基于机器学习建立老年慢性心力衰竭患者1年全因死亡预测模型

陈东升[1]，周泓屹[2]，姜 帆[1]

1 首都医科大学附属北京潞河医院 全科医学科 北京 101149；2 北京市通州区妇幼保健院 麻醉科 北京 101101

摘要：目的 基于机器学习建立老年慢性心力衰竭（CHF）患者1年全因死亡预测模型。**方法** 回顾性分析902例老年CHF患者（782例训练集和120例验证集）1年全因死亡的影响因素，随访1年，根据是否全因死亡分为死亡组和存活组，并建立5种机器学习预测模型。采用受试者操作特征（ROC）曲线、校准曲线、决策曲线对模型进行验证。**结果** 训练集782例老年CHF患者1年全因死亡率为12.28%；验证集120例老年CHF患者1年全因死亡率为10.83%。LASSO回归和多因素Logistic回归分析显示：体重指数增加、左室射血分数升高为老年CHF患者1年全因死亡的独立保护因素，纽约心脏协会心功能分级Ⅳ级、C反应蛋白升高、D−二聚体升高、氨基末端脑钠肽前体（NT-proBNP）升高为独立危险因素（$P<0.05$）。ROC曲线显示，训练集和验证集极端梯度提升（XGBoost）预测模型的曲线下面积（AUC）分别为0.897、0.864，均优于多因素Logistic回归（0.860、0.822）、决策树（0.767、0.761）、随机森林（0.875、0.818）、支持向量机（0.859、0.788）预测。验证集校准曲线显示，XGBoost预测模型的预测概率与实际曲线贴合，预测概率>0.10可为临床带来净收益。**结论** 体重指数、纽约心脏协会心功能分级、左室射血分数、C反应蛋白、NT-proBNP为老年CHF患者1年全因死亡的独立影响因素，基于此建立的XGBoost预测模型的预测效能最佳。

关键词：老年；慢性心力衰竭；机器学习；全因死亡；预测模型；极端梯度提升算法；左室射血分数

中图分类号：R541.6 **文献标识码：**A DOI：10.13885/j.issn.2097-681X.M20251161

Development of a machine learning-based model to predict one-year all-cause mortality in elderly patients with chronic heart failure

CHEN Dongsheng[1], ZHOU Hongyi[2], JIANG Fan[1]

1 Department of General Medicine, Beijing Luhe Hospital, Capital Medical University, Beijing 101149, China;

2 Department of Anesthesiology, Tongzhou District Maternal and Child Health Care Hospital, Beijing 101101, China

Abstract: Objective To develop a machine learning (ML)-based model for predicting 1-year all-cause mortality in elderly patients with chronic heart failure (CHF). **Methods** A ML-based model was built for predicting 1-year all-cause mortality in elderly patients with CHF. **Results** The 1-year all-cause mortality rate was 12.28% in the training set and 10.83% in the validation set. LASSO regression and multivariate logistic regression (LR) identified body mass index (BMI) and left ventricular ejection fraction (LVEF) as independent protective factors, while New York Heart Association (NYHA) class Ⅳ, elevated C-reactive protein (CRP), D-dimer (D-D), and N-terminal pro-brain natriuretic peptide (NT-proBNP) were independent risk factors ($P<0.05$). Receiver operator characteristic (ROC) curve showed that the extreme gradient boosting (XGBoost) model had the highest area under the curve (AUC) in both training and validation sets (0.897 and 0.864, respective-

收稿日期：2025−02−21 **修回日期：**2025-04-21 **接受日期：**2025-04-23

基金项目：首都全科医学与社区卫生研究专项资助项目（2023-2Y-014）

作者简介：姜 帆，女，副主任医师，研究方向为心血管疾病，e-mail：jiangfan11072019@163.com，通信作者

ly), outperforming LR (AUC=0.860, 0.822), decision tree (DT) (AUC=0.767, 0.761), random forest (RF) (AUC=0.875, 0.818), and support vector machine (SVM) (AUC=0.859, 0.788). The calibration curve in the validation set indicated that the XGBoost model closely aligned predicted and observed probabilities. DCA showed clinical benefit when predicted probability exceeded 0.10. **Conclusion** BMI, NYHA class, LVEF, CRP, and NT-proBNP were independent predictors of one-year all-cause mortality in elderly CHF patients. The XGBoost model based on these variables demonstrated superior predictive performance.

Keywords: elderly; chronic heart failure; machine learning; all-cause mortality; predictive model; extreme gradient boosting algorithm; left ventricular ejection fraction

心力衰竭为糖尿病、高血压、冠心病、心肌病、心脏瓣膜病、心律失常等各种原因引起心脏结构和/或功能异常，导致心室充盈或心室射血功能下降而无法满足机体代谢需求的一组复杂的临床综合征[1]。其主要临床表现为疲乏、呼吸困难、下肢水肿、体循环淤血和肺淤血，若未及时诊治可能导致心源性休克、严重心律失常、多器官功能衰竭、猝死等一系列严重后果，从而危及生命[2]。伴随老龄化趋势的加剧，近年来老年慢性心力衰竭（chronic heart failure, CHF）发病率也呈逐年攀升趋势，其死亡率也随年龄增长而增加[3]。据最新统计报道，当前中国心力衰竭患者约1 370万例，1年全因死亡率为13.7%，而18~49、50~64、65~79、≥80岁1年全因率分别为6.3%、9.2%、13.2%、21.6%，导致了巨大的疾病负担[4]。因此，准确预测老年CHF患者预后非常重要。随着人工智能的发展，机器学习（machine learning, ML）在医疗领域的应用日益广泛，可通过对大量临床数据进行深度挖掘与建模，实现个体化预测与干预决策支持[5]。本研究基于ML建立老年CHF患者1年全因死亡预测模型，旨在为临床医务人员提供便捷、准确的风险评估工具。

1　资料与方法

1.1　临床资料

本研究为回顾性研究，计算机随机数法选取2021年1月—2023年12月首都医科大学附属北京潞河医院收治的老年CHF患者782例为训练集。纳入标准：（1）年龄≥65岁；（2）CHF符合《老年人慢性心力衰竭诊治中国专家共识（2021）》[6]诊断标准；（3）有完整的临床资料和1年随访资料；（4）纽约心脏协会（New York Heart Association, NYHA）心功能分级Ⅲ~Ⅳ级。排除标准：（1）住院期间死亡者；（2）精神疾病或严重认知功能障碍患者；（3）近2周内感染者；（4）心脏移植或心脏起搏器置入或器械治疗患者；（5）恶性肿瘤患者；（6）血液系统疾病患者。本研究为匿名化的既有数据，免除知情同意。

1.2　方法

1.2.1　资料收集

收集老年CHF患者临床资料，包括人口统计学[性别、年龄、体重指数（body mass index, BMI）、民族]、CHF病程、NYHA心功能分级、既往病史、不良嗜好、入院时基本生命体征（心率、PR间期、校正QT间期、QRS时限、血压）、入院时心脏超声[左心房内径、左室舒张末期内径、左室射血分数（left ventricular ejection fraction, LVEF）]、入院时实验室指标[白细胞计数、红细胞计数、中性粒细胞百分比、血红蛋白、血小板计数、C反应蛋白（C-reactive protein, CRP）、谷丙转氨酶、谷草转氨酶、谷氨酰转移酶、直接胆红素、间接胆红素、白蛋白（albumin, ALB）、球蛋白、血肌酐、尿素氮、尿酸、总胆固醇、甘油三酯、高密度脂蛋白胆固醇（high-density lipoprotein cholesterol, HDL-C）、低密度脂蛋白胆固醇（low-density lipoprotein cholesterol, LDL-C）、血糖、D−二聚体（D-dimer, D-D）、氨基末端脑钠肽前体（N-terminal pro-brain natriuretic peptide, NT-proBNP）、肌酸激酶、估算肾小球滤过率]和用药情况[血管紧张素转换酶抑制剂/血管紧张素Ⅱ受体拮抗剂/血管紧张素受体−脑啡肽酶抑制剂（angiotensin-converting enzyme inhibitors/angiotensin Ⅱ receptor blockers/angiotensin receptor-neprilysin inhibitors, ACEI/ARB/ARNI）、β受体阻滞剂、盐皮质激素受体拮抗剂、袢利尿药、钠−葡萄糖协同转运蛋白2（sodium-dependent glucose transporters 2, SGLT-2）抑制剂]。

1.2.2 机器学习算法模型建立

经单因素分析后，采用LASSO回归和多因素Logistic回归（Logistic regression, LR）分析筛选和确定老年CHF患者1年全因死亡的独立影响因素。分别采用R语言的“stats包”“rpart.plot包”“randomForest包”“e1071包”“xgboost包”建立5种ML预测模型，分别为LR、决策树（decision tree, DT）、随机森林（random forest, RF）、支持向量机（support vector machine, SVM）、极端梯度提升（extreme gradient boosting, XGBoost）。

1.3 分组方法

本研究的随访起点为患者首次入院确诊并开始接受规范化治疗的时间，随访期为1年，2021年1月开始，2024年12月终止。随访方式主要为电话结合门诊复查。终点事件定义为随访期间的全因死亡。根据是否发生终点事件，将患者分为死亡组和存活组。

1.4 统计学分析

研究数据采用SPSS 28.0和RStudio 4.5.0软件处理。性别、民族、NYHA心功能分级等计数资料以例(%)表示并采用χ^2检验；经K-S检验呈正态分布计量资料PR间期、校正QT间期、QRS宽度等以平均数±标准差（$\bar{x}\pm s$）表示并采用t检验；偏态分布计量资料年龄、BMI、CHF病程等以中位数（四分位数）[$M(P_{25}, P_{75})$]表示并采用Mann-Whitney U检验；LASSO回归及多因素LR分析老年CHF患者1年全因死亡的影响因素；“pROC包”绘制受试者操作特征（receiver operator characteristic, ROC）曲线，分析各ML预测模型的预测能力，计算准确率、敏感度、特异度、精确率[真阳性/(真阳性＋假阳性)]、$F1$分数[2(精确率×敏感度)/(精确率＋敏感度)]；“rms包”绘制校准曲线，分析ML预测模型的准确性；“rmda包”绘制决策曲线，分析ML预测模型的临床效益；“shapviz包”沙普利可加性特征解释方法（Shapley additive explanations, SHAP）确定各变量对预测模型的具体贡献；检验水准α=0.05。

2 结果

2.1 一般资料

782例老年CHF患者年龄65~88岁，平均为76.00岁；女284例、男498例；BMI为16.33~31.72 kg/m^2，平均为23.64 kg/m^2；汉族717例，其他民族65例；CHF病程2~13年，平均5.00年；NYHA心功能分级Ⅲ级者559例、Ⅳ级者223例。

2.2 老年CHF患者1年全因死亡的单因素分析

782例老年CHF患者1年全因死亡率为12.28%（96/782）。单因素方差分析显示，年龄、BMI、CHF病程、NYHA心功能分级、左心房内径、左室舒张末期内径、LVEF、白细胞计数、CRP、谷丙转氨酶、直接胆红素、ALB、球蛋白、尿素氮、HDL-C、LDL-C、血糖、D-D、NT-proBNP、肌酸激酶、估算肾小球滤过率、ACEI/ARB/ANRI、β受体阻滞剂与老年CHF患者1年全因死亡有关（P<0.05），见表1。

2.3 老年CHF患者1年全因死亡的LASSO回归变量筛选

以年龄、BMI、CHF病程、NYHA心功能分级(Ⅳ级/Ⅲ级=1/0)、左心房内径、左室舒张末期内径、LVEF、白细胞计数、CRP、谷丙转氨酶、直接胆红素、ALB、球蛋白、尿素氮、HDL-C、LDL-C、血糖、D-D、NT-proBNP、肌酸激酶、估算肾小球滤过率、ACEI/ARB/ANRI（是/否=1/0）、β受体阻滞剂（是/否=1/0）为自变量（连续型变量均原值录入），老年CHF患者1年全因死亡（是/否=1/0）为因变量进行LASSO回归变量筛选。为获得更简洁的预测模型，选择lambda.1se（λ=−4.274 76 481）和十倍交叉验证作为惩罚系数，共筛选出9个具有非零特征的变量，分别为年龄、BMI、NYHA心功能分级、LVEF、CRP、ALB、尿素氮、D-D、NT-proBNP，见图1。

2.4 老年CHF患者1年全因死亡的多因素LR分析

以LASSO回归变量为自变量，老年CHF患者1年全因死亡为因变量建立LR模型，多重共线性较弱（VIF<5）。结果显示，BMI增加、LVEF升高为老年CHF患者1年全因死亡的独立保护因素，NYHA心功能分级Ⅳ级、CRP升高、D-D升高、NT-proBNP升高为独立危险因素（P<0.05），见表2。

表1　老年CHF患者1年全因死亡的单因素分析

组别	性别[*n*(%)]		年龄/岁	BMI/(kg·m^{-2})	民族[*n*(%)]		CHF病程/年
	男	女			汉族	其他	
死亡组(*n*=96)	63(65.6)	33(34.4)	79.50(71.25,87.00)	22.23(19.83,24.58)	86(89.6)	10(10.4)	6.00(5.00,7.00)
存活组(*n*=686)	435(63.4)	251(36.6)	76.00(68.00,83.00)	23.72(21.35,26.58)	631(92.0)	55(8.0)	5.00(4.00,6.00)
*P*值	0.673		0.005	<0.001	0.425		0.010

组别	NYHA心功能分级[*n*(%)]		既往病史[*n*(%)]					不良嗜好[*n*(%)]	
	Ⅲ级	Ⅳ级	高血压	糖尿病	冠心病	卒中史	房颤	吸烟	饮酒
死亡组(*n*=96)	49(51.0)	47(49.0)	49(51.0)	29(30.2)	50(52.1)	13(13.5)	34(35.4)	35(36.5)	16(16.7)
存活组(*n*=686)	510(74.3)	176(25.7)	349(50.9)	161(23.5)	337(49.1)	72(10.5)	206(30.0)	254(37.0)	137(20.0)
*P*值	<0.001		0.976	0.149	0.587	0.369	0.284	0.914	0.445

组别	心率/(次·min^{-1})	PR间期/ms	校正QT间期/ms	血压/mmHg		左心房内径/mm
				收缩压	舒张压	
死亡组(*n*=96)	87.00(70.00,99.75)	161.54±41.61	464.57±36.61	120.16±23.00	76.99±14.69	39.34±8.35
存活组(*n*=686)	86.00(70.75,101.00)	165.40±34.01	459.84±45.71	123.42±22.12	78.11±15.26	37.05±7.60
*P*值	0.932	0.387	0.254	0.178	0.498	0.007

组别	左室舒张末期内径/mm	LVEF/%	白细胞计数/(×10^{9}·L^{-1})	红细胞计数/(×10^{12}·L^{-1})	中性粒细胞百分比/%	血红蛋白/(g·L^{-1})	血小板计数/(×10^{9}·L^{-1})
死亡组(*n*=96)	52.14±8.19	49.75±12.81	9.39±4.07	3.81±0.81	75.82±13.83	113.89(102.35,131.45)	184.56±85.33
存活组(*n*=686)	49.43±6.56	54.50±12.35	8.33±2.60	3.91±0.64	74.53±14.17	112.45(68.20,144.93)	196.24±73.12
*P*值	0.002	<0.001	0.014	0.226	0.403	0.224	0.204

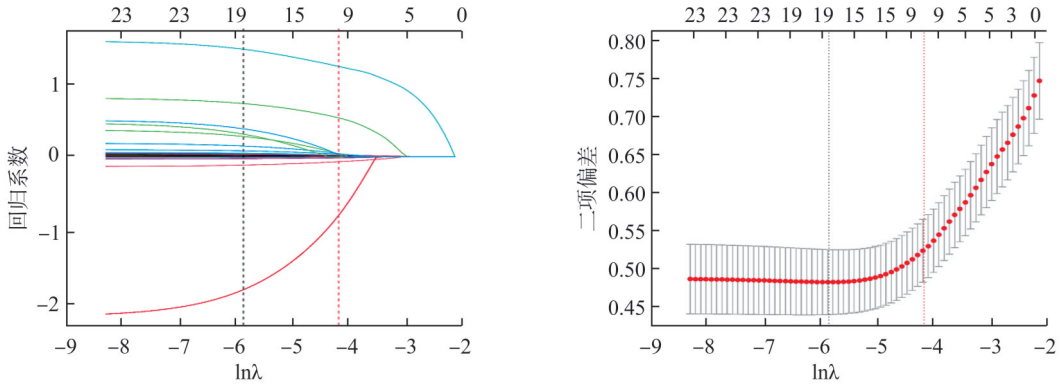
组别	C反应蛋白/(mg·L^{-1})	谷丙转氨酶/(U·L^{-1})	谷草转氨酶/(U·L^{-1})	谷氨酰转移酶/(U·L^{-1})	直接胆红素/(μmol·L^{-1})	间接胆红/(μmol·L^{-1})
死亡组(*n*=96)	44.36(31.08,57.38)	28.64(13.70,46.25)	33.25(19.08,43.39)	35.78(22.57,59.98)	4.72(1.98,8.61)	10.29(6.20,14.25)
存活组(*n*=686)	34.47(23.94,44.63)	24.04(14.08,34.30)	28.61(16.35,40.21)	30.92(17.06,47.14)	3.96(1.94,6.29)	9.73(6.24,13.41)
*P*值	<0.001	0.039	0.142	0.035	0.028	0.197

组别	白蛋白/(g·L^{-1})	球蛋白/(g·L^{-1})	血肌酐/(μmol·L^{-1})	尿素氮/(mmol·L^{-1})	尿酸/(μmol·L^{-1})	血糖/(mmol·L^{-1})	D-二聚体/(mg·L^{-1})
死亡组(*n*=96)	32.90(29.12,37.45)	30.63±6.06	96.75±46.10	9.05(4.22,12.811)	366.85±162.74	6.86±1.89	1.23(0.56,1.94)
存活组(*n*=686)	35.06(31.14,38.79)	29.26±5.93	89.35±33.27	7.46(5.20,9.69)	345.79±127.76	2.34±0.67	0.67(0.32,1.03)
*P*值	0.004	0.035	0.132	0.006	0.226	0.019	<0.001

组别	血脂4项/(mmol·L^{-1})				NT-proBNP/(ng·L^{-1})	肌酸激酶/(IU·L^{-1})
	总胆固醇	甘油三酯	HDL-C	LDL-C		
死亡组(*n*=96)	3.42(2.88,4.11)	1.33±0.64	0.92(0.76,1.08)	2.51±0.61	5 496.60(3 353.97,7 094.90)	195.31(82.19,370.74)
存活组(*n*=686)	3.62(3.14,4.15)	1.23±0.57	1.00(0.86,1.14)	2.34±0.67	195.31(82.19,370.74)	179.58(103.76,259.18)
*P*值	0.056	0.123	0.002	0.019	<0.001	0.035

组别	估算肾小球滤过率/(mL^{-1}·min^{-1}·1.73 m^{2})	用药情况[*n*(%)]				
		ACEI/ARB/ANRI	β受体阻滞剂	盐皮质激素受体拮抗剂	袢利尿药	SGLT2抑制剂
死亡组(*n*=96)	62.36(37.98,89.09)	80(83.3)	84(87.5)	88(91.7)	86(89.6)	20(20.8)
存活组(*n*=686)	73.92(52.22,91.14)	487(71.0)	533(77.7)	614(89.5)	597(87.0)	200(29.2)
*P*值	0.027	0.011	0.027	0.513	0.480	0.089

BMI　体重指数；CHF　慢性心力衰竭；NYHA　纽约心脏协会；LVEF　左室射血分数；HDL-C　高密度脂蛋白胆固醇；LDL-C　低密度脂蛋白胆固醇；NT-proBNP　氨基末端脑钠肽前体；ACEI/ARB/ARNI　血管紧张素转换酶抑制剂/血管紧张素Ⅱ受体拮抗剂/血管紧张素受体-脑啡肽酶抑制剂；SGLT-2　钠-葡萄糖协同转运蛋白2；1 mmHg≈0.133 kPa。



A LASSO回归路径图 B LASSO回归交叉验证曲线

图1 老年慢性心力衰竭患者1年全因死亡的LASSO回归变量筛选图

表2 老年慢性心力衰竭患者1年全因死亡的多因素Logistic回归分析

变量	回归系数	标准误	Wald χ^2值	*OR*值	95% *CI*	*P*值
年龄	0.028	0.017	2.907	1.029	[0.996,1.063]	0.088
体重指数	−0.136	0.040	11.810	0.873	[0.808,0.943]	0.001
纽约心脏协会心功能分级Ⅳ级	0.988	0.274	12.965	2.685	[1.568,4.597]	<0.001
左室射血分数	−0.034	0.011	9.866	0.967	[0.947,0.987]	0.002
C反应蛋白	0.037	0.008	18.711	1.037	[1.020,1.055]	<0.001
白蛋白	−0.035	0.023	2.357	0.966	[0.924,1.010]	0.125
尿素氮	0.695	0.516	1.811	2.003	[0.728,5.507]	0.178
D−二聚体	1.521	0.233	42.549	4.577	[2.898,7.228]	<0.001
氨基末端脑钠肽前体	0.000	0.000	32.411	1.000	[1.000,1.000]	<0.001
常数	−3.235	1.916	2.852	0.039		0.091

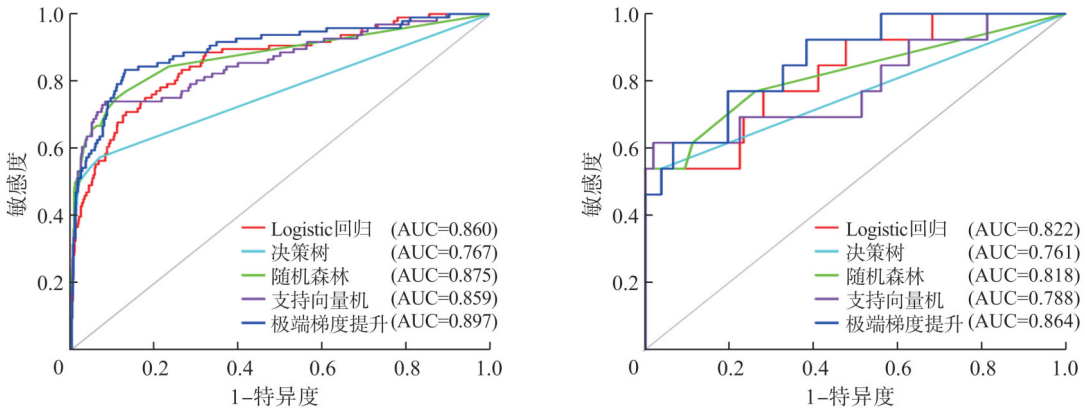
2.5 基于ML的老年CHF患者1年全因死亡预测模型建立与验证

以BMI、LVEF、NYHA心功能分级、CRP、D-D、NT-proBNP为特征变量，老年CHF患者1年全因死亡为预测变量建立LR、DT、RF、SVM、XGBoost预测模型。按照纳入和排除标准选取2024年1月—2024年6月收治的120例老年CHF患者为验证集，1年全因死亡率为10.83%（13/120）。ROC曲线显示，训练集和验证集XGBoost预测模型的曲线下面积（area under the curve, AUC）分别为0.897、0.864，均优于LR预测模型（AUC分别为0.860、0.822）、DT预测模型（AUC分别为0.767、0.761）、RF预测模型（AUC分别为0.875、0.818）、SVM预测模型（AUC分别为0.859、0.788），见表3和图2。校准曲线显示，验证集XGBoost预测模型的预测概率与实际曲线贴合；决策曲线显示，

表3 训练集与验证集中各ML预测模型的预测价值

预测模型	数据集	AUC	95% *CI*	*P*值	准确率	敏感度	特异度	精确率	*F*1分数
Logistic回归	训练集	0.860	[0.817,0.904]	<0.001	0.909	0.375	0.984	0.766	0.503
	验证集	0.822	[0.693,0.952]	<0.001	0.950	0.538	1.000	1.000	0.700
决策树	训练集	0.767	[0.714,0.820]	<0.001	0.923	0.500	0.983	0.800	0.615
	验证集	0.761	[0.614,0.907]	<0.001	0.950	0.538	1.000	1.000	0.700
随机森林	训练集	0.875	[0.830,0.919]	<0.001	0.910	0.302	0.996	0.906	0.453
	验证集	0.818	[0.672,0.964]	<0.001	0.933	0.385	1.000	1.000	0.556
支持向量机	训练集	0.859	[0.811,0.908]	<0.001	0.921	0.417	0.991	0.870	0.563
	验证集	0.788	[0.621,0.955]	<0.001	0.950	0.538	1.000	1.000	0.700
极端梯度提升	训练集	0.897	[0.857,0.936]	<0.001	0.914	0.333	0.996	0.914	0.489
	验证集	0.864	[0.759,0.969]	<0.001	0.933	0.385	1.000	1.000	0.556

AUC 曲线下面积。



A　训练集　　B　验证集

AUC　曲线下面积。

图2　各机器学习预测模型的受试者操作特征曲线

预测概率>0.10时，验证集XGBoost预测模型可为临床带来净收益，见图3、4。

2.6　XGBoost预测模型的变量重要性及SHAP分析

SHAP依赖图和基于平均绝对SHAP值绘制的特征重要性排序柱状图结果显示，在XGBoost预测模型中，各变量对老年CHF患者1年全因死亡风险的预测贡献由高到低依次表现为D-D、NT-proBNP、BMI、CRP、NYHA心功能分级、LVEF，见图5、6。

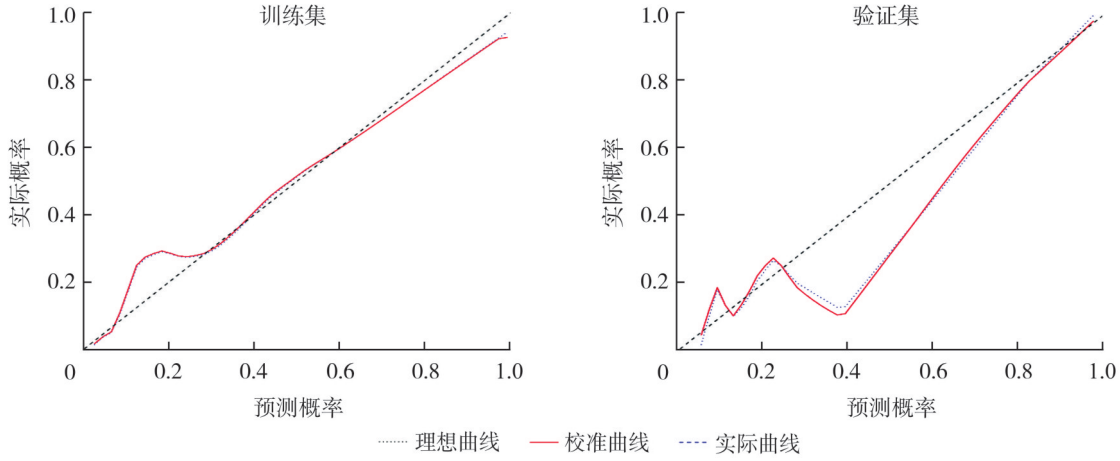


图3　极端梯度提升预测模型的校准曲线

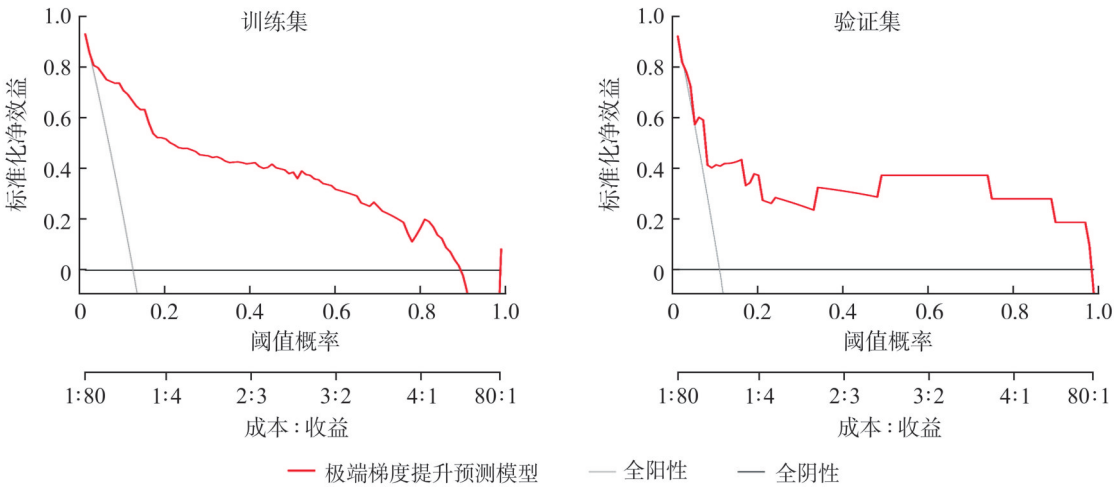
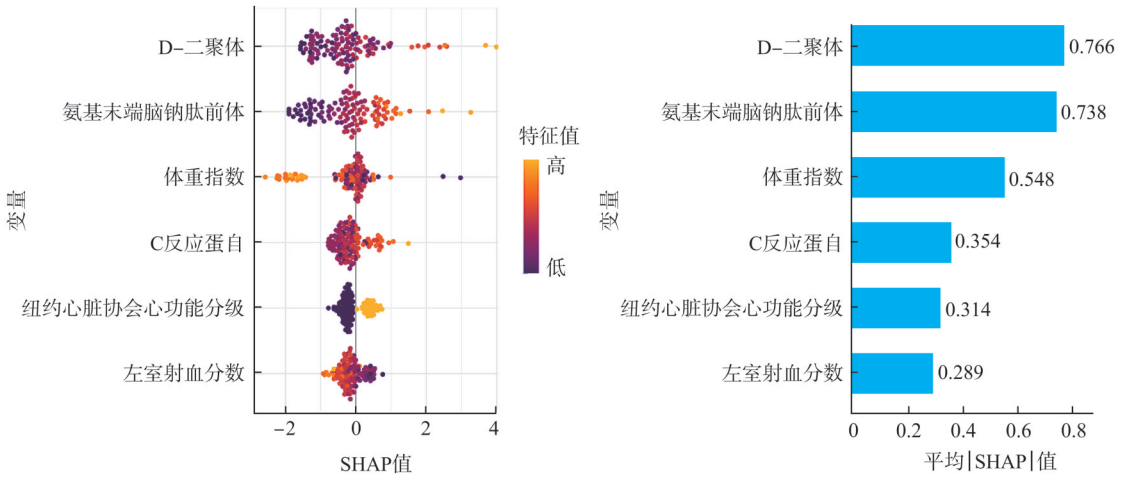


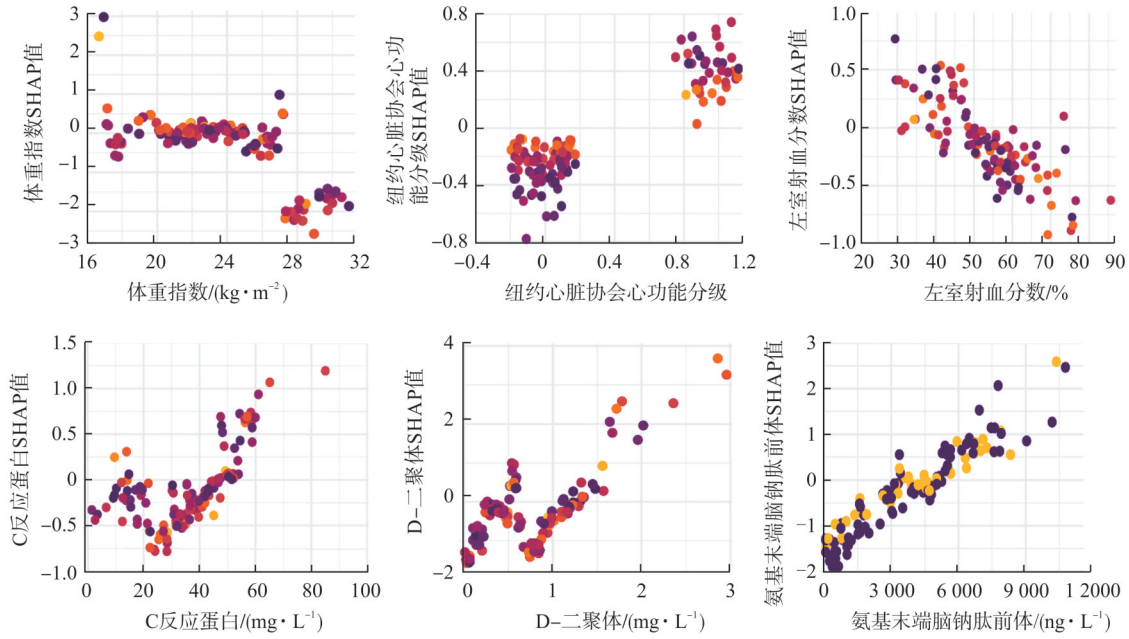
图4　极端梯度提升预测模型的决策曲线



A SHAP全局可解释性分析图 B SHAP总体特征重要性排序图

SHAP 沙普利可加性特征解释方法。

图5 各变量对极端梯度提升预测模型的沙普利可加性特征解释法依赖图



SHAP 沙普利可加性特征解释方法。

图6 单变量的沙普利可加性特征解释法依赖图

3 讨论

老年CHF患者由于自然衰老、多病共存、多重用药和伴有多种心血管疾病危险因素等原因，相较于非老年CHF群体更易出现急性失代偿、心律失常和多器官功能障碍，导致活动耐力下降和反复住院，并增加病死风险[6]。尽管目前已有研究分析了老年CHF患者预后影响因素，但多为单个指标或基于传统回归分析，难以准确捕捉多变量之间的复杂非线性关系[7-8]。ML方法能够自动处理高维数据、识别关键变量间的潜在关联，具有更强的拟合能力与泛化性能。

本研究结果显示，BMI增加为老年CHF患者1年全因死亡的独立保护因素。营养不良是CHF患者常见的问题，尤其在老年患者中更为突出，其机制涉及慢性炎症、食欲减退、营养吸收障碍、代谢亢进和心脏恶病质等多种因素[9]。营养不良可导致肌肉减少、免疫功能下降、运动耐力降低，进而加重心衰症状，增加感染风险与复发率，最终影响患者的长期生存率[10]。较高的BMI可能反映患者有更充足的营养储备，有助于维持肌肉质量和免疫功能，在CHF病程中提供一定的代偿能力，避

免因营养不良导致的心功能恶化，从而降低全因死亡风险[11]。赵灿等[12]也指出，低BMI（≤18.5 kg/m^2）为CHF患者2年全因死亡的独立影响因素。NYHA心功能分级是评估心力衰竭患者症状严重程度和日常活动耐力的重要临床工具，分级越高则病情更严重[1]。本研究结果显示，NYHA心功能分级Ⅳ级为老年CHF患者1年全因死亡的独立危险因素，这与既往研究[13]报道老年CHF患者死亡风险随着NYHA心功能分级增加而增加的结果相符。NYHA心功能分级Ⅳ级即休息时也会出现心力衰竭症状，反映患者心脏泵血能力严重降低，可引起全身器官灌注不足、肾功能恶化、肝功能损害、电解质紊乱等并发症，增加死亡风险[14]。LVEF是评估心脏收缩功能最常用的指标，反映左心室每次收缩向主动脉射出的血液比例，正常值通常为50%~70%[15]。LVEF升高反映老年CHF患者左心室泵血能力较好，心排血量维持相对充分，可有效保障全身器官组织的灌注，减少组织低灌注相关并发症发生，从而降低全因死亡风险。

本研究结果显示，CRP升高为老年CHF患者1年全因死亡的独立危险因素，这可能与炎症反应在CHF病程中发挥重要作用有关。CRP是经典的评估全身炎症水平的急性期反应蛋白，近年研究认为CHF是一种涉及慢性低度炎症激活的全身性疾病，CRP水平越高反映患者炎症反应越强，能通过心肌纤维化和心肌细胞凋亡等促进心室重构，进一步恶化心功能[16]；同时炎症反应也能通过血管内皮功能障碍和动脉粥样硬化等增加老年CHF患者多种主要不良心血管事件，进一步导致死亡风险增加[17]。本研究结果显示，D-D升高为老年CHF患者1年全因死亡的独立危险因素，这可能与血栓形成有关。D-D是纤维蛋白降解产物，是反映体内凝血与纤溶系统激活的重要指标，CHF患者由于血流瘀滞、血管内皮损伤及炎症反应等因素影响导致血液呈现高凝状态，易形成微血栓，引起肺栓塞、深静脉血栓、弥散性血管内凝血和心内血栓等并发症，导致死亡风险增加[18]。NT-proBNP是心肌细胞在压力负荷增加或容量负荷过重时释放的脑钠肽前体分子，相较于脑利钠肽不受神经内分泌因素和快速降解的影响，且稳定性较强、半衰期长，是评价心力衰竭严重程度和预后的常用生物标志物之一。NT-proBNP水平越高反映老年CHF患者心室壁张力越大、心功能越差，因此死亡风险更高[19]。

基于以上老年CHF患者1年全因死亡的独立影响因素，本研究构建了5种ML预测模型。结果显示，训练集和验证集XGBoost预测模型均优于LR预测模型、DT预测模型、RF预测模型、SVM预测模型，提示XGBoost预测模型表现最好。XGBoost具有强大的建模能力和对复杂非线性关系的适应性，能通过迭代构建多个决策树，逐步优化残差，具有出色的特征表达能力和泛化性能，使其在处理多变量复杂预测任务中具备明显优势[20]。进一步内部验证发现，验证集XGBoost预测模型的校准曲线预测概率与实际曲线贴合，决策曲线显示，预测概率>0.10时，验证集XGBoost预测模型可为临床带来净收益。表明该模型在实际应用中具有较高的临床实用价值，对临床早期识别高风险患者、制定个体化干预措施提供了有效工具。SHAP是一种基于博弈论的模型解释方法，用于量化每个特征对模型预测结果的贡献[21]。本研究结果显示，在XGBoost预测模型中，SHAP值排序变量依次为D-D、NT-proBNP、BMI、CRP、NYHA心功能分级和LVEF，D-D、NT-proBNP为前两位，提示凝血功能紊乱和心力衰竭程度是影响死亡风险的关键指标；BMI和CRP反映患者的营养与炎症状态，也显著影响预后；NYHA心功能分级和LVEF是心功能评估的重要临床指标，进一步强调了基础心功能在死亡风险预测中的核心地位。这一分析结果不仅提高了模型的可解释性，也为临床风险评估提供了变量优先级参考。但本研究仍存在一定局限性：首先为单中心回顾性研究，样本量有限，可能存在选择偏倚，影响模型的泛化能力；其次，仅纳入住院患者，缺乏社区或门诊CHF患者数据，限制了模型的外推性；再次，未结合影像组学、生物标志物或动态时间序列数据，模型信息维度仍不够全面。后续研究可通过多中心、大样本、前瞻性研究设计，提高模型的可靠性和代表性；同时结合多模态数据与深度学习技术，进一步优化预测性能，并开发基于模型的智能化评估工具，推动其在临床中的推广应用。

综上所述，BMI、NYHA心功能分级、LVEF、CRP、NT-proBNP为老年CHF患者1年全因死亡的独立影响因素，基于此建立的XGBoost预测模型的预测效能最佳。

伦理声明 本研究获得首都医科大学附属北京潞河医院伦理委员会批准(批件文号：2024-LHKY-073-02)。

利益冲突声明 本研究不存在研究者、伦理委员会成员、受试者监护人以及与公开研究成果有关的利益冲突。

作者贡献声明 陈东升：实施研究过程、进行分析，撰写文章；周泓屹：确定研究对象范围、收集数据、采集、清洗与分析数据；姜 帆：参与撰写论文，起草论文、修订论文、审核论文。

参考文献

[1] 郇轩,田梅香,尹江燕,等.心力衰竭合并糖尿病住院患者临床特点的回顾性分析及意义[J].兰州大学学报(医学版),2023,49(8):68-75.

[2] PAVÍA-LÓPEZ A A, MAGAÑA-SERRANO J A, CIGARROA-LÓPEZ J A, et al. Clinical practice guidelines for diagnostic and treatment of the chronic heart failure[J]. Archivos de cardiologia de Mexico, 2024, 94(Supl 1): 1-74.

[3] LIU Z Y, LI Z P, LI X Q, et al. Global trends in heart failure from 1990 to 2019: an age-period-cohort analysis from the global burden of disease study[J]. ESC heart failure, 2024, 11(5): 3264-3278.

[4] WANG H, LI Y Y, CHAI K, et al. Mortality in patients admitted to hospital with heart failure in China: a nationwide cardiovascular association database-heart failure centre registry cohort study[J]. The lancet global health, 2024, 12(4): e611-e622.

[5] SUN Z Y, WANG Z H, YUN Z J, et al. Machine learning-based model for worsening heart failure risk in Chinese chronic heart failure patients[J]. ESC heart failure, 2025, 12(1): 211-228.

[6]中华医学会老年医学分会心血管疾病学组,《老年慢性心力衰竭诊治中国专家共识》编写组.老年人慢性心力衰竭诊治中国专家共识(2021)[J].中华老年医学杂志,2021,40(5):550-561.

[7] TUERSUN R, ABUDOUWAYITI A, LI Y X, et al. Serum CA125: a prognostic biomarker for mortality in chronic heart failure[J]. BMC cardiovascular disorders, 2025, 25(1): 227.

[8] 李晨,江珊,汪欣,等.三种衰弱评估工具对老年慢性心力衰竭患者不良结局预测能力的比较[J].中华老年心脑血管病杂志,2024,26(10):1125-1129.

[9] MONTALTO M, D'IGNAZIO F, CAMILLI S, et al. Heart failure in older patients: an update[J]. Journal of clinical medicine, 2025, 14(6): 1982.

[10] LAU G H W, AUYEUNG T W, JOYCE E, et al. Consolidating concepts from recent international expert consensus on sarcopenia: implications for heart failure management[J/OL]. (2025-06-16)[2025-07-10]. Journal of cardiac failure, 2025: https://doi.org/10.1016/j.cardfail.2025.03.021.

[11] JENSEN G L, CEDERHOLM T, CORREIA M I T D, et al. GLIM consensus approach to diagnosis of malnutrition: a 5-year update[J]. Journal of parenteral and enteral nutrition, 2025, 49(4): 414-427.

[12] 赵灿,王刚,刘霄燕,等.BMI对慢性心力衰竭的预后价值[J].临床心血管病杂志,2024,40(3):194-198.

[13] XU Z, YUE Y B, XU M F, et al. The role of cut-off values for creatinine, blood urea nitrogen, and uric acid in prognostic assessment of chronic heart failure: a retrospective cohort study[J]. BMC cardiovascular disorders, 2025, 25(1): 209.

[14] 中国老年医学学会心电及心功能分会,中国医师协会心血管内科分会,中国心衰中心联盟专家委员会.慢性心力衰竭加重患者的综合管理中国专家共识2022[J].中国循环杂志,2022,37(3):215-225.

[15] 中华医学会超声医学分会超声心动图学组,国家超声诊断专业医疗质量控制中心专家委员会.经胸超声心动图检查规范化应用中国专家共识(2024版)[J].中华超声影像学杂志,2024,33(1):1-13.

[16] ROMAN-PEPINE D, SERBAN A M, CAPRAS R D, et al. A comprehensive review: unraveling the role of inflammation in the etiology of heart failure[J]. Heart failure reviews, 2025, 30(5): 931-954.

[17] FANG Z, RAZA U, SONG J, et al. Systemic aging fuels heart failure: molecular mechanisms and therapeutic avenues[J]. ESC heart failure, 2025, 12(2): 1059-1080.

[18] SEGRETI A, MASTROBERARDINO S, FRAU L, et al. Severe heart failure and intracardiac thrombosis: going beyond the appearance for diagnosis and treatments[J]. Monaldi archives for chest disease, 2025, 95(2): 1-10.

[19] 中国医师协会检验医师分会心血管专家委员会.B型利钠肽及N末端B型利钠肽前体实验室检测与临床应用中国专家共识[J].中华医学杂志,2022,102(35):2738-2754.

[20] SHANG S, WEI M, LV H S, et al. Construction and validation of a hospital mortality risk model for advanced elderly patients with heart failure based on machine learning[J]. International journal of general medicine, 2025, 18: 3277-3288.

[21] LI G J, ZHAO Z Q, YU Z L, et al. Machine learning for risk prediction of acute kidney injury in patients with diabetes mellitus combined with heart failure during hospitalization[J]. Scientific reports, 2025, 15(1): 10728.

(责任编辑：张 朗)