

肺结节恶性风险预测模型研究进展

王辉, 周云*

昆明医科大学第二附属医院胸外科, 云南 昆明

收稿日期: 2026年2月3日; 录用日期: 2026年2月26日; 发布日期: 2026年3月6日

摘要

随着低剂量CT (LDCT)筛查的广泛应用, 肺结节的检出率显著增加, 这使得肺结节良恶性鉴别成为临床诊断中的关键挑战。传统的影像学评估方法依赖于医师的经验, 存在较大主观性, 且难以在复杂病例中准确区分结节的良恶性。近年来, 数字诊断模型, 尤其是影像组学(radiomics)和深度学习模型, 逐渐成为肺结节管理的重要工具。这些模型通过提取和分析CT图像中的复杂特征, 能够在更高维度上进行结节评估, 显著提高了诊断的准确性。与此同时, 多模态融合方法, 结合临床数据、影像学特征和肿瘤标志物(如CEA、CYFRA 21-1), 为提高恶性预测能力提供了新的解决方案。同时, 新型液体活检标志物及模型校准度优化逐渐成为提升模型临床适用性的研究热点。然而, 尽管AI驱动模型在精确性和智能化方面展现出巨大潜力, 它们仍面临可解释性差、推广性有限以及数据偏移等问题。本文综述了肺结节风险评估模型的发展历程, 探讨了传统模型与新兴AI方法的优势与局限, 特别是在多模态融合与肿瘤标志物结合的潜力。最后, 展望了未来数字诊断工具的发展方向, 强调了提高可解释性、跨数据集验证和临床转化的必要性。

关键词

肺结节, 数字诊断模型, 肿瘤标志物, 液体活检, 人工智能

Advances in Predictive Models for Malignancy Risk in Pulmonary Nodules

Hui Wang, Yun Zhou*

Department of Thoracic Surgery, The Second Affiliated Hospital of Kunming Medical University, Kunming Yunnan

Received: February 3, 2026; accepted: February 26, 2026; published: March 6, 2026

*通讯作者。

文章引用: 王辉, 周云. 肺结节恶性风险预测模型研究进展[J]. 临床医学进展, 2026, 16(3): 1124-1133.
DOI: 10.12677/acm.2026.163888

Abstract

With the widespread application of low-dose CT (LDCT) screening, the detection rate of pulmonary nodules has significantly increased, making the accurate differentiation between benign and malignant nodules a critical clinical challenge. Traditional imaging evaluation methods are highly subjective and often inadequate in distinguishing the malignancy of nodules, especially in complex cases. In recent years, digital diagnostic models, particularly imaging omics (radiomics) and deep learning models, have become essential tools in the management of pulmonary nodules. These models extract and analyze complex features from CT images, enabling a higher-dimensional assessment of nodules and significantly improving diagnostic accuracy. Furthermore, multi-modal integration methods combining clinical data, imaging features, and tumor markers (e.g., CEA, CYFRA 21-1) offer new solutions to enhance malignancy prediction. Emerging liquid biopsy biomarkers and model calibration optimization are also becoming important directions to improve real-world clinical applicability. Despite the substantial potential of AI-driven models in accuracy and intelligence, challenges such as poor interpretability, limited generalizability, and data bias still remain. This review systematically examines the evolution of pulmonary nodule risk assessment models, comparing traditional methods with emerging AI approaches, and highlights the potential and challenges of integrating tumor markers. Finally, we discuss future directions for the development of digital diagnostic tools, focusing on improving interpretability, cross-dataset validation, and clinical translation.

Keywords

Pulmonary Nodules, Digital Diagnostic Models, Tumor Markers, Liquid Biopsy, Artificial Intelligence

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

肺结节相关基础概念

肺癌是所有癌症中死亡率最高的,是直肠癌死亡人数的 2.5 倍,但近年来在死亡率下降方面取得了显著进展,主要得益于早期检测的改进(如低剂量 CT 筛查) [1]。肺结节是指局限性、类圆形或不规则的肺实质占位影像学改变,其直径多为 3 cm 以下,并且被正常肺实质完全包绕,不伴随明显的肺不张、胸腔积液或肺门淋巴结肿大。这一定义在多篇经典文献中得到了确立与沿用,是当前临床与研究领域普遍接受的基础标准[2]-[4]。

早在 20 世纪 90 年代初, Gurney 等学者在关于孤立性肺结节(solitary pulmonary nodule, SPN)的早期综述中就指出, SPN 是一种“局灶性、界限相对清楚、直径小于 3 cm 的实性或亚实性肺内结节”,其影像特征是区分早期肺癌与良性病变的重要依据[2]。随着 CT 技术的发展,2000 年后大量影像学综述进一步完善了结节的现代分类方式,包括实性、部分实性(part-solid)和磨玻璃结节(ground-glass nodules, GGN)等不同密度类型,并强调密度特性在恶性风险分层中的核心价值[5] [6]。

在经典研究中, Swensen 等人在 1997 年发表的 Mayo Clinic 队列研究中,不仅使用这一标准定义 SPN,也基于这些定义构建了最具影响力的肺结节恶性概率模型之一[3]。此后,这一定义被广泛应用于各种风

险模型(如 Brock/PanCan 模型)及国际指南(如 Fleischner 指南)中, 用于规范化结节的描述和管理。并且在此基础上衍生出其他更多的区域化模型(如北大模型等)。

肺结节的良恶性鉴别长期以来是胸部影像学及临床决策中的关键挑战[7][8]。随着低剂量 CT (LDCT) 筛查的普及, 肺结节的检出率显著增加, 其中大多数为良性病变, 但仍有相当比例可能提示早期肺癌[7][9]。由于结节的影像表现往往缺乏特异性, 尤其是在小结节、磨玻璃结节(GGNs)及混合型结节中, 单纯依靠影像特征难以准确区分其良恶性[10][11]。这种判断的不确定性会带来两类临床风险: 一方面, 过高估计风险可能导致不必要的频繁随访、PET-CT 检查乃至穿刺或手术, 造成过度诊断和过度治疗[9][12]; 另一方面, 低估恶性风险则可能延误诊断, 错过早期发现与干预的关键时机[7][11]。因此, 提高肺结节风险评估的准确性与一致性具有重要临床意义。传统的结节判断方法高度依赖影像科医师和胸外科医师的经验, 关注结节大小、形态、边缘、钙化及生长情况等特征。然而, 这类经验性评估具有明显的主观性, 受操作者经验水平影响显著, 同时存在观察者间差异(interobserver variability) [13][14]。

此外, 在多特征并存或影像表现模糊的复杂病例中, 人工判断难以有效整合多维度信息, 并且随着筛查规模扩大和结节数量增加, 依赖经验的方式越来越难以满足现代精准医学和规范化管理的需求[14][15]。在此背景下, 基于大样本数据和统计或机器学习方法构建的数字化风险评估模型逐渐成为肺结节管理的核心工具[3][16]-[18]。该类模型通过整合患者年龄、吸烟史、既往肿瘤史、结节大小、密度、边缘特征等多变量, 能够量化计算恶性概率, 提供标准化、可重复、且远比人工经验更稳定的风险估计[3][18]。目前已有多个成熟模型, 如 Mayo Clinic 模型[3]、Brock/PanCan 模型[18][19]、北京大学团队构建的中国人人群模型[16], 以及复旦中山医院提出的多参数模型[17], 以及针对非实性小结节的模型[20]等, 它们在不同人群和临床场景中显示出重要价值。

数字化风险评估模型不仅有助于减轻经验依赖, 提高诊断一致性, 也为决策制定提供量化依据, 例如优化随访方案、筛选需要进一步检查或干预的患者[3][15]-[17]。随着人工智能的发展, 肺结节风险评估正朝着更加精准和智能化的方向演进[21]。本文旨在系统回顾并对经典肺结节风险评估模型(如 Mayo、Brock/PanCan 及中国本土模型)的演变逻辑、变量构成与适用场景; 梳理影像组学、深度学习及多模态融合等新兴方法的研究进展与模型效能; 重点探讨融合肿瘤标志物的多模态模型在提升诊断准确性方面的潜力与挑战; 同时, 随着液体活检等分子检测技术的发展, 新型生物标志物在肺结节风险评估中的潜在价值亦逐渐受到关注, 值得在未来模型构建中进一步探索。总结当前模型的局限性(如可解释性、泛化性问题), 并提出未来向可解释、强泛化、多模态整合方向发展的路径, 以期肺结节智能诊断体系的临床转化提供参考。

2. 肺结节的经典风险评估模型

肺结节良恶性风险评估模型是实现精准化管理的重要工具, 其目标是以患者信息和影像学特征为基础, 通过统计学或机器学习方法定量预测结节的恶性概率。经典模型主要构建于大规模筛查或高危人群数据集, 其核心优势在于能够标准化风险判断、减少观察者差异、提升决策一致性。本节将介绍临床上最常用且具有代表性的模型, 包括 Mayo Clinic 模型、Brock/PanCan 模型以及适应中国人人群的国产模型。

2.1. Mayo Clinic 模型(Swensen 模型)

Mayo 模型是最早的肺结节恶性风险预测工具之一, 由 Swensen 等基于美国高危人群构建[3]。模型包含六个核心变量: 年龄、吸烟史、既往恶性肿瘤史、结节大小、上叶位置和毛刺边缘等特征[3][22]。该模型计算方式简便, 在实性结节中具有较好的预测能力, 并被多部国际指南引用为基础风险判断方法[8]。

然而, Mayo 模型建立于以吸烟者为主的西方队列, 其在磨玻璃结节和东亚人群中预测性能下降, 外部验证显示其可能低估亚实性结节风险[16]。

2.2. Brock/PanCan 模型(McWilliams 模型)

Brock 模型源自加拿大 Pan-Canadian Early Detection of Lung Cancer 项目的大规模筛查数据, 被认为是目前预测性能最稳定的筛查人群模型之一[19]。该模型纳入的变量更为丰富, 包括年龄、性别、吸烟史、家族史、结节大小、密度(包含 GGN 和部分实性结节)、结节数量、边缘特征、形状及位置等[18] [19]。Brock 模型特别适用于 LDCT 筛查环境, 对多发结节与亚实性结节具有良好预测能力, 并被英国 BTS 指南推荐用于筛查阳性结节管理[23]。其主要局限是变量多、计算复杂, 需要依赖在线工具。

2.3. 面向中国人群的风险预测模型

由于中国人群肺癌的流行特征不同(非吸烟女性腺癌比例高、磨玻璃结节更常见), 直接套用国外模型会影响准确性, 因此国产模型逐渐出现并表现出最佳的适配性。

2.3.1. 北京大学人民医院模型(PKUPH 模型)

北大模型基于中国大样本临床队列构建, 更加关注磨玻璃结节的影像学特征, 变量包括年龄、性别、吸烟状态、结节大小、密度、分叶、毛刺等指标[16]。多项外部验证显示该模型在中国患者(特别 GGN)中的预测优于 Mayo 与 Brock 模型[16] [24]。

2.3.2. 复旦大学中山医院多参数模型(Fudan/Zhongshan 模型)

复旦团队基于大量外科手术病例提出了更精细化的多参数模型, 适用于亚实性结节和疑似腺癌患者[17]。该模型除传统影像学特征外, 还纳入: 结节密度分型(纯 GGN、部分实性)、毛刺、分叶、空泡、血管异常、胸膜牵拉等腺癌危险征象, 其在亚实性结节恶性预测中的曲线下面积(AUC)显著高于 Mayo 与 Brock, 在“是否手术”“是否继续随访”的临床决策中具有重要价值[17]。

3. 模型之间的比较

3.1. 变量比较(见表 1)

Mayo 模型变量较少, 主要包括年龄、吸烟、既往肿瘤病史、结节直径、毛刺(spiculation)、结节位置(upper lobe)——偏重基础临床信息与简单影像特征[3]。

Table 1. Comparison of variables across models

表 1. 各模型变量比较

Model	Age	Smoking	Nodule size	Density type	Spiculation	Location (upper lobe)	Lobulation	Vacuole sign	Pleural indentation	Vascular convergence
Mayo	Y	Y	Y	N	Y	Y	N	N	N	N
Brock/PanCan	Y	Y	Y	Y	Y	Y	N	N	N	N
PKUPH (Peking Univ.)	Y	Y	Y	Y	Y	Y	N	N	N	N
Fudan/Zhongshan	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: Y = included in the model; N = not included.

Brock/PanCan 模型(Brock)则基于加拿大大规模低剂量 CT (LDCT)筛查人群(Pan-Canadian Early Detection 项目)构建, 因此更贴合现代筛查环境、适用于体检或筛查中发现的肺结节, 尤其是小结节和亚实性/磨玻璃(subsolid/GGN)结节[19]。

中国本土模型常结合本地结节特点(高 GGN 比例)、临床人群特点(如亚洲非吸烟腺癌多)、并在某些多参数模型中引入更多 CT 定量特征或结合影像学 + 临床 + 生物标志(若有), 因而对国内结节评估更有潜在优势[16] [17]。

3.2. 模型构建人群与适用场景的差异

Mayo Clinic 模型(Mayo)建立于 20 世纪 90 年代美国的一组偶发单发肺结节(indeterminate SPN)患者队列, 其结节大小范围为 4~30 mm, 样本主要为吸烟者, 具有代表当时西方临床人群的特点。该模型适合传统门诊偶发结节的风险评估[3]。

Brock/PanCan 模型(Brock)则基于加拿大大规模低剂量 CT (LDCT)筛查人群(Pan-Canadian Early Detection 项目)构建, 因此更贴合现代筛查环境、适用于体检或筛查中发现的肺结节, 尤其是小结节和亚实性/磨玻璃(subsolid/GGN)结节[19]。

中国本土模型(例如 PKUPH 模型)和一些多参数模型则基于东方人群(高 GGN 比例、非吸烟腺癌高发)的临床/外科病人队列, 有助于提高模型对中国人群的适配性[16]。

不同模型对应不同的人群与场景——Mayo 适合传统偶发结节; Brock 适合筛查环境; 中国本土模型适合国内特殊流行病学背景。

3.3. 预测性能(AUC、敏感性、特异性)的比较

Mayo 模型原始报告 AUC \approx 0.83, 在其实验队列中表现良好[3]。对比而言, 多中心评估与 meta-分析发现 Brock 模型总体表现优异, 2024 年一项 meta 分析汇总其 52 项研究(85,558 名患者)的 pooled AUC 为 0.796 (95% CI: 0.771~0.820), 敏感性约 0.82、特异性约 0.80; 对于 1~8 mm 小结节池(full model)的 AUC 可达到 0.927, 说明其在小结节/筛查人群中识别恶性结节能力较强[25]。然而, 这项 meta 分析也指出 Brock 模型在亚洲人群(尤其是亚实性结节)中性能下降(AUC \approx 0.741), 提示跨种族/种群应用时需要谨慎[25]。对于中国本土模型(如 PKUPH)以及新建立的本土多参数模型(尤其考虑磨玻璃/亚实性结节), 虽然公开的大规模前瞻性验证相对少, 但已有研究和现实临床经验表明, 其在国内人群中可能优于西方模型[16] [17]。

Brock 模型在西方和筛查环境中表现稳定; 但在亚洲人群和亚实性结节中, 其性能可能受限, 本土模型具有潜在优势, 但仍需更多标准化验证。

3.4. 临床实用性与局限性

不同模型适用于不同场景(见表 2)。

Table 2. Target population, strengths and limitations of representative models

表 2. 各模型适用场景、优势与局限

Model	Population	Applicable nodule types	Strengths	Limitations
Mayo	Western high-risk smokers	Solid nodules	Simple and easy to use	Limited for subsolid/GGN
Brock/PanCan	Screening cohort	Solid, part-solid, GGN	Good overall accuracy; comprehensive variables	More complex; calculator often needed
PKUPH	Chinese clinical cohort	Including GGN	Better fit for East-Asian epidemiology	External validation still limited
Fudan/Zhongshan	Chinese surgical cohort	Subsolid nodules	Sensitive for adenocarcinoma-related signs	Potential selection bias; less suited for community screening

Note: GGN = ground-glass nodule.

Mayo 模型: 简单、变量少、计算方便—适合基层医院或资源有限场景。但对亚实性结节/筛查结节

(LDCT 发现)的适用性不足[3]。

Brock 模型: 适合筛查与体检环境、能够处理复杂结节类型(subsolid、多个结节); 但其对亚洲人群的泛化性不足, 且计算复杂, 需要工具支持[18] [19]。

中国本土模型/多参数模型: 更贴合国内结节谱(高 GGN、高腺癌比例)、更适合临床决策——特别是在考虑结节管理(随访 vs 手术)的情境中。但局限是多数为回顾/手术队列, 存在选择偏倚(高恶性率), 缺乏广泛前瞻性筛查队列验证[16] [17]。

模型选择应结合结节类型、人群特征与机构资源——没有“通用最佳模型”, 而是“适合场景 + 人群 + 目的”的模型。上述模型的分析, 主要基于其纳入的宏观临床与形态学变量。然而, 在临床实践中, 尤其是对于亚实性结节(subsolid nodules)这类影像表现特殊、恶性风险分层需求迫切的类型, 一些精细化、可量化的影像学特征已成为不可或缺的补充判断工具。其中, 实性成分占比(Consolidation-to-Tumor Ratio, CTR)作为评估亚实性结节侵袭潜力的核心指标, 在近年来的临床研究与模型构建中受到了广泛关注[26]。因此, 在系统比较各类风险评估模型后, 有必要对这一关键影像学特征进行专门探讨。

3.5. 关键影像学风险特征: 包括 CTR (Consolidation-to-Tumor Ratio)

Consolidation-to-Tumor Ratio (CTR)是亚实性肺结节(subsolid nodules)中用于量化实性成分比例的重要二维指标。CTR 定义为薄层 CT 肺窗下实性成分最大径与肿瘤总体最大径的比值, 已广泛应用于早期肺腺癌分级与术前决策。Suzuki 等在前瞻性研究中证实 CTR 与病理侵袭程度密切相关, $CTR \geq 0.5$ 的结节显著更可能为浸润性腺癌(IAC), 而 $CTR < 0.25$ 多为 AIS/MIA [26] [27]。大规模 Meta 分析显示, 高 CTR 与较差 OS 和 DFS 显著相关, 进一步支持 CTR 作为稳定预后指标的价值[28]。

在用于分级数据与恶性风险模型时, CTR 被视为最具可重复性的分类特征, 因此常用于构建良恶性区分模型, 以及 AIS/MIA/IAC 三分类预测模型[27] [28]。部分研究指出, 在 part-solid 结节中, CTR 与恶性程度仍呈显著线性关联, 即使其独立预后能力受队列影响亦保持预测价值[28] [29]。此外, 有研究显示将 CTR 与放射组学特征结合可显著提高对浸润性腺癌的识别性能[30]。

尽管少数研究提示 CTR 在多变量模型中可能不具独立预后地位[31], 但其测量简便、易标准化, 仍是当前恶性结节诊断分类中最核心、最常用的分级特征之一。

3.6. 模型校准度(Calibration)及其临床意义

除区分能力(discrimination, 如 AUC)外, 模型的校准度(calibration)同样是评价预测模型临床实用性的重要指标[32]。校准度反映的是模型预测风险与真实发生概率之间的一致程度。例如, 当模型预测某类结节恶性概率为 20%时, 若实际约有 20%的患者确实为恶性, 则说明模型具有良好校准度。

在实际应用中, 一些模型虽然 AUC 较高, 但在外部人群验证中常出现风险高估或低估现象, 即校准度下降[32]。这种情况多与构建模型时所使用的人群结构差异有关, 例如吸烟比例、结节类型分布以及筛查与临床就诊人群之间的差异。当模型被应用于不同地区或不同流行病学背景的人群时, 预测风险可能系统性偏高或偏低。

校准度不佳对临床决策可能产生重要影响。例如, 若模型高估恶性风险, 可能导致患者接受过度随访、PET-CT 检查甚至不必要的穿刺或手术干预, 增加医疗成本与患者负担; 相反, 若模型低估风险, 则可能延误诊断, 使患者错过早期治疗机会。因此, 在模型推广应用前进行外部验证和再校准(recalibration)对于保障临床安全具有重要意义[32]。

未来模型研究除报告 AUC 等区分性能外, 还应系统报告校准曲线、Brier 评分等校准指标, 并针对不同地区或筛查环境进行模型更新或动态校准, 以提高真实世界应用的可靠性[32]。

4. 新兴方法

4.1. 影像组学(Radiomics)

影像组学通过从 CT 图像中高吞吐量提取密度、纹理、形状及波形转换等数百至上千种特征,以量化结节的微观影像信息,从而捕捉传统肉眼无法识别的细微差异。通过特征筛选与建模(如 LASSO、随机森林、SVM 等),影像组学模型可用于预测肺结节的良恶性、基因突变特征以及生长趋势。多项研究表明,影像组学模型在结节风险分层中表现出良好准确性,尤其是在磨玻璃结节(GGNs)和早期肺癌鉴别方面,具有较高临床潜力[33] [34]。

4.2. 深度学习(CNN 模型)

深度学习,尤其是卷积神经网络(CNN),能够自动从原始 CT 图像中学习空间层级特征,无需人工特征工程。CNN 模型不仅能识别结节边缘、密度、纹理结构,还可以捕捉深层非线性模式,因此在恶性风险预测中表现优于传统统计模型。有研究显示,基于 3D CNN 的系统能达到接近甚至超过资深放射科医师的诊断性能,且在大型筛查数据集中具有良好的泛化性。深度学习的优势在于其端到端设计,但同时也面临可解释性较弱的问题[21] [35]。

4.3. 多模态融合模型(临床 + CT + AI)

近年来,多模态融合模型逐渐成为肺结节风险评估的前沿方向。相比单一影像或单一临床变量,融合模型可将年龄、性别、吸烟史、既往肿瘤史等临床信息与 CT 影像特征、影像组学或 CNN 特征进行综合分析,从而显著提升恶性预测能力。例如,有研究采用深度学习提取影像特征并与 Brock 模型或 Mayo 模型的临床变量结合,表现出优于任一单独模型的诊断性能。此类模型能更好适配真实临床场景,特别是在筛查人群及混合病灶上具有更高决策价值[36] [37]。

4.4. 现阶段限制(可解释性、推广性、数据偏移)

尽管 AI 模型在肺结节风险预测中展现出巨大潜力,但仍存在多项限制。首先,深度学习模型普遍缺乏可解释性,难以明确判断依据,不利于临床信任与验证。其次,模型推广性受到训练数据结构、采集设备、扫描参数等因素影响显著,跨机构与跨设备的泛化能力仍有限。此外,真实世界环境中常存在数据偏移问题,如民族差异、病种分布偏差或筛查 vs 就诊人群差异,可能导致模型性能下降。因此,提高可解释性、建立大规模多中心数据库并优化模型稳健性,是未来发展的关键方向[38]-[40]。

4.5. 肿瘤标志物融合模型(CT + 临床 + 生物标志物)

随着精准医学的发展,一些研究开始尝试将血清肿瘤标志物(如 CEA、CYFRA 21-1、NSE、ProGRP)纳入肺结节风险预测模型中,以提高复杂病灶的良恶性鉴别能力[22]。肿瘤标志物能够反映潜在的生物学行为,与影像特征具有互补价值,尤其适用于那些影像表现不典型、边界模糊或难以定性的肺结节。

多项研究表明,将 CEA 或 CYFRA 21-1 与临床信息及影像特征结合,可显著提高模型的 AUC,提高小结节及磨玻璃结节的恶性预测能力。如 Kammer [22] 等人构建的集成模型,结合了临床信息、CT 特征及 CEA、CYFRA 21-1 等标志物,在鉴别不确定性肺结节方面显示出优于传统临床模型(如 Mayo 模型)的 AUC 值,提示生物标志物在未来风险评估体系中可能发挥更大作用[22] [40]。

近年来,随着液体活检技术的发展,新型分子生物标志物在肺结节良恶性鉴别中的潜在价值逐渐受到关注。循环肿瘤 DNA (circulating tumor DNA, ctDNA)检测可通过外周血捕获来源于肿瘤细胞的 DNA 片段,从而反映潜在的恶性克隆活动。一些研究发现,即使在影像学表现为小结节或早期肺癌阶段,部分

患者体内亦可检测到特异性突变 ctDNA 信号, 为早期恶性识别提供可能[41]。此外, DNA 甲基化检测作为另一种重要液体活检方向, 通过分析外周血中肿瘤相关基因的异常甲基化模式, 可提高早期肺癌的检出率, 并在区分良恶性肺结节中显示出优于传统血清标志物的潜力[42]。

相比传统 CEA 或 NSE 等标志物主要反映肿瘤负荷的特点, ctDNA 与甲基化检测更接近肿瘤发生发展的分子层面改变, 具有更强的前瞻性和早期检测潜力[41][42]。未来, 结合影像学特征、临床信息及液体活检分子标志物的多模态模型, 可能进一步提升对影像学不确定肺结节的鉴别能力, 并减少不必要的侵入性检查。但目前相关检测成本较高, 且在筛查人群中的应用仍需更多前瞻性研究验证。

5. 未来方向(Discussion/Perspective)

随着人工智能及大数据技术的发展, 肺结节风险评估正朝着更精准、更自动化与可解释的方向演进。未来模型将更强调多模态融合, 结合 CT 影像、临床变量、生物标志物、基因组学及纵向随访数据, 构建更全面的恶性风险评估框架[33]。深度学习系统也将向端到端全流程模型发展, 实现从原始 CT 输入到恶性概率输出的自动化推断, 同时减少人工特征选择带来的偏差[21]。

与此同时, 提升模型的可解释性(explainability)是未来关键方向之一。使用显著性图、注意力机制或可视化热图的方式, 有助于临床医师理解 AI 决策过程, 提高其可接受性[43]。此外, 模型在真实世界中面临的数据异质性、成像协议差异及人群偏差, 仍然是性能下降的重要因素, 因此构建更大规模、多中心、多族群的数据集, 并加强外部验证, 将是提升模型可推广性(generalizability)的关键[39] (Park 2021; Obermeyer 2019)。

未来指南与临床路径也将更系统地引入 AI 工具, 使肺结节评估逐步从“影像主导”迈向 影像 + 生物学 + AI 综合判断 的新范式, 为精准诊断提供更强支持。

6. 结论(Conclusion)

肺结节的良恶性鉴别仍然是临床管理的重要挑战, 特别是在 LDCT 广泛应用后, 检出率大幅增加, 而影像特征本身常缺乏特异性。传统模型如 Mayo 模型、Brock (PanCan)模型、中国人群模型及复旦中山医院多参数模型, 为风险评估提供了可量化、可重复的工具, 在临床决策中得到广泛应用[3][16][17][19]。

近年来, 影像组学、深度学习以及临床-影像多模态融合模型进一步推动风险评估精细化, 显示出优于传统统计模型的预测能力[22][34]。它们能够自动整合高维影像特征及复杂临床信息, 为早期肺癌的筛查与鉴别提供更精准的支持。然而, 这些模型仍受到数据偏移、外部推广性不足及可解释性不强的限制, 尚未完全满足大规模临床应用要求[32][38]。同时, 新型液体活检标志物及模型校准度优化等方向的引入, 亦有望进一步提升预测模型在真实临床环境中的可靠性与适用性。

未来的发展方向应包括构建更大规模多中心数据库、提高模型透明度、强化外部验证, 并推动人工智能工具与临床指南的深度融合。总体而言, 肺结节风险评估正迈向数据驱动与智能化的新阶段, 有望进一步推动肺癌早筛与精准诊疗的临床实践。

参考文献

- [1] Siegel, R.L., Miller, K.D., Wagle, N.S. and Jemal, A. (2023) Cancer Statistics, 2023. *CA: A Cancer Journal for Clinicians*, **73**, 17-48. <https://doi.org/10.3322/caac.21763>
- [2] Lillington, G.A. (1991) Management of Solitary Pulmonary Nodules. *Disease-a-Month*, **37**, 269-318. [https://doi.org/10.1016/s0011-5029\(05\)80012-4](https://doi.org/10.1016/s0011-5029(05)80012-4)
- [3] Swensen, S.J., Silverstein, M.D., Ilstrup, D.M., et al. (1997) The Probability of Malignancy in Solitary Pulmonary Nodules. Application to Small Radiologically Indeterminate Nodules. *Archives of Internal Medicine*, **157**, 849-855.
- [4] Wyker, A., Sharma, S. and Henderson, W.W. (2025) Solitary Pulmonary Nodule. StatPearls Publishing.

- [5] Khan, A.N., Al-Jahdali, H.H., Irion, K.L., Arabi, M. and Koteyar, S.S. (2011) Solitary Pulmonary Nodule: A Diagnostic Algorithm in the Light of Current Imaging Technique. *Avicenna Journal of Medicine*, **1**, 39-51. <https://doi.org/10.4103/2231-0770.90915>
- [6] Mazzone, P.J. and Lam, L. (2022) Evaluating the Patient with a Pulmonary Nodule. *Journal of the American Medical Association*, **327**, 264. <https://doi.org/10.1001/jama.2021.24287>
- [7] Aberle, D.R., Adams, A.M., Berg, C.D., et al. (2011) Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. *The New England Journal of Medicine*, **365**, 395-409.
- [8] MacMahon, H., Naidich, D.P., Goo, J.M., Lee, K.S., Leung, A.N.C., Mayo, J.R., et al. (2017) Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017. *Radiology*, **284**, 228-243. <https://doi.org/10.1148/radiol.2017161659>
- [9] Henschke, C.I., Yankelevitz, D.F., Libby, D.M., et al. (2006) Survival of Patients with Stage I Lung Cancer Detected on CT Screening. *The New England Journal of Medicine*, **355**, 1763-1771.
- [10] Godoy, M.C.B. and Naidich, D.P. (2009) Subsolid Pulmonary Nodules and the Spectrum of Peripheral Adenocarcinomas of the Lung: Recommended Interim Guidelines for Assessment and Management. *Radiology*, **253**, 606-622. <https://doi.org/10.1148/radiol.2533090179>
- [11] Kim, Y.W., Kwon, B.S., Lim, S.Y., Lee, Y.J., Park, J.S., Cho, Y., et al. (2021) Lung Cancer Probability and Clinical Outcomes of Baseline and New Subsolid Nodules Detected on Low-Dose CT Screening. *Thorax*, **76**, 980-988. <https://doi.org/10.1136/thoraxjnl-2020-215107>
- [12] Yankelevitz, D.F. and Henschke, C.I. (2021) Overdiagnosis in Lung Cancer Screening. *Translational Lung Cancer Research*, **10**, 1136-1140. <https://doi.org/10.21037/tlcr-20-736>
- [13] Nair, A., Bartlett, E.C., Walsh, S.L.F., Wells, A.U., Navani, N., Hardavella, G., et al. (2018) Variable Radiological Lung Nodule Evaluation Leads to Divergent Management Recommendations. *European Respiratory Journal*, **52**, Article 1801359. <https://doi.org/10.1183/13993003.01359-2018>
- [14] Azour, L., Moore, W.H., O'Donnell, T., Truong, M.T., Babb, J., Niu, B., et al. (2022) Inter-Reader Variability of Volumetric Subsolid Pulmonary Nodule Radiomic Features. *Academic Radiology*, **29**, S98-S107. <https://doi.org/10.1016/j.acra.2021.01.026>
- [15] Bai, C., Choi, C.M., Chu, C.M., et al. (2016) Evaluation of Pulmonary Nodules: Clinical Practice Consensus Guidelines for Asia. *Chest*, **150**, 877-893.
- [16] 李运, 陈克终, 隋锡朝, 等. 孤立性肺结节良恶性判断数学预测模型的建立[J]. 北京大学学报(医学版), 2011, 43(3): 450-454.
- [17] Hu, F., Huang, H., Jiang, Y., Feng, M., Wang, H., Tang, M., et al. (2021) Discriminating Invasive Adenocarcinoma among Lung Pure Ground-Glass Nodules: A Multi-Parameter Prediction Model. *Journal of Thoracic Disease*, **13**, 5383-5394. <https://doi.org/10.21037/jtd-21-786>
- [18] Tammemagi, M.C., Schmidt, H., Martel, S., McWilliams, A., Goffin, J.R., Johnston, M.R., et al. (2017) Participant Selection for Lung Cancer Screening by Risk Modelling (the Pan-Canadian Early Detection of Lung Cancer [PanCan] Study): A Single-Arm, Prospective Study. *The Lancet Oncology*, **18**, 1523-1531. [https://doi.org/10.1016/s1470-2045\(17\)30597-1](https://doi.org/10.1016/s1470-2045(17)30597-1)
- [19] McWilliams, A., Tammemagi, M.C., Mayo, J.R., Roberts, H., Liu, G., Soghrati, K., et al. (2013) Probability of Cancer in Pulmonary Nodules Detected on First Screening CT. *New England Journal of Medicine*, **369**, 910-919. <https://doi.org/10.1056/nejmoa1214726>
- [20] Xiao, F., Yu, Q., Zhang, Z., et al. (2019) Establishment and Verification of a Novel Predictive Model of Malignancy for Non-Solid Pulmonary Nodules. *Chinese Journal of Lung Cancer*, **22**, 26-33.
- [21] Ardila, D., Kiraly, A.P., Bharadwaj, S., Choi, B., Reicher, J.J., Peng, L., et al. (2019) End-to-End Lung Cancer Screening with Three-Dimensional Deep Learning on Low-Dose Chest Computed Tomography. *Nature Medicine*, **25**, 954-961. <https://doi.org/10.1038/s41591-019-0447-x>
- [22] Kammer, M.N., Lakhani, D.A., Balar, A.B., Antic, S.L., Kussrow, A.K., Webster, R.L., et al. (2021) Integrated Biomarkers for the Management of Indeterminate Pulmonary Nodules. *American Journal of Respiratory and Critical Care Medicine*, **204**, 1306-1316. <https://doi.org/10.1164/rccm.202012-4438oc>
- [23] Graham, R.N.J., Baldwin, D.R., Callister, M.E.J. and Gleeson, F.V. (2016) Return of the Pulmonary Nodule: The Radiologist's Key Role in Implementing the 2015 BTS Guidelines on the Investigation and Management of Pulmonary Nodules. *The British Journal of Radiology*, **89**, Article 20150776. <https://doi.org/10.1259/bjr.20150776>
- [24] Wu, Z., Huang, T., Zhang, S., Cheng, D., Li, W. and Chen, B. (2021) A Prediction Model to Evaluate the Pretest Risk of Malignancy in Solitary Pulmonary Nodules: Evidence from a Large Chinese Southwestern Population. *Journal of Cancer Research and Clinical Oncology*, **147**, 275-285. <https://doi.org/10.1007/s00432-020-03408-2>
- [25] Chen, S., Lin, W.L., Liu, W.T., Zou, L.Y., Chen, Y. and Lu, F. (2025) Pulmonary Nodule Malignancy Probability: A

- Meta-Analysis of the Brock Model. *Clinical Radiology*, **82**, Article 106788. <https://doi.org/10.1016/j.crad.2024.106788>
- [26] 黄汉清, 叶波. 肿瘤实性成分占比在早期周围型肺癌诊疗中的研究进展[J]. 中国肺癌杂志, 2022, 25(10): 764-770.
- [27] Suzuki, K., Koike, T., Asakawa, T., Kusumoto, M., Asamura, H., Nagai, K., *et al.* (2011) A Prospective Radiological Study of Thin-Section Computed Tomography to Predict Pathological Noninvasiveness in Peripheral Clinical IA Lung Cancer (Japan Clinical Oncology Group 0201). *Journal of Thoracic Oncology*, **6**, 751-756. <https://doi.org/10.1097/jto.0b013e31821038ab>
- [28] Wu, Y., Song, W., Wang, D., Chang, J., Wang, Y., Tian, J., *et al.* (2023) Prognostic Value of Consolidation-to-Tumor Ratio on Computed Tomography in NSCLC: A Meta-Analysis. *World Journal of Surgical Oncology*, **21**, Article No. 190. <https://doi.org/10.1186/s12957-023-03081-y>
- [29] Yoon, D.W., Kim, C.H., Hwang, S., Choi, Y., Cho, J.H., Kim, H.K., *et al.* (2022) Reappraising the Clinical Usability of Consolidation-to-Tumor Ratio on CT in Clinical Stage IA Lung Cancer. *Insights into Imaging*, **13**, Article No. 103. <https://doi.org/10.1186/s13244-022-01235-2>
- [30] Wang, Y., Lyu, D., Yu, D., Hu, S., Ma, Y., Huang, W., *et al.* (2024) Intratumoral and Peritumoral Radiomics Combined with Computed Tomography Features for Predicting the Invasiveness of Lung Adenocarcinoma Presenting as a Subpleural Ground-Glass Nodule with a Consolidation-to-Tumor Ratio $\leq 50\%$. *Journal of Thoracic Disease*, **16**, 5122-5137. <https://doi.org/10.21037/jtd-24-243>
- [31] Kim, H., Goo, J.M., Kim, Y.T. and Park, C.M. (2019) Consolidation-to-Tumor Ratio and Tumor Disappearance Ratio Are Not Independent Prognostic Factors for the Patients with Resected Lung Adenocarcinomas. *Lung Cancer*, **137**, 123-128. <https://doi.org/10.1016/j.lungcan.2019.09.014>
- [32] Varoquaux, G. and Cheplygina, V. (2022) Machine Learning for Medical Imaging: Methodological Failures and Recommendations for the Future. *npj Digital Medicine*, **5**, Article No. 48. <https://doi.org/10.1038/s41746-022-00592-y>
- [33] Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R.G.P.M., Granton, P., *et al.* (2012) Radiomics: Extracting More Information from Medical Images Using Advanced Feature Analysis. *European Journal of Cancer*, **48**, 441-446. <https://doi.org/10.1016/j.ejca.2011.11.036>
- [34] Aerts, H.J.W.L., Velazquez, E.R., Leijenaar, R.T.H., Parmar, C., Grossmann, P., Carvalho, S., *et al.* (2014) Decoding Tumour Phenotype by Noninvasive Imaging Using a Quantitative Radiomics Approach. *Nature Communications*, **5**, Article No. 4006. <https://doi.org/10.1038/ncomms5006>
- [35] Setio, A.A.A., Traverso, A., de Bel, T., Berens, M.S.N., Bogaard, C.V.D., Cerello, P., *et al.* (2017) Validation, Comparison, and Combination of Algorithms for Automatic Detection of Pulmonary Nodules in Computed Tomography Images: The LUNA16 Challenge. *Medical Image Analysis*, **42**, 1-13. <https://doi.org/10.1016/j.media.2017.06.015>
- [36] Hawkins, S., Wang, H., Liu, Y., Garcia, A., Stringfield, O., Krewer, H., *et al.* (2016) Predicting Malignant Nodules from Screening CT Scans. *Journal of Thoracic Oncology*, **11**, 2120-2128. <https://doi.org/10.1016/j.jtho.2016.07.002>
- [37] Venkadesh, K.V., Setio, A.A.A., Schreuder, A., Scholten, E.T., Chung, K., W. Wille, M.M., *et al.* (2021) Deep Learning for Malignancy Risk Estimation of Pulmonary Nodules Detected at Low-Dose Screening CT. *Radiology*, **300**, 438-447. <https://doi.org/10.1148/radiol.2021204433>
- [38] Dutta, S., Prakash, P. and Matthews, C.G. (2020) Impact of Data Augmentation Techniques on a Deep Learning Based Medical Imaging Task. *Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications*, Houston, 2 March 2020, 113180M. <https://doi.org/10.1117/12.2549806>
- [39] Obermeyer, Z., Powers, B., Vogeli, C. and Mullainathan, S. (2019) Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations. *Science*, **366**, 447-453. <https://doi.org/10.1126/science.aax2342>
- [40] Yoo, J., Lee, M.W., Lee, D.H., Lee, J. and Han, J.K. (2020) Evaluation of a Serum Tumour Marker-Based Recurrence Prediction Model after Radiofrequency Ablation for Hepatocellular Carcinoma. *Liver International*, **40**, 1189-1200. <https://doi.org/10.1111/liv.14406>
- [41] Abbosh, C., Birkbak, N.J., Wilson, G.A., Jamal-Hanjani, M., Constantin, T., Salari, R., *et al.* (2017) Phylogenetic ctDNA Analysis Depicts Early-Stage Lung Cancer Evolution. *Nature*, **545**, 446-451. <https://doi.org/10.1038/nature22364>
- [42] Liu, M.C., Oxnard, G.R., Klein, E.A., Swanton, C., Seiden, M.V., Liu, M.C., *et al.* (2020) Sensitive and Specific Multi-Cancer Detection and Localization Using Methylation Signatures in Cell-Free DNA. *Annals of Oncology*, **31**, 745-759. <https://doi.org/10.1016/j.annonc.2020.02.011>
- [43] Rudin, C. (2019) Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence*, **1**, 206-215. <https://doi.org/10.1038/s42256-019-0048-x>