



· 论 著 ·

基于胸部增强CT影像组学模型用于胸腺瘤分类的研究

蒋佻宴^{1, 2}, 贾田颖², 张 琴²

1. 江苏大学医学院, 江苏 镇江, 212013;
2. 上海交通大学医学院附属胸科医院放疗科, 上海 200030

[摘要] 背景与目的: 胸腺瘤术前分类对治疗决策很有帮助, 但获取存在一定困难。本研究旨在建立基于增强对比计算机体层成像 (contrast-enhanced computed tomography, CECT) 影像组学的胸腺瘤风险分类训练模型, 并验证其性能、可靠性和泛化能力。方法: 本回顾性队列研究分析了在上海交通大学医学院附属胸科医院2008年1月—2017年12月接受胸腺瘤切除手术的患者 (Masaoka-Koga I~III期) 的临床资料。将患者随机分层为训练组 (80%) 和测试组 (20%), 将术后病理学检查结果作为金标准。低风险组织学类型包括A、AB和B1。高风险组织学类型包括B2和B3。手动分割术前CECT图像相关肿瘤区域并提取影像组学特征。使用最小绝对收敛和选择算子 (least absolute shrinkage and selection operator, LASSO) 回归进行特征选择, 并将临床特征添加到联合模型中。模型性能指标包括受试者工作特征 (receiver operating characteristic, ROC) 曲线的曲线下面积 (area under curve, AUC)、灵敏度和特异度。结果: 共纳入478例患者 (平均年龄51.3±12.3岁, 男性占48.1%)。临床模型、基于CECT的影像组学模型和在测试集上使用临床和CT特征的模型AUC分别为0.666、0.831和0.850。性能最佳的模型的灵敏度为0.829, 特异度为0.764。结论: 基于术前CECT的影像组学模型在胸腺瘤风险分类中表现良好。

[关键词] 纵隔腔; 胸腺瘤; 计算机体层成像; 机器学习

中图分类号: R734.2 文献标志码: A DOI: 10.19401/j.cnki.1007-3639.2024.06.006

Contrast-enhanced computed tomography-based radiomics models for the risk categorization of thymoma JIANG Tiaoyan^{1, 2}, JIA Tianying², ZHANG Qin² (1. Jiangsu University School of Medicine, Zhenjiang 212013, Jiangsu Province, China; 2. Department of Radiation Oncology, Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China)

Correspondence to: ZHANG Qin E-mail: zhangq0616@163.com

[Abstract] **Background and purpose:** Preoperative risk categorization of thymoma is useful for treatment decisions but remains challenging. This study focused on training radiomics models using contrast-enhanced computed tomography (CECT) images for thymoma risk categorization and validating the model's performance, reliability and generalizability in a relatively large cohort. **Methods:** This retrospective cohort study analyzed the clinical data of thymoma patients (Masaoka Koga I-III) who underwent thymectomy surgery at the Affiliated Chest Hospital of Shanghai Jiao Tong University School of Medicine from January 2008 to December 2017. The cohort was divided into a training group (80%) and a test group (20%) using stratified random selection. The gold standard for histologic types was based on surgically resected specimens. Low-risk histologic types included A, AB and B1. High-risk histologic types included B2 and B3. Radiomics features were extracted from manually segmented regions of interest on preoperative CECT images. Interobserver correlation and least absolute shrinkage and selection operator (LASSO) regression were used for feature selection. Model performance metrics included area under the curve (AUC) of receiver operating characteristic (ROC) curve, sensitivity and specificity. Clinical characteristics were added to the combined model. **Results:** A total of 478 patients (mean age 51.3±12.3 years, 48.1% was male) were included. The AUC of the clinical model, the CECT-based model, and the model using

第一作者: 蒋佻宴 (ORCID: 0009-0007-2470-5960), 住院医师。

通信作者: 张 琴 (ORCID: 0000-0003-2786-5277), 主任医师, 上海交通大学医学院附属胸科医院放疗科副主任, E-mail: zhangq0616@163.com。

both clinical and CECT features on the test set were 0.666, 0.831 and 0.850, respectively. The best performing model had a sensitivity of 0.829 and a specificity of 0.764. **Conclusion:** CECT-based radiomics models showed good performance in risk categorization of thymomas.

[**Keywords**] Mediastinal; Thymoma; Computed tomography; Machine learning

胸腺瘤起源于胸腺上皮细胞, 是前纵隔最常见的原发性恶性肿瘤。胸腺瘤有多种组织学类型, 侵袭性和患者预后差异较大^[1]。2015年世界卫生组织 (World Health Organization, WHO) 分类系统将胸腺瘤分为低风险 (类型A、AB和B1) 和高风险 (类型B2和B3) 两大类^[2]。高风险胸腺瘤侵袭性、转移风险高, 预后相对较差^[3-5]。

准确的组织学分型和疾病分期是影响治疗决策的关键因素。既往研究^[6-8]发现Ⅲ期胸腺瘤患者可能从新辅助治疗中获益。根据欧洲临床肿瘤学会对胸腺上皮肿瘤的临床实践指南, Ⅲ期胸腺瘤更常见于B2和B3型 (27%~28%) 而不是A、AB和B1型 (6%~9%)^[4]。因此, 患者术前评估组织学类型有利于治疗决策。胸腺瘤具有异质性, 常见混合不同的组织学类型 (例如AB和B1)。术前穿刺活检明确病理学检查在一些临床实践中存在操作风险或诊断困难。术前医学影像学检查具有无创且能提供肿瘤完整图像特征的优势, 并有可能提供胸腺肿瘤特异性的影像特征。Jeong等^[9]研究发现, 计算机体层成像 (computed tomography, CT) 图像的定性特征与胸腺上皮肿瘤 (thymic epithelial tumor, TET) 的风险评估及患者预后具有一定相关性; 某些CT特征在高风险胸腺瘤和胸腺癌 (thymic carcinomas, TC) 中比低风险胸腺瘤更常见, 并可能与复发及转移有关。

影像组学从医学图像中提取量化特征, 并构建机器学习模型来预测各种恶性肿瘤的组织学类型、分期、淋巴结转移、预后等^[10]。既往也有相关研究尝试用CT或磁共振成像 (magnetic resonance imaging, MRI) 的影像组学等方法来预测胸腺瘤的风险分类, 然而, 大多数研究基于小于300例患者的队列, 因此需要在更大的队列上验证使用影像组学来预测胸腺瘤风险分类

的可靠性和泛化能力。增强对比计算机体层成像 (contrast-enhanced computed tomography, CECT) 可提供良好的组织对比度和空间分辨率。本研究旨在建立影像组学分类模型, 利用CECT图像区分胸腺瘤的风险组, 并验证模型在相对较大的回顾性队列中的可靠性和泛化能力。

1 资料和方法

1.1 入排组标准

本回顾性队列研究分析了在上海交通大学医学院附属胸科医院2008年1月—2017年12月接受胸腺瘤切除手术的患者临床资料。收集患者相关临床信息: 年龄、性别、Masaoka-Koga分期、组织学类型、最大肿瘤直径和合并症等。纳入标准: ① 有术前CECT影像检查; ② 接受手术切除; ③ 术后病理学检查证实为胸腺瘤; ④ Masaoka-Koga分期 I~Ⅲ。排除标准: ① CT图像质量差; ② CT层厚大于5 mm; ③ 临床或病理学信息不完整。

1.2 临床信息

收集上海交通大学医学院附属胸科医院收治病例的病理学检查结果, 根据2015年WHO分类定义分为5种类型中的一种或组合 (A、AB、B1、B2和B3)。高风险类别包括具有B2或B3组病理学诊断结果的病例。如有混合组织学类型的胸腺瘤根据更具侵袭性的组分进行分类 (例如, 具有B1和B2组分的胸腺瘤被包括在高风险组中)。收集入组患者的临床特征: 年龄、性别、最大肿瘤直径和合并症。

1.3 CT扫描参数

本研究中主要采用CT机型: Brilliance 64 (荷兰飞利浦公司, $n=200$)、iCT 256 (荷兰飞利浦公司, $n=95$)、Ingenuity CT (荷兰飞利浦公司, $n=34$)、Discovery CT750 HD (通用电气

医疗系统, $n=50$) 和uCT S-160 (上海联影医疗有限公司, $n=99$)。管电压设定为120 kVp, 管电流150 mAs或自动管电流调控。所有图像在注射2 mL/kg体重的碘帕米醇注射液 (Iopamiro[®], 300 mg I/mL; 上海百浩信制药有限公司) 后60秒扫描, 扫描后重建矩阵 512×512 , 扫描层厚为5 mm, 准直器宽度 64×0.625 或 128×0.625 等, 螺距0.915或0.516等, 所有图像都采用标准算法重建。

1.4 感兴趣区域 (region of interest, ROI) 的分割

ROI的勾画基于Pinnacle[®]治疗规划系统 (treatment planning system, TPS) 在纵隔窗口 (宽度为400 HU, 30 HU) 时进行。肿瘤ROI被定义为总肿瘤体积 (the gross tumor volume, GTV)。所有ROI均由1名具有3年经验的放疗科或放射科医师手动分割, 并由另1名具有20年经验的影像诊断医师进行审查。随机选择了60个案例让两名放射治疗师分别勾画ROI, 行一致性观察测试。

1.5 CT图像特征提取和选择

影像学特征提取是使用开源软件3D Slicer^[11]及其SlicerRadiomics扩展完成的。在进行特征提取之前, CT图像被重新采样为 $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ 的分辨率。特征类别包括一阶统计特征、形状和纹理, 其中包括各种矩阵, 如灰度依赖矩阵、灰度共现矩阵、灰度运行长度矩阵、灰度大小区域矩阵和相邻灰度差异矩阵。此外, 对原始图像应用了8个小波滤波器, 柱宽度设置为25, 执行对称GLCM检查, 没有应用高斯拉普拉斯滤波器。

1.6 统计学处理

使用卡方检验、 t 检验和Wilcoxon秩和检验分析低风险和高风险组之间的差异。 $P < 0.05$ 为差异有统计学意义。使用R语言的Rstatix包 (版本0.7.0) 进行统计分析。年龄和肿瘤大小被视为定量变量。对于正态分布且方差相等的变量使用 t 检验。对于其他定量变量使用Wilcoxon秩和检验。性别和合并症的存在是两个可能值 (0和1) 的定性变量。卡方检验比较这两个变量之

间的组别, 研究队列使用分层随机选择分为训练集 (80%) 和测试集 (20%)。分层因素为Masaoka-Koga分期和组织学类型。

通过随机选择的60个扫描图像评估影像学特征的观察者一致性。组间相关系数 (interclass correlation coefficient, ICC) 大于0.75被视为一致性良好, 而大于0.90被视为一致性好。特征选择是在训练集上使用最小绝对值收敛和选择算子 (least absolute shrinkage and selection operator, LASSO) 与交叉验证 (cross validation, CV) 进行的。影像学特征被缩放为均值为0, 方差为1, 进行LASSO回归 (glmnet包, 版本4.1-4)。通过在训练集上使用逻辑回归 (logistic regression, LogReg) 和五折交叉验证进行50次重复来评估每个特征组合的区分能力, 其中平均准确性和过度拟合的鲁棒性作为主要性能评估标准。

模型的训练和性能验证是使用MLR包 (版本2.19.0) 完成的。采用的分类算法包括线性判别分析 (linear discriminant analysis, LDA)、二次判别分析 (quadratic discriminant analysis, QDA)、LogReg、支持向量机 (support vector machine, SVM) 和随机森林 (random forest, RF)。使用随机搜索和五折交叉验证调整了超参数。在测试集上评估了分类模型的性能, 评估指标包括95% CI的受试者工作特征 (receiver operating characteristic, ROC) 曲线的曲线下面积 (area under curve, AUC)、最佳诊断决策点处的灵敏度和特异度。使用“pROC”包 (版本1.18.0) 生成ROC曲线, 并使用DeLong的检验比较了ROC曲线。详细工作流程见图1。

2 结 果

2.1 患者临床资料分析

从2008年1月—2017年12月, 筛选了共1 110例接受手术治疗的胸腺瘤患者。这些患者均接受了R0切除。术后分期为Masaoka-Koga I~III期。筛选出3例伴有胸腺癌的胸腺瘤患者, 43例伴有其他癌症的患者, 筛选出586例缺失CT图像、无增强CT或CT图像质量差的患者。最终478

例胸腺瘤患者符合条件纳入本研究。其中273例 (57.1%) 属于低风险组, 205例属于高风险组。高风险组与低风险组在以下方面存在显著差异: ① 平均年龄较年轻, 标准差较大; ② 更多的较晚的Masaoka-Koga分期; ③ 更多的合并症, 包括重症肌无力、再生障碍性贫血、纯红细胞再生障碍性贫血、干燥综合征、全身性红斑性狼

疮、强直性脊柱炎和肾病综合征。两个风险组在性别和最大肿瘤直径上差异无统计学意义 ($P > 0.05$)。患者的临床特征见表1。在进行特征选择和模型训练之前, 根据分层因素Masaoka-Koga分期和组织学类型随机选择, 将研究队列分为训练集 ($n=382$) 和测试集 ($n=96$)。这两组的临床特征差异无统计学意义 ($P > 0.05$, 表2)。

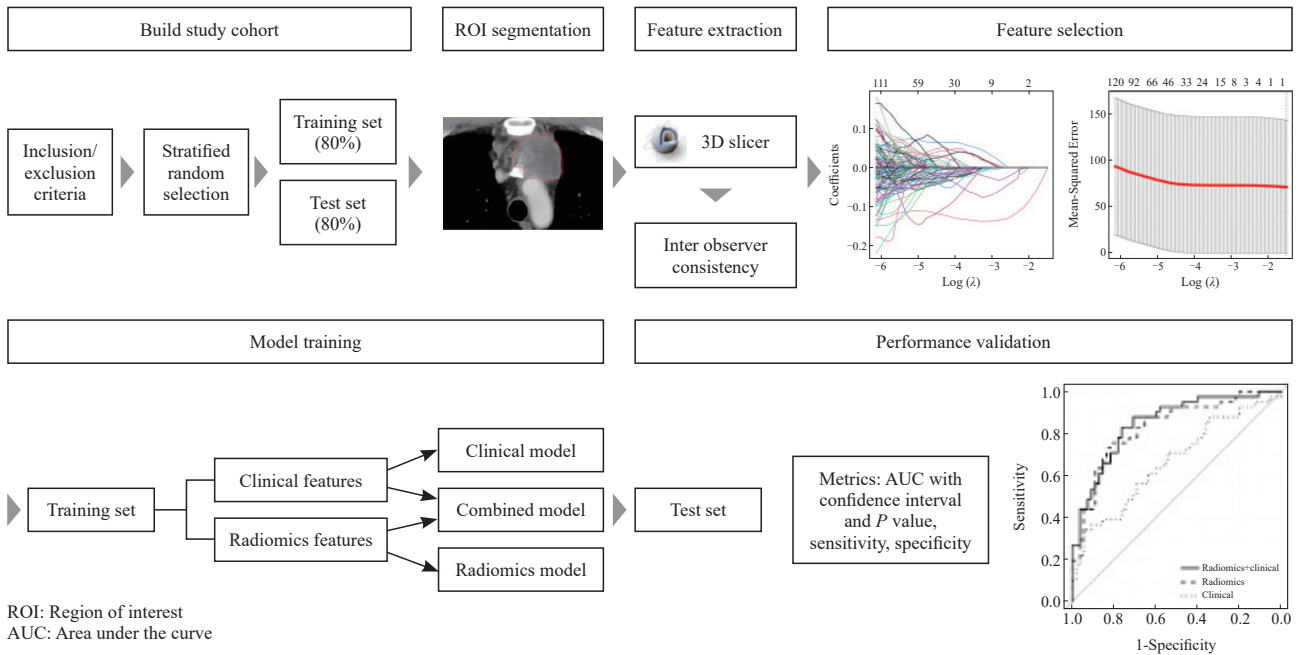


图1 研究队列、影像学特征提取和选择、分类模型训练和性能验证的工作流程

Fig. 1 Workflow of study cohort, radiomics feature extraction and selection, classification model training and performance validation

表1 患者的临床特征

Tab. 1 Clinical characteristics of patients

Characteristics	Low-risk thymomas ($n=273$)	High-risk thymomas ($n=205$)	P value
Gender n (%)			
Male	122 (44.69)	107 (52.2)	0.10
Female	151 (55.31)	98 (47.8)	
Age/year			
$\bar{x} \pm s$	54.5 \pm 10.8	50.9 \pm 13.7	0.01
Maximal diameter D/cm			
$\bar{x} \pm s$	6.18 \pm 2.97	6.07 \pm 2.57	0.98
Comorbidity n (%)			
Present	9 (3.3)	36 (17.56)	<0.001
Absent	264 (96.70)	169 (82.44)	
Masaoka-Koga stage n (%)			
I	110 (40.29)	28 (13.66)	<0.001
II	148 (54.21)	117 (57.07)	
III	15 (5.49)	60 (29.27)	

2.2 影像学特征选择

从每个肿瘤ROI中提取851个影像学特征。717个特征显示了良好的观察者一致性 ($ICC \geq 0.75$)，581个特征显示了非常好的一致性 ($ICC \geq 0.9$)。当 λ 值从0.01变化到0.1时，所选择的特征数量从44减少到3。在CV中，LogReg模型的平均准确率范围为0.737~0.774，并在4个和14个特征处显示了两个局部极大值（图2）。分别使用4个和14个特征训练了分类模型。最终的组合选择了4个特征，充分考虑其对过度拟合的鲁棒性。

2.3 模型性能

模型性能在测试集 ($n=96$) 上进行评估。表现最佳的临床模型采用了QDA算法，在训练集和测试集上分别实现了AUC值为0.652 (95% CI: 0.596 ~ 0.707) 和0.666 (95% CI: 0.550 ~ 0.780)，准确率为68.8%。在测试集上最佳诊断决策点的灵敏度和特异度分别为0.341和0.945。表现最佳的14个特征的影像学模型采用了LDA算法，在训练集和测试集上分别实

现了AUC值为0.849 (95% CI: 0.810 ~ 0.889) 和0.800 (95% CI: 0.711 ~ 0.890)。表现最佳的4个特征的影像学模型采用了QDA算法，在训练集和测试集上分别实现了AUC值为0.831 (95% CI: 0.790 ~ 0.872) 和0.831 (95% CI: 0.747 ~ 0.914)。与14个特征的LDA模型相比，4个特征的QDA模型对过度拟合更加稳健，并具有更好的灵敏度、特异度和准确率。选用的4个特征是：原始图像的形状特征“球度”、第一阶特征“中位数”（来自Wavelet-LHL和Wavelet-LLL）、灰度级运行长度矩阵特征（来自Wavelet-HLL）。采用4个临床特征和4个影像学特征的logistic回归模型在训练集和测试集上AUC值分别为0.830 (95% CI: 0.788 ~ 0.871) 和0.850 (95% CI: 0.773 ~ 0.927)。对3个模型比较显示，基于CECT的QDA模型明显优于临床模型 ($P=0.019$)。临床信息的添加略微提高了模型性能，但差异无统计学意义 ($P=0.3466$ ，图3，表3)。

表2 训练集和试验集临床特征比较

Tab. 2 Comparison of clinical characteristics between training and test sets

Characteristics	Training set ($n=382$)	Test set ($n=96$)	<i>P</i> value
Gender			
Male	218 (57.07)	55 (57.29)	0.97
Female	164 (42.93)	41 (42.71)	
Age/year			
$\bar{x} \pm s$	52.9 \pm 12.3	53.1 \pm 12.0	0.66
Maximal diameter <i>D</i> /cm			
$\bar{x} \pm s$	6.03 \pm 2.74	6.36 \pm 3.09	0.38
Concomitant disease <i>n</i> (%)			
Present	35 (9.16)	10 (10.42)	0.71
Absent	347 (90.84)	86 (89.58)	
Masaoka-Koga stage <i>n</i> (%)			
I	111 (29.06)	27 (28.13)	0.91
II	211 (55.24)	54 (56.25)	
III	60 (15.71)	15 (15.63)	
Risk category <i>n</i> (%)			
Low risk	218 (57.07)	55 (57.29)	0.97
High risk	164 (42.93)	41 (42.71)	

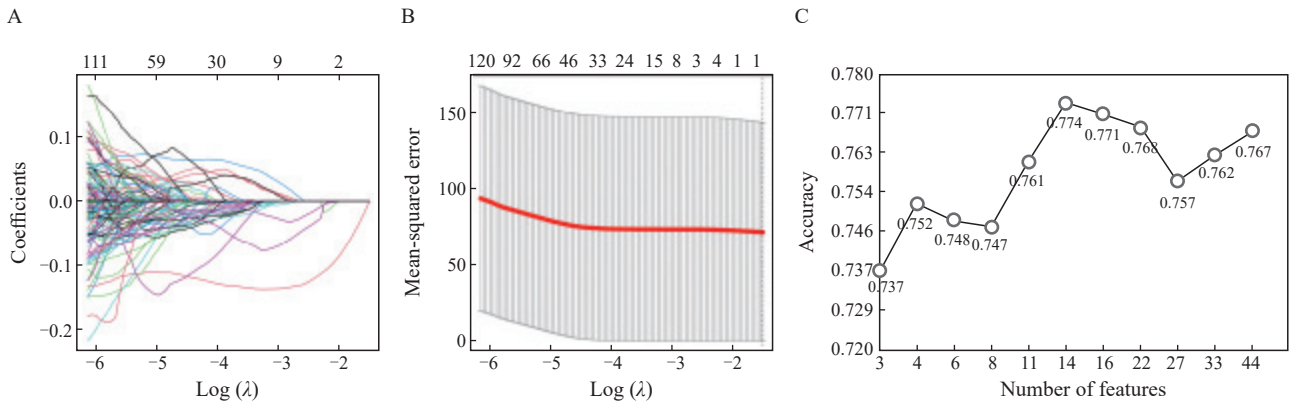


图2 使用LASSO回归算法进行特征选择以及基线 logistic 回归模型的性能

Fig. 2 Feature selection using the LASSO regression algorithm and the performance of the baseline logistic regression model

A: The LASSO coefficients of the 581 features; B: Mean square error with respect to $\log(\lambda)$. The average binominal deviance values for each model at a given λ were indicated by the dashed red curve; C: Mean accuracy of the baseline logistic model at a given number of features. Two regional maxima were found at 4 and 14 features.

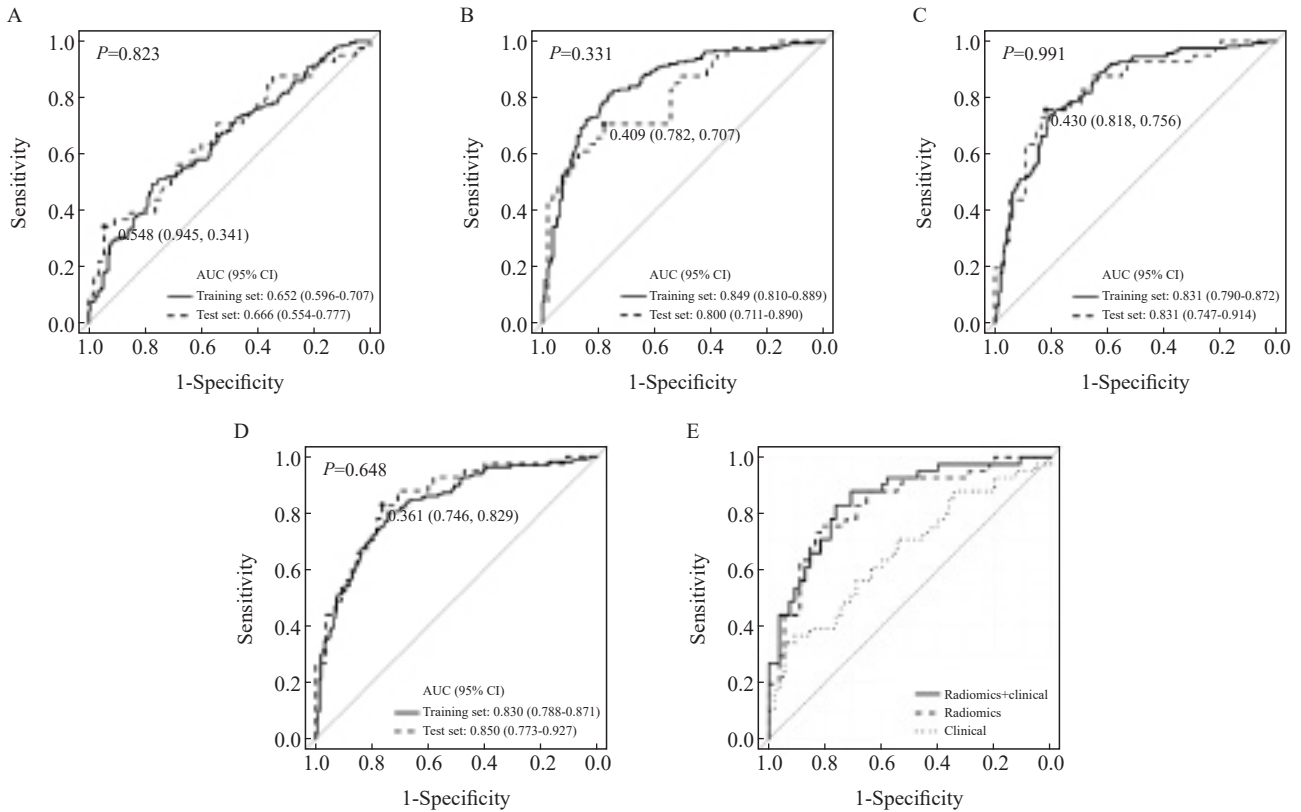


图3 训练集和测试集上风险组分类模型的ROC曲线

Fig. 3 ROC curves of risk group classification models on the training and test sets

A: ROC curve of the QDA model based on clinical features with an area under the ROC curve (AUC) of 0.652 (95% CI: 0.596-0.707) and 0.666 (95% CI: 0.554-0.777) on the training and test sets. The sensitivity and specificity at the best decision point were 0.341 and 0.945 on the test set. B: LDA model using fourteen radiomics features with an AUC of 0.849 on the training set and 0.800 on the test set. Sensitivity and specificity were 0.707 and 0.782 on the test set. C: QDA model using four radiomics features with an AUC of 0.831 on both the training and test sets. Sensitivity and specificity were 0.756 and 0.818, respectively, on the test set. D: Logistic regression model based on clinical and radiomic features with an AUC of 0.830 (95% CI: 0.788-0.871) and 0.850 (95% CI: 0.773-0.927) on the training and test set, respectively. Sensitivity and specificity were 0.829 and 0.764 on the test set. E: Comparison of the ROC curves of the three models.

表3 在测试集上测量模型性能指标

Tab. 3 Model performance measured on the test set

Data	AUC	95% CI	P value	Sensitivity	Specificity	Accuracy
Clinical	0.666	0.554-0.777	0.0028	0.341	0.945	0.688
CECT-1*	0.800	0.711-0.890	<0.001	0.707	0.782	0.750
CECT-2**	0.831	0.747-0.914	<0.001	0.756	0.818	0.792
Combined***	0.850	0.773-0.927	<0.001	0.829	0.764	0.792

*: Model constructed using 14 radiomics features; **: Model constructed using 4 radiomics features; ***: Model constructed using 4 radiomics features and clinical features.

3 讨 论

纵隔肿瘤特别是胸腺瘤常见于年轻患者，有症状就诊时已经是中晚期，手术是主要的根治性方法，术前辅助治疗可提高患者完全切除的机会，准确的组织学分型和疾病分期是影响治疗决策的关键因素。胸腺瘤的病理学类型复杂，不同特性对应不同的风险等级，与患者的复发、转移和预后密切相关。不同类型和分期的胸腺瘤并不一定均适合活组织病理学检查；其次胸腺瘤分型种类多且有部分为混合病理学类型，通过活组织病理学检查获得的标本不能很好地显示肿瘤内和肿瘤间的异质性，可提供的肿瘤病理学信息有限；活组织病理学检查本身是有创性检查，存在操作风险、取材标本阴性结果或者部分患者无法耐受等情况。此外，重复的活检在整个治疗过程中虽然可能有利于监测肿瘤的改变或重新进行病理学诊断等，但实际应用时易受患者依从性和实际操作中遇到的问题影响。

近年来，人工智能深度学习等技术试图解决这些问题。影像组学分析作为一种非侵入性的方法来补充活组织病理学检查，已被应用于肿瘤分子病理学诊断中，如判断是否存在特定基因突变等方面。成像分析的基本假设是，先进的成像技术可以捕获以宏观图像为基础的特征表达的基因组和蛋白质组学模式，肿瘤内蛋白表达模式的差异已被证明与影像学表现相关^[12-14]。

本研究基于胸腺瘤患者术前的CECT的影像学信息，构建了用于胸腺瘤风险分组分类的分类模型，并从回顾性队列中划分的训练集和测试集

上进行了测试，共包括478例患者。基于CECT的影像学模型显示出与临床模型相比更好的性能。构建性能最佳的4个影像学特征模型在训练集和测试集上AUC值为0.831，表明其具有抗过度拟合的鲁棒性。临床特征（年龄、性别、肿瘤直径和合并症的存在）对于组合模型的价值很小（AUC为0.850 vs 0.831， $P=0.3466$ ）。

在关于TET风险分类的影像学或深度学习研究领域目前研究报道颇多^[14-17]。许多基于CECT的影像学模型的研究使用的队列患者数少于300例^[18-19]。许多研究^[20-25]报道的样本量小于100例。这些研究的AUC值范围从0.621到0.90。并非所有研究都使用了独立的测试队列。基于这些小样本构建的模型可能相当不稳定且不可靠，结果可能无法复制。Feng等^[26]入组了509例患者的队列，这可能是目前样本量最大的研究之一。该队列包括238例低风险胸腺瘤、232例高风险胸腺瘤和39例胸腺癌。该研究表明，基于无增强CT（non-enhanced, NECT）和临床特征的SVM可以区分不同的TET。该模型在测试集上实现了0.844的AUC值，但他们的模型基于NECT，其缺点是组织对比度较低^[9]。

本研究的模型基于CECT表现出较好的性能。选择CECT而不是NECT的主要原因是其对于区分肿瘤与正常组织（如血管）具有良好的组织对比度。相对较大的研究队列包括更多种类的情况，使得模型更加可靠且可复制。这为CECT图像对于分类胸腺瘤风险的能力提供了额外的支持。

目前为止，影像组学分析生物信息还不能取代活组织病理学检查和液体活检。与活组织病

理学检查相比, 图像分析最有希望的优势在于通过图像获取的生物信息可以描述整个肿瘤的基因型和表型, 甚至将生物信息投射到图像的每个像素上, 以反映肿瘤内的异质性。由于肿瘤间的异质性, 每种病变的分子变化不同。图像分析可以弥补这一不足, 辅助指导更精准的个体化综合治疗。因此, 值得发展影像组学辅助诊断胸腺瘤的病理学特征, 特别是在分期较晚、风险较高的胸腺瘤的诊治中, 可能有助于更加精准的临床治疗。

本研究有其局限性。首先, 我们的研究使用了来自单中心的回顾性队列, 如果能使用外部数据作为独立的测试集可能会使模型的性能更加可靠。其次, CECT图像是在十多年的时间内由5台来自4个不同制造商的CT扫描仪获取。这些患者的图像质量、CT采集参数、CT品牌和重建算法并不一致, 这些变量的影响未纳入分析。再次, 我们使用的CT图像的体层厚度为5 mm, 并提取了具有1 mm×1 mm×1 mm空间分辨率的3D影像。插值过程引入了4个时间估计的数据。此外, CT图像的定性特征, 如分叶轮廓、纵隔脂肪浸润、大血管侵犯、不规则轮廓和椭圆形状, 在本研究中未进行讨论, 尽管这些特征已被证明与TET的组织学类型及患者预后密切相关。此外, 有研究^[27-28]报道了多参数影像组学及其他诊断方式的研究结果, 值得我们进一步对比分析。

本研究结果显示, 基于CECT的影像学特征可能成为一种可靠的辅助工具, 在临床诊断困难时, 可帮助区分高风险和低风险胸腺瘤。CT影像学分析提供了一种无创的诊断方式, 有利于在术前辅助临床决策。未来的工作可以集中在: ① 通过引入NECT和MRI图像来提高分类的准确性; ② 预测患者预后, 包括无进展生存、复发转移和总生存; ③ 进一步深入研究来辅助区分胸腺瘤与胸腺癌及良性囊肿。

利益冲突声明: 所有作者均声明无利益冲突。

作者贡献声明:

蒋兆宴和贾田颖进行文献调研与整理、收集

数据、采集、清洗与分析数据、绘制图谱、起草论文、修订论文、审核论文。张琴提出本研究的基本框架、设计研究思路、设计研究方案、研究方案可行性调查分析等, 参与论文修订、论文最终版本修订。

[参 考 文 献]

- [1] PRIOLA A M, PRIOLA S M, CARDINALE L, et al. The anterior mediastinum: diseases [J]. *Radiol Med*, 2006, 111(3): 312-342.
- [2] MARX A, CHAN J K, COINDRE J M, et al. The 2015 World Health Organization classification of tumors of the thymus: continuity and changes [J]. *J Thorac Oncol*, 2015, 10(10): 1383-1395.
- [3] CHEN G, MARX A, CHEN W H, et al. New WHO histologic classification predicts prognosis of thymic epithelial tumors: a clinicopathologic study of 200 thymoma cases from China [J]. *Cancer*, 2002, 95(2): 420-429.
- [4] GIRARD N, RUFFINI E, MARX A, et al. ESMO Guidelines Committee. Thymic epithelial tumours: ESMO clinical practice guidelines for diagnosis, treatment and follow-up [J]. *Ann Oncol*, 2015, 26(Suppl 5): v40-55.
- [5] SCORSETTI M, LEO F, TRAMA A, D'ANGELILLO R, et al. Thymoma and thymic carcinomas [J]. *Crit Rev Oncol Hematol*, 2016, 99: 332-350.
- [6] VENUTA F, RENDINA E A, LONGO F, et al. Long-term outcome after multimodality treatment for stage III thymic tumors [J]. *Ann Thorac Surg*, 2003, 76(6): 1866-1872; discussion: 1872.
- [7] VENUTA F, RENDINA E A, PESCARMONA E O, et al. Multimodality treatment of thymoma: a prospective study [J]. *Ann Thorac Surg*, 1997, 64(6): 1585-1591; discussion: 1591-1592.
- [8] OTANI Y, YOSHIDA I, ISHIKAWA S, et al. Neoadjuvant intra-arterial infusion chemotherapy for invasive thymoma [J]. *Oncol Rep*, 1997, 4(1): 23-25.
- [9] JEONG Y J, LEE K S, KIM J, et al. Does CT of thymic epithelial tumors enable us to differentiate histologic types and predict prognosis? [J]. *AJR Am J Roentgenol*, 2004, 183(2): 283-289.
- [10] LIU J, YIN P, WANG S C, et al. CT-based radiomics signatures for predicting the risk categorization of thymic epithelial tumors [J]. *Front Oncol*, 2021, 11: 628534.
- [11] FEDOROV A, BEICHEL R, KALPATHY-CRAMER J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network [J]. *Magn Reson Imaging*, 2012, 30(9): 1323-1341.
- [12] FRIEDMAN J, HASTIE T, TIBSHIRANI R. Regularization paths for generalized linear models via coordinate descent [J]. *J Statistical Software*, 33(1): 1-22.

- [13] BISCHL B, LANG M, KOTTHOFF L, et al. MLR: machine learning in R [J] . J Machine Learning Res, 2016, 17(1): 5938–5942.
- [14] ABDEL RAZEK A A, KHAIRY M, NADA N. Diffusion-weighted MR imaging in thymic epithelial tumors: correlation with world health organization classification and clinical staging [J] . Radiology, 2014, 273(1): 268–275.
- [15] XIAO G, HU Y C, REN J L, et al. MR imaging of thymomas: a combined radiomics nomogram to predict histologic types [J] . Eur Radiol, 2021, 31(1): 447–457.
- [16] WANG X H, SUN W, LIANG H Y, et al. Radiomics signatures of computed tomography imaging for predicting risk categorization and clinical stage of thymomas [J] . Biomed Res Int, 2019, 2019: 3616852.
- [17] XIAO G, RONG W C, HU Y C, et al. MRI radiomics analysis for predicting the pathologic classification and TNM staging of thymic epithelial tumors: a pilot studyB [J] . AJR Am J Roentgenol, 2020, 214(2): 328–340.
- [18] DONG W T, XIONG S T, LEI P G, et al. Application of a combined radiomics nomogram based on CE–CT in the preoperative prediction of thymomas risk categorization [J] . Front Oncol, 2022, 12: 944005.
- [19] YU C, LI T, YANG X, ZHANG R, et al. Contrast-enhanced CT-based radiomics model for differentiating risk subgroups of thymic epithelial tumors [J] . BMC Med Imaging, 2022, 22(1): 37.
- [20] LIU W, WANG W, ZHANG H Y, et al. Development and validation of multi-omics thymoma risk classification model based on transfer learning [J] . J Digit Imaging, 2023, 36(5): 2015–2024.
- [21] SHANG L, WANG F, GAO Y, et al. Machine-learning classifiers based on non-enhanced computed tomography radiomics to differentiate anterior mediastinal cysts from thymomas and low-risk from high-risk thymomas: a multi-center study [J] . Front Oncol, 2022, 12: 1043163.
- [22] KAYI CANGIR A, ORHAN K, KAHYA Y, et al. CT imaging-based machine learning model: a potential modality for predicting low-risk and high-risk groups of thymoma: "impact of surgical modality choice" [J] . World J Surg Oncol, 2021, 19(1): 147.
- [23] OZKAN E, ORHAN K, SOYDAL C, et al. Combined clinical and specific positron emission tomography/computed tomography-based radiomic features and machine-learning model in prediction of thymoma risk groups [J] . Nucl Med Commun, 2022, 43(5): 529–539.
- [24] RAJAMOHAN N, GOYAL A, KANDASAMY D, et al. CT texture analysis in evaluation of thymic tumors and thymic hyperplasia: correlation with the international thymic malignancy interest group (ITMIG) stage and WHO grade [J] . Br J Radiol, 2021, 94(1128): 20210583.
- [25] IANNARELLI A, SACCONI B, TOMEI F, et al. Analysis of CT features and quantitative texture analysis in patients with thymic tumors: correlation with grading and staging [J] . Radiol Med, 2018, 123(5): 345–350.
- [26] FENG X L, WANG S Z, CHEN H H, et al. Optimizing the radiomics-machine-learning model based on non-contrast enhanced CT for the simplified risk categorization of thymic epithelial tumors: a large cohort retrospective study [J] . Lung Cancer, 2022, 166: 150–160.
- [27] NAKAJO M, TAKEDA A, KATSUKI A, et al. The efficacy of ¹⁸F-FDG-PET-based radiomic and deep-learning features using a machine-learning approach to predict the pathological risk subtypes of thymic epithelial tumors [J] . Br J Radiol, 2022, 95(1134): 20211050.
- [28] MAHMOUDI S, GRUENEWALD L D, EICHLER K, et al. Multiparametric evaluation of radiomics features and dual-energy CT iodine maps for discrimination and outcome prediction of thymic masses [J] . Acad Radiol, 2023, 30(12): 3010–3021.

(收稿日期: 2024-04-23 修回日期: 2024-06-18)