DOI:10.13602/j.cnki.jcls.2022.03.08

·“迈瑞杯”临检发展专栏·

利用常见血清肿瘤标志物构建肺癌病理分型的风险预测模型*

王森钰,赖良,冯阳春(新疆医科大学附属肿瘤医院 & 第三临床医学院医学检验中心,乌鲁木齐 830011)

摘要:目的 联合运用 6 种常见血清肿瘤标志物建立预测肺癌常见病理分型的风险模型。**方法** 回顾性分析 2012 年 5 月至 2013 年 5 月新疆医科大学附属肿瘤医院收治的 342 例肺癌患者和 91 例肺部影像学可疑并确诊为非肺癌患者的临床资料。对 6 种肿瘤标志物进行热图分析,得到肿瘤标志物和肺癌各病理亚型的关系。Logistic 回归分析筛选各病理亚型独立预测指标,并构建预测模型。通过 ROC 曲线下面积(AUC^{ROC})评估各个模型的预测能力。运用 Bootstrap 重抽样法对模型的预测能力进行内部验证。**结果** 热图分析结果显示,Pro-GRP 和 NSE 预测小细胞肺癌,SCC 和 CYFRA21-1 预测肺鳞癌,CA125 和 CEA 预测肺腺癌的组合更好。小细胞肺癌、肺鳞癌和肺腺癌预测模型对应的 AUC^{ROC} 分别为 0.938、0.965 和 0.965。模型预测能力的内部验证结果表明 3 个模型预测效果理想。**结论** 血清肿瘤标志物联合患者一般资料构建针对肺癌患者病理分型的风险预测模型预测效果较好,验证后可以展示在检验报告单上,为临床医生诊疗提供帮助。

关键词:血清肿瘤标志物;肺癌;病理分型;预测模型

中图分类号:R734.2 **文献标志码:**A

Establishment of risk prediction models for pathological classification of lung cancer by common serum tumor markers

WANG Senyu, LAI Liang, FENG Yangchun (Department of Clinical Laboratory, Affiliated Cancer Hospital & Third Clinical Medical College of Xinjiang Medical University, Urumqi 830011, Xinjiang, China)

Abstract: Objective To establish a risk model for predicting the common pathological classification of lung cancer by the combination of six serum tumor markers. **Methods** The clinical data of 342 patients with lung cancer and 91 patients with suspicious lung imaging and diagnosed as non-lung cancer, who were treated in the Affiliated Cancer Hospital of Xinjiang Medical University during May 2012 and May 2013, were analyzed retrospectively. The relationship between tumor markers and pathological subtypes of lung cancer was obtained by the heat map analysis of 6 tumor markers. The independent predictive indicators of each pathological subtype were screened by the Logistic regression analysis, and a prediction model was constructed. The predictive ability of each model was evaluated by the area under the ROC curve (AUC^{ROC}). The Bootstrap resampling method was used to validate the predictive ability of each model. **Results** The heat map analysis showed that the combination of pro-GRP and NSE was better in predicting small cell lung cancer, SCC and CYFRA21-1 in predicting lung squamous cell carcinoma, and CA125 and CEA in predicting lung adenocarcinoma. The areas under the ROC curves of prediction models for small-cell lung cancer, lung squamous carcinoma and lung adenocarcinoma were 0.938, 0.965 and 0.965, respectively. Internal validation of the predictive ability of the three models showed that their predictive effects were ideal. **Conclusion** The risk prediction models for pathological classification of lung cancer patients constructed by serum tumor markers combined with general data of patients have good predictive effects. After further verification in the future, they could be displayed on the test report to provide help for clinicians in further diagnosis and treatment.

Key words: serum tumor marker; lung cancer; pathological classification; prediction model

《中华医学会肺癌临床诊疗指南(2019 版)》指出肺癌的诊断通常需要结合影像学和病理学检查[1]。肺癌血清肿瘤标志物的敏感性和特异性还不够高,但其升高可早于某些临床症状的出现,临床诊断时可根据需要检测肺癌相关的肿瘤标志物,行辅助诊断和早期鉴别诊断,并了解肺癌可能的病理类型。目前诊疗指南推荐常用的原发性肺癌标志物有癌胚抗原(carcino embryonic antigen, CEA)、神经元特异性烯醇化酶(neuron-specificenolase, NSE)、细胞角蛋白片段 19(cytokeratin 19 fragment, CYFRA21-1)、胃泌素释放肽前体(pro-gastrin-releasing peptide, ProGRP)、鳞状上皮细胞癌抗原(squa-

* **基金项目:**新疆医科大学附属肿瘤医院人文社会科学基金项目改革与发展专项(2019SK02003)。

作者简介:王森钰,1989 年生,女,主管技师,博士研究生,从事临床检验诊断学工作。

通信作者:冯阳春,副教授,博士,E-mail:paopao1987123@163.com。

mous cell carcinoma antigen, SCC)等[1]。本研究拟将上述 5 种血清肿瘤标志物及报道较多的糖类抗原 125(CA125)作为研究指标[2-4],构建血清肿瘤标志物预测肺癌病理分型的风险预测模型,以期为临床医生准确评估肺癌的病理类型提供实验依据。

1　材料和方法

1.1　资料来源　收集 2012 年 5 月至 2013 年 5 月新疆医科大学附属肿瘤医院初次入院就诊且确诊为原发性肺癌的患者相关资料。纳入标准:(1)就诊时有肺部结节并按照《原发性肺癌诊疗规范(2011 年版)》最终确诊;(2)纳入原发性肺癌中常见的病理类型,包括小细胞肺癌、肺鳞癌及肺腺癌;(3)患者均未进行任何放化疗以及手术治疗;(4)一般临床资料齐全。排除标准:(1)医院复诊的肺癌患者;(2)腺鳞癌、大细胞癌等肺癌不常见病理类型的肺癌患者;(3)肺部转移癌。按照以上标准,最终共计 342 例原发性肺癌患者纳入本次研究,男 209 例,女 133 例,年龄 28~82 岁,中位年龄 62 岁。其中小细胞肺癌 195 例,肺鳞状细胞癌 63 例,肺腺癌 84 例。另外纳入同期医院收治的 91 例肺部影像学可疑并确诊为非原发性肺癌患者作为对照组,包括肺结核 11 例,肺部炎症等肺部良性疾病 80 例。对照组年龄 37~70 岁,中位年龄 53 岁。本研究经新疆医科大学附属肿瘤医院伦理委员会批准(No.G-2019045)。

1.2　主要仪器及试剂　ARCHITECT i2000SR 全自动化学发光免疫分析仪(美国 Abbott 公司),Cobas e601 型全自动电化学发光免疫分析仪(德国 Roche 公司)。Pro-GRP 检测试剂盒、CEA 检测试剂盒、CA125 检测试剂盒、SCC 检测试剂盒、CYFRA21-1 检测试剂盒(美国 Abbott 公司),NSE 检测试剂盒(德国 Roche 公司)

1.3　标本采集及检测方法　分别采集原发性肺癌患者及良性患者晨起空腹(禁食 8~12 h)静脉血 3 mL,3 500×*g* 离心 15 min,血清样本置-20 ℃保存。采用 ARCHITECT i2000SR 全自动化学发光免疫分析仪,按照检验科制定的各项目的标准操作规程检测血清中 Pro-GRP、CEA、CA125、SCC 和 CYFRA21-1 的含量。采用 Cobas e601 型全自动电化学发光免疫分析仪,按照 NSE 检测试剂盒的检测标准操作规程检测血清中 NSE 的含量。

1.4　统计学分析　采用 R 包(来源于 http://cran.r-project.org/web/packages)进行热图分析及绘制 ROC 曲线,获得各指标预测肺癌病理分型的预测模型,使用 Bootstrap 重复采样法对模型的准确性进行内部验证,进行 1 000 次重复采样,计算模型的一致性指数,即 C 指数(index of concordance, C-index)。以 $P<0.05$ 为差异有统计学意义。

2　结果

2.1　6 种肺癌相关肿瘤标志物的热图分析　对 342 例肺癌患者的肿瘤标志物进行热图分析,结果显示,Pro-GRP 和 NSE 在小细胞肺癌中呈高表达,CA125 和 CEA 在肺腺癌中呈高表达,SCC 和 CYFRA21-1在肺鳞癌中呈高表达,见图 1。同时对 6 种常用血清肿瘤标志物中的 Pro-GRP 和 NSE,CA125 和 CEA,SCC 和 CYFRA21-1 分为 3 个初级分类;其次 CA125 和 CEA,SCC 和 CYFRA21-1 构成的 2 个初级分类又组成 1 个次级分类,而 Pro-GRP 和 NSE 构成的初级分类被分为另一类。综合上述结果可以认为 Pro-GRP 和 NSE 预测小细胞肺癌,SCC 和CYFRA21-1预测肺鳞癌,CA125 和 CEA 预测肺腺癌的组合更好。

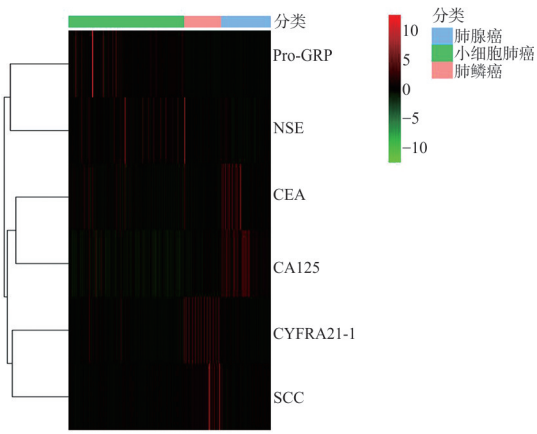


图 1　6 种常用血清肿瘤标志物在各病理类型肺癌表达的热图分析

2.2　6 种血清肿瘤标志物预测不同肺癌病理分型的模型　该研究纳入患者的一般资料,包括年龄、性别和吸烟状态,同时结合结果 2.1 部分血清肿瘤标志物的组合对小细胞肺癌、肺鳞癌和肺腺癌进行风险预测模型的构建,结果见图 2、表 1 和表 2。小细胞肺癌、肺鳞癌和肺腺癌预测模型对应的 AUC^{ROC} 分别为 0.938、0.965 和 0.965。

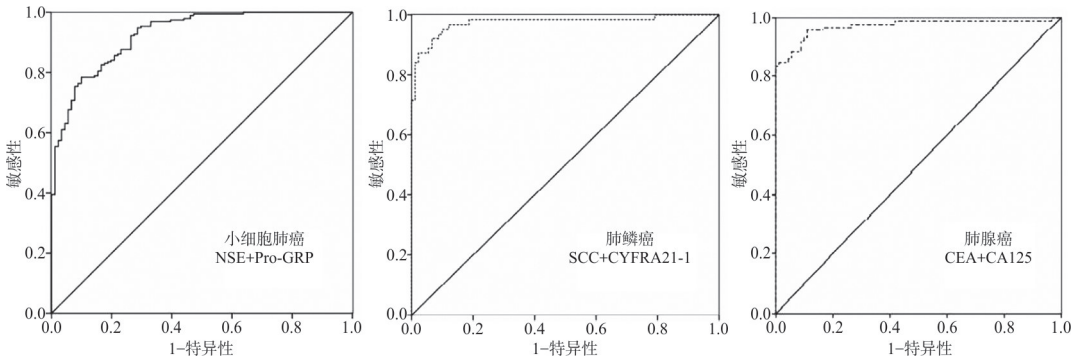


图 2　血清肿瘤标志物联合预测各种病理类型肺癌的 ROC 曲线

表 1　血清肿瘤标志物预测肺癌各种病理类型的多因素分析

分组	*n*	因素	回归系数	Wald X^2 值	*P* 值	*OR* 值	95%*CI*
小细胞肺癌	195	年龄	0.082	13.562	<0.001	1.086	1.039~1.134
		NSE	0.441	46.766	<0.001	1.554	1.370~1.763
		Pro-GRP	0.033	13.729	<0.001	1.034	1.016~1.052
		性别*	−0.020	0.002	0.966	0.980	0.380~2.525
		吸烟**	−0.438	0.826	0.363	0.645	0.251~1.659
		常量	−9.999	33.957	<0.001	—	—
肺鳞癌	63	年龄	0.114	10.062	0.002	1.121	1.045~1.203
		SCC	1.324	5.977	0.014	3.759	1.300~10.866
		CYFRA-211	0.877	9.336	0.002	2.405	1.370~4.222
		性别*	0.152	0.048	0.827	1.164	0.298~4.541
		吸烟**	−0.689	0.617	0.432	0.502	0.090~2.801
		常量	−11.255	20.145	<0.001	—	—
肺腺癌	84	年龄	0.171	17.632	<0.001	1.186	1.095~1.284
		CEA	0.456	7.084	0.008	1.578	1.128~2.209
		CA125	0.100	19.705	<0.001	1.105	1.057~1.154
		性别*	−0.905	1.522	0.217	0.404	0.096~1.704
		吸烟**	2.017	5.169	0.023	7.517	1.321~42.790
		常量	−15.200	26.464	<0.001	—	—

注：*，男性相比于女性；**，吸烟相比于不吸烟。

表 2　血清肿瘤标志物联合预测肺癌各种病理类型的模型

肺癌类型	预测模型
小细胞肺癌	$P=e^X/[1+e^X]$（$X=-9.999+[0.082\times$年龄$]+[0.441\times$NSE$]+[0.033\times$Pro-GRP$]$）
肺鳞癌	$P=e^X/[1+e^X]$（$X=-11.255+[0.114\times$年龄$]+[1.324\times$SCC$]+[0.877\times$CYFRA21-1$]$）
肺腺癌	$P=e^X/[1+e^X]$（$X=-15.200+[0.171\times$年龄$]+[0.456\times$CEA$]+[0.100\times$CA125$]+[2.017\times$吸烟$]$）

2.3　预测模型的内部验证结果　Bootstrap 重复采样法对 3 个模型的预测能力进行内部验证，结果证实 3 个模型的指数 C 分别为 0.938、0.965 和 0.965；原始曲线和校正曲线拟合良好。见图 3。表明预测模型的预测效果理想。

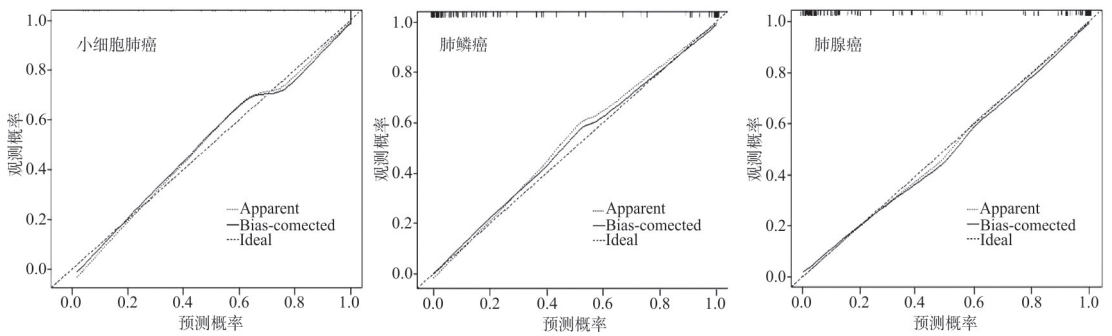


图 3　小细胞肺癌、肺鳞癌和肺腺癌各预测模型的校准曲线

3　讨论

血清肿瘤标志物作为传统的检测指标,在各临床实验室已经普遍开展。先前的研究主要聚焦于单个肿瘤标志物或者多个肿瘤标志物联合检测对恶性肿瘤的诊断价值,而随着研究的不断发展,应用血清肿瘤标志物对出现影像学特征等恶性肿瘤高危人群进行风险评估的模型已成为研究的热点,目前在临床检验领域应用最为广泛的风险评估模型是采用 CA125 和 HE4 血清肿瘤标志物计算的 ROMA 风险评估模型,该模型主要应用于女性盆腔肿块的良恶性风险评估,并已经出现在检验报告单上,对临床医生和患者帮助很大,体现了检验医学的价值[5-6]。对肺部影像学可疑的患者进行风险评估,构建类似于 ROMA 模型的风险评估模型,一直是肺部血清肿瘤标志物研究的热点。

在此次研究中,笔者首先利用热图分析的方法对实验室常用的 6 种血清肿瘤标志物进行分析,用图形分析的方法直观显示 6 种肿瘤标志物和常见肺癌病理分型之间的关系,这种创新性的分析手段在国内为首次报道。本次研究结果表明,Pro-GRP 和 NSE 预测小细胞肺癌,SCC 和CYFRA21-1预测肺鳞癌,CA125 和 CEA 预测肺腺癌的组合更好。这一结果和目前报道的血清肿瘤标志物和肺癌各种病理类型的相关性相一致[7-9]。

目前,针对肺部影像学可疑的患者建立的风险预测模型有很多,但是多是基于人群危险因素构建的模型,譬如 Mayo 模型、VA 模型和北大模型等[10]。这些模型只考虑患者的一般资料和影像学结果,然而,比较容易获得的血清肿瘤标志物检测结果多数被忽视。最近几年,推广比较多的是雅培诊断和中国肺癌防治联盟共同合作开发的 LCBP 风险模型[11],该风险模型利用 493 例患者的年龄、性别、吸烟史等一般资料,CT 检查信息,Pro-GRP、CEA、CYFRA21-1 和 SCC 4 项肺癌血清标志物,首次构建了有血清肿瘤标志物参与的肺癌风险评估模型,并在临床广泛应用。而本次研究利用血清肿瘤标志物结合患者的一般资料,构建针对各种肺癌病理类型的风险预测模型在以前的报道中较为罕见,构建的 3 个模型的 AUC^{ROC} 均大于 0.90,说明预测的准确性较高。本研究对 3 个模型的预测能力进行内部验证,结果表明各个模型的预测效果均比较理想。

由于 CT 检查结果对于检验工作人员而言很难获得,因此,本次研究模型并未纳入 CT 检查结果。同时相比于 LCBP 模型,更多地纳入了 NSE 和 CA125 2 个常见的肺癌肿瘤标志物,并按照病理分型分别建立模型,力求未来像 ROMA 模型那样,把 3 种常见肺癌病理类型分别计算风险展示在检验报告单上,为临床医生进一步诊疗提供依据。

本研究尚存在一定的局限性。首先,本研究仅为单中心的回顾性研究,样本量仅为 433 例,还存在一定的数量不足;其次,本研究仅对各个模型进行了内部验证,未来采用前瞻性研究进行外部验证非常有必要。综上所述,本研究建立的 3 个风险预测模型虽然可以较好地预测肺癌常见病理类型的发生,但是该结论还需多中心及前瞻性研究验证。

4　参考文献

[1] 王丽, 杨婷婷, 高启翔.《中华医学会肺癌临床诊疗指南(2019 版)》正式发布[J]. 中华医学信息导报, 2020, 35(11): 20.

[2] 周新华, 张余杭, 马逸婷. 肺癌患者外周血 CA15-3 和 CA125 表达与其病理特征的相关性[J]. 实用癌症杂志, 2021, 36(2): 224-227.

[3] 李一鸣. 血清肿瘤标志物 CA125 对非小细胞肺癌患者化疗效果的评估作用[J]. 中国老年学杂志, 2021, 41(15): 3184-3187.

[4] 王茜, 夏睿, 董惠霞, 等. ProGRP、NSE、CEA、CYFRA21-1、SCC 单独及联合检测在肺癌鉴别诊断中的价值[J]. 临床检验杂志, 2020, 38(12): 930-932.

[5] 田明, 吴玉梅. ROMA 模型在卵巢上皮性癌诊断中的应用价值[J]. 肿瘤学杂志, 2021, 27(1): 9-14.

[6] 朱莉, 秦丽, 王瑞雅, 等. 血清 CA125、HE4、VEGF 及 ROMA 指数对卵巢癌诊断及预测术后复发的价值[J]. 广东医学, 2021, 42(7): 796-801.

[7] 吕鑫, 张翊玲, 姚红梅. 基于知识图谱分析肿瘤标志物在肺癌中应用的研究热点、方向及趋势[J]. 山东医药, 2020, 60(12): 1-4.

[8] 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.

[9] 冯阳春, 徐怡, 黄艳春. 聚类和判别分析法在肺癌六种肿瘤标志物诊断中的应用[J]. 肿瘤防治研究, 2015, 42(3): 266-269.

[10] 安杨, 李春笋, 赵微. 三种肺癌风险预测模型在肺结节患者中的验证性研究[J]. 解放军医学院学报, 2020, 41(12): 1193-1196.

[11] Yang DW, Zhang XJ, Powell CA, *et al*. Probability of cancer in high-risk patients predicted by the protein-based lung cancer biomarker panel in China: LCBP study[J]. Cancer, 2018, 124(2): 262-270.

(**收稿日期**:2021-08-20)

(**本文编辑**:许晓蒙)