

论 著

CT影像组学结合肺癌血清肿瘤标志物预测非小细胞肺癌亚型的初步研究*

张桐睿¹ 李 军² 申 敏²
王广丽^{2,*}1.山东第一医科大学研究生院
(山东 泰安 271016)2.山东第一医科大学第一附属医院/山东省
千佛山医院 (山东 济南 250014)

【摘要】目的 通过CT影像组学特征和肺癌血清肿瘤标志物(SCC、NSE、CEA、PRO-GRP、CYFRA21-1)联合检测,建立肺腺癌与鳞癌的鉴别预测模型。方法 回顾性分析2018年1月至2020年6月期间通过活检、纤维支气管镜或手术确诊的100例患者,其中所有患者在治疗前均进行胸部CT平扫和五个血清学指标(SCC、NSE、CEA、PRO-GRP、CYFRA21-1)的检测。在肺窗勾画感兴趣区(Region of Interest,ROI),使用最小绝对收敛和选择算子(least absolute shrinkage and selection operator, LASSO)算法来选择影像组学特征。使用多因素Logistic回归分析建立预测模型,通过曲线下面积(area under the curve, AUC)评估训练集和测试集的影像组学诺模图的性能。结果 从训练集中的1217个影像组学特征中经过LASSO降维筛选出11个特征,计算影像组学评分。并使用性别、年龄、影像组学评分和血清肿瘤标志物构建多因素Logistic模型。经ROC曲线分析,训练集(AUC = 0.944; 95%置信区间(confidence interval, CI): 0.894-0.993)和测试集(AUC = 0.944; 95%CI: 0.867-1.000)能够很好地区分肺鳞癌和腺癌。结论 CT影像组学和肺癌血清肿瘤标志物联合检测的结合可用于预测肺腺癌和肺鳞癌亚型,两者的结合为非小细胞肺癌的亚型诊断及治疗提供了更为可靠的临床依据。

【关键词】影像组学;血清肿瘤标志物;非小细胞肺癌;病理分型

【中图分类号】R734.2

【文献标识码】A

【基金项目】山东省自然科学基金(ZR2015HL096)

DOI:10.3969/j.issn.1672-5131.2023.03.024

Apreliminary Study of CT Radiomic Combined with Serum Tumor Markers in Predicting Subtypes of Non-Small Cell Lung Cancer*

ZHANG Tong-rui¹, LI Jun², SHEN Min², WANG Guang-li^{2,*}.

1.Graduate School, Shangdong First Medical University, Taian 271016, Shandong Province, China

2.Department of Radiology,Shandong Provincial Qianfoshan Hospital/the First Affiliated Hospital of Shandong First Medical University, Jinan 250014, Shandong Province, China

ABSTRACT

Objective A predictive model for differentiating lung adenocarcinoma from squamous carcinoma was established by the combination of CT radiomic features and five major serum tumor markers (SCC, NSE, CEA, PRO-GRP, CYFRA21-1) for lung cancer. **Methods** One hundred lung cancer patients diagnosed by biopsy, fiberoptic bronchoscopy, or surgery between January 2018 and June 2020 were retrospectively identified, in which all patients underwent chest CT scan and five serum tumor markers (SCC, NSE, CEA, PRO-GRP, CYFRA21-1) prior to treatment. Regions of Interest (ROI) were outlined in the lung window, and the least absolute shrinkage and selection operator (LASSO) algorithm was used to select radiomic features and radiomics score. Predictive models were developed using multi-factor logistic regression analysis, and the performance of the radiomic nomogram was evaluated by area under the curve (AUC) for the training and test sets. **Results** From the 1217 radiomic features in the training set, 11 features were selected by LASSO dimensionality reduction to calculate radiomics score. A multifactorial logistic model was also constructed using gender, age, radiomic scores and serum tumor markers. The training set (AUC = 0.944; 95% confidence interval (CI): 0.894-0.993) and the test set (AUC = 0.944; 95% CI: 0.867-1.000) were able to distinguish well between squamous lung cancer and adenocarcinoma by ROC curve analysis. **Conclusion** The combination of CT radiomic and lung cancer serum tumor markers can be used to predict lung adenocarcinoma and squamous lung cancer subtypes, and the combination of the two provides a more reliable clinical basis for the diagnosis and treatment of non-small cell lung cancer subtypes.

Keywords: Radiomics; Serum Tumor Marker; Non-Small-Cell Lung Cancer; Pathological Classification

肺癌作为全球癌症死亡的首要原因^[1],确诊后5年死亡率高达82.3%,临床肺癌可分为两组:小细胞肺癌(SCLC, 25%)和非小细胞肺癌(NSCLC, 75%)。NSCLC最常见的组织学亚型是腺癌(38%)和鳞癌(20%)。这两种亚型具有明显的组织学特征^[2]。不仅治疗方案显著不同,而且预后也有很大区别。早期确诊可延长患者生存时间,因而早发现、早治疗具有重要临床意义^[3]。因此,治疗前准确区分非小细胞肺癌的两种亚型不但有效的提高了治疗效果,还避免了一部分的毒副作用。而影像组学是利用数学算法从CT图像中提取数据进行图像分析的新兴技术,可获得多于传统CT图像价值的信息^[4]。因此,可能在早期就识别出对治疗有反应的标志物,以此来提高患者的生存期^[5]。病理诊断作为常见的肿瘤临床诊断方法,是鉴别肺腺癌和鳞癌的“金标准”。然而,目前尚无非侵入式诊断或分型方法。但是,新兴的影像组学技术为肺癌的术前诊断及鉴别提供了契机^[6-7]。一般来说,肺癌的早期通常不产生任何症状^[8],所以,血清肿瘤标志物对肺癌的检验、确诊及预后具有重要意义。但是,鲜少有人报道应用CT影像组学与血清肿瘤标志物结合对肺癌的诊断及病理分型进行预测的研究。故本研究将二者结合起来,旨在为病人的诊疗及预后提供更为可靠的依据。

1 资料与方法

1.1 病例资料 搜集2018年1月至2020年6月在本院住院的肺癌患者100例,其中腺癌40例,鳞癌60例;男性69例,女性31例;年龄范围37~86岁,中位年龄66岁。采用简单随机抽样法,将100例患者按照7:3的比例随机分配到训练集(70例)和测试集(30例)。所有患者均作术前血清肿瘤标志物(SCC、NSE、CEA、PRO-GRP、CYFRA21-1)检测和胸部CT平扫,均无其他恶性肿瘤病史及放化疗史,且经病理结果证实。

1.2 CT图像的采集 采用GE公司的Discovery CT590 HD、Discovery CT590 RT、Optima CT660、Optima CT520及Revolution CT,扫描参数:管电压120kV,管电流自动调节,层厚、层间距均为0.5mm。从每例患者肺窗图像中提取影像组学特征。由两名有7年以上CT阅片经验的影像科医生,通过RadCloud v2.2(Huiying Medical Technology Co.,Ltd,Beijing,China)平台在肺窗勾画感兴趣区(Region of Interest,ROI),选取在肿瘤CT图像的最大横截面及上下两层轮廓作为ROI,得到病灶的二维结构图像进行影像组学特征的提取。

1.3 五种血清肿瘤标志物的检测 抽取未经过抗肿瘤治疗的患者外周静脉血。静置20 min,然后3 000 r/min离心5 min分离血清,于-20℃冰箱保存。患者血清中CYFRA21-1和NSE水平的检测采用电化学发光法,使用仪器设备及试剂为Roche Cobas E602全自动电化学发光分析仪及试剂;ProGRP、CEA和SCC水平的检测采用化学发光法,所用仪器设

【第一作者】张桐睿,女,住院医师,主要研究方向:CT与MRI诊断。E-mail: 303219875@qq.com

【通讯作者】王广丽,女,副主任医师,主要研究方向:CT与MRI诊断。E-mail: gd-life@163.com

备和试剂为雅培i2000化学发光分析仪及试剂。在影像诊断结果未知的情况下由检验科人员按现有的判定标准对结果做出解释。

1.4 统计分析

1.4.1 影像组学特征的选择及评分的建立 通过最小绝对收敛和选择算子(least absolute shrinkage and selection operator, LASSO)模型,对这1217个CT特征进行筛选。 λ 为惩罚参数根据偏差最小标准选取对应的 λ 值,并保留当前 λ 值对应的非零系数特征。加和计算影像组学得分,影像组学评分公式如下:其中N表示所选特征的个数,coef_i为所选第i个特征的非零系数值,X_i为所选第i个特征的值。

1.4.2 统计描述和预测模型 分训练集和测试集描述年龄、性别、血清肿瘤标志物和影像组学评分。筛选出性别、CEA、SCC和影像组学评分作为自变量。在训练集中建立肺腺癌和鳞癌的预测模型。采用受试者工作特征(receiver operating characteristic, ROC)曲线下面积(area under the curve, AUC)在训练集和测试集中评价logistic回归模型的预测能力,分别计算模型的AUC、灵敏度和特异度,采用DeLong方法计算AUC的95%置信区间(confidence interval, CI)。基于训练集构建的组合模型(性别+CEA+SCC+影像组学评分)绘制诺模图。

1.4.3 采用R 4.0.2软件完成统计工作 统计分析方法均采用双侧检验, $P<0.05$ 表示差异具有统计学意义。

2 结果

2.1 患者特征 共纳入100例患者,随机抽取70%的患者为训练集,训练集共42例鳞癌,28例腺癌;30%的人为测试集,测试集共18例鳞癌,12例腺癌。训练集和测试集中根据肺癌分型患者的统计学结果和临床特征见表1。

2.2 影像组学分数的建立 从患者CT图像中共提取1409个影像组学特征,在训练和测试数据集中筛选出1217个方差不等于0的CT特征。在训练集中,LASSO-logistic回归模型依据十折交叉验证计算得偏差最小时的Log(λ)=-2.38795,并确定了11个具有非零系数的影像组学特征(图1和图2)。在训练集中,鳞癌和腺癌的影像组学评分分别为-0.39±0.62和0.58±0.59,差异具有统计学意义($P<0.001$);在测试集中,鳞癌和腺癌的影像组学评分分别为-0.35±0.71和0.53±0.62,差异具有统计学意义($P=0.002$)(表2)。

2.3 Logistic模型和影像组学诺模图的验证和性能 表3展示了训练集患者预测的单因素和多因素logistic回归分析结果。根据后退法logistic回归筛选出性别、CEA、SCC和影像组学评分为预测因素,性别的优势比(odds ratio, OR)和95%CI为3.27 (0.65-16.57), $P=0.152$; CEA的OR(95%CI)为1.12 (1.02-1.23), $P=0.024$; SCC的OR(95%CI)为0.66 (0.38-1.14), $P=0.136$; 影像

组学评分的OR(95%CI)为19.71 (2.11-184.12), $P=0.009$ 。表4展示了不同变量组合的logistic回归模型的预测AUC。其中,性别、CEA、SCC和影像组学评分同时纳入的logistic回归模型预测性能最好,训练集的AUC为0.944(95%CI: 0.894-0.993),测试集的AUC为0.944(95%CI: 0.867-1.000),其ROC曲线结果见图4。影像组学诺模图结果见图3。

3 讨论

影像组学作为一种高通量的数据挖掘方法,可以利用无创的特点来提取影像图像特征。它提取出大量的图像成像特征,并将图像特征进行快速分析和验证,进而提供详细的肿瘤亚型预测,可应用于临床,这有助于提高肺癌的诊断的准确性。因此,影像组学已被临床应用于预测肺癌的肿瘤亚型、临床分期和远处转移率^[9-10]。

之前有研究表明,LASSO-logistic回归模型和诺模图已被用于预测结、直肠癌和膀胱癌的淋巴结转移及肺癌的预后^[11]。这些研究结果证实了影像组学在类似研究中的有效性。普通胸部CT扫描,是单纯根据横断面的影像图像对病情作出诊断,所以也可能造成早期肺癌的漏诊,因此,肿瘤标志物的结合可作为提高肺癌临床诊断有效性的一种方法,也可以弥补CT在肺癌诊断、定性中的不足。五种血清肿瘤标志物的联合检测,不但提高早期肺癌诊断的敏感性,还有利于降低漏诊率,为肺癌的早期诊断和分型提供有效的临床信息^[12]。因此在本研究中,我们建立了结合血清肿瘤标志物(CYFRA21-1、SCC、CEA、pro-GRP、NSE)的logistic回归模型,我们发现:训练集[AUC = 0.944; 95%置信区间(confidence interval, CI): 0.894-0.993]和验证集(AUC = 0.944; 95%CI: 0.867-1.000)显示出影像组学结合肿瘤标志物对预测非小细胞肺癌的病理分型有较高的准确性。

从我们的结果中可以得出:CT影像组学结合肺癌血清肿瘤标志物的联合检测是诊断非小细胞肺癌病理分型的重要临床依据。但是,本研究中还存在一些局限:(1)本研究中ROI全部为手工勾画,因此在勾画时有可能将正常肺组织划为ROI,或将部分ROI边缘漏画,这些均可能对结果产生影响^[13]。(2)本研究的样本量较小^[14],需进一步增加样本量对研究结果进行验证。(3)我们的研究着重关注CT与病理诊断之间的联系,在收集数据时,其他关键的风险因素(如吸烟史、家族病史及饮酒史)存在缺失。在未来的研究中,我们希望加入这些变量以改进模型。(4)本文数据都来自同一家医院,并且扫描参数相同,这大大降低了模型的泛化能力。

综上所述,影像组学结合血清肿瘤标志物预测非小细胞肺癌亚型可以作为一种诊断方式,对提高提高肺癌的诊断、预后具有一定临床价值。

表1 训练集和测试集的统计学结果和临床特征

	训练集 (n=70)		P	测试集 (n=30)		P
	鳞癌 (n=42)	腺癌 (n=28)		鳞癌 (n=18)	腺癌 (n=12)	
年龄, year	64.14 ± 9.85	63.75 ± 9.82	0.870	65.83 ± 8.83	57.83 ± 10.10	0.036
性别						
男性	36 (85.71)	10 (35.71)	<0.001	17(94.44)	6(50.00)	0.005
女性	6 (14.29)	18 (64.29)		1(5.56)	6(50.00)	
ProGRP, pg/mL	40.75 ± 11.27	37.45 ± 11.44	0.238	36.78 [31.82, 43.06]	39.06 [37.42, 46.99]	0.285
CEA, ng/mL	3.10 [1.54, 5.82]	4.79 [2.60, 24.71]	0.023	2.65 [1.65, 3.30]	3.94 [2.16, 31.66]	0.128
CYFRA21-1, ng/mL	5.56 [2.86, 13.88]	2.71 [2.24, 3.40]	<0.001	5.24 [2.59, 6.41]	2.73 [1.97, 3.61]	0.035
NSE, ug/L	17.01 [10.96, 22.91]	11.47 [8.58, 17.96]	0.019	14.25 [11.26, 17.24]	10.14 [8.07, 15.76]	0.325
SCC, ug/L	2.95 [1.10, 5.93]	0.60 [0.40, 0.85]	<0.001	1.65 [0.42, 2.35]	0.70 [0.55, 1.02]	0.090

表2 训练集和测试集中肺癌分型的基线特征比较

	训练集 (n=70)		P	测试集 (n=30)		P
	鳞癌 (n=42)	腺癌 (n=28)		鳞癌 (n=18)	腺癌 (n=12)	
年龄, year	64.14 ± 9.85	63.75 ± 9.82	0.870	65.83 ± 8.83	57.83 ± 10.10	0.036
性别						
男性	36 (85.71)	10 (35.71)	<0.001	17(94.44)	6 (50.00)	0.005
女性	6 (14.29)	18 (64.29)		1(5.56)	6(50.00)	
ProGRP, pg/mL	40.75 ± 11.27	37.45 ± 11.44	0.238	36.78 [31.82, 43.06]	39.06 [37.42, 46.99]	0.285
CEA, ng/mL	3.10 [1.54, 5.82]	4.79 [2.60, 24.71]	0.023	2.65 [1.65, 3.30]	3.94 [2.16, 31.66]	0.128
CYFRA21-1, ng/mL	5.56 [2.86, 13.88]	2.71 [2.24, 3.40]	<0.001	5.24 [2.59, 6.41]	2.73 [1.97, 3.61]	0.035
NSE, ug/L	17.01 [10.96, 22.91]	11.47 [8.58, 17.96]	0.019	14.25 [11.26, 17.24]	10.14 [8.07, 15.76]	0.325
SCC, ug/L	2.95 [1.10, 5.93]	0.60 [0.40, 0.85]	<0.001	1.65 [0.42, 2.35]	0.70 [0.55, 1.02]	0.090
影像组学得分	-0.39 ± 0.62	0.58 ± 0.59	<0.001	-0.35 ± 0.71	0.53 ± 0.62	0.001

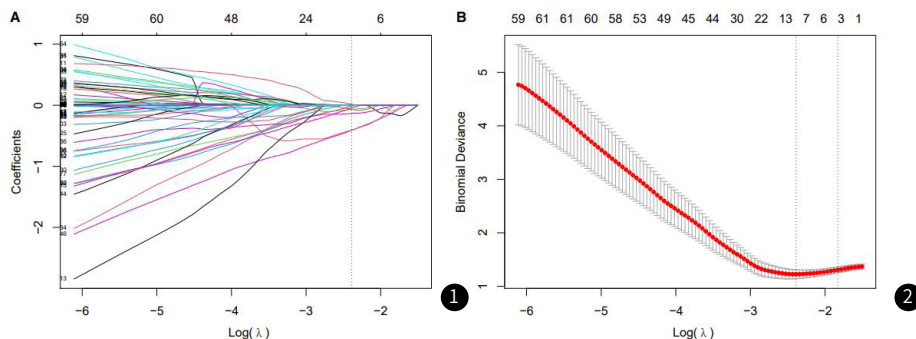


图1 Lasso模型中调节参数(λ)的选择。图2 1217个影像组学特征在模型中的系数。

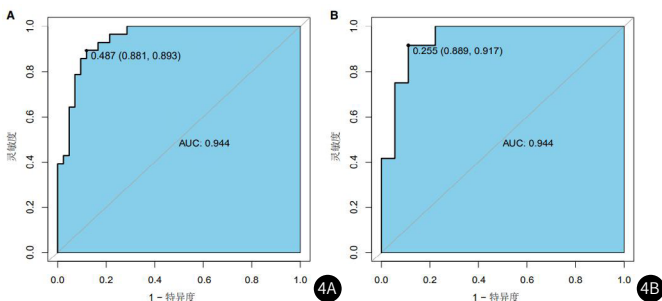
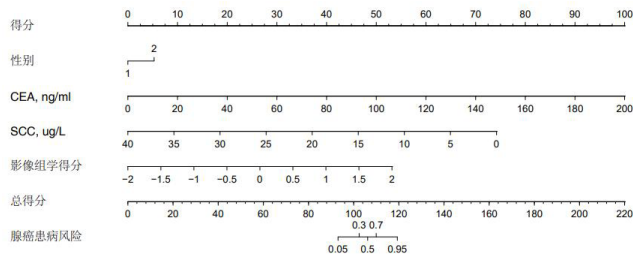


图3 影像组学诺模图。图4 A, B 受试者工作特征(ROC)曲线分析 图4A. 训练集ROC曲线、AUC、灵敏度和特异度; 图4B. 测试集ROC曲线、AUC、灵敏度和特异度。

表3 训练集患者预测的单因素和多因素logistic回归分析

变量名	单因素logistic回归		多因素logistic回归	
	OR (95% CI)	P	OR (95% CI)	P
年龄	1.00 (0.95-1.05)	0.868		
性别 男性	1		1	
女性	10.80 (3.39-34.44)	<0.001	3.27 (0.65-16.57)	0.152
ProGRP, pg/mL	0.97 (0.93-1.02)	0.234		
CEA, ng/mL	1.05 (1.01-1.10)	0.039	1.12 (1.02-1.23)	0.024
CYFRA21-1, ng/mL	0.78 (0.63-0.96)	0.019		
NSE, ug/L	0.92 (0.86-0.99)	0.018		
SCC, ug/L	0.41 (0.23-0.72)	0.002	0.66 (0.38-1.14)	0.136
影像组学评分	19.12 (4.59-79.58)	<0.001	19.71 (2.11-184.12)	0.009

表4 不同logistic回归模型在训练集和测试集的预测效果评价

变量名	训练集 (n=70)			测试集 (n=30)		
	AUC (95% CI)	灵敏度	特异度	AUC (95% CI)	灵敏度	特异度
性别	0.750 (0.645-0.855)	0.643	0.857	0.722 (0.565-0.880)	0.500	0.944
肿瘤标志物指标	0.865 (0.781-0.949)	0.964	0.690	0.787 (0.621-0.954)	1.000	0.556
影像组学得分	0.876 (0.798-0.954)	0.786	0.833	0.847 (0.704-0.990)	0.750	0.833
性别+肿瘤标志物指标	0.901 (0.830-0.971)	0.750	0.905	0.870 (0.743-0.998)	0.667	0.944
性别+影像组学评分	0.901 (0.833-0.970)	0.929	0.762	0.903 (0.792-1.000)	0.917	0.778
组合模型	0.944 (0.894-0.993)	0.893	0.881	0.944 (0.867-1.000)	0.917	0.889

注: 肿瘤标志物指标: CEA、SCC; 组合模型: 性别、CEA、SCC和影像组学评分为自变量

参考文献

- [1] 康书朝, 冯琪, 崔凤, 等. 基于影像组学良性肺结节定量特征的初步研究[J]. 实用放射学杂志, 2020, (06): 893-896.
- [2] He B, Zhao W, Pi JY, et al. A biomarker basing on radiomics for the prediction of overall survival in non-small cell lung cancer patients. Respir Res, 2018; 19(1): 199. Published, 2018, Oct 10.
- [3] 谭于飞, 李玲. 周围型肺癌与局灶性机化性肺炎的CT影像特点及鉴别价值分析[J]. 中国CT和MRI杂志, 2018, 16(4): 60-62, 76.
- [4] Feng B, Chen X, Chen Y, et al. Differentiating minimally invasive and invasive adenocarcinomas in patients with solitary sub-solid pulmonary nodules with a radiomics nomogram. Clin Radiol, 2019; 74(7): 570.e1-570.e11.
- [5] Chetan MR, Gleason FV. Radiomics in predicting treatment response in non-small-cell lung cancer: current status challenges and future perspectives. Eur Radiol, 2021; 31(2): 1049-1058.
- [6] Bousabarah K, Temming S, Hoevels M, et al. Radiomic analysis of planning computed tomograms for predicting radiation-induced lung injury and outcome in lung cancer patients treated with robotic stereotactic body radiation therapy. Radiomics-Analyse von Planung-Computertomogrammen zur Vorhersage von strahleninduzierter Lungenschädigung und onkologischem Ergebnis bei Lungentumorspatienten nach robotischer stereotaktischer Strahlentherapie. Strahlenther Onkol, 2019; 195(9): 830-842.
- [7] 王大鹏, 杨晓光, 赵磊, 等. 基于CT的影像组学分析在鉴别肺鳞癌与肺腺癌中的应用[J]. 临床放射学杂志, 2019, 38(11): 2055-2058.
- [8] 付亮, 王宏亮, 赵钰. 周围型肺癌在多层螺旋CT诊断中的影像学表现与术后病理学诊断的一致性分析[J]. 中国CT和MRI杂志, 2020, 18(4): 43-45.
- [9] 朱静, 肖欢, 周莹, 等. 影像组学在鉴别原发性肺癌与肉芽肿疾病中的应用[J]. 实用放射学杂志, 2020, (2): 202-206.
- [10] Chen XF, Fang MD, Dong D, et al. A radiomics signature in preoperative predicting degree of tumor differentiation in patients with non-small cell lung cancer. Acad Radiol, 2018; 25(12): 1548-1555.
- [11] Yang L, Yang J, Zhou X, et al. Development of a radiomics nomogram based on the 2D and 3D CT features to predict the survival of non-small cell lung cancer patients. Eur Radiol, 2019; 29(5): 2196-2206.
- [12] 宁爱东, 张林飞, 汪国伟. CT三维重建成像对孤立性肺结节形态特征的显示及良恶性的诊断价值探究[J]. 罕少疾病杂志, 2022, 29(8): 36-38.
- [13] Wang SZ, Zhou ML, Liu Z, et al. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. Med Image Anal, 2017; 40: 172-183.
- [14] 陈雅荣, 综述, 刘挨师, 审校. 影像组学在肺磨玻璃结节中的研究进展[J]. 实用放射学杂志, 2021, (6): 1032-1034, 1046.

(收稿日期: 2021-12-25)

(校对编辑: 朱丹丹)