

论著

To Constitute A Stratified Prediction Model of PD-L1 Expression in Primary Lung Adenocarcinoma Based on Artificial Intelligence Analysis Software

LUO Ya-ting, XU Wen-bei, XIE Zhi-wen, MENG Yan-kai, WANG Xiu-ling*.

Department of Radiology, The Affiliated Hospital of Xuzhou Medical University, XuZhou 221000, JiangSu Province, China

ABSTRACT

Objective To constitute a stratified prediction model of PD-L1 expression in primary lung adenocarcinoma based on imaging features obtained by artificial intelligence analysis software.

Methods The imaging data of 129 patients with primary lung adenocarcinoma confirmed by operation from June 2020 to June 2021 were analyzed retrospectively. The imaging features of the lesions and CT quantitative parameters and texture features were automatically obtained by the artificial intelligence lung nodule analysis software. It was then reviewed by a radiologist with 5 years experience in the diagnosis of chest CT, and the final decision was made by a radiologist with 25 years' experience in diagnosis. The expression of PD-L1 in pathological specimens was detected by immunohistochemical staining. The patients were divided into PD-L1 positive expression group ($TPS \geq 1\%$) and PD-L1 negative expression group ($TPS < 1\%$). SPSS26.0 software was used for statistical analysis. Logistic regression analysis was used to screen the predictive factors and construct predictive model. ROC analysis was used to evaluate the efficacy of the model. P values less than 0.05 were considered statistically significant. **Results** The volume and proportion of solid components in PD-L1 positive expression group were significantly higher than those in negative expression group ($P < 0.001$). The mass of solid components between the two groups was 5.57 and $1.59 \text{ mg} \times 10^3$ respectively. The difference was statistically significant ($P = 0.000$). The proportion of solid components in the positive expression group was also significantly higher than that in the negative group, which were 0.73 and 0.41 respectively ($P = 0.000$). The total mass of positive group and negative group was $6.40, 2.76 \text{ mg} \times 10^3$ respectively. The maximum and average CT values in PD-L1 positive expression group were higher than those in negative expression group ($P = 0.000$). There were significant differences in skewness and entropy between the two groups ($P = 0.001, 0.002$). The solid components, CT values and texture features were included in the Logistic regression analysis to constitute a prediction model. The ROC curve analysis showed that the AUC value of the combined variables was 0.887, the sensitivity was 74.5%, and the specificity was 90.2%. The difference was statistically significant ($P < 0.001$). **Conclusion** The imaging quantitative features based on artificial intelligence analysis software can be used to predict the expression level of PD-L1 in primary lung adenocarcinoma, and the prediction model of multi-parameter combination variables is the best.

Keywords: Lung Adenocarcinoma; PD-L1; CT; Quantitative Parameter; Texture Analysis

肺腺癌是肺癌最常见的病理类型^[1]。随着个体化精准治疗的发展，免疫检查点抑制剂(immune checkpoint inhibitors, ICIs)越来越多地应用于肺腺癌的治疗，其中以细胞程序性死亡受体-1(programmed cell death protein 1, PD-1)/细胞程序性死亡因子配体-1(programmed death ligand 1, PD-L1)抑制剂为代表的ICIs应用最为广泛。研究表明，PD-L1表达值 $\geq 1\%$ 的患者可通过PD-L1抑制剂治疗获益^[2-4]。

目前，临幊上PD-L1检测主要依赖免疫组化(Immunohistochemistry, IHC)获得，对于无法手术、穿刺活检的患者，由于无法获得病变标本，难以进行PD-L1检测，故限制了ICIs治疗的广泛应用^[5]。并且PD-L1表达在肺腺癌不同组织亚型中异质性较强，且活检需要较高成本^[6]。有学者尝试通过影像学、影像组学、机器学习的方法区分PD-L1表达水平，已经取得了较好的研究结果^[7-8]。但是，基于影像组学的方法需要医学与工科背景的研究人员进行多学科合作，同时图像分割耗时较多。随着人工智能分析软件在临幊更多的应用，并且可以快速的提供更多的定量参数特征，能否基于人工智能分析软件，用于构建PD-L1分层预测模型具有重要的临幊意义。

因此，本研究拟探讨基于人工智能分析软件的影像学特征在预测原发肺腺癌PD-L1表达中的价值。

1 资料与方法

1.1 研究资料 回顾性分析2020年6月至2021年6月于我院经手术切除的肺腺癌患者资料。

纳入标准：手术病理证实为原发肺腺癌；手术病理与胸部CT检查时间间隔小于1个月；患者行PD-L1免疫组化检查。排除标准：术前行肿瘤相关治疗，如免疫、靶向等；CT伪影明显，无法用于分析；合并明显的肺不张、胸腔积液等影响结节评估的因素，导致无法具体测量。

1.2 检查方法 所有患者均采用多层螺旋CT扫描，CT扫描仪：Siemens Somatom

基于人工智能分析软件 构建原发肺腺癌PD-L1 表达的分层预测模型

骆雅婷 徐文北 谢智雯

孟闫凯 汪秀玲*

徐州医科大学附属医院影像科
(江苏 徐州 221000)

【摘要】目的利用人工智能分析软件获取的影像学特征构建原发肺腺癌PD-L1表达的分层预测模型。**方法**回顾性分析2020年6月至2021年6月共129例经手术证实为原发肺腺癌的患者影像资料，由人工智能肺结节分析软件自动获得病灶影像表征、CT定量参数及纹理特征，再由1名具有5年胸部CT诊断经验的影像科医师进行复核，有争议时由一名具有25年诊断经验的医师最终阅片决定。采用免疫组织化学染色法检测患者病理标本中PD-L1的表达，将患者分为PD-L1阳性表达组($TPS \geq 1\%$)、PD-L1阴性表达组($TPS < 1\%$)。采用SPSS 26.0软件进行统计分析，Logistic回归分析筛选预测因子，构建预测模型，受试者工作特征曲线(ROC)分析评价模型效能。**结果**PD-L1阳性表达组实性体积、实性体积占比均明显高于阴性表达组，P值均 < 0.001 。两组间实性质量分别为5.57、 $1.59 \text{ mg} \times 10^3$ ，差异具有明显统计学意义($P=0.000$)，阳性表达组实性质量占比也明显高于阴性组，分别为0.73、0.41($P=0.000$)。阳性组、阴性组总质量分别为6.40、 $2.76 \text{ mg} \times 10^3$ ($P=0.015$)。PD-L1阳性表达组CT最大值、平均值均高于阴性表达组(P 值均 $=0.000$)。两组间偏度及熵值差异有统计学意义($P=0.001, 0.002$)。将实性成分、CT值、纹理特征纳入Logistic回归分析，构建组合变量，ROC曲线分析显示，组合变量预测高危组结节的AUC值为0.887，敏感度74.5%，特异度90.2%，差异具有统计学意义($P<0.001$)。**结论**基于人工智能分析软件的影像学定量特征能够用于预测原发肺腺癌PD-L1表达水平，多参数组合变量的预测模型最佳。

【关键词】肺腺癌；细胞程序性死亡因子配体-1；

计算机断层扫描；定量参数；纹理分析

【中图分类号】R734.2; R445.3

【文献标识码】A

DOI:10.3969/j.issn.1672-5131.2023.12.015

【第一作者】骆雅婷，女，住院医师，主要研究方向：影像诊断和研究工作。E-mail: 820414477@qq.com

【通讯作者】汪秀玲，女，主任医师，主要研究方向：影像诊断的临床教学与科研。E-mail: xuzhouwangxl@hotmail.com

Emotion 16和Somatom Definition CT机、Philips Ingenuity core 128及GE Optima 660 CT机。

扫描参数：管电压120kV，管电流为自动mAs，螺距为1.1.5，层厚为5.0mm，重组层厚为1.25mm或1.5mm，通过高分辨率骨算法重组。所有图像均上传至GE PACS系统、人工智能肺结节分析软件(Dr. Wise OCR 2.0)用于分析。

1.3 病理学分析 标本采用免疫组织化学染色，染色方法：Dako全自动免疫组织化学染色，抗体类型：PD-L1(22C3)，在显微镜下评价PD-L1染色得分。肿瘤比例评分(tumor proportion score, TPS)指肿瘤细胞中染色阳性的细胞比例。本实验将PD-L1 TPS $\geq 1\%$ 的病例定义为阳性表达组，PD-L1 TPS<1%为阴性表达组。由两名有经验的细胞病理医师独立判读，任何不一致的结果

均重新评价得到共识。

1.4 影像学分析 原发肺腺癌所有初始影像表征、CT平扫定量参数及纹理特征值由人工智能肺结节分析软件自动获得，再由1名具备5年胸部CT诊断经验的医生进行复核，有争议时由一名具有25年诊断经验的医师最终阅片决定。

病灶基线影像特征包括不规则、分叶征、支气管截断征、血管包埋征、胸膜牵拉^[9-10]。将阈值设定为-300HU^[11]获取CT定量参数包括实性体积、实性体积占比、实性质量、实性质量占比、总体积、总质量、CT最大值、CT最小值、CT平均值、CT值方差、球形度、最大表面积、表面积、3D长径、长短径平均值，纹理特征包括峰度、偏度、能量、紧凑度、熵(图1)。



图1A-图1C 男,50岁,左肺下叶浸润性腺癌,PD-L1 TPS评分<1%,表现为混杂磨玻璃密度结节,将薄层CT平扫图像上传至人工智能肺结节分析软件(Dr. Wise OCR 2.0)分析,将阈值设定为-300HU,图1B中深蓝色区域为自动识别的实性成分,图1C为分析软件自动得出的CT定量参数及纹理特征值。

1.5 统计学分析 采用SPSS 26.0软件对数据进行统计分析。所有定量资料以均数±标准差($\bar{x} \pm s$)表示，并进行正态分布及方差齐性分析，符合正态分布及方差齐性的采用独立样本T检验，不符合正态分布的采用Mann-Whitney U检验。分类计数资料均采用百分比(%)表示，组间比较采用 χ^2 检验。利用Logistic回归分析对有意义的参数进行多参数联合分析，绘制受试者工作特征(ROC)曲线，探究各项参数预测单独及联合预测PD-L1表达状态的能力。P<0.05为差异具有统计学意义。

2 结 果

2.1 人口统计学特征 本共纳入129例患者，其中男性51例，女性78例。年龄30-83岁，中位年龄61岁。根据PD-L1表达情况将所有患者分为阴性表达组(TPS<1%，n=82)和阳性表达组(TPS $\geq 1\%$ ，n=47)。PD-L1阳性表达组男性占比高于阴性表达组，差异有统计学意义(P<0.05)。两组患者年龄及影像表征差异均无统计学意义(P>0.05)(表1)。

表1 PD-L1阳性表达组和阴性表达组人口统计学特征

临床影像学特征	阳性表达组(n=47)(%)	阴性表达组(n=82)(%)	P值
性别			0.016
男性	25(53.2)	26(31.7)	
女性	22(46.8)	56(68.3)	
年龄(岁)	62.57±10.48	59.27±10.49	0.087
病灶影像特征			
不规则			0.579
有	10(21.3)	21(25.6)	
无	37(78.7)	61(74.4)	
分叶征			1.000
有	4(8.5)	6(7.3)	
无	43(91.5)	76(92.7)	
支气管截断征			0.197
有	9(19.1)	9(11.0)	
无	38(80.9)	73(89.0)	
血管包埋征			0.421
有	1(2.1)	6(7.3)	
无	46(97.9)	76(92.7)	
胸膜牵拉			0.747
有	7(14.9)	14(17.1)	
无	40(85.1)	68(82.9)	

2.2 实性成分差异 PD-L1阳性表达组实性成分体积、实性成分体积占比均明显高于阴性表达组，P值均<0.001。两组间实性成分质量分别为5.57、1.59mg $\times 10^3$ ，差异具有明显统计学意义(P=0.000)，阳性表达组实性成分质量占比也明显高于阴性组，分别为0.73、0.41(P=0.000)。阳性组、阴性组总质量分别为6.40、2.76mg $\times 10^3$ (P=0.015)(表2、图2)。

2.3 CT值及径线相关参数差异 PD-L1阳性表达组CT最大值、平均值均高于阴性表达组，分别为121.49、43.41HU和-238.64、-441.39HU(P值均=0.000)(表3、图2)。

2.4 纹理特征差异 PD-L1阳性表达组偏度绝对值小于阴性表达组，两组间偏度分别为-1.34、-1.88，差异具有统计学意义(P=0.001)。阳性组熵值低于阴性组，分别为8.96、9.65(P=0.002)(表4、图2)。

表2 PD-L1阳性表达组和阴性表达组间实性成分相关参数

实性成分相关参数	阳性表达组(n=47)	阴性表达组(n=82)	P值
实性成分体积(mm^3)($\times 10^3$)	5.69±10.98	1.75±3.01	0.000
实性成分体积占比	0.62±0.27	0.30±0.24	0.000
实性成分质量(mg)($\times 10^3$)	5.57±10.88	1.59±2.88	0.000
实性成分质量占比	0.73±0.23	0.41±0.26	0.000
总体积(mm^3)($\times 10^3$)	7.54±12.74	4.56±4.55	0.227
总质量(mg)($\times 10^3$)	6.40±11.63	2.76±3.39	0.015

表3 PD-L1阳性表达组和阴性表达组间CT值及径线相关参数比较

参数	阳性表达组(n=47)	阴性表达组(n=82)	P值
CT最大值(HU)	121.49±101.93	43.41±127.73	0.000
CT最小值(HU)	-892.02±100.57	-909.80±76.36	0.297
CT平均值(HU)	-238.64±175.63	-441.39±159.83	0.000
CT值方差($\times 10^4$)	4.73±1.81	4.18±1.62	0.080
最大面面积(mm^2)	354.17±348.73	278.22±198.95	0.117
表面积(mm^2)($\times 10^3$)	1.87±1.92	1.38±1.03	0.058
3D长径(mm)	28.49±13.41	25.25±10.74	0.135
长短径平均值(mm)	21.74±9.66	19.43±7.52	0.132

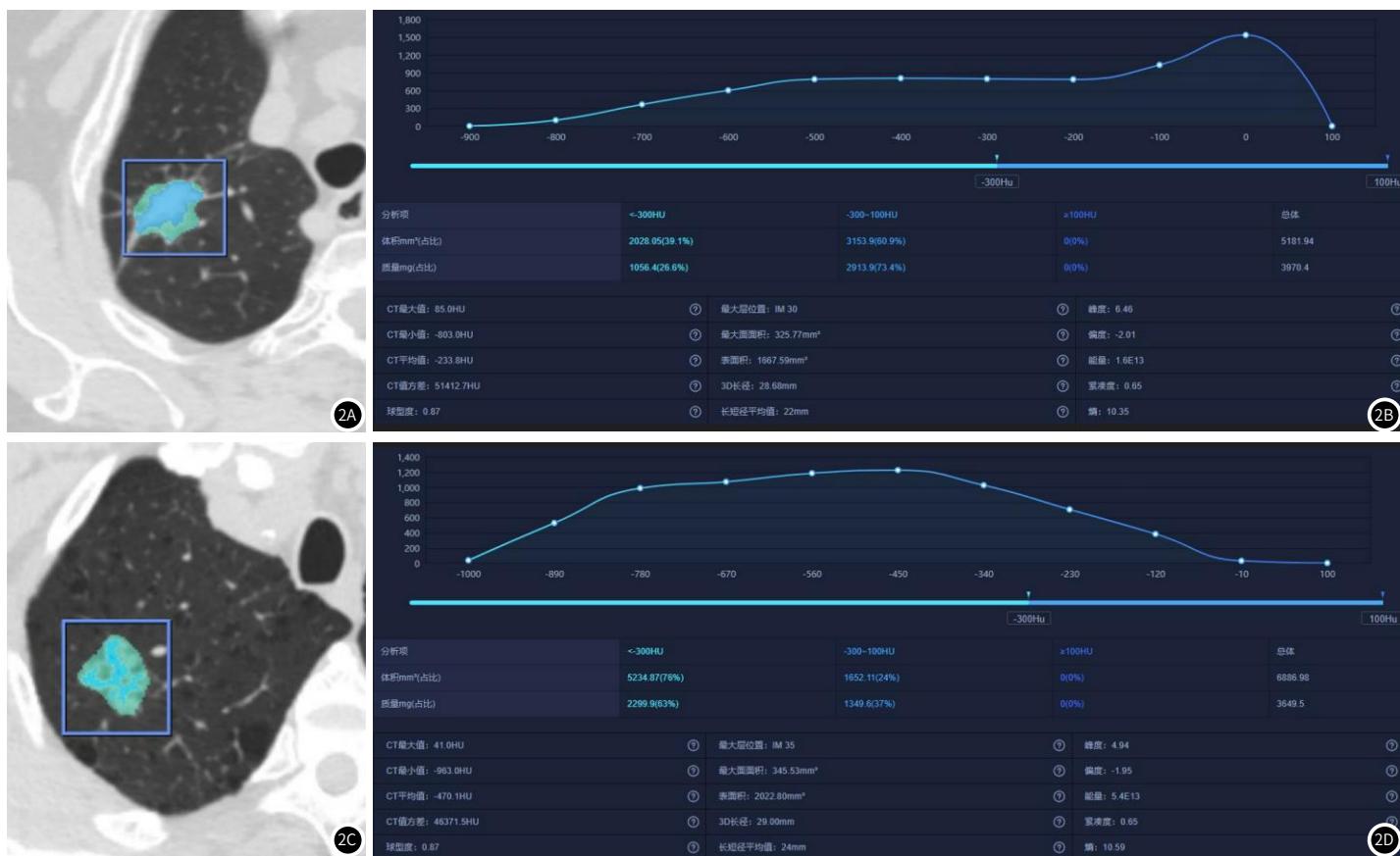


图2A-图2B 女, 72岁, 右肺上叶肺腺癌, PD-L1 TPS评分>1%, 表现为混杂磨玻璃密度结节, 其实性成分体积3153.9mm³、实性成分体积占比60.9%、实性成分质量2913.9mg、实性成分质量占比73.4%、总质量3970.4mg、CT最大值85.0HU、CT平均值-233.8HU、偏度-2.01、熵10.35;

图2C-图2D 男, 59岁, 右肺上叶肺腺癌, PD-L1 TPS评分<1%, 表现为混杂磨玻璃密度结节, 其实性成分体积1652.11mm³、实性成分体积占比24%、实性成分质量1349.6mg、实性成分质量占比37%、总质量3649.5mg、CT最大值41.0HU、CT平均值-963.0HU、偏度-1.95、熵10.59。

表4 PD-L1阳性表达组和阴性表达组间纹理特征比较

参数	阳性表达组(n=47)	阴性表达组(n=82)	P值
球型度	0.89±0.09	0.91±0.09	0.262
峰度	5.65±5.50	5.48±1.65	0.068
偏度	-1.34±0.96	-1.88±0.37	0.001
能量($\times 10^{13}$)	20.46±11.05	7.69±28.08	0.487
紧凑度	0.73±0.20	0.77±0.20	0.224
熵	8.96±1.26	9.65±0.96	0.002

表5 CT定量参数及纹理特征值预测PD-L1表达状态效能

参数	最佳截断值	AUC	敏感度(%)	特异度(%)	P值
实性体积(mm ³)	1232.865	0.710	68.1	64.6	0.000
实性体积占(%)	55	0.811	63.8	84.1	0.000
实性体积组合变量	0.459	0.811	63.8	84.1	0.000
实性质量(mg)	1092.85	0.715	68.1	64.6	0.000
实性质量占(%)	69.15	0.819	68.1	84.1	0.000
总质量(mg)	1392.4	0.629	78.7	46.3	0.015
实性质量组合变量	0.466	0.828	70.2	85.4	0.000
CT最大值(HU)	49	0.688	89.4	46.3	0.000
CT平均值(HU)	-309.9	0.793	66.0	80.5	0.000
CT值组合变量	0.394	0.795	68.1	80.5	0.000
偏度	-1.64	0.682	48.9	87.8	0.001
熵	8.97	0.670	74.4	63.8	0.001
纹理特征组合变量	0.340	0.681	63.8	75.6	0.001
所有参数组合变量	0.480	0.887	74.5	90.2	0.000

2.5 CT定量参数预测PD-L1表达状态效能 实性成分体积、实性成分体积占比、实性成分质量、实性成分质量占比、总质量、CT最大值、CT平均值、偏度、熵9个参数的曲线下面积(AUC)分别为0.710、0.811、0.715、0.819、0.629、0.688、0.793、0.682、0.670, 其中实性成分体积、体积占比、实性成分质量、质量占比、及CT平均值具有相对较好的预测效能。利用Logistic回归分析对9个参数进行多参数联合分析, 体积组合变量、质量组合变量、CT值组合变量、纹理特征组合变量及所有参数组合变量的AUC分别为0.811、0.828、0.795、0.681、0.887, 其中所有参数组合变量的预测效能最优, 阈值为0.480时, 其诊断敏感度为74.5%、特异度为90.2%(表5, 图3)。

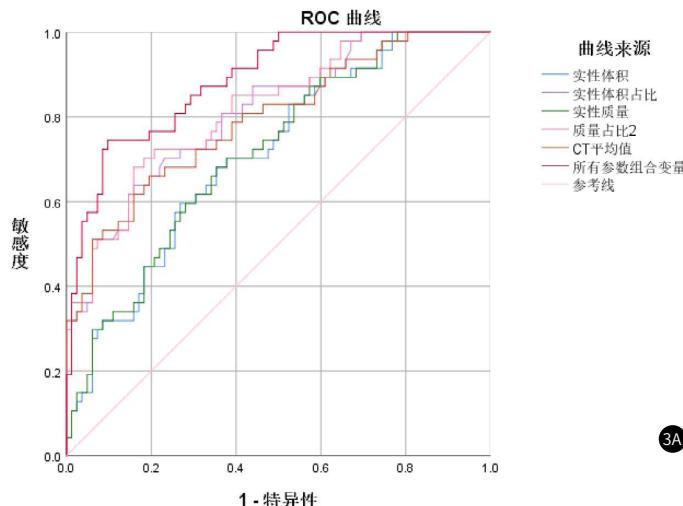
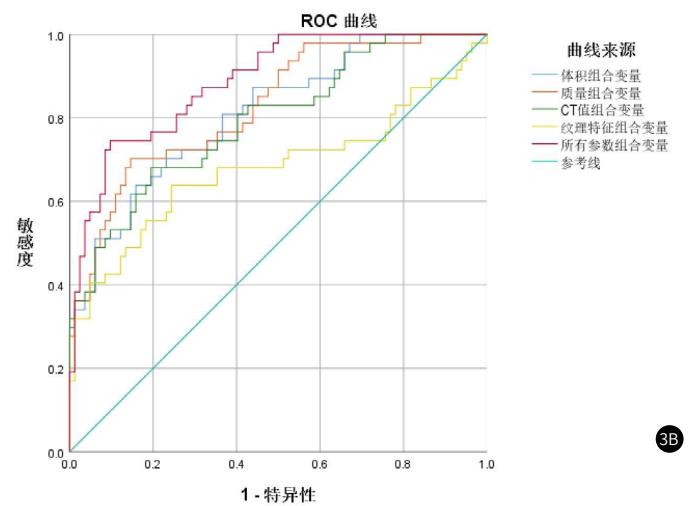


图3A 实性成分体积、体积占比、实性成分质量、质量占比、CT平均值及所有参数组合变量预测PD-L1表达状态的ROC曲线图，曲线下面积分别为0.710、0.811、0.715、0.793、0.887，所有参数组合变量的预测效能最优。

图3B 体积组合变量、质量组合变量、CT值组合变量、纹理特征组合变量、所有参数组合变量ROC曲线图，曲线下面积分别为0.811、0.828、0.795、0.681、0.887，所有参数组合变量的预测效能最优。



3 讨 论

据文献报道，PD-L1抑制剂显著提高了肺腺癌患者的长期生存率，而要实现其在患者中的准确应用，需要确定PD-L1的表达水平^[12-14]，因此基于人工智能分析软件获取的影像特征构建PD-L1表达分层模型十分有意义。

本研究结果表明实性成分质量、体积大小与PD-L1表达水平相关，病灶中实性成分越多PD-L1阳性表达的可能性越高，这与既往研究是一致的^[9-10,15]。病灶中实性成分与病理侵袭性、组织学亚型相关，实性和微乳头亚型为主的肺腺癌具有更大的恶性潜能，PD-L1表达水平更高^[16-17]。Dai等也报道了CT扫描图像上病灶实性成分占比与病理侵袭性正相关^[18]，因此，影像学上实性成分越多，肿瘤的侵袭性越强，PD-L1阳性表达水平更高。

本研究进一步将实性成分、CT值、纹理特征分别进行建模分析，相对于纹理特征，实性成分相关特征构建模型的预测效能最好，进一步表明肿瘤实性成分与PD-L1表达相关性更高。我们的结果中纹理特征构建的模型预测效能相对偏弱，这可能是由于本研究中提取的纹理特征相对较少，并且我们没有评估实性成分的纹理特征有关。而将多维度参数组合可以进一步提升模型的预测效能，且敏感度、特异度也有所提升。

同时，本研究还得出CT值相关特征与PD-L1表达显著相关。CT值反映了肿瘤对X射线的衰减和肿瘤组织的密度，受肿瘤实性成分的影响较大，病灶内实性成分越多，肿瘤CT值更大，肿瘤侵袭性更强，PD-L1表达更高^[19]。本研究中PD-L1阳性表达组男性占比更高，这与Wu T等^[9]的研究结果一致。本研究中PD-L1两组间影像表征均无统计学差异，这与Yoon J等^[10]的研究结果是一致的。我们认为这可能是由于本研究中入组的病例，T1期相对较多，肿瘤相对较小，形态学上的异质性并不明显导致的。

本研究存在的不足：(1)样本量较小；(2)本研究为单中心研究，缺乏外部对照；(3)没有对肺腺癌不同亚组患者PD-L1表达进一步研究。

综上所述，基于人工智能分析软件的影像学定量特征能够用于预测原发肺腺癌PD-L1表达水平。

参 考 文 献

- [1] Siegel RL, Miller KD, Fuchs HE, et al. Cancer statistics, 2022 [J]. CA Cancer J Clin, 2022; 72 (1): 7-33.
- [2] Deng H, Zhao Y, Cai X, et al. PD-L1 expression and tumor mutation burden as pathological response biomarkers of neoadjuvant immunotherapy for early-stage non-small cell lung cancer: a systematic review and meta-analysis [J]. Crit Rev Oncol Hematol, 2022, 170: 103582.
- [3] 凌徐心仪的, 张瑶, 钟华. 非小细胞肺癌免疫治疗获益人群筛选的研究进展 [J]. 上海交通大学学报(医学版), 2021, 41 (8): 1114-1119.
- [4] Mok TSK, Wu YL, Kudaba I, et al. Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): a randomised, open-label, controlled, phase 3 trial [J]. Lancet, 2019; 393 (10183): 1819-1830.
- [5] Shi W, Yang Z, Zhu M, et al. Correlation between PD-L1 expression and radiomic features in early-stage lung adenocarcinomas manifesting as ground-glass nodules [J]. Front Oncol, 2022, 12: 986579.
- [6] Tanco V, Grendár M, Farkaová A, et al. Programmed death ligand 1 protein expression, histological tumour differentiation and intratumoural heterogeneity in pulmonary adenocarcinoma [J]. Pathology, 2020; 52 (5): 538-545.
- [7] Jiang Z, Dong Y, Yang L, et al. CT-based hand-crafted radiomic signatures can predict PD-L1 expression levels in non-small cell lung cancer: a two-center study [J]. J Digit Imaging, 2021; 34 (5): 1073-1085.
- [8] 朱炳印, 汝晓睿, 张恒, 等. 双层探测器光谱CT对肺腺癌PD-L1表达水平分层研究价值 [J]. 临床放射学杂志, 2022, 41 (8): 1451-1455.
- [9] Wu T, Zhou F, Soodeen-Lalloo AK, et al. The association between imaging features of TSCT and the expression of PD-L1 in patients with surgical resection of lung adenocarcinoma [J]. Clin Lung Cancer, 2019; 20 (2): e195-e207.
- [10] Yoon J, Suh YJ, Han K, et al. Utility of CT radiomics for prediction of PD-L1 expression in advanced lung adenocarcinomas [J]. Thorac Cancer, 2020; 11 (4): 993-1004.
- [11] Yanagawa M, Tanaka Y, Kusumoto M, et al. Automated assessment of malignant degree of small peripheral adenocarcinomas using volumetric CT data: correlation with pathologic prognostic factors [J]. Lung Cancer, 2010; 70 (3): 286-294.
- [12] Paz-Ares L, Spira A, Raben D, et al. Outcomes with durvalumab by tumour PD-L1 expression in unresectable, stage III non-small-cell lung cancer in the PACIFIC trial [J]. Ann Oncol, 2020; 31 (6): 798-806.
- [13] Forde PM, Spicer J, Lu S, et al. Neoadjuvant nivolumab plus chemotherapy in resectable lung cancer [J]. N Engl J Med, 2022, 386 (21): 1973-1985.
- [14] Cascone T, William WN Jr, Weissferdt A, et al. Neoadjuvant nivolumab or nivolumab plus ipilimumab in operable non-small cell lung cancer: the phase 2 randomized NEOSTAR trial [J]. Nat Med, 2021, 27 (3): 504-514.
- [15] Toyokawa G, Takada K, Okamoto T, et al. Relevance between programmed death ligand 1 and radiologic invasiveness in pathologic stage I lung adenocarcinoma [J]. Ann Thorac Surg, 2017, 103 (6): 1750-1757.
- [16] Shimoji M, Shimizu S, Sato K, et al. Clinical and pathologic features of lung cancer expressing programmed cell death ligand 1 (PD-L1) [J]. Lung Cancer, 2016, 98: 69-75.
- [17] 郑雅琳, 邸吉廷, 熊焰, 等. I ~ II 期肺腺癌PD-L1表达的临床病理意义 [J]. 中华病理学杂志, 2021, 50 (7): 796-798.
- [18] Dai J, Shi J, Soodeen-Lalloo AK, et al. Air bronchogram: a potential indicator of epidermal growth factor receptor mutation in pulmonary subsolid nodules [J]. Lung Cancer, 2016; 98: 22-28.
- [19] Lee HY, Choi YL, Choi KS, et al. Pure ground-glass opacity neoplastic lung nodules: histopathology, imaging, and management [J]. AJR Am J Roentgenol, 2014, 202 (3): W224-W233.

(收稿日期: 2022-12-28)
(校对编辑: 孙晓晴)