J. Fd Hyg. Safety
Vol. 26, No. 4, pp. 302~307 (2011)

Journal of Food Hygiene and Safety
Available online at http://www.foodhygiene.or.kr

건강기능식품 부작용 원인분석을 위한 알고리즘

이경진[1] · 박경식[2]* · 김정훈[3] · 이영주[3] · 윤태형[3] · 노기미[3] · 박미선[3] · 임동길[3] · 윤창용[3] · 정자영[3]
[1]위해정보과, 식품의약품안전청, [2]건강기능식품기준과, 식품의약품안전청
[3]영양기능연구팀, 식품의약품안전평가원

Algorithms for Causality Evaluation of Adverse Events from Health/Functional Foods

Kyung-Jin Lee[1], Kyoung Sik Park[2]*, Jeong Hun Kim[3], Young Joo Lee[3], Taehyung Yoon[3], Ki-Mi No[3], Mi-Sun Park[3], Donggil Leem[3], Chang Yong Yoon[3], and Jayoung Jeong[3]
[1]*Risk Information Division, Korea Food & Drug Administration*
[2]*Functional Food Standards Division, Korea Food & Drug Administration*
[3]*Nutrition and Functional Food Research Team, National Institute of Food & Drug Safety Evaluation*
(Received August 19, 2011/Revised October 10, 2011/Accepted November 8, 2011)

ABSTRACT - One of the most important objectives of post-marketing monitoring of dietary supplements is the early detection of unknown and unexpected adverse events (AEs). Several causality algorithms, such as the Naranjo scale, the RUCAM scale, and the M&V scale are available for the estimation of the likelihood of causation between a product and an AE. Based on the existing algorithms, the Korea Food & Drug Administration has developed a new algorithm tool to reflect the characteristics of dietary supplements in the causality analysis. However, additional work will be required to confirm if the newly developed algorithm tool has reasonable sensitivity and not to generate an unacceptable number of false positives signals.

Key words: algorithm, causality, signal analysis, dietary supplements

서 론

건강기능식품에 대한 식약청의 시판 전 안전성 평가에도 불구하고 사용자의 오・남용, 위해성분의 혼입・오염, 개인별 특이한 생리반응 등에 의하여 부작용이 나타날 수 있다. 식약청에서는 건강기능식품의 섭취에 따른 부작용 발생사례를 다양한 채널로 수집하고 그 원인을 분석하며 분석결과를 공개하고 있다[1]. 특히, 건강기능식품의 섭취와 부작용 발생간의 인과성을 평가하기 위하여 알고리즘의 적용이 필요한데, 기존에 개발된 알고리즘은 의약품의 부작용 인과성 평가에 적용할 목적으로 개발된 것이므로 건강기능식품의 특성에 적합한 알고리즘의 개발이 필요하게 되었다.

*Correspondence to: Kyoung Sik Park, Functional Food Standards Division, Korea Food & Drug Administration, Osong Health Technology Administration Complex, 187 Osongsaengmyeong2(i)-ro, Gangoe-myeon, Cheongwon-gun, Chungcheongbuk-do 363-951, Korea
Tel: 82-43-719-2455, Fax: 82-43-719-2450
E-mail: parkks71@korea.kr

건강기능식품의 섭취와 부작용 발생간의 인과성 평가를 위한 알고리즘의 개발

알고리즘의 개요

알고리즘의 주요 목적은 사용자가 목적하는 결론에 도달하기 위하여 필요한 정보를 얻을 수 있도록 일련의 질문 형태로 기본 공통 특징을 공유하는 데 있다. 알고리즘은 부작용의 인과성을 계량화하기 위한 세부 실행 기준을 포함하는 몇 가지 질문 형태로 나타난다. 알고리즘은 부작용을 확인하는 체계화・표준화된 평가법으로서 발생 시점(time to onset), 시간적 순서(temporal sequence), 원료에 대한 기존 정보(previous information on the ingredient), 섭취중단(dechallenge)과 재섭취(rechallenge) 등과 같은 지표들을 질문 항목으로 사용한다. 인과성 평가의 비뚤림(bias), 재현성, 유효성 문제를 해결하기 위하여 다양한 알고리즘이 개발되고 발표되었지만 모든 알고리즘이 단점을 갖고 있어서 황금률로 받아들일 수 있는 알고리즘은 없는 실정이다[2].

주요 알고리즘

'French method'로 불리고 있는 평가법이 1977년부터 프랑스 규제기관에 의해 사용되어 왔다(Table 1). 이 평가법은 병력 평가, 임상적·생물학적 관찰, 이상반응 진단과 예후 등의 3단계로 이루어져 있다. 이 방법의 장점은 각 원료를 개별적으로 구분하여 따로 평가하므로 병용 약물에 의한 부작용 발생 가능성을 배제할 수 있다는 점이지만 평가하는 데 시간이 많이 소요된다는 단점을 갖는다[3].

Kramer 등은 기존의 알고리즘을 확장하여 이상반응을 평가하는 일련의 새로운 기준을 개발하였다[4]. 이 알고리즘은 단일 원료 섭취 후에 나타나는 단일 이상반응을 평가하는 데 적용될 수 있으며 여러 가지 원료가 관여될 때는 각각의 원료별로 평가한다. 이 알고리즘은 6개의 결정표로 구성되어 있으며 각 표의 축에 −1～+1점을 부여할 수 있는 점수 체계가 포함되어 있는데, 각 축에 부여된 점수를 모두 합하면 인과성을 평가할 수 있는 −7～+7 사이의 점수가 산출된다. 이 알고리즘의 장점은 투명성이다. 평가자 간의 불일치는 그 원인이 확인될 수 있으며 교정이 가능하지만 일정 수준의 전문가라도 이 방법을 효율적으로 사용하려면 경험과 시간이 필요할 정도로 사용하기가 어려운 단점이 있다.

Table 1. French method for causality evaluation of adverse events

The drug	
1. Knowledge about the drug and its possible role	
Hypothesis, still debated	+1
A matter of worry, diffused by publications or work in progress	+2
Causality established	+3
2. Communication about this knowledge	
Reassuring about a lack of danger	0
Relatively worrying	+2
Serious cause for concern about presence of danger	+3
The patient	
3. Clinical case: risk factors	
No risk factors	0
Risk factors hardly detectable	+2
Presence of risk factors, easy to detect	+3
4. Drug management	
Respect of recommendation(s) or lack of precaution(s) has played any role in this case	0
Recommendation(s) not applies easily in this patient	+2
Neglect of recommendation(s), easy to apply by the prescriber or the patient	+3
Prescriptions	
5. Conditions of prescription	
Prescription indispensible to the patient	-12
Questionable prescription but acceptable	-4
Needless or absolutely contra-indicated prescription (or inappropriate prescription)	+3
6. Management of the adverse reaction	
Excellent, with prevention of the aggravation of the adverse reaction	0
Inadequate	+2
Absent, with aggravation of the reaction	+3
총점	
11 초과	Certain
7~11	Probable
1~6	Possible
1 미만	Unlikely

또 다른 결정표가 Blanc 등에 의해 개발되었다[5]. 이 알고리즘은 원료와 부작용 사례 간 관계의 속성을 평가하는 데 주안점을 두었는데, 시간 순서, 반응 패턴, 보유 질환의 역할 등의 3가지 요인을 고려하였다. 이 3가지 결정 요인들을 점수화하고 이를 조합하여 인과 가능성을 5단계로 등급화 하였다.

Emanueli와 Sacchetti는 기존의 알고리즘 개념에 근거를 두고 대규모 임상시험에서 관찰된 이상반응을 분석하여 결정표 형태의 알고리즘을 개발하였다[6]. 이 알고리즘은 '예' 또는 '아니오'로 답하는 8개의 질문으로 이루어져 있는데, 이 방법은 기존의 '원료와의 관련성 여부'를 5척 점수화로 대체하였다. 이 알고리즘은 대규모 임상시험의 분석에 실용적이지만 임상시험 단계 또는 다른 치료법과 같은 다른 원인이 배제되지 않는 한 이 방법으로는 '가능성 있음' 보다 더 높은 등급을 부여하기 어려운 단점을 갖는다.

또 다른 인과 가능성 평가법이 Naranjo 등에 의해 개발되었다(Table 2). 이 평가법은 '예', '아니오', '알려진 바 없음', '적용할 수 없음' 으로 답할 수 있는 10개의 질문으로 구성되어 있으며 통상적인 범주를 이용하여 다양한 임상 상황에서의 인과성을 평가할 수 있도록 개발되었다[7]. 개별 질문에 부여되는 점수는 −1～+2 의 범위에 있는데, 각각의 부작용 사례는 총점에 의해 '확실함', '가능성 높음', '가능성 있음', '가능성 낮음' 으로 분류된다. 총점이 9점 이상이면 '확실함', 5~8점이면 '가능성 있음', 1~ 4점이면 '가능성 있음', 0점 이하이면 '가능성 낮음' 으로 분류된다. Naranjo 알고리즘은 잠재적인 원료 상호작용을 고려하지 않았음에도 불구하고 이를 고려한 Kramer 알고리즘과 평가결과가 잘 일치한다.

미국 FDA는 기존 알고리즘을 검토하면서 상대적으로 단순한 이상반응 평가 알고리즘을 개발하였다[8]. 4가지 특성을 이용하여 인과성을 3가지 범주로 결정하는 이 방법은 빠르고 사용하기 쉽지만 원료에 대한 기존 정보를 고려하지 못한다는 단점이 있다. 또한 하나의 범주가 다른 범주와 겹치기도 하고 인과성이 명확하게 평가되지 못하는 결함이 있다.

'summary time plot'은 다른 알고리즘처럼 일련의 질문을 사용하지 않았다[9]. 이 알고리즘은 원료 섭취와 이상반응 간의 시간적 관련성에 초점을 맞추었는데, 섭취와 이상반응

Table 2. Naranjo scale for causality evaluation of adverse events

Question	Yes	No	Do not know
1. Are there conclusive reports on this reaction?	+1	0	0
2. Did the adverse event appear after the suspected drug was administrated?	+2	-1	0
3. Did the adverse reaction improve when the drug was discontinued or a specific antagonist was administrated?	+1	0	0
4. Did the adverse reaction appear when the drug was readministrated?	+2	-1	0
5. Are there alternative causes (other than the drug) that could on their own have caused the reaction?	-1	+2	0
6. Did the reaction reappear when a placebo was given?	-1	+1	0
7. Was the drug detected in the blood (or other fluid) in concentrations known to be toxic?	+1	0	0
8. Was the reaction more severe when the dose was increased, or less severe when the dose was decreased?	+1	0	0
9. Did the patient have a similar reaction to the same or similar drugs in any previous exposure?	+1	0	0
10. Was the adverse event confirmed by any objective evidence?	+1	0	0

총점	
9 이상	definite
5~8	probable
1~4	posiible
0 이하	doubtful

Table 3. RUCAM scale for causality evaluation of adverse events

1.Time to onset	
Highly suggestive	+3
Suggestive	+2
Compatible	+1
Inconsistent	0
2.Course of the reaction	
Highly suggestive	+3
Suggestive	+2
Compatible	+1
Against the role of the drug	-2
Inconclusive or not available	0
3.Risk factors for drug reaction	
Presence	+1
Absence	0
4.Concomitant drug(s)	
Time to onset incompatible	0
Time to onset incompatible but unknown reaction	-1
Time to onset incompatible and known reaction	-2
Role proven	-3
None or no information	0
5.Non drug-related cause(s)	
Ruled out	+2
Possible or not (depending on the nature of reaction)	+1 to -2
Investigated or labelled	-3
6. Previous information on drug	
Reaction unknown	0
Reaction published but unlabelled	+1
Reaction labelled	+2
7. Response to rechallenge	
Positive	+3
Compatible	+1
Negative	-2
Not available or not Interpretable	0
Or plasma concentration of drug known to be toxic	+3
Or validated laboratory test with high specificity, sensitivity and predictive value	
Positive	+3
Negative	-3
Not interpretable or not available	0

총점	
8 초과	highly probable
6~8	probable
3~5	possible
1~2	unlikely
1 미만	excluded

의 기간을 가로축으로 하고 이상반응의 심각성을 세로축으로 하여 플로팅을 하였다. 이 방법은 시간 요인만을 평가한 것이므로 인과 가능성에 대한 결론에 도달하지 못한다. 그러나 이 방법은 사용하기 쉽고 최소한의 정보만 갖고 있어도 적용이 가능한 장점을 갖는다.

이상반응의 정의와 인과성 평가기준에 대한 국제 전문가 회의 결과 RUCAM (the Roussel Uclaf Causality Assessment Method)가 제안되었다(Table 3). 각각의 인과성 평가기준에는 –3～+3 사이의 점수가 부여되고 총점으로 인과 가능성이 결정된다. 이 방법은 평가자간의 불일치가 설명되며 각 기준에 대한 점수가 전문가 패널에 의해 합의된 것이므로 이전의 다른 평가법에 비하여 재현성이 높다. 이 방법은 사용하기 쉬운 장점이 있지만 간 또는 피부 손상 이외의 다른 임상 분야에 이 방법을 적용하기 위해서는 전문가들의 합의가 필요하다[10].

Maria와 Victorino는 약물에 의한 간 손상(Drug-induced liver injury: DILI)을 진단하는 지표로서 M&V 척도(Table 4)를 개발했다[11]. 5개 항목으로 인과 가능성을 평가하여 합산된 총점은 –6～20 사이의 값을 갖게 되며 총점으로 인과성을 평가한다. DILI의 진단은 복잡하며 정확성을 기하기 위해서는 숙련된 임상 경험자가 필요하다. M&V 척도는 특정 원료에 의한 간 손상의 가능성을 평가하는데 매우 유용하지만 2개 이상의 원료가 관여하는 경우에는 개별 원료별로 척도를 전산화 하는 작업이 필요하게 된다.

Table 4. M&V scale for causality evaluation of adverse events

1. Temporal relationship between drug intake and the onset of clinical picture	
A. Time from drug intake until the onset of first clinical or laboratory manifestation	
4 days to 8 weeks (or less than 4 days in cases of reexposure)	+3
Less than 4 days or more than 8 weeks	+1
B. Time from withdrawal of the drug until the onset of manifestation	
0 to 7 days	+3
8 to 15 days	0
More than 15 days	-3
C. Time from withdrawal of the drug normalization of laboratory values	
Less than 6 months (cholestatic or mixed patterns) or 2 months (hepatocellular)	+3
More than 6 months (cholestatic or mixed) or 2 months (hepatocellular)	0
2. Exclusion of alternative causes	
Viral hepatitis (HAV, HBV, HCV, CMV, EBV), Alcoholic liver disease, Bilary tree obstruction, Preexisting liver disease, Other (pregnancy, acute hypotension)	
Complete exclusion	+3
Partial exclusion	0
Possible alternative cause detected	-1
Probable alternative cause detected	-3
3. Extrahepatic manifestations	
Rash, fever, arthralgia, eosinophilia (>6%), cytopenia	
4 or more	+3
2 or 3	+2
1	+1
none	0
4. Internal or accidental reexposure to the drug	
Positive rechallenge test	+3
Negative or absent rechallenge test	0
5. Previous report in the literature of cases of DILI associated with the drug	
Yes	+2
No (drugs marketed for up to five years)	0
No (drugs marketed for more than five years)	-3
총점	
17 초과	definite
14~17	probable
10~13	possible
6~9	unlikely
6 미만	excluded

Table 5. Comparison for the characteristics of major algorithms

종류	French method	Naranjo scale	RUCAM scale	M&V scale
주요 특징	병력 평가, 임상적생물학적 관찰, 이상반응 진단과 예후 등의 3 단계 평가	통상적인 질문 항목을 이용하여 다양한 임상 상황에서의 인과성을 평가	간 독성 및 피부 손상의 인과관계 평가에 적용	특정 원료에 의한 간 손상의 인과관계 평가에 적용
장단점	병용 약물에 의한 부작용 발생 가능성을 배제할 수 있으나 시간이 많이 소요됨	쉬운 질문으로 구성되어 있어 사용이 편리하고 기존 알고리즘의 결과와 일치성이 높음	사용이 쉽고 재현성이 높으나 간 또는 피부 손상 이외의 다른 임상 분야에 이 방법을 적용하기 위해서는 전문가들의 합의가 필요	평가 구조가 복잡하므로 정확성을 높이기 위해서는 숙련된 임상 경험자가 필요

2005년에는 대부분의 부작용 발생사례로부터 얻을 수 있는 정보로는 답할 수 없는 질문들을 연결, 수정, 제거하는 방식으로 새로운 알고리즘이 개발되었다[12]. 기존 알고리즘의 56개 질문들이 8개의 질문과 새로운 점수판으로 교체되었다. 새로운 알고리즘은 통상적으로 수집되는 부작용 보고정보 이외의 다른 데이터를 요구하지 않으므로 쉽고 빠르게 이용할 수 있지만 인과 가능성을 평가하는 데 도움이 되는 다른 정보들이 결여되어 있다는 단점을 갖는다.

최근에 Horn 등은 약물 상호작용을 평가하기 위하여 Drug Interaction Probability Scale (DIPS)을 제안한 바 있다[13]. DIPS는 10개의 질문을 이용하는데, 양성 반응에는 점수를 더하고 음성 반응에는 점수를 빼는 방식으로 약물 상호작용 가능성을 추정한다. 이 평가법은 약물 상호작용에 의한 부작용 발생 가능성을 추정하는 데 도움을 줄 수 있다.

상기에서 열거한 주요 알고리즘의 특징 및 장단점을 요약하여 Table 5에 나타내었다.

건강기능식품 부작용 인과성 평가를 위한 알고리즘의 개발

전 세계적으로 건강기능식품의 섭취와 부작용 발생간의 인과성을 평가하기 위한 알고리즘이 전무한 상황에서, 식품의약품안전청에서는 기존에 의약품의 부작용 인과성에 적용하기 위해 개발되었던 알고리즘들을 참조하고 건강기능식품의 특성에 적합하도록 Table 6과 같은 건강기능식품 부작용 인과성 평가를 위한 알고리즘을 개발하였다.

구체적인 알고리즘 개발 과정으로는, ① 알고리즘의 전체 틀은 10개의 통상적인 질문 항목으로 구성되어 있어 사용이 편리한 Naranjo scale의 플랫폼을 근간으로 하였으며 ② 질문 항목은 기존 알고리즘에서 사용하고 있는 질문 항목을 총망라한 후 인과관계 평가시의 중요도, 건강기능식품 특성과의 적합성, 기존 알고리즘에서의 사용빈도 등을 고려하여 10개의 질문 항목을 도출하였다. 또한 ③ 각 질문 항목에 대한 평가점수는 가능하면 기존 알고

Table 6. Korean algorithm for causality evaluation of adverse events from Health/Functional Foods

질문	답변 선택	점수
1.섭취한 건강기능식품과 발생한 부작용간의 인과관계에 대하여 알려진 정보가 있는가?	발생한 부작용이 제품에 주의사항으로 표시되어 있다	+3
	제품에 주의사항으로 표시되어 있지 않으나 논문으로 발표된 바 있다	+2
	인과관계에 대한 논란 또는 가설이 제기된 바 있다	+1
	인과관계에 대하여 알려진 정보가 없다	0
2.건강기능식품 섭취에 있어서 제품에 표시된 사용법을 준수하였는가?	준수하였다	0
	준수하기 곤란하였다	+2
	준수하였다	+3
3.건강기능식품 섭취 후 부작용이 나타났는가?	명백히 섭취 후 부작용이 발생하였다	+2
	명확하지는 않으나 섭취 후 부작용이 발생한 것 같다	+1
	섭취 전에 이미 부작용이 발생한 것 같다	0
4.건강기능식품 섭취 중단 후 부작용 증상이 호전되었는가?	섭취를 중단하였더니 부작용 증상이 호전되었다	+3
	섭취를 중단하여도 부작용 증상이 그대로 유지되었다	-2
	섭취를 중단하지 않았다	0
5.건강기능식품을 재섭취 하였을 때 똑같은 부작용이 나타났는가?	재섭취 하였을 때 똑같은 부작용이 발생하였다	+3
	재섭취 하였을 때 똑같지는 않으나 유사한 부작용이 나타났다	+1
	재섭취 하였을 때 부작용이 나타나지 않았거나 다른 형태의 부작용이 나타났다	-2
	재섭취 하지 않았다	0
6.건강기능식품의 특정성분이 혈액검사에서 유해용량 수준으로 검출되었는가?	혈액검사 결과 그렇게 나타났다	+3
	혈액검사 하지 않았거나 혈액검사 결과 그렇지 않았다	0
7.제품 섭취량을 가감했을 때 부작용이 악화/완화되었는가?	그렇다	+1
	제품 섭취량을 가감하지 않았거나 가감했을 때에도 부작용에 변화가 없었다	0
8.기존에 동일한 또는 유사한 제품을 섭취하였을 때 비슷한 반응이 나타났는가?	그렇다	+1
	기존에 동일한 또는 유사한 제품을 섭취하지 않았거나 섭취하였을 때에도 비슷한 반응이 나타나지 않았다	0
9.발생한 부작용의 원인이 객관적 증거로 확인된 바 있는가?	예	+1
	아니오	0
10.동시에 섭취한 의약품 또는 다른 식품에 의한 부작용 발생 가능성은?	가능성이 있음을 입증할 수 있다	-3
	가능성이 없음을 입증할 수 있다	+2
	알 수 없다	0

총점	
13 이상	Certain (확실함)
7~12	Probable (가능성 높음)
3~6	Possible (가능성 낮음)
2 이하	Unlikely (가능성 없음)

리즘에서 사용하고 있는 점수를 준용하였으며 ④ 총점도 기존 알고리즘에서 정하고 있는 수준을 종합적으로 고려하여 설정하였다. 건강기능식품의 특성을 반영한 이러한 알고리즘의 개발은 의약품과는 특성을 달리하는 건강기능식품의 부작용 추정사례에 적용할 수 있는 최초의 시도라는 점에서 큰 의의를 갖는다고 하겠다.

식품의약품안전청에서 개발한 Table 6과 같은 건강기능식품 부작용 인과성 평가를 위한 알고리즘은 2010년 10월부터 '건강기능식품 부작용신고센터(www.hfcc.or.kr)' 에 적용되고 있어, 소비자가 건강기능식품의 부작용 추정사례를 신고한 후 자신이 신고한 부작용 추정사례가 섭취한 건강기능식품과 어느 정도의 인과성을 갖는지에 대해 자가진단 할 수 있도록 운영되고 있다.

참고문헌

1. http://hfoodi.kfda.go.kr/safeinfo/monitoring_result.jsp.
2. Macedo, A.F., Marques, F.B., Ribeiro, C.F., et al. : Causality assessment of adverse drug reactions: comparison of the results obtained from published decisional algorithms and from the evaluations of an expert panel, according to different levels of imputability. *J. Clin. Pharm. Ther.*, **28,** 137-143 (2003).
3. Begaud, B. : Standardized assessment of adverse drug reactions: the method used in France. *Drug Inf. J.*, **18**, 275-281 (1984).
4. Kramer, M.S. : Assessing causality of adverse drug reactions: global introspection and its limitations. *Drug Inf. J.*, **20**, 433-437 (1986).
5. Blanc, S., Leuenberger, P., Berger, J.P., et al. : Judgments of trained Computerobservers on adverse drug reactions. *Clin Pharmacol. Ther.*, **25**, 493-498 (1979).
6. Emanueli, A., Sacchetti, G. : An algorithm for the classification of untoward events in large scale clinical trials. *Agents Actions Suppl.*, **7**, 318-322 (1980).
7. Naranjo, C.A., Busto, U., Sellers, E.M., et al. : A method for estimating the probability of adverse drug reactions. *Clin Pharmacol. Ther.*, **30**, 239-245 (1981).
8. Jones, J.K. : Adverse drug reactions in the community health setting: approaches to recognizing, counseling, and reporting. *Fam. Community Health.*, **5**, 58-67 (1982).
9. Castle, W.M. : Assessment of causality in industrial settings. *Drug Inf. J.*, **18**, 297-302 (1984).
10. Benichou, C., Danan, G. : Causality assessment in the European pharmaceutical industry: presentation of the preliminary results of a new method. *Drug Inf. J.*, **26**, 589-592 (1992).
11. Maria, V.A., Victorino, R.M. : Development and validation of a clinical scale for the diagnosis of drug-induced hepatitis. *Hepatology.*, **26**, 664-669 (1997).
12. Koh, Y., Shu, C.L. : A new algorithm to identify the causality of adverse drug reactions. *Drug Saf.* **28**, 1159-1161 (2005).
13. Horn, J.R., Hansten, P.D., Chan, L.N. : Proposal for a new tool to evaluate drug interaction cases. *Ann. Pharmacother.*, **41**, 674-680 (2007).