·临床研究·

# 基于 MRI 影像组学的胶质瘤 *TERT* 基因突变状态预测模型的构建与验证

戴许豪，马瑞爽，葛小琴，胡雨晴，朱挺，任江平，杨继明，周瑛瑛，陶庆松

**【摘要】目的**　构建并验证用于术前预测胶质瘤端粒酶逆转录酶（*TERT*）基因突变状态的可解释机器学习模型。**方法**　回顾性收集 2019 年 1 月至 2025 年 1 月宁波大学附属第一医院收治的高级别胶质瘤患者 100 例，按 7∶3 比例分为训练集和验证集，分别应用 XGB、RF、LGBM、KNN 和 GBDT 机器学习算法进行建模，并通过 SHAP 方法评估模型可解释性。**结果**　5 种模型准确率 0.667 ～ 0.740，其中 LGBM 最高（0.740），XGB 和 RF 次之（0.733），XGB 精确率最佳（0.741）。综合各项指标，XGB 模型整体性能最优，LGBM 和 RF 也具有较好预测能力，而 GBDT 和 KNN 相对表现较弱。**结论**　XGB 可解释机器学习模型可在术前无创预测高级别胶质瘤 *TERT* 基因突变状态。

**【关键词】** 胶质瘤；MRI；机器学习；*TERT*

DOI:10.3969/j.issn.1671-0800.2026.01.006

**【中图分类号】** R739.41　**【文献标志码】** A　**【文章编号】** 1671-0800(2026)01-0024-04

胶质瘤是中枢神经系统最常见的原发性恶性肿瘤之一，按 WHO 分级可分为低级别和高级别胶质瘤[1-2]。胶质母细胞瘤（glioblastoma，GBM）是最常见且预后极差的高级别胶质瘤，具有高度的基因和细胞异质性，中位生存期仅约 14 个月[3-4]。分子标志物在胶质瘤诊断和预后评估中发挥重要作用，主要包括染色体臂 1p/19q 共缺失、异柠檬酸脱氢酶（*IDH*）和端粒酶逆转录酶（*TERT*）启动子突变、表皮生长因子受体（*EGFR*）扩增及+7/−10 染色体改变等[5]。约 70%的 GBM 患者存在 *TERT* 启动子突变，是 GBM 分子分型中的关键基因[6-8]。最新研究发现，精氨酸琥珀酸裂解酶（ASL）通过调控 *TERT* 表达和端粒酶活性影响肿瘤生长，为针对 ASL 和相关代谢产物富马酸的靶向治疗提供理论基础[9]。因此，对于高级别胶质瘤患者，尤其是老年、体弱无法耐受手术或拒绝手术者，*TERT* 基因突变的无创预测具有重要的临床价值和广阔的应用前景。本研究通过基于术前 MRI 影像组学特征构建多种可解释机器学习模型，系统比较不同算法在高级别胶质瘤 *TERT* 基因突变预测中的性能差异，并结合 SHAP 方法揭示关键特征对模型决策的贡献，旨在为 *TERT* 突变状态的术前无创评估提供可靠工具和理论依据，现报道如下。

## 1　资料与方法

1.1　一般资料　回顾性收集 2019 年 1 月至 2025 年 1 月于宁波大学附属第一医院手术治疗的高级别胶质瘤患者 100 例，均具备完整术前 MRI 及术后 1p/19q、*IDH*、*TERT* 等分子分型检测结果。纳入标准：18 ～ 80 岁，组织学确诊，术前 1 个月内行 MRI，未行新辅助治疗，分子检测完整。排除标准：病理信息缺失，病灶过小，影响 MRI 影像组学特征的术前治疗史，复发再手术，影像质量不佳。按 7∶3 的比例随机分为训练集和验证集。本研究获得宁波大学附属第一医院医学伦理委员会批准（2025 研第 145RS 号-01），豁免/免除知情同意。

1.2　图像获取与感兴趣区域（ROI）勾画　所有患者在相同条件下接受 Siemens Magnetom Trio 3.0 T 钆增强 $T_1$WI 扫描，所有扫描均在患者接受任何可能改变肿瘤生物学行为或治疗反应的干预措施之前进行。患者的肉眼肿瘤区域（gross tumor volume，GTV）被确定为 ROI。为保证 GTV 勾画的准确性与一致性，由两名具有 10 年及以上经验的放疗科医师使用 3D-Slicer 精确勾画，确保一致性与数据质量。

1.3　特征提取　采用 PyRadiomics 提取 7 类影像组

**基金项目：** 浙江省医药卫生科技计划项目（2024KY328）；浙江省基础公益研究计划项目（LY23H160008）

**作者单位：** 315010 宁波，宁波大学附属第一医院（戴许豪，马瑞爽，葛小琴，胡雨晴，朱挺，任江平，杨继明，周瑛瑛，陶庆松）；宁波大学医学部（胡雨晴，陶庆松）

**通信作者：** 陶庆松，Email：tqs3932@163.com

学特征：(1)一阶统计量(first order statistics)；(2)形状(shape)；(3)灰度共生矩阵(Glcm)；(4)灰度及带矩阵(Glszm)；(5)灰度游程长度矩阵(Glrlm)；(6)邻域灰度差分矩阵(Ngtdm)；(7)灰度依赖性矩阵(Gldm)。

1.4　特征筛选　从 1 410 个组学特征中，经最小冗余最大相关(mRMR)算法，筛选出前 100 个相关特征。使用 LASSO 算法进行进一步降维与筛选，通过 L1 正则化将不重要特征的系数压缩至零，再结合 10 折交叉验证，保留少数重要特征，同时绘制特征相关性热力图以剔除强相关冗余特征，优化模型性能，见图 1。

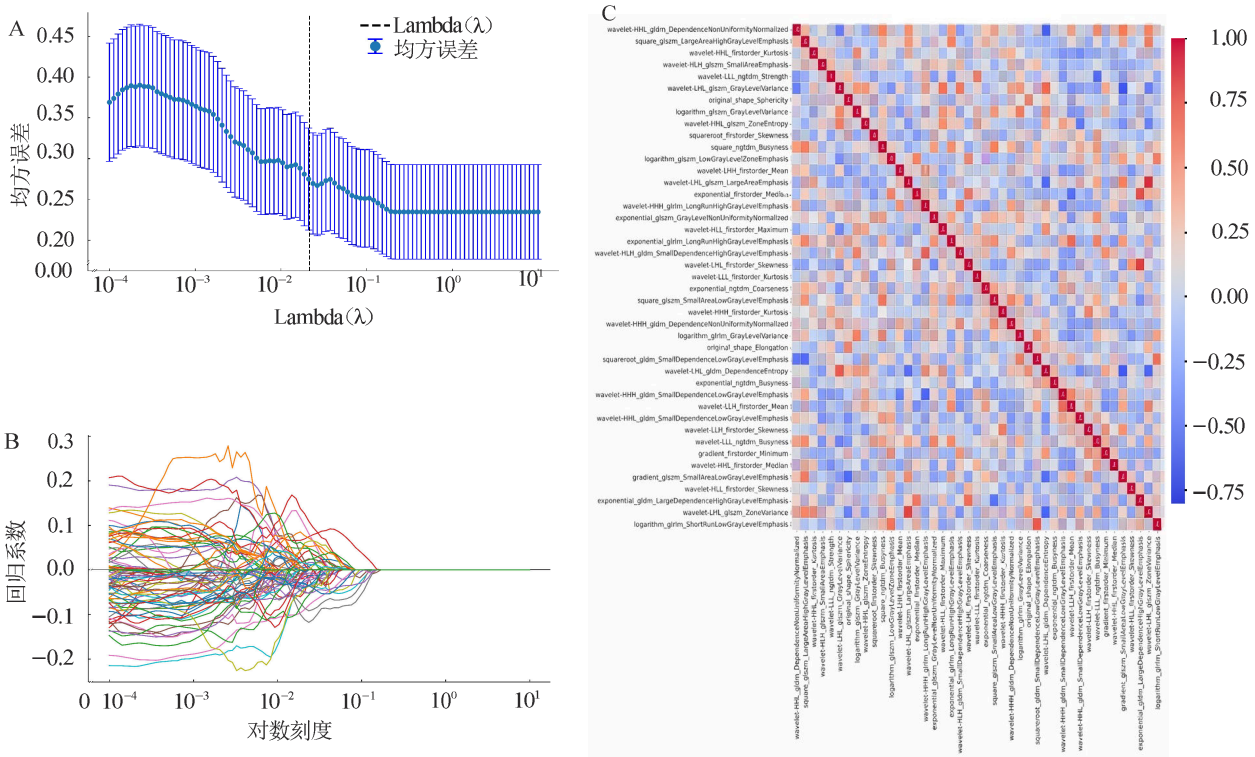
1.5　模型构建与评估　使用 XGB、RF、LGBM、KNN 和 GBDT 机器学习模型分别对钆增强快速稳态梯度回波 $T_1$ WI 图像中提取的最相关特征，构建 *TERT* 突变状态预测模型。调节模型参数(如 L1 正责化参数 C 值和最大迭代次数等)使模型达到最优的性能。最后，通过受试者工作特征(ROC)曲线下面积(*AUC*)、准确率、特异性和敏感性等指标评估各模型的性能，并筛选出最优的预测模型。

## 2　结果

构建并评估了 RF、LGBM、KNN、GBDT 和 XGB 机器学习模型用于胶质瘤 *TERT* 基因突变状态预测。模型准确率 0.667 ～ 0.740，其中 LGBM 最高(0.740)，XGB 和 RF 次之(0.733)，XGB 精确率最佳 (0.741)。综合各项指标，XGB 模型整体性能最优，LGBM 和 RF 也具有较好预测能力，而 GBDT 和 KNN 相对表现较弱，见图 2 ～ 4、表 1。

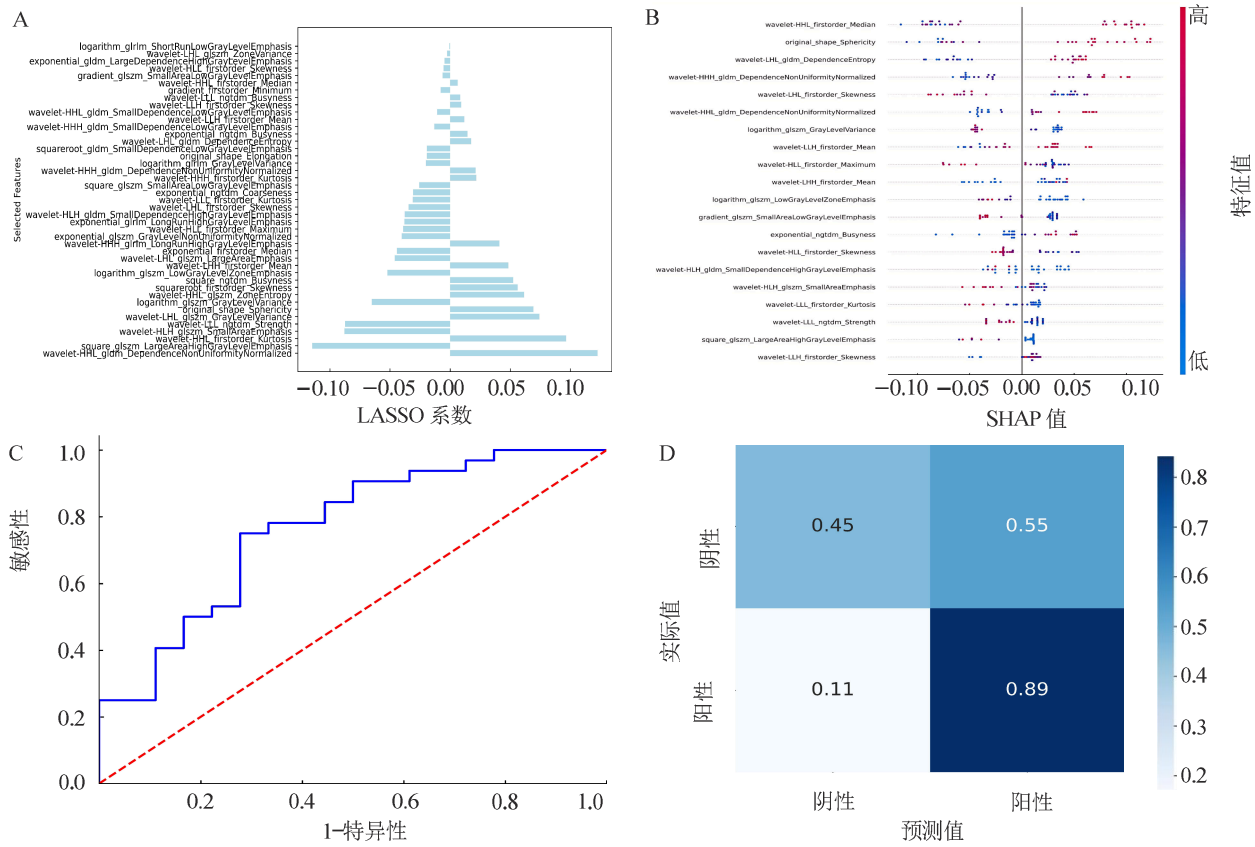
## 3　讨论

分子分型在胶质瘤的预后评估、治疗决策及个体化管理中作用显著，但现有检测依赖术后组织病理和分子检测，创伤性强且对深部或功能区邻近病灶患者风险高，可能引发神经功能损伤，并存在费用高、结果延迟等问题，限制了术前分子信息获取和精准治疗方案制定。随着人工智能的发展，基于影像组学结合机器学习/深度学习的胶质瘤预测研究不断增多[10-15]。已有学者探讨 *TERT* 突变的临床意义，并在低级别胶质瘤中利用影像组学预测其状态[16-17]，但 *TERT* 突变在高级别胶质瘤的预后评估及 GBM 诊断中价值更突出。Zhang 等[18]利用多参数 MRI 构建深度学习模型预测 GBM 的 *TERT* 突变状态，但其对大量数据的依赖及特征可解释性不足的“黑箱”问



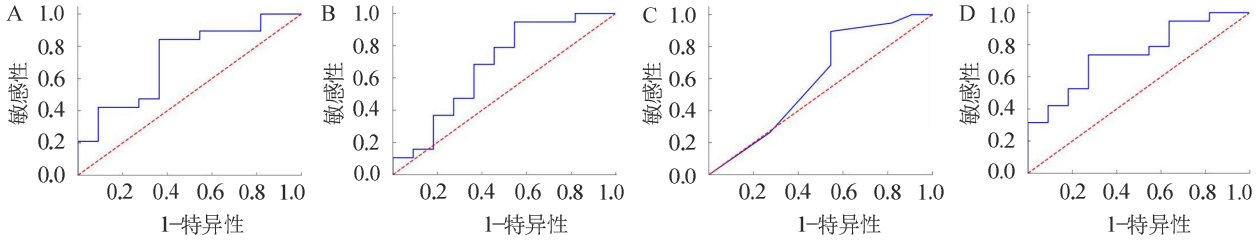

注：A 为不同正则化参数λ下的均方误差变化趋势，确定最优λ；B 为各特征回归系数随α(对数刻度)变化路径；C 为最终特征相关性热力图，颜色表示相关系数，高相关(＞ 0.9)特征可考虑剔除以降低多重共线性

图 1　LASSO 回归特征选择与相关性可视化



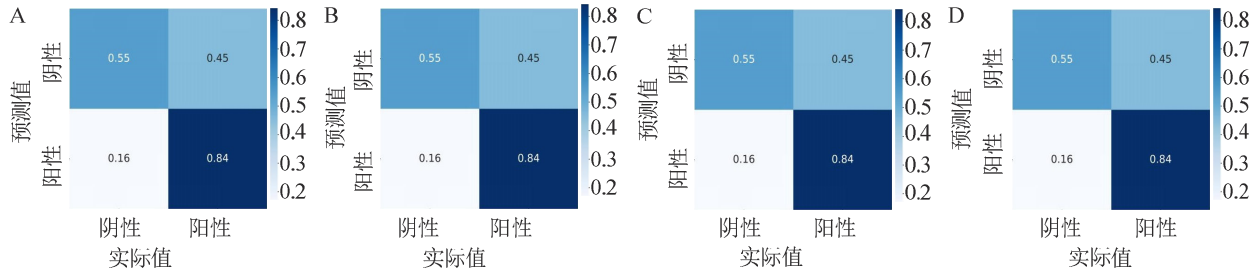

注：A 为 LASSO 筛选的重要影像组学特征及系数；B 为 SHAP 汇点图显示特征对模型输出的方向与强度；C 为 XGB 模型 ROC 曲线（*AUC*=0.764），提示良好区分能力；D 为验证集混淆矩阵，颜色深浅代表分类正确率

**图 2**　高级别胶质瘤 TERT 突变状态预测最优 XGB 模型可视化结果




注：A、B、C、D 分别代表 RF、LGBM、KNN 和 GBDT 模型的 ROC 曲线

**图 3**　其余各模型的 ROC 曲线




注：A、B、C、D 分别代表 RF、LGBM、KNN 和 GBDT 模型的混淆矩阵

**图 4**　其余各模型的混淆矩阵图

题，限制了临床推广。

本研究构建并评估了 RF、LGBM、KNN、GBDT 和 XGB 机器学习模型用于 *TERT* 突变状态的预测。结果显示，XGB 模型在准确率、精确率及综合性能指标上均表现最佳，准确率达 0.733，精确率高达 0.741，体现了其较强的预测能力。LGBM 和 RF 表现也较为稳定，准确率分别为 0.740 和 0.733，但在精确率和综合指标上略逊于 XGB。GBDT 和 KNN 的

表 1　各机器学习模型性能评估表

| 模型 | 准确率 | 精准率 | 召回率 | F1 分数 | 曲线下面积 |
|---|---|---|---|---|---|
| RF | 0.733 | 0.714 | 0.694 | 0.700 | 0.718 |
| LGBM | 0.740 | 0.725 | 0.688 | 0.696 | 0.675 |
| KNN | 0.733 | 0.727 | 0.675 | 0.683 | 0.596 |
| GBDT | 0.667 | 0.634 | 0.603 | 0.603 | 0.742 |
| XGB | 0.733 | 0.741 | 0.750 | 0.732 | 0.764 |

整体性能相对较弱。基于 XGB 模型，本研究进一步利用 SHAP 值对关键特征进行了可解释性分析，揭示了各特征对预测结果的具体贡献，增强了模型的透明度和临床应用价值。整体而言，XGB 模型结合关键特征的解释，为 *TERT* 突变的精准预测提供了有效工具和理论支持。

本研究仍存在以下局限，需在后续工作中加以改进。首先，样本量较小，可能影响模型的泛化能力与稳定性，尤其在医学图像分析中，扩大数据规模通常可显著提升模型性能。未来应引入来自不同中心的深度学习模型，以增强模型的性能与鲁棒性。此外，本研究仅基于 MRI 图像，单一影像模态亦难以全面反映肿瘤微环境及分子分型生物学特征，后续可考虑引入分析肿瘤代谢物成分的磁共振波谱（MRS）以及显示白质纤维束的扩散张量成像（DTI）等多模态影像，有望进一步提升 *TERT* 突变预测模型的性能。

综上所述，本研究构建了一种具备良好可解释性、可适应不同临床需求的胶质瘤 *TERT* 基因突变状态的机器学习预测模型。基于 MRI 影像组学特征，分别采用 RF、LGBM、KNN、GBDT 和 XGB 机器学习算法进行建模，并通过 SHAP 值分析实现特征贡献的可视化解释。研究强调根据不同预测任务选择最合适模型的重要性，提出多模型联合评估策略以提升预测准确性与临床适用性。未来，通过多中心合作进一步扩大样本规模、融合多模态数据并引入深度学习方法，有望进一步增强模型性能与临床推广价值。

**利益冲突**　所有作者声明无利益冲突

**作者贡献声明**　戴许豪：模型构建，论文撰写；马瑞爽：研究指导；葛小琴：临床指导，ROI 肿瘤区勾画；胡雨晴：数据收集；朱挺：ROI 勾画，核查；任江平：数据审查；杨继明：写作指导；周瑛瑛：数据分析；陶庆松：论文审阅，经费支持

## 参　考　文　献

[1] Weller M, Wen P Y, Chang S M, et al. Glioma[J]. Nat Rev Dis Primers, 2024, 10(1): 33.

[2] van Solinge T S, Nieland L, Chiocca E A, et al. Advances in local therapy for glioblastoma—taking the fight to the tumour[J]. Nat Rev Neurol, 2022, 18(4): 221-236.

[3] Yeo A T, Rawal S, Delcuze B, et al. Single-cell RNA sequencing reveals evolution of immune landscape during glioblastoma progression [J]. Nat Immunol, 2022, 23(6): 971-984.

[4] Louis D N, Perry A, Wesseling P, et al. The 2021 WHO classification of tumors of the central nervous system: A summary[J]. Neuro Oncol, 2021, 23(8): 1231-1251.

[5] Eckel-Passow J E, Lachance D H, Molinaro A M, et al. Glioma groups based on 1p/19q, *IDH*, and *TERT* Promoter mutations in tumors[J]. N Engl J Med, 2015, 372(26): 2499-2508.

[6] Diplas B H, He X J, Brosnan-Cashman J A, et al. The genomic landscape of TERT promoter wildtype-IDH wildtype glioblastoma [J]. Nat Commun, 2018, 9(1): 2087.

[7] Fujimoto K, Arita H, Satomi K, et al. TERT promoter mutation status is necessary and sufficient to diagnose IDH-wildtype diffuse astrocytic glioma with molecular features of glioblastoma[J]. Acta Neuropathol, 2021, 142(2): 323-338.

[8] 张科平,许洁,林丹义,等. TERT 启动子突变检测在胶质瘤中的意义[J]. 诊断病理学杂志,2022,29(7):628-630,634.

[9] Brennan C W, Verhaak R G, McKenna A, et al. The somatic genomic landscape of glioblastoma[J]. Cell, 2013, 155(2): 462-477.

[10] Zhou Q W, Yan X J, Guo Y W, et al. Machine learning algorithms for predicting glioma patient prognosis based on CD163+FPR3+ macrophage signature[J]. NPJ Precis Onc, 2024, 8: 201.

[11] Xu H S, Fan Z, Jiang S, et al. Integrating multiplex immunohistochemistry and machine learning for glioma subtyping and prognosis prediction[J]. MedComm, 2025, 6(5): e70138.

[12] Feng S Y, Zhu L, Qin Y, et al. Machine learning model reveals the role of angiogenesis and EMT genes in glioma patient prognosis and immunotherapy[J]. Biol Direct, 2024, 19(1): 113.

[13] Shi Z M, Ge X, Li M D, et al. Argininosuccinate lyase drives activation of mutant TERT promoter in glioblastomas[J]. Mol Cell, 2022, 82(20): 3919-3931.e7.

[14] Ghadimi D J, Vahdani A M, Karimi H, et al. Deep learning-based techniques in glioma brain tumor segmentation using multi-parametric MRI: A review on clinical applications and future outlooks[J]. J Magn Reson Imaging, 2025, 61(3): 1094-1109.

[15] Mahootiha M, Tak D, Ye Z Z, et al. Multimodal deep learning improves recurrence risk prediction in pediatric low-grade gliomas[J]. Neuro Oncol, 2025, 27(1): 277-290.

[16] 钟昱珏,马笑笑,段祺,等. 脑胶质瘤 TERT 启动子突变的影像学研究进展[J]. 四川大学学报(医学版),2024,55(6):1350-1356.

[17] 卢俊,李祥,黎海亮. 基于 ADC 和增强 MRI 的影像组学模型预测低级别胶质瘤 TERT 启动子突变状态[J]. 放射学实践,2022,37(5):538-542.

[18] Zhang H B, Zhang H W, Zhang Y Z, et al. Deep learning radiomics for the assessment of telomerase reverse transcriptase promoter mutation status in patients with glioblastoma using multiparametric MRI [J]. J Magn Reson Imaging, 2023, 58(5): 1441-1451.

**收稿日期：2025-08-09**

**（本文编辑：吴迪汉）**