

论 著

基于CT图像影像组学模型对甲状腺结节良恶性预测的研究

杨 鹏 武志峰*

山西白求恩医院(山西医学科学院)放射科
(山西太原 030012)

【摘要】目的 本研究旨在探讨基于CT影像组学模型对甲状腺结节良恶性的鉴别价值。**方法** 回顾性收集甲状腺结节患者185例患者的胸部CT纵膈窗影像资料,共186枚结节,良性结节104枚,恶性结节82枚。将显示病灶的薄层CT图像以DICOM格式上传至达尔文医准科研平台,在病灶的中心层面及其上下层面勾画感兴趣区(ROI),通过最小、最大值归一化预处理,对每一维度特征的线性进行拉伸。然后采用Lasso Glogistic 回归模型进行模型选择,对特征的重要性进行评估。最后以筛选出的特征维度为特征参数构建 Lasso Glogistic 回归模型,使用受试者工作特性曲线(ROC曲线)分析模型的诊断效能,并计算敏感度、特异度和诊断符合率。**结果** 经过特征筛选,8个影像组学特征用于构建甲状腺结节良恶性鉴别模型。训练组中预测模型的 ROC曲线下面积(AUC)为 0.83(95%可信区间: 0.73-0.93),灵敏度和特异度分别为 88.7%、82.0%,诊断准确率 75%;测试组中 AUC 为 0.81(95%可信区间: 0.6-0.99),灵敏度和特异度分别为 88.5%、84.6%,诊断准确率 76.13%。**结论** 基于胸部CT纵膈窗的影像组学模型对于甲状腺结节的良恶性预测具有较高的诊断效能。

【关键词】 影像组学; 甲状腺结节; 体层摄影技术; 诊断; 鉴别

【中图分类号】 R322.5+1

【文献标识码】 A

DOI:10.3969/j.issn.1672-5131.2023.01.016

Study on the Malignant Prediction of Thyroid Nodules Based on CT Image Imaging Method

YANG peng, WU Zhi-feng*

Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Taiyuan 030012, Shanxi Province, China

ABSTRACT

Objective This study aims to explore the identification value of the viciousness of thyroid nodules based on CT imaging method. **Methods** Retrospectively collects image information of 185 patients with thyroid nodules, a total of 186 nodules, 104 benign nodules, and 82 vicious nodules. The thin layer CT image of the lesion is uploaded to the DICOM format to the Darwin Medical Research Platform, and the center level of the lesion and the upper and lower levels are hidden in the upper and lower levels. The minimum, maximum is normalized, for each The linearity of dimensional characteristics is stretched. Then use the Lasso Glogistic regression model to perform model selection, and evaluate the importance of features. Finally, the Lasso Glogistic regression model is constructed with the characteristic parameters of the feature dimension, and the diagnostic performance of the model is used to analyze the model of the model, and the sensitivity, specificity, and diagnostic compliance rate are calculated. **Results** Extra characteristic screening, 8 imaging groups are used to construct a thyroid nodule and malignant identification model. The ROC curve of the predicted model in the training group (AUC) is 0.83 (95% trusted section: 0.73-0.93), the sensitivity and specificity are 88.7%, 82.0%, and the diagnostic accuracy is 75%; the AUC in the test group is 0.81 (95% Credit Zone: 0.6-0.99), sensitivity and specificity are 88.5%, 84.6%, and the diagnostic accuracy is 76.13%. **Conclusion** The imaging method based on CT scan has a high degree of diagnosis performance for the good and malignant prediction of thyroid nodules.

Keywords: Imaging Group; Thyroid Nodule; Body Layer Photography; X-ray Computer; Diagnosis; Identification

近年来,甲状腺结节发病率呈上升趋势^[1]。甲状腺结节非常常见:可触及结节的患病率约为4%~18.8%。尸检病理检查发现的甲状腺结节患病率接近50%,判定结节良恶性对后续治疗及疾病转归至关重要^[2]。但在影像检查中,良恶性结节诊断缺乏特异性,术前误诊率高达 40%~70%^[3]。目前对于甲状腺结节的主要检查手段,首选仍然是超声,超声检查的优点有价格低、无创及实时成像等,但是超声结果的准确性与操作者的经验及手法有比较密切的关系,同时对于胸骨后病变,可能评估受限。

影像组学是近年来新兴的影像分析技术,是从传统的影像中提取高通量特征来量化肿瘤病灶,并通过深度学习的方法挖掘与病理有潜在关联的特征,用于疾病的诊断及预后评估。本研究是基于胸部CT获得的图像建立影像组学模型,以期通过胸部CT机会性筛查,提高对甲状腺结节良恶性的鉴别能力。

1 资料和方法

1.1 一般资料 回顾性搜集2020年9月到2021年12月山西白求恩医院经手术病理证实的甲状腺结节患者。

纳入标准:术前接受胸部 CT检查,且检查与手术时间间隔 1~15 d;无其他肿瘤病史;术前超声检查甲状腺影像报告和数据库(thyroid imaging-reporting and data system, TI-RADS)分级为 II~IV级;病灶直径>0.6cm;初诊患者,病灶未经穿刺及放化疗。排除标准:图像质量较差;CT检查前接受穿刺活检或其他治疗手段。

1.2 影像检查方法 采用德国 Siemens Definition 双源 64层 CT扫描仪,管电压 120 kV,管电流 250 mAs,螺距1.0,矩阵512×512,层厚1.0mm,扫描前训练患者呼吸屏气,扫描时嘱患者仰卧、颈过仰,肩部尽量外扩,双手伸过头顶,禁吞咽动作。扫描范围自下颌至肺底。

1.3 图像分析和特征提取 本研究选取患者的CT平扫纵膈窗1mm薄层图像,将显示病灶的薄层图像以DICOM的格式上传至达尔文医准科研平台,分层随机以8:2的比例分隔为训练组及测试组。首先由一位具有5年颈部疾病影像诊断经验的医师,根据胸部CT图像上甲状腺结节中心层面进行ROI的勾画,ROI应避免坏死、囊变及钙化区域,并在相邻的上一层及下一层手动勾画ROI;再由一位有10年以上颈部疾病诊断经验的主任医师确认ROI勾画结果,见图1。

采用达尔文科研平台内置的R语言算法完成特征筛选,对CT图像进行最小、最大值归一化预处理,对每一维度特征的线性进行拉伸。然后采用Lasso Glogistic 回归模型进行模型选择,对特征的重要性进行评估,通过这些模型给出的重要性分数,筛选出低于阈值的特征。最后以筛选出的特征维度为特征参数构建 Lasso Glogistic 回归模型,模型计算

【第一作者】 杨 鹏,男,主治医师,主要研究方向:胸腹部影像诊断。E-mail: yangpeng504@126.com

【通讯作者】 武志峰,男,主任医师,主要研究方向:胸腹部影像诊断。E-mail: wuzhifeng2004@126.com

公式见表1, 采用ROC曲线评价模型的诊断效能, 计算敏感度、特异度和诊断符合率。

2 结果

2.1 基本资料 回顾性收集2020年9月到2021年12月山西白求恩医院经手术病理证实的甲状腺结节患者。共收集甲状腺结节患者185例, 根据2017版WHO甲状腺肿瘤分类简表^[4], 排除23例病理类型为滤泡性肿瘤、恶性潜能未定的滤泡性肿瘤(FT-UMP)、恶性潜能未定的高分化肿瘤(WT-UMP)等交界性或不确定性肿瘤的病例。共纳入162例患者, 其中男36例、女126例, 年龄25~74 (49±11)岁; 共186枚结节, 其中良性结节104枚, 包括结节性甲状腺肿102枚, 滤泡性腺瘤1枚, 嗜酸细胞腺瘤1枚; 恶性结节82枚, 其中乳头状癌75枚, 滤泡性癌3枚, 未分化癌1枚, 肉瘤1枚, 淋巴瘤1枚, 低分化鳞状细胞癌1枚。

2.2 特征降维及模型选择 运用达尔文平台(Darwin research platform)软件内置的特征降维组件进行降维。共提取1316个特征, 首先经过最大绝对值归一化和最小最大值归一化, 数据预处理

组件用于对原始的特征向量进行拉伸、标准化或者归一化处理, 将每一维度线性拉伸到(0, 1)区间内, 共获得34个特征, 采用Lasso Glogistic 回归模型进行模型选择, 对特征的重要性进行评估, 通过这些模型给出的重要性分数, 筛去低于阈值的特征, 共获得8个选择特征维度。包括, 依赖不均匀性 (dependence Non-uniformity normalized, DNN)、小依赖低灰度优势(small dependence low gray level emphasis, SDLGLE)、相关性关联措施(imc1)量化纹理的复杂性、最大概率(maximum probability, MP)、依赖方差(dependence variance)、区域大小不均匀归一化(size zone non-uniformity normalized, SZNN)、伸长(elongation)、低灰度优势(low gray level emphasis, LGLE), 最后以筛选出的8个特征维度为特征参数构建 Lasso Glogistic 回归模型, 模型公式如表1, 并进行ROC曲线分析。最终得出结论, 训练组模型的曲线下面积为0.83(95%CI: 0.73, 0.93), 灵敏度和特异度分别为 88.7%、82.0%, 诊断准确率 75%; 和测试组曲线下曲线下面积为0.81(95%CI: 0.6, 0.99), 灵敏度和特异度分别为 88.5%、84.6%, 诊断准确率 76.13%, 见图2。

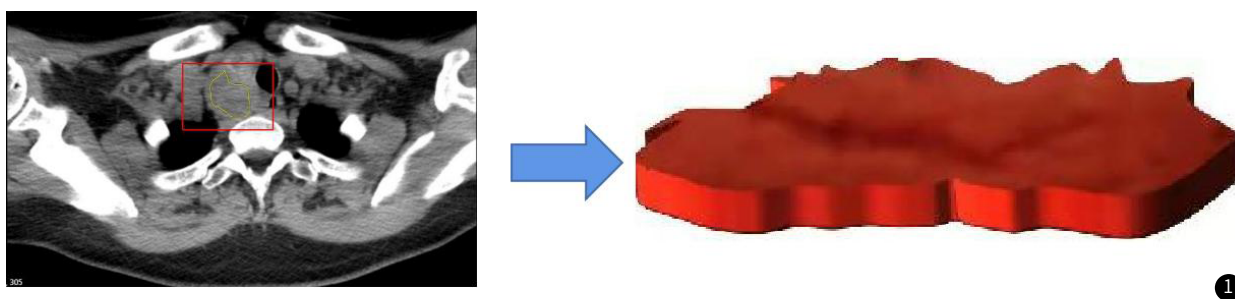


图1 男, 65岁, 甲状腺右侧叶乳头状甲状腺癌, 左图在达尔文科研平台上薄层病灶中心层面手动勾画ROI(红色实线区域), 右图为勾画感兴趣区的三维立体显示。

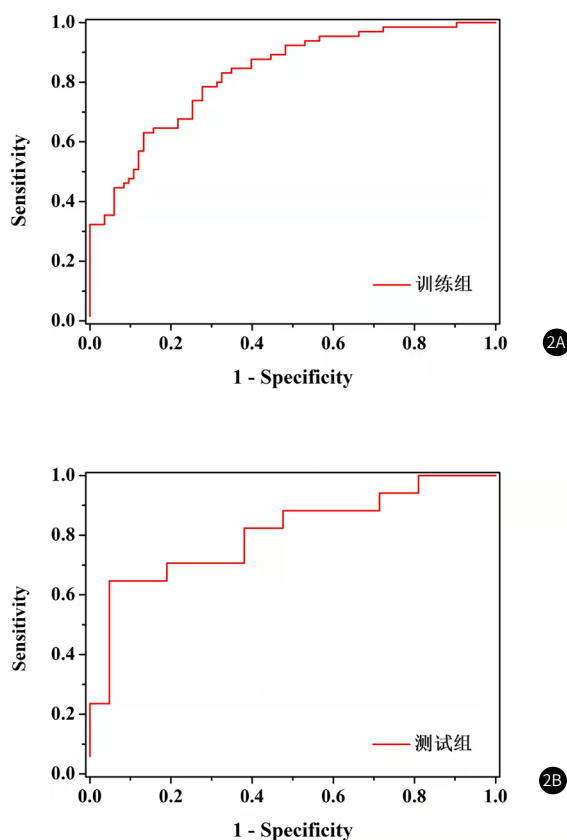


图2 图2A为训练组ROC曲线, AUC值为0.83(95%可信区间: 0.73~0.93), 图2B为测试组ROC曲线, AUC值为0.81(95%CI: 0.6, 0.99)。

表1 预测模型模型公式

模型公式= $-0.643 \times \text{pingsao_wavelet-LHH_gldm_DependenceNonUniformityNormalized} + 1.618 \times \text{pingsao_squareroot_gldm_SmallDependenceLowGrayLevelEmphasis} - 1.255 \times \text{pingsao_gradient_glcm_lmc1} + 1.190 \times \text{pingsao_wavelet-HLH_glcm_MaximumProbability} + 1.112 \times \text{pingsao_original_gldm_DependenceVariance} + 1.000 \times \text{pingsao_wavelet-LLH_glszm_SizeZoneNonUniformityNormalized} - 0.780 \times \text{pingsao_original_shape_Elongation} - 0.441 \times \text{pingsao_exponential_gldm_LowGrayLevelEmphasis} + 0.238$.

3 讨论

近年来随着低剂量 CT 的推广应用,安全剂量范围内CT检查在甲状腺疾病中的临床应用越来越多^[5]。本研究基于CT平扫图像为基础,通过影像组学模型的建立完成对甲状腺结节的良恶性的预测。虽然超声检查仍然是甲状腺结节诊断的重要检查手段,但近年来CT及MR运用于甲状腺结节良恶性的诊断及甲状腺癌淋巴结转移的研究逐渐增多。赵泓博,尹映丽等关于影像组学模型预测甲状腺乳头状癌颈部淋巴结转移的研究中,对于淋巴结转移的预测,AUC值达85%^[6]。

影像组学最早于2012年提出,作为一种新兴的影像分析方法,其目的是利用大数据分析技术,从影像数据中挖掘可量化病变的海量特征,并构建特征性影像组学标签,以期解析影像与病理、临床的潜在关联,尤其是一些高阶纹理特征,它们不仅反映了肿瘤组织学的异质性,甚至还在一定程度上反映出相应的遗传学差异^[7]。纹理特征是结构内部体素分布差异的表现,是基于特定像素的联合概率分布。恶性肿瘤内部结构往往比较复杂,主要原因是肿瘤内出现坏死、肿瘤血管的异常增殖以及细胞壁破坏所致通透性的改变,进而导致影像灰度的差异,所以其纹理特征会出现差异。上述变化肉眼不易察觉,但通过纹理特征可以发现,且不受主观影响^[8]。基于CT图像的纹理分析已被发现与病理信息相关^[9-10]。

本研究共研究基于胸部纵膈窗CT平扫图像影像组学对甲状腺结节进行评估,经预处理和特征降维,结果显示8个纹理特征均具有预测价值。其中灰度依赖矩阵特征4个、灰度共生矩阵特征2个、灰度区域大小矩阵特征1个及一阶特征1个。本研究中,纹理特征数量最多,说明肿瘤的内部结构及异质性与结节良恶性密切相关,但单个特征所占权重不大,可能与单个纹理特征仅能反映病灶的部分内部特征有关,需综合多个纹理特征才能提高其诊断效能。本研究通过 Lasso Glogistic 回归模型训练及预测甲状腺结节的良恶性,得出测试组AUC面积为0.81,有较高的诊断效能。LASSO是目前应用最多的特征筛选方法,它通过增加约束条件、构建惩罚函数的方式对模型中的系数进行压缩,一些影响较小或没有影响的自变量系数被压缩为0,仅保留系数非0的自变量。其中,由参数 λ 控制惩罚力度, λ 越小表示惩罚力度越小,容易导致过拟合,反之惩罚力度大会造成拟合不足,故而 λ 值非常重要^[11]。

影像组学的纹理分析在肿瘤影像中的应用主要是为了揭示肿瘤的临床相关特征,而目前大家对于影像组学的开发主要集中于基于纹理分析的计算机辅助诊断(CADx)系统,可以帮助区别甲状腺结节的良恶性,自动甲状腺结节检测和经典化系统的目标是实现至少可与标准细针抽吸检查相媲美的性能^[12]。本研究创新之处在于运用了胸部CT纵膈窗图像,不仅可以充分利用机会性筛查的影像数据,又能尽早发现甲状腺病变,为患者的诊断、治疗及预后起到重要作用。然而由于胸部CT扫描时双手位置的差异所导致图像质量参差

不齐的问题,仍需要解决;对于患者手臂摆放位置,仍需要不断探索,从而获得最佳图像质量;其次甲状腺结节病灶基于手工绘出,存在较大误差,随人工智能的不断发展,可通过深度学习等技术手段实现病灶的绘画,能够进一步降低误差^[13]。虽然目前影像组学正在兴起,并且赢得广大影像工作者的追捧,但由于其尚处于起步阶段,仍然不太成熟,可重复性较差,多中心数据采集与单中心数据相比,其提取的特征仍存在一定差异^[14]。

参考文献

- [1] Bomeli S R, LeBeau S O, Ferris R L. Evaluation of a thyroid nodule[J]. *Otolaryngol Clin North Am*, 2010, 43 (2): 229-238.
- [2] 陈玉恒, 郭兰伟, 张玥, 等. 中国2008年甲状腺发病、死亡和患病情况的估计和预测[J]. *中华疾病控制杂志*, 2014, 18 (3): 200-203.
- [3] 谢榜昆, 关玉宝, 袁小平, 等. 甲状腺癌的CT表现与病理相关性研究[J]. *癌症*, 2003, 22 (2): 192-197.
- [4] Lloyd R V, Osamura R Y, Klöppel G, et al. WHO classification of tumours: pathology and genetics of tumours of endocrine organs [M]. 4th ed. Lyon: IARC Press, 2017.
- [5] 中华医学会放射学分会头颈学组. 甲状腺结节影像检查流程专家共识[J]. *中华放射学杂志*, 2016, 50 (12): 911-915.
- [6] 赵泓博, 尹映丽, 刘畅, 等. 基于CT双期增强影像组学预测甲状腺乳头状癌淋巴结转移[J]. *放射学实践*, 2021, 36 (4): 458-463.
- [7] 张利文, 方梦捷, 臧亚丽, 等. 影像组学的发展与应用[J]. *中华放射学杂志*, 2017, 51 (1): 75-77.
- [8] Sun K, Huang S H, Wong DS, et al. Design and application of a variable selection method for multilayer perceptron neural network with LASSO[J]. *IEEE Trans Neural Netw Learn Syst*, 2017, 28 (6): 1386-1396.
- [9] Fave X, Zhang L, Yang J, et al. Impact of image preprocessing on the volume dependence and prognostic potential of radiomics features in non-small cell lung cancer[J]. *Transl Cancer Res*, 2016, 5: 349-363.
- [10] Basu S, Hall L O, Goldgof D B, et al. Developing a classifier model for lung tumors in CT-scan images. In: *Proceedings of systems, man, and cybernetics* [J]. Piscataway, NJ: IEEE, 011: 1306-1312.
- [11] Prediction of Immunohistochemistry of Suspected Thyroid Nodules by Use of Machine Learning-Based Radiomics[J]. *AJR Am J Roentgenol*, 2019 Dec, 213 (6): 1348-1357.
- [12] Parekh V, Jacobs M A. Radiomics: a new application from established techniques[J]. *Expert Rev Precis Med Drug Dev*, 2016, 1 (2): 207-226.
- [13] 孟玉柱, 潘昌杰, 等. 基于CT双期增强图像的深度学习模型对甲状腺良恶性结节的分类研究[J]. *放射学实践*, 2021, 36 (8): 971-975.
- [14] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. *Eur J Canc*, 2012, 48 (4): 441-446.

(收稿日期: 2022-09-17)

(校对编辑: 阮 靖)