基于可解释性机器学习构建脑卒中患者卒中相关性肺炎风险预测模型*

李春标[1] 王 婷[2△] 袁 娟[1] 袁琳丽[1] 刘 艺[1] 李大进[1] 尤 敏[2]

【提 要】 **目的** 基于可解释性机器学习方法构建卒中相关性肺炎（stroke-associated pneumonia，SAP）风险预测模型，为临床预防SAP提供参考。**方法** 回顾性收集2023年12月1日至2024年8月31日就诊于安徽中医药大学第二附属医院的763例脑卒中患者的资料，将其按照6∶4的比例随机分为模型训练集（*n*=457）和模型验证集（*n*=306）。首先对纳入的数据进行整理，随后使用LASSO回归筛选纳入的变量。使用决策树（decision tree，DT）、逻辑回归（logistic regression，LR）、极限梯度提升算法（extreme gradient boosting，XGBoost）、支持向量机（support vector machine，SVM）、轻量级梯度提升机算法（light gradient boosting machine，lightGBM）5种机器学习方法构建模型，通过网格搜索5折交叉验证进行超参数优化，优化模型的性能指标。模型的预测性能通过AUC、校准曲线和决策曲线分析（decision curve analysis，DCA）等指标进行评估，最后使用Shapley加性解释（Shapley additive explanation，SHAP）来解释最优模型的预测结果。**结果** 本研究中SAP的发生率为31.72%（242/763），通过LASSO回归筛选出6个变量，包括使用鼻胃管治疗、年龄、日常生活活动能力量表（activity of daily living scale，ADL）评分、白蛋白、超敏C反应蛋白和血红蛋白。在验证集中表现最佳的模型为XGBoost模型，AUC为0.926，精确度为0.914，*F*1分数为0.889，其校准曲线和DCA均展示出良好的性能，SHAP算法显示重要性排序排名第一的为ADL评分。**结论** 使用XGBoost构建的模型具有良好的预测性能和临床适用性，有望为临床决策提供支持，改善患者的预后。

【关键词】 脑卒中 卒中相关性肺炎 机器学习 预测模型 Shapley加性解释
【中图分类号】 R743.3；R563.1 【文献标志码】 A **DOI** 10.11783/j.issn.1002-3674.2026.04.002

Construct a Risk Prediction Model for Stroke-associated Pneumonia in Stroke Patients based on Interpretable Machine Learning

Li Chunbiao，Wang Ting，Yuan Juan，et al（*School of Nursing*，*Anhui University of Chinese Medicine*，*Hefei* 230012）

【**Abstract**】 **Objective** The risk prediction model of stroke-associated pneumonia（SAP）was constructed based on interpretable machine learning method to provide reference for clinical prevention of SAP. **Methods** Data of 763 stroke patients admitted to the Second Affiliated Hospital of Anhui University of Chinese Medicine from December 1，2023 to August 31，2024 were retrospectively collected and randomly divided into model training set（*n*=457）and model validation set（*n*=306）according to a ratio of 6∶4. Firstly，the included data were sorted out，and then the included variables were screened using LASSO regression. Five machine learning methods，including decision tree（DT），logistic regression（LR），extreme gradient boosting（XGBoost），support vector machine（SVM）and lightweight gradient boosting machine（lightGBM），were used to construct the model. Hyperparameter optimization was carried out through grid search and 5-fold cross-validation to optimize the performance index of the model. The predictive performance of the model was evaluated by AUC，calibration curve and decision curve analysis（DCA），and finally Shapley additive explanation（SHAP）was used to explain the prediction results of the optimal model. **Results** In this study，the incidence of SAP was 31.72%（242/763）. Six variables were screened by LASSO regression，including nasogastric tube therapy，age，activity of daily living scale（ADL）score，albumin，hypersensitive C-reactive protein，and haemoglobin. The model with the best performance in the verification set is XGBoost model，with AUC of 0.926，accuracy of 0.914 and *F*1 score of 0.889. Its calibration curve and DCA both show good performance. The SHAP algorithm shows that the model with the highest ranking in importance ranking is the ADL score. **Conclusion** The model built using XGBoost has good predictive performance and clinical applicability，which is expected to support clinical decision making and improve patient outcomes.

【**Key words**】 Stroke；Stroke-associated pneumonia；Machine learning；Prediction model；Shapley additive explanation

脑卒中是一种严重的神经系统疾病，对患者的生活质量和生存率产生重大影响。根据全球疾病负担调查数据显示，2019年脑卒中导致了约655万例死亡案例，位列所有死亡原因中的第二位，仅次于心血管疾病[1]。卒中后，患者常常面临多种并发症，其中卒中相关性肺炎（stroke-associated pneumonia，SAP）是最常见的并发症之一[2]。据调查显示，SAP的发生率为7%~38%[3-6]，SAP不仅延长了患者的住院时间、增加其经济负担，甚至可严重影响患者的病死率[4-5]。目前临床针对SAP的治疗方案主要为抗感染治疗[7]，但有研究表明，预防性应用抗生素并不能有效降低SAP的发生风险[8]。因此，临床工作者需要及时识别SAP的高危人群并给予相应预防措施，这对于改善脑卒中患者的预后具有重要意义。

* 基金项目：安徽省高校科研项目（2023AH050784）；安徽省科研编制计划项目（2022AH050463）；安徽省卫生健康科研项目（AHWJ2024Aa30205）
1.安徽中医药大学护理学院（230012）
2.安徽中医药大学第二附属医院护理部
△通信作者：王婷，E-mail：994340400@qq.com

风险预测模型的建立可以帮助临床工作者识别疾病的高风险人群，从而采取早期干预措施，降低疾病的发生率[9]。近年来，国内外学者已建立了多个预测SAP发生风险的模型，并以评分系统、打分卡等形式展示模型，如Kwon评分[10]、A2DS2评分[11]、急性缺血性卒中相关肺炎评分[12]等，为SAP的防治工作提供了较为可靠的评估工具。然而，即便是经过良好验证的模型，由于疾病危险因素、治疗措施及治疗背景等随时间变化，仍有可能导致其性能的下降，因此，模型需不断动态更新[13]。此外，使用可解释性机器学习构建SAP风险预测模型的研究较为少见，基于此，本研究考虑将新的预测因素与已知的预测因素相结合，使用机器学习方法构建模型，并采用SHAP算法对模型进行解释，以期提高SAP风险预测的准确性和可解释性。

资料与方法

1. 研究设计和研究对象

本研究为单中心、回顾性研究。回顾性收集2023年12月1日至2024年8月31日就诊于安徽中医药大学第二附属医院的763例脑卒中患者临床资料。纳入标准：①符合脑卒中诊断的患者；②年龄≥18岁；③脑卒中后7天内未进行机械通气。排除标准：①入院后24h内出院、转院或死亡；②入院前并发肺部感染；③收集数据的缺失比例超过30%。

2. 候选预测因子的确定

通过文献回顾和专家咨询确立本次研究的候选预测因子，共计27个，包括：①患者一般资料及临床治疗资料：性别、年龄、吸烟、饮酒、入院日常生活活动能力量表（activity of daily living scale，ADL）[14]评分、脑卒中类型、脑卒中部位、吞咽障碍（使用洼田饮水试验[15]评估，1级表示吞咽功能正常，2~5级表示存在吞咽障碍）、意识障碍（使用格拉斯哥昏迷评分[16]评估，低于13分为存在意识障碍）、高血压、糖尿病、脑卒中史、肺部基础疾病史、留置鼻胃管治疗、抑酸药、留置尿管；②实验室检查指标（于患者入院后24h内收集）：白蛋白（albumin，Alb）、甘油三酯（triglyceride，TG）、超敏C反应蛋白（hypersensitive C-reactive protein，hs-CRP）、白细胞（white blood cell，WBC）、中性粒细胞与淋巴细胞比值（neutrophil-to-lymphocyte ratio，NLR）、单核细胞与淋巴细胞比值（monocyte-to-lymphocyte ratio，MLR）、血红蛋白（haemoglobin，Hb）、血小板（platelet，PLT）、D-二聚体（D-Dimer）、高密度脂蛋白（high density lipoprotein，HDL）和低密度脂蛋白（low density lipoprotein，LDL）。

3. 结局变量的定义和诊断

参照英国多学科专家组成的SAP共识小组发表的共识[2]，将结局指标SAP定义为“非机械通气的卒中患者在发病7d内新出现的肺炎”；参照改良美国疾病控制预防中心[17]标准作为SAP的诊断标准。

4. 样本量估算

本研究采用Riley等[18]于2020年推出的一种专用于开发风险预测模型的样本量估算方法，该方法考虑到多分类、交互作用、非线性关系等方面的影响，最大限度地降低模型过拟合的可能性，同时对关键参数进行精确估计，以确定构建预测模型所需的合适样本量。具体运算可通过R软件中的pmsampsize包实现。结合文献资料可知，现有预测模型的C统计量平均为0.827[19]，SAP发生率约为7%~38%[3-6]，本研究预计纳入27个预测因子，通过计算得出所需样本量为701~1272例，预测结局事件发生数为179~253例。

5. 数据的收集和预处理

通过查阅电子病历中的入院记录、出院记录、护理记录、实验室检查等结果以确定数据来源，主要采集患者入院第一次的数据。若某变量缺失值的数量少于10%，采用随机森林法对数据进行插补，以保持数据的完整性和预测模型的准确性。为防止信息损失，所有的连续性变量均采用原值输入。数据处理完毕后，将数据按照6∶4的比例划分为训练集和验证集，训练集数据用来训练模型参数，验证集数据用来评估模型的预测性能。采用斯皮尔曼等级相关检验以评估变量之间的相关性，并以热图展示结果。热图的颜色深浅表示相关性强弱，蓝色代表变量间为正相关，橙色代表变量间为负相关，“*”表示相关性检验的P值<0.05，“**”表示P值<0.01，“***”表示P值<0.001。

6. 模型的构建和评价

以脑卒中患者是否发生SAP为因变量，备选预测因子为自变量，进行LASSO回归筛选变量，并进行10折交叉验证，以距离最小均方误差一个标准误差所对应的λ（lambda 1se）最优。随后基于tidymodels框架构建模型，本研究选取了以下5种机器学习算法构建模型：决策树（decision tree，DT）、逻辑回归（logistic regreesion，LR）、极限梯度提升算法（extreme gradient boosting，XGBoost）、支持向量机（support vector machine，SVM）和轻量级梯度提升机算法（light gradient boosting machine，lightGBM）。在训练集上进行了5折交叉验证作为重采样方法，并通过网格搜索对超参数进行了优化。在验证集上采用AUC来评估模型的区分度，ROC曲线的凸起程度越高，越偏向于左上角，提示该模型的预测性能越好；AUC取值介于0.5与1之间，越接近1，说明模型的区分度越高。绘制校准曲线以评估模型的拟合度，校准曲线越接近对角线，提示模型的预测概率与实际发生的概率越匹配。采用决策曲线分析（decision curve analysis，DCA）评估

模型的临床实用性，曲线越偏向右上角，提示模型的净获益值越大；多条曲线在同一阈值概率下，曲线越高表示模型的净获益值越大，模型的临床实用性越高。最后计算模型的准确度、灵敏度、特异度、*F*1分数、精确度、召回率等指标以综合评估模型的性能。

7. 模型的可解释性

Shapley加性解释（Shapley additive explanation，SHAP）是一种流行的模型解释方法，它基于博弈论中的Shapley值，用于解释任何模型的预测。SHAP的核心思想是将每个特征对预测的贡献分解开来，使得模型预测可以表示为特征贡献的加权和，从而提高模型的透明度和信任度，支持更合理的决策制定。本研究对最优模型使用SHAP算法，以获得特征变量的重要性排序，并以瀑布图、条形图、蜂群图的形式直观展示特征变量对预测模型的贡献度。

8. 统计学方法

数据的统计分析基于R 4.3.0软件实现。服从正态分布的计量资料以均数±标准差（$\bar{x} \pm s$）表示，组间比较采用独立样本*t*检验，不服从正态分布的计量资料以中位数和四分位间距[$M(P_{25}, P_{75})$]表示，组间比较采用Mann Whitney *U*检验；计数资料以例数和构成比[*n*（%）]表示，无序分类资料组间比较采用χ^2检验或Fisher's精确概率法。所有统计学检验均采用双侧检验，以*P*<0.05为差异有统计学意义。

结　果

1. 数据缺失情况

本研究共有7个变量存在缺失值，缺失的变量数目及占总样本量的比例分别为：Alb（22/2.88%）、TG（26/3.41%）、LDL（33/4.33%）、HDL（33/4.33%）、脑卒中部位（37/4.85%）、hs-CRP（53/6.95%）、D-Dimer（65/8.52%），缺失值均通过随机森林法进行插补。

2. 患者特征

根据患者是否发生SAP，将763例患者分为SAP组（*n*=242）和非SAP组（*n*=521）。纳入的763例患者中，男性患者为504例（66.06%），女性患者为259例（33.94%），年龄平均为67.35±13.00岁，SAP的发生率为31.72%。患者一般资料及临床治疗资料结果显示，两组的年龄、ADL评分、脑卒中类型、吞咽困难、意识障碍、糖尿病、肺部基础疾病史、留置鼻胃管情况、使用尿管情况比较差异具有统计学意义（*P*<0.05），见表1；两组实验室检查结果显示，两组的Alb、hs-CRP、WBC、NLR、MLR、Hb、D-Dimer、HDL比较差异具有统计学意义（*P*<0.05），见表2。

3. 变量相关性分析及变量筛选

绘制变量间的相关性热图，结果显示，两两变量间的相关系数多在0~0.6之间，提示变量间的共线性处于中度水平。年龄与Alb呈显著负相关，WBC与NLR、NLR与MLR呈显著正相关，相关系数在0.6以上，其余变量间的相关系数均小于0.6（图1）。使用LASSO回归对纳入的变量进行筛选，结果显示，当log（λ）= lambda 1se时，模型变量减少到6个，通过交叉验证确定最佳的λ值为0.39。最终纳入的变量为鼻胃管治疗、年龄、ADL评分、Alb、hs-CRP、Hb（图2A和图2B）。

4. 模型的构建和评价

基于上述筛选的6个变量构建模型。首先在训练集中进行交叉验证和超参数调优训练模型，随后在验证集中对模型的性能进行验证和评价。模型在验证集的ROC曲线显示，各模型的AUC均大于0.8，提示模型均具有良好的区分度，其中XGBoost模型的AUC值（0.926）最高，DT模型的AUC值（0.834）最低，见图3。校准曲线显示，除DT、SVM模型外，其余模型的平均预测概率和实际发生概率基本保持一致，见图4。DCA显示，相较于其他模型而言，在同一阈值概率范围内，XGBoost模型的净获益值最大，见图5。使用Delong检验对5种机器学习模型在验证集的AUC值进行比较，结果显示，LR、XGBoost、SVM和lightGBM模型的区分度优于DT模型（*P*<0.05），见表3；各模型性能指标比较结果显示，XGBoost模型的准确率、精确度、*F*1分数等优于其余模型，见表4。综合以上结果，XGBoost模型的预测性能和临床实用性更优。

5. 基于SHAP对XGBoost模型进行解释

基于SHAP绘制特征（即变量）的重要性排序图，结果显示，ADL评分的平均SHAP值最高，对SAP预测贡献度最大，见图6。绘制蜂群图以分析特征对模型预测产生的正负影响。图中*X*轴代表SHAP值，*Y*轴代表不同的特征，蜂群图由大量的点组成，每个点代表一个样本，图中颜色不同反映了特征值的高低，橙色表示特征高值，紫红色表示特征低值。以ADL评分为例，该特征显示了大量的正向和负向影响，表明其贡献度较大，且低ADL水平对结局预测是负向影响，高ADL水平对结局预测是正向影响，见图7。通过瀑布图展示某个样本中不同特征对SAP发生风险预测结果的影响。E[$f(x)$]代表SHAP基准值，即整个验证集的平均预测值，$f(x)$代表该样本最终的SHAP值，图中的橙色箭头代表该特征对预测有正向影响，紫红色箭头则相反，图中的值表示每个特征对模型输出结果的贡献程度。以数据中第一个样本为例，见图8。

讨　论

本研究基于鼻胃管治疗、年龄、ADL、Alb、hs-CRP、Hb 6个变量，使用5种机器学习方法构建了脑卒中患者发生SAP的风险预测模型，其中XGBoost模型

表1　SAP组和非SAP组患者一般资料及临床治疗资料比较

变量	全部患者(n=763)	SAP(n=242)	非SAP(n=521)	$\chi^2/Z/t$	P
性别[n(%)]				0.093	0.761
男性	504(66.06)	158(65.29)	346(66.41)		
女性	259(33.94)	84(34.71)	175(33.59)		
年龄(岁,$\bar{x}\pm s$)	67.35±13.00	72.08±11.84	65.15±12.94	7.301	<0.001
ADL评分[分,$M(P_{25},P_{75})$]	55.00(29.00,79.50)	20.00(10.00,40.00)	65.00(50.00,85.00)	−17.199	<0.001
脑卒中类型[n(%)]				5.594	<0.001
缺血性脑卒中	595(77.98)	159(65.70)	436(83.69)		
出血性脑卒中	124(16.25)	60(24.79)	64(12.28)		
混合性脑卒中	44(5.77)	23(9.50)	21(4.03)		
脑卒中部位[n(%)]				0.475	0.635
大脑左侧	223(29.23)	74(30.58)	149(28.60)		
大脑右侧	184(24.12)	48(19.83)	136(26.10)		
多发性	356(46.66)	120(49.59)	236(45.30)		
吞咽困难[n(%)]				103.251	<0.001
否	525(68.81)	106(43.80)	419(80.42)		
是	238(31.19)	136(56.20)	102(19.58)		
意识障碍[n(%)]				95.891	<0.001
否	708(92.79)	192(79.34)	516(99.04)		
是	55(7.21)	50(20.66)	5(0.96)		
高血压[n(%)]				0.003	0.955
否	141(18.48)	45(18.60)	96(18.43)		
是	622(81.52)	197(81.40)	425(81.57)		
糖尿病[n(%)]				4.339	0.037
否	460(60.29)	159(65.70)	301(57.77)		
是	303(39.71)	83(34.30)	220(42.23)		
脑卒中史[n(%)]				3.504	0.061
否	413(54.13)	119(49.17)	294(56.43)		
是	350(45.87)	123(50.83)	227(43.57)		
吸烟[n(%)]				1.357	0.244
否	703(92.14)	227(93.80)	476(91.36)		
是	60(7.86)	15(6.20)	45(8.64)		
饮酒[n(%)]				0.541	0.462
否	708(92.79)	227(93.80)	481(92.32)		
是	55(7.21)	15(6.20)	40(7.68)		
肺部基础疾病史[n(%)]				14.413	<0.001
否	733(96.07)	223(92.15)	510(97.89)		
是	30(3.93)	19(7.85)	11(2.11)		
使用鼻胃管治疗[n(%)]				327.611	<0.001
否	602(78.90)	96(39.67)	506(97.12)		
是	161(21.10)	146(60.33)	15(2.88)		
使用尿管[n(%)]				126.589	<0.001
否	664(87.02)	162(66.94)	502(96.35)		
是	99(12.98)	80(33.06)	19(3.65)		
使用抑酸剂治疗[n(%)]				2.480	0.115
否	561(73.53)	169(69.83)	392(75.24)		
是	202(26.47)	73(30.17)	129(24.76)		

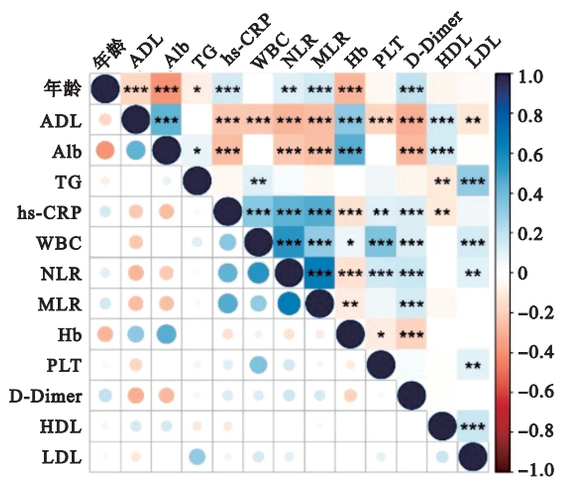
验证集中的预测性能和临床实用性更优，并且均优于先前研究的模型[10-12]。同时，为了提高模型的可解释性，研究使用了SHAP算法对XGBoost模型进行了深入解释，并揭示了各个变量对最终预测结果的贡献程度。这使得临床工作者能够更清晰地理解模型的预测逻辑，洞察患者的具体风险状况。

SHAP算法结果显示，ADL评分对预测SAP的贡献度最大。研究发现，与非SAP患者相比，SAP患者往往具有更长的住院时间和更低的ADL评分[20]。ADL评分低通常意味着患者在自我护理方面的能力受限，这可能导致他们花费更多时间卧床，而卧床可能会增加肺部感染的风险[21]。另外，ADL评分低可能与较差的营养状况有关，这会削弱免疫系统的功能，使患者更易受到感染。早期的康复锻炼有助于提高

表 2 SAP组和非SAP组患者实验室检查指标比较[$M(P_{25},P_{75})$]

实验室检查指标	全部患者(n=763)	SAP(n=242)	非SAP(n=521)	Z/t	P
Alb(g/L)*	38. 69±4. 05	36. 39±4. 24	39. 76±3. 47	−10. 806	<0. 001
TG(mmol/L)	1. 28(0. 95,1. 69)	1. 27(0. 90,1. 64)	1. 28(0. 97,1. 73)	−1. 630	0. 103
hs−CRP(mg/L)	2. 86(1. 04,10. 07)	9. 45(3. 38,19. 40)	1. 72(0. 83,4. 08)	12. 436	<0. 001
WBC(10^9/L)	6. 38(5. 16,7. 78)	6. 75(5. 28,9. 39)	6. 24(5. 10,7. 38)	4. 232	<0. 001
NLR	2. 38(1. 69,3. 38)	2. 87(1. 95,4. 71)	2. 20(1. 60,2. 98)	6. 593	<0. 001
MLR	0. 32(0. 24,0. 45)	0. 42(0. 27,0. 59)	0. 30(0. 23,0. 40)	7. 061	<0. 001
Hb(g/L)*	124. 74±17. 70	116. 09±18. 88	128. 76±15. 59	−9. 102	<0. 001
PLT(10^9/L)	208. 00(171. 50,255. 00)	214. 50(171. 50,278. 75)	204. 00(172. 00,247. 00)	1. 719	0. 086
D−Dimer(mg/L)	0. 57(0. 28,1. 39)	1. 14(0. 67,2. 33)	0. 38(0. 23,0. 82)	11. 922	<0. 001
HDL(mmol/L)	0. 93(0. 78,1. 11)	0. 89(0. 75,1. 07)	0. 94(0. 79,1. 13)	−2. 722	0. 006
LDL(mmol/L)	2. 14(1. 72,2. 71)	2. 22(1. 75,2. 81)	2. 12(1. 71,2. 63)	1. 603	0. 109

*:数据用$\bar{x}\pm s$表示。



*:P<0. 05, **:P<0. 01, ***:P<0. 001

图 1 变量间的相关性

脑卒中患者的自理能力,减少肺部感染的风险,临床工作者应针对行动不便的脑卒中患者进行及时干预。

在本研究中,脑卒中患者留置鼻胃管是发生SAP的一个危险因素。这一结果与现有文献中的报道一致[22],Brogan等[23]研究发现,鼻胃管比吞咽困难预测SAP的价值更高。研究显示,在实施鼻胃管营养支持过程中,胃腔体积显著扩张,可能导致胃部肌肉产生痉挛和食物积聚。同时,接受鼻胃管的患者普遍面临颅内压升高、呕吐和食物反流的风险,这些因素共同增加SAP的发生概率[24]。然而,一项针对急性卒中患者的研究发现,在患者发病48小时内放置鼻胃管并不会增加SAP的发生率、死亡率和不良功能结果[25],这与本研究结果相悖,因此有关鼻胃管留置时间与SAP之间的关系仍有待研究。

高龄脑卒中患者发生SAP的可能性更高,与以往的研究结果相似[11-12]。随着年龄的增长,老年人群体的身体机能和免疫系统功能逐渐衰退,这导致他们的呼吸系统防护功能有所降低。此外,相比于其他年龄段的脑卒中患者,老年脑卒中患者在吞咽和咳嗽反射方面的能力通常较弱,因此他们在脑卒中后更容易受到肺部感染的侵袭[26]。

在实验室检查指标中,高hs−CRP水平是影响SAP发生的因素,与Huang等[27]的研究结果一致。hs−CRP是一种由肝脏合成的急性期反应蛋白,在敏感性和早期鉴别性能方面表现更佳,即使在炎症初期或炎症水平较低的情况下也能被有效检测[28]。研究显示,脑梗死病人血清hs−CRP水平变化与继发肺部感染密切相关,hs−CRP水平越高,提示机体炎症反应越严重[29]。另一有效预测SAP的指标是Hb,Hb水平的下降通常意味着患者有贫血的风险,而贫血可显著

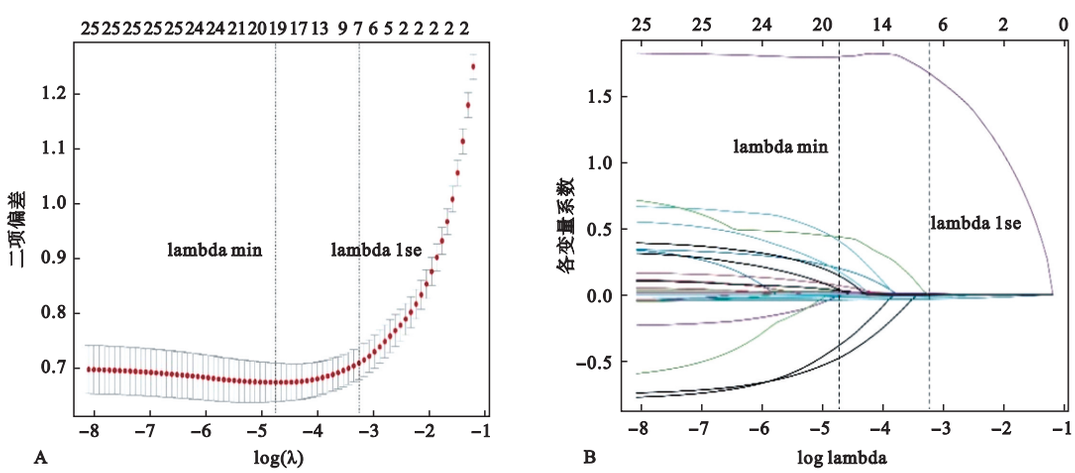


图 2 LASSO回归筛选变量过程

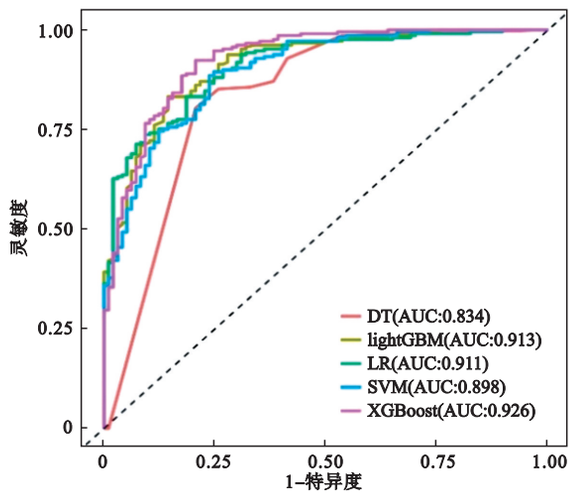


图3　5种机器学习模型的ROC曲线

增加患者的死亡率和发生肺炎的风险[30]，贫血有可能导致免疫系统功能受损，从而降低患者对感染的抵抗力，使得脑卒中患者更容易受到细菌或病毒的侵袭[31]。已有研究证实，低Alb水平的脑卒中患者更容易在住院期间发生感染或者肺炎[27,32-33]，一种可能的原因是，Alb具有抗炎、抗氧化、抗凝以及调节微血管通透性等作用，低Alb水平可能会导致全身性炎症反应从而引发肺炎[33]。此外，低Alb水平会引发患者发生营养不良，而营养不良会降低患者的整体健康状况和对感染的抵抗力，从而增加肺炎的风险[34]。

相比传统的回归方法，机器学习算法在处理高维数据、特征选择、预测精度等方面具有显著优势。XGBoost是目前较为流行的机器学习算法之一，具有正则化、运算速度快、减少过拟合等优点[35]，在风险预测等领域具有广泛的应用前景[36-38]。虽然在本研究中XGBoost表现出了良好的预测性能，但其AUC值和LR、SVM和lightGBM模型比较并无显著差异，这可能与各模型经过超参数调优、模型的预测性能均得到了提升有关，还可能与数据的随机抽样误差有关。未来研究可扩大样本量或使用其他方法更全面地评估模型的性能。此外，本研究运用SHAP算法解决了机器学习无法揭示预测结局与特征之间关系的难题，通过可视化特征的贡献度直观地展示预测过程，有助于临床工作者识别关键风险因素。

本研究仍存在一些局限性。首先，本研究为回顾性研究，数据的收集依赖于既往电子病历的信息，这有可能导致部分数据缺失或忽略了一些可能对结局事件影响较大的预测因子。其次，本模型尚未在其他人群中进行外部验证，模型的泛化性能有待进一步考察。最后，机器学习算法虽可较好地提高预测性能，但其如何与临床实际相结合仍是一个值得思考的问题。今后研究可进一步优化模型，包括优化实验设计、挖掘更多的潜在预测因子以及探讨模型在临床上的实际应用等。

综上所述，本研究构建的模型在预测SAP风险方面展现出良好的预测性能和临床适用性，其中XGBoost模型表现最优。在此基础上，本研究借助

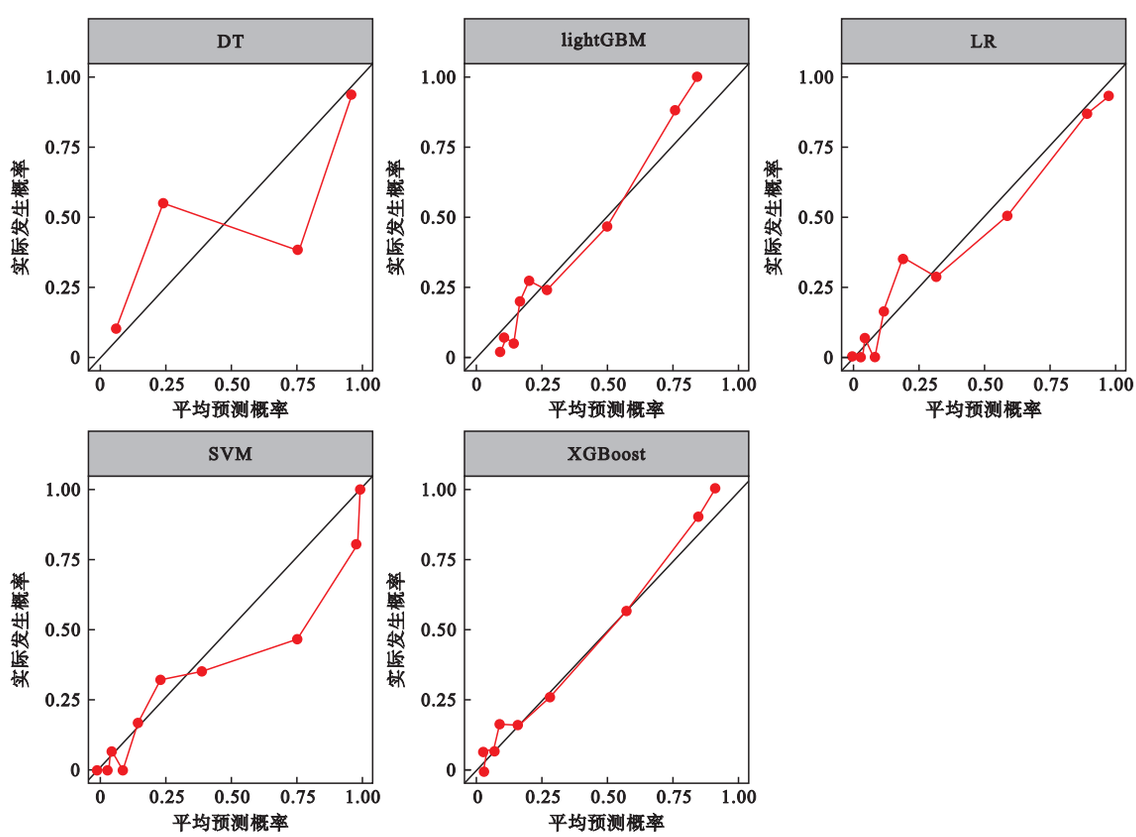


图4　5种机器学习模型的校准曲线

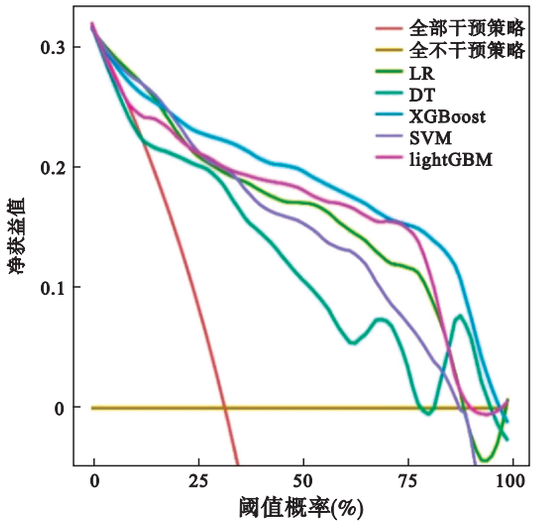


图 5　5 种机器学习模型的 DCA

表 3　5 种机器学习模型的 AUC 值比较

模型比较	*Z*	*P*	模型比较	*Z*	*P*
DT vs LR	−3. 713	0. 001	LR vs SVM	−0. 081	0. 935
DT vs XGBoost	−3. 031	0. 001	LR vs lightGBM	0. 257	0. 797
DT vs SVM	−3. 668	0. 001	XGBoost vs SVM	0. 629	0. 529
DT vs lightGBM	−2. 283	0. 020	XGBoost vs lightGBM	0. 903	0. 367
LR vs XGBoost	−0. 651	0. 515	SVM vs lightGBM	0. 268	0. 789

表 4　5 种机器学习模型的性能指标比较

	DT	LR	XGBoost	SVM	lightGBM
准确率	0. 817	0. 833	0. 853	0. 781	0. 840
Kappa 系数	0. 584	0. 614	0. 670	0. 529	0. 635
灵敏度	0. 852	0. 880	0. 866	0. 775	0. 871
特异度	0. 742	0. 732	0. 825	0. 794	0. 773
阳性预测值	0. 877	0. 876	0. 914	0. 890	0. 892
阴性预测值	0. 699	0. 740	0. 741	0. 621	0. 735
精确度	0. 877	0. 876	0. 914	0. 890	0. 892
召回率	0. 852	0. 880	0. 866	0. 775	0. 871
*F*1 分数	0. 864	0. 878	0. 889	0. 829	0. 881
AUC	0. 834	0. 911	0. 926	0. 898	0. 913

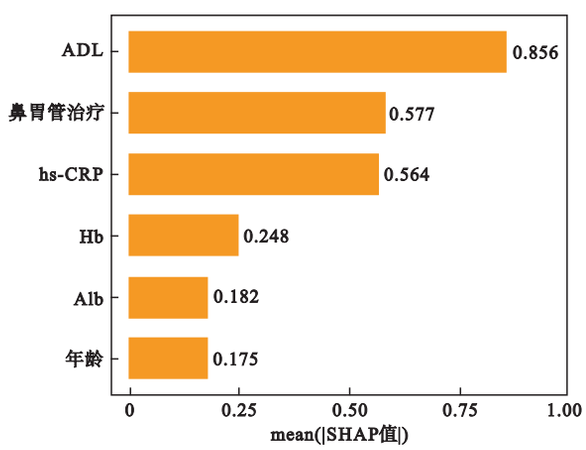


图 6　基于 SHAP 算法分析变量的重要性排序

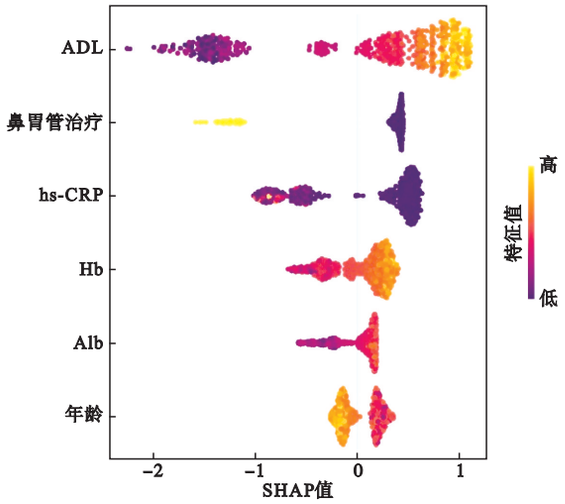


图 7　基于 SHAP 算法分析 XGBoost 模型的蜂群图

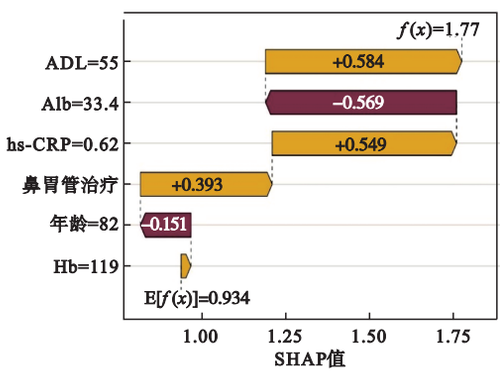


图 8　基于 SHAP 算法分析单样本发生 SAP 的风险

SHAP 算法对模型进行了深入解释，揭示了影响 SAP 发生的关键因素，为临床决策提供了参考依据。

参　考　文　献

[1] Feigin VL, Stark BA, Johnson CO, et al. Global, regional, and national burden of stroke and its risk factors, 1990—2019: a systematic analysis for the Global Burden of Disease Study 2019 [J]. Lancet Neurol, 2021, 20(10): 795-820.

[2] Smith CJ, Kishore AK, Vail A, et al. Diagnosis of Stroke-Associated Pneumonia: Recommendations from the Pneumonia in Stroke Consensus Group [J]. Stroke, 2015, 46(8): 2335-2340.

[3] Ji R, Wang D, Shen H, et al. Interrelationship among common medical complications after acute stroke: pneumonia plays an important role [J]. Stroke, 2013, 44(12): 3436-3444.

[4] Smith CJ, Bray BD, Hoffman A, et al. Can a Novel Clinical Risk Score Improve Pneumonia Prediction in Acute Stroke Care? A UK Multicenter Cohort Study [J]. J Am Heart Assoc, 2015, 4(1): e001307.

[5] Teh WH, Smith CJ, Barlas RS, et al. Impact of stroke-associated pneumonia on mortality, length of hospitalization, and functional outcome [J]. Acta Neurol Scand, 2018, 138(4): 293-300.

[6] Yu YJ, Weng WC, Su FC, et al. Association between pneumonia in acute stroke stage and 3-year mortality in patients with acute first-ever ischemic stroke [J]. J Clin Neurosci, 2016, 33: 124-128.

[7] Houck PM, Bratzler DW, Nsa W, et al. Timing of antibiotic administration and outcomes for Medicare patients hospitalized with community-acquired pneumonia [J]. Arch Intern Med, 2004, 164

(6): 637-644.
[8] Vermeij JD, Westendorp WF, Dippel DW, et al. Antibiotic therapy for preventing infections in people with acute stroke [J]. Cochrane Database Syst Rev, 2018, 1(1): CD008530.
[9] Ranstam J, Cook JA, Collins GS. Clinical prediction models [J]. Br J Surg, 2016, 103(13): 1886.
[10] Kwon HM, Jeong SW, Lee SH, et al. The pneumonia score: A simple grading scale for prediction of pneumonia after acute stroke [J]. Am J Infect Control, 2006, 34(2): 64-68.
[11] Hoffmann S, Malzahn U, Harms H, et al. Development of a Clinical Score (A2DS2) to Predict Pneumonia in Acute Ischemic Stroke [J]. Stroke, 2012, 43(10): 2617-2623.
[12] Ji R, Shen H, Pan Y, et al. Novel Risk Score to Predict Pneumonia After Acute Ischemic Stroke [J]. Stroke, 2013, 44(5): 1303-1309.
[13] 谷鸿秋, 周支瑞, 章仲恒, 等. 临床预测模型:基本概念、应用场景及研究思路[J]. 中国循证心血管医学杂志, 2018, 10(12): 1454-1456.
[14] Collin C, Wade DT, Davies S, et al. The Barthel ADL Index: A reliability study [J]. Int Disabil Stud, 1988, 10(2): 61-63.
[15] Osawa A, Maeshima S, Tanahashi N. Water-Swallowing Test: Screening for Aspiration in Stroke Patients [J]. Cerebrovasc Dis, 2013, 35(3): 276-281.
[16] Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale [J]. Lancet Lond Engl, 1974, 2(7872): 81-84.
[17] Horan TC, Andrus M, Dudeck MA. CDC/NHSN surveillance definition of health care-associated infection and criteria for specific types of infections in the acute care setting [J]. Am J Infect Control, 2008, 36(5): 309-332.
[18] Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model [J]. BMJ, 2020: m441.
[19] 曹燚, 曾茜, 赵信飞, 等. 卒中相关性肺炎风险预测模型的系统评价[J]. 中国循证医学杂志, 2023, 23(11): 1259-1268.
[20] Maeshima S, Osawa A, Hayashi T, et al. Elderly Age, Bilateral Lesions, and Severe Neurological Deficit Are Correlated with Stroke-associated Pneumonia [J]. J Stroke Cerebrovasc Dis, 2014, 23(3): 484-489.
[21] Matsusaka K, Kawakami G, Kamekawa H, et al. Pneumonia risks in bedridden patients receiving oral care and their screening tool: Malnutrition and urinary tract infection-induced inflammation [J]. Geriatr Gerontol Int, 2018, 18(5): 714-722.
[22] Schwarz M, Coccetti A, Murdoch A, et al. The impact of aspiration pneumonia and nasogastric feeding on clinical outcomes in stroke patients: A retrospective cohort study [J]. J Clin Nurs, 2018, 27(1-2): e235-e241.
[23] Brogan E, Langdon C, Brookes K, et al. Respiratory infections in acute stroke: nasogastric tubes and immobility are stronger predictors than dysphagia [J]. Dysphagia, 2014, 29(3): 340-345.
[24] Wang Q, Liu Y, Han L, et al. Risk factors for acute stroke-associated pneumonia and prediction of neutrophil-to-lymphocyte ratios [J]. Am J Emerg Med, 2021, 41: 55-59.
[25] Kalra L, Hodsoll J, Irshad S, et al. Association between nasogastric tubes, pneumonia, and clinical outcomes in acute stroke patients [J]. Neurology, 2016, 87(13): 1352-1359.
[26] Patel UK, Kodumuri N, Dave M, et al. Stroke-Associated Pneumonia: A Retrospective Study of Risk Factors and Outcomes [J]. The Neurologist, 2020, 25(3): 39-48.
[27] Huang L, Zhang R, Ji J, et al. Hypersensitive C-reactive protein-albumin ratio is associated with stroke-associated pneumonia and early clinical outcomes in patients with acute ischemic stroke [J]. Brain Behav, 2022, 12(7): e2675.
[28] Banait T, Wanjari A, Danade V, et al. Role of High-Sensitivity C-reactive Protein (hs-CRP) in Non-communicable Diseases: A Review [J]. Cureus, 2022, 14(10): e30225.
[29] Wang L, Wu L, Lang Y, et al. Association between high-sensitivity C-reactive protein levels and clinical outcomes in acute ischemic stroke patients treated with endovascular therapy [J]. Ann Transl Med, 2020, 8(21): 1379.
[30] Wei CC, Zhang ST, Tan G, et al. Impact of anemia on in-hospital complications after ischemic stroke [J]. Eur J Neurol, 2018, 25(5): 768-774.
[31] Song X, He Y, Bai J, et al. A nomogram based on nutritional status and A2DS2 score for predicting stroke-associated pneumonia in acute ischemic stroke patients with type 2 diabetes mellitus: A retrospective study [J]. Front Nutr, 2022, 9: 1009041.
[32] Yang X, Wang L, Zheng L, et al. Serum Albumin as a Potential Predictor of Pneumonia after an Acute Ischemic Stroke [J]. Curr Neurovasc Res, 2020, 17(4): 385-393.
[33] Lin G, Hu M, Song J, et al. High Fibrinogen to Albumin Ratio: A Novel Marker for Risk of Stroke-Associated Pneumonia [J]. Front Neurol, 2022, 12: 747118.
[34] Chen L, Xu M, Huang Q, et al. Clinical significance of albumin to globulin ratio among patients with stroke-associated pneumonia [J]. Front Nutr, 2022, 9: 970573.
[35] 卢娅欣, 黄月, 李康. XGBoost算法在二分类非平衡高维数据分析中的应用[J]. 中国卫生统计, 2021, 38(1): 21-24.
[36] Zhang Y, Xiang T, Wang Y, et al. Explainable machine learning for predicting 30-day readmission in acute heart failure patients [J]. iScience, 2024, 27(7): 110281.
[37] Islam MM, Rahman MJ, Rabby MS, et al. Predicting the risk of diabetic retinopathy using explainable machine learning algorithms [J]. Diabetes Metab Syndr, 2023, 17(12): 102919.
[38] Hou N, Li M, He L, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using Xgboost [J]. J Transl Med, 2020, 18(1): 1-14.

（责任编辑：林美艺）