

基于 PET/CT 的深度学习-影像组学-临床模型预测非小细胞肺癌的隐匿性淋巴结转移

于军¹ 何超² 李洋¹ 陈聪¹ 任春玲¹ 黄磊¹ 任东栋¹

¹宁波明州医院核医学科, 宁波 315100; ²苏州弘慈血液病医院核医学科, 苏州 215100

通信作者: 任东栋, Email: rendongdong1977@126.com

【摘要】 目的 探讨基于治疗前¹⁸F-FDG PET/CT 的影像组学、深度学习模型及临床特征预测非小细胞肺癌 (NSCLC) 患者隐匿性淋巴结转移 (OLM) 的价值。**方法** 回顾性队列研究。连续纳入 2019 年 1 月至 2024 年 6 月间于宁波明州医院行 PET/CT 检查、临床分期为 T1~3N0M0 的 NSCLC 患者 283 例[男 167 例、女 116 例, 年龄 67(60, 73) 岁]。采用简单随机抽样按 8:2 将患者分为训练集($n=226$)和验证集($n=57$)。另纳入同期 121 例苏州弘慈血液病医院 NSCLC 患者[男 72 例、女 49 例, 年龄 66(60, 72) 岁]作为测试集。逐层勾画肿瘤 ROI 提取 PET/CT 影像组学特征, 并从原发肿瘤面积最大层面及其上下各 2 层提取深度学习模型特征。基于临床特征、影像组学特征及深度学习特征构建 7 种预测模型, 行 ROC 曲线分析评估模型对 OLM 的诊断效能。根据最佳预测模型的最大约登指数确定最佳阈值, 并据其将患者分为高、低风险组, 比较 2 组间生存差异(log-rank 检验)。**结果** 在 7 种模型中, 深度学习-影像组学-临床(DRC)模型(采用 logistic 回归分类器)在验证集与测试集中表现最优, AUC 分别为 0.917 与 0.879, 准确性分别为 82.5%(47/57) 与 84.3%(102/121)。亚组分析显示, DRC 模型在 T1 期及 T2~3 期患者中的 AUC 分别为 0.903 与 0.842, 在肺腺癌、肺鳞状细胞癌患者中的 AUC 分别为 0.925 与 0.773。生存分析显示, 高风险组(≥ 0.210 ; $n=29$) 患者较低风险组(< 0.210 ; $n=87$) 的无进展生存(PFS)与总生存(OS)差(χ^2 值: 8.51、8.83, P 值: 0.004、0.003)。**结论** 融合 PET/CT 影像组学、深度学习及临床特征的 DRC 模型可有效预测 NSCLC 患者的 OLM 与预后。

【关键词】 癌, 非小细胞肺; 淋巴转移; 深度学习; 影像组学; 正电子发射断层显像术; 体层摄影术, X 线计算机; 氟脱氧葡萄糖 F18

DOI: 10.3760/cma.j.cn321828-20250721-00255

PET/CT-based deep learning-radiomics-clinical model for predicting occult lymph node metastasis of non-small cell lung cancer

Yu Jun¹, He Chao², Li Yang¹, Chen Cong¹, Ren Chunling¹, Huang Lei¹, Ren Dongdong¹

¹Department of Nuclear Medicine, Ningbo Mingzhou Hospital, Ningbo 315100, China; ²Department of Nuclear Medicine, Suzhou Hongci Hematology Hospital, Suzhou 215100, China

Corresponding author: Ren Dongdong, Email: rendongdong1977@126.com

【Abstract】 Objective To investigate the value of pre-therapy ¹⁸F-FDG PET/CT-based radiomics and deep learning (DL) models and clinical features in predicting occult lymph node metastasis (OLM) in patients with non-small cell lung cancer (NSCLC). **Methods** This was a retrospective cohort study. A total of 283 patients (167 males, 116 females; age 67(60, 73) years) with clinically staged T1-3N0M0 NSCLC who underwent ¹⁸F-FDG PET/CT from January 2019 to June 2024 in Ningbo Mingzhou Hospital were included consecutively. Using simple random sampling, patients were divided into training ($n=226$) and validation ($n=57$) sets at a ratio of 8:2. An additional 121 patients with NSCLC (72 males, 49 females; age 66(60, 72) years) from Suzhou Hongci Hematology Hospital were enrolled as a test set. Radiomics features were extracted from manually contoured tumor ROIs on PET/CT images, and DL features were extracted from the slice with the largest primary tumor area and two adjacent slices (± 2). Seven prediction models were constructed by integrating clinical features, radiomics features, and DL features. ROC curve analysis was performed to evaluate the diagnostic performance of different models for OLM prediction. Using the optimal predictive model, patients were stratified into high-risk and low-risk groups based on the maximum Youden index threshold, and the survival differences were compared using log-rank test. **Results** Among the 7 models, DL-radiomics-clinical (DRC) model (logistic regression classifier) demonstrated optimal performance in both validation set and test set, with AUCs of 0.917 and 0.879, respectively. Subgroup analysis revealed that the DRC model achieved AUCs of 0.903 for T1 stage and 0.842 for T2-3 stage, respectively. For pathological subtypes, the model showed AUCs of 0.925 for adenocarcinoma and 0.773 for squamous

cell carcinoma, respectively. Survival analysis showed that patients in the high-risk group (≥ 0.210 ; $n=29$) had poorer progression-free survival and overall survival compared to those in the low-risk group (<0.210 ; $n=87$; χ^2 values: 8.51, 8.83, P values: 0.004, 0.003). **Conclusion** The DRC model integrating PET/CT radiomics features, DL features, and clinical features, can effectively predict OLM and prognosis in NSCLC patients.

【Key words】 Carcinoma, non-small-cell lung; Lymphatic metastasis; Deep learning; Radiomics; Positron-emission tomography; Tomography, X-ray computed; Fluorodeoxyglucose F18

DOI:10.3760/cma.j.cn321828-20250721-00255

肺癌是全球癌症死亡的首要原因,其中非小细胞肺癌(non-small cell lung cancer, NSCLC)约占85%^[1]。淋巴结转移(lymph node metastasis, LNM)是NSCLC的主要转移途径,直接影响疾病分期、治疗策略及预后。¹⁸F-FDG PET/CT是评估淋巴结状态的重要无创手段,但其诊断LNM的灵敏度仍有不足;在PET/CT诊断为N0期的NSCLC患者中,仍有14.3%~23.1%存在LNM,即隐匿性LNM(occult LNM, OLM)^[2]。影像组学可通过量化分析肿瘤异质性信息提升LNM预测能力^[3-4],而深度学习(deep learning, DL)能够自动提取影像中的深层特征。研究表明,融合影像组学与DL特征可能产生协同效应,其预测性能优于单一方法^[5-7]。然而,如何有效整合PET/CT的影像组学特征、DL特征与临床特征,构建多模态模型用于早期NSCLC患者的OLM预测,相关研究仍有限。因此,本研究旨在开发一种DL-影像组学-临床(DL-radiomics-clinical, DRC)模型,以实现早期NSCLC患者OLM状态更准确、可靠的无创术前评估。

资料与方法

1.研究对象。回顾性队列研究。连续纳入2019年1月至2024年6月来自宁波明州医院(中心1)和苏州弘慈血液病医院(中心2)的NSCLC患者。纳入标准:(1)接受手术切除+系统性淋巴结清扫(≥ 3 组肺门淋巴结+3组纵隔淋巴结);(2)术后病理确诊为NSCLC[腺癌或鳞状细胞癌(鳞癌)],且淋巴结转移经病理证实;(3)术前1个月内行¹⁸F-FDG PET/CT检查;(4)临床分期为T1~3N0M0期(淋巴结短径 <1.0 cm, ¹⁸F-FDG摄取 $SUV_{max} < 2.5$ ^[8-9]且无远处转移);(5)临床数据完整。排除标准:(1)术前曾接受任何系统性抗肿瘤治疗(如化疗、放疗、靶向或免疫治疗等);(2)合并其他恶性肿瘤;(3)存在肺内多发肿瘤病灶(≥ 2 个)或肺原发肿瘤病灶长径 <1.0 cm;(4)图像不完整或质量差;(5)影像组学或DL特征提取失败。

最终纳入404例患者,男239例、女165例,年龄67(60,73)岁,年龄范围:26~85岁。采用简单随

机抽样法,将来自中心1的283例患者[男167例、女116例,年龄67(60,73)岁]按8:2分为训练集($n=226$)和验证集($n=57$);来自中心2的121例患者[男72例、女49例,年龄66(60,72)岁]作为测试集。本研究已获得上述2家医院医学伦理委员会批准[批准号:明伦202301052,(2025)伦研批第009号]。

2. ¹⁸F-FDG PET/CT显像和临床代谢参数获取。所有患者在术前1个月内接受¹⁸F-FDG PET/CT检查。其中,中心1的患者使用Biograph mCT仪(德国Siemens公司),中心2的患者使用Discovery IQ仪(美国GE公司)。PET图像按照标准方案采集:患者检查前禁食至少6 h,确保血糖水平低于11.1 mmol/L。按患者体质量经静脉注射¹⁸F-FDG(南京江原安迪科正电子研究发展有限公司产品,放化纯 $>95\%$)3.70~5.55 MBq/kg后,嘱其安静平躺休息60 min,随后进行PET/CT扫描。扫描范围为股骨中段至颅顶。先行CT扫描,层厚2 mm,用于衰减校正;随后行PET扫描,每个床位扫描2 min,共采集7~10个床位。PET图像采用有序子集最大期望值迭代法重建。

由2位经验丰富的核医学科医师采用盲法独立评估患者临床资料及影像数据,有分歧时协商解决。记录及分析的指标包括:(1)临床病理参数,包括年龄、性别、吸烟史、术前癌胚抗原(carcinoembryonic antigen, CEA)水平、组织学类型(腺癌/鳞癌)、肿瘤长径、肿瘤位置(中央型与周围型);(2)影像代谢参数,包括 SUV_{max} 、 SUV_{mean} 、SUV峰值(peak of SUV, SUV_{peak})等。

3.图像预处理和肿瘤分割。在特征提取前,对所有PET和CT图像进行预处理。首先,使用SimpleITK软件包(2.0.2版)将所有图像重采样至统一的体素空间(1 mm $\times$ 1 mm $\times$ 1 mm),并对图像灰度值进行标准化处理以消除体素尺寸和设备差异的影响。随后采用三维(three-dimensional, 3D)Slicer软件(5.8.1版;https://www.slicer.org)进行肿瘤区域的勾画。由1位核医学科医师采用盲法,在CT和PET图像上,沿肿瘤边界逐层进行手动ROI勾画,过程中注意避开正常肺组织、心脏等。PET图像肿瘤边界难以

确定时参照 CT 图像。为评估不同医师手动勾画的可重复性,采用简单随机抽样法从中心 1 抽取 50 例患者,由另 1 位核医学科医师盲法勾画。计算组内相关系数(intraclass correlation coefficient, ICC),仅保留 ICC>0.75 的特征用于后续分析。

4. 影像组学及 DL 特征提取与筛选。(1) 特征提取。①影像组学特征:使用 PyRadiomics 软件包(3.1.0 版),分别从 CT 及 PET 图像 ROI 中提取影像组学特征。提取的特征类别包括一阶特征、形状特征及纹理特征等。提取的影像组学特征均符合图像生物标志物标准化倡议(image biomarker standardization initiative, IBSI)的定义。② DL 特征:采用 9 种在 ImageNet 数据集上预训练的模型提取特征,包括 DenseNet121、EfficientNet、GoogleNet、MobileNet_V2、ResNet18、ResNet50、ResNet101、ShuffleNet_V2 和 VGG16 模型。首先,选择原发肿瘤面积最大的层面(第 n 层)及其上($n+2$)和下($n-2$)层面作为原始图像,将其分辨率归一化为 224×224 像素。然后,从每个模型倒数第 2 层中提取 DL 特征。

(2) 特征筛选。所有提取的影像组学特征及 DL 特征均进行 Z-score 归一化处理。针对训练集中类别不平衡的问题,采用合成少数类过采样技术进行数据平衡。特征选择流程:①通过 Mann-Whitney U 检验筛选与 OLM 状态相关的特征;②对 $r>0.9$ 的特征进行去冗余,仅保留其一;③使用最小冗余最大相关性算法(minimum redundancy maximum relevance, mRMR)对剩余特征进行重要性排序,保留排序前 50 位的特征;④将上述 50 个特征输入最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)分类器,通过 10 折交叉验证优化参数,并选择出与 OLM 状态相关及系数非零的影像组学特征、DL 特征。训练集的临床资料行单因素及多因素 logistic 回归分析,筛选出影响 OLM 状态的独立影响因素。

5. 模型构建及亚组分析。基于筛选出的影像组学特征、DL 特征及临床独立影响因素,采用 7 种机器学习方法(包括 logistic 回归、K 最近邻、支持向量机、极限梯度提升、随机森林、轻量级梯度提升及多层感知器),分别构建 7 种预测模型:临床模型、影像组学模型、DL 模型、影像组学-临床模型、DL-临床模型、DL-影像组学模型,以及 DRC 模型。另外,在验证集与测试集中,根据肿瘤 T 分期将患者分为 T1 组和 T2~3 组,并根据病理类型分为腺癌组和鳞癌组,以评估 DRC 模型在各亚组中的诊断效能。

6. 随访和生存分析。通过门诊记录、健康档案及电话随访获取验证集及测试集患者随访数据,并进行生存分析。随访截止日期为 2025 年 6 月 30 日。主要生存结局包括总生存(overall survival, OS)和无进展生存(progression-free survival, PFS)。OS 定义为从手术日期至因任何原因死亡日期或末次随访日期的时间间隔;PFS 定义为从手术日期至肿瘤复发、转移或因任何原因死亡的时间段。根据 DRC 模型生成的得分,通过 ROC 曲线最大约登指数确定最佳阈值,据其将患者分为高风险组和低风险组。采用 Kaplan-Meier 法绘制不同风险组患者的 OS 与 PFS 曲线,行 log-rank 检验比较不同分组患者的生存差异。

7. 统计学处理。采用 IBM SPSS Statistics 24.0、Python 3.9.21、R 4.5.0 进行统计学分析。不符合正态分布的定量资料采用 $M(Q_1, Q_3)$ 表示,采用 Kruskal-Wallis 秩和检验比较组间差异;定性变量采用频数(百分比)表述,行 χ^2 检验比较组间差异。采用 ROC 曲线评估并计算不同预测模型对 OLM 状态的诊断效能,并采用 Delong 检验比较 AUC 的差异。采用校准曲线评估模型预测概率与实际结果的一致性,采用决策曲线分析(decision curve analysis, DCA)评估不同模型的净获收益。 $P<0.05$ 为差异有统计学意义(双侧检验)。

结 果

1. 患者临床特征分析。不同数据集患者组间比较见表 1,各集患者的临床特征差异无统计学意义(H 值:0.28~4.58, χ^2 值:0.10~1.21,均 $P>0.05$)。训练集临床特征的 logistic 回归分析结果显示,年龄、肿瘤位置、术前 CEA、 SUV_{max} 及 SUV_{mean} 是 OLM 状态的独立预测因子(表 2),用于构建临床模型。

2. 特征提取及筛选。在影像组学特征提取中,从每例患者的 PET 和 CT 图像中分别提取 1 688 个特征,包括一阶统计特征 18 个、形状特征 14 个、纹理特征 75 个,以及基于 LoG、Wavelet、Square 等滤波器生成的特征 1 581 个。在 DL 特征提取中,基于前述 9 种方法预训练网络,从每个层面分别提取了 1 024、1 280、1 024、1 280、512、2 048、2 048、1 024 和 4 096 个特征。经 mRMR 和 LASSO 方法筛选后,分别有 28、35 和 22 个特征用于影像组学模型、DL 模型及 DL-影像组学模型的构建。

3. 预测模型的诊断效能(表 3)。针对 7 种预测模型,不同机器学习方法中,logistic 回归的 DRC 模型 AUC 较高:验证集 AUC 为 0.917,明显高于临床模

表 1 不同数据集 NSCLC 患者临床资料及代谢参数的比较

组别	例数	年龄 [岁; $M(Q_1, Q_3)$]	性别 男/女	吸烟史[例(%)]		病理类型[例(%)]		肿瘤位置[例(%)]	
				无	有	腺癌	鳞癌	中央型	周围型
训练集	226	67(60,73)	137/89	113(50.00%)	113(50.00%)	162(71.68%)	64(28.32%)	40(17.70%)	186(82.30%)
验证集	57	67(61,71)	30/27	31(54.39%)	26(45.61%)	38(66.67%)	19(33.33%)	10(17.54%)	47(82.46%)
测试集	121	66(60,72)	72/49	64(52.89%)	57(47.11%)	89(73.55%)	32(26.45%)	23(19.01%)	98(80.99%)
检验值		$H=0.28$	$\chi^2=1.21$	$\chi^2=0.49$		$\chi^2=0.91$		$\chi^2=0.10$	
P 值		0.871	0.546	0.784		0.635		0.950	

组别	肿瘤长径	CEA	SUV _{max}	SUV _{mean}	SUV _{peak}	MTV	TLG
	[cm; $M(Q_1, Q_3)$]	[$\mu\text{g/L}; M(Q_1, Q_3)$]	[$M(Q_1, Q_3)$]	[$M(Q_1, Q_3)$]	[$M(Q_1, Q_3)$]	[$\text{cm}^3; M(Q_1, Q_3)$]	[g; $M(Q_1, Q_3)$]
训练集	2.60(2.00,3.58)	2.72(1.72,5.13)	8.59(4.91,13.58)	5.08(2.73,7.97)	5.82(3.10,9.33)	4.78(2.46,9.46)	20.30(7.95,56.63)
验证集	2.40(1.80,3.20)	3.80(2.10,6.23)	8.18(4.05,13.48)	4.88(2.26,7.77)	5.35(2.65,9.43)	4.01(2.19,7.70)	14.80(7.22,35.50)
测试集	2.60(2.00,3.10)	2.60(1.50,5.10)	8.44(4.57,13.84)	4.79(2.56,8.99)	5.60(2.86,10.70)	3.95(2.46,6.64)	21.25(7.70,49.88)
检验值	$H=1.61$	$H=4.58$	$H=0.39$	$H=0.42$	$H=0.61$	$H=2.99$	$H=2.12$
P 值	0.447	0.101	0.823	0.812	0.736	0.224	0.346

注: NSCLC 为非小细胞肺癌, CEA 为癌胚抗原, SUV_{peak} 为 SUV 峰值, MTV 为肿瘤代谢体积, TLG 为病灶糖酵解总量; 鳞癌为鳞状细胞癌

表 2 Logistic 回归分析显示有统计学意义的预测训练集 NSCLC 患者($n=226$)隐匿性淋巴结转移状态的变量

变量	单因素分析		多因素分析	
	OR (95% CI)	P 值	OR (95% CI)	P 值
年龄 (岁)	0.950 (0.923~0.977)	<0.001	0.924 (0.891~0.955)	<0.001
肿瘤位置(0 为中央型, 1 为周围型)	0.184 (0.104~0.314)	<0.001	0.214 (0.110~0.404)	<0.001
病灶最大径 (cm)	1.325 (1.073~1.645)	0.010	1.123 (0.760~1.655)	0.557
CEA ($\mu\text{g/L}$)	1.089 (1.039~1.155)	0.002	1.154 (1.079~1.247)	<0.001
SUV _{max}	1.131 (1.086~1.181)	<0.001	0.384 (0.217~0.662)	0.001
SUV _{mean}	1.230 (1.151~1.320)	<0.001	6.862 (2.377~21.597)	0.001
SUV _{peak}	1.181 (1.119~1.252)	<0.001	0.901 (0.481~1.678)	0.743
TLG (g)	1.004 (1.002~1.007)	0.002	0.997 (0.993~1.002)	0.211

注: OR 为比值比

表 3 不同模型对 NSCLC 患者隐匿性淋巴结转移的预测效能

模型	数据集	例数	AUC	95% CI	灵敏度	特异度	准确性	约登指数
临床模型	训练集	226	0.821	0.766~0.876	88.5%(54/61)	64.8%(107/165)	71.2%(161/226)	0.533
	验证集	57	0.695	0.535~0.856	9/16	85.4%(35/41)	77.2%(44/57)	0.417
	测试集	121	0.717	0.607~0.828	76.2%(16/21)	64.0%(64/100)	66.1%(80/121)	0.402
影像组学模型	训练集	226	0.843	0.791~0.896	72.1%(44/61)	81.8%(135/165)	79.2%(179/226)	0.539
	验证集	57	0.704	0.554~0.853	10/16	82.9%(34/41)	77.2%(44/57)	0.454
	测试集	121	0.734	0.642~0.827	90.5%(19/21)	54.0%(54/100)	60.3%(73/121)	0.445
深度学习模型	训练集	226	0.967	0.947~0.986	95.1%(58/61)	89.1%(147/165)	90.7%(205/226)	0.842
	验证集	57	0.877	0.774~0.981	12/16	87.8%(36/41)	84.2%(48/57)	0.628
	测试集	121	0.771	0.666~0.876	71.4%(15/21)	74.0%(74/100)	73.6%(89/121)	0.454
影像组学-临床模型	训练集	226	0.861	0.811~0.910	72.1%(44/61)	87.3%(144/165)	83.2%(188/226)	0.594
	验证集	57	0.771	0.638~0.903	11/16	82.9%(34/41)	78.9%(45/57)	0.517
	测试集	121	0.770	0.684~0.857	90.5%(19/21)	63.0%(63/100)	67.8%(82/121)	0.535
深度学习-临床模型	训练集	226	0.964	0.943~0.984	96.7%(59/61)	88.5%(146/165)	90.7%(205/226)	0.852
	验证集	57	0.897	0.812~0.982	13/16	85.4%(35/41)	84.2%(48/57)	0.667
	测试集	121	0.825	0.730~0.919	81.0%(17/21)	70.0%(70/100)	71.9%(87/121)	0.510
深度学习-影像组学模型	训练集	226	0.850	0.799~0.901	82.0%(50/61)	75.8%(125/165)	77.4%(175/226)	0.578
	验证集	57	0.825	0.720~0.930	14/16	70.7%(29/41)	75.4%(43/57)	0.582
	测试集	121	0.771	0.673~0.869	76.2%(16/21)	68.0%(68/100)	69.4%(84/121)	0.442
深度学习-影像组学-临床模型	训练集	226	0.965	0.945~0.985	90.2%(55/61)	92.7%(153/165)	92.0%(208/226)	0.829
	验证集	57	0.917	0.848~0.986	15/16	78.0%(32/41)	82.5%(47/57)	0.718
	测试集	121	0.879	0.801~0.957	76.2%(16/21)	86.0%(86/100)	84.3%(102/121)	0.622

型、影像组学模型及影像组学-临床模型(Z 值:2.32~2.96, 均 $P<0.05$); 测试集 AUC 最高, 为 0.879(Z 值:2.00~3.09, 均 $P<0.05$)。DRC 模型准确性分别为 92.0% (208/226)、82.5% (47/57)、84.3% (102/121)。校准曲线结果(图 1)显示,DRC 模型的预测概率和实际结果一致性较好;DCA 结果(图 2)显示,训练集、验证集和测试集中风险阈概率分别为 0.56~0.84、0.33~0.73、0.38~0.78 时,DRC 模型净获益优于其余预测模型。

亚组分析中,DRC 模型预测 T1 组(130 例)、T2~3 组(48 例)、腺癌组(127 例)、鳞癌组(51 例)患者 OLM 的 AUC 分别为 0.903、0.842、0.925 和 0.773, 相应准确性分别为 76.2% (99/130)、83.3% (40/48)、77.2% (98/127)、76.5% (39/51)。其中,DRC 在腺癌组的 AUC 明显高于鳞癌组($Z=2.02, P=0.048$)。

4. DRC 模型与患者预后的分析。在验证集及测试集共 178 例患者中,将 116 例具有 1 年以上完整随访数据的患者纳入生存分析。图 3 展示了不同组别患者的生存曲线。其中,OLM 阳性组(30 例)患者的中位 PFS 与 OS 分别为 22.7 和 25.8 个月,低于

OLM 阴性组(86 例;29.3 和 31.2 个月; χ^2 值:21.87、16.68, 均 $P<0.001$)。在 DRC 评分风险分组(最佳阈值 0.210)中,高风险组(≥ 0.210 ;29 例)的中位 PFS 和 OS 分别为 23.1 和 26.7 个月,低于低风险组(<0.210 ;87 例;28.4 和 30.7 个月; χ^2 值:8.51、8.83, P 值:0.004、0.003)。上述结果表明,不同 OLM 状态及不同 DRC 评分风险分组患者间的生存情况不同。

讨 论

NSCLC 的淋巴结转移机制复杂,其发生与原发肿瘤的生物学特性密切相关^[10]。准确识别 OLM 对于患者预后评估、个体化治疗策略制定(尤其是系统性淋巴结清扫的决策)至关重要。本研究发现,年龄越小、中央型肺癌、血清 CEA 水平升高以及原发灶 SUV_{max} 与 SUV_{mean} 增高,均与 OLM 风险增加相关。其中,年轻患者风险较高可能与特定基因突变或肿瘤侵袭性强有关^[11];中央型肺癌因解剖位置邻近主要淋巴引流途径,OLM 风险亦显著高于周围型肺癌^[12]。高 SUV 通常反映肿瘤代谢活跃,提示更高的恶性潜能和转移倾向。然而仅依靠上述临床及代

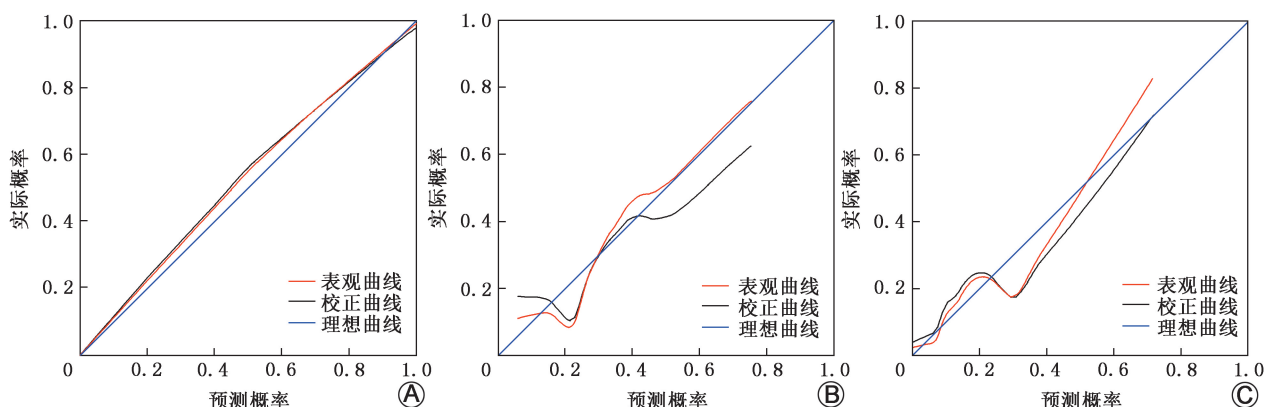


图 1 深度学习-影像组学-临床(DRC)模型在训练集(A;226 例)、验证集(B;57 例)、测试集(C;121 例)预测非小细胞肺癌隐匿性淋巴结转移的校准曲线 可见 DRC 模型的预测概率和实际结果具有较好的一致性

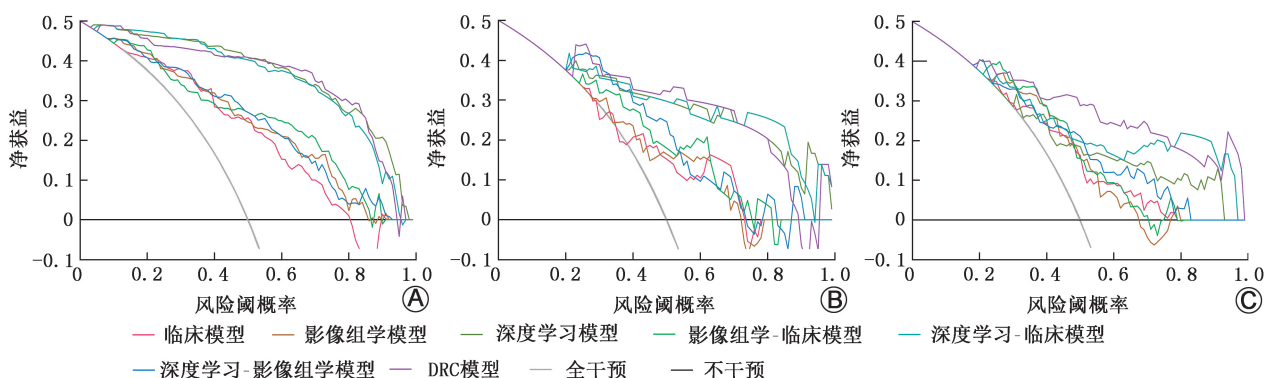


图 2 DRC 模型在训练集(A;226 例)、验证集(B;57 例)、测试集(C;121 例)预测非小细胞肺癌隐匿性淋巴结转移的决策曲线分析 当风险阈值分别为 0.56~0.84、0.33~0.73、0.38~0.78 时,DRC 模型净获益优于其余预测模型

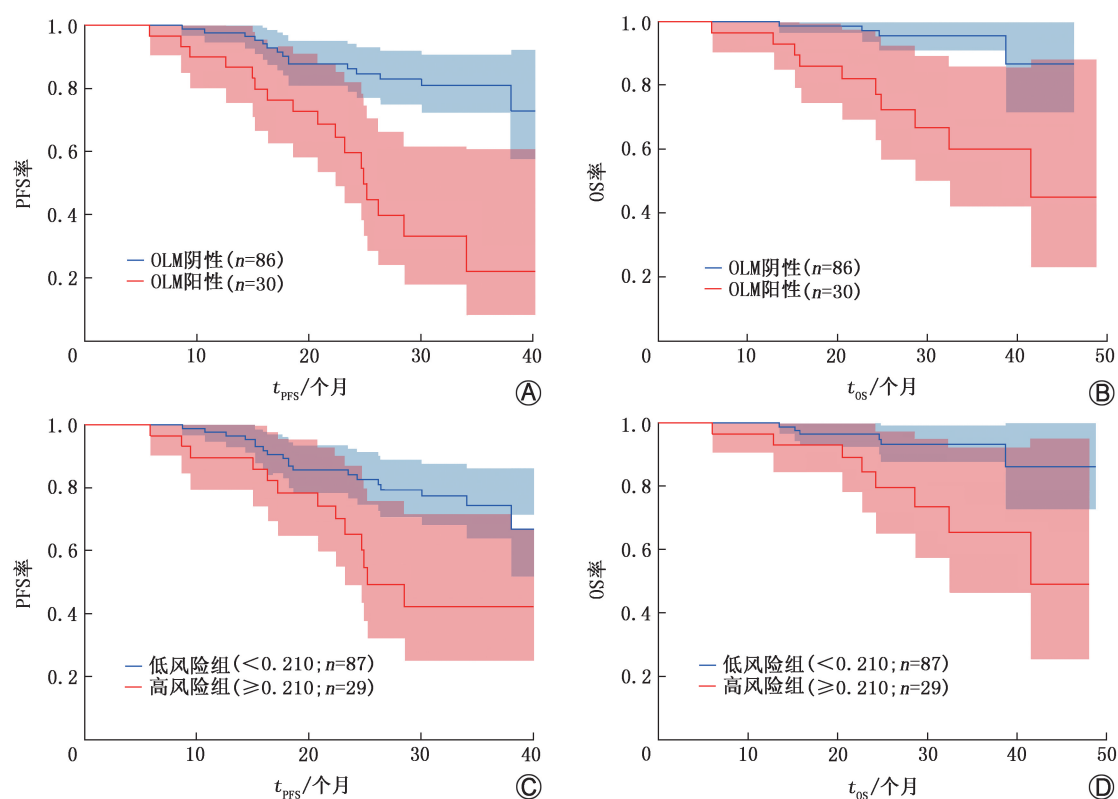


图3 不同隐性淋巴结转移(OLM)状态组(A,B)及不同风险分层[基于深度学习-影像学-临床(DRC)模型评分;C,D]非小细胞肺癌患者的Kaplan-Meier曲线 PFS为无进展生存,OS为总生存

谢参数预测 OLM,其诊断效能仍有限。值得注意的是,本研究单因素 logistic 回归分析结果显示, SUV_{max} ($OR=1.131$)及 SUV_{mean} ($OR=1.230$)越高 OLM 风险越高;而在多因素 logistic 回归分析中, SUV_{max} 与 SUV_{mean} 的 OR 值分别为 0.384 和 6.862, SUV_{max} 表现出相反的作用方向。导致这一现象的原因可能在于: SUV_{max} 与 SUV_{mean} 之间存在较强的共线性,在多因素分析中, SUV_{mean} 表现出更强的独立效应,当 SUV_{mean} 水平固定时, SUV_{max} 的相对变化呈现出负向关联。此外, SUV_{max} 表示靶区内葡萄糖摄取水平最高的单个像素 SUV,易受统计涨落及图像噪声的影响;而 SUV_{mean} 反映靶区范围内的平均 SUV,更能体现肿瘤的整体代谢负荷,因而表现出更强的预测效应。尽管 SUV_{max} 在多因素分析中呈现为保护因素,但这并不否定其在单因素分析中所提示的风险,而是提示在控制 SUV_{mean} 的影响后, SUV_{max} 的独立效应方向发生了反转,进一步表明在多因素模型中 SUV_{mean} 是更为关键的危险因素。

影像学在量化肿瘤异质性方面显著优于传统影像的肉眼评估^[13-14]。已有研究证实^[9-10],基于 PET/CT 影像学预测 LNM 具有较高的诊断效能 ($AUC:0.847\sim0.884$)。DL 的核心优势在于强大的特

征自动提取和学习能力^[15],能够从复杂的原始影像数据中挖掘出人眼难以识别的深层信息。既往研究显示,在预测早期 NSCLC 患者 LNM 方面,DL 模型优于传统影像学模型 ($AUC:0.834$ 与 0.761),而两者结合可进一步提升预测效能 ($AUC=0.861$)^[16]。

本研究构建了一种融合 DL 特征、影像学特征及临床特征的多模态 DRC 模型。该模型在验证集与测试集中均表现出最优 AUC。亚组分析显示,DRC 模型对肺腺癌 OLM 的预测效能 ($AUC=0.925$) 明显优于肺鳞癌 ($AUC=0.773$),可能与两者在影像学表现及生物学行为方面的差异有关^[17]。本研究显示,存在 OLM 的患者的 PFS 和 OS 均短于无 OLM 患者。基于 DRC 模型的风险分层显示,高风险患者的生存结局明显更差。这提示,DRC 模型能有效区分患者预后,有助于指导个体化手术方案的制定。既可避免对低风险患者进行不必要的过度清扫,有效减少手术创伤及相关并发症;还能准确识别 OLM 高风险患者,并实施充分清扫,这对于改善患者长期预后及优化临床管理策略具有重要意义。

本研究存在局限性:(1)回顾性研究设计,样本量有限,可能存在选择偏倚,未来需通过大规模、前瞻性、多中心研究进一步验证;(2)仅纳入了肺腺癌

与鳞癌患者,结论对其他类型(如大细胞癌、小细胞癌等)的适用性尚不明确;(3)淋巴结清扫范围可能存在局限,无法完全覆盖所有潜在转移区域;(4)随访时间相对较短,可能不足以充分评估患者的远期复发和转移风险。未来研究需纳入更广泛、更具代表性的患者群体,并尝试融合多组学数据,以构建更为全面和精准的淋巴结转移预测与预后评估体系。

综上,本研究证实了融合 DL 与影像组学方法,在深度挖掘 PET/CT 图像信息、精准预测 NSCLC 患者 OLM 方面具有重要潜力。构建的 DRC 模型诊断效能较好,可为临床个体化治疗决策提供客观依据;在避免对无转移患者可能的过度治疗同时,确保转移患者获得及时、充分的干预,最终提升治疗效果、改善患者生存质量。

利益冲突 所有作者声明无利益冲突

作者贡献声明 于军:研究设计与实施、数据采集与分析、统计分析、论文撰写;何超、李洋、黄磊:数据采集与分析;陈聪、任春玲、任东栋:研究设计与指导、论文修改

参 考 文 献

- [1] Siegel RL, Kratzer TB, Giaquinto AN, et al. Cancer statistics, 2025[J]. CA Cancer J Clin, 2025, 75(1): 10-45. DOI: 10.3322/caac.21871.
- [2] Beyaz F, Verhoeven RLJ, Schuurbiens OCJ, et al. Occult lymph node metastases in clinical N0/N1 NSCLC; A single center in-depth analysis[J]. Lung Cancer, 2020, 150: 186-194. DOI: 10.1016/j.lungcan.2020.10.022.
- [3] Dai X, Zhao B, Zang J, et al. Diagnostic performance of radiomics and deep learning to identify benign and malignant soft tissue tumors: a systematic review and meta-analysis[J]. Acad Radiol, 2024, 31(10): 3956-3967. DOI: 10.1016/j.acra.2024.03.033.
- [4] Li Y, Deng J, Ma X, et al. Diagnostic accuracy of CT and PET/CT radiomics in predicting lymph node metastasis in non-small cell lung cancer[J]. Eur Radiol, 2025, 35(4): 1966-1979. DOI: 10.1007/s00330-024-11036-4.
- [5] Li B, Su J, Liu K, et al. Deep learning radiomics model based on PET/CT predicts PD-L1 expression in non-small cell lung cancer[J]. Eur J Radiol Open, 2024, 12: 100549. DOI: 10.1016/j.ejro.2024.100549.
- [6] Yuan P, Huang ZH, Yang YH, et al. A ¹⁸F-FDG PET/CT-based deep learning-radiomics-clinical model for prediction of cervical lymph node metastasis in esophageal squamous cell carcinoma[J]. Cancer Imaging, 2024, 24(1): 153. DOI: 10.1186/s40644-024-00799-0.
- [7] 张乐乐, 陆路, 葛昭, 等. 影像组学与深度学习模型预测甲状腺乳头状癌¹³¹I 治疗疗效的可行性分析[J]. 中华核医学与分子影像杂志, 2025, 45(9): 543-548. DOI: 10.3760/cma.j.cn321828-20240904-00312.
- [8] Zhang LL, Lu L, Ge Z, et al. Feasibility analysis of radiomics and deep learning models in predicting the efficacy of ¹³¹I therapy for papillary thyroid cancer[J]. Chin J Nucl Med Mol Imaging, 2025, 45(9): 543-548. DOI: 10.3760/cma.j.cn321828-20240904-00312.
- [9] Deng J, Zhong Y, Wang T, et al. Lung cancer with PET/CT-defined occult nodal metastasis yields favourable prognosis and benefits from adjuvant therapy: a multicentre study[J]. Eur J Nucl Med Mol Imaging, 2022, 49(7): 2414-2424. DOI: 10.1007/s00259-022-05690-3.
- [10] Qiao J, Zhang X, Du M, et al. ¹⁸F-FDG PET/CT radiomics nomogram for predicting occult lymph node metastasis of non-small cell lung cancer[J]. Front Oncol, 2022, 12: 974934. DOI: 10.3389/fonc.2022.974934.
- [11] Huang Y, Jiang X, Xu H, et al. Preoperative prediction of mediastinal lymph node metastasis in non-small cell lung cancer based on ¹⁸F-FDG PET/CT radiomics[J]. Clin Radiol, 2023, 78(1): 8-17. DOI: 10.1016/j.crad.2022.08.140.
- [12] Han CB, Ma JT, Li F, et al. EGFR and KRAS mutations and altered c-Met gene copy numbers in primary non-small cell lung cancer and associated stage N2 lymph node-metastasis[J]. Cancer Lett, 2012, 314(1): 63-72. DOI: 10.1016/j.canlet.2011.09.012.
- [13] Miao H, Shaolei L, Nan L, et al. Occult mediastinal lymph node metastasis in FDG-PET/CT node-negative lung adenocarcinoma patients: risk factors and histopathological study[J]. Thorac Cancer, 2019, 10(6): 1453-1460. DOI: 10.1111/1759-7714.13093.
- [14] 冯源, 兰晓莉. 影像组学介绍[J]. 中华核医学与分子影像杂志, 2023, 43(1): 55-60. DOI: 10.3760/cma.j.cn321828-20211130-00427.
- [15] Feng Y, Lan XL. Introduction to radiomics[J]. Chin J Nucl Med Mol Imaging, 2023, 43(1): 55-60. DOI: 10.3760/cma.j.cn321828-20211130-00427.
- [16] 牟玮, 田捷. PET/CT、SPECT/CT 影像组学: 沟通宏观影像和微观分子的桥梁[J]. 中华核医学与分子影像杂志, 2024, 44(2): 65-67. DOI: 10.3760/cma.j.cn321828-20231212-00134.
- [17] Mu W, Tian J. Radiomics in PET/CT and SPECT/CT: the bridge between macroscopic images and microscopic molecules[J]. Chin J Nucl Med Mol Imaging, 2024, 44(2): 65-67. DOI: 10.3760/cma.j.cn321828-20231212-00134.
- [18] Zheng YM, Che JY, Yuan MG, et al. A CT-based deep learning radiomics nomogram to predict histological grades of head and neck squamous cell carcinoma[J]. Acad Radiol, 2023, 30(8): 1591-1599. DOI: 10.1016/j.acra.2022.11.007.
- [19] Zhang H, Liao M, Guo Q, et al. Predicting N2 lymph node metastasis in presurgical stage I - II non-small cell lung cancer using multiview radiomics and deep learning method[J]. Med Phys, 2023, 50(4): 2049-2060. DOI: 10.1002/mp.16177.
- [20] Liang B, Tong C, Nong J, et al. Histological subtype classification of non-small cell lung cancer with radiomics and 3D convolutional neural networks[J]. J Imaging Inform Med, 2024, 37(6): 2895-2909. DOI: 10.1007/s10278-024-01152-4.

(收稿日期: 2025-07-21)