

✧胸部影像学

CT radiomics model for evaluation on pathologic types of lung adenocarcinoma in situ combined with minimally invasive adenocarcinoma and invasive adenocarcinoma

YE Dingli¹, JIANG Wen², WU Jiani¹, HUANG Zhicheng^{1*}

(1. Department of Radiology, Jilin Cancer Hospital, Changchun 130021, China; 2. Department of Radiology, School Hospital of Shenzhen Technology University, Shenzhen 518118, China)

[Abstract] **Objective** To investigate the value of CT radiomics model for predicting pathologic types of adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA) and invasive adenocarcinoma (IAC) among lung adenocarcinoma. **Methods** Data of 542 patients with pathologically confirmed lung adenocarcinoma and clear subtypes were retrospective analyzed. AIS and MIA were classified as group 1 while IAC as group 2. The gender and age were compared between 2 groups. Feature extraction software was used to extract three-dimensional texture feature parameters of each lesion, and the imaging omics features obviously different between 2 groups were retained, then the optimal imaging omics features were selected to build a predictive model. All the data were divided into training set and validation set in a ratio of 2:1. Six machine learning algorithms were used to classify the five-fold cross-training sets to select the best classifier. Then, the five-fold cross-training data set, training set and validation set were analyzed with the prediction model to obtain the ROC curves of the model in predicting pathological subtypes of lung adenocarcinoma as well as the relative AUC, accuracy, specificity and sensitivity. **Results** There were 235 patients in group 1 and 307 in group 2. No statistical difference of gender nor age was found between 2 groups ($\chi^2=0.56$, $t=-0.19$, $P=0.63$, 0.98). A total of 1 766 three-dimensional texture feature parameters were extracted from the lesions, including 988 imaging omics features significantly different between 2 groups. Finally, 10 optimal imaging omics features were retained to construct the prediction model. Perceptron classifier was the best classifier. AUC of the predictive model in predicting pathological subtypes of validation set was 0.95, and the relative accuracy, specificity and sensitivity was 0.88, 0.87 and 0.84, respectively. **Conclusion** CT radiomics model could effectively predict pathological subtypes of AIS, MIA and IAC among lung adenocarcinoma.

[Keywords] lung neoplasms; tomography, X-ray computed; radiomics; pathology

DOI:10.13929/j.issn.1003-3289.2020.09.017

基于 CT 影像组学模型预测肺原位腺癌及微浸润腺癌与浸润性腺癌

叶钉利¹, 姜雯², 吴佳妮¹, 黄志成^{1*}

(1. 吉林省肿瘤医院放射线科, 吉林 长春 130021; 2. 深圳技术大学附属医院放射线科, 广东 深圳 518118)

[摘要] **目的** 探讨基于 CT 影像组学预测肺腺癌中的原位癌(AIS)和微浸润腺癌(MIA)以及浸润性腺癌(IAC)的价值。
方法 回顾性分析 542 例经手术病理确诊且病理亚型明确的肺腺癌患者, 将 AIS 及 MIA 归为第 1 组, IAC 为第 2 组。比较 2

[基金项目] 吉林省卫生与健康技术创新项目(2018J026)。

[第一作者] 叶钉利(1993—), 男, 吉林吉林人, 学士, 技师。研究方向: 胸部影像组学。E-mail: yedl1911@mails.jlu.edu.cn

[通信作者] 黄志成, 吉林省肿瘤医院放射线科, 130021。E-mail: 33088401@qq.com

[收稿日期] 2019-11-21 **[修回日期]** 2020-05-13

组患者性别和年龄差异。采用特征提取软件提取病灶三维纹理特征参数,分析组间差异明显的影像组学特征,筛选最佳影像组学特征构建预测模型。按 2:1 比例将数据分为训练集和验证集,采用 6 种机器学习算法对 5 倍交叉验证数据集进行分类,选择最佳分类器;以之分析 5 倍交叉验证数据集、训练集和验证集,获得模型预测肺腺癌病理分型的 ROC 曲线及相应 AUC、特异度、敏感度及准确率。**结果** 第 1 组 235 例,第 2 组 307 例,组间性别和年龄差异均无统计学意义($\chi^2=0.56$ 、 $t=-0.19$, $P=0.63, 0.98$)。共提取病灶 1 766 个三维纹理特征参数,其中 988 个影像组学特征存在明显组间差异,最终以 10 个最佳影像组学特征构建预测模型。以 Perceptron 分类器为最佳分类器。模型预测验证集病理分型的 AUC 为 0.95,准确率、特异度、敏感度分别为 0.88、0.87、0.84。**结论** 基于 CT 影像组学模型能有效预测肺腺癌中的 AIS 及 MIA 与 IAC。

【关键词】 肺肿瘤; 体层摄影术, X 线计算机; 影像组学; 病理学

【中图分类号】 R734.2; R814.42 【文献标识码】 A 【文章编号】 1003-3289(2020)09-1345-05

肺癌是世界上最常见的恶性肿瘤,发病率和死亡率呈逐年上升趋势^[1];随着低剂量薄层 CT 在肺癌筛查中的应用,早期非小细胞肺癌占比越来越高^[2]。治疗肺癌以手术为主要手段,术前准确区分其病理亚型有助于制定个体化手术方案,从而最大程度保留肺功能,提高手术效果^[3]。肺腺癌是肺癌组织学分类中数量最多者^[4],2011 年肺腺癌多学科分类标准^[5]将肺腺癌分为非典型腺瘤样增生(atypical adenomatous hyperplasia, AAH)、原位癌(adenocarcinoma in situ, AIS)、微浸润腺癌(minimally invasive adenocarcinoma, MIA)和浸润性腺癌(invasive adenocarcinoma, IAC)。借助计算机软件,影像组学于高通量医学影像数据中提取大量量化影像学特征,并采用机器学习方式筛选出最有价值的纹理特征,可用于评估肺腺癌病理分型^[6]。本研究探讨基于 CT 影像组学模型预测肺腺癌病理分型的价值。

1 资料与方法

1.1 一般资料 回顾性分析 2015 年 7 月—2019 年 2 月 542 例于吉林省肿瘤医院接受手术治疗的肺腺癌患者,男 257 例,女 285 例,年龄 29~88 岁,平均(61.2±9.9)岁。纳入标准:①均于术前接受同台 CT 胸部检查;②经 CT 引导下穿刺活检病理确诊为肺腺癌,并明确其病理亚型;③病灶长径 1.0~5.0 cm;④图像质量满足诊断需求。排除标准:①病理结果为非腺癌;②无法进一步确定病理分型;③图像有明显伪影。

1.2 仪器与方法 采用 GE Lightspeed 16 排 CT 机,嘱患者仰卧,双臂上举,头先进,深吸气后屏气接受扫描,参数:管电压 120 kV,管电流 215 mA,层厚 5 mm、层间距 5 mm,螺距 1.375:1,转速 27.5 mm/rot,探测器排数 16 排,重建矩阵 512×512。

1.3 提取影像组学特征与机器学习 根据病理结果将患者分为 2 组,第 1 组包含 AIS 和 MIA,第 2 组为 IAC。由 2 名具有 5 年以上胸部 CT 诊断经验的主治医师采用盲法分析图像并确定病灶边界,以 Philips_Radiomics_

v94_distribution 软件沿肿瘤边界逐层手动勾画病灶 ROI(图 1),提取病灶三维纹理特征参数;之后采用 Min-Max 标度算法进行归一化处理,方程如下:

$$X_{\text{normal}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

以 Spearman 相关分析观察组间影像组学特征差异,排除 $r < 0.2$ 或 $P > 0.05$ 者,保留组间差异明显者。应用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)算法进行降维^[7],筛选最佳影像组学特征,用于构建预测模型。

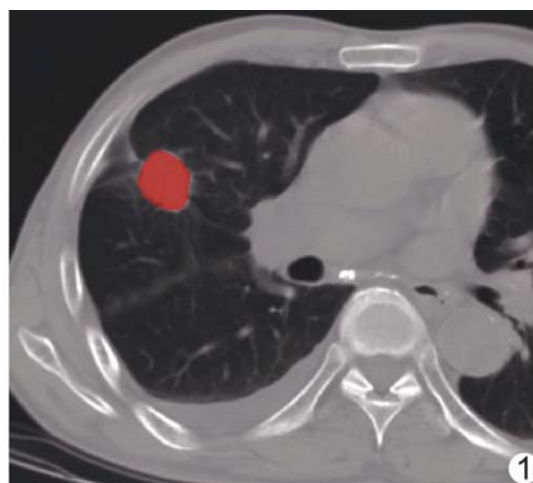


图 1 沿肿瘤边界逐层手动勾画 ROI(红色区域)示意图

按 2:1 比例将数据集分为训练集和验证集。采用 6 种机器学习模型对训练集数据进行处理,以 5 倍交叉验证结果为检测 6 种机器学习分类器(Perceptron Classifier、SGD Classifier、Logistic Regression、Passive Aggressive Classifier、Ridge Classifier、Linear SVC)效能的指标,以“准确率”为标准选择最佳分类器,之后对 5 倍交叉验证数据集、训练集和验证集进行分类,获得模型预测肺腺癌病理亚型的 ROC 曲线及相应 AUC、特异度(specificity)、敏感度(sensitivity)及准确率(accuracy)。

1.4 统计学分析 采用 SPSS 25.0 统计分析软件。计量资料以 $\bar{x} \pm s$ 表示,计数资料以频数表示。采用 χ^2 检验比较组间性别差异,以独立样本 t 检验比较组间年龄差异。 $P < 0.05$ 为差异有统计学意义。

2 结果

第 1 组 235 例,男 135 例,女 100 例,年龄 29~83 岁,平均 (60.3 ± 9.9) 岁;第 2 组 307 例,男 170 例,女 137 例,年龄 31~88 岁,平均 (61.9 ± 11.6) 岁。2 组间性别及年龄差异均无统计学意义 ($\chi^2 = 0.56$ 、 $t = -0.19$, $P = 0.63, 0.98$)。

共提取病灶 1 766 个三维纹理特征,其中 988 个组间差异明显,最终保留 10 个最佳影像组学特征以构建预测模型,见表 1。模型训练结果示 Perceptron 分类器的准确率最高(图 2),故以之为最佳分类器。采用该分类模型对 5 倍交叉验证数据集、训练集和验证集进行分类,获得模型预测 AIS、MIA 与 IAC 的 ROC 曲线(图 3),相应 AUC 分别为 0.94、0.95 和 0.95,准确率分别为 0.88、0.90 和 0.88,特异度分别为 0.81、0.82 和 0.87,敏感度分别为 0.90、0.93 和 0.84,见图 4。

表 1 最佳影像学特征

影像组学特征	系数
WaveletGLRLM-wavelet-LLH-RunEntropy	3.62
WaveletNGTDM-wavelet-HHH-Complexity	2.65
WaveletGLCM-wavelet-HHL-JointEnergy	1.93
WaveletGLCM-wavelet-HLL-InverseVariance	0.72
WaveletFirstOrder-wavelet-HHL-Range	-0.49
WaveletGLDM-wavelet-HLL-SDHGLLevelEmphasis	-1.26
WaveletGLSZM-wavelet-HLL-HGLevelZoneEmphasis	-1.35
WaveletGLRLM-wavelet-HHL-RLNUNormalized	-1.45
WaveletGLCM-wavelet-HLL-JointAverage	-1.76
WaveletGLCM-wavelet-HHL-JointEntropy	-1.86
Intercept	1.23

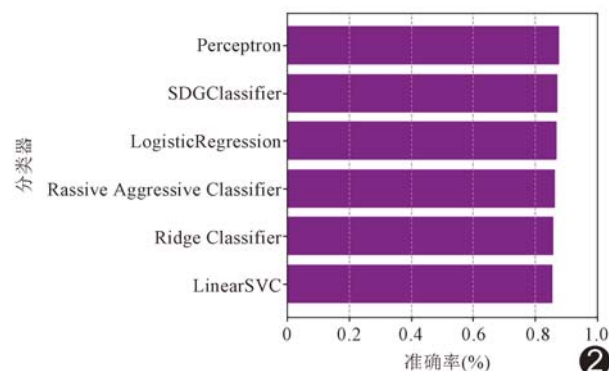


图 2 各分类器对 5 倍交叉验证数据集分类的准确率,其中 Perceptron 分类器的准确率最高

3 讨论

CT 在肺癌诊治过程中具有重要作用。随着影像组学和人工智能在医疗领域的发展,CT 不仅能提供病灶的形态学征象和密度值信息,结合影像组学后,还可提供更丰富的组织学和病理学信息。多项研究^[8-11]表明,影像组学在预测病灶性质、淋巴结转移、患者预后等方面均可取得较好效果。

影像组学通过采集图像、分割图像、提取特征和构建模型而实现鉴别诊断,已用于鉴别良恶性肺结节、肺癌分期、肺癌病理学分型、评估肺癌基因表达及预后。研究^[12]发现 AIS 和 MIA 患者 5 年无病生存率均为 100%,预后良好,而 IAC 的 5 年生存率却并不乐观。肿瘤的病理亚型除与预后密切相关外,对选择手术方式及决定是否需要辅助治疗亦有影响^[13]。本研究通过影像组学结合机器学习的方式探讨基于胸部 CT 预测肺腺癌病理亚型的效能,从临床干预的角度将可行亚肺叶切除、不需淋巴结清扫术和辅助治疗且预后良好的 AIS 和 MIA 合并为一组,将需切除肺叶、淋巴结

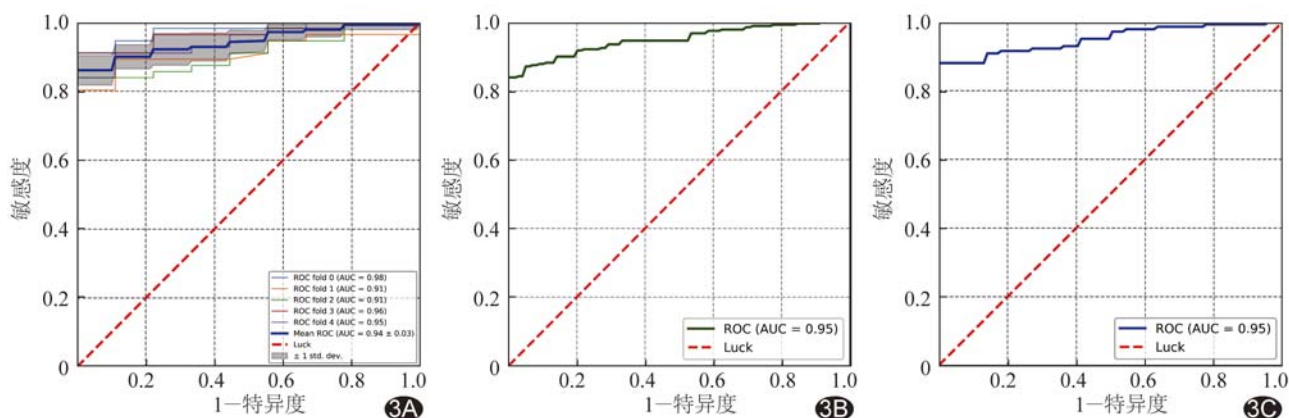


图 3 分类模型预测各数据集病理分型的 ROC 曲线 A. 5 倍交叉验证数据集; B. 训练集; C. 验证集

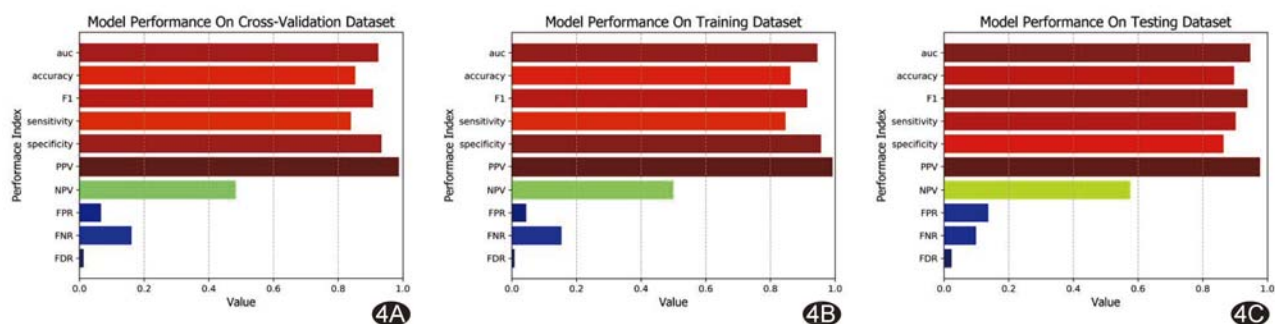


图 4 分类模型预测各数据集病理分型的效能 A. 5 倍交叉验证数据集; B. 训练集; C. 验证集

清扫或淋巴结采样且预后效果不理想的 IAC 归入另一组,筛选组间差异有统计学意义的最佳影像组学特征,实现术前区分病理亚型,以期术前制定手术策略提供更丰富的信息。

既往研究^[14-17]表明,不同病理亚型肺癌之间影像学表现存在差异。景瑞等^[4]发现不同病理类型病变密度有所差别,浸润前病变与浸润性腺癌的病变密度不同。肺结节大小、密度、血管、分叶、毛刺等影像学征象有助于区分病变病理阶段^[16-18]。而申磊磊等^[14]的研究结果显示,不同病理亚型肺癌间病灶数量和边缘毛刺征象存在差异,而实性成分、形状及反映边缘特征的分叶、支气管充气征等无明显差异;推测可能由于样本量不同或观察者主观因素导致不同研究结果之间存在差异。

多数影像组学肺癌病理亚型研究关注鉴别诊断 AIS 与 MIA 或 MIA 与 IAC,较少见从临床干预角度出发制定分组策略的鉴别诊断研究。本研究结果显示,联合 CT 影像组学与机器学习算法对 5 倍交叉验证数据集、训练集和验证集进行分析,鉴别 AIS、MIA 与 IAC 具有较好效能,有助于术前决策提供更多参考依据。

本研究的主要局限性:①纳入病例均为术后患者,可能存在选择性偏倚;②CT 扫描参数固定,导致研究结果可能缺乏拓展性;③未分析其他临床信息和影像学征象;④未采用神经网络算法等深度学习算法进行分析。

综上所述,基于 CT 影像组学模型能有效预测肺癌中的 AIS 和 MIA 与 IAC,可为临床制定治疗方案提供参考依据。

[参考文献]

[1] 曾庆鹏,赵峻.基于国际新分类肺癌亚型的临床研究进展[J].

肿瘤学杂志,2019,25(5):387-393.

- [2] NATIONAL LUNG SCREENING TRIAL RESEARCH TEAM, CHURCH T R, BLACK W C, et al. Results of initial low-dose computed tomographic screening for lung cancer [J]. N Engl J Med, 2013,368(21):1980-1991.
- [3] 赵扬,耿峻峰,澹台冀徽,等.术前使用 CT 引导下 Hook-wire 定位肺部小结节的临床价值及肺小结节恶性病变的危险因素分析[J].中国医学前沿杂志(电子版),2016,8(12):130-134.
- [4] 景瑞,金鑫,高洁,等.不同密度早期肺癌的影像特征与病理学分类[J].医学影像学杂志,2019,29(6):937-940,952.
- [5] TRAVIS W D, BRAMBILLA E, NOGUCHI M, et al. International association for the study of lung cancer/American thoracic society/European respiratory society: International multidisciplinary classification of lung adenocarcinoma [J]. J Thorac Oncol, 2011,6(2):244-285.
- [6] LAMBIN P, RIOS-Velazquez E, LEIJENAR R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis[J]. Eur J Cancer, 2012,48(4):441-446.
- [7] TIBSHIRANI R. Regression shrinkage and selection via the lasso [J]. J R Stat Soc B, 1996,58(1):267-288.
- [8] 杨晓东,刘屹,郭妍,等.CT 影像组学鉴别最大径 ≥ 1 cm 良恶性肺囊性肉[J].中国医学影像技术,2019,35(12):1842-1846.
- [9] 李武超,陈琦,蒋仪,等.基于 CT 影像组学术前预测胃癌淋巴管侵犯[J].中国医学影像技术,2019,35(7):1057-1060.
- [10] 李齐英,张广风,王欣如,等.影像组学在胃癌中的研究进展[J].中国医学影像技术,2019,35(2):286-289.
- [11] 肖磊,顾潜彪,张堃,等.CT 影像组学标签预测肺腺癌表皮生长因子受体基因敏感突变[J].中国介入影像与治疗学,2019,16(4):220-224.
- [12] 李娜,赵珩,张杰,等.2056 例手术切除肺腺癌的临床病理分析[J].中华胸心血管外科杂志,2014,30(12):715-718.
- [13] KEENAN R J, LANDRENEAU R J, MALEY R H, et al. Segmental resection spares pulmonary function in patients with stage I lung cancer [J]. Ann Thorac Surg, 2004,78(1):228-233.
- [14] 申磊磊,林吉兴,王柏霖,等.肺原位腺癌和微浸润性腺癌的影像学表现与临床病理学、分子基因特征及预后[J].南方医科大学学报,2019,39(9):1107-1112.

- [15] 彭燕,何杰,尹玉,等.肺微浸润性腺癌影像学特征及临床病理分析[J].安徽医学,2017,38(7):840-843.
- [16] 陈耿春,李阳,邱士军.肺部小结节与浸润性肺癌的影像学相关性研究[J].现代诊断与治疗,2018,29(24):4010-4012.
- [17] 左玉强,冯平勇,孟庆春,等.肺纯磨玻璃结节微浸润腺癌与浸润性腺癌的 CT 鉴别诊断[J].临床放射学杂志,2017,36(4):495-498.
- [18] 王萍.不同病理阶段肺腺癌性磨玻璃结节 PET/CT 表现差异对比[J].影像研究与医学应用,2019,3(14):125-126.

《中国介入影像与治疗学》杂志 2021 年征订启事

《中国介入影像与治疗学》杂志创刊于 2004 年,是由中国科学院主管,中国科学院声学研究所主办的国家级学术期刊,主编为邹英华教授、肖越勇教授。刊号:ISSN 1672-8475, CN 11-5213/R。为中国精品科技期刊(第 2 届)、中国科技论文统计源期刊(中国科技核心期刊)、中国科学引文数据库核心期刊、《中文核心期刊要目总览》收录期刊、WHO《西太平洋地区医学索引》(WPRIM)来源期刊、荷兰《医学文摘》收录源期刊、俄罗斯《文摘杂志》收录源期刊。

《中国介入影像与治疗学》杂志以报道介入影像与治疗学、介入超声学、介入材料学、药物学与护理学等方面的临床研究、基础研究以及医、理、工结合的成果与新进展为主,是介入影像、介入治疗专业人员学习、交流的平台。

《中国介入影像与治疗学》为月刊,64 页,大 16 开,彩色印刷。单价:20 元,全年定价 240 元。订户可随时向当地邮局订阅,邮发代号:80-220;亦可向编辑部直接订阅,免邮寄费(银行、支付宝转账均可,附言栏请注明订阅杂志名称)。

网 址 www.cjiit.com

编辑部地址 北京市海淀区北四环西路 21 号,中科院声学所大猷楼 502 室 邮编 100190

联系人 杜老师 联系电话 010-82547903 传真 010-82547903

银行账户名 《中国医学影像技术》期刊社

开 户 行 招商银行北京分行清华园支行

账 号 110907929010201

支付宝账号 cjmit@mail.ioa.ac.cn(账户名同银行账户名)

