

AI+多模态融合在肺结节良恶性鉴别中的应用



童琳^{1,2,3,4}, 白春学^{1,2,3,4*}

1. 复旦大学附属中山医院呼吸与危重症医学科, 上海 200032
2. 上海市呼吸物联网医学工程技术研究中心, 上海 200032
3. 上海市呼吸病研究所, 上海 200032
4. 复旦大学附属中山医院 AI+肺癌防治中心, 上海 200032

[摘要] **目的** 系统综述人工智能(artificial intelligence, AI)与多模态融合技术在肺结节良恶性鉴别中的研究进展, 重点阐述影像、临床资料及血液生物标志物整合决策的理论基础、关键技术、临床价值与应用边界, 为肺结节精准分层管理及临床转化提供参考。 **方法** 基于肺结节管理相关国际指南、经典风险预测模型、近年 AI 影像学、多组学与液体活检研究以及方法学规范文献, 从多模态技术评估肺结节的意义、影像数据与临床资料融合、血液标志物/循环肿瘤 DNA(circulating tumor DNA, ctDNA)/循环染色体异常细胞(circulating genetically abnormal cells, CAC)/蛋白组学协同价值、多模态模型的临床预测潜力、“辅助决策”而非“替代决策”的边界, 以及当前挑战与解决路径等方面进行归纳、分析与综合。 **结果** 当前肺结节管理仍以结节大小、体积、密度、边缘特征、增长动力学及患者年龄、吸烟史、既往肿瘤史等临床危险因素为基础, 并依托 Fleischner 学会、英国胸科学会(British thoracic society, BTS)、肺结节分级及数据系统(lung imaging reporting and data system, Lung-RADS)、美国胸外科协会(American association for thoracic surgery, AATS)等指南及 Mayo、Herder、Brock 等经典模型完成风险分层。然而, 在亚厘米结节、亚实性结节、多发结节、炎症背景结节及中间风险结节中, 单一影像征象或传统模型的校准能力和临床净获益仍有限。AI, 尤其是影像组学(radiomics)、深度学习和多模态机器学习, 可从 CT 中提取人眼难以稳定识别的高维特征; 与年龄、吸烟状态、肺气肿、既往恶性肿瘤史等临床变量融合后, 可提升风险评估的稳定性和重分类能力。同时, 游离细胞 DNA (cell-free DNA, cfDNA)/ctDNA 甲基化、片段组学、CAC 及蛋白分类器等液体活检手段, 为中间风险结节提供了额外的分子与细胞学证据, 有助于减少不必要的侵入性操作, 并促进真正高危病灶更早进入精准诊断流程。当前多模态模型的主要价值体现在中间风险结节的二次分层、专病门诊流程优化和多学科团队(multidisciplinary team, MDT)决策支持, 但其临床落地仍受数据异质性、外部泛化能力不足、前分析标准化不充分、高维低样本、监管与支付机制等因素制约。 **结论** 肺结节良恶性鉴别正由单模态影像判断向“影像—临床—生物标志物”多模态整合决策转变。AI 在当前阶段最合理的定位仍是辅助决策而非替代决策。未来真正可能改变临床实践的, 将是经前瞻性验证、具备可解释性与可审计性、符合指南逻辑并可嵌入肺结节专病中心和 MDT 工作流的多模态辅助决策系统。

[关键词] 肺结节; 人工智能; 多模态融合; 影像组学; 深度学习; cfDNA 甲基化; CAC; 蛋白组学; 辅助决策

[中图分类号] R 734.2 **[文献标志码]** A

Application of AI and multimodal fusion in the differential diagnosis of benign and malignant pulmonary nodules

Tong Lin^{1,2,3,4}, Bai Chunxue^{1,2,3,4*}

1. Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Fudan University, Shanghai 200032, China
2. Shanghai Engineer & Technology Research Center of Internet of Things for Respiratory Medicine, Shanghai 200032, China
3. Shanghai Respiratory Research Institution, Shanghai 200032, China
4. AI+ Lung Cancer Prevention and Treatment Center, Zhongshan Hospital, Fudan University, Shanghai 200032, China

[Abstract] **Objective** To systematically review recent advances in artificial intelligence (AI) and multimodal fusion for differentiating benign from malignant pulmonary nodules, with a focus on the theoretical basis, key technologies, clinical utility, and practical boundaries of integrated decision-making based on imaging, clinical data, and blood-based biomarkers. **Methods** International guidelines for pulmonary nodule management, classic risk prediction models, recent AI-based imaging studies, multi-

[收稿日期] 2026-03-12

[接受日期] 2026-03-28

[基金项目] 四大慢病重大专项(2024ZD0529300). Supported by Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0529300).

[作者简介] 童琳, 博士, 副主任医师. E-mail: tong.lin@zs-hospital.sh.cn

*通信作者(Corresponding author). 白春学, 博士, 教授, 主任医师. E-mail: bai.chunxue@zs-hospital.sh.cn

omics and liquid biopsy studies, and methodological consensus documents were reviewed. Evidence was synthesized from six perspectives: the significance of multimodal assessment, integration of imaging and clinical variables, synergistic value of blood biomarkers including ctDNA and circulating genetically abnormal cells (CAC), clinical potential of multimodal models, the boundary between decision support and decision replacement, and current challenges with possible solutions. **Results** Current pulmonary nodule management still relies primarily on nodule size, volume, density, margin characteristics, growth dynamics, and conventional clinical risk factors such as age, smoking history, and prior malignancy, under the framework of established guidelines and prediction models. However, in subcentimeter nodules, subsolid nodules, multiple nodules, inflammation-related nodules, and intermediate-risk nodules, single-modality imaging features and conventional models remain inadequate in calibration and net clinical benefit. AI-based radiomics, deep learning, and multimodal machine learning can extract high-dimensional CT features beyond human visual recognition and improve risk stratification when combined with clinical variables. Meanwhile, liquid biopsy approaches, including cfDNA/ctDNA methylation, fragmentomics, CAC, and proteomic classifiers, provide additional molecular and cellular evidence for intermediate-risk nodules, thereby helping reduce unnecessary invasive procedures and accelerating precision diagnosis in truly high-risk cases. Nevertheless, real-world implementation remains limited by data heterogeneity, insufficient external validation, lack of assay standardization, high-dimensional low-sample-size issues, and regulatory and reimbursement barriers. **Conclusion** The differential diagnosis of pulmonary nodules is evolving from single-modality imaging judgment toward multimodal integrated decision-making based on imaging, clinical data, and biomarkers. At the current stage, AI should be positioned as a decision-support tool rather than a decision-replacement tool. Future practice-changing systems will likely be prospectively validated, interpretable, auditable, guideline-concordant multimodal platforms that can be seamlessly embedded into pulmonary nodule clinics and multidisciplinary workflows.

[Key Words] pulmonary nodule; artificial intelligence; multimodal fusion; radiomics; deep learning; cfDNA methylation; circulating genetically abnormal cells; proteomics; decision support

肺结节是肺癌早诊早治时代最常见的临床入口之一。随着低剂量螺旋CT (low-dose computed tomography, LDCT) 筛查持续推进、体检胸部CT普及以及胸部影像检查量快速增加,肺结节已成为现代呼吸与胸部肿瘤临床实践中最常见、最复杂的影像学问题之一。美国国家肺癌筛查试验 (National Lung Screening Trial, NLST) 与荷兰-比利时肺癌筛查试验 (Nederlands-Leuvens Longkanker Screenings Onderzoek, NELSON) 等里程碑研究证实,LDCT可降低肺癌死亡率,使“早发现”成为高危人群管理的重要策略;但与之并行的是筛查结节和偶发结节数量持续增加,临床关注重点也由“是否发现结节”进一步转向“如何对发现的结节进行准确、分层、低创且可持续的良恶性鉴别”^[1-6]。

当前肺结节评估主要建立在“影像征象+临床危险因素+指南化随访/干预阈值”的框架上。Fleischner学会2017指南将实性与亚实性结节纳入统一管理体系,并对多发结节、危险因素和随访问隔作出细化;Lung-RADS v2022更新了筛查场景下的分级与管理建议;AATS亚实性结节专家共识进一步强调了亚实性结节的独特自然史和个体化处置逻辑。同时,Mayo、Herder、Brock等经典模型长期用于前测恶性风险估计^[3-10]。这些框架构成了当前肺结节管理的基本底盘,但更多反映群体层面的经验规则,在高度异质的个体病例中仍存在解释力不

足的问题^[10-12]。

肺结节管理的困难并不止于“良恶性二分类”,更在于这一判断直接影响随访频率、检查路径、取材方式、手术时机及患者心理负担。对患者而言,“肺结节”常迅速引发对肺癌的高度焦虑;对医生而言,则意味着必须在有限信息下作出既不能延误、又不能过度的平衡判断。传统路径中,这种平衡主要依赖影像科医师经验、呼吸/胸外科临床判断及既有指南框架,但随着结节数量显著增加,单纯依赖经验的模式已越来越难以满足高通量、标准化与个体化并存的现实需求^[3-10]。

此外,肺结节管理受明显的地区差异和疾病谱差异影响。在真菌感染、结核或慢性炎症性肺病较常见地区,炎性结节和瘢痕样结节比例更高,影像学假阳性问题更突出;而在高筛查覆盖地区,更多被发现的是体积更小、密度更淡、自然史更长的早期病灶,传统“见结节即高度警惕”的思维也需相应调整^[3-6]。因而,肺结节良恶性鉴别从来不是一个脱离背景的单纯影像问题,而是一个嵌入人群谱系、临床场景和资源配置条件中的综合问题。

正是在这一背景下,人工智能与多模态整合迅速进入肺结节研究与转化视野。影像组学将CT从“可视图像”转化为“高维可计算数据”,深度学习实现了从原始体素到高阶潜在表征的端到端学习,而cfDNA甲基化、CAC和蛋白分类器等液体活检工具,

则为影像难以定性的结节提供了额外的分子与细胞学证据^[13-40]。越来越多研究提示,肺结节管理未来最关键的升级,并不是再增加一个孤立的影像评分,而是构建能够整合影像、临床与生物标志物的辅助决策体系,使不同证据在同一患者体内完成相互校正、相互补充与共同赋能^[16,25-33]。

本文围绕多模态技术评估肺结节的意义、影像数据与临床资料融合、血液标志物与多组学协同价值、多模态模型在临床预测中的潜力、“辅助决策”而非“替代决策”的边界,以及挑战与解决路径等方面,系统综述AI+多模态融合在肺结节良恶性鉴别中的研究进展与转化方向,为后续建立更加精准、可解释、可落地的肺结节智能管理体系提供参考。

1 多模态技术评估肺结节的意义

LDCT筛查的临床价值已在大型随机研究中得到明确验证,但筛查获益并不是由“检出率”单独决定,而是由“检出后的管理质量”共同决定^[1-2]。如果一个筛查体系只能发现更多结节,却无法准确区分哪些值得继续随访、哪些需要进一步检查、哪些应尽快取材或手术,那么其总体收益将被大量假阳性、过度检查和患者焦虑所削弱^[1-4]。因此,肺结节管理的核心已由“能否看到病灶”转向“能否理解病灶”,即在影像发现之后尽快回答病灶的恶性概率、生长潜力及最合适的下一步路径。

单一CT形态学征象虽仍是肺结节评估的起点,但其本质上主要反映病灶宏观表型,难以完全对应底层肿瘤生物学。炎症性结节、肉芽肿、机化灶、局灶性纤维化、非典型腺瘤样增生、原位腺癌与早期浸润性腺癌,在影像上可能共享相似的大小、密度和边界特征^[3,5-6,13-18]。对亚实性结节而言,这种局限更为突出,因为其自然史较长、异质性更高,单次CT往往难以完整呈现其演化趋势^[5]。

多模态融合之所以受到重视,首先在于不同数据源之间具有信息互补性。影像提供空间结构信息,包括病灶形态、纹理、边界、内部异质性及其时间演化;临床资料提供先验概率信息,包括年龄、吸烟暴露、肺气肿、家族史和既往恶性肿瘤史等;血液与分子标志物则更接近病灶的生物学活动层面^[9,16,25-33]。其次,多模态还具有误差互补性。影像学判断易受炎症背景、机器参数和观察者经验影响,而液体活检则可能受肿瘤负荷、样本处理、检测平台和阈值选择影响。联合后有望降低单一模态的系统偏差。更重要的是,多模态整合的目的并非简单提高模型分数,而是帮助临床把模糊的中间风

险病例更可靠地向低风险或高风险两端移动,从而改变后续管理动作^[16,25,31-33]。

从临床可用性看,多模态评估最适合用于证据链不完整的结节,而非不分场景地应用于全部肺结节。典型场景包括:亚厘米结节,尤其是直径5~10 mm的实性或部分实性结节;亚实性结节和混合磨玻璃结节;传统模型评估为中间风险的病例;以及不适宜贸然侵入性操作的人群^[5,8-9,28-33]。对于这些患者,多模态整合不仅有望提高恶性风险判别,更重要的是可提供更符合个体条件的路径选择。

从更宏观的角度看,多模态评估的价值不能仅以准确率定义,还应从医疗资源配置角度理解。若一个模型能够更可靠地识别真正低风险人群,就可能减少不必要的高成本影像和有创操作;若能更早识别真正高风险人群,则可能缩短诊断延迟,提高早期可切除肺癌的发现率^[3-10]。因而,多模态评估的重要性并不只在于“更先进”,而在于其可能改变管理强度分布,使有限资源更集中于真正需要强化干预的人群。

需要强调的是,多模态并不等于把尽可能多的数据一股脑塞进模型。真正有临床意义的多模态,应体现为不同信息层级之间的有序整合:基础影像和临床危险因素可作为第一层,影像组学或深度学习特征作为第二层,血液标志物和 omics 信号作为第三层^[13-40]。这种分层设计更符合真实 workflow,也更有助于控制成本和提高可实施性。

2 影像数据与临床资料融合

无论AI如何发展,结节大小、体积、部位、数目、密度、边缘特征及增长动力学,仍然是肺结节风险评估中最基础、最有解释力的一类输入变量^[3-10]。同样,年龄、吸烟暴露、肺气肿、既往恶性肿瘤史和家族史等临床变量,也不会因为AI的出现而失去价值。Mayo、Herder和Brock模型的长期沿用,正体现了影像与临床联合判断的基本逻辑^[7-10]。

影像组学的本质,是把原本由肉眼观察和语言描述的影像学表型转化为可大规模提取、量化和建模的数值特征^[13-18,22-23]。这些特征包括灰度分布、纹理粗细、形状不规则性、边界复杂度、内部异质性、周围肺实质特征以及病灶与宿主界面模式等。其价值首先体现在对“人眼难以稳定识别差异”的捕捉能力上,但同时也面临对扫描协议、图像重建和分割方法敏感等局限^[22-24]。

深度学习进一步推动影像特征提取由“手工特征工程”走向“端到端表征学习”^[16,41-44]。它能够从

原始CT体素或多切面图像中自动学习高阶潜在模式,在复杂背景、边缘模糊病灶、微小结节和时序影像建模中显示出明显潜力。尤其在结节形态不典型、人工定义特征难以充分表达差异的场景中,深度学习可能更具优势。但与此同时,其黑箱性、外部泛化问题和解释困难,也决定了它难以脱离临床场景独立使用^[19-24]。

肺结节是高度依赖先验概率的临床问题。相同大小和外观的结节,在老龄重度吸烟者与年轻非吸烟者中的恶性含义并不相同^[7-10]。因此,影像AI模型若脱离患者背景单独输出“高风险”或“低风险”,其结果往往不具备足够稳定的临床解释力。已有研究提示,将影像AI风险分数与年龄、吸烟状态、肺气肿、既往恶性肿瘤史等变量联合后,模型的判别力、校准度和净获益通常更加稳定^[16,25,45-48]。

未来更重要的方向,是将影像-临床融合进一步扩展为时序化、动态化风险评估。肺结节不是静态对象,而是随时间演化的过程。体积倍增时间、实性成分增加、边界变化和长期稳定性,往往比单次直径更能反映其真实生物学行为^[4-6,10,44,48]。因此,高水平模型应尽量纳入纵向随访影像,使风险评估由单个时间点的静态分类升级为动态更新的风险曲线。

同时,临床变量不只是协变量,更是模型输出的解释框架。医生在接收模型结果时,天然会追问:这一输出是否已考虑高龄、吸烟史、肺气肿或既往肿瘤史?若缺乏这些语境,模型即使性能较高,也难以真正获得信任。影像分割质量同样决定后续建模上限,尤其在亚实性结节中,边界模糊和对比度使分割偏差更易被放大^[22-24]。此外,影像-临床融合模型在评价时应更加重视校准,而不只是区分能力,因为临床真正使用的不是“排序”,而是“概率是否可信”^[20-21]。

3 血液标志物、ctDNA、CAC、蛋白组学等协同价值

液体活检被纳入肺结节研究视野,并不是因为其已达到可以替代病理的程度,而是因为它提供了与影像和临床先验不同维度的证据^[25-27]。影像主要反映形态表型,临床变量主要反映背景风险,而液体活检则试图捕捉肿瘤相关分子、细胞或宿主反应信号。这种层次差异,决定了液体活检最适合做补证,而非单独定论。

CAC是近年来肺结节液体活检研究中的亮点之一。Ye等^[28]研究聚焦于 ≤ 10 mm肺结节,发现CAC对识别小结节肺癌具有附加诊断价值。这类

研究的重要性在于,其瞄准了影像最难分层、病理又不易获得、同时临床最担心延误早癌的一类病灶。后续研究进一步提示,CAC与CT特征及临床变量联合时,其判别能力优于单独应用^[29-30]。因而,CAC更适合作为影像和临床评估后的二次校正工具。

在分子层面,cfDNA/ctDNA甲基化和片段组学成为近年肺结节研究热点^[31-34]。早期肺癌和亚厘米结节由于肿瘤负荷较低,单纯突变型ctDNA检测常面临灵敏度不足问题;相比之下,甲基化和片段组学更可能在早期阶段保留可用信号。其优势在于更接近肿瘤发生早期的表观遗传异常,并不完全依赖高丰度肿瘤DNA释放。相关研究提示,将cfDNA甲基化与LDCT影像及临床变量联合,可形成更贴近实际应用场景的风险分层模型^[31-34]。

蛋白组学和蛋白分类器则代表另一条更易与门诊流程衔接的路线。传统血清肿瘤标志物在肺结节良恶性鉴别中的独立表现有限,因此研究重点已由单一指标转向多蛋白面板与整合分类器^[35-40]。PANOPTIC及其后续分析显示,整合若干血浆蛋白和临床因素的分类器,在某些低至中等风险结节人群中,具有辅助“风险下调”和减少不必要侵入性操作的潜力^[35-37]。

从现有证据看,未来肺结节生物标志物发展的最优方向,并不是寻找一个适用于所有场景的“万能检测”,而是根据不同临床问题整合不同层面的信号。对亚厘米结节和筛查场景,cfDNA甲基化与CAC可能更有潜力;对低至中等风险偶发结节,蛋白分类器与临床模型结合可能更现实;对复杂病例,则可能需要临床+CT+omics的深度整合模型^[25-34]。

同时必须指出,肺结节生物标志物的理想定位应是“二次分层”工具,而不是开局即替代影像作出决定。对肺结节而言,最现实的应用场景,是在影像和临床判断之后,为“尚未足以决定下一步动作”的患者提供额外证据^[25-27]。甲基化与片段组学的潜力在于“早期信号放大”,而蛋白组学的优势则在于更容易嵌入门诊流程。未来开发组合biomarker时,也应警惕单纯堆砌变量带来的冗余和不可解释性。对临床而言,一个可重复、可解释、可支付的简洁模型,往往比一个极其复杂但难以落地的模型更有价值。

4 多模态模型在临床预测中的潜力

肺结节模型研究中,AUC、敏感度和特异度仍是

常见指标,但从临床视角看,更重要的问题是:模型是否真正改变了患者管理^[16,20-21,25]。如果一个模型仅在统计学指标上略有提高,却未减少不必要穿刺、未降低良性手术比例、也未使高危结节更快进入治疗流程,那么其临床意义仍然有限。

在肺结节管理中,低风险结节通常可按指南随访,高风险结节往往已具备相对明确的侵入性诊断或手术倾向^[3-10]。真正最需要多模态辅助的,是处于中间风险区间的患者。对这部分人群而言,传统影像和临床信息往往难以给出确定结论,而任何额外的有价值证据都可能影响下一步动作。多模态模型若能将这部分患者更稳健地向低风险或高风险两端推移,其临床价值就远大于在总体样本中单纯追求更高分。

从 workflow 角度看,肺结节管理可分为3层:入口层完成结节检出与初步风险提示;专病门诊层完成影像、临床及必要生物标志物的进一步重分类;MDT层作出最终的随访、PET-CT、支气管镜、经皮穿刺或手术决策^[4-5,16]。多模态模型最适合嵌入第二层与第三层之间,作为结构化证据整合器。这样既能提高路径标准化程度,又能保留对复杂病例的人工裁量。

近年的代表性研究已逐步由“CT+年龄+吸烟史”的简化模式,发展到 clinic-radiomics-omics 的联合建模^[16,31-33]。这意味着肺结节 AI 的竞争重点正在由“谁的单模态模型更高分”转向“谁能构建更符合临床真实决策流程的整合系统”。这类系统的价值,不仅在于理论上更全面,更在于其更接近真实世界中的问题解决方式:医生从来不是只看一张 CT 作决定,而是在不同层面证据间进行综合判断。

从转化角度看,多模态模型更值得关注的终点,还包括“动作改变率”,即模型在多大比例患者中真正改变了医生原本的管理计划。例如,原本拟行经皮穿刺的患者,在综合风险下调后转为影像随访;原本计划半年后复查的患者,在综合风险上调后提前进入 MDT 或取材流程。只有当模型能在这些具体动作层面带来合理改变,才说明它真正参与了临床决策。

此外,多模态模型落地还必须回答一个现实问题:多高的风险需要进一步检查,多低的风险可以安心随访。不同场景下,这些阈值并不相同。筛查项目中,为控制漏诊,阈值可能更保守;在高龄合并症患者中,为避免过度干预,阈值可能更强调净获益;在亚实性结节中,短期风险与长期风险又不一定一致。因此,一个真正实用的肺结节模型,不应

只告诉医生“恶性概率 18%”,而应尽量提示该风险在当前场景下更倾向于哪一种管理路径。

随着模型能力提升,多模态系统还有望推动肺结节随访策略由“固定频率复查”转向“个体化风险驱动复查”。不同结节即使大小相似,其综合风险和增长潜力也可能差异明显。若低风险结节可适当延长复查间隔,而风险升高者提前进入更密集评估,则既可减少不必要辐射暴露和医疗负担,也更符合精准医学思路。

5 “辅助决策”而非“替代决策”的边界

无论模型表现多好,肺结节良恶性鉴别在相当长时期内仍将是一种概率判断^[3-10]。“高风险”意味着更应加快进一步验证,并不等于病理已确诊;“低风险”意味着更倾向于观察,也不等于绝对安全。因此,AI 与多模态模型的最佳定位,是帮助医生更精准地量化风险和优化路径,而不是脱离病理、随访和 MDT 独立作出不可逆决策。

同时,不同人群和场景下,同一模型分数的含义并不相同。筛查人群、偶发结节人群、既往恶性肿瘤患者、炎症高发地区居民及非吸烟女性,具有明显不同的疾病谱和先验概率^[4-6,9,16]。因此,同一个模型分数在不同人群中的含义不应被机械等同。任何模型在落地前都必须明确目标人群、阈值设定和不适用边界。

近年来,CLAIM、TRIPOD+AI、PROBAST+AI 等框架不断完善,强调医疗 AI 模型必须具备透明报告、偏倚评估、外部验证、可解释性和持续性能监测^[19-21]。对肺结节模型而言,这些要求尤为重要,因为此类模型高度依赖阈值、场景和人群适配。

在肺结节场景中,更符合实际的路径,是将人机协同设计为分工优化:AI 负责高通量、标准化、对微弱模式敏感的任务,如结节检出、定量测量、风险初筛和证据整理;临床医生负责病史整合、异常情境识别、价值判断和与患者共同决策。二者不是彼此竞争,而是在不同层面发挥优势。尤其对中间风险患者而言,其之所以困难,正是由于信息不完整、证据存在冲突、不同路径各有代价。若将这类情况简单外包给单一模型,最容易造成过度相信单一分数的误区。正确做法应是把模型结果作为强化讨论的工具,而不是压缩讨论的借口。

6 挑战与解决方案

肺结节影像数据在设备品牌、层厚、重建核、剂量、分割方式和标注标准上差异显著,这些差异会

直接影响影像组学和深度学习模型的稳定性^[15-24]。解决之道包括多中心标准化采集、稳健特征筛选、域适配、联邦学习和充分外部验证。只有当模型在不同医院、不同设备和不同人群中均能维持稳定性,才具备真实世界应用价值。

液体活检虽然非侵入,但其结果高度依赖采血时间、样本处理、保存方式、检测平台和阈值设定^[25-40]。不同实验室之间若缺乏统一SOP,结果的可比性和可重复性都将受到影响。因此,未来必须强化样本库建设、跨平台验证和流程标准化,明确不同标志物在不同结节场景下的最优使用方式。

多模态模型常同时纳入大量影像特征、临床变量和分子指标,但真实样本数和恶性事件数并不总是足够,易形成典型的高维低样本问题^[19-21]。这会带来训练集高性能、外部验证失效的风险。因此,未来研究应更加重视降维、嵌套验证、校准度和临床净获益,而不是片面追求高AUC。

即使模型研究结果理想,若不能满足监管审批要求、缺乏成本效果证据、无法与PACS/HIS/结节随访平台对接,其临床价值仍难以充分实现^[16,19-21,25]。真正决定多模态模型转化深度的,不只是技术表现,还包括其是否可解释、可审计、可部署、可支付,并能否自然嵌入现有肺结节专病中心工作流。

此外,许多肺结节AI或biomarker研究采用的是经过筛选的理想化队列,但真实临床中,患者流往往更加复杂。一个模型若只能在理想化数据集中工作,而不能适应真实世界复杂性,就很难真正进入日常诊疗。因此,未来研究需要更多地由“精选样本验证”转向“连续患者流验证”。

肺结节管理不仅是技术问题,也是沟通问题。即使模型准确,若医生无法向患者解释“为什么建议继续观察而不是立刻手术”“为什么虽然是结节但目前总体风险仍低”,患者仍可能因焦虑而倾向过度检查。因此,多模态系统未来若想真正提升临床质量,还应考虑如何把复杂模型结果转化为便于医生与患者沟通的风险表达方式。

一个真正可推广的肺结节多模态系统,至少应形成以下闭环:(1)在多中心队列中完成开发与外部验证;(2)明确目标人群、阈值和不适用场景;(3)在真实世界前瞻性研究中证明其能改变管理路径并改善资源使用;(4)形成与医院信息系统和结节随访平台兼容的可部署形态。缺少其中任一步,模型都可能停留在“科研可发表”而非“临床可使用”的阶段。

7 结论

肺结节良恶性鉴别正处于由单模态影像判断向多模态证据整合决策过渡的关键阶段^[16,25-33]。现行指南和传统临床风险模型仍是管理底座,但单纯依赖影像形态学和静态临床变量,已难以满足筛查时代对精细分层和低创诊疗的要求^[3-10]。AI尤其是影像组学、深度学习、多模态机器学习,以及cfDNA甲基化、CAC、蛋白分类器等液体活检工具的加入,使“影像—临床—生物标志物”联合决策成为现实可行的发展方向^[16,25-37]。

其中,中间风险结节将是多模态模型最重要、也最能体现临床价值的应用人群。对于这部分患者,单一模态证据常常不足,而多模态整合有望通过风险重分类减少不必要侵入性检查、降低良性结节手术比例,并促使真正高危病灶更快进入精准诊断和治疗路径。与此同时,也必须清醒认识到,AI在当前阶段最合理的角色仍是辅助决策而非替代决策。真正能够改变实践的,不是某一个孤立算法,也不是某一个单一生物标志物,而是经前瞻性验证、符合指南逻辑、具有可解释性和可审计性、并能嵌入肺结节专病中心和MDT工作流的多模态辅助决策系统^[16,19-21,25-33]。

还应进一步强调,肺结节多模态融合的终极目标,并不是把所有患者都送入更复杂、更昂贵的检测流程,而是建立一种“按风险逐层升级”的管理体系:低风险患者得到更安心、更简化的随访;中间风险患者获得更有针对性的补充证据;高风险患者更快进入精准诊断与治疗路径。只有当多模态系统能够同时服务于“减少过度”与“避免延误”这两个目标时,它才真正体现出临床价值。

从更长远的角度看,肺结节可能成为整个呼吸肿瘤智能管理体系的一个入口场景。随着影像组学、液体活检、纵向随访建模和专病中心平台化建设的持续推进,未来肺结节评估有望由一次性判断演变为持续更新的数字化风险管理过程。在这一过程中,AI不应被视为孤立工具,而应被视为连接影像、临床、生物标志物与MDT的桥梁。只有这样,肺结节管理才可能真正由“发现病灶”升级为“整合证据、精准分流、长期优化”。

未来研究应进一步强化多中心真实世界验证、时序数据融合、跨平台标准化和临床净获益评估,推动肺结节评估从“看见结节”真正迈向“看懂结节、分准风险、选对路径”。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 童琳:撰写、修改、核对文献;白春学:选题、撰写、定稿。

参考文献

- [1] National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening [J]. *N Engl J Med*, 2011, 365(5): 395-409.
- [2] de Koning H J, van der Aalst C M, de Jong P A, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial[J]. *N Engl J Med*, 2020, 382(6): 503-513.
- [3] MacMahon H, Naidich D P, Goo J M, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the fleischner society 2017[J]. *Radiology*, 2017, 284(1): 228-243.
- [4] Christensen J, Prosper A E, Wu C C, et al. ACR lung-RADS v2022: assessment categories and management recommendations [J]. *CHEST*, 2024, 165(3): 738-753.
- [5] Chen H, et al. The 2023 AATS expert consensus document: management of subsolid lung nodules. *J Thorac Cardiovasc Surg*. 2024.
- [6] Callister M E J, Baldwin D R, Akram A R, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules: accredited by NICE [J]. *Thorax*, 2015, 70(Suppl 2): ii1-ii54.
- [7] Swensen S J, Silverstein M D, Ilstrup D M, et al. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules [J]. *Arch Intern Med*, 1997, 157(8): 849-855.
- [8] Herder G J, van Tinteren H, Golding R P, et al. Clinical prediction model to characterize pulmonary nodules: validation and added value of 18F-fluorodeoxyglucose positron emission tomography[J]. *Chest*, 2005, 128(4): 2490-2496.
- [9] McWilliams A, Tammemagi M C, Mayo J R, et al. Probability of cancer in pulmonary nodules detected on first screening CT [J]. *N Engl J Med*, 2013, 369(10): 910-919.
- [10] Choi H K, Ghobrial M, Mazzone P J. Models to estimate the probability of malignancy in patients with pulmonary nodules [J]. *Ann Am Thorac Soc*, 2018, 15(10): 1117-1126.
- [11] Maldonado S G, Tanner N T, Silvestri G A, et al. Evaluation of prediction models for identifying malignancy in pulmonary nodules detected via low-dose CT screening. *JAMA Netw Open*. 2020;3(10):e2023121.
- [12] Chen S, Lin W L, Liu W T, et al. Pulmonary nodule malignancy probability: a meta-analysis of the Brock model[J]. *Clin Radiol*, 2025, 82: 106788.
- [13] Ather S, Kadir T, Gleeson F. Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications[J]. *Clin Radiol*, 2020, 75(1): 13-19.
- [14] Kim R Y. Radiomics and artificial intelligence for risk stratification of pulmonary nodules: Ready for primetime? [J]. *Cancer Biomark*, 2025, 42(1): CBM230360.
- [15] Warkentin M T, Al-Sawaihey H, Lam S, et al. Radiomics analysis to predict pulmonary nodule malignancy using machine learning approaches[J]. *Thorax*, 2024, 79(4): 307-315.
- [16] Wang C D, Shao J, He Y C, et al. Data-driven risk stratification and precision management of pulmonary nodules detected on chest computed tomography[J]. *Nat Med*, 2024, 30(11): 3184-3195.
- [17] Du S, Wu T W, Wang H, et al. Advances in modelling the risk of benign and malignant lung nodules[J]. *Front Oncol*, 2025, 15: 1648548.
- [18] Li X, et al. Narrative review of the application of artificial intelligence in the diagnosis and management of pulmonary nodules. *J Thorac Dis*. 2025.
- [19] Mongan J, Moy L, Kahn C E Jr. Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers[J]. *Radiol Artif Intell*, 2020, 2(2): e200029.
- [20] Collins G S, Moons K G M, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods [J]. *Bmj*, 2024: e078378.
- [21] Moons K G M, Damen J A A, Kaul T, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods[J]. *BMJ*, 2025, 388: e082505.
- [22] Gillies R J, Kinahan P E, Hricak H. Radiomics: images are more than pictures, they are data [J]. *Radiology*, 2016, 278(2): 563-577.
- [23] Zwanenburg A, Vallières M, Abdalah M A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping [J]. *Radiology*, 2020, 295(2): 328-338.
- [24] Avanzo M. Robust radiomics: a review of guidelines for reproducible and clinically translatable radiomic studies [J]. *Front Radiol*, 2025.
- [25] Paez R, Kammer M N, Tanner N T, et al. Update on biomarkers for the stratification of indeterminate pulmonary nodules[J]. *CHEST*, 2023, 164(4): 1028-1041.
- [26] McGann K C, Khalil T A, Kammer M N, et al. Biomarkers for the diagnosis of indeterminate pulmonary nodules: are we there yet?[J]. *J Thorac Dis*, 2025, 17(8): 6265-6282.
- [27] Afridi W A, Picos S H, Bark J M, et al. Minimally invasive biomarkers for triaging lung nodules—challenges and future perspectives[J]. *Cancer Metastasis Rev*, 2025, 44(1): 29.
- [28] Ye M S, Zheng X X, Ye X, et al. Circulating genetically abnormal cells add non-invasive diagnosis value to discriminate lung cancer in patients with pulmonary nodules ≤ 10 mm [J]. *Front Oncol*, 2021, 11: 638223.
- [29] Yang H, Chen H J, Zhang G R, et al. Diagnostic value of circulating genetically abnormal cells to support computed tomography for benign and malignant pulmonary nodules [J]. *BMC Cancer*, 2022, 22(1): 382.
- [30] Tahvilian S, Kuban J D, Yankelevitz D F, et al. The presence

- of circulating genetically abnormal cells in blood predicts risk of lung cancer in individuals with indeterminate pulmonary nodules [J]. *BMC Pulm Med*, 2023, 23(1): 193.
- [31] Liang W H, Tao J S, Cheng C, et al. A clinically effective model based on cell-free DNA methylation and low-dose CT for risk stratification of pulmonary nodules [J]. *Cell Rep Med*, 2024, 5(10): 101750.
- [32] Zhao M M, Xue G, He B X, et al. Integrated multiomics signatures to optimize the accurate diagnosis of lung cancer [J]. *Nat Commun*, 2025, 16: 84.
- [33] Yang M, Yu H S, Feng H X, et al. Enhancing the differential diagnosis of small pulmonary nodules: a comprehensive model integrating plasma methylation, protein biomarkers, and LDCT imaging features [J]. *J Transl Med*, 2024, 22(1): 984.
- [34] Li Y, Xie F F, Zheng Q, et al. Non-invasive diagnosis of pulmonary nodules by circulating tumor DNA methylation: a prospective multicenter study [J]. *Lung Cancer*, 2024, 195: 107930.
- [35] Silvestri G A, Tanner N T, Kearney P, et al. Assessment of plasma proteomics biomarker's ability to distinguish benign from malignant lung nodules: results of the PANOPTIC (pulmonary nodule plasma proteomic classifier) trial [J]. *Chest*, 2018, 154(3): 491-500.
- [36] Tanner N T, Porter A, Gould M K, et al. Assessment of integrated classifier's ability to improve the evaluation of indeterminate pulmonary nodules. *Chest*. 2021; 159(6): 2360-2369.
- [37] Long K J, Pitcher T, Kurman J S, et al. Using a blood biomarker to distinguish benign from malignant pulmonary nodules a subgroup analysis comparing screen detection, sex, smoking history, and nodule size [J]. *CHEST*, 2023, 164(6): 1572-1575.
- [38] Jia B, Wang T T, Pan L X, et al. An integrated proteomic classifier to distinguish benign from malignant pulmonary nodules [J]. *Clin Proteom*, 2025, 22(1): 11.
- [39] Li X J, Hayward C, Fong P Y, et al. A blood-based proteomic classifier for the molecular characterization of pulmonary nodules [J]. *Sci Transl Med*, 2013, 5(207): e3007013.
- [40] Kuang M Y, Tao X T, Peng Y Z, et al. Proteomic analysis of plasma exosomes to differentiate malignant from benign pulmonary nodules [J]. *Clin Proteom*, 2019, 16(1): 5.
- [41] Sun Y, Ge X Y, Niu R, et al. PET/CT radiomics and deep learning in the diagnosis of benign and malignant pulmonary nodules: progress and challenges [J]. *Front Oncol*, 2024, 14: 1491762.
- [42] Wang H. Risk-stratified classification of pulmonary nodule malignancy and invasiveness using multimodal machine learning [J]. *Lancet Reg Health West Pac*, 2025.
- [43] Sohn J H, Fields B K K. Radiomics and deep learning to predict pulmonary nodule metastasis at CT [J]. *Radiology*, 2024, 311(1): e233356.
- [44] Ohno Y, Aoyagi K, Yaguchi A, et al. Differentiation of benign from malignant pulmonary nodules by using a convolutional neural network to determine volume change at chest CT [J]. *Radiology*, 2020, 296(2): 432-443.
- [45] van Riel S J, Ciompi F, Winkler Wille M M, et al. Malignancy risk estimation of pulmonary nodules in screening CTs: Comparison between a computer model and human observers [J]. *PLoS One*, 2017, 12(11): e0185032.
- [46] Hammer M M, Palazzo L L, Kong C Y, et al. Limited utility of pulmonary nodule risk calculators for low-dose CT lung cancer screening [J]. *J Am Coll Radiol*, 2018, 15.
- [47] Nair V S, Sundaram V, Desai M, et al. Accuracy of models to identify lung nodule cancer risk in the national lung screening trial [J]. *Am J Respir Crit Care Med*, 2018, 197(9): 1220-1223.
- [48] Tammemagi M, Ritchie A J, Atkar-Khattra S, et al. Predicting malignancy risk of screen-detected lung nodules-mean diameter or volume [J]. *J Thorac Oncol*, 2019, 14(2): 203-211.

引用本文

童琳,白春学. AI+多模态融合在肺结节良恶性鉴别中的应用 [J]. 元宇宙医学, 2026, 3(1): 64-71.

Tong L, Bai C X. Application of AI and multimodal fusion in the differential diagnosis of benign and malignant pulmonary nodules [J]. *Metaverse Med*, 2026, 3(1): 64-71.