· 临床研究 ·

基于可解释机器学习构建脑卒中患者日常生活自理能力风险预测模型

叶 倩[1,2]，杨 云[1,3]，徐文韬[2]，刘玲玲[1*]

[1]南京医科大学附属第一医院康复医学中心，江苏 南京 210029；[2]南京中医药大学针灸推拿学院，江苏 南京 210029；[3]南京师范大学心理学院，江苏 南京 210023

［摘 要］ **目的**：利用机器学习算法预测影响脑卒中患者日常生活自理能力（activities of daily living，ADL）的风险因素，为其ADL管理决策提供参考。**方法**：对2015年1月—2019年2月在南京医科大学附属第一医院康复医学中心治疗的423例脑卒中患者进行回顾性分析。根据Barthel指数（Barthel index，BI）评定量表，将患者分为ADL较好组（BI≥60分）和ADL较差组（BI<60分），并进行数据预处理。采用共线性诊断及最小绝对收缩和选择算子（least absolute shrinkage and selection operator，LASSO）筛选特征变量。选择逻辑回归、支持向量机、随机森林（random forest，RF）、极限梯度提升及K最近邻5种机器学习算法进行预测建模，十倍交叉验证后，使用受试者工作特征曲线、受试者工作特征曲线下面积（area under curve，AUC）、精确召回率曲线、精确召回率曲线下的面积（area under the precision recall curve，PRAUC）、准确率、灵敏度、特异度分别对模型进行综合评估，引入Shapley加性解释（Shapley additive explanation，SHAP）对最优机器学习模型进行可解释化处理。**结果**：经LASSO回归分析后，确定16个特征变量用于构建机器学习模型。RF模型具有最高的AUC（0.74）、PRAUC（0.64）、准确率（0.97）、灵敏度（0.75）和特异度（0.97）。SHAP模型解释性分析显示，对ADL贡献度前5的特征中，Brunnstrom分期（下肢）的影响最为显著，其次是Brunnstrom分期（上肢）、D-二聚体、血清白蛋白水平及年龄。**结论**：RF模型预测脑卒中患者ADL的效能最优，为脑卒中患者ADL管理决策提供了有价值的参考。

［关键词］ 机器学习；预测模型；脑卒中；日常生活自理能力

［中图分类号］ R743.3 ［文献标志码］ A ［文章编号］ 1007-4368（2024）05-672-09

doi：10.7655/NYDXBNSN240009

Constructing a prediction model for stroke patients' activities of daily living risk based on interpretable machine learning

YE Qian[1,2]，YANG Yun[1,3]，XU Wentao[2]，LIU Lingling[1*]

[1]Center of Rehabilitation Medicine，the First Affiliated Hospital of Nanjing Medical University，Nanjing 210029；[2]College of Acupuncture and Massage，Nanjing University of Chinese Medicine，Nanjing 210029；[3]School of Psychology，Nanjing Normal University，Nanjing 210023，China

［**Abstract**］ **Objective**：To utilize machine learning algorithms to predict risk factors affecting the activities of daily living（ADL）of stroke patients，providing references for their ADL management decisions. **Methods**：A retrospective analysis was conducted on 423 stroke patients treated at the Rehabilitation Medicine Center of the First Affiliated Hospital of Nanjing Medical University from January 2015 to February 2019. Patients were categorized into a better ADL group（BI ≥ 60 points）and a poorer ADL group（BI <60 points）based on the Barthel Index（BI）assessment scale，and data preprocessing was performed. Feature variables were selected using colinearity diagnostics and the least absolute shrinkage and selection operator（LASSO）. Logistic regression（LR），support vector machine（SVM），random forest（RF），extreme gradient boosting（XGBoost），and K nearest neighbor（KNN）were selected as the five machine learning algorithms for predictive modeling. Afterten-fold cross-validation，the models were comprehensively evalutated using receiver operating characteristic（ROC）curves，area under aerue（AUC），precision recall（PR）curves，area under the precision recall

［基金项目］ 国家自然科学基金（82104993）

*通信作者（Corresponding author），E-mail：ch600lll@163.com

curve（PRAUC），accuracy，sensitivity，and specificity. The Shapley additive interpretation（SHAP）was introduced to interpret the optimal machine learning model. **Results:** After LASSO regression analysis，16 feature variables were identified for constructing the machine learning model. The RF model demonstrated superior performance with the highest AUC（0.74），PRAUC（0.64），accuracy（0.97），sensitivity（0.75），and specificity（0.97）. Interpretive analysis of the SHAP model revealed that among the top 5 features contributing to ADL，Brunnstrom stage（lower limb）exerted the most significant effect，followed by Brunnstrom stage（upper limb），D-dimer，serum albumin level，and age. **Conclusion:** The RF model emerged as the most effective in predicting ADL in stroke patient，providing valuable references for ADL management decisions in stroke patients.

[**Key words**]　machine learning；predictive modeling；stroke；activities of daily living

[J Nanjing Med Univ，2024，44（05）：672-680]

脑卒中是一种由于脑部血管突然破裂或因血管阻塞导致血液不能流入大脑的急性脑血管疾病，超过75%的脑卒中患者会留下不同程度的功能障碍，其中超过42.3%的患者会留下残疾导致生活不能自理，是我国成人致残、致死的首要病因[1-2]。及时早期识别影响脑卒中患者日常生活自理能力（activities of daily living，ADL）的风险因素，对提高患者脑卒中后生活自理能力，改善生活质量，促进患者更好地回归家庭及社会至关重要[3]。

影响脑卒中患者ADL的因素众多，包括年龄、性别、损伤部位、肢体功能、既往史、有无静脉血栓、营养状况等[4-8]。目前脑卒中后患者的ADL大都依靠人工判断，医生的临床经验和一些主观因素均会影响最终的判断结果，因此开发简单、易于推广的脑卒中后患者ADL风险预测模型具有重要的临床意义。随着大数据、人工智能等技术在医学领域的广泛应用，机器学习（machine learning，ML）技术目前已被应用于探讨影响康复效果的因素、预测康复效果等[9-12]。ML中的算法非常有助于做出准确的预测和给出正确的分析，利用ML技术，医生可以更快地确认患者病症，为患者的后续治疗争取时间，因此该技术具有重要的研究意义。本研究旨在通过ML算法精准识别脑卒中患者ADL的高风险因素，以更好地辅助临床决策和治疗。

1　对象和方法

1.1　对象

本研究是一项回顾性、观察性研究，以2015年1月—2019年2月在南京医科大学附属第一医院康复医学中心的脑卒中患者为研究对象。纳入标准：①经头颅计算机断层扫描（computed tomography，CT）或磁共振成像（magnetic resonance imaging，MRI）扫描诊断为脑卒中（缺血性或出血性）；②年龄18~85周岁；③临床及实验室医疗记录相对完整。排除标准：①合并下肢骨折病史；②合并严重的心、肺、肝脏等其他系统严重疾病；③既往有下肢深静脉血栓；④临床电子医疗记录不完整。本研究经南京医科大学附属第一医院伦理委员会批准通过（伦理号：2023-SR-104），并知情同意。

1.2　方法

1.2.1　数据收集

依据专家意见及相关文献查阅[4-8]，筛选与ADL相关的41个预测变量，除患者下肢深静脉血栓检查为入院后3 d内完成，其余变量采集时间均为入院24 h内（所有康复评估指标均由高年资康复专业人员评定）。主要包括：①人口统计学特征：性别、年龄；②现病史：脑卒中部位、脑卒中类型、病程等3个指标；③既往史：高血压病史、糖尿病病史、既往脑卒中病史等3个指标；④个人史：吸烟史、饮酒史；⑤康复评估指标：Brunnstrom分期（上肢、手、下肢）、小腿三头肌肌张力等4个指标；⑥临床用药指标：抗凝药物服用史、抗血小板聚集药物服用史；⑦实验室检查指标：血清载脂蛋白（lipoprotein，Lpa）、低密度脂蛋白（low density lipoprotein cholesterol，LDL-C）、高密度脂蛋白（high density lipoprotein cholesterol，HDL-C）、总胆固醇（total cholesterol，TC）、甘油三酯（triglyceride，TG）、血清白蛋白（albumin，ALB）、血清总蛋白（total protein，TP）、血小板压积（plateletcrit，PCT）、血小板体积分布宽度（platelet distribution width，PDW）、血小板平均容积（mean platelet volume，MPV）、血小板计数（blood platelet count，PLT）、红细胞平均血红蛋白浓度（mean corpuscular hemoglobin concentration，MCHC）、平均红细胞血红蛋白（mean corpuscular hemoglobin，MCH）、血红蛋白（hemoglobin，HGB）、红细胞分布宽度（coefficient of variation of erythrocyte distribution width，RDW-CV）、红细胞平均容量（mean

corpuscular volume，MCV）、红细胞压积（hematocrit，HCT）、红细胞计数（red blood count，RBC）、凝血酶原时间（prothrombin time，PT）、部分凝血活酶时间（activated partial thromboplastin time，APTT）、纤维蛋白原（fibrinogen，FIB）、国际标准化比值（international normalized ratio，INR）、凝血酶时间（thrombin time，TT）、D-二聚体（D-dimer）、下肢深静脉血栓形成等25个指标。

1.2.2 模型构建与评估

利用最低绝对收缩和选择算子（least absolute shrink age and selection operator，LASSO）回归进行特征降维筛选。基于K近邻模型（K-nearest neighbor，KNN）、支持向量机（support vector machine，SVM）、极端梯度提升（extreme gradient boosting，XGB）、随机森林（random forest，RF）、逻辑回归（logistic regression，LR）5种ML算法进行预测建模。

1.3 统计学方法

对于缺失比例<5%的指标，使用随机森林填补缺失数据。正态分布的连续变量表示为均数±标准差（$\bar{x} \pm s$），组间比较采用两独立样本t检验；非正态分布的连续变量表示为中位数（四分位数）[$M(P_{25}, P_{75})$]，组间比较采用秩和检验。分类变量用例数（百分率）表示，组间比较采用χ^2检验。采用受试者工作特征（receiver operating characteristic，ROC）曲线、受试者工作特征曲线下面积（area under curve，AUC）、精确召回率（precision recall，PR）曲线、精确召回率曲线下的面积（area under the precision recall curve，PRAUC）、准确率、灵敏度和特异度进行模型评估，使用十倍交叉验证进行模型训练。

采用R软件（版本4.2.2，https://www.r-project.org）完成数据预处理、特征选择、超参数优化、ML模型构建和综合评估。使用Python（版本3.11.3，https://www.python.org）绘制SHAP解释。$P < 0.05$为差异有统计学意义。

2 结 果

2.1 基线比较

根据纳除标准，最终纳入研究对象423例患者，其中，BI<60分351例（82.98%），≥60分72例（17.02%）。从基线结果（表1）中可以看出ADL与脑卒中类型（脑出血/脑梗死）、有无下肢深静脉血栓、是否使用抗血小板聚集药物、Brunnstrom分期（上肢、手、下肢）、小腿三头肌肌张力、ALB、RDW-CV、INR以及D-二聚体有关（P均< 0.05）。

2.2 特征变量筛选及模型构建

经LASSO回归分析后，共筛选了16个特征（图1），包括年龄、卒中类型、吸烟史、饮酒史、既往卒中史、Brunnstrom分期（上肢）、Brunnstrom分期（下肢）、抗血小板聚集药服用史、HDL-C、PCT、HGB、D-二聚体、RDW-CV、ALB、RBC、INR。基于上述16个特征构建5个ML模型。对ROC曲线、AUC、PR曲线、PRAUC、准确率、灵敏度和特异度进行模型评估后发现RF模型为最优模型（图2），其具有最大的AUC值（0.74），PRAUC（0.64），准确性（0.97），敏感度（0.75）和特异度（0.97）（表2）。因此，本研究最终选择了RF模型进行下一步的分析和应用。

2.3 最优模型解释及可视化

在基于SHAP框架对影响卒中后患者ADL的风险预测模型结果进行解释性分析。图3为SHAP特征摘要，根据特征重要性对影响ADL的风险因素进行分析。如图3所示，贡献度前5的特征分别为Brunnstrom分期（下肢）、Brunnstrom分期（上肢）、D-二聚体、ALB、年龄。采用SHAP对RF模型中重要特征如何影响预测结果进行分析，如图4所示，当下肢、上肢的Brunnstrom分期达到Ⅲ期时ADL开始提高，Ⅳ期后ADL显著提高；当D-二聚体增高时，脑卒中后患者的ADL会下降，其值接近2 mg/L时，则趋于稳定；ALB接近35 g/L后，脑卒中后患者的ADL显著提高；随着年龄增大，脑卒中后患者ADL整体呈下降趋势。

3 讨 论

脑卒中后患者ADL受多重因素影响，为避免患者卒中后ADL丧失，保证其生活质量，构建针对卒中后患者ADL的预测模型具有较高的临床价值。本研究纳入的423例脑卒中患者，通过LASSO回归筛选了16个广泛使用的临床特征，包括年龄、脑卒中类型、吸烟史、饮酒史、既往脑卒中史、Brunnstrom分期（下肢）、Brunnstrom分期（上肢）、抗血小板聚集药服用史、HDL-C、PCT、HGB、D-二聚体、RDW-CV、ALB、RBC、INR，进而构建了一种基于RF算法的可解释的ML模型，预测脑卒中后患者ADL风险发生的综合性能评价最好，可为提高脑卒中后ADL决策提供理论支持。

SHAP值表明，在脑卒中后患者ADL风险的16个特征中，排在前5位的分别是Brunnstrom分期（下肢）、Brunnstrom分期（上肢）、D-二聚体、ALB、年龄。目前许多研究表明脑卒中后约2/3的患者会遗

表1　基线特征比较

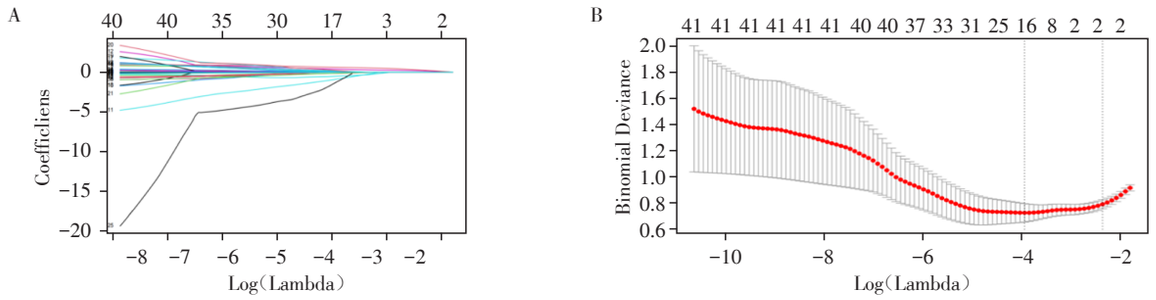
Table 1　Comparison of baseline characteristics

Characteristic	BI<60 points group（n=351）	BI≥60 points group（n=72）	P
Age（years，$\bar{x} \pm s$）	60.46 ± 13.48	57.28 ± 13.37	0.07
Sex[n（%）]			0.67
Male	246（70.09）	48（66.67）	
Female	105（29.91）	24（33.33）	
Course[months，M（P_{25}，P_{75}）]	1.00（1.00，2.00）	1.00（1.00，2.00）	0.32
Position of stroke [n（%）]			0.91
Basal ganglia region	181（51.57）	36（50.00）	
Non-basal ganglia region	170（48.43）	36（50.00）	
Type of stroke [n（%）]			0.01
Hemorrhage	145（41.31）	18（25.00）	
Ischemic	206（58.69）	54（75.00）	
Hypertension [n（%）]			0.35
Yes	256（72.93）	48（66.67）	
No	95（27.07）	24（33.33）	
Diabete mellitus [n（%）]			0.54
Yes	92（26.21）	22（30.56）	
No	259（73.79）	50（69.44）	
History of smoking [n（%）]			0.09
Yes	107（30.48）	30（41.67）	
No	244（69.52）	42（58.33）	
History of alcohol [n（%）]			0.05
Yes	64（18.23）	21（29.17）	
No	287（81.77）	51（70.83）	
History of stroke [n（%）]			0.44
Yes	38（10.83）	5（6.94）	
No	313（89.17）	67（93.06）	
Lower limb vein thrombosis [n（%）]			0.02
Yes	108（30.77）	12（16.67）	
No	243（69.23）	60（83.33）	
History of anticoagulant[n（%）]			0.07
Yes	36（10.26）	2（2.78）	
No	315（89.74）	70（97.22）	
History of antiplatelet[n（%）]			0.01
Yes	153（43.59）	47（65.28）	
No	198（56.41）	25（34.72）	
Brunnstrom stage（upper limb）[n（%）]			<0.01
Ⅰ	70（19.94）	1（1.39）	
Ⅱ	139（39.60）	4（5.56）	
Ⅲ	80（22.79）	23（31.94）	
Ⅳ	28（7.98）	14（19.44）	
Ⅴ	26（7.41）	15（20.83）	
Ⅵ	8（2.28）	15（20.83）	
Brunnstrom stage（hand）[n（%）]			<0.01
Ⅰ	70（19.94）	1（1.39）	
Ⅱ	139（39.60）	4（5.56）	
Ⅲ	80（22.79）	23（31.94）	
Ⅳ	28（7.98）	14（19.44）	
Ⅴ	34（9.67）	30（41.67）	

（续表1）

Characteristic	BI<60 points group（n=351）	BI≥60 points group（n=72）	P
Brunnstrom stage（lower limb）[n（%）]			<0.01
Ⅰ	47（13.39）	0（0）	
Ⅱ	90（25.64）	2（2.78）	
Ⅲ	127（36.18）	18（25.00）	
Ⅳ	57（16.24）	23（31.94）	
Ⅴ	29（8.26）	21（29.17）	
Ⅵ	1（0.28）	8（11.11）	
Triceps surae（Ashworth scale）[n（%）]			<0.01
0	207（58.97）	42（58.33）	
1	91（25.92）	18（25.00）	
2	37（10.54）	8（11.11）	
3	14（3.99）	4（5.56）	
4	2（0.57）	0（0）	
Lpa[mg/L，$M(P_{25},P_{75})$]	215.00（110.50，470.00）	155.00（93.25，446.25）	0.14
LDL-C（mmol/L，$\bar{x}\pm s$）	2.46 ± 0.86	2.29 ± 0.84	0.13
HDL-C（mmol/L，$\bar{x}\pm s$）	0.98 ± 0.43	1.07 ± 0.37	0.08
TC（mmol/L，$\bar{x}\pm s$）	3.96 ± 1.32	3.82 ± 1.03	0.38
TG（mmol/L，$\bar{x}\pm s$）	1.48 ± 0.83	1.57 ± 0.94	0.40
ALB（g/L，$\bar{x}\pm s$）	36.67 ± 3.96	38.31 ± 3.13	<0.01
TP（g/L，$\bar{x}\pm s$）	63.71 ± 5.34	64.68 ± 4.51	0.15
PCT（%，$\bar{x}\pm s$）	0.25 ± 0.07	0.24 ± 0.06	0.39
PDW（%，$\bar{x}\pm s$）	13.16 ± 3.86	13.13 ± 2.44	0.94
MPV（fL，$\bar{x}\pm s$）	10.76 ± 1.14	10.75 ± 1.08	0.94
PLT（/L，$\bar{x}\pm s$）	235.69 ± 72.00	227.24 ± 62.72	0.36
MCHC（g/L，$\bar{x}\pm s$）	334.48 ± 11.30	335.97 ± 10.93	0.31
MCH（pg，$\bar{x}\pm s$）	30.27 ± 2.53	30.28 ± 1.87	0.97
HGB（g/L，$\bar{x}\pm s$）	128.53 ± 17.76	127.59 ± 19.11	0.69
RDW-CV（%，$\bar{x}\pm s$）	13.33 ± 1.26	12.92 ± 1.04	0.01
MCV（fl，$\bar{x}\pm s$）	90.28 ± 6.29	90.14 ± 5.78	0.86
HCT（%，$\bar{x}\pm s$）	38.36 ± 5.35	38.49 ± 4.21	0.85
RBC（$\times10^9$/L，$\bar{x}\pm s$）	4.29 ± 0.86	4.47 ± 1.33	0.15
PT（s，$\bar{x}\pm s$）	12.76 ± 3.99	12.32 ± 2.96	0.38
APTT（s，$\bar{x}\pm s$）	28.20 ± 5.20	27.88 ± 3.63	0.62
FIB[g/L，$M(P_{25},P_{75})$]	1.12（1.04，2.42）	1.06（1.02，2.39）	0.17
INR[$M(P_{25},P_{75})$]	2.73（1.15，3.77）	2.10（1.10，2.94）	<0.01
TT [s，n（%）]	18.69（6.81）	19.23（6.54）	0.53
D-dimer[mg/L，$M(P_{25},P_{75})$]	0.78（0.39，1.69）	0.37（0.22，0.72）	<0.01

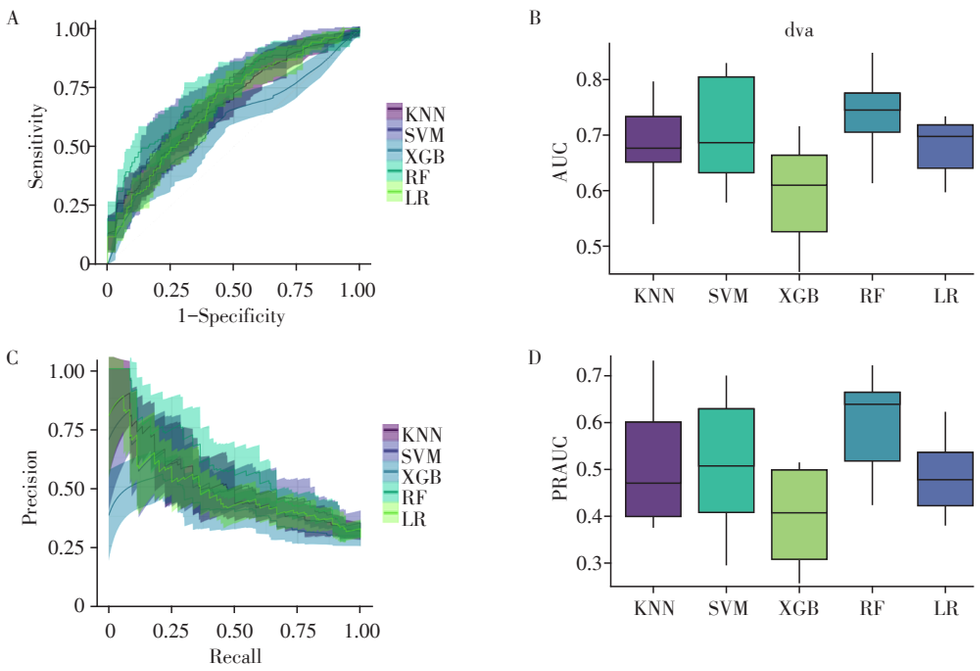
留患侧肢体运动功能障碍，严重影响患者的生活质量和自理能力[13-14]。本研究表明肢体功能的好坏，尤其是下肢功能，直接影响了患者ADL，与既往相关文献的结果一致[15-16]。这可能是因为ADL的步行和移动能力是日常生活中的基本要求，若下肢功能受损，患者将面临行动不便的问题，限制他们的活动范围和频率，进而影响生活质量，上肢在大脑皮层分区中的投射区面积更大，且上肢运动较下肢也更精细复杂，所承担的功能更多，患者的自我修复能力更弱[17]。上肢较下肢恢复更缓慢，康复难度更大，对脑卒中后患者ADL的影响也很大[18-19]，在日常康复训练中应该探索有益的康复方法促进脑卒中患者上肢的运动功能。其次影响因素为D-二聚体，D-二聚体主要反映纤维蛋白溶解功能，只要机体血管内有



A：Log（Lambda）values of the 41 features in the LASSO model. A coefficient profile plot was produced against the log（Lambda）sequence. B：Parameter selection in the LASSO model used five-fold cross validation via minimum criterion. Partial likelihood deviation（binomial deviation）curves and logarithmic（Lambda）curves were plotted.

图1　LASSO回归特征变量筛选

Figure 1　LASSO regression feature variable screening



A：ROC curve of ML models，with the y-axis representing 1-specificity and the x-axis representing sensitivity. B：Boxplot of AUC values of ML models，with the x-axis representing different ML models and the y-axis representing AUC value. Larger AUC values indicate better model performance. C：PR plot of ML models with the x-axis representing recall and the y-axis representing precision. D：Boxplot of PRAUC of ML models，with the x-axis representing different ML models and the y-axis representing PRAUC value. Larger PRAUC value indicate better model performance.

图2　ML模型综合分析

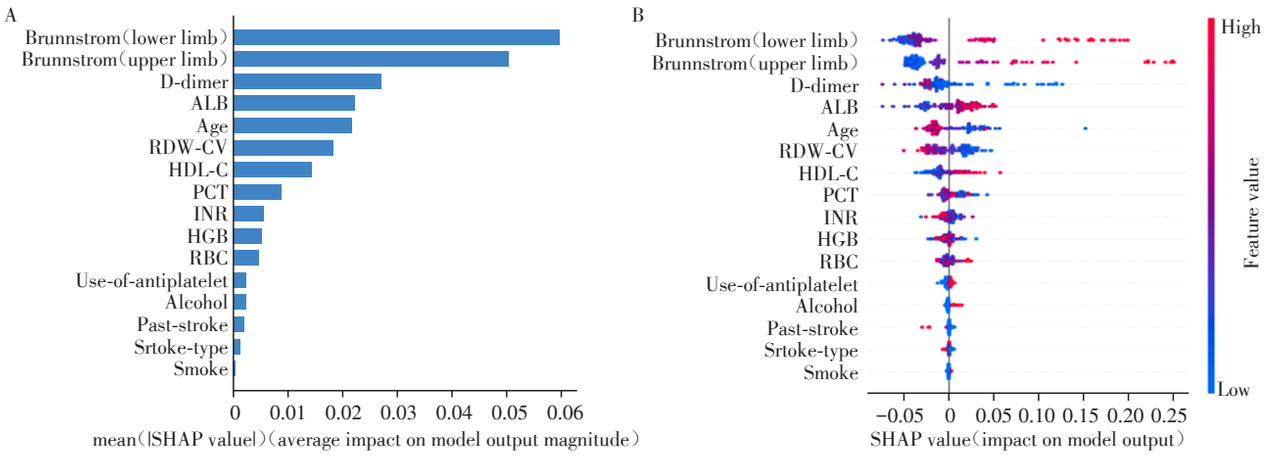
Figure 2　Comprehensive analysis of ML model

表2　ML模型性能综合评估

Table 2　Comprehensive evaluation of ML model performance

Model	AUC	PRAUC	Accuracy	Sensitivity	Specificity
KNN	0.67	0.47	0.93	0.63	0.97
RF	0.74	0.64	0.97	0.75	0.97
SVM	0.68	0.51	0.90	0.60	0.97
XGB	0.61	0.41	0.93	0.60	0.97
LR	0.69	0.47	0.95	0.67	0.97

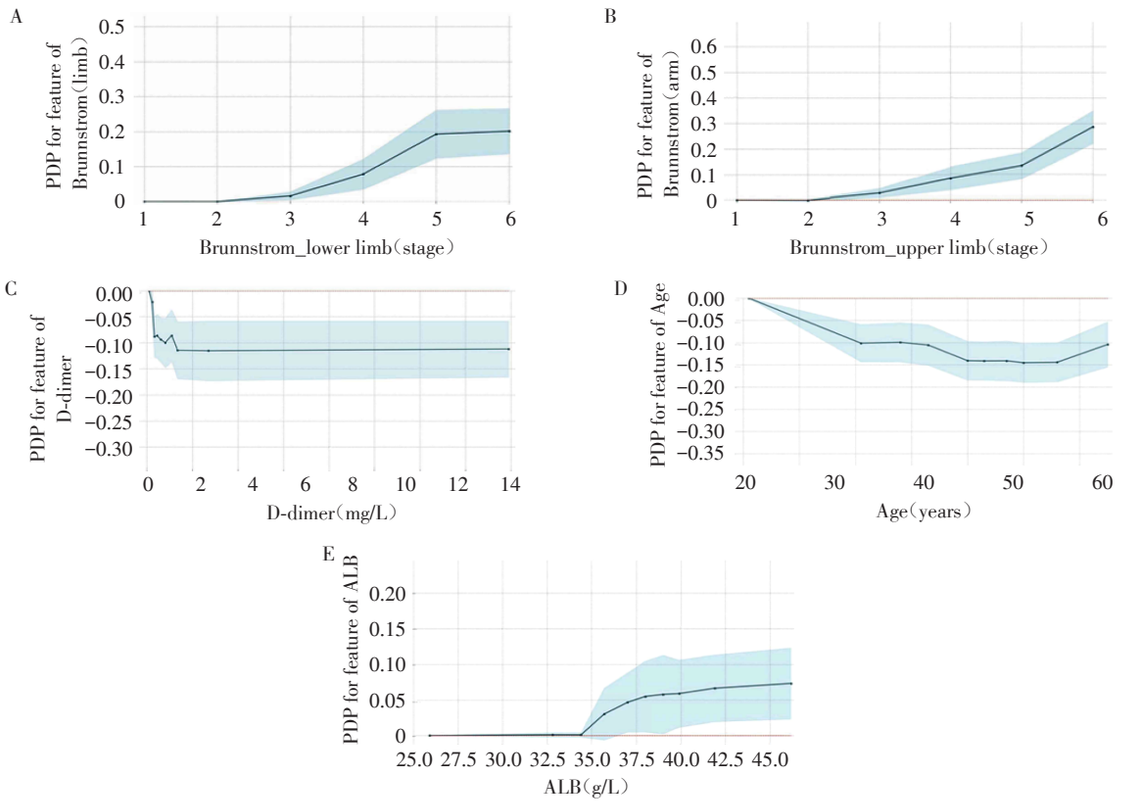
活化的血栓形成或纤维溶解活动，D-二聚体会增高，临床上主要用于排除静脉血栓性疾病[20-21]。患者发生脑卒中后首先进入ICU或者神经内科接受治疗，治疗过程中以卧床为主，活动量减少，增加了下肢静脉血栓的风险性[22-24]。但在对ADL影响方面，本研究最终模型并未纳入下肢深静脉血栓形成，而D-二聚体成为影响ADL的第3大因素，且D-二聚体接近2 mg/L时对ADL的影响趋于稳定。这提



A: Each row represents a feature, with higher SHAP value indicating greater contribution. B: The SHAP value of each feature, where each point represents a sample. The greater the dispersion of points, the more significant the impact of the feature. Colors represent feature values (red for high, blue for low).

图3　RF模型SHAP特征分析

Figure 3　SHAP feature analysis of RF model



The x-axis represents the actual value of the feature, while the y-axis represents the corresponding SHAP value; SHAP value >0 indicates an increase in the ability to care for oneself in ADL. A: Partial dependence plot of SHAP features for Brunnstrom stage (lower limb). The unique grio points are six. B: Partial dependence plot of SHAP features for Brunnstrom stage (upper limb). The unique grio points are six. C: Partial dependence plot of SHAP features for D-dimer. The unique grio points are ten. D: Partial dependence plot of SHAP features for age. E: Partial dependence plot of SHAP features for ALB. The unique grio points are ten.

图4　RF模型SHAP特征部分依赖图

Figure 4　SHAP feature partial dependence diagrams of RF model

示近2 mg/L以内的血清D-二聚体水平对ADL的影响更有预测意义。当D-二聚体在此范围内进行性增加即需要警惕，及时改变卧床状态，适当增加活动量，这对提高ADL能力具有重要意义。ALB是临床上评价

蛋白质营养状况常用的指标之一，可用于评价营养不良的程度，目前ALB与脑卒中患者预后的关系无法给出统一结论[25-26]。本研究表明，ALB水平接近35 g/L后患者ADL显著提高，当ALB达到40 g/L后对脑卒中后患者ADL的提高趋于平缓，提示在临床康复过程中应加强蛋白摄入，尽量将患者ALB提高到40 g/L，至少达到35 g/L。排在第5位的是年龄，尤其是20~40岁之间，随着年龄的增大，ADL大幅度下降，60岁之后，年龄对ADL的影响较前下降。

本研究围绕脑卒中后ADL相关影响因素进行分析，基于ML构建了5个模型，对于预测卒中患者ADL的最佳模型进行了解释性分析。另外，传统ML可以处理海量、高维的复杂数据，但无法获知预测结果与各个特征变量之间的量效关系，本研究借助Shapley法着重解决了这个问题，通过把每例患者的预测过程可视化来使医师和患者更清楚地了解主要的危险因素，进而实施个性化的二级预防措施。本研究也存在一定局限性。首先，本研究为单中心回顾性研究，仍需多中心前瞻性临床实验验证模型的准确性；其次，在本研究的模型中，可以看出ADL可能受多种因素影响，不同区域、患者家庭水平（收入、学历）等可能对结果产生一定影响。因此本研究结果仍需要引进外部数据进行验证。

综上所述，RF模型预测卒中后患者ADL效能最优，可为提高脑卒中后患者ADL的决策提供参考。

[参考文献]

[1] MA Q F，LI R，WANG L J，et al. Temporal trend and attributable risk factors of stroke burden in China，1990-2019：an analysis for the Global Burden of Disease Study 2019[J]. Lancet Public Health，2021，6(12)：e897-e906

[2] 《中国脑卒中防治报告》编写组.《中国脑卒中防治报告2019》概要[J]. 中国脑血管病杂志，2020，17(5)：272-281

[3] SINGER J，GUSTAFSON D，CUMMINGS C，et al. Independent ischemic stroke risk factors in older Americans：a systematic review[J]. Aging，2019，11(10)：3392-3407

[4] LIN W Y，CHEN C H，TSENG Y J，et al. Predicting post-stroke activities of daily living through a machine learning-based approach on initiating rehabilitation[J]. Int J Med Inform，2018，111：159-164

[5] 黄志萍，苏 娅，陈秀红，等. 脑卒中患者日常生活活动能力及其影响因素[J]. 中国老年学杂志，2015，35(10)：2663-2665

[6] LANGHORNE P，BAYLAN S，EARLY SUPPORTED DISCHARGE TRIALISTS. Early supported discharge services for people with acute stroke[J]. Cochrane Database Syst Rev，2017，7(7)：CD000443

[7] VEERBEEK J M，KWAKKEL G，VAN WEGEN E E，et al. Early prediction of outcome of activities of daily living after stroke：a systematic review[J]. Stroke，2011，42(5)：1482-1488

[8] FAST L，TEMUULEN U，VILLRINGER K，et al. Machine learning-based prediction of clinical outcomes after first-ever ischemic stroke[J]. Front Neurol，2023，14：1114360

[9] NANAYAKKARA S，FOGARTY S，TREMEER M，et al. Characterising risk of in-hospital mortality following cardiac arrest using machine learning：a retrospective international registry study[J]. PLoS Med，2018，15(11)：e1002709

[10] 陈思玎，刘 欢，黄馨莹，等. 基于机器学习算法构建缺血性卒中3个月死亡预测模型研究[J]. 中国卒中杂志，2020，15(6)：595-599

[11] HEO J，YOON J G，PARK H，et al. Machine learning-based model for prediction of outcomes in acute stroke[J]. Stroke，2019，50(5)：1263-1265

[12] CAMPAGNINI S，ARIENTI C，PATRINI M，et al. Machine learning methods for functional recovery prediction and prognosis in post-stroke rehabilitation：a systematic review[J]. J Neuroeng Rehabil，2022，19(1)：54

[13] 宋华隆，牛博真，曹 颖. 针刺结合运动再学习康复疗法对脑卒中后神经功能恢复的疗效观察[J]. 中华中医药学刊，2024，42(1)：78-81

[14] KAMONO A，OGIHARA N. Weight-shift ability significantly correlates with walking velocity in post-acute stroke patients[J]. Proc Inst Mech Eng H，2018，232(4)：361-370

[15] 姜璧珺，陈 瑶，于培雁，等. 下肢评估训练康复系统提高脑卒中患者下肢平衡功能的疗效评估[J]. 中国康复医学杂志，2021，36(6)：728-731

[16] POLESE J C，ADA L，TEIXEIRA-SALMELA L F. Relationship between oxygen cost of walking and level of walking disability after stroke：an experimental study[J/OL]. Physiother Res Int，2018[2018-03-22]. DOI：10.1002/pri.1688

[17] 曲斯伟，朱 琳，严 莉，等. 基于改良分级运动想象的康复训练对脑卒中患者上肢功能改善的疗效分析[J]. 中国康复医学杂志，2023，9(8)：1084-1089

[18] LU Q，MÅRTENSSON J，ZHAO Y，et al. Living on the edge：family caregivers' experiences of caring for post-stroke family members in China：a qualitative study[J]. Int J Nurs Stud，2019，94：1-8

[19] GHAZIANI E，COUPPÉ C，SIERSMA V，et al. Electrical somatosensory stimulation in early rehabilitation of arm paresis after stroke：a randomized controlled trial[J]. Neu-

rorehabil Neural Repair, 2018, 32(10): 899-912
[20] KRUGER P C, EIKELBOOM J W, DOUKETIS J D, et al. Deep vein thrombosis: update on diagnosis and management[J]. Med J Aust, 2019, 210(11): 516-524
[21] 中国研究型医院学会血栓与止血专业委员会, 周 洲, 华 潞, 唐 宁, 等. D-二聚体实验室检测与临床应用中国专家共识[J]. 中华医学杂志, 2023, 103(35): 2743-2756
[22] 刘雅鑫, 蒋运兰, 刘芯君, 等. 脑卒中患者下肢深静脉血栓形成的预测因子: 基于风险预测模型的Meta分析[J]. 实用心脑肺血管病杂志, 2023.31(6): 97-101, 109
[23] 陈 静, 倪朝民, 吴 鸣, 等. 早期康复对重症监护病房脑出血患者下肢静脉血栓形成的预防作用[J]. 中华物理医学与康复杂志, 2020, 42(3): 220-222
[24] LIU L, ZHOU J, ZHANG Y, et al. A nomogram for individualized prediction of calf muscular vein thrombosis in stroke patients during rehabilitation: a retrospective study[J]. Clin Appl Thromb Hemost, 2022, 28: 10760296221117991
[25] 曹 莉, 袁良津, 石 力, 等. 血清白蛋白及球蛋白水平与急性缺血性脑卒中患者出院结局的关系[J]. 中国老年学杂志, 2016, 36(21): 5330-5332
[26] XIANG W, CHEN X, YE W, et al. Prognostic nutritional index for predicting 3-month outcomes in ischemic stroke patients undergoing thrombolysis[J]. Front Neurol, 2020, 11: 599

[收稿日期] 2024-01-04
(本文编辑: 戴王娟)

(上接第665页)

my[J]. J Cancer, 2021, 12(2): 305-315
[10] HU J, LI H, HE T, et al. A nomogram incorporating PD-L1, NLR, and clinicopathologic features to predict inguinal lymph node metastasis in penile squamous cell carcinoma[J]. Urol Oncol, 2020, 38(7): 641.e619-641.e629
[11] LI Z, LI X, ZHANG X, et al. Prognostic significance of common preoperative laboratory variables in penile squamous cell carcinoma[J]. Int J Urol, 2020, 27(1): 76-82
[12] NØST T H, ALCALA K, URBAROVA I, et al. Systemic inflammation markers and cancer incidence in the UK Biobank[J]. Eur J Epidemiol, 2021, 36(8): 841-848
[13] JAKOBSEN J K. Sentinel node methods in penile cancer - a historical perspective on development of modern concepts[J]. Semin Nucl Med, 2022, 52(4): 486-497
[14] PEYRAUD F, ALLENET C, GROSS-GOUPIL M, et al. Current management and future perspectives of penile cancer: an updated review[J]. Cancer Treat Rev, 2020, 90: 102087
[15] KROON B K, HORENBLAS S, LONT A P, et al. Patients with penile carcinoma benefit from immediate resection of clinically occult lymph node metastases[J]. J Urol, 2005, 173(3): 816-819
[16] SADEGHI R, GHOLAMI H, ZAKAVI S R, et al. Accuracy of ^{18}F-FDG PET/CT for diagnosing inguinal lymph node involvement in penile squamous cell carcinoma: systematic review and meta-analysis of the literature[J]. Clin Nucl Med, 2012, 37(5): 436-441
[17] ZEKAN D S, DAHMAN A, HAJIRAN A J, et al. Prognostic predictors of lymph node metastasis in penile cancer: a systematic review[J]. Int Braz J Urol, 2021, 47(5): 943-956
[18] WANG Q, ZHU S R, HUANG X P, et al. Prognostic value of systemic immune-inflammation index in patients with urinary system cancers: a meta-analysis[J]. Eur Rev Med Pharmacol Sci, 2021, 25(3): 1302-1310
[19] ALDAWSARI H M, GORAIN B, ALHAKAMY N A, et al. Role of therapeutic agents on repolarisation of tumour-associated macrophage to halt lung cancer progression[J]. J Drug Target, 2020, 28(2): 166-175
[20] QUE H, FU Q, LAN T, et al. Tumor-associated neutrophils and neutrophil-targeted cancer therapies[J]. Biochim Biophys Acta Rev Cancer, 2022, 1877(5): 188762
[21] PAIJENS S T, VLEDDER A, DE BRUYN M, et al. Tumor-infiltrating lymphocytes in the immunotherapy era[J]. Cell Mol Immunol, 2021, 18(4): 842-859
[22] OLINGY C E, DINH H Q, HEDRICK C C. Monocyte heterogeneity and functions in cancer[J]. J Leukoc Biol, 2019, 106(2): 309-322
[23] TAO D L, TASSI YUNGA S, WILLIAMS C D, et al. Aspirin and antiplatelet treatments in cancer[J]. Blood, 2021, 137(23): 3201-3211
[24] FERRO M, MUSI G, SERINO A, et al. Neutrophil, platelets, and eosinophil to lymphocyte ratios predict gleason score upgrading in low-risk prostate cancer patients[J]. Urol Int, 2019, 102(1): 43-50

[收稿日期] 2023-11-13
(本文编辑: 唐 震)