

基于临床病理特征和影像学及血清生物指标分析对肺结节性质预测模型的构建与验证

袁 瑞^a, 汪桃利^b, 余文辉^a, 张书楠^a, 罗胜华^a, 李运雷^a, 王向荣^c, 王家传^d, 郭海涛^e (广州中医药大学第四临床医学院 / 深圳市中医院 a. 检验科; b. 肿瘤科; c. 放射科; d. 病理科; e. 胸外科, 广东深圳 518033)

摘要: **目的** 基于临床病理特征和影像学及血清生物指标分析构建肺结节 (pulmonary nodules, PN) 性质预测模型并验证, 为肺癌早诊断和早治疗提供科学决策。**方法** 对2019年1月~2023年2月深圳市中医医院胸外科和肿瘤科816例行手术切除或肺活检病理诊断明确的PN患者进行回顾性分析。其中, 剔除不符合纳入标准者113例, 余下703例纳入研究。该研究基于PN患者临床病理特征 (年龄、性别、吸烟史、戒烟史、癌症家族史)、胸部影像学 (结节最大直径、病变位置、边界清晰、分叶、毛刺、空泡、血管集束征、钙化、空气支气管征、肺气肿、结节性质及胸膜凹陷、结节数量) 和血清癌胚抗原 (carcinoembryonic antigen, CEA)、细胞角蛋白19片段 (cytokeratin 19 fragment, CYFRA21-1)、鳞状细胞癌抗原 (squamous cell carcinoma antigen, SCCA), 将上述病例随机分为建模组 ($n=552$, 良性237例, 恶性315例) 和验证组 ($n=151$, 良性85例, 恶性66例)。首先, 对研究对象进行单变量分析以筛选有统计学意义的PN性质预测因子。然后, 进行多变量回归分析以筛选PN性质的独立预测因子。最后, 采用logistic回归分析构建PN性质的预测模型。再将验证组数据分别代入该模型与梅奥诊所 (Mayo clinic, Mayo) 模型、退伍军人事务 (veterans affairs, VA) 模型、Brock大学 (Brock University, Brock) 模型、北京大学 (Peking University, PKU) 模型和广州医科大学 (Guangzhou Medical University, GZMU) 模型计算PN恶性概率, 绘制受试者工作特征 (receiver operating characteristic, ROC) 曲线。根据曲线下面积 (area under the curve, AUC) 比较各模型的诊断效能。**结果** 单变量分析筛选的具有统计学意义的变量包括年龄、癌症家族史、结节最大直径、结节性质、肺上叶、钙化、血管集束征、分叶、边界清晰、毛刺以及血清CEA, SCCA, CYFRA21-1等。多变量回归分析显示年龄、CEA, 边界、CYFRA21-1, SCCA, 肺上叶、结节最大直径、癌症家族史、毛刺、结节性质等为PN恶性的独立预测因子。该研究构建的PN性质预测模型方程如下: $f(x) = e^x / (1 + e^x)$, $X = (-6.318 + 0.020 \times \text{年龄} + 0.527 \times \text{CEA} - 0.928 \times \text{边界} + 0.294 \times \text{Cyfra21-1} + 0.294 \times \text{结节最大直径} + 1.220 \times \text{癌症家族史} + 0.573 \times \text{肺上叶} + 0.064 \times \text{SCCA} + 1.461 \times \text{毛刺} + 1.497 \times \text{结节性质})$ 。该模型与Mayo模型和VA模型比较, AUC (0.799 vs 0.659, 0.650) 差异具有统计学意义 ($Z=3.029, 2.638, P=0.003, 0.008$)。然而, 该模型与Brock模型、PKU模型、GZMU模型比较, AUC (0.799 vs 0.762, 0.773, 0.769) 差异无统计学意义 ($Z=1.063, 0.686, 0.757, P=0.288, 0.493, 0.449$)。**结论** 该研究构建的PN性质预测模型较为准确可靠, 可帮助临床实现早诊断和早干预, 值得推广应用。

关键词: 肺结节; 临床病理特征; 影像学; 生物指标; 预测模型

中图分类号: R446; R563 **文献标识码:** A **文章编号:** 1671-7414 (2024) 01-146-07

doi: 10.3969/j.issn.1671-7414.2024.01.027

Construction and Validation of A Prediction Model for Pulmonary Nodule Nature Based on Clinicopathological Features, Imaging and Serum Biomarkers

YUAN Rui^a, WANG Taoli^b, YU Wenhui^a, ZHANG Shunan^a, LUO Shenghua^a, LI Yunlei^a, WANG Xiangrong^c, WANG Jiachuan^d, GUO Haitao^e (a. Department of Clinical Laboratory; b. Department of Oncology; c. Department of Radiology; d. Department of Pathology; e. Department of Thoracic Surgery, Shenzhen Traditional Chinese Medicine Hospital / the Fourth Clinical College of Guangzhou University of Chinese Medicine, Guangdong Shenzhen 518033, China)

Abstract: Objective The study aimed to construct and validate a predictive model for pulmonary nodules (PN) nature based on clinicopathological features, imaging, and serum biomarkers, so as to provide scientific decision-making for early diagnosis and treatment of lung cancer. **Methods** A retrospective was performed on 816 PN patients with defined pathological diagnosis who received surgical resection analysis or lung biopsy in the Department of Thoracic Surgery and Oncology of Shenzhen Traditional Chinese Medicine Hospital from January 2019 to February 2023. Among them, 113 cases that did not meet the

基金项目: 深圳市科技计划项目 (JCYJ20190812180001770): β -细辛醚调控 LncRNA SNHG5/miR-205-5p/ PTEN 信号轴干预 EGFR-TKIs 获得性耐药相关机制研究。

作者简介: 袁瑞 (1988-), 男, 硕士研究生, 主管技师, 临床生化检验, E-mail: 13809891469@163.com。

通讯作者: 余文辉 (1964-), 男, 主任技师, 硕士生导师, 临床生化检验与肿瘤学研究, E-mail: whyuchina@163.com。

inclusion criteria were excluded, and the remaining 703 cases were included in the study. The study based on the clinicopathologic features (age, gender, smoking history, smoking cessation history and family history of cancer), chest imaging (maximum diameter of nodule, location of lesion, clear border, Lobulation, spiculation, vascular convergence sign, vacuole, calcification, air bronchial sign, emphysema, nodule type and pleural indentation, nodule number) and serum carcinoembryonic antigen (CEA), cytokeratin 19 fragment (CYFRA21-1), squamous cell carcinoma antigen (SCCA) in patients with PN. These cases were randomly divided into a modeling group ($n=552$, 237 benign, 315 malignant) and a validation group ($n=151$, 85 benign, 66 malignant). First, univariate analysis was performed to screen for statistically significant predictors of nodules nature. Then, multivariate regression analysis was performed to screen for independent predictors of nodules nature. Finally, the prediction model of PN nature was constructed by logistic regression analysis. Subsequently, the validation group data were entered into the proposed model and Mayo clinic (Mayo) model, veterans affairs (VA) model, Brock University (Brock) model, Peking University (PKU) model and Guangzhou Medical University (GZMU) model, respectively. PN malignancy probability was calculated. The receiver operating characteristic (ROC) curves were plotted. The diagnostic efficiency of each model was compared according to the area under the curve (AUC). **Results** There were statistically significant variables including age, family history of cancer, maximum nodule diameter, nodule type, upper lobe of lung, calcification, vascular convergence sign, lobulation, clear border, spiculation, and serum CEA, SCCA, CYFRA21-1 using univariate analysis. Multiple regression analysis showed that age, CEA, clear border, CYFRA21-1, SCCA, upper lobe of lung, maximum nodule diameter, family history of cancer, spiculation and nodule type were independent predictors of PN nature. The prediction model equation constructed in this study is as follows: $f(x) = e^x / (1 + e^x)$, $X = (-6.3188 + 0.0208 \times \text{Age} + 0.5274 \times \text{CEA} - 0.9284 \times \text{clear border} + 0.2946 \times \text{Cyfra21-1} + 0.294 \times \text{maximum nodule diameter} + 1.2201 \times \text{family history of cancer} + 0.5732 \times \text{upper lobe of lung} + 0.0648 \times \text{SCCA} + 1.4615 \times \text{Spiculation} + 1.4976 \times \text{nodule type})$. The AUC (0.799 vs 0.659, 0.650) of the proposed model was significantly higher compared with Mayo model and VA model, and there were statistically significant differences ($Z=3.029, 2.638, P=0.003, 0.008$). However, compared with Brock model, PKU model and GZMU model, the differences of AUC (0.799 vs 0.762, 0.773, 0.769) were not statistically significant ($Z=1.063, 0.686, 0.757, P=0.288, 0.493, 0.449$). **Conclusion** The prediction model for PN nature established in this study is accurate and reliable, which can help clinics with early diagnosis and early intervention, and this prediction model deserves to be popularized.

Keywords: pulmonary nodules; clinicopathological features; imaging; biomarkers; prediction model

全球癌症 (GLOBOCAN) 2020 数据显示, 中国肺癌发病数和死亡数分别占全球的 37.0% 和 39.8%, 肺癌已成为全球恶性肿瘤发病率及癌症相关死亡的主要原因, 严重威胁人类健康^[1]。临床试验已证实低剂量 CT (Low-dose computerized tomography, LDCT) 筛查可降低肺癌相关死亡率^[2], 提高肺癌早期检出率, 降低晚期肺癌发病率^[3]。肺癌防治已经成为我国恶性肿瘤防控面临的重大挑战。肺结节 (pulmonary nodules, PN) 性质预测模型, 旨在帮助临床对 PN 精准管理, 指导 PN 风险分层, 实现肺癌早诊断。常见模型如梅奥诊所 (Mayo Clinic, Mayo) 模型^[4]、退伍军人事务 (Veterans Affairs, VA) 模型^[5]、Brock 大学 (Brock University, Brock) 模型^[6]和北京大学 (Peking University, PKU) 模型^[7]等是基于患者临床特征和影像学特征构建的, 并未纳入生物指标进行研究。广州医科大学 (Guangzhou Medical University, GZMU) 模型^[8]只包含一种生物指标细胞角蛋白 19 片段 (cytokeratin 19 fragment, Cyfra 21-1)。目前国内尚缺少经过可靠外部验证且预测准确度高的模型。因此, 创建新的 PN 性质预测模型显得尤为迫切。本研究旨在综合患者临床病

理特征、影像信息和生物指标构建准确可靠的 PN 性质预测模型, 协助医生对 PN 进行风险分层, 制定个体化诊疗方案。

1 材料和方法

1.1 研究对象 收集 2019 年 1 月 ~ 2023 年 2 月在深圳市中医院胸外科和肿瘤科接受手术切除或肺活检, 明确病理诊断的 816 例 PN 患者病历资料进行回顾性分析。入选标准及排除标准参考中国肺癌低剂量螺旋 CT 筛查指南 (2018 年版)^[9]。剔除不符合纳入标准者共 113 例。其中, 病历资料不完整者 37 例, 胸部 CT 提示 PN 最大直径 > 30mm 者 25 例, 伴胸腔积液 17 例, 伴肺不张 11 例, 五年内患有恶性肿瘤并接受治疗或既往接受肺部手术 15 例, 伴肺结核活动期 8 例。余下 703 例随机分成建模组 ($n=552$, 良性 237 例, 恶性 315 例) 和验证组 ($n=151$, 良性 85 例, 恶性 66 例)。本研究经深圳市中医院伦理学委员会批准, 签署知情同意书。

1.2 仪器与试剂 西门子 X 线电子计算机断层扫描仪 SOMATOM Force 开源 CT (德国西门子公司); 雅培 i 4000 SR 化学发光分析仪 (美国雅培公司); 试剂为美国雅培公司配套试剂包括癌胚抗

原 (CEA), 鳞状细胞癌抗原 (SCCA), 细胞角蛋白 19 片段 (CYFRA21-1) 试剂盒、激发液、预激发液、冲洗液等。

1.3 方法

1.3.1 资料采集: 临床资料采集包括年龄、性别、吸烟史、戒烟史和癌症家族史。影像学资料包括 CT 肺窗上结节最大直径、病变位置 (肺上叶/中叶/下叶, 左肺/右肺)、边界清晰、分叶、毛刺、血管集束征、空泡、钙化、支气管充气征、肺气肿、结节性质 (密度) 及胸膜凹陷、结节数目。血清生物学指标包括 CEA, SCCA, CYFRA21-1 等。

1.3.2 单变量分析: 将建模组所有变量包括临床病理特征、CT 信息和生物指标等进行单变量分析, 筛选有统计学意义的变量。

1.3.3 多变量回归分析: 将上述筛选的变量进行多元回归分析, 筛选独立预测因子。

1.3.4 建模及验证: 将上述独立预测因子纳入 logistic 回归分析, 得出各变量相关系数和常量, 建立回归方程, 构建 PN 性质预测模型。根据函数 $f(x) = e^x / (1 + e^x)$, 计算肺结节恶性概率。再将验证组数据分别代入本模型 (proposed model) 与 Mayo, VA, Brock, PKU, GZMU 模型计算 PN 恶性概率, 绘制受试者工作特征 (receiver operating characteristic, ROC) 曲线。根据曲线下面积 (area under the curve, AUC) 比较各模型的诊断效能。选取最佳临界值确定灵敏度和特异度。

1.4 统计学分析 采用 SPSS 22.0 和 MedCalc17.6 进行统计分析。计数资料表示为 $n(\%)$, 组间比较采用 Pearson 卡方检验。计量资料表示为均数 \pm 标准差 ($\bar{x} \pm s$) 或中位数 (四分位数间距, IQR) [$M(P_{25}, P_{75})$]。首先对计量资料进行正态检验, 再根据检验结果, 组间比较正态分布采用独立样本 t 检验, 非正态分布采用 Mann-Whitney U 秩和检验。利用 AUC 评价预测模型诊断效能。以 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 临床特征 建模组与验证组的临床特征包括年龄 ($t=1.805$, $P=0.581$)、性别 ($\chi^2=1.462$, $P=0.144$)、吸烟史 ($\chi^2=0.632$, $P=0.428$)、戒烟史 ($Z=1.465$, $P=0.143$) 和癌症家族史 ($\chi^2=1.343$, $P=0.180$) 进行比较, 差异均无统计学意义。

2.2 建模组和验证组的单变量分析 见表 1, 2。将各变量进行赋值: 性别 (女性 = 1, 男性 = 0), 过去或现在吸烟 (是 = 1, 否 = 0), 癌症家族史 (是 = 1, 否 = 0), 位置 (左肺 = 1, 右肺 = 0; 上叶 = 1, 中叶或下叶 = 0), 边界清晰 (是 = 1, 否 = 0), 分叶 (是 = 1, 否 = 0), 毛刺 (是 = 1, 否 = 0),

血管集束征 (是 = 1, 否 = 0), 空泡 (是 = 1, 否 = 0), 钙化 (是 = 1, 否 = 0), 支气管充气征 (是 = 1, 否 = 0), 肺气肿 (是 = 1, 否 = 0), 结节性质 (部分实性或磨玻璃样 = 1, 实性 = 0) 及胸膜凹陷 (是 = 1, 否 = 0)。单变量分析显示, PN 良性与恶性患者除性别、肺左叶、结节数目、胸膜凹陷、空泡、支气管充气征、肺气肿、吸烟史和戒烟年限外 ($P > 0.05$), 其他影响因素包括年龄、癌症家族史、结节最大直径、结节性质、肺上叶、钙化、血管集束征、分叶、边界清晰、毛刺以及血清 CEA, SCCA, CYFRA21-1 差异均具有统计学意义 ($P < 0.05$)。

表 1 建模组的单变量分析 [$\bar{x} \pm s$, $n(\%)$, $M(P_{25}, P_{75})$]

类别	良性 ($n=237$)	恶性 ($n=315$)	$\chi^2/t/z$	P
年龄 (岁)	61.5 \pm 13.4	68.9 \pm 12.3	7.10	<0.001
性别 (女)	124 (52.3)	178 (56.5)	0.88	0.379
吸烟史 >5 年	46 (19.4)	73 (23.2)	2.28	0.131
戒烟	12(7.21)	13(8.15)	0.58	0.562
癌症家族史	7 (3.0)	28 (8.9)	8.01	0.005
结节数目 (个)	3(1, 3)	3(1, 3)	1.69	0.090
结节最大直径 (mm)	6.1 \pm 2.3	12.3 \pm 6.7	13.21	<0.001
结节性质 (部分实性或磨玻璃样)	42(17.7)	180(57.1)	87.41	<0.001
钙化	47 (19.8)	52 (16.5)	1.01	0.043
血管集束征	53 (22.4)	135 (42.9)	25.3	<0.001
肺上叶	103 (43.5)	174 (55.2)	7.51	0.006
肺左叶	109 (46.0)	165 (52.4)	2.21	0.138
胸膜凹陷	47 (19.8)	73 (23.2)	0.89	0.346
空泡	48(20.3)	82 (26.0)	2.51	0.113
分叶	51(21.5)	109 (34.6)	11.23	0.004
边界清晰	136 (57.4)	143(45.4)	7.77	0.005
肺气肿	55 (23.2)	81 (25.7)	0.46	0.499
支气管充气征	43 (18.1)	70(22.2)	1.38	0.240
毛刺	47 (19.8)	184 (58.4)	82.72	<0.001
CEA (ng/ml)	1.8(1.2, 2.7)	2.9(1.7, 4.5)	8.25	<0.001
SCCA(ng/ml)	1.0(0.6, 1.7)	1.4(0.7, 2.6)	3.84	<0.001
CYFRA21-1(ng/ml)	2.7(2.0, 3.8)	3.6(2.5, 4.8)	6.30	<0.001

2.3 多变量回归分析 将单变量分析筛选具有统计学意义的变量包括年龄、癌症家族史、结节最大直径、结节性质、肺上叶、钙化、血管集束征、分叶、边界清晰、毛刺以及血清 CEA, SCCA, CYFRA21-1 进行多变量回归分析, 筛选出 10 种独立预测因子包括年龄、癌症家族史、边界、结节最大直径、结节性质、肺上叶, 毛刺以及血清 CEA, SCCA, CYFRA21-1 等。

2.4 Logistic 回归分析及预测模型构建 将上述筛选的 10 种独立预测因子纳入 Logistic 回归分析,

得出各变量相关系数和常量,见表3。建立回归方程, $X = (-6.318\ 8 + 0.020\ 8 \times \text{年龄} + 0.573\ 2 \times \text{肺上叶} + 0.527\ 4 \times \text{CEA} - 0.928\ 4 \times \text{边界清晰} + 0.294\ 6 \times \text{CYFRA21-1} + 0.294\ 0 \times \text{结节最大直径} + 1.220\ 1 \times \text{癌症家族史} + 0.064\ 8 \times \text{SCCA} + 1.461\ 5 \times \text{毛刺} + 1.497\ 6 \times \text{结节性质})$ 。再根据函数 $f(x) = e^x / (1 + e^x)$, 计算肺结节恶性概率。

表2 验证组的单变量分析 [$\bar{x} \pm s$, $n(\%)$, $M(P_{25}, P_{75})$]

类别	良性 ($n=85$)	恶性 ($n=66$)	$\chi^2/t/z$	P
年龄(岁)	61.0 ± 12.6	67.9 ± 10.5	3.56	<0.001
性别(女)	42 (47.7)	38 (57.5)	0.98	0.320
吸烟史>5年	13 (15.3)	17 (25.7)	2.54	0.111
戒烟	22(14, 25)	12(12, 22)	1.53	0.127
癌症家族史	4 (4.7)	12 (18.2)	7.08	0.008
结节数目(个)	3(2, 3)	3(2, 3)	0.64	0.520
结节最大直径(mm)	8.0 ± 3.2	11.1 ± 3.4	5.54	<0.001
结节性质 (部分实性或磨玻璃样)	23(27.0)	43(65.2)	7.69	0.005
钙化	13 (15.3)	4 (6.1)	3.15	0.076
血管集束征	18 (21.2)	35 (53.0)	16.44	<0.001
肺上叶	30 (35.3)	38 (57.5)	7.40	0.007
肺左叶	37 (43.5)	23 (34.8)	1.16	0.281
胸膜凹陷	20(23.5)	24 (36.4)	2.94	0.086
空泡	13(15.3)	17 (25.7)	2.54	0.111
分叶	19(22.4)	27 (40.9)	5.99	0.014
边界清晰	41 (48.2)	15(22.7)	10.29	0.001
肺气肿	22 (25.8)	21 (31.8)	0.64	0.424
支气管充气征	9 (10.5)	11(16.6)	1.18	0.276
毛刺	9(10.5)	25 (37.8)	15.76	<0.001
CEA (ng/ml)	2.1(1.3, 3.0)	2.7(1.7, 3.6)	2.56	0.011
SCCA(ng/ml)	1.3(0.8, 1.8)	2.5(1.1, 3.2)	4.92	<0.001
CYFRA21-1(ng/ml)	2.6(1.9, 3.7)	3.7(3.1, 4.7)	4.79	<0.001

表3 肺结节性质独立预测因子的 Logistic 回归分析

类别	β	OR	95%CI	t	P
年龄	0.020 8	1.021 0	1.001 5 ~ 1.040 9	2.614	0.009
CEA	0.527 4	1.694 5	1.410 2 ~ 2.036 0	3.956	0.000
边界清晰	-0.928 4	0.395 2	0.234 8 ~ 0.665 3	2.811	0.005
CYFRA21-1	0.294 6	1.342 6	1.131 5 ~ 1.593 0	3.379	0.001
结节最大直径	0.294 0	1.341 8	1.226 0 ~ 1.468 4	4.909	<0.001
癌症家族史	1.220 1	3.387 4	1.050 0 ~ 10.927 7	2.533	0.012
肺上叶	0.573 2	0.563 7	0.293 3 ~ 1.083 5	2.092	0.037
SCCA	0.064 8	1.066 9	0.961 9 ~ 1.183 4	1.999	0.046
毛刺	1.461 5	4.312 5	2.342 6 ~ 7.938 8	6.668	<0.001
结节性质	1.497 6	4.471 1	2.590 0 ~ 7.718 5	4.807	<0.001

注: β (Coefficient): 相关系数; OR (odds ratio): 优势比; CI (confidential intervals): 可信区间。

2.5 本模型与其他模型的诊断效能比较 见图1和表4。本模型AUC值高于其他5种模型,本模型与Mayo模型和VA模型比较, $AUC(0.799 \text{ vs } 0.659, 0.650)$ 差异有统计学意义 ($Z=3.029, P=0.003; Z=2.638, P=0.008$)。然而,本模型与Brock模型、PKU模型、GZMU模型比较, $AUC(0.799 \text{ vs } 0.762, 0.773, 0.799 \text{ vs } 0.769)$ 差异均无统计学意义 ($Z=1.063, P=0.288; Z=0.686, P=0.493; Z=0.757, P=0.449$)。诊断灵敏度最高为Brock模型(90.9%),最低为VA模型(50.0%); 特异度最高为VA模型(76.5%),最低为Brock模型(55.3%)。阳性似然比(positive likelihood ratio, +LR)最高为本模型(3.25); 最低为Mayo模型(1.86)。阴性似然比(negative likelihood ratio, -LR)最高为VA模型(0.66),最低为PKU模型(0.23)。诊断效能从大至小依次为: 本模型、PKU, GZMU, Brock, Mayo, VA模型。

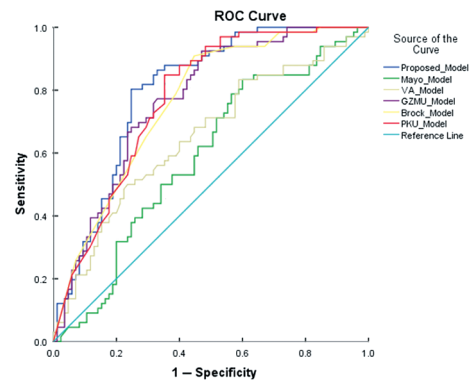


图1 本模型与其他模型的ROC曲线比较

3 讨论

预测模型是一种具有客观、精确、稳定及可重复等优点的评估工具,能够协助首诊医生给予肺结节患者的肺癌风险评估,根据恶性概率将患者进行风险分层,制定个体化诊疗方案,实现合理配置医疗资源的同时,达到患者利益最大化^[10]。21世纪全球开展两项大型随机对照试验,包括美国国立肺癌筛查试验(NLST)^[11]与荷兰-比利时肺癌筛查试验(NELSON)^[12],研究表明,肺癌早筛查与早诊治是降低人群肺癌死亡率的有效措施。最近,《柳叶刀》(THE LANCET)发表年度重磅肺癌筛查综述^[13],汇集当前全球最新医学证据,针对全球当前肺癌筛查现状与未来方向,展开荟萃分析。随机对照试验表明,用于肺癌筛查的LDCT有助于降低肺癌死亡率。指出肺癌发病率的地域差异,例如亚洲国家从未吸烟的人群中肺癌患者比例较高,需要特别注意改进选择标准和风险模型。本研究提示,是否吸烟对PN恶性风险无明显影响,差异无统计学意义。这有别于欧美国家人群,可能因种族和地区差异所致。

表4 本模型与其他模型的诊断效能比较

类别	最佳临界点	AUC	灵敏度 (%)	特异度 (%)	阳性似然比	阴性似然比	阳性预测值 (%)	阴性预测值 (%)	Z	P
Proposed	0.623 6	0.799	80.3	75.3	3.25	0.26	71.6	83.1	8.32	<0.001
Brock	0.659 6	0.762	90.9	55.3	2.03	0.35	61.2	78.8	7.00	<0.001
GZMU	0.622 5	0.769	68.2	75.3	2.76	0.42	68.2	75.3	7.05	<0.001
Mayo	0.282 1	0.659	74.2	60.0	1.86	0.43	59.0	75.0	3.57	<0.001
PKU	0.597 5	0.773	84.9	64.7	2.40	0.23	65.1	84.6	7.27	<0.001
VA	0.336 7	0.650	50.0	76.5	2.17	0.66	62.7	66.0	3.30	0.001

本研究发现年龄、癌症家族史、边界清晰、肺上叶、结节直径、结节性质、毛刺及血清CEA, CYFRA21-1, SCCA, 分叶、血管集束征、钙化等为肺结节的影响因子。其中, 前面10个变量为独立预测因子。除边界清晰为保护性因素外, 其余变量为肺结节恶性危险因素。独立预测因子的优势比(odds ratio, OR)依大小排列为结节性质、毛刺、癌症家族史、CEA, CYFRA21-1, SCCA, 直径、年龄、肺上叶、边界清晰。OR值大于3的分别是结节性质、毛刺和癌症家族史。磨玻璃结节(ground glass nodules, GGN), 也被称为毛玻璃结节, 指的是CT影像表现为局限性的模糊密度增高影。如果肺泡腔内存在少许黏液和包含的细胞量少时, CT影像表现为GGN; 如果结节内细胞或组织碎片大量堆积聚集、刺激胶原纤维成分大量增生时, CT影像则表现为混合磨玻璃结节(mixed ground glass nodules, mGGN)或部分实性结节(part-solid nodules, PSN)。研究发现PSN或GGN的恶性概率可能升高3.5倍以上(OR=4.5, 95%CI: 2.6 ~ 7.7), 在建模组中, 性质为PSN或GGN的良性与恶性结节分别占17.7%与57.1%, 差异有统计学意义($P<0.001$)。PSN与肺腺癌相关, 其恶性率高于纯GGN和实性结节^[14]。病理上, PSN主要被证实为微浸润和浸润性腺癌, 而纯GGN多为原位腺癌, 有时为微浸润腺癌^[15]。

边缘出现毛刺的结节恶性概率可能升高3.3倍以上(OR=4.3, 95%CI: 2.3 ~ 7.9), 在建模组中, 良性与恶性结节具有毛刺患者分别占19.8%和58.4%, 差异有统计学意义($P<0.001$)。毛刺多由两种因素形成, 一是肿瘤周围血管增生、增粗、牵拉引起, 多表现为细短毛刺, 二是病灶周围肺纤维化。年龄较大, 肺气肿和毛刺的存在与恶性肿瘤有关^[16]。具有癌症家族史患者恶性概率可能升高2.4倍(OR=3.4, 95%CI: 1.1 ~ 10.9), 在建模组中, 良性与恶性结节具有癌症家族史患者分别占3.0%和8.9%, 差异有统计学意义($P=0.005$)。肺癌家族史的个体比没有肺癌家族史的患肺癌的风险高2 ~ 3倍。一级亲属中有癌症家族史的个体与无癌

症家族史的个体相比, 患肺癌的风险增加约50%, 且这种关联不受性别、种族、组织学类型和其他已知肺癌危险因素的影响^[17]。ANG等^[18]研究提示, 有肺癌家族史的人群患肺癌的风险明显升高, 可能因存在肺癌遗传基因易感位点。研究显示, 在建模组中, 良性与恶性肺结节平均直径(6.1 mm vs 12.3 mm)差异有统计学意义($P<0.001$)。其中, 结节直径大于10.0 mm在良性与恶性组分别占12.0%和88.0%。LARICI等^[19]研究指出, 结节大小和生长速度仍然是最广泛使用的预测因子, 用于评估结节恶性概率。结节的初始大小、生长速度与恶性风险之间存在直接正比关系。这与本研究结果一致。尽管最近的指南引入了结节体积作为指标, 到目前为止, 结节管理一直基于结节直径大小。

研究最多的肺癌生物标志物包括CYFRA 21-1, CEA, SCCA和NSE等^[20]。本研究显示, 生物标志物OR值依大小排列为CEA, CYFRA21-1, SCCA。生物标志物在不同病理分型中的表现不同, 腺癌的CEA水平明显高于其他类型。而非小细胞癌CYFRA21-1具有较高表达。SCCA异常高表达常见于鳞癌, 包括肺癌、宫颈癌、卵巢癌等^[21]。李辉等^[22]研究发现, 与肺良性疾病组比较, 肺癌组患者CEA, CYFRA21-1, NSE, ProGRP, SCCA, CA125和HE4的水平均明显升高。诊断性能评价中, 单项检测时, 诊断肺癌AUC最大的为CYFRA21-1, 其次为CEA。而CEA, SCCA, NSE分别对腺癌、鳞癌和小细胞肺癌具有最佳诊断效能, 这与本研究结果相近。彭瑛等^[23]研究提示, NSE, CYFRA21-1, CEA, CA125, CA153, CA199和SCCA等7种标志物中, SCCA对肺癌诊断效能最高(AUC=0.817)。联合标志物的诊断效能优于单一指标, CEA+CA125+CYFRA21-1联合应用诊断肺癌效能更高。中华医学会肺癌临床诊疗指南^[24]指出, 胸部CT可有效检出早期周围型肺癌、明确病变所在的部位和累及范围。检测肺癌相关的生物标志物, 有助于辅助诊断和早期鉴别诊断并预测肺癌病理类型。

研究显示, PN恶性风险与年龄呈正比, PN恶性风险率从小至大依次为≤50岁[恶性33例/(良

性64例+恶性33例),恶性风险率34.0%,下同],>50岁[348/(257+348),56.6%],>60岁[283/(182+283),60.7%],>70岁[170/(85+170),66.7%]。肺上叶结节发病率较高,原因可能为吸入的有毒或有害物质首先进入并聚集于肺上叶,引起局部组织慢性炎症、增生、恶变。

似然比(LR)是反映诊断真实性的一种指标,属于同时反映灵敏度和特异度的复合指标。本研究构建模型+LR最高(3.25)。与Mayo,VA等模型比较,本模型具有较高PPV和NPV,且本模型AUC值高于其他5种模型。这可能归因于本研究参数更为全面,构建的模型更科学合理。研究提示,根据回归分析得出的OR值不同,经计分确认OR的权重如下:临床特征24.4%(4.76/19.55),CT信息54.6%(10.68/19.55),生物指标21.0%(4.11/19.55)。可见,CT信息权重最大,尤其结节性质,如果体检发现部分实性结节(PSN)或肺磨玻璃结节(GGN),应该根据恶性概率高低确定其危险度定期复查。其次,存在癌症家族史和(/或)生物指标异常,应根据预测模型计算结果分析判断,确定是否干预。

诚然,本模型存在一定局限,比如样本量偏小,缺乏大量外部数据验证。由于神经元烯醇化酶(NSE)检测结果受样本溶血及样本周转时间等因素干扰较大,胃泌素释放肽前体目前尚未开展,故这两个项目未纳入研究。再者,影像信息尚未将空洞、卫星灶、支气管截断征、晕征、结节周围支扩征、结节周围炎症等纳入研究。

综上所述,本预测模型具有较高的PN性质预测准确度,可帮助临床提高诊断准确度,使患者得到早诊断和早干预,避免过度医疗,值得推广应用。

参考文献:

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA-A Cancer Journal for Clinicians, 2021, 71(3): 209-249.
- [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]. New England Journal of Medicine, 2020, 382(6): 503-513.
- [3] POTTER A L, ROSENSTEIN A L, KIANG M V, et al. Association of computed tomography screening with lung cancer stage shift and survival in the United States: quasi-experimental study[J]. BMJ, 2022, 376: e069008.
- [4] 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]. Archives of Internal Medicine, 1997, 157(8): 849-855.
- [5] GOULD M K, ANANTH L, BARNETT P G, et al. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules[J]. Chest, 2007, 131(2): 383-388.
- [6] MCWILLIAMS A, TAMMEMAGI M C, MAYO J R, et al. Probability of cancer in pulmonary nodules detected on first screening CT[J]. New England Journal of Medicine, 2013, 369(10): 910-919.
- [7] 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.
LI Yun, CHEN Kezhong, SUI Xichao, et al. Establishment of a mathematical prediction model to evaluate the probability of malignancy or benign in patients with solitary pulmonary nodules[J]. Journal of Peking University(Health Sciences), 2011, 43(3): 450-454.
- [8] ZHANG Man, ZHUO Na, GUO Zhanlin, et al. Establishment of a mathematic model for predicting malignancy in solitary pulmonary nodules[J]. Journal of Thoracic Disease, 2015, 7(10): 1833-1841.
- [9] 周清华,范亚光,王颖,等.中国肺癌低剂量螺旋CT筛查指南(2018年版)[J].中国肺癌杂志, 2018, 21(2): 67-75.
ZHOU Qinghua, FAN Yaguang, WANG Ying, et al. China national lung cancer screening guideline with low-dose computed tomography(2018 version)[J]. Chinese Journal of Lung Cancer, 2018, 21(2): 67-75.
- [10] 中华医学会呼吸病学分会肺癌学组,中国肺癌防治联盟专家组.肺结节诊治中国专家共识(2018年版)[J].中华结核和呼吸杂志, 2018, 41(10): 763-771.
Lung Cancer Group of Chinese Thoracic Society, Expert Team, Union of Chinese Lung Cancer Management. Chinese expert consensus on diagnosis and treatment in pulmonary nodules(2018 version)[J]. Chinese Journal of Tuberculosis and Respiratory Diseases, 2018, 41(10): 763-771.
- [11] National Lung Screening Trial Research Team, ABERLE D R, ADAMS A M, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening[J]. New England Journal of Medicine, 2011, 365(5): 395-409.
- [12] VAN IERSEL C A, DE KONING H J, DRAISMA G, et al. Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON)[J]. International Journal of Cancer, 2007, 120(4): 868-874.
- [13] ADAMS S J, STONE E, BALDWIN D R, et al. Lung cancer screening[J]. Lancet, 2023, 401(10374): 390-408.
- [14] LEE H J, GOO J M, LEE C H, et al. Nodular ground-glass opacities on thin-section CT: size change during follow-up and pathological results[J]. Korean Journal of Radiology, 2007, 8(1): 22-31.

(下转第157页)

- [13] LIN Mingchun, LIU Meixiu, ZHANG Shulian, et al. Different types of minor blood group incompatibility causing haemolytic disease of neonates in one of the national children's medical centre in China[J]. *Journal of Blood Medicine*, 2021, 12: 497-504.
- [14] EL ALAOUI K, BENGHIAT F S, COLARD M. An unusual case of delayed hemolytic transfusion reaction with hyperhemolysis syndrome due to anti-Jk^b and anti-Fy^a alloantibodies[J]. *Journal of Hematology*, 2022, 11(2): 66-70.
- [15] 张澜, 刘灿华. 酶-IAT技术在红细胞抗体鉴定中的应用研究[J]. *中国输血杂志*, 2019, 32(1): 34-37.
ZHANG Lan, LIU Canhua. Application of two-stage enzyme test in the identification of red cell irregular antibodies[J]. *Chinese Journal of Blood Transfusion*, 2019, 32(1): 34-37.
- [16] 江涛, 杜娟, 安宁, 等. 16例Duffy血型不规则抗体血清学检测结果分析[J]. *细胞与分子免疫学杂志*, 2018, 34(6): 552-554.
JIANG Tao, DU Juan, AN Ning, et al. Analysis of serological detection results of irregular antibodies of 16 Duffy blood group[J]. *Chinese Journal of Cellular and Molecular Immunology*, 2018, 34(6): 552-554.
- [17] 陈春霞, 魏曾珍, 谭金哲, 等. 国内抗体筛查细胞试剂比对比[J]. *中国输血杂志*, 2017, 30(11): 1243-1245.
CHEN Chunxia, WEI Zengzhen, TAN Jinzhe, et al. The comparison of red blood cell reagents to screen unexpected antibodies[J]. *Chinese Journal of Blood Transfusion*, 2017, 30(11): 1243-1245.
- [18] PETIT F, MINNAI F, CHIARONI J, et al. The radial expansion of the Diego blood group system polymorphisms in Asia: mark of comigration with the Mongol conquests[J]. *European Journal of Human Genetics*, 2019, 27(1): 125-132.
- [19] EI DEWI D M, METWALLY T. Adsorption technique in pre-transfusion testing for patients with warm type autoimmune hemolytic anemia[J]. *Egyptian Journal of Immunology*, 2017, 24(2): 47-51.
- [20] DARA R C, SHARMA R, BHARDWAJ H. Severe drug-induced immune hemolysis due to ceftriaxone[J]. *Asian Journal of Transfusion Science*, 2020, 14(2): 187-191.
- 收稿日期: 2023-07-22
修回日期: 2023-10-25

(上接第151页)

- [15] LI Wangjia, LÜ Fajin, TAN Yiwen, et al. Benign and malignant pulmonary part-solid nodules: differentiation via thin-section computed tomography[J]. *Quantitative Imaging in Medicine and Surgery*, 2022, 12(1): 699-710.
- [16] ZAHARUDIN N, JAILAINI M F M, ABEED N N, et al. Prevalence and clinical characteristics of malignant lung nodules in tuberculosis endemic area in a single tertiary centre[J]. *BMC Pulmonary Medicine*, 2022, 22(1): 328.
- [17] KANWAL M, DING Xiaoji, CAO Yi. Familial risk for lung cancer[J]. *Oncology Letters*, 2017, 13(2): 535-542.
- [18] ANG L, CHAN C P Y, YAU W P, et al. Association between family history of lung cancer and lung cancer risk: a systematic review and meta-analysis [J]. *Lung Cancer*, 2020, 148: 129-137.
- [19] LARICI A R, FARCHIONE A, FRANCHI P, et al. Lung nodules: size still matters[J]. *European Respiratory Review*, 2017, 26(146): 170025.
- [20] NOORELDEEN R, BACH H. Current and future development in lung cancer diagnosis[J]. *International Journal of Molecular Sciences*, 2021, 22(16): 8661.
- [21] 王旋, 崔立春, 党升强. 非小细胞肺癌放疗化疗联合靶向治疗对血清肿瘤标志物、免疫功能及Cyclin D3水平影响的相关研究[J]. *现代检验医学杂志*, 2021, 36(4): 25-30.
WANG Xuan, CUI Lichun, DANG Shengqiang. Effects of radiotherapy and chemotherapy combined with targeted therapy on serum tumor markers, immune function and CyclinD3 level in non-small cell lung cancer[J]. *Journal of Modern Laboratory Medicine*, 2021, 36(4): 25-30.
- [22] 李辉, 汪春新, 秦明明, 等. 肺癌患者血清7项肿瘤标志物联合检测在病理分型及临床分期中的应用价值研究[J]. *现代检验医学杂志*, 2021, 36(4): 5-9, 121.
LI Hui, WANG Chunxin, QIN Mingming, et al. Study on the application value of combined detection of 7 tumor markers in serum of patients with lung cancer in pathological classification and clinical staging[J]. *Journal of Modern Laboratory Medicine*, 2021, 36(4): 5-9, 121.
- [23] 彭瑛, 邓正华, 温先勇. 国内13种血清肿瘤标志物对肺癌诊断价值的Meta分析[J]. *现代检验医学杂志*, 2016, 31(1): 96-100.
PENG Ying, DENG Zhenghua, WEN Xianyong. Diagnostic value of thirteen types of serum tumor markers for lung cancer in China: a meta-analysis[J]. *Journal of Modern Laboratory Medicine*, 2016, 31(1): 96-100.
- [24] 中华医学会肿瘤学分会, 中华医学会杂志社. 中华医学会肺癌临床诊疗指南(2022版)[J]. *中华医学杂志*, 2022, 102(23): 1706-1740.
Oncology Society of Chinese Medical Association, Chinese medical association Publishing House. Chinese medical association guideline for clinical diagnosis and treatment of lung cancer(2022 edition)[J]. *National Medical Journal of China*, 2022, 102(23): 1706-1740.
- 收稿日期: 2023-06-14
修回日期: 2023-09-08