 선정된 임상시험들에 대한 타당성을 제시한다.

- 이질성 요인 측면 (4.4.1 와 4.4.2장 참고)
- 공통대조군의 결과 측면 (4.5장의 결과제시 단계에서 확인)
- 효과 측정치 선택이 통계적 이질성에 미치는 영향 (4.4.3장 참고)

3) 간접비교 계산방식에 대한 적절하고 명확한 기술과 신뢰구간을 제시한다 (4.5장 참고)

4) 잠재적으로 논란의 여지가 있는 임상시험의 포함여부에 대하여 민감도 분석을 추가적으로 수행하여 제시한다.

본문에서 구체적으로 제시된 바와 같이 각 세부단계에서는 입증되기 불가능한 가정을 전제로 간접비교를 진행하게 되는 경우가 있으며, 증명될 수 없는 가정 하에서 소수의 임상시험만으로 수행된 간접비교를 통해 얻어진 결과를 해석하고 수용여부를 판단할 때는 그 근거수준을 낮추어 고려하여야 할 것이다.

4. 임상적 유용성 평가를 위한 간접비교 수행단계별 자료제출 지침

4.1. 서론

신청약제와 비교약제 간 주 적응증에 대한 최선의 임상 근거를 확인하고 제시함에 있어, 직접비교 무작위배정임상시험이 존재하는 경우에는 이에 근거한 임상적 유용성 평가가 선호되며, 이 경우 임상적 유용성 평가 및 자료제출을 위한 방법으로는 직접비교 무작위배정임상시험의 수행 및 결과 제시 또는 이에 대한 메타분석의 수행 및 결과 제시에 필요한 원칙을 따른다.

본 지침은 임상적 유용성 평가를 위해 이용 가능한 직접비교 무작위배정임상시험이 부재한 경우, 이에 대한 대안으로서 두 개 혹은 그 이상의 무작위배정 비교임상시험 및 하나 이상의 공통대조군이 존재하는 상황에서 이를 이용하여 간접비교를 수행하고 결과를 제시함에 있어서 필요한 절차와 제시방법을 포함한다.

이용 가능한 직접비교 무작위배정임상시험이 부재한 경우라 함은, 직접비교 무작위배정임상시험이 부재하거나 존재한다고 할지라도 연구수행이나 보고된 내용에 대하여 비둘림에 대한 우려 또는 너무 적은 표본 수 등의 문제로 인해 직접비교 임상시험을 신뢰할 수 없는 경우도 포함한다.

근거의 수준에 따라 신청약제와 비교약제를 직접적으로 비교한 무작위배정임상시험에 근거한 평가를 가장 선호하며, 평가에 적절히 활용될 수 있는 직접비교 무작위배정임상시험이 부재한 경우에는 본 지침에 따른 하나이상의 공통대조군을 포함한 두 세트 이상의 무작위배정임상시험에 근거한 간접비교를 고려하되, 본 지침이 요구하는 검색전략 및 검색과정을 제시함으로써 기저 근거수준에 대한 명확한 제시가 우선되어야 할 것이다.

본 맥락에서의 간접비교의 수행과 결과제시를 위한 본 지침의 작성은, 신청약제 및 설정된 비교약제에 대한 평가에 활용할 수 있는 직접비교 무작위배정임상시험의 부재를 확인함을 전제로 하여, 아래의 순서를 따라 요구되는 사항을 제시하는 것으

로 이루어졌다. 즉, 신청약제와 비교약제를 포함하는 무작위배정 비교임상시험에 대한 검색을 수행하여 적절한 직접비교 무작위배정임상시험이 존재하는 경우 직접비교를 수행하여야 하며, 직접비교 무작위배정임상시험이 존재하지 않는 경우 이를 직접비교 무작위배정임상시험 검색결과를 통해 제시하고, 직접비교 무작위배정임상시험이 존재하지 사용가능하지 않은 경우 부적격 사유를 제시한 후, 다음의 절차를 따라 간접비교를 수행한다.

1. 신청약제 또는 비교약제를 포함하는 잠재적인 무작위배정 비교임상시험에 대한 검색전략 및 검색결과 제시

- ▶ 그림 4-2 및 그림 4-3 제시

2. 공통대조군을 통한 간접비교에 포함될 모든 무작위배정임상시험에 대한 선정

- ▶ 표 4-1을 이용하여 간접비교를 위해 검색된 후보 무작위배정 임상시험 목록을 작성

- ▶ 표 4-4 부터 표 4-9 까지 작성을 통해 간접비교분석에 포함 가능성을 최종 판정하여 비교가능 여부를 표 4-1에 표기

- ✓ 표 4-4를 작성하여 간접비교에 잠재적으로 포함 가능한 후보 임상시험들에 관한 비뚤림 위험 평가

- ✓ 표 4-5를 작성하여 간접비교에 잠재적으로 포함 가능한 후보 임상시험 별 자료의 완결성에 대한 평가

- ✓ 표 4-6을 작성하여 각 후보 임상시험 별 선정 제외기준 요약

- ✓ 표 4-7을 작성하여 각 후보 임상시험의 피험자 특성을 군 별로 요약

- ✓ 표 4-8을 작성하여 각 후보 임상시험에 사용된 약제를 군 별로 요약

- ✓ 표 4-9를 작성하여 각 후보 임상시험에 사용된 결과변수를 요약

- ▶ 표 4-2를 작성하여 제외문헌에 대한 제외사유제시

- ▶ 표 4-3를 작성하여 간접비교에 최종 포함된 신청약제, 비교약제, 공통대조군 약제를 임상시험별로 요약

3. 표 4-10을 작성하여 간접비교 분석결과를 결론과 함께 제시.

- ▶ 이때 표에 제시된 공통대조군 결과값을 평가하여 2 단계에서 평가하였던 비교가능성을 최종 확인

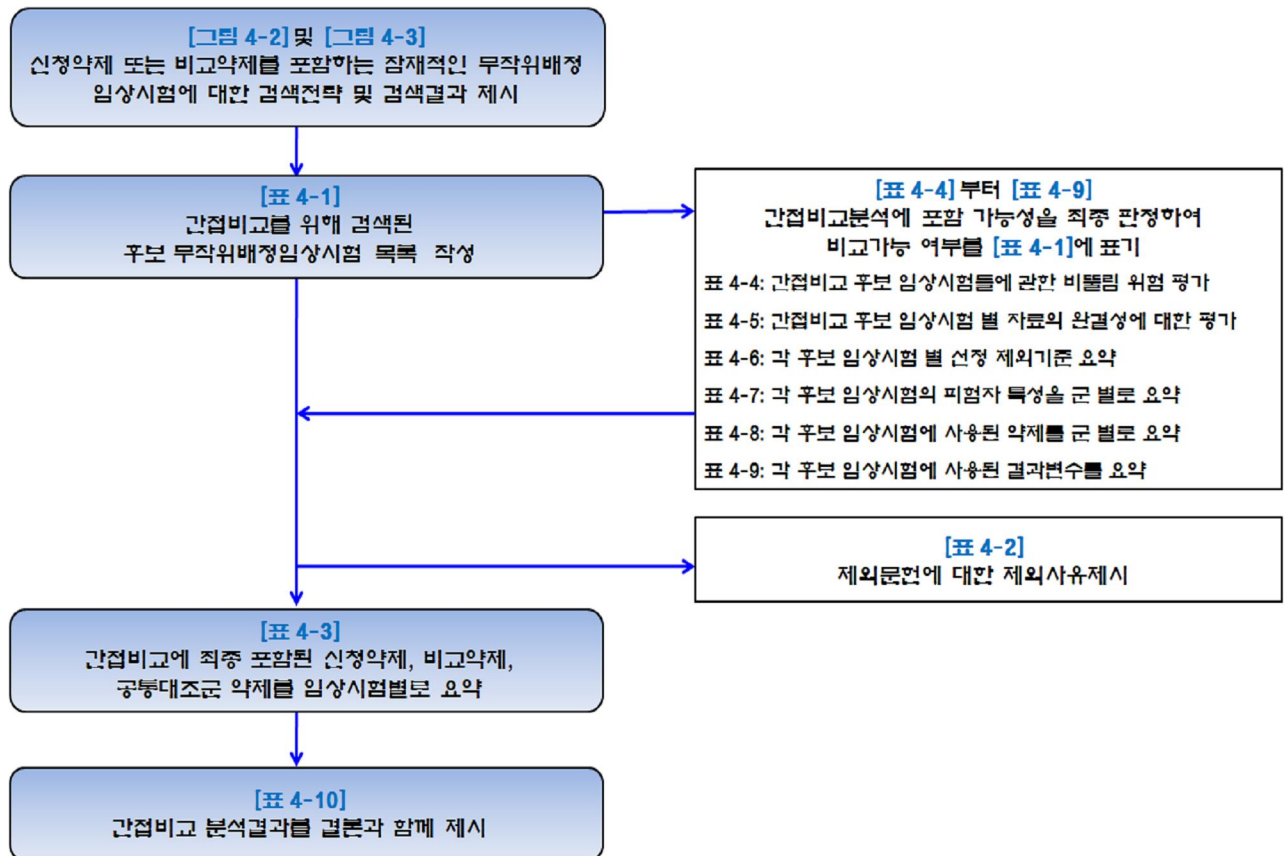


그림 4-1 제출자료 작성 흐름도

4.2. 검색전략에 대한 기술

4.2.1. 요구되는 정보

1. 이용 가능한 직접비교 무작위배정임상시험이 부재함을 제시한다. 이를 위한 임상 시험에 대한 검색전략은 일반적인 체계적 문헌고찰 에서와 같은 방식을 따른다.

검색으로 직접비교 무작위배정임상시험을 찾았으나 이용 불가능한 경우 각 임상 시험 별로 그 이유를 명확히 제시하여야 한다.

2. 직접비교 무작위배정임상시험이 부재함을 보인 후, 간접비교에 이용될 수 있는 모든 무작위배정임상시험을 검색하기 위하여 보다 폭넓은 검색을 실시하고, 출판된 문헌, 임상시험 register, 스폰서의 미출판 자료원 등을 통해 잠재적인 관련 임상시험 문헌들을 찾아내기 위하여 이용된 접근법 및 검색전략을 기술한다.

2단계에서의 검색에 대한 접근은 가능한 공통대조군들을 확인하기 위해 1단계에서 보다 광범위 하게 이루어져야 한다.

4.2.2. 검색전략

검색전략수립 방법은 일차적으로 일반적인 체계적 문헌고찰을 위한 검색 수행에서와 같은 원칙을 따르며 이에 대한 구체적인 접근방법을 기술하여 제시하여야 한다. 신청약제와 비교약제 외에 공통대조군으로 활용할 수 있는 약제는 사전에 구체적으로 결정할 수 있는 것이 아니므로, 신청약제 및 비교약제를 주 비교군으로 하고 신청 약제의 대상이 될 환자들과 같은 특성을 가진 피험자들을 대상으로 수행된 모든 관련 무작위배정임상시험을 찾아낼 수 있도록 검색이 수행되어야 한다.

검색과정은 적어도 아래의 접근방식을 포함해야 하며 검색 전략에 대한 기술은 제3자가 재현할 수 있도록 충분히 구체적으로 기술되어야한다.

- 출판된 문헌에 관한 검색
 - 임상시험 register에 대한 검색
 - 판매 승인을 위해 제약회사가 제출한 문건에 대한 조사와 제약회사의 본사 및 자회사들을 상대로 미출판된 임상시험이 존재하는지에 대한 확인
 - 그 외 경로로 얻은 관련 문헌과 그 참고문헌 목록에 대하여 수작업을 통한 확인
- 출판문헌에 대한 검색전략은 전반적인 검색의 완성도를 평가하는데 있어서 가장 핵

심적이므로, 아래와 같이 검색전략에 대한 특성을 구체화 한다.

- 이용된 전자 데이터베이스 검색원과 임상시험 register에 대하여 구체적으로 제시: 적어도 MEDLINE/PUBMED, EMBASE, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials를 기본적으로 포함하고 특정 약제에 관련한 검색원을 추가
- 해당약제의 회사 자료원과 해당 분야의 관련 임상시험 register들에 대한 자료원 검색 추가
- 검색이 수행된 날짜
- 검색이 적용된 기간(가장 최근 문헌이 검색될 수 있도록)
- 검색어(keyword 및 MeSH 포함)와 이들에 대한 연결 및 검색 방식이 제 3자에 의해 재현 가능하도록 구체적으로 기술된 최종 검색전략안 제시
- 데이터베이스 검색을 통해 확인된 문헌들의 참고문헌에 대한 수작업 확인 등 보충적인 검색에 대한 기술

4.3. 간접비교에 포함될 무작위배정임상시험의 선정

4.3.1. 선정과정 요약

1. 출판된 문헌, 임상시험 registry, 제약사의 미출판 자료원 등 확장된 검색으로 찾은 문헌들을 평가하여 공통대조군을 포함하고 주 적응증에 적합한 환자들과 같은 특성을 가진 피험자를 대상으로 한 신청약제나 비교약제에 대한 무작위배정 임상시험의 선정과정을 그림 4-2과 그림 4-3로 제시한다.
2. 검색된 문헌의 목록을 제시할 때는 (a) 연구 설계에서 무작위 절차가 포함된 시험, (b) 신청약제와 공통대조군 또는 비교약제와 공통대조군을 비교한 임상시험, (c) 주 적응증에 적합한 환자들과 같은 특성을 가진 피험자를 대상으로 했던 임상시험에 대해 각 문헌에 주석을 달아 제외한 문헌의 제외이유를 파악 가능하도록

록 한다.

3. 각각의 공통대조군에 대하여 신청약제와 비교약제에 대한 최종 확인된 무작위배정임상시험의 목록을 제시하고, 일관성을 위해서 각각의 시험에 대한 연번을 명시하고 제출서에 공통으로 사용한다.
4. 신청약제 A와 비교약제 B를 간접비교하기 전에, 신청약제 A와 공통대조군 C를 비교한 임상시험들에 대해 이질성을 평가하고, 비교약제 B와 공통대조군 C를 비교한 임상시험들의 이질성을 평가한다. 그리고 신청약제 A와 공통대조군 C를 비교한 임상시험들(A vs C)과 비교약제 B와 공통대조군 C를 비교한 임상시험들(B vs C) 간의 비교가능여부 뿐만 아니라, 공통대조군 간의 비교가능여부를 검토하고 비교가 가능하지 않은 무작위배정임상시험들은 제외해야 한다. 이때 제외하는 문헌에 대해서는 제외이유를 제시해야 한다.
5. 최종적으로 간접비교에 포함된 비교가능한 무작위배정임상시험을 약제별로 요약하고, 각 무작위배정임상시험의 문헌을 첨부한다. 이때 선정은 되었지만 비교가능하지 않아서 제외된 무작위배정임상시험에 대한 문헌도 제출서류의 부록으로 제출해야 한다.
6. 간접 비교 수행을 위한 관련 무작위배정임상시험이 없다면, {해당사항 없음}으로 표시한다.

4.3.2. 검색결과의 제시

다음 선정 기준을 만족하는 모든 무작위배정임상시험을 추출하기 위해, 확장된 검색 전략을 수립하여 찾은 모든 문헌을 평가해야 한다.

- 연구 설계에서 무작위 절차가 포함된 무작위배정임상시험
- 신청약제와 공통대조군 또는 비교약제와 공통대조군을 비교한 임상시험
- 주 적응증에 적합한 환자들과 같은 특성을 가진 피험자를 대상으로 했던 임상시험

검색 결과의 제시는 신청약제를 포함하고 있는 문헌을 검색한 경우와 비교약제를 포함하고 있는 문헌을 검색한 경우로 나누어서 제시해야 한다. 먼저 신청약제를 포함하고 있는 출판된 문헌 검색결과를 그림 4-2과 같이 정리하여 제시하고, 비교약제를 포함하고 있는 출판된 문헌 검색결과도 그림 4-2과 같이 정리하여 제시한다. 이 중 직접비교 무작위배정임상시험을 통한 임상적 유용성을 확인하는데 이용 가능한 문헌이 있는지 확인한다.

신청약제에 대해서 제약사의 미출판 자료원이 있는 경우 출판된 문헌 검색결과인 그림 4-2 뿐만 아니라 그림 4-3로 제약사 미출판 자료원까지 정리하여 제시한다. 이때 신청약제에 대해서 출판된 문헌 검색결과 최종 확인된 무작위배정임상시험의 수를 n_1 이라고 하고, 제약사의 미출판 자료원에서 최종 확인된 무작위배정임상시험의 수를 n_2 라고 하기로 한다.

그러나 비교약제의 경우 제약사의 미출판 자료원에 대한 확보가 어렵기 때문에 그림 4-2만 제시하며, 이때 최종 확인된 무작위배정임상시험의 수를 n_3 라고 기입한다.

3개의 흐름도를 통해 간접비교 포함될 무작위배정임상시험의 문헌의 수 ($N=n_1+n_2+n_3$)를 도출하고 이를 표 4-1로 정리한다.

이용 가능한 직접비교 임상시험 문헌이 부재한 경우에도, 신청약제에 대한 그림 4-2과 그림 4-3의 흐름도 및 비교약제에 대한 흐름도를 제시해야 한다.

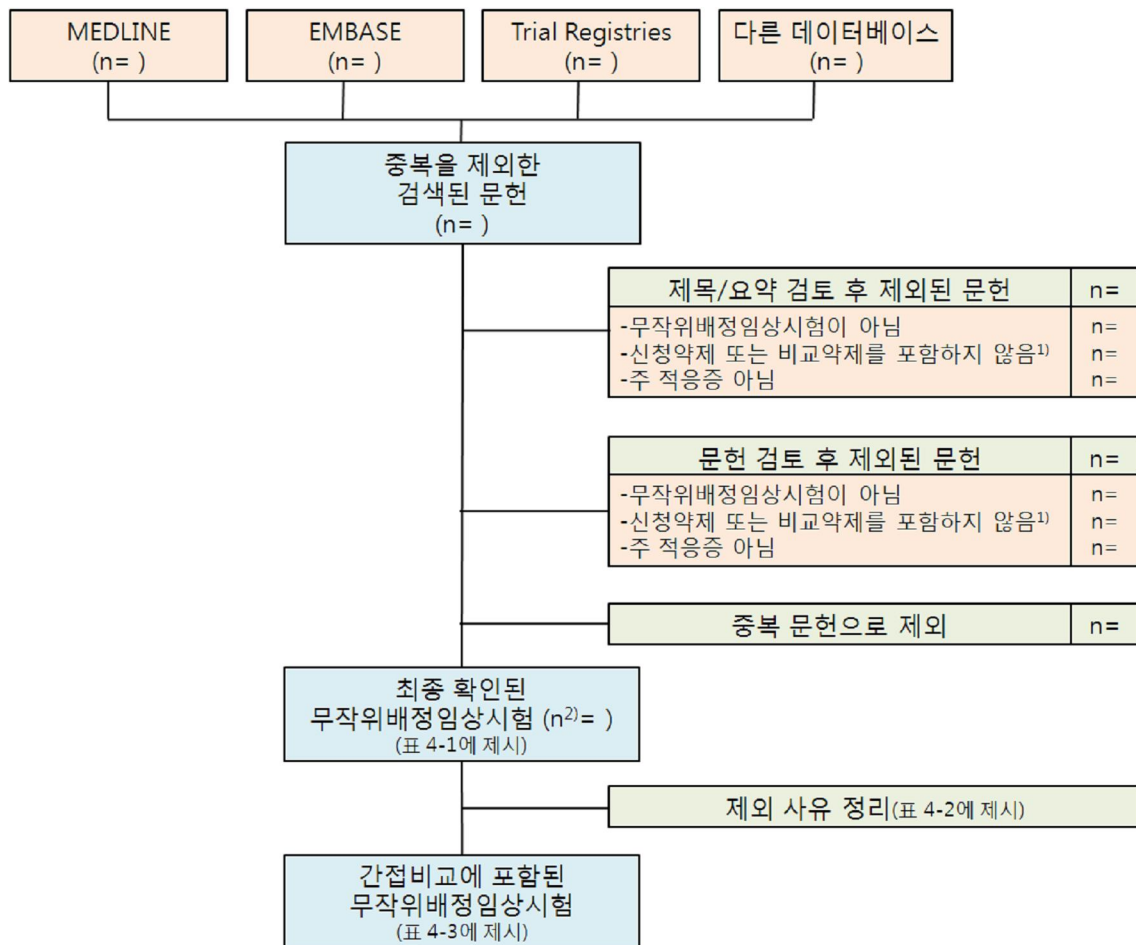


그림 4-2 출판된 문헌 검색 및 선정 흐름도

- 주: 1) 직접비교 무작위배정임상시험이 부재한 경우, 이 흐름도를 신청약제에 대한 무작위배정임상시험 검색 결과 제시에 사용하는 경우 "신청약제를 포함하지 않음"으로 변경하여 제시하고, 비교약제의 경우 "비교약제를 포함하지 않음"으로 변경하여 제시
- 2) 신청약제에 대한 무작위배정임상시험 검색 결과 제시에 사용하는 경우 n_1 , 비교약제의 경우 n_3 로 제시

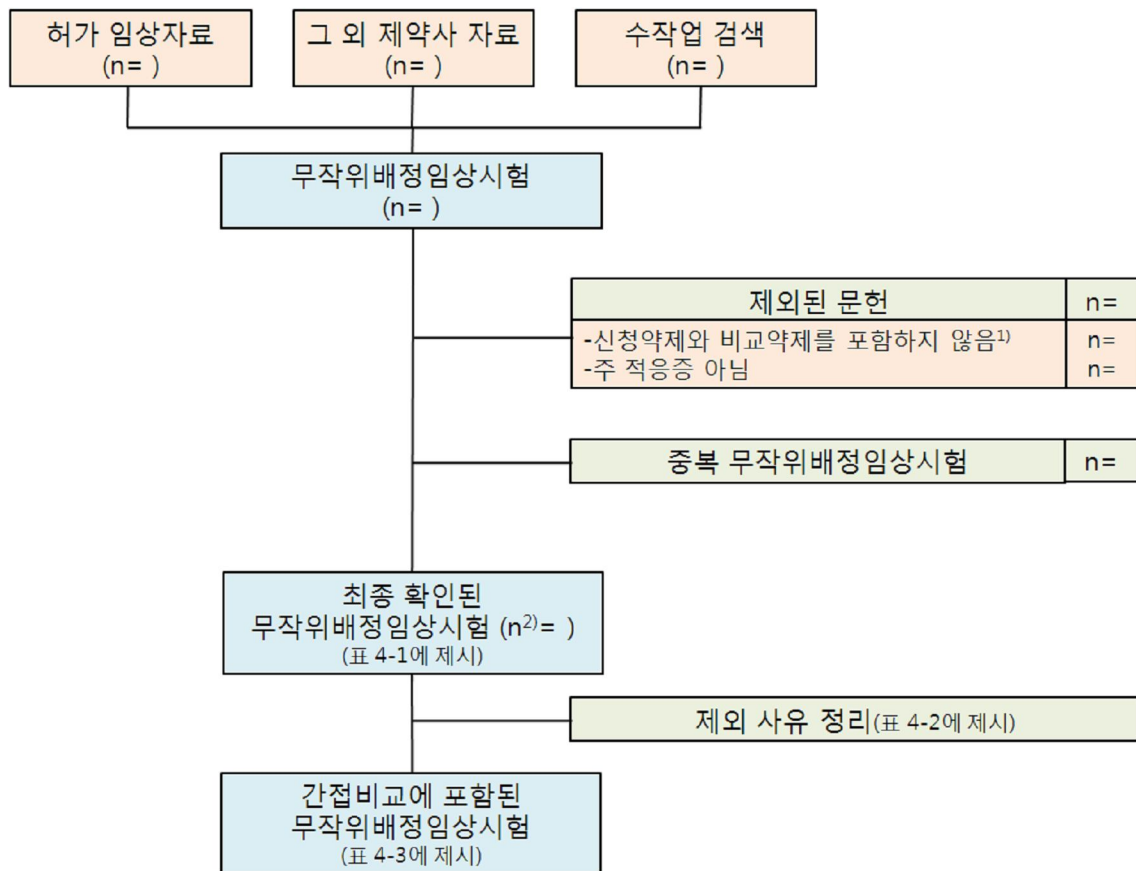


그림 4-3 제약사의 미출판 자료원 선정 흐름도

- 주: 1) 직접비교 무작위배정임상시험이 부재한 경우, 이 흐름도를 신청약제에 대한 무작위배정임상시험 검색 결과 제시에 사용하는 경우 "신청약제를 포함하지 않음"으로 변경하여 제시하고, 비교약제의 경우 "비교약제를 포함하지 않음"으로 변경하여 제시.
- 2) 신청약제에 대한 무작위배정임상시험 검색 결과 제시에 사용하는 경우 n_2 로 제시.

4.3.3. 최종 확인된 무작위배정임상시험의 목록 제시

신청약제와 비교약제에 대한 확장된 검색의 결과에 대한 위의 두 개의 흐름도로부터, 무작위배정 임상시험의 관련 문헌을 확인 할 수 있다. 간접 비교를 위해 검색을 통해 찾은 최종 확인된 무작위배정 임상시험의 목록은 표 4-1과 같이 제시한다. 이때 일관성을 위해서 각각의 시험을 식별을 위한 시험연번 등을 명시하고 이후 제출서에 공통으로 사용한다.

표 4-1 간접비교를 위해 검색된 후보 무작위배정임상시험 목록

시험연번	임상시험 요약 ¹⁾	참고문헌 ²⁾	비교가능여부 ³⁾
공통대조군: C			
신청약제: A			
RCT1			Y/N
RCT2			Y/N
비교약제: B			
RCT3			Y/N

주: 1), 비교군, 주적응증에 대한 정보를 포함하고 (RCT 아닌 경우 무작위여부 포함) 비열등성(또는 우위성)임상시험인지 등에 대한 정보를 포함해야 한다. 비열등성 임상시험의 경우, 비열등성 인정한계 등의 정보를 포함해야 함

2) 내부 연구 보고서인 경우 보고서 제목, 날짜를 제시하고, 출판 문헌인 경우 저자명(1저자 등), 논문제목, 학회지명(문건명), 출판 연도; 권(번호):페이지를 기재하도록 함

3) 표 4-1부터 표 4-9까지 완성하여 일차적으로 평가하고, 표 4-10의 공통대조군 결과를 평가한 후 최종적으로 비교가능성 여부를 편정 함

4.3.4. 간접비교를 위한 무작위배정임상시험의 선정/제외

신청약제와 비교약제를 직접 비교한 무작위배정 임상시험이 없는 경우, 간접비교의 결과가 직접비교 임상시험과 같은 비교가능성을 확립하기 위해, 간접비교에 포함될 무작위배정 임상시험을 선정해야 한다. 이때 무작위배정 임상시험의 질(quality), 환자의 특성 및 임상시험의 환경, 결과변수 등을 고려하여 간접비교에 포함하기 부적절한 무작위배정 임상시험을 제외하는 방법으로 비교가능성을 확인함으로써 간접비교에 최종적으로 포함될 무작위배정 임상시험을 선정한다.

간접비교에서 무작위배정임상시험의 제외이유는 다음과 같다.

- 무작위배정임상시험의 질에서 중요한 차이가 있는 경우(예, 불충분한 추적기간)
- 환자의 기저특징에서 중요한 차이가 있는 경우(예, 심한 질환을 가진 환자들이 포함된 시험의 치료효과는 가벼운 질환을 가진 환자들이 포함된 시험의 치료효과와 비교가능하지 않을 수 있음)
- 보고된 결과변수의 차이(예, 다른 시험에서는 고려하지 않은 결과를 보고한 시험이 있을 경우)
- 공통대조군에서의 차이(예, 활성(active) 공통대조군을 사용한 경우 시험들 간의

용량이 다를 경우)

간접비교가 공통대조군의 비교가능여부에 많이 의존하고 있기 때문에 공통대조군에서의 차이를 검토하는 것은 매우 중요하며, 추가적으로, 무작위배정 임상시험들이 다른 시기 또는 다른 지역에서 수행되었을 경우, 의료행위와 환자 특성의 변화는 비슷한 치료가 비교가능하지 않다는 것을 의미할 수 있기 때문에 이러한 측면에서 검토하여 비교가능성 여부를 고려하여야 한다.

간접비교에 포함된 무작위배정 임상시험 사이의 관찰할 수 있는 차이는 최소화하거나 시험간의 이질성에 기여하는 정도를 확인한 후 가능하면 보정해야 한다. 반면 관찰되지 않는 차이는 최소화되거나 보정될 수 없으며, 간접비교와 관련된 불확실성의 원인이 된다.

간접비교에서 무작위배정 임상시험을 제외할 때 적용될 수 있는 명확한 지침을 제공하는 것은 불가능하나 적어도 4.4장과 4.5장에서 기술한 절차를 따를 것을 권고한다. 무작위배정 임상시험을 선정할 경우에는 비교가능여부에 대해 검토해야 하며, 무작위배정 임상시험을 제외할 경우에는 선택 비뒀림(selection bias)문제가 일어날지에 대해 고려해야 한다. 만일 특정 무작위배정 임상시험의 제외 또는 선정이 논쟁의 여지가 있을 경우, 포함여부가 전체적인 임상적인 평가로부터의 결론에 차이를 만드는지 여부에 대해 민감도 분석을 수행하는 것을 권고한다.

하나 이상의 무작위배정 임상시험이 간접비교에서 제외되었다면, 표 4-2와 같이 각 무작위배정 임상시험별로 제외된 이유가 시험의 질, 환자 특성과 임상시험 환경, 결과변수와 관련이 있는지 표시해야 한다. 이때 시험의 질 측면에서는 무작위, 맹검(blinding), 추적관찰 등을 고려하여 기술해야 하며, 제외이유가 여러 개인 경우 제외이유에 따라 정리해야 한다.

표 4-2 간접비교에서 무작위배정임상시험의 제외 이유

시험 연번	제외 이유							
	시험의 질			교환가능성				기타
	무작위 배정	맹검	추적 관찰	환자 특성	임상시험 환경	결과 변수	공통 대조군 결과	
RCT1								
RCT2								
RCT3								

주: 1) 제외 이유에 해당하는 칸에 표기한 후, 구체적 사유를 본문에 기술하여 제시함

최종적으로 간접비교에 포함된 무작위배정임상시험은 표 4-3과 같이 약제별로 용량
까지 포함하여 요약하여 제시하여야 한다.

표 4-3 간접비교에 포함된 약제의 무작위배정임상시험별 요약

시험연번	신청약제	공통대조군	비교약제
RCT1	용량	용량	-
RCT2	용량		-
RCT3	-	용량	용량
RCT4	-	-	용량
RCT5	-	-	용량

주: 1) 병용요법 용량이 있는 경우 용량을 함께 표시

표 4-1에 제시된 간접비교를 위해 검색된 후보 무작위배정임상시험들에 대해 4.4장
과 4.5장에서 제시된 방법에 따라 교환가능성 검토과정을 본문에 제시하고, 각 무작
위배정 임상시험의 문헌을 첨부로 제출해야 한다. 이때 대조군 2개가 동시에 임상
시험에 포함된 경우, 4.4장에 따라 교환가능성을 검토한 후 적절한 하나의 공통대조
군을 정의한 후 간접비교를 수행해야 한다. 또한 선정은 되었지만 비교 가능하지
않은 무작위배정 임상시험에 대한 문헌도 제출서류의 부록으로 제출해야 한다. 무
작위배정 임상시험을 본문에 제시할 때 첨부된 보고서의 페이지, 표, 그림 번호 등
참조를 분명히 제시해야 한다.

4.4. 교환가능성 평가

4.4.1. 간접비교에 고려된 무작위배정임상시험에서 비뚤림 위험의 요약평가

1. 표 4-4와 표 4-5를 이용하여 간접비교를 위해 고려된 무작위배정임상시험들의 비뚤림 위험을 기술한다.
2. 각 무작위배정임상시험들의 비뚤림 위험을 부록 지침을 참고하여 평가한다.
3. 무작위배정임상시험들 중, 비교성이 확보되었다고 생각되는 무작위배정임상시험들에 대하여 존재할 수 있는 연구 질의 차이점을 명시화한다.

표 4-4 간접비교를 위한 후보임상시험별 비뚤림 위험 평가결과의 요약

시험연번	무작위배정표 생성	배정순서 은폐	눈가림	자료의 완결성	선택적 보고없음	기타 비뚤림없음
RCT1						
RCT2						
RCT3						

주: [부록3] 비뚤림 위험 평가 지침을 참조하여 iYes/No/Unclearj에 대해 평가

표 4-5 간접비교를 위한 후보임상시험별 자료의 완결성에 대한 평가

시험연번	신청약제	공통대조군 후보				비교약제
		위약	약제 1	약제 2	약제 3	
RCT1						
배정된 환자수	N	N	N	N	N	N
추적 손실	n(%)	n(%)	n(%)	n(%)	n(%)	n(%)
중단	n(%)	n(%)	n(%)	n(%)	n(%)	n(%)
분석에 포함	n(%)	n(%)	n(%)	n(%)	n(%)	n(%)
비고						
RCT2						
배정된 환자수	N	N	N	N	N	N
추적 손실	n(%)	n(%)	n(%)	n(%)	n(%)	n(%)
중단	n(%)	n(%)	n(%)	n(%)	n(%)	n(%)
분석에 포함	n(%)	n(%)	n(%)	n(%)	n(%)	n(%)
비고						

주: 필요한 경우 i비고j란을 추가하여 기입하며 구체적인 내용은 별도로 첨부

4.4.2. 간접비교에 고려된 무작위배정임상시험의 요약

1. 표 4-6, 표 4-7, 표 4-8을 이용하여 간접비교에 고려된 무작위배정임상시험들을 요약하여 기술한다.
2. 각 무작위배정임상시험들의 연구시점과 연구 수행기관에 대하여 기술한다.
3. 간접비교를 위해 고려된 무작위배정임상시험들을 위와 같은 관점에서 비교하고, 중요한 차이점에 대하여 평가한다.

간접비교는 연구대상의 특성이 치료결과에 영향을 미치지 않도록 무작위배정 절차를 수행하지 않으므로 무작위배정임상시험의 특성에 대한 기술을 통하여 간접비교에 포함된 무작위배정임상시험 간 비교를 가능하게 한다. 간접비교에 포함된 연구들의 비교가능성을 고려하기 위해서는 각 무작위배정임상시험에서 선정된 연구대상들의 기저위험에 대한 평가와 공통대조군에서 사용된 용량 요법을 평가하는 것이 필요하다. 이와 더불어 연구시점과 연구 수행기관/국가/지역도 평가한다. 의료기술과 연구대상 특성의 변화, 예를 들어 각 무작위배정임상시험이 수행되었던 연구시점과 지리적 위치에 차이가 존재할 경우, 치료법들간 비교 가능하지 않을 수 있기 때문이다.

표 4-6 간접비교를 위한 후보임상시험별 선정 제외기준의 요약

시험연번	선정 기준	제외 기준	추적관찰기간
RCT1(연도, 국가)			
RCT2(연도, 국가)			
RCT3(연도, 국가)			

표 4-7 간접비교를 위한 후보임상시험별 피험자의 특성

시험연번	신청약제	공통대조군 후보				비교약제
		위약	약제 1	약제 2	약제 3	
RCT1						
나이						
성별						
진단						
기타						
RCT2						
나이						
성별						
진단						
기타						

주: 표 3-1 을 참조하여 i이질성에 영향을 미칠 수 있는 피험자의 특성 i 추가

표 4-8 간접비교를 위한 후보임상시험별 비교되는 중재

시험연번	신청약제	공통대조군 후보				비교약제
		위약	약제 1	약제 2	약제 3	
RCT1						
용량 요법						
추적관찰기간						
RCT2						
용량 요법						
추적관찰기간						

4.4.3. 간접비교에 고려된 무작위배정임상시험별 결과변수

1. 표 4-9를 이용하여 간접비교에 고려된 무작위배정임상시험들의 결과변수들을 제시하되, 각 결과변수들의 고유의 측정단위, 측정기간 등 결과변수의 정의를 구체적으로 제시한다.
2. 각 무작위배정임상시험에서 측정된 결과변수들 간의 중요한 차이점을 비교하고 평가한다.

같은 결과변수라도 측정 방법은 시험에 따라 다를 수 있다는 점을 고려한다. 결과 변수에 대한 구체적인 기술을 통하여 신청약제와 공통대조군을 비교한 임상시험들 간, 비교약제와 공통대조군을 비교한 임상시험들간의 비교가 가능한지를 평가한다. 또한 신청약제와 공통대조군을 비교한 임상시험들과 비교약제와 공통대조군을 비교한 임상시험들간 비교 가능한지를 평가한다.

표 4-9 간접비교를 위한 후보임상시험별 결과변수

시험연번	결과변수의 정의	통계 분석 방법
RCT1		
RCT2		
기타		

주: 임상시험별 1차 결과변수는 "*" 표시

4.5. 간접비교 결과 제시 및 해석

1. 각 공통대조군에 대한 결과들을 (간접비교에 사용된) RCT 군들 간에 중대한 차이가 있는지 평가한다.
2. 다음과 같이 결과를 제시한다.
 - 이분형 결과들에 대해서는 3.2.4.1장에 따라 각 무작위배정임상시험의 치료효과 추정치를 선정하여 95% 신뢰구간과 함께 제시하고, Time to event 결과들에 대해서는 각 무작위배정임상시험의 치료효과 추정치를 HR과 이에 대한 95% 신뢰구간 제시하며, 연속형 결과들에 대하여는 3.2.4.2장에 따라 각 무작위배정임상시험의 치료효과 추정치를 선정하여 95% 신뢰구간과 함께 제시
 - 한 군에서 무작위배정임상시험이 두 개 이상이면 공통대조군과 신청약제와 공통대조군을 비교한 임상시험들과 비교약제와 공통대조군을 비교한 임상시험들에서 보고한 치료효과들을 각각 random effect 모형으로 메타분석 수행
3. 간접비교의 결과에서 합성한 상대측정치에 따른 효과의 크기와 95% 신뢰구간을 계산하기 위해 사용한 방법을 기술하고 사용 이유를 명시한다.

첫째 간접비교에 사용된 무작위배정임상시험들의 공통대조군에 대한 결과들이 중대한 차이가 있는지 평가한다. 이것은 무작위배정임상시험들이 비교가능한지 확인하기 위함이다. 이상적으로 비슷한 환자군에서 같은 공통대조군을 이용하여 비슷한 성과를 본 결과는 비슷하여야 한다.

개별 RCT와 RCT군들의 통합효과 추정치를 제시할 때는 상대적인 치료효과를 계산해야 한다. RCT군들 간의 간접비교효과 추정은 상대적인 치료효과의 비(95% 신뢰구간과 함께)를 사용한다. 상대적인 치료효과를 사용하는 것이 공통대조군 결과의 차이를 보정하는데 도움을 줄 수도 있다. 상대적인 치료효과는 절대적인 치료효과보다(추적기간의 차이도 포함하여)인구집단간의 변화의 정도가 더 작다. 추천하는 결과제시표는 아래의 표와 같다.

표 4-10 간접비교 결과의 요약

시험연번	신청약제가 포함된 임상시험			비교약제가 포함된 임상시험			간접비교를 통한 상대치료 효과 추정치 ²⁾ (95% CI)
	치료효과 추정치 ²⁾ (95% CI)	신청 약제 ³⁾	공통 대조군 ³⁾	치료효과 추정치 ²⁾ (95% CI)	비교 약제 ³⁾	공통 대조군 ³⁾	
RCT1							
RCT2							
RCT3							
⋮							
메타분석 ¹⁾							
이질성검증							

CI: 신뢰구간

주: 1) Random Effect 모형을 이용한 메타분석 결과(이질성 결과 제시)

2) RD, RR, OR 등 지침에 따라 적절하게 선택하여 사용

3) 이분형 자료는 n/N(n 사건의 수, N 각 군에 포함된 수) 형태로 표시하고 연속형 자료는 평균과 표준편차로 표시

간접비교의 결과를 계산하기 위해 사용한 추가적인 방법이나 기타 방법은 기술하고 근거를 밝힐 때는 사용한 방법이 재현가능하고 독립적으로 확인 가능함을 보장하여야 한다. 이를 위해 1) 컴퓨터를 이용한 분석 자료(재현을 위한 분석프로그램 및 자료)와; 2) 사용된 소프트웨어에 대한 정보 및 시스템 정보(CPU, OS version 등)를 제공하여야 한다.

필요하다면 덜 비교가능한(예를 들어 시험에 포함된 인구집단의 차이나 약물의 용량 등의 차이) 시험들을 배제하여 간접비교결과를 추가로 제시하여야 한다. 대안으로 간접비교에 보정방법을 사용하였다면 이에 대해 설명하고 제시하여야 한다.

가능하다면 기저의 가정에 대해 상대치료효과에 거의 변화가 없다고 할 수 있는 통계학적 배경이 있는지를 평가하여야 한다. 예를 들어 몇몇 질병들에 대해서는 치료효과의 차이가 기저위험에 따라 달라지는 것이 관찰되는데 이것은 절대효과의 차이이지 상대효과의 차이가 아니다. 따라서 상대위험이 일정하다는 가정의 배경이 될 수 있다.

위해보고는 간접비교나 RCT의 틀을 벗어나 상대적인 위해를 제시하여야 한다(관찰연구나 보고연구의 경우도 고려하여 약물의 부작용이나 의존성을 제시하여야 함).

간접비교를 할 때 선택 비뚤림과 같은 중요한 비뚤림을 최소화하지 못하기 때문에 간접비교의 결과와 해석은 조심스럽게 논의해야 한다. 임상평가의 결과에 근거하여 신청약제를 가장 잘 요약할 수 있도록 결론을 제시할 것을 권고한다.

4.6. 단순비교

공통대조군을 통한 간접비교를 수행할 수 있는 무작위배정임상시험이 부재하여 부득이 단순간접비교를 수행하여 제시하여야 한다면, 공통대조군을 통한 간접비교 수행이 불가능함을 제시한 후 앞의 절차를 다시 따라가며 표 4-1부터 표4-9를 작성하여 각 임상시험과 군을 요약하고 평가하며 단순비교과정을 제시할 수 있다. 다만, 이때에는 공통대조군 관련 부분이 제외될 것이며, 표 4-10의 결과표는 다음과 같이 (표 4-11) 수정하여 제시될 수 있을 것이다.

이때에는 공통대조군 비교가 불가능하였으며 교환가능성에 대한 평가 또한 충분히 이루어지지 못하였으므로, 단순비교결과만으로는 최종 결과에 따르는 비뚤림에 대한 잠재적 방향과 정도에 대한 파악이 불가능하므로 결과를 수용함에 있어 공통대조군을 통한 간접비교의 결과와 같은 근거수준이 적용될 수는 없다.

항암제 연구 등과 같이 대부분 단일군 2상 임상시험만 존재 하여 공통대조군을 통한 간접비교 수행이 불가능하나 분석에 포함할 수 있는 충분한 수의 연구가 존재 한다면, 앞서 제시된 절차를 통하여 결과에 비뚤림을 초래할 수 있는 요소들을 파악하고 간접비교에 메타회귀분석을 이용한 보정을 수행하여 단순비교를 통한 상대 효과 추정에 따른 문제의 극복을 시도할 수 있을 것이다.

표 4-11 단순비교 결과의 요약

시험연번	신청약제가 포함된 임상시험 결과	비교약제가 포함된 임상시험 결과	간접비교를 통한 상대치료 효과 추정치 ²⁾ (95% CI)
	치료효과 추정치 (95% CI)	치료효과 추정치 (95% CI)	
문헌1			
문헌2			
문헌3			
⋮			
메타분석 ¹⁾			
이질성검증			

CI: 신뢰구간

주: 1) Random Effect 모형을 이용한 메타분석 결과(이질성 결과 제시)

2) 분석방법을 구체적으로 제시하고 보정을 수행하였다면 보정 방법에 대한 구체적 기술을 추가로 제시

5. 간접비교 평가 사례 제시

5.1. 간접비교 사례 선정

본 장에서는 간접비교를 수행하여 출판된 기존의 문헌에서 포함되었던 간접비교 예제를 활용하여 간접비교를 선정하고, 해당 간접비교에 포함되었던 문헌들을 추출하여 이를 대상으로 본 자료제출지침서가 안내하고 있는 절차에 따라 평가수행 절차를 따라 간접비교를 재수행하는 과정을 제시하여 예시하였다.

이미 수행된 간접비교를 대상으로 수행되었으므로, 문제의 선정 과 문헌의 검색 과정 및 이를 통한 잠재적 문헌에 대한 선정 과정은 예시에서 생략되었으며 이는 각 해당 상황에서 전향적으로 고려하여야 할 것이다. 본 예시는 자료제출지침의 4.3.3 검색된 문헌목록 제시에서 출발한다.

본 사례는 Hochberg 등(2003)의 문헌에서 소개된 간접비교를 예제로 삼았으며, 본 문헌은 methotrexate에 불완전한 반응을 보인 류머티스성 관절염 환자를 대상으로 또다른 DMARD(Disease Modifying Antirheumatic Drugs)를 추가하여 사용한 효과를 평가하고자 수행된 분석으로서, methotrexate에 adalimumab, etanercept 또는 infliximab을 추가하는 것을 상호 비교하였다.

본 사례제시를 위한 목적으로, 상기 문헌에서 제시되었던 약물 중 판매 승인이 먼저 이루어진 Etanercept를 비교약제로 하고 Adalimumab을 신청약제로서 비교하는 형태를 취하였으며, 두 약제를 포함하고 있는 세 개의 직접비교 무작위 배정임상시험 문헌들(Keystone(2002); Weinblatt(1999); Weinblatt(2003))을 평가 대상 문헌으로 삼아 예시를 진행하였다(부록의 [별첨자료] 참고).

5.2. 임상시험의 요약 및 교환가능성 평가

표 5-1 간접비교를 위해 검색된 후보 무작위배정임상시험 목록

시험연번	임상시험 요약 ¹⁾	참고문헌 ²⁾	비교가능 여부 ³⁾
공통대조군: 위약			
신청약제: Adalimumab			
Weinblatt 2003	6개월간 Methotrexate를 투여한 불완전하게 반응하는 RA환자를 대상으로 피하로 Adalimumab을 20mg, 40mg, 80mg을 주입하는 군과 위약군을 비교한 임상시험 24주 시점에서 ACR20을 만족한 시험자의 분율차이를 일차 유효성 결과변수로 설정	Weinblatt et al. Adalimumab, a fully humanized anti-tumor necrosis factor a monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate. Arthritis Rheum 2003;48:35?45.	Y
Keystone 2002	- Methotrexate를 투여한 불완전하게 반응하는 RA환자를 대상으로 매주 20mg Adalimumab과 MTX군과 격주로 40mg Adalimumab과 MTX군, 매주 위약과 MTX 투여받는 군을 비교한 임상시험 - Adalimumab의 유효성, 안전성과 radiographic 결과를 평가	Keystone et al. Adalimumab (D2E7), a fully human anti-TNF-a monoclonal antibody, inhibits the progression of structural joint damage in patients with active RA despite concomitant methotrexate therapy [abstract]. Arthritis Rheum 2002; 46(suppl):S205 (abstract 468).	N
비교약제: Etanercept			
Weinblatt 1999	6개월간 Methotrexate를 투여한 불완전하게 반응하는 RA환자를 대상으로 Methotrexate(10~25mg)와 Etanercept 25mg 병용요법과 Methotrexate(10~25mg)와 위약군을 비교한 임상시험 24주 시점에서 ACR20을 만족한 시험자의 분율차이를 일차 유효성 결과변수로 설정	Weinblatt et al. A trial of etanercept, a recombinant tumor necrosis factor receptor: Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. N Engl J Med 1999;340:253?9	Y

주: 1) 비교군, 주적응증에 대한 정보를 포함하고 (RCT 아닌 경우 무작위여부 포함) 비열등성(또는 우위성)임상시험인지 등에 대한 정보를 포함해야 한다. 비열등성 임상시험의 경우, 비열등성 인정한계 등의 정보를 포함해야 함

2) 내부 연구 보고서인 경우 보고서 제목, 날짜를 제시하고, 출판 문헌인 경우 저자명(1저자 등), 논문제목, 학회지명(문헌명), 출판 연도; 권(번호):페이지를 기재하도록 함.

3) 비교가능여부는 표 4-1~표 4-10을 완성한 후, 최종적으로 비교가능여부를 평가 함.

표 5-2 간접비교에서 무작위배정임상시험의 제외 이유

시험연번	제외 이유							
	시험의 질			교환가능성				기타
	무작위 배정	맹검	추적 관찰	환자 특성	임상시험 환경	결과 변수	공통대 조군 결과	
Keystone 2002						O		O

주: 1) 제외 이유에 해당하는 칸에 표기한 후, 구체적 사유를 본문에 기술하여 제시함

Keystone(2002)는 결과변수가 52주 시점의 결과만 제시되어 다른 문헌결과와 비교 불가능하며(24주) 초록으로만 문헌이 존재하여 추가적인 내용과 연구방법론의 구체적인 사항의 확인이 불가능하여 제외하였다.

표 5-3 간접비교에 포함된 약제의 무작위배정임상시험별 요약

시험연번	신청약제 (Adalimumab)	대조군	비교약제 (Etanercept)
		위약	
Weinblatt 2003	40mg(격주)+MTX ¹⁾	위약+MTX ¹⁾	-
Weinblatt 1999	-	위약+MTX ¹⁾	25mg(주2회)+MTX ¹⁾

주: 1) 병용요법 용량이 있는 경우 용량을 함께 표시

표 5-4 간접비교를 위한 후보임상시험별 비뚤림 위험 평가결과의 요약

시험연번	무작위배정 표 생성	배정순서 은폐	눈가림	자료의 완결성	선택적 보고없음	기타 비뚤림없음
Weinblatt 2003	Yes	Unclear	Yes	Yes	Unclear	Unclear
Keystone 2002	Unclear	Unclear	Yes	Yes	No	No
Weinblatt 1999	Unclear	Unclear	Yes	Yes	Unclear	Unclear

주: 부록 지침서를 참고하여 iYes/No/Uncleari에 대해 평가

표 5-5 간접비교를 위한 후보임상시험별 자료의 완결성에 대한 평가

시험연번	신청약제 (Adalimumab)			대조군	비교약제 (Etanercept)
				위약	
용량	20mg	40mg	80mg		25mg
Weinblatt 2003					
배정된 환자수	69	67	73	62	
추적 손실	2(0.7%)				
중단	16(5.9%)				
분석에 포함	69(100%)	67(100%)	73(100%)	62(100%)	
비고	결측치는 non-responder로 처리하여 ITT 분석				
Keystone 2002					
배정된 환자수	212	207		200	
추적 손실	0	0		0	
중단	42(10%)			13(6.5%)	
분석에 포함	212(100%)	207(100%)		200(100%)	
비고	결측치 처리방법에 대한 언급은 없음				
Weinblatt 1999					
배정된 환자수				30	59
추적 손실				1(3.3%)	0(0%)
중단				5(20%)	2(3.4%)
분석에 포함				30(100%)	59(100%)
비고				결측치는 non-responder로 처리하여 분석	

주: 필요한 경우 i비고i란을 추가하여 기입하며 구체적인 내용은 별도로 첨부

Weinblatt(1999), Weinblatt(2003)에서는 임상시험에서 환자의 참여중단의 주요 사유로 부작용 또는 효과 부족이라고 보고하고 있으며 Keystone(2002) 또한 임상시험에서 중단의 주요 사유가 부작용 때문인 것으로 보고하였다. 따라서 이들을 non-responder로 간주한 것은 타당한 것으로 판단하였다.

세 문헌 모두 배정순서 은폐에 대한 기술이 부재하였으며, 초록만 존재하는 Keystone(2002) 문헌의 경우 선택적 보고에 의한 비뚤림 또는 그 외 비뚤림의 가능성을 배제할 수 없는 것으로 판단하였다.

표 5-6 간접비교를 위한 후보임상시험별 선정 제외기준의 요약

시험연번	선정 기준	제외 기준	추적관찰기간
Weinblatt 2003 (미국/캐나다)	18세 이상 ACR로 진단받은 RA 환자 적어도 6개월이상 MTX를 투여하였으나 충분히 반응하지 않은 환자	-RA 생물학제제 임상시험의 표준제외기준 적용 -Anti-CD4 therapy 또는 TNF antagonists 처치를 받은 환자 -Active listeriosis 또는 mycobacterial infection 병력이 있는 환자 -입원이 필요한 감염 또는 스크리닝 30일 이내에 정맥재 항생제 처치 또는 14일 이내에 경구 항생제를 처치 받은 환자	24주
Keystone 2002 (알수없음)	incomplete responders to MTX	NS	52주
Weinblatt 1999 (미국/캐나다)	18세 이상 ACR로 진단받은 RA 환자 적어도 6개월이상 MTX를 투여하였으나 충분히 반응하지 않은 환자	-임상시험 시작 2주전까지 sulfasalazine와 hydroxychloroquine 사용을 중지하지 않은 환자 -임상시험 시작 4주전까지 methotrexate 이외의 disease-modifying antirheumatic drug들을 사용한 환자	24주

주: NS(Not Stated) = 언급 없음

표 5-7 간접비교를 위한 후보임상시험별 피험자의 특성

시험연번	신청약제 (Adalimumab)			대조군	비교약제
				위약	(Etanercept)
용량	20mg	40mg	80mg		25mg
Weinblatt 2003					
나이(세)	53.5 _{12.4}	57.2 _{11.4}	55.5 _{11.7}	56.0 _{10.8}	
여성 비중(%)	75.4	74.6	75.3	82.3	
유병기간(년)	13.1 _{8.1}	12.2 _{11.1}	12.8 _{9.9}	11.1 _{8.0}	
RF(IU/liter)	264.6 _{462.7}	269.3 _{390.0}	318.8 _{536.0}	321.2 _{518.2}	
평균 MTX dose(mg/주)	16.9 _{4.4}	16.4 _{4.1}	17.2 _{4.7}	16.5 _{5.0}	
No. of prior DMARD	3.0	2.9	3.1	3.0	
No. of tender joints	28.5 _{14.4}	28.0 _{12.7}	30.3 _{15.7}	28.7 _{15.2}	
No. of Swollen joints	17.6 _{8.7}	17.3 _{8.6}	17.0 _{8.2}	16.9 _{9.5}	
Patient's assessment	57.6 _{21.6}	56.9 _{21.1}	58.8 _{24.9}	58.0 _{23.2}	
Physician's assessment	60.5 _{17.3}	58.7 _{15.8}	62.6 _{14.7}	58.9 _{15.3}	
Disability Index(HAQ)	1.52 _{0.62}	1.55 _{0.61}	1.55 _{0.66}	1.64 _{0.63}	
C-reactive protein	2.8 _{3.1}	2.1 _{1.8}	2.8 _{2.7}	3.1 _{3.9}	
Keystone 2002					
나이(세)				57	
여성 비중(%)				75	
유병기간(년)				11	
RF+					
평균 MTX dose(mg/주)				16.6	
No. of prior DMARD					
No. of tender joints				28	
No. of Swollen joints				19	
Patient's assessment					
Physician's assessment					
Disability Index(HAQ)				1.5	
C-reactive protein				1.7	
Weinblatt 1999					
나이(세)				53	48
여성 비중(%)				73	90
유병기간(년)				13	13
RF+(%)				90	84
평균 MTX dose(mg/주)				18	19
No. of prior DMARD				2.8	2.7
No. of tender joints ¹⁾				28	28
No. of Swollen joints ¹⁾				17	20
Patient's assessment ¹⁾				60	60
Physician's assessment ¹⁾				65	60
Disability Index(HAQ) ¹⁾				1.5	1.5
C-reactive protein ¹⁾				2.6	2.2

주: 1) 중앙 값

각 임상시험 별 피험자의 특성을 비교할 때 연령이 Adalimumab군에서 다소 높은 것을 확인하였으며 여성의 비중이 Etanercept에서 높으며 C-reactive protein의 경우 Adalimumab 40mg군이 Etanercept군이 가장 유사하였다. 그 외 전반적인 유사성과 이의 유의미성에 대하여 임상적으로 전반적인 검토가 필요할 것이다.

신청약제와 비교약제의 각 직접비교 무작위배정임상시험 별 대조군의 C-reactive protein 수치는 3.1과 2.6으로 차이가 존재하였으며, 이에 임상적 의미와 교환가능성이 성립하는지에 대해서는 임상적인 검토가 필요할 것으로 사료된다.

Ketystone(2002) 문헌은 각 군별 자료가 보고되지 않아 구체적인 검토가 불가능하였으나, C-reactive protein level이 다른 임상시험 들에 비해 현저히 낮은 것으로 관찰되었다.

표 5-8 간접비교를 위한 후보임상시험별 비교되는 중재

시험연번	신청약제 (Adalimumab)	대조군	비교약제 (Etanercept)
		위약	
Weinblatt 2003			
용량 요법	20/40/80mg every other week 2 subcutaneous injections of 1.6ml per injection	every other week	
추적관찰기간	24주	24주	
Keystone 2002			
용량 요법	20mg weekly 40mg every other week subcutaneous injections	weekly	
추적관찰기간	52주	52주	
Weinblatt 1999			
용량 요법		twice weekly	25mg twice weekly subcutaneous injections
추적관찰기간		24주	24주

두 개의 Adalimumab을 평가한 문헌에서 공통적으로 동일 용량 및 용법을 사용한 경우는 40mg eow였으나 추적관찰기간은 24주 및 52주로 두 개의 문헌 간 상이하였다.

표 5-9 간접비교를 위한 후보임상시험별 결과변수

시험연번	결과변수의 정의 ¹⁾	통계 분석 방법
Weinblatt 2003	• 24주시점에서의 ACR20 반응률* • 24주시점에서의 ACR50 반응률 • 24주시점에서의 ACR70 반응률 • 24주시점에서의 ACR component들의 absolute change, percent change	Dunnett's test, unadjusted t-test
Keystone 2002	• 52주시점에서의 ACR20 반응률 • 52주시점에서의 ACR50 반응률 • 52주시점에서의 ACR70 반응률 • Radiographic endpoint - Modified sharp X-ray score - Erosion score - Joint space narrowing score • no new erosions • AE	
Weinblatt 1999	• 24주시점에서의 ACR20 반응률* • 12주시점에서의 ACR20 반응률 • 12/24주시점에서의 ACR50 반응률 • 12/24주시점에서의 ACR70 반응률 • Disease and QoL at baseline and at 12 and 24 weeks.	Chi-square test (Fisher's exact test)

*: 1차 결과변수

1) 결과변수에 대하여 같은 용어를 사용한다고 하여도 정의 및 측정방법에 있어서 표준화 되어있지 않은 경우라고 한다면 이를 연구별로 표에 구체적으로 제시하여 비교가능성을 고려하여야 한다.

Keystone(2002)의 경우 1차 결과변수가 분명하게 제시되지 않았으며 Weinblatt(2003)과 Weinblatt(1999)의 1차 결과변수는 24주시점에서의 ACR20 반응률로 동일하였다.

따라서, 간접비교가 가장 적합한 신청약제와 비교약제의 조합은 Adalimumab 40mg eow와 Etanercept 25mg tw인 것으로 다음과 같은 근거에 따라 판단하였다.

- 연령이 Adalimumab군에서 다소 높은 것을 확인하였으며 여성의 비중이 Etanercept에서 높은 등 인구학적 특성에서 다소 차이가 있으나 효과평가에 가장 영향을 미칠 수 있을 것으로 보이는 C-reactive protein의 경우 Adalimumab 40mg군과 Etanercept군에서 가장 유사하였다(그러나 이러한 전반적인 유의미성에 대해서는 임상적인 전문가의 검토 하에 판단되어야 할 것이다).

- Adalimumab을 공통적으로 평가한 동일 용량 및 용법을 사용한 경우는 40mg eow 이었다(그러나 적정용량의 정의는 약물에 따라 질병에 따라 임상적인 전문가의 소견에 따라 검토하고 판단해야 할 것이다).

신청약제(Adalimumab)가 포함된 임상시험과 비교약제(Etanercept)가 포함된 임상시험 간의 효과평가시점이 가장 유사한 경우는 24주로 동일하였고, 또한 신청약제(Weinblatt2003)와 비교약제(Weinblatt1999)에서 1차 결과변수의 평가시점이 24주로 동일하였다. 따라서 24주 시점에서 평가된 효과를 기준으로 간접비교를하기로 결정하였다(24주에서 이루어지는 평가의 적절성에 대하여 임상적인 검토와 판단이 추가되어야 할 것이다.)

따라서, 본 사례제시의 측면에서 정리된 결과에 따라 관찰된 바에 의한 비교가능성이 가장 높고 따라서 교환가능성 가정에 가장 근접할 수 있을 것으로 판단되는 상황은 비교약제인 Adalimumab의 용량 및 용법을 40mg every other week인 것으로 선정하고 추적기간은 24주로 결정하여, 각 문헌에서 공통적으로 1차 결과변수로 정의되었던 ACR20의 분율을 결과변수로 간접비교를 수행하였다.

그러나 비교용량의 최종선택 시 이의 적절성에 대한 임상적인 판단 및 검토가 이루어져야 할 것이다.

5.3. 간접비교 결과 제시 및 해석

공통대조군의 결과값을 비교해볼 때, Adalimumab의 대조군에서의 분율은 9/62 이고 Etanercept의 대조군에서의 분율은 8/30으로 각 공통대조군에서의 결과가 1.8배 가량 차이가 나는 것을 관찰하였다. 이러한 차이의 정도가 두 대상 집단에 대한 기저위험의 차이 측면에서 어떠한 의미를 갖는지 유의미성에 대한 임상적인 고찰이 필요할 것이다.

대조군 값의 상이함에 비교적 영향을 적게 받는 RR과 OR을 이용하여 간접비교를 수행하였을 때, 두 지표에 의하여 합성된 상대치료효과 크기는 상당히 유사한 결과를 보였으며 Adalimumab을 사용한 경우의 반응률이 높음을 제시하였으나 이의 통

계적 유의성은 확보되지 않았다. 역시 통계적 유의성은 확보되지 않았으나, RD를 사용한 경우에도 같은 방향의 결과가 얻어짐을 확인할 수 있었다.

표 5-10 간접비교 결과의 요약(24주에서의 ACR20 반응률)

시험연번	신청약제(Adalimumab)가 포함된 임상시험			비교약제(Etanercept)가 포함된 임상시험			간접비교를 통한 상대치료 효과 추정치 ²⁾ (95% CI)
	치료효과 추정치 ²⁾ (95% CI)	비교 약제 (n/N)	공통 대조군 (n/N)	치료효과 추정치 ²⁾ (95% CI)	신청 약제 (n/N)	공통 대조군 (n/N)	
Weinblatt 2003	12.05(5.04,28.79)	45/67	9/62				
	4.63(2.47,8.66)						
	0.53(0.38,0.67)						
Weinblatt 1999				6.79(2.53,18.21)	42/59	8/30	
				2.67(1.44,4.94)			
				0.45(0.25,0.64)			
메타분석1)	해당사항 없음			해당사항 없음			1.78(0.48,6.62)
							1.73(0.72,4.18)
							0.08(-0.16,0.32)
이질성검증	해당사항 없음			해당사항 없음			

CI: 신뢰구간; n 사건의 수; N 각 군에 포함된 수
주: 1) Random Effect 모형을 이용한 메타분석 결과
2) 순서대로 OR, RR, RD의 결과값을 제시함

1차 평가변수로 선택한 ACR20의 반응률 외에 추가적으로 ACR50의 반응률에 대하여도 간접비교분석을 수행하였다.

ACR50의 경우에는 Adalimumab의 대조군에서의 반응률이 Etanercept 임상시험의 공통대조군에서보다 2.7배가량 높음을 관찰하였다. 합성된 상대치료효과크기를 살펴볼 때, OR과 RR의 경우는 일관된 방향을 보이며 Etanercept를 지지하는 방향을 제시하였으나, RD의 결과는 오히려 반대임을 알 수 있다. 어느 지표에 대해서도 치료효과차이의 통계적 유의성이 제시되지 못하였으나, 절대적으로 측정된 지표와 상대적으로 측정된 지표간의 결과의 상이함은 공통대조군의 결과값의 차이에서 기인한 것일 수 있을 것으로 사료된다.

표 5-11 간접비교 결과의 요약(24주에서의 ACR50 반응률)

시험연번	신청약제(Adalimumab)가 포함된 임상시험			비교약제(Etanercept)가 포함된 임상시험			간접비교를 통한 상대치료 효과 추정치 ²⁾ (95% CI)
	치료효과 추정치 ²⁾ (95% CI)	비교 약제 (n/N)	공통 대조군 (n/N)	치료효과 추정치 ²⁾ (95% CI)	신청 약제 (n/N)	공통 대조군 (n/N)	
Weinblatt 2003	14.06(5.00,39.51)	37/67	5/62				
	6.85(2.88,16.31)						
	0.47(0.33,0.61)						
Weinblatt 1999				18.53(2.36,145.51)	23/59	1/30	
				11.69(1.66,82.47)			
				0.36(0.22,0.50)			
메타분석1)	해당사항 없음			해당사항 없음			0.79(0.08,7.61)
							0.59(0.07,4.96)
							0.11(-0.09,0.31)
이질성검증	해당사항 없음			해당사항 없음			

CI: 신뢰구간; n 사건의 수; N 각 군에 포함된 수

주: 1) Random Effect 모델을 이용한 메타분석 결과(이질성 결과 제시)

2) 순서대로 OR, RR, RD의 결과값을 제시함

본 예제에서는 신청약제로 가정한 Adalimumab 에 관한 또하나의 공통대조군 연구였던 Keystone 2002 가 간접비교의 전제조건을 해치는 임상적 이질성에 의해서 메타분석과 간접비교에서 제외되었다. 만일 신청약제 대비 공통대조군 연구 또는 비교약제 대비 공통대조군 연구가 2개 이상 분석에 남아 있었다면 연구간 동질성에 대한 추가 평가가 필요하다. 본 예제에서 Keystone 2002 연구가 배제되지 않고 남아 분석에 포함되었을 경우를 가상하여 1차 평가편수에 대한 최종결과인 표 5-10 를 다시 산출하여 보면 표 5-12 와 같이 작성 되었을 것이다.

표 5-12 간접비교 결과의 요약(ACR20 반응률) - Keystone 포함

시험연번	신청약제(Adalimumab)가 포함된 임상시험			비교약제(Etanercept)가 포함된 임상시험			간접비교를 통한 상대치료 효과 추정치 ²⁾ (95% CI)
	치료효과 추정치 ²⁾ (95% CI)	비교 약제 (n/N)	공통 대조군 (n/N)	치료효과 추정치 ²⁾ (95% CI)	신청 약제 (n/N)	공통 대조군 (n/N)	
Weinblatt 2003	12.05(5.04,28.79)	45/67	9/62				
	4.63(2.47,8.66)						
	0.53(0.38,0.67)						
Keystone 2002	4.55(2.97,6.96)	122 /207	48 /200				
	2.46(1.87,3.22)						
	0.35(0.26,0.44)						
Weinblatt 1999				6.79(2.53,18.21)	42/59	8/30	
				2.67(1.44,4.94)			
				0.45(0.25,0.64)			
메타분석1)	6.85(2.67,17.60)			6.79(2.53,18.21)			1.49(0.51,4.37)
	3.16(1.71,5.84)			2.67(1.44,4.94)			1.18(0.50,2.83)
	0.43(0.26,0.60)			0.45(0.25,0.64)			-0.02(-0.28,0.24)
이질성검증	CHI ² =3.88,df=1(p=0.05),I ² =74%			해당사항 없음			
	CHI ² =3.37,df=1(p=0.07),I ² =70%						
	CHI ² =4.29,df=1(p=0.04),I ² =77%						

CI: 신뢰구간; n 사건의 수; N 각 군에 포함된 수

주: 1) Random Effect 모형을 이용한 메타분석 결과

2) 순서대로 OR, RR, RD의 결과값을 제시함

이와 같이 작성된 결과를 통해 Adalimumab 를 포함한 공통대조군 연구인 Weinblatt 2003 과 Keystone 2002 연구 간에 공통대조군 결과값이 유사한지 확인 하여야 하고, 비교약제인 Etanercept 를 포함한 공통대조군 연구인 Weinblatt 1999 에서의 공통대조군 결과 값과도 일관되게 유사한지 확인하는 작업이 이루어져야 한다.

이 경우, Etanercept 공통대조군 연구의 공통대조군 값과 비교할 때 공통대조군 값 이 더욱 유사한 쪽은 Weinblatt 2003 보다 Keystone 2002 연구이다. 우리는 앞서 공통대조군 결과값으로 각 군의 기저위험율의 유사한 정도를 판단하고자 하는 배경에 대해 설명하였으며, 공통대조군 값의 유사성만으로 의존하지 않아야 하는 이유에 대하여서도 설명하였다. 본 예제 역시 단지 공통대조군 값만을 비교해서 환자군 기저위험의 유사한 정도를 판별하고자 한다면 오류가 있을 수 있는 부분에 해당한다. Keystone 2002 연구는 52주 시점에서의 결과이며 Weinblatt 2002 연구는 24주

시점에서의 결과임을 생각한다면 결과값이 유사하게 얻어졌다는 것은 오히려 환자의 기저위험이 달랐다는 것으로도 설명될 수 있을 것이다. 앞서의 평가과정에서 보았듯 Keystone 2002 연구는 다른 연구들과 환자 특성 및 연구 설계적인 차이가 있었으므로 공통대조군 결과값이 매우 유사하다고 할지라도 앞서의 평가과정 중에서도 이미 제외되었어야 하는 연구인 것이다.

또한, 각 연구로부터 계산된 치료효과 추정치가 adalimumab 를 포함한 공통대조군 연구인 Weinblatt 2003 과 Keystone 2002 연구 간에 일관되게 동질한지도 확인하여야 한다. 두 연구에 대한 메타분석에서 나타나는 이질성 정도는 70%가 넘으며 통계적으로도 유의함을 이질성검증란에 표기된 통계량에서 확인 가능하다. 이 경우 두 연구가 제시한 치료효과크기가 매우 달랐기 때문에 이와 같은 이질성검정결과를 얻었으나, 일반적으로 2개의 연구와 같이 매우 소수만의 연구결과로 메타분석이 이루어질 때는 이질성검증을 위한 통계검정결과가 유의하지 않더라도 이를 참고적으로만 해석하고 [받아들여질 만한] 정도 안에서 동질한지에 대한 정성적인 기술이 부가되어야 한다.

결과제시 표기방식에 대한 예시의 목적으로 메타분석 결과를 활용한 간접비교를 상대치료효과추정치로 표 5-12에 함께 제시하였으나, 본 분석과 같이 매우 이질한 자료로 수행된 메타분석결과가 얻어졌을 때 이를 이용하여 간접비교를 위한 상대추정치를 산출하는 것은 원칙적으로 옳지 않다. 연구별로 이루어진 세심한 정성적 평가과정을 거쳐서 통과된 연구가 분석에서 통계적으로 이질한 결과를 제시하였다면, 이에 대한 원인 탐색과정이 추가로 이루어져야 할 것이며 이질한 결과를 일으키는 해당연구를 분석에 포함하여하는가에 대한 충분한 논의와 합당한 접근이 추가로 제시되어야 할 것이다. 예를 들어, 포함된 연구 수가 많은 경우에는 전문가의 도움을 얻어 메타회귀분석을 수행하여 차이를 보정하는 등의 접근이 고려될 수도 있을 것이다.

참고문헌

V. 체계적 문헌고찰. In : 박병주 등. 근거중심 보건의료. 고려의학; 2010

Briefing paper for methods review workshop on evidence synthesis (indirect and mixed treatment comparisons). National Institute for Health and Clinical Excellence; 2007

Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50: 683-691.

Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005; 331: 897-900.

Egger M et al. Systematic reviews in health care: meta-analysis in context. BMJ Books, 2001

Fiore MC, Bailey WC, Choen SJ, et al. Smoking Cessation, Clinical Practice Guideline No. 18 (AHCPR Publication no. 96-0692), Rockville, MD: Agency for Health Care Policy and Research, U.S. Department of Health and Human Services; 1996

Glenny AM, Altman DG, Song F et al. Indirect comparisons of competing interventions. Health Technol Assess 2005; 9: 1-134, iii-iv

Guide to the methods of technology appraisal. UK National Institute for Health and Clinical Excellence; June 2008.

Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Commonwealth of Australia Pharmaceutical Benefits Advisory Committee; December 2008

Guidelines for the economic evaluation of health technologies: Canada [3rd

- Edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
- Hasselblad V. Meta-Analysis of Multi-Treatment Studies. *Medical Decision Making*. 1998;18:37-43.
- Higgins J, Green S (Eds). *Cochrane handbook for systematic reviews of interventions*. Wiley & Sons, 2008
- Hochberg MC, Tracy JK, Hawkins-Holt M, Flores RH. Comparison of the efficacy of the tumour necrosis factor α blocking agents adalimumab, etanercept, and infliximab when added to methotrexate in patients with active rheumatoid arthritis. *Ann Rheum Dis* 2003;62(Suppl II):ii13?-ii16
- ICWG Report. Report of the Indirect Comparisons Working Group to the Pharmaceutical Benefits Advisory Committee: assessing indirect comparisons. 2008. Available from the URL: [http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/\\$File/ICWG%20Report%20FINAL2.pdf](http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/$File/ICWG%20Report%20FINAL2.pdf) [Last accessed Sep 24, 2010].
- Keystone E, Kavanaugh AF, Sharp J, Hua Y, Teoh L, Firschkoff S, et al. Adalimumab (D2E7), a fully human anti-TNF- α monoclonal antibody, inhibits the progression of structural joint damage in patients with active RA despite concomitant methotrexate therapy [abstract]. *Arthritis Rheum* 2002;46(suppl):S205 (abstract 468).
- Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association*. 2006;101:447-59.
- Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine*. 2004;23:3105-24.
- Lumley T. Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine*. 2002;21:2313-24.

Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published metaanalyses. *BMJ* 2003; 326: 472.

Song F, Harvey I, Lilford R. Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol* 2008; 61: 455-463.

Song F, Loke YK, Walsh T et al. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: a survey of published systematic reviews. *BMJ* 2009;338: 1-7.

Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI, et al. A trial of etanercept, a recombinant tumor necrosis factor receptor: Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med* 1999;340:253-9.

Weinblatt ME, Keystone EC, Furst DE, Moreland LW, Weisman MH, Birbara CA, et al. Adalimumab, a fully humanized anti-tumor necrosis factor a monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate. *Arthritis Rheum* 2003;48:35-45.

Wells GA, Sultan SA, Chen L, Khan M, Coyle D. Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009

부록

[1] 혼합비교방법

1.1. 이론적 배경

Lu and Ades(2004)가 제시한 혼합비교방법은 베이지안 계층적 모형을 통해 관심있는 상대치료효과를 추정하는 방법이다. 베이지안 방법이란 자료로부터 얻은 모수(parameter)에 대한 정보뿐만 아니라 과거의 경험 등과 같은 사전정보를 활용하여 불확실성을 수량화 한 사전분포(prior distribution)를 결합하여, 모수에 대한 사후분포(posterior distribution)를 추정하고, 이를 활용하여 결론을 도출하는 통계적인 분석방법이다.

Lu and Ades(2004)가 제시한 베이지안 계층적 모형은 고정효과 혼합비교 모형과 랜덤효과 혼합비교 모형으로 크게 나눌 수 있다. 고정효과 혼합비교 모형은 상대치료효과가 각 임상시험별로 동일하다고 가정할 수 있을 경우 사용할 수 있는데, 4개의 치료가 있을 경우 j 번째 임상시험에서 치료 b 에 대한 치료 k 의 상대치료효과에 관심이 있는 경우 베이지안 모형은 다음과 같다.

$$\log it(p_{jk}) = \begin{cases} \mu_{jb} & b=A, B, C; \quad \text{if } k=b \\ \mu_{jb} + d_{tk} = \mu_{jb} + d_{Ak} - d_{Ab} & b=A, B, C; k=B, C, D \quad \text{if } k > b \end{cases}$$

여기서 μ_{jb} 는 j 번째 임상시험의 baseline parameter, d_{tk} 는 치료 b 에 대한 치료 k 의 상대치료효과를 나타낸다.

이때 사전분포로 모수에 대한 사전정보가 미흡할 경우 사용하는 모호사전분포(vague prior)로 다음을 사용할 수 있다.

$$d_{AB}, d_{AC}, d_{AD} \sim Normal(0, 10^4)$$

$$\mu_A, \mu_B, \mu_C \sim Normal(0, 10^4)$$

랜덤효과 혼합비교 모형은 각 임상시험별로 상대치료효과가 다를 경우 사용할 수 있다. 랜덤효과 혼합비교 모형의 경우 분산-공분산 구조에 따라 여러 가지 복잡한 모형(Lu and Ades, 2004, 2006)으로 확장이 가능하나, 본 지침서에서는 분산-공분산 구조를 고려하지 않고 짝을 이룬 각각의 상대치료효과들이 서로 독립이라고 가정하는 단순 랜덤효과 혼합비교 모형을 소개한다.

$$\log it(p_{jk}) = \begin{cases} \mu_{jb} & b=A, B, C; \quad \text{if } k=b \\ \mu_{jb} + \delta_{jkk} & b=A, B, C; k=B, C, D \quad \text{if } k > b \end{cases}$$

$$\delta_{AB} \sim Normal(d_{AB}, \sigma^2), \delta_{AC} \sim Normal(d_{AC}, \sigma^2), \delta_{AD} \sim Normal(d_{AD}, \sigma^2)$$

$$d_{tk} = d_{Ak} - d_{Ab}$$

이때 d_{tk} 는 치료 b 에 대한 치료 k 의 상대치료효과인 log odds ratio이며, 사전분포는 모호사전분포(vague prior)로 다음을 사용할 수 있다.

$$d_{AB}, d_{AC}, d_{AD} \sim Normal(0, 10^4)$$

$$\mu_A, \mu_B, \mu_C \sim Normal(0, 10^4)$$

$$\sigma \sim Unif(0, 2)$$

위에서 제시된 베이지안 계층적 모형에 대한 추정을 위해서는 마코프 연쇄 몬테칼로(Markov chain monte Carlo)방법을 사용해야 하는데, 베이지안 모형을 추정할 수 있는 무료 통계패키지인 WinBUGS(<http://www.mrc-bsu.cam.ac.uk/bugs/>)를 사용하여 계산이 가능하다. 참고로 이 패키지는 필요한 요소들을 설치하면 다른 관련 통계패키지인 R에서 불러 와서도 사용이 가능하며, WinBUGS의 기본적인 사용방법에 관하여는 <http://www.mrc-bsu.cam.ac.uk/bugs/>를 참고할 수 있다.

1.2. 혼합비교 사례

금연 프로그램의 효과를 비교하기 위한 AHCPR Smoking Cessation guideline Panel (Fiore et al. 1996, Hasselblad, 1998) 자료를 사용하여 비접촉 (A)에 대한 자기 정신 요법 (B), 개별 상담 (C), 집단상담 (D)의 상대치료효과를 비교하고자 한다. 24개 연구의 치료별 금연 성공 환자들의 분포는 다음 표와 같다.

<i>Baseline</i>	<i>Study number</i>	<i>No contact (A)</i>	<i>Self-help (B)</i>	<i>Individual counseling (C)</i>	<i>Group counseling (D)</i>
$G_{(A)}$	1	9/140		23/140	10/138
$G_{(B)}$	2		11/78	12/85	29/170
$G_{(A)}$	3	79/702	77/694		
	4	18/671	21/535		
	5	8/116	19/146		
	6	75/731		363/714	
	7	2/106		9/205	
	8	58/549		237/1,561	
	9	0/33		9/48	
	10	3/100		31/98	
	11	1/31		26/95	
	12	6/39		17/77	
	13	95/1,107		134/1,031	
	14	15/187		35/504	
	15	78/584		73/675	
$G_{(B)}$	16	69/1,177		54/888	
	17	64/642		107/761	
	18	5/62		8/90	
	19	20/234		34/237	
	20	0/20			9/20
	21		20/49	16/43	
	22		7/66		32/127
$G_{(C)}$	23			12/76	20/74
	24			9/55	3/26

출처: Lu and Ades, 2006

1.2.1. 고정효과 혼합비교 모형

치료 A에 대한 치료 B, C, D의 상대치료효과를 추정하기 위한 고정효과 혼합비교 모형을 BUGS 언어를 사용하여 표현하면 다음과 같다. 여기서 $s[i]$ 는 임상시험번호, $t[i]$ 는 치료, $b[i]$ 는 baseline 치료 A를 나타내며, d 는 log odds ratio를 나타낸다. 24개 임상시험에 대한 baseline과 log odds ratio에 대해서는 모호사전분포를 고려한다.

```
model # 모형 정의
{
  for(i in 1:50)
  {
    r[i] ~ dbin(p[i],n[i])
    logit(p[i])<-mu[s[i]]+ d[t[i]] - d[b[i]]
  }
# Priors for 24 trial baselines
  for(j in 1:24)
  {
    mu[j]~dnorm(0,.0001)}
# Priors for basic parameters
  d[1]<-0
  for (k in 2:4)
  {
    d[k] ~ dnorm(0,.0001) }
```

24개의 임상시험 중 치료 A를 포함하고 있는 19개 임상시험의 baseline parameter의 평균을 baseline parameter의 추정값으로 정의하고, 각 치료의 성공률의 logit을 계산한다.

```

# Absolute log odds on Treatment A based on 19 trials in which it was used
for (i in 1: 50)
{
  mu1[i] <- mu[s[i]]*equals(t[i],1)}
m <- sum(mu1[])/19

# Calculate treatment effects, T[k], on natural scale
for (k in 1:4)
{
  logit(T[k]) <- m + d[k]}

```

베이지안 모형을 사용하여 혼합비교 수행 시, 가장 효과가 좋은 치료에 대한 순위를 다음과 같이 추정할 수 있다.

```

# Rank the treatment effects (with 1=best) & record the best treatment
for(k in 1:4)
{
  rk[k]<- 5 - rank(T[],k)
  best[k]<-equals(rk[k],1)}

```

치료 A에 대한 각 치료의 상대치료효과를 추정하기 위하여, log odds ratio의 차이를 계산하고, 이를 이용하여 odds ratio를 추정한다.

```

# All pairwise log odds ratios and odds ratios
for (c in 1:3)
{
  for (k in (c+1):4)
  {
    lor[c,k] <- d[k] - d[c]
    log(or[c,k]) <- lor[c,k]
  }
}

```

혼합비교를 위한 자료의 입력은 임상시험의 각 치료별로 한 행으로 입력하는데, 여기서 $s[]$ 는 임상시험, $t[]$ 는 치료, $r[]$ 는 금연을 성공한 사람수, $n[]$ 는 각 치료별 환자수, $b[]$ 는 각 임상시험에서의 baseline 치료를 나타낸다.

# 자료				
s[]	t[]	r[]	n[]	b[]
1	1	9	140	1
1	3	23	140	1
1	4	10	138	1
2	2	11	78	2
2	3	12	85	2
2	4	29	170	2
3	1	75	731	1
3	3	363	714	1
4	1	2	106	1
4	3	9	205	1
5	1	58	549	1
5	3	237	1561	1
6	1	0	33	1
6	3	9	48	1
7	1	3	100	1
7	3	31	98	1
8	1	1	31	1
8	3	26	95	1
9	1	6	39	1
9	3	17	77	1
10	1	79	702	1
10	2	77	694	1
11	1	18	671	1
11	2	21	535	1
12	1	64	642	1
12	3	107	761	1
13	1	5	62	1
13	3	8	90	1
14	1	20	234	1
14	3	34	237	1
15	1	0	20	1
15	4	9	20	1
16	1	8	116	1
16	2	19	149	1
17	1	95	1107	1
17	3	143	1031	1
18	1	15	187	1
18	3	36	504	1
19	1	78	584	1

19	3	73	675	1
20	1	69	1177	1
20	3	54	888	1
21	2	20	49	2
21	3	16	43	2
22	2	7	66	2
22	4	32	127	2
23	3	12	76	3
23	4	20	74	3
24	3	9	55	3
24	4	3	26	3
END				

베이지안 모형의 각 모수에 대한 초기값은 아래와 같이 임의로 입력할 수도 있고, WinBUGS 프로그램에서 자동으로 생성할 수도 있다. 그리고 두 가지 이상의 초기값을 사용하여 다중체인을 사용하면 겔만-루빈의 수렴여부를 진단할 수 있다.

```
# 초기값 1
list(d=c(NA,0,0,0),mu=c(0,0,0,0,0, 0,0,0,0,0, 0,0,0,0,0, 0,0,0,0,0))

# 다중연쇄를 위한 초기값 2
list(d=c(NA,0.1,-1,-0.2),mu=c(1,-1,-2,0,0,-2,1,0,2,2,1,-1,-2,0,0,-2,1,0,2,2,-2,-0.5,-3,0.5))
```

위 고정효과 혼합비교모형을 WinBUGS 프로그램에서 수행한 결과는 다음과 같다. 이때 MCMC를 사용하여 10,000개의 표본을 생성하고 초기값의 영향을 제거하기 위하여 처음 5,000개의 표본을 "burn-in" 과정으로 제외하였다.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
best[1]	0.0	0.0	1.414E-12	0.0	0.0	0.0	5001	5000
best[2]	0.0	0.0	1.414E-12	0.0	0.0	0.0	5001	5000
best[3]	0.3472	0.4761	0.01353	0.0	0.0	1.0	5001	5000
best[4]	0.6528	0.4761	0.01353	0.0	1.0	1.0	5001	5000
d[2]	0.2241	0.1259	0.003829	-0.01829	0.2207	0.4738	5001	5000
d[3]	0.767	0.05918	0.002037	0.6527	0.7685	0.8803	5001	5000
d[4]	0.8377	0.1752	0.006183	0.4942	0.8398	1.18	5001	5000
or[1,2]	1.261	0.1597	0.004883	0.9819	1.247	1.606	5001	5000
or[1,3]	2.157	0.1274	0.004383	1.921	2.157	2.412	5001	5000
or[1,4]	2.347	0.4159	0.01463	1.639	2.316	3.253	5001	5000
or[2,3]	1.736	0.2293	0.006691	1.325	1.724	2.226	5001	5000
or[2,4]	1.881	0.3649	0.01182	1.281	1.845	2.691	5001	5000
or[3,4]	1.09	0.1914	0.006411	0.7651	1.071	1.507	5001	5000

위 결과에 위하면 치료 A에 대한 치료 B의 OR은 1.261, 치료 A에 대한 치료 C의 OR은 2.157, 치료 A에 대한 치료 D의 OR은 2.347이며, 네 가지 치료중에서 가장 금연 성공률이 높은 치료는 치료 D임을 알 수 있다.

1.2.2. 단순 랜덤효과 혼합비교 모형

각 임상시험별 상대치료효과 ($\delta_{AB}, \delta_{AC}, \delta_{AD}$)가 서로 독립이라고 가정하는 단순 랜덤 효과 혼합비교 모형을 사용해서 치료 A에 대한 치료 B, C, D의 상대치료효과를 추정하고자 한다. 이때 모형정의 및 초기값 입력에 대한 BUGS 언어는 다음과 같으며, 상대치료효과 추정 및 치료효과의 순위, 자료의 입력 부분은 고정효과 혼합비교 모형과 동일하다.

모형정의 부분에서 상대치료효과 delta에 대해 사전분포를 정의하고, 사전분포가 가지는 초모수(hyperprior) $d_{AB}, d_{AC}, d_{AD}, \sigma^2$ 에 대해 초사전분포(hyperprior distribution)를 모호사전분포로 정의한다.

```

model # 모형 정의
{
  for(i in 1:50){
    r[i] ~ dbin(p[i],n[i])
    logit(p[i])<-mu[s[i]]+ delta[i]*(1-equals(t[i],b[i]))

# Random effects model for log-odds ratios
    delta[i] ~ dnorm(md[i],prec)
    md[i] <- d[t[i]] - d[b[i]]
  }

# priors for 24 trial baseline
for(j in 1:24){mu[j]~dnorm(0,.0001)}

# priors for between-study SD
prec <- 1/(sd*sd)
sd~dunif(0,2)

# priors for basic parameter
d[1]<-0
for (k in 2:4){d[k] ~ dnorm(0,.0001)}

```

페이지안 계층적 모형에서 추가된 모수 delta와 sd에 대한 초기값을 추가로 입력한다.

[illegible]

위 단순 랜덤효과 혼합비교 모형을 WinBUGS 프로그램에서 수행한 결과는 다음과 같다. 이때 MCMC를 사용하여 10,000개의 표본을 생성하고 초기값의 영향을 제거하기 위하여 처음 5,000개의 표본을 "burn-in" 과정으로 제외하였다.

node	mean	sd	MC error 2.5%		median	97.5%	start	sample
best[1]	0.0	0.0	1.414E-12	0.0	0.0	0.0	5001	5000
best[2]	0.0484	0.2146	0.00318	0.0	0.0	1.0	5001	5000
best[3]	0.17	0.3756	0.006685	0.0	0.0	1.0	5001	5000
best[4]	0.7816	0.4132	0.007419	0.0	1.0	1.0	5001	5000
d[2]	0.5258	0.3813	0.0081	-0.2113	0.5275	1.29	5001	5000
d[3]	0.8139	0.2314	0.005773	0.3645	0.8115	1.289	5001	5000
d[4]	1.172	0.4339	0.009454	0.3196	1.159	2.059	5001	5000
or[1,2]	1.82	0.7327	0.01509	0.8095	1.695	3.632	5001	5000
or[1,3]	2.319	0.5556	0.01387	1.44	2.251	3.629	5001	5000
or[1,4]	3.552	1.667	0.03591	1.377	3.187	7.836	5001	5000
or[2,3]	1.443	0.5981	0.01336	0.6084	1.335	2.912	5001	5000
or[2,4]	2.145	1.132	0.02508	0.7588	1.887	5.095	5001	5000
or[3,4]	1.567	0.7077	0.01654	0.6104	1.425	3.285	5001	5000
sd	0.8112	0.1777	0.006029	0.5291	0.7911	1.227	5001	5000

위 결과에 위하면 치료 A에 대한 치료 B의 OR은 1.82, 치료 A에 대한 치료 C의 OR은 2.319, 치료 A에 대한 치료 D의 OR은 3.552이며, 네 가지 치료 중에서 가장 금연 성공률이 높은 치료는 치료 D임을 알 수 있다.

[2] PBAC | 교신 자료



SEOUL NATIONAL UNIVERSITY
SAN 56-1, SHILIM-DONG, KWANAK-GU, SEOUL, 151-742, KOREA

The Secretary
Pharmaceutical Benefits Advisory Committee
GPO Box 9848
Canberra ACT 2601
Australia

10 December 2010

Dear Sir/Madam,

RE: Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3)

I am writing this letter in asking HIRA to communicate with your relevant party with regard to the above guideline. I am the principle investigator who is in charge of a project commissioned by the Korean Health Insurance Review & Assessment Service (HIRA) to develop a step-by-step instruction plan to present an indirect comparison of head-to-head trials in preparation for submissions to the Drug List Division for drug benefit coverage.

In building up the idea, the project team has primarily reviewed a few existing guidelines to gain an understanding of what recommendations other relevant authorities have provided from their perspectives. We have found the guideline offered by the Australian Pharmaceutical Benefits Advisory Committee very helpful as it is fairly specific with clear descriptions. We closely reviewed the background arguments and the actual guideline. They were particularly helpful in terms of how they highlighted the assumption of exchangeability with the provision of a specific guide for procedural validation. We have therefore developed our draft guideline by utilizing the basic framework of the PBAC guideline, although there are also some areas that we could not adopt from our point of view. For example, we did not limit performing indirect comparisons only to dichotomous outcomes or recommend any particular type of measure, such as RR or OR. We have nevertheless highlighted possible associated problems.

As we have borrowed some ideas from several exiting guidelines and PBAC's was one of the main references, it would be useful to have some feedback from the relevant department of PBAC before the HIRA disseminate the guideline for actual application. The HIRA will be willing to provide a copy of the report we have prepared for them, if you request, in which we suggested the contents of the forthcoming HIRA guideline. The report contains chapters regarding definitions of relevant terminology, assumptions for indirect comparison, considerations required when indirectly synthesizing a comparative effect estimate, step-by-step instructions for the preparation of a submission involving an indirection comparison with some illustrative demonstration. Although the draft is written in Korean, we presume that the diagrams and tables can give one a general skeletal idea of the outline.

A handwritten signature in black ink, appearing to read 'Seokyoung Hahn'.

Seokyoung Hahn, PhD
Seoul National University College of Medicine
28 Yongon-dong Chongno-gu Seoul 110-744, Korea
Tel: +82-2-740-8911 Fax: +82-2-743-8361
Email: hahns@snu.ac.kr

[3] 비뚤림 위험 평가 지침

무작위배정순서가 적절한 방법에 의해 생성되었는지 여부

- 예
- 난수표 이용 또는 컴퓨터를 이용한 난수 생성
 - 동전던지기, 카드섞기(꺼낸 카드는 다시 집어넣어야함), 주사위 던지기, 심지뽑기 등 사용. 그러나 i난수i임이 보장되는 수행과정 확인할 수 있어야 적절한 방법으로 볼 수 있음. 예를 들어, 동전을 던져서 앞면이 나오면 뒷면이 나올 때 까지 다시 던지지 않았음을 확인할 수 있어야 함 (배정순서 은폐와 연결됨).
 - 최소화법은 엄밀히 말해서 무작위로 순서가 배정 되는 것이라고 볼 수는 없으나 적절한 과정에 의해 수행된 최소화 법은 제대로 수행된 무작위배정으로 간주함.
-

- 아니오
- 무작위방법을 시행하지 않았거나 부적절한 방법의 사용
예를들어,
- 환자 등록번호 또는 병록번호의 홀수 짝수 등 규칙을 이용한 배정
 - 생년월일, 내원일 등의 규칙을 이용한 배정
 - 임상가의 판단에 따른 배정
 - 배정자가 임의로 배정
 - 환자의 선호도에 따른 배정
 - 검사결과에 의한 배정
 - 검사결과 순 또는 약제가 준비되는 순 등 이용가능 순에 의한 배정
-

불확실 무작위배정방법에 대한 언급이 없거나 모호

배정순서가 적절히 은폐되었는지 여부

- 예
- 적절한 방법에 의해 배정순서가 은폐됨으로써 연구자가 배정내용을 알 수 없어야 함
예를 들어,
- 독립적인 중앙 무작위배정 및 관리(웹기반, 전화, 제 삼의 관리기관에 의한 무작위배정 통제 등)
-

	○ 무작위배정순서에 의해 일련번호가 기록되어 있는 동일한 모양의 포장 사용 ○ 일련번호가 기록된 불투명하고 봉해진 봉투에 의한 배정순서 보관 및 개봉
아니오	배정순서가 은폐될 수 있는 방법을 사용하지 않았거나 부적절한 방법의 사용에 의해 배정순서가 은폐되지 않음 예를 들어, ○ 난수 또는 무작위배정순서가 기재된 표를 이용한 무작위 이행 ○ 밀봉되지 않거나 투명하거나 일련번호가 없는 등 안전장치가 없는 무작위배정 봉투를 사용 ○ 교대 혹은 순환법 등의 순서를 사용 ○ 생일, 병록번호 등을 이용
불확실	은폐방법이 기술되어 있지 않거나 구체적이지 않아서 판단하기 어려운 경우, 예를 들어서 봉투에 의해 배정되어 있다고 했으나 일련번호, 밀봉, 투명 여부에 대한 기술이 없을 때
참여자, 연구자, 결과평가자에 대한 눈가림이 적절한지 여부	
예	○ 결과의 비뚤림을 최소화할 수 있는 눈가림법을 채택하여 수행하였고 피험자와 연구자에 대한 눈가림이 깨지지 않았을 것으로 확인되는 경우 ○ 피험자와 연구중재수행자에 대해서는 눈가림을 사용하지는 않았으나 결과평가자에 대한 눈가림이 이루어졌고 이로써 결과 측정과 평가에 비뚤림을 초래하지 않을 것으로 판단할 수 있는 경우 ○ 의료장비에 의한 객관적인 방법으로 결과변수를 측정하는 경우 등 눈가림이 결과의 측정과 평가에 영향을 미치지 않는 경우
아니오	○ 눈가림이 결과에 영향을 미칠 수 있는 경우임에도 눈가림법을 시행하지 않았거나, 눈가림을 시도하였으나 방법이 부적절하고 결과측정에 영향을 받았을 것으로 판단되는 경우 ○ 피험자와 연구자에 대한 눈가림이 시도되었으나 눈가림이 유지되지 않았을 것으로 판단되는 경우
불확실	눈가림에 대한 정보가 충분하지 않은 경우

분석을 위한 결과자료가 충분히 제공되는지 여부

- 예
- 결측치가 없는 경우
 - 결측치가 결과에 영향을 미치지 않는 경우(생존분석에서는 결측이 절단값으로서 다루어짐)
 - 결측치가 중재 군간에 유사하게 발생하고 결측치가 생긴 이유도 유사함
 - 이분형 변수의 경우 결측치 분율이 관찰도수를 비추어볼 때 중재효과 추정에 임상적으로 유의한 차이를 낼 것으로 보이지 않는 경우
 - 연속변수의 경우 결측값 들로부터 예견되는 중재효과의 크기가 관찰된 효과의 크기 추정에 임상적으로 유의한 영향을 미칠 것으로 보이지 않는 경우
 - 적절한 통계적 방법을 사용하여 결측치를 대체한 경우

- 아니오
- 상당수의 결측치가 존재하고 실제 결과에 영향을 미칠 수 있는 경우 - 중재군 간의 불균형한 결측치 수 차이 자체 또는 결측이 생긴 이유가 결과에 비뚤림을 초래할 수 있는 경우
 - 이분형 변수의 경우 결측치 분율이 결과변수의 관찰 도수에 비추어 상당수일 때
 - 연속변수의 경우, 결측결과들로부터 예견되는 군간 중재효과 차이가 (평균의 차이 혹은 표준화 평균의 차이)가 효과크기 추정결과에 임상적으로 유의한 비뚤림을 초래하기에 충분한 경우
 - 무작위배정된 중재를 받지 않은 사람이 상당수 임에도 중재받은대로만 분석을 수행하여 결과자료를 제시한 경우
 - 부적절한 방법으로 결측치를 대체한 경우

- 불확실
- 배제/탈락에 대한 보고가 불충분한 경우(예, 무작위수 언급 없음, 결측 이유에 대한 언급 없음)

분석에 사용될 자료가 선택적 보고에 따른 결과일 가능성이 없는지 여부

- 예
- 연구 프로토콜이 존재하여 연구에서 미리 정의해놓은 일차, 이차 결과변수들의 정의 및 분석이 미리 정해진 방법대로 다루어졌음을 확인할 수 있는 경우
 - 연구 프로토콜은 없으나 보고된 정보를 미루어 보고된 결과자료가 사전
-

	에 예정된대로 수행된 분석결과임을 확신할 수 있는 경우
아니오	메타분석에 고려하고자 하는 결과변수에 대한 다음 경우에는 선택적 보고가 이루어졌을 가능성이 있음: ○ 미리 정해진 결과변수이었음에도 결과가 보고되지 않은 것이 있는 경우 ○ 보고된 결과 중 미리 정해지지 않은 측정, 분석방법, 소그룹 분석으로 이루어진 것이 있는 경우(이런 분석과 보고가 이루어진 데 대한 명백한 이유와 설명-예들 들어, 예상치 못한 부작용 등-이 있는 경우는 예외) ○ 불완전한 결과보고로 인해 분석에 포함시킬 수 없는 경우 ○ 현 연구에서 당연히 분석되었을 것으로 보이는 핵심결과에 대한 보고가 없는 경우
불확실	i예i, i아니오에 대한 판단을 위한 정보가 충분하지 않은 경우 (대다수의 연구들이 이 범주에 포함될 가능성이 있음)
비뚤림 위험을 높일 수 있는 다른 명백한 문제는 없는지 여부	
예	다른 비뚤림의 원인이 없는 것으로 보임
아니오	아래와 같은 경우에 해당될 때 추가 비뚤림의 위험이 있는 것으로 판단 ○ 연구 설계와 관련된 특정 비뚤림 위험이 있음 ○ 결과의 방향에 따라 연구가 조기에 중단되었음 ○ 피험자의 기저상태에 군간 심각한 불균형이 있음 ○ 연구에 부정이 있었다는 주장이 제기된 바 있음
불확실	추가 비뚤림 가능성에 대한 여지가 있으나 비뚤림의 위험이 어느 정도일지 평가할만한 충분한 정보나 근거가 없는 경우
	○ 예: 비뚤림의 위험이 낮음 ○ 아니오: 비뚤림의 위험이 높음 ○ 불확실: 비뚤림에 대한 위험을 판단하기 어려움. 불확실한 경우는 i아니오i의 경우에 준하여 다루는 것이 권고됨

[] 간접비교 사례 제시 문헌