

약물이상반응의 인과성 평가를 위한 한국형 알고리즘의 개발

서울대학교병원 임상약리실¹, 서울대학교병원 임상시험센터²,
서울대학교 의과대학 예방의학교실³, 식품의약품안전청⁴

홍경섭^{1,2,3}, 박병주^{2,3}, 신상구^{1,2}, 양재석⁴, 이승미³, 김윤이³, 서화정³, 김주한³

=Abstract=

Development of a Korean Algorithm for Causality Assessment of Adverse Drug Reactions

Kyoung-Sup Hong^{1,2,3}, Byung-Joo Park^{2,3}, Sang-Goo Shin^{1,2}, Jae-Seok Yang⁴,
Seung-Mi Lee³, Yooni Kim³, Hwa-Jeong Seo³, Ju Han Kim³

*Clinical Pharmacology Unit, Seoul National University Hospital¹,
Clinical Trial Center, Seoul National University Hospital²,*

*Department of Preventive Medicine, Seoul National University College of Medicine³,
Korea Food and Drug Administration⁴*

Background : According to the study conducted on hospitalized patients in the United States, there were more than 2,000,000 serious ADRs in hospitalized patients, causing over 100,000 deaths annually. If true, then ADR is the 4th leading cause of death in the United States. In Korea, ADR monitoring system has not been activated. ADR monitoring system in the Seoul National University Hospital and Korean Medical Association, representing the nationwide clinicians, was initiated on May 2001 and on May 2002, respectively. Until Oct 1st 2002, only 30 cases, 16 cases from the Seoul National University Hospital and 14 cases from the Korean Medical Association were reported. At present, rapid and valid review, resulting in helpful feedback to the reporters is essential for the activation of ADRs monitoring system in Korea. In the same context, the algorithm, easy-to-use and in Korean, for the initial causality assessment became necessary.

Methods : In this study, a draft algorithm was constructed after reviewing the three current algorithms: Naranjo, French, and RUCAM. Then a study was conducted to evaluate the validity and reliability of the algorithm. The test was done with three target users, a physician majoring in internal medicine, a general physician and pharmacist majoring in Pharmacoepidemiology, assessed the causality of the reported 26 ADR cases with the algorithm as test reviewers. Finally, the results were compared with the gold standard, the decision of the Committee for Drug Safety within the Korean Medical Association.

Results : The average sensitivity and specificity was 96%, 38% respectively. According to the result of the test-retest reliability with Pearson correlation coefficient, two of 3 test reviewers showed significant correlations.

Conclusion : The Korean Algorithm seemed to be valid and reliable. But it is important to make an additional effort to modify and update the Korean Algorithm. The development of the Korean Algorithm must be the first step to establish the Adverse Drug Monitoring System in Korea.

KEY WORDS : Adverse drug reactions, Algorithm, Causality assessment, Validity, Reliability.

교신저자 : 박병주

소 속 : 서울대학교 의과대학 예방의학교실

주 소 : 서울특별시 종로구 연건동 28, 110-799

전 화 : 02) 740-8325 Fax : 02) 747-4830 E-mail : bjpark@snu.ac.kr

서 론

Lazarou 등은 1966년에서 1996년에 걸친 전향적 메타 분석 연구를 통해 미국 입원 환자에서의 심각한 약물이상반응의 빈도 및 그로 인한 사망률을 산출하였고, 그 결과를 1998년 JAMA(the Journal of the American Medical Association)에 발표하였다¹⁾. 심각한 약물이상반응이란 사망 혹은 생명의 위협을 초래하거나, 입원 혹은 입원 기간의 연장, 기형아 출산, 영구적인 불구를 초래하는 경우를 의미한다²⁾. 이 연구를 통해 미국 입원 환자의 약 6.7%는 심각한 약물이상반응을 경험하며, 입원 환자의 0.32%는 그로 인해 사망하는 것으로 나타났고, 이 결과를 적용하여 미국에서만 연간 2,216,000건 이상의 심각한 약물이상반응이 발생하고, 그로 인해 연간 106,000명이 사망하는 것으로 추정하였다. 이는 폐질환, 당뇨병, 에이즈(AIDS), 사고로 인한 사망자 수보다 많고, 미국인에서의 사망 원인 중 4위에 해당함을 보고하였다. 더구나 이는 외래 환자를 포함하지 않은 결과로서 전체 발생 건수는 여전히 불확실하였다. 약물이상반응의 경제적인 측면에 관한 연구에서, Johnson 등(1995)은 약물이상반응에 의한 이환 및 사망으로 인한 연간 총 비용이 1,360억불로, 심혈관 질환 혹은 당뇨병에 의한 총 비용보다 많은 것으로 추정하였다³⁾. Classen 등(1997)은 약물이상반응을 경험한 입원 환자들이 경험하지 않은 환자에 비해 2배 이상의 재원기간 및 비용, 사망률을 보이는 것으로 보고하였다⁴⁾.

약물이상반응의 심각성은 최근 시장에서 출시된 terfenadine, astemizol, troglitazone, cisapride, phenylpropanolamine(PPA), cerivastatin 등의 약물을 통해서도 잘 알 수 있다. 이중 2세대 항히스타민제인 terfenadine은, 1985년 미국에서 시판 승인을 받은 이후 10년 이상 환자들에게 널리 사용된 약제로서, 시판 후 약물이상반응 보고체계 구

축 및 개별 증례보고에 대한 정확한 평가의 중요성을 보여 주는 단적인 예라 할 수 있다. 항진균제, 경구피임제, terfenadine의 3개 약물을 복용하던 39세 여자 환자에서 발생한, 치명적일 수 있는 심부정맥인 Torsades de Pointes와 그로 인한 간헐적인 실신이 1990년 Monahan 등에 의해 JAMA에 보고되었고⁵⁾, 뒤이은 약 7년 간의 철저한 조사를 통해 원인이 terfenadine임을 밝혀냈고, terfenadine은 1997년 미국 시장에서 철수되었다. 결과적으로, 치명적인 심부정맥으로 인한 추가적인 사망을 막을 수 있었고, 단 하나의 증례보고가 갖는 중요성 및 그에 대한 정확한 평가의 중요성을 일깨워주었음은 물론, 개별 증례에서의 인과성 평가를 위한 알고리즘의 개발이 필요함을 일깨워 준 계기가 되었다.

현재 외국에서 널리 사용되고 있는 알고리즘으로는 Naranjo algorithm(1981)⁶⁾, French algorithm(1982)⁷⁾, RUCAM algorithm(1993)⁸⁾, UMC causality categories²⁾ 등을 들 수 있다. 이들은 언어는 물론이고 목표 평가자, 증례보고 방법 및 정보량, 약물이상반응에 대한 인식 수준 등이 우리나라와는 다른 외국에서 사용되는 알고리즘으로서, 이들 알고리즘을 한국에서 그대로 사용하기에는 무리가 있는 것이 현실이다.

현재 시작 단계인 병원내 이상반응 모니터링 체계 구축 및 전국의 개업 및 봉직 의사를 대상으로 한 자발적 약물이상반응 보고체계의 정착 및 활성화를 위해서는 신속하고 정확하게 약물과 이상반응간의 인과관계를 평가하고, 이를 보고자에게 신속하게 회신하는 것이 필수적이다. 그러나, 평가를 위한 인력 및 자원이 일천한 현실에서 부족한 정보로도 이용 가능하고 간편한 초기 인과성 평가법의 개발이 절실하게 되었고, 그에 따라 한국형 약물이상반응 인과성 평가 알고리즘의 개발을 본 연구의 목표로 설정하였다.

표 1. 서울대학교병원 및 대한의사협회를 통한 월별 약물이상반응 보고 건수 추이
(2002년 10월 1일 현재)

기간		서울대학교병원*	대한의사협회**	계
년	월			
2001	1-3	-	-	-
	4-6	-	-	-
	7-9	3	-	3
	10-12	3	-	3
2002	1-3	2	-	2
	4-6	2	-	2
	7-9	6	14	20
		16	14	30

*2001. 1. 약물이상반응소위원회 발족, 활동 시작

**2002. 2. 약사위원회 산하 약물안전관리소위원회 발족

연구 대상 및 방법

오래 전부터 미국과 유럽 각국에서는, 드물지만 심각한 약물이상반응의 발견이 어려운 시판 전 임상시험의 한계를 깨닫고, 시판 후 조사의 일환으로 약물이상반응 보고체계를 구축하였고, 이를 통해 심각한 약화 사고로부터 자국 국민을 보호해 왔다. 2002년 10월 현재, 우리나라에는 선진 각국과 같은 약물이상반응 보고 제도가 마련되어 있으나, 활성화되지 못하고 있는 것이 현실이다. 그 주요한 원인으로 약물이상반응의 심각성에 대한 낮은 인식도와 홍보 부족으로 인하여 환자 진료에 대한 일차적인 책임을 갖는 의사들에 의한 자발적 약물이상반응 보고가 저조하다는 점이 지적되어 왔다.

2000년 10월 식품의약품안전청에서 전국 14개의 대학병원과 1개의 공단 일산병원에 대하여 시범적으로 병원내 의약품 이상반응 모니터링 체계를 구축하도록 협조를 요청하였다. 2001년 1월 서울대학교병원 약사위원회 산하 약물이상반응소위원회가 구성되었고, 동년 5월에 임상 각과 전임의와 선임전공의, 약물역학팀, 임상약리팀, 약제부 약사 등으로 구성된 실무팀이 정식 출범하였다. 병원 전산오더시스템에 정식 검사항목으로서 약물이상반응 모

니터링이 추가됨으로써, 원내 모든 주치의를 대상으로 자발적 약물이상반응 보고체계가 가동되었다. 2001년 7월부터 보고가 접수되었고, 2002년 10월 1일 현재, 총 16건이 보고되었다. 보고된 증례는 실무팀원에 의해, 병원 전산망을 이용한 초기 평가가 수행되었고, 이를 토대로 월 1회 열리는 약물이상반응소위원회를 통해 최종 평가가 수행되어 왔다. 한 단계 더 나아가 2002년 5월 대한의사협회 종합학술대회 이후 대한의사협회지에 자발적 약물이상반응 보고체계를 홍보하고 보고 서식을 첨부함으로써 명실상부한 전국의 개업 및 봉직 의사를 대상으로 한 보고체계가 구축되었다. 보고의 편의성과 인과성 평가를 위한 충분한 정보량의 확보는 상호 상충되는 부분으로 보고체계의 정착을 위한 시작 단계임을 감안하여 보고의 편의성에 주안점을 둔 간략한 보고 서식을 활용하였다. 추후 모니터링 대상이 의약품 이외로 확대될 가능성과, 식품의약품안전청 웹사이트에서 사용하고 있는 의약품 부작용이라는 용어를 사용하고 있음을 감안한 의약품 등 부작용 사례 보고서는 환자의 연령, 성별, 임신 여부 등의 인구학적 정보와 의심되거나, 병용투여 중인 의약품의 명칭, 용법, 투여 기간, 투여 목적 등의 의약품 정보, 부작용의 내용 및 경과,

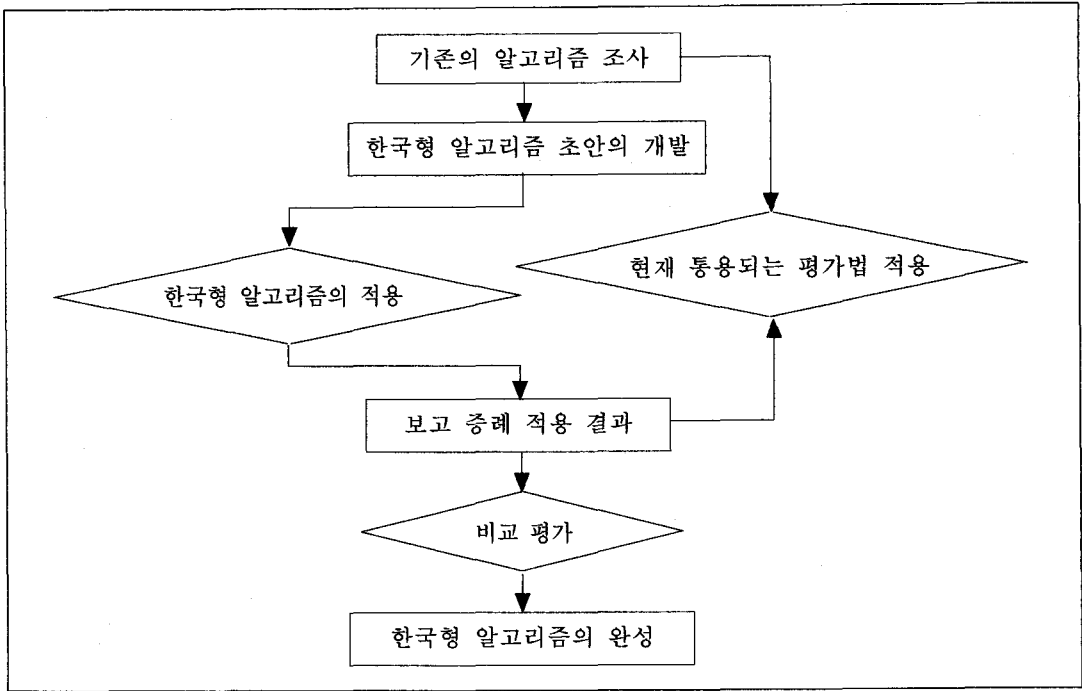


그림 1. 한국형 알고리즘의 개발을 위한 연구 과정.

재투약 여부와 결과 등의 부작용 정보, 참고 의견 등으로 간략하게 이루어져 있다. 당해 7월부터 보고가 접수되었고, 점차 건수가 늘어 2002년 10월 1일 현재 총 14건이 보고되었다. 서울대학교병원 및 대한의사협회를 통한 이상반응 보고 건수의 추이는 표 1과 같다.

2002년 10월 1일까지 보고된 총 30증례 중, 서울대학교병원의 초기 4개 증례의 경우, 현재와 같은 인과성 평가가 도입되기 이전으로, 황금기준으로 활용하기에는 부적절한 것으로 판단되어 연구 대상에서 제외하였고, 결과적으로 서울대학교병원에 보고된 12증례와, 대한의사협회 약물안전관리소위원회에 보고된 14증례, 총 26증례를 연구대상으로 삼았다.

연구의 첫 단계로, 현재 통용되는 알고리즘에 대한 장단점 분석은, 약물이상반응 평가 관련 문헌 및 저널 검색, WHO 산하 국제약물감시협력센터

(WHO Collaborating Center for International Drug Monitoring)인 Uppsala Monitoring Center (이하 UMC) 등 관련 단체의 웹사이트 검색을 통해 수행하였고, 구체적으로 Naranjo algorithm, French algorithm, RUCAM algorithm, UMC causality categories의 장단점을 분석하였다. 연구의 두 번째 단계로는 기존 알고리즘의 장단점 분석 결과를 기초로 한국형 알고리즘 개발을 위한 표준 알고리즘을 선택하였고, 한국의 실정에 맞게 보완하여 한국형 알고리즘 초안을 개발하였다. 세 번째 단계에서는 개발된 한국형 알고리즘 초안을 내과 전임의 1인, 약물역학 전공 의사 1인 및 약물역학 전공 약사 2인 등 향후 목표 사용자를 대표하는 4인이 시험 적용하였다. 2002년 10월 1일까지 보고된 30건의 증례 중, 최종 인과성 평가가 수행된 26증례를 대상으로 한국형 알고리즘 초안을 적용하였고, 그 결과에 대한 타당도 및 신뢰도 평가를 수

Question	Answer			Score
	Yes	No	Unknown	
1. Previous report?	+1	0	0	—
2. Event after drug?	+2	-1	0	—
3. Event abate on drug removal?	+1	0	0	—
4. +Re-challenge?	+2	-1	0	—
5. Alternative causes?	-1	+2	0	—
6. Reaction with placebo?	-1	+1	0	—
7. Drug blood level toxic?	+1	0	0	—
8. Reaction dose-related?	+1	0	0	—
9. Past history of similar event?	+1	0	0	—
10. ADR confirmed objectively?	+1	0	0	—

그림 2. Naranjo algorithm (Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, Janecek E, Domecq C, Greenblatt DJ. *A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther.* 1981 Aug;30:239-45).

행하였다. 동시에 약물역학 전공 의사인 1인의 평가자가 현재 세계적으로 널리 사용되고 있는 WHO 산하 국제약물감시협력센터인 UMC 에서 개발된 인과성 분류 기준(causality categories)²⁾을 적용하였고, 그 타당도를 평가하였다.

타당도 평가를 위해 서울대학교 의과대학 교수 19인으로 구성된 약사위원회 산하 약물이상반응소 위원회의 평가결과와, 의대 교수 11인으로 구성된, 대한의사협회 약사위원회 산하 약물안전관리소 위원회의 인과성 평가결과를 황금기준으로 삼았다. 최종 개발할 한국형 알고리즘의 목표 사용자는 궁극적으로, 모든 의사 및 간호사 그리고 임상 수련을 받은 약사 등으로 확대하는 것이 바람직하나, 타당도 및 신뢰도 평가를 통하여 한국형 알고리즘 개발을 최종 목표로 하는 본 연구의 성격상, 임상 의사를 대표하는 내과 전임의(전문의) 1인과, 약물역학 전공 의사 1인, 약물역학 전공 약사 2인 등 총 4인의 다른 배경을 가진 시험 평가자를 연구 대상으로 하여 알고리즘 초안을 이용한 인과성 평가를 수행하였다. 평가자 3인의 각 증례별 점수에 의해 인과성을 판정하였고, 그 결과를 황금기준과

비교하였다. 인과성 판정 결과가 Possible 이상인 증례에서의 일치도로 민감도를, 인과성 판정 결과가 Unlikely 이하인 증례에서의 일치도로 특이도를 평가하였다. 신뢰도는 동일한 증례 보고서를 대상으로 동일 시험 평가자가 1개월 간격을 두고 반복 평가한 결과간의 일치도를 상관 분석을 통해 평가하였다. 네 번째 단계에서는, 기존의 평가 결과 비교 및 평가 경험에 기초하여 알고리즘을 수정 보완하여 한국형 알고리즘을 완성하였다(그림 2).

결 과

1. 기존의 알고리즘

현재 외국에서 널리 사용되고 있는 알고리즘으로는 Naranjo algorithm(1981)⁶⁾, French algorithm(1982)⁷⁾, RUCAM algorithm(1993)⁸⁾, UMC causality categories²⁾ 등을 들 수 있다. Naranjo algorithm⁵⁾(그림 3)은 간략한 10개의 질문으로 구성되어 있고, 응답 및 점수도 예/아니오/알 수 없음의 3단계와 -1~+2점으로 간략히 구분되어 있

Re-administration Course	CHRONOLOGY (Time to onset)						incompatible
	very suggestive			compatible			
	(R+)	(R0)	(R-)	(R+)	(R0)	(R-)	
-Suggestive	C3	C3	C1	C3	C2	C1	C0
-Inconclusive	C3	C2	C1	C3	C1	C1	C0
-Not suggestive	C3	C2	C1	C3	C1	C1	C0

C0=incompatible chronology
C1=doubtful chronology
C2=possible chronology
C3=suggestive chronology
R=re-administration (positive, negative, inconclusive)

Alternative Non-drug causes	SEMIOLOGY					
	suggestive of the role of the drug (or as favoring factors)			other cases		
	L(+)	L(0)	L(-)	L(+)	L(0)	L(-)
None (after a search)	S3	S3	S1	S3	S3	S1
Possible or present (or not searched for)	S3	S2	S1	S3	S1	S1

S1=doubtful S2=possible S3=suggestive L=reliable and specific lab. test

CHRONOLOGY	SEMIOLOGY		
	S1	S2	S3
C0	10	10	10
C1	11	11	12
C2	11	12	13
C3	13	13	14

10=unlikely 11=doubtful 12=possible 13=likely 14=very likely

그림 3. French algorithm (Moore N, Biour M, Paux G, Loupi E, Begaud B, Boismare F, Royer RJ. *Adverse drug reaction monitoring: doing it the French way. Lancet. 1985 Nov 9;2:1056-8*).

다. Naranjo algorithm의 장점은 무엇보다도 간편하며 적용이 쉽고, 극히 적은 정보 및 지식으로도 평가가 가능하다는 점을 들 수 있다. 단점은 인과성 평가에 가장 중요한 시간관계에 대한 상세한 기술이 누락(2번)된 점과, 실제로 거의 이루어지지 않는 위약 투여 후의 반응 평가항목(6번), 구체적이지 않은 문항(10번) 등을 들 수 있고, 결과적으로 평가자에 따른 차이가 커질 수 있다. 실제로 일

원 환자에서의 약물이상반응을 대상으로 적용한 연구에서, 연구자간 큰 차이를 보일 수 있음을 보여 주었다^{9,10)}.

French algorithm⁶⁾(그림 4)은 Naranjo 알고리즘과 달리, 약물 투여 후 이상반응 발생까지의 시간, 재투약 결과, 임상 경과 등을 포함하는 연대기적 기준(chronological criteria)과, 증상이 약물에 의할 수 있는 포괄적인 개연성, 신뢰도와 특이도가 좋은

1. Time to Onset

Highly suggestive(+3), suggestive(+2), compatible(+1), inconclusive(0)

*If incompatible, “unrelated”

*If information not available, “insufficiently documented”

2. Course/Disappearance

Highly suggestive(+3), suggestive(+2), compatible(+1), against the drug(-2), inconclusive(0)

3. Risk Factor

Presence(+1 per each validated factors), absence(0)

4. Concomitant Drug

Time to onset incompatible(0), compatible but unknown(-1), compatible and known(-2), role proved in this case(-3)

5. Non-Drug causes

Ruled out(+2), possible or not investigated(-1~+2), probable(-3)

6. Previous Information on the Drug

Reaction unknown(0), published but unlabelled(+1), published and labeled(+2)

7. Response to Re-administration

Positive(+3), compatible(+1), negative(-2), not available(0)

*Toxic concentration of the drug(+3)

*Validated lab. Test positive(+3), negative(-3), not available(0)

*Attribution : **Highly probable(9~), Probable(6~8), Possible(3~5), Unlikely(1~2)**
Unrelated/Unclassified(~0)

그림 4. RUCAM algorithm (Danan G, Benichou C. Causality assessment of adverse reactions to drugs--I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. J Clin Epidemiol. 1993 Nov;46:1323-30).

검사결과, 질병경과 등의 비약물 요인에 대한 평가를 포함하는 징후학적 기준(semiological criteria)으로 구분되어 있고, 두 기준의 평가결과를 이용하여 최종 인과성을 평가한다. 장점으로는 시간관계에 대한 상세한 기술을 들 수 있고, 단점으로는 용량-반응 관계 및 투약 중단 후의 반응 등이 구체화 되어 있지 않고, 전체적으로 명확하지 않은 항목이 많아 해당 증례의 평가를 위해서는 해당 증례 및 임상 의학에 대한 상당한 경험과 지식을 요한다는 점을 들 수 있다. 또, 두 가지 다른 평가결과의 조합으로 최종 인과성을 평가함에 따라 개별 항목의 추가 및 평가결과 재조정 등의 개선이 상대적으로 곤란하다는 점을 들 수 있다.

가장 최근에 발표된 RUCAM algorithm⁸⁾(그림 5)

은 Naranjo algorithm과 유사한 문항으로 구성되어 있으나, 각 항목을 보다 상세히 기술하였고, 명확히 분류할 수 있도록 개선되었다. 장점으로는 시간관계, 병용약물, 특정 약물과 이상반응에 대하여 기존에 알려진 정보에 대한 평가항목이 상세히 보강되었다는 점과, 위험요인에 대한 평가가 추가되었다는 점을 들 수 있다. 단점으로는 임상경과 및 비약물 요인에 대한 평가가 여전히 명확하지 않다는 점과, 시간관계에 대한 기술이 구체화되었으나 여전히 그 평가에 있어서는 평가자의 임상 경험 및 지식 정도에 의해 개인차가 클 소지가 있다는 점을 들 수 있다.

WHO 국제약물감시협력기관인 UMC에서는 위에 언급한 점수화 평가를 위한 알고리즘과는 다른

표 2. Causality categories described by the UMC (WHO Collaborating Center for International Drug Monitoring)

Causality Terms	Description
1. Certain	약물 투여와 이상반응 발생이 적절한 시간 관계(plausible time relationship)를 보이고, 동반 질환이나 다른 의약품으로는 설명할 수 없고, 투약 중단 후 임상적으로 적절히 호전된 경우(clinically plausible de-challenge) 혹은 재투약 시 동일한 이상반응이 재현된 경우(satisfactory re-challenge procedure).
2. Probable/Likely	추론 가능한(reasonable) 시간 관계를 보이고, 동반 질환 및 다른 의약품으로 인한 가능성이 낮고(Unlikely), 투약 중단 후 임상적으로 적절히 호전된 경우.
3. Possible	추론 가능한 시간 관계를 보이지만 동반 질환 및 다른 의약품에 의해서도 설명가능하고, 투약 중단 후의 반응이 불확실하거나 정보가 부족한 경우.
4. Unlikely	적절하지 않은 시간 관계를 보이고, 동반 질환 및 다른 의약품에 의해서도 적절한 설명이 가능한 경우.
5. Conditional/Unclassified	평가를 위해서, 더 많은 정보가 필수적인 경우.
6. Inaccessible/Unclassified	보고된 정보가 모순되거나, 보충할 수 없는 경우 혹은 평가 불가능한 경우.

방식의 평가법으로, 인과성 용어 정의를 통해 인과성 평가를 수행하고 있다. 그 분류기준은 표 2와 같다. 모든 경우에서 비정상 검사결과를 보이는, 하나의 임상적 사건(clinical event)을 평가 대상으로 한다.

위의 평가법을 적용하기 위해서는 의심 약물 투여와 이상반응 발생간의 시간관계 및 투약 중단 후의 반응이 임상적으로 적절한지, 혹은 추론 가능한 지를 구별할 수 있어야 하며, 그 두 가지 기준과 비약물 요인에 대한 확인만으로도 인과성을 Certain으로 평가할 수 있다. 따라서, 위의 평가법은 간략하지만, 해당 약물 및 이상반응에 대한 지식과 경험이 풍부한 전문가에 의한 평가를 전제로 함을 알 수 있다. 따라서, 해당 분야 전문가에 의한 명확한 평가가 가능한 반면, 임상 비전문가에 의한 1차 인과성 평가에는 부적절하지만, 향후 개발될 인과성 평가용 알고리즘의 평가기준 설정에 도움이 될 것으로 판단하였다.

2. 한국형 알고리즘 초안

알고리즘 초안의 요건으로 주치의 이외의 약물 이상반응 모니터링의 실무경험이 있거나, 임상의학 및 약물역학을 수련받은 의사, 약사, 간호사 등에 의한 평가가 가능하도록 명확하고, 구체적인 항목으로 구성된 점수 평가법으로서, 이해하기 쉽고 명확한 한국어 표현이어야 하고, 각 평가항목은 보고된 정보의 손실이 없도록 세분화되어야 하며, 수정 및 보완이 쉬운 개별 항목별 평가법이 바람직하다고 판단하였고, 이에 근거하여 앞에서 비교 평가한 Naranjo algorithm, French algorithm, RUCAM algorithm 등 3개의 알고리즘 중 가장 최근에 개발된 RUCAM algorithm을 기초로 하여, 나머지 알고리즘의 장점을 반영한 한국형 알고리즘 초안을 작성하였다.

개선 내용으로는 우선 RUCAM 알고리즘의 단점으로 지적한 임상경과에 대한 명확하지 않은 기

표 3. 시험 평가자 4인에 적용한 알고리즘의 타당도 평가 결과

평가자	1	2	3	4	평균
Possible 이상 증례 수	23/24	22/24	23/24	24/24	23/24
민감도(%)	96	92	96	100	96
Unlikely 이하 증례 수	0/2	2/2	1/2	0/2	0.75/2
특이도(%)	0	100	50	0	38

술을 비례적인 용량-반응 관계와, 투약중단 시 이상반응의 종료 여부 및 그 양상을 추가함으로써 보다 구체화한 점을 들 수 있다. 그리고 투약시점과 이상반응 발생까지의 시간, 위험요인, 비약물요인 등의 항목에 일관된 평가를 위하여 예를 추가하였다. 그 외에, 단점으로 지적된 투약시점과 이상반응 발생까지의 시간 및 경과, 비약물요인에 대한 평가 항목은 개별 약물 / 환자의 병력 / 이상반응 양상 등에 의해 달라질 수 밖에 없고, 임상비전문가를 위한 구체적인 평가기준을 제시하는 것은 너무나 방대한 작업으로, 현실적으로 어려운 것으로 판단하였다.

3. 인과성 평가 결과

3.1. 한국형 알고리즘 초안 적용 결과

타당도 평가를 위해, 본 연구에서 황금기준으로 설정한 서울대학교병원 약물이상반응소위원회 및 대한의사협회 약물안전관리소위원회의 인과성 평가결과와, 목표 사용자를 대표하는 4인에서의 시험평가결과간의 일치도를 평가하였다. 연구 대상은 2002년 10월 1일 현재 서울대학교병원 약물이상반응 모니터링 소위원회에 보고된 총 16증례 중 인과성 평가가 수행된 12증례(증례 번호 1~12, 이하 서울대병원 증례)와 대한의사협회 약물안전관리소위원회에 보고된 14증례(증례 번호 13~26, 이하 대한의사협회 증례)이며, 총 26증례에 대한 소위원회의 최종 평가결과(황금기준) 및 4인의 목표 사용

자에서의 인과성 평가결과는 표 3에 요약하였다. 대한의사협회 약물안전관리소위원회의 평가는 알고리즘의 개발과 시기적으로 일치하였고, 시험평가자 4인 중 약물역학을 전공한 의사(평가자 3)의 평가결과를 대상으로, 중요 항목에 대한 조연도 받을 수 있었다. 전체적으로 인과성에 대한 Probable / Certain 등의 평가가 상대적으로 적었다는 점과, 이와 같은 저평가의 주요한 요인으로서 전형적인 임상경과(2번 항목), 시간관계(1번 항목) 등을 보이는 경우에도 평가항목의 최고 점수를 배점하지 않은 오류를 지적하였다. 또, 평가항목 중 위험요인(3번 항목)이 보다 구체화되어야 할 것을 지적하였다.

전체적인 결과의 분포를 보면, 최종 인과성 평가결과가 Possible 이상인 증례는 전체 26증례 중 24증례로 92%(24/26)이었다. 인과성 평가결과가 Probable 이상인 증례는 17증례로 65%(17/26), Possible 이하인 증례는 9증례로 35%(9/26)이었다. 특히, 인과성 평가결과가 Unlikely 이하인 증례는 26증례 중 2예에 불과하였다.

타당도 평가를 위한 항목 중 민감도는 최종 인과성 평가결과가 Possible 이상인 증례에 대한 일치도로 평가하였고, 특이도는 최종 인과성 평가결과가 Unlikely 이하인 증례에 대한 일치도로 평가하였다.

$$\text{민감도 (Sensitivity)} = \frac{\text{분모의 증례 중, 시험 평가 결과 Possible 이상인 증례 수}}{\text{최종 인과성 평가 결과 Possible 이상인 증례 수}}$$

$$\text{특이도 (Specificity)} = \frac{\text{분모의 증례 중, 시험 평가 결과 Unlikely 이하인 증례 수}}{\text{최종 인과성 평가 결과 Unlikely 이하인 증례 수}}$$

민감도 평가결과, 인과성 평가결과가 Possible 이상인 증례에 대한 일치도는 96%(23/24)이었고, 특이도인 인과성 평가결과가 Unlikely 이하인 증례에 대한 일치도는 38%(0.75/2)이었다. 평가자별 결과를 표 3에 요약하였다.

신뢰도는 최초 평가결과와, 약 1개월 경과 후에 3인의 동일한 평가자(시험 평가자 1, 2, 3)를 대상으로 반복 평가한 결과간의 상관성 검증을 수행함으로써 test-retest reliability를 평가하였다. 피어슨 상관 계수 산출 및 유의성 검정결과, 평가자 3인 중 2인에서 통계적으로 유의한 상관성이 있음을 확인할 수 있었다(표 4).

3.2. UMC 인과성분류기준 적용 결과

시험평가자 3을 대상으로 UMC 인과성 분류기준을 이용한 결과에 대한 타당도 평가결과 민감도는 100%(24/24), 특이도는 0%(0/2)로 나타났다.

한국형 알고리즘 초안을 이용한 평가와 비교시, 민감도는 95.8%에서 100%로 좋아진 반면, 특이도는 50%에서 0%로 낮아졌다(표 5).

4. 한국형 알고리즘

이상의 결과를 종합하여 볼 때, 한국형 알고리즘 초안의 민감도는 96%로 상당히 높았고, 특이도는 38%였으나, 증례수의 부족으로 인해 정확한 평가를 할 수는 없었다. 현재 국제적으로 통용되고 있는 UMC 인과성 분류기준은 시험평가자 1인의 결과에서, 근소하나마 더 좋은 민감도를 보였고, 호환성을 감안하여 그 평가기준이 한국형 알고리즘으로 통합될 수 있도록 개선이 필요하다고 판단하였다. 구체적으로 UMC 인과성 분류기준(표 2)에 부합되도록 시간관계, 임상적 경과, 비약물 요인에 대한 평가만으로도 Certain으로 인과성을 판정할 수 있어야 한다는 점과, 대한의사협회 약물안전관리소위원회의 지적 내용 중, 알고리즘 2번 항목인 임상적 경과에 대한 체계적인 저평가 가능성에 대한 개선이 필요하였다. 임상적 경과 중 현실적으로 확인하기가 쉽지 않은 용량-반응 관계의 확인 여부에 따른 점수 구분을 체계적 저평가의 요인으로 판단하였다. 따라서, 임상적 경과와 판단에서 비례적인 용량-반응 관계를 보이는 경우 이외에, 투약 중단 후 임상적으로 적절한 호전 양상을 보이는 경우에도 3점, 전형적이지는 않으나 추론 가능한 호전 양상을 보이는 경우 2점, 보존적 처치나 약물 치료 등으로 인해 경과를 정확히 파악할 수는 없으나, 투약 중단 후 호전된 경우 1점으로 배점기준

표 4. 시험 평가자 3인에 적용한 알고리즘의 신뢰도 평가 결과

평가자	1	2	3
상관 계수	0.7350	0.0835	0.8706
p-value	< 0.01	> 0.2	< 0.01

표 5. 평가자 3에서의 한국형 알고리즘과 UMC 인과성분류 기준을 적용한 타당도 평가결과 비교

평가법	한국형 알고리즘	UMC 인과성분류기준
민감도(%)	96 (23/24)	100 (24/24)
특이도(%)	50 (1/2)	0 (0/2)

1. 투약 시점과 이상반응 발생까지의 시간 (Time to Onset)

예) (0일/0주/0개월간의) 투약 (시작/종료) 후 0분/0시간/0일/0주

- 약물 및 이상반응의 특성과 잘 부합됨. (+3점)
- 약물 및 이상반응의 특성에 의해 설명 가능함. (+2점)
- 명확한 시간적 선후 관계. (+1점)
- 약물과 이상반응의 특성에 대한 추가 정보가 필요함. (0점)
- *투약과 이상반응 발생의 시간 관계를 알 수 없다면, 인과성 평가 불가.

2. 임상경과 (Clinical Course): 악화, 호전, 종료 등

- 비례적인 용량-반응 관계(dose-response relationship)를 보이거나, 투약 중단 후 임상적으로 적절한 (clinically plausible) 호전 양상을 보임. (+3점)
- 전형적이지는 않으나 추론 가능한(reasonable) 호전 양상을 보임. (+2점)
- 투약 중단 후 호전됨. (+1점)
- 경과에 대한 정보 부족. (0점)
- 위의 3가지 경우와 상반된 결과를 보임. (-2점)

3. 위험요인 (Risk Factor) 예) 가족력, 과거 병력, 간.신장 질환 등

- 있음. (1개당 +1점, 최대 2점)
- 없음. (0점)

4. 병용약물 (Concomitant Drug)의 기여도

- 투약과 이상반응 발생 시점이 모순되는 경우. (0점)
- 시간적으로 가능성이 있으나, 보고된 바 없음. (-1점)
- 시간적으로 가능성이 있고, 보고된 바 있음. (-2점)
- 본 증례에서의 역할이 밝혀져 있음. (-3점)

5. 비약물 요인 (Non-Drug Cause) 예) 질병 경과, 비약물 치료 등

- 배제 할 수 있음. (+3점)
- 가능성이 있음(-1점), 혹은 평가되지 않았거나 가능성 낮음. (0~+1점)
- 유력함; 질병의 특이적인 검사 결과 양성 혹은 전형적인 임상 경과를 보이는 경우. (-3점)

6. 약물에 대한 알려진 정보 (Previous Information on the Drug)

- 본 이상반응에 대하여 알려진 바 없음. (0점)
- 보고되었으나, 제품 허가 사항(FDA product label etc)에 반영되지 않았음. (+1점)
- 보고되었고, 허가 사항에 반영되어 있음. (+2점)

7. 재투약 시의 반응 (Response to Re-administration)

- 동일한 이상반응이 재현(Re-challenge +)됨. (+3점)
- 동일한 기전으로 설명 가능한 이상반응의 발생. (+1점)
- 반응 없음. (-2점)
- 재투약 하지 않았음. (0점)

8. 기 타

- 특이적인 검사 / 약물 농도 검사 결과; 독성 농도. (+3점)
- 최고 점수; 19점
- 최저 점수; -8점
- 점수별 인과성 평가
 - 9점 이상; 거의 확실함 (Certain) > 90%
 - 6~8점; 가능성 높음 (Probable/likely) > 70%
 - 3~5점; 가능성 있음 (Possible) 50%
 - 1~2점; 가능성 낮음 (Unlikely) < 30%
 - 0점 이하; 거의 무관함 (Contradictory) < 10%
 - 혹은, 추가 정보 필요(Conditional/unclassified)

그림 5. 약물이상반응의 초기 인과성 평가를 위한 한국형 알고리즘 (V 1.0).

을 완화하였다. 5번 항목인 비약물 요인에 대한 평가는 민감도를 높이기 위함은 물론, 평가자 간 차이를 줄이기 위하여도 수정이 필요하지만, 이 두 가지 목적은 서로 상충되는 측면이 있고, 항목의 성격상 방대한 임상정보, 충분한 임상지식 및 경험을 필요로 하므로 항목의 내용을 수정하는 것만으로 해결될 수는 없는 것으로 판단하였다. 다만, 비약물 요인을 배제할 수 있는 경우에, 초안의 2점에서 3점으로 배당 점수치를 높임으로써 시간관계, 임상적 경과, 비약물 요인에 대한 평가만으로도 최고 9점이 되어, Certain으로 판정할 수 있도록 개선하였다. 알고리즘 초안을 수정한 한국형 알고리즘은 그림 5와 같다.

고 찰

시험평가자 4인의 결과만으로, 다른 배경을 가진 목표이용자들 간의 일관된 경향성을 판단하는 것은 무리이고, 상당 부분이 평가자 개인의 태도에 따라 달라질 수 있다. 평가기준의 변경에 따른 결과의 차이를 평가하기 위한 Sensitivity Analysis를 수행하였다. 한국형 알고리즘의 타당도 평가를 위한 민감도를 최종 인과성 평가결과가 Possible 이상이 아닌, Probable 이상인 증례로 평가하였을 때, 내과 전문의인 임상 의사(평가자 1)의 경우, 민감도가 88%(Probable 이상; 15/17, cf: Possible 이상; 96%, 23/24)이었고, 약물역학 전공 의사(평가자 3)는 53%(Probable 이상; 9/17, cf: Possible 이상; 96%, 23/24), 약물역학을 전공하는 약사 2인(평가자 2, 4)은 각각 47%(8/17), 41%(7/17) (cf: Possible 이상; 92%(22/24), 100%(24/24))의 민감도를 보였다. 결과에서 알 수 있듯이, 인과성 평가결과가 Possible 이상인 증례에서의 민감도는 평가자 간 차이가 적은 반면, Probable 이상인 증례에서의 민감도는 내과 전문의인 평가자 1에서 높았고, 상대적으로 유사한 결과를 보였으나, 평가자 2, 3, 4

의 경우 민감도가 50% 내외로 저하되는 결과를 보였다. 따라서, 알고리즘을 이용한 정확한 인과성 평가를 위해서는 임상지식 및 경험이 필요할 것으로 판단하였다. 또, 한국형 알고리즘의 타당도를 더욱 개선하기 위하여는 Probable 이상의 인과성을 나타내는 증례에 대한 일치도를 높이기 위한 노력도 필요하며, 이는 UMC 인과성 분류기준을 한국형 알고리즘에 반영함으로써 어느 정도는 보완되었을 것으로 추정되나, 추가적인 검증 및 개선이 필요할 것으로 판단된다.

1991년 Classen 등은 입원환자에서 나타난 731건의 약물이상반응을 대상으로 Naranjo algorithm을 적용한 결과 99.3%의 증례가 Probable 이상임을 보고하였고⁹⁾, 2000년 Dormann 등은 입원환자에서 관찰된 46건의 약물이상반응을 대상으로 Naranjo algorithm을 적용한 결과 54.3%의 증례만이 Probable 이상이었음을 보고하였다¹⁰⁾. 현재까지도 약물이상반응의 인과성 평가에 사용되는 알고리즘의 효용성에 대한 체계적인 연구가 수행되지 않았고, 역학적으로 타당성이 입증된 알고리즘이 있는 것도 아니다. 현재 사용되는 여러 알고리즘에 대한 비교 분석을 통해, 한국형 알고리즘 초안을 제시하고, 타당도 및 신뢰도 평가 및 전문가의 자문을 통해 제기된 문제점을 보완하여 한국형 알고리즘을 완성한 본 연구의 성격상, 기존의 알고리즘을 이용한 결과와의 직접적인 비교를 통해 실제적인 개선을 확인할 수는 없었다. 하지만, 체계적으로 타당도 및 신뢰도 평가를 수행하였고, 황금기준인 소위원회의 최종 인과성 평가결과와 비교했을 때, 96%의 민감도를 보였다는 점은 특기할 만하다. 또, 내과 전문의인 평가자 1은 Probable 이상 증례를 대상으로 한 민감도 평가에서도 88.2%라는 높은 일치도를 보임으로써 임상 전문의를 대상으로 한 알고리즘의 활용 가능성도 엿볼 수 있었다.

본 연구의 장점은 임상 의사, 약물역학 전공 의사 및 약사 등, 다양한 배경을 가진 시험평가자를 대

상으로 하였다는 점을 들 수 있고, 결과적으로 평균 96%의 높은 민감도를 보임으로써, 본 한국형 알고리즘이 쉽고 효율적인 초기 인과성 평가와, 이를 위한 체계적 접근법으로 활용 가능성이 있다는 것을 확인한 것이다.

반면, 본 연구의 제한점으로는 4인이라는 적은 평가자 수로 인해 각 분야의 대표성이 결여되었다는 점과, 26예의 적은 증례를 대상으로 평가함으로써 다양한 결과를 얻지 못했다는 점을 들 수 있다. 이는 향후, 약물이상반응 보고체계의 정착 및 활성화에 따라 추가적으로 확인되어야 할 부분이다. 또, 대한의사협회를 통해 보고된 증례의 경우, 보고의 편의성에 주안점을 둔 간략한 보고서식을 활용하였고, 더구나 보고된 정보가 이에도 못 미치는 경우도 많았다. 추후에 보다 정확하고 일관된 평가를 위해서는 보고된 정보의 양적, 질적 평가가 필요하고, 이를 위한 구체적인 평가기준 및 방법도 개발되어야 할 것이다.

보다 정확한 인과성 평가를 위한 제안으로서, 본 한국형 알고리즘의 개선을 위하여 보다 많은 평가자와 증례를 대상으로 적용함으로써 문제점 파악 및 지속적인 수정이 필요하고, 인과성 평가에 필요한 증례보고서의 질 향상을 위하여 지속적인 홍보가 필요할 것이다.

현 단계에서 알고리즘의 일반화 가능성을 논하는 것은 어렵지만, 본 연구를 통해 개발된 한국형 알고리즘을 지속적으로 활용하고 개선해 나간다면, 한국의 약물이상반응 보고체계와 의료 현실 하에서, 적은 자원과 인력으로도 많은 보고를 신속하게 처리하기 위한 초기 인과성 평가법으로 적절히 활용될 수 있을 것이다. 또, 평가자의 입장에서는 인과성 평가를 위한 체계적인 접근을 할 수 있고, 더 나아가서 보고 서식에 알고리즘이 포함되고, 의료인 및 일반인들을 대상으로 약물이상반응에 대한 홍보 및 교육이 진행된다면, 보다 효율적으로 약물이상반응 모니터링을 수행할 수 있게 되어, 한국인에

서의 드물지만 심각한 약물이상반응의 조기 발견이라는 궁극적인 목표를 달성하기 위한 첫 걸음이 될 것으로 본다.

감 사 문

약물이상반응 증례의 인과관계 평가에 대한 지도를 아끼지 않으신 장인진 교수님과 약물이상반응소위원회 위원님으로서 임상적 평가에 많은 도움을 주신 서울대학교병원 권준수, 박준동, 정진호, 김주성 교수님과 간사로 수고해주신 정선희 선생님, 그리고 대한의사협회 약물안전관리소위원회 위원님으로서 많은 조언을 해주신 이애영, 안용민, 노규정, 배종면, 박건우, 김영식 교수님께 감사드립니다. 또, 세심하게 교정을 도와 준 오달석 선생님께도 다시 한 번 감사드립니다.

참 고 문 헌

1. Lazarou J, Pomeranz BH, Corey PN. *Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies.* JAMA 1998;279:1200-5
2. Uppsala Monitoring Center web site (www.who-umc.org/umc.html)
3. Johnson JA, Bootman JL. *Drug-related morbidity and mortality. A cost-of-illness model.* Arch Intern Med 1995;155:1949-56
4. Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. *Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality.* JAMA 1997;277:301-6
5. Monahan BP, Ferguson CL, Killeavy ES, Lloyd BK, Troy J, Cantilena LR Jr. *Torsades de pointes occurring in association*

- with terfenadine use. *JAMA* 1990;264:2788-90
6. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, Janecek E, Domecq C, Greenblatt DJ. *A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther* 1981;30:239-45
 7. Moore N, Biour M, Paux G, Loupi E, Begaud B, Boismare F, Royer RJ. *Adverse drug reaction monitoring: doing it the French way. The Lancet* 1985;2:1056-8
 8. Danan G, Benichou C. Causality assessment of adverse reactions to drugs--I. *A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. J Clin Epidemiol* 1993;46:1323-30
 9. Classen DC, Pestotnik SL, Evans RS, Burke JP. *Computerized surveillance of adverse drug events in hospital patients. JAMA* 1991;266:2847-51
 10. Dormann H, Muth-Selbach U, Krebs S, Criegee-Rieck M, Tegeder I, Schneider HT, Hahn EG, Levy M, Brune K, Geisslinger G. *Incidence and costs of adverse drug reactions during hospitalization: computerized monitoring versus stimulated spontaneous reporting. Drug Saf* 2000;22:161-8
 11. U.S. FDA web site (www.fda.gov)
 12. Strom BL. *Pharmacoepidemiology, 3rd ed. John Willey & Sons, West Sussex. p.525-38, 2000*