· 临床研究 ·

Machine learning models based on dual-energy CT quantitative parameters for predicting pathological response of gastric cancer to neoadjuvant immunochemotherapy

ZHENG Yue[1,2,3,4], *FAN Songmei*[1,2,3,4], *YUAN Mengchen*[1,2,3,4], *LIU Yiyang*[1,2,3,4], *YOU Yaru*[1,2,3,4], *KOU Songzi*[5], *CHENG Zhen*[1,2,3,4], *LI Liming*[1,2,3,4], *GAO Jianbo*[1,2,3,4*]

(1.*Department of Radiology*, 5.*Department of Pathology*, *the First Affiliated Hospital of Zhengzhou University*, *Zhengzhou* 450052, *China*; 2.*Henan International Joint Laboratory of Medical Imaging*, *Zhengzhou* 450052, *China*; 3.*Henan Key Laboratory of Image Diagnosis and Treatment for Digestive System Tumor*, *Zhengzhou* 450052, *China*; 4.*Henan Key Laboratory of CT Imaging*, *Zhengzhou* 450052, *China*)

[**Abstract**] **Objective** To observe the value of machine learning (ML) models based on dual-energy CT quantitative parameters for predicting pathological response of gastric cancer to neoadjuvant immunochemotherapy (nICT). **Methods** Totally 125 gastric cancer patients who underwent nICT were retrospectively enrolled. Seventy-five cases were randomly assigned to training set and further stratified according to tumor regression grade (TGR) into good response (GR) subgroup (TRG 0—1 grade, n=35) and poor response (PR) subgroup (TRG 2—3 grade, n=40), and the rest 50 cases were taken as validation set. Dual-energy CT quantitative parameters were measured, calculated and compared between subgroups in training set, and those associated with GR were selected to construct ML models using multilayer perceptron (MLP), random forest (RF), decision tree (DT), support vector machine (SVM), extreme gradient boosting (XGBoost) and logistic regression (LR). Then receiver operating characteristic curves were drawn, and the area under the curve (AUC) was used to identify the best ML model. **Results** Venous phase (VP) normalized iodine concentration (NIC), arterial phase(AP) and VP 100 keV CT values, and AP and VP 140 keV CT values of tumors in GR subgroup were all lower than those in PR subgroup (all P<0.05). After Spearman correlation analysis, AP 140 keV CT values, VP 140 keV CT values and VP NIC were selected for constructing models. MLP and LR models achieved AUC greater than 0.7 for predicting pathological response of gastric cancer to nICT in both training and validation sets, hence both being the optimal models, whereas RF, DT, SVM and XGBoost models exhibited overfitting or inconsistent performance. **Conclusion** MLP and LR models constructed based on dual-energy CT quantitative parameters could be used to predict pathological response of gastric cancer to nICT.

[**Keywords**] stomach neoplasms; tomography, X-ray computed; neoadjuvant therapy; machine learning

DOI: 10.13929/j.issn.1672-8475.2025.12.004

[**基金项目**] 国家自然科学基金(81971615)、河南省科技发展计划项目(242102311026)。

[**第一作者**] 郑月(1999—),女,四川成都人,在读硕士。研究方向:腹部影像诊断。E-mail: yueaza@163.com

[**通信作者**] 高剑波,郑州大学第一附属医院放射科,450052;河南省医学影像国际联合实验室,450052;河南省消化肿瘤影像重点实验室,450052;河南省CT影像重点实验室,450052。E-mail: cjr.gaojianbo@vip.163.com

[**收稿日期**] 2025-08-28 [**修回日期**] 2025-10-05

基于双能CT定量参数机器学习模型预测胃癌新辅助免疫化疗病理反应

郑　月[1,2,3,4]，范松梅[1,2,3,4]，袁梦晨[1,2,3,4]，刘译阳[1,2,3,4]，尤亚茹[1,2,3,4]，寇松子[5]，
程　震[1,2,3,4]，李莉明[1,2,3,4]，高剑波[1,2,3,4*]

（1.郑州大学第一附属医院放射科，5.病理科，河南 郑州　450052；2.河南省医学影像国际联合实验室，河南 郑州　450052；3.河南省消化肿瘤影像重点实验室，河南 郑州　450052；4.河南省CT影像重点实验室，河南 郑州　450052）

［**摘　要**］ **目的**　观察基于双能CT定量参数机器学习（ML）模型预测胃癌新辅助免疫化疗（nICT）病理反应的价值。**方法**　回顾性纳入125例接受nICT的胃癌患者，随机以75例为训练集并根据肿瘤退缩分级（TGR）分为病理缓解良好（GR）亚组（TRG 0～1级，n=35）与病理缓解不良（PR）亚组（TRG 2～3级，n=40），以另外50例为验证集。比较训练集内亚组间双能CT定量参数，筛选与GR相关者，并分别以多层感知器（MLP）、随机森林（RF）、决策树（DT）、支持向量机（SVM）、极限梯度提升（XGBoost）及逻辑回归（LR）建立ML模型；以受试者工作特征曲线下面积（AUC）筛选预测效能最佳者。**结果**　GR亚组肿瘤静脉期（VP）标准化碘浓度（NIC）、动脉期（AP）及VP 100 keV CT值、AP及VP 140 keV CT值均小于PR亚组（P均＜0.05）。经Spearman相关分析后将AP 140 keV CT值、VP 140 keV CT值及VP NIC纳入模型，结果显示MLP及LR模型预测训练集和验证集胃癌nICT病理反应的AUC均＞0.7，均为最优模型，而RF、DT、SVM、XGBoost模型均存在过拟合或效能不稳定。**结论**　基于双能CT定量参数构建的MLP及LR模型可用于预测胃癌nICT病理反应。

［**关键词**］ 胃肿瘤；体层摄影术，X线计算机；新辅助治疗；机器学习

［**中图分类号**］ R735.2；R814.42　［**文献标识码**］ A　［**文章编号**］ 1672-8475（2025）12-0769-06

多数胃癌患者确诊时已处于局部晚期阶段[1]；相比单纯化学治疗（化疗），此时新辅助免疫化疗（neoadjuvant immunochemotherapy，nICT）更有助于改善预后[2-3]，但部分病例可能疗效欠佳或出现治疗相关毒性。目前常以阳性评分法或肿瘤退缩分级（tumor regression grade，TRG）法预测或评估nICT用于胃癌效果，结果异质性高而滞后性强[4-6]。利用双能CT能有效预测新辅助治疗反应[7-12]。本研究观察基于双能CT定量参数机器学习（machine learning，ML）模型预测胃癌对于nICT病理反应的价值。

1　资料与方法

1.1　研究对象　回顾性纳入2022年7月—2025年4月郑州大学第一附属医院125例接受nICT及其后根治性胃切除术的胃腺癌患者，男103例、女22例，年龄22～79岁、平均（62.1±9.2）岁；nICT方案为2～8个周期奥沙利铂＋替吉奥、奥沙利铂＋卡培他滨或5-氟尿嘧啶＋亚叶酸钙＋奥沙利铂，联合免疫检查点抑制剂治疗；随机以75例为训练集，根据第8版美国癌症联合委员会（American Joint Committee on Cancer，AJCC）[13]提出的TRG法划分病理缓解良好（good responder，GR）亚组（TRG 0～1级，n=35）与病理缓解不良（poor responder，PR）亚组（TRG 2～3级，n=40）；以其余50例为验证集。纳入标准：①于治疗前2周内接受腹部双能CT增强扫描；②治疗前临床分期T2NxM0～T4aNxM0。排除标准：①图像质量差；②检查前接受抗肿瘤治疗。本研究经院伦理委员会批准（2022-KY-1447-002）。

1.2　仪器与方法　嘱患者检查前禁食6 h，扫描前10 min饮水1.0～1.5 L。采用Philips IQon或GE Discovery/APEX CT机进行腹部扫描，管电压瞬时切换80/140 kVp或120 kVp，管电流120～550 mA，旋转时间0.5或0.6 s，SFOV Large body，层厚及层间距1.25或1.0 mm。完成平扫后经肘静脉以流率3 ml/s注射对比剂碘海醇（350 mgI/ml）1.5 ml/kg体质量并跟注20 ml生理盐水冲管，以腹主动脉达100 HU为触发阈值，分别延迟11及30 s进行动脉期（arterial phase，AP）及静脉期（venous phase，VP）扫描。

1.3　图像分析　将AP及VP图像导入GE AW4.7或Philips IntelliSpace Portal工作站，重建虚拟单能级图像及碘图。由1名具有7年腹部影像诊断经验的放射科医师（医师1）独立基于混合能量图评估肿瘤T分期、厚度及位置；由另2名分别具有3及5年经验的放射科医师（医师2、3）于AP及VP 40～140 keV虚拟单能级图和碘图显示肿瘤最大切面及其相邻上、下层面中尽

量避开坏死、血管和钙化区域于病灶及腹主动脉处放置圆形ROI，测量病灶CT值、碘值及腹主动脉碘值，均测量3次取平均值；计算AP及VP病灶标准化碘浓度（normalized iodine concentration，NIC）、能谱曲线斜率（λ），以及动脉增强分数（arterial enhancement fraction，AEF）：NIC＝碘浓度（iodine concentration，IC）$_{肿瘤}$/IC$_{腹主动脉}$；λ＝（40 keV CT值$_{肿瘤}$－70 keV CT值$_{肿瘤}$）/30；AEF＝IC$_{AP}$/IC$_{VP}$×100%；以2名医师测值的均值为最终结果。见图1。

1.4 统计学分析 以SPSS 27.0、R软件4.2.1和Deepwise多模态研究平台（2.8.0版本）分析数据。以$\bar{x}\pm s$表示正态分布计量资料，行独立样本t检验；以中位数（上下四分位数）表示偏态分布计量资料，行Mann-Whitney U检验。以χ^2检验或Fisher精确概率法比较计数资料。基于组内相关系数（intra-class correlation coefficient，ICC）行观察者间一致性检验，ICC>0.75为一致性良好。采用Spearman相关分析评估双能CT定量参数间的共线性，剔除r_S>0.9者，以最小绝对收缩和选择算子（least absolute shrinkage and selection operator，LASSO）算法对剩余参数进一步降维；通过5折交叉验证确定误差最小λ值，获得最终特征子集；基于该特征子集采用6种ML算法，包括多层感知器（multilayer perceptron，MLP）、随机森林（random forest，RF）、决策树（decision tree，DT）、支持向量机（support vector machine，SVM）、极限梯度提升（extreme gradient boosting，XGBoost）和逻辑回归（logistic regression，LR）构建预测模型；以5折交叉验证评估模型效能。绘制受试者工作特征（receiver operating characteristic，ROC），计算曲线下面积（area under the curve，AUC），据以筛选预测效能最佳的ML模型，并以之联合组间差异有统计学意义的临床资料构建联合模型，于验证集进行验证。采用DeLong检验比较AUC差异。以校准曲线评估模型拟合优度，并利用决策曲线分析（decision curve analysis，DCA）评估模型临床净获益。P<0.05为差异有统计学意义。

2 结果

2.1 训练集内亚组间比较 2名医师测量参数的一致性良好（ICC为0.853～0.991）。GR亚组肿瘤临床T分期、VP NIC、AP及VP 100 keV CT值、AP及VP 140 keV CT值均小于PR亚组（P均<0.05），见表1。

2.2 模型预测效能 经Spearman相关分析筛选后，将AP 140 keV CT值、VP 140 keV CT值及VP NIC纳入模型；MLP及LR模型预测训练集和验证集胃癌nICT病理反应的AUC均>0.7，为最优模型。见表2。

2.3 联合模型预测效能 分别以MLP及LR模型联合临床T分期构建cMLP与cLR模型，其预测训练集胃癌nICT病理反应的AUC为0.791及0.794，在验证集分别为0.752及0.747，与MLP及LR模型差异均无统计学意义（Z=0.186～1.201，P均>0.05）；且在训练集和验证集的拟合优度均良好，临床净获益均显著。见图2～4。

3 讨论

随着胃癌诊疗模式的不断改善，nICT用于局部进展期胃癌获得显著临床收益[14]。双能CT衍生的碘图能直观描绘碘分布，定量测量IC，反映肿瘤新生血管和血供。本研究基于双能CT多参数联合临床T分期

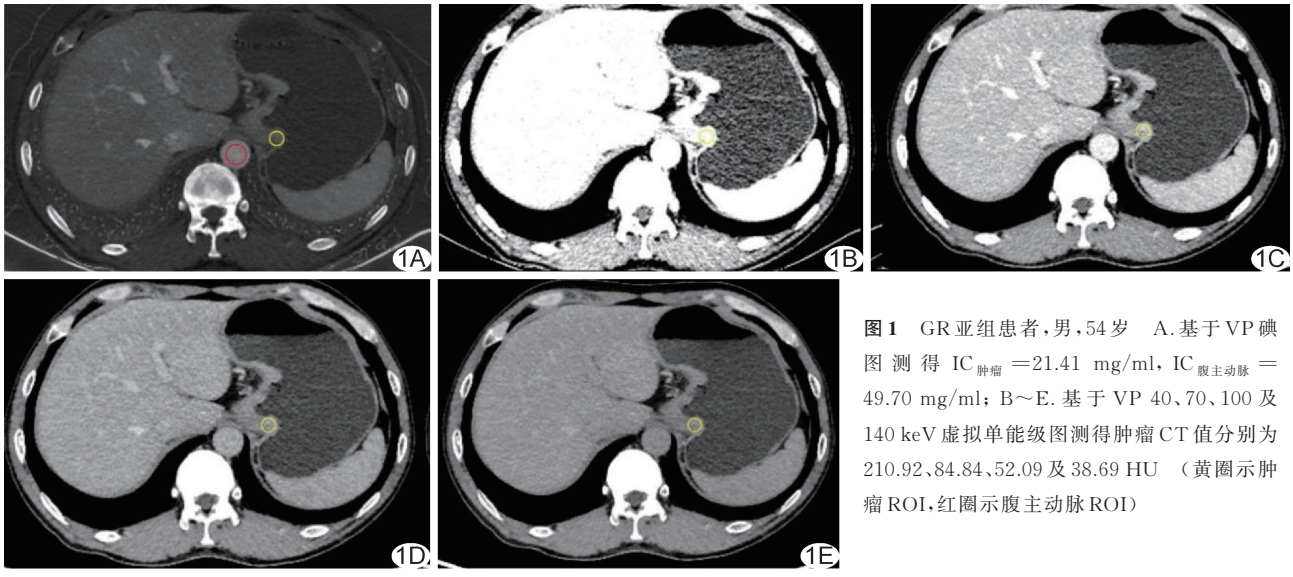

图1 GR亚组患者，男，54岁 A.基于VP碘图测得IC$_{肿瘤}$=21.41 mg/ml，IC$_{腹主动脉}$=49.70 mg/ml；B～E.基于VP 40、70、100及140 keV虚拟单能级图测得肿瘤CT值分别为210.92、84.84、52.09及38.69 HU （黄圈示肿瘤ROI，红圈示腹主动脉ROI）

表1　训练集患者临床资料及病灶双能CT定量参数比较

亚组别	年龄(岁)	性别(例)		CA125(例)		CA724(例)		CA199(例)	
		男	女	≤35 U/ml	>35 U/ml	≤6.9 U/ml	>6.9 U/ml	≤35 U/ml	>35 U/ml
GR亚组(n=35)	63.20±8.03	27	8	32	3	24	11	27	8
PR亚组(n=40)	61.47±8.38	33	7	38	2	30	10	35	5
$t/Z/\chi^2$值	0.907	0.335		—		0.383		1.397	
P值	0.368	0.563		0.877*		0.536		0.237	

亚组别	CEA(例)		临床T分期(例)			肿瘤厚度(mm)	肿瘤位置(例)			
	≤5 ng/ml	>5 ng/ml	T2	T3	T4a		上部	中部	下部	≥2/3部
GR亚组(n=35)	28	7	8	16	11	20.23±8.60	20	6	1	8
PR亚组(n=40)	31	9	1	26	13	20.82±5.86	28	1	5	6
$t/Z/\chi^2$值	0.070		—			−0.351	—			
P值	0.792		0.021*			0.727	0.056*			

亚组别	IC(mg/ml)		NIC(%)		40 keV CT值(HU)		70 keV CT值(HU)	
	AP	VP	AP	VP	AP	VP	AP	VP
GR亚组	14.27±	19.79	12.00	36.00	142.72	192.66	66.31±	83.75
(n=35)	6.97	(14.40,22.67)	(8.50,15.50)	(29.50,43.50)	(108.02,175.60)	(147.28,224.13)	16.57	(63.39,93.29)
PR亚组	15.47±	20.55	14.00	41.00	149.11	202.22	74.25±	88.71
(n=40)	6.68	(16.97,26.19)	(10.80,17.00)	(34.00,55.50)	(129.76,192.57)	(171.50,245.84)	18.03	(78.40,98.05)
$t/Z/\chi^2$值	−0.763	−1.635	−1.423	−2.443	−1.200	−1.412	−1.977	−1.827
P值	0.448	0.102	0.155	0.015	0.230	0.158	0.052	0.068

亚组别	100 keV CT值(HU)		140 keV CT值(HU)		λ		AEF(%)
	AP	VP	AP	VP	AP	VP	
GR亚组(n=35)	44.45(38.58,51.95)	51.69(47.13,59.25)	37.13±8.56	40.78±8.65	2.76±1.31	3.64(2.79,4.39)	73.00(58.50,83.50)
PR亚组(n=40)	50.79(47.44,54.28)	59.09(53.38,63.18)	43.69±7.33	47.91±9.46	3.01±1.27	3.85(3.13,4.92)	70.50(55.80,86.00)
$t/Z/\chi^2$值	−2.825	−2.692	−3.577	−3.386	−0.826	−1.221	0.234
P值	0.005	0.007	0.001	0.001	0.412	0.222	0.815

注：*：采用Fisher精确概率法；CA：糖类抗原(carbohydrate antigen)；CEA：癌胚抗原(carcinoembryonic antigen)

表2　各模型预测胃癌nICT病理反应的效能

模型	集别	敏感度(%)	特异度(%)	准确率(%)	阳性预测值(%)	阴性预测值(%)	AUC
MLP模型	训练集	62.90	75.00	69.30	68.80	69.80	0.772
	验证集	54.20	76.90	66.00	68.40	64.50	0.744
RF模型	训练集	97.10	97.50	97.30	97.10	97.50	0.993
	验证集	58.30	73.10	66.00	66.70	65.50	0.571
DT模型	训练集	91.40	57.50	73.30	65.30	88.50	0.846
	验证集	75.00	50.00	62.00	58.10	68.40	0.663
SVM模型	训练集	54.30	92.50	74.70	86.40	69.80	0.809
	验证集	41.70	88.50	66.00	76.90	62.20	0.639
XGBoost模型	训练集	94.30	95.00	94.70	94.30	95.00	0.988
	验证集	54.20	61.50	58.00	56.50	59.30	0.617
LR模型	训练集	74.30	60.00	66.70	61.90	72.70	0.761
	验证集	79.20	53.80	66.00	61.30	73.70	0.729

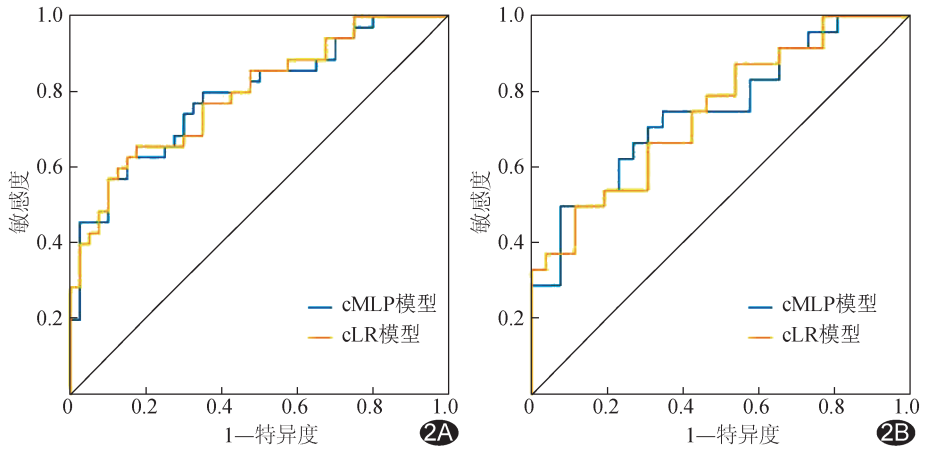

图2　cMLP与cLR模型预测胃癌nICT病理反应的ROC曲线　A.训练集；B.验证集

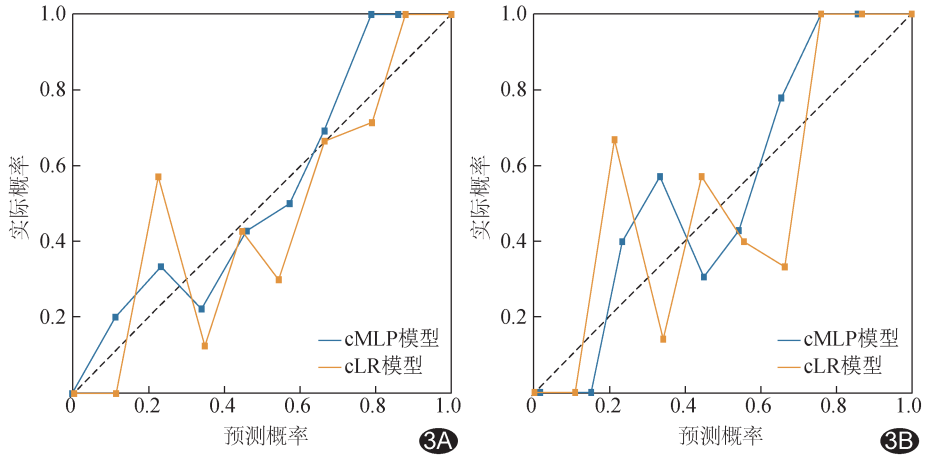

图3　cMLP与cLR模型预测胃癌nICT病理反应的校准曲线　A.训练集；B.验证集

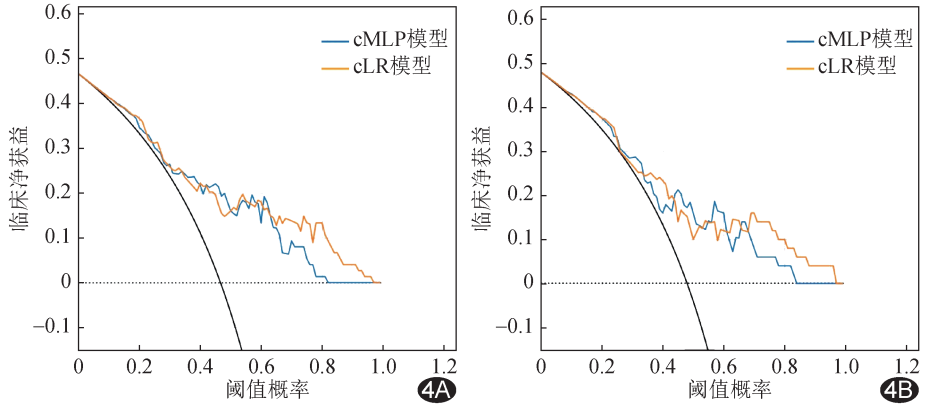

图4　cMLP与cLR模型预测胃癌nICT病理反应的DCA图　A.训练集；B.验证集

构建多种ML模型，观察其预测胃癌nICT病理反应的价值；125例中，59例（59/125，47.20%）达病理GR，与以往报道[2]一致；训练集GR亚组VP NIC显著低于PR亚组，主要原因在于低VP NIC胃癌血供相对差、血管生成水平较低，nICT后预后更佳[9]；PR亚组在AP及VP 100及140 keV虚拟单能级图中的CT值较高，可能与高能级图的低噪声特性使其表征组织非增强结构更佳、反映肿瘤异质性更好[15-16]有关。

本研究基于AP 140 keV CT值、VP 140 keV CT值及VP NIC构建的模型中，MLP及LR模型预测训练集和验证集胃癌nICT病理反应的性能最为稳健（AUC均＞0.7），可归因于其采取不同泛化控制策略应对有限样本量：MLP模型可自动学习建模特征间复杂的非线性关系、高阶关系及交互效应，通过结构约束展现其泛化能力；LR模型则能清晰解释系数，直接量化各特征贡献，且其先行决策边界在小样本数据中不易过拟合，证实基于双能CT定量参数ML模型可通过整合虚拟单能级图像和碘图信息而更全面地分析肿瘤生物学特征。本研究SVM及DT模型均可见明显训练集和验证集内敏感度和特异度不均衡及跨数据集性能不稳定：训练集中，DT模型呈高敏感度（91.40%）、低特异度（57.50%），易将PR误判为GR；SVM模型则相反，特异度高（92.50%）但敏感度不足（54.30%），表明其识别GR能力有限；而RF和XGBoost模型在验证集的预测效能出现明显下降，推测主要与过拟合有关：XGBoost引入正则化，但对于有限样本仍难以稳定泛化，RF采用多数投票机制[17]，用于有限样本时易出现过拟合。

本研究所获联合模型预测胃癌nICT病理反应的效能与MLP及LR模型均无明显差异，提示临床T分期可能与双能CT参数间存在非线性关联。cMLP模型在验证集的AUC略高于cLR模型（0.752 *vs.* 0.747），可能由于MLP能捕捉各临床T分期与双能CT参数间的特异性模式，而LR仅能为所有样本赋予固定权重却无法实现条件性建模，即MLP算法在处理随疾病分期变化的关系时更具灵活性，与既往研究[18]一致。

综上，基于双能CT定量参数构建MLP及LR模型可用于预测胃癌nICT病理反应。但本研究为单中心、回顾性分析，样本量有限，可能限制模型泛化能力并导致选择偏倚；未来需通过大规模、多中心、前瞻性研究进一步验证模型临床适用性。

利益冲突：全体作者声明无利益冲突。

作者贡献：郑月研究实施、数据分析、撰写文章；范松梅研究实施、数据分析；袁梦晨指导、修改文章；刘译阳、尤亚茹、程震统计分析；寇松子图像和数据分析；李莉明和高剑波指导、研究设计、审阅文章。

[参考文献]

[1] FILHO A M, LAVERSANNE M, FERLAY J, et al. The GLOBOCAN 2022 cancer estimates: Data sources, methods, and a snapshot of the cancer burden worldwide[J]. Int J Cancer, 2025, 156(7): 1336-1346.

[2] SUN Y Q, ZHONG Q, LV C B, et al. The safety and efficacy of neoadjuvant immunochemotherapy following laparoscopic gastrectomy for gastric cancer: A multicentre real-world clinical study[J]. Int J Surg, 2024, 110(8): 4830-4838.

[3] SHITARA K, RHA S Y, WYRWICZ L S, et al. Neoadjuvant and adjuvant pembrolizumab plus chemotherapy in locally advanced gastric or gastro-oesophageal cancer (KEYNOTE-585): An interim analysis of the multicentre, double-blind, randomised phase 3 study[J]. Lancet Oncol, 2024, 25(2): 212-224.

[4] LIU Y Q, HU P B, XU L, et al. Current progress on predictive biomarkers for response to immune checkpoint inhibitors in gastric cancer: How to maximize the immunotherapeutic benefit? [J]. Cancers (Basel), 2023, 15(8): 2273.

[5] KANG W D, QIU X, LUO Y G, et al. Application of radiomics-based multiomics combinations in the tumor microenvironment and cancer prognosis[J]. J Transl Med, 2023, 21(1): 598.

[6] MCGRAIL D J, PILIE P G, RASHID N U, et al. High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types[J]. Ann Oncol, 2021, 32(5): 661-672.

[7] TAN X Y, YANG X, HU S D, et al. A nomogram for predicting postoperative complications based on tumor spectral CT parameters and visceral fat area in gastric cancer patients[J]. Eur J Radiol, 2023, 167: 111072.

[8] GE H T, CHEN J W, WANG L L, et al. Preoperative prediction of lymphovascular and perineural invasion in gastric cancer using spectral computed tomography imaging and machine learning[J]. World J Gastroenterol, 2024, 30(6): 542-555.

[9] 李靖，王艺，魏莹莹，等. 能谱CT术前诊断胃腺癌周围神经侵犯[J]. 中国医学影像技术，2022，38(12)：1843-1847.

[10] LI J, XU S N, WANG Y, et al. Spectral CT *vs.* diffusion-weighted imaging for the quantitative prediction of pathologic response to neoadjuvant chemotherapy in locally advanced gastric cancer[J]. Eur Radiol, 2024, 34(9): 6193-6204.

[11] DENG J, ZHANG W, XU M, et al. Value of spectral CT parameters in predicting the efficacy of neoadjuvant chemotherapy for gastric cancer[J]. Clin Radiol, 2024, 79(1): 51-59.

[12] LI J, CHEN X J, XU S N, et al. Predicting pathologic response to neoadjuvant chemotherapy in locally advanced gastric cancer: The establishment of a spectral CT-based nomogram from prospective datasets[J]. Eur J Surg Oncol, 2024, 50(4): 108020.

[13] AMIN M B, GREENE F L, EDGE S B, et al. The eighth edition AJCC cancer staging manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging[J]. CA Cancer J Clin, 2017, 67(2): 93-99.

[14] VERSCHOOR Y L, van de HAAR J, van den BERG J G, et al. Neoadjuvant atezolizumab plus chemotherapy in gastric and gastroesophageal junction adenocarcinoma: The phase 2 PANDA trial[J]. Nat Med, 2024, 30(2): 519-530.

[15] LIU Y Y, ZHANG H, WANG L, et al. Predicting response to systemic chemotherapy for advanced gastric cancer using pre-treatment dual-energy CT radiomics: A pilot study[J]. Front Oncol, 2021, 11: 740732.

[16] 尤亚茹，刘译阳，袁梦晨，等. 能谱CT定量参数评估进展期胃癌T分期[J]. 中国医学影像技术，2024，40(11)：1704-1709.

[17] YOU Y R, LIANG Y, CHEN L H, et al. Radiomics analysis of dual-energy CT-derived iodine maps for differentiating between T1/2 and T3/4a in gastric cancer: A multicenter study[J]. Eur J Radiol, 2025, 186: 112054.

[18] TAN X Y, YANG X, HU S D, et al. Prediction of response to neoadjuvant chemotherapy in advanced gastric cancer: A radiomics nomogram analysis based on CT images and clinicopathological features[J]. J Xray Sci Technol, 2023, 31(1): 49-61.