

论 著

基于CT图像的活检可帮助预测非小细胞肺癌患者HOPX表达状态和预后的研究*

张学林* 王 杨 卓诗宇

河北北方学院附属第二医院肿瘤放疗科
(河北 张家口 075100)

【摘要】目的 本研究旨在阐明具有放射基因组特征的基于CT图像的活检以预测非小细胞肺癌(NSCLC)患者仅同源域蛋白同源盒(homeodomain-only protein homeobox, HOPX)基因表达状态和预后。**方法** 根据HOPX的表达将患者标记为HOPX阴性或阳性,并将其分为训练数据集(n=92)和测试数据集(n=24)。在对116例患者进行基因与图像特征的相关性分析中,从1218个图像特征中选出了8个与HOPX表达相关的显著特征作为放射基因组特征候选,以预测HOPX的表达状态和预后。**结果** 通过叠加集成学习模型建立具有放射基因组特征的影像活检模型,在测试数据集中,模型显示对HOPX表达的预测能力,ROC曲线下的面积为0.873, Kaplan-Meier曲线的预测能力(P=0.0066)。**结论** 具有放射基因组特征的基于CT图像的活检可以帮助医生预测HOPX在非小细胞肺癌中的表达状况和预后。

【关键词】 CT图像; 非小细胞肺癌;
同源域蛋白同源盒

【中图分类号】 R563

【文献标识码】 A

【基金项目】 河北省卫生健康委科研
基金项目(20231469)
张家口市重点研发计划
项目(2221174D)

DOI:10.3969/j.issn.1672-5131.2024.09.016

Biopsy Based on CT Images Can Help Predict HOPX Expression Status and Prognosis of Patients with Non-small Cell Lung Cancer*

ZHANG Xue-lin*, WANG Yang, ZHUO Shi-yu.

Department of Radiation Oncology, The Second Affiliated Hospital of Hebei North University, Zhangjiakou 075100, Hebei Province, China

ABSTRACT

Objective This study aimed to elucidate CT image-based biopsies with radiogenomic features to predict homeodomain-only protein homeobox (HOPX) gene expression status and prognosis in patients with non-small cell lung cancer (NSCLC). **Methods** Patients were labeled as HOPX-negative or HOPX-positive based on HOPX expression and divided into a training dataset (n=92) and a test dataset (n=24). In the analysis of correlation between gene and image features in 116 patients, 8 significant features related to HOPX expression were selected from 1218 image features as candidates for radiogenomic features to predict the expression status and prognosis of HOPX. **Results** An image biopsy model with radiogenomic features was established by superposition ensemble learning model. In the test data set, the model showed the predictive ability of HOPX expression, the area under ROC curve was 0.873, and the predictive ability of Kaplan-Meier curve was 0.0066 (P=0.0066). **Conclusion** CT image-based biopsies with radiogenomic signatures can help physicians predict HOPX expression status and prognosis in non-small cell lung cancer.

Keywords: CT Image; Non-small Cell Lung Cancer; Homeodomain Protein Homeobox

肺癌是世界上第二种最常见的癌症,它仍然是癌症死亡的主要原因。放射组学具有将医学图像转换为肿瘤内感兴趣区(ROI)的定量信息的前景,通过肿瘤三维成像,放射组学可以完全捕获肿瘤的异质性作为成像表型^[1]。放射基因组学是对活组织检查的补充或比较,它可以非侵入性地识别关于肿瘤的空间异质性基因组信息,并提高诊断准确性^[2]。尽管大多数肺癌放射基因组分析是基于从FDG PET/CT图像提取的图像特征,但根据指南和补偿要求,低剂量CT筛查仍然是推荐的肺癌诊断临床服务,以标准化实践并优化收益和风险之间的平衡。因此,这项研究试图揭示基于简单CT图像的放射基因组分析的可能性,以创建成像活检,以帮助医生预测肺癌的基因表达状态和进展^[3]。此外,最近许多关于放射基因组学的研究集中在EGFR突变上,并取得了良好的预测,这表明了将基因与图像联系起来的可能性。考虑到肿瘤微环境的复杂性,以及应该对患者进行更精确的治疗以获得更好的生存,基于CT图像的EGFR突变状态的预测可以扩展到其他基因^[4]。

最近, RNA测序分析的研究人员发现, HOPX基因可能在非小细胞肺癌中发挥肿瘤抑制作用^[5]。在肺癌细胞系中, HOPX完全或部分甲基化,导致基因沉默或其基因表达下调。作为一种特异的分化转录调节因子, HOPX的低表达可能导致肺癌患者的低分化肿瘤和预后不良^[6]。因此,在治疗前预测HOPX基因的表达状态,对于预测肺癌患者的总体生存期,为个性化治疗方案的决策过程起着至关重要的作用。我们假设, HOPX的表达状态可以通过基于CT影像的放射基因组学或影像活检来预测,而与HOPX表达相关的放射基因组学特征可以为我们提供将NSCLC患者分为高风险组和低风险组的预后力量^[7]。本研究旨在阐明一种基于CT图像的放射基因组学特征来预测非小细胞肺癌患者HOPX的表达状况和预后。

1 资料与方法

1.1 研究资料 本研究依据《赫尔辛基宣言》进行,我们对2018年8月至2023年8月非小细胞肺癌患者HOPX基因表达116名患者进行了回顾性分析,患者被随机分成训练组(n=92, 79%)和测试组(n=24, 21%),因此训练组和测试组之间在年龄、性别、组织学、分期和HOPX状态方面没有统计学上的显著差异。预测模型是从训练数据集开发的,并针对测试数据集进行了评估。扫描参数为:患者CT图像仰卧位,像素大小0.59-0.98mm(中位数:0.78mm),层厚0.625-2.5mm(中位数:1.25mm),X射线管电流124-699mA(平均:220mA),X射线管电压80-140千伏(平均:120千伏)。

从训练数据集中提取了1218个CT图像特征。研究了10674个基因表达数据与1218个CT图像特征的相关性。在这项研究中, HOPX基因被发现与图像特征有关。在建立HOPX预测模型的步骤中,使用差异基因表达分析(DEA)将所有患者标记为HOPX阴性和HOPX阳性。从训练数据集中提取的1218个CT图像特征中,选择了8个与HOPX表达相关的显著特征(Spearman系数绝对值>0.5)作为放射基因组特征。影像活检的预测模型是从堆叠集合学习中建立的,用于预测HOPX的表达状态和患者预后,并通过ROC分析对该模型进行评估。

【第一作者】 张学林,男,副主任医师,主要研究方向:医学影像和放射治疗学。E-mail: wangdan198277@163.com

【通讯作者】 张学林

1.2 研究方法 本研究从三种类型的图像中提取了1218个图像特征^[3]：原始CT图像、高斯拉普拉斯(LOG)特征和小波滤波。图像特征包括三组：一阶(直方图)特征(描述体素强度直方图的统计特性)；形状特征(测量肿瘤尺寸和量化三维形态)；纹理特征(描述灰度图像中联合概率和量化体素相关性)，包括灰度共生矩阵(GLCM)特征、灰度大小区域矩阵(GLSZM)特征、灰度游程矩阵(GLRLM)特征、相邻灰度色调差矩阵(NGTDM)特征和灰度依赖矩阵(GLDM)特征。使用z-Score标准化对所有图像特征进行归一化，以减少在尺度、范围和统计分布方面的基本差异。

肿瘤组织样本是在手术过程中从未接受治疗的受试者身上采集的，对肿瘤组织样品进行RNA测序，然后根据Illumina TruSeq样品制备试剂盒方案，将总RNA转换到Illumina HiSeq上的双端测序文库中，计算了10,674个基因表达数据和来自CT图像的1218个图像特征的所有组合的Spearman系数^[5]。最后发现HOPX与CT图像特征相关。

所有患者均采用基于RNA测序的DEA标记HOPX状态(阳性或阴性)。根据DEA获得的HOPX表达的对数二倍变化(Log2FC)将116例患者分为HOPX阴性组和HOPX阳性组。在DEA中，根据两年的生存时间(截止时间)将患者分为短期和长期生存组。将数据归一化后，使用PYTHON软件包“SciPy(Version 1.10.0)”对长生存组和短生存组进行Fisher精确检验，计算出p值。利用了Km曲线的生存分析，用于机器学习模型对患者进行分类。如果Log2Fc<Hopx1.57(低表达-)，则认为该患者Hopx阴性。否则(HOPX正常和过度表达)，患者被认为HOPX阳性。

1.3 统计学方法 所有数据分析用SPSS 26.0软件进行。P值<0.05被认为具有统计学意义。应用该方法对ROC曲线下面积(AUC)、准确度、灵敏度和特异性进行预测。使用Kaplan曲线和Meier曲线之间的对数秩检验的P值对模型分类的两组具有不同放射基因组特征的成像活检模型的预后能力进行评估。P值<0.05认为在本

研究中具有统计学意义。

2 结果

2.1 HOPX基因分析 具体如图1所示，图1A显示了10,674个基因的短存活组和长存活组获得的差异表达(Log2FC)和平均表达差异(HOPXlog10)之间的关系的火山图，HOPX低表达的患者更有可能表现出较差的总体存活率。HOPX表达状态的预测能力在本研究的训练数据集中得到了验证。如图1B所示，两组的存活率在统计学上有显著差异。

2.2 放射基因组学特征 具体如图2所示，计算10,674个基因表达数据和1218个CT图像特征的所有组合的Spearman系数，以识别图像特征相关的基因，显示19个基因(HOPX、A2M、ITGBL1、FMO5、SELENBP1、SUSD2、CYP4V2、GPD1L、SCN7A、C7、IL6ST、NBEAL1、PIGA、USP54、TTC38、LONRF3、MAOA、LAMA5和RPL17)与17个图像特征之间较高的Spearman系数。在所有这些基因中，只有HOPX基因显示大多数图像特征的Spearman系数绝对值高于0.5，8个特征，并且在差异表达分析中也检测到低表达水平的特征。在Spearman相关分析中，HOPX的P值和调整后的P值均低于0.05。

2.3 HOPX表达状态与预后的预测能力 具体如图3A-图3D所示，显示了由测试数据集的四个成像活检预测的Kaplan-Meier生存曲线。橙色和蓝色曲线分别代表预测的HOPX阴性(存活率差)和HOPX阳性(存活率好)组。影像活检D被选为最终的放射基因组学特征，显示出最显著的统计学差异(P=0.0066)。应用影像活检D的预测模型对HOPX的表达状态和总生存期具有预测能力。

2.4 具有放射基因组特征的最终图像活检 具体如图4A-4B所示，显示了不同HOPX表达状态的患者的最终放射基因组特征中的两个显著图像特征图，HOPX阳性患者和HOPX阴性患者之间的平均值也显示出统计学上的显著差异。

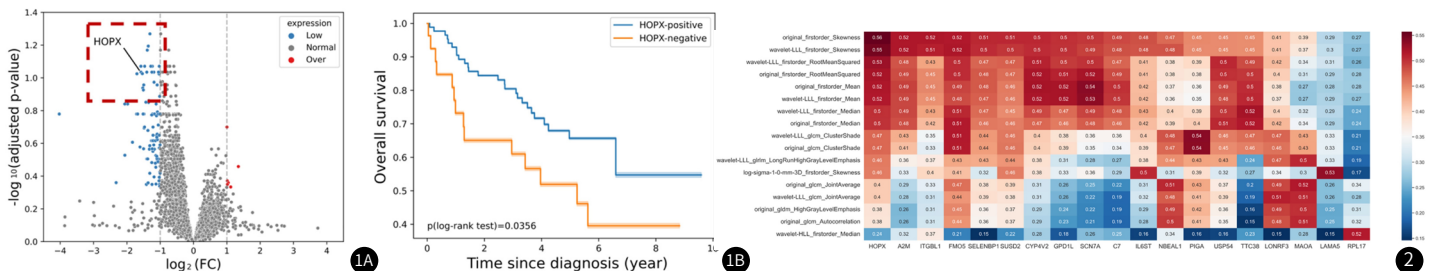


图1A-图1B 1A包括HOPX在内的10674个基因，短生存组和长生存组差异表达量(log2FC)与平均表达量(-log10)之间的关系火山图(P=0.092)具有统计学意义。蓝点表示校正前低表达基因，未经校正P<0.05。红点表示校正前的过表达基因，未经校正P<0.05；1B 训练数据集中生存的Kaplan-Meier曲线(P=0.0356)。橙色和蓝色曲线分别代表HOPX阴性(生存差)和HOPX阳性(生存好)组。

图2 基因表达(HOPX)与图像特征的相关系数(绝对值)热图。红色表示相关度高。

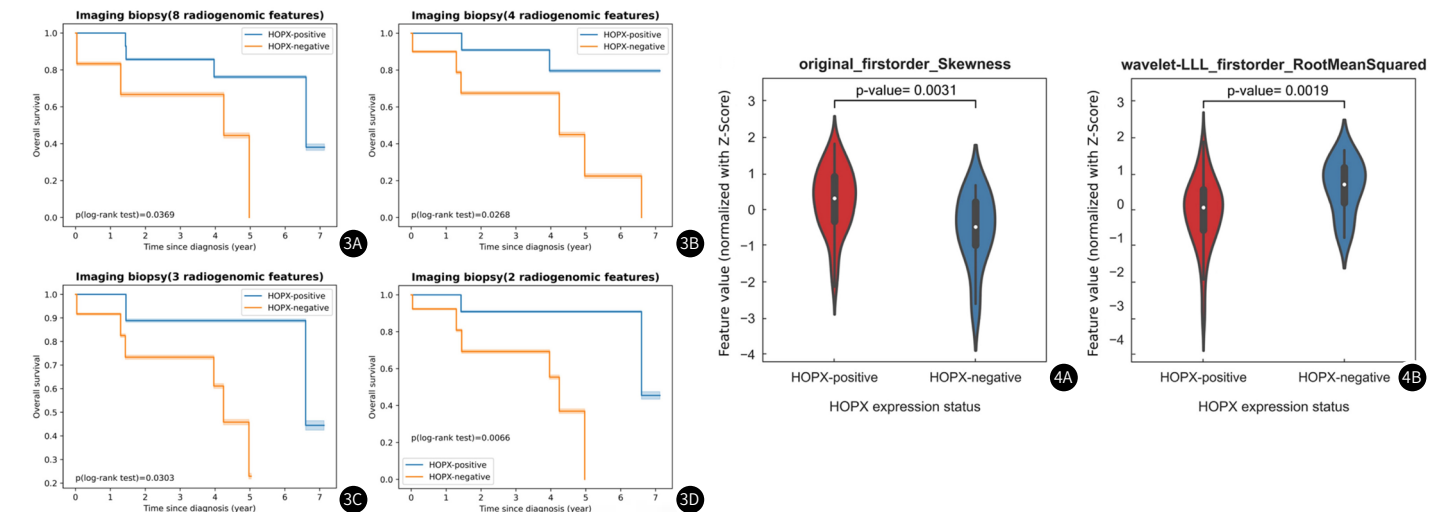


图3A-图3D 在测试数据集集中的Kaplan-Meier生存曲线：(3A)由8个放射基因组特征组成的成像活检生存曲线(P=0.0369)；(3B)由4个放射基因组特征组成的成像活检生存曲线(P=0.0268)；(3C)由3个放射基因组特征组成的成像活检生存曲线(P=0.0303)；(3D)由2个放射基因组特征组成的成像活检生存曲线(P=0.0066)。

图4A-图4B 4A 所有hopx阳性患者和hopx阴性患者图像特征的小提琴曲线图，数据分布有显著差异(P=0.0031)；(4B)hopx阳性患者和hopx阴性患者图像特征小提琴图，数据分布差异有统计学意义(P=0.0019)。

3 讨论

这项研究发现,在10674个基因中,HOPX基因与CT影像特征有关。基于HOPX基因放射基因组学特征的不同成像活检建立的机器学习模型显示了预测患者肺癌预后的能力^[8],这表明可以将基因表达和图像特征结合起来创建基于图像的活检,以帮助预测癌症的预后^[6]。影像活检包括8个放射基因组学特征(一阶特征),它们描述了CT图像上肿瘤内部体素强度的异质性。我们推测这一现象的原因可能与HOPX在肺癌和肺癌中的生物学功能有关。HOPX阴性患者(C)的肿瘤边界附近的偏斜度像素值高于HOPX阳性患者(G),这可能表明HOPX阴性患者的细胞侵袭性更强,分化程度更低^[7]。本研究中HOPX阴性患者和HOPX阳性患者之间偏度的一阶特征的数据分布也显示出类似的现象:HOPX阴性患者显示出更离散的数据分布^[9]。本研究中,HOPX阴性患者瘤内区域的均方根像素值高于HOPX阳性患者,其中HOPX阴性患者的均方根相对高于HOPX阳性患者的RMS。这些结果表明,HOPX阴性患者的异质性大于HOPX阳性患者^[10]。

近年来,一些研究试图阐明NSCLC的肿瘤促进和肿瘤抑制途径,以实现更有效的治疗^[11]。HOPX在调节心肺发育中扮演着特别重要的角色。通过追踪HOPX的信号通路,研究人员发现^[12],HOPX在抑制肺癌细胞增殖、抑制肿瘤细胞迁移和侵袭方面具有显著影响,导致NSCLC中表达降低,患者预后不良。HOPX还参与了转录的一部分该程序与远端气道上皮分化和肺腺癌(LUAD)进展有关,其中HOPX被认为通过激活Ras/MAPK信号通路和抑制Akt通路诱导细胞衰老^[13]。它是谱系选择通路(GATA6、HOPX和NKX2-1)中的重要节点之一,直接将气道上皮特异性效应物与LUAD亚型的转移抑制联系起来^[14]。最近的研究发现^[15],肺泡2型(AT2)细胞可能是人类LUAD的起源细胞,它导致去分化产生干细胞样状态,启动并保持癌症进展。HOPX的低表达可能加剧肺泡细胞向远端气道干细胞去分化的过程,导致存活不良^[16]。细胞分化与一阶特征的表型之间可能存在一定的关系,这可以从我们研究中选择的用于建立放射基因组特征的图像特征中看出^[17]。Spearman相关分析和DEA均发现与图像特征相关的基因为SELENBP1和SUSD2,而不是HOPX^[18]。硒结合蛋白1(SELENBP1)是NSCLC起源和发展过程中重要的抑瘤基因,可能是一种新的靶向生物标志物^[19]。SUSD2 (Sushi Domain Containing 2)被认为是肺腺癌新的预后和潜在的治疗靶点^[20]。

综上所述,调节肺癌肿瘤抑制功能的HOPX被发现与图像特征有关。本研究阐述了一种基于CT图像的具有放射基因组特征的影像活检,以帮助预测非小细胞肺癌患者HOPX的表达状况和预后。我们对放射基因组学特征的研究揭示了帮助预测肺癌预后的潜力,并证明了将成像数据与遗传信息联系起来的可能性分子水平上。从CT图像中检测基因表达状态可以提高湿活检的准确性。

参考文献

- [1]Saad M B,Hong L,Aminu M,et al.Predicting benefit from immune checkpoint inhibitors in patients with non-small-cell lung cancer by CT-based ensemble deep learning:a retrospective study[J].Lancet Digit Health,2023,5(7):e404-e420.
- [2]Serfling S E,Hartrampf P E,Zhi Y,et al.Somatostatin receptor-directed PET/CT for therapeutic decision-making and disease control in patients affected with small cell lung cancer[J].Clin Nucl Med,2023,48(4):309-314.

- [3]Sun Y,Xiao L,Wang Y,et al.Diagnostic value of dynamic (18)F-FDG PET/CT imaging in non-small cell lung cancer and FDG hypermetabolic lymph nodes[J].Quant Imaging Med Surg,2023,13(4):2556-2567.
- [4]Thanoon M A,Zulkifley M A,Mohd Z M,et al.A review of deep learning techniques for lung cancer screening and diagnosis based on CT Images[J].Diagnostics (Basel),2023,13(16).
- [5]Tisi S,Creamer A W,Dickson J,et al.Prevalence and clinical characteristics of non-malignant CT detected incidental findings in the SUMMIT lung cancer screening cohort[J].BMJ Open Respir Res,2023,10(1).
- [6]Tu W,Zhou T,Zhou X,et al.Nomograms using CT morphological features and clinical characteristics to identify COPD in patients with lung cancer:a multicenter study[J].Int J Chron Obstruct Pulmon Dis,2023,18:1169-1185.
- [7]Upperton S,Bradley C,Bhartia B,et al.The radiology quality assurance process in the Yorkshire Lung Screening Trial, and findings from the baseline round of low dose CT screening for lung cancer[J].Br J Radiol,2023:20230126.
- [8]Ding Y,Feng H,Yang Y,et al.Deep-learning based fast and accurate 3D CT deformable image registration in lung cancer[J].Med Phys,2023.
- [9]Chae K J,Lim S,Seo J B,et al.Interstitial lung abnormalities at CT in the Korean National Lung Cancer Screening Program:prevalence and deep learning-based texture analysis[J].Radiology,2023,307(4):e222828.
- [10]Wang M L,Zhang H,Yu H J,et al.An initial study on the comparison of diagnostic performance of (18)F-FDG PET/MR and (18)F-FDG PET/CT for thoracic staging of non-small cell lung cancer:focus on pleural invasion[J].Rev Esp Med Nucl Imagen Mol (Engl Ed),2023,42(1):16-23.
- [11]Wang M,Lin S,He N,et al.The introduction of low-dose CT imaging and lung cancer overdiagnosis in Chinese women[J].Chest,2023,163(1):239-250.
- [12]Wang S,Mahon R,Weiss E,et al.Automated lung cancer segmentation using a PET and CT dual-modality deep learning neural network[J].Int J Radiat Oncol Biol Phys,2023,115(2):529-539.
- [13]Wei Y,Ma L,Li P,et al.FAPI Compared with FDG PET/CT for diagnosis of primary and metastatic lung cancer[J].Radiology,2023,308(2):e222785.
- [14]Zhang X,Dong X,Saripan M,et al.Deep learning PET/CT-based radiomics integrates clinical data:A feasibility study to distinguish between tuberculosis nodules and lung cancer[J].Thorac Cancer,2023,14(19):1802-1811.
- [15]Weikert T,Jaeger P F,Yang S,et al.Automated lung cancer assessment on 18F-PET/CT using Retina U-Net and anatomical region segmentation[J].Eur Radiol,2023,33(6):4270-4279.
- [16]Zukotynski K A,Gerbaudo V H.Understanding the value of FAPI versus FDG PET/CT in primary and metastatic lung cancer[J].Radiology,2023,308(2):e231768.
- [17]Yang Q,Huang D,Wu J,et al.Performance of [(18)F]FDG PET/CT versus FAPI PET/CT for lung cancer assessment:a systematic review and meta-analysis[J].Eur Radiol,2023.
- [18]Yuan H,Tan X,Sun X,et al.Role of (18)F-FDG PET/CT and sarcopenia in untreated non-small cell lung cancer with advanced stage[J].Jpn J Radiol,2023,41(5):521-530.
- [19]Zhang T,Wang K,Cui H,et al.Topological structure and global features enhanced graph reasoning model for non-small cell lung cancer segmentation from CT[J].Phys Med Biol,2023,68(2).
- [20]Zhang W,Liu Z,Duan X,et al.Differentiating malignant and benign pleural effusion in patients with lung cancer: an (18)F-FDG PET/CT retrospectively study[J].Front Oncol,2023,13:1192870.

(收稿日期:2023-10-10)

(校对编辑:赵望淇)