

人工智能促进医学影像临床应用与研究

陈冲, 陈俊, 夏黎明

【摘要】 基于人工智能(AI)的医学影像分析技术被应用于疾病病灶的自动分割、辅助诊断以及治疗规划等方面,在深化医学影像临床研究与应用,减轻放射科医生临床读片负担以及挖掘新的影像生物标志物等方面起到了重要作用。尽管 AI 在医学影像中的应用取得了显著进展,但仍面临一些挑战,包括算法的泛化能力、模型可解释性以及大规模临床验证等。因此,医学影像在拥抱 AI 技术的同时也需要加强人工专家最终环节的审核和把控。为了促进 AI 在医学影像领域的进一步发展,需要医学、工程和研究领域的紧密合作。

【关键词】 医学影像学; 人工智能; 影像组学; 自动分割; 深度学习

【中图分类号】 R814.42; R445.2 **【文献标志码】** A **【文章编号】** 1000-0313(2024)01-0012-05

DOI:10.13609/j.cnki.1000-0313.2024.01.003

开放科学(资源服务)标识码(OSID):



在当今临床工作实践中,医学影像技术已成为日常诊断重要工具。从 X 射线到 CT、MRI,再到超声和核医学影像,这些影像技术不仅帮助医生进行疾病诊断,还在病情的监测和疗效评估中发挥着至关重要的作用。但是医生的阅片精准度、操作效率以及定量分析方面仍然存在局限性,往往依赖于放射科医生的经验、专业知识以及医生所处工作状态。这些主观因素均可能导致诊断结果的不稳定,从而影响患者的疾病诊断、手术治疗以及预后监测。

近年来随着医学影像成像技术的规范及其广泛使用,基于数据驱动和端到端学习的人工智能(artificial intelligence, AI)算法逐渐在医学影像辅助诊断中大放异彩。人工智能的医学影像分析算法在海量数据中学习训练,自适应地匹配不同的影像决策任务,这些技术不仅提高了疾病诊断的精确性和效率,还减少了放射科医生的工作负担,使他们能够更专注于复杂案例的分析。其中基于人工智能算法的诊断和分割是其应用在医学影像领域的两个重要研究方向,本文将重点介绍人工智能算法在医学影像自动分割、诊断中的研究现状以及未来的应用前景。

AI 算法发展与介绍

21 世纪以来,随着计算机软件、互联网技术以及微电子集成电路的发展,困扰 AI 领域多年的数据以及算力等问题得到了很大程度的解决,2006 年图灵奖

获得者 Geoffrey Hinton 利用深度神经网络训练自编码器以降低数据维度^[1],文章中提出的很多思想为后续的研究提供了新的视角和方法。2012 年 Hinton 和他的学生 Alex Krizhevsky 设计 AlexNet 获得 ImageNet 图像分类的冠军^[2]。相比传统机器学习特征的浅层判别,卷积神经网络通过不断堆叠的卷积层和池化层对原始数据进行自主地特征提取和筛选,由浅层图像纹理特征到深层语义特征转变,“深度学习”(deep learning, DL)由此得名。从 2012 年开始,在算法层面,数据科学家们在深度学习基础网络的理论研究,计算机视觉以及自然语言的应用等方面有了重大突破,深度学习算法能够适配越来越多的落地应用场景,特别是医学影像领域;在硬件层面,以 NVIDIA 公司、Intel 公司为代表的 AI 芯片公司在人工智能模型的训练端芯片、云服务器的推理芯片以及边缘计算的部署芯片等取得了多样性的发展,在硬件的加持下,人工智能产品和解决方案能够更快发展;在数据层面,万物互联理念和大数据技术的发展将越来越多的图像数据以及文本分享到互联网上,进一步推动深度学习技术的发展。2022 年,openAI 公司分别在 GPT-3^[3]的基础上提出智能对话机器人 chatGPT,在 CLIP^[4]和扩散模型(diffusion model)^[5]的基础上提出了多模态绘图产品 DALL-E 系列,这些产品极大地推动了大模型以及弱人工智能模型向强人工智能模型的发展。在过去的十年里,这些技术也逐渐应用到医学影像领域中,给放射科医生的工作模式以及疾病的诊疗预后优化带了巨大的革新。

人工智能在医学影像中的研究进展

1. 影像组学的应用

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(陈冲、夏黎明);武汉,拜耳医药保健有限公司(陈俊)

作者简介: 陈冲(1994—),男,河南信阳人,博士,住院医师,主要从事人工智能在肺结节、新冠肺炎、冠脉 CTA 中的应用研究工作。

通讯作者: 陈俊, E-mail: chencharrel@gmail.com; 夏黎明, E-mail: lmxia@tjh.tjmu.edu.cn

影像组学作为一种高通量的影像特征提取建模方法,能够将高维的原始影像学数据转换为所研究区域的影像学特征,主要包括形态学特征、一阶特征、纹理特征以及高阶特征^[6]。相比于传统放射学参数,影像组学特征通过客观的图像描述以及信息论原理等方法更加全面地体现所研究区域的形态学参数、影像信号强度分布、统计学参量、影像纹理定量指标、影像变换的全局-局部特征以及频域变换的特征等。机器学习作为人工智能的重要分支,在影像组学的后续研究中起到了关键性的作用。目前影像组学作为一种无创的疾病评估方法,在疾病诊疗与预后的全流程中都有广泛研究和应用。

在疾病诊断方面,研究人员从不同的医学影像以及影像模态上提取影像组学特征,通过 mRMR、LASO 以及特征差异性分析等方法进行影像特征筛选后,利用多因素逻辑回归、支持向量机、随机森林以及 XGBoost 等方法构建影像学风险评分,将其与临床因素结合构建联合模型,应用于疾病诊断并通过模型的判别性、校正性以及临床应用验证其临床收益^[7-10],研究表明相比临床资料以及影像的半定量诊断方法,基于机器学习的影像组学方法能够提升疾病的诊断效能,辅助临床诊断决策。乳腺癌腋窝淋巴转移状态是乳腺癌患者治疗规划以及预后复发风险评估的重要生物学标志物,Yu 等^[11]在 MRI 影像上提取腋窝淋巴和乳腺癌区域的影像学特征,通过随机森林进行特征筛选后,构建基于支持向量机的影像组学模型无创预测乳腺癌患者腋窝淋巴转移状态,并在 803 例训练集、179 例外部验证集以及前瞻性-回顾性临床实验收集的 106 例研究队列得到了验证,AUC 为 0.90~0.93,影像学特征不仅能够描述乳腺癌内部的定量特征,也能够反映浸润性乳腺癌免疫微环境的改变情况。此外,乳腺癌特征结合淋巴结特征能够有效提高模型的诊断效能,为后续其他肿瘤淋巴转移问题提供了新的思路和方法^[12-14]。

在疾病的预后风险评估方面,多项关于脑出血预后研究表明基于平扫 CT 的影像组学特征构建的机器学习模型能够有效评估脑出血患者短期不良预后的情况,辅助医生评估脑出血患者的术后不良风险^[15-18]。Yu 等^[19]从多参数 MRI(T1ce、T₂WI、DWI-ADC)提取乳腺癌以及腋窝淋巴结的影像特征用来预测和分析乳腺癌患者腋窝淋巴结转移风险与无疾病生存期(DFS),该研究纳入了中国 4 个中心 1214 例乳腺癌患者,构建了多机器学习模型,采用 LASSO 和随机森林进行重要特征筛选,支持向量机和逻辑回归进行建模分析,该研究表明所构建的临床决策诺莫图能够帮助乳腺癌患者无创诊断腋窝淋巴结转移以及术后无疾病

生存情况,这对早期乳腺癌患者的手术计划以及治疗规划有重要意义。

由此看来,基于机器学习的影像组学方法已经广泛应用于疾病鉴别诊断、治疗前规划、治疗后疾病反应以及患者后续生存状态等方面,为个性化、精准治疗提供重要依据。

2. 智能诊断

相比基于机器学习的影像组学方法,基于深度学习的智能诊断能够通过端到端数据驱动的方式自主学习特征,进行医学影像诊断分析。深度学习智能诊断研究的标准流程包括:①数据集:建立标准化的医学影像数据集,包含配对的图像数据、勾画 ROI 数据以及患者匿名化的临床数据;②图像预处理:通过 Z-score 以及重采样的方法对医学影像数据进行灰度和空间归一化,减小设备、扫描方式的差异性;③模型选择:根据数据量情况,选择合适的深度学习模型,包括 ResNet^[20]、VGGNet^[21]、InceptionNet^[22] 及 3DConv 等^[23];④模型训练:根据数据量情况,基于 ImageNet 预训练参数对模型部分或全部权重使用较小的学习率进行微调^[24];⑤模型验证:将微调好的模型在训练集、测试集以及外部验证集中测试验证并与其他传统诊断结果(如人工诊断及影像组学模型等)对比其性能优劣。在多项基于深度学习的医学影像诊断研究中发现基于深度学习的影像学模型拥有更好的诊断效能并且在多个外部验证集得到充分验证^[25-28]。在医学影像临床实践中,多任务的医学影像分析也是目前研究的热点,研究表明基于多任务的深度学习模型在提高单一任务的预测效能、减缓模型的过拟合现象以及提高深度学习可解释性方面有独特优势^[29-31]。

随着 Transformer 架构在自然语言研究中的广泛应用^[32],相对卷积神经网络逐层提高感受野,Transformer 通过自注意力机制策略能够无监督地获取文本数据的全局信息,提高网络特征的表达能力,能够很好地处理数据分布异制性及长尾问题,目前成为人工智能大模型研究的基础。在医学影像智能诊断的研究中,Transformer 网络架构也逐渐受到学者们的关注。多项研究表明,基于 ImageNet 预训练以及自监督学习预训练的 Transformer 深度学习模型能够更好地拟合医学影像全局的信息^[33],提高其诊断效能^[34-36]。值得注意的是,在一项大规模病理影像智能分析研究中,Huang 等^[37]构建了包含 20 万张病理影像的数据集 OpenPath,该数据集包含了病理影像以及对应的影像学人类语言描述。在 OpenPath 数据库和 CLIP 模型的基础上^[4],构建了病理影像的基础预训练大模型 PLIP,在多个下游任务(图像分类、图像检索等)中都有非常高效的表现。研究还表明相对于传统的分类类

别数量固定的模式,基于 CLIP 的大模型在不需要微调下能够快速拟合新的数据类别分布。

由此可见,基于人工智能的智能诊断研究由传统的影像组学、卷积神经网络以及多任务模型向预训练医学影像基础大模型发展,其诊断效能、模型的泛化能力越来越高,应用场景越来越多样化。但是,随着模型越来越复杂,特征层次逐渐加深,其算法的可解释性也受到多方面挑战。因此,大规模、规范化的医学影像数据集、大模型算法的可解释性是未来人工智能诊断的难点和发展方向。

3. 自动分割

2015 年 Ronneberger 等^[38]在卷积神经网络的基础上提出 UNet 架构应用于医学影像自动分割上,网络结构主要由编码器-解码器构成。其编码器逐层卷积提取图像的语义信息,特征图尺度逐层减小;解码器再通过反卷积操作逐层恢复特征图尺度并与编码特征融合分割目标区域,在医学图像分割领域有较高的效能。在后续的多个研究中分别从网络结构、损失函数以及分割目标部位等方面进行模型优化以更好地适配医学影像不同分割部位、不同模态数据分布的特点(CT、MRI、US 以及胃肠镜等),以达到更好的分割效果^[39-42]。但在医学影像分割领域中,图像重建参数、三维图像的层厚、扫描范围以及不同设备、不同中心的影像等都会影响模型的分割效能。

一方面,如何对多样性的医学影像进行高效预处理以及算法输出结果的后处理是人工智能算法应用在医学影像分割领域首先要考虑的问题。2020 年 Isensee 等^[43]提出一种一站式医学影像预处理、网络架构自动配置、自动训练、多分割模型分割融合以及模型评估的 nnUNet 架构,并在多个公开医学影像数据集中验证其分割效能^[44-47]。在数据预处理阶段,通过一系列自动配置方案根据使用者显卡情况对分割医学影像数据集进行自动分析,选择合适的层厚、扫描范围、预处理方案以及 UNet 编码器各个维度下采样方案对原始医学影像预处理;在模型训练阶段,采用基于空间位置的增强策略(如旋转、镜像翻转等)与基于灰度强度的增强策略(如高斯模糊、色度变换等)提高分割模型的泛化能力与分割精度;在模型应用阶段,nnUNet 采用窗口的方式进行滑动推理,并对窗口推理结果采用高斯融合提高分割精度;在模型评估阶段,采用 Dice 系数、Jaccard 系数以及真阳性率等多个指标综合评估模型预测效能。由于 nnUNet 架构易于使用及较高的分割效能,目前已在医学影像分割领域被广泛使用。

另一方面,如何结合最新人工智能技术改进 Unet 架构以提高多脏器结构化分割效能和泛化性是当前研

究的难点,目前已经有多项研究^[48-49]结合 Transformer^[50]、扩散概率模型^[5]、预训练大模型的思想^[48,51]以及自监督学习的策略^[33]对 UNet 网络结构进行改进,以通过自然语言描述、少量的人工交互以及全自动的方式实现模型自主的目标区域分割,让智能体模型更加理解医生的意图,进行大范围多模态的医学影像精确分割。

除了在预处理以及模型方面的改进之外,利用人工智能分割技术辅助构建标准化的医学影像数据集也是医学影像精细化分割的另一发展方向。为了提高模型的分割泛化性和精度,相关学者使用 nnUNet 框架在来自于不同中心、不同层厚、不同空间分辨率以及不同增强期相的 1204 个 CT 影像上训练了多个分割模型,能够分割 104 个人体解剖结构,其中包括 27 个器官、59 种骨骼结构、10 种肌肉组织以及 8 种血管,推动了人工智能分割模型从算法研究到落地应用的发展。目前人工智能医学影像软件分割算法正在改变放射科医生临床科研的工作方式。在一项大规模医学影像分割的研究中,在人工智能分割模型的基础上建立迭代式人工辅助半自动分割方案,在该方案中“自动分割-专家矫正-模型迭代”的模式能够帮助医生节省大量的手工标注时间,并且模型的分割精度随着迭代的进行,其效能不断提升以拟合新的数据分布,该方案仅用 3 个星期便可以完成 8000 例腹部 CT 的多脏器分割任务,而一名经验丰富的腹部放射科医生可能需要 30 年的时间才能完成,极大提高了标注效率。

由此可见,大规模的医学影像自动分割工具和软件是未来人工智能在医学影像应用的重要方向和研究趋势。

总 结

随着人工智能技术、计算机算力以及医学影像数据库标准化的发展,人工智能技术正在改变放射科医生日常的工作模式。人工智能新技术在疾病的早期筛查、临床诊断、病灶检测与分割、治疗方式的选择与规划以及治疗后的疗效评估与生存分析等临床工作中起到了不可忽视的作用,为临床和科研工作带来了巨大的便利。然而,人工智能作为一项新兴的技术,依然在不断发展和完善。医学影像相关算法模型仍不成熟,存在系统性故障以及模型误判漏判的风险,可能造成临床评估、诊断、方案规划等工作的误判。因此放射科医生在拥抱人工智能的同时仍需保持清醒客观的认识,不能形成完全依赖,AI 辅助的最终环节仍需医学专家严格地加以人工审核。未来人工智能技术要想进一步与医学影像融合,其模型的可解释性、算法的泛化性和鲁棒性、如何克服医学影像的多样性和复杂性以

及大规模数据的验证是人工智能医学影像领域研究和落地的难点和发展方向。

参考文献:

- [1] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural Networks[J]. Science, 2006, 313(5786): 504-507.
- [2] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks[J]. Commun ACM, 2017, 60(6): 84-90.
- [3] Brown TB, Mann B, Ryder N. Language models are few-shot learners[J]. arXiv, 2020. DOI: 10.48550/arXiv.2005.14165.
- [4] Radford A, Kim JW, Hallacy C. Learning transferable visual models from natural language supervision[J]. arXiv, 2021. DOI: 10.48550/arXiv.2103.00020.
- [5] Nichol A, Dhariwal P. Improved denoising diffusion probabilistic models[J]. arXiv, 2021. DOI: 10.48550/arXiv.2102.09672.
- [6] Lambin P, Rios-Velazquez E, Leijenaar R. Radiomics: extracting more information from medical images using advanced feature analysis[J]. Eur J Cancer, 2012, 48(4): 441-446.
- [7] Liu A, Wang Z, Yang Y. Preoperative diagnosis of malignant pulmonary nodules in lung cancer screening with a radiomics nomogram[J]. Cancer Commun, 2020, 40(1): 16-24.
- [8] Zheng Y, Zhou D, Liu H. CT-based radiomics analysis of different machine learning models for differentiating benign and malignant parotid tumors[J]. Eur Radiol, 2022, 32(10): 6953-6964.
- [9] Song T, Zhang QW, Duan SF. MRI-based radiomics approach for differentiation of hypovascular non-functional pancreatic neuroendocrine tumors and solid pseudopapillary neoplasms of the pancreas[J]. BMC Med Imaging, 2021, 21(1): 36.
- [10] Ma Y, Xu X, Pang P. A ct-based tumoral and mini-peritumoral radiomics approach: differentiate fat-poor angiomyolipoma from clear cell renal cell carcinoma[J]. Cancer Manag Res, 2021, 13(2): 1417-1425.
- [11] Yu Y, He Z, Ouyang J. Magnetic resonance imaging radiomics predicts preoperative axillary lymph node metastasis to support surgical decisions and is associated with tumor microenvironment in invasive breast cancer: A machine learning, multicenter study[J]. EBioMedicine, 2021, 69(6): 103460.
- [12] Yang J, Wang L, Qin J. Multi-view learning for lymph node metastasis prediction using tumor and nodal radiomics in gastric cancer[J]. Phys Med Biol, 2022, 67(5): 055007.
- [13] Liu Y, Li X, Zhu L. Preoperative prediction of axillary lymph node metastasis in breast cancer based on intratumoral and peritumoral DCE-MRI radiomics nomogram[J]. Contrast Media Mol Imaging, 2022, 2022(8): 6729473.
- [14] Ma M, Jiang Y, Qin N. A Radiomics model for preoperative predicting sentinel lymph node metastasis in breast cancer based on dynamic contrast-enhanced MRI[J]. Front Oncol, 2022, 12(6): 884599.
- [15] Xie Y, Chen F, Li H. Development and validation of a clinical-radiomics nomogram for predicting a poor outcome and 30-day mortality after a spontaneous intracerebral hemorrhage [J]. Quant Imaging Med Surg, 2022, 12(10): 4900-4913.
- [16] Fang C, An X, Li K. A nomogram based on CT radiomics and clinical risk factors for prediction of prognosis of hypertensive intracerebral hemorrhage[J]. Comput Intell Neurosci, 2022, 2022(12): 9751988.
- [17] Huang X, Wang D, Zhang Q. Radiomics for prediction of intracerebral hemorrhage outcomes: A retrospective multicenter study[J]. Neuroimage Clin, 2022, 36: 103242.
- [18] Li H, Xie Y, Liu H. Non-contrast CT-based radiomics score for predicting hematoma enlargement in spontaneous intracerebral hemorrhage[J]. Clin Neuroradiol, 2022, 32(2): 517-528.
- [19] Yu Y, Tan Y, Xie C. Development and validation of a preoperative magnetic resonance imaging radiomics-based signature to predict axillary lymph node metastasis and disease-free survival in patients with early-stage breast cancer[J]. JAMA Netw Open, 2020, 3(12): e2028086.
- [20] He K, Zhang X, Ren S. Deep residual learning for image recognition[J]. arXiv, 2015. DOI: 10.1109/CVPR.2016.90.
- [21] Simonyan K, Zisserman A. Very deep convolutional Networks for large-scale image recognition[J]. arXiv, 2015. DOI: 10.48550/arXiv.1409.1556.
- [22] Szegedy C, Vanhoucke V, Ioffe S. Rethinking the inception architecture for computer vision[J]. arXiv, 2015. DOI: 10.1109/CVPR.2016.308.
- [23] Tran D, Bourdev L, Fergus R. Learning spatiotemporal features with 3D convolutional Networks[J]. arXiv, 2015. DOI: 10.1109/ICCV.2015.510.
- [24] Russakovsky O, Deng J, Su H. ImageNet large scale visual recognition challenge[J]. arXiv, 2015. DOI: 10.1007/s11263-015-0816-y.
- [25] She Y, He B, Wang F. Deep learning for predicting major pathological response to neoadjuvant chemoimmunotherapy in non-small cell lung cancer: A multicentre study[J]. EBioMedicine, 2022, 86(12): 104364.
- [26] Zhang H, Liao M, Guo Q. Predicting N2 lymph node metastasis in presurgical stage I - II non-small cell lung cancer using multiview radiomics and deep learning method[J]. Med Phys, 2023, 50(4): 2049-2060.
- [27] Tian T, Gan T, Chen J. Graphic intelligent diagnosis of hypoxic-ischemic encephalopathy using MRI-based deep learning model[J]. Neonatology, 2023, 120(4): 441-449.
- [28] Wang S, Yu H, Gan Y. Mining whole-lung information by artificial intelligence for predicting EGFR genotype and targeted therapy response in lung cancer: a multicohort study[J]. Lancet Digit Health, 2022, 4(5): e309-e319.
- [29] Jiang Y, Zhang Z, Yuan Q. Predicting peritoneal recurrence and disease-free survival from CT images in gastric cancer with multitask deep learning: a retrospective study [J]. Lancet Digit Health, 2022, 4(5): e340-e350.
- [30] Zhong L, Dong D, Fang X. A deep learning-based radiomic nomogram for prognosis and treatment decision in advanced nasopharyngeal carcinoma: A multicentre study[J]. EBioMedicine, 2021, 70(8): 103522.
- [31] Fu S, Lai H, Li Q. Multi-task deep learning network to predict future macrovascular invasion in hepatocellular carcinoma [J]. EclinicalMedicine, 2021, 42(12): 101201.
- [32] Vaswani A, Shazeer N, Parmar N. Attention is all you need[J]. arXiv, 2023. DOI: 10.48550/arXiv.1706.03762.

- [33] He K, Chen X, Xie S. Masked autoencoders are scalable vision learners[J]. arXiv, 2021. DOI: 10.48550/arXiv.2111.06377.
- [34] Ma X, Xia L, Wan W. Development and validation of a deep learning signature for predicting lymph node metastasis in lung adenocarcinoma: comparison with radiomics signature and clinical-segmental model[J]. Eur Radiol, 2022, 33(3): 1949-1962.
- [35] Kang Q, Gao J, Li K. Deblurring masked autoencoder is better recipe for ultrasound image recognition[J]. arXiv, 2023. DOI: 10.48550/arXiv.2206.10207.
- [36] Huang Z, Shu Z, Zhu RH. Deep learning-based radiomics based on contrast-enhanced ultrasound predicts early recurrence and survival outcome in hepatocellular carcinoma[J]. World J Gastrointest Oncol, 2022, 14(12): 2380-2392.
- [37] Huang Z, Bianchi F, Yuksekgonul M. A visual-language foundation model for pathology image analysis using medical Twitter[J]. Nat Med, 2023, 29(9): 2307-2316.
- [38] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[J]. arXiv, 2015. DOI: 10.48550/arXiv.2111.00534.
- [39] Oktay O, Schlemper J, Folgoc LL. Attention U-Net: learning where to look for the pancreas[J]. arXiv, 2018. DOI: 10.48550/arXiv.1804.03999.
- [40] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N. Unet++: A nested u-net architecture for medical image segmentation[J]. Deep Learn Med Image Anal Multimodal Learn Clin Decis Support, 2018, 11045(12): 3-11.
- [41] Jha D, Smedsrud PH, Riegler MA. ResUNet++: an advanced architecture for medical image segmentation[J]. arXiv, 2019. DOI: 10.1109/ACCESS.2021.3071754.
- [42] Chen J, Lu Y, Yu Q. TransUNet: transformers make strong encoders for medical image segmentation[J]. Life, 2023, 13(4): 976.
- [43] Isensee F, Jäger PF, Kohl SAA. Automated design of deep learning methods for biomedical image segmentation[J]. Nat Methods, 2021, 18(2): 203-211.
- [44] Antonelli M, Reinke A, Bakas S. The medical segmentation decathlon[J]. Nat Commun, 2022, 13(1): 4128.
- [45] Litjens G, Toth R, Van De Ven W. Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge[J]. Med Image Anal, 2014, 18(2): 359-373.
- [46] Bernard O, Lalande A, Zotti C. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? [J]. IEEE Trans Med Imaging, 2018, 37(11): 2514-2525.
- [47] Bilic P, Christ P, Li HB. The liver tumor segmentation benchmark (LiTS)[J]. Med Image Anal, 2023, 84(2): 102680.
- [48] Li Y, Jing B, Li Z. nnSAM: plug-and-play segment anything model improves nnUNet performance[J]. arXiv, 2023. DOI: arxiv.org/abs/2309.1696.
- [49] Guo S, Chen Q, Wang L. nnUnetFormer: an automatic method based on nnUnet and transformer for brain tumor segmentation with multimodal MR images[J]. Phys Med Biol, 2023, 68(24): 245012.
- [50] Dosovitskiy A, Beyer L, Kolesnikov A. An image is worth 16×16 words: transformers for image recognition at scale[J]. arXiv, 2020. DOI: 10.48550/arXiv.2010.11929.
- [51] Ramesh A, Pavlov M, Goh G. Zero-Shot text-to-image generation[J]. arXiv, 2021. DOI: 10.48550/arXiv.2102.12092.

(收稿日期: 2024-01-06)

· 影像组学与人工智能专题 ·

影像组学模型对肺腺癌谱病变病理侵袭性的诊断价值

孙希子, 周舒畅, 夏黎明

【摘要】 目的:探讨针对肺腺癌谱病变的影像组学模型对其病理侵袭性的诊断效能。**方法:**回顾性分析我院经手术病理证实的 172 例肺腺癌谱病变(5~30 mm)患者的术前 CT 影像资料。采用 Pyradiomics 包提取术前 CT 图像病灶的影像组学特征,通过组间相关系数和带 L2 惩罚项的逻辑回归进行特征筛选,根据所选特征建立逻辑回归、随机森林、极致梯度提升分类器模型并绘制 ROC 曲线。由两位高年资放射科医生在不知道病理结果的情况下对结节侵袭性概率进行评分。采用 Delong 检验将三种分类器的诊断效能与年龄、ROI 体积、高年资医生的诊断效能进行比较。**结果:**从每一个感兴趣区提取、筛选后得到 420 个影像组学特征。逻辑回归、随机森林和极致梯度提升分类器模型在测试集上的 ROC 曲线下面积分别为 0.921、0.956 和 0.958。年龄、ROI 体积和高年资放射科医生在测试集的 ROC 曲线下面积分别为 0.620、0.863 和 0.896。Delong 检验提示三种分类器间的诊断效能差异无统计学意义($P>0.05$),三种分类器与 ROI 体积、高年资医生的诊断效能差异亦无统计学意义($P>0.05$)。**结论:**影像组学分类器模型进行肺腺癌谱病变的术前病理诊断具有较高的准确性,其诊断效能与 ROI 体积以及高年资医生相当。

【关键词】 肺腺癌; 肺结节; 影像组学; 体层摄影术, X 线计算机; Delong 检验; 人机竞赛

【中图分类号】 R734.2; R814.42 **【文献标志码】** A **【文章编号】** 1000-0313(2024)01-0017-05

DOI:10.13609/j.cnki.1000-0313.2024.01.004

开放科学(资源服务)标识码(OSID):



Diagnostic value of radiomics models for histological invasiveness of adenocarcinoma spectrum lesions of lung SUN Xi-zi, ZHOU Shu-chang, XIA Li-ming, Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

【Abstract】 Objective: To investigate the diagnostic performance of the radiomics model for pathological invasiveness of adenocarcinoma spectrum lesions of the lung. **Methods:** Preoperative CT imaging data of 172 patients with pathologically proven adenocarcinoma spectrum lesions of the lung in our hospital were retrospectively analyzed. Radiomics features were extracted from the preoperative CT images via Pyradiomics package and selected via intra-class correlation and logistic regression with L2 penalty. Logistic regression, random forest and extreme gradient boosting (XGboost) classifier model were established according to the selected features, and receiver operating characteristic (ROC) curves were plotted. Two senior radiologists scored the probability of nodule invasiveness without knowing the pathological findings. Delong test was used to compare the diagnostic performance of the three classifiers with that of age, volume of region of interest (ROI) and the senior radiologists. **Results:** 420 radiomics features were obtained from each ROI after feature extraction and selection. On test set, the area under the ROC curve of logistic regression, random forest and XGboost was 0.921, 0.956 and 0.958, respectively, and that of age, ROI volume and the senior radiologists was 0.620, 0.863 and 0.896, respectively. Delong test indicated that there was no significant difference among the three radiomics classifiers and no significant difference in diagnostic performance between the three classifiers and ROI volume or the senior radiologists either ($P>0.05$). **Conclusion:** The radiomics classifier models have high accuracy in preoperative pathological diagnosis of adenocarcinoma spectrum lesions of the lung, the diagnostic performance of which is comparable to that of ROI volume and senior radiolo-

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

作者简介: 孙希子(1998-), 女, 湖北荆州人, 博士研究生, 主要从事胸部影像学研究工作。

通讯作者: 夏黎明, E-mail: cjr.xialiming@vip.163.com

基金项目: 国家自然科学基金资助项目(82001785)

gists.

【Key words】 Lung adenocarcinoma; Lung nodule; Radiomics; Tomography, X-ray computed; Delong test; Man-machine competition

随着低剂量肺部 CT 筛查的广泛开展,肺结节的检出率越来越高。现有肺结节管理指南都是基于临床信息、影像学随访资料、PET/CT 结果、活检结果预测恶性概率而决定进一步管理方式^[1]。对于高度怀疑恶性的结节,活检和手术是需要选择的管理方式。通过活检和血清学肿瘤标志物检测,可鉴别小细胞肺癌(small cell lung cancer, SCLC)与非小细胞肺癌(non-small cell lung cancer, NSCLC),并可通过免疫组织化学染色进一步区分肺鳞癌与肺腺癌。然而不管是细针抽吸活检还是切割针活检都无法区分肺腺癌疾病谱。根据世界卫生组织(WHO)2021 年第五版肺肿瘤分类^[2],病理分型上肺腺癌谱分为腺体前驱病变-非典型腺瘤样增生(atypical adenomatous hyperplasia, AAH)、原位腺癌(adenocarcinoma in situ, AIS)和腺癌-微浸润性腺癌(minimally invasive adenocarcinoma, MIA)、浸润性腺癌(invasive adenocarcinoma, IAC),腺体前驱病变可逐步演变为肺腺癌^[3]。AAH 常在 5 mm 以内,很少被切除送检,在此不作讨论。据文献报道, AIS 和 MIA 患者术后的 5 年无病生存率(disease free survival, DFS)为 100%^[4,5],而 IAC 术后 5 年无病生存率为 54.0%~94.9% ($P < 0.001$)^[6],取决于占主导的组织学亚型。由于有极佳的预后, AIS 和 MIA 被研究者称为“惰性病变”^[7]。尽管尚有争议,大多数研究者认为定期随访或亚肺叶切除(楔形切除或节段性切除)适用于 AIS 和 MIA,而对 IAC 患者应该施行肺叶切除术^[8-10]。因此术前非侵入性地鉴别“惰性病变”与 IAC,对于存在肺腺癌谱病变患者的治疗指导和预后预测都有重要意义。

影像组学的概念首先由荷兰学者 Lambin 于 2012 年提出^[11],是指应用自动化数据特征化算法,将从影像特征中提取的医学影像数据转化为高分辨率可挖掘的特征空间数据^[12],即将医学影像转化为可进一步分析的高维数据,探索医学影像特征与疾病诊断或预后之间的关联。本研究旨在探讨肺腺癌谱病变的影像组学模型对其侵袭性的预测价值,以及影像组学模型与一般指标、人工判读能力的比较。

材料与方法

1. 研究对象

回顾性分析我院 2019 年 1 月至 2023 年 2 月肺腺癌切除术后 169 例患者(172 个结节)的临床、病理及影像资料,其中男 56 例,女 113 例,年龄 21~75 岁,平

均(51.87 ± 11.06)岁。病例纳入标准:①薄层 CT 图像上表现为 5~30 mm 的肺结节;②术前 1 个月内进行直接增强 CT 检查;③术中或术后病理证实为 AIS、MIA、IAC;④年龄 > 18 岁;⑤不限制有无淋巴结转移;⑥不限制有无毛刺征、分叶征、胸膜牵拉征等恶性征象。

2. CT 扫描和图像分割

所有患者入院后均常规行胸部薄层 MSCT 检查,层厚 1 mm,矩阵 512×512 ,标准算法重建。为增加研究的鲁棒性,纳入本研究的图像摄自不同厂家不同扫描参数的设置。将最终获得的数据和图像,通过 3D Slicer 软件完成结节的手动逐层分割。上述操作均由本院具有多年胸部影像学诊断经验的医师完成(工作年限 3 年左右),并进行观察者内和观察者间的可重复性研究 6 个月后由该医师和具备 5 年临床经验的医生再次分割,以评估组内和组间的一致性。

3. 特征提取

提取的特征可分为肿瘤强度、纹理、形状及小波特征四类。对体素大小重采样至 $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$,将体素强度值使用 64 HU 的固定间隔宽度进行离散化,对像素进行归一化处理后,提取包括形状特征、一阶特征、灰度共生矩阵(gray level cooccurrence, GLCM)特征、灰度区域大小矩阵(grey level size zone matrices, GLSZM)特征、灰度游程长度矩阵(gray level run length matrices, GLRLM)特征、邻域灰度差矩阵(neighbourhood gray-tone difference matrix, NGTDM)特征、灰度依赖矩阵(gray level dependence matrix, GLDM)特征和小波特征在内的特征用于进一步统计分析。采用组内相关系数(intra-class correlation coefficient, ICC)评估组间和组内一致性($\text{ICC} > 0.75$ 说明具有较好的一致性)。

4. 影像组学模型构建和人机竞赛

数据预处理:随机选择 60% 的样本用于训练模型,40% 的样本用于测试。划分好数据集后,首先对各数据集分别标准化,然后在训练集上通过带 L2 惩罚项的逻辑回归进行嵌入式特征筛选。

模型构建与评价:构建模型所使用的 3 种分类器分别为逻辑回归(logistic regression)、随机森林(random forest)和极致梯度提升(extreme gradient boosting, XGboost)。绘制对应的受试者工作特征(receiver operating characteristic, ROC)曲线并计算曲线下面积(area under curve, AUC)以评价模型的诊断效

能。通过 Delong 检验比较不同分类器之间 AUC 的差异。

人机竞赛:由另 2 位事先不知道病理结果的高年资放射科医生对所有图像进行侵袭性概率评分(0~1 分)。医生对于侵袭性的评价主要结合患者年龄、实性成分比例、毛刺征、胸膜凹陷等,根据临床经验对侵袭性概率进行主观评分,医生对侵袭性概率的最终评分为两高年资医生的评分均值。通过 Delong 检验比较医生与最优分类器的诊断效能。

5.统计学分析

采用 SPSS 25.0 软件进行统计学分析。*t* 检验用于正态分布的两独立样本的均值比较,皮尔逊卡方检验用于样本间构成比的比较,曼-惠特尼 U 检验用于非正态独立样本中位数的比较。采用组内相关系数(ICC)以及带 L2 惩罚项的逻辑回归进行特征筛选。通过逻辑回归(LG)、随机森林(RF)、极致梯度提升(XGboost)分别构建鉴别肺腺癌谱分型的影像组学模型。采用 ROC 曲线及 AUC 评价单个特征及模型的预测效能,Delong 检验比较不同诊断途径的诊断效能。以 $P<0.05$ 为差异有统计学意义。

结 果

1.一般资料

本研究共纳入 169 例患者,172 个结节,其中 AIS37 个,MIA47 个,IAC88 个。惰性病变组(AIS/MIA)与 IAC 组的患者年龄差异有统计学意义($P<0.05$),而两组在性别、有无吸烟史、病灶部位上差异无统计学意义($P>0.05$)。其中,惰性病变组无实性结节,而 IAC 组有 16 个实性结节,纯磨玻璃结节(pure ground glass nodule,pGGN)在惰性病变组中的比例显著高于 IAC 组,两组的病灶性质差异有统计学意义($P<0.05$)。惰性病变组的 ROI 体积显著小于 IAC 组,差异有统计学意义($P<0.05$,表 1)。

2.分类器预测效能

通过特征提取和 ICC 筛选共得到 863 个特征,通过进一步嵌入式特征筛选(对 $ICC>0.75$ 的特征进行带 L2 惩罚项的逻辑回归),共筛选得到 420 个特征。三种分类器的 ROC 曲线见图 1,逻辑回归、随机森林以及 XGboost 的 AUC 分别为 0.921、0.956 和 0.958。XGboost 模型在测试集上的混淆矩阵见图 2。Delong 检验结果提示三种分类器的诊断效能差异无统计学意义($P>0.05$,表 2)。

表 1 惰性病变组与 IAC 组的一般资料比较 (n,%)

特征	AIS/MIA (n=84)	IAC (n=88)	统计量	P 值
年龄(岁)	48.71±11.48	55.09±9.60	-3.959(<i>t</i>)	0.022
性别			1.188(χ^2)	0.276
男	24(28.6%)	32(36.4%)		
女	60(71.4%)	56(63.6%)		
吸烟史			1.320(χ^2)	0.251
有	10(11.9%)	16(18.2%)		
无	74(88.1%)	72(81.8%)		
病灶部位			1.627(χ^2)	0.202
右肺	52(61.9%)	46(52.3%)		
左肺	32(38.1%)	42(47.7%)		
病灶性质			58.13(χ^2)	0.000
磨玻璃结节	64(76.2%)	16(18.2%)		
实性/部分实性结节	20(23.8%)	72(81.8%)		
病灶体积(mm ³)	398.44(274.60~627.93)	1525.62(807.15~2960.57)	7.640(<i>Z</i>)	0.000

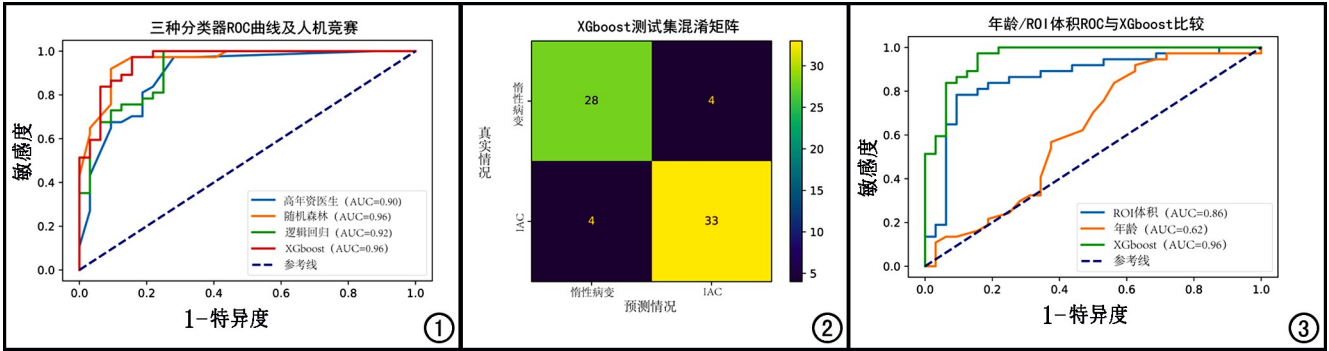


图 1 三种分类器的 ROC 曲线和人机竞赛。图 2 XGboost 在测试集上的混淆矩阵。图 3 年龄、ROI 体积、XGboost 诊断的 ROC 曲线。

表 2 Delong 检验结果

诊断方法 1 (AUC 更大)	诊断方法 2	Delong 检验 P 值
XGboost	年龄	0.000
ROI 体积	年龄	0.001
XGboost	ROI 体积	0.062
XGboost	逻辑回归	0.090
XGboost	高年资医生	0.111
随机森林	高年资医生	0.149
随机森林	逻辑回归	0.166
逻辑回归	高年资医生	0.569
XGboost	随机森林	0.936

3.各种诊断方法比较与人机竞赛

单独用年龄以及 ROI 体积诊断的 ROC 曲线见图 3,AUC 分别为 0.620 和 0.863,两者诊断效能差异有统计学意义($P<0.05$)。ROI 体积与 XGboost 的诊断效能差异无统计学意义($P>0.05$)。高年资医生诊断的 AUC 为 0.896,Delong 检验结果提示高年资医生与三种分类器的诊断效能差异均无统计学意义(P 均 >0.05 ,表 2)。

讨 论

本研究探讨了 CT 影像组学对肺腺癌谱病变侵袭性的诊断价值,并将其与一般诊断指标以及人工判读进行客观比较,研究结果表明 CT 影像组学用于术前判断肺腺癌谱病变侵袭性的准确性较高,与 ROI 体积及人工判读的诊断效能相当。

CT 作为临床上应用最为广泛的诊断技术之一,可为肺腺癌的诊断提供瘤体形态、边缘形态、瘤-肺界面、内部结构、瘤周征象等信息。传统的诊断方法依赖于这些肉眼可分辨的特征,而影像组学则利用大量可重复性强且精确的特征数据进行准确诊断。已有的对肺腺癌谱病变侵袭性诊断的影像组学研究集中于磨玻璃结节^[13-15]或部分实性结节^[16,17]。将实性结节和亚实性结节综合在一起的研究较少-She 等^[7]通过最小绝对收缩选择算子(least absolute shrinkage and selection operator,LASSO)-逻辑回归构建了基于 402 例肺腺癌谱病变的病理分型 CT 影像组学模型,发现模型评分和性别是预测肺腺癌谱病变侵袭性的独立因素。该研究模型的诊断效能较高,也更适用于临床情境。不过有基于分层分析的研究结果表明影像组学在预测磨玻璃结节/肿块浸润性中的效能高于在预测实性结节良恶性中的效能^[18]。因此,影像组学在不同类型肺结节/肿块中的独立及综合诊断效能有待进一步研究。

多项研究表明 CT 图像中肺腺癌谱结节实性成分含量与肿瘤浸润性显著相关^[19,20],这与本研究中纯磨玻璃结节(pGGN)在惰性病变中比例显著高于 IAC 组的结论相符。根据 Fleischner 指南^[21],临床上常使用

肺结节同一平面的长径和短径的均值作为球的直径计算结节体积。影像组学则是将 ROI 内体素的个数与体素的体积相乘得到结节体积,这种原理得出的体积更接近于结节的实际体积。本研究首次将影像组学中的 ROI 体积单独用于诊断效能评估,并得到 ROI 体积 AUC 与最优机器学习算法 AUC 差异无统计学意义的结论。

本研究所采用的三种分类器-逻辑回归、随机森林、XGboost 均属于常规广泛应用的算法,其中 XGboost 被最晚提出(2016 年),也表现出了最优的性能。Delong 检验是用于不同 ROC 曲线下面积差异显著性检验的算法,该算法不断被优化以适用于更多的科研场景^[22]。基于此,笔者在 Python 3.7.6 中编写了相应的程序以实现不同诊断方法之间 AUC 客观便捷的比较。

本研究存在以下局限性:①样本量较少且为单中心回顾性研究,而提取分析的特征相对较多;②图像分割均由人工手动实现,耗时耗力;③从 XGboost 在测试集上的混淆矩阵来看,影像组学的预测准确率还不够理想,但随机森林和 XGboost 算法在训练集上均无一例误判,说明这两种影像组学模型可能存在一定过拟合的情况;④图像来自多组参数的多台设备,对图像质量有一定影响。在今后的研究中,需要扩大样本量开展多中心前瞻性研究,并开发自动勾画感兴趣区的软件以及更优化的算法,更加深入便捷地探讨影像组学特征在肺腺癌谱病变术前病理诊断中的价值,早日实现影像组学在临床诊断中的应用。

综上所述,影像组学分类器模型进行肺腺癌谱病变的术前病理诊断具有较高的准确性,其诊断效能与 ROI 体积及高年资医生相当,能为临床诊疗提供一定依据。

参考文献:

[1] 刘春全,崔永.肺结节评估四大指南比较分析[J].中国肺癌杂志,2017,20(7):490-498.

[2] Nicholson AG,Tsao MS,Beasley MB,et al.The 2021 WHO classification of lung tumors:impact of advances since 2015[J].J Thorac Oncol,2022,17(3):362-387.

[3] Succony L,Rassl DM,Barker AP,et al.Adenocarcinoma spectrum lesions of the lung:detection, pathology and treatment strategies [J].Cancer Treat Rev,2021,99(9):102237.

[4] Yoshizawa A, Motoi N, Riely GJ, et al. Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases[J].Mod Pathol,2011,24(5):653-664.

[5] Woo T, Okudela K, Mitsui H, et al. Prognostic value of the IASLC/ATS/ERS classification of lung adenocarcinoma in stage I disease of Japanese cases[J].Pathol Int,2012,62(12):785-791.

- [6] Yanagawa N, Shiono S, Abiko M, et al. New IASLC/ATS/ERS classification and invasive tumor size are predictive of disease recurrence in stage I lung adenocarcinoma[J]. J Thorac Oncol, 2013, 8(5):612-618.
- [7] She Y, Zhang L, Zhu H, et al. The predictive value of CT-based radiomics in differentiating indolent from invasive lung adenocarcinoma in patients with pulmonary nodules[J]. Eur Radiol, 2018, 28(12):5121-5128.
- [8] Behera M, Owonikoko TK, Gal AA, et al. Lung adenocarcinoma staging using the 2011 IASLC/ATS/ERS classification: a pooled analysis of adenocarcinoma in situ and minimally invasive adenocarcinoma[J]. Clin Lung Cancer, 2016, 17(5):e57-e64.
- [9] Van Schil PE, Asamura H, Rusch VW, et al. Surgical implications of the new IASLC/ATS/ERS adenocarcinoma classification[J]. Eur Respir J, 2012, 39(2):478-486.
- [10] Dembitzer FR, Flores RM, Parides MK, et al. Impact of histologic subtyping on outcome in lobar vs sublobar resections for lung cancer: a pilot study[J]. Chest, 2014, 146(1):175-181.
- [11] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. Eur J Cancer, 2012, 48(4):441-446.
- [12] 吴珊珊, 沈桂权, 高波. 肺癌影像组学研究进展[J]. 中华放射学杂志, 2017, 51(12):986-989.
- [13] Zhao FH, Fan HJ, Shan KF, et al. Predictive efficacy of a radiomics random forest model for identifying pathological subtypes of lung adenocarcinoma presenting as ground-glass nodules[J]. Front Oncol, 2022, 12(12):872503.
- [14] Fan L, Fang M, Li Z, et al. Radiomics signature: a biomarker for the preoperative discrimination of lung invasive adenocarcinoma manifesting as a ground-glass nodule[J]. Eur Radiol, 2019, 29(2):889-897.
- [15] Son JY, Lee HY, Lee KS, et al. Quantitative CT analysis of pulmonary ground-glass opacity nodules for the distinction of invasive adenocarcinoma from pre-invasive or minimally invasive adenocarcinoma[J]. PloS one, 2014, 9(8):e104066.
- [16] Weng Q, Zhou L, Wang H, et al. A radiomics model for determining the invasiveness of solitary pulmonary nodules that manifest as part-solid nodules[J]. Clin Radiol, 2019, 74(12):933-943.
- [17] Chae HD, Park CM, Park SJ, et al. Computerized texture analysis of persistent part-solid ground-glass nodules: differentiation of preinvasive lesions from invasive pulmonary adenocarcinomas[J]. Radiology, 2014, 273(1):285-293.
- [18] 黄雪梅, 孙英丽, 高盼, 等. 分层分析影像组学模型在肺腺癌诊断中的价值[J]. 放射学实践, 2022, 37(2):191-199.
- [19] 张耀允, 余建群. 影像组学在磨玻璃结节型肺腺癌的应用研究进展[J]. 中国医疗设备, 2021, 36(2):167-171.
- [20] 胡红梅, 冯峰. 肺磨玻璃结节 CT 影像组学研究进展[J]. 放射学实践, 2020, 35(11):1472-1475.
- [21] Bankier AA, Macmahon H, Goo JM, et al. Recommendations for measuring pulmonary nodules at CT: a statement from the fleischner society[J]. Radiology, 2017, 285(2):584-600.
- [22] Sun X, Xu WC. Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves[J]. IEEE Signal Process Lett, 2014, 21(11):1389-1393.

(收稿日期:2023-04-13 修回日期:2023-09-20)

· 影像组学与人工智能专题 ·

钆塞酸二钠增强 MRI 肝胆期影像组学对乙肝患者肝储备功能的评估价值

胡景卉, 王芳, 侯承师, 黄京城, 杨鑫, 王文剑, 陈磊, 叶靖, 孙骏, 罗先富

【摘要】 目的:探讨钆塞酸二钠(Gd-EOB-DTPA)增强 MRI 肝胆期影像组学对乙型肝炎病毒感染患者肝储备功能的评估价值。**方法:**将 73 例慢性乙型肝炎肝功能正常者、136 例 Child-Pugh A 级和 70 例 Child-Pugh B/C 级乙型肝炎患者按照 8:2 随机分成训练组和测试组。从肝胆期 MRI 图像中勾画全肝实质进行影像组学特征提取。将得到的影像组学得分值(Rad-score)和多因素回归分析筛选出的临床变量联合,建立临床-影像组学联合模型。采用受试者工作特征(ROC)曲线分析评估效能。另一时间段 64 例(17 例肝功能正常、30 例 Child-Pugh A 级和 17 例 Child-Pugh B/C 级)乙肝患者用来验证。**结果:**多因素回归分析筛选出血小板、总胆红素和凝血酶原时间国际化比值用以建立鉴别肝功能正常与 Child-Pugh A 级的临床模型,利用 4 个最相关的影像组学特征构建影像组学模型,训练组、测试组和验证组中临床模型鉴别肝功能正常与 Child-Pugh A 级的曲线下面积(AUC)分别为 0.897、0.884、0.780,组学模型的 AUC 分别为 0.890、0.914、0.824,临床-影像组学联合模型的 AUC 分别为 0.951、0.978、0.886。联合模型与临床模型、组学模型相比,净重新分类改善指数(NRI)和综合判别改善指数(IDI)的检验值均 <0.05 。血小板和总胆红素被筛选出构建鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的临床模型,7 个最相关的影像组学特征构建影像组学模型,训练组、测试组和验证组中影像组学模型鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的 AUC 分别为 0.862、0.865、0.818,联合模型(AUC=0.940、0.934、0.951)与临床模型(AUC=0.916、0.893、0.914)的评估效能差异无统计学意义($P>0.05$)。**结论:**基于 Gd-EOB-DTPA 增强 MRI 肝胆期影像组学对乙肝患者的肝储备功能具有良好的评估价值,联合临床变量后可以提高 Child-Pugh A 级的早期检出效能,但对于 Child-Pugh A 级与 Child-Pugh B/C 级的评估,单独的临床模型已经表现出了良好的效能。

【关键词】 钆塞酸二钠; 乙型肝炎病毒; 肝储备功能; 磁共振成像; 影像组学; 列线图

【中图分类号】 R512.6; R445.2 **【文献标志码】** A **【文章编号】** 1000-0313(2024)01-0022-08

DOI:10.13609/j.cnki.1000-0313.2024.01.005

开放科学(资源服务)标识码(OSID):



Radiomics based on Gd-EOB-DTPA enhanced MR hepatobiliary phase images for evaluating liver reserve function of HBV patients HU Jing-hui, WANG Fang, HOU Cheng-shi, et al. The Clinical Medicine College of Yangzhou University, Jiangsu 225001, China

【Abstract】 Objective: To investigate the value of radiomics based on the hepatobiliary phase images of Gd-EOB-DTPA enhanced liver MR in evaluating liver reserve function of hepatitis B virus (HBV) patients. **Methods:** 73 chronic HBV patients with normal liver function, 136 HBV patients with Child-Pugh A liver function and 70 patients with Child-Pugh B/C liver function were randomly divided into the training group and the test group in the ratio of 8:2. Radiomics features were extracted from the hepatobiliary phase MR images via delineating the whole liver parenchyma. The clinic-radiomics model combining the radiomics score (Rad-score) and clinical variables selected by multivariate regression analysis was built. Evaluation performances of the model were analyzed using receiver operating characteristic (ROC) curve. 64 HBV patients were used for validation from another period, including 17 cases with normal liver function, 30 with Child-Pugh A and 17 with Child-Pugh B/C. **Results:**

作者单位: 225001 江苏,扬州大学医学院(胡景卉、王文剑); 225001 江苏,苏北人民医院医学影像科(胡景卉、侯承师、黄京城、杨鑫、王文剑、叶靖、孙骏、罗先富); 200232 上海,上海联影智能有限公司研发部(王芳、陈磊); 116031 辽宁,大连医科大学临床医学院(侯承师、黄京城、杨鑫)

作者简介: 胡景卉(1994—),女,安徽滁州人,硕士研究生,主要从事腹盆腔疾病影像诊断工作。

通讯作者: 罗先富, E-mail: xianfu-luo@hotmail.com

基金项目: 江苏省“333 高层次人才培养工程”(2022-3-6-139)

Platelet (PLT), total bilirubin (TBIL) and international normalized ratio of prothrombin time (INR) were selected by multivariate regression analysis to construct clinical model for distinguishing normal liver function from Child-Pugh A. The radiomics model was constructed using the four most relevant radiomics features. The areas under the curve (AUCs) of clinical model for distinguishing normal liver function from Child-Pugh A were 0.897, 0.884 and 0.780 in the training, test and validation cohorts, respectively. The AUCs of radiomics model were 0.890, 0.914 and 0.824 in the training, test and validation cohorts, respectively. The AUCs of clinic-radiomics combined model were 0.951, 0.978 and 0.886 in the training, test and validation cohorts, respectively, whose net reclassification index (NRI) and integrated discrimination improvement (IDI) were significantly different from the other two models ($P < 0.05$). PLT and TBIL were selected to construct clinical model for distinguishing Child-Pugh A from Child-Pugh B/C. The radiomics model was constructed using the seven most relevant radiomics features. The AUCs of radiomics model for distinguishing Child-Pugh A from Child-Pugh B/C were 0.862, 0.865 and 0.818 in the training, test and validation cohorts, respectively. There was no significant difference between the clinic-radiomics combined model (AUC=0.940, 0.934 and 0.951, respectively) and the clinical model (AUC=0.916, 0.893 and 0.914, respectively) ($P > 0.05$). **Conclusions:** Radiomics based on the hepatobiliary phase images of Gd-EOB-DTPA enhanced MR has a high evaluation value for liver reserve function of HBV patients, which can improve the early detection performance of Child-Pugh A when combined with clinical variables. However, the single clinical model has shown good evaluation efficacy for Child-Pugh A vs. Child-Pugh B/C.

【Key words】 Gd-EOB-DTPA; Hepatitis B virus; Liver reserve function; Magnetic resonance imaging; Radiomics; Nomogram

慢性乙肝是亚洲地区导致肝硬化的最常见原因,在早期阶段,临床表现常常较为隐蔽,当症状出现时,往往伴有严重的并发症。全球每年约有 200 万人死于肝病,其中超过 100 万人死于肝硬化并发症^[1]。随着肝炎的进展,肝细胞出现变性和坏死、纤维结缔组织增生和假小叶再生,导致正常肝细胞数量减少,影响肝脏的合成、分泌及代谢功能,从而导致肝储备功能的下降,临床上表现为总胆红素和转氨酶的升高、凝血障碍、脾肿大等。对慢性肝病患者进行早期的评估、管理和治疗至关重要。

目前,临床上主要通过血液生化指标来评估肝储备功能,但这些指标不仅不能显示肝脏的解剖结构或提供血液灌注信息,并且容易受急性感染等因素的影响。由于肉眼可见的结构变化往往在慢性疾病的发生过程中出现较晚,其他影像学检查如超声、CT 和常规 MRI 对于早期慢性肝病的检测能力有限^[2]。钆塞酸二钠 (gadolinium-ethoxybenzyl-diethylenetriamine-pentaacetic acid, Gd-EOB-DTPA) 作为一种肝细胞特异性对比剂,可在肝胆期被正常肝细胞通过肝细胞膜表面的有机阴离子转运系统阴离子转运多肽 (organic anion transporting polypeptide, OATP) 特异性摄取,并且肝胆期图像的多个测量参数已被证实可以用来评估肝功能,如肝实质相对强化程度、肝胆期的肝门静脉比等^[3-5]。然而,通过信号强度测量计算得出的参数来

评估肝功能是有限的,因为它们仅仅通过勾画肝脏几个层面的局部感兴趣区,这种方法并不能评估完整的肝实质储备功能。

影像组学是近几年新兴的影像研究领域,它可以通过勾画整个病变区域的体积,再通过机器学习分析大量的高维数据,提取出重要特征并定量表示,从而提高预测、诊断和预后的准确性^[6],这对于早期评估肝功能具有重要价值。周玮等^[7]仅从肝胆期影像组学方面评估了 Child-Pugh A 与 B/C 级肝硬化患者的肝储备功能,未联合临床进行分析是否临床已经具有较高的评估价值。张智星等^[8]采用了 3 期增强 CT 图像对 Child-Pugh 进行分级评估。本研究旨在探讨是否能够单独基于 Gd-EOB-DTPA 增强 MRI 肝胆期图像来建立临床-影像组学联合模型,对慢性乙肝患者肝功能进行更好的分级评估。

材料与方法

1. 病例资料

回顾性搜集 2018 年 12 月至 2021 年 7 月就诊于苏北人民医院行 Gd-EOB-DTPA 增强 MRI 检查的 1205 例存在慢性乙型肝炎感染的患者。病例纳入标准:①实验室证据表明存在乙型肝炎病毒感染;②具有完整的临床数据资料,包括是否有腹水、肝性脑病、机体白蛋白水平以及 MRI 检查前后 1 周内的总胆红素

(total bilirubin, TBIL)、丙氨酸氨基转移酶(alanine aminotransferase, ALT)、天冬氨酸氨基转移酶(aspartate transaminase, AST)、血小板(platelet, PLT)、凝血酶原时间(prothrombin time, PT)、凝血酶原时间国际化比值(international normalized ratio, INR);③肾功能正常。病例排除标准:①肝脏存在较大的病灶(最大直径 >5 cm)或多个病灶(数量 >5),导致正常肝实质范围较小影响测量准确性;②有肝介入治疗史或肝切除手术史;③门静脉存在栓子,影响血流动力学者^[9];④合并有其他肝炎病毒感染;⑤肝胆期图像质量不佳。根据收集的临床及实验室指标评估出 Child-Pugh 分级,包括总胆红素、白蛋白水平、凝血酶原延长时间、腹水量级以及是否有肝性脑病^[10]。采用相同的纳入和排除标准,从 2022 年 7 月至 12 月就诊的慢性乙型肝炎感染患者中筛选病例。本研究为回顾性临床研究,经苏北人民医院伦理委员会批准(批准文号:2021ky219),免除受试者知情同意。

2. MRI 检查方法

MRI 检查均采用 3.0T MRI 扫描仪(GE Discovery 750, 750W, 美国),8 通道体部相控阵线圈。扫描前嘱患者禁食 6~8 h,采用呼吸门控技术,对患者进行平静呼吸及屏气训练,扫描范围从膈顶至双肾下缘。所有患者均进行 MRI 平扫及钆塞酸二钠增强扫描。 T_1 WI 平扫及增强采用肝脏容积加速采集(liver acquisition with volume acceleration, LAVA)序列,分别于注射对比剂后 20 s(动脉期)、60 s(门静脉期)、3 min(移行期)及 20 min(肝胆特异期)进行扫描得到 4 期图像。LAVA 序列扫描参数:重复时间 3.7 ms,回波时间 1.7 ms,层厚 5.0 mm,层间距 2.5 mm,矩阵 320×224 ,视野 $40 \text{ cm} \times 40 \text{ cm}$,翻转角 18° 。对比剂采用钆塞酸二钠(Gd-EOB-DTPA, Primovist, 德国拜耳医药保健有限公司),注射流率 2.0 mL/s(剂量 0.025 mmol/kg),注射对比剂后再以相同流率注射 20 mL 0.9%氯化钠溶液进行冲洗。

3. 研究方法

临床模型构建:将临床变量(包括年龄、性别、身体质量指数、TBIL、ALT、AST、PLT、PT、INR)进行单因素和多因素逻辑回归分析,筛选出与 Child-Pugh 分

级相关的临床变量,即独立危险因素,将其联合构建出临床评估模型。采用受试者工作特征(receiver operating characteristic, ROC) 曲线的曲线下面积(area under the curve, AUC) 分析模型的评估能力。

肝半自动分割:由一位具有 5 年工作经验的放射科医师将所有纳入研究患者的肝胆期图像导入“uAI 科研平台”(uRP, United Imaging Healthcare Co, 中国上海)。一个名为 V-Net 的器官分割深度学习模型用于自动分割肝脏的感兴趣体积(volume of interest, VOI)^[11],再进一步进行手动修正,以排除每个层面上的脂肪、空气、肝脏病变、大血管和胆管区域(图 1、2)。慢性肝炎肝功能正常、Child-Pugh A 级和 Child-Pugh B/C 级的患者分别标记为 0、1 和 2。然后,由另一位具有 10 年工作经验的放射科医生检查核对 VOI 以及标记的标签。

影像组学特征提取与选择:使用 Python 3.7 中的 Pyradiomics 工具箱,从 Gd-EOB-DTPA 肝脏增强 MRI 肝胆期图像中提取符合国际生物标记物标准化倡议的 2600 个肝实质区域的影像特征。这些特征包括了 4 个特征组:18 个一级特征、14 个体积和形状特征、72 个纹理特征和 2496 个基于滤波器的特征。采用随机抽样方法将数据集按 8:2 的比例分成训练组和测试组,另一时间段的数据集构建验证组。为了避免

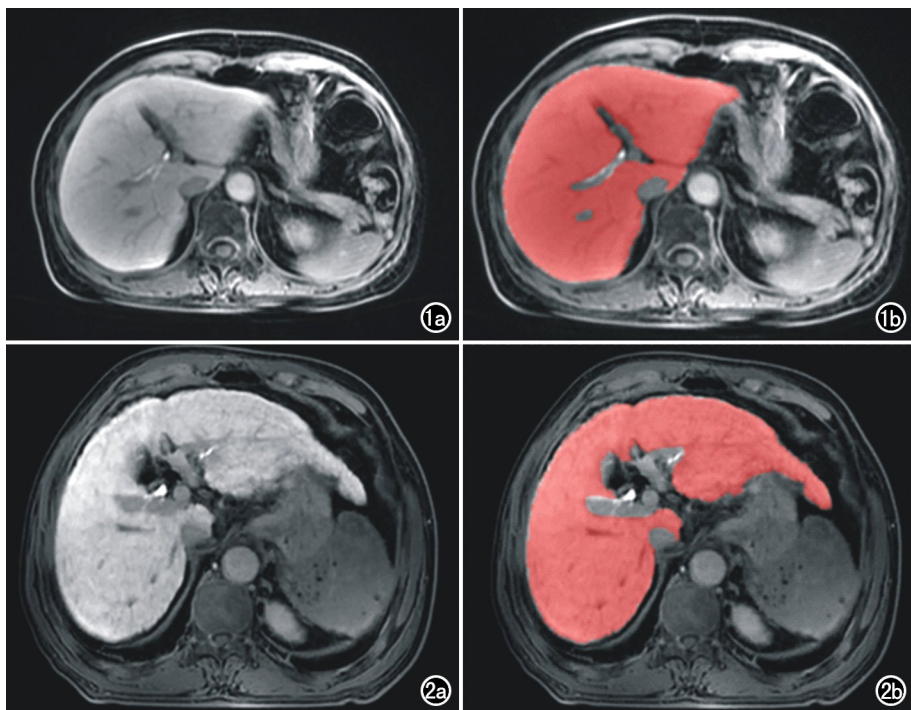


图 1 慢性乙肝肝功能正常者的 Gd-EOB-DTPA 增强 MRI 肝胆期图像, Rad-score 值为 0.565。a) 单个层面原图;b) 该层面勾画示例图。

图 2 Child-Pugh A 级患者的 Gd-EOB-DTPA 增强 MRI 肝胆期图像, Rad-score 值为 0.809。a) 单个层面原图,肝包膜稍欠光整;b) 该层面勾画示例图。

特征之间的维度影响,使用 Z 分数预处理方法进行特征归一化。应用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO) 算法进行降维处理,筛选出最相关的系数非 0 的影像组学特征。将得到的组学特征进行线性组合计算,得出每例患者对应的影像组学得分(Rad-score)。

影像组学模型及临床-组学联合模型构建:AUC 用于量化 Rad-score 的评估能力,在训练组、测试组和验证组中,分别建立 Rad-score 影像组学评估模型。将临床模型中的独立危险因素和 Rad-score 联合进行多因素逻辑回归分析,构建出临床-影像组学联合评估模型,用于对慢性肝炎患者进行肝储备功能评估。并根据训练组的约登指数确定临界值,以得到相应的敏感度、特异度、准确度以及阳性预测值、阴性预测值。净重新分类改善指数(net reclassification index, NRI)、综合判别改善指数(integrated discrimination improvement, IDI)用以评价联合模型较其他模型的检出效能改善情况。

列线图的建立和评估:利用训练组建立并绘制整合了临床变量和影像组学 Rad-score 的组学列线图。绘制列线图的校准曲线,通过 Hosmer-Lemeshow 检验来评估拟合优度。决策曲线分析(decision curve analysis, DCA)用于评估不同阈值概率下列线图的净收益。

4.统计学分析

采用 SPSS 24.0 软件和 R 语言(version 3.3.3, Vienna, Austria)进行统计学分析。根据正态性检验结果,将定量数据以中位数(四分位数间距,不符合正态

分布)或平均值±标准差(符合正态分布)表示,定性数据以百分数(%)表示。定量数据采用 Mann-Whitney U 检验或 Student's t 检验进行组间比较,定性数据采用卡方检验进行组间比较。以 $P<0.05$ 为差异有统计学意义。

结 果

1.病例资料

1205 例慢性乙肝患者经纳入、排除标准,排除临床数据资料不完整者 42 例,肝脏存在较大病灶(最大直径>5 cm)者 287 例,多个病灶(数量>5)者 336 例,有肝脏介入治疗史或肝切除手术史者 114 例,门静脉存在栓子者 56 例,合并丙肝者 55 例,戊肝者 17 例,肝胆期图像质量不佳者 19 例。最终 279 例患者纳入本研究,其中男 171 例,女 108 例,年龄中位数为 60 岁,按照临床 Child-Pugh 分级标准将 279 例患者分为 73 例慢性乙型肝炎肝功能正常者、136 例 Child-Pugh A 级患者和 70 例 Child-Pugh B/C 级患者。肝功能正常和 Child-Pugh A 级的 209 例患者被随机分配到训练组($n=167$)和测试组($n=42$)中,Child-Pugh A 和 Child-Pugh B/C 级的 206 例被随机分配到训练组($n=164$)和测试组($n=42$)中。患者的临床资料见表 1。另一时间段筛选出的 64 例患者用于验证,包括 17 例慢性乙肝肝功能正常者、30 例 Child-Pugh A 级和 17 例 Child-Pugh B/C 级乙肝患者。

2.临床模型

单因素及多因素逻辑回归分析结果显示,PLT、TBIL 和 INR 是乙肝 Child-Pugh A 级的独立危险因

表 1 患者基线资料 (n,%)

指标	肝功能正常和 Child-Pugh A 级患者				Child-Pugh A 和 Child-Pugh B/C 级患者			
	训练组 (n=167)	测试组 (n=42)	统计量	P 值	训练组 (n=164)	测试组 (n=42)	统计量	P 值
年龄(岁) ^a	60.00(18.00)	59.50(19.00)	-0.221 ^d	0.825	57.50(18.50)	62.50(14.75)	-1.669 ^d	0.095
性别(%) ^b			0.097 ^e	0.755			0.104 ^e	0.748
男性	103(61.68)	27(64.29)			101(61.59)	27(64.29)		
女性	64(38.32)	15(35.71)			63(38.41)	15(35.71)		
身高(cm) ^a	165.00(10.00)	165.00(12.00)	-1.069 ^d	0.285	165.00(10.00)	168.00(10.00)	-0.687 ^d	0.492
体重(kg) ^a	65.00(12.00)	65.00(10.50)	-1.767 ^d	0.077	65.00(11.75)	65.00(10.25)	-0.553 ^d	0.580
BMI(kg/m ²) ^c	24.11±3.25	23.61± 3.10	0.897 ^f	0.371	24.22±2.99	23.67±3.81	1.027 ^f	0.306
TBIL(μmol/L) ^a	13.80(9.10)	13.50(8.38)	-0.508 ^d	0.611	20.15(17.88)	17.60(11.73)	-1.559 ^d	0.119
ALT(U/L) ^a	24.00(19.00)	27.50(25.78)	-0.545 ^d	0.586	33.00(31.60)	28.45(24.25)	-2.131 ^d	0.033
AST(U/L) ^a	27.00(17.00)	28.00(19.00)	-0.258 ^d	0.796	38.00(37.75)	35.50(30.75)	-0.775 ^d	0.438
PLT(10 ⁹ /L) ^a	141.00(102.00)	113.50(55.00)	-2.956 ^d	0.003	91.00(86.25)	102.50(54.25)	-1.371 ^d	0.170
PT(s) ^c	13.20(1.70)	12.95(1.85)	-0.251 ^d	0.802	14.15(2.85)	14.10(2.43)	-0.680 ^d	0.496
INR ^a	1.05(0.14)	1.05(0.12)	-0.247 ^d	0.805	1.13(0.25)	1.16(0.14)	-0.791 ^d	0.429
Cr(μmol/L) ^a	74.20(14.00)	76.00(16.50)	-1.402 ^d	0.161	75.00(14.00)	76.50(22.25)	-0.892 ^d	0.372
Child-Pugh 分级 ^b			0.014 ^e	0.905			0.010 ^e	0.921
肝功能正常	58(34.73)	15(35.71)			—	—		
Child-Pugh A	109(65.27)	27(64.29)			108(65.85)	28(66.67)		
Child-Pugh B/C	—	—			56(34.15)	14(33.33)		

注: BMI 身体质量指数, TBIL 总胆红素, ALT 丙氨酸氨基转移酶, AST 天冬氨酸氨基转移酶, PLT 血小板, PT 凝血酶原时间, INR 凝血酶原时间国际标准化比值, Cr 肌酐。 ^a中位数(四分位间距); ^b百分数(%); ^c平均值±标准差; ^dZ 值; ^e卡方值; ^ft 值。

素,PLT、TBIL 是乙肝 Child-Pugh B/C 级的独立危险因素(表 2、3),分别进行联合,构建出临床评估模型,训练组、测试组和验证组中临床模型鉴别肝功能正常与 Child-Pugh A 级的 AUC 分别为 0.897、0.884、0.780,训练组、测试组和验证组中临床模型鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的 AUC 分别为 0.916、0.893、0.914。

3.影像组学特征及影像组学模型的评估效能

在训练组中,慢性肝炎肝功能正常者和 Child-Pugh A 级患者的数量分别为 58 例和 109 例。通过 LASSO 回归分析后,选择出 4 个相关特征,包括 2 个灰度依赖性矩阵(gray-level dependence matrix, GLDM)、1 个灰度大小区域矩阵(gray-level size zone matrix, GLSZM)和 1 个灰度行程矩阵(gray-level run length matrix, GLRLM)。所有特征的定量值在两组中均具有统计学差异($P<0.05$)。根据以上 4 个特征及其系数、截距值计算得到影像组学标签 Rad-score (图 3):

$$\text{Rad-score}_{\text{肝功能正常 vs. Child-Pugh A}} = 0.038 \times \text{normalize_glrlm_ShortRunHighGrayLevelEmphasis} + 0.026 \times \text{normalize_gldm_SmallDependenceEmphasis} - 0.005 \times \log_gldm_log\text{-sigma-1-5mm-3D-LargeDependenceHighGrayLevelEmphasis} - 0.083 \times \text{wavelet_glszm_wavelet-HHH-ZoneEntropy} + 0.651$$

在训练组、测试组和验证组中,Rad-score 鉴别肝

功能正常与 Child-Pugh A 级的 AUC 分别为 0.890、0.914、0.824(表 4)。

Child-Pugh A 级与 Child-Pugh B/C 级患者通过 LASSO 回归分析后,选择出 7 个相关特征,包括 4 个一阶特征(first order)、1 个灰度共生矩阵(gray-level cooccurrence matrix, GLCM)和 2 个 GLRLM。所有特征的定量值在两组中均具有统计学差异($P<0.05$)。根据以上 7 个特征及其系数、截距值计算得到影像组学标签 Rad-score:

$$\text{Radscore}_{\text{Child-Pugh A vs. B/C}} = 0.067 \times \text{normalize_glrlm_ShortRunHighGrayLevelEmphasis} + 0.028 \times \text{wavelet_firstorder_wavelet-LLH-Kurtosis} + 0.003 \times \text{normalize_glrlm_RunLength-NonUniformityNormalized} - 0.006 \times \text{wavelet_glcm_wavelet-HHH-InverseVariance} - 0.010 \times \text{normalize_firstorder_Maximum} - 0.011 \times \text{normalize_firstorder_TotalEnergy} - 0.074 \times \text{normalize_firstorder_RootMeanSquared} + 0.340$$

在训练组、测试组和验证组中,Rad-score 鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的 AUC 分别为 0.862、0.865、0.818。

4.临床-影像组学联合模型的评估效能

对于肝功能正常与 Child-Pugh A 级患者,将 PLT、TBIL、INR 与 Rad-score 进行多因素回归分析后,所有参数均纳入临床-影像组学联合模型(表 3)。训练组中,联合模型鉴别肝功能正常与 Child-Pugh A

表 2 临床参数的单因素和多因素回归分析结果

变量	肝功能正常和 Child-Pugh A 级				Child-Pugh A 和 Child-Pugh B/C 级患者			
	单因素回归分析		多因素回归分析		单因素回归分析		多因素回归分析	
	OR (95%CI)	P 值	OR (95%CI)	P 值	OR (95%CI)	P 值	OR (95%CI)	P 值
年龄	0.987(0.948~1.026)	0.503			1.015(0.966~1.067)	0.558		
性别	0.943(0.298~2.989)	0.921			1.459(0.451~4.715)	0.528		
BMI	1.084(0.906~1.296)	0.379			0.837(0.693~1.010)	0.063		
TBIL	1.119(1.019~1.229)	0.019	1.149(1.051~1.256)	0.002	1.116(1.043~1.194)	0.001	1.140(1.082~1.200)	<0.001
ALT	1.059(0.998~1.124)	0.060	1.001(0.983~1.019)	0.918				
AST	1.041(0.979~1.108)	0.200	1.004(0.984~1.024)	0.678				
PLT	0.990(0.982~0.998)	0.018	0.992(0.986~0.998)	0.012	0.986(0.975~0.999)	0.028	0.983(0.973~0.994)	0.002
PT	0.976(0.550~1.731)	0.934			1.695(0.993~2.895)	0.053		
INR	6.361(2.474~16.352)	<0.001	5.418(2.718~10.802)	<0.001	1.622(0.913~2.880)	0.099		

表 3 临床和影像组学参数的多因素回归分析结果

变量	肝功能正常与 Child-Pugh A 级		Child-Pugh A 与 Child-Pugh B/C 级患者	
	OR (95%CI)	P 值	OR (95%CI)	P 值
TBIL	1.140(1.039~1.250)	0.005	1.104(1.043~1.169)	0.001
PLT	0.990(0.982~0.998)	0.012	0.982(0.971~0.993)	0.001
INR	4.290(1.904~9.666)	<0.001	—	—
Rad-score	6.101(2.915~12.768)	<0.001	2.188(1.420~3.373)	<0.001

表 4 肝功能正常与 Child-Pugh A 级组的临床模型、影像组学模型和临床-组学联合模型评估效能

指标	临床模型			影像组学模型			联合模型		
	训练组	测试组	验证组	训练组	测试组	验证组	训练组	测试组	验证组
AUC (95%CI)	0.897 (0.849~0.945)	0.884 (0.787~0.981)	0.780 (0.649~0.911)	0.890 (0.838~0.942)	0.914 (0.824~1.000)	0.824 (0.707~0.940)	0.951 (0.920~0.982)	0.978 (0.944~1.000)	0.886 (0.789~0.984)
敏感度(%)	89.0	77.8	86.7	76.1	81.5	63.3	91.7	88.9	90.0
特异度(%)	79.3	73.3	47.1	87.9	80.0	76.5	89.7	86.7	76.5
准确度(%)	85.6	76.2	72.3	80.2	81.0	68.1	91.0	88.1	85.1
阳性预测值(%)	89.0	84.0	74.3	92.2	88.0	82.6	94.3	92.3	87.1
阴性预测值(%)	79.3	64.7	66.7	66.2	70.6	54.2	85.2	81.3	81.3

级的 AUC 为 0.951, 较临床模型 ($\text{NRI}=1.022, P<0.001$; $\text{IDI}=0.172, P<0.001$) 和影像组学模型 ($\text{NRI}=1.217, P<0.001$; $\text{IDI}=0.197, P<0.001$), 显著提高了评估效能。测试组中, 联合模型鉴别肝功能正常与 Child-Pugh A 级的 AUC 为 0.978, 同样显著提高了评估效能 (与临床模型比较: $\text{NRI}=1.452, P<0.001$; $\text{IDI}=0.331, P=0.001$ 。与组学模型比较: $\text{NRI}=1.644, P<0.001$; $\text{IDI}=0.224, P=0.002$)。验证组中, 联合模型鉴别肝功能正常与 Child-Pugh A 级的 AUC 为 0.886, 较临床模型 ($\text{NRI}=0.302, P=0.007$; $\text{IDI}=0.170, P=0.002$) 和组学模型 ($\text{NRI}=0.639, P<0.001$; $\text{IDI}=0.311, P<0.001$) 也显著提高了评估效能 (表 4、图 4)。

对于 Child-Pugh A 与 Child-Pugh B/C 级患者, 将 PLT、TBIL 与 Rad-score 进行多因素回归分析后, 所有参数均纳入临床-影像组学联合模型 (表 3)。在训练组、测试组和验证组中, 联合模型鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的 AUC 分别为 0.940、0.934、0.951, 相较于临床模型 ($\text{AUC}=0.916、0.893、0.914$) 并未有明显提高, 差异无统计学意义 (Delong 检验, P 均 >0.05)。

5. 列线图的建立和评估

对于肝功能正常与 Child-Pugh A 组, 临床-影像组学联合模型的列线图如图 5 所示, Hosmer-Lemeshow 拟合优度检验的 P 值无统计学意义 (训练组 $P=0.443$, 测试组 $P=1.000$, 验证组 $P=0.343$, 图

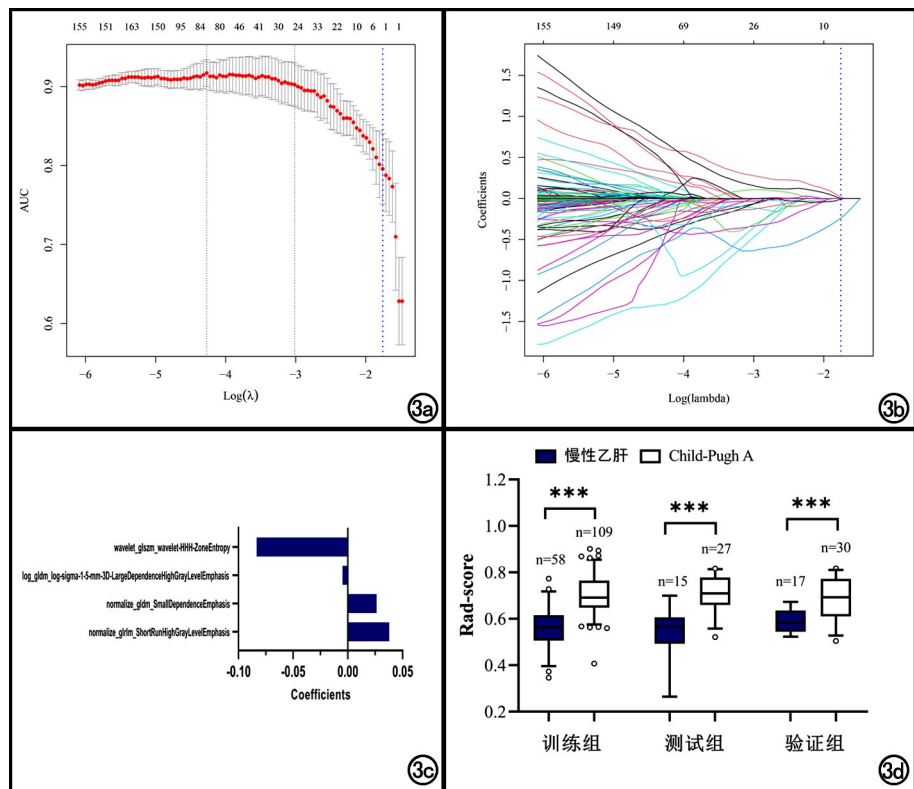


图 3 肝功能正常与 Child-Pugh A 级组采用 LASSO 算法进行影像组学特征筛选。a) 调整参数 λ 使拟合损失值二项偏差最小, 以筛选出最优的影像组学特征; b) 筛选出最优影像组学特征的系数收敛图, 在最佳 λ 值处画一条垂直线, 得到 4 个系数非零的特征; c) 选定的 4 个最优特征及其相关系数; d) 训练组、测试组和验证组中每例患者的影像组学得分, 在鉴别慢性乙肝与 Child-Pugh A 级患者中差异有统计学意义 ($P<0.05$)。

6), 表明列线图具有良好的校准性能。在 DCA 中的所有阈值概率下, 临床-影像组学联合模型的列线图表现出优于其他模型的临床净收益 (图 7)。

讨论

本研究从肝脏 Gd-EOB-DTPA 增强 MRI 肝胆期图像中提取出与慢性乙型肝炎肝储备功能最相关的影像组学特征, 并通过计算得到每例患者对应的影像组学标签即 Rad-score, 结果显示影像组学模型具有良好

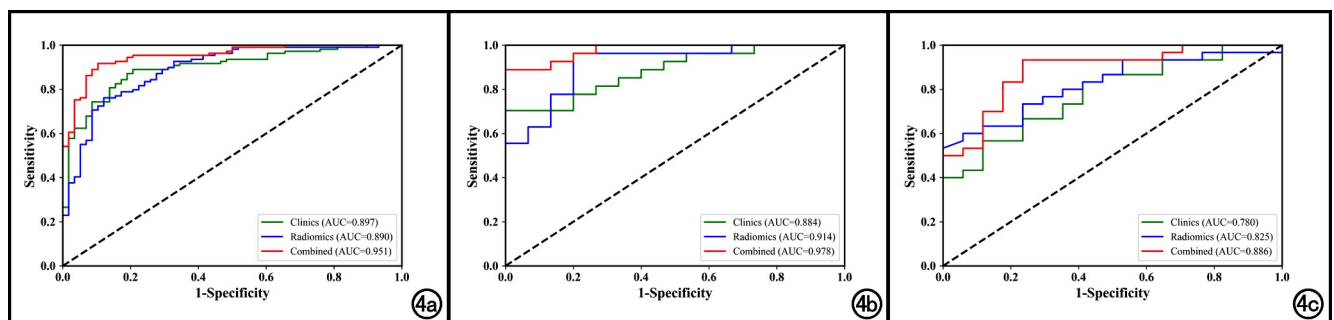


图 4 肝功能正常与 Child-Pugh A 级组中临床模型、影像组学模型和临床-组学联合模型的 ROC 曲线。a) 训练组; b) 测试组; c) 验证组。

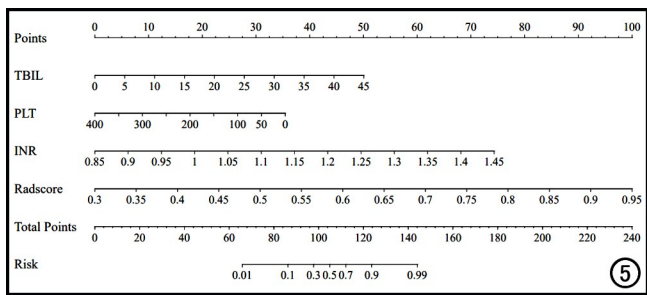


图5 肝功能正常与 Child-Pugh A 级组联合模型的列线图。

的评估效能。肝功能正常与 Child-Pugh A 级组在联合了影像组学和临床模型后,联合模型对于评估效能有所提高,优于单纯的临床模型和影像组学模型。但在 Child-Pugh A 级与 Child-Pugh B/C 级组中,临床模型已经表现出了较好的评估价值。

经过单因素和多因素回归分析后,肝功能正常与 Child-Pugh A 级组筛选出了 3 个临床参数纳入临床模型,分别是 TBIL、PLT 和 INR。Child-Pugh A 级与 Child-Pugh B/C 级组则筛选出 TBIL 和 PLT 建立临床模型。TBIL 是肝功能中的一个重要指标,肝炎进展过程中,肝功能下降,机体对于胆红素的代谢能力也下降,导致其在血液中积聚、增高。相关研究表明 PLT 与肝脏的再生功能密切相关,肝功能差的患者往往由于正常肝细胞减少,PLT 数量也相应下降,导致脾肿大、脾功能亢进,PLT 单因素对于肝功能的预测

就具有一定意义^[12,13]。同样,肝功能的下降,凝血酶原相对也会缺乏,凝血因子的合成发生障碍,INR 也是一种反映肝脏合成能力的重要指标^[14,15]。两组的临床模型测试组 AUC 分别达到了 0.884、0.893,验证组为 0.780、0.914,具有良好的评估效能。

本研究中肝功能正常与 Child-Pugh A 级组提取出了 4 个最相关的肝胆期影像组学纹理特征,包括 2 个 GLDM 特征、1 个 GLRLM 特征和 1 个 GLSZM 特征。GLSZM 纹理特征可以用来描述灰度值的分布,来自于 GLSZM 类的区域熵特征具有最高的绝对值相关性系数,可以度量灰度不均匀性或随机性,值越高表示肝实质组织的异质性越高^[16]。其他包括 GLDM 纹理特征可以量化肝胆期 MRI 图像的灰度依赖性,GLRLM 纹理特征则提供了具有相同灰度的连续像素运行的空间分布信息,它们都可以反映出一阶纹理特征无法显示的空间异质性改变^[17]。Child-Pugh A 级与 Child-Pugh B/C 级组提取出 7 个最相关的肝胆期影像组学纹理特征,包括 4 个一阶特征、1 个 GLCM 和 2 个 GLRLM 特征。一阶直方图特征描述了图像区域内体素强度的分布,晚期肝硬化有大小不一的结节和纤维间隔,导致体素强度分布不均;GLCM 则可以捕获具有预定义灰度强度的像素对或体素之间的空间关系。病理上,随着肝炎的进展,肝细胞变性坏死导致正常的肝小叶被破坏,假小叶内再生的肝细胞不具备正常肝细胞的功能,从而对于 Gd-EOB-DTPA 的摄取减少,肝胆期图像信号强度降低,肝脏组织的异质性也随

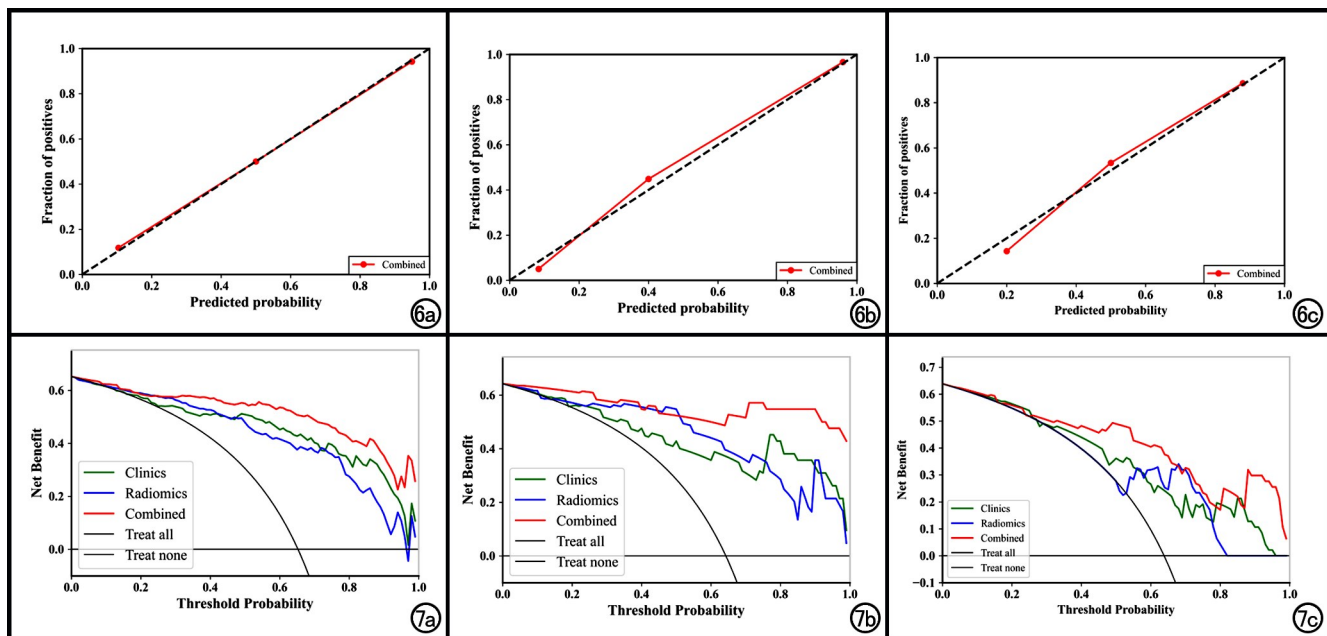


图6 校准曲线表明肝功能正常与 Child-Pugh A 级组中训练组、测试组和验证组的列线图校准良好,Hosmer-Lemeshow 拟合优度检验的 P 值均 >0.05 。a)训练组;b)测试组;c)验证组。图7 决策曲线分析结果表明,与临床模型和影像组学模型相比,肝功能正常与 Child-Pugh A 级组中训练组、测试组和验证组中联合模型列线图在各概率阈值下均具有更高的临床净效益。a)训练组;b)测试组;c)验证组。

着肝脏硬度的增加而增加,逐渐演变形成肝硬化^[18-20]。影像组学在肝胆期图像上进行全肝组织的体积分割,准确捕捉细微的组织变化,提取出高维的成像特征,通过筛选出的这些纹理特征计算得到 Rad-score 组学模型,也具有较高的评估效能,测试组分别为 0.914、0.865,验证组分别为 0.824、0.818。在联合了影像组学和临床模型后,诊断效能得到了显著提高,并且在两组的多因素回归分析中,Rad-score 都具有最高的权重系数(OR=6.101、2.188),这充分说明了影像组学在肝储备功能评估中的重要作用。肝功能正常与 Child-Pugh A 级组的联合模型 AUC 值,训练组为 0.951,测试组为 0.978,验证组为 0.886,并且 NRI 和 IDI 表明均改善了模型的评估能力,但对于 Child-Pugh A 级与 Child-Pugh B/C 级组,联合模型相对于临床模型并没有明显优势。

既往多项研究表明,影像组学可以更好地评估肝脏功能储备。相关研究表明通过评估术前肝功能,可以预测肝癌患者术后肝功能衰竭^[21,22];还可以用来预测、识别慢性肝病患者是否发生晚期肝纤维化以及准确分期患者的显著纤维化($\geq F2$)水平^[23,24]。然而,肝纤维化程度并不能直接反映整体的肝功能。Nitsch 等^[25]曾利用基于肝脏和脾脏的 MRI 影像特征开发了一种肝硬化疾病严重程度的预测模型,但它们仅限于临床失代偿的患者。本研究更全面地评估了乙肝患者的肝储备功能,为临床迅速采取干预措施提供了有价值的辅助信息。

本研究存在以下局限性:首先,本研究仅针对乙肝患者,然而在西方国家,丙肝和酒精性脂肪性肝炎也是肝病的常见原因,我们将在未来扩大纳入标准;其次,可能是由于选择偏倚或样本量较小,临床基线资料中肝功能正常与 Child-Pugh A 级组中的 PLT 和 Child-Pugh A 级与 Child-Pugh B/C 级组中的 ALT 在训练组和测试组中差异无统计学意义,并且由于 Child-Pugh C 级的病例数较少而和 B 级患者一起进行分析,未来将扩大样本量进行更全面研究。

综上所述,Gd-EOB-DTPA 增强 MRI 肝胆期影像组学模型可以对慢性乙肝患者进行较好的肝功能分级,临床-影像组学联合模型在乙肝患者肝功能损伤早期 Child-Pugh A 级的检出方面表现出了更好的性能。未来,采用基于影像组学的模型来评估肝功能,可能可以优化对于乙肝患者的管理及早期发现,更好地服务于精准医疗。

参考文献:

- [1] Asrani SK,Devarbhavi H,Eaton J,et al.Burden of liver diseases in the world[J].J Hepatol,2019,70(1):151-171.
- [2] Van Beers BE,Daire JL,Garteiser P.New imaging techniques for

- liver diseases[J].J Hepatol,2015,62(3):690-700.
- [3] Poetter-Lang S,Bastati N,Messner A,et al.Quantification of liver function using gadoxetic acid-enhanced MRI[J].Abdom Radiol (NY),2020,45(11):3532-3544.
- [4] 毕新军,张学琴,陆健,等.钆塞酸二钠增强 MRI 肝细胞摄取定量参数评估肝功能[J].放射学实践,2021,36(4):514-519.
- [5] Yang M,Zhang Y,Zhao W,et al.Evaluation of liver function using liver parenchyma,spleen and portal vein signal intensities during the hepatobiliary phase in Gd-EOB-D TPA-enhanced MRI [J].BMC Med Imaging,2020,20(1):119-128.
- [6] 徐向东,罗诗维,韦瑞丽,等.CT 影像组学预测肾透明细胞癌核分级:扫描期相 ROI 勾画策略[J].放射学实践,2022,37(12):1542-1547.
- [7] 周玮,胡红杰,沈博,等.基于钆塞酸二钠增强磁共振成像影像组学定量评估肝硬化患者肝脏储备功能的应用价值[J].中国医学科学院学报,2020,42(4):459-467.
- [8] 张智星,黄忠江,何生,等.基于增强 CT 影像组学评估肝硬化患者肝储备功能的应用[J].放射学实践,2022,37(6):676-682.
- [9] 陆力坚,黄仲奎,龙莉玲.磁共振门静脉血流动力学改变与肝功能相关性[J].放射学实践,2010,25(10):1125-1127.
- [10] Oliveira MS,Silva RPM,Valle SDCND,et al.Chronic hepatitis B and D:prognosis according to Child-Pugh score[J].Rev Bras Enferm,2017,70(5):1048-1053.
- [11] Zhang L,Zhang J,Shen P,et al.Block level skip connections across cascaded V-net for multi-organ segmentation [J].IEEE Trans Med Imaging,2020,39(9):2782-2793.
- [12] Takamoto T,Nara S,Ban D,et al.Application of albumin-bilirubin grade and platelet count to indocyanine green-based criteria for hepatectomy:predicting impaired liver function and postoperative outcomes of hepatocellular carcinoma[J].J Surg Oncol,2022,126(4):680-688.
- [13] Amygdalos I,Czigany Z,Bednarsch J,et al.Low postoperative platelet counts are associated with major morbidity and inferior survival in adult recipients of orthotopic liver transplantation[J].J Gastrointest Surg,2020,24(9):1996-2007.
- [14] Kaiser T,Kinny-Köster B,Bartels M,et al.Cholesterol esterification in plasma as a biomarker for liver function and prediction of mortality[J].BMC Gastroenterol,2017,17(1):57-63.
- [15] Gao F,Cai MX,Lin MT,et al.Prognostic value of international normalized ratio to albumin ratio among critically ill patients with cirrhosis[J].Eur J Gastroenterol Hepatol,2019,31(7):824-831.
- [16] Jajodia A,Gupta A,Prosch H,et al.Combination of radiomics and machine learning with diffusion-weighted MR imaging for clinical outcome prognostication in cervical cancer [J].Tomography,2021,7(3):344-357.
- [17] Mayerhoefer ME,Materka A,Langs G,et al.Introduction to radiomics[J].J Nucl Med,2020,61(4):488-495.
- [18] 吴娇艳,刘影.探讨肝脏特异性对比剂评价肝功能的价值[J].临床放射学杂志,2018,37(6):962-966.
- [19] Roehlen N,Crouch E,Baumert TF.Liver fibrosis:mechanistic concepts and therapeutic perspectives[J].Cells,2020,9(4):875-917.
- [20] Tamada T,Ito K,Higaki A,et al.Gd-EOB-DTPA-enhanced MR imaging:evaluation of hepatic enhancement effects in normal and

- cirrhotic livers[J].Eur J Radiol,2011,80(3):e311-316.
- [21] Zhu WS,Shi SY,Yang ZH,et al.Radiomics model based on pre-operative gadoxetic acid-enhanced MRI for predicting liver failure[J].World J Gastroenterol,2020,26(11):1208-1220.
- [22] Cai W,He B,Hu M,et al.A radiomics-based nomogram for the preoperative prediction of posthepatectomy liver failure in patients with hepatocellular carcinoma[J].Surg Oncol,2019,28(3):78-85.
- [23] Lu X,Zhou H,Wang K,et al.Comparing radiomics models with different inputs for accurate diagnosis of significant fibrosis in chronic liver disease[J].Eur Radiol,2021,31(11):8743-8754.
- [24] Zheng W,Guo W,Xiong M,et al.Clinic-radiological features and radiomics signatures based on Gd-BOPTA-enhanced MRI for predicting advanced liver fibrosis[J].Eur Radiol,2023,33(1):633-644.
- [25] Nitsch J,Sack J,Halle MW,et al.MRI-based radiomic feature analysis of end-stage liver disease for severity stratification[J].Int J Comput Assist Radiol Surg,2021,16(3):457-466.
- (收稿日期:2023-01-07 修回日期:2023-03-20)

《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 2020 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》。

《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!

· 影像组学与人工智能专题 ·

基于数字化乳腺 X 线影像组学预测浸润性乳腺癌腋窝淋巴结转移的多中心研究

谢玉海, 马培旗, 王小雷, 韩剑剑, 马文俊, 曹雪花, 张宁宁, 杨杨, 胡东

【摘要】 目的:探讨基于多中心数字化乳腺 X 线影像组学预测浸润性乳腺癌腋窝淋巴结转移的临床应用价值。**方法:**回顾性搜集 728 例经病理证实的浸润性乳腺癌患者,按照 7:3 的比例将皖南医学院第一附属医院弋矶山医院 413 例浸润性乳腺癌患者随机拆分为训练组 289 例(淋巴结转移阴性 197 例,淋巴结转移阳性 92 例)和验证组 124 例(淋巴结转移阴性 85 例,淋巴结转移阳性 39 例),将阜阳市人民医院和太和县人民医院浸润性乳腺癌患者共计 315 例(淋巴结转移阴性 221 例,淋巴结转移阳性 94 例)作为外部测试组。对比分析双乳内外斜位(MLO)和头尾位(CC)图像,选取病变面积较大的数字化乳腺 X 线图像使用深睿医疗多模态科研平台进行图像分割及影像组学特征提取,采用特征间线性相关性分析与最小绝对收缩和选择算法(LASSO)对组学特征进行降维并使用支持向量机(SVM)分类器构建影像组学预测模型。采用受试者工作特征(ROC)曲线和决策曲线分析(DCA)评价模型性能。**结果:**最终筛选出 8 个影像组学特征构建预测模型用于预测浸润性乳腺癌腋窝淋巴结转移,该模型的预测效能训练组、验证组和外部测试组分别为 0.807、0.790 和 0.753,敏感度、特异度分别为 84.8%和 61.4%、79.5%和 69.4%、44.7%和 92.8%。决策曲线证实了该模型的临床实用性。**结论:**基于数字化乳腺 X 线影像组学对浸润性乳腺癌腋窝淋巴结转移的预测具有高效能,对患者制定个性化的治疗方案和预后评估有着重要的临床应用价值。

【关键词】 乳腺癌; 数字乳腺 X 线摄影; 腋窝淋巴结转移; 影像组学

【中图分类号】 R737.9; R814.41 **【文献标志码】** A **【文章编号】** 1000-0313(2024)01-0031-06

DOI:10.13609/j.cnki.1000-0313.2024.01.006

开放科学(资源服务)标识码(OSID):



Radiomics nomogram based on digital mammography for predicting axillary lymph node metastasis of invasive breast cancer: a multicenter study XIE Yu-hai, MA Pei-qi, WANG Xiao-lei, et al. Department of Radiology, Taihe People's Hospital, Taihe Hospital affiliated to Wannan Medical College, Anhui 236600, China

【Abstract】 Objective: To investigate the clinical value of radiomics nomogram based on multi-center digital mammography in predicting axillary lymph node metastasis of invasive breast cancer. **Methods:** A total of 728 patients with pathologically proven invasive breast cancer were retrospectively collected. 413 patients with invasive breast cancer from Yiji Mountain Hospital, the First Affiliated Hospital of Wannan Medical College, were randomly divided into a training group of 289 patients (197 cases with negative lymph node metastasis and 92 cases with positive lymph node metastasis) and a validation group of 124 patients (85 cases with negative lymph node metastasis and 39 cases with positive lymph node metastasis) in the ratio of 7:3. A total of 315 patients with invasive breast cancer (221 cases with negative lymph node metastasis and 94 cases with positive lymph node metastasis) from Fuyang People's Hospital and Taihe County People's Hospital were selected as the external test groups. The images from medio-lateral oblique (MLO) and cranio-caudal (CC) views of two breasts were compared and analyzed, and the digital mammogram images with larger lesion area were selected

作者单位: 236600 安徽, 太和县人民医院/皖南医学院附属太和医院放射影像科(谢玉海、王小雷、马文俊、曹雪花、张宁宁、杨杨、胡东); 236000 安徽, 安徽省阜阳市人民医院放射影像科(马培旗); 241000 安徽, 皖南医学院第一附属医院/弋矶山医院放射科(韩剑剑)

作者简介: 谢玉海(1985-), 男, 安徽省太和县人, 硕士, 副主任医师, 主要从事放射影像诊断工作。

通讯作者: 胡东, E-mail: 13866238585@163.com

基金项目: 皖南医学院科研项目(项目编号: JXY202139); 北京医学奖励基金会睿影科研基金(编号项目 YXJL-2022-0105-0116)

for image segmentation and radiomics feature extraction using the DeepMed multimodal research platform. The dimension of radiomics features was reduced by Linear correlation analysis between features and Least absolute shrinkage and selection operator (LASSO) and a prediction model was constructed by a support vector machine (SVM) classifier. The performance of the model was evaluated by Receiver Operating Characteristic (ROC) and Decision Curve Analysis (DCA). **Results:** Eight radiomics features were finally selected to construct a model for predicting axillary lymph node metastasis in invasive breast cancer. The prediction performance of the model was 0.807, 0.790 and 0.753 in the training group, validation group and external test group, respectively, with sensitivities and specificities of 84.8% and 61.4%, 79.5% and 69.4%, 44.7% and 92.8%, respectively. The decision curve confirmed the clinical practicability of the model. **Conclusion:** Radiomics based on digital mammography has high efficiency in predicting axillary lymph nodes metastasis in invasive breast cancer and has important clinical application value for the formulation of individualized treatment plans and prognosis assessment for patients.

【Key words】 Breast cancer; Digital mammography; Axillary lymph node metastasis; Radiomics

乳腺癌(breast cancer, BC)是最常见的癌症之一,也是全世界女性死亡的主要原因^[1],我国女性乳腺癌的发病率和死亡率均居世界首位^[2]。腋窝淋巴结转移(axillary lymph node metastasis, ALNM)在乳腺癌患者中很常见,与疾病的临床分期、治疗方案、手术方式和患者预后密切相关^[3]。目前,乳腺癌患者的腋窝淋巴结转移状况是通过前哨淋巴结活检(sentinel lymph node biopsy, SLNB)或腋窝淋巴结清扫(axillary lymph node dissection, ALND)来证实的,但这些手术有一定风险,可能会导致肿瘤植入性转移、邻近血管神经损伤以及上肢淋巴水肿等^[4]。数字化乳腺 X 线摄影(digital mammography, DM)是乳腺癌体检筛查的常用检查方法,但评估腋窝淋巴结是否存在转移的准确性较低^[5]。影像组学通过提取肉眼无法观察的高通量图像特征来预测恶性肿瘤的生物学行为^[6]。基于数字乳腺 X 线影像中包含的影像组学特征组合或单一影像组学特征的分析可靠、准确、非侵入性和经济有效地评估乳腺癌的特征,在预测组织学分级、ALNM、侵袭性、受体状态、Ki-67 增殖状态、分子亚型以及复发和生存方面具有很大潜力^[7-11]。虽然已有研究基于多参数 MRI 影像组学特征预测乳腺癌腋窝淋巴结转移^[12],但基于 DM 影像组学预测浸润性乳腺癌腋窝淋巴结转移的报道较少,且多为单中心研究^[13-14]。因此,本研究旨在探讨基于多中心 DM 影像组学预测浸润性乳腺癌腋窝淋巴结转移的临床应用价值。

材料与方法

1. 病例资料

回顾性分析 2016 年 12 月至 2021 年 12 月间皖南医学院第一附属医院弋矶山医院(机构 1)、2018 年 1 月至 2021 年 6 月间阜阳市人民医院(机构 2)、2016 年

7 月至 2022 年 3 月间太和县人民医院(机构 3)经病理证实的女性乳腺癌患者 728 例,其中皖南医学院第一附属医院弋矶山医院 413 例、阜阳市人民医院和太和县人民医院共计 315 例。病例纳入标准:①经病理证实为浸润性癌;②X 线检查前无穿刺、手术或新辅助化疗史等;③乳腺 X 线检查后 2 周内行穿刺或手术切除。病例排除标准:①病变显示不清或瘤体显示不全,无法行感兴趣区(region of interest, ROI)勾画;②有其他恶性肿瘤病史。将机构 1 乳腺癌患者随机拆分为训练组($n=289$)和验证组($n=124$),将机构 2 和机构 3 乳腺癌患者视作外部测试组($n=315$)。本研究为回顾性研究,经医院伦理委员会审批通过免除患者知情同意。

2. 检查方法

机构 1 采用美国 HOLOGIC 数字化乳腺 X 线机,型号为 selenia dimensions,图像尺寸为 $24\text{ cm} \times 30\text{ cm}$ 。机构 2 采用 GE 数字化乳腺 X 线机,型号为 Senographe Ess entail,图像尺寸为 $24\text{ cm} \times 31\text{ cm}$ 。机构 3 采用联影数字化乳腺 X 线机,型号为 uMammo 590i,图像尺寸为 $24\text{ cm} \times 30\text{ cm}$ 。拍摄体位均包括双侧乳腺头尾位(cranial caudal, CC)和内外侧斜位(mediolateral oblique, MLO),曝光条件采用全自动曝光模式。本研究在特征提取前对所有图像均进行分辨率重采样以避免因扫描机器参数不同导致的误差,重采样像素间距设为(1.0, 1.0, 1.0)。

3. 病灶分割与特征提取

将所有患者 X 线图像从 PACS 工作站以 DICOM 格式导出后导入深睿医疗多模态科研平台(<https://keyan.deepwise.com>)进行图像分割及特征提取、降维。图像分割由 2 位具有 5 年诊断经验的放射科医师 A 和 B 完成。在进行分割时,2 位放射科医生在被告

知肿瘤的确切位置后手动勾画分割病灶,结合 CC 位和 MLO 位选择病灶显示最清晰且面积最大的体位进行 ROI 勾画(图 1),医师 A 于 2 周后再次对 DM 图像进行 ROI 勾画。图像通过均一化处理后自动提取 1562 个组学特征,包括一阶特征 324 个、形状特征 14 个、灰度共生矩阵(gray level co-occurrence matrix, GLCM)特征 396 个、灰度区域矩阵(gray level co-occurrence matrix, GLCM)特征 288 个、灰度游程矩阵(gray level run length matrix, GLRLM)特征 288 个、灰度相依矩阵(gray level dependence matrix, GLDM)特征 252 个。

4.特征筛选与建模

第一步将所有样本量的特征参数通过特征剔除(剔除缺失值≥10%的特征)和组间相关系数(intra-class correlation coefficient, ICC)筛选出一致性较好的非零组学特征(ICC>0.85)。第二步使用深睿医疗

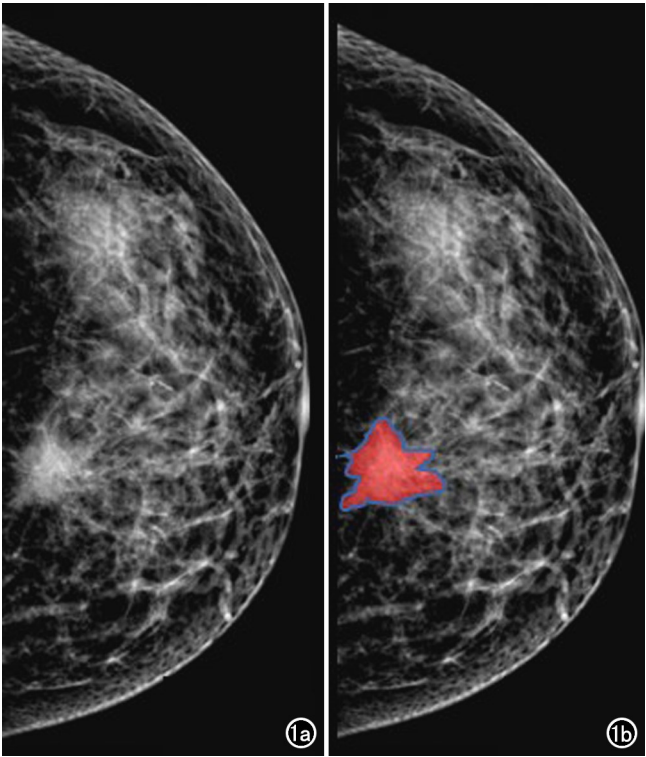


图 1 a) 乳腺癌肿块原始图像;b) 手动勾画 ROI 示意图(红色区域)。

多模态科研平台,将机构 1 样本量按照 7:3 随机拆分为训练组和验证组,将机构 2 和机构 3 样本量作为外部测试组,在训练组中通过特征间线性相关性分析,相关系数(coefficient, C)=0.8 与最小绝对收缩和选择算法(least absolute shrinkage and selection operator, LASSO)方法进行特征降维筛选出最优组学特征,使用支持向量机(support vector machine, SVM)分类器参与构建影像组学模型,验证组和外部测试组对模型的稳定性与可重复性进行验证。

5.统计学分析

采用 R 软件(Version3.6.0)进行统计学分析。采用 ICC 评估影像组学特征提取的观察者间和观察者内的一致性。使用 MedCalc 软件(Version19.0.2)绘制受试者工作特征(receiver operating characteristic, ROC)曲线,采用曲线下面积(area under curve, AUC)和决策曲线分析(decision curve analysis, DCA)评价模型性能。以 $P<0.05$ 为差异有统计学意义。

结 果

1.病例资料

本研究纳入的皖南医学院第一附属医院弋矶山医院 413 例浸润性乳腺癌患者随机拆分为训练组 289 例(淋巴结转移阴性 197 例,淋巴结转移阳性 92 例)和验证组 124 例(淋巴结转移阴性 85 例,淋巴结转移阳性 39 例);阜阳市人民医院和太和县人民医院浸润性乳腺癌患者共计 315 例(淋巴结转移阴性 221 例,淋巴结转移阳性 94 例)作为外部测试组。训练组年龄 30~84 岁,平均(53.55±9.70)岁;验证组年龄 33~78 岁,平均(55.19±9.83)岁;外部测试组年龄 23~84 岁,平均(53.04±10.20)岁。

2.影像组学特征筛选及模型构建

第一步通过特征剔除(剔除缺失值≥10%的特征)和组间相关系数(ICC>0.85)分析剔除 558 个特征。剩余 1004 个组学特征通过特征间线性相关性分析($C=0.8$)和 LASSO 回归进行特征降维后筛选出 8 个与浸润性乳腺癌腋窝淋巴结转移预测最相关的影像组学特征(表 1),经支持向量机分类器构建预测模型。

表 1 影像组学预测浸润性乳腺癌腋窝淋巴结转移的最优特征

特征名称	特征系数	t/Z 值	P 值
exponential_glcm_ClusterShade	0.3121	-3.191	0.001
exponential_gldm_SmallDependenceLowGrayLevelEmphasis	0.0225	-3.882	<0.001
exponential_glrml_ShortRunLowGrayLevelEmphasis	0.7682	-1.993	0.046
gradient_glcm_Imc1	-0.2977	-2.140	0.032
lbp-2D_gldm_SmallDependenceEmphasis	0.2386	-2.593	0.010
lbp-2D_glrml_RunLengthNonUniformity	0.8265	-4.088	0.001
logarithm_glcm_Imc2	0.141	-3.076	0.002
wavelet-LLL_glcm_Contrast	0.4043	-4.060	<0.001

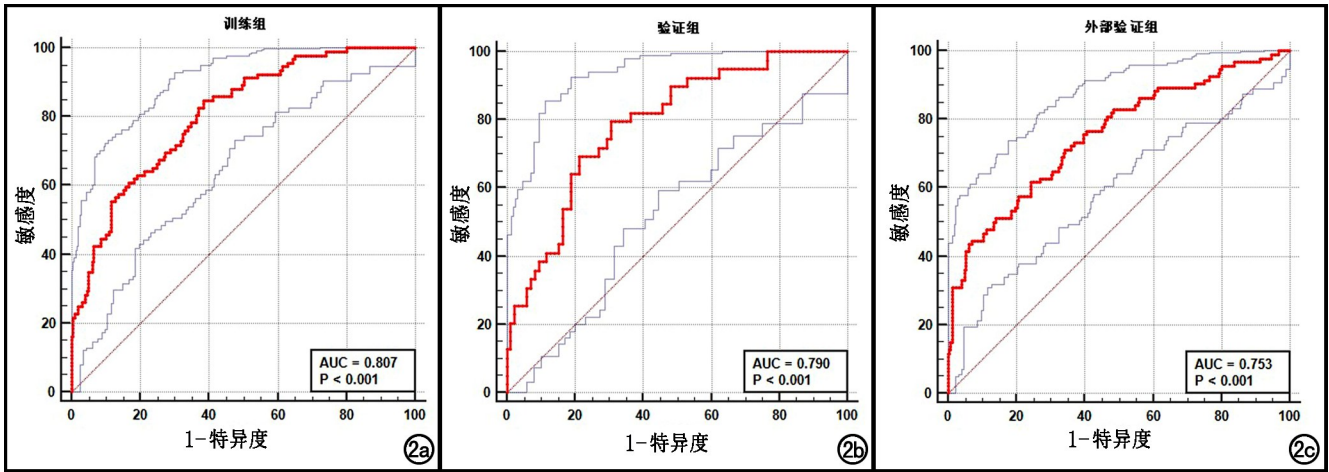


图 2 预测模型训练组、验证组及外部测试组 ROC 曲线。a) 训练组;b) 验证组;c) 外部测试组。

3.影像组学预测模型的诊断效能

预测模型在训练组、验证组、外部测试组中预测乳腺癌腋窝淋巴结转移的 AUC 分别为 0.807、0.790、0.753,敏感度、特异度分别为 84.8%和 61.4%、79.5%和 69.4%、44.7%和 92.8%(表 2、图 2)。决策曲线分析结果显示在概率值为 15%~80%时,预测模型鉴别乳腺癌腋窝淋巴结转移具有较好的净收益(图 3)。

表 2 预测模型在三组间的诊断效能

组别	AUC	95%CI	最佳界值	敏感度	特异度
训练组	0.807	0.755~0.858	0.245	0.848	0.614
验证组	0.790	0.707~0.873	0.291	0.795	0.694
外部验证组	0.753	0.692~0.814	0.423	0.447	0.928

讨 论

乳腺 X 线摄影是目前临床乳腺癌筛查最常用的影像学方法之一。本研究基于 DM 影像组学研究方法,通过多中心大样本构建影像组学模型用于预测浸润性乳腺癌发生腋窝淋巴结转移的可能性,从而进一步挖掘 DM 的潜能,使患者和临床受益。

腋窝淋巴结转移对于评估乳腺癌患者的预后和制定治疗方案具有重要意义^[15],腋窝淋巴结清扫术是确定腋窝淋巴结状况的最可靠方法,但手术会引起术区

感染、神经血管损伤、功能区障碍以及植入性转移等并发症^[16]。因此,为了减少不必要的手术损伤及过度医疗给患者造成的负担,迫切需要一种新的术前预测方式,在提高诊断准确性的同时非侵入性地判断腋窝淋巴结是否存在转移。虽然影像学检查可以显示肿大的淋巴结形态学特征,但是无法与一些炎性淋巴结区分^[5]。本研究通过乳腺 X 线摄影图像构建的影像组学模型在训练组、验证组及外部测试组中的 AUC 分别为 0.807、0.790 和 0.753,诊断效能较传统影像学检查方法得到进一步提高。陈春发等^[17]研究表明,乳腺 X 线摄影诊断乳腺癌 ALNM 与病理诊断的一致性检验 Kappa 值仅为 0.354。

影像组学可以提取高通量的组学特征,并通过分析获取一些肉眼无法识别的纹理特征及高阶特征^[18-19]。本研究最终筛选出 8 个与腋窝淋巴结转移相关的影像组学特征用于构建预测模型,其中 1 个灰度依赖矩阵特征(exponential_gldm_SmallDependenceLowGrayLevelEmphasis)、2 个灰度游程矩阵特征(lbp-2D_glrilm_RunLengthNonUniformity、exponential_glrilm_ShortRunLowGrayLevelEmphasis)、4 个灰度共生矩阵特征(exponential_glcmm_ClusterShade、wavelet-LLL_glcmm_Contrast、gradient_glcmm_Imc1、logarithm_glcmm_Imc2)。灰度共生矩阵(glcmm)中的 Cluster Shade 是衡量偏度和均匀度的纹理特征。李翠平等^[20]研究表明宫颈癌 Ki-67 低表达组的 Cluster Shade 值低于高表达组,而本研究结果也表明浸润性乳腺癌腋窝淋巴结转移组的 Cluster Shade 值高于非转移组。glcmm 中的 Contrast、Imc1 和 Imc2 主要反映纹理的粗糙性和复杂性,其绝对值越高,说明图像越粗糙、纹理越复杂,肿瘤的异质性越高。本研究腋窝淋巴结转移组的 glcmm_Contrast、glcmm_Imc1 和 glcmm_Imc2 的绝对值均高于非转移组,说明腋窝淋巴结转移

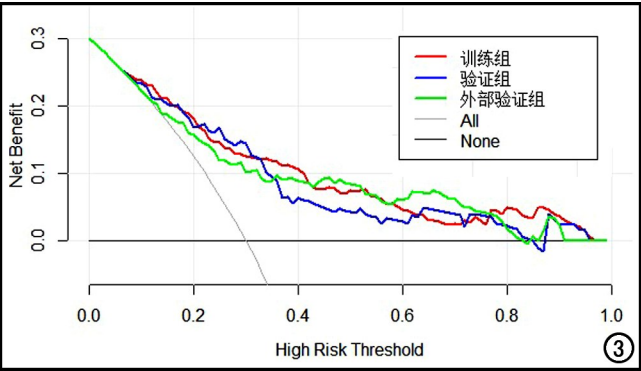


图 3 预测模型的决策曲线。

组的肿瘤异质性较高。朱海涛等^[21]的研究结果也证实 glcm_Contrast 值越高,肿瘤的恶性程度越高。灰度依赖矩阵(GLDM)参数是描述图像中的灰度相关性,是基于中心像素或体素与其邻近区域之间的灰度关系,其中 Small Dependence Low Gray Level Emphasis 是计算 ROI 中小相关性与图像中较暗部分的分散情况。本研究发现腋窝淋巴结转移组的 Small Dependence Low Gray Level Emphasis 值高于非转移组,说明 ALNM 组的图像均匀性差,笔者认为这可能是由于 ALNM 组的恶性程度高、瘤内易发生坏死所致。灰度游程矩阵(glrlm)中的 Run Length Non Uniformity 和 Short Run Low Gray Level Emphasis 主要描述像素值的分布情况,前者描述整个图像中游程长度的相似性,后者测量图像中较短游程长度与较低灰度值的联合分布,两者的值越大,像素分布越不均匀,纹理越粗糙。本研究发现 ALNM 组的 Run Length Non Uniformity 和 Short Run Low Gray Level Emphasis 值高于非转移组,说明 ALNM 组的 DM 图像纹理较非转移组更粗糙。此外,本研究最终采用支持向量机分类器(SVM),因为它是一种常用而有效的算法,该算法具有精度高、计算能力小等优点,SVM 在识别复杂数据集中的微妙模式方面非常强大^[22]。刘畅等^[23]使用支持向量机分类器构建预测模型,能够在术前无创评估和预测肝细胞癌微血管浸润,可作为指导临床后续个性化治疗的有效工具。

本研究存在以下局限性:①影像组学特征是在二维图像上进行手动勾画后提取,存在一定的人为误差;②在选择病灶方面仅从 MLO 位和 CC 位两者之一选择病灶最大层面进行 ROI 勾画,缺乏一定的对比,可能遗漏部分特征,影响最终预测效能;③未将基于 DM 影像组学标签与 MRI 影像组学特征预测腋窝淋巴结转移的效能进行比较。本研究结果与 Han 等^[24]和 Mao 等^[25]基于磁共振动态增强图像的影像组学预测乳腺癌 ALNM 的效能一致,而 Yu 等^[26]联合磁共振增强 T₁WI、T₂WI 和扩散加权成像 ADC 图构建的影像组学模型预测乳腺癌 ALNM 的效能在训练集和验证集分别为 0.88 和 0.85,高于本研究结果。因此,今后将通过前瞻性设计,开展比较 DM 影像组学标签与 MRI 影像组学标签预测乳腺癌 ALNM 效能的后续研究。

综上所述,基于乳腺 X 线摄影的影像组学预测模型可作为一种非侵入性工具,对浸润性乳腺癌腋窝淋巴结转移具有较高的诊断效能,对乳腺癌患者制定个性化的治疗方案和预后评估有着重要的临床应用价值。

参考文献:

- [1] Wang X, Xie T, Luo J, et al. Radiomics predicts the prognosis of patients with locally advanced breast cancer by reflecting the heterogeneity of tumor cells and the tumor microenvironment[J]. Breast Cancer Res, 2022, 24(1): 20.
- [2] 张雪, 董晓平, 管雅喆, 等. 女性乳腺癌流行病学趋势及危险因素研究进展[J]. 肿瘤防治研究, 2021, 48(1): 87-92.
- [3] Zhang J, Li L, Zhe X, et al. The diagnostic performance of machine learning-based radiomics of DCE-MRI in predicting axillary lymph node metastasis in breast cancer: a meta-analysis[J]. Front Oncol, 2022, 12(4): 799209.
- [4] Kootstra J, Hoekstra-Weebers JE, Rietman H, et al. Quality of life after sentinel lymph node biopsy or axillary lymph node dissection in stage I/II breast cancer patients: a prospective longitudinal study[J]. Ann Surg Oncol, 2008, 15(9): 2533-2541.
- [5] 乔江华, 朱立元, 韦伟. 数字化钼靶检查在判断乳腺癌腋窝淋巴结转移中的价值探讨[J]. 临床外科杂志, 2007, 15(11): 751-752.
- [6] 杨采薇, 刘曦娇, 魏毅, 等. CT 影像组学联合传统影像特征对胃肠道间质瘤 Ki-67 指数表达的预测价值[J]. 放射学实践, 2022, 37(9): 1068-1073.
- [7] Siviengphanom S, Gandomkar Z, Lewis SJ, et al. Mammography-based radiomics in breast cancer: a scoping review of current knowledge and future needs[J]. Acad Radiol, 2022, 29(8): 1228-1247.
- [8] Niu S, Jiang W, Zhao N, et al. Intra- and peritumoral radiomics on assessment of breast cancer molecular subtypes based on mammography and MRI[J]. J Cancer Res Clin Oncol, 2022, 148(1): 97-106.
- [9] Jiang T, Song J, Wang X, et al. Intratumoral and peritumoral analysis of mammography, tomosynthesis, and multiparametric MRI for predicting Ki-67 level in breast cancer: a radiomics-based study[J]. Mol Imaging Biol, 2022, 24(4): 550-559.
- [10] Wang G, Shi D, Guo Q, et al. Radiomics based on digital mammography helps to identify mammographic masses suspicious for cancer[J]. Front Oncol, 2022, 12(1): 843436.
- [11] 张玉姣, 宋德领, 王燕飞, 等. 基于乳腺 X 线图像影像组学列线图对乳腺癌腋窝淋巴结转移的预测价值[J]. 放射学实践, 2022, 37(1): 48-54.
- [12] 姜原, 马明明, 程元甲, 等. 基于乳腺癌动态增强 MRI 图像的影像组学特征预测腋窝淋巴结转移的可行性研究[J]. 中华放射学杂志, 2022, 56(6): 631-635.
- [13] Tan H, Wu Y, Bao F, et al. Mammography-based radiomics nomogram: a potential biomarker to predict axillary lymph node metastasis in breast cancer[J]. Br J Radiol, 2020, 93(1111): 20191019.
- [14] Yang J, Wang T, Yang L, et al. Preoperative prediction of axillary lymph node metastasis in breast cancer using mammography-based radiomics method[J]. Sci Rep, 2019, 9(1): 4429.
- [15] Hu J, Xia X, Yang H, et al. Dissection of level III axillary lymph nodes in breast cancer[J]. Cancer Manag Res, 2021, 13(26): 2041-2046.
- [16] 马芹芹, 冯雯, 陈圆圆, 等. 术前 MRI 预测乳腺癌腋窝淋巴结转移的研究进展[J]. 磁共振成像, 2022, 13(9): 151-155.
- [17] 陈春发, 张玉铃, 林豪雨, 等. 乳腺彩超和 X 线摄影在乳腺癌早期淋巴结的评估作用[J]. 汕头大学医学院学报, 2016, 29(2): 109-110, 115.

- [18] Yan BC, Li Y, Ma FH, et al. Radiologists with MRI-based radiomics aids to predict the pelvic lymph node metastasis in endometrial cancer: a multicenter study[J]. Eur Radiol, 2021, 31(1): 411-422.
- [19] Rogers W, Thulasi Seetha S, Refaee TAG, et al. Radiomics: from qualitative to quantitative imaging [J]. Br J Radiol, 2020, 93(1108): 20190948.
- [20] 李翠平, 郑明雪, 董江宁, 等. 基于 MR T₂WI 的纹理分析术前预测宫颈鳞癌 Ki-67 表达水平的研究[J]. 实用放射学杂志, 2022, 38(5): 786-790, 800.
- [21] 朱海涛, 张晓燕, 李晓婷, 等. 基线 MRI 中直肠癌原发灶与淋巴结交互特征预测新辅助治疗后淋巴结状态[J]. 肿瘤影像学, 2022, 31(2): 113-120.
- [22] Huang S, Cai N, Pacheco PP, et al. Applications of support vector machine (SVM) learning in cancer genomics[J]. Cancer Genomics Proteomics, 2018, 15(1): 41-51.
- [23] 刘畅, 赵泓博, 黄京城, 等. 基于 CT 增强影像组学的支持向量机模型术前预测肝细胞癌微血管浸润的研究[J]. 临床放射学杂志, 2021, 40(12): 2390-2396.
- [24] Han L, Zhu YB, Liu Z, et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer[J]. Eur Radiol, 2019, 29(7): 3820-3829.
- [25] Mao N, Dai Y, Lin F, et al. Radiomics nomogram of DCE-MRI for the prediction of axillary lymph node metastasis in breast cancer [J]. Front Oncol, 2020, 10(27): 541849.
- [26] Yu YF, Tan YJ, Xie CM, et al. Development and validation of a preoperative magnetic resonance imaging radiomics-based signature to predict axillary lymph node metastasis and disease-free survival in patients with early-stage breast cancer [J]. JAMA Netw Open, 2020, 3(12): e2028086.

(收稿日期: 2022-12-29 修回日期: 2023-04-13)

欢迎订阅 2024 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 39 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 2020 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

· 影像组学与人工智能专题 ·

基于胸部 X 线片的深度迁移学习模型早期诊断儿童肺炎支原体肺炎

孟名柱, 潘昌杰, 张浩, 陈晓霞

【摘要】 目的:探讨基于 VGG19 的深度迁移学习模型在儿童胸部 X 线片(胸片)早期诊断肺炎支原体肺炎(MPP)中的价值。**方法:**搜集常州市第二人民医院儿科就诊的患儿 3763 例,其临床及影像资料齐全。将患儿胸片分成细菌性肺炎组、MPP 组和病毒性肺炎组,并按照 9:1 随机分为训练集和测试集。另外选择 150 例患儿胸片作为验证集(细菌性肺炎、MPP 和病毒性肺炎各 50 例)。模型性能评价指标包括在训练集和测试集中的准确率(Ac)及在验证集中的精确度(Pr)、召回率(Rc)、F1 评分(F1)和 ROC 曲线下面积(AUC)。**结果:**VGG19 在训练集和测试集中的最高 Ac 分别为 0.99 和 0.95。细菌性肺炎组的 Pr、Rc、F1、AUC 分别为 0.87、0.82、0.85、0.92, MPP 组分别为 0.85、0.74、0.78、0.90, 病毒性肺炎组分别为 0.76、0.88、0.81、0.93。VGG19 对三组图像分类诊断的总体 AUC 为 0.92。**结论:**基于 VGG19 的深度迁移学习模型是儿童胸部 X 线片诊断 MPP 的可靠方法,能够帮助临床早期诊断 MPP。

【关键词】 肺炎支原体肺炎; 胸片; 儿童; 深度迁移学习; 人工智能

【中图分类号】 R725.6; R814.41 **【文献标识码】** A **【文章编号】** 1000-0313(2024)01-0037-05

DOI:10.13609/j.cnki.1000-0313.2024.01.007

开放科学(资源服务)标识码(OSID):



Early diagnosis of mycoplasma pneumoniae in children based on chest X-ray images using deep transfer learning method MENG Ming-zhu, PAN Chang-jie, ZHANG Hao, et al. Department of Radiology, the Second Hospital of Changzhou Affiliated to Nanjing Medical University, Jingsu 213164, China

【Abstract】 Objective: This study was aimed to investigate the diagnostic value of a deep transfer learning method based on chest X-ray images for mycoplasma pneumoniae (MPP) in children. **Methods:** A total of 3763 cases from the Department of Pediatrics of the Second People's Hospital of Changzhou were recruited in this study. Clinical data and radiographic data of each patient were completely recorded. The chest X-ray images were divided into bacterial, MPP, and viral groups. Each group was randomly divided into a train set and test set by a ratio of 9:1. Deep transfer learning based on the VGG19 model was established. Also, another 150 chest X-ray image samples with bacterial pneumonia, MPP and viral pneumonia were selected as a validation set; 50 cases each. Metrics for model performance evaluation include accuracy in the train set and test set, precision (Pr), recall rate (Rc), F1 score (F1) and area under the ROC curve (AUC) in the validation set. **Results:** Accuracies of VGG19 on the train set and the test set were 0.99 and 0.95, respectively. The values of Pr, Rc, F1 and AUC were bacterial group (0.87, 0.82, 0.85, 0.92), MPP group (0.85, 0.74, 0.78, 0.90) and viral group (0.76, 0.88, 0.81, 0.93), respectively. The VGG19 model achieved an overall AUC of 0.92 for three groups. **Conclusion:** The VGG19 model is a reliable method to classify chest X-ray images of MPP children, which is a promising AI-assisted diagnosis method for early diagnosis of MPP.

【Key words】 Mycoplasma pneumoniae pneumonia; Chest films; Child; Deep transfer learning; Artificial intelligence

肺炎支原体肺炎(mycoplasma pneumoniae pneu-

monia, MPP)是由于肺炎支原体(mycoplasma pneumoniae, MP)所引起的呼吸道感染及肺部的急性炎症性改变,以儿童和青年为主^[1]。一项前瞻性、多中心的儿童肺炎支原体流行病学监测研究显示,在新型冠状病毒肺炎流行期间,有 9.0% 的患儿检测到有 MP 感

作者单位:213164 江苏,南京医科大学附属常州第二人民医院医学影像科

作者简介:孟名柱(1984-),男,安徽凤阳人,硕士研究生,主治医师,主要从事人工智能辅助影像诊断工作。

通讯作者:陈晓霞, E-mail: czeycxs2022@163.com

基金项目:常州市科技局应用基础研究计划(CJ20220260)

染^[2]。MP 对呼吸道上皮细胞具有高亲和力,易引发机体毒性作用及严重的炎症反应,而且病程较长,可导致患儿免疫功能降低,对患儿生长发育造成一定影响。因此,MPP 的早期诊断及治疗非常重要。MPP 患者的临床表现缺乏特异性,目前确诊还需要实验室检查^[3]。MP 的检测方法主要分病原学检测和血清学检测两大类,不同的检测方法各有其特点,应依据不同需求和目的选择不同检测方法,其中聚合酶链反应(polymerase chain reaction,PCR)被认为是金标准,但该方法操作复杂,而且对实验人员及仪器的要求较高,对于一些基层医院及一般实验室难以开展。此外,PCR 比较耗时,不利于 MPP 的早期诊断。胸部 X 线评估对儿童 MPP 诊疗管理具有重要意义^[4]。但是由于不同病原体所致的肺炎在胸部 X 线片(简称胸片)中的表现具有相似性,即使是经验丰富的影像科医生,也难以通过胸片做出 MPP 的诊断^[3]。

因此,如果能够运用人工智能技术辅助 MPP 的诊断,在节约人力的同时还可以早期诊断 MPP。随着人工智能的不断发展,深度学习在医学图像分析领域的应用逐渐增多^[5]。其中,医学影像与人工智能的结合是最具发展前景的领域^[6]。Thakur 等^[7]应用深度学习方法在胸片和 CT 混合图像的数据集中对 COVID-19 患者进行分类诊断,结果显示准确率达98.28%。戴正行等^[8]的研究结果显示,相较于人工诊断,应用人工智能技术对肺结节的良、恶性进行诊断具有良好、可靠的诊断准确性。本文提出了一种基于 VGG19 的深度迁移学习(deep transfer learning,DTL)模型算法对 MPP 胸部 X 线片进行诊断和鉴别诊断,旨在实现 MPP 的自动检测,以期辅助影像科医师在胸片中对 MPP 进行早期诊断,这有助于 MPP 患儿及时有效的治疗,减少耐药性的发生。

材料与方法

1.病例资料

搜集 2017 年 1 月至 2021 年 12 月在常州市第二人民医院儿科就诊的患儿共 3763 例,临床资料及影像资料(胸部 X 线片)齐全。将患儿胸片分为细菌性肺炎组(简称细菌组,1273 例,其中男 673 例,女 600 例)、MPP 组(1246 例,其中男 962 例,女 284 例)和病毒性肺炎组(简称病毒组,1244 例,其中男 752 例,女 492 例)。本研究为回顾性研究,通过医院伦理审查委员会批准[编号(2020)KY234-01],免除患者知情同意。

本研究 MPP 组病例的纳入标准:即临床有肺炎的表现和胸部 X 线改变,且 MP 病原学检查阳性的患儿。MPP 诊断标准参照《儿童肺炎支原体肺炎诊治专

家共识(2015 年版)》^[9]。本研究细菌性肺炎和病毒性肺炎诊断标准参照《儿童社区获得性肺炎管理指南(2013 修订)(上)》^[10]。

2.影像学检查

胸片的数据采集分别在 6 台 X 线设备上,进行,后前位摄片(PA)占比 73.82%(表 1)。因胸部侧位 X 线片较少,不纳入本研究。

表 1 三组患儿数据采集 X 线设备分布

设备	PA	AP	管电压(kV)	管电流(mAS)
飞利浦 DR 4 台	1260	425	81	自动
AGFA DR 1 台	826	330	68	自动
西门子 DR 1 台	692	230	66	11.2
合计	2778	985		

注:PA,后前位摄片;AP,前后位摄片。

3.训练集、测试集和验证集

训练集和测试集图像由两部分构成。第一部分为网络公开数据集(<https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>):经过筛选,共获得此部数据图像 9882 幅^[11]。第二部分为自建数据集,共获得图像 3763 幅。运用程序指令将 3 组图像随机打乱,并按照 9:1 随机分为训练集(细菌组 6837 幅,MPP 组 1122 幅,病毒组 4327 幅)和测试集(细菌组 756 幅,MPP 组 124 幅,病毒组 480 幅)。另外选择 150 例患儿胸片作为验证集(细菌组、MPP 组和病毒组各 50 幅),验证集的图像没有在训练集中出现过。数据集总体分布见图 1。

4.计算机环境配置

采用 Windows 10 企业版 64 位操作系统,处理器为酷睿 i7 10700F,8 核心 16 线程,图形处理器(graphics processing unit, GPU)为 NVIDIA GeForce RTX2060,显存 6GB,程序运行时关闭其它所有应用程序。模型参数设置见表 2,DTL 模型结构设计见图 2。

表 2 模型参数设置

模型参数	参数值
旋转角度	60°
剪切范围	0.2
扩大范围	0.2
镜像翻转	True
垂直翻转	True
充填模式	Nearest
保存最好模型	True
早停	False
激活函数	Relu
分类函数	Softmax
丢失值	0.5
训练轮回	250

5.模型评估指标

模型评估指标中的准确率(accuracy, Ac)、损失值、精确度(precision, Pr)、召回率(recall rate, Rc)、F1

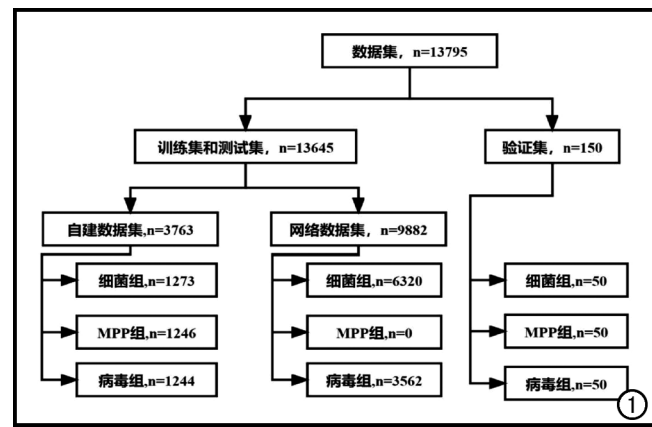


图 1 数据集分布图。

评分(F1 score, F1)和受试者工作特征(receiver operating characteristic, ROC)曲线的曲线下面积(area under curve, AUC)为模型运算得出。平均 Ac、平均 Rc 和平均 F1 为模型对验证集中三组数据验证结果的平均值,而不是模型在验证集中三组数据 Ac、Rc、F1 的平均数。计算公式分别如下:

$$Ac = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Pr = \frac{TP}{TP + FP} \quad (2)$$

$$Rc = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2 \times PPV \times Rc}{PPV + Rc} \quad (4)$$

在本研究中,TP 和 TN 分别表示模型正确诊断的三组图像数量,FP 和 FN 分别表示模型错误诊断的三组图像数量,AUC 为模型运算得出。F1 分数值为 0~1 之间,同时兼顾了分类模型的 Ac 和 Rc,可以认为是模型 Ac 和 Rc 的一种加权平均。

6. 统计学分析

采用 SPSS 23.0 软件进行统计学分析。患者年龄以均值±标准差表示,组间比较采用方差分析。以 $P < 0.05$ 差异有统计学意义。

结 果

1. 患儿年龄比较

三组患儿的年龄比较差异有统计学意义($F = 90.78, P < 0.01$),其中细菌组患儿年龄 $[(4.86 \pm 2.45)$ 岁]与病毒组 $[(4.94 \pm 2.52)$ 岁]比较差异无统计学意义($P = 0.384$),细菌组、病毒组患儿年龄与 MPP 组 $[(3.75 \pm 2.46)$ 岁]比较差异均有统计学意义(P 值均 < 0.01)。

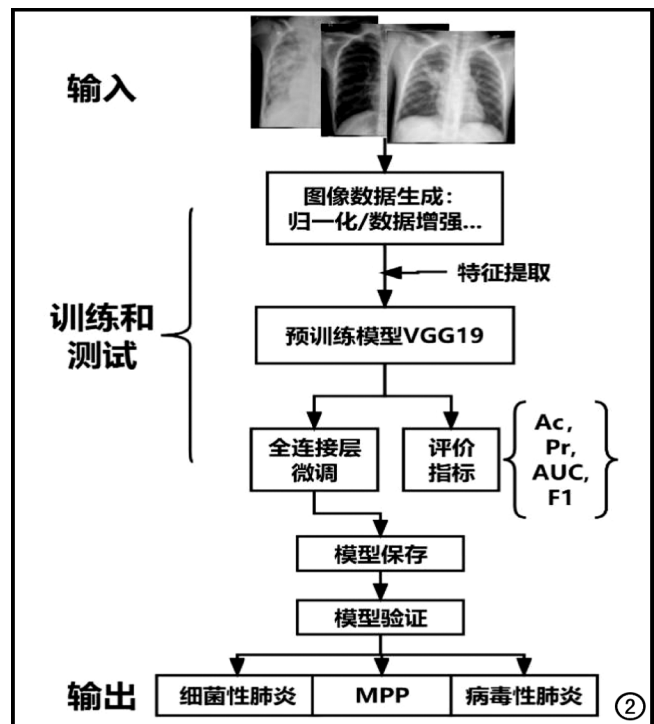


图 2 深度学习模型架构图。

2. 数据训练结果

学习曲线图(图 3)显示训练集准确率随着 epoch 的增加而逐渐增加,最高准确率为 0.99,且保持相对稳定。测试集准确率变化幅度较训练集大,最高准确率为 0.95。训练集的损失值随着 epoch 增加而逐渐减小,并趋于稳定。测试集的损失值在第 20 个 epoch 时最低,在此之后有逐渐增高的趋势。以上结果表明所建立的迁移学习模型在数据集中已经出现了过拟合现象。

3. 类激活热力图

类激活热力图是一种深度学习的可视化方法,它有助于了解一幅图像的哪一部分让卷积神经网络做出了最终的分类决策。这有助于对卷积神经网络的决策过程进行调试,特别是出现分类错误的情况下。类激活热力图是与特定输出类别相关的二维分数网格,对任何输入图像的每个位置都要进行计算,它表示每个

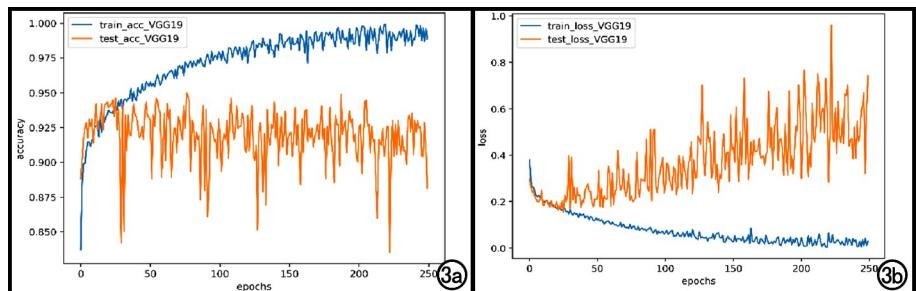


图 3 VGG19 在训练集和测试集中的学习曲线图。a)在训练集和测试集中的准确率;b)在训练集和测试集中的损失值。accuracy: 准确率;epochs: 训练的回合数,1 个 epoch 表示训练了 1 遍训练集中的所有样本;loss: 损失值。

位置对该类别的重要程度。本研究结果显示卷积神经网络对细菌组胸片的激活部位与激活程度明显多于 MPP 组和病毒组(图 4)。

4.DTL 模型验证结果

混淆矩阵图(图 5)分析结果显示在验证集中,VGG19 对 MPP 组的诊断准确率为 0.74(37/50),低于细菌组的 0.82(41/50)和病毒组的 0.88(44/50)。

VGG19 模型在验证集中对细菌组、MPP 组和病毒组胸部 X 线片的平均 Pr、Rc 和 F1 值分别为 0.81、0.81 和 0.81。VGG19 在验证集中对细菌组的 Pr 值和 F1 值最高,分别为 0.87、0.85。VGG19 对病毒组的 Rc 值最高,为 0.88(表 3)。

VGG19 对三组图像分类诊断的总体 AUC 为 0.92,对病毒组分类诊断的 AUC 最高,为 0.93,对细菌组和 MPP 组分类诊断的 AUC 分别为 0.92 和 0.90(图 6)。

讨 论

不同病原所致的儿童肺炎胸片表现有所不同,细菌性肺炎以肺部实质性改变为主,MPP 胸片可

表 3 验证集验证结果分类报告

分类	Pr	Rc	F1	支持图像数 (幅)
细菌组	0.87	0.82	0.85	50
MPP 组	0.82	0.74	0.78	50
病毒组	0.76	0.88	0.81	50
平均/总数	0.81	0.81	0.81	150

见支气管周围斑片状影、网状影及局灶性实变影,但是由于影像征象的重叠以及人类视觉能力的局限性,通过胸片无法对细菌性肺炎与 MPP 进行鉴别诊断。本研究建立了基于 VGG19 的胸部病变 X 线三分类 DTL 模型,并在验证集中对此模型的鲁棒性进行了进一步验证。通过对性能指标的评估和比较,结果显示 VGG19 模型对三组病变均有较好的分类效果,总体分类诊断 AUC 达 0.92。

Jia 等^[12]建立的 DTL 模型对胸片进行分类诊断,在训练集中的准确率达 99.6%,这说明了 DL 在辅助医学影像诊断中的巨大潜力。本研究结果显示,基于 VGG19 模型在训练集和测试集中均得到了较高的准

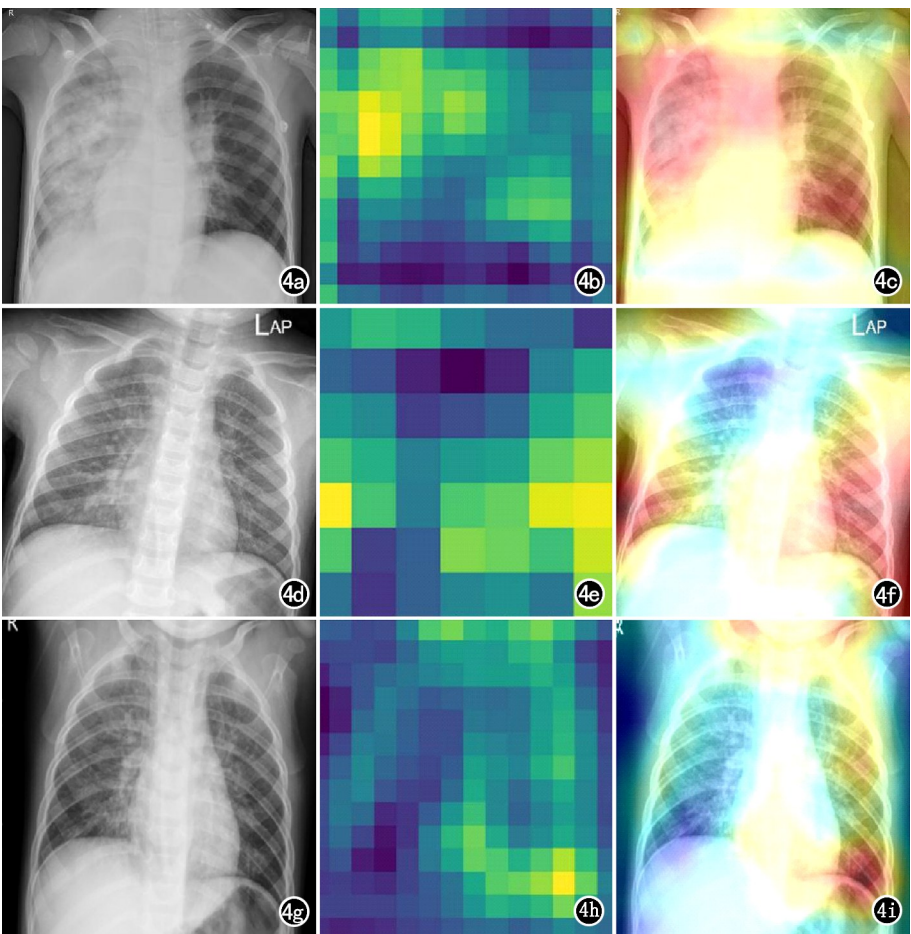


图 4 类激活热力图。a)细菌性肺炎的胸部 X 线图像;b)细菌性肺炎对应的类激活热力图;c)细菌性肺炎对应的类激活热力图融合图像;d)MPP 的胸部 X 线图像;e)MPP 对应的类激活热力图;f)MPP 对应的类激活热力图融合图像;g)病毒性肺炎的胸部 X 线图像;h)病毒性肺炎对应的类激活热力图 ;i)病毒性肺炎对应的类激活热力图融合图像。

确率,分别为 0.99 和 0.95,但在测试集中的损失值却有逐渐增高的趋势,这是模型过拟合现象。造成模型过拟合的主要原因有以下两点:首先为新任务的数据集较小,其次为所用的模型过于复杂。模型的过拟合是导致模型泛化能力差的重要原因。为了进一步验证本研究训练出的模型的泛化能力,我们又设置了验证集,而且此部分的数据图像没有在训练集和测试集中出现过。结果显示,在细菌组、MPP 组和病毒组的 AUC 分别为 0.92、0.90 和 0.93,而且总体 AUC 为 0.92。在临床实践中,放射科医生通过儿童胸部 X 片来对儿童肺炎类型进行分类诊断是比较困难的。这提示我们,DL 在辅助儿童胸部 X 片肺炎分类诊断方面的能力高于人工诊断。MPP 的早期正确诊断是 MPP 患儿得到及时有效治疗的前提,也是减少耐药性发生的重要基础。

此外,本研究还通过类激活热力图形式进一步展示了深度学习模型的过程可视化,通过类激活热力图

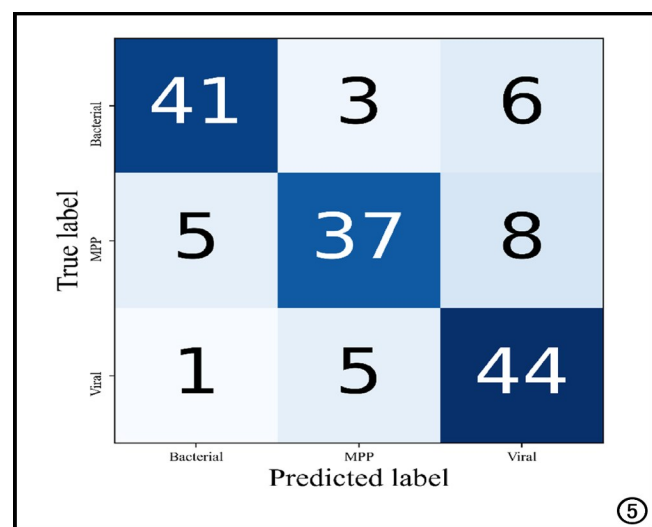


图5 VGG19 在验证集中的混淆矩阵图,图中的数字为模型正确分类的胸部 X 线片的数量。Bacteria:细菌组;MPP:MPP 组;Viral:病毒组

可直观地观察到卷积神经网络对不同图像的激活区域和方式各不相同,因此能够对细菌组、MPP 组和病毒组图像进行正确分类。这也为进行进一步的相关研究打下了基础。

本研究存在以下局限性:首先,本研究数据集 73.82% 为胸部后前位 X 线片,这导致网络对前后位胸片的学习不足,会影响模型的鲁棒性;其次,本研究数据为单中心,这也会影响模型的泛化能力。多中心来源的大样本数据集可以构建高泛化能力的深度学习模型^[13];再次,社区获得性肺炎中除细菌性肺炎及肺炎支原体肺炎以外,还有病毒性肺炎、真菌性肺炎及混合型感染肺炎^[14],本研究没有将后者纳入分析;最后,尚有其它深度学习模型可用于胸部 X 线片的分类诊断,这需要花费更多的资源和更多的精力。在此,笔者呼吁更多的专业技术人员参与这项研究。

综上所述,在通过胸部 X 线片对儿童细菌性肺炎、MPP 与病毒性肺炎的分类诊断中,基于 VGG19 的迁移学习模型是一种可靠的方法。该方法可以有效帮助放射科医生做出 MPP 诊断,为 MPP 的早期诊断奠定了基础,能够为临床医师制定合理的治疗策略提供可靠的循证医学证据,尤其是在实验室条件相对较差的医院将会起到更好的辅助诊断作用。

参考文献:

- [1] 中华医学会儿科学分会临床检验学组.儿童肺炎支原体呼吸道感染实验室诊断中国专家共识[J].中华检验医学杂志,2019,42(7):507-513.
- [2] Chen J,Zhang J,Lu Z,et al.Mycoplasma pneumoniae among Chinese Outpatient Children with Mild Respiratory Tract Infections during the Coronavirus Disease 2019 Pandemic[J].Microbiol Spectr,2022,10(1):e01550-01521.

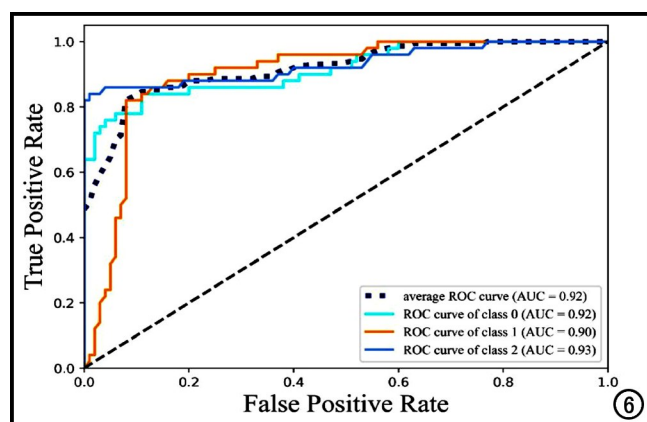


图6 VGG19 对三组图像分类诊断 ROC 曲线。average:模型对细菌组、MPP 组和病毒组 3 组图像分类的平均 ROC 曲线;class0:细菌组;class1:MPP 组;class2:病毒组。

- [3] 赵德育,邓欢.肺炎支原体肺炎实验室诊断的意义及进展[J].中华实用儿科临床杂志,2018,36(16):1205-1209.
- [4] Cho YJ,Han MS,Kim WS,et al.Correlation between chest radiographic findings and clinical features in hospitalized children with Mycoplasma pneumoniae pneumonia[J].PloS one,2019,14(8):e0219463.
- [5] Abumalloh RA,Nilashi M,Youssof Ismail M,et al.Medical image processing and COVID-19:A literature review and bibliometric analysis[J].J Infect Public Health,2022,15(1):75-93.
- [6] 朱文珍,胡琼洁.人工智能与医学影像融合发展:机遇与挑战[J].放射学实践,2019,34(9):938-941.
- [7] Thakur S,Kumar A.X-ray and CT-scan-based automated detection and classification of COVID-19 using convolutional neural networks (CNN)[J].Biomed Signal Process Control,2021,69(8):102920.
- [8] 戴正行,胡春洪,王希明,等.基于 DenseNet 网络深度学习法 CT 图像人工智能分析技术判断肺结节良恶性[J].放射学实践,2020,35(4):484-488.
- [9] 中华医学会儿科学分会呼吸学组.《中华实用儿科临床杂志》编辑委员会.儿童肺炎支原体肺炎诊治专家共识(2015 年版)[J].中华实用儿科临床杂志,2015,30(17):1304-1308.
- [10] 中华医学会儿科学分会呼吸学组.儿童社区获得性肺炎管理指南(2013 修订)(上)[J].中华儿科杂志,2013,51(10):745-752.
- [11] Loey M,El-Sappagh S,Mirjalili S.Bayesian-based optimized deep learning model to detect COVID-19 patients using chest X-ray image data[J].Comput Biol Med,2022,142(3):105213.
- [12] Jia G,Lam HK,Xu Y.Classification of COVID-19 chest X-ray and CT images using a type of dynamic CNN modification method[J].Comput Biol Med,2021,134(7):104425.
- [13] 安超,张晨,郑广平,等.基于深度学习的胸部 X 线肺结核检测研究及多中心临床验证[J].放射学实践,2022,37(6):704-709.
- [14] 赵登峰,徐慧,陈恒,等.不同类型病原体感染社区获得性肺炎临床与影像特征及转归[J].中华医院感染学杂志,2021,31(18):2768-2772.

(收稿日期:2022-09-26 修回日期:2023-02-16)

· 影像组学与人工智能专题 ·

影像组学在胰腺炎中的应用研究进展

王莎, 张艳利, 陈圆圆, 马芹芹, 雷军强

【摘要】 影像组学领域在过去几年中发展迅速,影像组学代表了以高通量方式提取定量特征的过程,该过程是将影像图像转换为高维数据并对数据进行分析以用于决策支持。影像组学能够以一种非侵入性的方式深入捕捉肉眼无法观察到的大量关于病变异质性的信息,有望弥补传统影像学局限于形态学和解剖学评估的不足。本文基于 CT 和 MRI 的影像组学在胰腺炎诊治中的应用进展进行综述。

【关键词】 胰腺炎; 影像组学; 体层摄影术, X 线计算机; 磁共振成像

【中图分类号】 R576; R445.2; R814.42 **【文献标识码】** A

【文章编号】 1000-0313(2024)01-0042-07

DOI:10.13609/j.cnki.1000-0313.2024.01.008

开放科学(资源服务)标识码(OSID):



胰腺炎是最常见的胰腺外分泌疾病^[1],其中急性胰腺炎(acute pancreatitis, AP)最为常见,全球发病率为 34/100,000 人^[1,2]。2017 年全球约有 160 万新发 AP 患者,其中约 10 万人死亡^[3]。慢性胰腺炎(chronic pancreatitis, CP)是由于各种原因导致的胰腺局部或弥漫性的慢性进展性炎症,对于胰腺内外分泌功能具有不可逆的损伤。AP 有转变为 CP 的可能,而 CP 会增加胰腺癌(pancreatic carcinoma, PC)的发生风险。因此,对胰腺炎的早期发现对临床干预有重要的指导意义。鉴于此类良性疾病更推荐无创检查,CT 及 MRI 等在胰腺炎的诊断、严重程度、并发症及鉴别诊断等方面发挥了重要作用。然而,传统影像学主要反映病变的大小、形态及强化方式,对于疾病的异质性反映不佳。影像组学在过去几年中发展迅速,能够以一种非侵入性的方式深入捕捉肉眼无法观察到的大量关于病变异质性的信息,然后将许多患者的成像数据与复杂的生物信息学工具相结合,开发出一种能够显著提高诊断和预后预测准确性的模型。即使疾病早期的传统影像学表现不明显,不同严重程度胰腺炎患者的实质改变也存在差异,这种差异也可能通过影像组学特征反映出来,所以影像组学有望弥补传统影像学的不足。

影像组学概述

影像组学是由 Lambin 等^[4]提出用于医学图像分析的方法,能够从图像中提取大量的、肉眼无法识别的

多种图像特征。现已广泛应用于超声、CT、MRI 及正电子发射计算机断层显像(positron emission tomography, PET)等医学图像的分析。其一般流程包括:①图像获取;②图像分割:通常通过手动或半自动地在图像上绘制感兴趣区来执行;③特征提取及筛选:影像组学特征包括形态(语义)、一阶(直方图)、纹理和高阶特征^[5-7]。由于影像组学特征的高维性,使用这些特征建立的模型可能会受到过拟合的影响。此外,影像组学特征之间存在高度相关性,可能会产生多重共线性问题。因此,要建立稳健的决策或预测模型,特征选择是必不可少的。常用的特征选择方法包括基于统计方法的逐步特征选择或最小绝对收缩和选择算子等自动算法^[7-9];④数据分析:即使在特征选择过程之后,仍会有大量的影像组学特征被选择,而不同于典型临床研究所评估的有限数量的变量。传统的统计方法可能不适合处理大量的影像组学特征,通常使用机器学习方法或正则化函数回归分析来构建影像组学分类器或模型^[8,9]。在模型构建之后,使用内部或外部测试数据验证模型进而用于疾病的诊断及预测等。

影像组学在 AP 中的应用

1. 预测 AP 复发

随着 AP 诊治指南的完善及临床救治水平的日趋成熟,AP 在治疗后基本可以痊愈,但 AP 转归后容易复发,研究报道 10%~30% 的 AP 患者首次发作后会复发进而转变为复发性急性胰腺炎(recurrent acute pancreatitis, RAP)^[10],而且 AP 反复发作会增加进展为 CP 及 PC 的风险^[11,12]。这不仅会导致临床治疗难度增加,还会给患者带来身心及经济负担,成为临床治疗的困扰,因此对 AP 复发的预测至关重要。近年来临床上主要探讨了 RAP 的危险因素^[13-15]以期通过危

作者单位: 730000 兰州,兰州大学第一临床医学院(王莎、马芹芹、陈圆圆);730000 兰州,兰州大学第一医院放射科(张艳利、雷军强)

作者简介: 王莎(1998—),女,甘肃武威人,硕士研究生,主要从事腹部影像诊断工作。

通讯作者: 雷军强, E-mail: leijq2011@126.com

基金项目: 甘肃省放射影像医学临床医学研究中心基金资助项目(20JR10FA668)

险因素的去预防 RAP 的发生,虽然此类措施达到了一定效果,但缺乏特异性,加之高达 30% 的 RAP 病例的病因仍然未知^[16],难以进行精准预测及预防。传统影像学对于 RAP 则主要是复发后的诊断,AP 两次发作间期在传统影像学上表现通常是正常的。因此,仅从传统影像学及临床角度预测 RAP 存在一定困难。有研究显示,从 AP 到 RAP 的发展存在一个连续变化的过程,从而产生不可逆转的胰腺结构或功能变化^[17]。如果可以在 AP 发生后检测到此病理变化的过程,则可以在 AP 发生后便进行干预以预防 RAP 的发生,可以提取图像定量信息及揭示病变异质性的影像组学应用而生。

Chen 等^[10]从 389 例首次出现 AP 患者的 CT 动脉和静脉期图像中提取了 412 个影像组学特征,最终选择了 10 个特征来建立影像组学预测模型,平均临床随访时间为 (62.8 ± 6.7) 个月,以确定 AP 是否复发,并建立临床模型、影像组学结合临床特征模型(radiomics combined with clinical feature model, COMB)以探讨最佳模型。结果显示临床特征中高脂血症是其最主要的危险因素,这与之前的研究结果一致^[18-20],然而当作者建立包含临床特征的临床模型预测 AP 复发时,发现高脂血症在两个队列中的预测能力均有限。相比之下影像组学模型的预测效能远高于临床模型,COMB 模型可以进一步提高影像组学模型的预测能力。此结果表明影像组学是预测 RAP 的潜在定量工具,相比临床模型,影像组学更加简单、精准,特异性更强。

考虑到 CT 具有辐射性且提供的信息有限,而 MRI 具有无辐射及软组织分辨率高等优点,Hu 等^[21]建立了基于 T_2WI 的影像组学模型预测 RAP,并比较了影像组学模型、临床模型和 COMB 模型对 RAP 的预测能力,最后建立评估 AP 患者复发风险的诺模图。结果显示影像组学模型和 COMB 模型的曲线下面积(areas under curve, AUC)均大于临床模型($P = 0.008, 0.007$),但影像组学模型与 COMB 模型之间的 AUC 差异无统计学意义($P = 0.067$)。此研究同样发现高脂血症是 RAP 的危险因素,但影响 AP 复发的临床因素有很多,仅凭临床特征很难预测 RAP;而且作者在建立诺模图时发现影像组学特征对 AP 复发风险的预测比例大于临床特征,提示影像组学特征在预测 AP 复发方面优于独立的临床特征。因此,基于 CT 及 MRI 的影像组学均可以很好地预测 AP 的复发情况^[22]。此外,基于 T_2WI 的影像组学还可以清楚显示胰腺结构的解剖范围及其炎症范围,并且可以结合其他参数判断胰腺的出血水肿情况,为临床精准治疗提供很大的帮助^[23]。值得注意的是,与 CT 相比,MRI

的检查时间更长、费用更高,基于 MRI 的影像组学模型带来的附加价值应该与其相对较大的资源消耗相平衡,才是有利于临床有益于患者的理想模型。然而,目前基于 MRI 所建立的模型似乎并未显示出明显的优势,这可能与其序列选择之间存在一定关系,所有基于 MRI 的模型均只纳入 T_2WI 图像,未将其他序列纳入进行对比分析,如可以提供定量参数的扩散加权成像(diffusion weighted imaging, DWI)序列,未来的研究可以纳入更多序列进行对比探讨,以期建立最佳影像组学预测模型。

2. 预测评估 AP 严重程度

根据国际共识 2012 年修订的亚特兰大分类和定义(2012-RACD),AP 可分为三类,按其临床严重程度分为轻度急性胰腺炎(mild acute pancreatitis, MAP)、中度重度急性胰腺炎(moderate severe acute pancreatitis, MSAP)和重度急性胰腺炎(severe acute pancreatitis, SAP)^[24]。文献报道 SAP 的死亡率高达 36%~50%,持续性器官衰竭并发感染性坏死患者的死亡率非常高^[24]。因此,对 AP 严重程度的早期预测至关重要,这不仅有利于 MSAP 和 SAP 患者的早期诊断和治疗,也有利于 MSAP 和 SAP 患者的早期转诊。目前,早期预测 AP 严重程度的方法主要基于临床特征、实验室检查以及传统影像学检查结果^[25-28]。但由于 AP 的复杂性和多样性,其病程仍难以预测^[29]。

虽然几种传统的评分系统可以预测 SAP^[30,31],但它们都是复杂且耗时的^[32],如临床最常用的预测评分系统是急性生理学和慢性健康检查(APACHE II)^[33,34],此评分系统需要收集大量临床数据,在临床中较难实现。针对这一复杂的评分系统,有学者提出了 BISAP 评分^[35],此评分相对简单,但存在主观评价的指标即对精神状态的评估,容易导致一定的误差。基于 CT 的评分系统 CTSI、MCTSI 均存在类似的缺陷^[36],对于炎症和坏死的评估依赖影像科医师的肉眼评估。此外,有研究表明 MAP 在发病的最初几小时内伴有高血流灌注,而 SAP 则伴有进行性组织缺血和血流灌注减少^[37],即 AP 早期即存在微循环障碍及其导致的组织损伤,然而传统影像学胰腺形态改变往往并不明显,尤其是胰腺的坏死,这可能会低估疾病的严重程度^[38,39]。影像组学可以挖掘隐藏在常规图像中的信息,从而揭示潜在的生物学基础^[40],并非依赖其图像形态改变进行评估,可以弥补传统影像学滞后性的缺陷。

Lin 等^[41]在一项 259 例 AP 患者的回顾性研究中报道,基于增强 MRI 的影像组学特征可以预测 AP 的严重程度。基于 MRI 门静脉期图像,Lin 等从整个胰腺感兴趣区提取了 353 个影像组学特征,最后选择了

11 个特征来建立支持向量机模型。在训练组中,影像组学模型诊断的敏感度、特异度、阳性预测值(positive prediction value, PPV)、阴性预测值(negative prediction value, NPV)、准确度和 AUC 分别为 77.8%、91.9%、88.7%、83.5%、85.6% 和 0.917;在验证组中,相应指标分别为 75.0%、86.0%、81.8%、80.4%、81.0% 和 0.848。影像组学的 AUC 均明显高于现有的一些临床和影像学评分系统(所有 P 值均 <0.05)。该研究表明与现有的一些临床和影像学评分系统相比,影像组学模型在预测 AP 的严重程度方面可能更准确。影像组学可以挖掘出不同于传统影像的定量信息,在 AP 严重程度的分类中具有广阔的应用前景。值得注意的是,目前基于影像组学在预测 AP 严重程度的研究中^[42],均未探讨 AP 患者的持续性器官衰竭,而持续性器官衰竭是 AP 患者死亡的重要因素。因此,在以后的研究中预测 AP 患者是否会发生持续性器官衰竭是一个值得研究的方向,存在重要的临床指导意义。此外,此类研究只探讨了影像组学模型的预测能力,均未结合临床因素进行预测,未来的研究中将临床因素及影像组学特征相结合,或许可以提供一种更加精准的预测模型。

3. 预测 AP 并发症

根据国际共识的 2012-RACD^[24],AP 局部并发症包括 4 种,分别为急性胰周液体积聚(acute peripancreatic fluid collection, APFC)、胰腺假性囊肿、急性坏死性积聚及包裹性坏死。2012-RACD 推荐判断 AP 局部并发症的影像学表现是基于增强 CT 图像。有学者认为 MRI 对 AP 并发症的显示更佳^[43],DWI 作为 MRI 检查的一部分,对组织成分各个方面的变化都很敏感,而且可以提供一个定量指标即表观扩散系数(apparent diffusion coefficient, ADC)值。先前的研究表明 ADC 值与 AP 所致的肝损害之间存在关联,而且并发肝损害组中不同严重程度 AP 患者的肝脏 ADC 值差异有统计学意义,随着 AP 严重程度的增加 ADC 值降低^[44]。Iranmahboob 等^[45]对 41 例 AP 患者整个胰腺的 ADC 值进行了直方图分析,结果显示虽然 ADC 值差异无统计学意义,但 ADC 直方图的一些一阶特征与新并发症的发生之间存在关联,包括新的小室积液、胰腺坏死、静脉血栓或动脉假性动脉瘤,因此,ADC 直方图特征可以提供传统 ADC 值无法提供的组织异质性信息。其中峰度预测新并发症的敏感度为 75.0%,特异度为 91.9%,AUC 为 0.784,但出现并发症的患者峰度升高的确切依据尚不清楚。有推测认为峰度的增加是由于信号强度异质性的增加,这反映了潜在的炎症和黏液实质改变,这些改变可能使患者容易发生并发症。

在 AP 早期阶段,通常是指在发病后第一周结束,但可能会延续到第二周^[24],此阶段形态学变化与并发症严重程度不成正比,尤其是前 3 天内,AP 早期的胰腺外坏死(extrapaneatronic necrosis, EXPN)和 APFC 在影像上均表现为液体信号或密度,仅通过传统影像学很难鉴别。Zhou 等^[46]建立基于 T_2 WI 图像胰腺外周聚集和胰腺实质 MRI 增强动脉晚期图像的影像组学模型,以预测 AP 早期 EXPN,并比较影像组学模型、临床模型及传统影像学评分系统的预测效能。结果显示影像组学模型的预测效能均高于临床模型及传统影像学评分系统,基于胰腺外周聚集 T_2 WI 的影像组学模型对 AP 早期 EXPN 的预测效能最高,训练组和验证组的 AUC 均大于 0.950,其次是基于胰腺实质的影像组学模型,AUC 也超过 0.900。究其原因,影像组学可以提供疾病早期常规图像中不可见的潜在信息,从复杂的内容物中提取信息进行鉴别,从而在临床及传统影像学均未做出诊断时预测 AP 并发症的发生。

AP 全身性并发症包括器官功能衰竭、胰性脑病及脓毒症等^[47],其中以器官功能衰竭最为严重,如肾衰竭及呼吸衰竭等。文献报道 AP 合并急性肾损伤会导致死亡率的增高^[48]。因此,AP 患者在肾小管功能受损之前的精准诊断极为重要。然而,AP 合并急性肾损伤的传统影像学表现并无特异性,仅表现为肾脏皮髓质分界不清。临床常用的预测指标血清肌酐在肾小球滤过率减少 50% 才会出现异常,很容易错过治疗的最佳时间窗。因此,基于临床及传统影像学均很难预测 AP 合并急性肾损伤。纹理分析作为影像组学的技术手段之一,通过分析图像中像素或体素灰度级的分布和关系,提供对病灶异质性的客观、定量评估^[49]。贾颖等^[50]探讨了基于增强 CT 的纹理分析技术对 AP 合并急性肾损伤的诊断价值,从动脉期及静脉期分割出的双侧肾脏三维容积感兴趣区进行纹理分析,共提取出 396 个纹理特征,最终纳入 6 个纹理特征,其中以平均像素值的诊断效能最佳,其诊断准确度、特异度及 AUC 分别为 93.3%、100% 及 0.983。此研究中筛选出的纹理特征与其传统影像表现相对应,如不合并急性肾损伤的患者局部灰度相关性较大,相关性是反映病灶局部像素灰度差异性大小的纹理特征,相关性越大差异越小,相关性较大在传统影像学上则表现为图像清晰,边界清楚。而并发急性肾损伤时皮髓质强化减低,边界模糊,纹理分析特征可见其成分不均,局部差异大。因此,纹理分析有望成为诊断 AP 合并急性肾损伤的影像标记物,可在血清肌酐达到现有临床诊断标准之前诊断急性肾损伤,以期在治疗时间窗内及时治疗以防病情恶化。然而,由于目前此类研究数量较

少,纳入研究的纹理特征只是众多纹理特征中的一部分,未来的研究有望探讨更加多样的纹理特征进而提供更多信息。

影像组学在 CP 中的应用

1. CP 与 PC 的鉴别诊断

肿块形成性胰腺炎(mass-forming pancreatitis, MFP)是一种特殊类型的 CP,包括自身免疫性胰腺炎(autoimmune pancreatitis, AIP)及肿块形成性慢性胰腺炎(mass-forming chronic pancreatitis, MFCP)等^[51,52]。MFP 是一种良性疾病,但其临床表现及传统影像学表现与胰腺导管腺癌(pancreatic ductal adenocarcinoma, PDAC)很难鉴别。有研究发现 PDAC 与 MFCP 的胰腺扩张程度及肿块大小并无明显差异,基于肿块大小及和胰管扩张鉴别两者的 AUC 仅为 0.697 和 0.589~0.622^[53,54],鉴别效能不佳。然而两者的治疗方式存在很大的差异,术前误诊常导致患者病情延误或治疗不当。相关研究表明有 1%~6% 的 PDAC 患者被误诊为 MFP 而延误治疗,5%~10% 的 MFP 患者误诊为 PDAC 而进行了胰十二指肠切除术,这会导致胰腺功能不全^[54-57]。虽然穿刺活检发挥了一定作用,但由于穿刺标本大小有限,即使是超声引导的活检也显示出 12%~14% 的假阴性^[51],而且穿刺活检为有创检查。

Ma 等^[58]探讨了增强 CT 影像组学模型在鉴别 MFP 与 PC 中的作用,研究纳入了 151 例 PC 患者和 24 例 MFP 患者(18 例 AIP 患者和 6 例 MFCP),分别建立动脉期影像组学模型、静脉期影像组学模型、动脉期结合静脉期影像组学模型、临床特征模型、COMB 模型进行鉴别,结果显示上述 5 种模型的 AUC 值分别为 0.905、0.941、0.941、0.822 和 0.980。COMB 模型的 AUC 明显高于动脉期影像组学模型($P=0.004$)和临床特征模型($P<0.001$),鉴别诊断效果最佳。动脉期结合静脉期影像组学模型与静脉期影像组学模型的 AUC 值一致。COMB 模型的敏感度和特异度分别达 0.947 和 0.917。因此,COMB 模型可以成为鉴别 PC 与 CP 的潜在工具,从而有助于临床决策。影像组学特征反映的是病变的病理变化及图像的异质性,PC 的病理特征主要是癌细胞、间质星形细胞和间质纤维化,病变间质纤维化阻碍了对比剂的穿透,导致 PC 增强扫描均表现为低密度。胰腺炎的病理特征主要是炎症细胞浸润和纤维组织增生,增强程度与纤维组织相似,动脉期轻度强化,静脉期和延迟期强化^[59]。虽然 PC 和 MFP 在动脉期中都呈低密度,但静脉期和延迟期的表现有助于鉴别这些肿瘤。因此,静脉期的影像组学特征比动脉期更有诊断价值。在此研究中,静脉期

影像组学模型的 AUC 高于动脉期组,虽然两组的 AUC 差异无统计学意义,但该模型的敏感性和准确性都高于动脉期模型,说明静脉期图像提取特征的诊断性能优于动脉期模型。Deng 等^[60]基于增强 MRI 建立影像组学模型同样显示静脉期模型的 AUC 高于动脉期模型,分别为 0.997 和 0.958。然而张晶晶等^[61]的研究显示动脉期模型的 AUC(0.801)高于静脉期(0.769),导致此差异的原因可能是由于该研究探讨的是胰头部的 PC 与 MFCP 的鉴别,而且肿瘤之间本身存在异质性及扫描仪器之间也会存在差异所致。未来的研究或许可以分部位进行对比探讨,以进一步验证存在差异的原因,也有助于更加精准地进行鉴别。

2. 区分功能性腹痛、RAP 及 CP

RAP 是指两次或两次以上的 AP 发作,但疾病发作间期病变完全恢复正常。AP 转归后患者再次出现腹痛和胰酶升高时,即使影像学表现正常,临床也可以诊断为 RAP。然而,腹痛及胰酶升高并不是 RAP 的特异性表现,当 AP 转变为 CP 时也可出现腹痛及胰酶升高,加之 AP 在发作缓解后传统影像学表现通常是正常的,导致 AP 转归后患者出现腹痛时的诊断存在一定挑战。虽然 RAP 与 CP 可以通过病理变化鉴别,RAP 主要是胰腺的出血水肿及坏死,而 CP 主要病理变化是其纤维化,但考虑到病理活检的有创性,并不推荐用于良性疾病的诊断。影像组学可以很好地从图像中挖掘出病理信息,或许是此类患者的福音。Mashayekhi 等^[62]探讨了基于 CT 影像组学特征鉴别功能性腹痛、RAP 与 CP,研究纳入 56 例患者(RAP 20 例,功能性腹痛 19 例,CP 17 例),建立支持向量机模型进行分类预测。结果显示模型的预测准确率为 82.1%,虽然存在 CP 误诊为 RAP 的病例,但不存在 CP 及 RAP 误诊为功能性腹痛的病例。因此,对于仅根据腹痛症状和实验室检查无法确诊为 RAP 的患者,影像组学可能是有用的辅助诊断手段。

不足与展望

影像组学在胰腺炎的诊治中显现出极大的潜力。然而,目前研究相对较少,且大多数研究为单中心回顾性研究,这会导致预测诊断模型的泛化及不稳定性增加,未来的研究有望采用前瞻性多中心研究来提高模型的稳定性,同时对其适用性及普遍性进行评估。此外,基于 MRI 建立的模型都未进行多序列比较,只纳入部分序列进行建模,尤其对于 AP 的 MRI 模型中只有一项研究纳入了 DWI 进行建模。对于胰腺炎入院常规更推荐成本低、时间短的 CT 检查。使用高成本的 MRI 进行建模,是否得到了与之相匹配的效益,是模型应用于临床需要考虑的问题。因此,未来基于

MRI 的研究可以纳入更多参数以挖掘其更高的价值,而非局限于常规序列。

目前,随着影像技术及人工智能的不断发展,各种新技术应用于临床,如能量 CT、深度学习等都在疾病诊疗中显示出了极大的潜力。人工智能与影像技术的结合也被逐步应用于临床^[63,64],并显示出良好的效果。在未来的研究中,有效地将各种新技术相结合可能是潜在的方向^[5,6,65-69]。

参考文献:

- [1] Xiao AY, Tan ML, Wu LM, et al. Global incidence and mortality of pancreatic diseases: a systematic review, meta-analysis, and meta-regression of population-based cohort studies[J]. *Lancet Gastroenterol Hepatol*, 2016, 1(1): 45-55.
- [2] Boxhoorn L, Voermans RP, Bouwense SA, et al. Acute pancreatitis [J]. *Lancet*, 2020, 396(10252): 726-734.
- [3] GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017 [J]. *Lancet*, 2018, 392(10159): 1736-1788.
- [4] Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine[J]. *Nat Rev Clin Oncol*, 2017, 14(12): 749-762.
- [5] Parakh A, Tirkes T. Advanced imaging techniques for chronic pancreatitis[J]. *Abdom Radiol (NY)*, 2020, 45(5): 1420-1438.
- [6] Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping[J]. *Radiology*, 2020, 295(2): 328-338.
- [7] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data[J]. *Radiology*, 2016, 278(2): 563-577.
- [8] Park HJ, Park B, Lee SS. Radiomics and deep learning: hepatic applications[J]. *Korean J Radiol*, 2020, 21(4): 387-401.
- [9] Park JE, Park SY, Kim HJ, et al. Reproducibility and generalizability in radiomics modeling: possible strategies in radiologic and statistical perspectives[J]. *Korean J Radiol*, 2019, 20(7): 1124-1137.
- [10] Chen Y, Chen TW, Wu CQ, et al. Radiomics model of contrast-enhanced computed tomography for predicting the recurrence of acute pancreatitis[J]. *Eur Radiol*, 2019, 29(8): 4408-4417.
- [11] Sankaran SJ, Xiao AY, Wu LM, et al. Frequency of progression from acute to chronic pancreatitis and risk factors: a meta-analysis[J]. *Gastroenterology*, 2015, 149(6): 1490-1500. e1491.
- [12] Gandhi S, de la Fuente J, Murad MH, et al. Chronic pancreatitis is a risk factor for pancreatic cancer, and incidence increases with duration of disease: a systematic review and meta-analysis[J]. *Clin Transl Gastroenterol*, 2022, 13(3): e00463.
- [13] Magnusdottir BA, Baldursdottir MB, Kalaitzakis E, et al. Risk factors for chronic and recurrent pancreatitis after first attack of acute pancreatitis[J]. *Scand J Gastroenterol*, 2019, 54(1): 87-94.
- [14] Yu B, Li J, Li N, et al. Progression to recurrent acute pancreatitis after a first attack of acute pancreatitis in adults[J]. *Pancreatol*, 2020, 20(7): 1340-1346.
- [15] Sun Y, Jin J, Zhu A, et al. Risk factors for recurrent pancreatitis after first episode of acute pancreatitis[J]. *Int J Gen Med*, 2022, 15(9): 1319-1328.
- [16] Testoni PA. Acute recurrent pancreatitis: etiopathogenesis, diagnosis and treatment[J]. *World J Gastroenterol*, 2014, 20(45): 16891-16901.
- [17] Whitcomb DC. Genetic risk factors for pancreatic disorders[J]. *Gastroenterology*, 2013, 144(6): 1292-1302.
- [18] Vippera K, Somerville C, Furlan A, et al. Clinical profile and natural course in a large cohort of patients with hypertriglyceridemia and pancreatitis[J]. *J Clin Gastroenterol*, 2017, 51(1): 77-85.
- [19] Qiu L, Sun RQ, Jia RR, et al. Comparison of existing clinical scoring systems in predicting severity and prognoses of hyperlipidemic acute pancreatitis in chinese patients: a retrospective study [J]. *Medicine*, 2015, 94(23): e957.
- [20] Huang YX, Jia L, Jiang SM, et al. Incidence and clinical features of hyperlipidemic acute pancreatitis from Guangdong, China: a retrospective multicenter study[J]. *Pancreas*, 2014, 43(4): 548-552.
- [21] Hu Y, Liu N, Tang L, et al. Three-dimensional radiomics features of magnetic resonance T₂-weighted imaging combined with clinical characteristics to predict the recurrence of acute pancreatitis [J]. *Front Med (Lausanne)*, 2022, 9(10): 777368.
- [22] 胡云涛, 黄小华, 刘念, 等. 基于磁共振 T₂WI 序列影像组学预测急性胰腺炎复发的价值[J]. *磁共振成像*, 2021, 12(10): 12-15, 21.
- [23] 彭铮堃, 吴娜珊, 陈馨, 等. 基于磁共振 T₂WI 序列影像组学预测急性胰腺炎复发的价值研究[J]. *影像研究与医学应用*, 2022, 6(8): 101-103.
- [24] Banks PA, Bollen TL, Dervenis C, et al. Classification of acute pancreatitis—2012: revision of the Atlanta classification and definitions by international consensus[J]. *Gut*, 2013, 62(1): 102-111.
- [25] Kuo DC, Rider AC, Estrada P, et al. Acute pancreatitis: what's the score? [J]. *J Emerg Med*, 2015, 48(6): 762-770.
- [26] van den Berg FF, de Bruijn AC, van Santvoort HC, et al. Early laboratory biomarkers for severity in acute pancreatitis: A systematic review and meta-analysis[J]. *Pancreatol*, 2020, 20(7): 1302-1311.
- [27] Yan G, Li H, Bhetuwal A, et al. Pleural effusion volume in patients with acute pancreatitis: a retrospective study from three acute pancreatitis centers[J]. *Ann Med*, 2021, 53(1): 2003-2018.
- [28] Zhou T, Chen Y, Wu JL, et al. Extra pancreatic inflammation on magnetic resonance imaging for the early prediction of acute pancreatitis severity[J]. *Pancreas*, 2020, 49(1): 46-52.
- [29] Zerem E. Treatment of severe acute pancreatitis and its complications[J]. *World J Gastroenterol*, 2014, 20(38): 13879-13892.
- [30] Kim BG, Noh MH, Ryu CH, et al. A comparison of the BISAP score and serum procalcitonin for predicting the severity of acute pancreatitis[J]. *Korean J Intern Med*, 2013, 28(3): 322-329.
- [31] Wan J, Shu W, He W, et al. Serum creatinine level and APACHE-II score within 24th of Admission are effective for predicting persistent organ failure in acute pancreatitis[J]. *Gastroenterol Res Pract*, 2019, 2019(10): 8201096.

- [32] Cho SK, Kim JW, Huh JH, et al. Atherogenic index of plasma is a potential biomarker for severe acute pancreatitis: a prospective observational study[J]. *J Clin Med*, 2020, 9(9):2982.
- [33] Yeung YP, Lam BY, Yip AW. APACHE system is better than Ranson system in the prediction of severity of acute pancreatitis[J]. *Hepatobiliary Pancreat Dis Int*, 2006, 5(2):294-299.
- [34] Iezzoni LI, Hotchkin EK, Ash AS, et al. MedisGroups databases. The impact of data collection guidelines on predicting in-hospital mortality[J]. *Med Care*, 1993, 31(3):277-283.
- [35] Wu BU, Johannes RS, Sun X, et al. The early prediction of mortality in acute pancreatitis: a large population-based study[J]. *Gut*, 2008, 57(12):1698-1703.
- [36] Banday IA, Gattoo I, Khan AM, et al. Modified computed tomography severity index for evaluation of acute pancreatitis and its correlation with clinical outcome: a tertiary care hospital based observational study[J]. *J Clin Diagn Res*, 2015, 9(8):TC01-TC05.
- [37] Knoefel WT, Kollias N, Warshaw AL, et al. Pancreatic microcirculatory changes in experimental pancreatitis of graded severity in the rat[J]. *Surgery*, 1994, 116(5):904-913.
- [38] Shinagare AB, Ip IK, Raja AS, et al. Use of CT and MRI in emergency department patients with acute pancreatitis[J]. *Abdom Imaging*, 2015, 40(2):272-277.
- [39] Spanier BW, Nio Y, van der Hulst RW, et al. Practice and yield of early CT scan in acute pancreatitis: a dutch observational multicenter study[J]. *Pancreatol*, 2010, 10(2-3):222-228.
- [40] Grossmann P, Stringfield O, El-Hachem N, et al. Defining the biological basis of radiomic phenotypes in lung cancer[J]. *Elife*, 2017, 6(21):e23421.
- [41] Lin Q, Ji YF, Chen Y, et al. Radiomics model of contrast-enhanced MRI for early prediction of acute pancreatitis severity[J]. *J Magn Reson Imaging*, 2020, 51(2):397-406.
- [42] 范海云, 陈基明, 陈亮亮, 等. 基于胰腺周围脂肪间隙 CT 影像组学预测早期急性胰腺炎进展的价值[J]. *放射学实践*, 2022, 37(6):683-689.
- [43] Xiao B, Zhang XM, Tang W, et al. Magnetic resonance imaging for local complications of acute pancreatitis: a pictorial review[J]. *World J Gastroenterol*, 2010, 16(22):2735-2742.
- [44] 杨茹, 胡富碧, 唐伟, 等. 急性胰腺炎肝损害的 DWI 研究[J]. *放射学实践*, 2015, 30(3):255-259.
- [45] Iranmahboob AK, Kierans AS, Huang C, et al. Preliminary investigation of whole-pancreas 3D histogram ADC metrics for predicting progression of acute pancreatitis[J]. *Clin Imaging*, 2017, 42(3-4):172-177.
- [46] Zhou T, Xie CL, Chen Y, et al. Magnetic resonance imaging-based radiomics models to predict early extrapancreatic necrosis in acute pancreatitis[J]. *Pancreas*, 2021, 50(10):1368-1375.
- [47] 杜奕奇, 陈其奎, 李宏宇, 等. 中国急性胰腺炎诊治指南(2019 年, 沈阳)[J]. *临床肝胆病杂志*, 2019, 35(12):2706-2711.
- [48] Kes P, Vucicević Z, Ratković-Gusić I, et al. Acute renal failure complicating severe acute pancreatitis[J]. *Nephron*, 1996, 73(3):488-489.
- [49] Ganeshan B, Miles KA. Quantifying tumour heterogeneity with CT[J]. *Cancer Imaging*, 2013, 13(1):140-149.
- [50] 贾颖, 黄子星, 汪翊, 等. 基于 CT 增强图像纹理分析对急性胰腺炎合并急性肾损伤的诊断价值[J]. *中国普外基础与临床杂志*, 2019, 26(7):865-869.
- [51] Matsubayashi H, Matsui T, Yabuuchi Y, et al. Endoscopic ultrasonography guided-fine needle aspiration for the diagnosis of solid pancreaticobiliary lesions: clinical aspects to improve the diagnosis[J]. *World J Gastroenterol*, 2016, 22(2):628-640.
- [52] Lee H, Lee JK, Kang SS, et al. Is there any clinical or radiologic feature as a preoperative marker for differentiating mass-forming pancreatitis from early-stage pancreatic adenocarcinoma? [J]. *Hepatogastroenterology*, 2007, 54(79):2134-2140.
- [53] Ren S, Zhang J, Chen J, et al. Evaluation of texture analysis for the differential diagnosis of mass-forming pancreatitis from pancreatic ductal adenocarcinoma on contrast-enhanced CT images[J]. *Front Oncol*, 2019, 9(5):1171.
- [54] Sandrasegaran K, Nutakki K, Tahir B, et al. Use of diffusion-weighted MRI to differentiate chronic pancreatitis from pancreatic cancer[J]. *AJR Am J Roentgenol*, 2013, 201(5):1002-1008.
- [55] Balthazar EJ. Pancreatitis associated with pancreatic carcinoma. Preoperative diagnosis: role of CT imaging in detection and evaluation[J]. *Pancreatol*, 2005, 5(4-5):330-344.
- [56] Birgin E, Hablawetz P, Téoule P, et al. Chronic pancreatitis and resectable synchronous pancreatic carcinoma: A survival analysis[J]. *Pancreatol*, 2018, 18(4):394-398.
- [57] Kirkegård J, Mortensen FV, Cronin-Fenton D. Chronic pancreatitis and pancreatic cancer risk: a systematic review and meta-analysis[J]. *Am J Gastroenterol*, 2017, 112(9):1366-1372.
- [58] Ma X, Wang YR, Zhuo LY, et al. Retrospective analysis of the value of enhanced CT radiomics analysis in the differential diagnosis between pancreatic cancer and chronic pancreatitis[J]. *Int J Gen Med*, 2022, 15(6):233-241.
- [59] Liu Y, Wang M, Ji R, et al. Differentiation of pancreatic ductal adenocarcinoma from inflammatory mass: added value of magnetic resonance elastography[J]. *Clin Radiol*, 2018, 73(10):865-872.
- [60] Deng Y, Ming B, Zhou T, et al. Radiomics model based on MR images to discriminate pancreatic ductal adenocarcinoma and mass-forming chronic pancreatitis lesions[J]. *Front Oncol*, 2021, 11(3):620981.
- [61] 张晶晶, 李庆祝, 王建华, 等. 增强 CT 联合纹理分析鉴别胰头肿块型胰腺炎与胰头癌价值分析[J]. *中华医学杂志*, 2019, 105(33):2575-2580.
- [62] Mashayekhi R, Parekh VS, Faghih M, et al. Radiomic features of the pancreas on CT imaging accurately differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis[J]. *Eur J Radiol*, 2020, 123(2):108778.
- [63] Ohira S, Koike Y, Akino Y, et al. Improvement of image quality for pancreatic cancer using deep learning-generated virtual monochromatic images: Comparison with single-energy computed tomography[J]. *Phys Med*, 2021, 85(5):8-14.
- [64] Noda Y, Goshima S, Miyoshi T, et al. Assessing chemotherapeutic response in pancreatic ductal adenocarcinoma: histogram analysis of iodine concentration and CT number in single-source dual-energy CT[J]. *AJR Am J Roentgenol*, 2018, 211(6):1221-1226.
- [65] Ghandili S, Shayesteh S, Fouladi DF, et al. Emerging imaging techniques for acute pancreatitis[J]. *Abdom Radiol (NY)*, 2020, 45(5):1299-1307.

[66]

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial intelligence for the management of pancreatic diseases[J]. Dig Endosc, 2021, 33(2):231-241.

[67]

Goyal H, Mann R, Gandhi Z, et al. Application of artificial intelligence in pancreaticobiliary diseases[J]. Ther Adv Gastrointest Endosc, 2021, 14(2):2631774521993059.

[68]

Tong T, Gu J, Xu D, et al. Deep learning radiomics based on contrast-enhanced ultrasound images for assisted diagnosis of pancreatic ductal adenocarcinoma and chronic pancreatitis[J]. BMC Med, 2022, 20(1):74.

[69]

Ziegelmayr S, Kaissis G, Harder F, et al. Deep convolutional neural Network-assisted feature extraction for diagnostic discrimination and feature visualization in pancreatic ductal adenocarcinoma (PDAC) versus autoimmune pancreatitis (AIP)[J]. J Clin Med, 2020, 9(12):4013.

(收稿日期:2022-09-07 修回日期:2022-11-16)

本刊可直接使用的医学缩略语

医学论文中正确、合理使用专业名词可以精简文字,节省篇幅,使文章精炼易懂。现将放射学专业领域为大家所熟知的专业名词缩略语公布如下(按照英文首字母顺序排列),以后本刊在论文中将对这一类缩略语不再注释其英文全称和中文。

- ADC (apparent diffusion coefficient):表观扩散系数
- ALT:丙氨酸转氨酶;AST:天冬氨酸转氨酶
- BF (blood flow):血流量
- BOLD (blood oxygenation level dependent):血氧水平依赖
- BV (blood volume):血容量
- b:扩散梯度因子
- CAG (coronary angiography):冠状动脉造影
- CPR (curve planar reformation):曲面重组
- CR(computed radiography):计算机 X 线摄影术
- CT (computed tomography):计算机体层成像
- CTA (computed tomography angiography):CT 血管成像
- CTPI(CT perfusion imaging):CT 灌注成像
- DICOM (digital imaging and communication in medicine):医学数字成像和传输
- DR(digital radiography):数字化 X 线摄影术
- DSA (digital subtraction angiography):数字减影血管造影
- DWI (diffusion weighted imaging):扩散加权成像
- DTI (diffusion tensor imaging):扩散张量成像
- ECG (electrocardiography):心电图
- EPI (echo planar imaging):回波平面成像
- ERCP(endoscopic retrograde cholangiopancreatography):经内镜逆行胰胆管造影术
- ETL (echo train length):回波链长度
- FLAIR (fluid attenuation inversion recovery):液体衰减反转恢复
- FLASH (fast low angel shot):快速小角度激发
- FOV (field of view):视野
- FSE (fast spin echo):快速自旋回波
- fMRI (functional magnetic resonance imaging):功能磁共振成像
- IR (inversion recovery):反转恢复
- Gd-DTPA:钆喷替酸葡甲胺
- GRE (gradient echo):梯度回波

- HE 染色:苏木素-伊红染色
- HRCT(high resolution CT):高分辨率 CT
- MPR (multi-planar reformation):多平面重组
- MIP (maximum intensity projection):最大密(强)度投影
- MinIP (minimum intensity projection):最小密(强)度投影
- MRA (magnetic resonance angiography):磁共振血管成像
- MRI (magnetic resonance imaging):磁共振成像
- MRS (magnetic resonance spectroscopy):磁共振波谱学
- MRCP(magnetic resonance cholangiopancreatography):磁共振胰胆管成像
- MSCT (multi-slice spiral CT):多层螺旋 CT
- MTT (mean transit time):平均通过时间
- NEX (number of excitation):激励次数
- PACS (picture archiving and communication system):图像存储与传输系统
- PC (phase contrast):相位对比法
- PET (positron emission tomography):正电子发射计算机体层成像
- PS (surface permeability):表面通透性
- ROC 曲线(receiver operating characteristic curve):受试者操作特征曲线
- SPECT (single photon emission computed tomography):单光子发射计算机体层摄影术
- PWI (perfusion weighted imaging):灌注加权成像
- ROI (region of interest):兴趣区
- SE (spin echo):自旋回波
- STIR(short time inversion recovery):短时反转恢复
- TACE(transcatheter arterial chemoembolization):经导管动脉化疗栓塞术
- T₁ WI (T₁ weighted image):T₁ 加权像
- T₂ WI (T₂ weighted image):T₂ 加权像
- TE (time of echo):回波时间
- TI (time of inversion):反转时间
- TR (time of repetition):重复时间
- TOF (time of flight):时间飞跃法
- TSE (turbo spin echo):快速自旋回波
- VR (volume rendering):容积再现
- WHO (World Health Organization):世界卫生组织
- NAA(N-acetylaspartate):N-乙酰天门冬氨酸
- Cho(choline):胆碱
- Cr(creatine):肌酸

(本刊编辑部)