

基于肺部超声影像组学分析联合机器学习 评估重症患者血管外肺水指数

蒙伟宇 张驰 胡军涛 汤展宏

广西医科大学第一附属医院重症医学科, 南宁 530021

通信作者: 汤展宏, Email: tangzhanhong139@sina.com

【摘要】 目的 探讨与血管外肺水指数 (EVLWI) 相关的肺部超声影像组学特征, 采用基于肺部超声的影像组学方法联合机器学习预测重症患者的 EVLWI 并进行效能验证。方法 采用回顾性病例对照研究方法, 收集 2021 年 11 月至 2022 年 10 月广西医科大学第一附属医院重症医学科收治的重症患者肺部超声视频和脉搏指示连续心排量 (PiCCO) 监测结果, 按照 8 : 2 的比例随机分为训练集与验证集。从肺部超声视频取帧得到对应图像并提取影像组学特征, 以 PiCCO 测得的 EVLWI 为“金标准”, 通过统计分析和 LASSO 算法对训练集影像组学特征进行筛选。采用经过筛选的影像组学特征训练 8 种机器学习模型, 包括随机森林 (RF)、极限梯度提升 (XGBoost)、决策树 (DT)、朴素贝叶斯 (NB)、多层感知器 (MLP)、K-近邻 (KNN)、支持向量机 (SVM) 和 Logistic 回归 (LR); 绘制受试者工作特征曲线 (ROC 曲线), 评估上述模型在验证集中对 EVLWI 的预测效能。结果 最终共 30 例患者 151 组样本 (包括 906 份肺部超声视频和 151 份 PiCCO 监测结果) 纳入分析, 其中训练集 120 组样本, 验证集 31 组样本; 两项数据集的性别、年龄、体质量指数 (BMI)、平均动脉压 (MAP)、中心静脉压 (CVP)、心率 (HR)、心排量指数 (CI)、心功能指数 (CFI)、每搏量指数 (SVI)、全心舒张期末容积指数 (GEDVI)、全身血管阻力指数 (SVRI)、肺血管通透性指数 (PVPI)、EVLWI 等主要基线资料差异均无统计学意义。151 份 PiCCO 监测结果中整体 EVLWI 范围为 3.7 ~ 25.6 mL/kg; 分层分析显示, 两项数据集 EVLWI 均集中于 7 ~ 15 mL/kg 区间, EVLWI 分布差异无统计学意义。通过 LASSO 算法筛选出 2 个影像组学特征, 即灰阶不均匀性 (权重为 -0.006464) 和复杂度 (权重为 -0.167583), 并用于建模; ROC 曲线分析显示, MLP 模型具有较好的预测效能, 其预测验证集 EVLWI 的 ROC 曲线下面积 (AUC) 高于 RF、XGBoost、DT、KNN、LR、SVM、NB 模型 (0.682 比 0.658、0.657、0.614、0.608、0.596、0.557、0.472)。结论 肺部超声灰阶不均匀性和复杂度是与 PiCCO 测得的 EVLWI 相关性最高的影像组学特征; 基于肺部超声灰阶不均匀性和复杂度构建的 MLP 模型可用于半定量预测重症患者 EVLWI。

【关键词】 肺部超声; 影像组学; 血管外肺水指数; 机器学习

基金项目: 国家自然科学基金 (81960342)

DOI: 10.3760/cma.j.cn121430-20230209-00077

Evaluation of extravascular lung water index in critically ill patients based on lung ultrasound radiomics analysis combined with machine learning

Meng Weiyu, Zhang Chi, Hu Juntao, Tang Zhanhong

Department of Critical Care Medicine, the First Affiliated Hospital of Guangxi Medical University, Nanning 530021, Guangxi Zhuang Autonomous Region, China

Corresponding author: Tang Zhanhong, Email: tangzhanhong139@sina.com

【Abstract】 **Objective** To explore lung ultrasound radiomics features which related to extravascular lung water index (EVLWI), and to predict EVLWI in critically ill patients based on lung ultrasound radiomics combined with machine learning and validate its effectiveness. **Methods** A retrospective case-control study was conducted. The lung ultrasound videos and pulse indicated continuous cardiac output (PiCCO) monitoring results of critically ill patients admitted to the department of critical care medicine of the First Affiliated Hospital of Guangxi Medical University from November 2021 to October 2022 were collected, and randomly divided into training set and validation set at 8 : 2. The corresponding images from lung ultrasound videos were obtained to extract radiomics features. The EVLWI measured by PiCCO was regarded as the "gold standard", and the radiomics features of training set was filtered through statistical analysis and LASSO algorithm. Eight machine learning models were trained using filtered radiomics features including random forest (RF), extreme gradient boost (XGBoost), decision tree (DT), Naive Bayes (NB), multi-layer perceptron (MLP), K-nearest neighbor (KNN), support vector machine (SVM), and Logistic regression (LR). Receiver operator characteristic curve (ROC curve) was plotted to evaluate the predictive performance of models on EVLWI in the validation set. **Results** A total of 151 samples from 30 patients were enrolled (including 906 lung ultrasound videos and 151 PiCCO monitoring results), 120 in the training set, and 31 in the validation set. There were no statistically significant differences in main baseline data including gender, age, body mass index (BMI), mean arterial pressure (MAP), central venous pressure (CVP), heart rate (HR), cardiac index (CI), cardiac function index (CFI), stroke volume index (SVI),

global end diastolic volume index (GEDVI), systemic vascular resistance index (SVRI), pulmonary vascular permeability index (PVPI) and EVLWI. The overall EVLWI range in 151 PiCCO monitoring results was 3.7–25.6 mL/kg. Layered analysis showed that both datasets had EVLWI in the 7–15 mL/kg interval, and there was no statistically significant difference in EVLWI distribution. Two radiomics features were selected by using LASSO algorithm, namely grayscale non-uniformity (weight was $-0.006\ 464$) and complexity (weight was $-0.167\ 583$), and they were used for modeling. ROC curve analysis showed that the MLP model had better predictive performance. The area under the ROC curve (AUC) of the prediction validation set EVLWI was higher than that of RF, XGBoost, DT, KNN, LR, SVM, NB models (0.682 vs. 0.658, 0.657, 0.614, 0.608, 0.596, 0.557, 0.472). **Conclusions** The gray level non-uniformity and complexity of lung ultrasound were the most correlated radiomics features with EVLWI monitored by PiCCO. The MLP model based on gray level non-uniformity and complexity of lung ultrasound can be used for semi-quantitative prediction of EVLWI in critically ill patients.

【Key words】 Lung ultrasound; Radiomics; Extravascular lung water index; Machine learning

Fund program: National Natural Science Foundation of China (81960342)

DOI: 10.3760/cma.j.cn121430-20230209-00077

肺水肿是重症监护病房(intensive care unit, ICU)的常见并发症,表现为血管外肺水(extravascular lung water, EVLW)增加。及时监测 EVLW 可指导临床决策,改善重症患者预后^[1-3]。目前主要通过经肺热稀释法测定 EVLW,可以准确监测 EVLW^[4],但该方法存在侵入性强、成本高等缺点,监测的实施取决于患者的血管状况^[5]。

肺部超声不受血管状况的影响,其诊断疾病或评估病情的准确性不劣于 X 线胸片及 CT^[6-7],在诊断急性心源性肺水肿和急性呼吸窘迫综合征(acute respiratory distress syndrome, ARDS)^[8]、预测脓毒症^[9]或 ARDS^[10]预后等方面均具有良好表现。肺水肿严重程度由床旁急诊肺部超声(bedside lung ultrasound emergency, BLUE)方案包含的超声征象(A 线、B 线、火箭征、碎片征等)决定,其中 B 线征在形态和数量上与 EVLW 表现出明显的相关性^[11-13]。“定量肺部超声”已成为评估肺水肿的主要方法,研究人员通过量化 B 线实现肺水肿诊断和分层^[14-17]。

随着超声影像组学的发展,通过分析超声图像的非语义特征提取病变信息,并建立临床或影像组学模型,从而预测疾病严重程度,已成为一种趋势。影像组学特征符合本课题组对定量和客观的要求。为提高预测稳定性,有必要将肺部超声影像组学与机器学习相结合。目前关于肺部超声影像组学在重症医学中应用的研究较少,尤其在 EVLW 预测方面鲜有报道^[18-20]。本研究旨在探讨与 EVLW 相关的肺部超声影像组学特征,并试图构建机器学习模型,直接预测重症患者 EVLW。

1 资料和方法

1.1 研究对象:采用回顾性病例对照研究方法,收集 2021 年 11 月至 2022 年 10 月本院外科 ICU 收治的接受脉搏指示连续心排量(pulse indicated

continuous cardiac output, PiCCO)监测的重症患者的临床资料。排除存在胸部创伤、烧伤和皮肤病, PiCCO 测定基线不稳定,合并或存在可能影响图像分析的肺部疾病[如肺气肿、肺栓塞、慢性阻塞性肺疾病(chronic obstructive pulmonary disease, COPD)、哮喘和肺部肿瘤],因体位限制、其他监护仪器覆盖而仅行局部超声检查,年龄 < 14 岁,以及妊娠期或存在腹部巨大肿瘤的患者。

1.2 伦理学:本研究方案符合医学伦理学标准,获得了广西医科大学第一附属医院医学伦理委员会的批准(审批号:2022-E321-01),数据采集与处理符合相关指南和规范。

1.3 数据收集:收集患者的基本信息,包括性别、年龄、体质量指数(body mass index, BMI);PiCCO 监测数据,包括平均动脉压(mean arterial pressure, MAP)、中心静脉压(central venous pressure, CVP)、心率(heart rate, HR)、心排量指数(cardiac index, CI)、心功能指数(cardiac function index, CFI)、每搏量指数(stroke volume index, SVI)、全心舒张期末容积指数(global end-diastolic volume index, GEDVI)、全身血管阻力指数(systemic vascular resistance index, SVRI)、肺血管通透性指数(pulmonary vascular permeability index, PVPI)、血管外肺水指数(extravascular lung water index, EVLWI)、体温;肺部超声视频。

1.4 数据预处理与特征提取:用 Python 3.10 软件提取超声视频中最具代表性的帧并进行灰度变换,用 Labelme 手动勾画感兴趣区域(region of interest, ROI),将帧图像与 ROI 掩模导入 3DSlicer 5.0.3 软件(Python 3.9.10, numpy 1.22.1, pyradiomics 3.0.1, simpleitk 2.2.0, 归一化灰度范围为 0~255),并提取影像组学特征:①用于描述 ROI 基本尺寸的形态特征:平坦度、延展性、最小轴长、长轴长等;②描述

ROI 频率分布特征的一阶灰度直方图：能量、熵、四分位数间距、平均绝对偏差、偏度等；③ 描述灰度空间分布特征的二阶 / 高阶纹理特征，包括灰度共生矩阵 (gray level co-occurrence matrix, GLCM)、灰度依赖矩阵 (gray level dependence matrix, GLDM)、灰度游程矩阵 (gray level run length matrix, GLRLM)、灰度大小区域矩阵 (gray level size zone matrix, GLSZM) 和邻域灰度差矩阵 (neighbourhood gray-tone difference matrix, NGTDM)。体素重采样 “resampling” 大小、“LoG kernel” 大小及初始宽度 “bin width” 分别设定为 3*3*3、(4,5)、25。每个样本所有特征取 6 个位点平均值。

1.5 特征选择：将所有样本混合后，填补缺失值并进行标准化，按照 8 : 2 的比例随机分为训练集和验证集。训练集用于特征提取、模型训练和内部验证，验证集则用于外部验证。每个数据集均根据患者 EVLWI 水平分为 3 组 (<7 mL/kg、7 ~ 15 mL/kg、>15 mL/kg)^[21]。

采用简单最小二乘拟合训练集，通过简单线性回归得到 R^2 ，以 $R^2 > 0.8$ 及 $P < 0.05$ 为标准，初步筛选与 EVLWI 呈简单线性相关的所有影像组学特征。

采用 8 折交叉验证、100 000 次迭代 LASSO 算法拟合训练集，筛选最佳影像组学特征。训练集随机分为 8 个子集，其中 1 个子集作为测试集，其他 7 个子集用于拟合，整个过程只重复 1 次。

1.6 统计学分析与模型构建：使用含有 sklearn (0.0 base) 及 Scipy (1.8.0) 包的 Python 3.10 软件进行统计学分析。经正态性检验提示所有数据均不符合正态分布，以中位数 (四分位数) [$M(Q_L, Q_U)$] 描述，采用 Mann-Whitney U 检验和 Kruskal-Wallis 检验；分类数据以频数 (百分比) 描述，采用 χ^2 检验。采用随机森林 (random forest, RF)、极限梯度提升 (extreme gradient boost, XGBoost)、决策树 (decision tree, DT)、朴素贝叶斯 (Naive Bayes, NB)、多层感知器 (multi-layer perceptron, MLP)、K-近邻 (K-nearest neighbor, KNN)、支持向量机 (support vector machine, SVM)、Logistic 回归 (Logistic

regression, LR) 8 种常见机器学习算法构建模型。以 PiCCO 测得的 EVLWI 为 “金标准”，根据肺部超声影像学特征绘制机器学习模型预测 EVLWI 的受试者工作特征曲线 (receiver operator characteristic curve, ROC 曲线)，并计算 ROC 曲线下面积 (area under the ROC curve, AUC)，采用 DeLong 检验对模型的 AUC 进行两两比较。AUC、召回率、精确度、准确度和 F1 得分用于评估模型效能，选择 AUC 最大的模型作为优化模型。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 基线特征：共入选 219 例患者，排除未行 PiCCO 监测 176 例、PiCCO 基线不稳定 1 例、皮肤烧伤 1 例、体位限制 2 例、超声未做满 6 个位点 4 例、PiCCO 结果不全 5 例，最终 30 例患者的 151 组样本被纳入分析，包括 906 份肺部超声视频、151 份 PiCCO 监测结果。在病因学方面，入组患者中腹腔感染的比例最高 [30.0% (9/30)]，其次是呼吸道感染 [23.4% (7/30)]，此外还包括其他部位感染 [13.4% (4/30)]、移植后感染 [10.0% (3/30)]、颅内出血 [6.7% (2/30)]、烧伤 [3.3% (1/30)]、外伤 [3.3% (1/30)]、肝衰竭 [3.3% (1/30)]、胰腺炎 [3.3% (1/30)] 及病因不明的腹痛 [3.3% (1/30)]。随机分配后数据集基线特征见表 1，性别、年龄和体质量指数 (body mass index, BMI) 作为加权指标；两项数据集的性别、年龄、BMI、MAP、CVP、HR、CI、CFI、SVI、GEDVI、SVRI、PVPI、EVLWI 差异均无统计学意义 (均 $P > 0.05$)；仅训练集样本的血温显著高于验证集 ($P < 0.05$)。

表 1 不同数据集重症患者基线特征比较 [$M(Q_L, Q_U)$]

指标	训练集 (n=120)	验证集 (n=31)	Z 值	P 值
MAP (mmHg)	82.00 (73.25, 96.75)	90.00 (73.00, 96.00)	-0.454	0.650
CVP (cmH ₂ O)	10.50 (7.00, 13.00)	10.00 (7.00, 13.00)	-0.245	0.807
HR (次/min)	95.00 (85.00, 108.00)	92.00 (81.00, 99.00)	-1.297	0.195
体温 (°C)	36.80 (36.40, 37.60)	36.55 (35.68, 37.05)	-2.083	0.037
CI (mL · s ⁻¹ · m ⁻²)	63.68 (53.01, 76.68)	68.18 (50.01, 79.35)	-0.539	0.590
CFI (L/min)	5.40 (4.50, 6.40)	5.20 (4.00, 6.70)	-0.150	0.881
SVI (mL/m ²)	38.00 (31.00, 52.00)	40.00 (35.00, 51.00)	-0.769	0.442
GEDVI (mL/m ²)	710.50 (634.25, 832.50)	766.00 (667.00, 861.00)	-1.751	0.080
SVRI (kPa · s · L ⁻¹ · m ⁻²)	153.85 (122.30, 205.85)	143.90 (117.70, 198.90)	-0.896	0.370
PVPI	1.70 (1.50, 2.10)	1.80 (1.50, 2.50)	-1.315	0.189
EVLWI (mL/kg)	10.00 (7.00, 12.50)	11.00 (8.20, 15.00)	-1.687	0.092

注：将 30 例重症患者共 151 组样本 [包括 906 份肺部超声视频和 151 份脉搏指示连续心排血量 (PiCCO) 监测结果] 按 8 : 2 的比例随机分为训练集和验证集；MAP 为平均动脉压，CVP 为中心静脉压，HR 为心率，CI 为心排血指数，CFI 为心功能指数，SVI 为每搏量指数，GEDVI 为全心舒张期末容积指数，SVRI 为全身血管阻力指数，PVPI 为肺血管通透性指数，EVLWI 为血管外肺水指数；1 mmHg ≈ 0.133 kPa，1 cmH₂O ≈ 0.098 kPa；对于同一患者不同时间点获得的样本，考虑到重症患者血管外肺水 (EVLW) 和超声图像的频繁变化，将每个样本视为独立样本，并在随后的统计分析中通过随机分组和权重分配减少偏差

151份PiCCO监测结果中,整体EVLWI范围为3.7~25.6 mL/kg。对EVLWI进行分层分析显示,两项数据集EVLWI均集中于7~15 mL/kg区间,分布差异无统计学意义($\chi^2=3.200, P=0.200$;表2)。

数据集	例数(例)	EVLWI [例(%)]		
		<7 mL/kg	7~15 mL/kg	>15 mL/kg
训练集	120	36(30.0)	68(56.7)	16(13.3)
验证集	31	5(16.1)	19(61.3)	7(22.6)

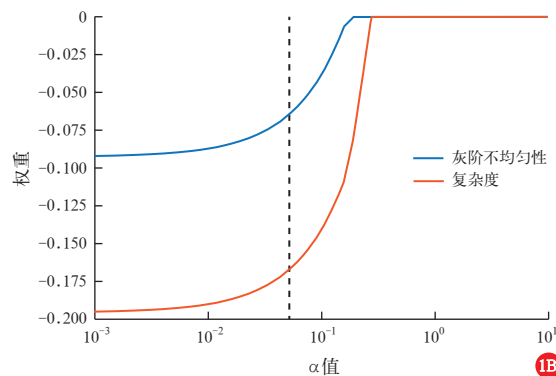
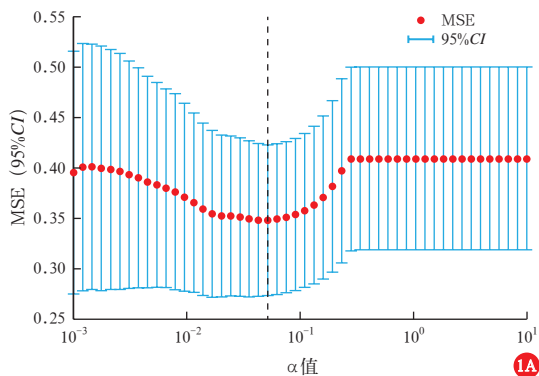
注:将30例重症患者共151组样本[包括906份肺部超声视频和151份脉搏指示连续心排血量(PiCCO)监测结果]按8:2的比例随机分为训练集和验证集;EVLWI为血管外肺水指数;3组间比较, $\chi^2=3.200, P=0.200$

2.2 提取的影像组学特征:通过简单线性回归拟合训练集影像组学特征,获得的所有 R^2 均小于0.6,提示未发现与EVLWI呈显著线性相关的影像组学特性。

两项数据集均提取出112个影像组学特征。经Kruskal-Wallis检验和单因素方差分析,从训练集中

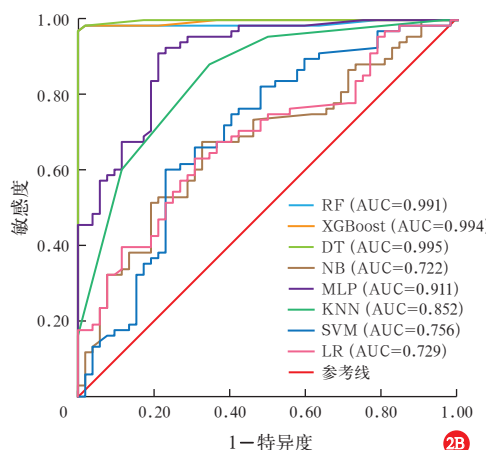
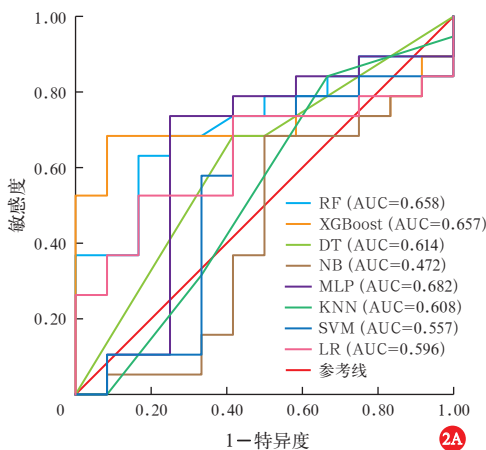
共筛选出50个具有显著差异的特征;基于最小均方误差(mean square error, MSE),通过LASSO算法($\alpha=0.051\ 797\ 467\ 923\ 121$)进一步筛选出2个显著特征,即GLRLM灰阶不均匀性(权重为-0.006464)和NGTDM复杂度(权重为-0.167583;图1),且二者存在强相关(Spearman秩相关系数0.85, $P<0.05$)。

2.3 模型的构建与验证:除MLP参数“max_iter”被设置为5000以确保模型的完全收敛外,其他模型均使用默认参数构建。验证集内ROC曲线分析显示(图2A),MLP模型在预测EVLWI方面的表现最佳(AUC=0.682),而NB模型的预测效能最差(AUC=0.472)。DeLong检验未发现任意两个模型间AUC存在显著差异。预测效能指标分析显示(表3),MLP模型和KNN模型的准确度及召回度均最高,XGBoost模型的精确度和F1得分最高。训练集内ROC曲线分析显示(图2B),DT模型的预测效能最佳(AUC=0.995),NB模型的预测效能最差(AUC=0.722)。



注:黑色虚线分别表示最小均方误差(MSE)时的LASSO算法参数 α 值(A)及该 α 值对应的两种特征的权重(B);95%CI为95%可信区间

图1 重症患者肺部超声影像组学特征的LASSO回归交叉验证



注:将30例重症患者共151组样本[包括906份肺部超声视频和151份脉搏指示连续心排血量(PiCCO)监测结果]按8:2的比例随机分为训练集和验证集;RF为随机森林,XGBoost为极限梯度提升,DT为决策树,NB为朴素贝叶斯,MLP为多层感知器,KNN为K-近邻,SVM为支持向量机,LR为Logistic回归,EVLWI为血管外肺水指数,ROC曲线为受试者工作特征曲线,AUC为ROC曲线下面积

图2 验证集(A)与训练集(B)内不同机器学习模型预测重症患者EVLWI的ROC曲线

表3 验证集内不同机器学习模型预测重症患者 EVLWI 的效能分析

机器学习模型	准确度	精确度	召回度	F1 得分	AUC
RF	0.55	0.48	0.55	0.51	0.658
XGBoost	0.55	0.54	0.55	0.53	0.657
DT	0.52	0.52	0.51	0.51	0.614
NB	0.48	0.43	0.48	0.45	0.472
MLP	0.58	0.45	0.58	0.50	0.682
KNN	0.58	0.45	0.58	0.50	0.608
SVM	0.55	0.43	0.55	0.48	0.557
LR	0.48	0.38	0.48	0.43	0.596

注:将30例重症患者共151组样本[包括906份肺部超声视频和151份脉搏指示连续心排量(PiCCO)监测结果]按8:2的比例随机分为训练集和验证集;RF为随机森林,XGBoost为极限梯度提升,DT为决策树,NB为朴素贝叶斯,MLP为多层感知器,KNN为K-近邻,SVM为支持向量机,LR为Logistic回归,EVLWI为血管外肺水指数,AUC为受试者工作特征曲线下面积

3 讨论

临床上使用肺部超声B线征进行肺水肿定量评估有如下缺点:①不同操作者在B线数量上存在观测差异,这种主观差异可能影响B线量化;②重症患者的B线图像经常发生变化,如B线融合和模糊,不利于B线量化;③少数重症患者存在B线特点(如数量和有无)与EVLW不匹配的情况。此外,很少有人提议将量化目标直接放在EVLW上。本研究中将肺部超声影像组学纳入重症患者EVLWI半定量预测,以B线以外的超声特征作为突破口,寻找其他客观特征量化EVLWI,经验证,肺部超声影像组学可以半定量预测EVLWI。与床旁超声相比,PiCCO具有耗材昂贵、创伤大、可能因管道堵塞或血管脆性增加而被迫停止使用等缺点。因此,使用肺部超声影像组学作为PiCCO的替代方法评估重症患者EVLWI具有潜在的临床价值。

本研究中发现肺部超声本身也存在一定缺点,如肺区视野不足,通常见于皮下组织厚度过大(如肥胖或强壮)及存在胸腔积液的患者。肺部区域视野减少导致接收到的信息量减少,但可以通过调整监测深度(约6cm)或强化学习来解决。

影像组学,即从超声图像中提取的基于形状、像素强度和纹理的特征,是一个多阶段的过程^[22]。本研究结合以往的统计差异分析,在重症患者肺部超声视频中筛选出2个差异明显且有代表性的影像特征,即灰阶不均匀性和复杂度,且二者可能与EVLWI有关。当EVLWI<7 mL/kg时,肺部超声图像灰阶不均匀性和复杂度增加;但EVLWI为7~15 mL/kg或>15 mL/kg时,二者变化并不显著。表明随着超声

肺部纹理紊乱程度降低,测量EVLWI得到异常值的可能性逐渐增大。其他研究也证实,除超声信号外,肺部超声图像特征也可以用于预测EVLW。Corradi等^[15]使用计算机辅助定量超声测量EVLW,表明肺部回声强度与高水平EVLWI之间存在强相关性。这些基于像素点阵的微特征具有区分EVLW差异的能力,但无法用肉眼观测。尽管EVLWI增加与超声肺部纹理变化关系的机制尚未明确,但仍认为肺区存在一种隐藏于A线与回声阴影后的无法直接识别的纹理,随着该纹理紊乱程度降低,肺部超声图像随之表现出“A线—A线紊乱、模糊—A线消失、B线出现”这一过程,EVLWI处于异常值范围内的概率也随之增加。

本研究结果显示,在所使用的机器学习模型中,XGBoost、KNN、RF和MLP对EVLWI的预测效能高于其他模型,但它们之间无显著差异。有研究者使用XGBoost、RF、LR等7种机器学习算法预测脓毒症继发性肾损伤的风险,也发现这些算法具有相似的预测效能(AUC为0.664~0.817)^[23]。Zhu等^[24]使用7种分类算法预测重症患者是否需要机械通气治疗,同样发现这些算法的预测效能差异较小(AUC为0.743~0.821)。本研究中所使用的机器学习模型虽然能有效划分EVLWI的区间,但其AUC与基于B线的人工分辨效果相当^[15],而模型在训练集中具有良好的预测效能,考虑模型存在过拟合。经过减少特征的数量(即减少到1个特征)、用常见超参数优化算法(如网格搜索和贝叶斯优化)、重新调整两个数据集(使数据分布更加相似)和交叉验证后,过拟合无明显改善。本研究显示,两项数据集间血温差异存在统计学意义,但不具有临床意义,无法用于解释过拟合。因此,考虑不存在过拟合,两项数据集间AUC差异可能源于样本量不足或数据质量较低。本研究中肺部超声视频均以AP4和AVI格式存储,视频帧的分辨率低可能导致提取特征的分布变得分散,造成与EVLWI低相关并降低模型效能。为改进这一问题,可考虑在进一步研究中增加样本量或改用DICOM格式存储视频。

本研究的优势:①在超声探头选择中,与BLUE方案^[25]推荐的凸阵探头相比,本研究使用了线性探头作为主要检查部件,后者在4~6cm的深度下对胸膜和外周肺实质具有更高的分辨率,形成更清晰的图像,从而获得更准确的图像信息。Baldi等^[26]研究表明,仅探测浅胸膜下区域即可描述EVLW变化,

其准确度与 CT 相近。② 采用各位点特征取平均值以应对因重力导致的全肺各区域含水量不均等状况。③ 为提高模型的稳定性,使用不同的超声仪器和通气策略提高训练集与验证集的异质性。

本研究的局限性:① 本研究为单中心研究,仍在内部验证中,所选样本均来自本院患者,样本量有限。② 本研究中的模型仅适用于重症患者,为确保稳定性,有必要根据研究所采用的方法收集、处理和分析样本。本研究中的模型在以下情况无法保证其实际应用的准确性:存在肺气肿、皮下气肿、肺部肿瘤等其他影响超声成像的合并症;另一种超声方案[如肺部超声介导的限制性液体管理(fluid administration limited by lung sonography, FALLs)方案];另一种 ROI 勾画或特征提取方法(如 LabelImg、ITK-SNAP 等软件算法)。上述两项局限性均需扩展数据集并开展多中心研究进一步补充。③ 影像组学特征与 B 线组成之间关系的可解释性有待进一步探究。

此外,本研究可向其他方向扩展。通过应用深度学习,自动识别和绘制肺野实现一系列图像处理(如特征提取、特征分析),或直接使用基于 U-net 或 Res-net 等现有神经网络的迁移学习,开发和优化基于肺部超声的 EVLWI 测定方法。

4 结论

本研究表明,重症患者肺部超声影像组学特征可作为半定量评估 EVLWI 的标志,可能为重症患者的肺含水量预测提供一种新的非侵入性方法,作为经肺热稀释技术的替代方案。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] PiCCO Pulmonary Edema Study Group. Relationship between extravascular lung water and severity categories of acute respiratory distress syndrome by the Berlin definition [J]. *Crit Care*, 2013, 17 (4): R132. DOI: 10.1186/cc12811.
- [2] Jozwiak M, Silva S, Persichini R, et al. Extravascular lung water is an independent prognostic factor in patients with acute respiratory distress syndrome [J]. *Crit Care Med*, 2013, 41 (2): 472-480. DOI: 10.1097/CCM.0b013e31826ab377.
- [3] Vieillard-Baron A, Matthay M, Teboul JL, et al. Experts' opinion on management of hemodynamics in ARDS patients: focus on the effects of mechanical ventilation [J]. *Intensive Care Med*, 2016, 42 (5): 739-749. DOI: 10.1007/s00134-016-4326-3.
- [4] Tagami T, Kushimoto S, Yamamoto Y, et al. Validation of extravascular lung water measurement by single transpulmonary thermodilution: human autopsy study [J]. *Crit Care*, 2010, 14 (5): R162. DOI: 10.1186/cc9250.
- [5] Loosen G, Conrad AM, Hagman M, et al. Transpulmonary thermodilution in patients treated with veno-venous extracorporeal membrane oxygenation [J]. *Ann Intensive Care*, 2021, 11 (1): 101. DOI: 10.1186/s13613-021-00890-w.
- [6] 朱海云,段军,孙艳文,等.急诊床旁肺部超声在诊断 ICU 术后低氧血症中的价值 [J]. *中国中西医结合急救杂志*, 2019, 26 (3):

- 293-295. DOI: 10.3969/j.issn.1008-9691.2019.03.009.
- [7] 吕杰,吕娜,郭晓夏,等.肺部超声指导危重症患者肺部病变诊疗的效果评价 [J]. *中国中西医结合急救杂志*, 2021, 28 (3): 324-328. DOI: 10.3969/j.issn.1008-9691.2021.03.016.
- [8] Lichtenstein DA. Lung ultrasound in the critically ill [J]. *Ann Intensive Care*, 2014, 4 (1): 1. DOI: 10.1186/2110-5820-4-1.
- [9] 张晓勤,李春玲,潘灵爱,等.经胸肺部超声彗尾征及血管外肺水指数在脓毒症患者预后评估中的价值 [J]. *中华危重病急救医学*, 2021, 33 (8): 985-989. DOI: 10.3760/cma.j.cn121430-20200414-00294.
- [10] 张磊,俞万钧,周成杰,等.肺部超声评估急性呼吸窘迫综合征患者血管外肺水及预后的临床应用价值 [J]. *中华危重病急救医学*, 2020, 32 (5): 585-589. DOI: 10.3760/cma.j.cn121430-20200102-00088.
- [11] Platz E, Merz A, Silverman M, et al. Association between lung ultrasound findings and invasive exercise haemodynamics in patients with undifferentiated dyspnoea [J]. *ESC Heart Fail*, 2019, 6 (1): 202-207. DOI: 10.1002/ehf2.12381.
- [12] Lung Ultrasound Consortium. Ultrasound versus computed tomography assessment of focal lung aeration in invasively ventilated ICU patients [J]. *Ultrasound Med Biol*, 2021, 47 (9): 2589-2597. DOI: 10.1016/j.ultrasmedbio.2021.05.019.
- [13] Cox EGM, Wiersema R, Wong A, et al. Six versus eight and twenty-eight scan sites for B-line assessment: differences in examination time and findings [J]. *Intensive Care Med*, 2020, 46 (5): 1063-1064. DOI: 10.1007/s00134-020-06004-7.
- [14] Brusasco C, Santori G, Bruzzo E, et al. Quantitative lung ultrasonography: a putative new algorithm for automatic detection and quantification of B-lines [J]. *Crit Care*, 2019, 23 (1): 288. DOI: 10.1186/s13054-019-2569-4.
- [15] Corradi F, Brusasco C, Vezzani A, et al. Computer-aided quantitative ultrasonography for detection of pulmonary edema in mechanically ventilated cardiac surgery patients [J]. *Chest*, 2016, 150 (3): 640-651. DOI: 10.1016/j.chest.2016.04.013.
- [16] Picano E, Pellicka PA. Ultrasound of extravascular lung water: a new standard for pulmonary congestion [J]. *Eur Heart J*, 2016, 37 (27): 2097-2104. DOI: 10.1093/eurheartj/ehw164.
- [17] Tierney DM, Boland LL, Overgaard JD, et al. Pulmonary ultrasound scoring system for intubated critically ill patients and its association with clinical metrics and mortality: a prospective cohort study [J]. *J Clin Ultrasound*, 2018, 46 (1): 14-22. DOI: 10.1002/jcu.22526.
- [18] Jiao J, Du YR, Li XK, et al. Prenatal prediction of neonatal respiratory morbidity: a radiomics method based on imbalanced few-shot fetal lung ultrasound images [J]. *BMC Med Imaging*, 2022, 22 (1): 2. DOI: 10.1186/s12880-021-00731-z.
- [19] Du Y, Fang Z, Jiao J, et al. Application of ultrasound-based radiomics technology in fetal-lung-texture analysis in pregnancies complicated by gestational diabetes and/or pre-eclampsia [J]. *Ultrasound Obstet Gynecol*, 2021, 57 (5): 804-812. DOI: 10.1002/uog.22037.
- [20] Du YR, Jiao J, Ji C, et al. Ultrasound-based radiomics technology in fetal lung texture analysis prediction of neonatal respiratory morbidity [J]. *Sci Rep*, 2022, 12 (1): 12747. DOI: 10.1038/s41598-022-17129-8.
- [21] Tagami T, Ong MEH. Extravascular lung water measurements in acute respiratory distress syndrome: why, how, and when? [J]. *Curr Opin Crit Care*, 2018, 24 (3): 209-215. DOI: 10.1097/MCC.000000000000503.
- [22] Rogers W, Thulasi Seetha S, Refaee TAG, et al. Radiomics: from qualitative to quantitative imaging [J]. *Br J Radiol*, 2020, 93 (1108): 20190948. DOI: 10.1259/bjr.20190948.
- [23] Yue SR, Li SS, Huang XY, et al. Machine learning for the prediction of acute kidney injury in patients with sepsis [J]. *J Transl Med*, 2022, 20 (1): 215. DOI: 10.1186/s12967-022-03364-0.
- [24] Zhu YB, Zhang J, Wang GW, et al. Machine learning prediction models for mechanically ventilated patients: analyses of the MIMIC-III database [J]. *Front Med (Lausanne)*, 2021, 8: 662340. DOI: 10.3389/fmed.2021.662340.
- [25] Lichtenstein DA. Lung ultrasound in the critically ill: the BLUE protocol [M]. Switzerland: Springer, 2016: 31-41.
- [26] Baldi G, Gargani L, Abramo A, et al. Lung water assessment by lung ultrasonography in intensive care: a pilot study [J]. *Intensive Care Med*, 2013, 39 (1): 74-84. DOI: 10.1007/s00134-012-2694-x.

(收稿日期: 2023-02-09)
(责任编辑: 孙茜 张耘菲)