

论著·临床研究

机器学习预测乳腺癌新辅助治疗后炎症代谢状态改变的模型评价

吴其臻*, 刘启明*, 柴烨子, 陶政宇, 王依楠, 郭欣宁, 姜 萌#, 卜 军#

上海交通大学医学院附属仁济医院心内科, 上海 200127

[摘要] **目的**·通过机器学习方法,采用临床常见实验室指标及心脏彩色多普勒超声指标,在乳腺癌患者中探究早期识别并预测新辅助治疗后发生与代谢状态改变相关的心血管疾病高风险患者的方案。**方法**·连续入选2020年9月—2022年9月在上海交通大学医学院附属仁济医院乳腺外科确诊的原发性浸润性乳腺癌女性患者。收集并记录患者的一般情况、实验室检查结果及心脏彩色多普勒超声结果。经过特征提取后,分别应用梯度增强 (gradient boost, GB)、支持向量机 (support vector machine, SVM)、决策树 (decision tree, DT)、K-近邻 (K-nearest neighbour, KNN) 及随机森林 (random forest, RF) 5种机器学习方法构建新辅助治疗后患者炎症代谢状态改变预测模型,并比较5种模型的预测性能。**结果**·最终纳入232例有效临床数据,其中135例为新辅助治疗前,97例为完成4个周期的新辅助治疗后。特征提取筛选出白细胞计数、血红蛋白、高密度脂蛋白、白细胞介素-2受体和白细胞介素-8这5项特征。在多特征分析中,白细胞计数+血红蛋白+高密度脂蛋白的受试者操作特征曲线下面积高于白细胞介素-2受体+白细胞介素-8 (RF: 0.928 vs 0.772; GB: 0.900 vs 0.792; SVM: 0.941 vs 0.764; KNN: 0.907 vs 0.762; DT: 0.799 vs 0.714),并且在RF、SVM、GB模型中的曲线下面积 (0.928、0.941、0.900) 及准确率 (0.914、0.897、0.776) 较高;与RF、GB模型 ($P=0.122$, $P=0.097$) 相比,SVM模型在训练集数据上校准度较好 ($P=0.394$)。**结论**·SVM模型可通过纳入白细胞计数、血红蛋白、高密度脂蛋白、白细胞介素-2受体、白细胞介素-8这5项临床常见指标,在乳腺癌患者中建立早期预测新辅助治疗后代谢状态改变相关心血管疾病预防风险的预测模型,可能有助于临床上建立基于患者炎症代谢状态的个体化筛查方案。

[关键词] 乳腺癌; 新辅助治疗; 机器学习; 支持向量机**[DOI]** 10.3969/j.issn.1674-8115.2024.09.012 **[中图分类号]** R541.9 **[文献标志码]** A

Evaluation of machine learning prediction of altered inflammatory metabolic state after neoadjuvant therapy for breast cancer

WU Qizhen*, LIU Qiming*, CHAI Yezi, TAO Zhengyu, WANG Yinan, GUO Xinning, JIANG Meng#, PU Jun#

Department of Cardiology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200127, China

[Abstract] **Objective**·To develop a machine learning approach for early identification of metabolic syndromes associated with inflammatory metabolic state changes in breast cancer patients after neoadjuvant therapy, using common laboratory and transthoracic echocardiography indices. **Methods**·Female patients with primary invasive breast cancer diagnosed at the Department of Breast Surgery, Renji Hospital, Shanghai Jiao Tong University School of Medicine, between September 2020 and September 2022, were included. General patient information, laboratory test results, and transthoracic echocardiography data were collected. After feature extraction, five machine learning algorithms, including random forest (RF), gradient boosting (GB), support vector machine (SVM), K-nearest neighbor (KNN), and decision tree (DT), were applied to construct a prediction model for the changes of the patients' metabolic state after neoadjuvant therapy, and the prediction performances of the five models were compared.

[基金项目] 国家自然科学基金 (U21A20341, 81971570); 上海市科学技术委员会基金 (20Y11910500); 上海市优秀学术/技术带头人计划项目 (21XD1432100); 上海申康医院发展中心临床三年行动计划项目 (SHDC2020CR2025B); 上海市肿瘤研究所基金 (ZZ-20-22SYL); 上海交通大学医学院“双百人”项目 (20172014)。

[作者简介] 吴其臻 (1997—), 女, 博士生; 电子信箱: wuqizhen@sjtu.edu.cn。刘启明 (1996—), 男, 硕士; 电子信箱: 090503liu@sjtu.edu.cn。*为共同第一作者。

[通信作者] 姜 萌, 电子信箱: jiangmeng0919@163.com。卜 军, 电子信箱: pujun310@hotmail.com。#为共同通信作者。

[Funding Information] National Natural Science Foundation of China (U21A20341, 81971570); Project of Science and Technology Commission of Shanghai Municipality (20Y11910500); Advanced Technology Leader of Science and Technology Commission of Shanghai Municipality (21XD1432100); Three-year Action Plan of Shanghai Shenkang Hospital Development Center (SHDC2020CR2025B); Project of Shanghai Cancer Institute (ZZ-20-22SYL); “Two-hundred Talents” Program of Shanghai Jiao Tong University School of Medicine (20172014).

[Corresponding Author] JIANG Meng, E-mail: jiangmeng0919@163.com. PU Jun, E-mail: pujun310@hotmail.com. #Co-corresponding authors.



Results A total of 232 cases with valid clinical data were included, comprising 135 cases before neoadjuvant therapy and 97 cases after completing 4 cycles of neoadjuvant therapy. Feature extraction identified five key features: white blood cell count, hemoglobin, high-density lipoprotein (HDL), interleukin-2 receptor, and interleukin-8. In the multi-feature analysis, the area under the receiver operating characteristic curve (AUC) was higher in the combination of white blood cell count, hemoglobin and HDL compared to the combination of interleukin-2 receptor and interleukin-8 (RF: 0.928 vs 0.772, GB: 0.900 vs 0.792, SVM: 0.941 vs 0.764, KNN: 0.907 vs 0.762, DT: 0.799 vs 0.714). The RF, SVM, and GB models showed higher AUC (0.928, 0.941, 0.900) and accuracy (0.914, 0.897, 0.776). The SVM model exhibited superior accuracy in the training data compared to the RF and GB models ($P=0.394$, 0.122 and 0.097, respectively). **Conclusion** The SVM model can be used to establish a prediction model for identifying breast cancer patients at high risk of developing inflammatory metabolic state-related metabolic syndrome after neoadjuvant therapy by incorporating five common clinical indicators, namely, white blood cell count, hemoglobin, high-density lipoprotein, interleukin-2 receptor, and interleukin-8. SVM modeling may be useful for clinicians to establish individualized screening protocols based on a patient's inflammatory metabolic state.

[Key words] breast cancer; neoadjuvant therapy; machine learning; support vector machine

2021年及2024年世界卫生组织发布的全球癌症负担数据^[1-2]显示,近年来乳腺癌发病率和病死率居高不下,全球发病率仅次于肺癌,已经成为女性第一大恶性肿瘤。蒽环类、紫杉类及人表皮生长因子受体-2 (human epidermal growth factor receptor-2, HER-2) 靶向药物新辅助治疗的开展延长了早期可手术乳腺癌女性患者的无病生存期和总生存期。然而, RAMIN等^[3]的研究显示,与未患癌者相比,乳腺癌患者发生心血管疾病的风险和死亡率明显升高。该研究在前瞻性 CLUE II 队列中对 628 例乳腺癌女性患者及 3 140 名未患癌女性进行了长达 25 年的随访,发现与未患癌女性相比,心血管疾病相关的死亡风险在确诊乳腺癌 8 年后升高 1.65 倍,是乳腺癌患者的第二大死因。因此,早期识别患者在新辅助治疗后的代谢特征,有助于及时发现心血管疾病风险升高的患者,并定制个体化早期保护方案,以降低潜在继发于代谢状态改变的心血管疾病的发生率。

肿瘤幸存者患心血管疾病的风险受多种因素的影响。目前构建预测心血管事件风险模型的研究,多纳入患者基线时的心血管危险因素、年龄、用药种类及剂量、癌症分级等,但尚未将治疗前后参数纳入模型比较。理论上患者在新辅助治疗后出现代谢状态改变的概率增加,从而进一步提高了患心血管疾病的概率,但在基线临床指标上变化可能不明显。

机器学习是一种通过使用算法模拟学习和归纳过程,并基于现有数据实现数据分类或预测的数据挖掘方法,已广泛用于医学图像分类和基于临床数据的疾病预测与预后风险评估^[4]。近年来已有在一般人群、子痫前期、代谢综合征等不同患病人群中使用机器学习方法预测罹患心血管疾病风险的回顾

性研究^[5-6]。YANG 等^[6]利用人口学与血清学指标,使用差分的方法进行特征筛选与建模,用于预测代谢综合征; WANG 等^[7]则利用子痫前期患者的基本信息与临床指标,实现了对子痫后患者心血管风险的评估。这些研究都集中于利用患者基线的人口学及血清学特征,并结合机器学习算法为临床服务。

本研究拟通过机器学习方法,基于乳腺癌队列和治疗前后数据,找到合适的指标组合,帮助快速判断患者目前炎症代谢状态,探究早期识别乳腺癌患者可能在新辅助治疗后出现的继发于代谢状态改变的心血管疾病高风险的机器学习方案,建立较为经济、简洁的预测模型,辅助临床决策,为制定肿瘤患者个性化随访及治疗方案提供参考。

1 对象与方法

1.1 研究对象

连续纳入 2020 年 9 月—2022 年 9 月在上海交通大学医学院附属仁济医院乳腺外科确诊的原发性浸润性乳腺癌女性患者。

入选标准: ① 在乳腺外科初诊为原发性浸润性乳腺癌的女性患者。② 年龄 ≥ 18 岁且 < 70 岁。③ 经乳腺外科及其他多学科评估可耐受新辅助治疗者。④ 在开始新辅助治疗前,经超声心动图评估左心室射血分数 $> 50\%$ 者。⑤ 签署知情同意书,愿意接受随访者。

排除标准: ① 冠心病患者。② 心肌病患者。③ 中重度心脏瓣膜病患者。④ 不稳定心律失常患者。⑤ 曾接受过抗癌药物治疗。⑥ 严重急性或慢性

肾功能衰竭(估算肾小球滤过率 $<30\text{ mL}/(\text{min}\cdot 1.73\text{ m}^2)$)者。⑦既往有免疫缺陷或器官移植史。⑧患有经研究者评估认为不适合参与研究的任何其他病症。

1.2 新辅助治疗方案

入组患者经过乳腺外科评估后接受4种新辅助治疗方案中的1种:①静脉给予紫杉醇($80\text{ mg}/\text{m}^2$,第1、8、15、22日)、顺铂($25\text{ mg}/\text{m}^2$,第1、8、15日)。②静脉给予紫杉醇($80\text{ mg}/\text{m}^2$,第1、8、15、22日)、顺铂($25\text{ mg}/\text{m}^2$,第1、8、15日)联合曲妥珠单抗抗体(首次剂量 $4\text{ mg}/\text{kg}$,其后每周 $2\text{ mg}/\text{kg}$)。③静脉给予紫杉醇($80\text{ mg}/\text{m}^2$,第1、8、15、22日)、顺铂($25\text{ mg}/\text{m}^2$,第1、8、15日)联合每日口服甲磺酸阿帕替尼 425 mg 。④静脉给予紫杉醇($80\text{ mg}/\text{m}^2$,第1、8、15、22日)、顺铂($25\text{ mg}/\text{m}^2$,第1、8、15日)联合每日口服 400 mg 吡咯替尼。

1.3 临床数据采集

新辅助治疗前后为同一批患者,收集的临床数据包括患者的一般情况和实验室检查结果,如身高、体质量、体质量指数(body mass index, BMI)、肿瘤位置、是否有淋巴结转移、新辅助治疗方案、白细胞计数(white blood cell count, WBC)、红细胞压积(hematocrit, HCT),以及血红蛋白(hemoglobin, HB)、高敏C反应蛋白(high sensitivity C-reactive protein, hsCRP)、肌钙蛋白I(troponin I, TNI)、B型钠尿肽(B-type natriuretic peptide, BNP)、N末端B型利钠肽原(blood N-terminal B-type natriuretic peptide proteins, NT-proBNP)、可溶性生长刺激表达基因2蛋白(soluble growth-stimulating expressed gene 2 protein, ST2)、三酰甘油(triacglycerol, TAG)、总胆固醇(total cholesterol, TC)、高密度脂蛋白(high-density lipoprotein, HDL)、低密度脂蛋白(low-density lipoprotein, LDL)、非高密度脂蛋白(non-high-density lipoprotein, NHDL)、空腹血糖(fasting plasma glucose, FPG)、谷草转氨酶(glutamic oxaloacetic transaminase, GOT)、谷丙转氨酶(glutamic pyruvic transaminase, GPT)、 γ -谷氨酰转肽酶(γ -glutamyltransferase, γ -GT)、白细胞介素-1 β (interleukin-1 β , IL-1 β)、白细胞介素-2受体(interleukin-2 receptor, IL-2R)、白细胞介素-6

(interleukin-6, IL-6)、IL-8、IL-10、肿瘤坏死因子 α (tumor necrosis factor- α , TNF- α)水平。

1.4 超声数据采集

入组患者均接受经超声心动图检查。纳入指标包括主动脉根部内径(aortic root internal diameter, AOD)、左心房内径(left atrial internal diameter, LAD)、左心室舒张末期内径(left ventricular end-diastolic internal diameter, LVEDD)、左心室收缩末期内径(left ventricular systolic internal diameter, LVESD)、左心室射血分数(left ventricular ejection fraction, EF)、左心室后壁厚度(left ventricular posterior wall thickness, LVPWT)、室间隔厚度(interventricular septal thickness, IVS)、左室心内膜缩短分数(fraction of shortened endocardium, FS)。

1.5 划分数据集

将数据经过匿名化处理后,进一步用Python的random模块中的sample函数,使用随机抽样方法将数据分为训练集和测试集。训练集用于模型构建与验证,测试集用于模型表现评估。

1.6 缺失数据处理

首先根据患者接受化学治疗(化疗)的状态分为新辅助治疗前与新辅助治疗后。由于存在缺失数据的变量为连续型变量,且缺失数据小于该变量下总数据的5%,遂对缺失值进行中位数插值处理。随后将2组合并,并进行Z-score标准化处理(转化函数为 $z=(x-\mu)/\sigma$),使数据分布呈均值为0,标准差为1的标准正态分布。

1.7 特征提取

计算各特征之间的Pearson相关系数(ρ), $\rho>0.8$ 的特征定义为具有高度共线性并被移除;使用5折交叉最小绝对值收敛和选择算子(least absolute shrinkage and selection operator, LASSO),以最小化均方误差(mean standard error, MSE)为目标进一步筛选函数特征;使用内核为随机森林(random forest, RF)的Boruta算法进一步优化剩余临床数据特征。

1.8 机器学习模型设置

本研究采用梯度增强 (gradient boost, GB)、支持向量机 (support vector machine, SVM)、决策树 (decision tree, DT)、K-近邻 (K-nearest neighbour, KNN) 及 RF 共 5 种机器学习模型进行实验。使用网格搜索 (grid search) 方法进行模型超参数的调节 [7]。

1.9 模型评价

使用 5 种机器学习模型对筛选后的单特征表现进行训练与验证。随后根据单特征所属组别进一步分类, 评估多特征模型的表现。使用混淆矩阵 (confusion matrix) 评估模型的性能, 主要评估指标包括准确率 ($ACC = \frac{TP + TN}{TP + TN + FP + FN}$)、精确率 ($Precision = \frac{TP}{TP + FP}$)、受试者操作特征曲线下面积 (area under the receiver operating characteristic curve, AUC)、F1 值 ($\frac{2}{F1} = \frac{1}{P} + \frac{1}{R}$) 等 (公式中的 TP、TN、FP、FN 分别表示真阳性率、真阴性率、假阳性率和假阴性率, P 和 R 分别表示精确率和召回率)。对于最终模型, 进一步采用决策曲线评估其临床实用性。

1.10 统计学分析

使用 Python (版本 3.7.10) 进行数据处理及统计分析。AUC 用 DeLong 检验比较, $P < 0.05$ 表示差异有统计学意义。使用 Hosmer-Lemeshaw 检验评估决策曲线, $P > 0.05$ 视为模型具有较好的校准效果。

2 结果

2.1 研究人群的基本情况

符合入选标准的乳腺癌新辅助治疗患者共 242 例, 经过数据预处理后, 最终纳入 232 例有效临床数据; 其中, 135 例为接受新辅助治疗前, 97 例为完成 4 个周期的新辅助治疗后 (图 1)。入组的新辅助治疗前后患者身高中位数均为 1.60 m; 治疗前后 2 组体质质量中位数分别为 59 及 60 kg; 心血管危险因素在 2 组间的差异无统计学意义。新辅助治疗后患者的 WBC ($P=0.000$)、HCT ($P=0.000$)、HB ($P=0.000$)、HDL ($P=0.000$)、NHDL ($P=0.034$)、IL-8 ($P=0.002$) 均低于新辅助治疗前; 而 TAG ($P=0.000$)、FGP ($P=0.010$)、IL-2R ($P=0.000$)、IL-6 ($P=0.003$)、TNF- α ($P=0.000$) 则高于新辅助治疗前, 差异均具有统计学意义 (表 1)。

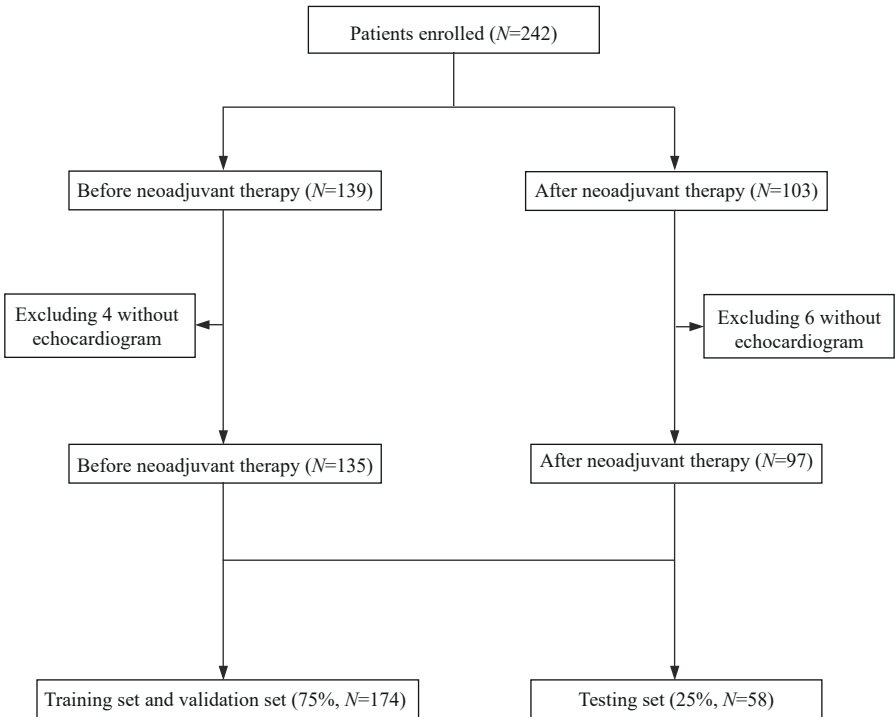


图 1 患者入组及数据预处理流程图
Fig 1 Flow chart of patient enrollment and data preprocessing

表 1 乳腺癌新辅助治疗前及新辅助治疗后患者所有特征变量信息

Tab 1 All feature variable information of patients with breast cancer before and after neoadjuvant therapy

Indicator	All (n=232)	Patient before neoadjuvant therapy (n=135)	Patient after neoadjuvant therapy (n=97)
Demographic parameter			
Age/year	47.00 (40.00, 55.00)	46.00 (39.00, 54.50)	48.00 (40.00, 56.00)
Height/m	1.60 (1.58, 1.65)	1.60 (1.58, 1.65)	1.60 (1.58, 1.65)
Weight/kg	60.00 (54.00, 67.78)	59.00 (55.00, 66.25)	60.00 (54.00, 68.00)
BMI/ (kg·m ⁻²)	22.87 (21.10, 25.39)	22.66 (21.09, 24.96)	23.10 (21.10, 25.50)
Systolic blood pressue/mmHg	122.70±13.65	124.77±13.47	119.81±13.45
Diastolic blood pressure/mmHg	77.15±8.72	77.59±7.84	76.54±9.83
Heart rate/ (beat ·min ⁻¹)	78.19±5.56	77.63±5.32	78.96±5.87
Accompanied by lymph node metastasis/n(%)	72 (31.0)	44 (32.6)	28 (28.9)
Tumor location/n(%)			
Left side	113 (48.7)	68 (50.4)	45 (46.4)
Right side	116 (50.0)	65 (48.1)	51 (52.6)
Bilateral	3 (1.3)	2 (1.5)	1 (1.0)
Neoadjuvant treatment/n(%)			
Paclitaxel+cisplatin	26 (26.8)	—	26 (26.8)
Paclitaxel+cisplatin+trastuzumab	21 (21.6)	—	21 (21.6)
Paclitaxel+cisplatin+apatinib	35 (36.1)	—	35 (36.1)
Paclitaxel+cisplatin+pyrotinib	15 (15.5)	—	15 (15.5)
Cardiovascular risk factor/n(%)			
Coronary heart disease	0 (0)	0 (0)	0 (0)
Hypertension	52 (22.4)	21 (15.6)	31 (31.0)
Type 2 diabetes	14 (6.0)	6 (4.4)	8 (8.2)
Hyperlipidemia	26 (11.2)	10 (7.4)	16 (16.5)
Smoking	0 (0)	0 (0)	0 (0)
Electrocardiographic parameter/n(%)			
ST-T change	0 (0)	0 (0)	2 (2.0)
QTc prolongation	0 (0)	0 (0)	0 (0)
Echocardiographic parameter			
AOD/mm	30.00 (28.00, 32.00)	30.00 (28.00, 32.00)	31.00 (28.00, 32.00)
LAD/mm	33.44±4.31	33.30±4.14	33.64±4.54
LVEDD/mm	44.62±3.30	44.66±3.35	44.57±3.24
LVESD/mm	28.00 (27.00, 30.00)	28.00 (27.00, 30.00)	29.00 (27.00, 30.00)
IVS/mm	8.00 (7.00, 9.00)	8.00 (7.00, 9.00)	8.00 (7.00, 9.00)
LVPWT/mm	8.00 (7.00, 8.00)	8.00 (7.00, 8.00)	8.00 (7.00, 9.00)
FS/%	36.00 (34.00, 38.00)	36.00 (34.00, 38.00)	37.00 (35.00, 38.00)
EF/%	65.00 (63.00, 68.00)	65.00 (63.00, 68.00)	66.00 (63.00, 69.00)
Laboratory examination parameter			
WBC/ (×10 ⁹ ·L ⁻¹)	5.11 (4.04, 6.56)	5.65 (4.68, 6.86)	4.44 (3.50, 5.79) ^①
HB/ (g·L ⁻¹)	120.00 (105.00, 129.00)	127.00 (120.00, 133.00)	105.00 (97.00, 117.00) ^①
HCT	0.36 (0.32, 0.39)	0.38 (0.36, 0.40)	0.32 (0.29, 0.35) ^①
ST2/ (ng·mL ⁻¹)	18.14 (13.83, 24.86)	18.38 (13.75, 25.36)	18.11 (14.25, 24.13)
BNP/ (pg·mL ⁻¹)	19.00 (12.00, 32.00)	19.00 (12.00, 31.00)	21.00 (15.00, 35.00)
NT-proBNP/ (pg·mL ⁻¹)	24.04 (10.00, 42.58)	24.42 (10.00, 41.58)	23.95 (10.48, 46.28)
TNI/ (ng·mL ⁻¹)	0 (0, 0)	0 (0, 0)	0 (0, 0.01)

Continued Tab

Indicator	All (n=232)	Patient before neoadjuvant therapy (n=135)	Patient after neoadjuvant therapy (n=97)
hsCRP/ (mg·L ⁻¹)	0.71 (0.32, 1.41)	0.70 (0.32, 1.42)	0.86 (0.34, 1.40)
TC/ (mmol·L ⁻¹)	4.48 (3.91, 5.16)	4.52 (3.98, 5.24)	4.42 (3.85, 4.99)
TAG/ (mmol·L ⁻¹)	1.35 (0.86, 2.04)	1.02 (0.74, 1.66)	1.75 (1.27, 2.40) ^①
HDL/ (mmol·L ⁻¹)	1.10 (0.87, 1.32)	1.24 (1.04, 1.45)	0.87 (0.75, 1.03) ^①
LDL/ (mmol·L ⁻¹)	3.34 (2.82, 3.98)	3.21 (2.76, 3.92)	3.47 (2.94, 4.12)
NHDL/ (mmol·L ⁻¹)	2.52 (2.16, 3.13)	2.60 (2.21, 3.20)	2.41 (2.05, 2.91) ^②
FPG/ (mmol·L ⁻¹)	4.95 (4.74, 5.26)	4.88 (4.69, 5.24)	5.05 (4.83, 5.29) ^④
GOT/ (mmol·L ⁻¹)	20.00 (16.00, 26.00)	17.00 (14.00, 23.00)	22.00 (19.00, 29.75)
GPT/ (mmol·L ⁻¹)	19.00 (13.00, 29.50)	15.00 (11.00, 27.00)	24.00 (17.00, 34.00)
γ-GT/ (mmol·L ⁻¹)	21.00 (14.00, 40.00)	18.00 (12.00, 32.00)	33.50 (17.00, 47.00)
IL-1β/ (pg·mL ⁻¹)	5.00 (5.00, 5.00)	5.00 (5.00, 5.00)	5.00 (5.00, 5.00)
IL-2R/ (U·mL ⁻¹)	330.00 (270.75, 424.75)	295.00 (249.50, 369.50)	399.00 (322.00, 508.00) ^①
IL-6/ (pg·mL ⁻¹)	3.15 (2.21, 4.25)	2.97 (2.12, 3.74)	3.66 (2.68, 4.89) ^⑤
IL-8/ (pg·mL ⁻¹)	11.30 (7.72, 18.95)	13.20 (8.81, 21.10)	9.84 (7.06, 15.20) ^③
IL-10/ (pg·mL ⁻¹)	5.00 (5.00, 5.00)	5.00 (5.00, 5.00)	5.00 (5.00, 5.00)
TNF-α/ (pg·mL ⁻¹)	6.53 (5.26, 7.86)	6.01 (5.00, 7.02)	7.33 (6.03, 8.60) ^①

Note: 1 mmHg=0.133 kPa. ^①P=0.000, ^②P=0.034, ^③P=0.002, ^④P=0.010, ^⑤P=0.003 compared with patient before neoadjuvant therapy.

2.2 特征提取与特征分组

经过3个步骤的特征提取，有5项特征仍保留，分别为WBC、HB、HDL、IL-2R和IL-8，所有心脏超声特征均被剔除（图2）。考虑到结果的临床可获

得性与经济价值，进一步将上述特征中WBC、HB和HDL作为一组，IL-2R和IL-8作为一组，全部特征作为一组，用于后续多特征模型分析。

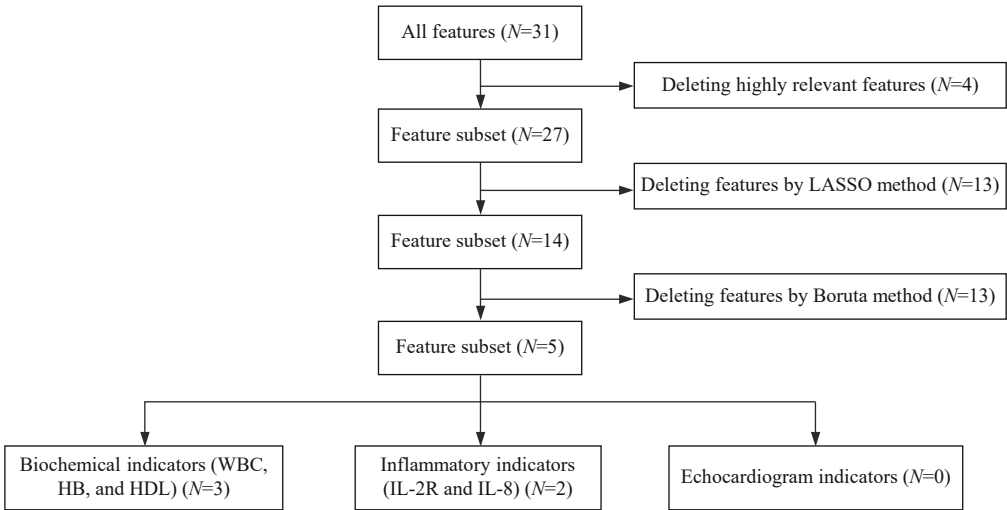
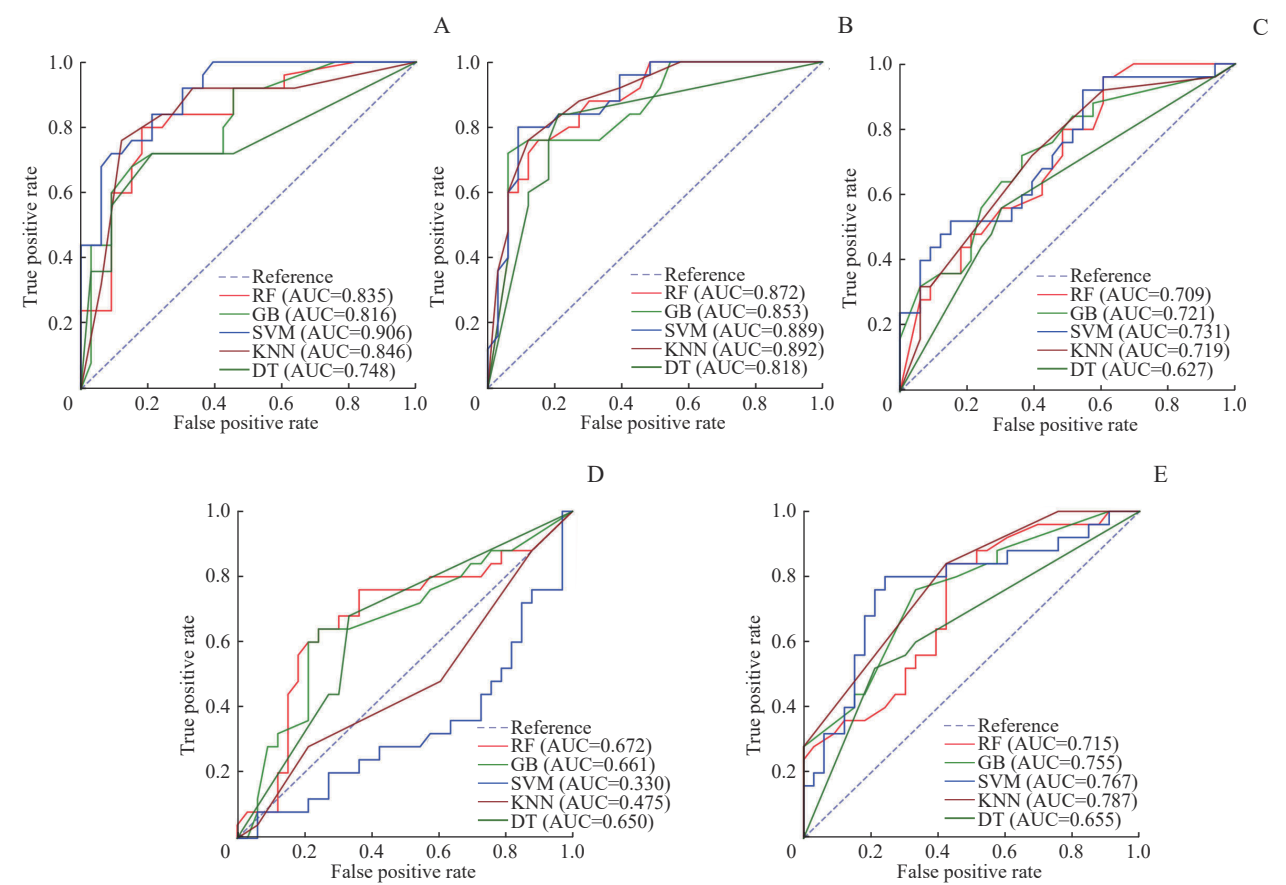


图2 乳腺癌新辅助治疗前后患者特征提取与特征分组流程图
Fig 2 Flowchart of feature extraction and feature grouping of patients with breast cancer before and after neoadjuvant therapy

2.3 单特征模型表现

在提取出的5项特征中，HB、HDL组在5种机器学习模型中的AUC值较高（HB 0.748~0.906，HDL 0.818~0.892）（图3A、B），其中基于HDL的

SVM方法准确率（84.5%）及F1值（0.870）最高（表2）。IL-2R、IL-8在5种机器学习模型中的AUC值均低于前者（IL-2R 0.627~0.731，IL-8 0.330~0.672，WBC 0.655~0.787），见图3C~E。



Note: A. HB single-feature ROC curve. B. HDL single-feature ROC curve. C. IL-2R single-feature ROC curve. D. IL-8 single-feature ROC curve. E. WBC single-feature ROC curve.

图3 乳腺癌新辅助治疗前后患者5种模型在测试集中的单特征分析ROC曲线

Fig 3 ROC curves of single-feature analysis for five models in the test set for patients with breast cancer before and after neoadjuvant therapy

表2 乳腺癌新辅助治疗前后患者5种模型在测试集中的单特征变量预测性能比较

Tab 2 Comparison of the predictive performance of the five models for single-feature variables in the test set for patients with breast cancer before and after neoadjuvant therapy

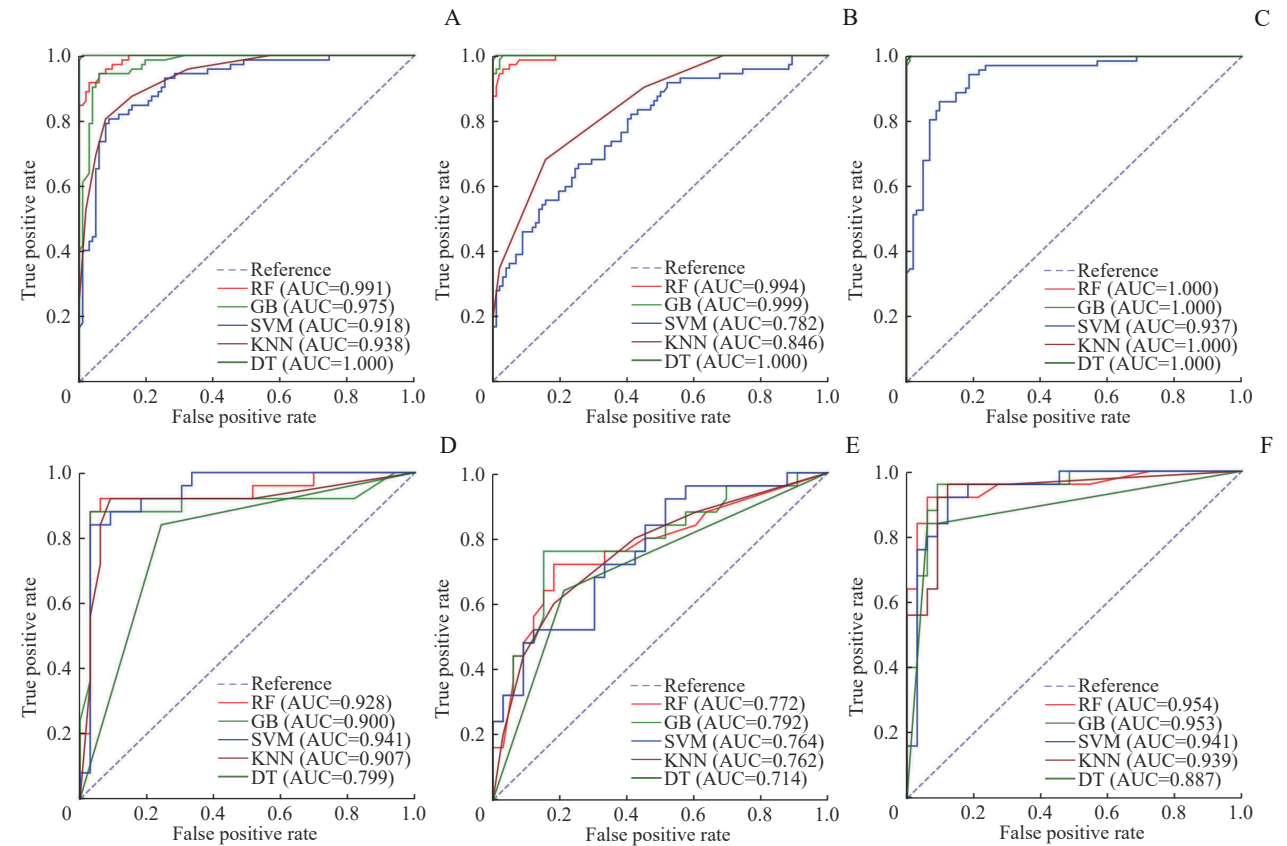
Single feature	Model	AUC	Accuracy	Precision rate	Recall	F1 score
WBC	RF	0.715	0.603	0.656	0.636	0.646
	GB	0.755	0.655	0.659	0.818	0.730
	SVM	0.767	0.690	0.683	0.848	0.757
	DT	0.655	0.569	0.643	0.545	0.590
	KNN	0.787	0.690	0.703	0.788	0.743
HB	RF	0.835	0.845	0.853	0.879	0.866
	GB	0.816	0.759	0.788	0.788	0.788
	SVM	0.906	0.810	0.824	0.848	0.836
	DT	0.748	0.759	0.788	0.788	0.788
	KNN	0.846	0.793	0.862	0.758	0.806
HDL	RF	0.872	0.810	0.824	0.848	0.836
	GB	0.853	0.828	0.829	0.879	0.853
	SVM	0.889	0.845	0.833	0.909	0.870
	DT	0.818	0.759	0.771	0.818	0.794
	KNN	0.892	0.828	0.829	0.879	0.853
IL-2R	RF	0.709	0.672	0.684	0.788	0.732
	GB	0.721	0.655	0.710	0.667	0.688

Continued Tab

Single feature	Model	AUC	Accuracy	Precision rate	Recall	F1 score
IL-8	SVM	0.731	0.690	0.683	0.848	0.757
	DT	0.627	0.638	0.676	0.697	0.687
	KNN	0.719	0.655	0.651	0.848	0.737
	RF	0.672	0.603	0.692	0.545	0.610
	GB	0.661	0.690	0.727	0.727	0.727
	SVM	0.330	0.569	0.569	1.000	0.725
	DT	0.650	0.672	0.733	0.667	0.698
	KNN	0.475	0.500	0.553	0.636	0.592

2.4 多特征模型表现

在 5 种机器学习模型的测试集中，WBC+HB+HDL 组的 AUC 高于 IL-2R+IL-8 组（RF： 0.928 vs 0.772； GB： 0.900 vs 0.792； SVM： 0.941 vs 0.764； KNN： 0.907 vs 0.762； DT： 0.799 vs 0.714），见图 4；在 RF、SVM、GB 模型中 WBC+HB+HDL 组的 AUC（0.928、0.941、0.900）及准确率（0.914、0.897、0.776）均较高（表3）。



Note: A–C. ROC curves of training group (A. WBC+HB+HDL multi-feature ROC curve; B. IL-2R+IL-8 multi-feature ROC curve; C. All features multi-feature ROC curve). D–F. ROC curves of testing group (D. WBC+HB+HDL multi-feature ROC curve; E. IL-2R+IL-8 multi-feature ROC curve; F. All features multi-feature ROC curve).

图4 乳腺癌新辅助治疗前后患者的5种模型的多特征分析ROC曲线

Fig 4 ROC curves of multi-feature analysis for five models for patients with breast cancer before and after neoadjuvant therapy

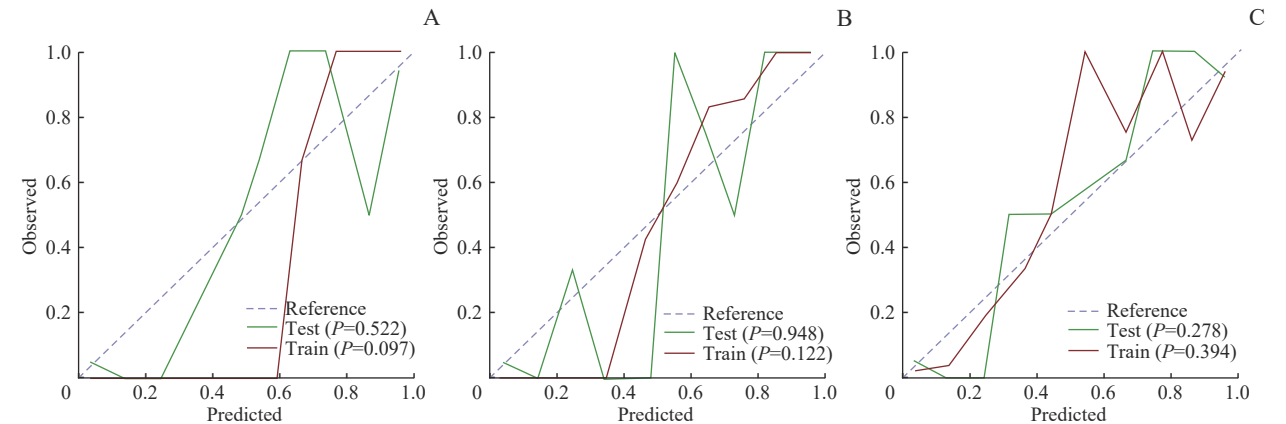
对上述3种方法绘制校准曲线（图5），结果显示，与RF、GB模型（ $P=0.122$ ， $P=0.097$ ）相比，SVM模型在训练集数据上校准度较好（ $P=0.394$ ）。结合所有特征的训练集和测试集的ROC曲线（图4C、F），可见RF与GB模型存在过拟合（RF：1.000 vs 0.954；GB：1.000 vs 0.953），提示SVM模型相比其他2种方法具有更好的泛化性能。

对SVM模型绘制决策曲线、ROC曲线与混淆

表3 乳腺癌新辅助治疗前后患者的5种模型在测试集中的多特征变量预测性能比较

Tab 3 Comparison of the predictive performance of the five models for multi-feature variables in the test set for patients with breast cancer before and after neoadjuvant therapy

Multi-feature	Model	AUC	Accuracy	Precision rate	Recall	F1 value
WBC+HB+HDL	RF	0.928	0.914	0.938	0.909	0.923
	GB	0.900	0.776	0.885	0.697	0.780
	SVM	0.941	0.897	0.909	0.909	0.909
	DT	0.799	0.793	0.862	0.758	0.806
	KNN	0.907	0.897	0.886	0.939	0.912
IL-2R+IL-8	RF	0.772	0.776	0.763	0.879	0.817
	GB	0.792	0.793	0.818	0.818	0.818
	SVM	0.764	0.707	0.690	0.879	0.773
	DT	0.714	0.707	0.735	0.758	0.746
	KNN	0.762	0.724	0.730	0.818	0.771
All features	RF	0.954	0.897	0.909	0.909	0.909
	GB	0.953	0.914	0.967	0.879	0.921
	SVM	0.941	0.879	0.882	0.909	0.896
	DT	0.887	0.862	0.857	0.909	0.882
	KNN	0.939	0.862	0.903	0.848	0.875



Note: A. GB model calibration curve. B. RF model calibration curve. C. SVM model calibration curve.

图5 所有特征组GB、RF、SVM模型的校准曲线

Fig 5 Calibration curves of the three models of GB, RF and SVM for all feature sets

矩阵，DeLong 检验证实 WBC+HB+HDL 组、所有特征组 AUC 均显著优于 IL-2R+IL-8 组 ($P=0.009$, $P=0.004$)，而前 2 组之间差异无统计学意义。WBC+HB+HDL 组与 IL-2R+IL-8 组的决策曲线下面积也有显著差异 (图 6)。混淆矩阵显示 WBC+HB+HDL 组、所有特征组均表现良好，而 IL-2R+IL-8 组产生了较多的假阴性病例 (图 7)。对以上 5 个特征绘制列线图，可知每个特征在模型中的权重 (图 8)。SVM 模型中 3 个多特征分析的 ROC 曲线见图 9。

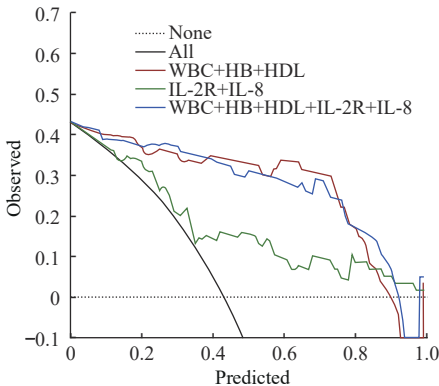
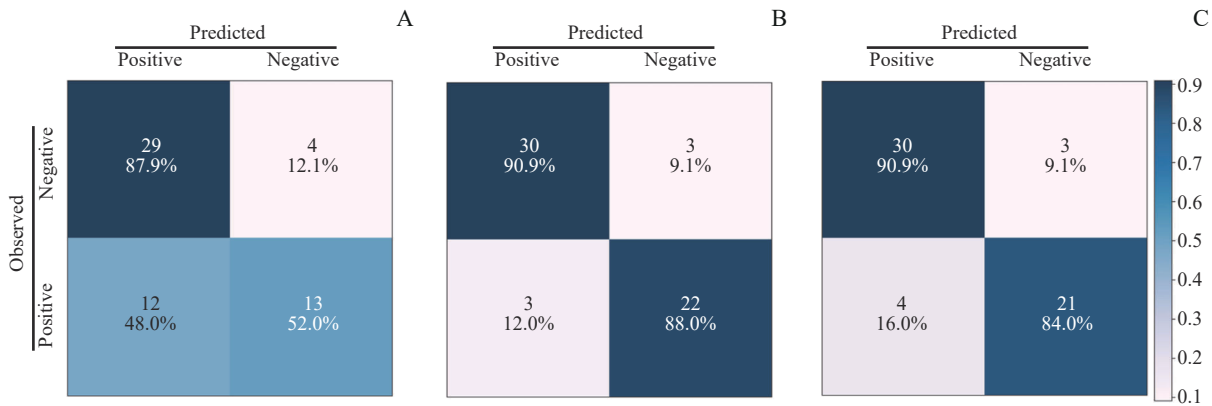


图6 SVM模型中3个多特征分析组的决策曲线

Fig 6 Decision curves for the three multi-featured analysis groups in SVM model



Note: A. Confusion matrix of IL-2R+IL-8 in SVM model. B. Confusion matrix of WBC+HB+HDL in SVM model. C. Confusion matrix of all features (WBC+HB+HDL+IL-2R+IL-8) in SVM model.

图7 SVM模型中3个多特征分析组的混淆矩阵

Fig 7 Confusion matrices for the three multi-feature analysis groups in SVM model

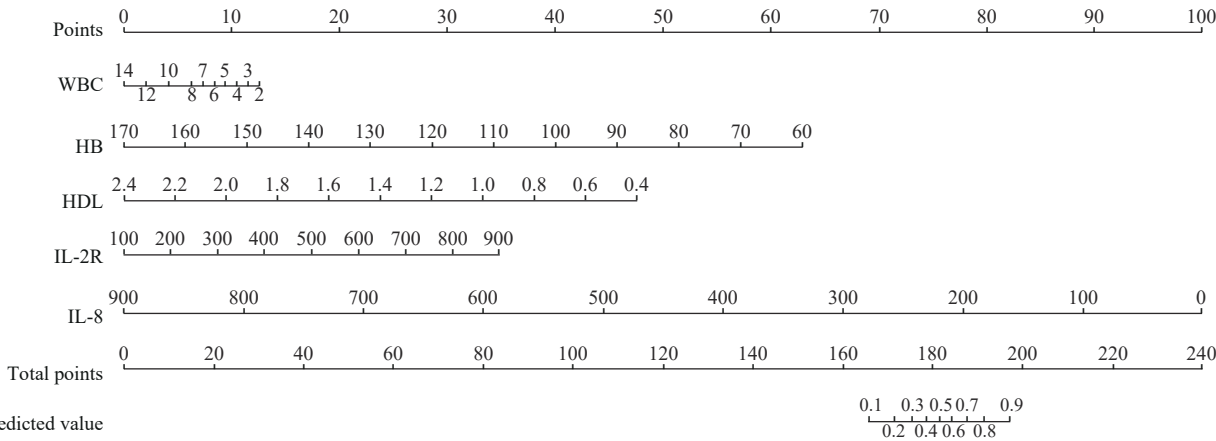


图8 乳腺癌患者新辅助治疗前后SVM预测模型列线图

Fig 8 Column line graph of SVM prediction model for breast cancer patients before and after neoadjuvant therapy

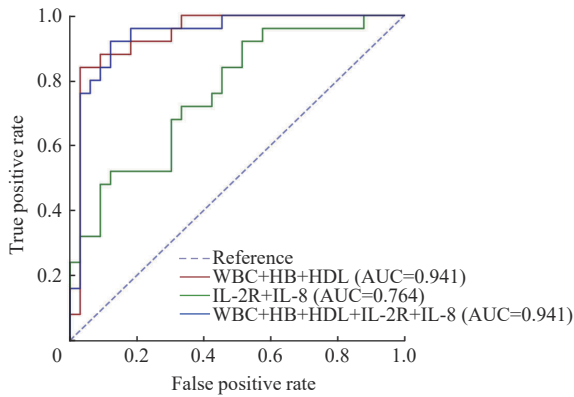


图9 SVM模型中3个多特征分析组的ROC曲线

Fig 9 ROC curves of subjects in the three multiple characterization groups in SVM model

3 讨论

机器学习进入医学领域后为快速诊断、辅助决策、同质化评价带来了前所未有的变革。在辅助临床

决策方面, BETANCUR等^[8]纳入1 638例可疑冠心病患者, 基于患者的单光子发射计算机断层成像(singlephoton emission computed tomography, SPECT)图像构建了深度学习模型, 并证实了该模型诊断冠心病的敏感度和特异度高于半定量方法。ALAA等^[9]基于UK Biobank数据库中的43 604名在基线时无冠心病史的参与者, 建立了共473个参数的心血管疾病风险预测模型, 包括年龄、性别等人口学数据, 吸烟史、家族史等临床危险因素, 以及血液、尿液实验室检测指标等参数, 预测致死及非致死性心血管事件(AUC 0.774, 95%CI 0.768~0.780), 并验证了该模型对一般人群及糖尿病人群患心血管疾病的预测能力高于Framingham评分(一般人群: AUC=0.724, 95%CI 0.720~0.728, $P<0.001$; 糖尿病人群: AUC=0.713, 95%CI 0.703~0.723)。PUJADA等^[10]使用UK Biobank中在2006至2010年间发生过房颤、

高血压、心肌梗死或卒中的30 926例患者的心脏磁共振参数、血管风险因素及心脏磁共振图像建立了7个综合模型,发现影像组学特征对心血管疾病事件具有更高的预测价值。

机器学习也被尝试应用于心血管事件的预测。TAWFIQ等^[11]对14 263例肿瘤幸存者进行回顾性分析,将心血管疾病风险预测方程预测的5年心血管事件风险与实际观测到的心血管事件对比,发现无论是在低风险组(<5%)还是高风险组(>15%),心血管事件风险被低估或高估者仅约2%,提示基于机器学习方法在心血管疾病风险预测模型中应用的可行性。由此,研究者提出了基于不同临床风险预测分组的干预方式:对于5年心血管事件风险小于5%的肿瘤幸存者,推荐生活方式干预;对于5年心血管事件风险大于5%且小于15%的肿瘤幸存者,推荐生活方式干预并可以考虑开始心血管疾病药物治疗;对于5年心血管事件风险大于15%的肿瘤幸存者,强烈建议药物和生活方式干预。

本研究尝试在乳腺癌新辅助治疗患者中,使用机器学习方法,根据治疗后临床或实验室检查特征提取来评价治疗前状态,预测治疗后出现炎症代谢状态改变的患者,通过特征差异筛选出未来可能出现炎症代谢状态改变的患者。结果发现,纳入WBC、HB、HDL、IL-2R和IL-8这5项特征的SVM模型具有更好的识别能力。

本研究纳入的是早期乳腺癌患者治疗前后的血清学检测指标和心脏超声指标,对使用机器学习方法识别新辅助治疗前后患者进行了探索。由于模型需要,根据数据的分布,找出高度共线性的特征并剔除,这有助于模型构建。选择特征数降维的常用方法(LASSO方法),通过统计可以获得对结局有重要价值的特征,同时排除数学上认为价值较低的特征。以RF为内置筛选方法的Boruta算法也在本研究中作为特征筛选方式,能进一步从统计学上找到有价值的特征。本研究经过筛选得到了5项指标,这5项指标在表1的统计分析中都展示了小于0.005的 P 值,反映了经过上述流程所选的特征在2组人群的分布差异显著,更有可能构建一个鲁棒性佳的模型。研究发现,在5种机器学习模型中,SVM算法的表现最佳。SVM属于有监督学习算法的一种,主要用于解决模式识别领域中的数据分类问题,主要通过最大化超平面与数据点之间的间隔找到一个有效的分类器,使

SVM能够在高维度空间中进行复杂的数据分离,并且在一定程度上容忍数据中的噪声和异常值。此前,其他研究^[12]已经验证,SVM在处理小样本数据时具有较好的准确性。在本研究的单特征和多特征模型中,SVM模型结果在校准曲线、ROC曲线和混淆矩阵等关键评价指标上都有较好的表现,提示了SVM模型在提高临床预测准确性方面的潜力。在多特征模型中,以WBC、HB和HDL为特征子集组和包含所有特征组表现出了相同的性能。这一发现提示了此模型具有通过更简洁特征获取同等分类表现的潜力,在不影响预测准确性的前提下,简化模型构建,提高效率,并提升可解释性。

虽然新辅助治疗前后患者的IL-2R和IL-8多特征模型分析提示其在识别患者上的表现不及包含WBC、HB和HDL的特征子集,但仍可提示患者在接受新辅助治疗后出现以IL-2R升高和IL-8降低为特征的炎症代谢状态变化。

IL-2R是由 α 、 β 、 γ 3条链组成的异聚体。研究^[13]发现,IL-2R的 α 亚基在恶性肿瘤细胞中表达上调,在包括直肠癌、前列腺癌、乳腺癌和肺癌在内的多种癌症中都可观察到可溶性IL-2R和IL-2R α 水平升高。在肿瘤靶向治疗领域,可溶性IL-2R可诱导肿瘤细胞凋亡,抑制肿瘤血管生成并干扰致癌途径^[14]。IL-8又称趋化因子CXCL8,其受体为趋化因子受体(CXCR1/2)^[15]。在肿瘤微环境中,肿瘤细胞、内皮细胞、肿瘤相关成纤维细胞和免疫浸润细胞均可分泌IL-8。已经在多种实体肿瘤中发现,IL-8介导CD8⁺T细胞耗竭^[16],其上调与肿瘤细胞增殖、血管生成、侵袭和转移有关^[17]。与正常乳腺组织相比,IL-8在乳腺癌中过度表达;IL-8可诱导上皮-间质转化,从而维持乳腺癌细胞的干细胞特性,且IL-8水平与肿瘤干细胞的活性具有相关性^[18]。此外,AHMED等^[19]发现,在HER-2阳性的局部进展性乳腺癌中,肿瘤细胞分泌的IL-8导致了拉帕替尼的耐药性。在心血管领域,近年来的研究发现,在心肌梗死患者人群中,血清高IL-2R水平与IL-8水平对患者未来1年内的主要心血管不良事件具有显著的预测价值^[20]。已有研究^[21]证实,IL-2通过促进小血管生成,在心肌梗死患者中具有保护心脏的功能。IL-8与促进动脉粥样硬化形成有关^[22],并可能是肥胖相关代谢并发症如2型糖尿病、高脂血症的潜在标志物^[23],但在心肌梗死患者中的作用尚不明确。已有研究^[24]发现,肥胖

相关代谢障碍与心肌水肿和纤维化有关。血清 IL-8 水平与代谢障碍导致的心肌水肿和纤维化的关系,有待进一步研究。值得注意的是,研究发现,新辅助治疗后手术前以 LDL 升高和 HDL 降低为特征的脂代谢紊乱与乳腺癌患者远期预后不良有关^[25],且基线时血清 HDL 水平较高是乳腺癌患者新辅助治疗效果的独立预测因子^[26]。因此我们推断,本研究中被 SVM 模型判定为“治疗后状态”的,实际上为新辅助治疗前的患者,可能存在与脂代谢紊乱和炎症代谢激活有关的远期不良预后。

与基于乳腺磁共振成像和病理图像的预测乳腺癌新辅助治疗效果及预后的机器学习模型^[26-27]相比,本研究使用的模型曲线下面积更佳 (AUC 0.941),准确率更高 (0.897)。而且,最终筛选出的特征均为拟接受新辅助治疗患者的常规检查项目,在实际场景中获成本低、可得性高,提高了该方法用于患者风险筛查的可行性与便捷性,更利于推广。另外,使用 SVM 模型的优势在于,可以使用较小样本的数据库建立及训练模型;并且,由于数据的可获得性高,随着样本量增加,可以不断迭代模型,实现更准确的判别。

本研究作为机器学习在识别新辅助治疗患者领域的探索研究,有一定的不足。本研究为病例对照研究,可能出现选择偏倚及信息偏倚。本研究使用的样本量中等,尽管包含了较多的患者特征,但仍可能出现选择偏倚,这可能限制了研究结果在更广泛人群中的推广。此外,本研究假设经算法认定为“治疗后状态”的新辅助治疗前患者,存在与不良心血管预后相关的炎症代谢状态,该结果能否在不同患者群体和不同的医疗环境中应用仍需要进一步外部验证。

综上所述, SVM 模型纳入了包含 WBC、HB、HDL、IL-2R 和 IL-8 的 5 项临床常见特征,建立了较为经济、简洁的预测模型,在乳腺癌患者中辅助预测早期新辅助治疗后哪些患者更可能出现与炎症代谢状

态相关的不良预后。对于模型识别出的可能存在不良炎症代谢状态的患者,可能需要进行代谢综合征的早期筛查,并提示乳腺外科医师可能需要增加该类患者的随访频率,以提高总体预后。

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

伦理批准和知情同意/Ethics Approval and Patient Consent

本研究涉及的所有试验均已通过上海交通大学医学院附属仁济医院科学伦理委员会的审核批准 (审批号 KY2020-189)。所有试验过程均遵照《赫尔辛基宣言》的条例进行。受试对象或其亲属已经签署知情同意书。

All experimental protocols in this study were reviewed and approved by the Scientific Ethics Committee of Renji Hospital, Shanghai Jiao Tong University School of Medicine (Approval Letter No. KY2020-189), and all experimental protocols were carried out by following the guidelines of *Declaration of Helsinki*. Consent letters have been signed by the research participants or their relatives.

作者贡献/Authors' Contributions

吴其蓁、刘启明、姜萌、卜军参与试验设计、论文写作及修改;吴其蓁、刘启明、柴烨子、陶政宇、王依楠、郭欣宇参与临床数据收集。所有作者均阅读并同意了最终稿件的提交。

WU Qizhen, LIU Qiming, JIANG Meng and PU Jun participated in the experimental design, paper writing and revision. WU Qizhen, LIU Qiming, CHAI Yezi, TAO Zhengyu, WANG Yinan and GUO Xinning participated in the clinical data collection. All the authors have read the last version of paper and consented for submission.

临床试验注册/Clinical Trial Registry

本研究已于美国临床试验注册平台 (ClinicalTrials.gov) 注册,注册号为 NCT04510532。

This study has been registered on the United States Clinical Trial Registry website (ClinicalTrials.gov), and the registration number is NCT04510532.

- Received: 2024-01-29
- Accepted: 2024-06-04
- Published online: 2024-09-28

参 · 考 · 文 · 献

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2021, 71(3): 209-249.
- [2] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2024, 74(3): 229-263.
- [3] RAMIN C, SCHAEFFER M L, ZHENG Z, et al. All-Cause and cardiovascular disease mortality among breast cancer survivors in CLUE II, a long-standing community-based cohort[J]. J Natl Cancer Inst, 2021, 113(2): 137-145.
- [4] JOHNSON K W, TORRES SOTO J, GLICKSBERG B S, et al. Artificial intelligence in cardiology[J]. J Am Coll Cardiol, 2018,

- 71(23): 2668-2679.
- [5] OIKONOMOU E K, KHERA R. Machine learning in precision diabetes care and cardiovascular risk prediction[J]. *Cardiovasc Diabetol*, 2023, 22(1): 259.
- [6] YANG H Z, YU B X, OUYANG P, et al. Machine learning-aided risk prediction for metabolic syndrome based on 3 years study[J]. *Sci Rep*, 2022, 12(1): 2248.
- [7] WANG G, ZHANG Y, LI S, et al. A machine learning-based prediction model for cardiovascular risk in women with preeclampsia[J]. *Front Cardiovasc Med*, 2021, 8: 736491.
- [8] BETANCUR J, COMMANDEUR F, MOTLAGH M, et al. Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT: a multicenter study[J]. *JACC Cardiovasc Imaging*, 2018, 11(11): 1654-1663.
- [9] ALAA A M, BOLTON T, DI ANGELANTONIO E, et al. Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423, 604 UK Biobank participants[J]. *PLoS One*, 2019, 14(5): e0213653.
- [10] PUJADAS E R, RAISI-ESTABRAGH Z, SZABO L, et al. Prediction of incident cardiovascular events using machine learning and CMR radiomics[J]. *Eur Radiol*, 2023, 33(5): 3488-3500.
- [11] TAWFIQ E, SELAK V, ELWOOD J M, et al. Performance of cardiovascular disease risk prediction equations in more than 14 000 survivors of cancer in New Zealand primary care: a validation study[J]. *Lancet*, 2023, 401(10374): 357-365.
- [12] MENG M, ZHAO C. Application of support vector machines to a small-sample prediction[J]. *Adv Petrol Explor Dev*, 2015, 2(10): 72-75.
- [13] MUHAMMAD S, FAN T, HAI Y, et al. Reigniting hope in cancer treatment: the promise and pitfalls of IL-2 and IL-2R targeting strategies[J]. *Mol Cancer*, 2023, 22(1): 121.
- [14] HERNANDEZ R, PÖDER J, LAPORTE K M, et al. Engineering IL-2 for immunotherapy of autoimmunity and cancer[J]. *Nat Rev Immunol*, 2022, 22(10): 614-628.
- [15] BAGGIOLINI M, CLARK-LEWIS I. Interleukin-8, a chemotactic and inflammatory cytokine[J]. *FEBS Lett*, 1992, 307(1): 97-101.
- [16] LIU Y Y, ZHOU N N, ZHOU L, et al. IL-2 regulates tumor-reactive CD8⁺ T cell exhaustion by activating the aryl hydrocarbon receptor[J]. *Nat Immunol*, 2021, 22(3): 358-369.
- [17] LIU Q, LI A, TIAN Y, et al. The CXCL8-CXCR1/2 pathways in cancer[J]. *Cytokine Growth Factor Rev*, 2016, 31: 61-71.
- [18] SINGH J K, SIMÕES B M, HOWELL S J, et al. Recent advances reveal IL-8 signaling as a potential key to targeting breast cancer stem cells[J]. *Breast Cancer Res*, 2013, 15(4): 210.
- [19] AHMED S, MOHAMED H T, EL-HUSSEINY N, et al. IL-8 secreted by tumor associated macrophages contribute to lapatinib resistance in HER2-positive locally advanced breast cancer via activation of Src/STAT3/ERK1/2-mediated EGFR signaling[J]. *Biochim Biophys Acta Mol Cell Res*, 2021, 1868(6): 118995.
- [20] PAN K, XU C, CHEN C, et al. Soluble interleukin-2 receptor combined with interleukin-8 is a powerful predictor of future adverse cardiovascular events in patients with acute myocardial infarction[J]. *Front Cardiovasc Med*, 2023, 10: 1110742.
- [21] BOUCHENTOUF M, WILLIAMS P, FORNER K A, et al. Interleukin-2 enhances angiogenesis and preserves cardiac function following myocardial infarction[J]. *Cytokine*, 2011, 56(3): 732-738.
- [22] APOSTOLAKIS S, VOGIATZI K, AMANATIDOU V, et al. Interleukin 8 and cardiovascular disease[J]. *Cardiovasc Res*, 2009, 84(3): 353-360.
- [23] KIM C S, PARK H S, KAWADA T, et al. Circulating levels of MCP-1 and IL-8 are elevated in human obese subjects and associated with obesity-related parameters[J]. *Int J Obes*, 2006, 30(9): 1347-1355.
- [24] ZHAO H, HUANG R, JIANG M, et al. Myocardial tissue-level characteristics of adults with metabolically healthy obesity[J]. *JACC Cardiovasc Imaging*, 2023, 16(7): 889-901.
- [25] QU F, CHEN R, PENG Y, et al. Assessment of the predictive role of serum lipid profiles in breast cancer patients receiving neoadjuvant chemotherapy[J]. *J Breast Cancer*, 2020, 23(3): 246-258.
- [26] TAHMASSEBI A, WENGERT G J, HELBICH T H, et al. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients[J]. *Invest Radiol*, 2019, 54(2): 110-117.
- [27] SHI Z, HUANG X, CHENG Z, et al. MRI-based quantification of intratumoral heterogeneity for predicting treatment response to neoadjuvant chemotherapy in breast cancer[J]. *Radiology*, 2023, 308(1): e222830.

[本文编辑] 吴 洋

