

推进人工智能大模型在医疗领域中的应用

朱文珍, 吕文志, 陈敏

【摘要】 人工智能大模型在医疗领域展现出广泛应用前景,尤其是在精准诊疗、药物研发和医学教育方面取得显著进展。在精准诊疗中,大模型通过自然语言处理和计算机视觉技术辅助病历管理、影像分析和病理诊断,显著提高效率与准确性,同时减少医生负担。在药物研发中,大模型通过知识图谱、基因靶点预测、蛋白质结构建模等手段加速药物发现,并优化临床试验流程。在医学教育中,大模型助力试题生成、虚拟案例学习和临床实践模拟,提升教学效率和学生能力。然而,大模型的医疗应用仍面临评估标准不统一、伦理风险、技术缺陷和国际技术封锁等挑战。未来需加强伦理规范建设、技术优化和国产医学大模型研发,以推动大模型技术更好地服务临床需求和医学进步。

【关键字】 人工智能大模型;精准诊疗;药物研发;医学教育;自然语言处理;计算机视觉

【中图分类号】 TP18 **【文献标志码】** A **【文章编号】** 1000-0313(2025)01-0005-04

DOI:10.13609/j.cnki.1000-0313.2025.01.002

开放科学(资源服务)标识码(OSID):



全球人口老龄化和慢性疾病高发正给医疗系统带来前所未有的挑战^[1,2]。而人工智能作为引领新一轮科技革命和产业变革的战略性技术,为现有医疗领域提供了新的解决方案^[3]。传统深度学习模型高度依赖精确标注的数据,受限于医疗数据标注困难及疾病样本稀缺。然而,以 Transformer 为基础的大模型通过利用海量无标注数据进行预训练,建立了通用基础模型,微调后可适配多种相关下游任务,克服了这一局限。其广泛覆盖医学需求的潜力使其更贴近临床实践,并成为下一代医疗技术的重要方向^[4,5]。当前,大模型已经能够通过跨领域、跨模态的学习,处理更加复杂的医学任务,如影像数据、基因组数据分析和临床决策支持等。此外,随着大模型在医学领域的不断进化,未来将会进入基于大数据的动态个性化治疗阶段,不仅通过医学大数据预测疾病发展趋势,还能根据患者的具体情况生成个性化的治疗方案,为精准医疗提供更强有力的支持^[6]。本文从人工智能大模型在医疗中的应用角度,阐述大模型在精准诊疗、药物研发和医学教育等方面的现状,旨在探讨大模型技术在医疗应用中的挑战和未来发展方向。

大模型在精准诊疗中的应用

随着电子病历的普及,临床记录的工作量显著增加,医师平均需要花费约 60% 的时间在总结和整理文本信息上^[7]。然而,以 ChatGPT 为代表的通用大语言

模型已展示出接近经验丰富的临床医师的文本处理能力,尤其是在用于支持敏感的临床决策时^[7]。在分诊过程中,大模型与护士协作可显著提升门诊接待的效率和准确性^[8]。在复杂临床诊断和疾病分析中,GPT-4 展现出优异的效能,并能节省约 60% 的诊断时间^[9]。尽管如此,通用大模型在医学诊疗信息的深入理解方面仍存在不足,因此许多研究者致力于构建专门的医疗数据语料库,如 MedQA、PubMedQA、MMLU 和 MultiMedQA 等。医学影像检查报告中包含大量文本信息,借助大模型的强大文本处理能力,可有效提升报告撰写的效率、规范性和准确性。多个研究团队对比了不同通用大模型(如 ChatGPT、LLaMa、Bard 等)在报告内容质控^[10]、自动结构化^[11]、影像所见生成诊断结论^[12]、报告结论简化^[13]和疾病诊断^[14]等方面的应用,结果表明大模型有助于优化放射科工作流程,可将诊断报告的生成时间有效降低约 58%^[15]。

医学影像占据临床数据量约 90%^[16],大模型在其中也有广泛的应用。例如,针对影像分割任务,基于通用视觉模型 SAM 微调的医学影像分割大模型 MedSAM 使分割准确性提升 52.3%^[17]。在疾病诊断方面,数字病理影像大模型发展迅速,基于社交媒体图像和知识的病理语言-图像预训练模型(PLIP),在零样本学习中对多种疾病的诊断准确性达到 0.565~0.832 的 F1 评分^[18]。此外,针对真实临床病理图像中的跨尺度、多模态、高噪声等挑战,研究人员基于 LongNet 算法采取两阶段级联结构,开发了全切片尺度的数字病理大模型 GigaPath,在 9 个癌症分型和 17 项病理学任务中取得 25 个任务的领先成绩^[19]。通过结合弱监督学习和 Transformer 技术,CHIEF 模型在癌症检

作者单位: 430030 武汉,华中科技大学同济医学院附属同济医院放射科(朱文珍、吕文志);100730 北京,北京医院放射科(陈敏)

作者简介: 朱文珍(1969—),女,湖北孝感人,博士,主任医师,主要从事中枢神经影像学、人工智能及分子影像研究工作。

测、肿瘤起源预测和生存预测等任务中表现优异,相较于 SOTA 方法提升了 36.1%^[20]。基于 MAE 架构的眼底大模型 RETFound 通过自监督学习,使用无标注的眼底图像进行训练,在多个公开数据集的眼部疾病诊断和预后预测中表现最佳,并节省了 90% 的数据标注量和 80% 的模型训练时间^[21]。此外,结合深度卷积神经网络与大语言模型的糖尿病护理和筛查系统 DeepDR-LLM 将糖尿病视网膜病变的转诊准确率从 81.0% 提升至 91.3%^[22]。类似方法还被应用于超声^[23]、内镜^[24]等领域。放射影像领域还出现了视觉指令自回归文本生成模型,通过指令调优实现了报告文本与 2D 或 3D 影像的无缝衔接,并开源了首个放射影像大模型 RadFM^[25]。此外,基于扩散模型构建的跨模态生成大模型 MINIM,可通过文本生成 CT、MRI 和眼底等多种影像,有助于缓解医学影像研究数据量不足的问题^[26]。

此外,在手术过程中,大模型可结合手术机器人提供实时的手术指导和辅助。例如,Embodied AI 多模态手术大模型和 MicroNeuro 微创脑手术柔性机器人系统,可以辅助医生进行微创脑手术,提高手术的精准度和安全性。在康复指导方面,智能机器人辅助设备用于脑血栓患者的手部运动和感知训练^[27]。而针对重大疾病引发的心理问题,基于心理学专业数据微调的大模型 Psy-LLM 可即时准确地响应患者^[28]。

综上所述,大模型可通过自然语言处理和计算机视觉分析协助病历记录、影像分析和病理诊断,帮助医生快速制定治疗方案。同时,大模型的应用让医生有更多时间与患者沟通,提供更具同理心的人文关怀,形成标准化、精准化的诊疗流程,显著提升患者的医疗体验。

大模型赋能药物研发的应用进展

药物在疾病治疗中具有不可替代的作用,而新药的研发通常耗资巨大,历时长久。为加速药物研发,不同研究者已从疾病机制、药物合成和临床试验等多个角度进行了深入探索。不同于通用大模型对视觉和文本的理解,药物研发领域的大模型还需具备特定领域的语言学习能力,如处理 SMILES 编码的分子结构及 FASTA 格式的基因序列等。

理解疾病机制是药物开发的关键起点,旨在确定潜在的药物靶点。大模型通过文献分析构建知识图谱,揭示基因与疾病之间的联系,有助于理解疾病变异并识别不同患者之间疾病机制的潜在差异^[29]。例如研究者通过双向编码预训练方式捕捉传统语言与 DNA 的语义联系,建立整合基因信息的靶点预测模型 DNA-BERT^[30]。迁移学习和掩码自监督学习方法也

被用以构建加速疾病治疗的靶点识别大模型 Geneformer,该模型通过计算机模拟成功识别了心肌病的候选治疗靶点^[31]。此外,通过微调可以精确建模基因网络,理解复杂的动态关系,并扩展到单细胞水平的注释,如 scMulan、scGPT 和 xTrimoscFoundation 等大模型。

靶点验证是靶点识别后的关键步骤。研究者结合大语言模型与互联网搜索、代码执行和实验自动化等工具,开发了化学合成智能体 Coscientist,能够自主设计和执行复杂化学实验,成功实现 Sonogashira 和 Suzuki-Miyaura 交叉偶联反应^[32]。在蛋白质结构预测领域,AlphaFold 模型引入 Evoformer 组件,即使在无相似结构的情况下也能产生原子级别的准确预测^[33]。另一项研究通过无监督学习,从大型蛋白质序列数据库中训练出人工蛋白质预测模型 ProGen,突破了天然蛋白质预测的瓶颈^[34]。此外,基于序列建模的混合模型 StripedHyena 基于整个原核基因组训练,形成了涵盖 DNA、RNA 和蛋白质三种模态的多模态大模型 Evo,具备跨模态的零样本预测能力^[35]。

临床试验是药物开发的最后阶段,对于评估候选药物的安全性和有效性至关重要。然而,将患者匹配到合适的临床试验具有挑战性,且容易出现人为错误。一些研究结合单次提示、选择推理和思维链技术微调 InstructGPT,用于自动分析临床试验的入组标准,协助医师筛选患者,可减少高达 90% 的工作量^[36]。另一智能体 TrialGPT 集成候选试验筛选、配对分析和重排序子模块,在患者招募中减少了 42.6% 的筛选时间,并显著提升了匹配效率^[37]。

综上所述,将大模型整合到药物研发领域标志着一个重大的范式转变,为理解疾病机制、促进药物发现和优化临床试验流程提供了新的方法。基于 AI 在其中的巨大应用价值,诺贝尔 2024 年化学奖颁发给 AI 领域的权威专家。随着大模型的持续发展,药物研发流程有望进一步加速并降低成本,从而提升人类健康水平。

大模型促进医学教育的应用进展

大模型在目前医学教育中发挥着重要的作用,主要体现在辅助教师准备教学试题、帮助医学生进行临床实践模拟学习。例如,基于 ChatGPT 的大模型能够显著提高临床案例相关多项选择题的生成效率^[38]。此外,大模型通过互动案例展示出在前临床阶段帮助医学生学习基础医学知识并加强专业学习的潜力^[39]。还有研究指出,GPT-4 在医学测试中表现优异,可能成为成绩较低的医学生的重要学习工具^[40]。与此同时,大模型能够提供灵活的虚拟临床场景,通过即时反

帮助帮助学生识别弱点和优势,从而提升临床实践能力^[41]。另一项研究将已发表的 36 个临床场景输入大模型,发现该模型在鉴别诊断、诊断测试、最终诊断及总体管理决策方面表现出色,这进一步证明了大模型在帮助医师学习和辅助临床决策中的作用^[42]。此外,大模型生成的病例场景可用于进行交互式测验,帮助医学生提高其制定诊断和治疗计划的能力^[43]。

大模型在医疗应用中面临的问题与展望

尽管大模型在医疗领域的多个方面取得了突破性进展,但在成为实际临床应用工具之前,仍面临诸多挑战:①评价标准不统一。目前,大模型的评估严重依赖人为设定的规则进行公开数据集测评或临床研究评估,缺乏临床随机对照试验的严格评估标准;②伦理风险。大模型在医疗应用中通常需要大量含有患者个人敏感信息的医疗数据,不当使用可能导致隐私泄露;③技术缺陷。与通用大模型类似,医疗大模型也存在知识幻觉的问题,且准确的回答通常需要精确的提示词设计;④技术封锁。欧美国家对高端 GPU 的封锁日益严峻,深刻影响了全球 AI 产业格局。我国凭借庞大的人口基数和丰富的医疗数据,具备开发大模型的独特优势,并且存在强烈的医学大模型应用需求,深入开展国产医学大模型的研究将成为当下的必然趋势。

参考文献:

- [1] Beard JR, Officer A, De Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing[J]. *Lancet*, 2016, 387(10033): 2145-2154.
- [2] Chen X, Giles J, Yao Y, et al. The path to healthy ageing in China: a Peking University-Lancet commission [J]. *Lancet*, 2022, 400(10367): 1967-2006.
- [3] 郭华源, 刘盼, 卢若谷, 等. 人工智能大模型医学应用研究[J]. *中国科学: 生命科学*, 2024, 54(3): 482-506.
- [4] Moor M, Banerjee O, Abad ZSH, et al. Foundation models for generalist medical artificial intelligence[J]. *Nature*, 2023, 616(7956): 259-265.
- [5] 陈晓红, 刘浏, 袁依格, 等. 医疗大模型技术及应用发展研究[J]. *中国工程科学*, 2023: 1-12.
- [6] Zhou H, Gu B, Zou X, et al. A survey of large language models in medicine: progress, application, and challenge [J]. *ArXiv*, 2023, abs/2311.05112.
- [7] Van Veen D, Van Uden C, Blankemeier L, et al. Adapted large language models can outperform medical experts in clinical text summarization[J]. *Nat Med*, 2024, 30(4): 1134-1142.
- [8] Wan P, Huang Z, Tang W, et al. Outpatient reception via collaboration between nurses and a large language model: a randomized controlled trial[J]. *Nat Med*, 2024, 30(10): 2878-2885.
- [9] Kanjee Z, Crowe B, Rodman A. Accuracy of a generative artificial intelligence model in a complex diagnostic challenge [J]. *Jama*, 2023, 330(1): 78-80.
- [10] Gertz RJ, Dratsch T, Bunck AC, et al. Potential of GPT-4 for detecting errors in radiology reports: implications for reporting accuracy[J]. *Radiology*, 2024, 311(1): e232714.
- [11] Lehnen NC, Dorn F, Wiest IC, et al. Data extraction from free-text reports on mechanical thrombectomy in acute ischemic stroke using ChatGPT: a retrospective analysis [J]. *Radiology*, 2024, 311(1): e232741.
- [12] Sun Z, Ong H, Kennedy P, et al. Evaluating GPT-4 on impressions generation in radiology reports [J]. *Radiology*, 2023, 307(5): e231259.
- [13] Doshi R, Amin KS, Khosla P, et al. Quantitative evaluation of large language models to streamline radiology report impressions: a multimodal retrospective analysis [J]. *Radiology*, 2024, 310(3): e231593.
- [14] Cozzi A, Pinker K, Hidber A, et al. BI-RADS category assignments by GPT-3.5, GPT-4, and Google Bard: A multilanguage study [J]. *Radiology*, 2024, 311(1): e232133.
- [15] Bhayana R, Nanda B, Dehkharghanian T, et al. Large language models for automated synoptic reports and resectability categorization in pancreatic cancer [J]. *Radiology*, 2024, 311(3): e233117.
- [16] 萧毅, 刘士远. 产学研用协作促进中国医学影像 AI 产业健康发展(英文) [J]. *Chinese Medical Sciences Journal*, 2019, 34(2): 84-88, 158.
- [17] Ma J, He Y, Li F, et al. Segment anything in medical images [J]. *Nat Commun*, 2024, 15(1): 654.
- [18] Huang Z, Bianchi F, Yuksekogonul M, et al. A visual-language foundation model for pathology image analysis using medical Twitter [J]. *Nat Med*, 2023, 29(9): 2307-2316.
- [19] Xu H, Usuyama N, Bagga J, et al. A whole-slide foundation model for digital pathology from real-world data [J]. *Nature*, 2024, 630(8015): 181-188.
- [20] Wang X, Zhao J, Marostica E, et al. A pathology foundation model for cancer diagnosis and prognosis prediction [J]. *Nature*, 2024, 634(8035): 970-978.
- [21] Zhou Y, Chia MA, Wagner SK, et al. A foundation model for generalizable disease detection from retinal images [J]. *Nature*, 2023, 622(7981): 156-163.
- [22] Li J, Guan Z, Wang J, et al. Integrated image-based deep learning and language models for primary diabetes care [J]. *Nat Med*, 2024, 30(10): 2886-2896.
- [23] Jiao J, Zhou J, Li X, et al. USFM: A universal ultrasound foundation model generalized to tasks and organs towards label efficient image analysis [J]. *Med Image Anal*, 2024, 96: 103202.
- [24] Boers TGW, Fockens KN, Van Der Putten JA, et al. Foundation models in gastrointestinal endoscopic AI: impact of architecture, pre-training approach and data efficiency [J]. *Med Image Anal*, 2024, 98: 103298.
- [25] Wu C, Zhang X, Zhang Y, et al. Towards generalist foundation model for radiology [J]. *ArXiv*, 2023, abs/2308.02463.
- [26] Wang J, Wang K, Yu Y, et al. Self-improving generative foundation model for synthetic medical image generation and clinical applications [J]. *Nat Med*, 2024. DOI: 10.1038/s41591-024-03359-y.
- [27] Ranzani R, Lambercy O, Metzger JC, et al. Neurocognitive robot-assisted rehabilitation of hand function: a randomized control trial on motor recovery in subacute stroke [J]. *J Neuroeng Rehabil*, 2020, 17(1): 115.

- [28] Lai T, Shi Y, Du Z, et al. Psy-LLM: scaling up global mental health psychological services with AI-based large language models[J]. ArXiv, 2023, abs/2307.11991.
- [29] Pun FW, Ozerov IV, Zhavoronkov A. AI-powered therapeutic target discovery[J]. Trends Pharmacol Sci, 2023, 44(9): 561-572.
- [30] Ji Y, Zhou Z, Liu H, et al. DNABERT: pre-trained bidirectional encoder representations from transformers model for DNA-language in genome[J]. Bioinformatics, 2021, 37(15): 2112-2120.
- [31] Theodoris CV, Xiao L, Chopra A, et al. Transfer learning enables predictions in network biology[J]. Nature, 2023, 618(7965): 616-624.
- [32] Boiko DA, Macknight R, Kline B, et al. Autonomous chemical research with large language models[J]. Nature, 2023, 624(7992): 570-578.
- [33] Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3[J]. Nature, 2024, 630(8016): 493-500.
- [34] Madani A, Krause B, Greene ER, et al. Large language models generate functional protein sequences across diverse families[J]. Nat Biotechnol, 2023, 41(8): 1099-1106.
- [35] Nguyen E, Poli M, Durrant MG, et al. Sequence modeling and design from molecular to genome scale with Evo[J]. Science, 2024, 386(6723): eado9336.
- [36] Hamer D, Schoor P, Polak TB, et al. Improving patient Pre-screening for clinical trials: assisting physicians with large language models[J]. ArXiv, 2023, abs/2304.07396.
- [37] Jin Q, Wang Z, Floudas CS, et al. Matching patients to clinical trials with large language models[J]. Nat Commun, 2024, 15(1): 9074.
- [38] Kiyak YS, Kononowicz AA. Case-based MCQ generator: A custom ChatGPT based on published prompts in the literature for automatic item generation[J]. Med Teach, 2024, 46(8): 1018-1020.
- [39] Safranek CW, Sidamon-Eristoff AE, Gilson A, et al. The role of large language models in medical education: applications and implications[J]. JMIR Med Educ, 2023, 9: e50945.
- [40] Nori H, King N, McKinney SM, et al. Capabilities of GPT-4 on medical challenge problems[J]. ArXiv, 2023, abs/2303.13375.
- [41] Fernandes AC, Souto M. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine[J]. N Engl J Med, 2023, 388(25): 2399-2400.
- [42] Rao A, Pang M, Kim J, et al. Assessing the utility of ChatGPT throughout the entire clinical workflow: development and usability study[J]. J Med Internet Res, 2023, 25: e48659.
- [43] Zhui L, Yhap N, Liping L, et al. Impact of large language models on medical education and teaching adaptations[J]. JMIR Med Inform, 2024, 12: e55933.

(收稿日期:2024-12-12)

基于 CT 影像组学评估胃腺癌人表皮生长因子受体-2 表达状态的价值

王素雅, 詹鹏超, 邢静静, 梁盼, 岳松伟, 张永高, 高剑波

【摘要】 目的:基于 CT 影像组学及临床特征建立可无创性评估胃腺癌人表皮生长因子受体 2 (HER-2) 表达状态的预测模型, 并验证其效能。**方法:**回顾性分析本院 263 例经病理确诊为胃腺癌患者的临床资料, 其中 HER-2 阳性患者 49 例, 阴性 214 例。将所有患者以 7:3 的比例随机分为训练集($n=185$)和验证集($n=78$)。基于 3D Slicer 软件和门脉期 CT 图像手动勾画病灶感兴趣区(ROI), 并提取影像组学特征。在训练集中, 比较 HER-2 阳性与阴性组患者的临床特征差异, 采用多因素 Logistic 回归确定临床独立预测因子, 建立临床模型。基于最小绝对收缩和选择算子(LASSO)回归算法构建影像组学模型, 计算影像组学分数(Radscore)。结合临床独立预测因子和 Radscore 构建联合模型。根据 ROC 曲线的曲线下面积(AUC)评估模型的预测效能, 绘制校准曲线评价模型预测概率与真实概率之间的一致性, 采用决策曲线(DCA)分析模型的临床价值。**结果:**肿瘤厚径($OR=1.04, P=0.033$)、cT 分期($OR=2.39, P=0.038$)、cN 分期($OR=2.15, P=0.046$)为 HER-2 阳性表达的临床独立预测因子。在训练集中, 临床模型、影像组学模型和联合模型预测胃癌 HER-2 阳性表达的 AUC 分别为 0.711 (0.626~0.795)、0.852 (0.787~0.917) 和 0.872 (0.808~0.936); 在验证集中, 临床模型、影像组学模型和联合模型预测胃癌 HER-2 阳性表达的 AUC 分别为 0.698 (0.534~0.861)、0.818 (0.698~0.938) 和 0.853 (0.747~0.959)。校准曲线显示联合模型预测概率与真实概率之间的一致性良好, DCA 结果显示联合模型可为胃腺癌患者提供临床净获益。**结论:**基于 CT 影像组学和临床特征构建的联合模型可用于治疗前无创性评估胃腺癌的 HER-2 表达状态。

【关键词】 胃肿瘤; 人表皮生长因子受体-2; 表达状态; 体层摄影术, X 线计算机; 影像组学

【中图分类号】 R735.2; R814.42 **【文献标志码】** A **【文章编号】** 1000-0313(2025)01-0009-07

DOI: 10.13609/j.cnki.1000-0313.2025.01.003

开放科学(资源服务)标识码(OSID):



The value of CT-based radiomics for evaluating the expression status of human epidermal growth factor receptor 2 in gastric adenocarcinoma WANG Su-ya, ZHAN Peng-chao, XING Jing-jing, et al. Department of Radiology, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China

【Abstract】 Objective: The aim of this study is to develop and validate a non-invasive prediction model for evaluating the expression status of human epidermal growth factor receptor 2 (HER-2) in gastric adenocarcinoma. This prediction model is based on CT radiomic features and clinical characteristics. **Methods:** In this study, a retrospective analysis was conducted on 263 clinical data of patients diagnosed with gastric adenocarcinoma at our hospital, out of which 49 cases were HER-2 positive and 214 cases were HER-2 negative. Patients were randomly divided into a training set ($n=185$) and a validation set ($n=78$) at a ratio of 7:3. Radiomic features were extracted by manually delineating the region of interest (ROI) on venous phase CT images using 3D slicer software. In the training set, clinical features differences between HER-2 positive and negative groups were compared. Independent clinical predictors were determined using multivariate logistic regression and were utilized to develop a clinical model. A radiomics model was also created using a least absolute shrinkage and selection operator (LASSO) algorithm, which calculated the radiomics score (Radscore) for each patient. An integrated model that combined the independent clinical predictors and Radscore was also constructed as a final

作者单位: 450052 郑州, 郑州大学第一附属医院放射科

作者简介: 王素雅(1989-), 女, 河南许昌人, 硕士研究生, 主治医师, 主要从事腹部影像诊断工作。

通讯作者: 高剑波, Email: cjr.gaojianbo@vip.163.com

model. The predictive performance of the model was evaluated using the area under the receiver operating characteristic curve (AUC), while calibration curves were used to assess the agreement between predicted and observed probabilities. Additionally, decision curve analysis (DCA) was conducted to examine the clinical utility of the models. **Results:** Tumor thickness ($OR=1.04, P=0.033$), cT stage ($OR=2.39, P=0.038$), and cN stage ($OR=2.15, P=0.046$) were identified as the independent clinical predictors for HER-2 positive expression. In the training set, the area under the receiver operating characteristic curve (AUC) values of the clinical, radiomics, and combined models were 0.711 (0.626~0.795), 0.852 (0.787~0.917), and 0.872 (0.808~0.936), respectively. Similarly, in the validation set, the respective AUC values were 0.698 (0.534~0.861), 0.818 (0.698~0.938), and 0.853 (0.747~0.959) for the clinical, radiomics, and combined models. The calibration curves demonstrated good agreement between predicted probabilities and actual observations for the combined model. Furthermore, the decision curve analysis (DCA) indicated that the combined model would provide clinical net benefit for gastric adenocarcinoma patients. **Conclusion:** The combined model built on CT radiomics and clinical features can be utilized for non-invasive assessment of HER-2 status in gastric adenocarcinoma before treatment.

【Key words】 Stomach neoplasm; Human epidermal growth factor receptor-2; Expression status; Tomography, X-ray computed; Radiomics

胃癌起源于胃黏膜上皮,在所有恶性肿瘤中,其全球发病率和死亡率均位列前五^[1]。胃癌早期发病隐匿,多数患者就诊时已至晚期或进展期,导致预后较差。根治性切除是治愈胃癌的唯一治疗方式,但术后复发率高达 50% 以上,长期预后仍不理想^[2]。近年来,随着分子靶向治疗、免疫治疗等新兴治疗手段的出现,胃癌的治疗模式发生了根本性改变。多项大型临床试验如 TOGA 等确立了以曲妥珠单抗为代表的抗人表皮生长因子受体-2(human epidermal growth factor receptor-2, HER-2)靶向药物作为 HER-2 阳性胃癌的一线治疗策略^[3,4]。

目前,HER-2 表达状态的主要检测方法为基于活检或手术标本的免疫组织化学(immunohistochemistry, IHC)分析和荧光原位杂交(fluorescence in situ hybridization, FISH)法,检测成本偏高,创伤大且耗时^[5]。因此,迫切需要开发一种无创、简便易行的新方法评估 HER-2 状态。CT 作为治疗前全面评估胃癌的主要影像学检查方法,近年来以 CT 影像组学为代表的人工智能技术在 HER-2 状态检测中的价值逐渐成为研究焦点,目前国内有关基于 CT 影像组学评估胃癌 HER-2 状态的报道较为少见。因此,本研究基于 CT 影像组学和临床特征开发一种可无创性评估胃癌 HER-2 状态的预测模型。

材料与方 法

1. 研究对象

回顾性搜集 2013 年 6 月至 2023 年 12 月在本医院确诊为胃癌患者的临床和 CT 资料。病例纳入标准:

①原发病灶经病理证实为胃腺癌;②经 IHC 或荧光原位杂交 FISH 法准确评估 HER-2 表达状态;③治疗前 2 周内接受了腹部增强 CT 扫描。病例排除标准:①增强 CT 扫描前接受了化疗或其它抗肿瘤治疗;②图像质量差,存在伪影;③胃充盈差等导致病灶显示不佳,无法准确勾画胃癌病灶。根据纳入和排除标准,本研究最终纳入 263 例胃癌患者,其中男 197 例,女 66 例,年龄 23~83 岁,平均(59.76±11.09)岁,HER-2 阳性 49 例,HER-2 阴性 214 例。按照 7:3 的比例将所有患者随机分为训练集和验证集,其中训练集 185 例,HER-2 阳性与阴性比为 35:150;验证集 78 例,HER-2 阳性与阴性比为 14:64。本研究经郑州大学第一附属医院伦理委员会批准(2021-KY-1070-002),并豁免签署知情同意书。

2. 病例资料及 HER-2 状态的判定标准

纳入的临床特征包括性别、年龄、肿瘤位置、Borrmann 分型、糖类抗原 72-4(carbohydrate antigen 72-4, CA724)、糖类抗原 19-9(carbohydrate antigen 19-9, CA19-9)和癌胚抗原(carcino embryonic antigen, CEA)。CT 特征包括平扫期 CT 值、动脉期 CT 值、门脉期 CT 值、肿瘤长径(多平面重建图像中沿胃壁最大径所见肿瘤长度)、肿瘤厚径(肿瘤黏膜面至浸润最深处的最大垂直距离)、肿瘤临床 TNM 分期(cTNM)和临床分期。CT 特征由 2 位放射科医师在不知病理结果的情况下进行测量、评估,肿瘤长径和肿瘤厚径取两者测量结果的平均值,cTNM 分期和临床分期根据第八版 UICC/AJCC 标准进行评估,两者意见不一致时经讨论达成一致。

通过 IHC 或 FISH 法检测 HER-2 表达状态, HER-2 状态判定标准: IHC 3+ 判读为 HER2 阳性, IHC 2+ 判读为不确定性, IHC 1+ 和 IHC 0 判读为阴性。对于 IHC 2+ 的患者行 FISH 检测进一步明确 HER-2 基因状态, HER-2 基因扩增阳性定义为肿瘤细胞的 HER-2/CEP17 比值 ≥ 2.0 , 其余为阴性。

3. 检查方法

所有患者均行腹部平扫及双期增强扫描, 嘱患者禁食 8~12 h, 扫描前饮用 800~1200 mL 温开水使胃充盈。扫描仪器及参数: GE Revolution CT, 管电压 120 kV, 自动管电流 240~450 mA, 球管旋转时间 0.6 s; Siemens Somatom Force CT, 管电压 120 kV, 自动管电流 100~250 mA, 球管旋转时间 0.5 s。增强扫描对比剂使用碘海醇(350 mg I/mL), 剂量 1.5 mL/kg, 流率 3.0~3.5 mL/s, 在注射对比剂后 30 s 和 70 s 后分别行动脉期和门脉期扫描。层厚、层间距均为 5 mm, 采用标准算法重建薄层图像。

4. 特征筛选及模型建立

鉴于门脉期有助于更清晰地区分胃癌组织与正常胃壁, 本研究选择了门脉期图像进行影像组学分析。

将门脉期 CT 图像上传至 3D Slicer 5.0.2 软件(<https://www.Slicer.org/>)并进行信息脱敏处理。首先由 1 位具有 5 年腹部诊断经验的放射科医师在肿瘤最大横截面及其相邻两层沿肿瘤边缘手动勾画 ROI, 勾画完成后基于 ROI 提取影像组学特征。从每个 ROI 中提取 851 个影像组学特征, 包括 18 个一阶特征、14 个形状特征、75 个纹理特征和 744 个小波特征。1 周后随机抽取 30 例患者的 CT 图像, 由这位医师和另外 1 位放射科医师(具有 10 年以上腹部诊断经验)再次进行 ROI 勾画和特征提取。对 2 位医师及同 1 位医师前、后两次提取的特征进行可重复性检验, 保留组内及组间相关系数(intra- and inter-class correlation coefficient, ICCs) ≥ 0.75 的稳定特征, 然后将筛选后的影像组学特征进行 Z-score 标准化处理, 以便于进一步分析。

在训练集中, 使用最大相关最小冗余(max-relevance and min-redundancy, mRMR)算法和最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)算法对影像组学特征进行降维, 经过 10 折交叉验证, 筛选出具有低共线性的最优

表 1 HER-2 阳性组与阴性组间临床、CT 特征的比较 (例)

特征	训练集(n=185)				验证集(n=78)			
	HER-2 阳性	HER-2 阴性	t/ χ^2 值	P 值	HER-2 阳性	HER-2 阴性	t/ χ^2 值	P 值
年龄(岁)	61.43±11.22	59.73±11.39	0.803 ^a	0.426	62.14±10.64	58.41±10.42	1.195 ^a	0.247
性别			0.001 ^b	0.972				
男/女	26/9	111/39			13/1	47/17	/	0.007 ^c
肿瘤位置			1.129 ^b	0.569			/	0.448 ^c
胃上部	18	63			10	33		
胃中部	10	55			2	14		
胃下部	7	32			2	17		
Borrmann 分型			1.402 ^b	0.236			0.021 ^b	0.884
I/II	18	96			9	37		
III/IV	17	54			5	27		
CA72-4			0.176 ^b	0.675			/	0.443 ^c
正常/升高	28/7	127/23			11/3	55/9		
CA19-9			1.838 ^b	0.175			1.072 ^b	0.301
正常/升高	28/7	135/15			9/5	52/12		
CEA			4.309 ^b	0.038			/	0.999 ^c
正常/升高	27/8	135/15			12/2	56/8		
平扫期 CT 值(HU)	36.07±7.87	34.93±7.95	0.775 ^a	0.442	32.98±7.88	35.83±7.32	1.241 ^a	0.230
动脉期 CT 值(HU)	70.01±15.26	67.44±17.5	0.870 ^a	0.388	65.26±20.04	65.74±15.64	0.083 ^a	0.935
门脉期 CT 值(HU)	80.69±20.65	78.92±20.76	0.456 ^a	0.650	68.52±21.57	80.75±18.56	1.968 ^a	0.065
肿瘤长径(mm)	51.14±16.93	51.92±14.62	-0.250 ^a	0.804	43.81±12.27	53.26±15.97	2.459 ^a	0.022
肿瘤厚径(mm)	32.22±9.42	27.36±10.07	2.711 ^a	0.009	27.63±10.50	21.69±7.48	-2.008 ^a	0.062
cT 分期			5.818 ^b	0.016			/	0.205 ^c
T1/T2	6	61			2	22		
T3/T4	29	89			12	42		
cN 分期			5.567 ^b	0.018			/	0.239 ^c
N0/N+	11/24	83/67			4/10	31/33		
cM 分期			3.747 ^b	0.053			/	0.660 ^c
M0/M+	27/8	136/14			12/2	57/7		
临床分期			2.432 ^b	0.119			0.129 ^b	0.720
I/II	12	76			5	29		
III/IV	23	74			9	35		

注: ^a 为 t 值; ^b 为 χ^2 值; ^c 为 Fisher 确切概率法检验

特征子集以构建影像组学模型,并计算影像组学分数(Radiomics score, Radscore)。

在训练集中,将 HER-2 阳性与阴性组间比较差异有统计学意义的临床特征纳入多因素逻辑回归,筛选出 HER-2 阳性表达的独立预测因子,并构建临床模型;采用逻辑回归算法基于临床独立预测因子和 Radscore 构建联合模型。

5. 统计学分析

采用 R4.2.0 软件进行统计学分析。采用 Shapiro-Wilk 检验对计量资料进行正态性检验,计量资料以均数±标准差或中位数(上、下四分位数)表示。计量资料的组间比较采用独立样本 *t* 检验或 Mann-Whitney *U* 检验,计数资料的组间比较采用 χ^2 检验或 Fisher 确切概率法。采用多因素逻辑回归分析构建预测模型并绘制列线图,绘制校准曲线评价模型预测概率与真实概率之间的一致性,并采用 Hosmer-Lemeshow 检验评价模型的拟合优度。根据受试者工作特征(receiver operating characteristic, ROC)曲线的曲线下面积(area under curve, AUC)等指标评估模型的预测效能,采用 Delong 检验比较不同模型之间的 AUC 值差异。绘制决策曲线评价模型的临床净收益。以 $P < 0.05$ 为差异有统计学意义。

结果

1. 临床资料

训练集和验证集中 HER-2 阳性组与阴性组之间各临床特征的比较结果见表 1。在训练集中,HER-2 阳性组与阴性组之间 CEA、肿瘤厚径、cT 分期、cN 分期的差异均有统计学意义($P < 0.05$);在验证集中,两组间性别、肿瘤长径的差异具有统计学意义($P < 0.05$),余临床资料差异无统计学意义。

2. 特征筛选及模型建立

经组内和组间 ICC 一致性分析后保留 658 个特征;在训练集中通过 mRMR 算法、LASSO 回归算法(图 1)进行特征筛选,最终确定 11 个最优特征构建影像组学模型,相关特征及其系数见表 2。根据特征系数加权相加计算所有患者的 Radscore。在训练集中,HER-2 阳性组的 Radscore 为 1.00 ± 0.41 ,阴性组为 1.60 ± 0.40 ,HER-2 阳性组的 Radscore 低于阴性组,差异有统计学意义($t = 7.815, P < 0.001$)。在验证集中,HER-2 阳性组的 Radscore 为 0.88 ± 0.34 ,阴性组为 1.33 ± 0.42 ,HER-2 阳性组的 Radscore 低于阴性组,差异有统计学意义($t = 4.254, P < 0.001$)。

3. 临床及联合模型的建立

将训练集中组间差异有统计学意义的临床特征纳入多因素逻辑回归分析,结果显示肿瘤厚径(OR =

表 2 影像组学标签

特征	系数
截距	1.483763148
wavelet.LLH_gldm_Low Gray Level Emphasis	0.072527285
original_gldm_Cluster Tendency	0.06748114
wavelet.LHH_glszm_Size Zone Non-Uniformity Normalized	-0.075066504
wavelet.LHL_firstorder_Kurtosis	0.014392359
wavelet.HLL_firstorder_10Percentile	0.104049663
wavelet.HLL_firstorder_90Percentile	-0.031913575
original_shape_Maximum2D Diameter Row	-0.111754509
wavelet.HLH_firstorder_Skewness	0.081780354
wavelet.HHL_firstorder_Median	0.128527762
original_gldm_Imc2	0.053668665
wavelet.HLL_glrml_Run Entropy	-0.212124581

1.04, $P = 0.033$)、cT 分期(OR = 2.39, $P = 0.038$)、cN 分期(OR = 2.15, $P = 0.046$)为 HER-2 阳性表达的独立预测因子,以此构建临床模型。结合临床独立预测因子和 Radscore 构建联合模型。

4. 各模型的诊断效能分析

在训练集中,临床模型预测胃癌 HER-2 阳性表达的 AUC 为 0.711 (95% CI: 0.626~0.795),影像组学模型的 AUC 为 0.852 (95% CI: 0.787~0.917),联合模型的 AUC 为 0.872 (95% CI: 0.808~0.936);在验证集中,临床模型预测胃癌 HER-2 阳性表达的 AUC 为 0.698 (95% CI: 0.534~0.861),影像组学模型的 AUC 为 0.818 (95% CI: 0.698~0.938),联合模型的 AUC 为 0.853 (95% CI: 0.747~0.959,图 2、表 3)。

表 3 不同模型在训练集和验证集中的诊断效能对比

模型/数据集	AUC(95%CI)	诊断效能		
		准确度	敏感度	特异度
临床模型				
训练集	0.711(0.626~0.795)	58.9%	52.7%	85.7%
验证集	0.698(0.534~0.861)	78.2%	81.3%	64.3%
影像组学模型				
训练集	0.852(0.787~0.917)	76.2%	74.0%	85.7%
验证集	0.818(0.698~0.938)	80.8%	81.3%	78.6%
联合模型				
训练集	0.872(0.808~0.936)	73.0%	68.7%	91.4%
验证集	0.853(0.747~0.959)	80.8%	79.7%	85.7%

Delong 检验结果显示,在训练集中,联合模型与影像组学模型之间的 AUC 值差异无统计学意义($Z = 1.497, P = 0.134$),联合模型、影像组学模型与临床模型之间的 AUC 值差异均有统计学意义($Z = 4.105, P < 0.001; Z = 2.904, P = 0.004$);在验证集中,联合模型与临床模型之间的 AUC 值差异有统计学意义($Z = 2.064, P = 0.039$),余各模型之间的 AUC 值差异均无统计学意义(P 均 > 0.05)。

绘制联合模型的列线图(图 3)。校准曲线显示联合模型在训练集及验证集中的预测结果与实际结果基本贴合,具有较好的一致性;Hosmer-Lemeshow 拟合优度检验显示,预测结果与实际结果间差异无统计学

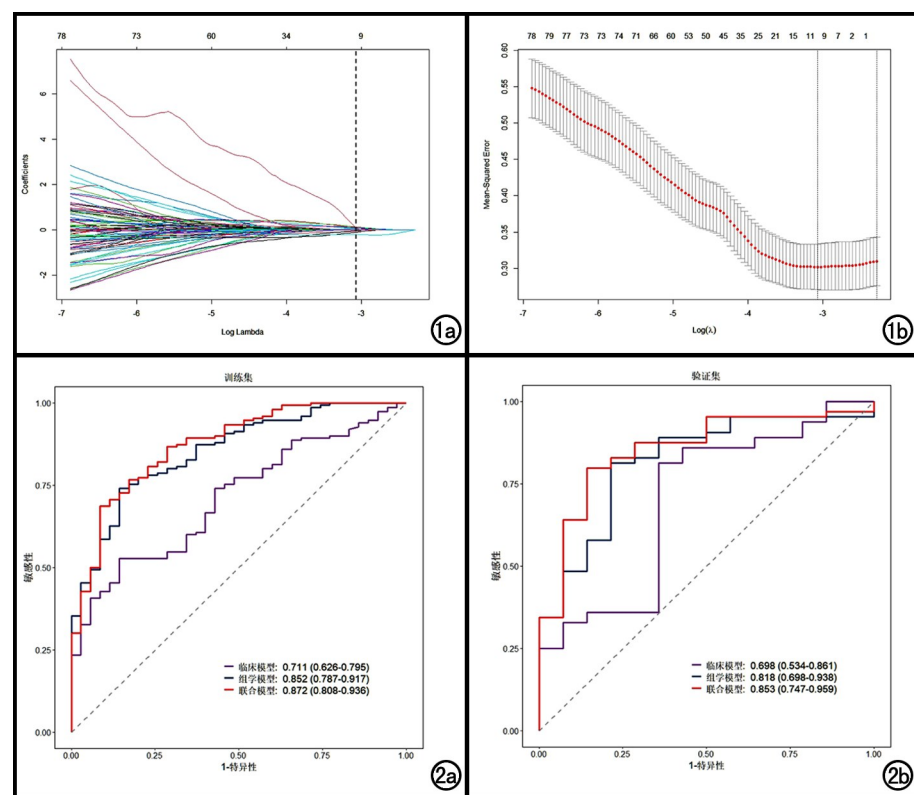


图1 LASSO回归。a)底部X轴表示log lambda,顶部X轴代表影像组学特征数量,Y轴为LASSO回归系数;b)底部X轴表示惩罚系数的对数 $\log(\lambda)$,顶部X轴代表非零系数特征的个数,Y轴为均方误差(Mean-Squared Error),Y轴值越小说明方程的拟合效果越好,两条虚线代表两个特殊的lambda值。左边虚线代表lambda min,右边虚线代表lambda. 1se,通过10折交叉验证确定LASSO回归中的调优参数lambda,基于lambda min可得到11个非零系数特征。图2 不同模型预测HER-2表达的ROC曲线。a)训练集;b)验证集。

意义,模型拟合良好(训练集: $P=0.938$;验证集: $P=0.910$,图4)。决策曲线显示联合模型在训练集、验证集中均可为胃癌患者提供临床净获益(图5)。

讨论

HER-2是重要的表皮生长因子受体通路成员,在细胞增殖、分化及存活中发挥关键调控作用,与肿瘤的发生与发展关系密切。在胃癌中,约13%~34%的患者存在HER-2基因扩增或蛋白过表达^[6]。多项研究显示,HER-2阳性胃癌使用曲妥珠单抗联合化疗,较单纯化疗可显著提高有效率和中位生存期,各大临床指南已将其作为HER-2阳性胃癌的标准一线治疗方法,准确评估HER-2状态对指导个体化靶向治疗决策意义重大^[7]。

CT检查在胃癌的临床分期、疗效评价及随访监测方面具有重要价值,是首选的影像学评估手段。目前关于CT形态学特征与HER-2状态的相关性研究

较少且存在分歧。潘圣宝等^[8]回顾性分析169例接受根治性切除术的进展期胃癌患者,结果显示浆膜面受侵、淋巴结转移等CT征象可初步评估HER-2表达状态。郑红伟等^[9]回顾性分析了433例胃癌患者的临床和影像资料,发现HER-2阳性组较阴性组出现肿瘤边界模糊、较低的N分期及门脉期明显强化的概率更高。陈龙等^[10]回顾性分析62例晚期胃癌患者的临床资料,发现与HER-2阴性胃癌相比,HER2阳性胃癌表现为强化程度高、更易发生肝脏转移。另有研究报道,HER-2状态与肿瘤厚度、发生部位、CT衰减值及cT分期相关^[5,11]。另有研究显示,碘浓度值等双能量CT参数在检测胃癌HER-2状态方面亦表现出较强的应用潜力^[5]。本研究结果显示,肿瘤厚径、cT分期及cN分期为HER-2阳性表达的独立预测因子,与相关研究结果一致。以此构建的临床模型在训练集和验证集中预测胃癌HER-2阳性表达的AUC分别为0.711(95%CI:0.626~0.795)、0.698(95%CI:0.534~0.861),效能尚可。总体而言,利用常规CT形态学特征或双能量CT新参数预测HER-2表达状态的可行性和效能还有待进一步研究。

常规CT无法深层次反映肿瘤的分子生物学信

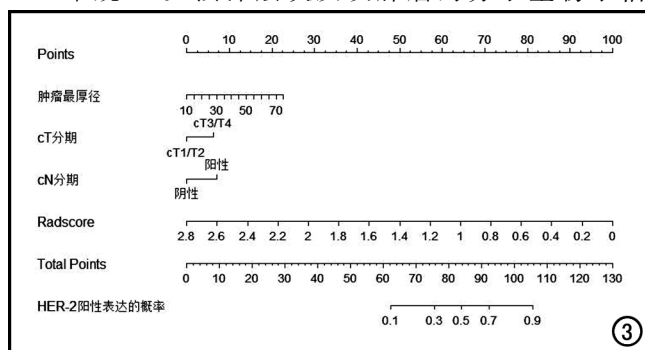


图3 联合模型列线图。每个变量对应应在模型中的得分,得分相加得到总分,再对应至下方的风险值。Points:得分;cT分期:临床T分期;cN分期:临床N分期;Radscore:影像组学分数;Total Points:总分;HER-2:人表皮生长因子受体-2。

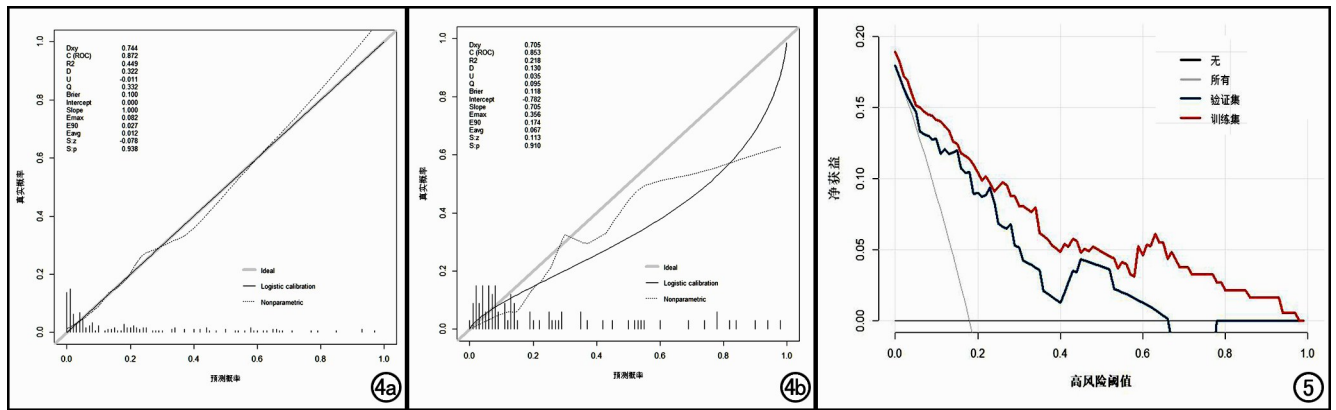


图 4 联合模型列线图预测胃癌 HER-2 表达的校准曲线。a) 训练集; b) 验证集。图 5 联合模型列线图预测胃癌 HER-2 表达的决策曲线。

息,影像组学可以更深入挖掘图像的生物学本质以支持临床决策,为评估胃癌 HER-2 状态提供了全新的思路和技术路径^[12,13]。Wang 等^[14]回顾性分析 101 例食管胃交界处腺癌患者的临床资料,基于 cT 分期和 CT 影像组学特征构建的列线图模型对评估 HER-2 表达状态具有较高的临床价值。另一项研究回顾性分析 132 例进展期胃癌根治术患者的临床资料,基于 CT 动脉期构建的影像组学模型可以较好地预测 HER-2 阴性胃癌,在训练队列中 AUC 为 0.756 (0.656~0.840),在验证队列中 AUC 为 0.830 (0.678~0.930)^[6]。此外,另有研究报道,针对 IHC 2+ 的 HER-2 弱阳性胃癌患者,基于门脉期 CT 影像组学特征构建的随机森林模型可进一步预测患者是否出现基因扩增,在验证集中 AUC 达到了 0.751 (0.636~0.862)^[15]。值得注意的是,目前有关 CT 影像组学预测 HER-2 表达状态的研究所采用的算法和模型不一,主要包括逻辑回归、支持向量机、随机森林等。不同的模型和算法可能会对预测结果产生影响。梁盼等^[16]对比了四种机器学习模型预测局部进展期胃癌脉管侵犯的效能,结果显示不同机器学习模型均具有较高的效能,逻辑回归模型的诊断效能最佳且较为稳定。本研究基于 LASSO 回归算法确定了 11 个重复性好、鲁棒性强的影像组学特征,根据其系数权重计算得到的 Radscore 与 HER-2 表达状态显著相关,在训练集和验证集中 HER-2 阳性患者的 Radscore 均低于阴性患者,差异有统计学意义。此外,采用逻辑回归算法结合临床独立预测因子及 Radscore 构建的列线图模型在训练集和验证集中的 AUC 值分别为 0.872 (0.808~0.936)、0.853 (0.747~0.959),初步实现了对胃癌 HER-2 表达状态的准确评估,效能优于既往所报道的预测模型。校准曲线和 DCA 曲线均显示列线图模型的预测结果与实际结果一致性较好,可为胃癌患者提供临床净获益。目前多数 HER-2 影像组学研究

仅采用了单一算法和模型,缺乏对比,未来应进一步探索不同算法、模型对预测效能的具体影响,并尝试结合多种算法优化模型,以期达到更高的准确性和临床应用价值。

本研究存在以下局限性:①本研究为单中心回顾性研究,可能存在病例选择偏倚,缺乏外部验证;②尽管较既往研究样本量有所提高,但未来仍需要扩大样本量进一步优化模型。

综上所述,本研究基于 CT 影像组学开发的列线图模型,可无创性预测胃癌 HER-2 表达状态,有望为靶向治疗决策提供有力支撑。

参考文献:

- [1] Xia C, Dong X, Li H, et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants [J]. Chin Med J (Engl), 2022, 135(5): 584-590.
- [2] 赵亚子, 翟建, 赵田英, 等. 术前腹部脂肪含量对胃癌根治术后远处转移的预测价值[J]. 放射学实践, 2023, 38(8): 1033-1038.
- [3] Bang YJ, Van Cutsem E, Feyereislova A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial [J]. Lancet, 2010, 376(9742): 687-697.
- [4] Qin S, Ji J, Xu RH, et al. Treatment patterns and outcomes in Chinese patients with gastric cancer by HER2 status: a noninterventional registry study (EVIDENCE) [J]. Oncologist, 2021, 26(9): e1567-e1580.
- [5] Zhao H, Li W, Huang W, et al. Dual-energy CT-based bomogram for decoding HER2 status in patients with gastric cancer [J]. AJR Am J Roentgenol, 2021, 216(6): 1539-1548.
- [6] Wang Y, Yu Y, Han W, et al. CT radiomics for distinction of human epidermal growth factor receptor 2 negative gastric cancer [J]. Acad Radiol, 2021, 28(3): e86-e92.
- [7] Lordick F, Carneiro F, Cascinu S, et al. Gastric cancer: ESMO clinical practice guideline for diagnosis, treatment and follow-up [J]. Ann Oncol, 2022, 33(10): 1005-1020.
- [8] 潘圣宝, 雷振. 进展期胃癌多层螺旋 CT 表现与 HER2 表达的相关

- 性[J].中国医学影像学杂志,2019,27(4):282-285.
- [9] 郑红伟,彭晓博,郑凌云,等.胃癌 CT 征象与 HER2 基因表达的相关性研究[J].中国 CT 和 MRI 杂志,2021,19(1):161-164.
- [10] 陈龙,叶薇,王康,等.晚期胃癌 HER2 表达与 MSCT 多参数影像表现相关性分析[J].影像技术,2023,35(6):27-33.
- [11] 梁盼,高剑波.胃癌靶向治疗标志物人类表皮生长因子受体-2 的影像学研究现状和进展[J].中华放射学杂志,2022,56(12):1392-1396.
- [12] 詹鹏超,刘珂衍,邱庆雅,等.基于 CT 影像组学预测胆管癌神经侵犯的价值[J].放射学实践,2023,38(7):910-915.
- [13] 张晰,单文莉,徐青青,等.基于增强 CT 的影像组学诺莫图预测进展期胃癌 HER-2 表达水平的研究[J].实用放射学杂志,2023,39(10):1635-1638,1651.
- [14] Wang S, Chen Y, Zhang H, et al. The value of predicting human Epidermal growth factor receptor 2 status in adenocarcinoma of the esophagogastric junction on CT-based radiomics nomogram [J]. Front Oncol, 2021, 11: 707686.
- [15] 黄文鹏,李林峰,刘晨晨,等.基于 CT 影像组学模型预测人表皮生长因子受体-2 弱阳性胃癌患者基因扩增的价值[J].中华消化病与影像杂志(电子版),2022,12(2):69-75.
- [16] 梁盼,雍刘亮,程铭,等.基于静脉期增强 CT 影像组学的机器学习模型术前预测局部进展期胃癌脉管侵犯[J].中华放射学杂志,2023,57(5):535-540.

(收稿日期:2024-02-28 修回日期:2024-06-18)

欢迎订阅 2025 年《放射学实践》

《放射学实践》是由国家教育部主管,华中科技大学同济医学院主办,与德国合办的全国性影像学学术期刊,创刊至今已 40 周年。遵照同行评议、价值导向、等效应用原则,国内各大学会、协会、组织机构通过科技工作者推荐、专家评议、结果公示等规定程序,《放射学实践》杂志入选中国科协发布 10 大领域《我国高质量科技期刊分级目录》业内认可的较高水平期刊。《放射学实践》杂志再次入选 2023 年版北京大学和北京高校图书馆期刊工作研究会共同主持的国家社会科学基金项目“学术期刊评价及文献计量学研究”研究成果——《中国核心期刊要目总览》。

本刊坚持服务广大医学影像医务人员的办刊方向,关注国内外影像医学的新进展、新动态,全面介绍 X 线、CT、磁共振、介入放射及放射治疗、超声诊断、核医学、影像技术学等医学影像方面的新知识、新成果,受到广大影像医师的普遍喜爱。

本刊为中国科学引文数据库(CSCD)核心库来源期刊、《中文核心期刊要目总览》收录期刊、中国科技核心期刊、中国高质量科技期刊分级目录临床医学领域医学影像学期刊,并在中国学术期刊分区中位列 Q1 区。

主要栏目:论著、继续教育园地、专家荐稿、研究生展版、图文讲座、本刊特稿、实验研究、传染病影像学、影像技术学、外刊摘要、学术动态、请您诊断、病例报道、知名产品介绍、信息窗等。

本刊为月刊,每册 25 元,全年定价 300 元。

国内统一刊号:ISSN 1000-0313/CN 42-1208/R 邮政代号:38-122

电话:(027)69378385

E-mail:fsxsjzz@163.com 网址:http://www.fsxsj.net

编辑部地址:430199 武汉市蔡甸区中法新城同济医院专家社区别墅 C 栋

基于 CT 影像组学预测不可切除性胃癌姑息性化疗疗效

程震, 尤亚茹, 詹鹏超, 郑月, 范儒阳, 吕培杰, 梁盼, 高剑波

【摘要】 目的:构建基于治疗前增强 CT 的临床影像组学模型,评估其用于个体化预测不可切除性胃癌姑息性化疗的临床反应。方法:纳入 256 例经病理证实的不可切除性胃癌患者,并在姑息性化疗前增强 CT 检查。根据实体肿瘤临床疗效评价标准确定治疗效果。对性别、年龄、肿瘤位置、实验室检查(CA19-9,CEA,CA724)、肿瘤最长径(横截面沿胃壁的最长轴)、肿瘤最厚径(垂直于横截面图像长轴的最大直径)、肿瘤临床 T 分期(cT)、临床 N 分期(cN)及临床 M 分期(cM)进行单、多因素逻辑回归分析,获得有统计学意义的临床独立预测因子以构建临床模型。使用 3D Slicer 软件在静脉期图像勾画感兴趣区(ROI),提取影像组学特征。选择最小绝对收缩及选择算子(LASSO)算法筛选并获得影像组学特征,构建影像组学模型。联合影像组学模型与临床模型构建影像组学列线图,使用诺莫图对模型进行可视化。采用 ROC 曲线的曲线下面积(AUC)等指标对构建模型的预测效果进行全方位评估。结果:95 例患者对姑息性化疗有反应,161 例无反应。肿瘤的临床分期是姑息性化疗疗效的临床独立预测因子($P < 0.05$),将临床分期作为独立预测因素以构建临床模型,该模型在训练集和验证集中预测姑息性化疗治疗反应的 AUC 分别为 0.591(95% CI:0.525~0.657)、0.674(95% CI:0.574~0.776)。在筛选得到 12 个最优影像组学特征的基础上构建影像组学模型,该模型在训练集和验证集中的 AUC 分别为 0.799(95% CI:0.733~0.865)、0.761(95% CI:0.656~0.865),均高于临床模型。影像组学列线图在训练集和验证集中的 AUC 分别为 0.814(95% CI:0.748~0.879)、0.785(95% CI:0.687~0.882)。校正曲线分析结果表明,模型表现出较好的拟合性,预测结果与实际情况相符。决策曲线分析结果显示,在大部分阈值概率范畴当中,不管在训练集还是验证集中均表现出较高的净收益,具备很好的预测效能。结论:基于治疗前 CT 的影像组学模型在预测不可切除性胃癌患者对姑息性化疗的临床反应方面表现良好,不但能够提升临床决策的科学性,同时还能够有效提高患者的生存率。

【关键词】 胃肿瘤;不可切除性胃癌;姑息性化疗;影像组学;体层摄影术,X 线计算机

【中图分类号】 R735.2;R814.42 **【文献标志码】** A **【文章编号】** 1000-0313(2025)01-0016-08

DOI:10.13609/j.cnki.1000-0313.2025.01.004

开放科学(资源服务)标识码(OSID):



A study on predicting the efficacy of palliative chemotherapy for inoperable gastric cancer based on CT imaging radiomics CHENG Zhen, YOU Ya-ru, ZHAN Peng-chao, et al. Department of Radiology, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China

【Abstract】 Objective: To build and assess a pre-treatment clinical radiomics model based on enhanced CT imaging for individualized prediction of clinical response to palliative chemotherapy in inoperable Gastric Cancer. **Methods:** This retrospective study included a total of 256 histologically confirmed cases of unresectable gastric cancer, and enhanced CT scans were performed before palliative chemotherapy. The treatment response was determined according to the RECIST criteria. Univariate and multivariate logistic regression analyses were conducted to identify independent predictive factors, including gender, age, tumor location, clinical stage, laboratory tests (CA19-9, CEA, CA724), CT image features [tumor longest diameter along the gastric wall, tumor thickest diameter perpendicular to the longest axis on cross-sectional images, clinical T stage (cT), clinical N stage (cN), and clinical M stage (cM)]. 3D Slicer software was used to delineate regions of interest (ROIs) on the venous phase images and extract radiomic features. The features were selected using the least absolute shrinkage and selection operator (LASSO) method, and radiomic labels were constructed to build a radiomics model. Fur-

作者单位:450052 郑州,郑州大学第一附属医院放射科

作者简介:程震(1998-),男,河南焦作人,硕士研究生,主要从事腹部影像诊断工作。

通讯作者:高剑波, E-mail: cjr.gaojianbo@vip.163.com

thermore, radiomic nomograms were constructed, and the model was visualized using nomograms. The predictive performance of the constructed model was comprehensively evaluated using metrics such as the area under the receiver operating characteristic curve (AUC). **Results:** Among the included patients, 95 patients showed a response to palliative chemotherapy. Clinical stage of the tumor was identified as a clinically independent predictive factor for the efficacy of palliative chemotherapy ($P < 0.05$). A clinical model was constructed using clinical stage as an independent predictor, and the AUC of this model in the training and validation sets were 0.591 (95% CI: 0.525~0.657) and 0.674 (95% CI: 0.574~0.776), respectively. On the basis of selecting 12 optimal radiomic features, a radiomics model was constructed, and its AUC in the training and validation sets were 0.799 (95% CI: 0.733~0.865) and 0.761 (95% CI: 0.656~0.865), respectively, which were higher than those of the clinical model. The AUC of the constructed model in the training and validation sets were 0.814 (95% CI: 0.748~0.879) and 0.785 (95% CI: 0.687~0.882), respectively. Calibration curves demonstrated a high consistency between the predicted and actual conditions of this model. Additionally, decision curve analysis (DCA) indicated a higher net benefit in both the training and validation sets, even within most threshold probability ranges. **Conclusion:** The CT-based radiomics model performs well in predicting the clinical response of patients with unresectable gastric cancer to palliative chemotherapy before treatment, which may contribute to clinical decision-making and improving patient survival rates.

【Key words】 Stomach neoplasms; Inoperable gastric cancer; Palliative chemotherapy; Radiomics; Tomography, X-ray computed

胃癌(Gastric cancer, GC)是世界上最常见的恶性肿瘤之一,其发病率和死亡率分别居世界第 5 位和第 3 位。2018 年,全球新增病例超过 100 万例,死亡人数约为 78.4 万人^[1];此外,大多数胃癌病例确诊时已为晚期^[2]。因此,选择有效的晚期不可切除性胃癌治疗方案对最大化提升整体治疗效益来说是至关重要的。姑息性化疗可以改善不可切除、局部晚期或转移性胃癌患者的生存和生活质量^[3]。此外,接受姑息性化疗的不可切除性胃癌患者的总生存期比单独接受最佳支持治疗的患者长 8 个月^[4-5]。然而,大多数治疗方案的肿瘤反应率低于 40%,并且化疗药物在一些患者中会产生严重的副作用^[6-7]。因此,在治疗前预测肿瘤对姑息性化疗的反应,可以为更精确的患者选择和个体化用药提供帮助,具有重要的临床意义。

作为一种无创非侵入性诊疗手段,CT 可以将细小病变显示得更直观清晰,不受时空异质性影响,目前 CT 已广泛应用于胃癌术前诊断及疗效评估^[8-9],不过在分子与基因水平上无法有效满足生物学相关信息。现阶段,临床对高价值的成像标记物的需求逐渐增加,其能够对肿瘤的生物学行为实现有效的无创预测,从而满足更高层次的医疗需求。在先进技术的支持下,医学成像技术越来越能发挥其作用。影像组学可以无创地分析肿瘤生物学,区分人眼无法分辨的细微差异,量化肿瘤异质性,监测肿瘤的发展和对治疗的反应^[10-12]。通过提取大量的定量特征,影像组学可以描述肿瘤异质性指标,这些指标可能反映与治疗反应相

关的病理生理特征^[13-14]。事实上,治疗前影像学与其他 CT 技术已被证实可以无创地预测 GC 的治疗反应^[15-16]。因此,本研究通过构建并验证一个基于 CT 的影像组学列线图,用于预测不可切除性胃癌患者对姑息性化疗的临床反应,为临床提供个体化治疗策略。

材料与方法

1. 研究对象

采用病例对照的研究方法,回顾性搜集郑州大学第一附属医院 256 例确诊为不可切除性胃癌患者的临床、CT 影像资料,病例纳入标准:①经病理活检确诊为胃癌;②术前或治疗之前全部进行了腹部增强 CT 检查;③在增强 CT 扫描前未接受放疗、化疗或其他抗肿瘤治疗;④CT 检查发现胃癌转移、不可切除而行姑息性化疗。病例排除标准:①临床资料不完整或 CT 图像缺失;②胃充盈不佳、肿瘤显示差;③合并其他恶性肿瘤或严重全身性疾病;④化疗不耐受。本研究为回顾性研究,已获得郑州大学第一附属医院伦理委员会批准(2021-KY-1070-002),并豁免知情同意书签署。

将所有患者随机分为训练集 167 例和验证集 89 例。纳入的临床特征包括年龄、性别、肿瘤位置、临床分期、CA19-9、CEA、CA724。CT 影像特征包括肿瘤最长径(横截面沿胃壁的最长轴)、肿瘤最厚径(垂直于横截面图像长轴的最大直径)、肿瘤临床 T 分期(cT)、临床 N 分期(cN)及临床 M 分期(cM),根据第八版的 UICC/AJCC 标准评估 cT/N/M 分期^[17]。

2. 姑息性化疗方案及治疗反应评价

在本研究中,入组患者主要接受卡培他滨+奥沙利铂(XELOX)方案或 S-1+奥沙利铂(SOX)方案治疗。根据 RECIST 1.1(姑息性化疗前 CT 为基线)标准进行治疗反应的疗效评估,分别为:①完全缓解(complete response, CR):靶病灶完全消失;②部分缓解(partial response, PR):靶病灶的最大径之和减少 $\geq 30\%$;③疾病稳定(stable disease, SD):靶病灶缩小的程度在 PD 和 PR 之间;④疾病进展(progressive disease, PD):靶病灶直径和相对增加 $\geq 20\%$ 且增加绝对值 ≥ 5 mm 或出现新病灶。其中 CR+PR 为治疗应答组(图 1),SD+PD 为无应答组。

3. 检查方法

所有患者术前均接受腹部三期增强扫描,患者在检查之前禁食 8~12 h,检查之前喝 800~1000 mL 温开水,以保证胃部充盈。GE Revolution CT 扫描参数:管电压 120 kVp,自动管电流 240~450 mA,球管旋转时间 0.6 s;Siemens Somatom Force CT 扫描参数:管电压 120 kVp,自动管电流 100~250 mA,球管旋转时间 0.5 s。首先获取平扫图像,之后以 3.0~3.5 mL/s 流率团注对比剂碘海醇(350 mg I/mL),剂量 1.5 mL/kg,经肘静脉通路注射,分别在注射后 30 s 和 70 s 行动脉期(AP)和静脉期(VP)扫描,层厚、层间距均为 5 mm,采用标准算法重建薄层 CT 图像。

4. 临床模型构建

利用电子病历系统,对患者的临床资料、相关检查数据进行整理,包括性别、年龄、临床分期以及 CA199、CEA、CA724 检测结果。由 2 位分别具有 20 年和 5 年腹部影像诊断经验的医生对检查获得的 CT 图像进行评估,评估保持一定的独立性。对 CT 特征进行分析,如肿瘤的部位、大小、TNM 分期等。若两者评估意见不一致,通过讨论达成统一意见。选择单、多因素逻辑回归分析,确定与姑息性化疗反应相关的临床独立预测因素,在此基础上进行临床模型的构建。

5. 病灶分割和影像组学特征提取

将所有患者的静脉期图像上传至图像开源处理软件 3D Slicer 5.0.2(<https://www.Slicer.org/>)。随机选择 30 例患者,由 2 位具有 5 年以上腹部疾病诊断经验的放射科医生完成 ROI 勾画,使用组内相关系数(intra-class correlation coefficient, ICC)评估 2 位观察者之间的一致性。剩余患者的 ROI 勾画由其中 1 位医生完成,于肿瘤最大横截面及其相邻两层沿边缘手动连续勾画完成(图 2)。由另 1 位医生检查,如果 ROI 被质疑,则在两者协调一致后重新分割。从每个 ROI 中提取了 851 个影像组学特征,包括 18 个一阶特征、14 个形状特征、75 个纹理特征和 744 个小波特征。

6. 影像组学模型及影像组学列线图的构建

采用一致性评估、单变量分析、最小绝对值收缩和选择算子(least absolute shrinkage and selection operator, LASSO)和多变量 Logistic 回归分析进行特征

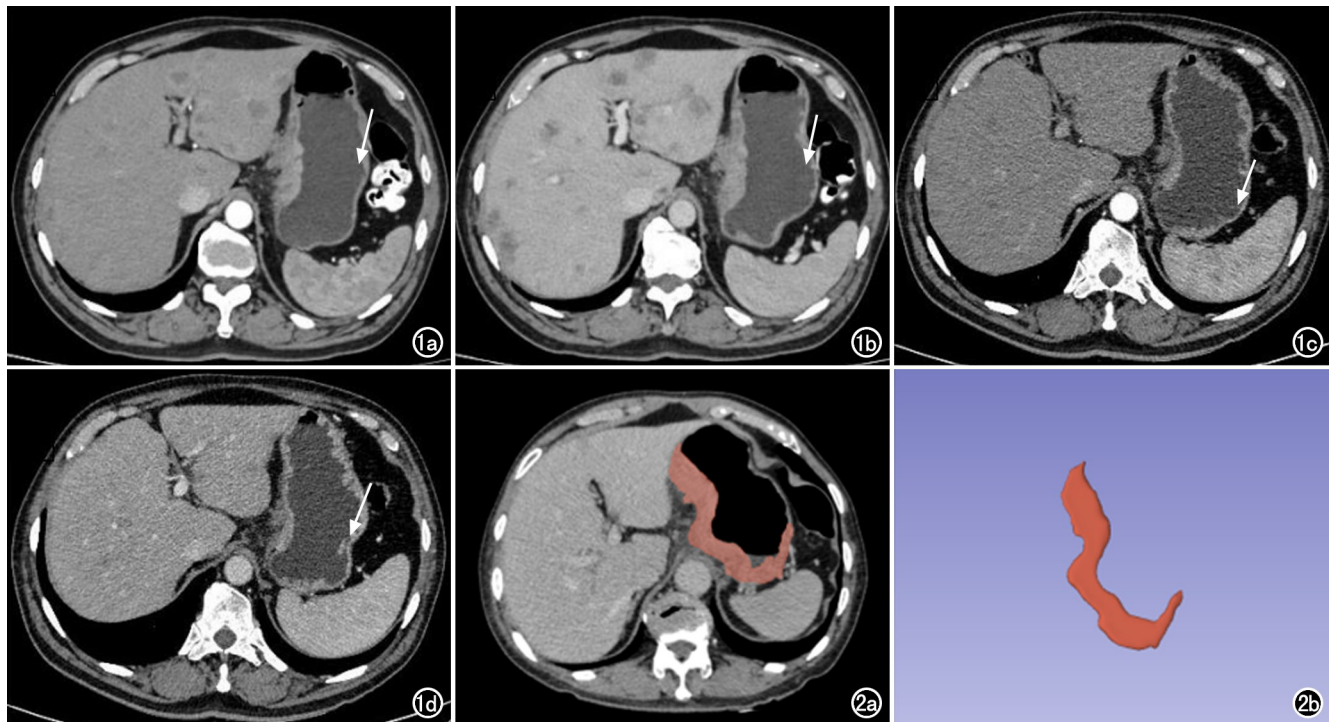


图 1 胃癌患者,男,66 岁,化疗有效的 CT 动脉期及门脉期图像。a~b)胃体可见局部胃壁增厚(箭);c~d)胃体处胃壁稍增厚,病变较前消退(箭)。图 2 胃癌 ROI 示意图。a)静脉期 CT 肿瘤最大层面,沿肿瘤边缘及相邻两层手动勾画 ROI;b)融合后的 3D ROI 示意图。

筛选。利用专业的十折交叉验证法实现数据的有效分析,进行 Lambda 值的优化,获得最为合适的 Lambda 值,尽可能地降低模型误差,这样能够提升模型的可靠性,最终留下系数非零的组学特征。选择多因素逻辑回归分析,基于影像组学特征实现影像组学模型的构建,同时计算影像组学分数(Radiomics score, Radscore),将影像组学模型与临床模型进行有效结合,最终实现影像组学列线图的构建。同时,将影像组学列线图可视化 nomogram 以预测个性化的反应概率。

7. 评估与验证模型性能

选择受试者工作特征(receiver operating characteristic, ROC)曲线分析评价各模型的预测效能,采用 DeLong 检验对不同模型的曲线下面积(area under curve, AUC)差异进行比较。根据预测响应概率与实际响应概率的一致性,绘制校准曲线来描述校准性能。基于不同阈值概率下的净收益,选择决策曲线分析(Decision curve analysis, DCA)实现模型在临床价值方面的评价。

使用 ROC 曲线进行各模型预测效果的有效评估,包括 AUC、敏感度(Sensitivity, SEN)和特异度(Specificity, SPE)等指标;通过 AUC 对临床模型及影像组学模型预测的准确性进行量化。利用校正曲线对模型的校正水平进行评估,通过 DCA 实现多种条件阈值概率范围内净效益的精准运算,从而对不同模型的临床作用有一个全方位的评估。相关流程图见图 3。

8. 统计学分析

采用 SPSS 25.0 软件和 R 语言(version, 4.1.3, <https://www.r-project.org>)进行统计学分析。采用单变量分析来确定组间差异,连续变量以均值±标准差表示,并运用 *t* 检验或 Mann-Whitney *U* 检验进行组间比较,分类变量采用卡方检验或 Fisher's 精确检

验进行组间比较。使用 R 语言进行 LASSO 回归分析并绘制 ROC 曲线、校正曲线、DCA 及 Nomogram。具有统计学意义的临床因素和影像组学特征作为多因素逻辑回归分析的输入变量,构建影像组学 Nomogram。利用 DeLong 检验比较各个模型之间的 AUC 差异。以 $P < 0.05$ 为差异有统计学意义。

结果

1. 临床、CT 影像特征的 Logistic 回归分析

本研究共纳入 256 例胃癌患者,其中男 178 例,女 78 例,年龄范围为 24~91 岁,平均年龄(60.25±12.30)岁,治疗应答组 95 例,治疗无应答组 161 例。在训练集中,治疗应答组(CR+PR)与无应答组(PD+SD)的临床分期及 Radscore 差异有统计学意义($P < 0.05$),其他临床特征差异无统计学意义($P > 0.05$)。在验证集中,两组患者的性别、肿瘤最厚径、临床分期及 Radscore 差异有统计学意义($P < 0.05$, 表 1)。

2. 临床模型构建

使用训练集数据构建临床模型。选择单因素逻辑回归分析方法对两组的临床、影像特征进行分析,找出预测姑息性治疗反应的潜在预测因子,包括年龄、性别、临床分期、肿瘤位置、实验室检查(CA19-9, CEA, CA724)、肿瘤最长径、肿瘤最厚径、肿瘤临床 T 分期(cT)、临床 N 分期(cN)及临床 M 分期(cM)。将单因素逻辑回归分析中具有统计学意义的因子纳入到多因素逻辑回归中,结果显示临床分期为预测姑息性化疗治疗反应的独立预测因子(表 2),建立以独立预测因子为基础的临床模型。

3. 影像组学模型构建

基于增强 CT 提取的影像组学特征使用 LASSO-Logistic 模型进行特征筛选,采用交叉验证确定 Lambda(λ)(图 4),最终筛选出 12 个最重要的特征构

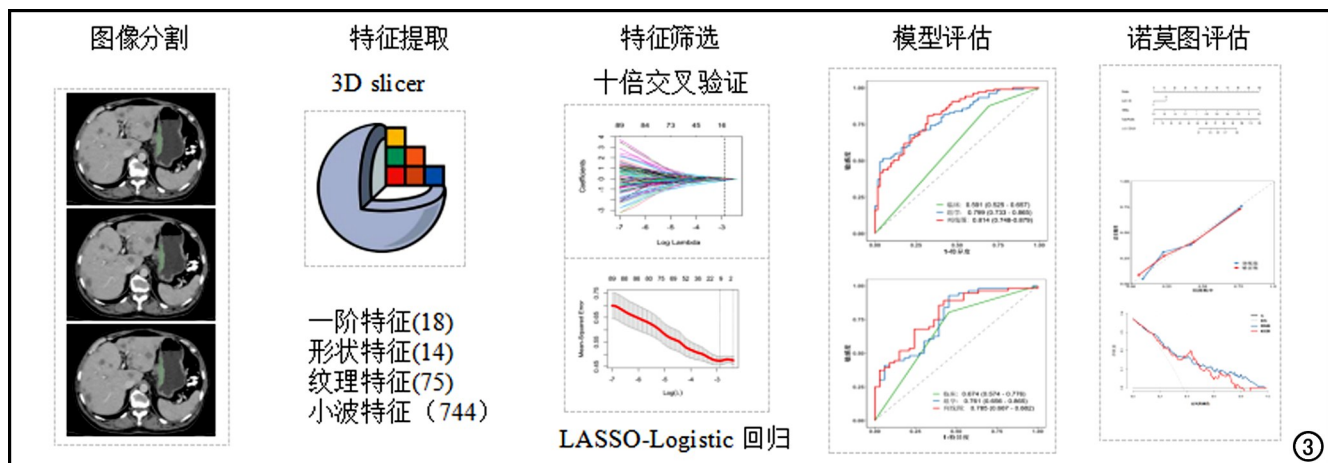


图 3 影像组学工作流程。影像组学过程主要包括以下 5 个步骤:感兴趣区图像分割、特征提取、特征筛选、模型评估和诺莫图评估。

表1 患者的基线资料 (例)

临床特征	训练集(n=167)				验证集(n=89)			
	CR+PR (n=62)	PD+SD (n=105)	t 值	P 值	CR+PR (n=33)	PD+SD (n=56)	t 值	P 值
年龄(岁)	59.87±9.81	60.71±13.54	0.464	0.643	62.70±10.84	58.34±13.21	1.603	0.113
性别			/	0.424			/	0.037
男性	40	74			28	36		
女性	22	31			5	20		
位置			/	0.620			/	0.585
胃上部	34	48			19	26		
胃中部	16	30			7	14		
胃下部	11	26			7	14		
超过 2/3 胃	1	1			0	2		
肿瘤最厚径(mm)	20.02±7.46	19.63±8.27	0.301	0.764	20.25±6.13	16.42±5.03	3.196	0.002
肿瘤最长径(mm)	65.87±27.48	62.59±26.06	0.770	0.442	68.95±24.54	59.51±28.79	1.574	0.119
CA19-9			/	0.810			/	0.895
阴性	42	73			24	40		
阳性	20	32			9	16		
CEA			/	0.463			/	0.851
阴性	33	62			17	30		
阳性	29	43			16	26		
CA724			/	0.540			/	0.248
阴性	39	61			19	39		
阳性	23	44			14	17		
临床分期			/	0.004			/	<0.001
III	19	13			18	11		
IV	43	92			15	45		
cT 分期			/	0.834			/	0.108
3	11	20			5	17		
4	51	85			28	39		
cN 分期			/	0.145			/	0.556
0	7	21			7	15		
1	55	84			26	41		
cM 分期			/	0.274			/	0.459
0	19	41			18	26		
1	43	64			15	30		
Radscore	0.35±0.23	0.65±0.28	7.059	<0.001	0.43±0.22	0.67±0.27	4.397	<0.001

表2 Logistic 回归分析结果

特征 (参考水平)	单因素逻辑回归			多因素逻辑回归		
	OR	95%CI	P 值	OR	95%CI	P 值
年龄	0.994	0.969~1.020	0.667	0.988	0.961~1.016	0.391
性别(女性)	1.313	0.673~2.561	0.425	1.248	0.610~2.551	0.544
位置						
胃上部	0.708	0.043~11.723	0.810			
胃中部	0.533	0.031~9.105	0.664			
胃下部	0.423	0.024~7.388	0.556			
超过 2/3 胃			0.623			
肿瘤最厚径	1.006	0.967~1.046	0.762	0.990	0.946~1.036	0.663
肿瘤最长径	1.005	0.993~1.016	0.441	1.002	0.989~1.016	0.765
CEA(阳性)	1.267	0.673~2.385	0.463	1.478	0.754~2.898	0.256
CA199(阳性)	1.086	0.553~2.134	0.810			
CA724(阳性)	0.818	0.429~1.558	0.540			
CT 分期				1.355	0.557~3.296	0.503
cT3						
cT4	1.091	0.484~2.461	0.834			
CN 分期						
cN0						
cN1	1.964	0.782~4.931	0.151			
CM 分期						
cM0						
cM1	1.450	0.744~2.826	0.275			
临床分期	0.320	0.145~0.707	0.005	0.279	0.120~0.647	0.003

表 3 基于增强 CT 筛选特征及系数

Active.Features	Active.coefficients
(Intercept)	0.541168353
wavelet.HHH_glszm_SmallAreaEmphasis	0.024865483
wavelet.LHH_glcm_Imc1	-0.005544035
wavelet.LLL_glcm_ClusterShade	0.040468482
original_glcm_Correlation	0.079168311
original_gldm_DependenceVariance	-0.004328878
wavelet.HHH_glszm_ZonePercentage	0.040808792
wavelet.LHL_glszm_LargeAreaLowGrayLevelEmphasis	-0.132350539
wavelet.HLL_gldm_DependenceVariance	-0.094033584
original_firstorder_InterquartileRange	0.073032525
original_firstorder_Skewness	0.080900507
wavelet.HLH_glszm_GrayLevelVariance	-0.170735646

建影像组学标签(表 3),在最终获得的影像组学特征的基础上构建影像组学模型。该模型在训练集、验证集中预测姑息性化疗治疗反应的 AUC 分别为 0.799 (95% CI: 0.733 ~ 0.865)、0.761 (95% CI: 0.656 ~

0.865),均高于临床模型(图 5)。

4. 构建与验证影像组学列线图

将临床分期和 Radscore 进行联合,选择 Logistic 回归分析,实现影像组学列线图的构建(图 6)。通过

列线图可得到临床分期和 Radscore 分别对应的评分,全部加在一起就能得到需要的 Nomogram,概率值即为治疗应答率。影像组学列线图在训练集、验证集中预测姑息性化疗治疗反应的 AUC 分别为 0.814 (95% CI: 0.748 ~ 0.879)、0.785 (95% CI: 0.687 ~ 0.882)。DeLong 检验结果显示,不管在训练集还是验证集中,影像组学列线图获得的诊断效能都略高于影像组学模型(训练集分别为 0.814 和 0.799,验证集分别为 0.785 和 0.761)。

通过绘制校正曲线(图 7)和 Hosmer-Lemeshow 拟合优度检验,结果显示在训练集和验证集中模型均表现出较好的拟合性(P 均 >0.05),说明模型预测结果与实际情况相符,验证了模型预测结果与临床的一致性。决策曲线分析结果显示,在大部分阈值概率范畴当中,不管在训练集还是验证集中均表现出较高的净收益,具备很好的预测效能(图 8)。使用影像组学 Nomogram 对患者进行临床干预,在训练集和验证集模型中均可使患者受益。

讨论

影像组学在医学影像分析方

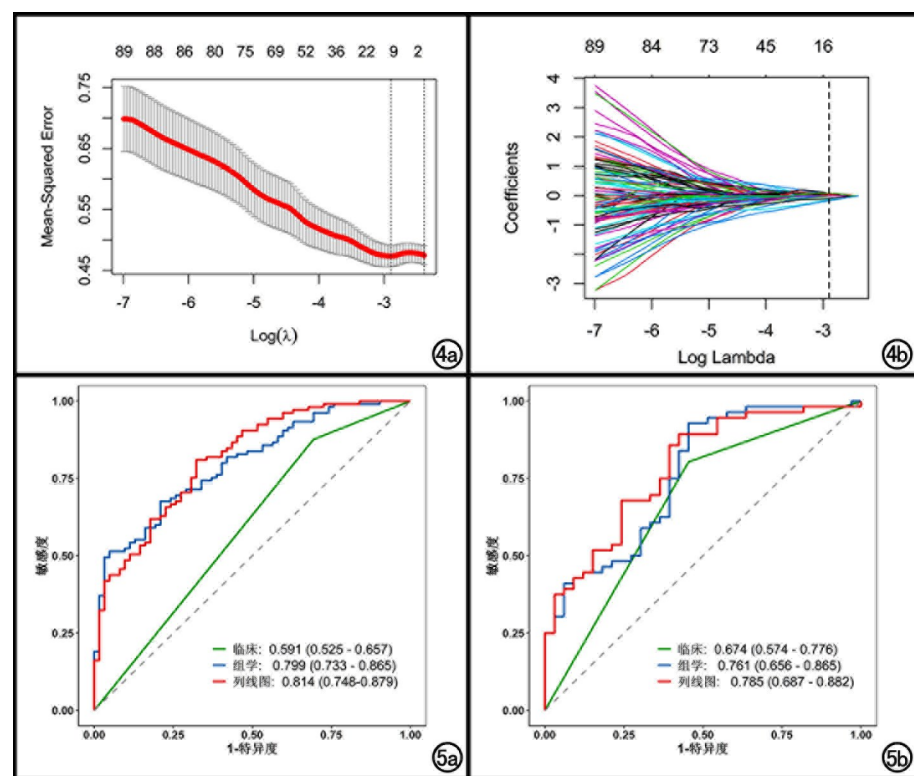


图 4 增强 CT 提取的影像组学特征使用 LASSO-Logistic 模型进行特征筛选。采用交叉验证确定 λ , 优化参数值对应模型的 AUC。a) 纵轴为 AUC, 横轴为 $\log(\lambda)$, 虚线表示最佳 λ 值对应的 $\log(\lambda)$ 值。以最小化标准及 1-SE 为选择标准; b) 纵轴为 89 个特征在模型中的系数, 横轴为 $\log(\lambda)$ 值, 垂直虚线表示采用十倍交叉验证选择后选出的非零特征。

图 5 训练集和验证集的 ROC 曲线。红色、绿色、蓝色分别代表影像组学列线图、临床模型、影像组学模型预测姑息性化疗治疗反应的 ROC 曲线。训练集中影像组学列线图、临床模型、影像组学模型的 AUC 分别为 0.814 (95% CI: 0.748 ~ 0.879)、0.591 (95% CI: 0.525 ~ 0.657)、0.799 (95% CI: 0.733 ~ 0.865); 验证集中影像组学列线图、临床模型、影像组学模型的 AUC 分别为 0.785 (95% CI: 0.687 ~ 0.882)、0.674 (95% CI: 0.574 ~ 0.776)、0.761 (95% CI: 0.656 ~ 0.865)。a) 训练集 ROC 曲线; b) 验证集 ROC 曲线。

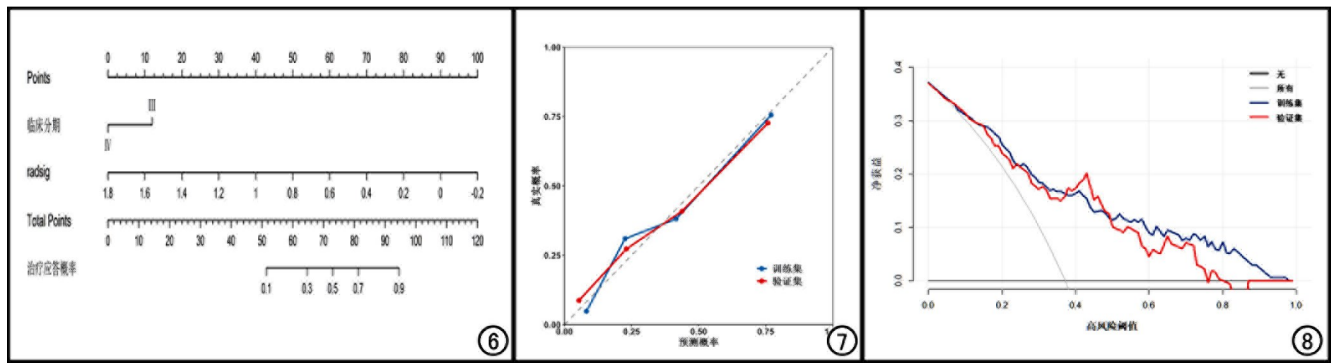


图 6 影像组学列线图预测不可切除性胃癌患者姑息性化疗治疗应答概率,依据每一个参数对应图最上方的量化数值,全部相加后的评分在图形最下方得到预测概率。图 7 校正曲线分析。x 轴和 y 轴分别表示影像组学模型的预测概率和实际概率。校准曲线表示根据预测化疗反应与观察到的化疗反应概率之间的一致性对影像组学进行校准。灰色对角线表示完美的预测,实线表示影像组学的预测性能。实线与虚线的拟合如果越处于接近的状态,那么说明预测的性能越突出。图 8 决策曲线分析。y 轴表示净收益,灰线表示所有患者对化疗都有反应的假设,蓝线表示训练集影像组学模型,红线表示验证集影像组学模型。

面是一个新兴的研究领域,它利用已有的自动化或半自动化的计算机成像技术软件,获取高通量量化的医学影像资料,并对其进行定量分析,建立相应的临床预测模型^[18]。已有研究表明,在预测肿瘤的组织学期、生物学行为、治疗和预后等方面,影像组学展现出了巨大的潜力^[19]。影像组学可以有效挖掘人体肉眼无法识别出来的影像细节,对肿瘤异质性的捕捉效能大大提升,这样就能够对存在预测价值的影像表型有一个更加精确的界定^[19]。既往研究发现,CT 在临床治疗上有着不错的应用,在胃癌疾病的治疗上,其能够对患者关于新辅助化疗的反应有一个全方位、精准的评估;增强扫描 CT 值被视为胃癌淋巴结转移的独立预测因素,通过对其影像特征的具体分析,可以实现患者预后的精准预估^[20-21]。相关学者通过大型局部晚期胃癌患者队列展开专业的 CT 影像分析,实现相应模型的构建,最终证实构建的模型能够很好地实现对胃癌患者新辅助化疗反应的有效预测。若将其应用在临床医学上,能够很好地区分潜在的治疗反应者与非反应者存在的差异^[22]。通过对双期增强 CT 图像的分析,能够对患者胃癌血管的分布情况进行全方位的观察,所以基于双期增强 CT 的模型设计能够有效强化影像组学分析,从而进一步提高其诊断能力^[20]。除此之外,还有学者在研究中提出,在 CT 图像的技术支持下,能够实现对患者肿瘤浸润免疫细胞的微观表型的深层次展示^[23]。在淋巴转移预测方面,应用淋巴结清扫能够实现对患者病情的有效控制,而 PET/CT 能够最大程度提升患者的手术成功率,在早期阶段对胃癌患者的预后评估中取得了不错的效果,并可预测术后复发^[24]。

本研究利用郑州大学第一附属医院 256 例患者的

病例资料,构建了基于增强 CT 的临床影像组学列线图,涉及临床因素和影像组学评分,用于预测不可切除性胃癌患者的姑息性化疗反应,显示出令人满意的预测能力,可用于对可能受益于姑息性化疗的患者进行分层。影像组学列线图在训练集(AUC 为 0.814)和验证集(AUC 为 0.785)中的预测能力优于临床模型(训练集 AUC 为 0.591,验证集 AUC 0.674)和影像组学模型(训练集 AUC 为 0.799,验证集 AUC 为 0.761),均表现出较高的预测能力。影像组学能够最大程度获得有价值的信息,具有识别通过视觉难以察觉的微小变化的能力,充分验证了其临床应用价值^[18]。影像组学已经广泛应用于实践及学术研究中,推动了不同类型疾病的术前诊断等方面的发展^[25]。

本研究基于治疗前静脉期 CT 图像筛选的 12 个重要影像组学特征代表着图像中不同的纹理和形状特征,较好地揭示了病灶内部肉眼观察不到的特征。影像组学特征基于最大层面连续 3 层的三维特征,可更好地反映病变异质性,相较于大多数研究采用的二维勾画,能更全面地描述肿瘤生物学特征;其在预测姑息性治疗反应方面起到了重要作用,为获得预测姑息性化疗治疗反应的满意模型,本研究有效构建三大模型,并且对这些模型进行全方位的比较,最终得出影像组学列线图的诊断效能更为优越的结论。

肿瘤厚度在预测胃癌的治疗效果方面起着重要作用。Wang 等^[26]发现肿瘤厚度减少是胃癌患者新辅助化疗(neoadjuvant chemotherapy, NAC)后病理完全缓解(pathologic complete response, pCR)的良好预测指标;然而,NAC 前的肿瘤厚度对预测 pCR 没有帮助。本研究发现治疗前肿瘤厚度与姑息性化疗反应无关。既往研究表明,临床治疗策略的选择、预后与临床

分期存在着一定关联性。同时,本研究也证实了临床分期与姑息性化疗反应显著相关,将临床分期添加到影像组学模型中可提高其预测能力。

本研究存在以下局限性:①本研究为回顾性研究,可能存在选择偏倚;②样本量偏少,缺乏来自独立机构的前瞻性验证队列来证明该模型的普遍性。未来仍需多中心大样本研究进一步验证模型的泛化能力,并转化为临床应用;③CT 扫描协议不统一,不过已通过多种图像预处理来降低其影响;④缺乏外部验证组,模型的效能缺乏说服力,后续研究会积极纳入外部验证组以提升模型的效能;⑤纳入的 CT 征象有限,未来笔者将进一步深入探讨,包括普通 CT 征象以及双能量 CT 多参数等在晚期胃癌姑息性化疗疗效预测中的应用价值。

综上所述,治疗前基于 CT 的影像组学模型在预测不可切除性胃癌患者对姑息性化疗的临床反应方面表现良好,不但能够提升临床决策的科学性,同时还能够有效提高患者的生存率。

参考文献:

[1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2018, 68(6):394-424.

[2] Necula L, Matei L, Dragu D, et al. Recent advances in gastric cancer early diagnosis[J]. World J Gastroenterol, 2019, 25(17):2029-2044.

[3] Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer[J]. Lancet, 2020, 396(10251):635-648.

[4] Fuchs CS, Shitara K, Di Bartolomeo M, et al. Ramucirumab with cisplatin and fluoropyrimidine as first-line therapy in patients with metastatic gastric or junctional adenocarcinoma (RAINFALL): a double-blind, randomised, placebo-controlled, phase 3 trial [J]. Lancet Oncol, 2019, 20(3):420-435.

[5] Shah MA, Bang YJ, Lordick F, et al. Effect of fluorouracil, leucovorin, and oxaliplatin with or without onartuzumab in HER2-negative, MET-positive gastroesophageal adenocarcinoma: the MET-Gastric randomized clinical trial [J]. JAMA Oncol, 2017, 3(5):620-627.

[6] Du M, Ouyang Y, Meng F, et al. Nanotargeted agents: an emerging therapeutic strategy for breast cancer [J]. Nanomedicine (Lond), 2019, 14(13):1771-1786.

[7] Van Cutsem E, Sagaert X, Topal B, et al. Gastric cancer [J]. Lancet, 2016, 388(10060):2654-2664.

[8] 黄钰迅, 李瑞, 张宝腾, 等. 基于临床-影像组学列线图模型术前预测进展期胃癌神经侵犯的价值 [J]. 放射学实践, 2022, 37(12):1548-1554.

[9] 黄裕存, 曹治, 陆少范, 等. 基于 CT 增强图像纹理特征预测可切除胃癌患者预后 [J]. 中国医学影像技术, 2020, 36(7):1046-1050.

[10] Ligerio M, Garcia-Ruiz A, Viaplana C, et al. A CT-based radiomics signature is associated with response to immune checkpoint inhibitors in advanced solid tumors [J]. Radiology, 2021, 299(1):

109-119.

[11] Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study [J]. Ann Oncol, 2020, 31(7):912-920.

[12] Dong D, Tang L, Li ZY, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer [J]. Ann Oncol, 2019, 30(3):431-438.

[13] Hu H, Gong L, Dong D, et al. Identifying early gastric cancer under magnifying narrow-band images with deep learning: a multicenter study [J]. Gastrointest Endosc, 2021, 93(6):1333-1341.e3.

[14] Meng L, Dong D, Chen X, et al. 2D and 3D CT radiomic features performance comparison in characterization of gastric cancer: a multi-center study [J]. IEEE J Biomed Health Inform, 2021, 25(3):755-763.

[15] Sun KY, Hu HT, Chen SL, et al. CT-based radiomics scores predict response to neoadjuvant chemotherapy and survival in patients with gastric cancer [J]. BMC Cancer, 2020, 20(1):468.

[16] Zhang L, Dong D, Zhang W, et al. A deep learning risk prediction model for overall survival in patients with gastric cancer: A multicenter study [J]. Radiother Oncol, 2020, 150:73-80.

[17] 唐磊. 从 UICC/AJCC 第 8 版 TNM 分期看胃癌影像学 T 分期的发展方向 [J]. 中华胃肠外科杂志, 2017, 20(7):735-739.

[18] Ma T, Cui J, Wang L, et al. A multiphase contrast-enhanced CT radiomics model for prediction of human epidermal growth factor receptor 2 status in advanced gastric cancer [J]. Front Genet, 2022, 13:968027.

[19] 刘娟娟, 刘炜, 薛华丹, 等. 基于增强 CT 的影像组学定量特征模型在预测胃癌分期中的价值 [J]. 放射学实践, 2022, 37(9):1061-1067.

[20] He J, Shi H, Zhou Z, et al. Correlation between apparent diffusion coefficients and HER2 status in gastric cancers: pilot study [J]. BMC Cancer, 2015, 15:749.

[21] Ji C, Zhang Q, Guan W, et al. Role of intravoxel incoherent motion MR imaging in preoperative assessing HER2 status of gastric cancers [J]. Oncotarget, 2017, 8(30):49293-49302.

[22] Chen R, Zhou X, Liu J, et al. Relationship Between ¹⁸F-FDG PET/CT Findings and HER2 Expression in gastric cancer [J]. J Nucl Med, 2016, 57(7):1040-1044.

[23] Celli R, Colunga M, Patel N, et al. Metabolic signature on ¹⁸F-FDG PET/CT, HER2 status, and survival in gastric adenocarcinomas [J]. J Nucl Med Technol, 2016, 44(4):234-238.

[24] Park S, Ha S, Kwon HW, et al. Prospective evaluation of changes in tumor size and tumor metabolism in patients with advanced gastric cancer undergoing chemotherapy: association and clinical implication [J]. J Nucl Med, 2017, 58(6):899-904.

[25] Zhou N, Liu C, Guo X, et al. Impact of ⁶⁸Ga-NOTA-MAL-MZHER2 PET imaging in advanced gastric cancer patients and therapeutic response monitoring [J]. Eur J Nucl Med Mol Imaging, 2021, 48(1):161-175.

[26] Wang ZL, Li YL, Li XT, et al. Role of CT in the prediction of pathological complete response in gastric cancer after neoadjuvant chemotherapy [J]. Abdom Radiol (NY), 2021, 46(7):3011-3018.

影像组学在透明细胞型肝癌诊断中的价值

刘斌, 陈枫, 张岩岩, 朱海峰, 李宏军

【摘要】 目的:通过比较不同模型的诊断效能评估影像组学诊断透明细胞型肝癌的价值。**方法:**42 例透明细胞型肝癌和 109 例普通型肝细胞癌患者以 2:1 的比率随机分为训练集和验证集,使用训练集建立模型,使用验证集进行验证。**结果:**在训练集中,临床模型诊断透明细胞型肝癌的 AUC 值、敏感度、特异度和准确度分别为 0.797、0.808、0.662 和 0.700,影像组学融合模型分别为 0.878、1.000、0.703 和 0.780,组合模型分别为 0.940、1.000、0.757 和 0.820;在验证集中,临床模型诊断透明细胞型肝癌的 AUC 值、敏感度、特异度和准确度分别为 0.647、0.625、0.600 和 0.608,影像组学融合模型分别为 0.812、0.688、0.857 和 0.804,组合模型分别为 0.868、0.688、0.886 和 0.824。**结论:**影像组学对于诊断透明细胞型肝癌具有一定临床价值,影像组学评分是诊断透明细胞型肝癌的独立影响因素。

【关键词】 透明细胞型肝癌;普通型肝细胞癌;影像组学;体层摄影术,X 线计算机

【中图分类号】 R735.7;R814.42 **【文献标志码】** A **【文章编号】** 1000-0313(2025)01-0024-05

DOI:10.13609/j.cnki.1000-0313.2025.01.005

开放科学(资源服务)标识码(OSID):



The value of radiomics in the diagnosis of primary clear cell carcinoma of the liver LIU Bin, CHEN Feng, ZHANG Yan-yan, et al. Department of Radiology, Civil Aviation General Hospital, Beijing 100123, China

【Abstract】 Objective: The aim of this study was to assess the value of radiomics in diagnosing primary clear cell carcinoma of the liver by comparing the diagnostic efficacy of different models. **Methods:** We randomly assigned 42 cases of primary clear cell carcinoma of the liver (PCCCL) and 109 cases of common hepatocellular carcinoma (CHCC) into training and validation sets at a ratio of 2:1. The training set was used to establish different models while the validation set was used to evaluate the performance of these models. **Results:** In the training set, the clinical model demonstrated an AUC value of 0.797, sensitivity of 0.808, specificity of 0.662, and accuracy of 0.700 in diagnosing PCCCL. The imaging fusion model showed improved performance with an AUC value of 0.878, sensitivity of 1.000, specificity of 0.703, and accuracy of 0.780. Additionally, the combined model demonstrated the highest performance with an AUC value of 0.940, sensitivity of 1.000, specificity of 0.757, and accuracy of 0.820. In the validation set, the clinical model had an AUC value of 0.647, sensitivity of 0.625, specificity of 0.600, and accuracy of 0.608. The imaging fusion model had an AUC value of 0.812, sensitivity of 0.688, specificity of 0.857, and accuracy of 0.804. The combined model continued to perform well with an AUC value of 0.868, sensitivity of 0.688, specificity of 0.886, and accuracy of 0.824. **Conclusion:** Our findings suggest that radiomics has significant clinical value in accurately diagnosing PCCCL. Importantly, our analysis revealed that Radscore, a radiomics-based score, is an independent factor in distinguishing between PCCCL and CHCC, demonstrating its potential for use in clinical practice.

【Key words】 Primary clear cell carcinoma of the liver; Common hepatocellular carcinoma; Radiomics; Tomography, X-ray computed

作者单位:100123 北京,民航总医院放射科(刘斌、朱海峰);100069 北京,首都医科大学附属北京佑安医院放射科(刘斌、陈枫、张岩岩、李宏军)

作者简介:刘斌(1987-),男,河北涿州人,博士,副主任医师,主要从事腹部疾病影像诊断工作。

通讯作者:李宏军, E-mail:lihongjun00113@ccmu.edu.cn

基金项目:国家自然科学基金项目(项目批准号:61936013)

肝细胞癌是肝脏最常见的原发性恶性肿瘤,包括普通型、透明细胞型、硬化型、脂肪性肝炎样型、粗梁实体型等多个亚型^[1]。透明细胞型肝癌(primary clear cell carcinoma of the liver, PCCCL)的预后好于普通型肝细胞癌(common hepatocellular carcinoma,

CHCC)^[2-4], 并且 PCCCL 对程序性细胞死亡蛋白 1 (programmed cell death protein 1, PD-1) 或细胞程序性细胞死亡配体 1 (programmed cell death-ligand 1, PD-L1) 等免疫治疗手段更敏感^[5], 因此两类肿瘤的精准诊断对于临床具有重要意义。然而, 通过传统影像检查方法鉴别 PCCCL 与 CHCC 较为困难^[6], 影像组学的兴起为我们提供了新思路。本研究通过比较临床模型、影像组学融合模型和组合模型的诊断效能, 旨在探讨影像组学在诊断 PCCCL 中的价值。

材料与方法

1. 研究对象

回顾性分析 2012 年 6 月至 2022 年 5 月 62 例 PCCCL 和 1017 例 CHCC 患者的临床和影像资料, 根据本研究排除标准排除 PCCCL 患者 20 例和 CHCC 患者 375 例。由于两组样本量严重失衡可能影响预测模型的诊断效能, 因此随机抽取 109 例 CHCC 患者参与研究, 42 例 PCCCL 患者全部参与研究, 所有患者以 2:1 的比例随机分为训练集和验证集, 另外在所有入组的病例中随机抽取 30 例作为测试集用于影像组学表型特征的筛选。

病例纳入标准: ① 外科手术后病理结果为 PCCCL 或 CHCC; ② 肿瘤病灶为单发; ③ 患者术前 2 周内接受过 CT 平扫和多期增强扫描。病例排除标准: ① 外科手术前接受过介入手术干预; ② 临床、影像或病理资料不符合研究要求。

2. 检查方法

CT 检查采用 GE64 Lightspeed VCT 机(美国通用电器公司), 所有患者接受检查时均取仰卧位, 扫描范围为膈肌顶部至肝脏下缘。扫描参数: 管电压 120 kV, 自动管电流, 探测器准直宽度 0.984 mm, 视野 36 cm×36 cm, 矩阵 512×512, 旋转时间 0.6 s, 层间距 5.0 mm, 层厚 5.0 mm, 重建层厚 0.625 mm。增强扫描对比剂为碘普罗胺(370 mg I/mL), 经肘前静脉注入, 注入对比剂后使用 30 mL 生理盐水冲管, 流率均为

3.0 mL/s。增强扫描动脉期采用 Smart Prep 对比剂跟踪技术, 门脉期延迟时间为 65 s, 平衡期延迟时间为 180 s。

3. 分析方法

由 2 位高年资影像科医生对肿瘤影像特征进行分析, 记录肿瘤位置、最大径、形态、边界、肝硬化、密度是否均匀(图 1a)、瘤内脂肪密度(图 1b)、强化方式、假包膜样强化、瘤周异常灌注情况、瘤内动脉征(图 1c)、门静脉瘤栓、淋巴转移等, 分析过程中二者意见出现分歧时通过协商达成一致意见。

由 1 位有 8 年诊断经验的影像科医生使用 ITK-SNAP 软件对包含肿瘤层面的图像进行分割, 分割后根据不同期相保存图像, 使用开源 python 平台在训练集中提取肿瘤的影像组学特征。在随机抽取 30 例患者的测试集中, 由 1 位具有 8 年诊断经验和 1 位具有 6 年诊断经验的影像科医生使用 ITK-SNAP 软件对包含肿瘤层面的图像进行分割, 使用 Matlab 软件(2014 版)根据分割肿瘤感兴趣区的结果在原始图像和小波图像上提取重复性较好的影像组学表型特征, 计算所有表型特征的组间相关系数(intra-class correlation coefficient, ICC)和组内相关系数(inter-class correlation coefficient, CCC), 其中 ICC 和 CCC 均高于 0.8 的影像组学表型特征被保留。

4. 统计学分析

采用 SPSS 21.0 软件进行统计学分析。根据数据是否符合正态分布, 对定量资料的组间比较采用独立样本 *t* 检验或 Mann-Whitney *U* 检验, 定性资料的组间比较采用卡方检验, 以 $P > 0.1$ 作为逐步回归指标剔除的标准。利用差异有统计学意义的临床特征和影像特征构建临床模型, 使用决策树算法筛选单序列影像组学标签得到最优影像组学融合模型并形成影像组学评分, 对差异有统计学意义的临床特征、影像特征和影像组学评分进行逐步回归构建组合模型。使用 DeLong-test 检验比较不同模型之间的差异。以 $P < 0.05$ 为差异有统计学意义。

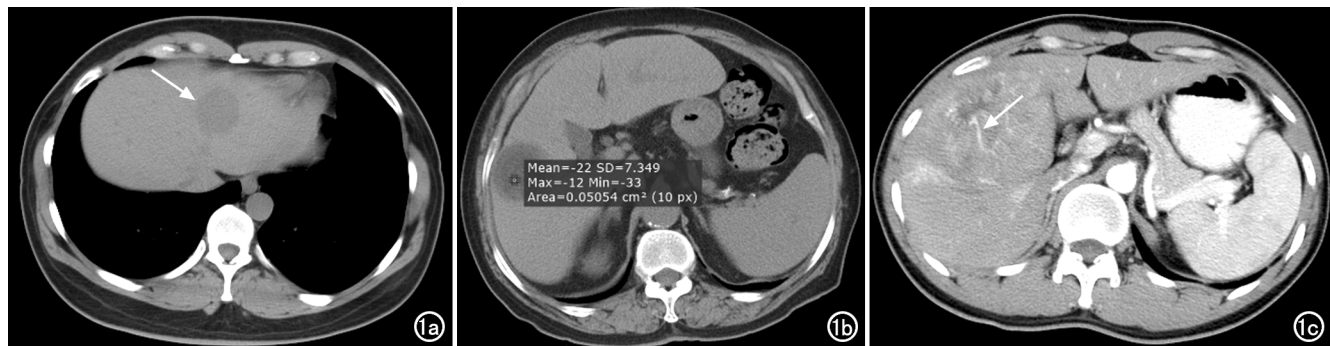


图 1 a) CT 平扫示肿瘤密度均匀(箭); b) CT 平扫示瘤内脂肪密度, 测量瘤内部分区域平均 CT 值约为 -22HU; c) CT 增强扫描动脉期瘤内可见动脉血管显影(瘤内动脉征, 箭)。

表 1 训练集两组患者的临床特征比较 (例)

临床特征	PCCCL 组	CHCC 组	$t/U/\chi^2$ 值	P 值
年龄(岁)	58.41±8.26	55.29±11.16	1.322	0.189
性别(男/女)	21/6	64/9	1.513	0.219
临床症状(有/无)	11/16	23/50	0.749	0.387
AFP(ng/mL)	33.36(5.70~24.60)	45.88(4.52~673.75)	936.000	0.701
丙肝(是/否)	7/20	4/69	8.417	0.004

表 2 训练集两组患者的影像特征比较 (例)

影像特征	PCCCL 组	CHCC 组	U/χ^2 值	P 值
位置(单叶/多叶)	26/1	68/5	0.346	0.557
最大径(cm)	4.50(2.50~5.60)	4.70(2.85~8.30)	814.500	0.184
形态(类圆形/不规则)	20/7	46/27	1.074	0.300
边界(清楚/不清)	25/2	66/7	0.115	0.735
密度均匀(是/否)	5/22	35/38	7.111	0.008
瘤内脂肪(有/无)	6/21	2/71	10.165	0.001
假包膜强化(有/无)	20/7	46/27	1.074	0.300
经典强化(是/否)	17/10	50/23	0.273	0.602
淋巴结转移(有/无)	0/27	1/72	0.374	0.541
门脉内瘤栓(有/无)	1/26	15/58	4.286	0.041
瘤内动脉征(有/无)	7/20	36/37	4.399	0.036
异常灌注(有/无)	5/22	8/65	0.996	0.318
肝硬化(是/否)	17/10	42/31	0.240	0.624

结果

1. 训练集患者临床特征分析

27 例 PCCCL 患者和 73 例 CHCC 患者的年龄、性别、临床症状、AFP 结果差异均无统计学意义,两组的丙肝结果差异有统计学意义($P < 0.05$,表 1)。

2. 训练集患者影像特征分析

训练集中 PCCCL 组与 CHCC 组患者的肿瘤位置、最大径、形态、边界、增强方式、淋巴结转移、瘤周异常灌注及肝硬化差异均无统计学意义(P 均 > 0.05),两组的肿瘤密度均匀性、瘤内脂肪密度、门静脉瘤栓、瘤内动脉征差异有统计学意义($P < 0.05$,表 2)。

3. 临床模型的构建

对差异有统计学意义的临床特征和影像特征进行

逐步回归,最终将丙肝、瘤内动脉征和密度均匀性纳入临床模型构建,使用验证集对建立的模型进行验证,临床模型在训练集和验证集的 AUC 值、敏感度、特异度及准确度见表 3。

4. 影像组学模型的构建

本研究比较了支持向量机、逻辑回归、决策树和随机森林的整体表现,结果显示决策树的表现优于其它 3 个分类器,因此使用决策树算法构建单序列影像组学模型。逐步回归结果显示 CT 平扫图像和动脉期图像 2 个单序列影像组

学标签组合为最优组合,利用以上组合构建影像组学融合模型并获取影像组学评分。对丙肝、瘤内动脉征、肿瘤密度均匀性和影像组学评分进行逐步回归,最终将丙肝、肿瘤密度均匀性和影像组学评分纳入组合模型。4 种单序列影像组学模型、影像组学融合模型及组合模型诊断 PCCCL 的 AUC 值、敏感度、特异度及准确度见表 3。

5. 不同模型的诊断效能比较

临床模型、影像组学融合模型和组合模型在训练集和验证集的 ROC 曲线(图 2)显示影像组学融合模型的诊断效能高于临床模型。3 种模型在训练集和验证集的小提琴图(图 3)显示影像组学融合模型和组合模型的概率密度分布峰值之间的距离均大于临床模型,表明影像组学融合模型和组合模型相对于临床模

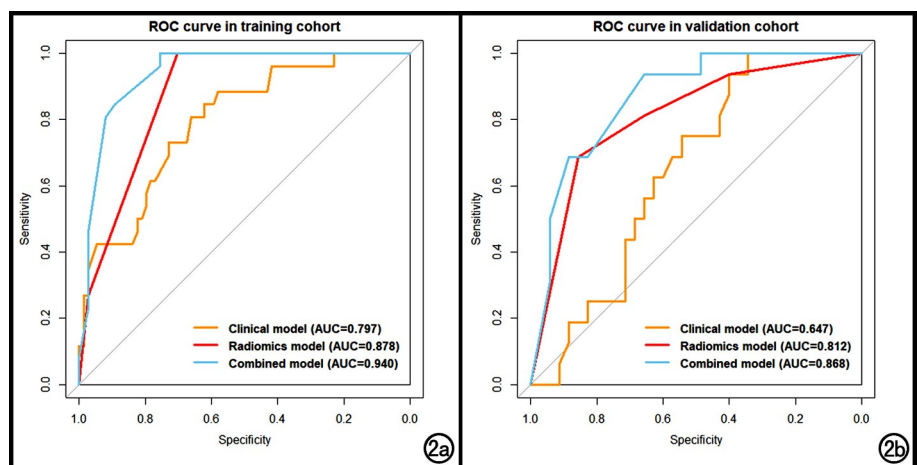


图 2 3 种模型在训练集和验证集中诊断透明细胞型肝癌的 ROC 曲线。a) 训练集; b) 验证集。

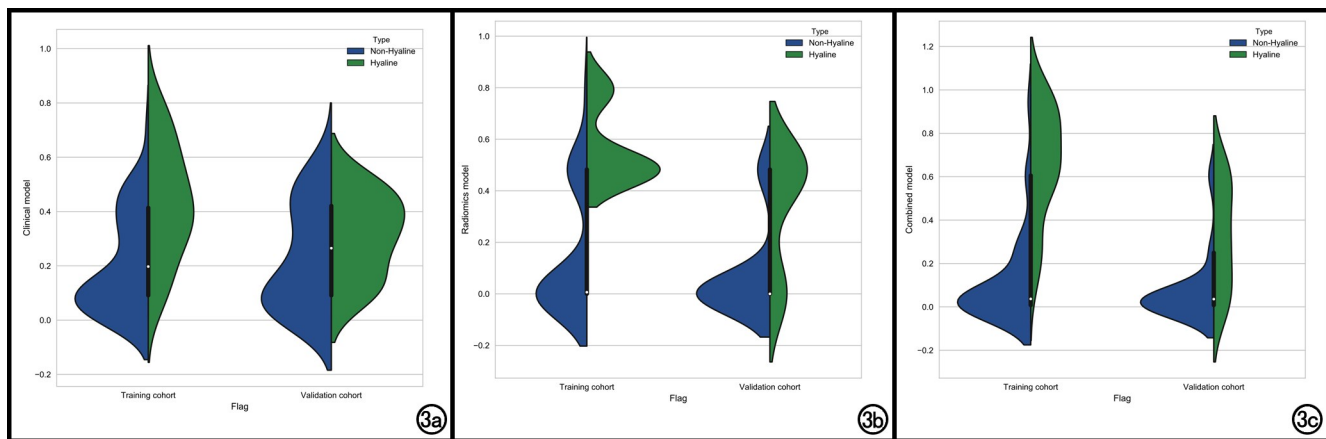


图 3 3 种模型的小提琴图。a) 临床模型; b) 影像组学模型; c) 组合模型。

表 3 不同模型的诊断效能

模型	训练集				验证集			
	AUC	敏感度	特异度	准确度	AUC	敏感度	特异度	准确度
平扫	0.811	0.720	1.000	0.622	0.735	0.706	0.813	0.657
动脉期	0.764	0.550	1.000	0.392	0.706	0.667	0.813	0.600
门脉期	0.773	0.680	0.923	0.595	0.697	0.647	0.750	0.600
平衡期	0.686	0.550	0.962	0.405	0.645	0.569	0.813	0.457
临床模型	0.797	0.808	0.662	0.700	0.647	0.625	0.600	0.608
影像组学融合模型	0.878	1.000	0.703	0.780	0.812	0.688	0.857	0.804
组合模型	0.940	1.000	0.757	0.820	0.868	0.688	0.886	0.824

型来说更容易区分 PCCCL 与 CHCC。预测模型校正曲线(图 4)显示组合模型预测结果与病理结果差异无统计学意义。DeLong-test 检验结果显示在训练集($P=0.124$)和验证集($P=0.090$)中,临床模型与影像组学融合模型之间差异无统计学意义;在训练集($P<0.001$)和验证集($P=0.004$)中,组合模型与临床模型之间差异有统计学意义,组合模型较临床模型增加了影像组学评分,因此影像组学评分是鉴别 PCCCL 与 CHCC 的独立影响因子。

讨论

CT 和 MRI 检查是诊断肝细胞癌常用的辅助检

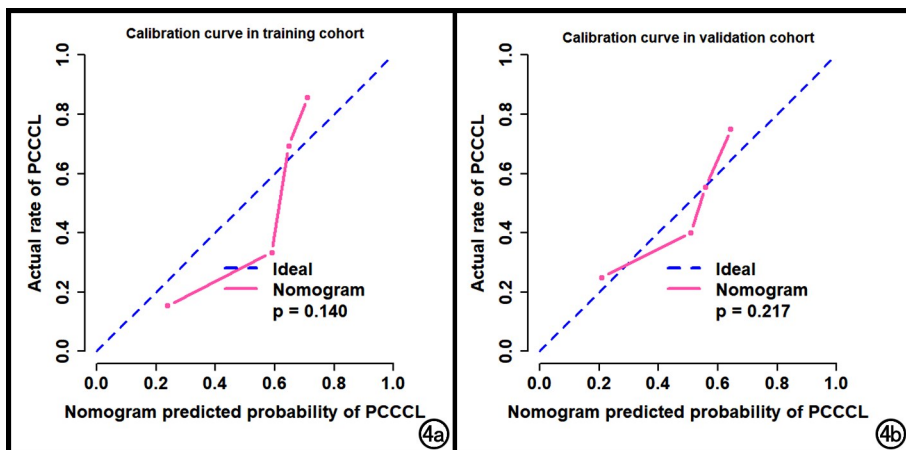


图 4 模型分别在训练集和验证集判定 PCCCL 情况的预测模型校正曲线。

a) 训练集; b) 验证集。

查手段^[7-9],然而单纯通过传统的肿瘤形态学特征区分 PCCCL 与 CHCC 较为困难^[10],影像组学可以深度挖掘影像图像中庞大的数据信息,发现肉眼看不到的病灶影像组学表型特征^[11-12]。本研究的影像组学表型特征主要包括形状大小特征、一阶统计特征、纹理特征和小波特征等 4 个类别。通过影像组学的深度挖掘,本研究最终提取原始图像的形状大小特征、原始及小波图像的一阶统计特征、纹理特征及小波特征共 851 个,筛选 ICC 和 CCC 较高的影像组学表型特征用于影像组学融合模型和组合模型的构建。DeLong-test 结果表明无论是训练集还是验证集中,临床模型与影像组学融合模型的诊断效能差异无统计学意义,而加入影

影像组学表型特征的组合模型在训练集和验证集中与临床模型比较差异均有统计学意义,因此影像组学对于精准诊断 PCCCL 和 CHCC 具有较高的临床价值。

PCCCL 患者的 1 年、3 年及 5 年总体生存率高于 CHCC 患者^[13],透明细胞比例高的肿瘤患者预后更好^[14]。相关研究报道 PCCCL 与 CHCC 患者的性别、年龄、术前 AFP 结果、患者临床症状差异均无统计学意义^[15],本研究结果显示 PCCCL 与 CHCC 患者的性别、年龄、临床症状、术前

AFP 结果差异均无统计学意义,与既往研究结果基本一致。相关文献报道 PCCCL 患者较 CHCC 患者丙肝感染率更高^[16],本研究与既往研究结果基本一致。

相关文献报道,PCCCL 与 CHCC 患者肿瘤的大小、强化方式、是否合并肿瘤破裂、门静脉瘤栓和淋巴转移等方面在 CT 或 MRI 图像上差异无统计学意义,PCCCL 患者假包膜出现概率较 CHCC 患者更高^[17];本研究结果显示,PCCCL 与 CHCC 患者假包膜样强化的结果差异无统计学意义,PCCCL 患者门静脉内瘤栓出现概率更低,这与既往研究结果不一致,以上差异可能与样本的数量和分布有关。PCCCL 瘤内出现脂肪变性较 CHCC 更常见^[18],本研究中 PCCCL 患者瘤内脂肪密度出现率较 CHCC 患者更高,差异有统计学意义,这与既往研究结果基本一致。本研究首次将瘤内密度均匀性和瘤内动脉征用于鉴别 PCCCL 与 CHCC,结果 2 个变量在 PCCCL 与 CHCC 比较的结果中差异均有统计学意义,这可能与 PCCCL 透明细胞含量更高、肿瘤血管生成及血供有关。

本研究首次将影像组学应用于 PCCCL 和 CHCC 的诊断,具有一定的创新性。然而,本研究 PCCCL 患者样本量偏小且为单中心研究,未来需要联合多中心扩大样本量以提高模型的诊断效能。

综上所述,影像组学对于诊断 PCCCL 具有一定临床价值,影像组学评分是诊断 PCCCL 的独立影响因素。

参考文献:

[1] 中华人民共和国国家卫生健康委员会医政医管局.原发性肝细胞癌诊疗规范(2022 年版)[J].中国实用外科杂志,2022,42(3):241-273.

[2] Xu W,Ge P,Liao W,et al.Edmondson grade predicts survival of patients with primary clear cell carcinoma of liver after curative resection:A retrospective study with long-term follow-up[J].Asia Pac J Clin Oncol,2017,13(5):e312-e320.

[3] Chen ZS,Zhu SL,Qi LN,et al.Long-term survival and prognosis for primary clear cell carcinoma of the liver after hepatectomy[J].Onco Targets Ther,2016,9(1):4129-4135.

[4] Bannasch P,Ribback S,Su Q,et al.Clear cell hepatocellular carcinoma:origin,metabolic traits and fate of glycogenotic clear and

ground glass cells[J].Hepatobiliary Pancreat Dis Int,2017,16(6):570-594.

[5] 尚玉奎.透明细胞型肝癌基因组突变特征研究[D].北京:北京交通大学,2018.

[6] Liu QY,Li HG,Gao M,et al.Primary clear cell carcinoma in the liver:CT and MRI findings[J].World J Gastroenterol,2011,17(7):946-952.

[7] 潘克华,张昭,贾秀芬,等.基于增强 CT 图像纹理特征术前预测肝细胞癌微血管侵犯状态及危险度等级[J].放射学实践,2023,38(9):1096-1100.

[8] 王海,连永伟,颜显杰,等.基于动态增强 MRI 纹理分析术前预测肝细胞癌 Ki-67 表达状态的价值[J].放射学实践,2022,37(6):729-733.

[9] 牟彬,熊婧彤,李雨娜,等.《乙型肝炎病毒相关早期肝细胞癌影像学检查与诊断标准共识》解读[J].新发传染病电子杂志,2023,8(3):90-96.

[10] 刘斌,孙君,陈枫,等.临床影像综合模型在透明细胞型肝癌精准诊断中的价值[J].中国医学影像学杂志,2021,29(9):807-901.

[11] Xia TY,Zhou ZH,Meng XP,et al.Predicting microvascular invasion in hepatocellular carcinoma using CT-based radiomics model[J].Radiology,2023,307(4):e222729.

[12] Feng Z,Li H,Liu Q,et al.CT radiomics to predict macrotrabecular-massive subtype and immune status in hepatocellular carcinoma[J].Radiology,2023,307(1):e221291.

[13] Li ZH,Bi XY,Zhao JJ,et al.Clinicopathological and prognostic analysis of primary clear cell carcinoma of the liver[J].Clin J Oncol,2013,35(2):140-143.

[14] Liu Z,Ma W,Li H,et al.Clinicopathological and prognostic features of primary clear cell carcinoma of the liver[J].Hepatol Res,2008,38(3):291-299.

[15] Emile JF,Lemoine A,Azoulay D,et al.Histological,genomic and clinical heterogeneity of clear cell hepatocellular carcinoma[J].Histopathology,2001,38(3):225-231.

[16] Jeon J,Park J,Lee E,et al.Huge primary clear cell hepatocellular carcinoma:a case report and review of radiologic findings[J].Curr Med Imaging,2023,20(1):1-5.

[17] Liu QY,Li HG,Gao M,et al.Primary clear cell carcinoma in the liver:CT and MRI findings[J].World J Gastroenterol,2011,17(7):946-952.

[18] 陈枫,赵晶,李宏军,等.原发性透明细胞型肝癌的影像表现及比较分析研究[J].临床放射学杂志,2019,38(8):1418-1423.

(收稿日期:2024-02-06 修回日期:2024-04-10)

影像组学和影像基因组学在胶质瘤中的研究进展

宁子睿, 周蓉艳, 白颖楠, 李卓琳

【摘要】 胶质瘤是成人中枢神经系统最常见的原发恶性肿瘤, 多呈侵袭性生长。随着分子遗传学的不断发展和分子生物学技术的不断进步, 越来越多的生物分子标志物被证实对脑胶质瘤的个体化诊断、治疗和预后评估具有重要的指导意义。近年来, 影像组学和影像基因组学领域有了显著的发展, 其通过先进的方法来挖掘医学影像背后的信息, 将这些图像信息转化为与肿瘤管理中的临床事件相关的可挖掘的数据, 其中大量的影像特征与肿瘤的分子遗传学特征相关, 使它能够在无创地预测胶质瘤的分级、基因分型、治疗反应和预后, 其最终目的是对患者进行分层, 为脑胶质瘤患者的个体化诊疗模式的临床应用提供科学依据, 现已成为精准医疗领域的研究热点。本文回顾了近年来有关成人脑胶质瘤影像组学和影像基因组学的研究, 对其在成人脑胶质瘤术前分级、基因分型、治疗和疗效评价、生存预测等方面的研究进展进行综述。

【关键词】 胶质瘤; 影像组学; 影像基因组学; 精准诊疗; 预后评估; 磁共振成像

【中图分类号】 R739.4; R445.2 **【文献标志码】** A **【文章编号】** 1000-0313(2025)01-0029-06

DOI: 10.13609/j.cnki.1000-0313.2025.01.006

开放科学(资源服务)标识码(OSID):



胶质瘤是成人最常见的恶性原发脑肿瘤, 占恶性脑肿瘤的 75%。不同分级、分子亚型胶质瘤的预后有很大差别, 如弥漫性星形细胞瘤(世界卫生组织分级 II 或 III 级)、异柠檬酸脱氢酶(isocitrate dehydrogenase, IDH)突变 II、III 级的中位生存期分别为 >10 年和 3~5 年; 胶质母细胞瘤(glioblastoma, GBM) WHO IV 级、IDH 野生型则预后不良, 中位生存期仅约为 1.3 年^[1]。这些分子表型术前常规影像学很难鉴别, 只能依靠有创的组织病检和分子检测进行鉴别^[2]。脑组织活检具有有创性和高成本的缺点, 而且在 7%~15% 的患者中, 病理诊断仍不能确定分子表型^[3]。如果能在术前获得准确的胶质瘤组织学分级和分子表型, 则有助于制定个体化的手术治疗方案, 最大限度地改善患者预后。

随着影像学的快速发展, 一些新的影像技术被应用到临床, 提高了胶质瘤术前分级诊断的准确性。然而, 准确预测肿瘤的分子分型仍是传统影像学的瓶颈。影像组学和影像基因组学工具实现了肿瘤微环境的无创检测, 即所谓的“虚拟活检”, 得以无创地对中枢神经系统肿瘤的区域异质性进行全面评估^[4], 通过对病灶的空间和分子异质性进行活体标记, 有可能将患者分层到更细致的诊断, 并对患者实现“实时”治疗监测^[5]。本文对目前影像组学和影像基因组学在脑肿瘤的临床

研究进展进行综述。

影像组学在胶质瘤鉴别诊断方面的应用

1. 影像组学在鉴别低级别胶质瘤与高级别胶质瘤中的价值

胶质瘤的分级与预后密切相关, 在治疗决策和预后评估中具有重要意义。相关研究表明, 纹理分析在鉴别低级别胶质瘤(low grade glioma, LGG)与高级别胶质瘤(high grade glioma, HGG)方面具有应用价值, 可以非侵入性地鉴别胶质瘤的异质性^[6,7], 以评估胶质瘤的预后、严重程度和治疗反应。Skogen 等^[8]运用过滤直方图技术来表征肿瘤的异质性, 纳入 95 例胶质瘤患者(II 级 27 例, III 级 34 例, IV 级 34 例), 在精细纹理比例内用标准差区分 LGG 与 HGG 的敏感度和特异度分别为 93% 和 81% (AUC = 0.91, $P < 0.0001$)。该研究对不同亚组(II~IV 级)胶质瘤的诊断能力略低, 但仍具有诊断意义。

不同级别和亚型的胶质瘤治疗策略不一致^[9,10], 因此需要确定胶质瘤的分级和亚型。Tian 等^[11]对 153 例患者(分别有 42、33 和 78 例 II、III 和 IV 级胶质瘤患者)的多参数 MRI 进行了纹理分析, 结果显示用支持向量机(support vector machine, SVM)分类器对 LGG 与 HGG 的分类准确率为 96.8%, 对 III 级与 IV 级的分类准确率为 98.1%。Xie 等^[12]评估了 42 例胶质瘤患者动态对比增强 MRI 的 5 个灰度共生矩阵特征后得出结论: 熵(AUC = 0.885)和逆差矩(AUC = 0.901)分别能够区分 III 级与 IV 级、II 级与 III 级胶质瘤。

作者单位: 650118 昆明, 昆明医科大学第三附属医院(云南省肿瘤医院)放射科

作者简介: 宁子睿(2000-), 男, 云南临沧人, 硕士研究生, 主要从事中枢肿瘤影像诊断工作。

通讯作者: 李卓琳, E-mail: lizhuolin0327@163.com

因此,通过影像量化胶质瘤病变的异质性来评估胶质瘤的恶性程度可能成为指导个体化、精确化治疗决策的有力工具,还可以优化利用稀缺的医疗资源。

2. 影像组学鉴别 LGG 的分子亚型

成人原发性脑肿瘤大多起源于胶质细胞,主要是 IDH 突变的星形细胞瘤, IDH 突变和 1p/19q 缺失的少突胶质细胞瘤,以及 IDH 野生型的 GBM^[13]。其中, IDH1 突变型 LGG 与 IDH1 野生型 LGG(即 IDH1 突变阴性)相比预后更好,生存时间更长,后者的行为更具侵袭性^[14]。1p/19q 基因缺失与化疗疗效好、预后好有关。非侵入性影像基因组学分子表型分析可能使 LGG 的个体化治疗成为可能。

多项研究已经评估了影像组学特征以确定胶质瘤的分子表型^[15]。Zhou 等^[16]使用自动纹理分析和伦勃朗视觉感受图像(visually accessible rembrandt images, VASARI)特征建立影像组学模型,以预测弥漫性低级别与中等级别胶质瘤患者($n=165$)的 IDH1 突变(AUC=0.86)、1p/19q 共缺失状态(AUC=0.96)、组织学分级(AUC=0.86)和肿瘤进展(AUC=0.80)。Zhang 等^[17]采用支持向量机递归特征消除法提取的 15 个最佳影像学特征($n=152$)能够检测 IDH 突变,准确率为 82.2%。Han 等^[18]从 42 例经组织病理证实的胶质瘤患者中提取 MRI 影像学特征,展示了来自 T_1 WI、 T_2 WI 的联合变量和 T_1 加权对比增强像(T_1 weighted imaging with contrast enhancement, T_1 WI+C)的成像直方图和灰度共生矩阵特征以用于检测 IDH1 突变型胶质瘤;在 T_2 WI 像上可发现 IDH1 突变,最高 AUC 值为 0.848, T_1 WI、 T_1 WI+C 联合变量预测 IDH1 突变的 AUC 值分别为 0.927、0.984;建模后, T_1 WI+C、 T_1 WI 和 T_2 WI 特征模型的准确率分别为 0.952、0.857 和 0.738。Jakola 等^[19]研究发现,纹理参数的同质性可以区分 IDH 突变型 LGG 患者与 IDH 野生型患者($P=0.005$)。联合参数(肿瘤同质性和肿瘤体积)预测 IDH 突变的 AUC 值为 0.940,但无法区分 IDH 突变型 LGG 是否有 1p/19q 编码缺失。Bahrami 等^[20]在液体衰减反转恢复(fluid attenuated inversion recovery, FLAIR)序列图像上测量了 61 例 IDH 患者的组织异质性和边缘对比度(edge contrast, EC),发现 IDH 野生型胶质瘤与 IDH 突变型胶质瘤相比,表现出更强的信号异质性($P=0.001$)和更低的 EC($P=0.008$);在 IDH 突变型胶质瘤患者中,1p/19q 共缺失的胶质瘤具有更强的信号异质性($P=0.002$)和较低的 EC($P=0.005$);与未甲基化组相比, O6-甲基鸟嘌呤-DNA-甲基转移酶(O6-Methylguanine-DNA Methyltransferase, MGMT)甲基化的肿瘤显示出较低的 EC($P=0.03$)。结合 FLAIR 序列的边缘清晰

度、异质性和像素相关性,可根据 IDH 状态对肿瘤进行最佳分类。作者得出结论, II/III 级胶质瘤的 FLAIR 异质性和边缘形状的定量成像特征可以在首次诊断成像时提供分子状态的独特信息,从而指导后续的手术和诊疗管理。

江少凡等^[21]利用 MRI 不同序列的影像组学特征构建机器学习模型以预测胶质瘤 IDH1 突变,搜集了经手术病理证实的 161 例胶质瘤患者(70 例 IDH1 突变型和 91 例野生型)的临床资料和 MRI 图像,临床资料包括 Karnofsky 功能状态(KPS)评分和肿瘤的病理分级, MRI 图像包括 T_2 WI、 T_2 -FLAIR、 T_1 WI+C 以及 ADC 图。根据 MRI 图像提取影像组学特征,利用机器学习算法构建预测模型并筛选出表现最佳的模型,然后结合临床参数构建联合模型,并评估这些联合模型的诊断效能。结果表明,基于 ADC 图和 T_1 WI+C 序列组合提取的组学特征构建的 4 种机器学习模型在测试集中的 AUC 分别为 0.888、0.872、0.896 和 0.887,均高于其它序列或序列组合构建的机器学习模型,此模型结合 KPS 评分和肿瘤病理分级所构建的联合模型在测试集中的 AUC 为 0.874(95%CI:0.722~0.985)。综合分析后认为,基于 ADC 图和 T_1 WI+C 序列组合提取的组学特征构建的机器学习模型对胶质瘤 IDH1 突变具有较好的预测效能,联合 KPS 评分和病理分级可提高预测效能。

大量研究证实非侵入性影像基因组学研究可以预测胶质瘤的不同分子亚型,并为临床的个体化治疗决策提供依据。

3. 影像组学鉴别 LGG 与 GBM 瘤周区域

大约 90% 的胶质瘤复发发生在瘤周区域^[22],了解该区域肿瘤细胞的异质性和细胞的微浸润有可能影响后续治疗。GBM 瘤周区域由肿瘤细胞和水肿组成,而浸润性的 LGG 瘤周区域仅有肿瘤, Malik 等^[23]使用 GBM 瘤周区域和 LGG 病例的 MRI 影像数据进行分割,使用 AdaBoost 分类器获得了所有鉴别性能最好的特征,得出敏感度、特异度、准确度和 AUC 值分别为 91%、86%、89% 和 0.96。在特征选择技术中,递归特征消除技术的结果最好, AUC 范围为 0.87~0.92;研究表明 MRI 常规序列的定量分析可以有效区分人眼无法区分的 GBM 瘤周区域与 LGG,确定 GBM 瘤周区域内的肿瘤浸润范围有助于个体化放疗计划的制定;目前该研究模型还可以扩展到使用类似 LGG 信号的特征信号独立识别 GBM 瘤周区域中的微小肿瘤。

评估预后

1. 预测 LGG 和 HGG 患者的生存预后

根据 LGG 边缘强化特征进行预后预测:既往研究表明,根据 VASARI 注释评分的 MRI 特征可以用来预测 GBM 患者的生存预后和分子表型^[24,25],然而还没有学者对 LGG 进行类似研究。Zhou 等^[16]研究发现,在 MRI 图像上,无强化和平滑的非强化边缘预示着较长的无进展生存期(progression free survival, PFS),而平滑的非强化边缘预示着更长的总生存期(all survival, OS)。

根据 LGG 免疫亚型进行预后预测:越来越多的研究表明,LGG 与 HGG 的免疫学有着明显差别,这突出了对 LGG 免疫学进行专门研究的重要性;并且 LGG 患者在接受免疫治疗前需要对 LGG 肿瘤免疫学和潜在的肿瘤免疫表型有详细的了解^[26]。陈燕生等^[27]对 133 例 LGG 患者的 MRI 影像特征和免疫相关基因数据进行影像基因组学分析,提取和筛选出与预后显著相关的 11 个影像组学特征并构建影像组学评分,影像组学风险评分中位值将患者分为高风险组($n=66$)和低风险组($n=67$),高分组患者的生存期预后明显较差。ROC 曲线分析结果显示影像组学预测 LGG 患者 1 年、5 年生存风险的 AUC 分别为 0.79 和 0.86。使用聚类分析将 LGG 患者分为免疫炎症型($n=13$)和免疫沙漠型($n=113$)两种亚型,数据分析结果显示影像组学评分可以有效预测 LGG 患者预后,影像组学评分与免疫分型存在一定平行关系,大部分 LGG 患者属于免疫沙漠型,少部分免疫炎症型患者的影像组学评分较高,异质性明显高于免疫沙漠型,免疫炎症型患者可能获益于免疫治疗。

根据 GBM 缺氧状态进行预后预测:超过 40% 的 HGG 患者对常规化疗和放射治疗无效,并在 6~9 个月内出现进展^[28]。多形性 GBM 缺氧是已知的促进肿瘤新生血管和侵袭健康组织的关键途径,也是导致治疗抵抗和不良预后的关键途径^[29]。细胞增殖、凋亡和血管生成增加等多种途径也是导致 PFS 较短的原因之一。目前,缺乏有效的生物标志物来监测低氧水平和预测抗血管生成药物的治疗反应^[30]。

Beig 等^[31]研究发现,影像组学可以捕捉到治疗前 MRI 上 GBM 的缺氧程度,并预测生存。该研究使用了来自 85 例 GBM 患者的微阵列表达数据来构建缺氧富集评分(hypoxia enrichment score, HES),并训练了一个与 HES 相关的影像组学模型,然后用它来根据 GBM 的 OS 进行分层。在 30 例患者的验证集上,与 HES 密切相关的影像组学特征也可以区分短期生存($OS < 7$ 个月)与长期生存($OS > 16$ 个月)($P = 0.003$)。这项研究中确定的影像组学特征可以用于监测缺氧,帮助确定治疗耐药的时间表,并评估抗血管生成治疗在 GBM 和其他肿瘤中的疗效。Mallick 等^[28]

的另一项研究($n=119$)提取了超过 12190 个 GBM 的影像特征,其中 11 个影像特征可以预测 PFS 和 OS。

通过免疫相关标志预测预后:GBM 相关的肿瘤微环境和免疫细胞浸润在肿瘤的发生、进展和预后中发挥着至关重要的作用;而 GBM 免疫治疗的疗效可能受到肿瘤或肿瘤微环境因素的影响,其中包括肿瘤细胞、基质细胞、免疫细胞等相关因素^[32]。

Liu 等^[33]从 3 个数据库中获取 GBM 患者的基因表型和生存数据,在测试集($n=400$)中为每例 GBM 患者制定个体化的免疫细胞浸润评分,并在独立验证集中($n=374$)进行独立验证。该研究从瘤内和瘤周区域共提取了 5915 个影像组学特征,使用递归特征消除和 SVM 方法来选择关键特征,并生成低或高免疫细胞浸润得分的预测模型。在一个独立的数据集($n=149$)中,使用影像和生存数据来检验已识别的影像基因组学模型的预后预测价值。研究结果显示免疫细胞浸润评分越高,患者预后越差(测试集和验证集的多变量风险比分别为 0.48 和 0.63),免疫检查点相关基因的表达水平也越高。结合 11 个影像学特征的模型可以很好地区分不同免疫细胞浸润评分的肿瘤(AUC 为 0.96,准确率为 94%),这些征象可以通过瘤内和瘤周区域的影像基因组学特征进行无创性评估。在精确医学时代,影像基因组学具有无创预测和评估 GBM 分子表型的潜力。

2. 影像基因组学预测 GBM 的 MGMT 甲基化状态

影像基因组学可以用来了解与特定影像学表现相关的任何可能的分子表型,或表明特定的基因变异如何影响肿瘤的影像表现^[34-36]。

许多研究表明,MGMT 启动子甲基化状态是预测多形性 GBM 患者替莫唑胺耐药性和 PFS 的重要生物标志物。因此,确定 MGMT 甲基化状态,对制定准确的治疗策略,并提高 GBM 治疗的成功率具有重要意义。尽管使用手术标本的分子检测是评估 MGMT 甲基化状态的标准,但该方法过于繁琐复杂,而其他技术如活性分析、免疫组织化学和甲基化芯片分析等,则有技术限制^[37]。此外,不完全活检采样的可能性、检测的高昂价格以及肿瘤内异质性的高度复杂性限制了这些侵入性技术在医院中的应用^[38]。

为无创性预测 GBM 的 MGMT 甲基化状态,Do 等^[39]使用了一种新的基于影像组学的机器学习用于术前无创预测 MGMT 甲基化状态。研究人员从标注了 MGMT 甲基化状态的 GBM 患者的 MRI 多模态图像中提取影像组学特征,并经过极端梯度提升算法(eXtreme Gradient Boosting, XGBoost)特征模型和基于遗传算法(genetic algorithm, GA)的封装模型的处理

理,提取最有预测意义的影像组学特征。交叉验证结果表明,基于 GA 的封装模型在预测 GBM 中的 MGMT 甲基化状态方面取得了较高的效能,敏感度为 0.894,特异度为 0.966,准确度为 0.925,该模型可能会很好地预测 GBM 患者的 MGMT 甲基化状态。研究人员还将提取的 GBM 影像组学特征应用于 LGG 的数据集,也达到了 0.780 的敏感度、0.620 的特异度和 0.750 的准确度,表明所选影像组学特征还有可能广泛地应用于 LGG 和 HGG。在胶质瘤的治疗中,临床医生便可根据肿瘤的特点和分级来优化治疗,包括化疗、放射治疗和手术切除。

3. 影像组学对脑胶质瘤术后肿瘤复发与治疗相关影响的鉴别

胶质瘤放疗后的一种典型的继发性改变是 MRI 影像上的异常强化,而真正的肿瘤复发的影像表现也与之相似,这使得真正的肿瘤复发与治疗相关效应看起来都是 MRI 影像上的异常强化,从而使它们难以迅速区分。因此,需要从治疗相关效应中及时鉴别真正的肿瘤复发,并根据需要调整治疗计划。但是,传统 MRI 特征很难区分胶质瘤真正的肿瘤复发与治疗相关效应。Ren 等^[40]开发了一种基于多模态 MRI 影像组学的机器学习模型,以提高识别胶质瘤真正肿瘤复发的效率。该研究将招募的 131 例患者随机分为训练集($n=91$)和测试集($n=40$),并从 4 个常规 MRI 序列的术后增强区域和水肿区域提取影像组学特征,对所提取的特征进行分析后选择影像组学特征构建 SVM 和 k-近邻(k-nearest neighbor, KNN)模型。经过验证,术后增强模型的 AUC 值显著高于水肿模型($P < 0.05$)。在使用单序列构建的模型中,使用 T_1WI+C 的术后增强区域特征的模型优于其他模型,SVM 的 AUC 值为 0.905,KNN 的 AUC 值为 0.899。在多模态模型中,术后增强模型优于水肿模型,SVM 的 AUC 为 0.931,KNN 的 AUC 为 0.896。结合常规序列和整个区域特征的多模态模型表现出更好的性能,SVM 的 AUC 为 0.965,KNN 的 AUC 为 0.955。根据以上结果得出结论,多模态影像组学可以很好地区分真正的肿瘤复发与治疗相关效应,还可以帮助临床医生更有效和可靠地评估患者的复发状态,并有助于对胶质瘤患者进行个性化管理和治疗,此外还可以帮助胶质瘤术后患者进行风险分层和生存预测。

影像组学在选择最佳治疗方案中的应用

抗血管生成治疗是治疗复发性 GBM 的首选方法。Kickingereder 等^[41]研究了可能能够预测治疗结果的成像生物标记物。在贝伐单抗抗血管生成治疗之前,对复发性 GBM 进行多序列 MRI 扫描,提取并分

析 172 例患者的 MRI 影像学特征,利用这些影像学特征建立了一个模型来预测低风险组与高风险组 PFS 和 OS 的中位数。作者得出结论,影像组学可能能够确定哪些患者将从贝伐单抗治疗中受益最多。

然而,在贝伐单抗等抗血管生成治疗后,确定肿瘤反应变得特别困难,因为这些药物在缺乏真正的肿瘤反应的情况下,往往会降低增强扫描和 FLAIR 序列的高信号^[42]。Bahrami 等^[43]对 33 例 HGG 患者在贝伐单抗治疗前后进行了评估,分割 FLAIR 高信号区内的感兴趣体积(volume of interest, VOI),并使用 3D-FLAIR 图像提取每个 VOI 的 EC 大小。研究结果显示,FLAIR 高信号区内更低的 EC 预示着 HGG 患者的 PFS($P=0.009$)和 OS 更差($P=0.022$)。FLAIR 图像边缘形状的定量评估具有作为可靠生物标志物的潜力,可对接受贝伐单抗治疗的 HGG 患者的无强化肿瘤的进展进行监测。相对于以往研究中使用的视觉评级来说,这种 3D 边缘提取方法可以提高 FLAIR 图像边缘形状评估的精度和可靠性。

展望与总结

限制影像组学和影像基因组学量化的一个主要因素是可重复性差,其次是可变性和缺乏一致性,这归因于缺乏标准化的采集参数和组学方法^[44]。此外,纹理分析软件可用性的变化增加了标准化的复杂性和重复性,不同研究使用的软件具有不同算法,使得这些研究几乎没有重复的可能。需要未来的研究来评估这些来自不同类型软件的结果的准确性,以帮助实现标准化。

影像组学和影像基因组学是一个集影像学、肿瘤学、分子生物学和机器学习于一体的新兴且发展迅速的领域,其提高了医学影像数据的利用率,可通过客观的肿瘤特征来提升医学影像的作用,能够以非侵入性的方式准确识别成人脑胶质瘤与其他颅内肿瘤,帮助临床专家发现肉眼无法观察到的肿瘤特征,并在预测成人脑胶质瘤的分级、基因分型和预后方面表现优异,为个体化诊疗模式的临床应用提供了科学依据。在神经外科领域,它还可以用于改进手术前和手术后的治疗计划。随着临床数据的增长和先进机器学习方法的出现,影像组学和影像基因组学在精确诊断和肿瘤学中将发挥越来越重要的作用。

参考文献:

- [1] Lapointe S, Perry A, Butowski NA. Primary brain tumours in adults[J]. Lancet, 2018, 392(10145): 432-446.
- [2] Dangelo L, Armocida D, Sampirisi L, et al. Role of endoscopic surgical biopsy in diagnoses of intraventricular/periventricular tumors: review of literature including a monocentric case series [J]. Acta Neurol Belg, 2020, 120(3): 517-530.
- [3] Gulsen S. Achieving higher diagnostic results in stereotactic brain

- biopsy by simple and novel technique[J]. *Open Access Maced J Med Sci*, 2015, 3(1): 99-104.
- [4] Marwa I, Ramon C, Kaustav B, et al. Spatial-and-context aware (Space) "virtual biopsy" radiogenomic maps to target tumor mutational status on structural MRI[M]. MICCAI, 2020; 305-314.
- [5] Davatzikos C, Barnholtz-Sloan JS, Bakas S, et al. Ai-based prognostic imaging biomarkers for precision neuro-oncology: the respond consortium[J]. *Neuro Oncol*, 2020, 22(6): 886-888.
- [6] Davnall F, Yip CS, Ljungqvist G, et al. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? [J]. *Insights Imaging*, 2012, 3(6): 573-589.
- [7] Miles KA, Ganeshan B, Hayball MP. CT texture analysis using the filtration-histogram method: what do the measurements mean? [J]. *Cancer Imaging*, 2013, 13(3): 400-406.
- [8] Skogen K, Schulz A, Dormagen JB, et al. Diagnostic performance of texture analysis on MRI in grading cerebral gliomas[J]. *Eur J Radiol*, 2016, 85(4): 824-829.
- [9] Weller M, Van Den Bent M, Tonn JC, et al. European association for neuro-oncology (Eano) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas[J]. *Lancet Oncol*, 2017, 18(6): e315-e329.
- [10] Cairncross G, Wang M, Shaw E, et al. Phase iii trial of chemoradiotherapy for anaplastic oligodendroglioma: long-term results of rtog 9402[J]. *J Clin Oncol*, 2013, 31(3): 337-343.
- [11] Tian Q, Yan LF, Zhang X, et al. Radiomics strategy for glioma grading using texture features from multiparametric MRI[J]. *J Magn Reson Imaging*, 2018, 48(6): 1518-1528.
- [12] Xie T, Chen X, Fang J, et al. Textural features of dynamic contrast-enhanced MRI derived model-free and model-based parameter maps in glioma grading[J]. *J Magn Reson Imaging*, 2018, 47(4): 1099-1111.
- [13] Van Den Bent MJ, Geurts M, French PJ, et al. Primary brain tumours in adults[J]. *Lancet*, 2023, 402(10412): 1564-1579.
- [14] Diamandis P, Aldape K. World health organization 2016 classification of central nervous system tumors[J]. *Neurol Clin*, 2018, 36(3): 439-447.
- [15] Soni N, Priya S, Bathla G. Texture analysis in cerebral gliomas: a review of the literature[J]. *AJNR Am J Neuroradiol*, 2019, 40(6): 928-934.
- [16] Zhou H, Vallieres M, Bai HX, et al. MRI features predict survival and molecular markers in diffuse lower-grade gliomas[J]. *Neuro Oncol*, 2017, 19(6): 862-870.
- [17] Zhang X, Qiang T, Yu XW, et al. Idh mutation assessment of glioma using texture features of multimodal MR images[C]. *Proceedings of SPIE*, 2017.
- [18] Han L, Wang S, Miao Y, et al. MRI texture analysis based on 3d tumor measurement reflects the idh1 mutations in gliomas—a preliminary study[J]. *Eur J Radiol*, 2019, 112: 169-179.
- [19] Jakola AS, Zhang YH, Skjulsvik AJ, et al. Quantitative texture analysis in the prediction of idh status in low-grade gliomas[J]. *Clin Neurol Neurosurg*, 2018, 164: 114-120.
- [20] Bahrami N, Hartman SJ, Chang YH, et al. Molecular classification of patients with grade ii/iii glioma using quantitative MRI characteristics[J]. *J Neurooncol*, 2018, 139(3): 633-642.
- [21] 江少凡, 宋阳, 薛蕴菁, 等. 基于 MRI 影像组学特征构建脑胶质瘤 idh1 突变预测模型[J]. *放射学实践*, 2023, 38(12): 1493-1499.
- [22] Petrecca K, Guiot MC, Panet-Raymond V, et al. Failure pattern following complete resection plus radiotherapy and temozolomide is at the resection margin in patients with glioblastoma[J]. *J Neurooncol*, 2013, 111(1): 19-23.
- [23] Malik N, Geraghty B, Dasgupta A, et al. MRI radiomics to differentiate between low grade glioma and glioblastoma peritumoral region[J]. *J Neurooncol*, 2021, 155(2): 181-191.
- [24] Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multi-forme: exploratory radiogenomic analysis by using quantitative image features[J]. *Radiology*, 2014, 273(1): 168-174.
- [25] Gutman DA, Cooper LA, Hwang SN, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the tcga glioblastoma data set[J]. *Radiology*, 2013, 267(2): 560-569.
- [26] Haddad AF, Young JS, Oh JY, et al. The immunology of low-grade gliomas[J]. *Neurosurg Focus*, 2022, 52(2): E2.
- [27] 陈燕生, 杨哲, 沈卓, 等. MRI 影像组学预测较低级别胶质瘤免疫亚型的临床研究[J]. *放射学实践*, 2023, 38(11): 1374-1379.
- [28] Mallick S, Benson R, Hakim A, et al. Management of glioblastoma after recurrence: a changing paradigm[J]. *J Egypt Natl Canc Inst*, 2016, 28(4): 199-210.
- [29] Louis DN, Aldape K, Brat DJ, et al. Cimpact-now (the consortium to inform molecular and practical approaches to cns tumor taxonomy): a new initiative in advancing nervous system tumor classification[J]. *Brain Pathol*, 2017, 27(6): 851-852.
- [30] Rapisarda A, Melillo G. Overcoming disappointing results with antiangiogenic therapy by targeting hypoxia[J]. *Nat Rev Clin Oncol*, 2012, 9(7): 378-390.
- [31] Beig N, Patel J, Prasanna P, et al. Radiogenomic analysis of hypoxia pathway is predictive of overall survival in glioblastoma[J]. *Sci Rep*, 2018, 8(1): 7.
- [32] Zhou X, Qu M, Tebon P, et al. Screening cancer immunotherapy: when engineering approaches meet artificial intelligence[J]. *Adv Sci (Weinh)*, 2020, 7(19): 2001447.
- [33] Liu D, Chen J, Ge H, et al. Radiogenomics to characterize the immune-related prognostic signature associated with biological functions in glioblastoma[J]. *Eur Radiol*, 2023, 33(1): 209-220.
- [34] Zinn PO, Mahmood Z, Elbanan MG, et al. Imaging genomics in gliomas[J]. *Cancer J*, 2015, 21(3): 225-234.
- [35] Pinker K, Shitano F, Sala E, et al. Background, current role, and potential applications of radiogenomics[J]. *J Magn Reson Imaging*, 2018, 47(3): 604-620.
- [36] Kuo MD, Jamshidi N. Behind the numbers: decoding molecular phenotypes with radiogenomics—guiding principles and technical considerations[J]. *Radiology*, 2014, 270(2): 320-325.
- [37] Hegi ME, Diserens AC, Gorlia T, et al. Mgmt gene silencing and benefit from temozolomide in glioblastoma[J]. *N Engl J Med*, 2005, 352(10): 997-1003.
- [38] Kong Z, Lin Y, Jiang C, et al. ¹⁸F-fdg-pet-based radiomics signature predicts mgmt promoter methylation status in primary diffuse glioma[J]. *Cancer Imaging*, 2019, 19(1): 58.
- [39] Do DT, Yang MR, Lam LHT, et al. Improving mgmt methylation status prediction of glioblastoma through optimizing radiomics features using genetic algorithm-based machine learning ap-

- proach[J].Sci Rep,2022,12(1):13412.
- [40] Ren J,Zhai X,Yin H, et al.Multimodality MRI radiomics based on machine learning for identifying true tumor recurrence and treatment-related effects in patients with postoperative glioma[J].Neurol Ther,2023,12(5):1729-1743.
- [41] Kickingreder P,Burth S,Wick A, et al.Radiomic profiling of glioblastoma:identifying an imaging predictor of patient survival with improved performance over established clinical and radiologic risk models[J].Radiology,2016,280(3):880-889.
- [42] Wen PY,Macdonald DR,Reardon DA, et al.Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group[J].J Clin Oncol,2010,28(11):1963-1972.
- [43] Bahrami N,Piccioni D,Karunamuni R, et al.Edge contrast of the flair hyperintense region predicts survival in patients with high-grade gliomas following treatment with bevacizumab[J].AJNR Am J Neuroradiol,2018,39(6):1017-1024.
- [44] Erickson B J,Korfiatis P,Akkus Z, et al.Machine learning for medical Imaging[J].Radiographics,2017,37(2):505-515.
- (收稿日期:2024-01-10 修回日期:2024-06-09)

《放射学实践》杂志微信公众平台开通啦!

遵照同行评议、价值导向、等效应用原则,国内各大学会、协会、组织机构通过科技工作者推荐、专家评议、结果公示等规定程序,《放射学实践》杂志入选中国科协发布 10 大领域《我国高质量科技期刊分级目录》业内认可的较高水平期刊。《放射学实践》杂志入选 2023 年版北京大学和北京高校图书馆期刊工作研究会共同主持的国家自然科学基金项目“学术期刊评价及文献计量学研究”研究成果——《中国核心期刊要目总览》。

《放射学实践》杂志微信公众平台立足于准确地传递医学影像领域的最新信息,致力于为关注医学影像领域的广大人士服务。欢迎大家通过微信平台,以文字、图片、音频和视频等形式与我们互动,分享交流最新的医学影像资讯。您还可以通过微信平台免费阅读及搜索本刊所有发表过的论文,投稿作者可以查询稿件状态等。

您可以通过以下方式关注《放射学实践》杂志微信公众平台:

1. 打开微信,通过“添加朋友”,在搜索栏里直接输入“放射学实践”进行搜索。
2. 在“查找微信公众号”栏里输入“放射学实践”即可找到微信公众号,点击“关注”,添加到通讯录。
3. 打开微信,点击“扫一扫”,手机镜头对准下面的二维码,扫出后点击关注即可。



关注有惊喜!