

预测 PCI 术后主要不良心血管事件的新模型探索 ——基于 Super Learner 算法*

朱祥¹ 张频² 涂嘉欣¹ 吴磊¹

[摘要] 目的:建立多种基于集成学习的经皮冠状动脉介入(PCI)预测模型,从而筛选出最佳的预测模型组合及变量。方法:本研究纳入 2018 年 1 月—2022 年 6 月所有符合 PCI 手术标准并行 PCI 术的心肌梗死患者,采用 Cox 回归对变量进行筛选,根据不同的单一模型组合通过五折交叉验证建立 Super Learner(SL)模型,并用 ROC 曲线和 PR 曲线对各个模型进行评价。结果:本次研究共收集到 3 880 例 PCI 患者,平均年龄为 64.46 岁,其中大部分是男性(73.8%)和高血压(57.2%)。Cox 回归共筛选出 24 个相关变量,最终共建立 4 个 SL 模型,预测能力在测试集上表现均较好,AUC 值均在 0.84 以上,而 PRC 值均在 0.72 以上,其中混合模型 1 表现最佳(AUC:0.846,PRC:0.729)。预测变量中住院次数、是否肾功能不全、病变冠脉支数等是各个混合模型中的重要变量,且与主要不良心血管事件(MACE)发生呈正相关。结论:这项研究为 PCI 预后影响因素研究及 SL 模型在 PCI 领域的应用增加了证据。住院次数、肾功能不全、病变冠脉支数、Cre 等对 MACE 的发生有一定影响,术前有效改善其肾功能降低 MACE 的发生。而 Super Learner 模型作为集成学习的实际应用之一,有效提高了整体预测模型的稳定性及适用性,同时证明了在 PCI 术预后上的应用价值。

[关键词] 集成学习;Super Learner;经皮冠状动脉介入术;主要不良心血管事件

DOI:10.13201/j.issn.1001-1439.2023.05.008

[中图分类号] R541.4 **[文献标志码]** A

A new model for predicting major adverse cardiovascular events after PCI surgery: based on Super Learner algorithm

ZHU Xiang¹ ZHANG Pin² TU Jiaxin¹ WU Lei¹

(¹Jiangxi Provincial Key Laboratory of Preventive Medicine, Nanchang University, Nanchang, 330000, China;²School of Public Health and Management, Nanchang Medical College)

Corresponding author: WU Lei, E-mail: leiwu@ncu.edu.cn

Abstract Objective: To establish a variety of PCI prediction models based on ensemble learning to screen the best combination of prediction models and variables. **Methods:** In this study, all patients who met the criteria for PCI and underwent PCI from January 2018 to June 2022 were included. The variables were screened by Cox regression, and the Super Learner model was established by 5-fold cross-validation according to different single model combinations. ROC curve and PR curve were used to evaluate each model. **Results:** A total of 3880 patients with PCI were included in the study, with the average age of 64.46 years old. Most of them were male(73.8%) and had hypertension(57.2%). Twenty-four related variables were screened out by Cox regression and four SL models were finally established. The prediction abilities of the four SL models were all good on the test set, with the AUC values above 0.84 and the PRC values above 0.72, among which the hybrid model 1 performed the best (AUC: 0.846 and PRC: 0.729). The number of hospitalizations, renal insufficiency or not, and the number of diseased coronary artery branches were important variables in each mixed model among the prediction variables, and they were positively correlated with the occurrence of MACE. **Conclusion:** This study provides additional evidence for that study of prognostic factors of PCI and the application of SL model in PCI. The times of hospitalization, renal insufficiency, the number of diseased coronary artery branches, and Cre have a certain influence on the occurrence of MACE. Effective improvement of renal function before surgery can reduce the occurrence of MACE. As one of the practical applications of ensemble learning, Super Learner model effectively improves the stability and applicability of the overall prediction model, and proves its application value in the prognosis of PCI.

Key words ensemble learning; Super Learner; PCI; MACE

*基金项目:国家自然科学基金地区基金项目(No:81960611);江西省研究生创新专项基金(No:YC2022-s096)

¹南昌大学江西省预防医学重点实验室(南昌,330000)

²南昌医学院公共卫生与管理学院

通信作者:吴磊,E-mail:leiwu@ncu.edu.cn

根据最新发布的《中国心血管健康与疾病报告 2021》^[1],农村、城市中心血管疾病分别占死因的 46.74%和 44.26%,即每 5 例死亡中就有 2 例死于心血管疾病,而冠心病以现患人数 1139 万居心血管疾病患病第 2 位。经皮冠状动脉介入术(PCI)作为治疗冠心病的有效治疗手段之一^[2-4],其技术手段已发展地日益成熟。根据第 25 届全国介入心脏病学论坛发布的中国大陆冠心病介入治疗数据,2021 年 PCI 术治疗数相较于 2020 年增长约 20.18%。然而 PCI 术后却依旧伴随着多种并发症,如出血、回流、夹层流量减少、血栓等^[5-7]。这不仅会引起 PCI 术对原有疾病治疗的失败,还有可能导致再发心肌梗死、死亡等严重后果。因此,为改善其手术预后情况,多项研究将主要心血管不良事件(MACE)作为 PCI 术预后的评价指标之一,并进行糖尿病、血管造影、植入支架类型等影响因素的研究^[8-9]。

而机器学习中的集成学习作为近年来各类疾病生存数据建立预测模型的优先项^[10-12],其中由 van der Laan 等^[13]提出的 Super Learner(SL)模型借助堆叠泛化原理将最小二乘、随机森林、人工神经网络和支持向量机等多个经典模型进行集成学习,并通过交叉验证保证预测模型的稳定性。研究显示,在产后感染、疾病发作、情感障碍等领域最终形成的 SL 模型预测能力均显著优于单一模型^[14-16]。而 PCI 目前建立的 MACE 模型预测曲线下面积大多在 0.7 左右^[17],且运用 Cox、人工神经网络等单一模型而较少应用 SL 的集成学习。因此,本研究通过建立基于集成学习的 SL 模型预测 PCI 术预后情况,并根据模型原理特征设立不同组合的 SL 模型进行评价,借助权重系数、均方误差对 SL 模型中的单一模型、变量贡献进行分析,从而筛选出适合 PCI 预后的预测模型组合及变量,为临床 PCI 术后患者个性化治疗方案提供参考。

1 对象与方法

1.1 对象

纳入 2018 年 1 月—2022 年 6 月江西省南昌大学第二附属医院所有符合 PCI 手术标准并行 PCI 术的心肌梗死患者,疾病诊断分类标准依据 ICD-10 的分类。手术指征包括慢性稳定型冠心病有较大范围心肌缺血证据、不稳定型心绞痛、非 ST 段抬高性心肌梗死及急性 ST 段抬高性心肌梗死。排除有 PCI 治疗史或因其他心脏病引起的并发症的患者。

1.2 数据收集

从医院管理信息系统(HIS)中获取患者的相关信息(人口学特征、病史、血液检查、手术相关信息等),收集到的临床数据及流行病学数据见表 1。同时对患者的手术预后情况进行随访,随访起点时

间为 2018 年 1 月 1 日—2021 年 6 月 30 日,末次随访时间为 2022 年 6 月 30 日。失访定义为截至 2022 年 6 月 30 日进行 3 次电话随访无回应则纳入失访数据。考虑到失访数据 < 5%,属于完全随机缺失,故直接进行删除处理。结局指标采用 MACE 和首次出现 MACE 的时间,MACE 定义为心源性死亡、心肌梗死、心绞痛发作、心力衰竭、血运重建、恶性心律失常、支架内血栓形成等^[18]。

1.3 SL

SL 是一种基于损失函数的估计器^[13],并通过多折交叉验证组合多个的参数、半参数模型或其他合适的模型,SL 首先根据提供的数据自动选择初始候选预测模型的函数形式,并使用损失函数均方误差对候选模型及组合模型进行评价,同时通过系数对各个模型赋予不同的权重,从而得出一个最佳的组合模型。目前 SL 中包括多个模型^[19],如人工神经网络、递归分割树、支持向量机、随机森林、极端随机树、xgboost、广义加性模型及 GBM 等。此外,训练后得到的 SL 模型可以通过逐一剔除自变量的方式计算 MSE 的变化,从而得出各个自变量在模型中的变量重要性。

我们在 SL 模型中纳入了以下几种算法:classification and regression training(caret)、random-Forest、cforest、earth(Multivariate Adaptive Regression Splines)、广义线性模型(glm)、广义相加模型(gam)、AIC 逐步回归(step)、岭回归(ridge)、正则化回归(glmnet)、xgboost 算法、Bagging 算法(ipredbagg)、梯度提升机(gbm)、非负最小二乘回归(nnl)、支持向量机(svm)、线性回归模型(lm)。

本研究将年龄、性别、BMI、Killip 分级、既往疾病(高血压、高脂血症、糖尿病、肾功能不全、肺部感染)、脉搏、收缩压、舒张压、吸烟、饮酒、家族病史(糖尿病、高血压、冠心病)、住院次数、病变冠脉支数、心电图(窦性心律、心房颤动、起搏心律、高度或三度房室传导阻滞、ST 段变化、完全性左束支传导阻滞、完全性右束支传导阻滞、异常 Q 波、左心室高电压、T 波改变)、血液检查[BNP、AST、CK、CKMB、血清血肌酐(Cre)、eGFR、肌酐清除率(Scr)、K_t、血管狭窄程度(左主干、左前降支、左回旋支、右冠)、TIMI 血流分级(左主干、左前降支、左回旋支、右冠)、PCI 手术信息(心脏骤停、发病至 PCI 术时间、介入途径、手术方法)等作为解释变量纳入研究。首先对收集到的患者数据进行整理清洗,剔除信息缺失较多(>5%)的研究对象,对缺失较少的数据进行数据插补。再通过 Cox 回归变量筛选进一步简化模型结构,并按照 75%的比例划分训练集和测试集。基于 5 折交叉验证训练不同组合的 SL 模型,并绘制 ROC 曲线及 PR 曲线进行模型评价。解释变量的重要性由逐一剔除解释变

量后模型的均方误差计算而得,剔除前后的 MSE 变化越大说明该变量的重要性更高。

1.4 统计学处理

所有的统计分析均基于 R 4.2.1 软件进行分析处理,对连续型变量采取 $M(P_{25}, P_{75})$ 表示,分类变量采用频数及构成比进行描述。通过 R 软件 mean 函数计算数据缺失比例为 3.78%, <5% 可

认为是完全随机缺失,故对缺失数据进行数据插补。连续型变量缺失值采用 mice 包^[20]中的预测均值匹配填补,而分类型变量则采用分类回归树方法填补。解释变量采用多因素 Cox 回归初步筛选 ($P < 0.05$),数据集通过 cart 包^[21]进行训练集和测试集的划分,借助 SL 包建立 SL 预测模型。

表 1 患者基本特征

Table 1 General data

例(%), $M(P_{25}, P_{75})$

项目	总人数(3 880 例)	MACE 组(706 例)	非 MACE 组(3 174 例)
年龄/岁	65(57,72)	65(58,73)	65(57,72)
男性	2 862(73.8)	543(76.9)	2 319(73.1)
既往病史			
高血压	2 220(57.2)	423(59.9)	1 797(56.6)
高脂血症	1 117(28.8)	192(27.2)	925(29.1)
糖尿病	1 151(29.7)	235(33.3)	916(28.9)
肾功能不全	523(13.5)	105(14.9)	418(13.2)
肺部感染	353(9.1)	69(9.8)	284(8.9)
吸烟史	1 324(34.1)	247(35.0)	1 077(33.9)
饮酒史	933(24.0)	172(24.4)	761(24.0)
家族史			
糖尿病	83(2.1)	21(3.0)	62(2.0)
高血压	161(4.1)	30(4.2)	131(4.1)
冠心病	71(1.8)	16(2.3)	55(1.7)
心电图			
窦性心律	3 701(95.4)	673(95.3)	3 028(95.4)
心房颤动	172(4.4)	30(4.2)	142(4.5)
起搏心律	374(9.6)	67(9.5)	307(9.7)
三度及高度房室传导阻滞	76(2.0)	15(2.1)	61(1.9)
ST 段改变	1 623(41.8)	296(41.9)	1 327(41.8)
完全性左束支传导阻滞	60(1.5)	13(1.8)	47(1.5)
完全性右束支传导阻滞	225(5.8)	43(6.1)	182(5.7)
异常 Q 波	1 058(27.3)	201(28.5)	857(27.0)
左心室高电压	545(14.0)	100(14.2)	445(14.0)
T 波倒置	1 131(29.1)	221(31.3)	910(28.7)
Kiliip 分级			
I	249(6.4)	43(6.1)	206(6.5)
II	1 772(45.7)	309(43.8)	1 463(46.1)
III	1 508(38.9)	303(42.9)	1 205(38.0)
IV	154(4.0)	28(4.0)	126(4.0)
住院次数			
1 次	2 319(59.8)	65(9.2)	2 254(71.0)
2 次	857(22.1)	344(48.7)	513(16.2)
3 次	434(11.2)	222(31.4)	212(6.7)
4 次	92(2.4)	36(5.1)	56(1.8)
病变冠脉支数			
1 支	923(23.8)	111(15.7)	812(25.6)
2 支	1 175(30.3)	181(25.6)	994(31.3)
3 支	1 568(40.4)	354(50.1)	1 214(38.2)
4 支	214(5.5)	60(8.5)	154(4.9)

续表 1

项目	总人数(3 880 例)	MACE 组(706 例)	非 MACE 组(3 174 例)
BMI/(kg/m ²)	23.9(22.0,25.7)	23.9(22.0,25.7)	23.9(22.0,25.7)
脉搏/(次/min)	87(80,94)	88(80,94)	87(80,94)
收缩压/mmHg	129(115,145)	129(115,145)	129(115,145)
舒张压/mmHg	70(63,77)	70(63,76)	70(63,78)
血液检查			
BNP/(pg/mL)	150.3(50.0,445.7)	186.4(62.7,527.5)	144.5(48.4,424.3)
AST/(U/L)	28.3(21.2,54.1)	27.9(20.8,50.5)	28.3(21.2,54.8)
CK/(U/L)	126.5(82.0,317.9)	124.3(78.2,325.1)	126.9(82.9,317.0)
CKMB/(U/L)	19.8(13.7,39.9)	18.9(13.2,39.1)	19.9(13.9,40.0)
Cre/(mmol/L)	81.8(68.6,99.7)	84.5(71.8,104.6)	81.1(67.8,98.6)
eGFR/[mL·min ⁻¹ ·(1.73m ²) ⁻¹]	82.4(65.1,99.0)	78.9(61.1,95.9)	83.1(65.8,99.8)
Scr/(mL/min)	66.9(48.0,87.5)	64.5(46.8,83.3)	67.4(48.2,87.9)
K/(mmol/L)	3.9(3.6,4.2)	3.9(3.6,4.2)	3.9(3.6,4.2)
PCI 手术信息			
心脏骤停	58(1.5)	16(2.3)	42(1.3)
从发病到 PCI 的时间			
<3 h	189(4.9)	39(5.5)	150(4.7)
3~6 h	202(5.2)	42(5.9)	160(5.0)
6~9 h	362(9.3)	75(10.6)	287(9.0)
9~12 h	359(9.3)	65(9.2)	294(9.3)
>12 h	2 630(67.8)	453(64.2)	2 177(68.6)
介入途径			
股动脉	178(4.6)	39(5.5)	139(4.4)
桡动脉	3 676(94.7)	662(93.8)	3 014(95.0)
尺动脉	26(0.7)	5(0.7)	21(0.7)
手术方式			
血栓抽吸	44(1.1)	9(1.3)	35(1.1)
PTCA	3 595(92.7)	657(93.1)	2 938(92.6)
支架植入	241(6.2)	40(5.7)	201(6.3)

2 结果

本次研究共收集到 3 880 例行 PCI 术患者,中位随访时间为 603 d,随访出现 MACE 的患者占 18.2%,中位年龄为 65 岁,其中大部分是男性(73.8%)和高血压(57.2%),少部分患有高脂血症(28.8%)、糖尿病(29.7%)、肾功能不全(13.5%)。部分患者存在吸烟(34.1%)和饮酒(24.0%)的习惯。而存在家族史的患者较少。心电图结果显示 95.4% 患者为窦性心律,且多数 Killip 分级为 II 级(45.7%)和 III 级(38.9%)。71.0% 的未发生 MACE 患者住院次数仅为 1 次,而仅有 9.2% 的发生 MACE 患者住院次数为 1 次。见表 1。

以随访时间和 MACE 作为结局变量,将变量纳入 Cox 回归中,共筛选出 24 个对结局有影响的变量,并将其纳入 SL 模型中。由于结局变量 MACE 占比过小,属于不平衡数据^[22],故在模型构建前对数据进行欠采样处理,最终纳入模型的样本

量为 2 000 例(非 MACE 组 1 294 例,MACE 组 706 例)。再随机划分训练集(1 501 例)和测试集(499 例),根据 SL 模型中包含的单一模型特征进行分组建立不同的混合模型(模型 1: caret、randomForest、cforest、earth;模型 2: glm、gam、step;模型 3: ridge、glmnet;模型 4: xgboost、ipredbag、gbm;模型 5: nnls、svm、lm),通过训练集对 5 个模型进行逐一训练,训练结果如图 1 所示,模型 1 中系数不为 0 的模型为 randomForest(coef=0.248)和 cforest(coef=0.7516),模型 2 中系数不为 0 的模型为 glm(coef=0.220)和 step(coef=0.780),模型 3 中系数不为 0 的模型仅为 glmnet,模型 4 中系数不为 0 的模型为 xgboost(coef=0.023)、ipredbag(coef=0.324)及 gbm(coef=0.653),模型 5 中系数不为 0 的模型为 nnls(coef=0.174)、svm(coef=0.444)及 lm(coef=0.382)。5 个模型的五折交叉验证风险及置信区间见图 1。

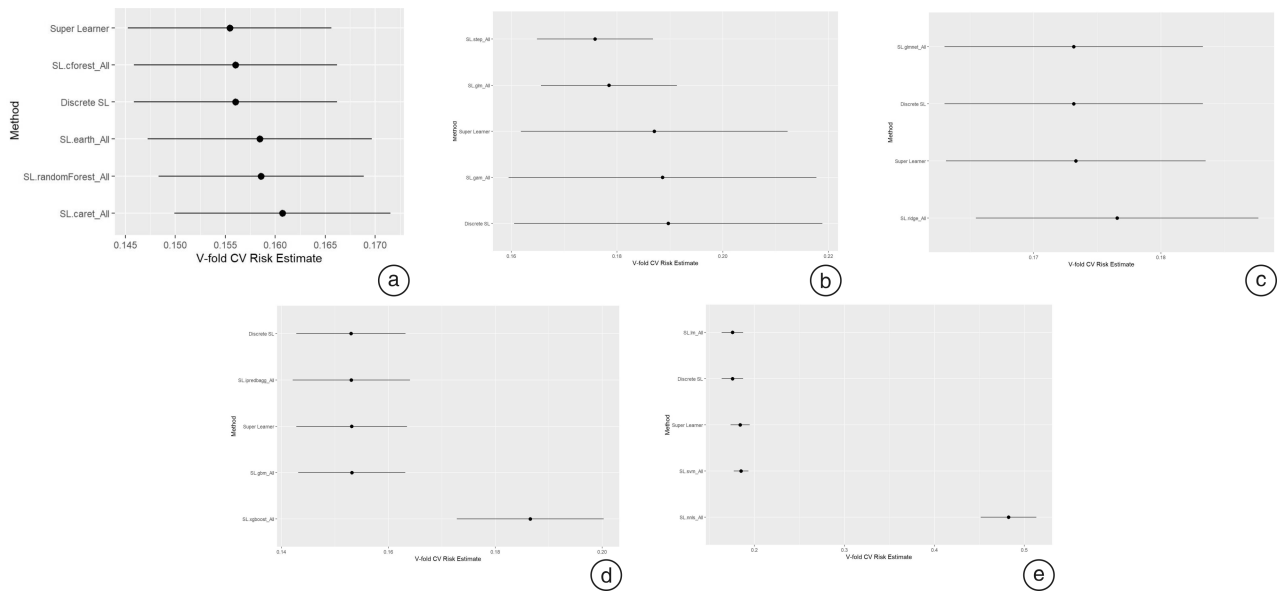


图 1 各训练模型的五折交叉验证风险图

Figure 1 Cross validation risk diagram of each training model

再根据训练集在各单一模型上的表现建立 4 个集成模型,集成模型 1 和集成模型 3 分别是选取系数最高的和系数不为 0 的单一模型组合,而集成模型 2 和集成模型 4 则依次是集成模型 1 和集成模型 3 在训练集中再次训练后剔除系数为 0 后的模型组合,具体集成模型所包含的单一模型见表 2。在测试集上分别测试 4 个集成模型的预测能力并绘制 ROC 图及 PR 图。结果显示,所有集成模型在测试集上表现均较好, AUC0.843~0.846, PRC0.721~0.729, 其中集成模型 1 表现最佳 (AUC:0.846, PRC:0.729), 见图 2。混淆矩阵结果显示,准确度最高的为集成模型 3(0.8136, 95% CI 0.7766~0.8469), 灵敏度最高的为集成模型 1 (0.8109), 特异度最高的为集成模型 3(0.8864) 和集成模型 4(0.8864)。

表 2 各集成模型的组成

Table 2

集成模型	Em1	Em2	Em3	Em4
randomForest	×	×	√	×
cforest	√	√	√	×
glm	×	×	√	×
step	√	×	√	×
glmnet	√	×	√	×
xgboost	×	×	√	√
ipredbag	×	×	√	√
gbm	√	√	√	√
nnls	×	×	√	√
svm	√	×	√	×
lm	×	×	√	×

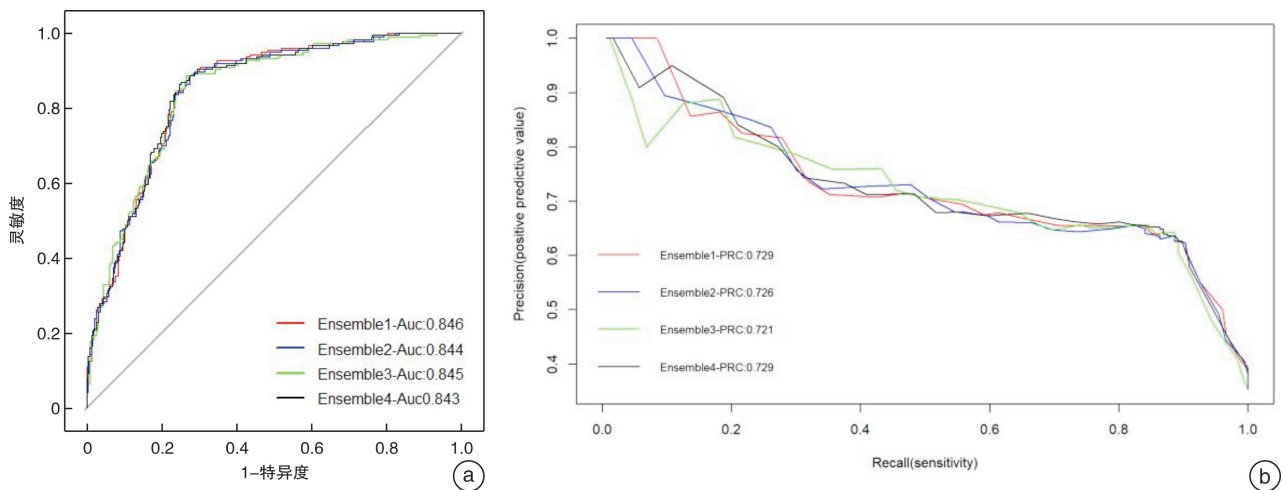
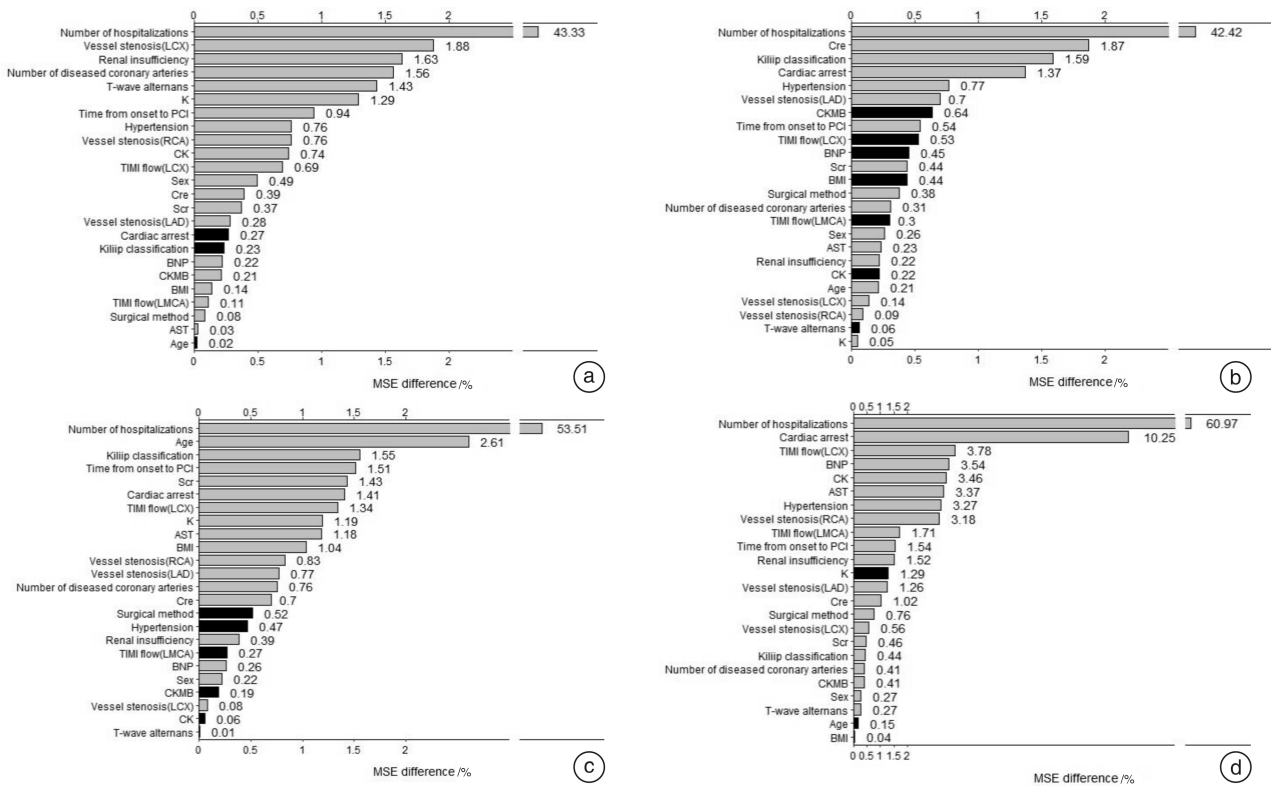


图 2 各集成模型在测试集中的 ROC 图和 PR 图

Figure 2 ROC diagram and PR diagram of each integration model in the test set

通过逐一剔除变量的方法计算各纳入变量的重要性,并进行排序,各集成模型的变量重要性排序见图 3。结果显示,24 个纳入变量中住院次数、左回旋支血管狭窄程度、肾功能不全、病变冠脉支数、Cre、Kiliip 分级、是否心脏骤停、就诊时年龄、发

病到 PCI 时间、左回旋支血流 TIMI 分级和 BNP 是重要的预测变量。4 个集成模型均认为住院次数为最重要的预测变量。住院次数、病变冠脉支数、左回旋支血管狭窄程度和是否肾功能不全是预测 MACE 发生的危险因素。



黑色标记为与 MACE 负相关的变量。

图 3 各集成模型的变量重要性排序

Figure 3 Importance ranking of variables in each integration model

3 讨论

本研究旨在探讨 SL 模型在 PCI 术预后上的应用价值及 PCI 术后患者 MACE 发生的相关危险因素。我们发现住院次数、病变冠脉支数、左回旋支血管狭窄程度和肾功能不全是 PCI 术后 MACE 发生的危险因素。

目前国内外的 PCI 预后研究将各类机器学习模型应用在预测 MACE 上。Blum 等^[23]运用调整后的 Cox 回归模型评估高敏 C 反应蛋白(hsCRP)对 PCI 术后 MACE 发生的影响。Liu 等^[24]分析比较了 6 种不同的模型(支持向量机、决策树、随机森林、梯度提升决策树、神经网络和逻辑回归)对 PCI 术长期预后进行预测。Kulkarni 等^[25]则建立了一个基于人工神经网络的预测模型,在测试集中的准确度均达到 80% 以上。不同类型的模型在预测 PCI 术后表现出不同效果,具有各自的优缺点。集成学习将单一模型进行混合学习,融合了各个单一模型的优点,有效提高预测模型的准确度和适用性。多项研究均表明,集成机器学习模型往往优于

单一预测模型^[26-27]。而 SL 属于集成方法中的堆叠泛化,即将几个不同的预测模型算法组合成一个集成模型,再使用 V 折交叉验证从候选算法库中构建预测的最佳加权组合,从而提高最终模型的预测准确性^[28]。研究显示 SL 在生存预测^[29]、疾病严重程度预测^[30]上表现均较佳。

本研究最终得到的 4 个集成 SL 模型 AUC 值均大于 0.84,且 PRC 值均在 0.72 以上,提示建立的这 4 种集成预测模型能有效预测 PCI 术后 MACE 的发生。在测试集的预测上,集成模型 1 和 3 表现较好,其中集成模型 1 在识别发生 MACE 病例上拟合较好,而集成模型 3 在识别不发生 MACE 病例上拟合较好。因此,在对不同模型进行集成学习时,可单独提取出相对更适合的模型,这样既减少了预测模型的复杂度又提高了模型的预测能力。而集成模型 1 和 3 预测表现均优于集成模型 2 和 4,说明在针对模型结构进行二次调整后并未对模型的整体预测能力有所提高,反而是降低了其预测准确性。

我们确定住院次数是PCI术后发生MACE的一个重要的危险因素。这与Biswas等^[31]在PCI前入院次数和PCI后心脏疾病发生的关系一致。造成这种关联的原因有很多。大部分研究显示,这可能是由于不同疾病导致的心脏问题处理前一次或多次入院,而这些既往疾病史也引起了患者健康状况的变化,提高了出院后心脏问题的复发率,从而增加了患者的再入院风险^[32-34]。此外,Dibben等^[35]发现随着住院次数的增加,患有心脏疾病的患者运动时间也在下降,从而引起心肌梗死的风险增加,并大幅度提高了全因住院及小幅度提高全因死亡率。

肾功能不全也被发现是PCI预后的一个危险因素。我们的研究结论与多项研究结果相符合^[36-40]。肾可致心脏病,心脏可致肾病。心脏和肾脏疾病在急性和慢性环境中以复杂的双向和相互依赖的方式相互作用。肾功能不全的患者随着肾功能的下降及肾功能衰竭的发生,常引起蛋白尿、血尿的症状,导致贫血、高血压等血管疾病的发生,从而提高了心血管疾病的发生风险^[41]。目前对于心脏和肾之间的相互作用认为存在3种主要机制^[42],一种是血流动力学机制,即人体内水和盐的储留导致了心脏和肾的静脉充血,而肾静脉的充血正是肾功能不全的关键;第2种是神经激素机制,如经典的心力衰竭代偿途径,通过同时激活肾素-血管紧张素-醛固酮系统(RAAS)和交感神经系统进行作用;第3种则是与心血管疾病相关的机制,包括局部炎症的发生、免疫系统的异常、贫血及酸碱失衡等。此外,PCI术中所使用的造影剂对肾功能也存在一定的影响,如造影剂肾病的发生。

与之前的研究相比,我们还额外发现了病变冠脉支数与MACE的发生之间的关联。一些研究并没有证实病变的冠脉数量与心血管不良事件之间的关系。我们认为可能存在以下几种原因。首先是该项随机试验纳入的研究对象是多血管冠状动脉疾病患者和ST段抬高型心肌梗死患者,而我们研究所纳入的患者是心肌梗死患者,通过收集病变的冠脉数量进行分析比较,患者之间的可比性可能更高。其次是比较的结局指标存在不同,我们研究的预后结局是MACE,而该项随机对照试验收集的是主要心脑血管不良事件。

此次研究依然存在以下几点不足之处。首先病例数据的收集是在PCI治疗方面经验丰富的三甲医院进行的,不具有很好的外部有效性,模型及变量在其他医疗机构适用程度可能会存在影响。其次是纳入变量大多为PCI术前检查数据,并未纳入术后数据,而PCI术后数值的变化是否对MACE的发生存在影响还有待讨论。最后,尽管对预测模型的预测性能有所提升,但所纳入的算法

多样性导致了其模型的可解释性不高,且不同模型中的各变量的排序存在变化,影响模型稳定性。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 国家心血管病中心. 中国心血管健康与疾病报告2021[M]. 北京:科学出版社,2022:100.
- [2] Al-Lamee RK, Nowbar AN, Francis DP. Percutaneous coronary intervention for stable coronary artery disease[J]. Heart, 2019, 105(1): 11-19.
- [3] Teoh Z, Al-Lamee RK, Courage O. Percutaneous Coronary Intervention for Stable Coronary Artery Disease [J]. Interv Cardiol Clin, 2020, 9(4): 469-482.
- [4] Groenland FTW, Neleman T, Kakar H, et al. Intravascular ultrasound-guided versus coronary angiography-guided percutaneous coronary intervention in patients with acute myocardial infarction: A systematic review and meta-analysis[J]. Int J Cardiol, 2022, 353: 35-42.
- [5] Abtan J, Wiviott SD, Sorbets E, et al. Prevalence, clinical determinants and prognostic implications of coronary procedural complications of percutaneous coronary intervention in non-ST-segment elevation myocardial infarction: Insights from the contemporary multinational TAO trial [J]. Arch Cardiovasc Dis, 2021, 114(3): 187-196.
- [6] 刘健迪, 龚韧, 夏华松, 等. 高密度脂蛋白胆固醇水平对药物洗脱支架植入后早期血管愈合的影响[J]. 临床心血管病杂志, 2022, 38(2): 123-131.
- [7] 徐慧慧, 谢艳辉, 宋恒良, 等. PCSK9 抑制剂对急性ST段抬高型心肌梗死PCI术后的微循环功能及左心室重构的影响[J]. 临床心血管病杂志, 2022, 38(1): 22-28.
- [8] Xu B, Tu S, Song L, et al. Angiographic quantitative flow ratio-guided coronary intervention (FAVOR III China): a multicentre, randomised, sham-controlled trial[J]. Lancet, 2021, 398(10317): 2149-2159.
- [9] Madhavan MV, Kirtane AJ, Redfors B, et al. Stent-related adverse events >1 year after percutaneous coronary intervention[J]. J Am Coll Cardiol, 2020, 75(6): 590-604.
- [10] Chen Z, Duan J, Kang L, et al. Class-imbalanced deep learning via a class-balanced ensemble [J]. IEEE Trans Neural Netw Learn Syst, 2021, 22: 100.
- [11] El Asnaoui K. Design ensemble deep learning model for pneumonia disease classification [J]. Int J Multimed Inf Retr, 2021, 10(1): 55-68.
- [12] Chen Z, Li Q, Li R, et al. Ensemble learning accurately predicts the potential benefits of thrombolytic therapy in acute ischemic stroke [J]. Quant Imaging Med Surg, 2021, 11(9): 3978-3989.
- [13] van der Laan MJ, Polley EC, et al. Super learner [J]. Stat Appl Genet Mol Biol, 2007, 6: 25.
- [14] Wardenaar KJ, Riese H, Giltay EJ, et al. Common and specific determinants of 9-year depression and anxiety course-trajectories: A machine-learning investigation

- in the Netherlands Study of Depression and Anxiety (NESDA)[J]. *J Affect Disord*, 2021, 293: 295-304.
- [15] Cartus AR, Naimi AI, Himes KP, et al. Can ensemble machine learning improve the accuracy of severe maternal morbidity screening in a perinatal database? [J]. *Epidemiology*, 2022, 33(1): 95-104.
- [16] Murnane PM, Ayieko J, Vittinghoff E, et al. Machine learning algorithms using routinely collected data do not adequately predict viremia to inform targeted services in postpartum women living with HIV[J]. *J Acquir Immune Defic Syndr*, 2021, 88(5): 439-447.
- [17] 徐芳, 徐俊杰, 吴春苑, 等. 左室心肌做功对急性心肌梗死患者急诊 PCI 术后 MACE 发生的预测价值[J]. *临床心血管病杂志*, 2022, 38(12): 960-966.
- [18] Yiannoullou P, Summers A, Goh SC, et al. Major adverse cardiovascular events following simultaneous pancreas and kidney transplantation in the united kingdom[J]. *Diabetes Care*, 2019, 42(4): 665-673.
- [19] Polley E, LeDell E, Kennedy C, van der Laan M. SuperLearner: SuperLearner Prediction [J]. *Rpackage Version*, 20210, 20: 100.
- [20] Stef van B, Karin GO. Multivariate imputation by chained equations in R[J]. *J Statist Software*, 2011, 45(3), 1-67.
- [21] Kuhn M. Classification and regression training [J]. *Rpackage Version*, 2022, 20: 100.
- [22] He H, Garcia EA. Learning from imbalanced data[J]. *IEEE Trans Knowl Data Eng*, 2009, 21(9): 1263-1284.
- [23] Blum M, Cao D, Chandiramani R, et al. Prevalence and prognostic impact of hsCRP elevation are age-dependent in women but not in men undergoing percutaneous coronary intervention[J]. *Catheter Cardiovasc Interv*, 2021, 97(7): E936-E944.
- [24] Liu S, Yang S, Xing A, et al. Machine learning-based long-term outcome prediction in patients undergoing percutaneous coronary intervention[J]. *Cardiovasc Diagn Ther*, 2021, 11(3): 736-743.
- [25] Kulkarni H, Thangam M, Amin AP. Artificial neural network-based prediction of prolonged length of stay and need for post-acute care in acute coronary syndrome patients undergoing percutaneous coronary intervention[J]. *Eur J Clin Invest*, 2021, 51(3): e13406.
- [26] Druchok M, Yarish D, Garkot S, et al. Ensembling machine learning models to boost molecular affinity prediction[J]. *Comput Biol Chem*, 2021, 93: 107529.
- [27] El Asnaoui K. Design ensemble deep learning model for pneumonia disease classification [J]. *Int J Multimed Inf Retr*, 2021, 10(1): 55-68.
- [28] Naimi AI, Balzer LB. Stacked generalization: an introduction to super learning [J]. *Eur J Epidemiol*, 2018, 33(5): 459-464.
- [29] Golmakani MK, Polley EC. Super learner for survival data prediction [J]. *Int J Biostat*, 2020, 20: 100.
- [30] Ehwerhemuepha L, Danioko S, Verma S, et al. A super learner ensemble of 14 statistical learning models for predicting COVID-19 severity among patients with cardiovascular conditions [J]. *Intell Based Med*, 2021, 5: 100030.
- [31] Biswas S, Dinh D, Lucas M, et al. Incidence and predictors of unplanned hospital readmission after percutaneous coronary intervention [J]. *J Clin Med*, 2020, 9(10): 3242.
- [32] Harjai KJ, Thompson HW, Turgut T, et al. Simple clinical variables are markers of the propensity for re-admission in patients hospitalized with heart failure [J]. *Am J Cardiol*, 2001, 87(2): 234-239.
- [33] Hummel SL, Katrapati P, Gillespie BW, et al. Impact of prior admissions on 30-day readmissions in medicare heart failure inpatients [J]. *Mayo Clin Proc*, 2014, 89(5): 623-630.
- [34] Kilkenny MF, Dewey HM, Sundararajan V, et al. Re-admissions after stroke: linked data from the Australian Stroke Clinical Registry and hospital databases [J]. *Med J Aust*, 2015, 203(2): 102-106.
- [35] Dibben G, Faulkner J, Oldridge N, et al. Exercise-based cardiac rehabilitation for coronary heart disease [J]. *Cochrane Database Syst Rev*, 2021, 11(11): CD001800.
- [36] Qi L, Liu H, Cheng L, et al. Impact of renal insufficiency on prognosis of patients with acute coronary syndrome [J]. *Int J Gen Med*, 2021, 14: 8919-8927.
- [37] Chiang CY, Huang SC, Chen M, et al. Effects of renal impairment on cardiac remodeling and clinical outcomes after myocardial infarction [J]. *Int J Med Sci*, 2021, 18(13): 2842-2848.
- [38] Deguchi Y, Saito Y, Nakao M, et al. Trajectory of renal function change and kidney injury after percutaneous coronary intervention in patients with stable coronary artery disease [J]. *Heart Vessels*, 2021, 36(3): 315-320.
- [39] Khoury S, Margolis G, Ravid D, et al. Outcomes of early and reversible renal impairment in patients with ST segment elevation myocardial infarction undergoing percutaneous coronary intervention [J]. *Eur Heart J Acute Cardiovasc Care*, 2020, 9(7): 684-689.
- [40] Sciahbasi A, Cuono A, Marrangoni A, et al. Acute kidney injury and multivessel percutaneous coronary interventions in chronic renal disease: the AMICI study [J]. *Minerva Cardiol Angiol*, 2021, 69(5): 491-498.
- [41] Ammirati AL. Chronic kidney disease [J]. *Rev Assoc Med Bras*, 2020, 66: s03-s09.
- [42] Schefold JC, Filippatos G, Hasenfuss G, et al. Heart failure and kidney dysfunction: epidemiology, mechanisms and management [J]. *Nat Rev Nephrol*, 2016, 12(10): 610-623.