分类号：R445单位代码：10159

密 级：公开学 号：2018120410

**硕士学位论文**

|  |  |  |
| --- | --- | --- |
| 中文题目： | 放射组学对鉴别纵膈淋巴结良恶性的研究 |  |
| 英文题目： | Preoperatively estimating the malignant potential of mediastinal lymph nodes: A pilot study towards establishing a robust radiomics model based on contrast-enhanced CT imaging |  |

论文作者： 董梦实

指导教师： 徐 克 教授

学科专业： 影像医学与核医学

完成时间： 2021年3月

中国医科大学硕士学位论文

放射组学对鉴别纵膈淋巴结良恶性的研究

Preoperatively estimating the malignant potential of mediastinal lymph nodes: A pilot study towards establishing a robust radiomics model based on contrast-enhanced CT imaging

论文作者 董梦实 　指导教师 徐克 教授

申请学位 医学硕士 　培养单位 第一临床学院

一级学科 临床医学 二级学科 影像医学与核医学

研究方向 放射组学

论文起止时间 20？？年？月—2021年3月

论文完成时间 2021年3月

中国医科大学（辽宁）

2021年3月

中国医科大学学位论文独创性声明

本人郑重声明：本论文是我个人在导师指导下独立进行的研究工作及取得的研究成果，论文中除加以标注的内容外，不包含其他人或机构已经发表或撰写过的研究成果，也不包含本人为获得其他学位而使用过的成果。对本研究提供贡献的其他个人和集体均已在文中进行了明确的说明并表示谢意。本人完全意识到本声明的法律结果由本人承担。

论文作者签名：

日 期： 年 月 日

中国医科大学学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交学位论文的原件、复印件和电子版，允许学位论文被查阅和借阅。本人授权中国医科大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编学位论文。

保密（ ），在 年后解密适用本授权书。（保密：请在括号内划“√”）

论文作者签名： 指导教师签名：

日期： 年 月 日 日期： 年 月 日

**摘 要**

**目的：**纵隔淋巴结为恶性肿瘤发生淋巴结转移的常见部位之一。且明确淋巴结转移情况在肿瘤分期、治疗方案的选择以及预后预测等方面起到了至关重要的作用。因此术前评估淋巴结的良恶性情况尤为重要。本研究拟建立并验证基于计算机断层扫描（Computed Tomography，CT）增强成像的纵隔淋巴结恶性程度的放射组学模型。

**方法：**本研究为回顾性研究，共纳入了129名被试的201个淋巴结，其中81个淋巴结为良性，120个淋巴结为恶性。所有入组的淋巴结均有病理诊断结果。201个淋巴结以7:3的比例被分为训练组和测试组。本研究选则在静脉期CT图像上进行感兴趣区（region of interest，ROI）的勾画。勾画ROI时，选取增大淋巴结的轴位最大层面，沿着淋巴结边缘进行手动绘制。随后对ROI进行放射组学特征的提取。继而使用最小绝对收缩和选择算子（least absolute shrinkage and selection operator，LASSO）二元逻辑回归的方法对提取的放射组学特征进行选择。在应用多元逻辑回归进行特征筛选时，采用了后退逐步淘汰法。最终本研究利用筛选出来的放射组学特征，构建了一个用来判别纵隔淋巴结良恶性的逻辑回归模型。本研究评估了该诊断模型的分类性能，并将其与利用其他5种机器学习算法（支持向量机算法、朴素贝叶斯算法、随机森林算法、决策树算法、K最邻近算法）构建的诊断模型的接收者操作特征曲线(receiver operating characteristic curve，ROC曲线)进行了比较，以对比这几种方法的分类性能。本研究使用校准曲线以及Hosmer-Lemeshow检验来评估逻辑回归模型的校准度。并且分别在训练组和测试组内进行了决策曲线分析（decision curve analysis，DCA），以评估逻辑回归模型的临床应用价值。为了明确不同品牌的CT扫描仪或不同层厚的CT扫描图像对诊断模型的分类性能是否有影响，本研究还进行了分层分析。

**结果：**在6种不同的机器学习算法构建的模型中，利用筛选出的8个放射组学特征训练出的逻辑回归模型在鉴别纵隔淋巴结良恶性方面得到了较好的诊断性能。在训练组/测试组中，其精确性、敏感性、特异性以及ROC曲线下的面积（area under the ROC curve，AUC）分别为0.850/0.803，0.821/0.806，0.893/0.800以及0.922/0.850。Hosmer-Lemeshow检验结果显示P值大于0.05，这一结果证明诊断模型拥有良好的校准度；且校准曲线表明诊断模型的分类结果与淋巴结的真实病理结果之间具有较好的一致性。DCA结果表明，当测试组的阈值概率在30%～90%之间时，放射组学模型将获得更大的收益。分层分析显示不同品牌的CT扫描仪和不同层厚的CT扫描图像对诊断模型的分类性能没有影响。DeLong检验结果得出任意两个亚组之间的P值均大于0.05，这说明任意两个亚组的诊断性能之间均无统计学差异；这一结果表明了放射组学评分在不同因素之间的泛化能力。

**结论：**本研究构建的基于增强CT图像的放射组学模型有助于对纵隔淋巴结的良恶性进行术前评估，并且该模型在不同品牌的CT扫描仪和不同层厚的CT扫描图像中具有较好的稳定性。

**关键词：**纵隔淋巴结；良性；恶性；放射组学；机器学习；分类

**Abstract**

**Purpose:** To establish and validate a radiomics model to estimate the malignancy of mediastinal lymph nodes (LNs) based on contrast-enhanced CT imaging.

**Method:** In total, 201 pathologically confirmed mediastinal LNs (81 benign, 120 malignant) from 129 patients were retrospectively enrolled and assigned to training and test sets (7:3 ratio). Radiomics features were extracted from the region of interest (ROI) delineated on contrast-enhanced CT imaging. The ROI was manually drawn along the boundary of the LN on the maximal 2D axial slice. Feature selection was performed with least absolute shrinkage and selection operator (LASSO) binary logistic regression. Multivariate logistic regression was performed with the backward stepwise elimination method. A logistic regression model was fitted to associate mediastinal LN malignancy with the selected features. The diagnostic performance of the model was assessed and compared to that of five other machine learning algorithms (support vector machine, naive Bayes, random forest, decision tree, K-nearest neighbor) using receiver operating characteristic (ROC) curves. Calibration curves and Hosmer-Lemeshow tests were used to assess the calibration degree. Decision curve analysis (DCA) was used to assess the clinical usefulness of the logistic regression model in both the training and test sets. Stratified analysis was performed for different CT scanners and slice thicknesses.

**Result:** Among the six machine learning methods, the logistic regression model with the eight strongest features showed a significant association with mediastinal LN status and the satisfactory diagnostic performance for distinguishing malignant LNs from benign LNs. The accuracy, sensitivity, specificity and area under the ROC curve (AUC) were 0.850/0.803, 0.821/0.806, 0.893/0.800 and 0.922/0.850 in the training/test sets, respectively. The Hosmer-Lemeshow test showed that the P value was > 0.05, indicating good calibration, and the calibration curves of the model showed good agreement between the classifications and actual observations. DCA showed that the radiomics model would obtain more benefit when the threshold probability was between 30% and 90% in the test set. Stratified analysis showed that the classification performance was not affected by different scanners or slice thicknesses. There was no significant difference (DeLong test, P > 0.05) between any two subgroups, which showed the generalization of the radiomics score across different factors.

**Conclusion:** The radiomics model we built could help assist the preoperative estimation of mediastinal LN malignancy based on contrast-enhanced CT imaging, with stability for different scanners and slice thicknesses.

**Key words**: Mediastinal lymph nodes, Malignant, Benign, Radiomics, Machine learning, Classification

英文缩略语

|  |  |  |
| --- | --- | --- |
| **英文缩写** | **英文全称** | **中文全称** |
| EBUS-TBNA  CT  MR  PET | Endobronchial ultrasound-guided transbronchial needle aspiration  Computed tomography  Magnetic resonance  Positron Emission Tomography | 超声内镜引导下的经支气管针吸活检  计算机断层扫描  磁共振  正电子发射计算机断层扫描 |

目 录

1 前言

1.1 研究背景及意义

纵隔淋巴结为恶性肿瘤（如肺癌、食管癌等）发生淋巴结转移的最常见的部位之一。明确恶性肿瘤患者的淋巴结转移情况对于肿瘤分期、治疗方案的选择以及对于患者预后的预测起到了至关重要的作用[1-4]。因此，对纵隔淋巴结的良恶性进行术前评估就变得尤为重要[5]。术前明确纵隔淋巴结的良恶性，有助于临床医生全面的了解患者的状况，从而做出更适合患者的诊断和治疗决策。然而，在临床工作中，术前评估恶性肿瘤患者的淋巴结转移情况尚且存在一定的难度。但恶性肿瘤发生纵隔淋巴结转移时，部分病例不会出现淋巴结肿大。有时，淋巴结最大直径小于1cm，但最终病理确诊为淋巴结转移；而有的淋巴结明显肿大，最终的病理结果却显示为淋巴结的慢性炎症或淋巴结的反应性增生。

超声内镜引导下的经支气管针吸活检（Endobronchial ultrasound-guided transbronchial needle aspiration，EBUS-TBNA）是评估纵隔淋巴结良恶性的最有效的诊断方法[6-9]。然而，EBUS-TBNA在发展中国家的普及率较差。而且EBUS-TBNA作为一种侵入性的诊断手段，可能会导致多种并发症[10-12]，并且可能存在假阴性的结果[13-16]。寻找到一种能够帮助临床医生在术前鉴别纵隔淋巴结良恶性的方法是亟待解决的问题。

1.2 研究现状

众所周知，计算机断层扫描（Computed Tomography，CT）是一种可以在术前进行的检查方法。近年来的一些研究人员指出放射组学在位于纵隔的病变的术前分期中表现出了很好的效果[5]。“放射组学”是一种采用高级的特征分析技术高通量的从数字医学图像中--包括CT、磁共振（Magnetic Resonance，MR）以及正电子发射计算机断层扫描（Positron Emission Tomography，PET）图像--提取大量的定量特征的方法[17]。多个研究证明在对肿瘤进行诊断、疗效评估以及预后预测时，放射组学是一种非常有效的方法[18-20]；除此之外，放射组学在肿瘤表型解码方面也呈现出了良好的效果[18, 21]。近期有证据表明放射组学特征能够对肺腺癌的远处转移[22]以及结直肠癌的淋巴结转移[23]进行较为准确的预测。数位研究人员尝试利用放射组学来评估原发性肺癌患者的纵隔淋巴结的良恶性[22, 24]；也有研究人员尝试使用纵隔淋巴结的纹理特征来鉴别其良恶性[25, 26],这些研究的结果都表明放射组学具有鉴别纵隔淋巴结良恶性的能力。一项近期的研究显示，基于CT图像的放射组学模型在诊断非小细胞肺癌患者的纵隔淋巴结转移方面得到了非常好的结果[27]。然而，在不同的研究中，其所采用的扫描设备和扫描方案不同，诊断模型的可重复性和适用性相对较低。

本研究拟基于纵隔淋巴结的CT增强扫描图像，采用放射组学的分析方法，构建了一个鉴别纵隔淋巴结良恶性的机器学习诊断模型；并且也将会对该模型进行分层分析[28]，从而验证其鲁棒性。我们假设这一模型可以有效的鉴别纵隔淋巴结的良恶性，并且具有很好的鲁棒性；可以广泛的应用在不同品牌的CT扫描仪和不同层厚的CT扫描图像中，且仍然可以得到稳定的分类结果。

2 材料和方法

2.1 研究对象

本研究为一项回顾性的研究；得到了中国医科大学附属第一医院机构评审委员会的批准，并免除了知情同意的要求。根据中国医科大学附属第一医院机构伦理审查委员会的指导方针，对病历进行了审查。本研究共纳入了2014年7月至2018年7月之间的129名患者，其中分别有73名男性和56名女性，年龄范围为18岁至93岁之间，平均年龄为56岁，中位年龄为57岁；且以上患者均具有纵隔淋巴结的EBUS-TBNA的病理诊断结果和胸部CT增强扫描图像。入组标准为：1.患者具有纵隔淋巴结的EBUS-TBNA的病理诊断结果；2.胸部CT增强扫描的扫描时间和纵隔淋巴结的EBUS-TBNA检查时间之间的时间间隔不超过2个星期。排除标准为：1.患者患有原发性恶性肿瘤，而纵隔淋巴结的EBUS-TBNA的病理诊断结果为阴性；2.纵隔淋巴结病灶难以勾画；3.具有EBUS-TBNA的病理诊断结果的淋巴结在胸部CT增强扫描图像中定位困难。具有EBUS-TBNA的病理诊断结果的淋巴结大部分位于4R组、4L组或7组（4R组，右侧下气管旁淋巴结；4L组，左侧下气管旁淋巴结；7组，气管隆突下淋巴结）。淋巴结的良恶性分类以EBUS-TBNA的病理诊断结果作为金标准。

所有的淋巴结（共有201个淋巴结）以7：3的比例被随机分入2个亚组[29-31]。70%的患者（共140个淋巴结）以分层抽样的方法被分入训练组，其中包括84个恶性淋巴结（60%，84/140）和56个良性淋巴结（40%，56/140）。余下的30%（共61个淋巴结）被分入了测试组，其中包括36个恶性淋巴结（59%，36/61）和25个良性淋巴结（41%，25/61）。

患者入组和数据分配的流程图如图1所示。



图1. 筛选患者和数据分配流程图

图中n代表患者数，N代表纵隔淋巴结数。

2.2 CT图像获取

本研究的流程如图2所示。患者所做的胸部CT增强扫描共来自于4种品牌的CT扫描仪，分别为Siemens (SOMATOM Force, 192-section dual-source CT), GE (Discovery CT 750 HD, 64 multidetector CT system), Toshiba (Aquilion One, 320 slice)以及Philips (Brilliance iCT, 256 slice)。进行胸部CT增强扫描时，患者呈仰卧位，双臂上举，深吸气后屏住呼吸。扫描参数如下：管电压120 kVp，管电流100-200 mA，扫描野350 mm×350 mm，矩阵大小512×512。

以2.5−3.0 mL/s的速度静脉注射70−90 mL非离子造影剂（Iproamine）和35-45mL生理盐水之后，延迟30−40 s后进行静脉期对比增强CT扫描。而后从图片存档和通信系统（Picture Archiving and Communication System，PACS）（IMPAX，AGFA，Belgium）检索并获取胸部CT增强扫描图像。



图2. 实验流程图

2.3 放射组学特征提取

从PACS中获取胸部CT增强扫描图像，并将其导入软件ITK-SNAP (version 3.8.0; http://www.itksnap.org)中。由一名拥有10年诊断经验的胸组放射科医师（阅片者1）在不知道病理诊断结果的情况下进行阅片，阅片时使用的是软件ITK-SNAP,选取增强扫描的静脉期图像，将胸片调整为纵隔窗（窗宽：350，窗位：50），并勾画出感兴趣区（region of interest，ROI）。ROI是沿着纵隔淋巴结的边沿，在淋巴结的轴位最大层面上进行手动勾画的。

在阅片者1初次勾画ROI的6个月后，再次随机选择了50个淋巴结，由另一名拥有7年诊断经验的胸组放射科医师（阅片者2）同样在不知道病理结果的情况下对这50个淋巴结重新进行第二次勾画ROI；目的是为了评估提取出来的放射组学特征在阅片者间的一致性。本研究进行了组内相关系数（intraclass correlation coefficient，ICC）的分析。

原始的DICOM（digital imaging and communications in medicine）图像和相对应的ROI图像都被导入到了软件A.K. (Artificial Intelligence Kit, GE Healthcare, China)中，以图像生物标志物标准化倡议（Image Biomarker Standardization Initiative，IBSI）为依据，对定量的放射组学特征进行自动化的提取[32]。

2.4 放射组学模型构建

为了减少过拟合，在进行放射组学模型的构建之前，本研究在训练组中采用了10折交叉验证的最小绝对收缩和选择算子（least absolute shrinkage and selection operator，LASSO）逻辑回归分析的方法，并重复了100次，以筛选可以用来鉴别纵隔淋巴结良恶性的相关系数不为0的特征。本研究利用后退逐步淘汰法的多元逻辑回归的方式来筛选与纵隔淋巴结恶性概率相关的相关性最强的特征。

将筛选出来的相关性最强的特征分别按其系数进行加权后，再进行线性组合，就可以算出每个淋巴结的放射组学评分。为了比较不同的机器学习算法在鉴别纵隔淋巴结良恶性方面的分类性能，本研究又构建了其它5种机器学习模型，包括支持向量机算法（support vector machine，SVM）、朴素贝叶斯算法（naive Bayes，NB）、随机森林算法（random forest，RF）、决策树算法（decision tree，DT）、K最邻近算法（K-nearest neighbors，KNN）。

2.5 放射组学模型评估

本研究通过训练组以及测试组的接收者操作特征曲线(receiver operating characteristic curve，ROC曲线)，并计算诊断模型的精确性，来评估不同机器学习模型鉴别纵隔淋巴结良恶性的诊断性能。在训练组和测试组内，通过辨别度、校准度和临床应用三个方面来评估本研究所构建的放射组学评分的诊断性能。

2.5.1 辨别度

辨别度，即衡量模型区分纵隔淋巴结良恶性的能力。本研究利用曲线下面积（are under the ROC curve，AUC）来评估诊断模型的辨别度。通过训练组计算出来的放射组学模型的最佳阈值点被应用到了测试组中。

2.5.2 校准度

校准度，即衡量实际恶性病理结果的频率和恶性分类结果的概率之间的一致性。本研究通过绘制校准曲线来评估诊断模型的校准的。并通过Hosmer-Lemeshow检验来确定诊断模型的拟合优度，当P值大于0.05时，被认为是校准效果较好。

2.5.3 临床应用

为了评估诊断模型的临床应用潜能，通过量化不同阈值下的净收益来进行决策曲线分析（decision curve analysis，DCA）。

本研究使用与上述相同的机器学习过程来分析胸部CT平扫图像。该过程包括勾画ROI、放射组学特征提取、放射组学特征构建以及放射组学特征评估。

2.6 分层分析

本研究在不同品牌的CT扫描仪或不同层厚的CT扫描图像的亚组之间进行了分层分析。基于不同品牌的CT扫描仪，201个淋巴结被分成了4个亚组：GE组（共计24个淋巴结，其中有17个恶性淋巴结、7个良性淋巴结）、Philips组（共计91个淋巴结，其中有43个恶性淋巴结、48个良性淋巴结）、Siemens组（共计14个淋巴结，其中有13个恶性淋巴结、1个良性淋巴结）以及Toshiba组（共计72个淋巴结，其中有47个恶性淋巴结、25个良性淋巴结）。基于不同层厚的CT扫描图像，201个淋巴结被分成了4个亚组：1 mm组（共计14个淋巴结，其中有7个恶性淋巴结、7个良性淋巴结）、2 mm组（共计10个淋巴结，其中有8个恶性淋巴结、2个良性淋巴结）、5 mm组（共计82个淋巴结，其中有40个恶性淋巴结、42个良性淋巴结）以及8 mm组（共计95个淋巴结，其中有30个恶性淋巴结、65个良性淋巴结）。本研究使用DeLong检验来比较所构建的放射组学评分在各亚组间的AUC值的差异；当P值大于0.05时，表明放射组学模型在各亚组间的诊断性能无统计学差异。

2.7 统计分析

本研究使用软件R Studio（version 1.2.1335）来完成所有的统计分析。利用 “glmnet” 软件包完成LASSO回归的步骤。利用“pROC”软件包进行ROC曲线的绘制。利用“DecisionCurve”功能来完成DCA的步骤。当P值小于0.05（双尾）时，被认为具有统计学差异。根据数据的分布类型，采用独立样本t检验来评估训练组和测试组中的恶性亚组和良性亚组的临床因素和恶性相关特征方面的差异。利用卡方检验来比较分类变量之间差异的显著性。采用DeLong检验来比较不同诊断模型的AUC值之间的差异。

3 结果

3.1 病理结果和临床资料

129个患者的201个淋巴结的病理结果和临床资料可见于表1和表2。训练组和测试组内的良性淋巴结和恶性淋巴结的分布没有统计学差异（P值大于0.05）。在训练组内部，良性组和恶性组的性别和年龄的分布没有统计学差异（P值大于0.05）。在测试组内部，恶性组的年龄与良性组相比较大（P值等于0.001），且男性的比例更多（P值小于0.001）。

表1. 训练组和测试组的人口统计学资料

|  |  |  |  |
| --- | --- | --- | --- |
|  | 训练组 | 测试组 | P值 |
| 年龄 (m, [Q1, Q3]) | 56.00 (50.00, 63.00) | 55.00 (47.70, 61.30) | 0.276 |
| 性别 (No, [%]) |  |  | 0.744 |
| 男 | 74 (52.86) | 30(49.18) |  |
| 女 | 66 (47.14) | 31(50.82) |  |
| 吸烟史 |  |  | 0.557 |
| 从不吸烟 | 74 | 34 |  |
| 既往吸烟 | 11 | 8 |  |
| 吸烟 | 35 | 12 |  |
| 未知 | 20 | 7 |  |

表中m代表中位年龄，Q1、Q3分别代表第1、3百分位数，No代表患者人数。

表2**.** 训练组和测试组内的良性亚组和恶性亚组的人口统计学资料

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 特征 | 训练组 | | P值 | | 测试组 | | | P值 | |
| 良性 | 恶性 | 良性 | 恶性 | |
| 年龄 (years, mean ± SD) | 55.14 ±11.85 | 56.92 ± 11.99 | | 0.390 | 47.12 ± 12.36 | | 58.33 ± 11.42 | | 0.001 |
| 性别 (No, [%])  男  女 | 34 (60.71)  22 (39.29) | 40 (47.62)  44 (52.38) | | 0.128 | 2 (0.08)  23 (0.92) | | 28 (0.78)  8 (0.22) | | <0.001 |
| 病理结果  肺结核  结节病  炎性病变  肺腺癌  肺鳞癌  小细胞肺癌  乳腺癌  肾癌 | 3  46  7  -  -  -  -  - | -  -  -  32  8  39  3  2 | | N/A | 3  19  3  -  -  -  -  - | | -  -  -  16  2  17  1  0 | | N/A |
| 吸烟史  从不吸烟  既往吸烟  吸烟  未知 | 31  5  17  3 | 43  6  18  17 | | 0.074 | 24  0  1  0 | | 10  8  11  7 | | <0.001 |
| 毒害物质接触史  是  否 | 3  53 | 5  79 | | 1.000 | 2  23 | | 4  32 | | 1.000 |
| 家族史  肺部肿瘤  肺外肿瘤 | 0  3 | 3  7 | | 0.528 | 0  1 | | 1  3 | | 1.000 |

表中year代表年，mean代表平均年龄，SD代表标准差， No代表患者人数，N/A代表无法获得。

3.2 放射组学特征提取

本研究共提起了396个定量的放射组学特征，包括一阶直方图特征（共42个）、高阶纹理特征，包括haralick特征（共10个）、灰度大小区域矩阵特征（gray level size zone matrix，GLSZM）特征（共11个），灰度共生矩阵（gray level cooccurrence matrix，GLCM）特征（共144个）和灰度游程长度矩阵（gray level run length matrix，GLRLM）特征（共180个）以及形状因子（form-factor）特征（共9个）。

3.3 放射组学模型构建

为了构建一个鲁棒的放射组学模型，我们只筛选出了稳定性较高的特征用于下一步的处理。共有274个ICC值高于0.75的放射组学特征被保留了下来（如图3所示）。



图3. 组内相关系数直方图

在放射组学特征筛选步骤中，使用LASSO逻辑回归的方法，筛选出了26个系数不为0的放射组学特征（如图4所示）。在使用后退逐步淘汰法的多元逻辑回归的方法后，从26个放射组学特征中筛选出了8个相关性最强的特征保留下来。随后，基于这8个放射组学特征，构建了放射组学诊断模型，模型如下所示：

放射组学评分 = 3.603 – Percentile20 × 10.941 – ClusterShade\_AllDirection\_offset1\_SD × 1.599 + GLCMEnergy\_angle135\_offset7 × 2.592 – Inertia\_AllDirection\_offset1\_SD × 1.760 – sumAverage × 1.645 + HighGrayLevelRunEmphasis\_AllDirection\_offset7\_SD × 2.479 + LongRunHighGrayLevelEmphasis\_angle135\_offset4 × 2.281 + SurfaceArea × 2.504



图4. 基于最小绝对收缩和选择算子（LASSO）的放射组学特征筛选

A：以LASSO回归为方法，以最小均方误为目标，进行10折交叉验证来调整参数（λ）。 绘制了二项式偏差与log（λ）之间的关系。 垂直的左虚线是使用最小均方误标准的最佳λ值，而右侧是使用最小均方误标准的1个标准误（SE）（1-SE标准）在最佳λ值。 根据10折交叉验证，选择λ值为0.01987634，log（λ）为-3.918。

B：396个特征的LASSO系数分布图。 绘制系数分布图与log（λ）坐标系。 在10折交叉验证所选择的值处绘制一条垂直线，其中最佳λ处存在26个系数非0的特征。

本研究另外又用了2种不同的方法来进行放射组学特征的筛选。这两种方法分别是单因素逻辑回归的Spearman相关分析，以及梯度提升决策树（Gradient Boost Decision Tree，GBDT）和多元逻辑回归的Spearman相关分析。这两种方法所建立的诊断模型的分类性能与LASSO回归所构建的诊断模型的分类性能相似。

3.4 放射组学模型评估

在测试组中，前文所述的6种不同的机器学习算法所构建的诊断模型的AUC值都超过了0.65（如表3所示）。

表3**.** 不同算法所构建的模型的分类性能

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 组别 | 逻辑回归 | 支持向量机 | 朴素贝叶斯 | 随机森林 | 决策树 | K-最邻近 |
| 训练组  曲线下面积  精确性  敏感性  特异性 | 0.922  0.850  0.821  0.893 | 0.907  0.843  0.869  0.804 | 0.808  0.771  0.810  0.714 | 0.999  0.986  0.976  1.000 | 0.765  0.757  0.726  0.804 | 0.914  0.836  0.845  0.821 |
| 测试组  曲线下面积  精确性  敏感性  特异性 | 0.850  0.803  0.806  0.800 | 0.853  0.803  0.833  0.760 | 0.751  0.738  0.778  0.680 | 0.789  0.672  0.833  0.440 | 0.661  0.672  0.722  0.600 | 0.743  0.705  0.722  0.680 |

本研究使用DeLong检验对比了六种机器学习算法的分类性能。结果显示利用支持向量机和逻辑回归两种算法所构建的诊断模型的分类性能相似，且两种算法都是广义线性模型。考虑到可解释性和泛化能力，最终选择了逻辑回归模型作为主要结果。放射组学评分鉴别纵隔淋巴结良恶性的分类性能如下。

3.4.1 辨别度

训练组和测试组的放射组学诊断模型的ROC曲线可见于图5A-B。放射组学诊断模型在训练组的精确性、敏感性、特异性和AUC分别为0.850、0.821、0.893以及0.922（95%置信区间：0.880-0.964）。放射组学诊断模型在测试组的精确性、敏感性、特异性和AUC分别为0.803、0.806、0.800以及0.850（95%置信区间：0.747-0.953）。基于阅片者2所勾画的ROI提取出的放射组学特征所构建的放射组学诊断模型的精确性、敏感性、特异性和AUC分别为0.900、0.870、0.926以及0.957（95%置信区间：0.906-1.000）。



图5. 逻辑回归模型鉴别纵隔淋巴结良恶性的接收者操作特征曲线(ROC曲线)

A：代表训练组的ROC曲线； B：代表测试组的ROC曲线；C：代表不同品牌CT扫描仪的ROC曲线；D：代表不同扫描层厚的ROC曲线。图中AUC代表曲线下面积，95% CI代表95%置信区间。

训练组和测试组的放射组学评分的分布如图6所示。



图6. 淋巴结的放射组学评分

A：代表训练组的每个淋巴结的放射组学评分； B：代表测试组的每个淋巴结的放射组学评分。图中y轴代表放射组学评分减去最佳阈值后得到的值。横线上下分别代表模型诊断为恶性和良性的淋巴结。蓝条和红条分别表示病理结果为恶性和良性的淋巴结。

3.4.2 校准度

在训练组和测试组中，放射组学模型所诊断的结果与实际病理结果之间均显示出了较好的一致性（如图7A-B所示）。Hosmer-Lemeshow检验结果显示训练组（P值：0.6898）和测试组（P值：0.1762）均无统计学差异。



图7. 放射组学模型的校准曲线和决策曲线分析（DCA）

A：代表训练组的校准曲线； B：代表测试组的校准曲线；对角线代表最佳分类。C：代表训练组的DCA；D：代表测试组的DCA；y轴代表净收益，净收益=真阳性率–（假阳性率×加权因子； 加权因子=阈值概率/（1-阈值概率），蓝线代表假设所有的淋巴结均为恶性，黑线代表假设所有的淋巴结均为良性。

3.4.3 临床应用

DCA结果显示，在训练组，当阈值概率在10%-100%的范围内时，应用放射组学模型诊断纵隔淋巴结的良恶性会获得更多的净收益；在测试组，当阈值概率在30%-90%的范围内时，应用放射组学模型诊断纵隔淋巴结的良恶性会获得更多的净收益（如图7 C-D所示）。

基于静脉期图像所构建的放射组学模型的诊断性能优于基于平扫图像所构建的放射组学模型的诊断性能。基于平扫图像所构建的放射组学模型的ROC曲线如图8所示。



图8. 基于平扫图像构建的放射组学模型的接收者操作特征曲线(ROC曲线)

A：代表训练组的ROC曲线； B：代表测试组的ROC曲线。图中AUC代表曲线下面积。

参考文献

1. Guo D, Ni Y, Lv X, Zhang Z, Ye P: **Distribution and prognosis of mediastinal lymph node metastases of nonsmall cell lung cancer**. *Journal of cancer research and therapeutics* 2016, **12**(5):120.

2. Hegde PV, Liberman M: **Mediastinal staging: endosonographic ultrasound lymph node biopsy or mediastinoscopy**. *Thoracic surgery clinics* 2016, **26**(3):243-249.

3. Alberts WM: **Diagnosis and management of lung cancer executive summary: ACCP evidence-based clinical practice guidelines**. *Chest* 2007, **132**(3):1S-19S.

4. Li L, Zhang H, Wang L, Xie C, Zhou Y, Zhong Y: **A retrospective analysis on metastatic rate of the internal mammary lymph node and its clinical significance in adjuvant radiotherapy of breast cancer patients**. *BMC cancer* 2020, **20**(1):1-7.

5. Sanz-Santos J, Call S: **Preoperative staging of the mediastinum is an essential and multidisciplinary task**. *Respirology* 2020.

6. Gu P, Zhao YZ, Jiang LY, Zhang W, Xin Y, Han BH: **Endobronchial ultrasound-guided transbronchial needle aspiration for staging of lung cancer: a systematic review and meta-analysis**. *Eur J Cancer* 2009, **45**(8):1389-1396.

7. Silvestri GA, Gonzalez AV, Jantz MA, Margolis ML, Gould MK, Tanoue LT, Harris LJ, Detterbeck FC: **Methods for staging non-small cell lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines**. *Chest* 2013, **143**(5 Suppl):e211S-e250S.

8. Varela-Lema L, Fernandez-Villar A, Ruano-Ravina A: **Effectiveness and safety of endobronchial ultrasound-transbronchial needle aspiration: a systematic review**. *Eur Respir J* 2009, **33**(5):1156-1164.

9. Yasufuku K, Chiyo M, Koh E, Moriya Y, Iyoda A, Sekine Y, Shibuya K, Iizasa T, Fujisawa T: **Endobronchial ultrasound guided transbronchial needle aspiration for staging of lung cancer**. *Lung Cancer* 2005, **50**(3):347-354.

10. Haas AR: **Infectious complications from full extension endobronchial ultrasound transbronchial needle aspiration**. *Eur Respir J* 2009, **33**(4):935-938.

11. Asano F, Aoe M, Ohsaki Y, Okada Y, Sasada S, Sato S, Suzuki E, Semba H, Fukuoka K, Fujino S *et al*: **Complications associated with endobronchial ultrasound-guided transbronchial needle aspiration: a nationwide survey by the Japan Society for Respiratory Endoscopy**. *Respir Res* 2013, **14**:50.

12. Steinfort DP, Johnson DF, Irving LB: **Infective complications from endobronchial ultrasound-transbronchial needle aspiration**. *Eur Respir J* 2009, **34**(2):524-525; author reply 525.

13. Tian Q, Chen LA, Wang RT, Yang Z, An YJTc: **The reasons of false negative results of endobronchial ultrasound‐guided transbronchial needle aspiration in the diagnosis of intrapulmonary and mediastinal malignancy**. 2013, **4**(2):186-190.

14. Medford AR, Bennett J, Free C, Agrawal SJPmj: **Endobronchial ultrasound guided transbronchial needle aspiration**. 2010, **86**(1012):106-115.

15. Kemp S, El Batrawy S, Harrison R, Skwarski K, Munavvar M, Roselli A, Cusworth K, Shah P: **Learning curves for endobronchial ultrasound using cusum analysis**. In*.*: BMJ Publishing Group Ltd; 2010.

16. Medford ARJTc: **False negative results and endobronchial ultrasound‐guided transbronchial needle aspiration**. 2013, **4**(4):484-484.

17. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D *et al*: **Radiomics: the process and the challenges**. *Magn Reson Imaging* 2012, **30**(9):1234-1248.

18. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D *et al*: **Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach**. *Nat Commun* 2014, **5**:4006.

19. Gillies RJ, Kinahan PE, Hricak H: **Radiomics: Images Are More than Pictures, They Are Data**. *Radiology* 2016, **278**(2):563-577.

20. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A *et al*: **Radiomics: extracting more information from medical images using advanced feature analysis**. *Eur J Cancer* 2012, **48**(4):441-446.

21. Zhao B, Tan Y, Tsai WY, Qi J, Xie C, Lu L, Schwartz LH: **Reproducibility of radiomics for deciphering tumor phenotype with imaging**. *Sci Rep* 2016, **6**:23428.

22. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RT, Hermann G, Lambin P, Haibe-Kains B, Mak RH, Aerts HJ: **CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma**. *Radiother Oncol* 2015, **114**(3):345-350.

23. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, Ma ZL, Liu ZY: **Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer**. *J Clin Oncol* 2016, **34**(18):2157-2164.

24. Cong M, Feng H, Ren JL, Xu Q, Cong L, Hou Z, Wang YY, Shi G: **Development of a predictive radiomics model for lymph node metastases in pre-surgical CT-based stage IA non-small cell lung cancer**. *Lung Cancer* 2019, **139**:73-79.

25. Bayanati H, R ET, Souza CA, Sethi-Virmani V, Gupta A, Maziak D, Amjadi K, Dennie C: **Quantitative CT texture and shape analysis: can it differentiate benign and malignant mediastinal lymph nodes in patients with primary lung cancer?** *Eur Radiol* 2015, **25**(2):480-487.

26. Shin SY, Hong IK, Jo YS: **Quantitative computed tomography texture analysis: can it improve diagnostic accuracy to differentiate malignant lymph nodes?** *Cancer Imaging* 2019, **19**(1):25.

27. Sha X, Gong G, Qiu Q, Duan J, Li D, Yin YJBMI: **Discrimination of mediastinal metastatic lymph nodes in NSCLC based on radiomic features in different phases of CT imaging**. 2020, **20**(1):12.

28. Dong D, Tang L, Li ZY, Fang MJ, Gao JB, Shan XH, Ying XJ, Sun YS, Fu J, Wang XX *et al*: **Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer**. *Annals of Oncology* 2019, **30**(3):431-438.

29. Xu X, Li H, Wang S, Fang M, Zhong L, Fan W, Dong D, Tian J, Zhao X: **Multiplanar MRI-Based Predictive Model for Preoperative Assessment of Lymph Node Metastasis in Endometrial Cancer**. *Front Oncol* 2019, **9**:1007.

30. Cong M, Feng H, Ren J-L, Xu Q, Cong L, Hou Z, Wang Y-y, Shi G: **Development of a predictive radiomics model for lymph node metastases in pre-surgical CT-based stage IA non-small cell lung cancer**. *Lung Cancer* 2020, **139**:73-79.

31. Zhang Y, Yu S, Zhang L, Kang L: **Radiomics Based on CECT in Differentiating Kimura Disease From Lymph Node Metastases in Head and Neck: A Non-Invasive and Reliable Method**. *Front Oncol* 2020, **10**:1121.

32. Zwanenburg A, Leger S, Vallières M, Löck S: **Image biomarker standardisation initiative**. *arXiv preprint arXiv:161207003* 2016.