화상환자에서 사망예측모델의 성능 평가에 관한 연구

우재연 · 김도현¹

한림대학교 강남성심병원 외과학교실, ¹한림대학교 한강성심병원 외과학교실

The Accuracy of Prediction Models in Burn Patients

Jaeyeon Woo, M.D. and Dohern Kym, M.D.¹

Department of Surgery, Hallym University Kangnam Sacred Heart Hospital, ¹Department of Surgery, Hallym University Hangang Sacred Heart Hospital, Seoul, Korea

Purpose: The purpose of this study was to evaluate the accuracy of four prediction models in adult burn patients.

Methods: This retrospective study was conducted on 696 adult burn patients who were treated at burn intensive care unit (BICU) of Hallym University Hangang Sacred Heart Hospital from January 2017 to December 2019. The models are ABSI, APACHE IV, rBaux and Hangang score.

Results: The discrimination of each prediction model was analyzed as AUC of ROC curve. AUC value was the highest with Hangang score of 0.931 (0.908 ~ 0.954), followed by rBaux 0.896 (0.867 ~ 0.924), ABSI 0.883 (0.853 ~ 0.913) and APACHE IV 0.851 (0.818 ~ 0.884).

Conclusion: The results of evaluating the accuracy of the four models, Hangang score showed the highest prediction. But it is necessary to apply the appropriate prediction model according to characteristics of the burn center. (*J Korean Burn Soc* 2021;24:1-6)

Key Words: Burn, Prediction model, Hangang score

서 론

중증 화상은 광범위한 피부의 손상으로 인한 전신성 염증 반응을 일으키고 감염에 취약하여 사망률과 합병증 빈도가 높은 중증 질환이다[1]. 화상의 치료법이 발전함에 따라 전 세계적으로 화상의 사망률은 감소하고 있다. 하지만 선진국 및 화상전문센터를 가진 나라를 제외하고는 여전히 중증화상환자의 사망률은 높다[2,3].

화상의 종류나 환자 요인 등 다양한 요소가 환자의 예후에 영향을 미치며, 이를 반영하여 고안된 사망예측모델은 중증 화상 환자의 치료방향을 선택하는데 중요한 역할을 한다. 현재까지 여러가지 사망예측모델들이 발표되고 개정

되어왔다[4,5].

The abbreviated burn severity index (ABSI)는 1982년 처음 발표되었으며 화상환자를 대상으로 한다. 성별, 연령, 화상과 관련된 지표(체표면적, 흡인 및 전층 화상 유무)로 점수를 산출하며 점수가 낮을수록 생존률(%)은 증가한다[6].

Acute physiology and chronic health evaluation (APACHE IV)은 화상 유무와는 상관없이 중환자실에 입원한 환자를 대상으로 24시간 이내에 측정된 변수들로 사망률을 예측한다[7]. APACHE는 1981년 처음 고안되었으며 2006년 APACHE IV로 개정되었다[8].

Revised Baux score (rBaux)는 기존의 Baux score에 흡인 화상 항목을 추가하여 2010년에 개정되었다. 연령+화상을 입은 체표면적(%TSBA)+17×(흡인 화상, 1=유, 0=무)로 계산되며 점수가 높을수록 사망률도 증가하는 지표이다[9].

Hangang score는 2019년 처음 고안되었으며 중증 화상을 입은 성인 환자를 대상으로 한다. 다양한 지표를 통해 점수를 산출하며 점수가 높을수록 사망률도 증가한다[10].

지금까지 많이 사용되었던 ABSI, APACHE IV, rBaux의

접수일: 2021. 3. 21, 수정일: 2021. 4. 10, 승인일: 2021. 4. 13

책임저자 : 김도현, 서울시 영등포구 버드나루로7길

☎ 07247, 한림대학교 한강성심병원 화상센터 외과학교실

Tel: 02-2639-5446, Fax: 02-2678-4386

E-mail: dohern@hallym.or.kr

정확도는 앞선 여러 연구에서 확인되었고, 최근에 발표된 Hangang score는 환자 요인 뿐 아니라 7가지 검사항목을 추가하여 뛰어난 정확도를 보여주었다[10]. 따라서 본 연구의 목적은 실제 환자의 사망률을 통해 각 모델의 정확도를 평가 및 비교해보는데 있다.

대상 및 방법

1. 환자 선택

2017년 1월부터 2019년 12월까지 한림대학교 한강성심 병원에 내원하여 Burn Intensive Care Unit (BICU)에서 치료받은 18세 이상의 성인 환자 696명을 대상으로 하였으며 의무기록(EMR)을 이용한 후향적 조사 연구로 진행되었다.

BICU 입원 치료의 기준은 1) 체표면적 20% 이상의 화상 또는 10% 이상의 화상을 입은 65세 이상의 환자 2) 흡인 화상 3) 고압의 전기 화상 4) 주요 외상 또는 기저질환이 동반되어 의학적으로 중환자실 입원이 필요하다고 판단된 경우이다. 입원 후 초기 수액 치료는 modified Parkland formula ($4 \text{ mL} \times \text{kg} \times \% \text{ TBSA burned}$)에 따라 시행하였고 시간당 소변량이 최소 0.5 mL/kg/hour 가 되도록 조절하였다.

2. 사망예측모델

ABSI score는 5가지 항목으로 계산된다. 연령(1~5점), 화상을 입은 체표면적(%TBSA) (1~10점), 여성(1점), 흡인 화상의 유무(1점), 전층 화상의 유무(1점). APACHE IV는 ICU 입원 후 24시간 이내에 측정된 변수로 계산된다. 17가지의 생리학적 지표와 더불어 기계 환기, 혈전 용해, Glasgow Coma Scale, $\text{PaO}_2/\text{FiO}_2$ ratio 그리고 질병의 종류를 세분화하여 반영하였다. rBaux는 연령+화상을 입은 체표면적(%TBSA)+17x (흡인 화상, 1=유, 0=무)로 계산하였다. Hangang score는 환자의 연령, 화상을 입은 체표면적(%TBSA), 흡인 화상 유무, lactate (mmol/L), PH, prothrombin time (sec), bilirubin (mg/dL), myoglobin (ng/mL), creatinine (mg/dL), lactate dehydrogenase (IU/L)로 계산된다.

흡인 화상은 밀집되고 폐쇄된 공간에서 화상을 입은 경우, 불에 그슬리거나 탄 얼굴의 털, 얼굴과 코 또는 입 주변의 화상, 탄소 가루가 섞인 가래 등의 진찰 소견이나 기관지 내시경, 일산화탄소 수치, 흉부 x선 촬영 등의 검사를 통해 진단하였다.

3. 통계분석방법

정규분포와 비정규분포의 연속변수들은 각각 평균±표준편차(SD)와 중앙값(사분위 범위[IQR])으로 표시하였다. 두

그룹을 비교하기 위해 정규분포를 따르는 데이터의 경우 등분산 검정을 통하여 이를 만족하는 경우 student's t-test, 만족하지 않는 경우에는 Welch's t-test로 분석하였다. 정규분포를 따르지 않는 데이터의 경우에는 Wilcoxon signed rank test를 사용하였다. 범주형 변수의 경우에는 백분율로 표시하였으며 카이-제곱 검정을 사용하여 차이를 검정하였다. Lethal area fifty percent (LA50)은 화상환자에서 50%의 사망률을 보이는 TBSA (%)로 logistic regression으로 계산하였다. Receiver Operating Characteristic curve (ROC)의 Area Under the Curve (AUC)를 이용하여 각 예측 모델의 사망에 대한 변별력을 확인하였으며, 이를 위해 DeLong's method를 사용하였다. AUC는 1에 가까워질수록 변별력이 증가했다[11]. 각 예측 모델의 전체적인 정확도 비교를 위해 Brier score를 산출하였으며 이는 0과 1사이의 값을 가지고 0에 가까울수록 정확한 예측도를 보인다[12]. Scaled Brier score는 즉각적인 비교를 위해서 Brier score의 최대값에 의해 조정되어 0에서 100% 사이의 값을 가지며, 수치가 커질수록 정확도가 더 좋다[13]. P-value가 0.05 이하인 경우 통계적으로 유의한 것으로 판단하였으며 통계 분석은 statistical computing R-project version 3.6.1을 사용했다.

결 과

1. 환자기본정보

전체 696명 환자 중에서 입원기간동안 생존한 환자는 522명(75%)이며 사망한 환자는 174명(25%)이었다. 생존한 환자의 나이의 중앙값은 52.0세이고 사망한 환자의 나이의 중앙값은 60.0세로 더 높았다. 전체 환자의 TSBA의 중앙값은 25.0이며 생존한 환자의 TSBA 중앙값은 20.0, 사망한 환자의 TSBA 중앙값은 57.5로 월등히 높았다. 전체 환자, 생존 환자, 사망 환자에서 APACHE IV, ABSI, rBaux, Hangang score를 계산한 결과 APACHE IV에서는 41.0 (29.0~56.0), 36.0 (26.0~48.0), 64.0 (50.0~79.0), ABSI에서는 7.0 (6.0~9.0), 7.0 (6.0~8.0), 11.0 (9.0~13.0)이며 rBaux의 경우 89.8 ± 29.9 , 78.8 ± 21.0 , 122.8 ± 28.4 로 나타났다. Hangang score에서는 130.5 (118.0~146.5), 124.0 (115.0~135.0), 159.0 (148.0~179.0)이며 4가지 모델 모두 사망한 환자 군에서 더 높은 점수를 보였다(Table 1).

2. 사망률과 LA50

전체 환자 중 사망한 환자는 174명(25%)이고 사망한 환자의 LA50은 53.4 (46.9~60.7)로 나타났다. 전체 환자를 나이에 따라 18~50세, 51~65세, 65세 이상의 세가지 군으로 분류하여 LA50과 사망률을 비교하였을 때 65세 이상 군에

Table 1. Baseline Characteristics

Variables	Total (N=696)	Survivors (N=522)	Non-survivors (N=174)	P-value
Age	54.0 [44.0~64.0]	52.0 [43.0~62.0]	60.0 [49.0~73.0]	<0.001
Sex				0.267
- Male	543 (78.0%)	413 (79.1%)	130 (74.7%)	
- Female	153 (22.0%)	109 (20.9%)	44 (25.3%)	
Type				<0.001
- FB	463 (66.5%)	313 (60.0%)	150 (86.2%)	
- SB	60 (8.6%)	51 (9.8%)	9 (5.2%)	
- EB	120 (17.2%)	117 (22.4%)	3 (1.7%)	
- CoB	39 (5.6%)	29 (5.6%)	10 (5.7%)	
- ChB	14 (2.0%)	12 (2.3%)	2 (1.1%)	
TBSA	25.0 [12.0~40.0]	20.0 [10.0~30.0]	57.5 [34.0~80.0]	<0.001
Inhalation				<0.001
- Yes	211 (30.3%)	129 (24.7%)	82 (47.1%)	
- No	485 (69.7%)	393 (75.3%)	92 (52.9%)	
LOS	43.0 [20.0~66.0]	53.0 [35.0~78.0]	10.0 [2.0~23.0]	<0.001
LOICU	9.0 [3.0~23.0]	7.5 [3.0~24.0]	10.0 [2.0~23.0]	0.285
APACHE_IV	41.0 [29.0~56.0]	36.0 [26.0~48.0]	64.0 [50.0~79.0]	<0.001
ABSI	7.0 [6.0~9.0]	7.0 [6.0~8.0]	11.0 [9.0~13.0]	<0.001
rBaux	89.8±29.9	78.8±21.0	122.8±28.4	<0.001
Hangang	130.5 [118.0~146.5]	124.0 [115.0~135.0]	159.0 [148.0~179.0]	<0.001

FB = flame burn, SB = scald burn, EB = electrical burn, CoB = contact burn, ChB = chemical burn, TBSA = total body surface area, LOS = length of stay, LOICU = length of intensive care unit.

Table 2. LA50 and Mortality

Age	LA50 (95% CI)	Mortality (%)
18~50 years	59.1 (52.0~67.2)	19.9
51~65 years	53.5 (45.2~63.4)	20.5
>65 years	29.5 (21.2~41.1)	42.4
Total	53.4 (46.9~60.7)	25.0

서 LA50이 가장 낮은 반면에 사망률은 가장 높게 나타났다 (Table 2).

또한 각 모델에 대한 LA50을 비교해보면, lethal Hangang score는 150.0 (147.7~152.4), lethal ABSI는 10.0 (9.6~10.4), lethal APACHE IV는 62.9 (59.2~67.0), lethal rBaux는 112.4 (108.6~116.2)로 나타났다.

3. AUC와 분포도

4가지 모델의 ROC curve는 Fig. 1에 나타나 있으며 AUC 값은 Hangang score가 0.931 (0.908~0.954)로 가장 높았으며 rBaux 0.896 (0.867~0.924), ABSI 0.883 (0.853~0.913) 그리고 APACHE IV 0.851 (0.818~0.884) 순으로 나타났다. 각 모델의 AUC 값을 서로 비교하였을 때 Hangang score와 ABSI, APACHE IV의 결과에서만 통계적으로 유의한 차이를 보였다($P<0.001$) (Table 3). 각 모델

Fig. 1. ROC curve for models.

에 따른 실제 환자들의 점수 분포도에서도 4가지 모델 모두 생존자와 사망자의 점수 분포에 차이를 보였다(Fig. 2).

4. 모델 성능

4가지 모델에 대하여 Brier score, Scaled Brier score, AUC값을 산출하였다. Brier score에서는 Hangang score가 0.080으로 가장 0에 가까웠으며 rBaux는 0.100, ABSI는 0.104, APACHE IV는 0.123의 값을 보였으며 0.25를 넘는 값을 가지는 모델은 없었다. Scaled Brier score와 AUC에서 Hangang

Table 3. Comparisons for AUC of Models

Prediction models (AUC)	Hangang (0.931)	rBaux (0.896)	ABSI (0.883)	APACHE IV (0.851)
Hangang (0.931)	-	0.004	<0.001	<0.001
rBaux (0.896)		-	0.133	0.025
ABSI (0.883)			-	0.141
APACHE IV (0.851)				-

Fig. 2. Distribution of models.**Table 4.** Model Performance of Models

	APACHE IV model	ABSI model	rBaux model	Hangang model
Brier score	0.123	0.104	0.100	0.080
Scaled Brier score (95% CI)	0.344 (0.264~0.429)	0.445 (0.371~0.534)	0.467 (0.381~0.547)	0.573 (0.49~0.653)
AUC (95% CI)	0.851 (0.813~0.885)	0.883 (0.85~0.912)	0.896 (0.863~0.922)	0.931 (0.905~0.952)

score가 각각 0.573 (0.49~0.653), 0.931 (0.905~0.952)로 가장 높은 값을 보였다. rBaux는 0.467 (0.381~0.547), 0.896 (0.863~0.922), ABSI에서는 0.445 (0.371~0.534), 0.883 (0.85~0.912), APACHE IV에서는 0.344 (0.264~0.429) 0.851 (0.813~0.885)로 나타났다(Table 4).

고찰

화상환자를 대상으로 한 사망 예측 모델은 앞서 소개한 바와 같이 여러가지가 있다. 어떤 모집단을 대상으로 하는 지에 따라 각 사망예측모델의 정확도에는 차이가 있을 수

있지만 유효성은 여러 연구를 통해 증명되었다. 하지만 모든 집단에 적용 가능한 이상적인 모델은 아직까지 없다. 대부분의 모델은 선진국의 모집단에 의해 검증되었고 Hangang score 또한 한림대학교 한강성심병원에 내원하여 Burn Intensive Care Unit (BICU)에 입원한 환자들을 대상으로 개발 및 검증되었다[10].

이 논문에서는 화상환자를 대상으로 한 3가지 모델과 일반 중환자를 대상으로 한 한가지 모델을 평가하였다. 화상환자를 대상으로 한 3가지 예측모델(ABSI, rBaux, Hangang score)의 공통점은 환자의 연령, TSBA (%), 흡인 화상이 중요한 요소로 작용한다는 것이다. 본 연구에서는 Hangang score가 가장 좋은 예측도를 보여주었는데 이는 정확성과 객관성을 보장하기 위해 일곱 가지 변수(lactate, pH, PT, bilirubin, myoglobin, creatinine, LD)를 추가하여 반영한 것이 기여했을 것이라 생각한다. 중증 화상에 의한 대사성 산증은 lactate, PT, bilirubin의 증가와 PH의 감소를 야기한다[10]. 또한 lactate, creatinine, LD는 중증 화상으로 인한 급성신부전을 확인하는데 유용하다[14]. Serum myoglobin과 LD는 중증 화상에 따른 횡문근융해증을 고려한 지표로서 다른 예측 모델에서 변수로 사용되지 않았으나 이전의 연구에서 myoglobin과 LD가 화상환자에서 사망률과 관련이 있다고 보고된 바 있다[15-17].

한편 APACHE IV가 가장 낮은 예측도를 보여주었다. 이는 다른 모델과 달리 일반 중환자를 대상으로 하여 앞서 언급한 TSBA (%), 흡인 화상과 같이 주요한 화상 관련 요소를 포함하고 있지 않기 때문이라고 여겨진다. 하지만 본 연구와 달리 APACHE score와 화상 환자의 사망률을 비교한 연구에서 APACHE II, III 모두 높은 예측도를 보여주기도 하였다[18,19]. 최근 연구에서 AHACHE II는 화상과 관련한 요소를 포함하지 않음에도 불구하고 다른 모델들과 비교하였을 때 좋은 예측도를 보여주었고[20], APACHE score에 화상과 관련된 요소를 추가하면 더 높은 예측도를 보여줄 수 있다[21]. 실제로 FLAMES score는 APACHE II를 기반으로 하여 화상 관련 위험인자 항목을 추가하여 계산되는 모델로 2008년에 발표되었으며 화상환자에서 APACHE II보다 뛰어난 예측도를 보여주었다[22].

본 연구에서는 각 모델 별로 예측도에 차이가 있지만 화상 치료의 전문성 및 객관성과 센터에서 확인 할 수 있는 검사의 종류에 따라 예측모델을 선택하여 평가 하는 것이 중요하다.

본 연구의 제한점으로는 첫째, 후향적 연구로 진행되었기 때문에 예측 모델의 항목 이외의 변수의 영향을 배제할 수 없다는 점이다. 둘째, 한강성심병원 화상중환자실(BICU)에 입원한 환자를 대상으로 한 단일 기관 연구이기 때문에

다른 집단에 적용시켰을 때 결과가 달라질 수 있다. 셋째, 성인을 대상으로 한 연구이기 때문에 소아에게 적용시키는데 한계가 있다. 각 예측 모델의 정확한 비교를 위해서는 더 큰 모집단에서 전향적인 다기관연구가 필요하다.

결 론

중증화상환자에서 사망률을 예측하는 것은 중요하다. 본 연구에서 4가지 사망예측모델을 비교한 결과 Hangang score의 예측도가 가장 정확하였으나 화상센터의 특징에 따라 적절한 모델을 선택하여 적용하는 것이 필요하다.

REFERENCES

1. Cho YS, Yim H, Yang HT, Hur J, Chun W, Kim JH, et al. Use of parenteral colistin for the treatment of multiresistant Gram-negative organisms in major burn patients in South Korea. *Infection*. 2012;40:27-33.
2. Smolle C, Cambiaso-Daniel J, Forbes AA, Wurzer P, Hundeshagen G, Branski LK, et al. Recent trends in burn epidemiology worldwide: A systematic review. *Burns* 2017; 43:249-57.
3. Seo DK, Kym D, Yim H, Yang HT, Cho YS, Kim JH, et al. Epidemiological trends and risk factors in major burns patients in South Korea: a 10-year experience. *Burns* 2015; 41:181-7.
4. Strassle PD, Williams FN, Napravnik S, van Duin D, Weber DJ, Charles A, et al. Improved Survival of Patients With Extensive Burns: Trends in Patient Characteristics and Mortality Among Burn Patients in a Tertiary Care Burn Facility, 2004-2013. *J Burn Care Res* 2017;38:187-93.
5. Salehi SH, As'adi K, Abbaszadeh-Kasbi A, Isfeedvajani MS, Khodaei N. Comparison of six outcome prediction models in an adult burn population in a developing country. *Ann Burns Fire Disasters* 2017;30:13-7.
6. Tobiasen J, Hiebert JM, Edlich RF. The abbreviated burn severity index. *Ann Emerg Med* 1982;11:260-2.
7. Korkmaz Toker M, Gülleroğlu A, Karabay AG, Biçer İ G, Demiraran Y. SAPS III or APACHE IV: Which score to choose for acute trauma patients in intensive care unit? *Ulus Travma Acil Cerrahi Derg* 2019;25:247-52.
8. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006;34:1297-310.
9. Osler T, Glance LG, Hosmer DW. Simplified estimates of the probability of death after burn injuries: extending and updating the baux score. *J Trauma* 2010;68:690-7.
10. Kim Y, Kym D, Hur J, Jeon J, Yoon J, Yim H, et al. Development of a risk prediction model (Hangang) and

- comparison with clinical severity scores in burn patients. *PLoS One* 2019;14:e0211075.
11. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.
 12. Brier GW. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 1950;78:1-3.
 13. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology (Cambridge, Mass)* 2010;21:128-38.
 14. Kym D, Cho YS, Yoon J, Yim H, Yang HT. Evaluation of diagnostic biomarkers for acute kidney injury in major burn patients. *Ann Surg Treat Res* 2015;88:281-8.
 15. Walsh MB, Miller SL, Kagen LJ. Myoglobinemia in severely burned patients: correlations with severity and survival. *J Trauma* 1982;22:6-10.
 16. GÜldoğan CE, Kendirci M, Gündoğdu E, Yastı A. Analysis of factors associated with mortality in major burn patients. *Turk J Surg* 2019;35:155-64.
 17. Liu ZJ, Zhang Y, Zhang XB, Yang X. Observation and identification of lactate dehydrogenase anomaly in a postburn patient. *Postgrad Med J* 2004;80:481-3.
 18. Wong TH, Tan BH, Ling ML, Song C. Multi-resistant *Acinetobacter baumannii* on a burns unit—clinical risk factors and prognosis. *Burns* 2002;28:349-57.
 19. Tanaka Y, Shimizu M, Hirabayashi H. Acute physiology, age, and chronic health evaluation (APACHE) III score is an alternative efficient predictor of mortality in burn patients. *Burns* 2007;33:316-20.
 20. Halgas B, Bay C, Foster K. A comparison of injury scoring systems in predicting burn mortality. *Ann Burns Fire Disasters* 2018;31:89-93.
 21. Moore EC, Pilcher DV, Bailey MJ, Cleland H, McNamee J. A simple tool for mortality prediction in burns patients: APACHE III score and FTSA. *Burns* 2010;36:1086-91.
 22. Gomez M, Wong DT, Stewart TE, Redelmeier DA, Fish JS. The FLAMES score accurately predicts mortality risk in burn patients. *J Trauma* 2008;65:636-45.