Open Access
DOI:10.3724/zdxbyxb-2025-0205

智能医药

·专题报道·

构建融合临床常规参数和肿瘤突变负荷的癌症患者免疫治疗预后预测模型

朱旭东[1],郝舒强[1],程　真[2],方维佳[1]

1. 浙江大学医学院附属第一医院肿瘤内科,浙江 杭州 310003
2. 东阳市人民医院肿瘤内科,浙江 金华 322100

［摘　要］ **目的**:构建融合临床常规参数及肿瘤突变负荷(TMB)的机器学习模型,并评估其对各类癌症患者程序性死亡受体1(PD-1)/程序性死亡受体配体1(PD-L1)抑制剂治疗应答的预测价值。**方法**:回顾性研究2019年11月至2022年10月在浙江大学医学院附属第一医院接受PD-1/PD-L1抑制剂治疗的146例晚期癌症患者的临床资料。采用分层随机抽样法按4∶1比例将患者分为训练集(116例)和验证集(30例)。整合患者年龄、性别、体重指数、TMB、全身治疗史、中性粒细胞与淋巴细胞比值及其他血常规参数,基于PyTorch框架构建多层感知机网络,利用AutoGluon进行超参数自动优化,通过五折交叉验证进行模型优化,并应用梯度沙普利可加性特征解释(SHAP)算法在训练集中对最优模型进行特征重要性分析。通过受试者操作特征曲线下面积(AUC)、准确率、F1分数、灵敏度和特异度等指标比较NNT9模型与单一TMB指标的预测性能,随后绘制混淆矩阵以直观评估模型的判别能力,最后分析模型预测的PD-1/PD-L1抑制剂治疗应答差异与患者无进展生存时间的关系。**结果**:最终构建的最优模型为NNT9,全身治疗史、TMB、血小板计数和体重指数为前四位预测因子。构建的NNT9模型在训练集和验证集中的AUC分别为0.949和0.851,均优于单一TMB(AUC分别为0.747和0.720)。在验证集中,NNT9模型的灵敏度(0.571)、准确率(0.867)、F1分数(0.667)、阳性预测值(0.800)和阴性预测值(0.880)均优于TMB。在验证集中,NNT9模型错误分类的患者例数仅为TMB的一半。Kaplan-Meier生存分析显示,在训练集和验证集中,NNT9模型预测为免疫治疗应答组的患者,其无进展生存时间显著长于非应答组的患者(均$P<0.01$)。**结论**:融合临床常规参数及TMB的机器学习模型可作为一种精准且可行的工具用于预测不同类型癌症患者的免疫治疗获益。

［关键词］ 恶性肿瘤;免疫治疗;免疫检查点抑制剂;治疗应答;预测模型;机器学习;肿瘤突变负荷

收稿日期(Received):2025-03-16　修改返回日期(Revised):2025-12-26　接受日期(Accepted):2026-01-20
基金项目(Funding):国家自然科学基金(82373428)
第一作者(First author):朱旭东,主治医师,主要从事消化道肿瘤全程治疗、实体肿瘤免疫治疗研究;E-mail:xudongzhu@zju.edu.cn;ORCID:0009-0003-6082-9072
通信作者(Corresponding author):方维佳,主任医师,博士生导师,主要从事实体肿瘤免疫治疗研究;E-mail:weijiafang@zju.edu.cn;ORCID:0000-0001-9849-347X

［中图分类号］ R730.5;TP18 ［文献标志码］ A

Construction of a prognosis forecasting model for immunotherapy response in cancer patients by integrating routine clinical parameters and tumor mutational burden

ZHU Xudong[1], HAO Shuqiang[1], CHENG Zhen[2], FANG Weijia[1] (*1. Department of Medical Oncology, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China; 2. Department of Medical Oncology, Dongyang People's Hospital, Jinhua 322100, Zhejiang Province, China*)

Corresponding author: FANG Weijia, E-mail: weijiafang@zju.edu.cn, ORCID: 0000-0001-9849-347X

［**Abstract**］ **Objective**: To develop a machine-learning model that integrates routine clinical parameters with tumor mutational burden (TMB) and to evaluate its performance in predicting responses to programmed death-1 (PD-1)/programmed death-ligand 1(PD-L1) inhibitors across various cancer types. **Methods**: We conducted a retrospective study of 146 patients with advanced solid tumors who were treated with PD-1/PD-L1 inhibitors. The cohort was randomly divided into a training set (n=116) and a validation set (n=30) at a 4:1 ratio. Using the PyTorch framework, we constructed a neural network model (designated NNT9) incorporating age, sex, body mass index (BMI), TMB, history of systemic therapy, neutrophil-to-lymphocyte ratio (NLR), and other routine blood parameters. The model employed a multilayer perceptron architecture. Hyperparameters were automatically optimized using AutoGluon, and the model was refined via 5-fold cross-validation. SHapley Additive exPlanations (SHAP) was used to perform feature importance analysis on the optimal model in the training set. Predictive performance was compared against TMB alone using metrics including the area under the receiver operating characteristic curve (AUC), accuracy, F1 score, sensitivity, and specificity. Confusion matrices were generated, and the association between model-predicted response groups and progress free survive (PFS) was analyzed. **Results**: NNT9 was identified as the optimal model, and the history of systemic therapy, TMB, platelet count, and BMI were the four most important predictive features. NNT9 achieved AUCs of 0.949 and 0.851 in the training and validation sets, respectively, outperforming TMB alone (AUCs: 0.747 and 0.720). In the validation set, NNT9 also demonstrated superior sensitivity (0.571), accuracy (0.867), F1 score (0.667), positive predictive value (0.800), and negative predictive value (0.880). The confusion matrix revealed that NNT9 misclassified only half as many patients as TMB alone in the validation set. Kaplan-Meier analysis showed that patients predicted to be responders by NNT9 had significantly longer PFS than non-responders in both training and validation sets (both P<0.01). **Conclusion**: The NNT9 model, which integrates readily available clinical parameters with TMB, represents an accurate and clinically feasible tool for predicting immunotherapy benefit in a pan-cancer cohort, and shows promise for clinical translation.

［**Key words**］ Malignant tumor; Immunotherapy; Immune checkpoint inhibitor;

Treatment response; Forecasting model; Machine learning; Tumor mutational burden

[J Zhejiang Univ (Med Sci), 2026, 55(1): 36-45.]

[缩略语] 程序性死亡受体1(programmed death-1,PD-1);程序性死亡受体配体1(programmed death-ligand 1,PD-L1);肿瘤突变负荷(tumor mutational burden,TMB);无进展生存(progress free survive,PFS);置信区间(confidence interval,CI);中性粒细胞与淋巴细胞比值(neutrophil-to-lymphocyte ratio,NLR);美国病理学家学会(College of American Pathologists,CAP);美国临床实验室改进修正案(Clinical Laboratory Improvement Amendments,CLIA);沙普利可加性特征解释(SHapley Additive exPlanations,SHAP);曲线下面积(area under the curve,AUC);最小绝对收缩和选择算子(least absolute shrinkage and selection operator,LASSO)

免疫检查点抑制剂如PD-1/PD-L1抑制剂的应用明显提升了抗肿瘤治疗的临床应答率,已成为癌症患者治疗的重要选择[1]。然而,患者对PD-1/PD-L1抑制剂的疗效因癌种及分子特征谱不同而存在较大差异,迫切需要探索可在人群中预测其疗效的新型生物标志物[2]。研究表明,TMB(即肿瘤外显子组中非同义突变的累积计数)是一个能够较好预测免疫治疗效果的生物标志物,但由于肿瘤异质性及其免疫微环境复杂性,单纯依靠TMB的预测能力有限[3-5]。多项研究尝试整合基因组变异、转录组特征和临床参数来构建预测模型,以期建立免疫治疗疗效评估的复合指标体系[6-8]。与TMB比较,此类模型虽展现出更优的预测性能,但其依赖多组学联合检测,临床可及性和实用性仍面临挑战。例如,在临床实践中,晚期癌症患者的组织标本多通过穿刺获取,其标本体积常不足以同步完成免疫组化、全外显子测序和RNA测序等多项检测。因此,在临床标本有限的情况下,开发出既精准又易于推广的预测模型,具有重要的临床意义。

研究证实,肿瘤浸润淋巴细胞的亚群分布及空间定位与免疫治疗敏感性显著相关[9]。外周血常规参数不仅可替代肿瘤浸润淋巴细胞丰度的生物标志物,还能动态反映宿主系统性炎症状态及适应性免疫激活程度[10]。多项研究证实,在PD-1/PD-L1抑制剂治疗前系统评估外周血细胞谱,可有效预测治疗应答模式和PFS时间[11-12]。而且,全血细胞计数作为抗肿瘤治疗期间的常规检测项目,具备标准化操作流程和广泛可及的优势。

近年来,机器学习在解析高维度生物医学数据(如影像组学等多组学特征参数)方面取得突破性进展[13-14],为基于有限特征变量构建高效预测模型提供了有力工具。与传统统计学方法聚焦于变量间因果推断不同,机器学习更注重通过特征工程和算法优化提升模型的预测性能[15]。其中,基于深度学习方法构建的预测框架已在肿瘤分子分型和治疗反应预测中展现优势,正逐步从实验研究向临床转化推进[16]。神经网络作为深度学习的核心算法模型,在建模复杂的非线性特征交互关系方面具有明显优势,尤其适用于生物医学数据中高维变量的集成和挖掘分析,其多层结构可实现从原始特征到高阶抽象表示的自动提取,有利于提升预测的准确性和泛化能力[13]。然而,神经网络要实现理想的特征自动提取通常需要较大的样本量。考虑到本研究样本量有限,若完全依赖模型自动挖掘特征极易产生过拟合,且提取出的特征往往缺乏临床解释性。因此,本研究采用“专家知识引导+深度学习建模”的策略:先基于临床经验筛选出有明确生物学意义的指标,再利用神经网络进行建模。该策略能兼顾模型稳定性和预测准确性。本研究拟通过融合癌症患者临床常规检测指标和TMB数据构建一个可用于精准筛选跨癌种PD-1/PD-L1抑制剂治疗获益人群的预测模型。

1 资料与方法

1.1 研究对象

本研究为回顾性队列研究,纳入2019年11月至2022年10月在浙江大学医学院附属第一医院接受PD-1/PD-L1抑制剂治疗的不同类型晚期癌症患者。纳入标准:①确诊时年龄18岁以上;②经

组织病理学确诊的晚期实体瘤；③可获得完整的基线TMB检测数据和治疗前外周血常规检测结果；④至少完成1个周期免疫治疗。排除标准：①基线影像学检查与首次给药时间间隔超过6周；②治疗期间未按计划（每6~8周）进行影像学随访评估；③接受新辅助或辅助免疫治疗等非姑息性治疗方案。

共146例患者纳入分析队列。恶性肿瘤类型分布为胆道癌26.7%（39/146）、结直肠癌19.2%（28/146）、胰腺癌16.4%（24/146）、肺癌13.7%（20/146），其余24.0%（35/146）为肾细胞癌、膀胱癌、胃癌、肝细胞癌及肉瘤等。全队列中位TMB为6.1（3.2，9.8）muts/Mb；8.2%的患者接受抗PD-1单药治疗，91.8%的患者接受抗PD-1联合化疗和/或靶向治疗；47.3%的患者将免疫治疗作为一线方案，52.7%的患者将免疫治疗作为二线及以上治疗。全队列应答（完全缓解或部分缓解）率为22.6%（图1），PFS中位时间达5.3个月（95%CI：4.1~6.5）。

1.2 资料收集

通过查阅电子病历系统，回顾性收集患者的临床资料，包括年龄、性别、体重指数、恶性肿瘤病理类型、全身治疗史及具体治疗方案。同时记录免疫治疗开始前的实验室检查指标，包括中性粒细胞计数、淋巴细胞计数、血红蛋白、血小板计数及白细胞计数，并计算NLR。患者的TMB数据来源于经CAP和CLIA认证的基因检测机构（南京世和基因生物技术股份有限公司、厦门艾德生物医药科技股份有限公司以及广州燃石医学检验所有限公司）提供的二代测序结果。

1.3 疗效评估

影像学随访采用标准化影像采集协议，每6~8周依据实体瘤疗效评价标准（RECIST 1.1）[17]盲法评估影像结果。基于患者在整个研究期间的最佳观察结果将疗效判定为：①应答，即完全缓解或部分缓解；②非应答，即仅达到疾病稳定或疾病进展。PFS时间定义为患者从接受首剂PD-1/PD-L1抑制剂治疗至疾病进展或死亡的时间。

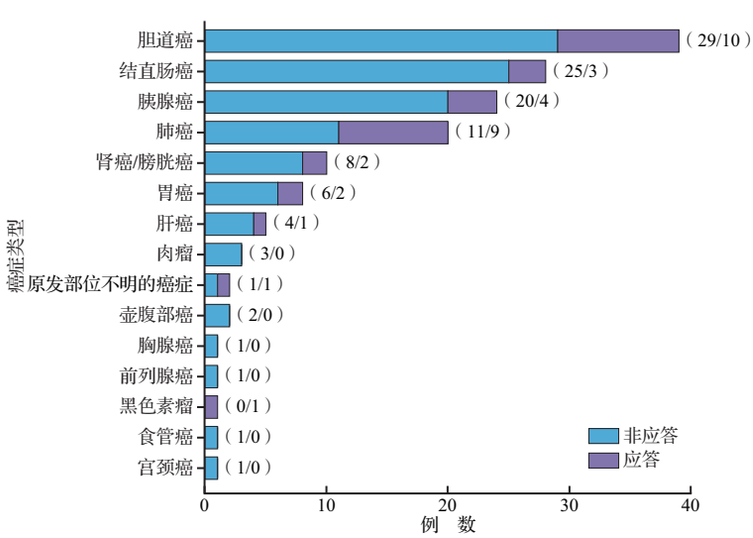
1.4 建模特征筛选

鉴于外周血参数可反映系统性炎症及免疫状态[11-12]，且既往全身治疗史与肿瘤免疫微环境重塑密切相关[18-19]，本研究综合考虑生物学合理性、临床可及性和数据完整性，选定九个特征纳入模型构建：年龄、性别、体重指数、TMB、全身治疗史、NLR、血红蛋白、血小板计数和白细胞计数。

1.5 模型构建

基于PyTorch框架构建用于不同类型癌症患者免疫治疗疗效预测的深度神经网络分类器。技术实现依托Python 3.9.18软件，集成scikit-learn和AutoGluon机器学习库。数据经唯一ID去重后，按4∶1比例随机划分为训练集和验证集。训练集内采用五折分层交叉验证指导早停和超参数选择，统一设置随机种子（NumPy、PyTorch、scikit-learn、AutoGluon均为42），并为PyTorch启用确定性算法，以保证重复性。在特征处理中，分类变量缺失值单独编码为新类别；连续变量缺失值在每个训练折内采用中位数填补并使用StandardScaler进行标化后再应用于对应验证，避免信息泄露。若探索性分析提示类别分布不均衡，则在损失函数中引入类别加权以优化对少数类别的学习。

神经网络采用多层感知机架构（输入层–隐藏层1–隐藏层2–输出层）。默认激活函数为ReLU（待搜索的候选超参数包含ReLU、SELU和Tanh）。隐藏层1和隐藏层2的单元数分别为64和32，各隐藏层之后依



根据实体瘤疗效评价标准（RECIST 1.1）评估[17].

图1 各类癌症患者治疗应答情况分布

Figure 1 Distribution of treatment response across cancer types

次连接批标准化(BatchNorm1d)和Dropout正则化(丢弃概率p=0.20)。输出层为1个输出单元,训练使用BCEWithLogitsLoss,推断时以Sigmoid得到概率。使用Adam优化器(参数:$\beta1$=0.9,$\beta2$=0.999,weight_decay=1e-4),初始学习率lr=0.002,使用ReduceLROnPlateau调度器(参数:mode="min",factor=0.5,patience=3,min_lr=1e-5)配合学习率调整,按验证损失自适应减速。训练设置批量大小为16、最大训练轮数为100,采用早停机制(验证损失patience=5、min_delta=1e-4)。

在AutoGluon的TabularPredictor(problem_type="binary",eval_metric="roc_auc")框架下进行自动训练和超参数优化。超参数搜索设置为:学习率LogUniform[1e-4,1e-2]、候选的激活函数{ReLU,SELU,Tanh}、丢弃率Uniform[0.0,0.5];并补充搜索hidden_size{32,64,128}、num_layers{1,2}、weight_decay LogUniform[1e-6,1e-3]。以训练集五折交叉验证进行最大试验次数为40及总时间限制为1800 s搜索;选取验证表现最优的前五个模型配置构建集成模型。采用自助聚合(Bagging)策略(设定重采样10次,每次训练采样率为67%),并利用袋外样本(即未参与当前轮次训练的样本)来评估模型的泛化性能。确定最优超参数和集成方案后,在完整训练集中重新训练最优模型。为增强模型构建阶段的可解释性,在确定最优网络结构和超参数后,应用梯度SHAP算法在训练集中对最优模型进行特征重要性分析。

1.6 模型评估

在训练集中进行模型拟合能力评估,在验证集中进行模型泛化能力评估。首先,以0.5为阈值,对训练集和验证集进行二分类预测(≥0.5为阳性,<0.5为阴性)。然后根据样本的真实属性,将样本分为四类,分别是真阳性、真阴性、假阳性、假阴性,据此计算准确率、灵敏度、特异度、阳性预测值、阴性预测值和F1分数。绘制受试者操作特征曲线,并计算AUC值以衡量模型的判别能力;采用非参数Bootstrap法(通过Bootstrap=1000次、由放回重采样且固定随机种子的方法获得AUC的95%CI)评估AUC的稳定性。为直观展示模型在不同类别上的判别能力,在训练集和验证集中分别绘制混淆矩阵。最后,将模型预测的应答状态纳入生存分析,采用Kaplan-Meier法绘制PFS时间曲线,并通过log-rank检验评估预测应答组与非应答组的预后差异。所有可视化过程由Matplotlib和scikit-learn实现,评估流程和参数(折数、重采样次数、阈值与网格等)均预先设定以确保结果可复现。

1.7 统计学方法

分类变量以例数(百分比)[n(%)]表示,组间比较采用χ^2检验。连续变量符合正态分布者以均数±标准差($\bar{x}\pm s$)表示,组间比较使用独立样本t检验;非正态分布者以中位数(上下四分位数)[$M(Q_1,Q_3)$]表示,组间比较采用Mann-Whitney U检验。所有统计检验均为双侧,P<0.05为差异有统计学意义。统计分析和数据可视化使用IBM SPSS Statistics 26.0和Python 3.9.18软件实现。

2 结 果

2.1 训练集和验证集基线特征

采用分层随机抽样,146例患者分为训练集(n=116)和验证集(n=30),训练集和验证集在年龄、性别、体重指数、全身治疗史、TMB、NLR、血红蛋白、血小板计数及白细胞计数等基线特征上的分布基本相似(表1),表明随机分层划分有效,适用于评估模型的泛化性能。

2.2 最优模型及SHAP方法的特征重要性分析

本研究最终构建的最优模型为NNT9神经网

表1 训练集与验证集患者基线特征比较

Table 1 Baseline characteristics of patients in the training and validation sets

[$\bar{x}\pm s$或n(%)或$M(Q_1,Q_3)$]

组 别	n	年龄(岁)	男 性	体重指数(kg/m^2)	全身治疗史	TMB(muts/Mb)
全队列	146	61.2±11.8	87(59.6)	21.6(19.4,23.7)	77(52.7)	6.1(3.2,9.8)
训练集	116	61.1±12.5	70(60.3)	21.6(19.3,23.8)	60(51.7)	6.3(3.2,9.4)
验证集	30	61.4±8.7	17(56.7)	21.0(19.4,23.5)	17(56.7)	5.4(1.4,9.4)

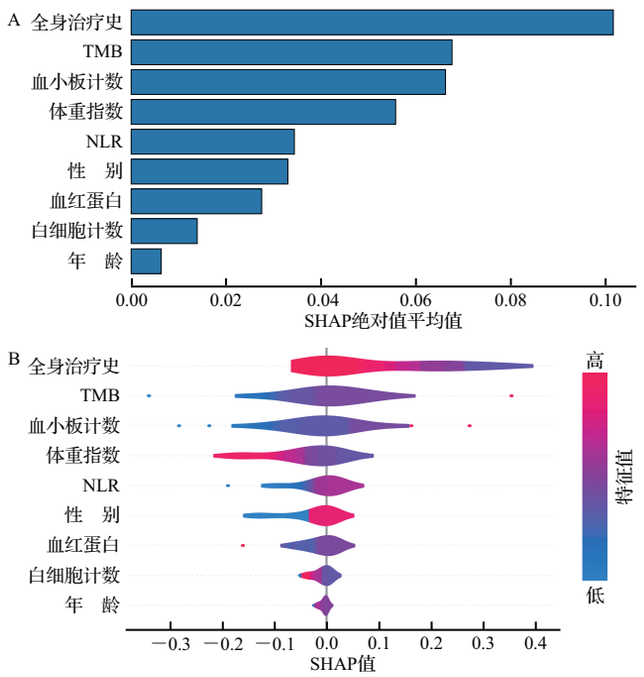
组 别	n	NLR	血红蛋白(g/L)	血小板计数(×10^9/L)	白细胞计数(×10^9/L)	抗PD-1/PD-L1应答
全队列	146	4.0±3.1	119.6±20.8	215.6±97.9	6.6±2.9	33(22.6)
训练集	116	4.2±3.2	118.8±21.5	214.5±92.3	6.6±2.6	26(22.4)
验证集	30	3.3±2.6	123.0±18.0	219.6±118.8	6.6±3.8	7(23.3)

训练集与验证集比较,差异均无统计学意义(均P>0.05). TMB:肿瘤突变负荷;NLR:中性粒细胞与淋巴细胞比值;PD-1:程序性死亡受体1;PD-L1:程序性死亡受体配体1.

络模型。SHAP特征重要性分析显示，全身治疗史、TMB、血小板计数和体重指数为排名前四的预测因子，见图2，提示恶性肿瘤本身的免疫原性、既往治疗带来的生物学改变及患者初治时良好的体能状态，可能与更好的免疫应答相关。

2.3 模型预测性能

在训练集和验证集中，NNT9模型的AUC值分别为0.949和0.851，而TMB的AUC值分别为0.747和0.720。在训练集中，NNT9模型的灵敏度、准确率、F1分数和阴性预测值均优于TMB；在验证集中，NNT9模型的灵敏度、准确率、F1分数、阳性预测值和阴性预测值均优于TMB。见表2和图3。混淆矩阵分析进一步显示了NNT9模型和TMB的预测性能差异：在训练集中，TMB错误分类22例患者，而NNT9模型仅错误分类16例患者（图4A）；在验证集中，NNT9模型错误分类的患者例数仅为TMB的一半（分别为4例和8例，见图4B）。上述结果提示NNT9模型在训练集和验证集中均保持稳定的性能，具有良好的泛化能力。此外，Kaplan-Meier生存分析显示，在训练集和验证集中，由NNT9模型预测的应答者，其PFS时间均显著长于非应答者（均P<0.01）；由TMB预测的应答者与非应答者之间PFS时间差异均无统计学意义（图5）。综上，NNT9模型在预测免疫治疗获益人群方面，性能稳定且明显优于单一TMB指标。



A：基于梯度SHAP算法的特征重要性排序；B：特征的SHAP值分布。红色代表特征值较高，蓝色代表特征值较低；SHAP值大于0表示该特征倾向于预测为“应答”，小于0表示该特征倾向于预测为“非应答”。TMB：肿瘤突变负荷；NLR：中性粒细胞与淋巴细胞比值；SHAP：沙普利可加性特征解释。

图2 NNT9模型中各特征的SHAP值及其分布

Figure 2 Feature importance and SHAP value distribution of the NNT9 model

表2 NNT9模型与TMB预测癌症患者对PD-1/PD-L1抑制剂治疗反应的性能比较

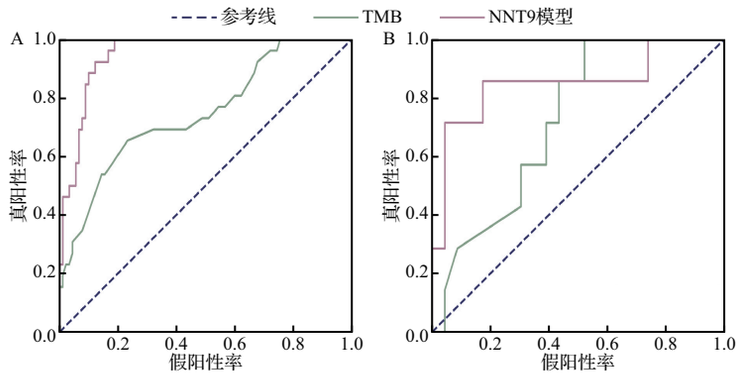
Table 2 Performance metrics of the NNT9 model and TMB for predicting response to PD-1/PD-L1 inhibitor therapy in cancer patients

预测方法	训练集						
	灵敏度	特异度	阳性预测值	阴性预测值	准确率	AUC	F1分数
TMB	0.231	0.978	0.750	0.815	0.810	0.747	0.353
NNT9模型	0.615	0.933	0.727	0.894	0.862	0.949	0.667
预测方法	验证集						
	灵敏度	特异度	阳性预测值	阴性预测值	准确率	AUC	F1分数
TMB	0.000	0.957	0.000	0.759	0.733	0.720	0.000
NNT9模型	0.571	0.957	0.800	0.880	0.867	0.851	0.667

TMB：肿瘤突变负荷；AUC：曲线下面积；PD-1：程序性死亡受体1；PD-L1：程序性死亡受体配体1.

3 讨 论

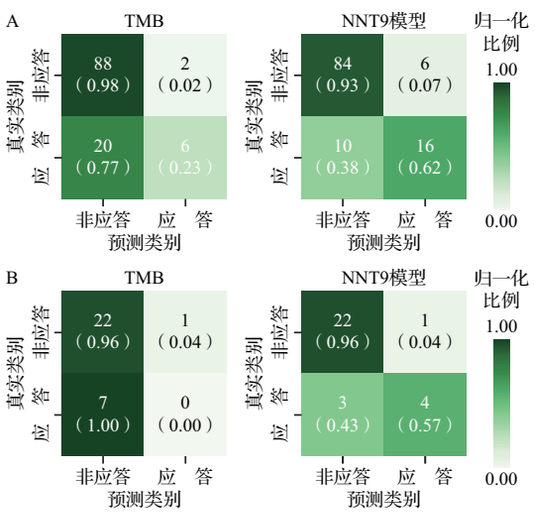
近年来，尽管以PD-1/PD-L1抑制剂为代表的免疫疗法明显改善了癌症患者生存期，但其客观缓解率有限，即使其与化疗或靶向药物联合使用也仅维持30%~60%的客观缓解率[2, 20]。值得注意的是，免疫检查点抑制剂的不良事件发生率虽低于传统化疗，但其一旦发生便可能导致严重后果[21]。因此，若能早期识别PD-1/PD-L1抑制剂治疗无效的患者，将有助于避免无效治疗及相关不良反应。本研究基于机器学习方法，融合临床常规参数和TMB构建了NNT9模型以预测癌症患者对PD-1/PD-L1抑制剂治疗的应答情况。结果显示，NNT9模型在训练集和验证集的AUC值分别高达0.949和0.851，其预测性能明显优于目前广



A:训练集;B:验证集. TMB:肿瘤突变负荷;PD-1:程序性死亡受体1;PD-L1:程序性死亡受体配体1.

图3 NNT9模型和TMB预测癌症患者对PD-1/PD-L1抑制剂治疗反应的受试者操作特征曲线

Figure 3 Receiver operating characteristic curves of the NNT9 model and TMB for predicting response to PD-1/PD-L1 inhibitor therapy in cancer patients



A:训练集;B:验证集. TMB:肿瘤突变负荷;PD-1:程序性死亡受体1;PD-L1:程序性死亡受体配体1.

图4 NNT9模型和TMB预测癌症患者对PD-1/PD-L1抑制剂治疗反应的混淆矩阵

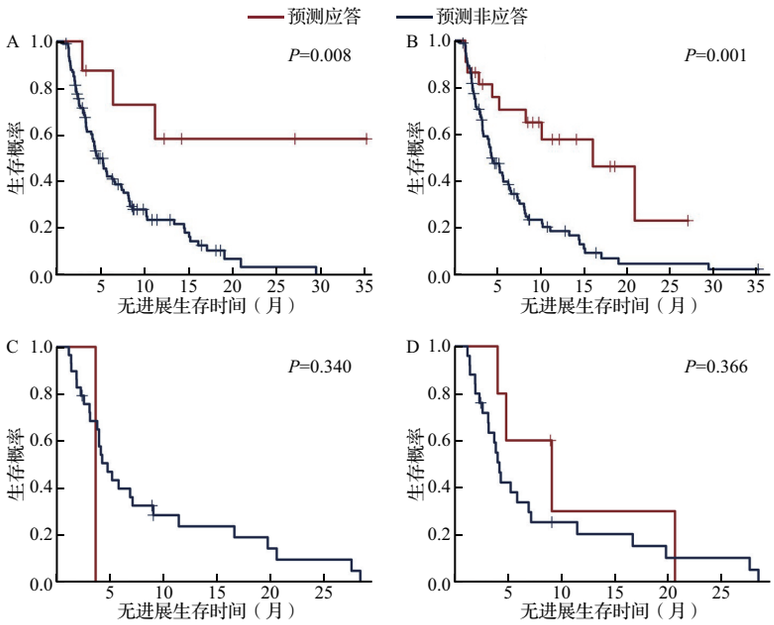
Figure 4 Confusion matrix of NNT9 model and TMB for predicting response to PD-1/PD-L1 inhibitor therapy in cancer patients

泛应用的免疫治疗疗效预测生物标志物TMB。

在精准医学时代,识别可能从PD-1/PD-L1抑制剂治疗中获益的癌症患者至关重要,不仅能优化临床决策(如患者筛选),还能为非应答者制订新的治疗策略提供依据。既往研究已基于不同类型数据,包括临床病理特征、影像组学、基因组学以及转录组学数据等构建了PD-1/PD-L1抑制剂疗效预测模型[6, 16, 22]。尽管这些模型展现出优异的预测性能,例如Chowell等[16]开发的泛癌种预测模型AUC达0.82,Chen等[22]开发的胃癌多维预测模型验证集AUC可达0.85,但此类模型构建依赖多组学数据,限制了其在临床实践中的推广应用。针对上述局限性,本研究构建了一个融合TMB和临床常规参数的预测模型,所选用的临床常规参数(如血常规指标)是临床实践中广泛可及的检查项目。而且,与传统的统计学方法(侧重于探究变量与结局之间的关联)比较,本研究采用的机器学习方法更擅长处理复杂的预测任务。

本研究最初尝试纳入40余项潜在变量,并利用自动化特征选择算法(如LASSO回归、基于树模型的特征重要性评分等)构建预测模型,但在样本量相对有限的前提下,模型表现出较大的性能波动和较差的泛化能力。可能与部分变量间存在较强共线性、数据缺失率高或临床意义不明确有关。鉴于此,我们最终采用以临床经验为主导的特征选择策略,优选了九个生物学意义明确、数据质量高且在实际工作中广泛可及的变量纳入模型。既往研究表明,这种专家驱动的变量选择方法在预测性能、结果稳定性及可解释性方面均具有良好表现,同时更契合临床决策的实际需求,有望为数据驱动的个体化治疗提供支持[23]。

本研究验证了临床常规参数对PD-1/PD-L1抑制剂疗效的预测价值,与既往研究结论一致,如基线NLR值及其在免疫治疗过程中的动态变化、血小板计数、血红蛋白以及体重指数均与免疫治疗应答有关[24-28]。患者既往接受的全身治疗线数(即全身治疗史)也是影响PD-1/PD-L1抑制剂疗效的重要因素,一般而言,治疗线数越多,应答率越低[19],本研究通过量化各参数的预测权重,发现全身治疗史是PD-1/PD-L1抑制剂疗效的最优预测因子,其重要性甚至超过TMB。全身治疗史对免疫应答的预测意义可能与其对肿瘤免疫微环境的影响有关。一方面,前线的化疗、靶向治疗或放疗会导致肿瘤抗原持续暴露,进而与免疫抑制通路共同促进T淋巴细胞耗



A、B：训练集（A）和验证集（B）基于NNT9模型分组的癌症患者生存曲线；C、D：训练集（C）和验证集（D）基于TMB分组的癌症患者生存曲线. TMB：肿瘤突变负荷.

图5 基于TMB和NNT9模型预测癌症患者无进展生存时间Kaplan-Meier曲线

Figure 5 Kaplan-Meier curves of progress free survive for cancer patients stratified by TMB and NNT9 model prediction

竭，导致$CD8^+$ T淋巴细胞功能耗竭，表现为细胞毒性功能受损及细胞毒性T淋巴细胞相关抗原4等耗竭分子表达上调[18]。另一方面，既往放化疗还可促进髓源性抑制细胞、调节性T细胞和肿瘤相关巨噬细胞等免疫抑制细胞在肿瘤微环境中的浸润，加重肿瘤微环境的免疫抑制[29]，这也解释了在非小细胞肺癌中，免疫检查点抑制剂单药一线治疗的疗效优于二线，进一步印证了全身治疗史是PD-1/PD-L1抑制剂的负向调控因素[19]。

鉴于TMB是预测免疫治疗疗效的公认生物标志物，本研究从多维度比较了NNT9模型与TMB的预测性能，评估融合临床常规参数的NNT9模型能否超越TMB的预测性能。本文资料中混淆矩阵分析显示，在训练集和验证集中，NNT9模型的假阴性率约为TMB的一半，同时其真阳性率明显高于TMB。上述结果表明，NNT9模型能够更灵敏地筛选出可能从PD-1/PD-L1抑制剂治疗获益的患者，减少患者错失免疫治疗的机会。值得注意的是，本模型的特异度超过90%，表明其在识别非应答患者方面可靠性高，可有效避免不必要的治疗风险。NNT9模型是基于TMB和高度可及的临床常规参数构建，具备较高的预测准确性和临床易用性，可为实现免疫治疗精准决策提供量化支持。具体而言，对于模型预测为“应答”的患者，可优先考虑使用PD-1/PD-L1抑制剂（单药或联合治疗策略），以提高治疗效率；对于预测为“非应答”的患者，模型结果可提示临床医生警惕免疫治疗可能无效的风险，从而引导进行更深入的分子检测（如微卫星不稳定性、PD-L1表达或免疫微环境特征分析）或优先考虑化疗、靶向治疗及双重免疫检查点抑制剂联合治疗等替代方案，并在后续适时评估是否联合免疫治疗或加入临床试验等。综上，NNT9模型不仅明显增强了TMB的单变量预测能力，而且可作为临床分层管理的重要工具，为晚期癌症患者提供更精准的个体化治疗方案。

本研究亦存在一定的局限性。首先，尽管采用了分层随机抽样方法，并在训练集和验证集中均观察到较为一致的模型性能（AUC分别为0.949和0.851），但总样本量仅146例，验证集仅30例，样本规模偏小，可能对模型的泛化能力产生一定影响。其次，本研究中癌症类型分布不均衡，消化系统癌症（如胆道癌、结直肠癌、胰腺癌）占比较高，提示该模型的适用范围可能更偏向于此类癌种。因此，未来研究将进一步优化癌种结构、纳入更多类型的癌症患者，以提升模型在癌症免疫治疗疗效预测中的适用性。

总之，本研究通过融合TMB与体重指数、外周血参数、全身治疗史等易获取的临床常规参数，构建了一个基于机器学习的癌症预测模型NNT9，可有效预测晚期癌症患者对PD-1/PD-L1抑制剂治疗的应答。该模型有助于精准筛选潜在获益人群，避免对非应答患者实施不必要的免疫治疗，从而优化临床决策。这一临床可及性较好的模型为实现精准免疫治疗，进而改善患者生存结局提供了新工具。

本文附加文件见电子版。

志谢 本研究得到国家自然科学基金(82373428)支持.浙江大学医学院附属第一医院呼吸内科周建娅主任医师,结直肠外科吴国生主任医师,肿瘤内科阮健主任医师、张晓琛主任医师、单建贞主任医师、马雁鸿主治医师,放射介入科应世红副主任医师、龚绍林主治医师提供部分患者资料.浙江大学医学院附属第一医院肿瘤内科秦浩康研究助理和孙旭琪主治医师协助数据复核及文字润色.《浙江大学学报(医学版)》编辑部余方和沈洁编辑对论文提出修改建议

Acknowledgements This study was supported by National Natural Science Foundation of China (82373428). ZHOU Jianya, WU Guosheng, RUAN Jian, ZHANG Xiaochen, SHAN Jianzhen, MA Yanhong, YING Shihong, GONG Shaolin from the First Affiliated Hospital of Zhejiang University School of Medicine provided part of the clinical patient data. QIN Haokang and SUN Xuqi from the First Affiliated Hospital of Zhejiang University School of Medicine assisted with data verification and manuscript polishing. Editor YU Fang and SHEN Jie from the Editorial Department of *Journal of Zhejiang University (Medical Sciences)* provided valuable suggestions for the revision of this manuscript

作者贡献 朱旭东、郝舒强、程真和方维佳参与论文选题和设计或参与资料获取、分析或解释,起草研究论文或修改重要智力性内容.所有作者均已阅读并认可最终稿件,并对数据的完整性和安全性负责.具体见电子版

Author Contributions ZHU Xudong, HAO Shuqiang, CHENG Zhen and FANG Weijia participated in brewing and designing experiments, or acquisition, analysis, or interpretation of data for the work; drafting the work, or revising it critically for important intellectual content. All authors have read and approved the final manuscript, and take responsibility for the integrity and security of the data. See the electronic version for details

数据可用性 本研究的相关数据可根据合理请求从通信作者(E-mail:weijiafang@zju.edu.cn)处获得

Data Availability All relevant data supporting the findings of this study can be accessed upon reasonable request from the corresponding author at weijiafang@zju.edu.cn

医学伦理 本研究通过浙江大学医学院附属第一医院ⅡT伦理审查委员会批准(浙大一院伦审2024研第0106号-快)并符合1964年《赫尔辛基宣言》及之后的修订版或类似的伦理标准.本研究豁免患者知情同意

Ethical Approval All procedures performed in this study involving human participants were in accordance with the ethical standards of the Clinical Research Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine (2024-Research-0106-Fast), and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The exemption from informed consent for this study was approved

利益冲突 所有作者均声明不存在利益冲突

Conflict of Interests The authors declare that there is no conflict of interests

©The author(s) 2026. This is an open access article under the CC BY-NC-ND 4.0 license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

参考文献(References)

[1] WANG Y, WANG M, WU H X, et al. Advancing to the era of cancer immunotherapy[J]. **Cancer Commun**, 2021, 41(9): 803-829.

[2] HAVEL J J, CHOWELL D, CHAN T A. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy[J]. **Nat Rev Cancer**, 2019, 19(3): 133-150.

[3] ALEXANDROV L B, NIK-ZAINAL S, WEDGE D C, et al. Signatures of mutational processes in human cancer[J]. **Nature**, 2013, 500(7463): 415-421.

[4] MALEKI VAREKI S. High and low mutational burden tumors versus immunologically hot and cold tumors and response to immune checkpoint inhibitors[J]. **J Immunother Cancer**, 2018, 6(1): 157.

[5] ANAGNOSTOU V, NIKNAFS N, MARRONE K, et al. Multimodal genomic features predict outcome of immune checkpoint blockade in non-small-cell lung cancer[J]. **Nat Cancer**, 2020, 1(1): 99-111.

[6] ZHANG Z, WANG Z X, CHEN Y X, et al. Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response[J]. **Genome Med**, 2022, 14(1): 45.

[7] ZOU Y, XIE J, ZHENG S, et al. Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery[J]. **Int J Surg**, 2022, 107: 106936.

[8] KONG J, HA D, LEE J, et al. Network-based machine learning approach to predict immunotherapy response in cancer patients[J]. **Nat Commun**, 2022, 13(1): 3703.

[9] GOODEN M J M, DE BOCK G H, LEFFERS N, et al. The prognostic influence of tumour-infiltrating lymphocytes in cancer: a systematic review with meta-analysis [J]. **Br J Cancer**, 2011, 105(1): 93-103.

[10] BALKWILL F, MANTOVANI A. Inflammation and

cancer: back to Virchow?[J]. **Lancet**, 2001, 357(9255): 539-545.

[11] HOPKINS A M, ROWLAND A, KICHENADASSE G, et al. Predicting response and toxicity to immune checkpoint inhibitors using routinely available blood and clinical markers[J]. **Br J Cancer**, 2017, 117(7): 913-920.

[12] BENZEKRY S, GRANGEON M, KARLSEN M, et al. Machine learning for prediction of immunotherapy efficacy in non-small cell lung cancer from simple clinical and biological data[J]. **Cancers**, 2021, 13(24): 6210.

[13] TOPOL E J. High-performance medicine: the convergence of human and artificial intelligence[J]. **Nat Med**, 2019, 25(1): 44-56.

[14] RAJKOMAR A, DEAN J, KOHANE I. Machine learning in medicine[J]. **N Engl J Med**, 2019, 380(14):1347-1358.

[15] BENZEKRY S. Artificial intelligence and mechanistic modeling for clinical decision making in oncology[J]. **Clin Pharmacol Ther**, 2020, 108(3): 471-486.

[16] CHOWELL D, YOO S K, VALERO C, et al. Improved prediction of immune checkpoint blockade efficacy across multiple cancer types[J]. **Nat Biotechnol**, 2022, 40(4): 499-506.

[17] EISENHAUER E A, THERASSE P, BOGAERTS J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)[J]. **Eur J Cancer**, 2009, 45(2): 228-247.

[18] NAIR R, SOMASUNDARAM V, KURIAKOSE A, et al. Deciphering T-cell exhaustion in the tumor microenvironment: paving the way for innovative solid tumor therapies[J]. **Front Immunol**, 2025, 16: 1548234.

[19] WU S, WANG L, LI W, et al. Comparison between the first-line and second-line immunotherapy drugs in the progression-free survival and overall survival in advanced non-small cell lung cancer: a systematic review and meta-analysis of randomized controlled trials[J]. **Ann Palliat Med**, 2021, 10(2): 1717-1726.

[20] PAZ-ARES L, LUFT A, VICENTE D, et al. Pembrolizumab plus chemotherapy for squamous non-small-cell lung cancer[J]. **N Engl J Med**, 2018, 379(21): 2040-2051.

[21] LI Y, LIANG X, LI H, et al. Comparative efficacy and safety of immune checkpoint inhibitors for unresectable advanced melanoma: a systematic review and network meta-analysis[J]. **Int Immunopharmacol**, 2023, 115: 109657.

[22] CHEN Y, JIA K, SUN Y, et al. Predicting response to immunotherapy in gastric cancer *via* multi-dimensional analyses of the tumour immune microenvironment[J]. **Nat Commun**, 2022, 13(1): 4851.

[23] KORNOWICZ J, THOMMES K. Algorithm, expert, or both? Evaluating the role of feature selection methods on user preferences and reliance[J/OL]. **PLoS One**, 2025, 20(3): e0318874.

[24] LALANI A A, XIE W, MARTINI D J, et al. Change in neutrophil-to-lymphocyte ratio (NLR) in response to immune checkpoint blockade for metastatic renal cell carcinoma[J]. **J Immunother Cancer**, 2018, 6(1): 5.

[25] SUN X, FENG Y, ZHANG B, et al. The role of neutrophil-to-lymphocyte ratio in predicting pathological response for resectable non-small cell lung cancer treated with neoadjuvant chemotherapy combined with PD-1 checkpoint inhibitors[J]. **Cancer Res Treat**, 2022, 54(4): 1017-1029.

[26] RUGAMBWA T K, ABDIHAMID O, ZHANG X, et al. Neutrophil-lymphocyte ratio and platelet-lymphocyte ratio as potential predictive markers of treatment response in cancer patients treated with immune checkpoint inhibitors: a systematic review and meta-analysis[J]. **Front Oncol**, 2023, 13: 1181248.

[27] AYERS K L, MA M, DEBUSSCHE G, et al. A composite biomarker of neutrophil-lymphocyte ratio and hemoglobin level correlates with clinical response to PD-1 and PD-L1 inhibitors in advanced non-small cell lung cancers[J]. **BMC Cancer**, 2021, 21(1): 441.

[28] DONNELLY D, BAJAJ S, YU J, et al. The complex relationship between body mass index and response to immune checkpoint inhibition in metastatic melanoma patients[J]. **J Immunother Cancer**, 2019, 7(1): 222.

[29] LIU S, WANG W, HU S, et al. Radiotherapy remodels the tumor microenvironment for enhancing immunotherapeutic sensitivity[J]. **Cell Death Dis**, 2023, 14(10): 679.

［本文编辑　余　方　沈　洁］