

DeepSurv model for predicting prognosis of stage I non-small cell lung cancer after CT-guided radiofrequency ablation

WANG Sibin^{1,2}, ZHAO Yi³, DU Peng⁴, ZHANG Jing², XIAO Yueyong², ZHANG Xiao^{2*}

(1. PLA Medical School, Beijing 100853, China; 2. Department of Radiology, the First Medical Center, Chinese PLA General Hospital, Beijing 100853, China; 3. Research and Development Department, Zoye Corporation, Qingdao 266100, China; 4. Department of Radiology, the Sixth Medical Center, Chinese PLA General Hospital, Beijing 100048, China)

[Abstract] **Objective** To observe the value of deep survival neural network (DeepSurv) model for predicting prognosis of stage I non-small cell lung cancer (NSCLC) after CT-guided radiofrequency ablation (RFA). **Methods** Totally 217 patients with stage I NSCLC who underwent CT-guided RFA were retrospectively enrolled and divided into training set ($n=109$), internal validation set ($n=46$) and external validation set ($n=62$). Pre-treatment gross tumor region (GTR) and post-treatment ablation region (PTZ) were automatically segmented on CT images using TotalSegmentator tool based on nnU-Net architecture. After deformable registration with DeedsBCV algorithm, a 5 mm isotropic expansion of GTR was used to construct the ideal ablation target zone, and three-dimensional target coverage ratio (TCR) was calculated. Optimal radiomics features were extracted from PTZ to derive radiomics score (Radscore). Multivariable Cox regression was performed to identify independent impact factors of local tumor progression (LTP). These factors, together with Radscore, were incorporated into DeepSurv model. Model discrimination and calibration were assessed using concordance index (C-index) and Brier score. **Results** Gender, fibrinogen (FIB), and three-dimensional TCR were all independent impact factors of LTP (HR [95% CI]=1.568 [1.020, 2.410], 1.540 [1.170, 2.027], and 0.110 [0.044, 0.273], respectively; all $P<0.05$) and were incorporated into DeepSurv model. C-index value of DeepSurv model for predicting LTP after CT-guided RFA for stage I NSCLC was 0.906, 0.809 and 0.861 in training, internal validation and external validation sets, respectively, the corresponding Brier score was 0.187, 0.177 and 0.182, indicating good predictive performance and calibration. **Conclusion** DeepSurv model integrating tumor radiomics features, three-dimensional TCR, gender and FIB could be used to accurately predict prognosis of stage I NSCLC after CT-guided RFA.

[Keywords] carcinoma, non-small-cell lung; radiofrequency ablation; prognosis; radiomics

DOI:10.13929/j.issn.1003-3289.2026.03.004

基于 DeepSurv 模型预测 I 期非小细胞肺癌经 CT 引导下射频消融后预后

王思斌^{1,2}, 赵毅³, 杜鹏⁴, 张晶², 肖越勇², 张肖^{2*}

(1. 中国人民解放军医学院, 北京 100853; 2. 中国人民解放军总医院第一医学中心放射诊断科, 北京 100853; 3. 山东卓业医疗科技有限公司研发部, 山东 青岛 266100; 4. 中国人民解放军总医院第六医学中心放射诊断科, 北京 100048)

[摘要] **目的** 观察基于深度生存神经网络(DeepSurv)模型预测 I 期非小细胞肺癌(NSCLC)经 CT 引导下射频消融(RFA)后预后的价值。**方法** 回顾性纳入 217 例接受 CT 引导下 RFA 的 I 期 NSCLC 患者, 分为训练集($n=109$)、内部

[基金项目] 军委后勤保障部卫生局技术产品研究项目(24BJZ16)。

[第一作者] 王思斌(1998—), 男(蒙古族), 内蒙古通辽人, 在读硕士。研究方向: 影像引导微创介入诊疗。E-mail: wsbin@foxmail.com

[通信作者] 张肖, 中国人民解放军总医院第一医学中心放射诊断科, 100853。E-mail: doczhang301@163.com

[收稿日期] 2025-12-30 **[修回日期]** 2026-03-14

验证集($n=46$)及外部验证集($n=62$)。采用 nnU-Net 架构 TotalSegmentator 工具分别于 CT 图像中自动分割肿瘤区, 获得治疗前肿瘤区(GTR)及治疗后消融区(PTZ), 经 DeedsBCV 算法弹性配准后, 以 GTR 外扩 5 mm 为理想消融靶区计算三维目标覆盖率(TCR); 于 PTZ 中提取并筛选最优影像组学特征, 计算影像组学评分(Radscore)。以多因素 Cox 回归分析筛选局部肿瘤进展(LTP)的独立影响因素, 联合 Radscore 输入 DeepSurv 模型, 以一致性指数(C-index)及 Brier 评分评估模型效能及校准度。**结果** 性别、纤维蛋白原(FIB)及三维 TCR 均为 LTP 独立影响因素[HR (95% CI) = 1.568 (1.020, 2.410)、1.540 (1.170, 2.027)、0.110 (0.044, 0.273), P 均 < 0.05]。将性别、FIB、三维 TCR 及 Radscore 输入 DeepSurv 模型, 其预测训练集、内部验证集及外部验证集 I 期 NSCLC 经 RFA 后 LTP 的 C-index 分别为 0.906、0.809 及 0.861, Brier 评分分别为 0.187、0.177 及 0.182, 提示其效能及校准度均良好。**结论** 基于 DeepSurv 模型融合肿瘤影像组学特征、三维 TCR 及患者性别、FIB 能准确预测 I 期 NSCLC 经 CT 引导下 RFA 后预后。

[关键词] 癌, 非小细胞肺; 射频消融; 预后; 影像组学

[中图分类号] R734.2; R814.42 [文献标识码] A [文章编号] 1003-3289(2026)03-0336-06

肺癌发病率与死亡率均居恶性肿瘤首位。受患者高龄或心肺功能不全等因素限制, 约 20%~30% 早期非小细胞肺癌(non-small cell lung cancer, NSCLC)无法耐受外科手术切除^[1-3], 使得 CT 引导下射频消融(radiofrequency ablation, RFA)等经皮热消融技术成为根治性治疗此类患者的重要选择^[4]; 而治疗后局部肿瘤进展(local tumor progression, LTP)为限制长期生存获益的关键^[5]。为降低消融靶区覆盖不足所致 LTP, 专家共识^[6-7]建议将治疗中所见磨玻璃密度(ground glass opacity, GGO)范围作为判断消融靶区的参考, 通过保留 > 5 mm 安全边界降低复发风险; 但临床多采用目视评估消融靶区二维覆盖程度, 易受呼吸运动形变及病灶空间复杂几何构型等的影响, 精确度不足且主观差异较大^[8-9]。本研究提出三维目标覆盖率(target coverage ratio, TCR), 旨在通过高精度图像配准量化消融区对目标靶区的覆盖程度^[10-13]。传统 Cox 比例风险模型为生存分析主流方法, 但对数线性假设及比例风险前提限制了其对高维特征及非线性交互关系的拟合能力^[14]。深度生存神经网络(deep survival neural network, DeepSurv)将 Cox 框架与神经网络表征学习相结合, 能有效处理复杂协变量, 显著提升非线性风险建模准确性^[15]。本研究观察基于 DeepSurv 模型预测 I 期 NSCLC 经 CT 引导下 RFA 后预后的价值。

1 资料与方法

1.1 研究对象 回顾性纳入 2019 年 1 月—2022 年 12 月中国人民解放军总医院第一医学中心(A 中心)及第六医学中心(B 中心)共 217 例接受 CT 引导下经皮 RFA 的 I 期 NSCLC 患者, 男 109 例、女 108 例, 年龄 25~86 岁、中位年龄 65 岁。纳入标准: ①依据国际肺癌研究协会(International Association for the Study of Lung Cancer, IASLC)第 9 版分期系统并经组织病理学证实为 I 期 NSCLC; ②接受 CT 引导下经皮根治性 RFA 治疗; ③RFA

前、后 CT 数据完整且质量好; ④临床资料及随访信息完整。排除标准: ①曾接受 NSCLC 相关治疗; ②合并严重感染或凝血功能障碍; ③存在远处转移; ④CT 图像质量欠佳、无法用于分割或配准分析。按 7:3 比例将 A 中心 155 例分为训练集($n=109$)与内部验证集($n=46$), 以 B 中心 62 例为外部验证集。本研究经医院医学伦理委员会批准(伦审第 S2023-695-01 号、HZKY-PJ-2023-37)并豁免知情同意书。

1.2 仪器与方法

1.2.1 RFA 程序与图像采集 采用 Cool-tip RFA 系统(Medtronic, 美国)于 CT(A 中心采用 Philips Brilliance、B 中心采用 GE Discovery CT 机)引导下实施 RFA。CT 扫描管电压 120 kV、管电流 250 mAs, 层厚及层间距均为 5 mm。消融过程中根据肿瘤大小及实时阻抗反馈动态调整消融参数, 功率 20~40 W, 持续时间 12~21 min。

1.2.2 图像后处理与三维量化 采用基于 nnU-Net 架构的 TotalSegmentator(v1.5.6)工具包自动于 CT 图像中分割治疗前肿瘤区(gross tumor region, GTR)及治疗后消融区(post-treatment ablation zone, PTZ); 采用 DeedsBCV 弹性配准算法对治疗前、后图像进行空间对齐, 确保其处于同一坐标系, 校正呼吸运动相关非线性形变; 利用形态学膨胀将 GTR 外扩 5 mm 作为理想消融靶区($PTV_{5\text{mm}}$)并计算三维 TCR:

$$\text{三维 TCR} = \frac{\text{Volume}(PTV_{5\text{mm}} \cap \text{PTZ})}{\text{Volume}(PTV_{5\text{mm}})}$$

对三维 TCR 进行二分类处理, 将 $\text{TCR} = 1$ 定义为完全覆盖、对应最小消融边界 ≥ 5 mm; $\text{TCR} < 1$ 定义为覆盖不足。见图 1。

1.3 提取影像组学特征 将全部 CT 图像重采样至 $1.0\text{mm} \times 1.0\text{mm} \times 1.0\text{mm}$ 各向同性体素, 采用 Z-score 进行归一化处理; 利用 PyRadiomics(v3.0.1)于

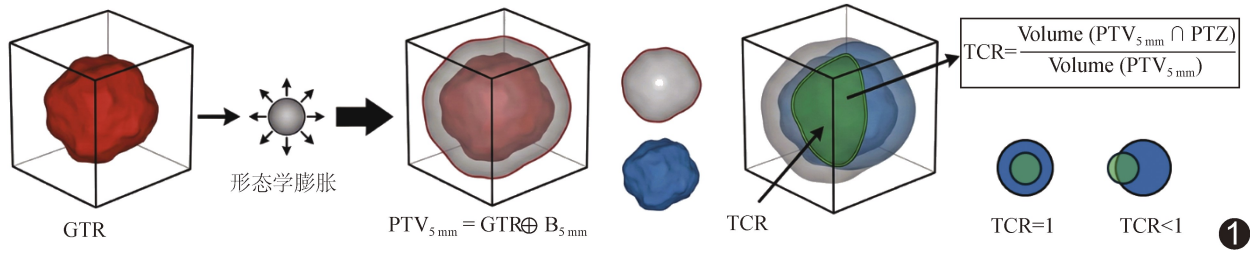


图 1 获取三维 TCR 流程图

PTZ 中提取 1 316 个影像组学特征,经 Pearson 相关性分析对高度相关特征去冗余后,采用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)算法结合 10 折交叉验证筛选出最优特征;计算影像组学评分(radiomics score, Radscore)。

1.4 构建 DeepSurv 模型 基于 PyTorch 框架构建 DeepSurv 模型,以最优影像组学特征、二分类三维 TCR 及临床资料作为输入,其中的连续变量经训练集标准化处理,于训练折内完成筛选特征与预处理。网络隐藏层采用引入双残差块的全连接结构,结合特征注意力模块以自适应调节不同模态特征的权重,以带 L2 正则化的 Cox 部分似然损失函数作为训练模型的优化目标。训练完成后于注意力模块分别导出特征的注意力系数及加权有效信号,用于评估不同特征对于预测 LTP 的相对贡献度。

1.5 随访 分别于治疗后 1、3、6 及 12 个月复查胸部 CT,此后每年随访 1 次。参考改良实体瘤疗效评价标

准,以目标病灶最大径之和增加 $\geq 20\%$ 或出现新发病灶为 LTP,并视为研究终点。以 RFA 至确诊 LTP 的时间间隔为无进展生存期(progression-free survival, PFS)。

1.6 统计学分析 采用 SPSS 26.0、R 4.5.0 及 Python 3.8 统计分析软件。以 $\bar{x} \pm s$ 表示正态分布计量资料,行独立样本 *t* 检验或单因素方差分析;以中位数(上下四分位数)表示偏态分布计量资料,行 Mann-Whitney U 检验或 Kruskal-Wallis H 检验。以 χ^2 检验比较计数资料。采用单因素及多因素 Cox 回归分析筛选 LTP 的独立影响因素,计算风险比(hazard ratio, HR);分别以一致性指数(concordance index, C-index)及 Brier 评分评估模型效能及校准度。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 临床资料及 CT 表现 对 217 例随访 23.0 (13.2,37.0)个月。各集间临床资料及 CT 表现差异均无统计学意义(P 均 > 0.05),见表 1。

表 1 217 例 I 期 NSCLC 患者临床资料及 CT 表现

集别	年龄(岁)	体质量指数(kg/m ²)	性别(例)		吸烟史(例)	高血压(例)	糖尿病(例)	高脂血症(例)	红细胞计数($\times 10^{12}/L$)	白细胞计数($\times 10^9/L$)
			男	女						
训练集(n=109)	64.0(56.0, 71.0)	24.00 \pm 3.13	53	56	48	52	13	23	4.20(4.00,4.60)	6.00(4.80,8.00)
内部验证集(n=46)	65.5(60.0, 69.8)	24.17 \pm 3.01	23	23	24	22	7	15	4.30(3.90,4.60)	6.10(4.80,6.90)
外部验证集(n=62)	67.0(61.0, 72.0)	24.36 \pm 3.27	33	29	30	33	11	15	4.30(4.00,4.60)	6.00(4.80,7.10)
F/H/ χ^2 值	2.846	0.262	0.336	0.926	0.533	1.133	2.323	0.817	0.409	
P 值	0.241	0.770	0.845	0.629	0.766	0.568	0.313	0.664	0.815	
集别	血小板计数($\times 10^9/L$)	凝血酶原时间(s)	活化部分凝血活酶时间(s)		凝血酶时间(s)	FIB(g/L)	D-二聚体(mg/L)			
训练集(n=109)	205.00(177.00,241.00)	13.00(12.50,13.40)	35.00(32.70,37.80)	16.60(15.90,17.10)	3.00(2.60,3.50)	0.30(0.20,0.50)				
内部验证集(n=46)	217.50(184.00,265.50)	13.20(12.70,13.70)	35.60(33.30,37.50)	16.60(16.00,17.00)	2.90(2.60,3.30)	0.30(0.20,0.40)				
外部验证集(n=62)	202.00(171.00,246.50)	13.00(12.50,13.60)	35.00(31.50,36.60)	16.40(15.90,16.90)	3.00(2.60,3.30)	0.20(0.20,0.40)				
F/H/ χ^2 值	2.984	2.602	0.806	1.476	0.188	2.809				
P 值	0.225	0.272	0.668	0.478	0.910	0.245				

待续

续表

集别	病灶直径(mm)	病理诊断(例)		病灶位置(例)					病灶性质(例)		
		腺癌	鳞癌	右肺上叶	右肺中叶	右肺下叶	左肺上叶	左肺下叶	实性	部分实性	磨玻璃
训练集 (n=109)	18.00(14.00,23.00)	99	10	39	10	16	32	12	47	33	29
内部验证集 (n=46)	17.50(15.00,25.80)	42	4	17	3	6	14	6	19	16	11
外部验证集 (n=62)	16.00(13.00,21.00)	54	8	24	3	12	15	8	25	18	19
F/H/ χ^2 值	3.849	0.737		2.626					0.855		
P 值	0.146	0.692		0.956					0.931		

注:FIB:纤维蛋白原(fibrinogen)

2.2 三维 TCR

217 例中,147 例三维 TCR<1,70 例三维 TCR=1。见图 2。

2.3 LTP 独立影响因素

Cox 回归分析结果示,性别、FIB、病灶性质及三维 TCR 均与 LTP 相关 [HR (95%CI)= 1.534 (1.008, 2.336)、1.490 (1.141, 1.944) 及 0.711(0.544, 0.928)、0.101(0.041, 0.248), P 均<0.05],其中性别、FIB 及三维 TCR 均为 LTP 的独立影响因素 [HR(95%CI)= 1.568(1.020, 2.410)、1.540(1.170, 2.027)、0.110(0.044, 0.273), P 均<0.05]。

2.4 模型效能

最终筛选出 11 个最优影像组学特征,包括 2 个一阶特征和 9 个纹理特征;按来源分为原始特征 3 个及高阶(小波变换)特征 8 个。将 Radscore 联合独立影响因素输入 DeepSurv 模型,可视化列线图见图 3。DeepSurv 模型

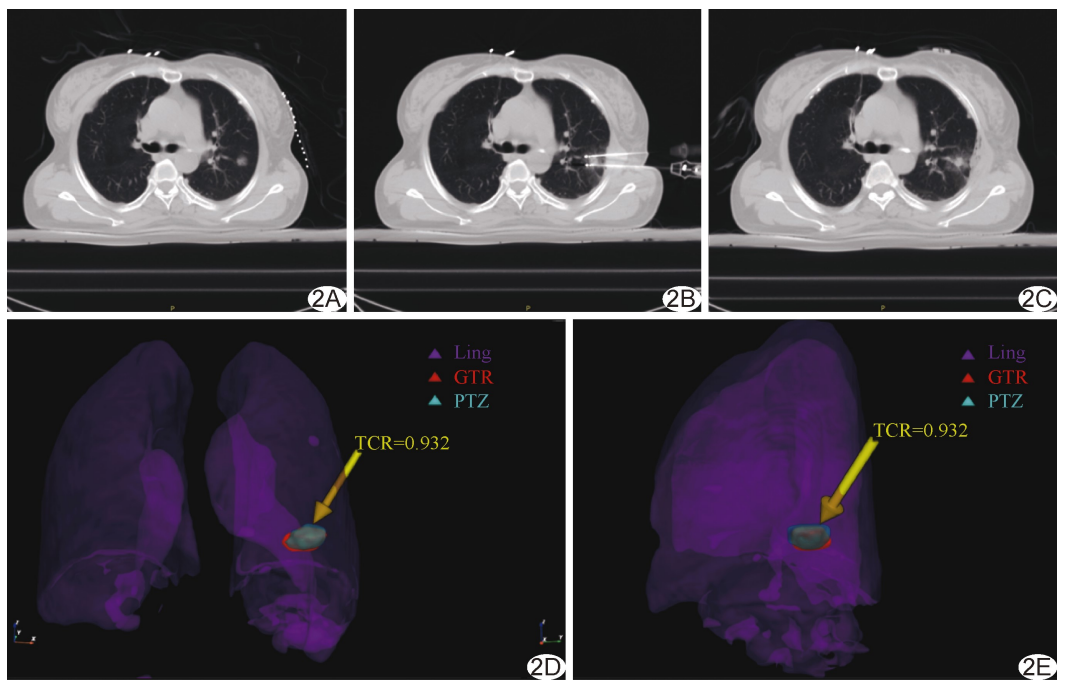


图 2 患者女,58 岁, I 期 NSCLC A. RFA 前胸部轴位 CT 图示左肺上叶尖后段混合磨玻璃密度结节; B. RFA 过程中双针夹击消融 CT 图; C. RFA 后即刻 CT 图示 GGO 形成; D、E. GTR 与 PTZ 三维融合图正位观(D)和侧位观(E),计算三维 TCR 为 0.932

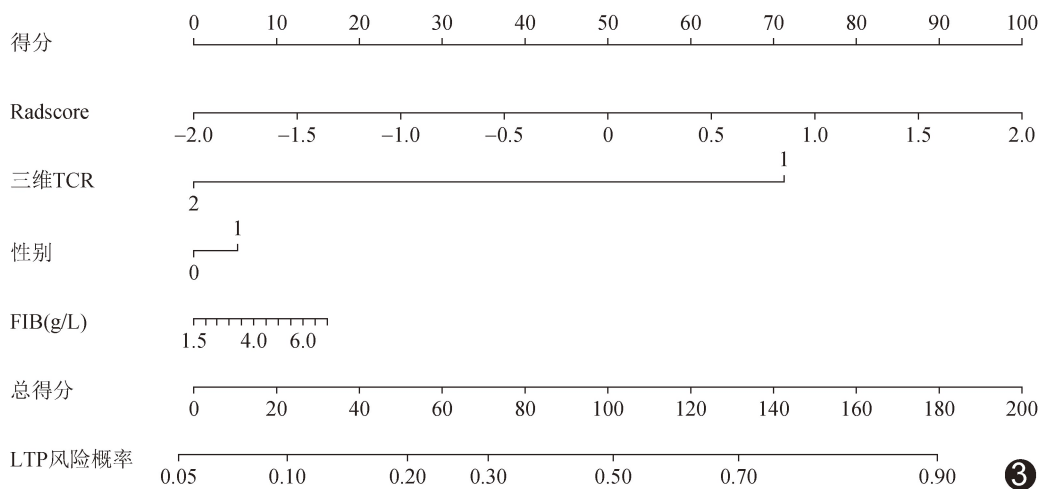


图 3 DeepSurv 模型可视化列线图

预测训练集、内部验证集及外部验证集 I 期 NSCLC 经 CT 引导下 RFA 后发生 LTP 的 C-index 分别为 0.906、0.809 及 0.861, Brier 评分分别为 0.187、0.177 及 0.182, 提示模型效能及校准度良好、预测误差较低且在集间相对一致。特征注意力分析显示, 影像组学主成分贡献最大, FIB 加权信号高于 TCR; TCR 注意力系数较稳定, 可作为物理消融安全边界的量化指标(图 4)。

3 讨论

DeepSurv 模型利用全连接结构与残差模块拟合协变量间的非线性交互, 可深度融合高维影像组学特征与临床数据。本研究基于 DeepSurv 框架, 整合三维 TCR、影像组学特征及临床资料构建的预测模型在跨中心验证中表现稳健, 提示其有助于实现精准风险分层。

本研究多因素 Cox 回归分析结果显示, 三维 TCR、FIB 及性别均为 I 期 NSCLC 经 RFA 后 LTP 的独立影响因素。三维 TCR 反映 PTZ 对 5 mm 外扩理想靶区的覆盖率, 其显著的保护效应[HR=0.110 (0.044, 0.273)]证实了足够的消融安全边界对于局部肿瘤控制的重要性。相比视觉评估传统二维图像, 基于全自动分割与弹性配准计算得出的三维 TCR 所受呼吸运动及复杂几何形态干扰显著降低, 能更客观地识别边缘覆盖不足; 三维 TCR<1 与 LTP 高风险密切相关, 表明评估三维边界对肿瘤预后分层具有重要意义。本研究发现基线 FIB 升高与高 LTP 风险有关, 可能由于 FIB 为凝血与炎症标志物, 其水平异常在一定程度上提示全身微环境有利于肿瘤进展, 可致消融边缘残留病灶复发^[16-18]。此外, 性别被纳入 DeepSurv 预测模型, 提示不同性别人群经 RFA 后肿瘤局部控制率存在差异, 考虑与基线肺功能、解剖结构差异或吸烟史等因素潜在相关。

综上, 基于 DeepSurv 模型融合肿瘤影像组学特征、三维 TCR 及患者性别、FIB 能准确预测 CT 引导下 RFA 治疗 I 期 NSCLC 预后。本研究的主要局限性: 为回顾性分析, 未纳入肿瘤病理分化程度、浸润特征、驱动基因突变状态及 Ki-67 增殖指数等重要生物学标志物, 限制了模型泛化能力; 样本量有限且腺癌占比较高。未来将通过前瞻性多中心队列进一步验证三

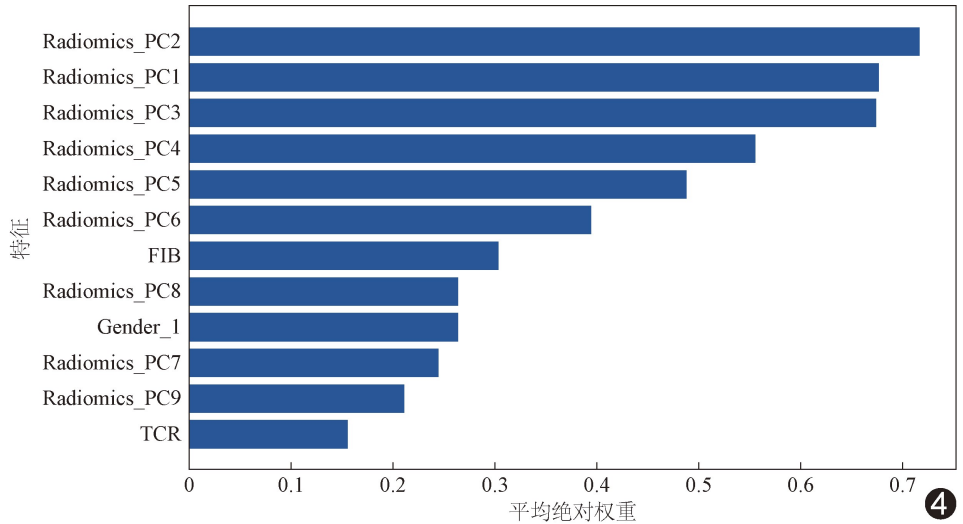


图 4 DeepSurv 模型输入变量贡献度排序图

维 TCR 驱动 DeepSurv 模型的临床可靠性, 并探索将该模型算法集成至微创治疗导航系统的可行性。

利益冲突: 全体作者声明无利益冲突。

作者贡献: 王思斌研究设计和实施、数据分析、撰写文章; 赵毅研究实施、修改文章; 杜鹏统计分析; 张晶图像处理; 肖越勇指导、审阅文章; 张肖研究设计、审阅文章、经费支持。

[参考文献]

- [1] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2024, 74(3):229-263.
- [2] CHOI J I. Medically inoperable stage I non-small cell lung cancer: Best practices and long-term outcomes[J]. Transl Lung Cancer Res, 2019, 8(1):32-47.
- [3] JONES G C, KEHRER J D, KAHN J, et al. Primary treatment options for high-risk/medically inoperable early stage NSCLC patients[J]. Clin Lung Cancer, 2015, 16(6):413-430.
- [4] YE X, FAN W, CHEN J H, et al. Chinese expert consensus workshop report: Guidelines for thermal ablation of primary and metastatic lung tumors[J]. Thorac Cancer, 2015, 6(1):112-121.
- [5] NAJAFI A, de BAERE T, PURENNE E, et al. Risk factors for local tumor progression after RFA of pulmonary metastases: A matched case-control study[J]. Eur Radiol, 2021, 31(7):5361-5369.
- [6] CHANG L K, YANG S M, CHIEN N, et al. 2024 multidisciplinary consensus on image-guided lung tumor ablation from the Taiwan Academy of Tumor Ablation [J]. Thorac Cancer, 2024, 15(20):1607-1613.
- [7] ZHANG T, AN C, LI W, et al. Consensus from the Asian conference on tumor ablation working group on image-guided

- thermal ablation for lung tumor: 10 key questions and answers[J]. *J Cancer Res Ther*, 2025, 21(6):1098-1109.
- [8] 上海医师协会介入医师分会, 上海市抗癌协会实体肿瘤聚焦诊疗专委会, 中国抗癌协会肿瘤微创治疗专委会, 等. 影像学引导下肺肿瘤多模态消融操作规范专家共识[J]. *中国介入影像与治疗学*, 2025, 22(6):369-374.
- [9] ZHU W, XU F, LOU K, et al. The impact of Inter-observation variation on radiomic features of pulmonary nodules[J]. *Front Oncol*, 2025, 15:1567028.
- [10] TANI S, TATLI S, HATA N, et al. Three-dimensional quantitative assessment of ablation margins based on registration of pre- and post-procedural MRI and distance map[J]. *Int J Comput Assist Radiol Surg*, 2016, 11(6):1133-1142.
- [11] KIM Y S, LEE W J, RHIM H, et al. The minimal ablative margin of radiofrequency ablation of hepatocellular carcinoma (> 2 and < 5 cm) needed to prevent local tumor progression: 3D quantitative assessment using CT image fusion[J]. *AJR Am J Roentgenol*, 2010, 195(3):758-765.
- [12] YOON J H, LEE J M, KLOTZ E, et al. Prediction of local tumor progression after radiofrequency ablation (RFA) of hepatocellular carcinoma by assessment of ablative margin using pre-RFA MRI and post-RFA CT registration [J]. *Korean J Radiol*, 2018, 19(6):1053-1065.
- [13] YAN P, TONG A N, NIE X L, et al. Assessment of safety margin after microwave ablation of stage I NSCLC with three-dimensional reconstruction technique using CT imaging [J]. *BMC Med Imaging*, 2021, 21(1):96.
- [14] KUITUNEN I, PONKILAINEN V T, UIMONEN M M, et al. Testing the proportional hazards assumption in cox regression and dealing with possible non-proportionality in total joint arthroplasty research: Methodological perspectives and review[J]. *BMC Musculoskelet Disord*, 2021, 22(1):489.
- [15] KATZMAN J L, SHAHAM U, CLONINGER A, et al. DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network[J]. *BMC Med Res Methodol*, 2018, 18(1):24.
- [16] WU X, YU X, CHEN C, et al. Fibrinogen and tumors[J]. *Front Oncol*, 2024, 14:1393599.
- [17] MITSUI S, TANAKA Y, DOI T, et al. Prognostic value of preoperative plasma fibrinogen levels in resected stage I non-small cell lung cancer [J]. *Thorac Cancer*, 2022, 13 (10): 1490-1495.
- [18] ZHAO H, ZHANG X, REN D, et al. Preoperative plasma fibrinogen combined with the platelet-to-lymphocyte ratio (F-PLR) serves as a prognostic indicator in patients with non-small cell lung cancer[J]. *Front Oncol*, 2025, 15:1587443.