

基于临床及CT深度影像特征构建胸部钝性损伤预后模型的研究

武鼎华 郑燕

西安市红会医院, 中国·陕西 西安 710000

摘要: 目的: 探讨整合临床与双期 CT 深度影像特征构建胸部钝性损伤 (BCT) 预后不良预测模型, 解决当前评估依赖主观经验、未覆盖动态损伤演变、泛化能力不足的问题。方法: 纳入医院 2020 年 1 月–2025 年 1 月收治的 284 例 BCT 患者, 分为训练集 200 例、外部验证集 84 例, 提取临床特征、传统影像组学及深度学习特征, 经多阶段筛选后构建多模态融合预后模型, 验证模型预测效能。结果: 训练集中急诊+临床+DLRscore 的 AUC 达 0.91, 优于单纯临床模型; 外部验证中该模型 AUC 为 0.78, 为各模型最高, delta 融合模型敏感度达 0.77, 为各模型最优。结论: 整合影像深度特征可提升 BCT 预后预测效能, 急诊融合模型泛化性最优, delta 融合模型更契合高危筛查需求, 具备临床应用潜力。

关键词: 胸部钝性损伤; CT 影像; 深度学习; 影像组学; 预后预测模型

Study on Constructing a Prognostic Model for Blunt Chest Trauma Based on Clinical and CT Deep Imaging Features

Wu Dinghua, Zheng Yan

Xi'an Red Cross Hospital, China Shaanxi Xi'an 710000

Abstract: Objective: To explore integrating clinical and dual-phase CT deep imaging features to build a model predicting poor prognosis in blunt chest trauma (BCT), addressing current issues where assessments rely on subjective experience, fail to cover dynamic injury changes, and have limited generalizability. Methods: 284 BCT patients admitted between January 2020 and January 2025 were included, divided into a training set (200 cases) and an external validation set (84 cases). Clinical features, traditional radiomics, and deep learning features were extracted, and after multi-stage screening, a multimodal fusion prognostic model was established to validate predictive performance. Results: In the training set, the emergency clinical DLRscore achieved an AUC of 0.91, outperforming the clinical-only model; in external validation, this model had an AUC of 0.78, the highest among all models. The delta fusion model had a sensitivity of 0.77, the best among all models. Conclusion: Integrating deep imaging features can enhance BCT prognosis prediction. The emergency fusion model shows the best generalizability, while the delta fusion model is better suited for high-risk screening, demonstrating clinical application potential.

Keywords: Blunt chest trauma; CT imaging; Deep learning; Radiomics; Prognostic prediction model

0 引言

胸部钝性损伤是创伤急诊高发危急症, 近年随交通事故与高处坠落事故增多, 发病率持续升高。该病损伤常呈进行性加重, 部分患者早期表现隐匿, 若不能早期识别预后不良高危人群, 易延误治疗升高死亡风险^[1]。目前临床评估多依赖经验性临床指标, 未深度挖掘 CT 影像信息, 也未关注损伤动态演变, 且多数研究缺乏高质量外部验证, 模型泛化能力不足, 因此本研究整合临床与双期 CT 深度特征构建预后模型, 填补该领域研究空白。

1 资料与方法

1.1 一般资料

本研究共纳入医院 2020 年 1 月–2025 年 1 月收治的 284 例 BCT 患者, 分为训练集 200 例、外部验证集 84 例。训练集平均年龄 60.13 ± 14.93 岁, 女性 85 例 (42.50%); 验证集平均年龄 54.30 ± 12.82 岁, 女性 22 例 (26.19%)。两组在年龄、性别、死亡率 (训练集 1 例 vs 验证集 13 例)、多发伤比例 (19.00% vs 40.48%)、肋骨骨折中位数 (4.00 vs 6.00)、急诊期肺挫伤体积比 (0.76% vs 4.70%) 及

delta 血气胸体积比 (0.34% vs -0.23%) 等方面差异均有统计学意义 ($P<0.05$)，提示队列间存在临床异质性，符合真实世界数据分布特征。全部患者均完成急诊与住院期两次 CT 扫描，影像覆盖范围自第 1 肋上 2 cm 至 L2 水平，确保病灶完整捕获。

1.2 纳入排除标准

纳入标准：①经临床与影像确诊的 BCT 患者；②急诊及住院期均行非增强胸部 CT 检查；③电子病历资料完整，可获取年龄、性别、多发伤史、住院时长及结局信息；④影像 DICOM 数据可上传至 uAI Research Portal 平台进行分割处理。

排除标准：①合并恶性肿瘤或活动性感染性肺炎（需与肺挫伤鉴别）；② CT 图像质量差致 ROI 无法可靠勾画；③随访缺失或死亡时间不明；④既往有严重肺部基础疾病影响体积测量。

1.3 模型构建方法

本研究采用多模态影像-临床融合策略构建预后预测模型。首先从电子病历提取人口学与创伤相关变量，再由两位医师（A：2 年经验；B：15 年经验）在肺窗下以伪点击交互式算法半自动勾画 ROI，B 对随机子集复核，A1 个月后重分割评估可重复性；所有 DICOM 图像经强度离散化、最大最小归一化与重采样三步预处理后输入分析平台。传统影像组学特征（RF）提取 4528 维，深度学习特征（DLF）基于 3D DenseNet 网络导出 4096 维，delta-RF/delta-DLF 为两期特征的相对净变化量（（随访值-基线值）/基线值），更稳健反映病灶进展速率。特征筛选分三阶段：① ICC>0.8 保留高重复性特征；② Spearman $|r|>0.8$ 剔除高度共线性特征；③结合 Select K-best ($P<0.05$)、mRMR 与 LASSO 回归（10 折 CV）确定最优子集，最终急诊模型选 5 RF+9 DLF，住院模型 8 RF+8 DLF，delta 模型 8 RF+7 DLF。构建单模态基模型后输出概率，通过 LR 元模型生成融合评分 DLRscore，整合临床独立危险因素（肋骨骨折数、多发伤）得到最终融合预后模型^[2]。

1.4 观察指标

主要结局为预后不良，定义为住院时间 >7 d 或院内死亡，阈值依据既往文献及本队列总体中位数设定，确保临床可操作性。次要指标包括各模型的 AUC、敏感度、特异度、准确率、阳性预测值（PPV）及 F1 分数；模型校准性通过校准曲线评估，临床效用以决策曲线分析量化，模型增量价值采用净重新分类指数（NRI）与综合判别改善指数（IDI）检验。所有特征均经标准化处理，避免量纲干

扰，指标体系覆盖从静态解剖到动态病理演化，契合 BCT 损伤进行性加重的临床特点。

1.5 统计学分析

全部分析在 R 4.3.0 环境完成，显著性水平设为 $\alpha=0.05$ 。连续变量正态性经 Kolmogorov-Smirnov 检验判定：符合正态分布者以均值 $\pm$ 标准差表示，组间比较采用两独立样本 t 检验；非正态者以中位数（四分位间距）描述，使用 Mann-Whitney U 检验。分类变量以频数（百分比）呈现，组间比较采用卡方检验或 Fisher 确切概率法。单因素 Logistic 回归初筛临床变量，多因素回归确认独立预测因子，机器学习超参数经 5 折交叉验证优化，以外部验证集评估最终泛化能力。

2 结果

2.1 患者基线特征分布

验证集患者更年轻、女性比例更低、死亡率显著升高，且肺挫伤与血气胸负荷明显高于训练集，反映外部队列整体病情更重；delta 肺挫伤体积比组间差异达统计学意义 ($P=0.027$)，变化方向相反（训练集中位 +0.34%，验证集 -0.23%），提示损伤演变模式存在中心间差异，凸显外部验证的必要性。所有变量无缺失值，保障后续建模完整性，基线特征分布见表 1。

2.2 影像特征筛选与模型构建参数

特征工程共提取 RF 4528 维、DLF 4096 维，经过滤后剩余 4359 RF 与 5267 DLF 进入最终筛选，LASSO 回归确定各模态最优特征数见表 2，其中 delta-RF 中 wavelet 域特征占比最高，反映纹理变化对动态损伤的敏感性，DLF 为数据驱动的网络激活特征，无需人为设计。本研究采用的分层融合架构兼顾了模型可解释性与预测能力，如表 2 所示。

2.3 多模型性能比较与外部验证结果

如表 3 所示，所有融合模型在训练集 AUC 均 >0.85，其中急诊-临床-DLRscore 达 0.91，显著优于单纯临床模型 ($\Delta AUC=0.16, P<0.001$)；但在外部验证集，AUC 普遍下降至 0.70-0.78 区间，提示过拟合风险存在，尤以住院模型最甚 (0.70)。需注意的是，delta-模型在验证集中敏感度最高 (0.77)，特异度略低 (0.65)，反映其对进展性恶化患者的识别优势；PPV 稳定在 0.79-0.84，临床实用性较强。

3 讨论

单纯临床模型仅纳入年龄、性别、肋骨骨折数、多发伤等公认危险因素，虽可初步区分预后，但存在明显局

表 1 患者基线特征分布

变量	总体 (n=284)	训练集 (n=200)	验证集 (n=84)	P 值
年龄 (岁)	58.40 ± 14.56	60.13 ± 14.93	54.30 ± 12.82	0.002
女性 [例 (%)]	107 (37.68)	85 (42.50)	22 (26.19)	0.010
死亡 [例 (%)]	14 (4.93)	1 (0.50)	13 (15.48)	<0.001
住院时长 (天)	7.00 (6.00,11.00)	7.00 (6.00,11.00)	7.00 (3.00,12.00)	0.002
多发伤 [例 (%)]	72 (25.35)	38 (19.00)	34 (40.48)	<0.001
肋骨骨折数量	5.00 (3.00,7.00)	4.00 (3.00,6.00)	6.00 (3.00,8.25)	<0.001
肺挫伤比 (%) 急诊	1.60 (0.24,5.19)	0.76 (0.14,2.80)	4.70 (1.87,10.42)	<0.001
肺挫伤比 (%) 住院	1.81 (0.66,5.08)	1.31 (0.60,2.91)	5.67 (1.46,10.69)	<0.001
肺挫伤 delta (%)	0.23 (-1.01,1.35)	0.34 (-0.48,1.29)	-0.23 (-3.41,1.52)	0.027
血气胸比 (%) 急诊	0.00 (0.00,1.41)	0.00 (0.00,0.66)	0.69 (0.00,5.01)	0.002
血气胸比 (%) 住院	2.02 (0.00,6.86)	1.81 (0.00,5.11)	3.96 (0.45,11.69)	<0.001
血气胸 delta (%)	1.57 (0.00,4.81)	1.50 (0.00,4.02)	2.11 (-0.20,7.60)	0.918

注：数据以均值 ± 标准差或中位数 (P25,P75) 表示；P 值基于两独立样本 t 检验或 Mann-Whitney U 检验。

表 2 影像特征筛选表

模态	特征类型	最优特征数量	代表性特征 (前 3 项)
急诊	RF	5	original_shape_Maximum2DDiameterRow, normalize_glszm_ZoneEntropy, mean_glcml_Autocorrelation
	DLF	9	channel0_54, channel2_51, channel4_49
住院	RF	8	original_gldm_DependenceNonUniformityNormalized, boxmean_gldm_DependenceNonUniformityNormalized, wavelet_glcml_wavelet-LLL-JointEnergy
	DLF	8	channel5_39, channel5_49, channel19_35
delta	RF	8	original_glrml_LongRunEmphasis, original_glrml_RunVariance, original_glrml_ShortRunHighGrayLevelEmphasis
	DLF	7	channel26_51, channel29_54, channel2_41

表 3 多模型性能比较与外部验证结果

模型组合	队列	AUC (95%CI)	敏感度	特异度	准确率	PPV	F1 Score
临床	训练集	0.75 (0.69 - 0.82)	0.77	0.68	0.73	0.70	0.73
	验证集	0.74 (0.63 - 0.84)	0.70	0.71	0.70	0.80	0.75
急诊 + 临床 +DLRscore	训练集	0.91 (0.87 - 0.95)	0.84	0.85	0.85	0.85	0.84
	验证集	0.78 (0.68 - 0.87)	0.66	0.71	0.68	0.80	0.72
住院 + 临床 +DLRscore	训练集	0.88 (0.84 - 0.93)	0.78	0.79	0.79	0.79	0.78
	验证集	0.70 (0.58 - 0.81)	0.68	0.65	0.67	0.77	0.72
delta+ 临床 +DLRscore	训练集	0.89 (0.84 - 0.94)	0.82	0.85	0.84	0.84	0.83
	验证集	0.74 (0.62 - 0.85)	0.77	0.65	0.73	0.79	0.78

限：一是依赖医师经验，主观性强；二是无法捕捉 CT 中肉眼不可见的细微病理变化，对早期隐匿进展患者识别能力差，也未体现病灶的动态演变趋势^[3]。本研究采用多模态融合策略，整合传统影像组学、深度学习深度特征与双期动态变化量，结合临床危险因素构建模型，兼顾了解释性与预测能力。

本研究结果显示，融合模型的整体预测性能优于传统单纯临床模型，与既往 AI 影像模型在创伤预后研究中的结论一致，原因在于融合模型既保留了临床常用预测信息，又挖掘了 CT 影像的隐藏特征，同时加入病灶动态变化信息，能更全面反映 BCT 的损伤严重程度与进展趋势^[4]。本研究中，急诊 + 临床 +DLRscore 在训练集 AUC 达 0.91，

较单纯临床模型提升 0.16，各指标均衡性良好；外部验证阶段，尽管受两组队列异质性影响（验证集患者病情更重，死亡率、肺挫伤负荷均显著高于训练集），所有模型性能均有一定下降，但急诊 + 临床 +DLRscore 的 AUC 仍达 0.78，为所有模型最高，证实该模型泛化能力最优，更适配真实世界临床应用场景^[5]。delta+ 临床 +DLRscore 虽 AUC 为 0.74，略低于急诊融合模型，但敏感度达 0.77、F1 分数 0.78，均为所有模型最高。临床实践中，漏诊高危患者的危害远高于误诊低危患者，该模型的高敏感度特性恰好契合临床高危人群早筛的需求，漏诊率更低，临床应用价值突出^[6]。

综上，本研究构建了整合临床特征、双期 CT 传统影

像组学与深度学习特征的 BCT 预后预测模型,证实加入影像深度特征可显著提升 BCT 预后不良的预测效能,其中急诊融合模型泛化能力最优,delta 融合模型识别敏感度最高,二者各有优势,具备临床应用潜力^[7]。后续仍需扩大样本量,纳入多中心不同层级医院的数据进一步验证优化,助力临床早期识别高危患者,改善 BCT 患者预后。

参考文献:

[1] 惠智艳,王小军,沙海旺等. ICU 严重钝性胸部损伤机械通气时间延长的危险因素分析[J]. 临床急诊杂志, 2025(6):363-368.

[2] 代霞. 胸部 CT 对老年稳定钝性胸部创伤患者临床诊断中的效果研究[J]. 中文科技期刊数据库(全文版)医药卫生, 2024(002):017-018.

[3] 王秋兰,姚猛飞,陈检明等. 不同年龄段胸部创伤

致伤原因及特征分析[J]. 创伤外科杂志, 2023, 25(12):929-933.

[4] 陈志强,徐欣,裴磊. 胸部 CT 在老年稳定钝性胸部创伤临床诊断中的应用[J]. 浙江创伤外科, 2023, 28(4): 758-761.

[5] 蒋如芳,杜建成,王胜峰等. 经尸体解剖病理证实钝性胸部损伤致心肌挫伤 1 例[J]. 心电与循环, 2023, 42(2):187-190.

[6] 燕鑫鑫,李文娟. 血必净治疗胸部钝性创伤所致肺挫伤的临床疗效及对患者免疫功能,炎症反应,氧化应激的影响[J]. 现代临床医学, 2025(3):17-18.

[7] 李树满,李建国,刘立志等. 胸腔延期置管引流治疗钝性创伤性中等量血胸的临床价值[J]. 天津医药, 2023, 51(2):213-215.