

# 중환자 중증도 평가도구의 타당도 평가 - APACHE III, SAPS II, MPM II

김은경, 권영대<sup>1)</sup>, 황정해<sup>2)</sup>

울지외과대학교 간호대학, 성균관대학교 의과대학 사회의학교실<sup>1)</sup>, 건강보험심사평가원 조사연구실<sup>2)</sup>

## Comparing the Performance of Three Severity Scoring Systems for ICU Patients: APACHE III, SAPS II, MPM II

Eun-Kyung Kim, Young-Dae Kwon<sup>1)</sup>, Jeong-Hae Hwang<sup>2)</sup>

School of Nursing, Eulji University; Department of Social and Preventive Medicine College of Medicine Sungkyunkwan University<sup>1)</sup>; Research Department, Health Insurance Review Agency<sup>2)</sup>

**Objectives :** To evaluate the predictive validity of three scoring systems; the acute physiology and chronic health evaluation(APACHE) III, simplified acute physiology score(SAPS) II, and mortality probability model(MPM) II systems in critically ill patients.

**Methods :** A concurrent and retrospective study conducted by collecting data on consecutive patients admitted to the intensive care unit(ICU) including surgical, medical and coronary care unit between January 1, 2004, and March 31, 2004. Data were collected on 348 patients consecutively admitted to the ICU(aged 16 years or older, no transfer, ICU stay at least 8 hours). Three models were analyzed using logistic regression. Discrimination was assessed using receiver operating characteristic(ROC) curves, sensitivity, specificity, and correct classification rate. Calibration was assessed using the Lemeshow-Hosmer goodness of fit H-statistic.

**Results :** For the APACHE III, SAPS II and MPM II systems, the area under the receiver operating characterist

ic(ROC) curves were 0.981, 0.978, and 0.941 respectively. With a predicted risk of 0.5, the sensitivities for the APACHE III, SAPS II, and MPM II systems were 81.1, 79.2 and 71.7%, the specificities 98.3, 98.6, and 98.3%, and the correct classification rates 95.7, 95.7, and 94.3%, respectively. The SAPS II and APACHE III systems showed good calibrations(chi-squared H=2.5838 p=0.9577 for SAPS II, and chi-squared H=4.3761 p=0.8217 for APACHE III).

**Conclusions :** The APACHE III and SAPS II systems have excellent powers of mortality prediction, and calibration, and can be useful tools for the quality assessment of intensive care units(ICUs).

J Prev Med Public Health 2005;38(3):276-282

**Key words :** Intensive care units(ICUs), Acute Physiology and Chronic Health Evaluation(APACHE), Calibration, Mortality

## 서론

중환자실에 입실된 환자는 임상적 병리 상태가 불안정하거나 자주 변화할 수 있으며, 이들 환자의 회복 가능성은 질병의 종류, 질병의 중증도, 환자의 나이, 과거의 건강상태 및 이용가능한 치료방법 등에 좌우되며, 특히 중환자실 입실 24시간 내의 질병의 중증도는 환자의 회복에 매우 중요한 영향을 미친다 [1]. 임상에서 중환자들을 적절하게 치료하기 위해서는 각 환자들의 중증도를 평가하고, 어떠한 조치를 행할 것인가에 대한 정확한 평가가

요구된다.

또한 중환자실의 치료효과를 평가하기 위해 환자의 실제 사망률에만 기초하는 경우 병원 간 중환자실 환자의 질병과 중증도 구성의 다양함으로 인하여 정확한 비교가 될 수 없다 [2,3]. 그러므로 효율적인 중환자관리를 위해서는 다양한 진료환경과 의료기술로 치료받는 환자를 비교 분석함에 있어 중증도 조절이 선행되어야 하고, 그 외 정확한 상태파악과 치료의 결과에 대한 평가가 이루어져야만 객관성을 부여하게 된다 [4].

중환자의 중증도를 정확히 평가하면 각

중환자실 간의 임상성적에 대한 정보교환 및 질적인 평가가 가능해지며, 새로운 치료법의 결과를 기존의 치료법에 의한 결과와 비교 분석하여 새로운 치료법에 대한 확실한 평가가 가능한 이점이 있다. 또한 중환자의 가족에게 환자결과와 가능성과 치료적 의사결정에 대한 정보제공에 유용할 뿐 아니라 중환자실의 의료 인력 및 비용의 적절한 운영을 위한 방향을 제시할 수 있어 중환자실의 임상적 관리와 자원할당에 대한 행정적인 관리에 중요한 지침을 제공하기도 한다 [5-11].

그동안 중환자의 예후를 측정하기 위해 여러 국가들에서 중증도 측정체계 (scoring system)를 개발하여 중환자실에 입실한 환

접수: 2004년 8월 29일, 채택: 2005년 2월 16일

이 논문은 2003년도 병석학술경과재단 연구비의 지원에 의하여 이루어진 것임.

책임저자: 김은경(대전광역시 중구 용두2동 143-5, 전화: 042-259-1716, 팩스: 042-259-1709, E-mail: kek@eulji.ac.kr)

자의 치료에 대한 예후와 건강상태를 평가하는데 사용하고 있다 [12-20]. 이러한 측정체계들은 중환자의 사망여부를 치료 결과로 잘 반영하는 도구들이라고 할 수 있다 [19,20].

중증도 평가도구 중에서 현재 APACHE (acute physiology and chronic health evaluation), SAPS (simplified acute physiology score), MPM (mortality prediction model) 등의 측정체계가 가장 보편적으로 사용되고 있다 [5,6,12,21-25]. 이들은 최근 각각 APACHE III와 SAPS II, MPM II로 개정되었고, 미국과 유럽에서는 각 모형이 중환자들의 사망률을 예측하는데 있어 유용성이 뛰어난 것으로 평가되었다 [6,13,23].

APACHE III와 SAPS II, MPM II는 각각 사망 예측치 계산을 위해 필요한 변수들을 포함하고 있으며, 모두 사망과 생존을 명확히 분별해 내는데 주안점을 두고 개발되었다. APACHE III는 다른 도구들에 비해 가장 많은 18개의 변수를 포함하고 있으며, 모두 134개의 가중치를 가지기 때문에 가장 복잡하고, 시간이 많이 걸리며, 점수의 범위가 0점에서 299점까지로 매우 크다. SAPS II는 많은 부분 APACHE III의 생리적 변수들을 포함하여 축약된 형태로 15개의 변수를 포함하여 55개의 가중치를 두고 있으며, 점수의 범위는 0점에서 163점이다. MPM II는 12개의 변수를 포함하며, 적용여부뿐만 판단하고 있어 가장 간단하고 시간이 걸리지 않는다는 특징이 있다.

만일 사망 예측치 계산을 위해 필요한 변수들을 최소화하면서도 사망과 생존을 명확히 분별해 낼 수 있는 모델이라면 가장 이상적인 평가도구가 될 것이다. 아직까지 우리나라에서는 응급실 내원 급성 심근경색증 환자를 대상으로 세 도구를 적용한 Koh [26]의 연구를 제외하고는 전체 동일 중환자에게 APACHE III, SAPS II, MPM II를 동시에 적용하여 중환자의 예후를 예측할 수 있는 도구로서의 타당도를 비교한 연구는 없었다.

이에 본 연구는 중환자의 예후 예측을 위해 보편적으로 사용되고 있는 측정체계 중 APACHE III, SAPS II, MPM II를 동일 중

환자들에게 적용하여 도구의 타당도를 비교·평가하고자 수행하였다.

## 연구대상 및 방법

### I 연구대상

일개 광역시에 소재한 800명상 규모의 종합전문요양기관에 2004년 1월 1일부터 3월 31일까지 내·외과계 중환자실에 입실한 환자 전수를 연구대상으로 하였다. 조사기간 동안 중환자실에 입원한 전체 연구대상자수는 조사기간 동안 중환자실에 입실하여 최소한 8시간 이상 머물렀던 만 16세 이상의 환자로 반복 입원한 경우는 마지막 입원을 조사대상으로 포함시켰다. 또한 진료결과가 자의퇴원이거나 다른 병원으로 전원한 환자도 제외시켰다. 조사기간 동안 중환자실에 입원한 전체 대상 건수는 419건이었으나, 조사완료시점까지 계속 입원하고 있는 환자 9명, 중환자실에 6시간 이내 머물렀던 환자 3명, 16세 미만 환자 1명, 자의퇴원 환자 13명, 다른 병원으로 전원한 환자 4명을 제외하였고, 동일 환자가 반복 입원한 경우 41건을 제외한 348건을 분석하였다.

### 2. 연구방법

#### 1) 자료수집

자료수집은 조사도구의 모든 항목에 대해 교육받은 중환자실 간호경력 5년 이상의 간호사 1인이 환자의 중환자실 입실기간 동안 매일 수행하였다. 결과 입력은 중환자실 입실 후 24시간 이내에 시행한 검사 결과 중에서 가장 중증도 점수가 높은 검사 결과를 채택하였다. 이는 환자의 상태가 가장 악화된 중증도 측정값으로 중증도가 높은 환자일수록 높은 점수를 받을 수 있도록 하였다. 진료결과는 퇴원 시 최종 상태(생존/사망)로 후향적인 의무기록조사로 수행되었다.

#### 2) 조사도구

##### (1) APACHE III

APACHE 도구는 중환자들에서 병태 생리학적으로 공통적으로 발생하는 특성을 정량화 하려는 시도로 가장 널리 인용되고 있다. 최초의 APACHE System은 1981년

Knaus 등에 의해 개발되었고 [8], 파악해야 하는 생리지표수가 많아 1985년 APACHE II로 개정되었다. 1991년에 개정된 [6] APACHE III는 APACHE II System에 생리지표들이 정상치로부터의 간격 정도가 커질수록 생리점수에 가중치를 더 부여하여 환자의 예후예측력을 향상시키고자, 기존의 생리적 변수(맥박수, 평균혈압, 체온, 호흡수, 맥산소압, 혈색소, 백혈구수, 크레아티닌, 산염기장애)에 5개의 생리지수 (BUN, urine output, serum, albumin, bilirubin, glucose)를 추가하고 Glasgow Coma Scale의 점수 부여방식을 변화시켰다. APACHE III 도구는 생리적 변수(0-252점), 만성건강상태(0-23점), 연령(0-24점)으로 구성된 도구로 총 점수는 최소 0점에서 최고 299점으로 점수가 높을수록 중증도가 높은 환자임을 의미한다.

APACHE III 도구에 대한 비교연구에서 APACHE III 도구가 APACHE II 도구에 비해 예후 예측에 대한 정확도가 더 보장되었음이 입증되었다 [27,28]. APACHE III를 적용한 연구들에서 0.701~0.918의 사망예측력을 보고하였다 [6,29-31].

##### (2) SAPS II

Le Gall [13]이 중환자실로의 입실 첫 24시간 이내의 자료들 중 15개의 변수를 이용하여 예후를 예측하고자 SAPS II를 고안하였다. 이는 중환자실로 입원 후 24시간 이내에 얻을 수 있는 34개의 생리적인 변수를 이상 정도에 따라 0~4점으로 채점하여 합산해서 구한 APS(acute physiology score)의 변형으로 자료수집 과정에서 결과치가 없어 소실되는 오차를 최소화하기 위해 간편화한 것이었다. 몇 가지 생리적 점수(serum K<sup>+</sup>, HCO<sub>3</sub><sup>-</sup>)를 제외하고는 APACHE III의 생리적 변수와 중복되고, 입원형태만 추가되었다. SAPS 점수가 증가할수록 중증도가 높음을 의미한다.

SAPS II는 간단하며 경제적인 장점이 있으나 사망률 예측에 있어서는 APACHE III와 비교할 때 정확성이 떨어진다는 지적도 있다 [32]. 그러나 SAPS II를 적용한 연구들에서 0.672~0.86의 사망예측력을 보고하였다 [22,24,33].

**Table 1.** Baseline characteristics of sample population unit, number (%)

Characteristic	Number (%)
Total number of patients	348(100.0)
Gender	
male	214 ( 61.5)
female	134 ( 38.5)
Age [years, median(range)]	63.0 (18-96)
Days at intensive care unit [median(range)]	3.0 ( 1-57)
Reason of admission	
medical admission	266 ( 76.4)
elective operation	46 ( 13.2)
emergency operation	26 ( 7.5)
other	10 ( 2.9)
Route of admission	
emergency room	205 ( 58.9)
ward room	80 ( 23.0)
operating room	59 ( 17.0)
other	4 ( 1.1)
Disease category	
Cardiovascular	82 ( 23.6)
Neurologic	47 ( 13.5)
Metabolic	42 ( 12.1)
Gastrointestinal	35 ( 10.1)
Respiratory	27 ( 7.8)
Infection	26 ( 7.5)
Toxic	25 ( 7.2)
Surgical	14 ( 4.0)
Other	7 ( 2.0)
Clinical result	
Survived	295 ( 84.8)
Death	53 ( 15.2)

(3) MPM II

MPMII는 MPM\*II로 조사되었다. MPM\*는 중환자실 입원 24시간 동안의 만성질환(간경화, 뇌내출혈, 전이된 악성종양)과 24시간 내 건강상태(혼수나 깊은 혼미상태, 크리아티닌치, 확진된 감염, 기계적 환기장치, 산소분압, 프로트롬빈 시간, 8시간 소변양, 혈관수축제 투약), 기타(내과적이거나 계획되지 않은 수술)로 모두 12개의 변수로 이루어져 있다. 중환자실 입실 1시간 전후의 상태를 측정하는 것으로 MPM\*도 있으나 본 연구에서는 다른 도구들과 동일기간의 결과 비교를 위해 MPM\*II를 적용하였다. MPM II는 모두 쉽게 얻을 수 있는 변수들로 구성되어 있고, 변수가 적어 자료수집에 부담이 적고 실수할 가능성이 줄어들어 자료의 신빙성이 높다는 장점이 있다. MPM II를 적용한 연구들에서 0.69~0.956의 사망예측력을 보고하였다 [23,34].

3) 분석방법

3가지 도구의 사망예측 타당도를 평가하기 위해 SAS 8.1을 이용하여 다음의 분석방법을 적용하였다. 각 도구의 사망예측 타당도를 보기 위해 ROC(receiver operating characteristics) curve를 구하였다. ROC 곡선

은 각 도구에 의해 사망환자비율을 정확하게 예측하고 있는지를 민감도와 1-특이도로 도시하며, ROC 곡선 아래의 영역인 c-통계량으로 평가하였다. ROC 곡선하의 면적을 계산하는 방법은 전체 판측된 값의 쌍 중 예측결과와 관찰결과가 실제 일치된 쌍이 차지하는 비율로 0~1의 값을 가지며, 0.5는 예측능력이 없고, 1은 완벽한 예측능력을 의미한다.

각 도구별 사망의 판단기준은 몇 개의 cutoff에 따른 2x2 분할표를 이용하여 민감도(sensitivity), 특이도(specificity), 정확도(accuracy)를 구하여 비교하였다.

각 모델의 적합도는 Hosmer-Lemeshow goodness-of-fit test(chi-square H)를 이용하였다. 이것은 각 모델들의 측정을 평가하고 전체 확률범위 이상의 예측결과와 측정결과 사이의 연관성의 강도를 나타내는데 사용된다. 측정결과는 사망형태를 기술하고, 각 모델에 의한 위험예측 정확도를 지적할 수 있는 각 도구의 능력을 정의한다.

연구결과

I. 대상군의 임상적 특성

중환자실에 입실한 연구 대상환자의 일반적 특성을 살펴본 결과는 다음과 같다 (Table 1). 대상자의 61.5%가 여성이었으며, 연령분포 범위는 18세에서 96세였고, 중환자실 입실기간은 최단 1일에서 최장 57일까지로 나타났다. 중환자실 입실의 이유는 내과적 입실이 266명(76.4%)으로 대부분을 차지하였으며, 다음으로 예정수술(13.2%), 응급수술(7.5%) 순으로 나타났다. 중환자실 입실 경로는 응급실을 경유하여 입원한 경우가 205명(58.9%)으로 가장 많았다. 또한 중환자실 입실환자의 질병을 분류한 결과, 심혈관계 질환이 전체 환자의 82명(23.6%)으로 가장 많았으며, 다음으로 신경계 질환(13.5%), 대사성 질환(12.1%), 소화기계 질환(10.1%), 감염(7.5%)의 순으로 나타났다. 퇴원 시 최종 임상결과로는 사망이 전체 대상 환자의 15.2%로 나타났다.

2. 사망예측 타당도 평가

세 측정도구의 사망예측 타당도는 ROC 곡선하의 면적과 민감도, 특이도, 정확도를 포함한 판별력과 적합도로 평가하였다.

측정도구가 실제 생존자와 사망자를 일

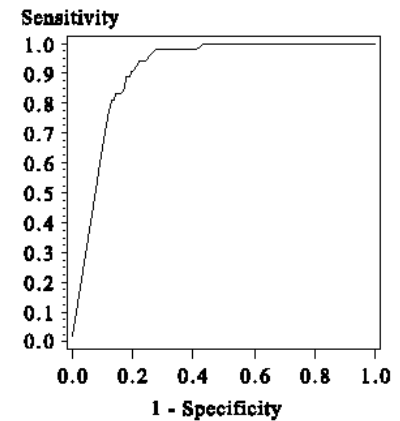


Figure 1. ROC curve of APACHE III score. (The area under the curve is 0.981)

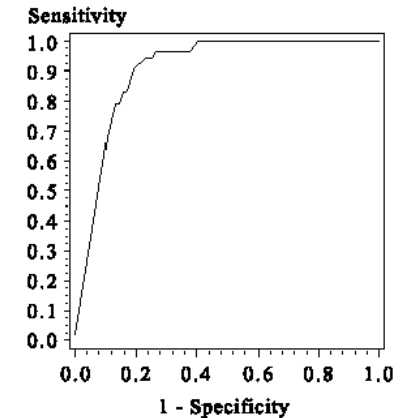


Figure 2. ROC curve of SAPS II score. (The area under the curve is 0.978)

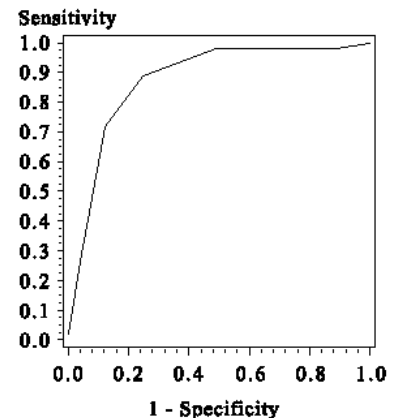


Figure 3. ROC curve of MPM II score. (The area under the curve is 0.941)

**Table 2** Comparing the sensitivity(Se), specificity(Sp) and accuracy(Ac) of the APACHE III, SAPS II and MPM II at decision thresholds of 0.3, 0.4, 0.5, 0.6 and 0.7. (values are percent)

Cutoff value	APACHE III			SAPS II			MPM II		
	Se	Sp	Ac	Se	Sp	Ac	Se	Sp	Ac
0.3	83.0	95.9	94.0	83.0	94.6	92.8	71.7	98.3	94.3
0.4	83.0	96.6	94.5	83.0	96.6	94.5	71.7	98.3	94.3
0.5	81.1	98.3	95.7	79.2	98.6	95.7	71.7	98.3	94.3
0.6	79.2	99.3	96.3	73.6	99.0	95.1	71.7	98.3	94.3
0.7	77.4	99.7	96.3	67.9	99.7	94.8	45.3	99.3	91.1

**Table 3.** Hosmer-Lemeshow's H chi-square test for APACHE III, SAPS II, and MPM II score

Score system	Predicted		Non-survivors		Survivors	
	death rate	N	Observed	Expected	Observed	Expected
APACHE III	0.0 ≤ p < 0.1	28	0	0.01	28	27.99
	0.1 ≤ p < 0.2	42	0	0.02	42	41.98
	0.2 ≤ p < 0.3	36	0	0.04	36	35.96
	0.3 ≤ p < 0.4	36	0	0.08	36	35.92
	0.4 ≤ p < 0.5	34	0	0.15	34	33.85
	0.5 ≤ p < 0.6	36	1	0.33	35	35.67
	0.6 ≤ p < 0.7	33	0	0.73	33	32.27
	0.7 ≤ p < 0.8	34	4	2.73	30	31.27
	0.8 ≤ p < 0.9	34	13	14.82	21	19.18
	0.9 ≤ p < 1.0	35	35	34.12	0	0.88
Chi-square 4.3761 with 8 DF (p=0.8217)						
SAPS II	0.0 ≤ p < 0.1	18	0	0.00	18	18.00
	0.1 ≤ p < 0.2	35	0	0.01	38	37.99
	0.2 ≤ p < 0.3	32	0	0.02	40	39.98
	0.3 ≤ p < 0.4	36	0	0.06	39	38.94
	0.4 ≤ p < 0.5	33	0	0.14	40	39.86
	0.5 ≤ p < 0.6	40	0	0.26	33	32.74
	0.6 ≤ p < 0.7	39	2	0.85	34	35.15
	0.7 ≤ p < 0.8	40	2	2.66	30	29.34
	0.8 ≤ p < 0.9	38	13	13.68	22	21.32
	0.9 ≤ p < 1.0	18	36	35.32	1	1.68
Chi-square 2.5838 with 8 DF (p=0.9577)						
MPM II	0.0 ≤ p < 0.1	40	1	0.06	39	39.94
	0.1 ≤ p < 0.2	138	0	1.28	138	136.72
	0.2 ≤ p < 0.3	84	5	4.46	79	79.54
	0.3 ≤ p < 0.4	43	9	10.84	34	32.16
	0.4 ≤ p < 0.5	43	38	36.36	5	6.64
	0.5 ≤ p < 0.6	0	0	0.00	0	0.00
	0.6 ≤ p < 0.7	0	0	0.00	0	0.00
	0.7 ≤ p < 0.8	0	0	0.00	0	0.00
	0.8 ≤ p < 0.9	0	0	0.00	0	0.00
	0.9 ≤ p < 1.0	0	0	0.00	0	0.00
Chi-square 16.4622 with 8 DF (p=0.0009)						

마나 잘 구별하는가에 대한 평가로 ROC 곡선을 구하였다. 본 연구에서 각 도구의 c-통계량은 APACHE III가 0.981, SAPS II가 0.978, MPM II가 0.941로 세 측정도구 모두 1에 근접한 높은 판별력을 나타내었다 (Figure 1-3).

각각의 모형에서 구한 사망의 판단기준 (decision criteria)은 임계값(cutoff value)을 변화시켰을 때의 민감도, 특이도, 정확도로 제시하였다 (Table 2). 세 도구 모두 정확도가 90% 이상으로 높았다. APACHE III는 임계값 0.6에서 정확도가 96.3%로 가장 높게 나타났으며, 이 때 기대되는 민감도는

79.2%, 특이도는 99.3%이었다. SAPS II는 임계값 0.5에서 정확도가 95.7%로 가장 높았으며, 이 때 기대되는 민감도는 79.2%, 특이도는 99.0%이었다. MPM II는 임계값 0.3, 0.4, 0.5, 0.6에서 모두 같은 정확도 (94.3%), 민감도(71.7%)와 특이도(98.3%)를 보였다.

적합도(calibration) 평가는 모형을 이용하여 예측한 사망률이 실제 사망률에 얼마나 가까운가에 대한 평가이다. 각 모형의 적합도 평가를 위해 Hosmer-Lemeshow의 적합도 검정(goodness-of-fit test, type H chi-square)을 이용하였다. 세 측정체계의 적합

도 판정결과는 APACHE III (chi-square 4.3761 p=0.8217), SAPS II (chi-square 2.5838 p=0.9577), MPM II (chi-square 16.4622 p=0.0009)으로 나타났다 (Table 3).

### 고찰

중환자실에서 수행된 기존의 많은 연구들은 중증도의 양적 도구로서 측정체계가 가치 있다는 것을 보여주고 있다. 특히 APACHE III, SAPS II, MPM II 등을 이용하여 측정된 중증도 점수가 나쁠수록 병원 사망의 위험도가 높은 환자들임을 밝히고 있다 [16,18,35-36].

중증도 측정체계의 사망예측 타당도를 보기 위해 ROC 곡선 아래의 영역인 c-통계량으로 평가한 결과 [37], 본 연구에서 나온 APACHE III (0.981), SAPS II (0.978), MPM II (0.941)의 c-통계량은 이전에 수행된 캐나다, 스페인, 영국, 미국, 이탈리아 포르투갈의 연구결과들보다 높았다 [14,29-32]. Castella [29]의 연구에서는 APACHE III (0.866), SAPS II (0.847), MPM II (0.805)로 모두 0.8 이상으로 나타나 모형의 예측능력이 우수하였다. 그밖에도 APACHE의 경우 Knaus [6]의 연구에서는 0.90, Beck [30]은 0.847로 나타났다. SAPS II의 경우는 Le Gall [22]의 연구에서는 0.86, Moreno [32]의 연구에서는 0.822로 나타났으며, MPM II를 적용한 Lemeshow [23] 연구에서는 0.844였다. 내과계 중환자만을 대상으로 한 Patel과 Grant [27]의 연구에서는 SAPS II (0.6721), APACHE II I(0.7016), MPM II (0.6953)로 나타났다. 국내의 경우 급성심근경색증 환자를 대상으로 APACHE III를 적용한 Park [38]의 연구에서는 0.845이었으며, 동일 질병군을 대상으로 한 Koh [31]의 연구에서는 0.918로 우수한 예측능력을 가진 것으로 나타났다. MPM II를 적용한 Lee [34]의 연구에서는 0.956으로 본 연구보다 높게 나타났다.

이러한 결과들은 기존 주요 중증도 측정체계의 사망예측 타당도가 높음을 보여준다. 주목할 점으로는 대부분의 연구결과에서 특정 질환자를 대상으로 하기보다는 중환자실 대상자 모두를 포함했을 때 더

높은 결과를 보이고 있다. 이는 중증도 측정체계가 개발 당시 일반 중환자실에 입원한 이질적이고 다양한 환자를 대상으로 개발되었기 때문에 모집단의 차이에서 온 것으로 생각된다. 동일 진단의 질환자를 대상으로 한 연구에서는 기존의 사망 예측변수의 영향력이 줄어들고, 다른 변수가 사망에 중요한 영향변수로 작용할 수 있다. 그 결과 이러한 변수들이 처음에 이질적인 대규모 환자들을 대상으로 개발된 모형의 중증도 측정체계에 부적절한 가중치로 작용될 수 있다. 일반적으로 예측사망률에 대한 특정 점수는 전체 중환자실 입실 환자를 전체로 하지만, 개개 환자의 예측사망률에 적용되지 않는 경우도 있다. 즉, 중환자실의 특정 환자들에 대해서는 이러한 측정체계가 맞지 않을 수 있다. 몇몇 연구자들은 사망예측모형의 완성도는 같은 질병으로 규정되거나 같은 이유로 중환자실에 입실한 환자들로 정의된 집단에 따라 개발해야 한다고 주장하기도 한다 [33-39]. 그래서 Groeger [40]는 중환자실 입실 환자들에게 MPM II를 기반으로 입실 환자의 사망예측력을 향상시키도록 ICMM(ICU Cancer Mortality Model)을 개발하기도 하였다. 따라서 특정 성격의 중환자실이나 특정 진단환자만을 대상으로 할 경우 예측사망률을 계산하는데 이용하려면 별도의 도구개발이 필요하다.

본 연구결과에서 c-통계량이 1에 가까운 우수한 예측능력을 보인 것은 대상자에 일반 중환자실의 이질적인 대상자를 모두 포함시켰고, 진료결과가 호전에 따른 퇴원과 병원내 사망환자만을 대상으로 포함시켰기 때문일 가능성이 높다. 즉, 진료결과가 불명확한 전원은 생존자에 포함될 수 있지만 대상자에서 제외시켰으며, 상태가 좋지 않아 사망할 가능성이 높음에도 불구하고 환자 및 보호자가 원하여 사망 없는 퇴원을 하는 경우를 대상에서 제외하였다.

측정체계별 사망의 판단기준에 따라 분류의 정확도를 평가하기 위해 민감도, 특이도와 정확도를 구하여 비교한 결과, APACHE II가 임계값 0.6에서 가장 정확하게 사망여부를 예측할 수 있는 능력은

96.3%이었으며, 가양성률 (false-positive rate)은 임계값 0.7에서 0.3%로 낮았다. 이러한 결과는 매우 우수한 것으로 내과계 환자만을 대상으로 한 Patel과 Grant [27]의 연구결과(임계값 0.4에서 APACHE II는 정확도 67%, MPM II는 65%, 임계값 0.5에서 SAPS II는 68%)와 비교해도 매우 높은 결과이다.

사망의 예측결과와 측정결과 사이의 연관성의 강도를 보기 위해 Lemeshow-Hosmer's goodness-of-fit test(chi-square C)를 이용하여 각 모형의 적합도를 구한 결과, APACHE III(chi-square 4.3761 p=0.8217)와 SAPS II(chi-square 2.5838 p=0.9577)는 사망률 예측 적합도가 우수하였으나 MPM II(chi-square 16.4622 p=0.0009)는 떨어지는 것으로 나타났다(p>0.05이면 모형의 적합도가 적합한 것으로 판단한다). 여기서 p값이 1에 가까울수록 예측 사망과 실제 사망이 서로 잘 일치함을 의미하므로 SAPS II와 APACHE III는 도구의 사망률 예측이 적합한 것으로 나타났으나 반면, MPM II는 p값이 0.0009로 예측 사망률과 실제 사망률에 유의한 차이가 있어 적합도 면에서 떨어지는 것으로 나타났다. 이는 MPM II는 다른 도구에 비해 변수의 숫자가 적고, 가중치가 없어 결과적으로 점수 폭(1~12점)이 좁아서 적용이 간단하다는 장점이 있으나 한 두 개의 변수 값도 사망예측에 큰 영향을 주었을 가능성이 높다. 실제 대상자수가 적고, 점수의 폭도 0점에서 9점의 범위에 있었던 점을 감안하면 그 영향은 더욱 컸을 것으로 생각된다.

모형의 판별력과 적합도 값이 우수하지 않은 경우는 다음의 두 가지 가능성이 있다. 첫째는 의료의 질(quality of care)이 기대한 것보다 더 낮거나 좋은 경우이다. 이 경우는 높은 질의 진료를 제공함으로써 측정체계가 생존을 예측한 환자보다 많은 수의 환자가 생존하여 측정체계의 예측 정확도를 떨어뜨렸거나, 낮은 질의 의료 서비스를 제공함으로써 생존가능성이 있는 환자들을 죽게 한 경우이다. 둘째는 적용 환자의 예외성 때문에 모형의 예측 정확도에 제한을 주는 문제이다. 이 도구들이 판별력이 좋음에도 불구하고, 의료의

질이 향상됨으로써 판별력과 적합도가 떨어질 수 있으며, 상황에 따라 중증도 점수 체계의 적용 가능성도 줄어들 수 있다. 이러한 제한들은 의료의 질의 변화와 생존율 향상을 자주 점검하여 재측정함으로써 극복할 수 있다 [27].

본 연구결과를 통해 세 측정체계가 중환자실 입실환자의 중증도에 따른 사망예측 도구로서 받아들일만한 가치가 있으며, 중환자실에 입실한 환자의 결과를 평가할 만하다는 것을 증명하고 있다. 특히 APACHE III와 SAPS II는 판별력, 적합도, 정확도 평가의 모든 면을 고려할 때, MPM II보다 더 정확한 예측력이 있는 것으로 나타났다. 이러한 결과는 변수가 많아 다양한 측면의 건강상태를 반영하고, 변수의 수에 따른 점수 폭이 커지면서 복잡한 측정체계가 예측력이 높음을 보여주는 것으로 나타났다.

세 측정체계 중에서 APACHE III는 변수의 숫자가 가장 많으며, SAPS II는 APACHE III를 기반으로 변수의 수를 줄인 도구이므로 단순성 면에서 본다면 SAPS II의 유용성이 더 높다고 할 수 있다. 그러나 현재 많은 병원들에서 임상정보시스템을 잘 갖추고 때문에 APACHE III에 포함된 검사 결과치들이 자동적으로 생성되도록 한다면 APACHE III를 사용하는 데도 큰 문제가 되지 않을 것이다. 또한 임상정보시스템에 이러한 측정체계를 일상적으로 적용하여 많은 환자를 대상으로 지속적인 연구가 이루어진다면 중환자 진료의 질 평가와 향상에도 도움이 될 것이다.

현대 사회는 질(quality)을 강조하고 있고, 의료분야도 예외는 아니다. 이러한 현실에서 중증도 측정도구를 활용한 중환자의 진료 결과 평가는 진료의 수행결과(outcome of care) 차이를 반영하는 지표로서도 중요한 의미를 가진다고 생각된다.

## 요약 및 결론

본 연구는 중환자의 사망률의 예측에 가장 많이 사용되고 있는 측정체계 중 APACHE III, SAPS II, MPM II를 동일 중환자에게 적용하여 중증도를 가장 잘 반영할 수

있는 도구의 타당도를 평가하고자 시도되었다.

측정체계가 실제 생존자와 사망자를 얼마나 잘 구별하는가를 나타내는 판별력 평가에서 각 도구의 c-통계량은 APACHE III(0.981), SAPS II(0.978), MPM I I(0.941)로 모두 1에 근접한 높은 판별력을 나타내었다. 측정체계별 사망예측률을 판단기준에 따라 민감도, 특이도와 정확도를 비교한 결과, 세 도구 모두 정확도가 90% 이상으로 높게 나타났다. 예측 사망률과 실제 사망률에 얼마나 가까운가에 대한 분류체계의 적합도 평가에서도 APACHE III, SAPS II는 사망률 예측에 적합한 도구로 나타났다. 결과적으로 APACHE III는 중환자의 사망률 예측에 가장 우수한 도구로 평가되었고, APACHE 항목을 축약한 SAPS II도 모든 면에서 우수한 도구임이 증명되었다. 이상의 결과에서 APACHE III와 SAPS II는 중환자의 사망 예측과 중환자 진료의 질 평가에 매우 유용하리라고 사료된다.

본 연구는 일개 종합전문요양기관 중환자실의 특정기간 동안의 자료를 이용하였고, 대상자의 수가 적어서 일반화하는데 제한점이 있다. 앞으로 대규모의 다기관 공동연구를 통한 검증이 뒤따라야 하겠다.

**참고문헌**

1. Shim A. Clinical survey of patients of the Intensive care unit in Wonju Christian Hospital. *Korean J Anesthesiol* 1994; 1(27): 84-89 (Korean)
2. Nunn JF, Milledge JS, Singaraya J. Survival of patients ventilated in an intensive therapy unit. *BMJ* 1979; 285: 1525
3. Thibault GE, Mulley AG, Barnett GO. Medical intensive care: Indications, interventions and outcome. *N Engl J Med* 1980; 302: 938
4. Cho KJ. The general analysis of the patients in intensive care unit by severity systems of an university hospital in Seoul [dissertation]. Koea: Seoul National University; 1991 (Korean)
5. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: A severity of disease classification system. *Crit Care Med* 1985; 13(10): 818-829
6. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG. The APACHE III prognostic system: Risk prediction of hospital mortality for critically ill hospitalized adult.

- Chest* 1991; 100(6): 1619-1636
7. Rafkin HS, Hoyt JW. Objective data and quality assurance programs: current and future trends. *Crit Care Clinics* 1994; 10(1): 157
8. Knaus WA, Zimmerman JE, Wagner DP. APACHE-acute physiology and chronic health evaluation: A physiologically based classification system. *Crit Care Med* 1981; 9: 591-597
9. Shortell SM, Zimmerman JE, Rousseau DM. The performance of intensive care units: does good management make a difference? *Med Care* 1994; 32: 508-525
10. Hall JB, Schmid GA, Wood LD. Principle of critical care. New York: McGraw-Hill Inc. ; 1992. p. 551-557
11. Cullen DJ, Chernow B. Predicting outcome in critically ill patients. *Crit Care Med* 1994; 22: 1345-1348
12. Lemeshow S, Teres D, Avrunin JS, Gage RW. Refining intensive care outcome prediction by using changing probabilities of mortality. *Crit Care Med* 1988; 16: 470-477
13. Le Gall JR, Loira P, Alperovitch A, Glaser P, Granthil C, Mathieu D, Mercier P, Thomas R, Viller D. A simplified acute physiology score for ICU patients. *Crit Care Med* 1984 ; 12 : 975-977
14. Wong DT, Crofts SL, Gomez M. Evaluation of predictive ability of APACHE II system and hospital outcome in Canadian intensive care unit patients. *Crit Care Med* 1995; 23: 1177-1183
15. Zimmerman JE. The APACHE III study design: Analytic plan for evaluation of severity and outcome. *Crit Care Med* 1989; 17: S169-221
16. Sculier JP, Paesmans M, Markiewicz E, Berghmans T. Scoring systems in cancer patients admitted for an acute complication in a medical intensive care unit. *Crit Care Med* 2000; 28: 2786-2792
17. Lee DH, No MY, Kim BS. Interhospital comparison of outcome from intensive care unit with APACHE III scoring system. *Korean J Prev Med* 1994; 27(3): 437-445 (Korean)
18. Lee KO, Shin HI, Park HA, Jung HM, Lee MH, Choi EH, Lee JM, Kim UI, Sim YK, Park GJ. Patient severity classification in a medical ICU using APACHE III and patient severity classification tool. *J Korean Acad Nurs* 2000; 30(5): 1243-1253 (Korean)
19. Seneff M, Knaus WA. Predicting patient outcome from intensive care: A guide to APACHE, MPM, SAPS, PRISM and other prognostic scoring system. *J Intensive Care Med* 1990; 5(1): 33-52
20. Laurent GG, Turner MO, Andrew D. Rating the quality of intensive care units: Is it a

- function of the intensive care unit scoring system? *Crit Care Med* 2002; 30(9): 1976-1982
21. Teres D, Lemeshow S, Avrunin JS. Validation of the mortality prediction model for ICU patients. *Crit Care Med* 1987; 15: 208-212
22. Le Gall JR, Lemeshow S, Saulnier F. A new simplified acute physiology score(SAPS II) based on a European/North American multicenter study. *JAMA* 1993; 270: 2957-2963
23. Lemeshow S, Teres D, Klar J. Mortality probability models(MPM II) based on an international cohort of intensive care unit patients. *JAMA* 1993; 270: 2957-2963
24. Sicignano A, Giudici D. Customization of SAPS II for the assessment of severity in Italian ICU patients. *Minerva Anesthesiol* 2000; 66: 139-145
25. Rué M, Artigas A, Álvarez M. Performance of the mortality probability models(MPM II) in assessing severity of illness during the first week in the intensive care unit. *Crit Care Med* 2000; 28: 2819-2824
27. Beck DH, Taylor BL, Millar B. Prediction of outcome from intensive care: A prospective cohort study comparing APACHE II and III prognostic systems in a United Kingdom intensive care unit. *Crit Care Med* 1997; 25: 9-15
28. Castella X, Artiga A, Bion J, Kari A. A comparison of severity of illness scoring systems for intensive care unit patients: results of multicenter, multinational study. *Crit Care Med* 1995; 23(8): 1327-1335
29. Castella K, Artigas A, Bion J, Kari A & The European/North American Severity Study Group. A comparison of severity of illness scoring systems for intensive care unit patients of multicenter, multinational study. *Crit Care Med* 1995; 23(8): 1327-1335
30. Patel PA, Grant BJB. Application of mortality prediction systems to individual intensive care units. *Intensive Care Med* 1999; 25: 977-982
31. Koh BY. Prediction model for in-hospital mortality of patients visited emergency room with acute myocardial infarction [dissertation]. Korea: Chung-ang University; 2002 (Korean)
32. Moreno R, Morians P. Outcome prediction in intensive care: Its of a prospective, multicenter, Portuguese study. *Intensive Care Med* 1997; 23: 177-186
33. Reina A, Vazquez G, Aguayo E, Bravo I, Colmenero M, Bravo M. Mortality discrimination in acute myocardial infarction: comparison between APACHE III & SAPS II prognosis system. *Intensive Care Med* 1997; 23: 326-330
34. Lee JH, Back KJ, Han SB, Ahn ST, Shin DW,

- Kim AJ, Kim JS. Mortality analysis of intensive care units patients using mortality probability models(MPM II). *J Traumatol* 2001; 14(2): 101-107 (Korean)
35. Guiguet M, Blot F, escudier B, Antoun S, Leclercq B, Nitenber G. Severity illness scores for neutropenic patients in an intensive care unit: which is the best predictor? Do multiple assessment times improve the predictive value? *Crit Care Med* 1998; 26: 488-493
36. Schellongowski P, Benesch M, Lang T, Traummuller F, Zauner C, Laczika K, Locker GJ, Frass M, Staudinge T. Comparison of three severity score for critically ill cancer patients. *Intensive Care Med* 2004; 30(3): 430-436
37. Iezzoni LI. The risk adjustment. *JAMA* 1997; 278(19): 1600-1607
38. Park YS. A study on correlation of clinical process indicator and mortality after severity adjustment [dissertation]. Korea: Korea University, 1998 (Korean)
39. Carson SS, Bach PB. Predicting mortality in patients suffering from prolonged critical illness: an assessment of four severity-of-illness measures. *Chest* 2001; 120: 928-933
40. Groeger JS, Lemeshow P, Nierman DM, White P Jr, Klar J, Granovsky S, Horak D, Kish SK. Multicenter outcome study of cancer patients admitted to the intensive care unit: a probability mortality model. *J Clin Oncol* 1998; 16: 761-770