

临床 I 期非小细胞肺癌诊断模型构建：基于临床影像学特征联合叶酸受体阳性循环肿瘤细胞检测的研究



孔德志, 刘傲, 崔健, 冷晓亮, 沃杨, 董燕亭, 矫文捷

青岛大学附属医院 胸外科 (山东青岛 266071)

【摘要】 目的 分析 FR⁺CTC (folate receptor-positive circulating tumor cells, 叶酸受体阳性循环肿瘤细胞) 与患者肺部病灶良恶性的相关性, 建立基于临床资料、影像学 and FR⁺CTC 的肺部肿瘤恶性预测模型。方法 回顾性分析青岛大学附属医院 2018 年 9 月至 2019 年 12 月 1 277 例入院进行了外周血 CTC 检查且病理结果明确的肺结节及肺肿瘤患者的临床资料, 其中男 518 例、女 759 例, 中位年龄为 57 (29~85) 岁。随机将患者分配到试验组和验证组, 进行单因素及多因素分析, 建立 Nomogram 预测模型, 对其进行内部及外部验证, 受试者工作特征 (ROC) 曲线检验模型区分度, 校准曲线检验模型一致性。结果 试验组非小细胞肺癌患者 925 例、良性疾病患者 113 例; 验证组非小细胞肺癌患者 219 例、良性疾病患者 20 例, 非小细胞肺癌患者外周血中 FR⁺CTC 明显高于肺部良性疾病患者 ($P<0.001$)。多因素分析显示年龄 ≥ 60 岁、女性、FR⁺CTC 值 >8.7 、胸膜牵拉征阳性、结节直径、毛刺征阳性、结节中实性成分直径比值 <1 是病灶直径 ≤ 4 cm 肺肿瘤良恶性病变的独立危险因素。建立 Nomogram 预测模型, 区分度检验试验组 ROC 曲线下面积 (AUC) 值为 0.918, 灵敏度为 86.36%, 特异性为 83.19%; 验证组 AUC 值为 0.903, 模型灵敏度为 79.45%, 特异性为 90.00%, 表明 Nomogram 模型区分度良好。校准曲线表明 Nomogram 预测模型一致性良好。结论 非小细胞肺癌患者外周血中 FR⁺CTC 显著高于肺部良性疾病患者。本研究建立的临床 I 期非小细胞肺癌诊断模型精准度良好, 可以为临床诊断提供依据。

【关键词】 非小细胞肺癌; 早期诊断; 列线图; 叶酸受体; 循环肿瘤细胞

Establishment of a diagnostic model for clinical stage I non-small cell lung cancer: A study based on clinical imaging features combined with folate receptor-positive circulating tumor cells tests

KONG Dezhi, LIU Ao, CUI Jian, LENG Xiaoliang, WO Yang, DONG Yanting, JIAO Wenjie

Department of Thoracic Surgery, Affiliated Hospital of Qingdao University, Qingdao, 266071, Shandong, P.R.China

Corresponding author: JIAO Wenjie, Email: jiaowj@qduhospital.cn

【Abstract】 Objective To analyze the correlation between folate receptor-positive circulating tumor cells (FR⁺CTC) of the benign and malignant lesions of the lung, and to establish a malignant prediction model for pulmonary neoplasm based on clinical data, imaging and FR⁺CTC tests. **Methods** A retrospective analysis was done on 1 277 patients admitted by the Affiliated Hospital of Qingdao University from September 2018 to December 2019, including 518 males and 759 females, with a median age of 57 (29-85) years. They underwent CTC examination of peripheral blood and had pathological results of pulmonary nodules and lung tumors. The patients were randomly divided into a trial group and a validation group. Univariate and multivariate analyses were realized on the data of the two groups. Then the Nomogram prediction model was established and verified internally and externally. Receiver operating characteristic (ROC) curve was used to test the differentiation of the model and calibration curve was used to test the consistency of the model. **Results** Totally 925 patients suffered non-small cell lung cancer and 113 patients benign diseases in the trial group; 219 patients non-small cell lung cancer and 20 patients benign diseases in the verification group. The FR⁺CTC in

DOI: 10.7507/1007-4848.202101072

基金项目: 山东省自然科学基金面上项目 (ZR2020MH234)

通信作者: 矫文捷, Email: jiaowj@qduhospital.cn

peripheral blood of non-small cell lung cancer patients was higher than that found in the lungs of the patients who were in favorite conditions ($P < 0.001$). Multivariate analysis showed that age ≥ 60 years, female, FR^+CTC value > 8.7 , positive pleural traction sign, nodule diameter, positive burr sign, consolidation/tumor ratio < 1 were independent risk factor for benign and malignant lung tumors with a lesion diameter of ≤ 4 cm. Thereby, the Nomogram prediction model was established. The AUC value under the ROC curve of the trial group was 0.918, the sensitivity was 86.36%, and the specificity was 83.19%. The AUC value of the verification group was 0.903, the sensitivity of the model was 79.45%, and the specificity was 90.00%, indicating Nomogram model discrimination was efficient. The calibration curve also showed that the Nomogram model calibration worked well. **Conclusion** FR^+CTC in peripheral blood of non-small cell lung cancer patients is higher than that found in the lungs of the patients who carry benign pulmonary diseases. The diagnostic model of clinical stage I non-small cell lung cancer established in this study owns good accuracy and can provide a basis for clinical diagnosis.

【Key words】 Non-small cell lung cancer; early diagnosis; nomogram; folate receptor; circulating tumor cell

在全球范围内,肺癌是人类健康和生命威胁最大的恶性肿瘤之一,2018年估计有210万新发肺癌病例和180万人死亡,占与癌症相关的死亡人数的近1/5(18.4%)^[1]。我国约75%的肺癌患者在诊断时已属晚期,5年生存率为15.6%,而早期肺癌患者的5年生存率为80%^[2]。这一现状不但与缺乏筛查有关,更与缺乏科学鉴别肺结节的方法有关。

目前针对肺癌高发人群筛查常通过低剂量螺旋计算机断层扫描(low-dose computed tomography, LDCT)。然而,有研究^[3]表明,LDCT假阳性率较高(18.5%),且存在辐射暴露、随访成本高以及发现中央型病灶的肺癌患者效果较差等问题,并不能完全达到有效的早期诊断和预防的目的^[4]。传统的肿瘤相关抗原在肺癌早期诊断中的敏感性和特异性均不理想。其它侵入性诊断方法,如针吸活检或支气管镜刷检,并发症发生率高,给患者带来极大的不适^[5]。

随着精准医学的发展和技术的进步,液体活检凭借其无创性、敏感性、动态性等特质,被寄予厚望^[6-7]。外周血标本易获得、侵入性小、可反复采集,是临床上常规检测较为理想的标本来源。循环肿瘤细胞(circulating tumor cell, CTC)是指从原发或转移病灶脱落到血液循环中的肿瘤细胞。CTC作为新一代的肿瘤标志物,大量研究^[8-10]已经证实CTC检测将有助于肿瘤早期诊断、复发转移监测、判断患者预后及指导术后辅助治疗。叶酸受体(folate receptor, FR)是一种细胞表面受体糖蛋白,在许多上皮来源的恶性肿瘤细胞中高度表达,在75.7%的非小细胞肺癌(NSCLC)患者中显著上调^[11]。

因此,我们致力于研究 FR^+CTC 在临床I期NSCLC患者的表达,并建立一个基于 FR^+CTC 的临床I期NSCLC诊断模型,为可疑肺结节的早期诊

断提供新的依据。

1 资料与方法

1.1 临床资料和分组

回顾性分析青岛大学附属医院2018年9月至2019年12月入院的1520例进行了CTC检测的患者的临床资料,其中男518例、女759例,中位年龄为57(29~85)岁。这些患者均接受了胸部CT扫描和手术治疗,并进行了病理检查。筛选出临床I期非小细胞肺癌患者及肺部良性疾病患者。纳入标准如下:(1)NSCLC或良性疾病患者术前胸部CT检查有肺结节。检查无淋巴结、区域或远处转移;(2)术前采集血样进行 FR^+CTC 检测;(3)CT上病灶直径 ≤ 4 cm;(4)病理组织学检查有明确的病理诊断结果;(5)患者有完整的临床病例资料。随后,为保证有足够数量的患者建立准确的模型,将符合标准的1277例患者采用随机数字表法按4:1的比例随机分为两组,其中试验组1038例、验证组239例。

1.2 血液样本中 FR^+CTC 的定量检测

临床检测样本为全血样本,每次采血时用加有EDTA抗凝剂的真空采血管抽取患者静脉血3mL,上下颠倒7~8次。样本采用“叶酸受体细胞检测试剂盒”检测样本中 FR^+CTC 的数量。根据试剂盒说明书,CTCs是通过红细胞的裂解和血液中白细胞的免疫磁性去除(通过抗CD45和抗CD14涂层的磁珠)来富集的。然后用肿瘤特异性配体叶酸和人工合成的寡核苷酸的结合物标记它,收集CTC进行定量PCR分析。扩增前,接产物首先在逆转录酶引物上退火延伸。免疫荧光染色后,CTCs被定义为表达叶酸配体和细胞角蛋白(8, 18和19)的细胞,以及4',6-二氨基-2-苯基吲哚染色的细胞

核。在本研究中,我们定义 1 个 CTC 单位为在 3 mL 血液中检测到的 CTC 数量^[12]。

样本一经采集,则尽快送至检测实验室。如若不能马上送检样本,则于 4~8℃ 暂存,并在 24 h 之内进行检测。如超出保存时间,则丢弃该血样,重新采集。样本如需运输,用 4~8℃ 冰袋保持低温运输。FR⁺CTC 值>8.7 定义为阳性。

1.3 数据收集

收集患者的临床资料、FR⁺CTC 值、影像学特征及肿瘤标志物。其中,临床资料包括年龄、性别、吸烟史、临床症状。影像学特征包括结节位置、结节直径、结节中实性成分直径比值 (consolidation/tumor ratio, CTR)、胸膜牵拉征、毛刺征、肺气肿。肿瘤标志物包括癌胚抗原 (carcinoembryonic antigen, CEA)、鳞状上皮细胞癌抗原 (squamous cell carcinoma antigen, SCC)、神经元特异性烯醇化酶 (neuron-specific enolase, NSE)、糖类抗原 199 (carbohydrate antigen 129, CA199) 和糖类抗原 125 (carbohydrate antigen 125, CA125)。

1.4 统计学分析

正态分布的连续变量采用均数±标准差 ($\bar{x}\pm s$) 表示,非正态分布的采用中位数 (上下四分位数) [(P_{25}, P_{75})] 表示,组间比较采用 *t* 检验、Mann-Whitney *U* 检验或 Kruskal-Wallis 检。分类变量分析采用 χ^2 检验和 Fisher 确切概率法。将数据进行单因素分析,将单因素分析结果中有统计学意义的因素进行 logistic 回归分析,找出独立危险因素,建立预测模型,绘制列线图 (Nomogram),对模型进行内部及外部验证,检测模型的预测性能。绘制受试者工作特征 (receiver operating characteristic, ROC) 曲线,计算曲线下面积 (area under the curve, AUC) 检验模型的区分度,校准曲线检验模型的一致性。在所有分析中,双侧 $P\leq 0.05$ 为差异有统计学意义。采用 SPSS 25.0 软件、Med-Calculator 19.1 软件、GraphPad Prism 8 软件和 R 3.6.1 软件及 RMS 包执行所有分析。

1.5 伦理审查

本研究已通过青岛大学附属医院伦理委员会批准,批准号为:QYFYWZLL 26169。

2 结果

2.1 研究人群的临床特征及 FR⁺CTC 分布

共 925 例经病理证实的 NSCLC 患者和 113 例良性结节患者纳入试验组。其中,NSCLC 患者组中 907 例病理证实为腺癌,良性结节中包括 50 例

为错构瘤、27 例为炎性假瘤、25 例为结核、11 例为硬化性肺泡细胞瘤。

在试验组中,患者的平均年龄为 59.15 岁,其中 NSCLC 患者的平均年龄为 60 岁,良性结节患者的平均年龄为 57 岁;验证组中,患者的平均年龄为 59.29 岁,其中 NSCLC 患者的平均年龄为 59.71 岁,良性结节患者的平均年龄为 54.35 岁。试验组中,患 NSCLC 的女性比例明显高于男性 (62.1% vs. 37.9%),没有吸烟史的患者 (77.2%) 占据大多数,有症状来就诊的患者多于没有症状来就诊 (多为查体发现) 的患者 (64.0% vs. 36.0%); 见表 1; 患 NSCLC 的外周血中 FR⁺CTC 值阳性 (>8.7) 的患者显著多于 FR⁺CTC 值阴性的患者 (78.8% vs. 23.9%, $P<0.001$); 同时,在肺良性结节的患者中,检测 FR⁺CTC 值阴性的患者比例则明显高于检测 FR⁺CTC 值阳性的患者 (76.1% vs. 23.9%)。影像学特征方面,肺恶性结节较良性结节更多的位于上叶 (57.9% vs. 52.2%),恶性结节位于右肺的比例也明显高于左肺 (60.5% vs. 39.5%)。肺恶性结节的直径的中位值为 17 mm,良性结节的直径的中位值为 13 mm,两组多数样本的结节直径大小都分布在 10~20 mm (45.2% vs. 54.9%),其次恶性结节更多分布在 20~30 mm (25.9%),而良性结节则分布在 <10 mm (27.4%)。恶性结节组胸部 CT 中结节胸膜牵拉征和毛刺征出现的频率显著高于良性结节组,分别为 40.2% vs. 19.5%, 62.9% vs. 23.9%。根据结节的 CTR 值,将结节分为实性结节、亚实性结节和纯磨玻璃样结节 (pure-ground glass opacity, pGGO),良性结节多数为实性结节 (78.8%),而超过一半的恶性结节含有磨玻璃样成份 (53.2%)。CEA 在肺恶性结节组的阳性率高于良性结节组 (23.1% vs. 14.2%),其余的肿瘤标志物如 SCC、CA 19-9、NSE 则没有太大差异。以上结果在验证组也得到了近似的分布趋势。

另外,我们也对全组患者的 FR⁺CTC 值进行了分析,如表 2 所示,NSCLC 患者组的 FR⁺CTC 值中位数为 11.5 个单位,而肺良性结节患者组的 FR⁺CTC 值中位数为 7.6 个单位, Mann-Whitney *U* 检验结果显示,差异有统计学意义 ($P<0.001$); 见图 1; 而 FR⁺CTC 值在年龄、性别、吸烟史的分布上,差异均无统计学意义。值得注意的是,虽然结节直径没有检验出差异,但是可以看出,随着结节直径逐渐增大,FR⁺CTC 值表现出来逐渐增多的趋势。

2.2 试验组的单因素和多因素分析

表 1 两组患者临床资料比较 [例 (%)]

临床资料	建模组 (n=1 038)			验证组 (n=239)		
	恶性 (n=925)	良性 (n=113)	P值	恶性 (n=219)	良性 (n=20)	P值
年龄			<0.001			0.052
≥60 岁	499 (53.9)	40 (35.4)		126 (57.5)	7 (35.0)	
<60 岁	426 (46.1)	73 (64.6)		93 (42.5)	13 (65.0)	
性别			<0.001			0.826
男	351 (37.9)	65 (57.5)		93 (42.5)	9 (45.0)	
女	574 (62.1)	48 (42.5)		126 (57.5)	11 (55.0)	
吸烟史			0.054			0.769
有	211 (22.8)	35 (31.0)		59 (26.9)	6 (30.0)	
无	714 (77.2)	78 (69.0)		160 (73.1)	14 (70.0)	
症状			0.029			0.484
有	591 (64.0)	84 (74.3)		136 (62.1)	14 (70.0)	
无	333 (36.0)	29 (25.7)		83 (37.9)	6 (30.0)	
FR ⁺ CTC 值			<0.001			<0.001
>8.7 FU/3 mL	729 (78.8)	27 (23.9)		175 (79.9)	4 (20.0)	
≤8.7 FU/3 mL	196 (21.2)	86 (76.1)		44 (20.1)	16 (80.0)	
肺气肿			0.167			0.836
有	199 (21.5)	18 (15.9)		43 (19.6)	3 (15.0)	
无	725 (78.5)	95 (84.1)		176 (80.4)	17 (85.0)	
结节位置			0.134			0.163
右上叶	326 (35.3)	27 (23.9)		65 (29.7)	3 (15.0)	
右中叶	54 (5.8)	5 (4.4)		9 (4.1)	3 (15.0)	
右下叶	179 (19.4)	27 (23.9)		39 (17.8)	5 (25.0)	
左上叶	109 (22.6)	32 (28.3)		57 (26)	3 (15.0)	
左下叶	156 (16.9)	22 (19.5)		49 (22.4)	6 (30.0)	
结节直径			<0.001			0.257
<10 mm	148 (16.0)	31 (27.4)		34 (15.5)	6 (30.0)	
10 ~ <20 mm	418 (45.2)	62 (54.9)		100 (45.7)	10 (50.0)	
20 ~ <30 mm	240 (25.9)	12 (10.6)		62 (28.3)	3 (15.0%)	
30 ~ ≤40 mm	119 (12.9)	8 (7.1)		23 (10.5)	1 (5.0)	
胸膜牵拉征			<0.001			0.144
有	372 (40.2)	22 (19.5)		86 (39.3)	4 (20.0)	
无	553 (59.8)	91 (80.5)		133 (60.7)	16 (70.0)	
毛刺征			<0.001			<0.001
有	581 (62.9)	27 (23.9)		140 (63.9)	4 (20.0)	
无	343 (37.1)	86 (76.1)		79 (36.1)	16 (80.0)	
CTR			<0.001			0.208
0	433 (46.8)	89 (78.8)		99 (45.2)	13 (65.0)	
>0 ~ <1	136 (14.7)	8 (7.1)		45 (20.5)	2 (10.0)	

续表 1

临床资料	建模组 (n=1 038)			验证组 (n=239)		
	恶性 (n=925)	良性 (n=113)	P值	恶性 (n=219)	良性 (n=20)	P值
1	356 (38.5)	16 (14.2)		75 (34.2)	5 (25.0)	
CEA (ng/mL)			0.030			0.476
>3.4	214 (23.1)	16 (14.2)		29 (13.2)	1 (5.0)	
≤3.4	711 (76.9)	97 (85.8)		190 (86.8)	19 (95.0)	
SCC			0.761			0.930
>2.5 ng/mL	73 (7.9)	8 (7.1)		10 (4.6)	1 (5.0)	
≤2.5 ng/mL	852 (92.1)	105 (92.9)		209 (95.4)	19 (95.0)	
CA19-9			0.335			0.132
>3.3 U/mL	200 (21.6)	20 (17.7)		48 (21.9)	1 (5.0)	
≤3.3 U/mL	725 (78.4)	93 (82.3)		171 (78.1)	19 (95.0)	
NSE			0.165			0.575
>17 ng/mL	85 (9.2)	15 (13.3)		26 (11.9)	1 (5.0)	
≤17 ng/mL	840 (90.8)	98 (86.7)		193 (88.1)	19 (95.0)	

FR⁺CTC: 叶酸受体阳性循环肿瘤细胞; CTR: 结节中实性成分直径比值; CEA: 癌胚抗原; SCC: 鳞状上皮细胞癌抗原; NSE: 神经元特异性烯醇化酶; CA199: 糖类抗原 199

表 2 良恶性肺结节患者外周血 FR⁺CTC 值 [M (P₂₅, P₇₅)]

因素	FR ⁺ CTC 值	P值
性质		<0.001
良性	7.60 (6.05, 8.60)	
恶性	11.50 (9.10, 16.10)	
年龄		0.443
≥60 岁	11.20 (8.23, 15.80)	
<60 岁	10.80 (8.40, 14.91)	
性别		0.108
男	10.65 (8.00, 15.30)	
女	11.40 (8.60, 15.60)	
吸烟史		0.435
有	10.70 (8.20, 15.40)	
无	11.20 (8.40, 15.49)	
结节直径		0.471
<10 mm	10.70 (8.70, 15.30)	
10 ~ <20 mm	10.80 (8.00, 15.10)	
20 ~ <30 mm	11.20 (8.50, 16.10)	
30 ~ 40 mm	11.90 (8.32, 15.75)	

将试验组内的数据对上述因素进行 logistic 单因素分析, 结果显示, 试验组患者中, NSCLC 患者和肺良性结节患者在年龄 (P<0.001)、性别 (P<0.001)、结节直径 (P<0.001)、临床症状 (P=0.030)、

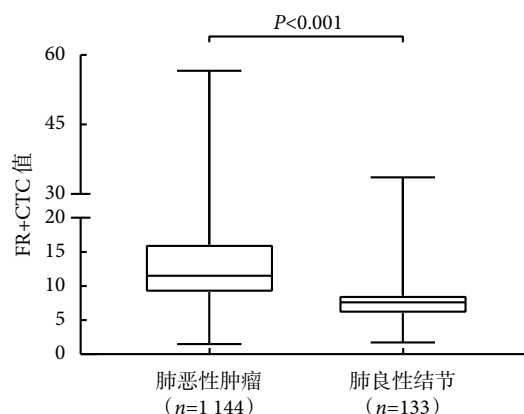


图 1 非小细胞肺癌和良性结节患者外周血中 FR⁺CTC 值差异

FR⁺CTC 值 (P<0.001)、胸膜牵拉征 (P<0.001)、毛刺征 (P<0.001)、CTR (P<0.001) 及 CEA (P=0.032) 方面差异均有统计学意义; 见表 3。

将年龄、性别、结节直径、临床症状、FR⁺CTC 值、胸膜牵拉征、毛刺征、CTR 及 CEA 共 9 项因素纳入 logistic 多因素分析, 结果显示, 年龄 ≥60 岁 (OR 1.700, 95%CI 1.002 ~ 2.884, P=0.049)、女性 (OR=2.617, 95%CI 1.541 ~ 4.442, P<0.001)、FR⁺CTC 值>8.7 (OR=16.992, 95%CI 9.728 ~ 29.682, P<0.001)、胸膜牵拉征阳性 (OR=2.199, 95%CI 1.195 ~ 4.405, P=0.011)、结节直径 (OR=1.728, 95%CI 1.244 ~ 2.402, P=0.001)、毛刺征阳性 (OR=6.322, 95%CI 3.561 ~ 11.225, P<0.001)、

表 3 试验组的单因素及多因素分析

因素		单因素分析			多因素分析		
		OR	95%CI	P值	OR	95%CI	P值
年龄(岁)	(≥60/<60)	2.138	1.423 ~ 3.210	<0.001	1.700	1.002 ~ 2.884	0.049
性别	(女/男)	2.215	1.490 ~ 3.290	<0.001	2.617	1.541 ~ 4.442	<0.001
吸烟史	(有/无)	0.659	0.430 ~ 1.010	0.055			
症状	(有/无)	0.613	0.393 ~ 0.954	0.030	0.641	0.361 ~ 1.137	0.129
FR ⁺ CTC值(FU/3 mL)	(>8.7/≤8.7)	11.847	7.477 ~ 18.771	<0.001	16.992	9.728 ~ 29.682	<0.001
肺气肿	(有/无)	1.449	0.855 ~ 2.455	0.169			
结节直径(mm)	(<10/10 ~ 20/20 ~ 30/30 ~ 40)	1.686	1.323 ~ 2.148	<0.001	1.728	1.244 ~ 2.402	0.001
胸膜牵拉征	(有/无)	2.783	1.715 ~ 4.514	<0.001	2.199	1.195 ~ 4.405	0.011
毛刺征	(有/无)	5.395	3.432 ~ 8.483	<0.001	6.322	3.561 ~ 11.225	<0.001
CTR	(1/0 ~ 1/0)	2.253	1.709 ~ 2.971	<0.001	3.172	2.264 ~ 4.446	<0.001
CEA (ng/mL)	(>3.4/≤3.4)	1.825	1.052 ~ 3.164	0.032	1.959	0.939 ~ 4.089	0.073
SCC (ng/mL)	(>2.5/≤2.5)	1.125	0.527 ~ 2.399	0.761			
CA19-9 (U/mL)	(>3.3/≤3.3)	1.283	0.772 ~ 2.132	0.337			
NSE (ng/mL)	(>17/≤17)	0.996	0.990 ~ 1.002	0.167			

FR⁺CTC: 叶酸受体阳性循环肿瘤细胞; CTR: 结节中实性成分直径比值; CEA: 癌胚抗原; SCC: 鳞状上皮细胞癌抗原; NSE: 神经元特异性烯醇化酶; CA199: 糖类抗原 199

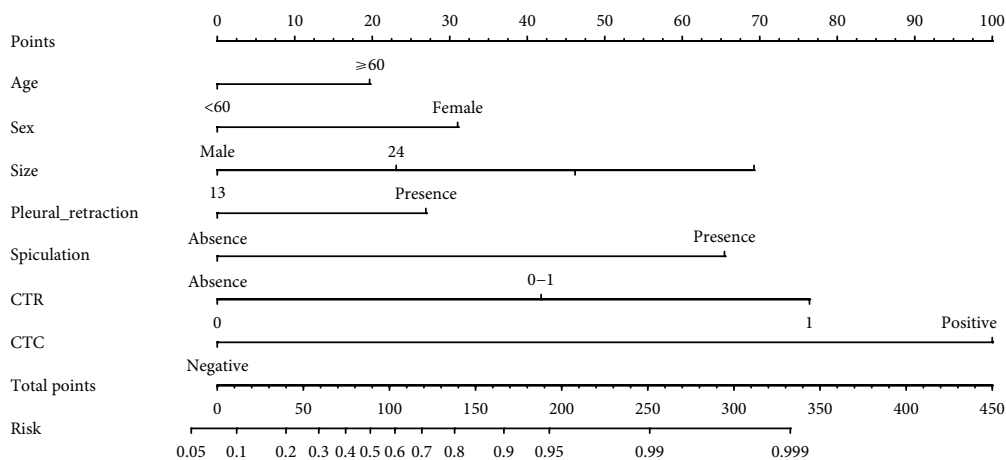


图 2 临床 I 期非小细胞肺癌诊断预测模型的列线图

CTR<1 (OR=3.172, 95%CI 2.264 ~ 4.446, P<0.001) 是病灶直径 ≤ 4 cm 肺部肿物良恶性病变的独立危险因素; 见表 3。

2.3 建立 Nomogram 预测模型

将 logistic 回归分析的结果纳入 Nomogram 预测模型, 即与病灶直径 ≤ 4 cm 肺部肿物良恶性病变的独立危险因素年龄、性别、结节直径、胸膜牵拉征、毛刺征、CTR、FR⁺CTC 值纳入预测模型, 绘出列线图; 见图 2。

2.4 Nomogram 预测模型的内部验证和外部验证

首先, 对建立的 Nomogram 预测模型进行内部

验证, ROC 曲线用于验证模型的区分度, 试验组 AUC 值为 0.918 (95%CI 0.900 ~ 0.934)。当截断点 T=0.695 5 时, 约登指数最大, 此时模型灵敏度为 86.36%, 特异性为 83.19%; 见图 3a。校准曲线用于验证 Nomogram 模型的预测一致性, 显示了肺结节良恶性的预测精度; 见图 3c。预测概率的校准曲线图表明, Nomogram 预测模型估计的预测值与实际观测结果有很大的一致性。

同时, 我们在验证队列评估了包括年龄、性别、结节直径、胸膜牵拉征、毛刺征、CTR、FR⁺CTC 值在内的 Nomogram 预测模型的预测价值, 即

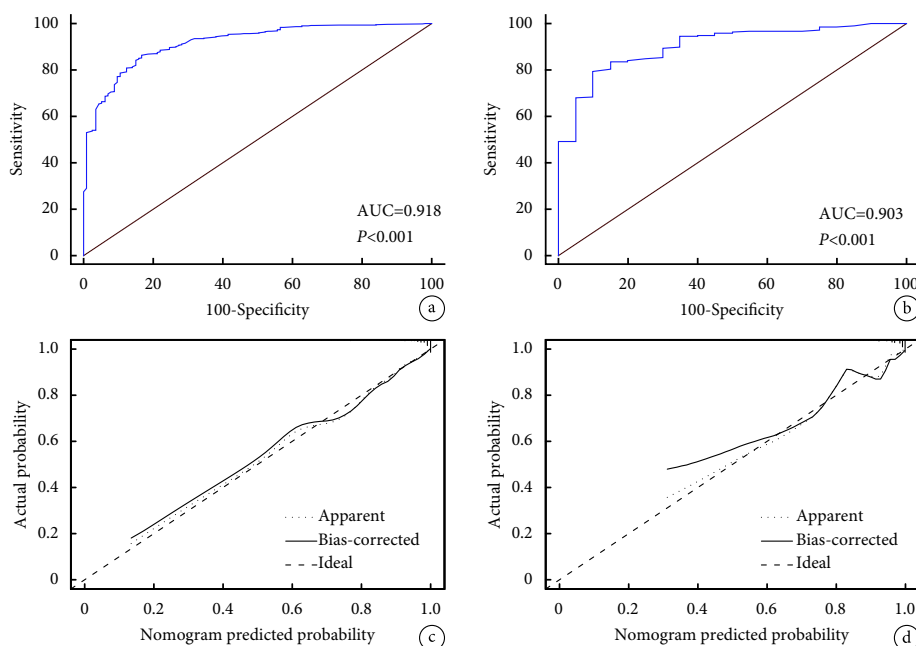


图3 临床 I 期非小细胞肺癌诊断预测模型的内部及外部验证

a: 试验组 ROC 曲线评价 Nomogram 模型的诊断价值; b: 验证组 ROC 曲线评价 Nomogram 模型的诊断价值; c: 内部验证 Nomogram 模型的校准曲线; d: 外部验证 Nomogram 模型的校准曲线

Nomogram 预测模型的外部验证。在验证队列中, 同样通过 ROC 曲线验证模型的区分度, 验证组 AUC 值为 0.903 (95%CI 0.859 ~ 0.938); 见图 3b。校准曲线验证 Nomogram 预测模型的预测一致性, 预测肺结节良恶性的校准曲线表明 Nomogram 预测模型一致性良好; 见图 3d。

3 讨论

我国肺恶性肿瘤发病率和死亡率高居全国首位, 其中 NSCLC 约占 85%。与之前相比, 肺腺癌发病率显著升高, 女性发病率和死亡率上升, 青年发病率显著上升^[13]。NSCLC 患者的预后主要取决于病理分期, I / II 期患者的 5 年生存率为 68% ~ 92%, 而晚期患者的 5 年生存率只有 1% ~ 13%。然而, 由于缺乏有效的方法, 只有 16% 的肺癌患者在早期被诊断出来^[8]。目前一般应用 LDCT 进行早期肺癌的筛查, 但是存在诊断小结节的假阳性率较高等缺陷, 对 LDCT 在肺癌筛查及指导治疗中的有效性存在一些争议^[14]。目前世界上常用对肺结节进行良恶性预测的模型, 包括 Mayo 模型^[15]、VA 模型^[16]和 Brock 模型^[17]等, 都包括一定时效性、地域性、人群分布等局限, 且大多只纳入了可疑肺结节的影像学特征, 并未将肿瘤标记物等资料纳入联合诊断模型。随着近年来高分辨率 CT (high resolution CT, HRCT) 的发展, 越来越多的肺结节, 特别是磨玻璃

结节被发现, 毛刺征、胸膜牵拉征等影像学特征也较 LDCT 更为清晰。因此, 迫切需要寻找新的联合诊断因素, 结合新的技术, 建立 NSCLC 的早期诊断预测模型, 以提高总体生存率。

1869 年, 澳大利亚医生 Thomas Ashworth 解除了 1 例晚期癌症患者, 意外地在他的外周血液发现了与原发肿瘤大小和形状相似的细胞, 随后提出了循环肿瘤细胞的概念^[18]。基于 Ashworth 的发现, 英国病理学家 Stephen Paget 于 1889 年提出了著名的种子和土壤假说, 成功地解释了肿瘤复发和转移的机制^[19]。2010 年, CTC 分析被写入美国癌症联合委员会 (AJCC) 癌症分期手册, 作为补充病理分期的新指标, 作为 cM0 (i⁺) 出现在 M0 和 M1 之间^[20], CTC 和循环肿瘤 DNA (CtDNA) 是主要的候选标志物^[9]。在先前的研究中, CTCs 和血清标志物被用作早期 NSCLC 的预后和预测生物标志物^[21]。CTC 在肺癌进展和预后中的作用, 以及评估疗效、开发靶向药物和个体化治疗也是一个普遍的研究方向^[22-24]。叶酸受体 (FR) 是一种跨膜单链糖蛋白, 含有 α 、 β 和 γ 三种亚型, 在人体大部分的肿瘤细胞中都有特异性的高表达, 如卵巢癌、肺癌和泌尿系统癌, 而在正常器官中很少表达。研究发现 72% ~ 83% 的肺肿瘤细胞表面高表达 FR, 因此 FR 被选为一种理想的肺肿瘤细胞筛选靶标^[25]。FR 还可用于识别活的 CTC, 其丰度不受上皮间质转化 (epithelial-

mesenchymal transition, EMT) 的影响^[25-26], 因为被选为 CTC 特异性标记物。一项“叶酸受体细胞检测试剂盒”多中心、随机、单盲上市前临床验证试验结果证明试剂盒检测的 FR⁺CTC 作为肺癌的辅助诊断手段具有高灵敏度 (79.6%) 和特异性 (88.2%), 对肺癌的诊断效能优于目前临床常用的其它血液标志物^[12, 27-28]。综上所述, 我们选择纳入 FR⁺CTC 值, 探讨结合临床资料和影像学特征, 构建临床 I 期 NSCLC 的诊断预测模型的可能性。

我们的研究显示, 年龄 (≥ 60 岁)、性别 (女)、FR⁺CTC 值 (>8.7)、胸膜牵拉征 (阳性)、病灶直径 (越大)、毛刺征 (阳性)、CTR (<1) 是判断 ≤ 4 cm 肺病灶良恶性的独立危险因素。与之前的研究结果一致^[12, 29], 我们发现早期 NSCLC 患者的外周血中 CTC 数量明显高于肺部良性疾病患者。其原因可能是尽管 HRCT 未见明显异常, 但早期肺癌可能存在肿瘤细胞从病灶中脱落并进入血液循环。在逃避免疫清除的情况下, 肿瘤细胞通常通过外周血聚集并定植于其它器官, 形成转移灶 (12 ~ 15 个)。研究^[30] 表明, 即使是从原发灶脱落的少量 CTC 在没有被免疫系统识别和清除的情况下存活下来, 也可能在某个时间或其它时候发生远处转移。此外, FR⁺CTC 在年龄较高的人群中也稍高, 这可能是因为老年人长期叶酸缺乏导致肿瘤组织中 FR 的过度表达, 从而促使 FR⁺CTC 水平升高^[31]。结节直径的大小与 CTC 值关系密切, 这与之前 Ding 等^[32] 的研究结果一致, 需要我们进一步挖掘 CTC 与肺癌诊断、预后的关系^[33]。考虑到亚洲地区空气污染状况、亚洲非吸烟女性腺癌的高患病率^[34], 与基于欧美人群为样本研究建立 Mayo 模型、VA 模型和 Brock 模型等相比, 我们的预测模型更适用于我国患者可疑肺部肿物良恶性的诊断。

至于此 Nomogram 诊断模型的应用, 假如胸部 CT 上有 1 个 15 mm 可疑肺结节的 55 岁男性患者, 胸部 CT 显示结节有胸膜牵拉, 无毛刺征, CTR 为 0.5, 外周血中 FR⁺CTC 值为阳性, 则利用 Nomogram 诊断模型计算该患者的肺结节恶性概率是 95%, 如图 2 所示, Nomogram 图中计算如下: 年龄 <60 岁、男性、无毛刺征, 对应于 0 分; 结节直径位于 10 ~ 20 区间, 对应于 22 分; 胸膜牵拉存在, 对应于 27 分; CTR=0.5, 相当于 41 分; 外周血中 FR⁺CTC 值为阳性, 相当于 100 分; 最后总分等于 190 分, 对应肺结节恶性概率为 95%。模型区分度由 ROC 曲线来检验。在试验队列中, Nomogram 诊断模型的 ROC 曲线的 AUC 值为 0.918 (95%CI

0.900 ~ 0.934), 而验证队列的 AUC 值为 0.903 (95%CI 0.859 ~ 0.938)。这表明, 此 Nomogram 诊断模型可以从没有事件发生的患者中分辨出有事件发生的患者, 其诊断效能大约是 91.9%。因此, 我们可以认为, 包含年龄、性别、结节直径、胸膜牵拉征、毛刺征、CTR、FR⁺CTC 值的诊断模型可以诊断临床 I 期 NSCLC。

我们的预测模型的灵敏度为 82.91% (86.36% vs. 79.45%), 特异性为 86.60% (83.19% vs. 90.00%), 均高于或者近似与此前应用 CTC 诊断肺癌的模型的结果^[12, 32, 35-36]。Moon 等^[35] 建立了一个基于 CTC 和 ct-DNA 的肺癌联合诊断模型, 灵敏度为 95%, 特异性为 16.7%, 与之相比我们纳入了影像学特征及其它临床因素, 尽可能地提高了模型的稳定性, 较其模型的特异性显著提高。Chen 等^[12] 建立了一个基于 FR⁺CTC 联合 CEA、NSE 和 Cyfra21-1 诊断肺癌的预测模型, AUC 值为 0.883 (0.841 ~ 0.925), 灵敏度为 84.21%, 特异性为 83.91%。与之相比, 我们的研究对象为早期肺癌患者, 且绝大多数为肺腺癌^[37], CEA 在单因素分析中被认为与 NSCLC 有关, 其它肿瘤标志物在分析中没有被验证为独立危险因素, 有必要进一步研究 FR⁺CTC 与肿瘤标志物表达在肺癌联合诊断中的作用。

我们的研究应用了 FR⁺CTC 联合影像学特征诊断临床 I 期 NSCLC, 并开发了 Nomogram 模型。目前这项研究的优势在于, 它专门调查了早期 NSCLC 患者, 并验证了结合 CTC 和影像特征的诊断准确性。结合液体活检的非侵入性等特点, 可以更好地优化个人的预测准确性, 帮助临床医生诊断肺结节, 有助于减少低危疾病的误诊和过度治疗。一些不能进行病理活检的患者, 将直接从本研究中获益。

本研究存在一些局限性。首先, 本研究为单中心、回顾性研究, 样本量较少, 模型需经过前瞻性、多中心、大样本研究进一步检验; 其次, 在本研究中, 研究对象人群为可疑恶性结节且有外科手术治疗倾向的患者, 存在一定的选择偏倚。同时, 经本模型诊断的恶性概率低的结节同样需要随访, 尤其是 CEA 高的患者, 要根据结节的动态变化由有经验的胸外科医师制定下一步诊疗计划^[38]。

综上所述, 我们研究了 FR⁺CTC 在临床 I 期 NSCLC 患者和肺良性结节患者血液中表达的差异, 发现 NSCLC 患者外周血中 FR⁺CTC 明显高于肺部良性疾病患者, 年龄、性别、结节直径、胸膜牵拉征、毛刺征、CTR、FR⁺CTC 值是临床 I 期 NSCLC

的独立危险因素,并建立了一个由此组成的新的可疑肺部肿物诊断预测模型,并且我们证实此预测模型是有意义的,可以为临床诊断提供依据。

利益冲突:无。

作者贡献:孔德志完善研究设计、查阅资料、数据整理与分析及论文初稿撰写与修改等;刘傲指导数据的统计学分析等;崔健、冷晓亮负责细化写作思路、对文章中图表进一步完善加工;沃杨、董燕亭负责收集数据,文章润色;矫文捷负责论文设计、指导、审阅与修改等。

参考文献

- Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2018, 68(6): 394-424.
- Gao S, Li N, Wang S, et al. Lung cancer in People's Republic of China. *J Thorac Oncol*, 2020, 15(10): 1567-1576.
- Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*, 2011, 365(5): 395-409.
- Boiselle PM. Computed tomography screening for lung cancer. *JAMA*, 2013, 309(11): 1163-1170.
- Heerink WJ, de Bock GH, de Jonge GJ, et al. Complication rates of CT-guided transthoracic lung biopsy: Meta-analysis. *Eur Radiol*, 2017, 27(1): 138-148.
- Hofman P. Liquid biopsy and therapeutic targets: Present and future issues in thoracic oncology. *Cancers (Basel)*, 2017, 9(11): 154.
- Rolfo C, Mack PC, Scagliotti GV, et al. Liquid biopsy for advanced non-small cell lung cancer (NSCLC): A statement paper from the IASLC. *J Thorac Oncol*, 2018, 13(9): 1248-1268.
- Mateo J, Gerlinger M, Rodrigues DN, et al. The promise of circulating tumor cell analysis in cancer management. *Genome Biol*, 2014, 15(8): 448.
- Calabuig-Fariñas S, Jantus-Lewintre E, Herreros-Pomares A, et al. Circulating tumor cells versus circulating tumor DNA in lung cancer-which one will win? *Transl Lung Cancer Res*, 2016, 5(5): 466-482.
- Mayo-de-Las-Casas C, Garzón Ibáñez M, Jordana-Ariza N, et al. An update on liquid biopsy analysis for diagnostic and monitoring applications in non-small cell lung cancer. *Expert Rev Mol Diagn*, 2018, 18(1): 35-45.
- Shivange G, Urbanek K, Przanowski P, et al. A single-agent dual-specificity targeting of FOLR1 and DR5 as an effective strategy for ovarian cancer. *Cancer Cell*, 2018, 34(2): 331-345.
- Chen X, Zhou F, Li X, et al. Folate receptor-positive circulating tumor cell detected by LT-PCR-based method as a diagnostic biomarker for non-small-cell lung cancer. *J Thorac Oncol*, 2015, 10(8): 1163-1171.
- Wei W, Zeng H, Zheng R, et al. Cancer registration in China and its role in cancer prevention and control. *Lancet Oncol*, 2020, 21(7): e342-e349.
- Horeweg N, Scholten ET, de Jong PA, et al. Detection of lung cancer through low-dose CT screening (NELSON): A prespecified analysis of screening test performance and interval cancers. *Lancet Oncol*, 2014, 15(12): 1342-1350.
- Swensen SJ, Silverstein MD, Ilstrup DM, et al. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Arch Intern Med*, 1997, 157(8): 849-855.
- Gould MK, Ananth L, Barnett PG, et al. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. *Chest*, 2007, 131(2): 383-388.
- McWilliams A, Tammemagi MC, Mayo JR, et al. Probability of cancer in pulmonary nodules detected on first screening CT. *N Engl J Med*, 2013, 369(10): 910-919.
- Hu B, Rochefort H, Goldkorn A. Circulating tumor cells in prostate cancer. *Cancers (Basel)*, 2013, 5(4): 1676-1690.
- Paget S. The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev*, 1989, 8(2): 98-101.
- Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*, 2010, 17(6): 1471-1474.
- Crosbie PA, Shah R, Summers Y, et al. Prognostic and predictive biomarkers in early stage NSCLC: CTCs and serum/plasma markers. *Transl Lung Cancer Res*, 2013, 2(5): 382-397.
- Wong SC, Chan CM, Ma BB, et al. Clinical significance of cytokeratin 20-positive circulating tumor cells detected by a refined immunomagnetic enrichment assay in colorectal cancer patients. *Clin Cancer Res*, 2009, 15(3): 1005-1012.
- Brown HK, Tellez-Gabriel M, Cartron PF, et al. Characterization of circulating tumor cells as a reflection of the tumor heterogeneity: Myth or reality? *Drug Discov Today*, 2019, 24(3): 763-772.
- Hofman V, Heeke S, Marquette CH, et al. Circulating tumor cell detection in lung cancer: But to what end? *Cancers (Basel)*, 2019, 11(2): 262.
- Nunez MI, Behrens C, Woods DM, et al. High expression of folate receptor alpha in lung cancer correlates with adenocarcinoma histology and EGFR [corrected] mutation. *J Thorac Oncol*, 2012, 7(5): 833-840.
- Parker N, Turk MJ, Westrick E, et al. Folate receptor expression in carcinomas and normal tissues determined by a quantitative radioligand binding assay. *Anal Biochem*, 2005, 338(2): 284-293.
- Lou J, Ben S, Yang G, et al. Quantification of rare circulating tumor cells in non-small cell lung cancer by ligand-targeted PCR. *PLoS One*, 2013, 8(12): e80458.
- Yu Y, Chen Z, Dong J, et al. Folate receptor-positive circulating tumor cells as a novel diagnostic biomarker in non-small cell lung cancer. *Transl Oncol*, 2013, 6(6): 697-702.
- Ning N, Zhan T, Zhang Y, et al. Improvement of specific detection of circulating tumor cells using combined CD45 staining and fluorescence in situ hybridization. *Clin Chim Acta*, 2014, 433: 69-75.
- Harper KL, Sosa MS, Entenberg D, et al. Mechanism of early dissemination and metastasis in Her2 + mammary cancer. *Nature*, 2016, 540(7634): 588-592.
- Zhu WY, Alliegro MA, Melera PW. The rate of folate receptor alpha (FR alpha) synthesis in folate depleted CHL cells is regulated by a translational mechanism sensitive to media folate levels, while stable overexpression of its mRNA is mediated by gene amplification and an increase in transcript half-life. *J Cell Biochem*,

- 2001, 81(2): 205-219.
- 32 Ding C, Zhou X, Xu C, *et al.* Circulating tumor cell levels and carcinoembryonic antigen: An improved diagnostic method for lung adenocarcinoma. *Thorac Cancer*, 2018, 9(11): 1413-1420.
- 33 Chang S, Hur JY, Choi YL, *et al.* Current status and future perspectives of liquid biopsy in non-small cell lung cancer. *J Pathol Transl Med*, 2020, 54(3): 204-212.
- 34 李媛秋, 么鸿雁. 肺癌主要危险因素的研究进展. *中国肿瘤*, 2016, 25(10): 782-786.
- 35 Moon SM, Kim JH, Kim SK, *et al.* Clinical utility of combined circulating tumor cell and circulating tumor DNA assays for diagnosis of primary lung cancer. *Anticancer Res*, 2020, 40(6): 3435-3444.
- 36 Zhong CH, Tong D, Zhou ZQ, *et al.* Performance evaluation of detecting circulating tumor cells and tumor cells in bronchoalveolar lavage fluid in diagnosis of peripheral lung cancer. *J Thorac Dis*, 2018, 10(Suppl 7): S830-S837.
- 37 Dai F, Wang J, An H, *et al.* Therapy of ¹²⁵I particles implantation inhibited the local growth of advanced non-small cell lung cancer: A retrospective clinical study. *Am J Transl Res*, 2019, 11(6): 3737-3749.
- 38 Zhou J, Diao X, Wang S, *et al.* Diagnosis value of combined detection of serum SF, CEA and CRP in non-small cell lung cancer. *Cancer Manag Res*, 2020, 12: 8813-8819.

收稿日期: 2021-01-21 修回日期: 2021-03-18

本文编辑: 董敏