

基于机器学习算法构建老年患者肺功能智能评估模型

陈翠妍^{1,2}, 张文娟^{1,2}, 蔡依琳^{1,2}, 欧阳媚^{1,2}, 梁叶^{2,3}, 梁铭标², 梁会营^{1,2}

¹南方医科大学公共卫生学院, 广东 广州 510515; ²南方医科大学附属广东省人民医院(广东省医学科学院), 广东 广州 510000; ³广东省心血管病研究所, 广东 广州 510000

摘要:目的 利用常规临床资料, 构建接受度高、普适性广的肺功能智能评估模型, 为老年患者肺功能异常的早期识别与诊疗提供参考。方法 选取2020年9月1日~2024年12月31日于广东省人民医院完成肺功能检查(PFT)的7720例老年患者, 整合临床资料, 经单因素分析及LASSO回归筛选变量, 按8:2比例随机分为训练集($n=6176$)和测试集($n=1544$), 分别构建线性回归、随机森林、极端梯度提升(XGBoost)、梯度提升机及支持向量回归模型以预测用力肺活量(FVC)和1秒用力呼气量(FEV₁), 采用Pearson相关系数(r)、组内相关系数(ICC)等评估性能, 确定最优模型。结合沙普利加性解释方法分析特征贡献, 并评估模型对呼吸高危人群的分类性能。结果 XGBoost为最优模型, FVC和FEV₁的 r 值分别为0.733、0.711, ICC为0.694、0.668。性别与年龄为核心预测变量, 男性与低龄对肺功能指标呈正向效应, 女性与高龄呈负向效应。FVC%、FEV₁%最佳截断值分别为82.93%、85.46%, AUC为0.700、0.706。结论 本研究构建的模型简便易行、适用性强, 可辅助老年患者肺功能评估, 具备良好的临床应用前景。

关键词:机器学习; 老年人; 肺功能; 评估模型; 极端梯度提升

Development of a machine learning-based model for assessing pulmonary function in elderly patients

CHEN Cuiyan^{1,2}, ZHANG Wenxian^{1,2}, CAI Yilin^{1,2}, OUYANG Mei^{1,2}, LIANG Ye^{2,3}, LIANG Mingbiao², LIANG Huiying^{1,2}

¹School of Public Health, Southern Medical University, Guangzhou 510515, China; ²Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510000, China; ³Guangdong Provincial Cardiovascular Institute, Guangzhou 510000, China

Abstract: Objective To develop a machine learning-based model with good generalization ability for pulmonary function assessment in elderly patients with impaired pulmonary function. **Methods** A total of 7720 elderly patients undergoing pulmonary function testing (PFT) in Guangdong Provincial People's Hospital between September 1, 2020 and December 31, 2024 were enrolled. After variable screening by univariate analysis and LASSO regression, the patients were randomly assigned in an 8:2 ratio to the training set ($n=6176$) and test set ($n=1544$). Five regression models were constructed for predicting forced vital capacity (FVC) and forced expiratory volume in one second (FEV₁), and their performance was evaluated based on Pearson correlation coefficient (r) and intraclass correlation coefficient (ICC). Feature contributions were interpreted using Shapley additive explanations, and the classification performance of the models for high-risk populations was assessed. **Results** XGBoost showed the best predictive performance in the elderly patients ($r=0.733$ for FVC and 0.711 for FEV₁; ICC=0.694 for FVC and 0.668 for FEV₁). Male gender and a younger age showed positive effects on pulmonary function, while female gender and an advanced age showed negative effects. The optimal cut-off values of FVC% and FEV₁% were 82.93% and 85.46%, with AUC of 0.700 and 0.706, respectively. **Conclusion** The XGBoost-based model developed in this study is simple and reliable with good generalization ability for pulmonary function assessment in elderly patients.

Keywords: machine learning; elderly; pulmonary function; model evaluation; extreme gradient boosting

根据2020年第7次人口普查数据, 我国≥60岁老年人口已达2.64亿, 占总人口18.7%^[1]。随着人口老龄化加剧, 老年人肺功能水平随年龄增长自然下降, 不仅与各类呼吸系统疾病密切相关, 还会增加心血管疾病、阿尔兹海默症及脑卒中等多种疾病的患病风险, 甚至与全因死亡风险呈负相关, 肺功能水平越低, 全因死亡率越高, 严重威胁老年居民健康^[2-5]。

肺功能检查(PFT)是评估肺功能水平的重要手段, 已被证实与改善患者预后、降低全因死亡率相关^[6]。中

华医学会明确建议, 将PFT纳入≥60岁人群的定期体检项目, 对肺功能减退的老年人群, 每3~6个月进行1次监测筛查^[7]。但临床实践中, PFT对受试者配合度要求较高, 部分老年患者常因自身疾病、认知障碍等原因而导致数据缺失, 给肺功能评估带来了挑战^[6, 8]。

机器学习作为人工智能核心分支, 凭借参数优化与特征工程优势, 已在医学领域展现出良好的应用前景^[9]。现有研究虽在特定呼吸系统疾病的诊断预测方面取得一定进展, 但针对老年患者的肺功能评估研究仍较为匮乏^[10, 11]。基于此, 本研究以用力肺活量(FVC)和1s用力呼气量(FEV₁)为目标, 整合易获取的临床指标, 利用机器学习算法构建接受度高、普适性广的肺功能评估模型, 为老年患者肺功能异常的早期识别、及时干预及个

收稿日期: 2026-02-13

基金项目: 广东省科技计划项目重点领域研发计划(2025B0101120008)

作者简介: 陈翠妍, 在读硕士研究生, E-mail: 17675629838@163.com

通信作者: 梁会营, 教授, 博士生导师, E-mail: lianghuiying@gdph.org.cn

体化诊疗决策提供参考。

1 资料和方法

1.1 研究对象

选取2020年9月1日~2024年12月31日在广东省人民医院进行PFT的7720例老年患者作为研究对象。根据中华医学会呼吸病学分会发布的《肺功能检查指南(第二部分)——肺量计检查》,肺量计对FVC和FEV₁的标准测量量程为0.5~8 L^[12]。纳入标准:年龄≥60岁;行PFT前未使用支气管扩张剂。排除标准:FVC或FEV₁检测数据缺失者;FVC或FEV₁超出仪器标准测量范围者。本研究已通过广东省人民医院伦理委员会审批(伦理批号:KY2024-1206-01)。

1.2 研究方法

1.2.1 肺功能检查 由经验丰富的技术人员严格遵循美国胸科学会和欧洲呼吸学会的相关标准规范操作^[13],采用Cosmed K4b2肺功能仪进行检测。呼吸高危人群的定义标准为:使用支气管扩张剂前FVC占正常预计值的百分比(FVC%)或FEV₁占正常预计值的百分比(FEV₁%)低于80%^[14]。

1.2.2 数据收集及预处理 从电子病历中收集研究对象的数据:基本信息:年龄、性别、婚姻状况、体质量指数(BMI)、吸烟史、饮酒史;疾病相关因素:住院次数、糖尿病、脑卒中、高血压、冠状动脉粥样硬化(CAD)、肺动脉高压以及呼吸系统疾病(其中包括肺部感染、肺炎、支气管炎、哮喘、慢性阻塞性肺病、支气管扩张、间质性肺炎、肺结核、肺气肿、肺不张和肺癌);实验室检查:中性粒细胞比例(NEUT%)、乳酸脱氢酶(LDH)、凝血酶原时间(PT)、凝血酶时间(TT)、嗜酸性粒细胞百分比(EOS%)、大血小板比率(P-LCR)、氯(Cl)、活化部分凝血活酶时间(APTT)、红细胞分布宽度变异系数(RDW-CV)、红细胞分布宽度标准差(RDW-SD)、红细胞压积(HCT)、红细胞计数(RBC)、葡萄糖(Glu)、血小板计数(PLT)、血浆纤维蛋白原(FIB)、血红蛋白(HGB)、钙(Ca)、钠(Na)、钾(K)。对缺失比例≥30%的临床变量进行删除,缺失比例<30%的变量使用多重插补法进行填补(表1)。

1.2.3 变量筛选与模型构建 首先通过单因素分析初步筛选预测变量,然后采用LASSO回归进一步筛选变量,通过十折交叉验证确定最优惩罚参数,以减少多重共线性。按照8:2的比例将数据集随机划分为训练集(6176例)和测试集(1544例)。使用线性回归(LR)、随机森林(RF)、极端梯度提升(XGBoost)、梯度提升机(GBM)和支持向量回归(SVR)构建筛查模型。采用Pearson相关系数(r)、组内相关系数(ICC)、均方误差(MSE)、均方根误差(RMSE)和平均绝对误差(MAE)评估模型的预测性能并筛选出最优模型。同时比较完整数据集与多重插补后数据集的模型性能差异,以评估缺失数据处理方

表1 变量缺失情况统计

Tab.1 Statistics of missing values of the variables

Variable	Missing (%)	Variable	Missing (%)
FT4	47.14 (deleted)	Number of hospitalizations	0.99
D-Dimer	37.98 (deleted)	TT	0.69
CK	37.11 (deleted)	APTT	0.69
Mg	35.85 (deleted)	FIB	0.69
Drinking	27.82	K	0.33
Smoking	27.72	Glu	0.28
PT	18.85	RDW-SD	0.27
Cl	18.10	NEUT%	0.15
Ca	18.10	EOS%	0.15
Na	18.10	RDW-CV	0.15
LDH	17.75	RBC	0.15
HCT	17.73	Marriage	0.14
HGB	17.73	BMI	0.05
P-LCR	12.13		

FT4: Free thyroxine; CK: Creatine kinase; PT: Prothrombin time; LDH: Lactate dehydrogenase; HCT: Hematocrit; HGB: Hemoglobin; P-LCR: Platelet-large cell ratio; TT: Thrombin time; APTT: Activated partial thromboplastin time; FIB: Fibrinogen; Glu: Glucose; RDW-SD: Red cell distribution width-standard deviation; PLT: Platelets; NEUT%: Neutrophil percentage; EOS%: Eosinophil percentage; RDW-CV: Red cell distribution width-coefficient of variation; RBC: Red blood cell count; BMI: Body mass index.

法对模型预测结果的影响。使用沙普利加性解释(SHAP)方法分析预测变量对最优模型的影响。此外,通过确定最佳截断值,比较模型预测值与真实测量值的受试者工作特征曲线(ROC)下面积(AUC)、敏感度和特异度,评估模型对呼吸高危人群的分类效能,并与临床标准截断值进行对比。

1.2.4 统计学分析 根据FVC和FEV₁的四分位数,将研究对象分为4个水平组进行基线特征分析。其中,连续变量以均数±标准差描述,组间比较采用t检验或方差分析,分类变量使用频数(百分比)描述,使用卡方检验或Fisher确切概率法比较组间差异。 $P<0.05$ 为差异有统计学意义,所有统计分析均使用R4.5.1进行。

2 结果

2.1 研究对象的基线特征及单因素分析

本研究共纳入7720例研究对象,其中,男性4197例(54.4%),女性3523例(45.6%);年龄60~96(68.56±6.27)岁。肺功能指标方面,FVC为0.5~5.19 L,2.45±0.69 L;FEV₁为0.5~3.93 L,1.99±0.54 L。单因素分析结果显示,FVC中共计28个变量的差异有统计学意义,FEV₁中27个变量的差异有统计学意义(表2、3)。

2.2 预测变量的LASSO回归筛选

LASSO回归结果显示,FVC的 $\lambda_{\min}$ 和 λ_{lse} 分别为

表2 根据FVC四分位数划分的基线特征
Tab.2 Baseline characteristics of the participants stratified by the quartiles of forced vital capacity (FVC)

Variable	Overall	Q1	Q2	Q3	Q4	P
Sample (n)	7720	1968	1911	1937	1904	
FVC (L)	2.45±0.69	1.6±0.28	2.19±0.13	2.66±0.15	3.36±0.34	<0.001
Age (year)	68.56±6.27	71.25±6.83	68.68±6.20	68.01±5.76	66.22±5.03	<0.001
Gender (men)	4197 (54.4%)	354 (18.0%)	688 (36.0%)	1339 (69.1%)	1816 (95.4%)	<0.001
BMI (kg/m ²)	25.09±7.38	25.1±7.63	25.17±7.65	25.04±7.15	25.04±7.09	0.94
Smoking	1762 (22.8%)	171 (8.7%)	324 (17%)	546 (28.2%)	721 (37.9%)	<0.001
Drinking	551 (7.1%)	46 (2.3%)	97 (5.1%)	158 (8.2%)	250 (13.1%)	<0.001
Marital status						<0.001
Never married	43 (0.6%)	15 (0.8%)	13 (0.7%)	8 (0.4%)	7 (0.4%)	
Married	7127 (92.3%)	1708 (86.8%)	1751 (91.6%)	1829 (94.4%)	1839 (96.6%)	
Divorced/seperated/widowed	550 (7.1%)	245 (12.4%)	147 (7.7%)	100 (5.2%)	58 (3.0%)	
Number of hospitalizations	1.81±2.37	1.81±2.22	1.73±1.98	1.89±2.93	1.81±2.22	0.22
Disease of respiratory system	626(8.1%)	157 (8.0%)	135 (7.1%)	152 (7.8%)	182 (9.6%)	0.04
Hypertention	2559 (33.1%)	731 (37.1%)	622 (32.5%)	644 (33.2%)	562 (29.5%)	<0.001
Diabetes	1189 (15.4%)	339 (17.2%)	286 (15.0%)	301 (15.5%)	263 (13.8%)	0.03
Stroke	392 (5.1%)	137 (7.0%)	86 (4.5%)	100 (5.2%)	69 (3.6%)	<0.001
Pulmonary arterial hypertension	271 (3.5%)	158 (8.0%)	57 (3.0%)	35 (1.8%)	21 (1.1%)	<0.001
CAD	837 (10.8%)	236 (12.0%)	199 (10.4%)	209 (10.8%)	193 (10.1%)	0.26
NEUT (%)	70.47±8.52	71.2±9.04	70.84±8.1	70.08±8.64	69.73±8.16	<0.001
LDH (U/L)	243.67±412.54	269.88±303.08	237.51±249.84	236.37±430.66	230.19±585.7	0.01
PT (s)	14.32±2.55	15.11±3.56	14.16±2.25	14.04±1.97	13.96±1.85	<0.001
TT (s)	18.6±8.47	19.93±11.69	18.29±7.56	18.31±7.32	17.83±5.95	<0.001
EOS (%)	2.1±2.03	1.92±1.93	1.92±1.83	2.24±2.34	2.33±1.94	<0.001
P-LCR (%)	27.64±24.86	28.19±9.05	28.44±47.8	26.95±8.08	26.96±7.81	0.12
Cl(mmol/L)	105.09±2.98	104.92±3.21	105.29±3.01	105.09±2.82	105.05±2.86	0.002
APTT (s)	38.9±7.35	40.27±9.51	38.3±6.85	38.55±6.39	38.43±5.87	<0.001
RDW-CV	0.14±0.02	0.14±0.03	0.14±0.02	0.14±0.02	0.14±0.02	<0.001
RDW-SD (fl)	44.9±5.62	45.68±6.18	44.57±5.58	44.76±5.25	44.58±5.33	<0.001
HCT (%)	36.27±5.05	34.55±5.00	35.78±4.76	36.79±4.92	38.01±4.86	<0.001
RBC (×10 ¹² /L)	4.03±0.64	3.87±0.64	3.99±0.62	4.08±0.64	4.19±0.63	<0.001
Glu (mmol/L)	6.3±1.85	6.54±2.02	6.33±1.82	6.22±1.81	6.08±1.69	<0.001
PLT (×10 ⁹ /L)	212.03±71.58	212.43±75.93	215.59±73.07	212.8±70.48	207.24±66.14	0.004
FIB (g/L)	3.74±1.00	3.87±0.99	3.75±0.99	3.73±1.00	3.6±0.98	<0.001
HGB (g/L)	119.59±17.28	113.18±16.87	117.71±16.14	121.49±16.8	126.17±16.57	<0.001
Ca (mmol/L)	2.23±0.12	2.23±0.12	2.23±0.12	2.23±0.11	2.22±0.12	0.04
Na (mmol/L)	139.21±3.19	139.36±3.25	139.34±3.36	139.17±2.95	138.97±3.15	<0.001
K (mmol/L)	3.85±0.46	3.87±0.40	3.83±0.44	3.83±0.39	3.86±0.59	0.007

FVC: Forced vital capacity; BMI: Body mass index; NEUT%: Neutrophil percentage; LDH: Lactate dehydrogenase; PT: Prothrombin time; TT: Thrombin time; EOS%: Eosimophil; P-LCR: Platelet-large cell ratio; Cl: Chloride; APTT: Activated partialthromboplastin time; RDW-CV: Red cell distribution width-coefficient of variation; RDW-SD: Red cell distribution width-standard deviation; HCT: Hematocrit; RBC: Red blood cell count; Glu: Glucose; PLT: Platelets; FIB: Fibrinogen; HGB: Hemoglobin; Ca: Calcium; Na: Sodium; K: Kalium

0.0013 和 0.0191, 分别筛选出 27 个和 14 个变量;FEV₁ 的 $\lambda_{\min}$ 和 λ_{lse} 分别为 0.0004 和 0.0099, 分别筛选出 29 个和 16 个变量。为平衡模型拟合度与复杂度,避免过拟合,本研究基于 λ_{lse} 筛选变量,最终FVC的预测变量为年龄、BMI、性别、呼吸系统疾病、肺动脉高压、脑卒中、PT、Na、Cl、APTT、RDW-CV、FIB、HGB 和 K,FEV₁ 的预测

表3 根据FEV₁四分位数划分的基线特征表

Tab.3 Baseline characteristics of the participants stratified by the quartiles of forced expiratory volume in 1 second (FEV₁)

Variable	Overall	Q1	Q2	Q3	Q4	P
Sample	7720	1971	1932	1904	1913	
FEV ₁ (L)	1.99±0.54	1.32±0.24	1.8±0.1	2.16±0.11	2.69±0.26	<0.001
Age (year)	68.56±6.27	71.82±6.75	68.82±6.17	67.59±5.56	65.91±4.83	<0.001
Gender (men)	4197 (54.4%)	474(24%)	728 (37.7%)	1193 (62.7%)	1802 (94.2%)	<0.001
BMI (kg/m ²)	25.09±7.38	24.81±7.5	25.1±7.42	25.23±7.48	25.22±7.12	0.24
Smoking	1762 (22.8%)	254 (12.9%)	379(19.6%)	462 (24.3%)	667 (34.9%)	<0.001
Drinking	551 (7.1%)	63 (3.2%)	100 (5.2%)	132 (6.9%)	256 (13.4%)	<0.001
Marital status						<0.001
Never married	43 (0.6%)	16 (0.8%)	13 (0.7%)	4 (0.2%)	10 (0.5%)	
Married	7127 (92.3%)	1716 (87.1%)	1759 (91%)	1799 (94.5%)	1853 (96.9%)	
Divorced/seperated/widowed	550 (7.1%)	239 (12.1%)	160 (8.3%)	101 (5.3%)	50 (2.6%)	
Number of hospitalizations	1.81±2.37	1.87±2.23	1.77±2.13	1.84±2.53	1.76±2.56	0.43
Disease of respiratory system	626 (8.1%)	212 (10.8%)	151 (7.8%)	125 (6.6%)	138 (7.2%)	<0.001
Hypertention	2559 (33.1%)	713 (36.2%)	645 (33.4%)	621 (32.6%)	580 (30.3%)	0.002
Diabetes	1189 (15.4%)	317 (16.1%)	295(15.3%)	294 (15.4%)	283 (14.8%)	0.74
Stroke	392 (5.1%)	132 (6.7%)	98 (5.1%)	98 (5.1%)	64 (3.3%)	<0.001
Pulmonary arterial hypertension	271 (3.5%)	150 (7.6%)	67 (3.5%)	30 (1.6%)	24 (1.3%)	<0.001
CAD	837 (10.8%)	238 (12.1%)	205 (10.6%)	200 (10.5%)	194 (10.1%)	0.22
NEUT (%)	70.47±8.52	71.11±9	70.74±8.48	70.26±8.39	69.73±8.1	<0.001
LDH(U/L)	243.67±412.54	269.32±327.14	251.35±393.6	225.9±301.33	227.18±573.27	0.002
PT (s)	14.32±2.55	15.03±3.5	14.22±2.33	13.99±1.93	14.02±1.94	<0.001
TT (s)	18.6±8.47	19.73±11.39	18.24±6.98	18.37±7.93	18.03±6.52	<0.001
EOS (%)	2.1±2.03	2.05±2.04	1.95±2.2	2.11±1.88	2.3±1.96	<0.001
P-LCR (%)	27.64±24.86	29.14±47.23	26.98±8.07	27.14±8.15	27.25±7.86	0.02
Cl (mmol/L)	105.09±2.98	104.84±3.2	105.25±3.04	105.18±2.78	105.09±2.88	<0.001
APTT (s)	38.9±7.35	40.38±9.38	38.4±6.8	38.35±6.37	38.41±6.11	<0.001
RDW-CV	0.14±0.02	0.14±0.03	0.14±0.02	0.14±0.02	0.14±0.02	<0.001
RDW-SD (fl)	44.9±5.62	45.86±6.32	44.72±5.56	44.62±5.19	44.4±5.19	<0.001
HCT (%)	36.27±5.05	34.59±5.08	35.77±4.85	36.73±4.88	38.06±4.73	<0.001
RBC (×10 ¹² /L)	4.03±0.64	3.86±0.65	4±0.64	4.06±0.62	4.21±0.62	<0.001
Glu (mmol/L)	6.3±1.85	6.46±1.99	6.4±1.82	6.18±1.8	6.14±1.76	<0.001
PLT (×10 ⁹ /L)	212.03±71.58	212.39±77.02	218.53±74.39	212.34±67.7	204.77±65.79	<0.001
FIB (g/L)	3.74±1	3.88±0.99	3.79±1.02	3.7±0.97	3.58±0.98	<0.001
HGB (g/L)	119.59±17.28	113.27±17.11	117.66±16.38	121.39±16.72	126.26±16.17	<0.001
Ca (mmol/L)	2.23±0.12	2.23±0.12	2.23±0.11	2.23±0.12	2.22±0.11	0.29
Na (mmol/L)	139.21±3.19	139.22±3.18	139.36±3.42	139.23±2.9	139.03±3.22	0.01
K (mmol/L)	3.85±0.46	3.87±0.4	3.83±0.4	3.83±0.42	3.86±0.6	0.01

FEV₁: Forced expiratory volume in the first second; BMI: Body mass index; NEUT%: Neutrophil percentage; LDH: Lactate dehydrogenase; PT: Prothrombin time; TT: Thrombin time; EOS%: Eosinophil; P-LCR: Platelet-large cell ratio; Cl: chloride; APTT: Activated partialthromboplastin time; RDW-CV: Red cell distribution width-coefficient of variation; RDW-SD: Red cell distribution width-standard deviation; HCT: Hematocrit; RBC: Red blood cell count; Glu: Glucose; PLT: Platelets; FIB: Fibrinogen; HGB: Hemoglobin; Ca: Calcium; Na: Sodium; K: Kalium.

变量为年龄、性别、吸烟史、呼吸系统疾病、肺动脉高压、脑卒中、PT、EOS%、P-LCR、Cl、APTT、RDW-CV、RDW-SD、PLT、FIB和HGB(图1)。

2.3 5种机器学习模型比较

在插补后数据集中构建5种机器学习模型,结果显示 XGBoost对 FVC 和 FEV₁的预测性能最佳,r值分别

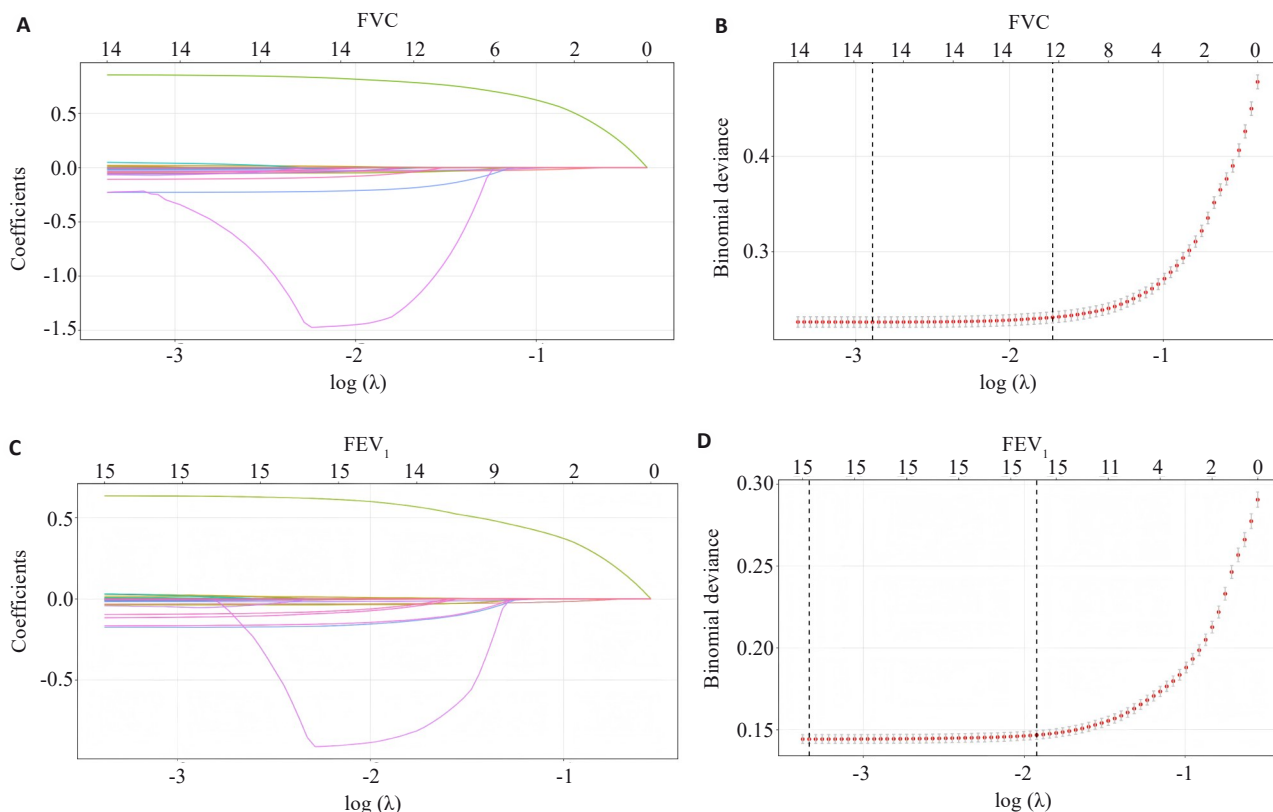


图1 预测变量的LASSO回归筛选过程和最佳参数 λ 的选择

Fig. 1 LASSO regression screening of predictive variables and selection of the optimal parameters λ . A: LASSO regression coefficient trajectory plot for FVC. B: Binomial deviance plot for optimal λ selection in LASSO regression (FVC). C: LASSO regression coefficient trajectory plot for FEV₁. D: Binomial deviance plot for optimal λ selection in LASSO regression (FEV₁).

为0.733和0.711, ICC值为0.694和0.668, MSE值为0.210L²和0.136L², RMSE值为0.459L和0.369L, MAE值为0.364L和0.291L(表4、图2)。

2.4 完整数据集与插补后数据集模型性能对比

结果显示,完整数据集下所有模型的r和ICC均略有降低, MSE、RMSE、MAE等误差指标较插补后数据集有所升高,提示多重插补对模型预测性能有所改善。各模型的预测性能差异幅度较小,且XGBoost在两种数据集中均保持了接近最优的表现,表明本研究基于插补后数据集构建的模型结论具有一定稳健性(表5)。

2.5 XGBoost的SHAP解释

采用SHAP方法对最优XGBoost模型进行可解释性分析。变量重要性排序显示(图3A、B),性别、年龄是影响FVC与FEV₁预测的共同核心因素。在FVC预测中,贡献度前10位的变量依次为性别、年龄、PT、HGB、RDW-CV、FIB、BMI、Cl、K、Na;在FEV₁预测中,前10位变量则为性别、年龄、HGB、FIB、PT、吸烟史、RDW-CV、Cl、APTT、呼吸系统疾病,提示FVC与FEV₁的影响因素既有共性,也存在特异性差异。蜂群图直观展示了各变量的影响方向与强度(图3C、D),其中男性均表现为显著正向影响,女性则表现为显著负向影响;年龄呈明确的负向影响,随着年龄升高,对FVC与FEV₁预测值的负

向影响持续增强。单样本SHAP力图进一步展示了变量对个体预测值的作用方向与幅度(图4)。

2.6 XGBoost模型对呼吸高危人群的分类性能

基于XGBoost模型分析得出,FVC%与FEV₁%的最佳截断值分别为82.93%、85.46%,对应AUC为0.700、0.706。进一步以临床标准截断值进行对照评估,结果显示,FVC%和FEV₁%的AUC均为0.688。整体对比可见,相较于临床截断值,模型最优截断值下的区分效能更佳,整体分类表现更稳定(表6、图5)。

3 讨论

PFT是临床评估肺功能的重要手段,但其检测对受试者配合度要求较高,老年患者常因认知、体力、疾病等因素难以完成规范检查,易导致数据缺失,进而阻碍肺功能评估^[6,8]。为此,本研究基于临床常规易获取的指标,使用机器学习算法构建接受度高、普适性好的肺功能评估模型,为临床早期干预与诊疗决策提供参考依据。

在5种机器学习算法中,XGBoost模型表现出最优预测效能。XGBoost模型基于决策树的集成学习框架,具有良好的泛化能力与拟合性能,已在疾病诊断、分型及预后预测等复杂临床任务中展现出优异的应用价

表4 5种模型预测FVC和FEV₁的性能比较(插补后数据集)

Tab.4 Performance comparison of the 5 models for predicting FVC and FEV₁ (imputed dataset)

Model	<i>r</i>	ICC	MSE	RMSE	MAE
FVC					
LR	0.728	0.695	0.212	0.461	0.365
RF	0.727	0.691	0.213	0.462	0.366
XGBoost	0.733	0.694	0.210	0.459	0.364
GBM	0.728	0.692	0.211	0.460	0.364
SVR	0.721	0.694	0.218	0.467	0.368
FEV ₁					
LR	0.709	0.666	0.141	0.376	0.292
RF	0.701	0.659	0.140	0.374	0.293
XGBoost	0.711	0.668	0.136	0.369	0.291
GBM	0.708	0.665	0.142	0.380	0.296
SVR	0.688	0.662	0.146	0.382	0.297

ICC: Intraclass correlation coefficient; MSE: Mean squared error; RMSE: Root mean square error; MAE: Mean absolute error; LR: Linear regression; RF: Random forest; XGBoost: Extreme gradient boosting; GBM: Gradient boosting machine; SVR: Support vector regression.

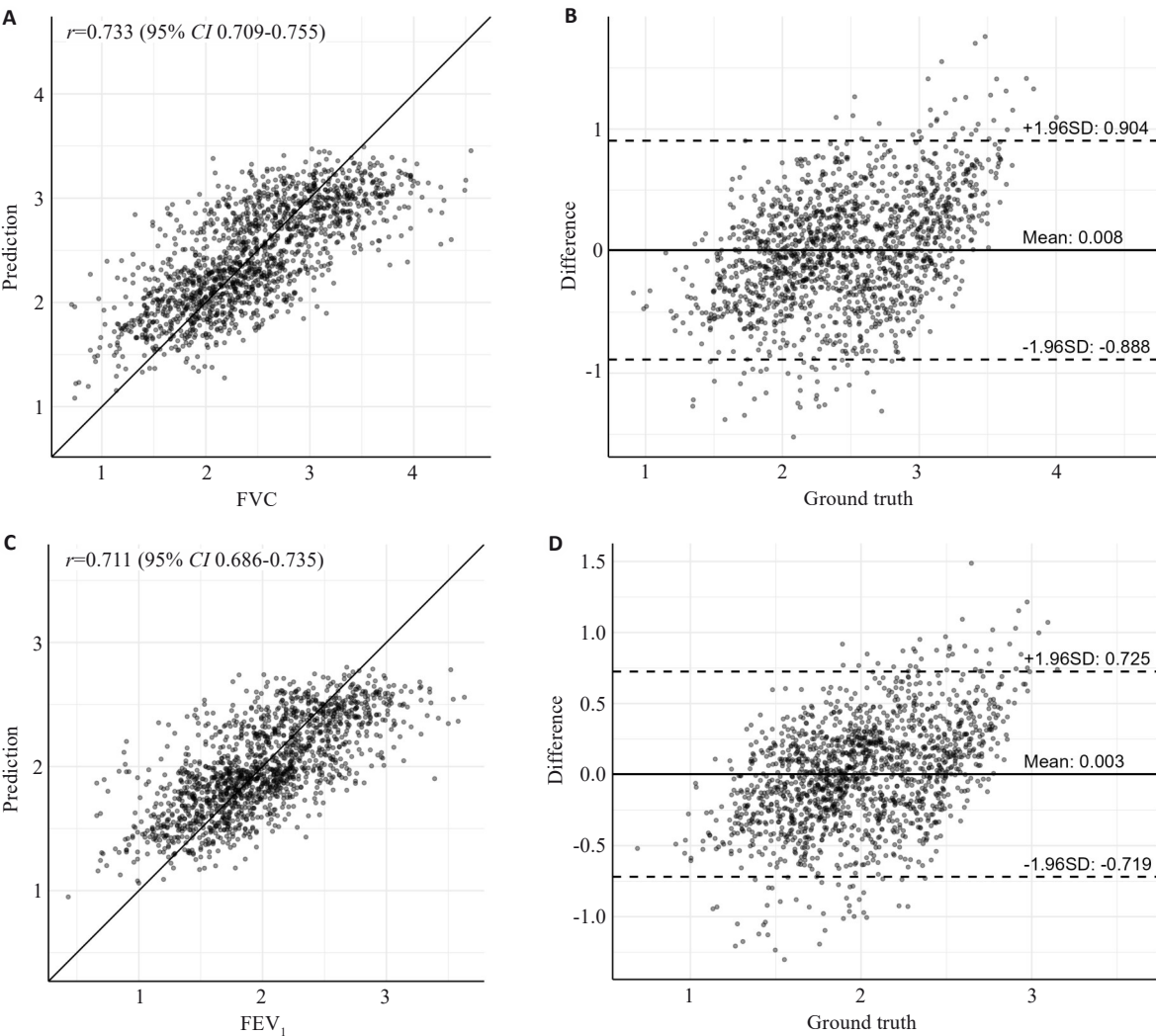


图2 XGBoost预测FVC和FEV₁的性能

Fig.2 Performance of XGBoost model for predicting FVC and FEV₁. **A:** Scatter plot of predicted vs measured FVC. **B:** Bland-Altman plot of FVC predictions. **C:** Scatter plot of predicted vs measured FEV₁. **D:** Bland-Altman plot of FEV₁ predictions.

表5 5种模型预测FVC和FEV₁的性能比较(完整数据集)
Tab.5 Performance comparison of the 5 models for predicting FVC and FEV₁ (complete dataset)

Model	<i>r</i>	ICC	MSE	RMSE	MAE
FVC					
LR	0.677	0.639	0.264	0.514	0.399
RF	0.674	0.595	0.268	0.518	0.403
XGBoost	0.668	0.621	0.269	0.519	0.403
GBM	0.671	0.628	0.268	0.517	0.402
SVR	0.659	0.636	0.281	0.530	0.410
FEV ₁					
LR	0.695	0.655	0.149	0.386	0.304
RF	0.679	0.634	0.155	0.394	0.312
XGBoost	0.683	0.633	0.154	0.39	0.311
GBM	0.685	0.639	0.153	0.391	0.308
SVR	0.700	0.668	0.147	0.384	0.304

FVC: Forced vital capacity; FEV₁: Forced expiratory volume in the first second; ICC: Intraclass correlation coefficient; MSE: Mean squared error; RMSE: Root mean square error; MAE: Mean absolute error; LR: Linear regression; RF: Random forest; XGBoost: eXtreme gradient boosting; GBM: Gradient boosting machine; SVR: Support vector regression.

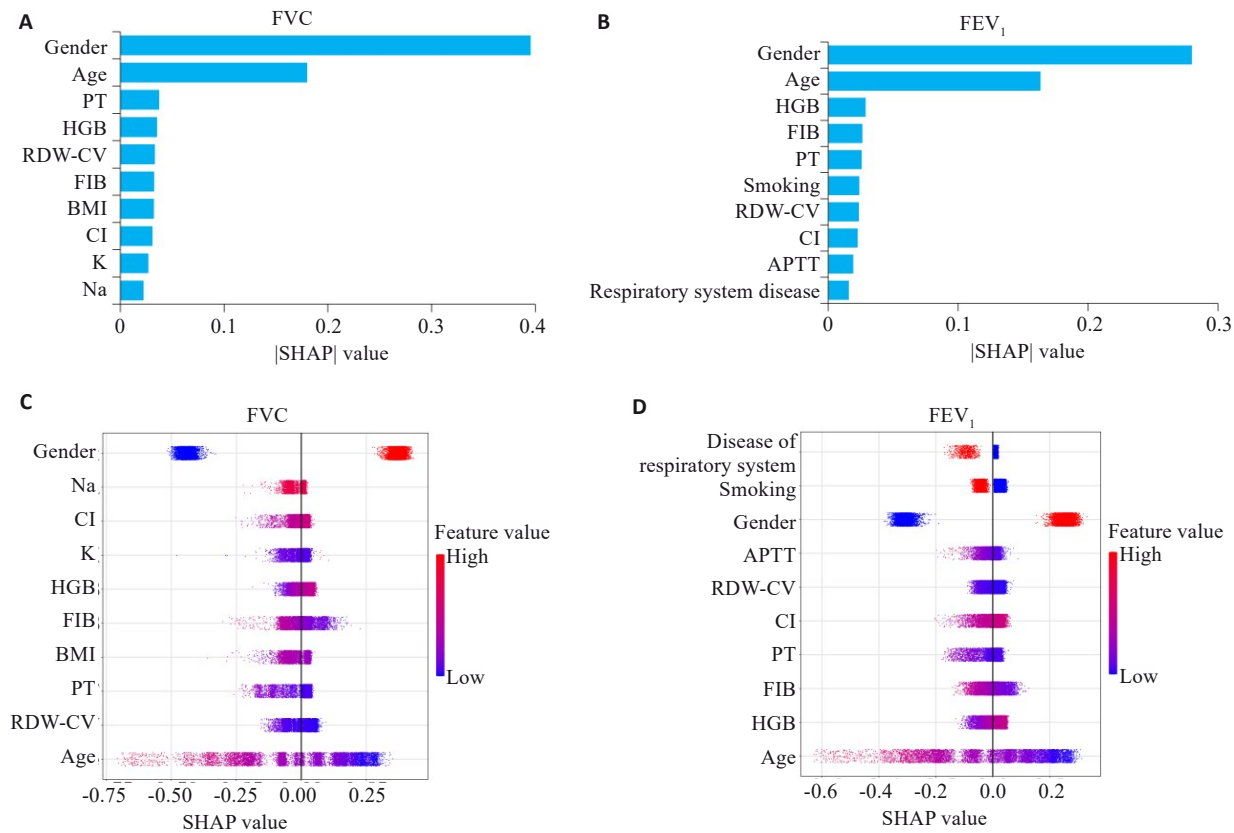


图3 XGBoost模型的特征重要性及蜂群图
Fig.3 Feature importance and beeswarm plots of the XGBoost model. A: Feature importance bar plot of FVC. B: Feature importance bar plot of FEV₁. C: Beeswarm plots of FVC. D: Beeswarm plots of FEV₁.

值^[15-17]。为提升模型的可解释性,本研究采用SHAP算法进行特征分析。其中,男性对FVC及FEV₁均呈正向影响,这可能与男性在肺部体积、气道管径及胸廓结构上所具备的天然解剖生理优势密切相关^[18]。年龄对肺功能的影响呈现负向趋势,高龄个体的肺功能预测值更低,其机制可能与肺功能的生理性衰退有关,包括肺组织弹性下降、肺泡表面积减少及呼吸肌力减弱等,这些变化共同削弱肺部的防御与代偿能力^[19, 20];也可能与环

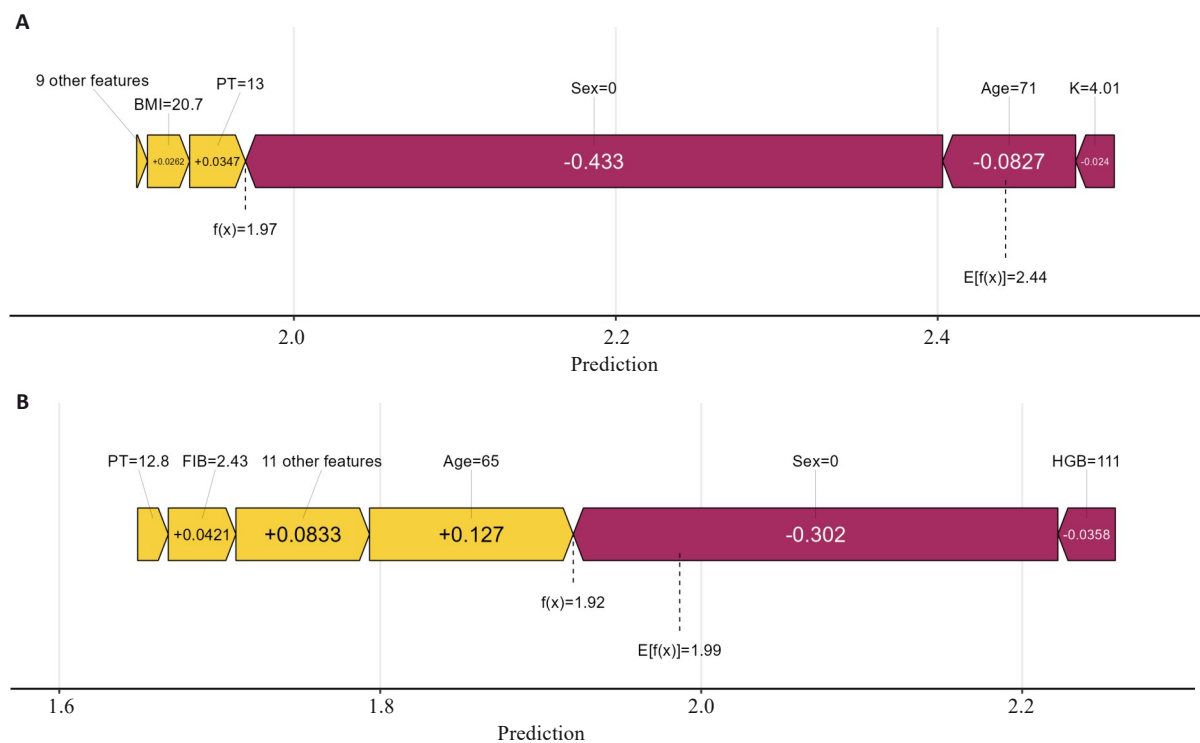


图4 XGBoost模型预测FVC和FEV₁的SHAP力图(单个样本解释)
Fig.4 SHAP force plot of XGBoost model for predicting FVC and FEV₁ (single sample explanation). **A**: SHAP contribution decomposition for FVC prediction. **B**: SHAP contribution decomposition for FEV₁ prediction.

表6 XGBoost的风险预测性能
Tab.6 Risk prediction performance of XGBoost

Index	Cut-off values (%)	Sensitivity	Specificity	AUC
Clinical cut-off value				
FVC%	80	0.540	0.740	0.688
FEV ₁ %	80	0.440	0.809	0.688
Model cut-off value				
FVC%	82.93	0.645	0.645	0.700
FEV ₁ %	85.46	0.651	0.651	0.706

境暴露的累积效应有关,随着年龄增长,个体接触烟草、职业粉尘及空气污染物等有害因素的总量增加,导致气道与肺实质的慢性进行性损伤^[20, 21]。同时,凝血指标(PT、APTT、FIB)也有重要的预测价值。从病理机制来看,血液高凝状态由肺微小血管炎性损伤、血管内皮受损、血小板活化及慢性缺氧等多种因素诱发,而这种高凝状态会进一步影响肺气体交换,加重肺功能损伤^[22, 23]。多项研究证实,哮喘患者的凝血功能指标与肺功能指标存在显著相关性^[24, 25]。此外,FEV₁水平与呼吸系统疾病密切相关。现有研究已表明,哮喘、慢性阻塞性肺疾病、肺炎等呼吸系统疾病均可显著影响FEV₁水平^[26, 27]。上述结果反映了不同肺功能指标的病理生理特征,可为临床开展个体化肺功能评估提供可解释性依据。

本研究进一步评估了XGBoost模型对呼吸高危人

群的分类效能。结果显示模型具备一定的高危人群识别能力,但整体效能处于中等水平,尚未达到高精度筛查与鉴别要求。此外,模型最佳截断值(FVC%: 82.93%,FEV₁%:85.46%)均高于临床80%的标准阈值,存在一定的阈值偏移。核心原因可能在于呼吸高危人群肺功能状态受多因素共同调控,年龄、性别、基础疾病、生活习惯及环境等变量间存在复杂交互关系,增加了模型精准区分的难度。同时,肺功能评估存在显著个体异质性,不同个体生理基础与疾病进展差异较大,这也是临床肺功能预测模型普遍面临的挑战。

本研究构建的模型具有一定优势:以FVC和FEV₁为预测目标,利用常规易获取的临床资料构建模型,无需患者积极配合即可完成肺功能预测,贴合临床实际需求;样本量充足,为模型构建提供了坚实的数据基础,有助于获得稳健的统计推断结果。

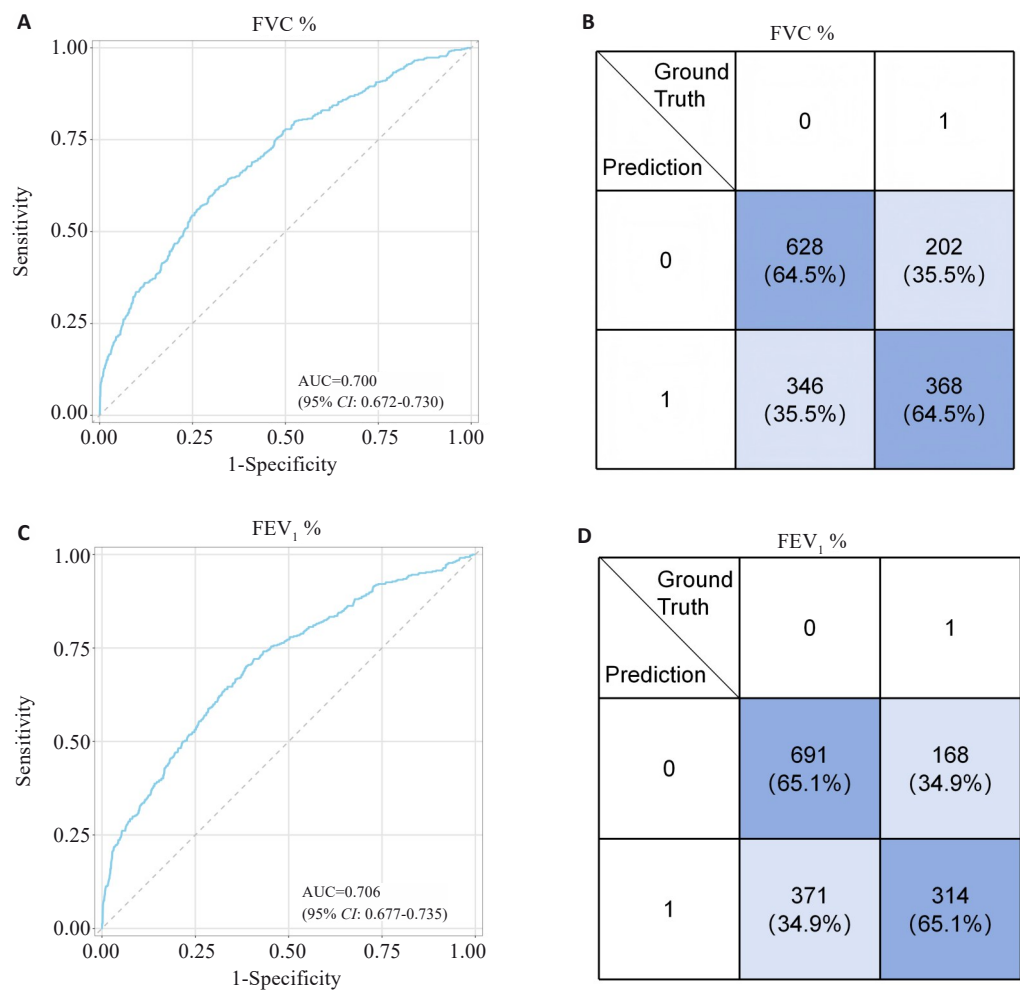


图5 基于XGBoost的ROC曲线和混淆矩阵
Fig.5 ROC and Confusion matrix based on XGBoost. **A:** ROC of FVC%. **B:** Confusion matrix of FVC%. **C:** ROC of FEV₁%. **D:** Confusion matrix of FEV₁%.

然而,本研究也存在一定的局限性:研究采用横断面及回顾性设计,虽总体样本量较大,但仅在单中心数据集进行验证,后续需开展多中心、多人群外部验证,以全面评估模型的泛化能力;模型对FVC、FEV₁的预测及高危人群分类效能均处于中等水平,未来可通过扩大样本量、增加预测变量及融合影像特征等方式,进一步提升预测准确性。

综上,本研究构建了一个肺功能评估模型,为老年患者肺功能异常的早期识别、及时干预及个体化诊疗决策提供参考。本研究结果也为后续模型优化提供了方向,为进一步完善肺功能评估体系、探索更精准的高危人群识别策略奠定了基础。

Declaration of interests: The authors declare no competing interests.

参考文献:

[1] 国家统计局,国务院第七次全国人口普查领导小组办公室. 第七次全国人口普查公报(第五号)[EB/OL]. (2021-05-11) [2025-01-08] [Z].
[2] Ding YQ, Hu JC, Yu CQ, et al. Spirometric pattern and

cardiovascular risk: a prospective study of 0.3 million Chinese never-smokers [J]. *Lancet Reg Health West Pac*, 2025, 54: 101274.
[3] Grande G, Li YJ, Trevisan C, et al. Lung function in relation to brain aging and cognitive transitions in older adults: a population-based cohort study [J]. *Alzheimers Dement*, 2024, 20(8): 5662-73.
[4] Backman H, Sawalha S, Nilsson U, et al. All-cause and cause-specific mortality by spirometric pattern and sex - a population-based cohort study [J]. *Ther Adv Respir Dis*, 2024, 18: 17534666241232768.
[5] Kokkinos P, Faselis C, Samuel IBH, et al. Cardiorespiratory fitness and mortality risk across the spectra of age, race, and sex [J]. *J Am Coll Cardiol*, 2022, 80(6): 598-609.
[6] Dehondt V, Vauterin D, Derom E, et al. Association of timely spirometry with lower all-cause mortality a nationwide obstructive cohort study [J]. *CHEST*, 2025, 168(3): 602-12.
[7] 中华医学会健康管理学分会, 中华健康管理学杂志编委会. 健康管理概念与学科体系的中国专家初步共识 [J]. *中华健康管理学杂志*, 2009, 3(3): 141-7.
[8] Maheshwari R, Singh A, Chaudhari PS, et al. Impulse oscillometry: an adjunct or alternative to spirometry in the diagnosis of obstructive airway disease. Results from an observational study [J]. *Indian J Tuberc*, 2026, 73(1): 21-4.

- [9] Lynch CJ, Liston C. New machine-learning technologies for computer-aided diagnosis[J]. *Nat Med*, 2018, 24(9): 1304-5.
- [10] Zhang BY, Sun D, Niu HT, et al. Development of a prediction model to identify undiagnosed chronic obstructive pulmonary disease patients in primary care settings in China[J]. *Chin Med J*, 2023, 136(6): 676-82.
- [11] Wu XN, Yin CS, Chen XQ, et al. Idiopathic pulmonary fibrosis mortality risk prediction based on artificial intelligence: the CTPF model[J]. *Front Pharmacol*, 2022, 13: 878764.
- [12] 中华医学会呼吸病学分会肺功能专业组. 肺功能检查指南(第二部分)——肺量计检查[J]. *中华结核和呼吸杂志*, 2014, 37(07): 481-6.
- [13] Graham BL, Steenbruggen I, Miller MR, et al. Standardization of spirometry 2019 update. an official American thoracic society and European respiratory society technical statement[J]. *Am J Respir Crit Care Med*, 2019, 200(8): e70-88.
- [14] Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease: 2026 Report [R]. Global Initiative for Chronic Obstructive Lung Disease, 2025.
- [15] Fernández-Blázquez MA, Ruiz-Sánchez de León JM, Sanz-Blasco R, et al. XGBoost models based on non imaging features for the prediction of mild cognitive impairment in older adults[J]. *Sci Rep*, 2025, 15(1): 29732.
- [16] Inoue T, Ichikawa D, Ueno T, et al. XGBoost, a machine learning method, predicts neurological recovery in patients with cervical spinal cord injury[J]. *Neurotrauma Rep*, 2020, 1(1): 8-16.
- [17] Schindele A, Krebold A, Heiß U, et al. Interpretable machine learning for thyroid cancer recurrence prediction: Leveraging XGBoost and SHAP analysis[J]. *Eur J Radiol*, 2025, 186: 112049.
- [18] Dominelli PB, Molgat-Seon Y. Sex, gender and the pulmonary physiology of exercise[J]. *Eur Respir Rev*, 2022, 31(163): 210074.
- [19] Wang C, Xu JY, Yang L, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study[J]. *Lancet*, 2018, 391(10131): 1706-17.
- [20] Chen SM, Kuhn M, Prettner K, et al. The global economic burden of chronic obstructive pulmonary disease for 204 countries and territories in 2020-50: a health-augmented macroeconomic modelling study[J]. *Lancet Glob Health*, 2023, 11(8): e1183-93.
- [21] Adenloye D, Song PG, Zhu YJ, et al. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis[J]. *Lancet Respir Med*, 2022, 10(5): 447-58.
- [22] 刘 领, 吴文杰, 杨 霁, 等. 慢性阻塞性肺疾病患者凝血状态临床观察[J]. *中国医药科学*, 2016, 6(2): 176-9.
- [23] 左晓红, 李庆红. 老年慢性阻塞性肺疾病急性加重期患者凝血功能的监测[J]. *血栓与止血学*, 2017, 23(1): 89-91, 94.
- [24] 李一龙, 景小莹. 支气管哮喘患者血清 LXA4 水平及凝血功能与肺功能及哮喘病程的相关性研究[J]. *河北医药*, 2019, 41(8): 1200-3.
- [25] 肉克亚木汗·艾尔肯. 支气管哮喘患者凝血功能与肺功能的相关性研究[D]. 乌鲁木齐: 新疆医科大学, 2022.
- [26] Díez-Sierra J, Medina-Gallardo JF, Muñoz-Sánchez B, et al. FEV(1) decline and clinical predictors in asthma after 10 Years of follow-up: BMI as a modifiable risk factor[J]. *Respir Med*, 2025, 246: 108247.
- [27] Lange P, Celli B, Agustí A, et al. Lung-function trajectories leading to chronic obstructive pulmonary disease[J]. *N Engl J Med*, 2015, 373(2): 111-22.

(编辑:经 媛)