

同济大学陈昶团队研究利用人工智能多组学技术开发早期肺癌无创诊断新方法

作者：writer 来源：自然·通讯

本文原地址：<https://www.iikx.com/news/progress/32012.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

近日，同济大学医学院、同济大学附属上海市肺科医院陈昶教授、赵德平教授牵头，利用人工智能技术开发了结合临床特征、影像组学特征以及表观修饰区域中的循环游离DNA片段组学特征的多组学诊断模型clinic-RadmC，提供了一种更有效且非侵入性的工具优化肺癌诊断，并进一步促进肺癌早期精准治疗干预。相关成果以“Integrated multiomics signatures to optimize the accurate diagnosis of lung cancer”为题发表于《自然·通讯》(Nature Communications)。

nature communications



Article

<https://doi.org/10.1038/s41467-024-55594-z>

Integrated multiomics signatures to optimize the accurate diagnosis of lung cancer

Received: 25 January 2024

Accepted: 14 December 2024

Published online: 02 January 2025

Check for updates

Mengmeng Zhao^{1,15}, Gang Xue^{2,3,15}, Bingxi He^{4,5,6,15}, Jiajun Deng¹, Tingting Wang⁷, Yifan Zhong¹, Shenghui Li¹, Yang Wang¹, Yiming He^{1,8}, Tao Chen¹, Jun Zhang⁹, Ziyue Yan⁹, Xinlei Hu^{2,3}, Liuning Guo¹⁰, Wendong Qu¹⁰, Yongxiang Song¹⁰, Minglei Yang¹¹, Guofang Zhao¹¹, Bentong Yu⁸, Minjie Ma¹², Lunxu Liu¹³, Xiwen Sun¹⁴, Yunlang She^{1,16}✉, Dan Xie^{2,3,16}✉, Deping Zhao^{1,16}✉ & Chang Chen^{1,16}✉

该项研究中，研究者纳入了来自5个不同临床中心的2,032名不确定性肺结节(IPL)参与患者，在随机分组后，训练集、验证集和内部测试集分别包含1030、344和344名参与者。外部测试集包含314名参与者。

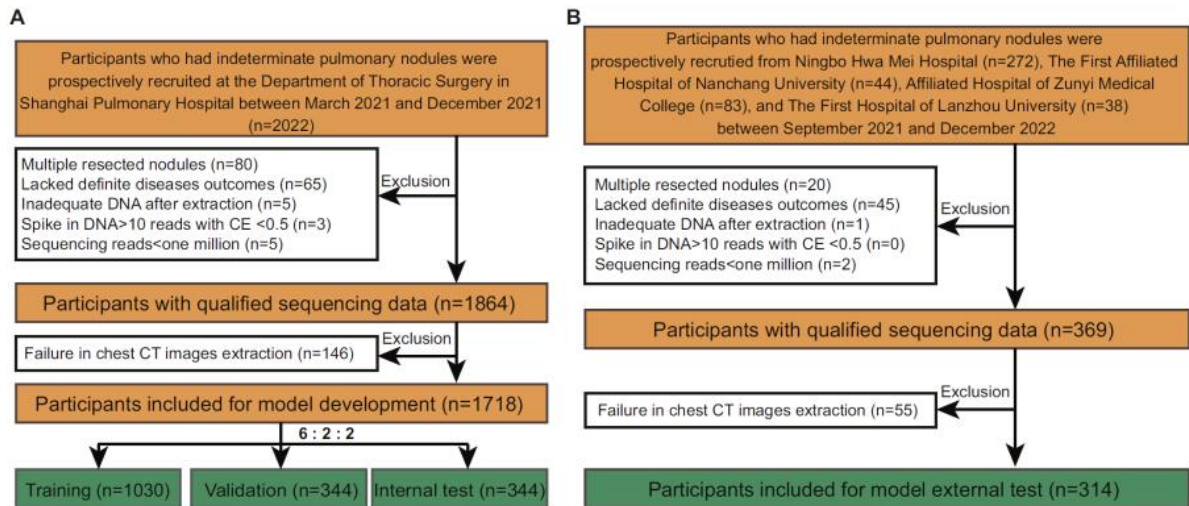
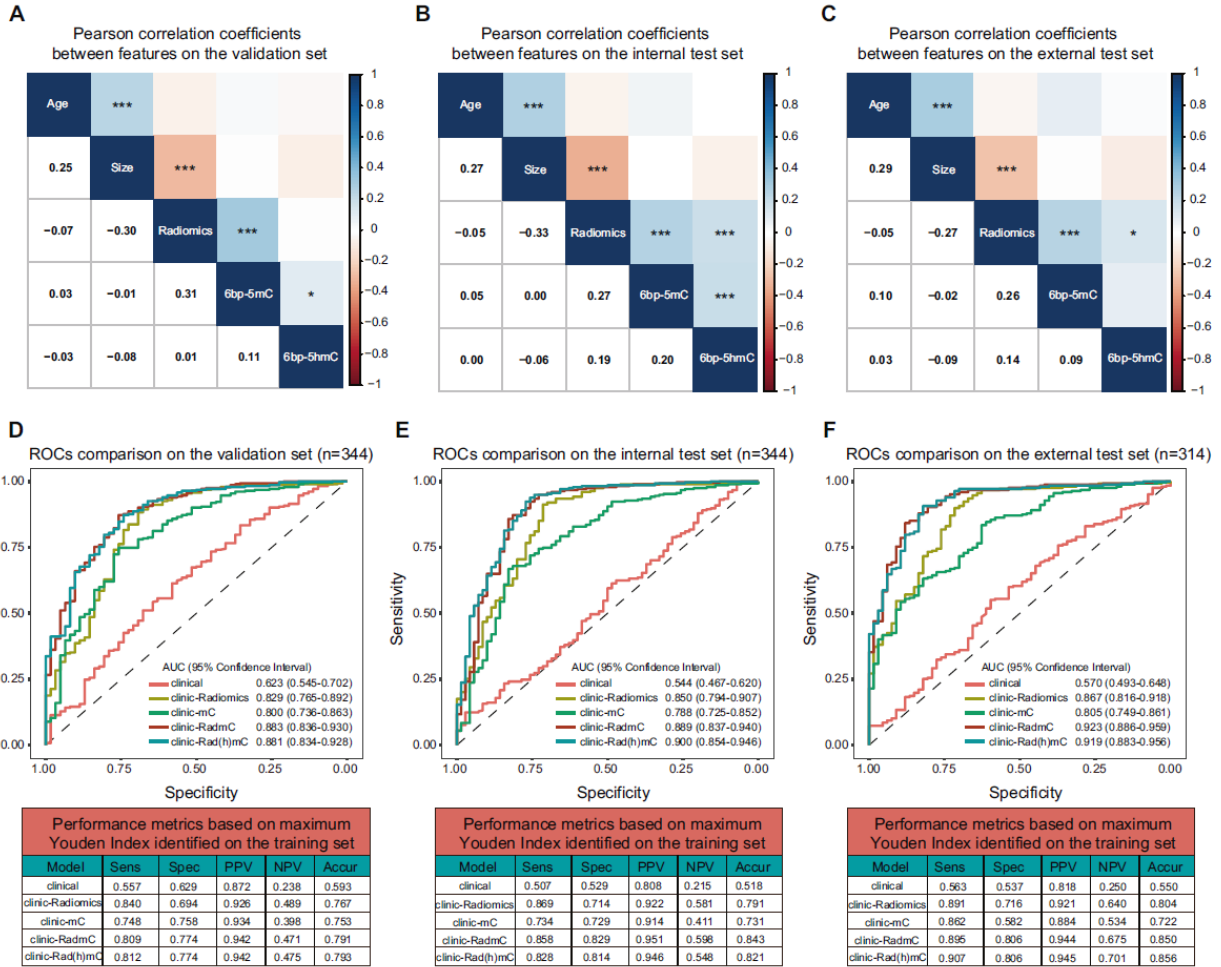
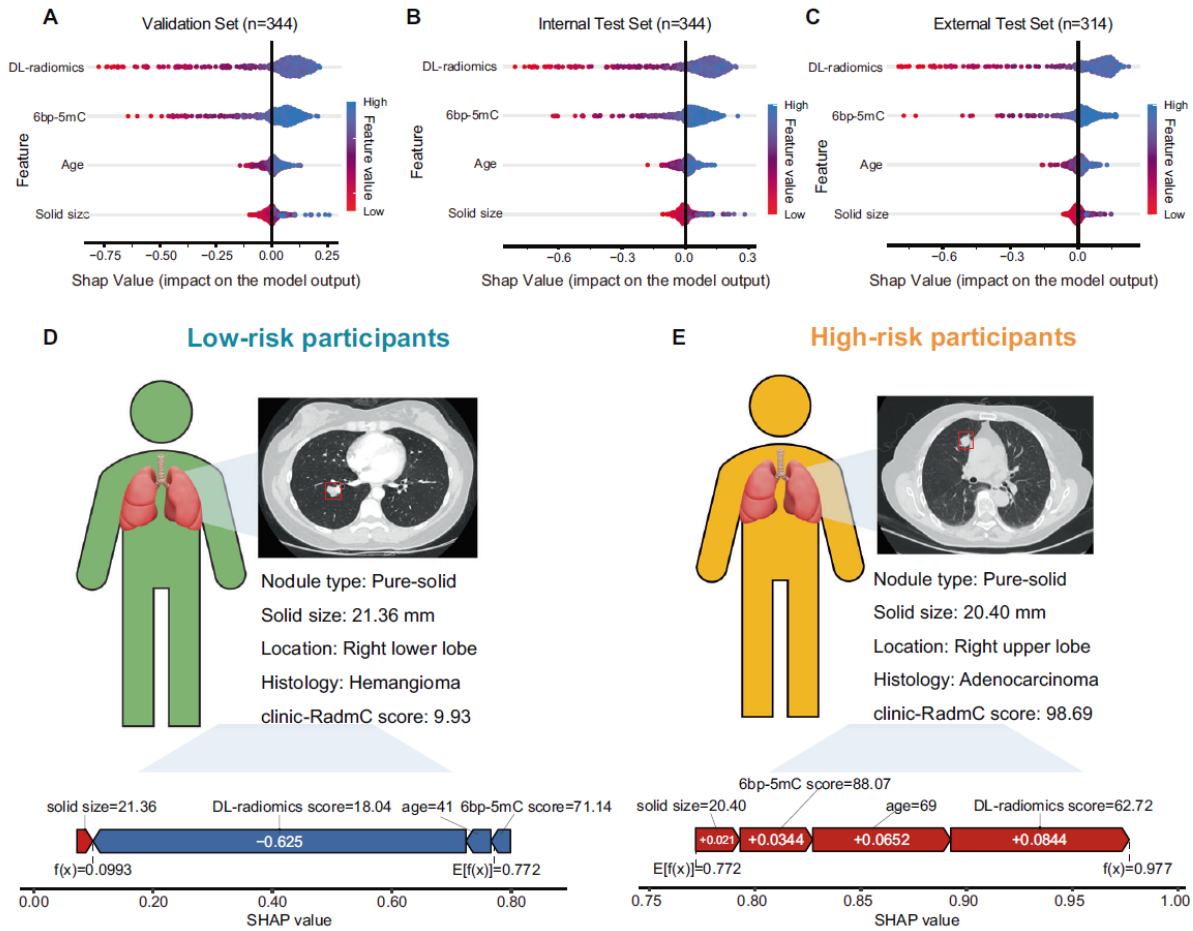


Fig. 1 | Flowchart illustrating the participants selection scheme. A participants selecting from Shanghai Pulmonary Hospital; **B** participants selecting from external 4 hospitals.

通过对影像组学、cfDNA不同表观修饰区域的片段组学进行分析，并结合建立的临床模型，发现多组学间存在诸多共性及互补性，并最终经筛选比较，建立了基于临床因素、深度学习影像组学模型及6bp-5mC模型的多组学模型clinic-RadmC。clinic-RadmC多组学模型在外部测试集上获得了0.923的曲线下面积(AUC)，在验证集中获得了显著提高的AUC值为0.883、在内部测试集中为0.889、在外部测试集中为0.923。校准曲线及决策曲线的结果均优于单一组学及仅结合临床特征与放射组学或5mC富集区域中的片段组学特征的模型(所有 $p < 0.050$)。



与既往研究相比，clinic-RadmC多组学模型有望提供更准确、更稳定、更具成本效益和无创的肺癌早期诊断工具。该研究具有以下优势：(1)前瞻性招募了不同临床中心的2,032名参与者，这是迄今为止该领域报告的最大样本队列。队列包含了临床上常见的磨玻璃结节和亚实性结节等各种常见的结节类型；(2)补充额外信息以支持cfDNA 6bp末端基序谱在5mC和5hmC富集区域对肺癌的诊断价值：团队研究了来自5mC测序数据的4bp和6bp末端基序的性能，发现其在区分恶性结节和良性结节方面均具有良好的性能(>0.80)。与既往报道的类似检测相比，4bp和6bp-5mC模型在肺癌检测方面都保持了相当且稳健的性能。此外，在研究团队之前的研究中，5hmC生物标志物已被用作肺癌检测的重要工具，然而在本研究中纳入甲基化信息后，从5hmC测序数据中添加诊断评分并不能为预测不确定性肺结节(IPL)的恶性肿瘤风险提供额外的改进。这些发现进一步证明，来自临床、影像组学和cfDNA 5mC测序数据的信息组合可以提供足够的性能，作为临床实践中诊断肺癌的更具成本效益的方法；(3)该研究中使用了人工智能技术，例如支持向量机、深度学习等来构建模型，并进行了模型的可解释性分析。将AI研究转化为临床工作流程的优化是这项最先进技术的最終目标，也是衡量AI模型临床价值的基准指标。研究团队对末端基序进行了功能分析，通过DeepLIFT探索基于深度学习的影像组学特征的重要性水平，并通过SHAP对clinical-RadmC模型中每个特征的重要性进行了排名，这些分析可能有助于提高预测模型在临床实践中的可信度和透明度。



综上，该研究基于迄今为止最大数量样本队列多维度数据，联合多学科团队构建的clinic-Radmc多组学模型，相较于单组学或其他多组学模型，是一种更高效、更稳健、更准确的肺癌诊断工具，为改善肺癌筛查项目和减轻全球肺癌负担提供了新途径。

同济大学附属上海市肺科医院陈昶教授、赵德平教授、余云浪副研究员，四川大学附属华西医院谢丹教授为论文共同通讯作者。同济大学附属上海市肺科医院赵蒙蒙博士、四川大学附属华西医院薛刚博士、北京航空航天大学何秉羲博士后为论文共同第一作者。

论文链接：<https://www.nature.com/articles/s41467-024-55594-z>

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发