

• 论著 •

基于机器学习模型的脓毒症患者 有创机械通气早期预警方法

刘万军¹ 肖文艳¹ 张金¹ 胡娟娟¹ 黄珊珊¹ 刘瑜² 华天凤¹ 杨旻¹

¹安徽医科大学第二附属医院重症医学科二科,心肺复苏与重症医学科实验室,合肥 230601;

²安徽大学智能计算与信号处理实验室,合肥 230601

通信作者:杨旻, Email: yangmin@ahmu.edu.cn

【摘要】 目的 建立一种用于识别脓毒症机械通气人群中高危患者的方法,并基于该方法对此类患者进行表型分析。**方法** 使用美国重症监护医学信息数据库(MIMIC-IV 2.0、MIMIC-III 1.4)、电子重症监护病房协作研究数据库 2.0(eICU-CRD 2.0),以及安徽医科大学第二附属医院的本地数据库,纳入符合脓毒症 3.0 诊断标准且首次入重症监护病房(ICU)首日接受有创机械通气(IMV)的成人患者。将数据完整度最高的 MIMIC-IV 数据集按照 6:1 的比例分为训练集和测试集,其他数据集作为验证集。提取患者人口统计学信息、合并症、实验室指标、ICU 常用评分及治疗措施。使用患者入 ICU 首日内临床资料计算序贯器官衰竭评分(SOFA),并使用 K 均值聚类法对患者 SOFA 评分子项进行聚类,采用平方误差和(SSE)及 Davies-Boulding 指数(DBI)确定最佳聚类数。针对聚类结果确定疾病表型数,采用归一化方法,以可视化方式分析比较不同表型患者间基线资料的差异,并采用 Kaplan-Meier 曲线比较不同表型的临床结局。**结果** 最终共纳入 MIMIC-IV 数据集 11 166 例患者、MIMIC-III 数据集 4 821 例患者、eICU-CRD 数据集 6 624 例患者、本地数据集 110 例患者,4 个数据集患者中位年龄相似,男性比例均超过 50%。以 MIMIC-IV 数据集 85% 样本作为训练集、15% 作为测试集,其他数据集作为验证集。基于 SOFA 的 6 个子项评分,采用 K 均值聚类法确定最佳聚类数为 3,并基于此将患者划分为 3 种表型。训练集分析结果显示,与表型 II 和表型 III 患者相比,表型 I 患者循环及呼吸功能损伤更重,使用血管活性药物比例更高,代谢性酸中毒、缺氧症状更明显,充血性心力衰竭发生率更高;表型 II 患者以呼吸功能障碍为主,表现出较高的内脏损伤;表型 III 患者器官功能相对稳定。上述特征在测试集和验证集中也有相同表现。感染相关指标分析显示,表型 I 患者入 ICU 7 d 内 SOFA 评分始终保持最高,血小板计数(PLT)呈先下降后上升趋势,中性粒细胞、淋巴细胞、单核细胞的绝对值亦高于表型 II 和表型 III 患者,血培养中革兰阳性菌、革兰阴性菌和真菌的阳性率显著高于表型 II 和表型 III 患者。Kaplan-Meier 曲线显示,在训练集、测试集、验证集中,表型 I 患者 28 d 累积病死率均显著高于表型 II 和表型 III 患者。**结论** 基于无监督机器学习方法推导出接受 IMV 的脓毒症患者存在 3 种独特表型,其中表型 I 以呼吸、循环衰竭为特征,可以用于对此类人群中高危患者的早期识别,且该人群更倾向为血流感染高危人群,预后较差。

【关键词】 脓毒症; 机械通气; 临床表型; 机器学习

基金项目:安徽省转化医学科研项目(2022zhyyx-C46)

DOI: 10.3760/cma.j.cn121430-20240422-00368

Early warning method for invasive mechanical ventilation in septic patients based on machine learning model

Liu Wanjun¹, Xiao Wenyan¹, Zhang Jin¹, Hu Juanjuan¹, Huang Shanshan¹, Liu Yu², Hua Tianfeng¹, Yang Min¹

¹The Second Department of Critical Care Medicine, the Second Affiliated Hospital of Anhui Medical University, the Laboratory of Cardiopulmonary Resuscitation and Critical Care Medicine, Hefei 230601, China; ²Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education, Anhui University, Hefei 230601, China

Corresponding author: Yang Min, Email: yangmin@ahmu.edu.cn

【Abstract】 Objective To develop a method for identifying high-risk patients among septic populations requiring mechanical ventilation, and to conduct phenotypic analysis based on this method. **Methods** Data from four sources were utilized: the Medical Information Mart for Intensive Care (MIMIC-IV 2.0, MIMIC-III 1.4), the Philips eICU-Collaborative Research Database 2.0 (eICU-CRD 2.0), and the Anhui Medical University Second Affiliated Hospital dataset. The adult patients in intensive care unit (ICU) who met Sepsis-3 and received invasive mechanical ventilation (IMV) on the first day of first admission were enrolled. The MIMIC-IV dataset with the highest data integrity was divided into a training set and a test set at a 6:1 ratio, while the remaining datasets were served as validation sets. The demographic information, comorbidities, laboratory indicators, commonly used ICU scores, and treatment measures of patients were extracted. Clinical data collected within first day of ICU admission were used to calculate the sequential organ failure assessment (SOFA) score. K-means clustering was applied to cluster SOFA score components, and the sum of squared errors (SSE) and Davies-Bouldin index (DBI) were used to determine the optimal number of disease subtypes. For clustering results, normalized methods were employed to compare baseline characteristics by visualization, and Kaplan-Meier curves were used to analyze clinical outcomes across phenotypes. **Results** This study enrolled patients

from MIMIC-IV dataset ($n = 11\,166$), MIMIC-III dataset ($n = 4\,821$), eICU-CRD dataset ($n = 6\,624$), and a local dataset ($n = 110$), with the four datasets showing similar median ages and male proportions exceeding 50%; using 85% of the MIMIC-IV dataset as the training set, 15% as the test set, and the rest dataset as the validation set. K-means clustering based on the six-item SOFA score was performed to determine the optimal number of clusters as 3, and patients were finally classified into three phenotypes. In the training set, compared with the patients with phenotype II and phenotype III, those with phenotype I had the more severe circulatory and respiratory dysfunction, a higher proportion of vasoactive drug usage, more obvious metabolic acidosis and hypoxia, and a higher incidence of congestive heart failure. The patients with phenotype II was dominated by respiratory dysfunction with higher visceral injury. The patients with phenotype III had relatively stable organ function. The above characteristics were consistent in both the test and validation sets. Analysis of infection-related indicators showed that the patients with phenotype I had the highest SOFA score within 7 days after ICU admission, initial decreases and later increases in platelet count (PLT), and higher counts of neutrophils, lymphocytes, and monocytes as compared with those with phenotype II and phenotype III, their blood cultures had a higher positivity rates for Gram-positive bacteria, Gram-negative bacteria and fungi as compared with those with phenotype II and phenotype III. The Kaplan-Meier curve indicated that in the training, test, and validation sets, the 28-day cumulative mortality of patients with phenotype I was significantly higher than that of patients with phenotypes II and phenotype III. **Conclusions** Three distinct phenotypes in septic patients receiving IMV based on unsupervised machine learning is derived, among which phenotype I, characterized by cardiorespiratory failure, can be used for the early identification of high-risk patients in this population. Moreover, this population is more prone to bloodstream infections, posing a high risk and having a poor prognosis.

[Key words] Sepsis; Mechanical ventilation; Clinical phenotype; Machine learning

Fund program: Anhui Provincial Translational Medicine Scientific Research Project (2022zhxy-C46)

DOI: 10.3760/cma.j.cn121430-20240422-00368

脓毒症患者存在异质性,因此即使诊断标准相同,预后和治疗需求也存在差异^[1-3]。有创机械通气(invasive mechanical ventilation, IMV)是脓毒症患者的常见治疗方式,超过 40% 的脓毒症患者在住院期间接受 IMV 治疗^[4-5]。在这组具有异质性的患者中,使用 IMV 应适应更精细的管理计划,有可能改善医疗资源分配及治疗效果^[6]。脓毒症致急性呼吸窘迫综合征(acute respiratory distress syndrome, ARDS)是患者接受 IMV 的常见原因^[7-8]。然而,脓毒症患者接受 IMV 时肺部的易感性更高,疾病进展和恶化速度更快,且并发症可能增加^[9-11]。尽管如此,关于脓毒症患者使用 IMV 的指征,以及在这些高危患者中实施早期预警和监护提升的有效方法仍缺乏共识。因此,早期识别及区分接受机械通气的脓毒症患者可能有利于治疗、疾病评估和精准重症监护。本研究假设基于脓毒症患者接受 IMV 的临床特征可以区分其异质性,并使用无监督机器学习模型从大量真实世界的临床数据中获取经验,以便在早期阶段更有效、更容易地区分和预测这些患者的预后。

1 资料与方法

1.1 数据来源:本研究基于 4 个数据源,包括 3 个大型公共重症监护数据库,即重症监护医学信息数据库(Medical Information Mart for Intensive Care, MIMIC-IV 2.0、MIMIC-III 1.4)、电子重症监护病房协作研究数据库(eICU Collaborative Research Database, eICU-CRD; 版本 2.0),以及安徽医科大学第二附属医院的本地数据库。本次研究者均接受了数据库官方网站培训,

并在开始研究前获得数据集的使用许可(证书编号:39691989)。由于这些数据集中的患者私人信息已加密,因此官方网站免除了伦理审批。在使用本地数据集之前,研究已通过安徽医科大学第二附属医院医学研究伦理委员会审批(审批号:YX2022-146)。本地病例信息从医院电子病历系统中收集,并已获得每位患者的知情同意。所有数据源的数据均经过授权批准,并依照相关政策和指南合法使用。

1.2 研究人群与定义:本研究纳入了 4 个数据库中符合脓毒症 3.0 诊断标准^[12]且首次入重症监护病房(intensive care unit, ICU)首日接受 IMV 治疗的成人患者。为减少偏差,未纳入 24 h 后接受 IMV 治疗的患者,对于多次接受 IMV 治疗的患者,仅将其首次 IMV 治疗的数据纳入分析。

1.3 数据集划分与变量提取:将 MIMIC-IV 数据集按照 6:1 的比例分为训练集和测试集,其他数据集作为验证集。在 3 个公开可用的数据库中,使用官方提供的标准化查询语言代码包提取变量;本地数据集则经过人工筛选,从本院电子病历系统中获得。

提取变量包括:① 人口统计学信息:年龄、性别及体质量指数(body mass index, BMI)。② 合并症:基于国际疾病分类代码提取,包括心肌梗死、慢性肺病、失代偿期肝病、恶性肿瘤、充血性心力衰竭和糖尿病。③ 实验室指标:白细胞计数(white blood cell count, WBC)、血红蛋白(hemoglobin, Hb)、血小板计数(platelet count, PLT)、血肌酐(serum creatinine, SCr)、pH 值、动脉血氧分压(arterial partial pressure of

oxygen, PaO_2)、动脉血二氧化碳分压(arterial partial pressure of carbon dioxide, PaCO_2)、氧合指数($\text{PaO}_2/\text{FiO}_2$)及碱剩余(base excess, BE),这些变量均取入 ICU 首日最差值;使用时间加权平均值计算持续监测指标(即值的时间曲线下面积除以时间),包括心率(heart rate, RR)、呼吸频率(respiratory rate, RR)、平均动脉压(mean arterial pressure, MAP)、脉搏血氧饱和度(pulse oxygen saturation, SpO_2)等。④ ICU 常用评分:查尔森合并症指数(Charlson comorbidity index, CCI)评分、格拉斯哥昏迷评分(Glasgow coma score, GCS)、简化急性生理学评分 II (simplified acute physiology score II, SAPS II)、牛津急性疾病严重程度评分(Oxford acute severity of illness score, OASIS)及序贯器官衰竭评分(sequential organ failure assessment, SOFA)。⑤ 对于连续性肾脏替代治疗(continuous renal replacement therapy, CRRT)、血管活性药物等治疗措施,根据是否使用提取为二分类变量。上述变量提取时间限于入 ICU 24 h 内;动态分析时,按上述方法提取每日变量值。

无监督机器学习模型需要具有代表性和等尺度变量。本研究将患者入 ICU 首日 SOFA 评分中 6 个子项作为聚类变量。分析不同表型的临床结局变量。

1.4 数据分析:连续变量均不符合正态分布,以中位数(四分位数)[$M(Q_1, Q_3)$]表示,组间比较采用 Mann-Whitney U 检验或 Kruskal-Wallis 检验;分类变量以频数(百分比)表示,组间比较采用 χ^2 检验或 Fisher 精确概率法。3 个数据库中部分数据存在异常值和缺失值。为了避免极端异常值的影响,本研究综合考虑了临床实践和数据库中的数据分布,排除了连续监测中超过 99.9% 分位数的数据。对于缺失比例超过一半的变量,也予以排除,以避免缺失值对分析结果造成影响;如缺失数据比例 $< 50\%$,则采用多重插补法。本地数据集没有缺失数据。

采用 K 均值聚类算法对纳入患者的器官功能障碍类型数据进行聚类。由于 K 均值算法存在随机性,故分别采用平方误差和(sum of squared errors, SSE)及 Davies-Boulding 指数(Davies-Boulding index, DBI)两种方法确定最佳聚类数。聚类完成后,使用 t 分布随机邻域嵌入算法进行降维,实现数据可视化。

绘制 Kaplan-Meier 曲线,分析 28 d 累积病死率及其 95% 置信区间(95% confidence interval, 95%CI),并使用对数秩检验(Log-rank 检验)进行比较。使用 Mann-Kendall 检验进行趋势分析。K 均值算法基于欧氏距离,通过比较新样本与原始聚类中心的最小

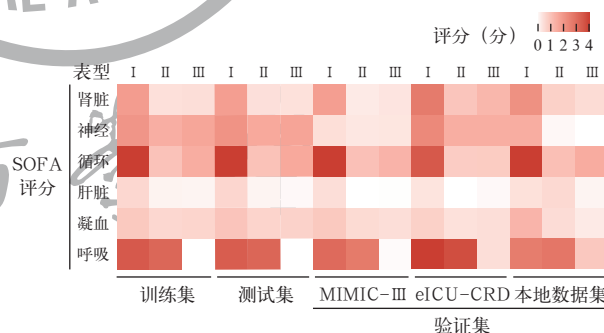
欧氏距离来分配。基于 Logistic 回归模型比较不同表型之间的病原体和感染部位特征,使用优势比(odds ratio, OR)进行描述。

使用 R (版本 4.1.3) 及 Python (版本 3.7.3) 软件建模、完成统计分析、绘制图形。检验水准 α 值取双侧 0.05。

2 结果

2.1 人群基线:最终纳入 MIMIC-IV 数据集 11 166 例患者、MIMIC-III 数据集 4 821 例患者、eICU-CRD 数据集 6 624 例患者、本地数据集 110 例患者。4 个数据集患者中位年龄相似,男性比例均超过 50%。由于 MIMIC-IV 数据集患者数量最多且缺失数据比例较低,故作为模型构建基础。MIMIC-IV 数据集的 85% 作为训练集、15% 作为测试集,其他数据集作为验证集。利用 SOFA 评分 6 个子项,采用 K 均值聚类算法、SSE 及 DBI 确定最佳聚类数。3 个聚类中,基于 K 均值聚类算法、SSE 和 DBI 瞬时变化保持在一个局部较小值,最终根据入 ICU 首日 SOFA 子项评分将患者群体划分为 3 种表型,分别定义为表型 I、II、III。

2.2 不同表型患者器官损伤特点(图 1):热图显示,不同表型患者 SOFA 评分 6 个子项间存在显著差异。训练集中,表型 I 患者循环、呼吸评分高于表型 II 和表型 III 患者;表型 II 患者呼吸评分较为突出,而表型 III 患者则器官损伤表现较少。测试集和验证集中 3 种表型患者也表现出相似的差异。



注:将重症监护医学信息数据库 IV (MIMIC-IV) 数据集按照 6:1 的比例分为训练集和测试集, MIMIC-III、电子重症监护病房协作研究数据库(eICU-CRD)及安徽医科大学第二附属医院本地数据库 3 个数据集作为验证集;IMV 为有创机械通气, ICU 为重症监护病房, SOFA 为序贯器官衰竭评分

图 1 不同数据集中 3 种表型脓毒症 IMV 患者入 ICU 首日 SOFA 评分 6 个子项均值热图

为了更详细地展示不同表型之间的差异,进行了 3 种表型之间的两两比较。训练集分析结果显示(图 2),与表型 II 和表型 III 患者相比,表型 I 患者始终表现出更高的血管活性药物使用率、更严重的器

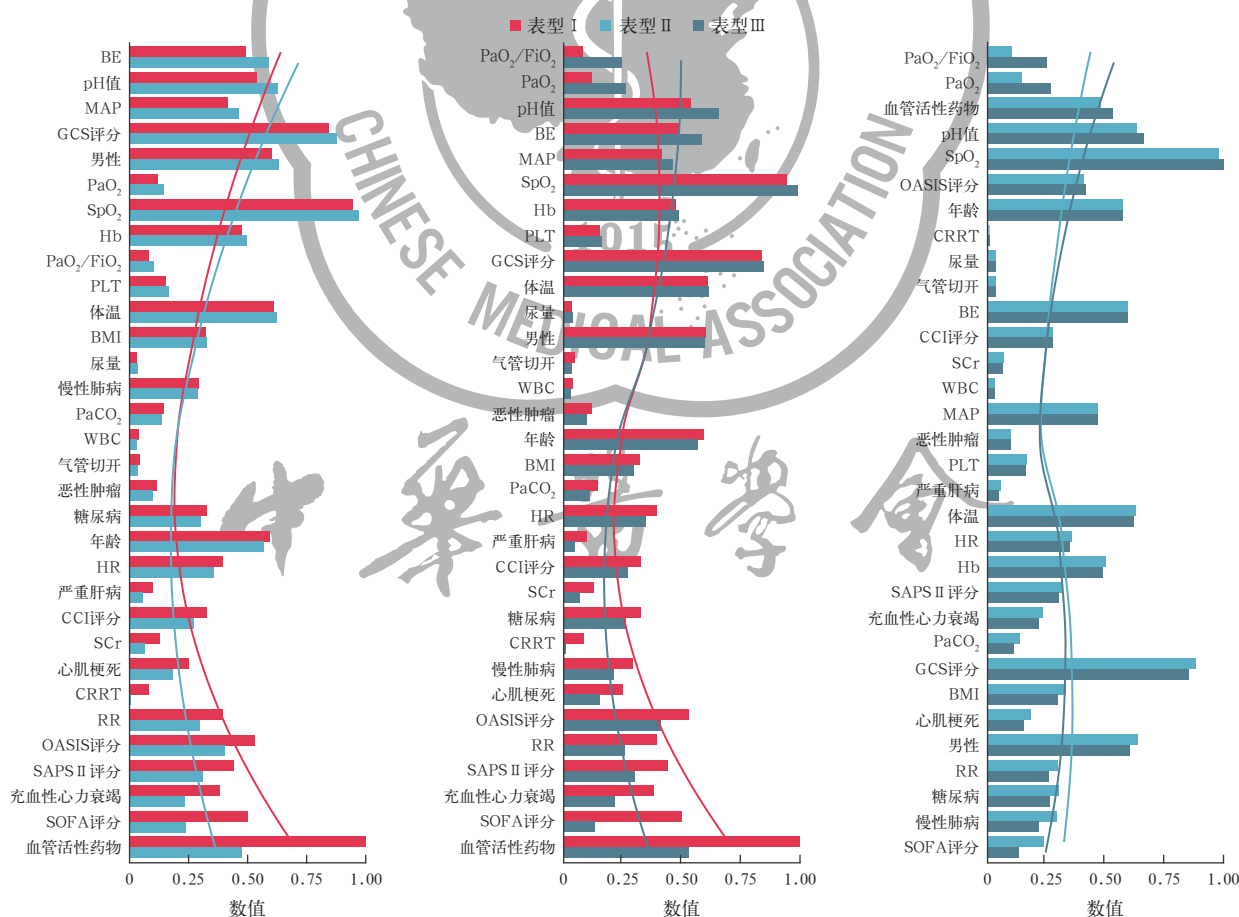
官损伤和更高的充血性心力衰竭发生率。此外,表型 I 患者还具有一些特征性生物指标,其内环境较表型 II 患者更偏酸性(基础过剩、pH 值),较表型 III 患者表现出更明显的缺氧症状。对表型 II 和表型 III 患者进行比较时发现,表型 II 患者可能表现出更高层次的内脏损伤,以呼吸相关指标($\text{PaO}_2/\text{FiO}_2$ 、 PaO_2 和 RR)为主要表现。这些特征在测试集和验证集中也有相同表现。

2.3 不同表型患者感染相关指标特征:对 MIMIC-IV 数据集中脓毒症患者入 ICU 7 d 内感染相关指标进行可视化,结果显示(图 3),表型 I 患者 SOFA 评分入 ICU 7 d 内始终保持最高,且呈现先上升后下降趋势。表型 I 患者 PLT 整体较低,并呈现先下降后上升趋势;表型 II 和表型 III 患者初始 PLT 高于表型 I 患者,且整体呈现上升趋势。在炎症细胞比较中,3 种表型患者入 ICU 7 d 内 WBC 均有所下降,但整体下降趋势差异无统计学意义。表型 I 患者入 ICU 7 d 内

中性粒细胞、淋巴细胞和单核细胞的绝对值均高于表型 II 和表型 III 患者。

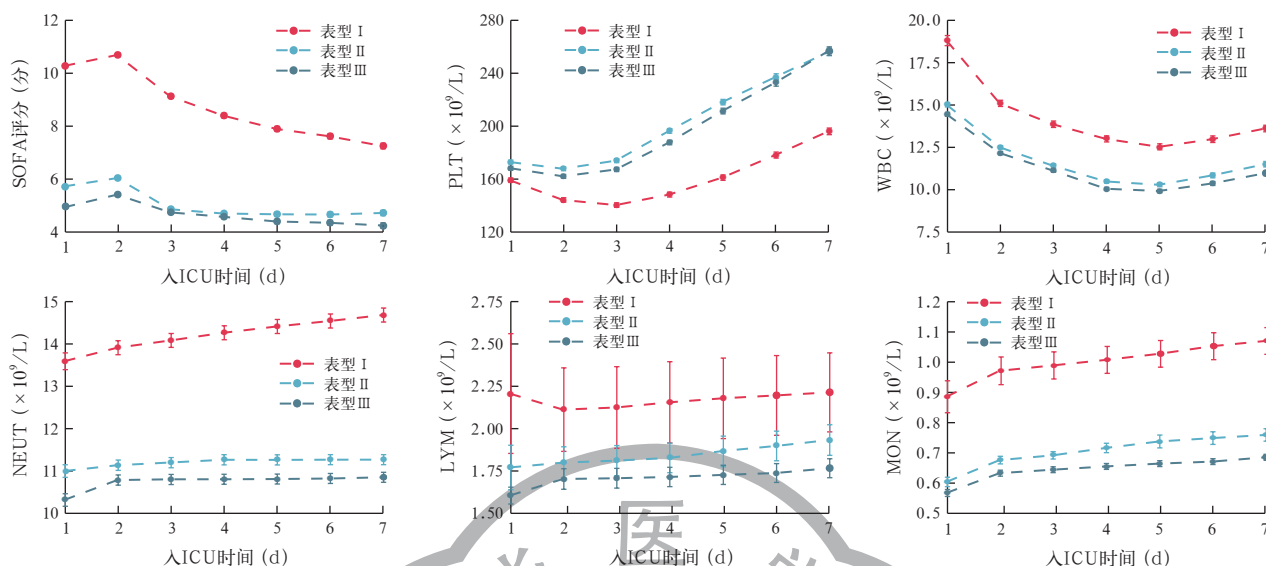
3 种表型患者病原体和感染部位的特征性分析显示(图 4),与表型 II 和表型 III 患者相比,表型 I 患者血培养中革兰阳性菌、革兰阴性菌和真菌的阳性率更高(均 $P < 0.05$)。说明表型 I 脓毒症患者更倾向为血流感染高危人群。在感染部位的比较中,表型 I 患者肺部感染、泌尿系感染和肠道感染的比例均高于表型 II 患者(均 $P < 0.05$);与表型 III 患者相比,表型 II 患者泌尿系感染比例更低,而肺部感染比例更高(均 $P < 0.05$)。说明表型 II 脓毒症患者更倾向为肺部感染高危人群。

2.4 不同表型患者临床结局特征:Kaplan-Meier 曲线显示(图 5),在训练集、测试集及验证集(MIMIC-III 数据集、eICU-CRD 数据集、本地数据集)中,表型 I 患者 28 d 累积病死率均显著高于表型 II 和表型 III 患者(均 $P < 0.05$)。



注: IMV 为有创机械通气, BE 为碱剩余, MAP 为平均动脉压, GCS 为格拉斯哥昏迷评分, PaO_2 为动脉血氧分压, SpO_2 为脉搏血氧饱和度, Hb 为血红蛋白, $\text{PaO}_2/\text{FiO}_2$ 为氧合指数, PLT 为血小板计数, BMI 为体质量指数, PaCO_2 为动脉血二氧化碳分压, WBC 为白细胞计数, HR 为心率, CCI 为查尔森合并症指数, SCr 为血肌酐, CRRT 为连续性肾脏替代治疗, RR 为呼吸频率, OASIS 为牛津急性疾病严重程度评分, SAPS II 为简化急性生理学评分 II, SOFA 为序贯器官衰竭评分;所有基线变量均归一化至 0~1

图 2 不同表型脓毒症 IMV 患者生物指标比较的侧向柱状图



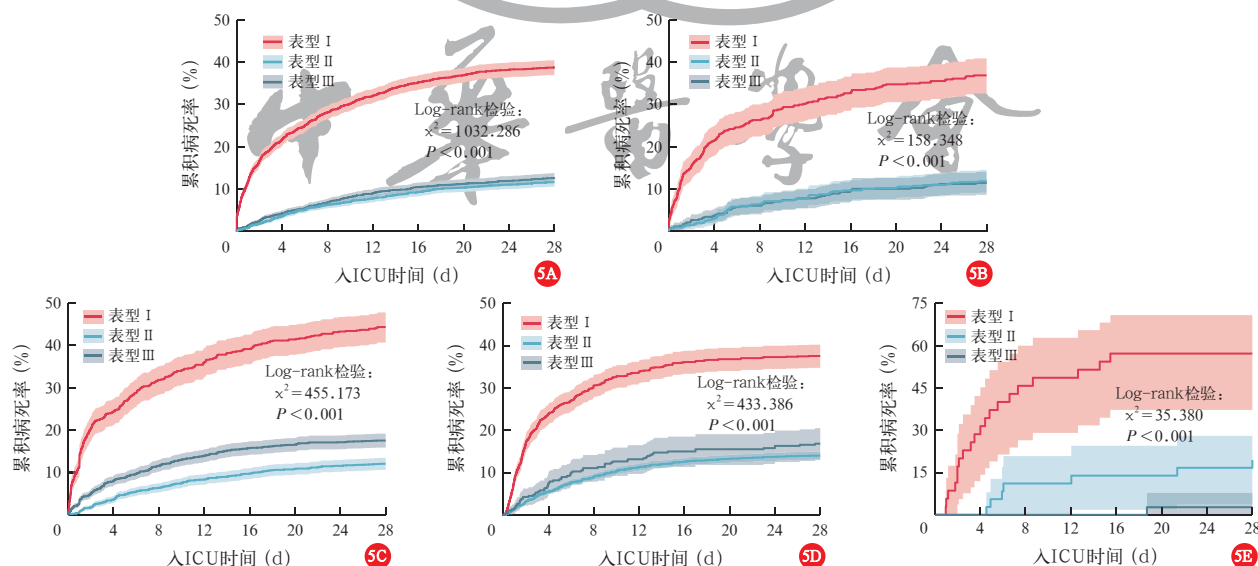
注: IMV 为有创机械通气, ICU 为重症监护病房, SOFA 为序贯器官衰竭评分, PLT 为血小板计数, WBC 为白细胞计数, NEUT 为中性粒细胞计数, LYM 为淋巴细胞计数, MON 为单核细胞计数

图3 不同表型脓毒症 IMV 患者入 ICU 7 d 内 SOFA 评分及感染相关指标的变化趋势

指标	表型 I 比表型 II		表型 I 比表型 III		表型 II 比表型 III	
	OR (95%CI)	P值	OR (95%CI)	P值	OR (95%CI)	P值
血培养结果						
革兰阳性菌	1.75 [1.40, 2.19]	<0.001	1.73 [1.38, 2.18]	<0.001	0.99 [0.77, 1.27]	0.931
革兰阴性菌	3.57 [2.40, 5.44]	<0.001	3.49 [2.33, 5.39]	<0.001	0.98 [0.59, 1.63]	0.935
真菌	4.11 [2.38, 7.56]	<0.001	4.71 [2.61, 9.25]	<0.001	1.15 [0.54, 2.50]	0.726
感染部位						
肺部感染	1.43 [1.30, 1.58]	<0.001	2.24 [2.01, 2.50]	<0.001	1.57 [1.40, 1.75]	<0.001
泌尿系感染	1.36 [1.18, 1.56]	<0.001	1.03 [0.90, 1.17]	0.706	0.76 [0.66, 0.87]	<0.001
肠道感染	1.74 [1.37, 2.23]	<0.001	1.99 [1.54, 2.59]	<0.001	1.14 [0.86, 1.51]	0.353

注: IMV 为有创机械通气, OR 为优势比, 95%CI 为 95% 置信区间

图4 不同表型脓毒症 IMV 患者血培养结果和感染部位比较的森林图



注: 将重症监护医学信息数据库Ⅳ (MIMIC-Ⅳ) 数据集按照 6:1 的比例分为训练集和测试集, MIMIC-Ⅲ、电子重症监护病房协作研究数据库 (eICU-CRD) 及安徽医科大学第二附属医院本地数据库 3 个数据集作为验证集; IMV 为有创机械通气, ICU 为重症监护病房

图5 训练集(A)、测试集(B)及验证集 [MIMIC-Ⅲ数据集(C)、eICU-CRD 数据集(D)、本地数据集(E)] 中不同表型脓毒症 IMV 患者 28 d 累积病死率的 Kaplan-Meier 曲线

3 讨论

本研究提出,接受 IMV 的脓毒症患者可划分为 3 种表型,分别定义为表型 I、II、III。在这些表型中,表型 I 患者预后最差,其特点是呼吸、心血管功能障碍及炎症反应升高。相比之下,表型 II 患者主要表现为呼吸功能障碍,而表型 III 患者器官损伤最轻,这两种表型患者预后相似。这些表型在不同数据来源的训练集、测试集和验证集中均表现出较强的通用性。此外,这 3 种表型在临床特征、器官功能障碍模式和预后方面显示出良好的区分度。这些表型可在早期阶段进行识别,为完善接受 IMV 的脓毒症患者的管理提供参考。根据脓毒症 3.0 定义,疑似感染被视为新的脓毒症诊断标准,诊断为脓毒症的患者并不总是发展为脓毒性休克或更加严重的器官功能衰竭^[12]。ARDS 可能是肺部感染继发的,也可能是由休克、创伤等肺外因素引起的,脓毒症与 ARDS 的关系错综复杂。研究表明,脓毒症相关 ARDS 患者表现出不同类型的炎症反应和器官功能障碍模式^[13-16]。基于过去的经验,区分和管理高风险人群有助于有效分配医疗资源并提高治疗效果^[1,17-18]。然而,目前临床上尚无可靠的方法确定这类患者的异质性。本研究涉及一个因感染、器官损伤和外力同时作用而表现出复杂表型的人群。表型模型为理解与区分这一人群的临床特征和预后提供了一种新方法。在本研究中,3 种表型是根据患者入 ICU 24 h 内 SOFA 评分 6 个子项确定的,因此可应用于有或无脓毒症的患者,并可进一步推广至所有接受机械通气的患者。该表型模型在重症监护数据库及急诊数据库中均得到了验证,临床特征和结局的分布一致,结果稳健。此外,在线工具的应用为临床医生提供了一个快速评估患者表型的新途径,使得在启动 IMV 时能够迅速识别患者表型,并根据表型制定个体化治疗方案。这不仅有助于早期发现高危患者,还可以提高医疗资源的利用效率。期待未来能继续深入探索和研究,进一步完善和优化这一表型分类方法,并将其应用于更广泛的临床实践中。

前期研究已确定了脓毒症和 ARDS 的 3~6 个临床表型,这些表型来源包括电子健康记录、随机对照试验和转录组学数据^[1-3,17-18]。在脓毒症患者中,2017 年的一项回顾性研究探索了被称为“Mars1”的脓毒症遗传亚型,并确定了与 28 d 死亡相关的基因型^[19]。这表明基因表达可能是脓毒症异质性的根本原因。然而,在基因水平上进行诊断和分类既耗时

又费力。Seymour 等^[1]提出了 γ 表型与 δ 表型作为脓毒症患者的超炎症表型和脓毒性休克表型,这些表型的临床特征和预后与本研究中的表型 I 患者相似。与之类似,Bhavani 等^[17]根据生命体征的轨迹将脓毒症患者分为 A、B、C、D 4 种表型,其中 A 型及 D 型患者特征也是超炎症反应、低血压和预后不良。本研究中表型 I 患者的特征性临床表现是高水平的炎症和循环功能障碍,与上述两种大规模脓毒症表型分类模型一致。此外,ARDS 患者之间也存在异质性。对于 ARDS,尽管其定义和诊断标准相对明确,但患者之间的临床表现和预后仍存在显著差异。许多研究者试图通过不同分类方法揭示 ARDS 的异质性,以期更好地指导治疗和预测预后。然而,目前尚缺乏被广泛接受和验证的 ARDS 表型分类方法。Liu 等^[18]基于 21 个临床变量提出了 3 种 ARDS 表型,其中一种表型与肾功能障碍和酸中毒相关的最高病死率有关。在本研究中,表型 I 患者内环境呈酸性并伴有肾损伤,这能解释脓毒症中的炎症和灌注障碍,并与 ARDS 研究中的表型特征一致。然而,在表型 I 患者中,SCr 升高(非尿量减少)是肾损伤的主要原因。在临床实践中,SCr 变化在危重患者中滞后于肾脏疾病的发展,这表明表型 I 患者的肾损伤可能是疾病发展的一个启动因素。为了深入了解表型 I 患者的肾损伤机制,可以进一步探索与 SCr 升高相关的潜在因素,如肾灌注不足、药物不良反应或潜在的慢性肾脏疾病。此外,通过长期随访可持续追踪这些患者的肾功能恢复情况,以及肾损伤对整体预后和长期生存的影响。值得注意的是,虽然本研究发现了与脓毒症和 ARDS 相关的不同表型,但患者的临床表现可能并非完全独立。在某些情况下,患者可能同时表现出多种表型的特征。因此,未来的研究可以考虑采用更复杂的分类方法,以更全面地描述患者的临床异质性。总之,通过对表型 I 患者的深入分析,能更好地理解脓毒症和 ARDS 的发病机制,并为临床决策提供更有针对性的建议,以提升疗效,改善生活质量,降低医疗成本。

本研究中另外两种表型也具有易于区分的临床特征。表型 II 患者的主要临床特征是呼吸功能不全, $\text{PaO}_2/\text{FiO}_2$ 较低,而其他器官损伤不太明显。在临床实践中,患者因呼吸衰竭而单独接受 IMV 是常见的;然而,这些患者死亡的主要原因并不是孤立的单一因素,而是机体的级联反应导致多器官损伤^[20-22]。在 3 种表型中,表型 III 患者的器官损伤最少,预后最

好,其病原体播散情况与表型Ⅱ患者相当。值得注意的是,表型Ⅲ患者的感染部位与表型Ⅱ患者不同;然而,两种表型之间的短期预后没有显著差异,与既往研究者发现脓毒症群体中存在临床特征不同但结局相似患者的现象一致^[23]。利用这些表型数据优化和融合预测模型,有助于提高对脓毒症患者病情变化的预测能力,进而改善患者的治疗效果和生活质量,为未来的临床研究提供了新的思路 and 方向。

本研究的局限性:首先,尽管在结果方面对接受 IMV 脓毒症患者的分类相对简单,但分类的计算涉及 SOFA 评分、数据处理和聚类算法,难以手动完成,本研究中基于网络的计算表型工具的应用仍然有限,未来需进一步完善和优化,提高其易用性和准确性,以便更广泛地应用于临床实践。其次,训练集、测试集和部分验证集仅基于美国公共数据库构建,且验证集样本量较小,需要获取更大的样本量进一步验证。再次,由于数据不足,本研究中病原体感染部位的解释并未得到完全验证。最后,本研究为回顾性研究,不同表型的临床结局评估可能受潜在混杂因素的影响,需要设计并实施更严格的临床对照研究,进一步验证不同表型的临床结局,并探索潜在的混杂因素对研究结果的影响。

4 结 论

本研究确定了接受 IMV 的脓毒症患者的 3 种独特表型。这 3 种表型各具特色,其中表型Ⅰ患者属于高风险群体,因此需要在 ICU 早期阶段重点关注。

利益冲突 所有作者声明不存在利益冲突

作者贡献声明 刘万军:实验操作、数据采集、样本分析、论文撰写等过程执行工作;肖文艳、张金、胡娟娟、黄珊珊、刘瑜、华天凤:数据可视化、文献调研、论文修改等支持辅助工作;杨旻:课题思路酝酿、实验方案设计、研究框架搭建等前期设计工作

参考文献

- [1] Seymour CW, Kennedy JN, Wang S, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis [J]. *JAMA*, 2019, 321 (20): 2003–2017. DOI: 10.1001/jama.2019.5791.
- [2] Sweeney TE, Azad TD, Donato M, et al. Unsupervised analysis of transcriptomics in bacterial sepsis across multiple datasets reveals three robust clusters [J]. *Crit Care Med*, 2018, 46 (6): 915–925. DOI: 10.1097/CCM.0000000000003084.
- [3] Davenport EE, Burnham KL, Radhakrishnan J, et al. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study [J]. *Lancet Respir Med*, 2016, 4 (4): 259–271. DOI: 10.1016/S2213–2600(16)00046–1.
- [4] Kaukonen KM, Bailey M, Suzuki S, et al. Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000–2012 [J]. *JAMA*, 2014, 311 (13): 1308–1316. DOI: 10.1001/jama.2014.2637.
- [5] Esteban A, Frutos-Vivar F, Ferguson ND, et al. Sepsis incidence and outcome: contrasting the intensive care unit with the hospital ward [J]. *Crit Care Med*, 2007, 35 (5): 1284–1289. DOI: 10.1097/01.CCM.0000260960.94300.DE.
- [6] Damiani E, Donati A, Serafini G, et al. Effect of performance improvement programs on compliance with sepsis bundles and mortality: a systematic review and meta-analysis of observational studies [J]. *PLoS One*, 2015, 10 (5): e0125827. DOI: 10.1371/journal.pone.0125827.
- [7] ARDS Definition Task Force. Acute respiratory distress syndrome: the Berlin definition [J]. *JAMA*, 2012, 307 (23): 2526–2533. DOI: 10.1001/jama.2012.5669.
- [8] Zhang P, Liu BY, Zheng WH, et al. Pulmonary microbial composition in sepsis-induced acute respiratory distress syndrome [J]. *Front Mol Biosci*, 2022, 9: 862570. DOI: 10.3389/fmolb.2022.862570.
- [9] Wellman TJ, Winkler T, Costa EL, et al. Effect of local tidal lung strain on inflammation in normal and lipopolysaccharide-exposed sheep [J]. *Crit Care Med*, 2014, 42 (7): e491–500. DOI: 10.1097/CCM.0000000000000346.
- [10] ERICC (Epidemiology of Respiratory Insufficiency in Critical Care) investigators. Clinical outcomes of patients requiring ventilatory support in Brazilian intensive care units: a multicenter, prospective, cohort study [J]. *Crit Care*, 2013, 17 (2): R63. DOI: 10.1186/cc12594.
- [11] Müller-Redetzky HC, Will D, Hellwig K, et al. Mechanical ventilation drives pneumococcal pneumonia into lung injury and sepsis in mice: protection by adrenomedullin [J]. *Crit Care*, 2014, 18 (2): R73. DOI: 10.1186/cc13830.
- [12] Sepsis Definitions Task Force. Developing a new definition and assessing new clinical criteria for septic shock: for the third international consensus definitions for sepsis and septic shock (Sepsis-3) [J]. *JAMA*, 2016, 315 (8): 775–787. DOI: 10.1001/jama.2016.0289.
- [13] Liu WJ, Tao G, Zhang YJ, et al. A simple weaning model based on interpretable machine learning algorithm for patients with sepsis: a research of MIMIC-IV and eICU databases [J]. *Front Med (Lausanne)*, 2022, 8: 814566. DOI: 10.3389/fmed.2021.814566.
- [14] Mirbeyk M, Saghaideh A, Rezaei N. A systematic review of pregnant women with COVID-19 and their neonates [J]. *Arch Gynecol Obstet*, 2021, 304 (1): 5–38. DOI: 10.1007/s00404–021–06049–z.
- [15] Gong FC, Li RR, Zheng XT, et al. OLFM4 regulates lung epithelial cell function in sepsis-associated ARDS/ALI via LDHA-mediated NF-κB signaling [J]. *J Inflamm Res*, 2021, 14: 7035–7051. DOI: 10.2147/JIR.S335915.
- [16] Reilly JP, Wang F, Jones TK, et al. Plasma angiopoietin-2 as a potential causal marker in sepsis-associated ARDS development: evidence from Mendelian randomization and mediation analysis [J]. *Intensive Care Med*, 2018, 44 (11): 1849–1858. DOI: 10.1007/s00134–018–5328–0.
- [17] Bhavani SV, Semler M, Qian ET, et al. Development and validation of novel sepsis subphenotypes using trajectories of vital signs [J]. *Intensive Care Med*, 2022, 48 (11): 1582–1592. DOI: 10.1007/s00134–022–06890–z.
- [18] Liu XW, Jiang YS, Jia XN, et al. Identification of distinct clinical phenotypes of acute respiratory distress syndrome with differential responses to treatment [J]. *Crit Care*, 2021, 25 (1): 320. DOI: 10.1186/s13054–021–03734–y.
- [19] MARS consortium. Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study [J]. *Lancet Respir Med*, 2017, 5 (10): 816–826. DOI: 10.1016/S2213–2600(17)30294–1.
- [20] Ketcham SW, Sedhai YR, Miller HC, et al. Causes and characteristics of death in patients with acute hypoxemic respiratory failure and acute respiratory distress syndrome: a retrospective cohort study [J]. *Crit Care*, 2020, 24 (1): 391. DOI: 10.1186/s13054–020–03108–w.
- [21] 崔秋敏, 孙小鸽, 马宁, 等. 血小板最大聚集率对脓毒性休克及脓毒性休克合并 DIC 的预测价值 [J]. *中华危重病急救医学*, 2023, 35 (3): 238–243. DOI: 10.3760/cma.j.cn121430–20221230–01134.
- [22] 叶溯文, 李飞飞, 杨宝华, 等. 探索脓毒症患者最佳脉搏血氧饱和度范围: 一项基于 MIMIC-IV 数据库的回顾性研究 [J]. *中华危重病急救医学*, 2024, 36 (8): 813–820. DOI: 10.3760/cma.j.cn121430–20231019–00885.
- [23] Miller PE, Van Diepen S, Metkus TS, et al. Association between respiratory failure and clinical outcomes in patients with acute heart failure: analysis of 5 pooled clinical trials [J]. *J Card Fail*, 2021, 27 (5): 602–606. DOI: 10.1016/j.cardfail.2021.01.018.

(收稿日期: 2024–03–29)

(本文编辑: 孙茜 马英)