

## Machine learning model based on photon-counting CT quantitative parameters for predicting pathological subtype of lung cancer

LI Rui<sup>1</sup>, YANG Xiaofei<sup>1</sup>, ZHAN Pengchao<sup>2</sup>, GU Yanbo<sup>1</sup>, CHEN Yan<sup>1</sup>, GAO Jianbo<sup>1\*</sup>

(1. Department of Radiology, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China;

2. Department of Radiology, the Third People's Hospital of Henan Province, Zhengzhou 450052, China)

[Abstract] **Objective** To investigate the value of machine learning (ML) model based on quantitative parameters of photon-counting CT (PCCT) for predicting pathological subtype of lung cancer. **Methods** A total of 96 patients with pathologically confirmed lung cancer were prospectively enrolled, including 51 cases of adenocarcinoma (AC) (AC group), 24 cases of squamous cell carcinoma (SCC) (SCC group) and 21 cases of small cell lung cancer (SCLC) (SCLC group). Based on pre-treatment contrast-enhanced PCCT, valuable parameters were selected using Boruta algorithm, and a decision tree (DT) ML model was constructed. Receiver operating characteristic curve was drawn, the area under the curve (AUC) was calculated to evaluate the performance of this model for identifying each pathological subtype of lung cancer. **Results** In SCC and SCLC groups, the arterial phase iodine concentration (IC), normalized IC (NIC), CT values at 40–70 keV,  $K_{40-80\text{ keV}}$ , as well as venous phase NIC and CT values at 40–80 keV were all lower than those in AC group (all  $P < 0.05$ ), but no significant difference was found between SCC and SCLC groups (all  $P > 0.05$ ). Venous phase IC and venous phase  $K_{40-80\text{ keV}}$  increased sequentially in SCLC, SCC and AC groups (all  $P < 0.05$ ). A total of 18 PCCT parameters were selected, and the overall diagnostic accuracy of DT model constructed based on these parameters for predicting pathological subtype of lung cancer was 80.21%, with AUC for identifying AC, SCC and SCLC of 0.912, 0.771 and 0.945, respectively. **Conclusion** DT model constructed based on PCCT quantitative parameters had high value for predicting pathological subtype of lung cancer.

[Keywords] lung neoplasms; tomography, X-ray computed; pathology; diagnosis, differential; prospective studies

DOI:10.13929/j.issn.1003-3289.2026.03.005

## 基于光子计数 CT 定量参数机器学习模型预测 肺癌病理亚型

李蕊<sup>1</sup>, 杨晓菲<sup>1</sup>, 詹鹏超<sup>2</sup>, 谷艳博<sup>1</sup>, 陈岩<sup>1</sup>, 高剑波<sup>1\*</sup>

(1. 郑州大学第一附属医院放射科, 河南 郑州 450052; 2. 河南省直第三人民医院影像科, 河南 郑州 450052)

[摘要] **目的** 探讨基于光子计数 CT(PCCT)定量参数构建的机器学习(ML)模型预测肺癌病理亚型的价值。**方法** 前瞻性纳入 96 例经术后病理证实的肺癌患者,包括 51 例腺癌(AC)(AC 组)、24 例鳞状细胞癌(SCC)(SCC 组)及 21 例小细胞肺癌(SCLC)(SCLC 组)。采用 Boruta 算法基于治疗前增强 PCCT 筛选组间差异有统计学意义的参数并构建决策树(DT)ML 模型;绘制受试者工作特征曲线,计算曲线下面积(AUC),评估模型预测不同病理亚型肺癌的效能。**结果** SCC 组和 SCLC 组病灶的动脉期碘浓度(IC)、标准化 IC(NIC)、40~70 keV CT 值、 $K_{40-80\text{ keV}}$ ,以及静脉期 NIC、40~80 keV CT 值均低于 AC 组( $P$ 均 $<0.05$ ),而前二者差异均无统计学意义( $P$ 均 $>0.05$ );SCLC 组、SCC 组及 AC 组静脉期 IC 及静脉

[基金项目] 河南省医学科技攻关计划联合共建项目(LHGJ20240246)。

[第一作者] 李蕊(1990—),女,河南周口人,硕士,主治医师。研究方向:胸腹部疾病影像学。E-mail: lirui\_0131@126.com

[通信作者] 高剑波,郑州大学第一附属医院放射科,450052。E-mail: cjr.gaojianbo@vip.163.com

[收稿日期] 2025-09-14 [修回日期] 2026-03-11

期  $K_{40\sim 80\text{ keV}}$  依次升高 ( $P$  均  $< 0.05$ )。基于 18 个 PCCT 参数构建的 DT 模型预测肺癌亚型的整体准确率为 80.21%，其预测 AC、SCC 及 SCLC 的 AUC 分别为 0.912、0.771 及 0.945。结论 基于 PCCT 定量参数构建的 DT 模型用于预测肺癌病理亚型具有较高价值。

[关键词] 肺肿瘤; 体层摄影术, X 线计算机; 病理学; 诊断, 鉴别; 前瞻性研究

[中图分类号] R734.2; R445.3 [文献标识码] A [文章编号] 1003-3289(2026)03-0342-05

肺癌发病率和死亡率均较高<sup>[1-2]</sup>, 不同病理亚型肺癌对于治疗的反应及预后差异显著, 故治疗前明确诊断对于指导个体化决策至关重要<sup>[3-4]</sup>。常规 CT 仅为影像学初步评估肺结节手段, 其鉴别不同病理亚型肺癌的能力有限<sup>[5]</sup>。双能量 CT (dual-energy CT, DECT) 用于鉴别诊断不同病理亚型肺癌的价值已初步显现<sup>[6]</sup>; 光子计数 CT (photon-counting CT, PCCT) 空间分辨率更高、对比度噪声比优异且 CT 值响应稳定, 有望为精准诊断肺癌提供更丰富、准确的定量信息<sup>[7]</sup>。本研究探讨基于 PCCT 定量参数构建机器学习 (machine learning, ML) 模型用于预测肺癌病理亚型的价值。

## 1 资料与方法

1.1 研究对象 前瞻性纳入 2024 年 7 月—2025 年 1 月于郑州大学第一附属医院 96 例治疗前接受增强 PCCT 扫描、之后经病理证实的原发性肺癌患者, 男 57 例、女 39 例, 年龄 43~77 岁、平均 (61.1±7.9) 岁; 肿瘤直径  $> 1\text{ cm}$ , 且 PCCT 前未接受相关治疗。排除碘对比剂过敏、心肾功能障碍、孕妇, 以及病理亚型不明确者。依据 2015 年 WHO 肺肿瘤分类标准, 96 例中, 腺癌 (adenocarcinoma, AC) 51 例 (AC 组)、鳞状细胞癌 (squamous cell carcinoma, SCC) 24 例 (SCC 组)、小细胞肺癌 (small cell lung cancer, SCLC) 21 例 (SCLC 组)。本研究通过本院伦理委员会批准 (2024-KY-0862-001); 检查前患者均签署知情同意书。

1.2 仪器与方法 采用 Siemens Naeotom Alpha PCCT 机行仰卧位胸部平扫及增强扫描, 范围自肺尖至肺底; 扫描参数: 管电压 120 kVp, 采用图像质量参考 mAs 39 (IQ 40) 并启用自动曝光控制, 准直宽度  $144 \times 0.4\text{ mm}$ , 转速 0.25 s/rot, 螺距 1.2, 扫描层厚 5 mm, 层间距 5 mm。完成平扫后以 2.0 ml/s 流率注入对比剂碘海醇 (320 mgI/ml) 1.0 ml/kg 体质量并跟注生理盐水 20 ml, 于注射后约 30 s 行动脉期扫描、约 60 s 行静脉期扫描。采用 QIR 3 迭代重建算法对原始数据进行重建, 重建核心 BI60/Br40、层厚 1.0 mm、层间距 0.7 mm, 之后将重建双能量图像传送至后处理工作站 (Siemens Syngo. via VBs)。

1.3 分析图像 由具备 5 及 10 年经验的影像科医师各 1 名采用盲法独立评估病灶, 基于混合能量图观察病灶常规 CT 表现, 根据第 8 版美国癌症联合委员会 (American Joint Committee on Cancer, AJCC) 肺癌分期指南进行 TNM 临床分期, 意见存在分歧时经协商达成一致; 于 PCCT 中勾画病灶 ROI (图 1), 测量病灶在 40~80 keV 虚拟单能量成像图中的平均 CT 值并计算能谱曲线斜率  $K_{40\sim 80\text{ keV}}$ ; 于碘图中测量病灶碘浓度 (iodine concentration, IC) 并计算标准化 IC (normalized IC, NIC), 即病灶 ROI 的 IC 与同层面主动脉或锁骨下动脉 IC 的比值。对上述参数均在选定的 3 个连续层面中各测量 1 次、取平均值; 以 2 名医师测值的算术平均值作为最终结果。

1.4 统计学分析 采用 SPSS 26.0 和 R 4.2.0 软件进行统计学分析。分别以  $\bar{x} \pm s$  及中位数 (上下四分位数) 表示符合及不符合正态分布的定量资料, 对正态分布的定量资料以方差齐性检验进行多组间比较和单因素方差分析, 采用最小显著差异法进行事后检验; 对未通过方差齐性检验或不符合正态分布的定量资料采用秩和检验进行多组间比较, 多重检验通过 Bonferroni 法校正。以例 (%) 表示定性资料, 采用  $\chi^2$  检验进行多组间比较, 多重检验通过 Bonferroni 法校正。采用 Boruta 算法筛选组间差异有统计学意义的参数以构建决策树 (decision tree, DT) ML 模型; 绘制受试者工作特征 (receiver operating characteristic, ROC) 曲线, 计算曲线下面积 (area under the curve, AUC), 评估模型预测效能。  $P < 0.05$  为差异有统计学意义。

## 2 结果

2.1 一般资料 SCLC 组和 SCC 组吸烟者占比及 TNM III 期者占比均高于 AC 组 ( $P$  均  $< 0.05$ )。3 组患者年龄及性别差异均无统计学意义 ( $P$  均  $> 0.05$ )。见表 1。

2.2 常规 CT 表现 SCC 组和 SCLC 组中央型肺癌占比均高于 AC 组 ( $P$  均  $< 0.05$ ), 而前二者差异无统计学意义 ( $P > 0.05$ ); SCC 组病灶左右径大于 AC 组 ( $P < 0.05$ ), 而此 2 组与 SCLC 组差异均无统计学意

义( $P$ 均 $>0.05$ ); SCC 组病灶前后径大于 AC 组及 SCLC 组( $P$ 均 $<0.05$ ),而后 2 组间差异无统计学意义( $P>0.05$ )。SCC 组和 SCLC 组毛刺征占比和胸膜牵拉征占比均低于( $P$ 均 $<0.05$ ),而血管受累、纵隔侵

犯占比均高于 AC 组( $P$ 均 $<0.05$ )。淋巴结融合征多见于 SCLC 组,其次为 SCC 组与 AC 组( $P$ 均 $<0.05$ )。3 组间空泡征占比及强化特点分布差异均无统计学意义( $P$ 均 $>0.05$ )。见表 1。

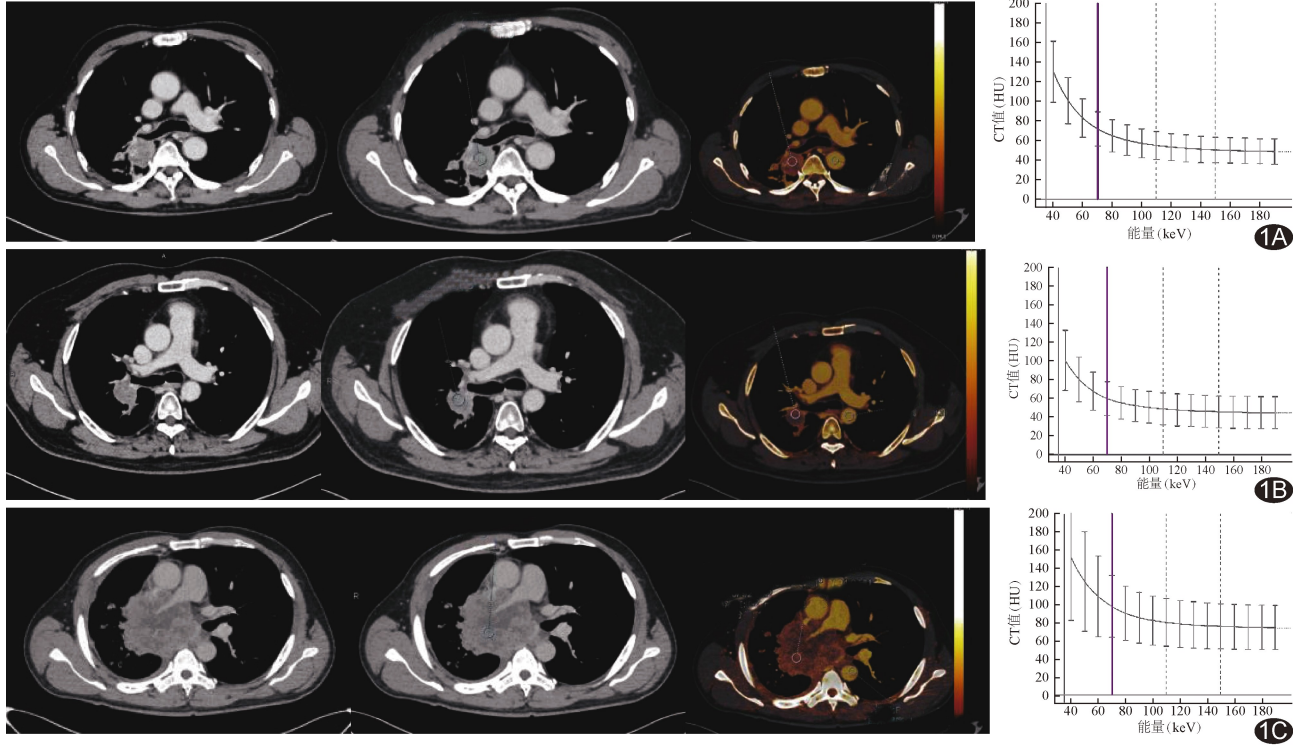


图 1 轴位 PCCT 图及能谱曲线图示原发性肺癌 A. AC 组患者,男,71 岁,右肺上叶后段 AC; B. SCC 组患者,男,45 岁,右肺上叶后段 SCC; C. SCLC 组患者,男,52 岁,右肺门部 SCLC (从左到右均依次为 120 kVp 混合能量图像、70 keV 单能级图像、碘基图及能谱曲线图)

表 1 不同肺癌亚型组患者临床资料及病灶常规 CT 表现比较

组别	性别[例(%)]		年龄(岁)	吸烟史[例(%)]	TNM 分期[例(%)]			
	男	女			I 期	II 期	III 期	IV 期
SCLC 组(n=21)	15(71.43)	6(28.57)	59.7±7.7	12(57.14)*	7(33.33)	2(9.52)	7(33.33)*	5(23.81)
SCC 组(n=24)	16(66.67)	8(33.33)	62.1±7.7	12(50.00)*	5(20.83)	2(8.33)	9(37.50)*	8(33.33)
AC 组(n=51)	26(50.98)	25(49.02)	61.3±8.1	12(23.53)	22(43.14)	2(3.92)	3(5.88)	24(47.06)
$\chi^2/F/H$ 值	3.284		0.535	9.304	16.34			
P 值	0.198		0.587	0.009	0.011			
组别	常规 CT 肿瘤类型[例(%)]		左右径(mm)	前后径(mm)	毛刺征[例(%)]	胸膜牵拉[例(%)]		
	中央型	外周型						
SCLC 组(n=21)	14(66.67)*	7(33.33)*	38.50(20.60,38.50)	29.40(13.70,29.50)#	12(57.14)*	10(47.62)*		
SCC 组(n=24)	15(62.50)*	9(37.50)*	46.59(31.05,50.55)*	44.26(27.12,51.17)*	15(62.50)*	15(62.50)*		
AC 组(n=51)	15(29.41)	36(70.59)	27.70(19.20,41.30)	22.80(15.00,35.65)	48(94.12)	45(88.24)		
$\chi^2/F/H$ 值	11.896		6.982	11.446	16.471	14.184		
P 值	0.002		0.030	0.003	<0.001	0.001		
组别	空泡征[例(%)]	强化特点[例(%)]		血管受累[例(%)]	淋巴结融合征[例(%)]	纵隔侵犯[例(%)]		
		不均匀	均匀					
SCLC 组(n=21)	2(9.52)	5(23.81)	16(76.19)	15(71.43)*	15(71.43)*#	12(57.14)*		
SCC 组(n=24)	3(12.50)	10(41.67)	14(58.33)	17(70.83)*	9(37.50)*	7(29.17)*		
AC 组(n=51)	7(13.73)	19(37.25)	32(62.75)	23(45.10)	12(23.53)	3(5.88)		
$\chi^2/F/H$ 值	0.240	1.722		6.613	14.561	22.834		
P 值	0.992	0.457		0.038	0.001	<0.001		

注: \* :与 AC 组比较  $P<0.05$ ; #:与 SCC 组比较  $P<0.05$

2.3 PCCT 定量参数比较 SCC 组和 SCLC 组病灶动脉期 IC、NIC、40~70 keV CT 值、K<sub>40~80 keV</sub>，以及静脉期 NIC、40~80 keV CT 值均低于 AC 组 (P 均 < 0.05)，而前 2 组间差异均无统计学意义 (P 均 > 0.05)；SCLC 组、SCC 组、AC 组静脉期 IC 及静脉期 K<sub>40~80 keV</sub> 均依次升高 (P 均 < 0.05)；3 组间动脉期 80 keV CT 值差异无统计学意义 (P > 0.05)。见表 2。

2.4 DT 模型 以 Boruta 算法筛选出 18 个 PCCT 参数 (图 2)，基于此构建的 DT 三分类模型预测肺癌亚型的整体准确率为 80.21%，诊断 AC、SCC 及 SCLC

的准确率分别为 86.46%、83.33% 及 83.33%，特异度分别为 77.78%、95.83% 及 78.67%，敏感度分别为 94.12%、45.83% 及 100%，AUC 分别为 0.912、0.771 及 0.945 (图 3)。

### 3 讨论

本研究构建的 DT 模型预测肺癌亚型的整体准确率达 80.21%，识别 AC、SCC 及 SCLC 的 AUC 分别为 0.912、0.771 及 0.945，提示 PCCT 定量参数用于预测肺癌病理亚型具有一定潜力。

本研究中不同亚型肺癌常规 CT 表现存在差异，与

表 2 不同病理亚型肺癌 PCCT 定量参数比较

组别	动脉期							
	IC (mg/ml)	NIC (%)	40 keV CT 值(HU)	50 keV CT 值(HU)	60 keV CT 值(HU)	70 keV CT 值(HU)	80 keV CT 值(HU)	K <sub>40~80 keV</sub>
SCLC 组 (n=21)	0.64±0.20*	10.15(9.70, 11.80)*	88.90(87.35, 91.20)*	72.60±9.93*	65.86±6.66*	60.89±6.97*	57.80±2.95	0.75±0.20*
SCC 组 (n=24)	0.82±0.26*	11.52(9.47, 14.60)*	91.67(79.88, 104.70)*	76.43±11.64*	67.04±9.53*	61.23±8.89*	58.15±15.20	0.86±0.30*
AC 组 (n=51)	0.94±0.28	12.70(10.35, 15.55)	99.70(93.10, 118.25)	84.72±12.11	72.85±9.57	66.00±9.15	61.00±11.70	1.07±0.41
F/H 值	10.517	9.722	14.219	9.717	11.642	3.878	3.129	13.754
P 值	<0.001	0.008	0.001	<0.001	0.003	0.024	0.209	0.001

组别	静脉期							
	IC (mg/ml)	NIC (%)	40 keV CT 值(HU)	50 keV CT 值(HU)	60 keV CT 值(HU)	70 keV CT 值(HU)	80 keV CT 值(HU)	K <sub>40~80 keV</sub>
SCLC 组 (n=21)	0.88±0.18*#	26.48±5.17*	102.22±10.28*	84.50±7.91*	73.51±6.84*	66.95(62.20, 69.10)*	62.10(57.40, 64.70)*	0.98±0.20*#
SCC 组 (n=24)	1.04±0.25*	29.41±7.31*	108.66±17.85*	86.19±12.31*	72.65±9.31*	62.90(60.32, 69.32)*	57.25(55.22, 62.25)*	1.25±0.34*
AC 组 (n=51)	1.21±0.30	35.89±8.92	125.82±21.11	99.28±14.76	83.07±12.37	71.30(67.45, 79.90)	66.50(60.75, 72.15)	1.48±0.47
F/H 值	20.994	20.433	24.789	22.384	19.845	17.355	15.987	21.213
P 值	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

注：\*：与 AC 组比较 P<0.05；#：与 SCC 组比较 P<0.05

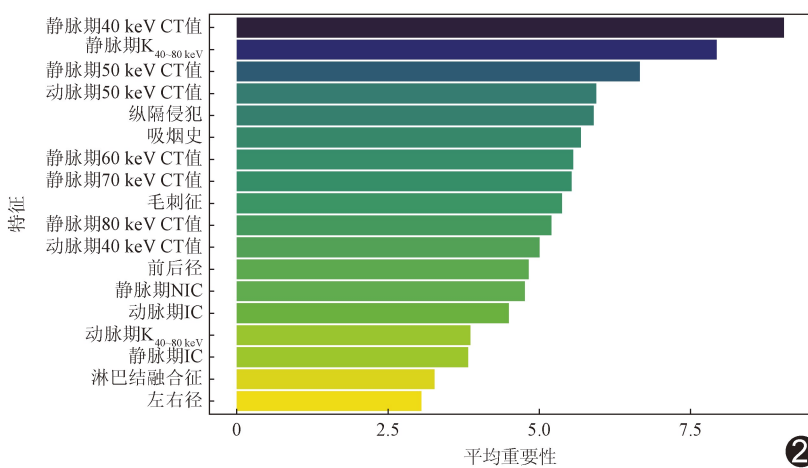


图 2 Boruta 算法筛选出的 18 个 PCCT 参数的重要性排序

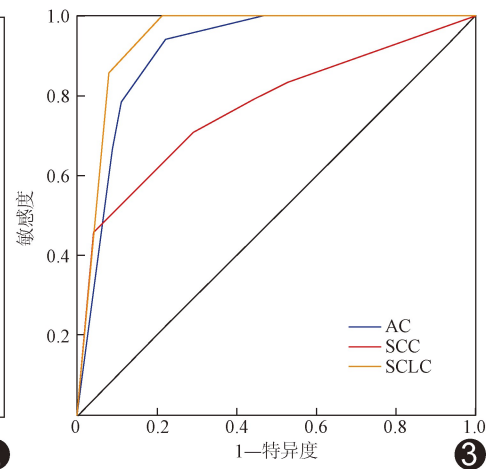


图 3 DT 模型预测 AC、SCC 及 SCLC 的 ROC 曲线

既往报道<sup>[8-10]</sup>一致;但各亚型肺癌常规 CT 表现存在重叠,需要依据客观、定量影像学指标辅助进行精确诊断。本研究 SCC 组和 SCLC 组病灶的动、静脉期低 keV CT 值和  $K_{40\sim 80\text{ keV}}$  均低于 AC 组。分析原因,低 keV 图像对于碘剂更敏感,AC 具有更高 CT 值和  $K_{40\sim 80\text{ keV}}$  提示其 IC 或组织密度更高,与其组织结构疏松、微血管密度大、利于对比剂渗入并滞留的病理特征相符<sup>[11]</sup>;SCLC 常伴广泛坏死,细胞排列紧密血供相对不足,导致以上参数值较低;而 SCC 常伴角化、中心坏死,细胞排列相对致密,故其 PCCT 参数值介于前两者之间<sup>[12]</sup>。此外,本研究 SCLC 组病灶的动、静脉期 IC 和 NIC 均最低,提示其相对乏血供或对比剂快速廓清;而 AC 组病灶在动脉期已表现较高的碘摄取,在静脉期的 IC 和 NIC 显著高于 SCC 及 SCLC,持续强化模式提示其微血管网络更丰富、利于对比剂积聚滞留<sup>[10,13]</sup>;SCC 组病灶的 IC 和 NIC 则介于 AC 与 SCLC 之间。

本研究以能充分利用多维度定量信息的 Boruta 算法<sup>[14]</sup>筛选出 18 个 PCCT 参数,以之构建的 DT 模型具有良好的整体诊断准确率,且区分不同亚型肺癌的 AUC 亦较好,表明基于 PCCT 定量参数构建的 ML 模型能于术前有效预测肺癌病理亚型,体现了 PCCT 作为数据源的技术优越性。

综上,基于 PCCT 定量参数构建的 DT 模型用于术前预测肺癌病理亚型具有较高价值。本研究的主要局限性:为单中心分析,仅针对 3 种常见病理亚型肺癌,样本量有限、分布不均衡,且缺乏外部验证;有待未来通过多中心、大样本研究进一步观察。

利益冲突:全体作者声明无利益冲突。

作者贡献:李蕊研究设计和实施、图像分析、撰写和修改文章;杨晓菲数据分析、经费支持;詹鹏超研究实施、统计分析;谷艳博和陈岩研究实施、图像处理;高剑波指导、审阅文章。

### [参考文献]

- [1] 雷洁, 杨阳, 邱晓辉. 基于动脉期增强 CT 影像组学模型预测小细胞肺癌无进展生存期[J]. 中国医学影像技术, 2024, 40(11): 1698-1703.
- [2] SIEGEL R L, GIAQUINTO A N, JEMAL A. Cancer statistics, 2024[J]. CA Cancer J Clin, 2024, 74(1):12-49.
- [3] HENDRIKS L E L, CORTIULA F, MARTINS-BRANCO D, et al. Updated treatment recommendations for systemic treatment: From the ESMO oncogene-addicted metastatic NSCLC living guideline[J]. Ann Oncol, 2025, 36(10):1227-1231.
- [4] HOWINGTON J, SOUTER L H, ARENBERG D, et al. Management of patients with early-stage non-small cell lung cancer: An American College of Chest Physicians clinical practice guideline[J]. Chest, 2025, 168(3):810-827.
- [5] CHEN B T, CHEN Z, YE N, et al. Differentiating peripherally-located small cell lung cancer from non-small cell lung cancer using a CT radiomic approach[J]. Front Oncol, 2020, 10:593.
- [6] ZHANG R, ZHENG H, LIN J, et al. Review of the application of dual-energy CT combined with radiomics in the diagnosis and analysis of lung cancer[J]. J Appl Clin Med Phys, 2025, 26(4):e70020.
- [7] MOHAMMADZADEH S, MOHEBBI A, KIANI I, et al. Direct comparison of photon counting-CT and conventional CT in image quality of lung nodules: A systematic review and meta-analysis[J]. Eur J Radiol, 2025, 183:111859.
- [8] WANG Z, LI M, HUANG Y, et al. Clinical and radiological characteristics of central pulmonary adenocarcinoma: A comparison with central squamous cell carcinoma and small cell lung cancer and the impact on treatment response[J]. Onco Targets Ther, 2018, 11:2509-2517.
- [9] 万绪明, 侯红军, 刘杰, 等. 双层探测器光谱 CT 在肺间质性病变合并肺癌病理亚型评估的应用价值[J]. 医学影像学杂志, 2023, 33(8):1373-1377.
- [10] 徐晓莉, 隋昕, 钟巍, 等. 双能量 CT 碘图定量参数联合 CT 征象建模在诊断肺癌病理亚型中的价值[J]. 中华放射学杂志, 2018, 52(11):823-828.
- [11] 徐艳秋, 杨晓光, 史洁. 双能 CT 定量参数在肺癌病理特征中的诊断价值[J]. 中国 CT 和 MRI 杂志, 2024, 22(9):168-169.
- [12] 吴冬梅, 梁腾飞, 赵合保, 等. 双源 CT 双能量多参数成像对肺癌病理类型诊断价值[J]. 医学影像学杂志, 2022, 32(9):1502-1506.
- [13] 皮亚文, 王广丽. 双能量 CT 在肺肿瘤诊断及治疗中的研究进展[J]. 国际医学放射学杂志, 2022, 45(6):700-705.
- [14] MAURYA N S, KUSHWAH S, KUSHWAHA S, et al. Prognostic model development for classification of colorectal adenocarcinoma by using machine learning model based on feature selection technique Boruta[J]. Sci Rep, 2023, 13(1):6413.