

· 临床研究 ·

急性生理和慢性健康状况 IV 评分模型在老年重症医学科中的验证

董家辉¹, 莫泽珣¹, 孙照琨¹, 孙 杰¹, 曾 安², 郭振辉^{1*}

(¹ 广州军区广州总医院老年重症医学科, 广州 510010; ² 广东工业大学计算机学院, 广州 510000)

【摘要】 目的 研究急性生理和慢性健康状况 (APACHE) IV 评分模型的精确度以及是否适用于老年重症患者。**方法** 前瞻性观察 2011 年 7 月至 2015 年 7 月广州军区广州总医院老年重症医学科收治的 421 例重症老年患者, 计算入住 ICU 24 h 内的 APACHE IV 评分, 通过受试者工作曲线下面积 (ROC AUC) 和 Hosmer-Lemeshow (HL) 检验分别检验模型的辨别力和校准, 标准化死亡比值 (SMR) 检验模型预测群体死亡率的准确性, 并计算模型的 Brier 评分 (BS)。**结果** 421 例老年重症患者, 住院死亡率为 33.3% (140/421)。APACHE IV 模型辨别力为良好到优秀 (AUC = 0.82, 95% CI: 0.78 ~ 0.86), 模型的 HL 检验显示校准不适 ($\chi^2 = 20.49, P = 0.009$)。BS 显示模型良好的精确度 (BS = 0.18), 但 APACHE IV 模型可能低估老年重症患者住院死亡率 (SMR = 1.50, 95% CI: 1.30 ~ 1.70)。**结论** 在老年重症患者中, APACHE IV 模型显示优秀的辨别力, 但模型校准不良同时低估住院死亡率。老年专科患者使用危险预测模型前应验证模型并使用定制化模型拟合预测群体。

【关键词】 急性生理与慢性健康状况评分; 老年; 住院病死率; 预后评估; 验证

【中图分类号】 R541.4

【文献标志码】 A

【DOI】 10.11915/j.issn.1671-5403.2017.06.096

Validation of the acute physiology and chronic health evaluation IV model for the prediction of hospital mortality in a geriatric intensive care unit

DONG Jia-Hui¹, MO Ze-Xun¹, SUN Zhao-Kun¹, SUN Jie¹, ZENG An², GUO Zhen-Hui^{1*}

(¹ Geriatric Intensive Care Unit, Guangzhou General Hospital of Guangzhou Military Command, Guangzhou 510010, China; ² College of Computing, Guangdong University of Technology, Guangzhou 510000, China)

【Abstract】 Objective To validate the accuracy and feasibility of acute physiology and chronic health evaluation (APACHE) IV model in critically ill elderly patients. **Methods** A total of 421 elderly critically ill patients admitted in our hospital from July 2011 to July 2015 were prospectively recruited in this study. Their APACHE IV scores were calculated within 24 h in ICU. Discrimination and calibration were respectively assessed by using area under the receiver operating characteristic curve (ROC AUC) and Hosmer-Lemeshow (HL) test. Standardized mortality ratio (SMR) was reported to quantify the mortality variation between observation and prediction and the brier score (BS) to represent the overall performance. **Results** The mortality rate of the 421 subjects was 33.3% (140/421). The discrimination of the APACHE IV model was good to excellent (AUC = 0.82, 95% CI: 0.78 - 0.86), but the calibration was inadequacy according to the results of HL test (Chi square = 20.49, $P = 0.009$). The BS showed good accuracy of APACHE IV model (BS = 0.18), but it may underestimate the in-hospital mortality (SMR = 1.50, 95% CI: 1.30 - 1.70) in elderly critically ill patients. **Conclusion** The APACHE IV model shows good discrimination in critically ill elderly patients, but has poor calibration and underestimated in-hospital mortality. Therefore, before we use the risk prediction models for geriatric patients, we should validate it and employ customized models to fit the population to be predicted.

【Key words】 acute physiology and chronic health evaluation; aged; hospital mortality; prognosis evaluation; validation

This work was supported by National Natural Science Foundation of China (6130010), the Program of Science and Technology Plan of Guangdong Province (2012A061400010), and the Innovative Major Project of Guangzhou City (201508020253).

Corresponding author: GUO Zhen-Hui, E-mail: micugzh@126.com

近年疾病严重程度评分、预测死亡率模型的研究与应用在危重症监护领域日益受到关注, 通过模

型可量化患者的生理状态, 从而便于临床医师评估疾病危重程度、预测预后及收集临床研究信息。急

收稿日期: 2017-01-17; 修回日期: 2017-04-27

基金项目: 国家自然科学基金(6130010); 广东省科技计划项目(2012A061400010); 广州市创新重大专项(201508020253)

通信作者: 郭振辉, E-mail: micugzh@126.com

性生理与慢性健康状况评分 (acute physiology and chronic health evaluation, APACHE) 系统在 2006 年更新为 APACHE IV 评分模型^[1]。重症监护病房 (intensive care unit, ICU) 中老年患者比例较高, 约 45% ~ 57%^[2]。老年患者由于器官功能的退化性改变以及伴随的多种基础疾病, 生理和病理生理方面具有特殊性, 因此对评估模型的要求也不同。APACHE IV 评分模型用于专科和综合 ICU 重症患者的研究结果表明其具有很好的预测价值^[3,4], 但它是否可用于评估老年重症患者病情严重程度和预后、以及评估的准确性尚没有得到验证, 为此我们进行了此方面的研究。

1 对象与方法

1.1 研究对象

入选 2011 年 7 月至 2015 年 7 月广州军区广州总医院老年重症医学科收治的 421 例重症老年患者, 回访至 2016 年 7 月。男性 359 例, 女性 62 例, 年龄 (81.5 ± 8.8) 岁。纳入标准: (1) 年龄 ≥ 60 岁; (2) 根据 1999 年美国危重病医学会 ICU 转入指南收住; (3) 治疗方法得到患者或家属知情同意。剔除标准: 入住 ICU 时间 < 24 h 或 ≥ 365 d; 重复入住 ICU; 器官捐献者。

1.2 方法

计算所有患者入住 ICU 24 h 内的 APACHE IV 评分, 利用 APACHE IV 评分模型评估生存和死亡患者人数。辨别力和校准被用于评估 APACHE IV 评分模型的精确度。辨别力为辨别生存与非生存患者的能力, 通过受试者工作特性 (receiver operating characteristic, ROC) 曲线下面积 (area under the

curve, AUC) 表示。校准为模型数据集中描述死亡样式的能力, 通过 Hosmer-Lemeshow 检验评估, 当预测模型的预测死亡式样与观察式样显著不一致, 则该预测模型的校准弱。标准化死亡比值 (standard mortality rate, SMR) 检验模型预测群体死亡率的准确性, 为实际死亡率和预测死亡率的比值。Brier 评分 (BS) 从模型整体评估精确度, 包含辨别力和校准, 为实际死亡率和预测死亡率两者的均方误差, 评分越接近 0, 预测模型越精准。Brier 评分和 SMR 分别计算, 用 95% 置信区间 (confidence interval, CI) 表示范围。根据年龄、入院类型、来源及是否脓毒症, 进行不同组的模型验证。来源中的院内和外院转入包含由普通病房和 ICU 转入的患者。

1.3 统计学处理

ROC 曲线下面积分析采用 MedCalc 12.7, 其余统计分析使用统计软件 SPSS19.0。计量资料用均数 ± 标准差 ($\bar{x} \pm s$) 表示, 两组比较采用 *t* 检验。计数资料用百分率表示, 组间比较用 χ^2 检验。*P* < 0.05 为差异有统计学意义。Hosmer-Lemeshow 检验若 *P* > 0.05 提示校准良好。

2 结果

2.1 患者一般资料

共入组 421 例患者, 140 例患者死亡 (33.3%)。生存患者和死亡患者的 APACHE IV 评分差异有统计学意义 (49.90 ± 21.89 *vs* 60.13 ± 27.15, *P* < 0.001)。入院类型、来源和器官支持需求差异有统计学意义 (*P* < 0.001)。慢性心功能不全 (NYHA 分级 ≥ 3 级)、痴呆、肿瘤组的生存患者与死亡患者人数差异有统计学意义 (*P* < 0.05; 表 1)。

表 1 患者的一般资料
Table 1 Baseline data of all subjects [n(%)]

Item	Survivors (n = 281)	Non-survivors (n = 140)	χ^2	<i>P</i> value
Comorbidities				
Chronic cardiac insufficiency	42 (14.9)	40 (28.6)	11.06	<0.001
Dementia	40 (14.2)	33 (23.6)	5.68	0.017
Neoplasm	30 (10.7)	27 (19.3)	5.92	0.015
Type of admission			31.88	<0.001
Department of internal medicine	196 (69.8)	128 (91.4)		
Emergency surgery	5 (1.8)	5 (3.6)		
Elective surgery	80 (28.5)	7 (5.0)		
Patient origin			220.83	<0.001
Emergency admission	103 (36.7)	20 (14.3)		
Department transference	162 (57.7)	112 (80.0)		
Hospital transference	16 (5.7)	8 (5.7)		
Sepsis	109 (38.8)	91 (65.0)	26.76	<0.001
Mechanical ventilation	86 (30.6)	119 (85.0)	110.67	<0.001
Renal replacement therapy	33 (11.7)	88 (62.9)	119.20	<0.001

2.2 APACHE IV 评分模型验证

APACHE IV 验证结果见表 2。该模型具有优秀的辨别力 ($AUC = 0.82, 95\% \text{ CI}: 0.78 \sim 0.86$; 图 1), 不适的校准 ($\chi^2 = 20.49, P = 0.009$; 图 2), 模型低估老年重症患者的住院死亡率 ($SMR = 1.50, 95\% \text{ CI}: 1.30 \sim 1.70$); 精确度良好 ($BS = 0.18$)。

APACHE IV 评估模型中的亚组结果表明, ≥ 80 岁患者和择期手术患者的校准不适 ($\chi^2 = 16.67, P = 0.034; \chi^2 = 17.61, P = 0.024$); 同时模型低估内科、院内转入以及脓毒症患者的住院死亡率 ($SMR = 1.53, 95\% \text{ CI}: 1.33 \sim 1.74; SMR = 1.51, 95\% \text{ CI}: 1.30 \sim 1.73; SMR = 1.48, 95\% \text{ CI}: 1.26 \sim 1.70$; 表 2)。

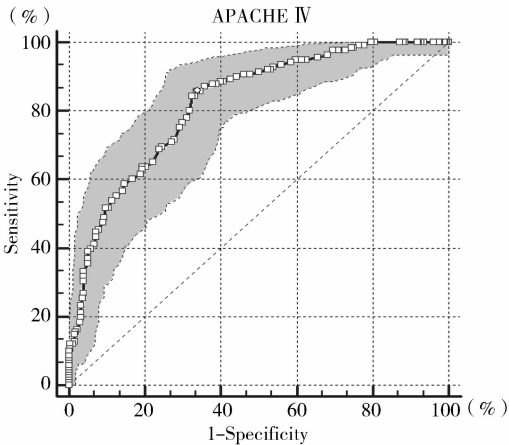


图 1 APACHE IV 的受试者工作特性曲线 (ROC) 下面积
Figure 1 Area under the receiver operating characteristic (ROC) curve of APACHE IV model
Shadow region is 95% confidence interval

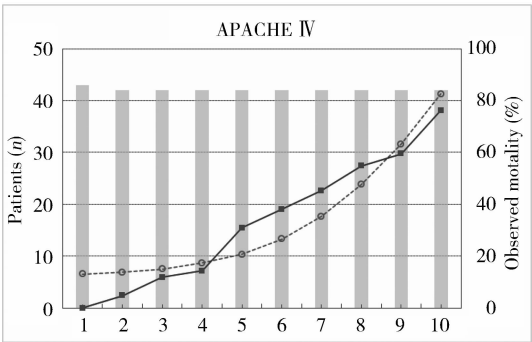


图 2 APACHE IV 的 Hosmer-Lemeshow (HL) 检验和校准曲线
Figure 2 Homer-Lemeshow goodness of fit test and calibration curves of APACHE IV model

Histogram is the number of patients in each group. The broken line demonstrates predicted mortality, and the solid line demonstrates actual mortality

3 讨论

APACHE IV 评分模型具有以下优势^[5]: 在旧版模型的基础上增加了更多参数; 统筹了 116 个诊断目录, 不同组有独立的系数, 使得能通过不同组推算死亡率; 可预测入住 ICU 时间; 为冠状动脉搭桥建立了专有预测目录。再加上大数据分析, 该模型的辨别力和校准功能变得更强^[3]。每一个新模型应用于新领域前需经过外部验证^[6]。国外不同国家、不同临床领域在应用 APACHE IV 评分模型前均进行模型验证^[4,7] 以检测模型的效能。

AUC 能表示模型评估住院死亡和生存患者的能力。美国、丹麦、韩国等的既往研究结果表明^[1,3,6,7], APACHE IV 评分模型的范围为 0.82 ~ 0.88,

表 2 APACHE IV 评分模型的表现
Table 2 Performance of APACHE IV model (n = 421)

Group	ROC curve		Hosmer-lemeshow test		Brier score(95% CI)	SMR (95% CI)
	AUC (95% CI)	P value	χ^2	P value		
Age						
60 - 79 years	0.84(0.77 - 0.91)	<0.001	10.79	0.214	0.19(0.15 - 0.24)	1.88(1.42 - 2.33)
≥ 80 years	0.81(0.76 - 0.86)	<0.001	16.67	0.034	0.18(0.15 - 0.21)	1.37(1.14 - 1.60)
Type of admission						
Department of internal medicine	0.79(0.74 - 0.84)	<0.001	9.71	0.286	0.18(0.15 - 0.21)	1.53(1.33 - 1.74)
Emergency surgery	0.80(0.49 - 1.00)	0.117	9.55	0.298	0.16(0.06 - 0.38)	2.21(0.84 - 3.57)
Elective surgery	0.84(0.68 - 1.00)	0.003	17.61	0.024	0.21(0.15 - 0.27)	0.92(0.27 - 1.58)
Patient origin						
Emergency new admission	0.83(0.74 - 0.92)	<0.001	8.22	0.412	0.22(0.16 - 0.27)	1.39(0.83 - 1.94)
Department transference	0.80(0.75 - 0.85)	<0.001	12.91	0.115	0.17(0.15 - 0.20)	1.51(1.30 - 1.73)
Hospital transference	0.79(0.58 - 1.00)	0.022	9.70	0.287	0.14(0.04 - 0.23)	1.62(0.71 - 2.54)
Sepsis	0.79(0.73 - 0.85)	<0.001	5.24	0.732	0.20(0.16 - 0.23)	1.48(1.26 - 1.70)

APACHE: acute physiology and chronic health evaluation; ROC: receiver operating characteristic; AUC: area under curve; SMR: standard mortality rate

本研究结果表明 APACHE IV 评分模型 AUC 为 0.82 (95% CI: 0.78 ~ 0.86), 辨别力为优秀。对于年龄 ≥ 80 岁的老年患者, 模型具有同等的辨别力 (AUC = 0.81)。而不同组分析结果表明模型对不同分类患者的辨别力并不一致, 对脓毒症患者的辨别力弱 (AUC = 0.79), 丹麦 Brinkman 等^[6]的研究则显示模型对心脏骤停患者辨别力弱 (AUC = 0.72)。

Hosmer-Lemeshow 检验 APACHE IV 评分模型的校准。结果表明它对于年龄 ≥ 80 岁和择期手术老年患者校准弱。很多验证研究结果^[3,6-10]也表明模型对专科人群校准不适, 可能是由于研究人群、地域、研究时间不同以及患者对姑息治疗的态度等因素造成^[6,11]。老年患者具有自身特点, 如合并症比例高、多种合并症并存、存在严重合并症、衰弱、生活能力下降、认知能力下降等^[12,13], 这些因素可导致其病情进展迅速、病死率高、需要长期器官功能支持概率高和救治医疗资源投入大, 也影响其预后。本研究也表明合并痴呆、肿瘤、慢性心功能不全等疾病的老年患者治疗预后差, 原有的 APACHE 系统并未全面的对这些因素加以合适的调整。

BS 综合评价模型辨别力和校准的结果表明^[3], APACHE IV 评分模型 (BS = 0.18) 具有良好的精确度, 但 BS 分值低于其他人的研究^[3,6,9]。SMR 检验模型预测群体住院死亡率的准确性结果表明, APACHE IV 评分模型低估了老年重症患者住院死亡率 (SMR = 1.50, 95% CI: 1.30 ~ 1.70), 与美国 (SMR = 0.997)^[1]、丹麦 (SMR = 0.87)^[6]、韩国研究结果 (SMR = 0.21)^[3] 不符。模型不同组的分析结果表明, 模型低估了不同年龄分层老年患者、内科疾病 (SMR = 1.53)、院内转入 (SMR = 1.51) 或脓毒症患者 (SMR = 1.48) 的死亡率, 符合笔者既往验证 APACHE IV 评分模型评价老年脓毒症患者的结果^[14], 模型低估老年重症患者的住院死亡率, 特别是对于某些分类群体。尽管国外研究人员针对旧、新模型的验证研究表明新版模型具有一些优势^[1,8,10,15], 但本研究表明模型的验证以及验证后的调整更重要^[3,6,10]。

本研究纳入超过 400 例患者, 但未探讨 APACHE IV 评分模型对于离退休军人 (62.0%) 住院预后的评价, 下一步我们将分析各种危险因素对预后的评估; 其次我们没有排除一些入院前状态、ICU 收治转出政策或离开 ICU 后的混杂因素干扰, 所以尽管 APACHE IV 评分模型对老年重症患者显示了优秀的辨别力, 但模型校准不良并低估某些亚组群体的住院死亡率。专科重症患者如使用该模型前

应验证模型, 并使用定制化模型拟合预测群体。

【参考文献】

- [1] Zimmerman JE, Kramer AA, McNair DS, *et al.* Acute physiology and chronic health evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients [J]. *Crit Care Med*, 2006, 34 (5): 1297 - 1310. DOI: 10.1097/01.CCM.0000215112.84523.F0.
- [2] Shum HP, Chan KC, Wong HY, *et al.* Outcome of elderly patients who receive intensive care at a regional hospital in Hong Kong [J]. *Hong Kong Med J*, 2015, 21 (6): 490 - 498. DOI: 10.12809/hkmj144445.
- [3] Lee H, Shon YJ, Kim H, *et al.* Validation of the APACHE IV model and its comparison with the APACHE II, SAPS 3, and Korean SAPS 3 models for the prediction of hospital mortality in a Korean surgical intensive care unit [J]. *Korean J Anesthesiol*, 2014, 67 (2): 115 - 122. DOI: 10.4097/kjae.2014.67.2.115.
- [4] Zhang Z, Chen K, Chen L. APACHE III outcome prediction in patients admitted to the intensive care unit with sepsis associated acute lung injury [J]. *PLoS One*, 2015, 10 (9): e139374. DOI: 10.1371/journal.pone.0139374.
- [5] Salluh JF, Soares M. ICU severity of illness scores [J]. *Curr Opin Crit Care*, 2014, 20 (5): 557 - 565.
- [6] Brinkman S, Bakhshi-Raiez F, Abu-Hanna A, *et al.* External validation of acute physiology and chronic health evaluation IV in Dutch intensive care units and comparison with acute physiology and chronic health evaluation II and simplified acute physiology score II [J]. *J Crit Care*, 2011, 26 (1): e11 - e18. DOI: 10.1016/j.jcrc.2010.07.007.
- [7] Wong RS, Ismail NA, Tan CC. An external independent validation of APACHE IV in a Malaysian intensive care unit [J]. *Ann Acad Med Singapore*, 2015, 44 (4): 127 - 132.
- [8] Sanchez-Casado M, Hostiguela-Martin VA, Raigal-Cano A, *et al.* Predictive scoring systems in multiorgan failure: a cohort study [J]. *Med Intensiva*, 2016, 40 (3): 145 - 153. DOI: 10.1016/j.medin.2015.03.005.
- [9] Keegan MT, Gajic O, Afessa B. Comparison of APACHE III, APACHE IV, SAPS 3, and MPM0 III and influence of resuscitation status on model performance [J]. *Chest*, 2012, 142 (4): 851 - 858. DOI: 10.1378/chest.11-2164.
- [10] Mann SL, Marshall MR, Woodford BJ, *et al.* Predictive performance of acute physiological and chronic health evaluation releases II to IV: a single New Zealand centre experience [J]. *Anaesth Intensive Care*, 2012, 40 (3): 479 - 489.
- [11] Nassar Junior AP, Mocelin AO, Andrade FM, *et al.* SAPS 3, APACHE IV or GRACE: which score to choose for acute coronary syndrome patients in intensive care units? [J]. *Sao Paulo Med J*, 2013, 131 (3): 173 - 178.
- [12] Heyland DK, Garland A, Bagshaw SM, *et al.* Recovery after critical illness in patients aged 80 years or older: a multi-center prospective observational cohort study [J]. *Intensive Care Med*, 2015, 41 (11): 1911 - 1920. DOI: 10.1007/s00134-015-4028-2.
- [13] Kenneth R, Guo ZH, Song XW. The frail older adult and the intensive care unit [J]. *Chin J Geriatr*, 2013, 32 (1): 3 - 8.
- [14] 董家辉, 孙杰, 陈蕊, 等. 急性生理学及慢性健康状况评分系统 II/IV 对老年脓毒症患者预后的预测价值 [J]. *中华危重病急救医学*, 2013, 25 (10): 594 - 599. DOI: 10.3760/cma.j.issn.2095-4352.2013.10.005.
- [15] Dong JH, Sun J, Chen R, *et al.* Assessment of the value of acute physiology and chronic health evaluation II/IV prognostic models in elderly patients with sepsis [J]. *Chin Crit Care Med*, 2013, 25 (10): 594 - 599. DOI: 10.3760/cma.j.issn.2095-4352.2013.10.005.
- [15] Xing X, Gao Y, Wang H, *et al.* Performance of three prognostic models in patients with cancer in need of intensive care in a medical center in China [J]. *PLoS One*, 2015, 10 (6): e131329.