

# 중증도 보정 모형 개발을 위한 가이드라인 개발



김경훈 부연구위원  
건강보험심사평가원 국제협력특계팀

## 1. 들어가며

중증도 보정 모형 개발은 의료제공자별 재입원, 사망과 같은 결과(outcome)를 공평하게 평가하기 위하여 결과에 영향을 주는 요인들을 선정하는 과정이다. 1986년 보건 의료재정청(Health Care Financing Administration, 현 Centers for Medicare & Medicaid Services)이 보정되지 않은 의료기관별 사망률을 발표함에 따라(Blumberg, 1987), 의료기관 성과평가에서 중증도 보정 모형(risk adjustment model) 개발의 중요성이 부각되었다.

의료기관 성과평가에 대한 결과 공개는 사회·경제적 파급 효과가 크다는 점에서 객관적인 평가결과 도출을 위한 노력이 요구된다. 일례로, 미국과 캐나다 등은 포괄적인 자료수집 정보를 토대로 결과를 잘 설명하는 중증도 보정 모형을 개발하여 객관적인 평가결과를 도출하고 있다(PHC4, 2012; CIHI, 2013). 또한 중증도 보정 모형 개발 과정을 기록한 기술 보고서(technical note)를 발간하여 평가 과정의 신뢰성을 확보하고 있다.

현재 건강보험심사평가원의 적정성 평가 영역은 확대되고 있으며, 이와 더불어 객관적인 평가결과 도출을 위한 중증도 보정 모형의 중요성이 강조되고 있다. 따라서 적정성 평가의 객관성을 담보하기 위해서는 보다 체계적이고 효율적인 중증도 보정 모형 개발을 위한 가이드라인 개발이 요구되고 있는 실정이다. 이에 본고에서는 중증도 보정 모형 개발 과정을 단

계별로 제시하고, 각 단계별 고려사항들을 정리함으로써 적정성 평가 분석의 가이드라인 개발 과정을 살펴보고자 한다.<sup>1)</sup>

## 2. 중증도 보정 모형이란?

의료서비스의 결과는 환자 요인(patient risk factor), 치료효과(treatment effectiveness), 의료의 질(quality of care), 랜덤효과(random chance)에 의해 결정된다(그림 1). 이러한 결정요인들 중에서 랜덤효과는 통계적인 방법을 통하여 통제가 가능하다. 그리고 치료효과가 동일하다고 가정할 경우, 환자 요인을 보정하면 의료의 질은 환자의 결과와 동일한 형태로 나타나게 된다. 따라서 중증도 보정 모형은 결과에 영향을 주는 다양한 환자 요인들 중에서 결과를 가장 잘 설명해주는 요인으로 구성된 모형이다.



그림 1. 중증도 보정의 개념적 틀

의료서비스의 결과를 공평하게 평가하는 것은 쉽지 않다. 그 이유는 첫째, 환자의 기저위험(baseline risk)이 다양하기 때문이다. 예를 들어, A병원과 B병원에서 관상동맥우회술을 받은 환자의 실제 사망률이 동일하다고 가정한다면, A병원이 B병원 보다 더 위험한 환자를 수술한다고 해도 객관적인 평가 결과는 두 병원의 의료의 질이 다르다고 나타날 것이다. 둘째, 일반적으로 이상 결과(adverse outcome)는 빈번하게 발생하지 않으므로, 실제 평가 결과는 랜덤변이(random variation)에 더 많은 영향을 받게 된다. 따라서 평가 결과를 비교하기 위해서는 통계적 방법을 통하여 환자의 특성 차이와 불확실성(uncertainty)을 반영할 수 있어야 한다. 이들 요소가 적절하게 통제된다면, 평가 결과의 차이는 환자 요인보다 의료 제공자의 행태(behaviors)에서 기인한 것이라고 볼 수 있다.

1) 이 글은 2013년 건강보험심사평가원 연구보고서 「적정성 평가자로 분석을 위한 가이드라인 개발」 연구 내용을 요약 발췌한 것으로, 자세한 내용은 보고서를 참고하시기 바랍니다.

### 3. 중증도 보정 모형 개발 과정

중증도 보정 모형 개발과정은 준비단계와 개발단계로 구분된다. 우선, 준비단계에는 모형개발 대상 선정, 사전 고려사항 검토, 후보요인 목록 작성, 단변량 분석이 수반된다. 그리고 개발단계에는 모형 선정과 타당도 평가를 통해 최종적으로 중증도 보정 모형을 선정하는 과정이 포함된다.



그림 2. 중증도 보정 모형 개발 과정

#### 가. 대상 선정

모형 개발 대상은 선정 및 제외기준을 만족하는 결과지표 산출 환자를 중심으로 선정한다.

#### 나. 사전 고려사항 검토

##### 1) 자료의 질 평가

자료의 질은 일반적으로 수집 변수의 평균과 빈도 등을 통하여 판단한다. 그리고 자료의 질 평가에서는 이상치 확인이 수행되며, 일반적으로 결측값 처리방법을 통해 확인한다.

결측값을 대체하는 방법으로는 ①결측이 있는 환자를 제외하고 완전한 자료로 분석하는 방법(complete case analyses 또는 listwise deletion), ②결측값을 하나의 범주로 고려한 방법(missing indicator), ③결측값을 하나의 값으로 대체하는 단일 대체 방법(single imputation), 그리고 ④결측값을 다양한 값으로 대체하는 다중 대체 방법(multiple imputation)이 있다. 이와 같이, 결측값을 대체하는 다양한 통계적 방법론들이 소개되고 있으나 결측값을 무작위(random)라고 가정하고 있으므로, 단순하게 적용하기에 적절하지 않은

부분이 존재한다. 왜냐하면, 평가 자료에 존재하는 결측값은 대부분 완전히 무작위가 아니며, 일반적으로 다른 변수들과 간접적 혹은 직접적으로 관련되어 있기 때문이다. 따라서 자료가 결측값을 포함하고 있을 경우, 하나의 결측값 대체 방법을 적용하여 모형을 개발하기 보다는 자료의 특성을 고려한 결측 대체방법론을 비교·분석함으로써 분석결과의 일치성(consistency)을 증명할 필요가 있다.

## 2) 변수형태 결정

변수형태를 결정한다는 것은 새로운 변수를 생성하거나, 변수의 속성을 변경하는 것에 해당된다. 수집된 변수를 통하여 새로운 변수를 생성하는 경우, 그 근거를 명확히 하고 변수통합에 따른 결과 변화를 파악해야 한다.

변수의 속성은 크게 연속형과 범주형 변수로 구분할 수 있다. 자료의 특성을 잘 반영한 모형을 개발하기 위해서는 변수의 속성은 그대로 유지하는 것이 바람직하다(Royston 등, 2006). 자료 분석과 해석이 용이하다는 이유로 변수 속성을 변경하는 경우가 있으나, 임상적 근거가 배제된 채 변수의 분포만을 고려하여 범주 구분을 위한 기준을 설정하는 것은 심각한 편향(bias)을 가져올 수 있다.

## 다. 후보요인 목록 작성

후보요인 목록 작성은 모형 개발의 기초 과정이다. 후보요인 목록은 수집된 조사자료 범위 내에서 결과에 영향을 줄 것으로 예상되는 변수로 구성된다. 그리고 후보요인 목록은 변수의 평균, 빈도, 결측률 등 기초 분석결과를 토대로 전문가의 합의과정을 거쳐 작성된다.

후보요인은 크게 일반적 특성, 동반질환, 수술 과거력, 임상정보, 환자상태, 질환 중증도 등으로 구분할 수 있다. 중증도 보정에 포함되는 요인은 입원 전 혹은 입원 초기 요인을 대상으로 한다. 후보요인은 결과와 관련 있는 모든 변수를 대상으로 하며, 관련성에 대한 인과관계(causal)는 고려하지 않는다(Moons 등, 2012).

후보요인 선정기준에는 자료수집의 신뢰성, 임상적 중요성, 통계적 유의성이 포함된다. 먼저, 자료수집 과정에서 신뢰성이 낮은 변수는 후보요인에서 제외하고, 의미가 유사한 변수는 통합한다. 특히, 빈도가 적고, 임상적으로 의미가 유사한 변수들은 하나의 변수로 통합할 수 있다. 이 과정을 거쳐 재작성된 변수 목록은 임상적 중요성과 통계적 유의성을 기준으로 후보요인으로 작성한다. 그러나 단변량 분석 결과 통계적으로 유의하지 않더라도 다른 변수와 상호작용(interaction)하여 결과에 유의하게 영향을 미칠 수 있는 요인이 존재할

수 있으므로(Sun 등, 1996; Steyerberg 등, 2001), 후보요인 목록 작성 시 임상적 판단이 우선적으로 이뤄져야 한다.

후보요인 목록 작성 과정은 체계적으로 정리되어야 하며, 후보요인에 대한 선정근거는 명확하게 제시되어야 한다. 마지막으로, 선정된 후보요인의 변수명, 변수 속성, 범주 설명 등이 포함된 후보요인 목록표를 작성한다.

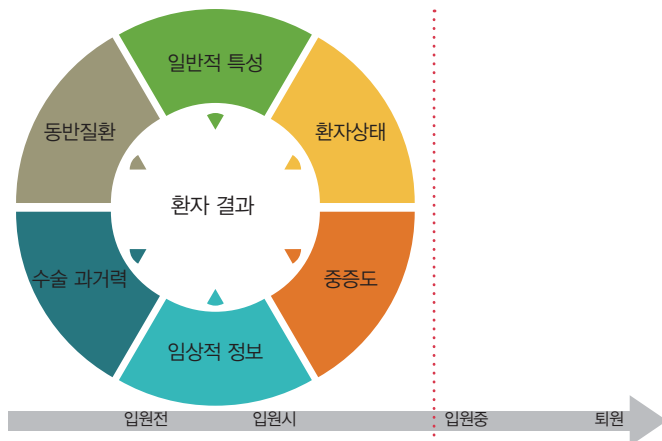


그림 3. 후보요인의 범위

### 라. 단변량 분석

후보요인 목록이 최종적으로 확정되면, 후보요인에 대한 기초 분석을 실시한다. 기초 분석 시 연속형 변수는 평균과 표준편차(혹은 중앙값)를 제시하고, 범주형 변수는 빈도와 퍼센트를 제시한다. 특히, 후보요인이 결측값을 포함하고 있는 경우에는 그 분포를 제시한다. 그리고 통계적 검정을 실시하여 후보요인과 결과와의 관련성을 분석한다.

표 1. 후보요인과 재입원 여부의 단변량 분석결과 정리 예시

변수	범주	전체	재입원 여부		P-값 <sup>2</sup>
			아니오	예	
연령 <sup>1</sup>		64.2(66.0)	64.0(65.0)	70.1(71.0)	<.0001
성별	남성	1,975(70.9)	1,922(97.3)	53(2.7)	0.0548
	여성	811(29.1)	778(95.9)	33(4.1)	
체질량지수 <sup>1</sup>		24.6(24.5)	24.7(24.6)	23.4(22.9)	0.0001
7일 내 AMI 발생여부	아니오	2,418(86.8)	2,355(97.4)	63(2.6)	<.0001
	예	265(9.5)	244(92.1)	21(7.9)	
	결측	103(3.7)	101(98.1)	2(1.9)	

주 1. 연령과 체질량지수는 평균과 중앙값을 제시하였다.

2. P-값은 독립표본 t-검정과 카이제곱 검정 결과이다.

## 마. 모형 선정

모형 선정은 후보요인 중에서 중증도 보정 모형에 포함될 결과를 잘 설명해주는 변수를 선정하는 과정으로, 자료의 수, 후보요인의 수, 결과 발생률 등을 토대로 결정된다. 변수선정 방법으로는 첫째, 모든 후보요인을 모형에 포함시키는 방법, 둘째, 임상적 판단에 근거하여 일부 후보요인을 선정하는 방법, 셋째, 통계적 방법이 있다.

임상적 판단에 근거하여 후보요인을 선정하는 방법은 임상적 측면을 고려했다는 점에서 타당한 방법이라고 할 수 있으나, 전문가들마다 서로 다른 의견이 존재할 수 있다. 변수를 선정하는 통계적 방법으로는 단변량 분석결과를 근거로 선정하는 방법과 변수선택법으로 선정하는 방법이 있다. 단변량 분석결과를 근거로 선정하는 방법은 변수들 개별적으로는 유의한 관련성이 없으나, 다른 변수와 결합하여 결과에 유의한 영향을 주는 변수를 누락시킬 가능성이 있다. 변수선택법에는 전진선택법(forward selection), 후진제거법(backward selection), 단계적 선택법(stepwise selection), 최적선택법(best subset)이 있다. 그러나 변수선택법은 미미한 자료의 변화(추가 또는 삭제)에도 불안정한 결과를 가져오기도 하고, 회귀계수의 표준오차가 낮게 편향되기도 하는 등 모형 설명력이 과추정될 가능성이 있다 (Steyerberg 등, 2000).

이와 같이, 변수선정은 후보 요인 중에서 결과를 잘 설명해주는 변수 혹은 변수의 집합을 선정하는 과정이다. 아직까지 모형에 포함될 변수를 선정하는 최선의 방법에 대한 합의는 이루어지지 않았으며, 단순히 통계적 접근에 의존하는 경향이 있다. 그러나 변수선정은 통계적 접근과 더불어 임상적 판단이 요구된다. 즉, 모형에 포함될 후보요인은 모형개발의 목적, 자료의 특성 등을 반영한 통계적 접근뿐만 아니라 임상적 판단을 근거로 선정되어야 한다.

## 바. 타당도 평가

중증도 보정 모형이 개발되면, 타당도 평가를 통하여 모형이 결과를 잘 설명하고 있는지에 대해 검토해야 한다. 모형 타당도 평가방법에는 명백한 내적 타당도(apparent internal validation), 내적 타당도(internal validation), 외적 타당도(external validation)가 있다. 명백한 내적 타당도 방법은 모형 개발에 사용된 원자료(raw data)를 통해 타당도를 평가하는 방법으로, 높은 타당도 평가 결과로 편향될 가능성이 높다. 따라서 개발된 모형이 자료를 얼마나 잘 설명하고 있는지, 자료의 변화에도 얼마나 안정적인지를 평가하기 위하여 내적 타당도 평가를 실시해야 한다. 일반적인 내적 타당도 평가방법에는 교차타당도, 붓스트

랩 타당도 평가 등이 있다.

개발된 모형이 자료를 얼마나 잘 설명하는지를 평가하는 기준으로는 전반적 모형 수행력(overall performance), 판별력(discrimination), 교정력(calibration) 등이 있다. 판별력은 결과가 발생한 집단과 결과가 발생하지 않은 집단이 얼마나 잘 구분되고 있는지를 측정하는 것으로, 일반적으로 C-통계량을 사용한다. 교정력은 실제 결과발생과 예측 발생률의 일치도를 평가하는 것으로 Hosmer-Lemeshow의 카이제곱 검정이 많이 사용된다. 이밖에도, 교정력을 평가하는 다른 척도로는 교정그림(calibration plot), 교정 기울기(calibration slope) 등이 있다.

표 2. 모형 수행력 측정 영역 및 기준

영역	기준	설명
전반적 모형 수행력	$R^2$ , Brier Score	실제 값과 예측 값의 차이
판별력	C-statistic Discrimination slope	결과가 발생한 그룹과 발생하지 않은 그룹을 짝을 지어 비교
교정력	Calibration-in-the-large	평균적인 예측 결과발생률과 평균적인 실제 결과발생률의 차이 비교
	Calibration slope	회귀계수와의 선형적인 형태
	Hosmer-Lemeshow test	예측결과 발생률을 10그룹으로 구분하여 분포 비교

## 4. 나가며

중증도 보정 모형 개발은 의료기관의 객관적인 성과평가에 있어 중요한 영역이다. 환자의 중증도를 제대로 반영하지 않은 의료기관의 평가결과는 사회경제적 파급 효과가 크다는 점에서 중증도 보정모형에 대한 관심이 증대되고 있다. 따라서 본 연구에서는 중증도 보정 모형 개발 과정의 효율성을 향상시키고, 객관적인 평가결과 도출을 위하여 모형 개발 가이드라인을 개발하였다.

본 가이드라인은 모형 개발과정을 단계별로 제시하고, 각 단계별로 고려해야 될 사항을 정리해 놓은 것으로, 단계에 따라 자료를 분석함으로써 분석 과정에서 누락되는 부분을 확인할 수 있다. 그리고 가이드라인에서는 모형 개발과정에서 발생할 수 있는 불확실성을 감소시키기 위한 통계적 방법론들을 소개하고 있으므로, 자료와 변수의 수, 결측값 분포 등 자료의 특성을 토대로 분석방법을 선택할 수 있다. 가이드라인은 내부적으로는 분석과정의 효

올성을 향상시키고, 외부적으로는 평가결과의 신뢰성을 담보하는 계기가 될 것으로 판단된다. 최근 의료기관 성과평가를 위한 다양한 통계방법론들이 개발되고 있다. 따라서 본 연구에서 개발한 가이드라인 역시 국내외 성과평가 방법론들을 모니터링하여 정기적으로 보완해야 할 것이다. 🌐

## 참고문헌

---

- Blumberg MS. Comments on HCFA hospital death rate statistical outliers. *Health Care Financing Administration, Health Serv Res.* 1987;21(6):715-39.
- Canadian Institute for Health Information. Canadian hospital reporting project: Technical Note: Clinical Indicator. 2013 CIHI.
- Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, Grobbee DE. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart.* 2012;98(9):683-90.
- Pennsylvania Health Care Cost Containment Council. Technical notes for the hospital performance report. 2012.
- Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med.* 2006;25(1):127-41.
- Steyerberg EW, Eijkemans, MJC, Harrell Jr FE, Habbema, JDF. Prognostic modeling with logistic regression analysis: a comparison of selection and estimation method in small data sets. *Statistics in Medicine* 2000; 1059-79.
- Steyerberg EW, Eijkemans MJ, Harrell FE Jr, Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making.* 2001;21(1):45-56.
- Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol.* 1996;49(8):907-16.