

경제성평가의 국가간 적용가능성 진단도구를 이용한 류마티스관절염 치료제 사례 연구

구혜민 · 권정미 · 박세영 · 강숙현 · 이의경

숙명여자대학교 임상약학대학원

(2009년 8월 8일 접수 · 2009년 9월 15일 수정 · 2009년 9월 30일 승인)

A Case Study using Transferability Diagnosis Tools Between Countries for Economic Evaluations of Rheumatoid Arthritis Drugs

Hyemin Ku, Jeongmi Kwon, Seyoung Park, Sukhyun Kang, and Eui Kyung Lee

Graduate School of Clinical Pharmacy, Sookmyung Women's University, Seoul, Korea

(Received August 8, 2009 · Revised September 15, 2009 · Accepted September 30, 2009)

Objectives: With more and more health economic evaluations being performed in many countries, the question of whether the results from the foreign studies can be transferred to other country is becoming important. The objective of this study was to conduct a case study using three transferability diagnosis tools for economic evaluation on rheumatoid arthritis.

Methods: Three diagnosis tools were used to assess transferability: Welte's tool, Boulenger's and Urdahl's. Five researchers who have experience on economic evaluation made consensus on the evaluation through the semi-delphi method.

Results: With Welte's tool, absolute and relative prices in healthcare, practice variation, health-status preference, productivity and work-loss time were evaluated as biased. The transferability information score by Boulenger's was 85.9%, which means good quality reporting. In case of Urdahl's, research questions were well defined and reporting was transparent and explicitly stated. However, both the relevance of data inputs to Korea and robustness of model were relatively low. In conclusion the UK study on rheumatoid arthritis could be partially transferable to Korea, and will need modeling-based adjustments.

□ Key words - Economic Evaluation, Transferability, Diagnosis, Rheumatoid Arthritis

선별 등재 제도가 도입된 이후 신약에 대한 경제성 평가 자료 제출이 의무화되었다. 이에 따라 의약품 경제성 평가에 대한 수요는 증가하고 있으나 경제성 평가의 모델 구축 등 심도 있는 연구를 수행할 수 있는 인력은 부족한 상태이다. 더욱이 경제성 평가 수행을 위해서는 많은 비용과 시간을 필요로 하므로 해당 기업으로서는 일정 부분 부담을 갖게 되는 것이 사실이다.

이러한 문제를 일부 해결할 수 있는 방안이 외국에서 수행한 경제성 평가 결과를 활용하는 방안이다. 예컨대 우리나라

라보다 일찍 경제성 평가 제도를 도입한 호주나 영국, 캐나다, 그리고 신약 개발을 가장 많이 하는 미국 등에서는 이미 신약 개발 및 보험 등재 과정에서 경제성 평가를 수행하여 의사 결정에 활용한 경우가 많이 있다. 따라서 우리나라의 경우 외국에서 이미 실시한 경제성 평가 결과를 일부 혹은 전부 활용하는 방안을 논의할 필요가 있다.

이미 외국에서도 다른 국가에서 수행한 경제성 평가 연구 결과를 해당 국가에 활용할 수 있는지 여부에 대하여 많은 논의가 있어 왔다.¹⁾ 선행 연구에 따르면 국가간 인구학적 차이, 역학 자료의 이질성, 의료 이용 및 치료 패턴, 가격 및 비용의 차이 등으로 상당 부분 한계가 있기는 하나,²⁾ 경제성 평가의 일부 항목에 대해서는 외국의 연구를 해당 국가에 적용하는 것이 가능한 것으로 알려져 있다.

이러한 적용가능성(transferability)을 국제의약품경제성평가학회인 ISPOR(International Society of Pharmacoeconomics & Outcomes Research)에서는 “어느 setting에서 진행된 연구

Correspondence to : 이의경

숙명여자대학교 임상약학대학원
서울특별시 용산구 청과동 2가 53-12
숙명여자대학교 약학대학 307호 (140-742)
Tel: +82-2-710-9799, Fax: +82-2-6395-1214
E-mail: ekyung@sm.ac.kr

의 결과를 특정 setting에 적용할 수 있고 참값을 보이는 정도”로 정의하고 있다. 이때 이와 유사하게 사용되는 용어인 일반화(generalizability)는 연구된 국가에서의 결과가 다른 국가에 수정이나 보정 없이 그대로 사용 가능한 것을 의미하여 적용가능성과는 차이를 보인다.

선행 연구를 살펴보면 외국의 경제성 평가 결과의 적용가능성을 측정하는 도구로써 3가지 방법이 제시되고 있는데 Welte 방법, Boulenger/EURONHEED 방법, 그리고 Urdahl 방법이 있다.³⁾⁶⁾ Welte 측정 방법은 2004년에 pharmacoeconomics에서 처음 소개된 방법으로 의사결정도(decision chart)에 따라서 진행되되, 정해진 체크리스트를 이용하여 각 요인간의 중요도(relevance)와 일치도(correspondence)를 바탕으로 도입 국가에서의 비용효과비(cost effectiveness ratio: CER)에 미치는 영향 정도를 평가하는 방식이다. 만약 양국간 해당 요인이 높은 일치도를 보여 도입국가에서의 CER에 영향을 미치지 않게 된다면(unbiased) 이는 적용가능성이 높아진다는 것을 의미한다. 평가를 완료하게 되면 이후 모델 보정(modeling adjustments)이 필요한지, 필요하다면 기존에 연구를 수행했던 국가의 자료를 어느 정도 이용할 수 있는지에 따라 적용가능성 여부와 그 정도를 평가하게 된다.

Boulenger 측정 방법은 2005년에 European Journal of Health Economics에서 처음 소개된 방법으로 11개의 큰 카테고리 구분된 42개 항목의 체크리스트를 이용하여 적용가능성 정도를 점수화하는 방법이다. 이 방법은 연구의 높은 질과 명확한 서술을 평가하여 적용가능성의 정도를 평가하는데 그 목적이 있다.

Urdahl 측정 방법은 2006년에 pharmacoeconomics에 처음 소개된 방법으로 4개의 큰 카테고리 구분된 직접적인 서술로써 평가하는 방법이다. 4가지 평가 항목은 첫째, 경제성 평가를 리뷰하고 결정할 결정권자 및 환경에 대한 서술, 둘째, 모형 및 가정에 대한 자세한 보고 및 서술, 셋째, 도입 국가에서 사용할 자료의 관련성, 넷째, 민감도 분석 결과의 확인이다.

따라서 본 연구에서는 위의 3가지 적용가능성 진단 도구를 비교하고자 하며, 이를 위하여 사례연구(case study)로서 외국에서 수행한 류마티스관절염 치료제 경제성 평가연구를 대상으로 한국으로의 적용가능성을 평가하였다. 류마티스관절염은 면역 질환의 일종으로 장기적인 치료를 필요로 하는 질환이다. 특히 류마티스관절염 치료제인 etanercept 등의 TNF-알파 제제들은 다른 치료제들에 비하여 매우 고가의 약물이다. 해당 치료제에 대한 보험급여 결정은 건강보험재정 관리 측면에 있어서 매우 중요한 문제로서 경제성 평가가 활발히 이루어지고 있는 질환 중 하나이다. 따라서 etanercept에 대한 경제성 평가 연구에 대하여 위에서 언급한 3가지 진단 도구를 이용하여 국내 도입의 가능 여부 및 한계점과 적용가능성의 정도를 평가하고 이를 통하여 3가지 적용가능성 도구의 장단점 등을 고찰하고자 한다.

연구 방법

경제성평가연구의 국가간 적용가능성 진단 도구들을 비교하기 위하여 우선 사례연구를 수행하기 위한 논문을 선정하였다. 대상은 류마티스관절염 치료제인 etanercept에 대한 경제성평가 연구로서, 2004년 이후에 외국에서 수행된 경제성평가 연구로 제한하였다. 성과변수는 ACR20 및 HAQ (Health Assessment Questionnaire)은 포함하였고, DAS (Disability Assessment Score)는 국내에 측정하는 경우가 적기 때문에 제외하였다. 국가 및 연구자는 류마티스관절염 경제성평가 연구 논문의 빈도가 높은 경우를 선정하였다. 이상의 기준에 따라 영국의 Brennan⁷⁻⁸⁾(2004) 논문을 선정하였는데, 본 연구에서는 영국 BSR(the British Society for Rheumatology) 가이드라인에 따라 최소 6개월 이상 methotrexate를 포함한 두 가지 이상의 DMARD(Disease Modifying anti-Rheumatic Drugs)에 실패한 환자를 대상으로 etanercept와 다른 DMARD 간의 경제성 평가를 실시하였다.

다음으로 영국에서 수행된 경제성평가 연구의 국내 적용가능성을 측정하는 도구로써 Welte, Boulenger/EURONHEED 그리고 Urdahl 방법 등 3가지 방법을 이용하였다.

Welte 방법의 경우 ‘적용가능성 의사결정도(transferability decision chart)’에 따라서 평가가 이루어지는데 먼저 3가지의 일반적인 실격기준 (general knock-out criteria) 질문에 의해 실격 여부를 결정한다. 질문 내용에 따라 평가 기술, 비교 대안군, 연구의 질 등을 평가하였다. 일반적인 실격기준 (general knock-out criteria)을 통과한 이후, 그 다음 단계로서 구체적 실격기준 (specific knock-out criteria)을 평가하였다. 구체적 실격기준은 방법론 특성 (methodological characteristic), 보건의료체계 특성 (healthcare system characteristic), 인구 특성 (population characteristic)의 3가지 큰 카테고리 구분되어 있는데, 구체적으로는 총 14가지의 항목을 평가하였다.

각 14가지 항목에 대하여 영국에서 수행된 연구 결과를 우리나라에 적용할 수 있는지 여부를 검토했는데, 첫째 영국 연구의 국내 적용가능성을 평가하는데 있어서 각 항목이 얼마나 중요한 지 그 중요도(relevance for transferability)를 평가하였다. 둘째로는 영국에서 수행한 연구 결과가 우리나라 상황과 어느 정도 일치도 (correspondence)를 보일지 여부를 평가하였다. 각 항목별 중요도 및 일치도는 “매우 높음(very high)”, “높음(high)”, “보통(medium)”, “낮음(low)”, “매우 낮음(very low)”, “관련 없음(not relevant)”으로 평가하였다. 그 다음으로는 각 항목별 중요도 및 일치도 평가 결과를 바탕으로 우리나라에서의 비용효과비(CER)에 미치는 영향을 평가하였는데, 파급영향 정도는 다음의 4가지 중 하나를 선택하였다: “영향을 미치지 않음(unbiased)”, “영향을 약간 미침 (slightly high or slightly low)”, “영향을 어느 정도 미침 (high or low)”, “영향을 많이 미침 (too high or too low)”.

Boulenger 측정 방법에서는 각 42개의 9항목에 대해 긍정

(yes), 부분 긍정(partially), 부정(no) 혹은 정보 없음(no information), 해당사항 없음(not applicable)의 4가지 중 하나로 평가하였다. 평가 이후 yes는 1점, partially는 0.5점, no 혹은 no information은 0점을 적용하여 적용가능성 점수를 계산하였다. 점수는 다음과 같은 공식으로써 산출하였으며, 백분율로써 제시하였다: $(1/n-x) * \sum_i (Six100)$. 여기서 n은 질문의 개수이며, x는 not applicable로 응답한 질문의 개수이다. 1는 1부터 n까지를 의미하며, S는 각 질문에 대한 점수이다.

Urdahl 측정 방법은 4개의 큰 카테고리로 구분된 체크리스트에 대하여 긍정(yes), 부정(no) 혹은 직접적인 서술로써 평가하는 방법이다. 4개의 카테고리 내에 세부 질문에 대하여 평가하였으며, 각 질문에 대한 답변을 종합하여 연구자가 최종적으로 적용가능성을 평가하였다. 4가지 평가 항목은 첫째, 연구질문의 정의에 대한 서술 양상, 둘째, 모형 및 가정에 대한 자세한 보고 및 서술 양상, 셋째, 영국 논문에서 사용된 데이터를 국내에 사용하기 위한 관련성, 넷째, 양 국가간 데이터 차이에도 불구하고 모델의 강건함(robustness)이 민감도분석을 통하여 입증되는지에 대한 평가였다.

적용가능성 평가는 5인의 연구진이 준델파이방법(semi-delphi)을 통하여 실시하였다. 5인의 연구진은 모두 경제성 평가에 대한 실제 경험이 있는 자들로써 약학부문의 전문적 소양을 갖추었으며, 병원약국에서 근무한 경험이 있는 자 2명, 제약기업에서 연구개발 근무 경험이 있는 자 2명, 전문연구직 종사자 1명으로 구성되었다. 이상의 5명 연구진은 각자 개별적으로 3개의 진단도구를 이용하여 국가간 적용가능성을 평가하였으며, 그 결과를 모두 취합하였다. 취합된 평가 결과를 바탕으로 각 진단도구별로 토론하였으며, 토론과정에서 본인의 평가 내용을 수정하고 합의를 도출하여 최종 평가 결과를 정리하였다.

연구 결과

Welte 도구를 통한 진단 결과

Welte 도구를 통한 진단의 첫번째 단계는 일반적 실격 기준(general knock out criteria)에 대한 평가로서, 첫째, 평가 약물이 도입 국가에서 사용하는 것과 다른지, 둘째, 비교 대안이 도입 국가에 등록되어 있지 않은지, 셋째, 연구가 질적으로 우수하지 않은지에 대해서 평가하였으며 위의 세가지 실격 기준 관련 질문에 모두 해당하지 않아(Table 1) 적용가

능성 의사결정도(transferability decision chart)에 따라 구체적 실격 기준(specific knock-out criteria) 평가로 진행하였다.

적용가능성과 관련한 각 14가지 항목에 대하여 중요도와 일치도를 평가한 결과, 만약 해당 항목을 그대로 적용 시 우리나라에서의 비용효과비에 영향을 미칠 것으로 평가된 항목은 비용과 관련된 전반적인 항목과 할인율, 생산성 손실, 치료방법 및 건강 상태에 대한 선호도였다 (Table 2).

우선 방법론적 특성(methodological characteristics)에 해당하는 관점(perspective), 할인율(discount rate), 의료 비용(medical cost) 산출 방법, 생산성 비용(productivity cost approach) 산출방법 항목들 중에서 ‘관전’만이 편향되지 않은 ‘unbiased’라는 결과를 도출하였다. Brennan의 연구는 류마티스관절염 환자를 대상으로 한 장기간(lifetime) 연구로 관점에 따라 비용의 차이가 많이 발생할 수 있기 때문에 관점의 중요도가 매우 높다(very high)고 평가하였는데, 본 논문과 우리나라 모두 경제성 평가의 관점이 사회적 관점으로 동일하기 때문에 일치성 또한 매우 높다(very high)고 평가하였고 그 결과 CER에 미치는 영향은 unbiased로 결론을 내렸다.

할인율은 우리나라에서는 효과, 비용에 대하여 모두 5%의 할인율을 기본 적용하지만, Brennan 연구의 시험 국가에서는 NICE 가이드라인을 근거로 비용에는 연간 6%, 효과에는 연간 1.5%의 할인율을 적용하였다.⁹⁾ 관점과 마찬가지로 할인율의 중요도는 매우 높으나(very high) 할인율에 대한 양국간의 차이가 크기 때문에 일치성은 낮은(low) 것으로 평가하였고, 이것은 도입 국가의 CER에 많은 영향을 미칠 것으로 판단하였다.

의료비용 산출 방법론 항목과 관련하여 영국의 Brennan 연구에서는 약값등은 직접적으로 비용 자료를 활용한 반면, 입원비 등에 대해서는 HAQ(Health Assessment Questionnaire) 값을 이용하여 추정하였다. 즉 HAQ 점수가 1점 악화될 때마다 직접 보진 의료 비용이 연간 860파운드가 증가하는 것을 모델에 적용하였다. 이러한 비용산출 방법론은 국내 적용성을 평가하는데 신중히 고려되어야 하므로 중요도는 ‘very high’로 평가하였고, 국내와의 일치도는 ‘low’로 하여 종합판정은 상당한 편향(too high~too low)로 결정하였다.

생산성 비용산출 방법론 항목에서는 평가항목의 중요도 및 양국간의 일치도 모두 보통(medium)으로 평가하여 한국의 비용효과비에는 약간의 영향(slightly high or low)만을 미칠 것으로 판단하였다. 이러한 판단은 Brennan의 연구에서 생산

Table 1. Transferability estimation by general knock-out criteria of Welte checklist

General knock-out criteria	assessment
1. The evaluated technology is not comparable to the one that shall be used in the decision country	Pass
2. The comparator is not comparable to the one that is relevant to the decision country. An example is a comparator drug that is not licensed in the as decision country.	Pass
3. The study does not possess an acceptable quality, i.e. it does not live up to the standards required in the decision context, e.g. there is double counting of costs.	Pass

Table 2. Transferability estimation by specific knock-out criteria of Welte checklist

	Transferability factor	Estimated relevance	Estimated correspondence between study and decision country	Estimation of CER of decision country based on CER of study country is:
Methodological characteristics	Perspective	Very high	Very high	Unbiased
	Discount rate	Very high	Low	Too high ~ too low
	Medical cost	Very high	Low	Too high ~ too low
	Productivity cost approach	Medium	Medium	Slightly high or low
Healthcare system characteristics	Absolute and relative prices in healthcare	Very high	Low	Too high ~ too low
	Practice variation	High	Medium	high or low
	Technology availability	Not relevant	NA	Unbiased
Population characteristics	Incidence/prevalence	Medium	high	Unbiased
	Case-mix	High	Very high	Unbiased
	Life expectancy	Medium	Very high	Unbiased
	Health-status preference	Very high	Medium	high or low
	Acceptance, compliance, incentive to patients	Not relevant	Medium	Unbiased
	Productivity and work loss time	Medium	Medium ~ low	high or low
	Disease spread	Not relevant	NA	Unbiased

성 비용은 기본 분석이 아닌 민감도 분석에만 들어갔다는 점, 생산성 산출과정에서 HAQ 점수에 따라 고용 상태를 구분하여 생산성손실비용을 산출하였는데, HAQ에 따른 생산성 손실경향은 양 국가간 유사할 것이라는 판단에 근거하였다.

보건의료체계특성 (healthcare system characteristics) 항목들 중에는 ‘보건의료관련 절대적 및 상대적 가격(absolute and relative price in healthcare)’, ‘진료 변이 (practice variation)’에서 적용가능성이 낮은 것으로 판정되었다.

우선 ‘보건의료관련 절대적 및 상대적 가격(absolute and relative price in healthcare)’의 경우, 영국과 한국의 보건의료 부문의 가격에는 상당한 차이가 있으므로 적용가능성 평가 항목으로서 매우 중요하며, 일치도가 낮은 것으로 평가하였다.

류마티스관절염 치료에서는 치료 순서 (treatment sequence)가 중요한데 환자의 특성 및 의사 특성, 국가에 따라 어느 정도 차이가 있으므로 진료 변이 (Practice variation)가 경제성평가에서 중요한 요인이 된다. 따라서 적용가능성 평가 항목으로서의 중요도는 높다고 평가하였다. 그러나 진료가이드라인은 국가별로 어느 정도 일치성을 보이는데, 예컨대 조기 환자에게는 DMARD를 투약하고, 여기에 반응성이 없을 경우 etanercept와 같은 생물학적 DMARD를 사용한다는 측면 등 큰 골격은 국가간 거의 유사하다. 즉 일정 수준까지는 주요 외국의 가이드라인이 존중되는 측면이 있기 때문에 ‘medium’으로 판단하였다.

기술 보유 (technology availability) 항목은 의학기술이나 검진 장비 등의 보급과 수준 등의 차이로 인해서 질병 치료

에 영향을 주는지 여부를 판단하는 것이다. 예컨대 국가간 신기술의 초기 보급 속도에 차이가 있으면 질병 확진 및 치료에 큰 차이가 나타날 수 있으며, 검진보급율이 낮을 경우 검진 등을 받기 위해 병원까지 먼 거리 여행을 하면 생산성 손실이나 교통비등이 많이 들 수 있다. Brennan의 연구에서는 시술, 수술, 특수 검사 등과 관계가 적기 때문에 중요도와 일치도에 해당사항이 없으므로 평가하고 우리나라의 비용 효과비에 영향을 미치지 않을 것으로 평가하였다.

인구특성 (population characteristics) 건강상태선호 (health-status preference)와 생산성손실 (productivity and work-loss time) 항목에서 적용가능성이 낮은 것으로 평가되었다. 우선 유병률 항목에 있어서 류마티스관절염은 유병률은 0.8~1%로 양국간 차이가 거의 없다고 판단하였다. 환자구성(case-mix) 항목에서는 류마티스관절염은 유전적 질환으로서 성별, 나이, 유전과도 관련이 있기 때문에 case-mix와는 중요도가 높다고 판단하였으나, 양국간 환자구성 내용은 거의 유사하다고 판단하였다. 기대여명(life expectancy)에서는 통상 류마티스관절염 경제성평가 연구가 생애에 걸친 장기 연구이기 때문에 life expectancy가 어느 정도 중요한 요인이라고 평가하였는데, 2009년 OECD에서 발표한 자료에 따르면 한국 여성의 평균 수명은 82.4세, 남성은 75.7세였으며, 영국 여성의 평균 수명은 81.1세, 남성은 77.1세로 모두 유사한 값을 보였다. 따라서 위의 세가지 요인은 모두 중요도에 상관없이 매우 높은(very high) 일치도를 보였기 때문에 unbiased 하다는 결론을 내렸다.

건강상태 선호도 (health-status preference) 항목은 본 연구가 최종적으로 EQ-5D를 이용하여 ICER 값을 구하고 있기 때문에 매우 관련이 높은 중요한 요인으로 평가하였다. 한편 일반적으로 utility는 국가별로 차이가 존재하는 것으로 알려져 있어 우리나라의 비용효과비에 어느 정도 영향을 미칠 것으로 판단하였다. 수용도(acceptance), 순응도(compliance), 환자에 대한 유인(incentive to patients)과 관련한 요인은 Brennan의 연구에 활용되지 않았으므로 본 항목의 중요도에는 ‘해당사항 없음’으로 판정하였고, 일치도는 중간 수준으로 평가하여 결국 적용가능성에 영향을 미치지 않을 것으로 판단하였다. 생산성손실 (Productivity and work-loss time)은 위의 생산성 비용 산출방법론 (productivity cost approach)과 유사하게 평가하였는데, 다만 생산성 손실비용 그 자체는 방법론 보다 국가간 차이를 더 많이 보일 것 같아서 일치도는

좀 더 낮게 평가하였다. 마지막으로 류마티스관절염은 감염성 질환이 아니므로 질병전파(disease spread)와는 관련이 없는 것으로 평가하였다.

Boulenger 도구를 통한 진단 결과

Boulenger 도구는 11개의 카테고리, 42개의 항목으로 구성되어 있는데, 평가 결과 긍정적인 항목이 31개, 부분 긍정이 5개, 부정적 항목 혹은 정보 부재가 3개, 비해당이 3개였다. 비해당 3항목을 제외한 39개에 대한 평균은 89.5점으로 상당히 높은 점수를 얻었는데 이는 Brennan의 연구가 높은 질임을 보여주는 것이다(Table 3). 다만 비교 대안, 인구 집단 특성, 분석 대상이나 연구 샘플에 대한 대표성에 대하여는 서술이 충분하지 않았고, 효과 자료와 재원에 대해서도 서술이 충분하지 않아 이에 대해서는 부분 긍정 (partially)으로 판정

Table 3. Transferability estimation by Boulenger method

	Question	Answer	Score
Study questions	Q1. Is the study question clearly stated?	Yes	1
	Q2. Are the alternative technologies justified by the author(s)?	Yes	1
Health technology	HT1a. Is the intervention described insufficient detail?	Yes	1
	HT2a. Is (are) the comparator(s) described in sufficient details?	Partially	0.5
Setting	SE1. Did the authors correctly specify the setting in which the study took place (e.g. primary care, community)?	Yes	1
	SE2a. Is (are) the country (ies) in which the economic study took place clearly specified?	Yes	1
Perspective	P1a. Did the authors correctly state which perspective they adopted for the economic analysis?	Yes	1
Study populations	SP1a. Is the target population of the health technology clearly stated by the authors or when it is not done can it be inferred by reading the article?	Yes	1
	SP2. Are the population characteristics described? (e.g. age, sex, health status, socio-economic status, inclusion/exclusion criteria)	Partially	0.5
	SP3a. Does the article provide sufficient detail about the study sample(s)?	Yes	1
	SP4. Does the paper provide sufficient information to assess the representativeness of the study sample with respect to the target population?	Partially	0.5
Modeling	M1. If a model is used is it described in detail?	Yes	1
	M2. Are the origins of the parameters used in the model given?	Yes	1
Effectiveness	E1. If a single study is used is the study design described (sample selection, study design, allocation, follow-up)?	Yes	1
	E2. If a single study is used are the methods of data analysis described (ITT/per protocol or observational data)?	Yes	1
	E3. If based on a review/synthesis of previous published studies, are review methods described (search strategy, inclusion criteria, sources, judgement criteria, combination, investigation of differences)?	Not applicable	-
	E4. If based on opinion, are the methods used to derive estimates described?	Not applicable	-
	E5a. Have the principal estimates of effectiveness measures been reported?	Yes	1
	E6. Are the side effects or adverse effects addressed in the analysis?	Yes	1
	E7a. Does the article provide the results of a statistical analysis of the effectiveness results?	Yes	1

Table 3. Transferability estimation by Boulenger method (continued)

Benefit measure	B1. Do the authors specify any summary benefit measure(s) used in the economic analysis?	Yes	1
	B2. Do the authors report the basic method of valuation of health states or interventions?	Yes	1
	B3. Do the authors specify the source(s) of health states (e.g. specific patient population or the general public)?	Yes	1
	B4. Do the authors specify the valuation tool used?	Yes	1
	B5a. Is the level of reporting of benefit data adequate(incremental analysis, statistical analyses)?	Partially	0.5
Cost	C1a. Are the cost components/items used in the economic analysis presented?	Yes	1
	C2. Are the methods used to measure costs components/items provided?	Yes	1
	C3. Are the sources of resource consumption data provided?	Yes	1
	C4. Are the sources of unit price data provided?	Yes	1
	C5a. Are unit prices for resources given?	Partially	0.5
	C6a. Are costs and quantities reported separately?	No/ No information	0
	C7a. Is the price year given?	Yes	1
	C8. Is the time horizon given for each element of the cost analysis?	Yes	1
	C9a. Is the currency unit reported?	Yes	1
	C10. Is a currency conversion rate given?	No/ No information	0
	C11. Does the article provide the results of a statistical analysis of cost results?	No/ No information	0
Discounting	D1. Was the summary benefit measure(s) discounted?	Yes	1
	D2. Were the cost data discounted?	Yes	1
	D3. Do the authors specify the rate(s) used in discounting costs and benefits?	Yes	1
	D4. Were discounted and not discounted results reported?	Yes	1
Discussion	S1a. Are quantitative and/or descriptive analysis conducted to explore variability from place to place?	Not applicable	-
	O1a. Did the authors discuss caveats regarding the generalisability of their results?	Yes	1
Transferability information score		(31+2.5+0)/(42-3)= 85.9%	

하였다. 또한 각 비용 항목에 대해 비용과 사용량을 구분하여 서술하지 않았고, 통화 환율이나, 비용 결과의 통계적 분석 결과를 제시하고 있지 않아 부정/정보부재(no/no information)로 판정하였다. 자세한 방법 서술, 가치 판단 도출 방법, 각 국가 간의 변이성 조사 분석 여부는 본 연구와 관련이 없는 항목으로 비해당(not applicable)으로 판정하였다.

Urdahl 도구를 통한 진단 결과

Urdahl 도구의 첫째 항목은 연구질문의 명확한 정의 등에 대한 평가로서 Brennan의 연구는 명확하고 충분히 서술하고 있는 것으로 평가하였다. 둘째 항목은 모형 및 가정에 대한 자세한 보고 및 서술로써 연구 setting, 인구 집단, 비교 대안, 모델의 구조 및 주요 가정 등에 대한 서술 내용을 평가한 결과 이에 대해서도 자세하고 충분히 서술 및 열거하고 있었다. 셋째 항목은 영국 데이터가 우리나라에 어느 정도 관련성을 보이는데 임상자료에서는 어느 정도 관련성이 높으나, 의료

이용 자료, 단위 비용, 효용에 대해서는 관련성이 낮을 것으로 판단되어 부분적으로 관련성이 있다고 평가하였다. 네 번째로는 양 국가간 데이터 차이에도 불구하고 모델의 강건함(robustness)이 민감도분석을 통하여 입증되는지에 대한 평가였는데, Brennan의 연구에서 실시했던 민감도 분석의 범위 내에 국내 관련 데이터의 값이 포함되지 않는 경우가 있어서 모델의 강건함이 일부만 입증된 것으로 판단하였다.

고찰 및 결론

류마티스관절염 치료제 경제성평가 사례연구를 통하여 국가간 적용가능성 평가도구를 비교한 결과 각각의 장단점을 다음과 같이 제시할 수 있다. 우선 Welte 진단 도구는 비용과 효과에 영향을 줄 수 있는 14개의 항목에 대하여 적용가능성 평가 항목으로서의 중요도 및 양 국가간의 일치도 분석을 통하여 적용 국가의 비용효과비에 미치는 영향을 평가

Table 4. Transferability estimation by Urdahl method

Research question	Specific question	evaluation
Definition of target decision-maker or jurisdiction	1) Decision maker 2) Economic study question	Yes
Transparent reporting of model specification	1) Study setting 2) Patient population 3) Alternative intervention 4) Model structure and key assumption	Yes
Relevance of data inputs to target decision-maker or jurisdiction	1) Clinical data 2) Resource use data 3) Unit costs 4) Preferences/Utility	In part
Assessment of robustness of model to variation in data inputs within and between jurisdictions	1) Robustness 2) Model accommodation for potential difference in compliance rate 3) Model accommodate for differences in utilities 4) External consistency	In part

하고 이에 따라 적용가능성을 판단하는 것이다. 이 과정에서 양국의 실정을 자세하게 파악해야 하기 때문에 연구 목적 및 디자인, 모델에 대한 상세한 이해가 필요한 것은 기본이며 이 외에도 자료를 평가하기 위하여 많은 노력을 필요로 한다. 그러나 적용가능성에 대하여 상세히 평가할 수 있는 기본적인 틀을 제공하고 있으며, 실제로 외국에서 수행한 연구결과를 적용할 때 비용과 효과에 어떠한 영향을 줄 것인지를 구체적으로 확인할 수 있다는 장점이 있다.

그러나 평가 자체에 구체적인 기준이 없고, 중요도와 일치도의 조합에 따른 도입 국가에서의 CER 변화 정도의 평가도 진행되는 연구자들의 판단에 기인하기 때문에 선입견이 들어갈 수 있고 연구자마다 다른 결론을 내릴 수 있다는 단점이 있다. 예를 들어, 류마티스관절염 연구에 대한 할인율의 중요도를 평가할 때 장기간의 연구라는 이유로 이를 매우 중요(very high)하다고 판단할 것인지 중요(high)하다고 판단할 것인지 기준이 명확하지 않다. 또한 Welte의 진단 도구에는 분석 기간, 분석 대상 인구 집단, 분석 기법, 결과 등에 대한 언급이 없고, 연구의 질이나 불확실성, 재정에 대해서도 구체적인 언급이 없기 때문에 Welte의 진단 도구에서 제시하고 있는 요인 이외에도 별도로 확인이 필요할 수 있다.

한편 Boulenger 진단 도구는 연구의 질이 높은지 그리고 각 항목의 서술이 명확한지에 중점을 두고 있다는 점에서 다른 도구와 차이를 보였다. 본 도구의 장점으로는 42개의 항목으로 구분되어 있어 다른 진단 도구에 비해 체크 항목이 자세하며, 항목 또한 전반적으로 중요한 항목들에 대해 고루 분포되어 있어 좀 더 자세하고 명확하게 체크해 볼 수 있다는 점이다. 평가 또한 4가지로 간단하게 분류하고 있고 대답에 따라서 적용가능성 정보 점수(transferability information score)를 백분율로써 측정할 수 있기 때문에 보다 명시화, 구체화하기가 용이하다.

그러나 이 점수를 적용가능성으로 직접적으로 해석하기에는 다소 무리가 있을 수 있다. 예를 들면, 비용 및 사용량,

자원, 자료원 등에 대해서 명확하고 자세하게 서술하고 있는 높은 질의 연구라 할지라도 시험 국가와 도입 국가간 비용적인 측면이나 그 현황이 현저하게 다르다면, 적용가능성 정보 점수는 높을지라도 현실적으로 도입하기에는 어려울 수 있다. 이러한 특징으로 인하여 오히려 다른 나라에 실제로 적용가능한지 여부에 대한 물음에 대해서는 답하기가 어려울 수 있다. 또한 전체적인 점수로써 평가하기 때문에 각 항목에 대하여 개별적으로 도입 여부를 결정할 수 없다는 단점이 있다.

Urdahl의 진단 도구는 결정권자, 모형, 자료의 관련성, 민감도 분석과 관련하여 크게 4가지 항목에 대해서 평가하는데, 이들 4가지 항목은 경제성평가의 전반적인 부분과 관련이 있으므로 기존 경제성 평가 결과를 전체적으로 검토하기에 적절하다. 그러나 포괄적이고 4가지 항목에 대한 세부 질문도 명확하게 명시하고 있지 않기 때문에 진단 도구 평가 후에 적용가능성의 여부를 평가하는데 평가자의 주관이 많이 개입될 우려가 있다. 또한 각 세부 질문에 대한 답변도 yes, no, in part 등의 간단한 대답으로 이루어지기 때문에 해당 질문 및 항목에 대한 답으로써는 연구의 질 등을 충분히 표현하기가 어렵고 평가 이후 연구의 수준도 가늠하기가 어렵다. 따라서 Urdahl의 진단 도구는 적은 노력을 들여 전반적인 내용에 대하여 다시 점검하고 큰 틀을 보고자 하기에는 좋은 도구라 할 수 있으나, 반면 이 도구만을 이용하여 적용가능성을 평가하는 것은 충분하지 않을 수 있다.

사례연구 결과를 살펴보면 영국에서 수행된 Brennan 연구의 국내 적용가능성 평가와 관련하여 Welte 진단 도구를 이용한 경우에는 비용 관련 항목과 할인율, 생산성 손실, 치료 방법, 건강 상태 선호도의 경우에는 적용가능성이 낮은 것으로 평가되었으므로 모델 보정(modeling adjustment)을 해야 할 것이라는 결론을 내릴 수 있다. Boulenger 진단 도구를 이용한 경우에는 적용가능성 정보 점수가 85.9%로써 상당히 점수가 높아 해당 연구의 질이 높고 서술이 명확하게 이루어

어졌음을 알 수 있었다. 또한 Urdahl 진단 도구를 이용한 경우에는 전반적으로 명확히 서술되어 있으나, 일부 자료에 대해서는 국가간 관련성이 부분적으로만 입증되었기 때문에 해당 자료의 도입은 좀 더 자세하게 확인하고 고려해야 한다는 결론을 내릴 수 있었다.

위의 세가지 결과를 종합하여 볼 때 영국에서 시행한 경제성 평가를 그대로 한국에 적용하고, 일반화하기는 어려우며 만약 이를 도입하고자 한다면 우리나라 특성에 맞게 각 항목을 합리적으로 조정할 필요가 있다. 특히 류마티스관절염에 대한 약물 투약 순서(treatment sequence) 및 비용 등은 국가에 따라 차이를 보일 수 있기 때문에, 이런 부분에 대하여 주의 깊게 비교 및 검토를 해야 한다. 또한 류마티스관절염 연구에서는 비용효용분석과 관련하여 효용가중치(utility)가 많이 이용되는데 외국자료를 이용하여 건강 상태에 대한 선호도를 반영할 때 특히 국가간 신중을 기해야 한다.

이상에서 살펴본 바와 같이 현재까지 개발된 진단 도구들은 각 도구별로 뚜렷한 특징을 보이고 있다. Boulenger 도구는 외국에 적용하고자 하는 연구가 그 자체로서 명확하게 기술되고 연구의 질은 우수한지 여부에 평가의 초점이 맞추어져 있다. 반면 Welte 도구에서는 각각의 평가 항목에 대한 중요도 및 국가간의 일치도를 파악함으로써 도입국가에서의 비용효과비에 미치는 정도를 평가하는데 중점을 두고 있다. Urdahl 진단 도구는 적용가능성에 대한 기본적 요건을 간결하게 평가하는데 좋은 도구이다. 즉 현재 개발된 도구들은 각각 적용가능성의 일부를 평가하는데에는 활용할 수 있으나 전체적인 평가의 측면에서는 미흡한 점이 있다. 따라서 경제성평가의 국가간 적용가능성을 보다 종합적으로 평가하기 위해서는 향후 적용가능성 평가의 다각적인 측면을 포괄할 수 있는 체계적인 도구의 개발이 필요하다.

감사의 말씀

본 연구는 일부 보건복지가족부 보건의료기술연구개발사업의 지원에 의하여 이루어진 것으로(A084794) 이에 감사드립니다.

참고문헌

1. Drummond MF, Bloom BS, Carrin G, et al.: Issues in the cross-national assessment of health technology. *Int J*

- Technol Assess Health Care.* 8, 671 (1992).
2. Francis Pang.: Design, Analysis and Presentation of Multinational Economic Studies: The Need for Guidance. *Pharmacoeconomics.* 20, 75 (2002).
3. Robert Welte, Talitha Feenstra, Hans Jager and Reiner Leidl.: A Decision Chart for Assessing and Improving the Transferability of Economic Evaluation Results Between Countries. *Pharmacoeconomics* 22, 857 (2004).
4. Stephanie Boulenger, John Nixon, Michael Drummond, Philippe Ulmann, Stephen Rice, Gerardde Pouvourville.: Can economic evaluations be made more transferable? *Eur J Health Econ.* 6, 334 (2005).
5. John Nixon, Stephen Rice, Michael Drummond, Stephanie Boulenger, Philippe Ulmann, Gerard de Pouvourville.: Guidelines for completing the EURONHEED transferability information checklist. *Eur J Health Econ.* 10, 157 (2009).
6. Hege Urdahl, Andrea Manca, and Mark J Sculpher.: Assessing Generalisability in Model-Based Economic Evaluation Studies: A Structured Review in Osteoporosis *Pharmacoeconomics.* 24, 1181 (2006).
7. New treatments in arthritis: the use of TNF- α blockers in adults with rheumatoid arthritis. Report of a working party of the British Society for Rheumatology, 2 April 2001.
8. Brennan, N. Bansback, A. Reynolds, P. Conway.: Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK *A. Rheumatology.* 43, 62 (2004);
9. National Institute for Clinical Excellence. Guidance for manufacturers and sponsors: Technology appraisals process series No. 5. London. *National Institute for Clinical Excellence.* (2001).
10. Saskia Knies, MPhil, André J. H. A. Ament, PhD, Silvia M. A. A. Evers, PhD, Johan L. Severens, PhD.: The Transferability of Economic Evaluations: Testing the Model of Welte. *Value in Health.* 12, 730 (2009).
11. Sculpher MJ, Pang FS, Manca A, Drummond MF, Golder S, Urdahl H, Davies LM, Eastwood A.: Generalisability in economic evaluation studies in healthcare: a review and case studies. *Health Technol Assess.* 49, 1 (2004) .