

Logistic 回归和人工神经网络在鉴别诊断肺癌性 胸腔积液中的应用研究

李锐成¹⁾, 千红维¹⁾, 范艳妮²⁾, 赵佩佩¹⁾, 魏 珊¹⁾, 景花荣¹⁾

(1)空军军医大学第二附属医院唐都医院检验科; 2)信息科, 陕西 西安 710038)

[摘要] 目的 应用 Logistic 回归分析方法和人工神经网络(artificial neural network, ANN)技术, 评估血清 (serum, S-) 和胸腔积液 (pleural effusion, P-) 中的癌胚抗原(CEA)、铁蛋白(FRT)、神经元特异性烯醇化酶(NSE)、鳞状细胞癌相关抗原(SCC)、糖类抗原 50(CA50)、糖类抗原 125(CA125)和细胞角蛋白19 片段(CY21-1) 在鉴别肺癌性胸腔积液(malignant pleural effusion of lung cancer, LC-MPE)与良性胸腔积液(benign pleural effusion, BPE)中的应用价值, 建立通过肿瘤标志物诊断 LC-MPE 的诊断模型。方法 对临床初诊的 LC-MPE 和 BPE 患者的血清和胸腔积液肿瘤标志物结果进行分析, 应用 Logistic 回归分析和 ANN 技术分别建立诊断 LC-MPE 的诊断模型。结果 S-NSE、S-CY21-1、P-CEA 和 P-NSE4 项指标被筛选出并用于建模, 研究建立的诊断 LC-MPE 的 Logistic 回归模型的灵敏度为 93.23%, 特异度为 97.46%, ROC 曲线下面积为 0.992。建立的 ANN 模型的灵敏度为 95.35%, 特异度为 97.22%, ROC 曲线下面积为 0.990 ($P < 0.05$)。结论 在通过肿瘤标志物诊断 LC-MPE 方面, 建立的 Logistic 回归模型和 ANN 模型均有较好的诊断性能, 上述 2 个模型均可辅助临床医生提高诊断准确率。

[关键词] 胸腔积液; 人工神经网络; 肿瘤标志物; 肺癌; 诊断模型

[中图分类号] R446.11 [文献标志码] A [文章编号] 2095-610X(2024)10-0055-06

Application of Logistic Regression and Artificial Neural Networks in the Differential Diagnosis of LC-MPE

LI Ruicheng¹⁾, QIAN Hongwei¹⁾, FAN Yanni²⁾, ZHAO Peipei¹⁾, WEI Shan¹⁾, JING Huarong¹⁾

(1) Dept. of Clinical Laboratory; 2) Dept. of Information, The 2nd Affiliated Hospital of Air Force Medical University, Xi'an Shaanxi 710038, China)

[Abstract] Objectives To evaluate the application value of carcinoembryonic antigen (CEA), ferritin (FRT), neuron-specific enolase (NSE), squamous cell carcinoma-related antigen (SCC), carbohydrate antigen 50 (CA50), carbohydrate antigen 125 (CA125), and cytokeratin 19 fragment (CY21-1) in serum (S-) and pleural effusion (P-) for differentiating malignant pleural effusion of lung cancer (LC-MPE) from benign pleural effusion (BPE). We aim to establish a diagnostic model for LC-MPE using tumor markers and analyze the data using logistic regression and artificial neural network (ANN) techniques. Methods The serum and pleural effusion tumor marker results of patients with newly diagnosed LC-MPE and BPE were analyzed, and diagnostic models for LC-MPE were established using Logistic regression analysis and ANN technology. Results The indicators S-NSE, S-CY21-1, P-CEA, and P-NSE were selected and used for modeling. The Logistic regression model for diagnosing LC-MPE established in this study had a sensitivity of 93.23% and a specificity of 97.46%, with an area under the ROC curve of 0.992. The established ANN model had a sensitivity of 95.35%, a specificity of 97.22%, and an area under the ROC curve of 0.990 ($P < 0.05$). Conclusions In diagnosing LC-MPE through tumor

[收稿日期] 2024-05-20

[基金项目] 国家自然科学基金资助项目(81772485)

[作者简介] 李锐成(1985~), 男, 河南商丘市, 医学硕士, 副主任技师, 主要从事肿瘤临床检验诊断方面的研究工作。

[通信作者] 景花荣, E-mail: 43277715@qq.com

markers, both the Logistic regression model and the ANN model established in this study showed good diagnostic efficacy. These two models can assist clinicians in improving diagnostic accuracy.

[**Key words**] Pleural effusion; Artificial neural network; Tumor markers; Lung cancer; Diagnostic model

胸腔积液待查在临床中较为常见, 因其产生原因较多, 病因复杂, 对其有效鉴别一直是较为棘手的难题^[1-2]。多项研究报道, 引起渗出性胸腔积液最多见的原因因为肺癌、结核和肺炎^[3-5]。疾病的良恶性决定着致病机理、治疗方案及预后转归。及早明确诊断、采取针对性治疗对患者具有重要的临床意义^[6]。目前, 关于肿瘤标志物与肺癌性胸腔积液 (malignant pleural effusion associated with lung cancer, LC-MPE) 及良性胸腔积液 (benign pleural effusion, BPE) 鉴别诊断的相关性研究较多, 但多为 1 个或几个指标的逐项经验性总结, 且多局限于血清肿瘤标志物, 缺乏胸腔积液多指标的、量化的系统研究。本研究试图运用 Logistic 回归分析和人工神经网络 (artificial neural network, ANN) 方法, 对血清和胸腔积液共计 14 项肿瘤标志物进行多指标研究, 建立诊断模型, 提高肿瘤标志物对 LC-MPE 的诊断准确率。

1 资料与方法

1.1 研究对象

研究对象来源于 2020 年 08 月至 2023 年 02 月期间唐都医院胸腔外科和呼吸科收治的不明原因胸腔积液患者。纳入标准: (1) 出院最终诊断明确, 符合胸腔积液诊断标准^[7]; (2) 均为 PE 初诊患者, 尚未采取任何针对性治疗; (3) 无严重肝、肾、心功能异常; (4) 未合并其它部位肿瘤; (5) 临床资料完整。排除标准: (1) 过往肿瘤病史或结核病史; (2) 淋巴管阻塞; (3) 伴严重肝、肾、心疾病; (4) 临床资料不完善。诊断经由临床症状、影像学检查、实验室检查并最终由组织或胸水病理学或试验性诊断确诊, 筛选出临床资料完备的 251 例胸腔积液患者。251 例患者中, LC-MPE 组 133 例 (肺腺癌 106 例, 肺鳞状细胞癌 15 例, 肺小细胞癌 12 例), 女性 56 例, 男性 77 例, 平均年龄 (64.06 ± 10.84) 岁; BPE 组 118 例 (结核性胸腔积液 77 例, 肺炎性胸腔积液 41 例), 女性 37 例, 男性 81 例, 平均年龄 (62.51 ± 12.81) 岁。性别和年龄差异无统计学意义 ($P > 0.05$), 见表 1。本项研究取得了唐都医院伦理委员会的批准, 并获得了受试者知情同意的豁免 (TDLL-20170924)。

表 1 2 组胸腔积液患者基本资料 [n(%)]

Tab. 1 Basic characteristics of patients with pleural effusion in 2 groups [n(%)]

指标	LC-MPE组 (n=133)	BPE组 (n=118)	χ^2/t	P
性别			3.098	0.078
男	77(57.90)	81(68.64)		
女	56(42.10)	37(31.36)		
年龄 (岁)			1.029	0.305
($\bar{x} \pm s$)	37 ~ 88 64.06 ± 10.84	35 ~ 89 62.51 ± 12.81		

1.2 标本收集及检测

血清和胸腔积液肿瘤标志物结果的收集均在初次诊断时。样本要求新鲜、足量, 留取后立即送检。血清和胸腔积液中的肿瘤标志物均包括癌胚抗原 (CEA)、铁蛋白 (FRT)、神经元特异性烯醇化酶 (NSE)、鳞状细胞癌相关抗原 (SCC)、糖类抗原 50 (CA50)、糖类抗原 125 (CA125) 和细胞角蛋白 19 片段 (CY21-1), 共计 7 种 14 项指标。为了以下内容表述方便, 特将血清肿瘤标志物记为 “S-”, 胸腔积液肿瘤标志物记为 “P-”。2 种标本类型的肿瘤标志物检测均采用电化学发光法, 仪器采用德国罗氏 E-601 型电化学发光自动分析仪, 试剂为德国罗氏配套产品, 操作严格按照标准程序进行。

1.3 统计学处理

采用 IBM SPSS 软件 18.0 进行分析, 2 组间的性别比例比较采用 χ^2 检验, 年龄以均数 ± 标准差 ($\bar{x} \pm s$) 表示并采用 The Student's *t*-test, 肿瘤标志物数据以中位数和四分位数间距表示并进行 Nonparametric Mann-Whitney U test; $P < 0.05$ 为差异具有统计学意义。以胸腔积液类型为因变量, 以 2 组间差异有统计学意义的肿瘤标志物指标为协变量, 应用二元 Logistic 多因素回归分析, 采用 “逐步法” 筛选出对构建模型具有统计学意义的指标, 构建诊断 LC-MPE 的 Logistic 诊断模型。同时将上述筛选出的指标作为输入端, 胸腔积液类型作为输出端, 在此基础上构建肿瘤标志物指标数据库, 随机选取其中的 172 例构成测试集 (96 例 MPE, 76 例 BPE), 其余的 79 例构成验证集 (37 例 MPE, 42 例 BPE), 建立 ANN 诊断模型。

2 结果

2.1 肿瘤标志物在 2 组中的检测结果比较

为了明确血清和胸腔积液共计 7 种 14 项肿瘤

标志物指标在 LC-MPE 组和 BPE 组间的水平差异, 本研究分别对各指标进行了非参数检验, 结果发现仅有血清和胸腔积液中 SCC 两项指标在 2 组间无统计学意义 ($P > 0.05$), 其它 12 项指标结果比较, 差异均有统计学意义 ($P < 0.05$), 见表 2。

表 2 各单项肿瘤标志物在两组中的结果比较 [$M(P25, P75)$]

Tab. 2 Comparison of results for individual tumor markers between the two groups [$M(P25, P75)$]

指标	LC-MPE组($n=133$)	BPE组($n=118$)	Z	P
S-CEA (ng/mL)	9.70 (3.64, 49.01)	1.54 (1.03, 2.33)	10.155	<0.001*
S-FRT (ng/mL)	179.00 (109.30, 307.25)	299.85 (152.93, 517.55)	4.443	<0.001*
S-NSE (ng/mL)	17.84 (12.86, 25.85)	13.79 (10.51, 19.21)	4.450	<0.001*
S-SCC (ng/mL)	0.48 (0.34, 0.65)	0.46 (0.36, 0.56)	1.227	0.220
S-CA50 (U/mL)	8.16 (5.16, 16.45)	5.44 (3.65, 7.99)	4.841	<0.001*
S-CA125 (U/mL)	99.89 (47.13, 185.30)	76.25 (39.10, 127.20)	2.682	0.007*
S-CY21-1 (ng/mL)	5.64 (3.21, 12.19)	1.83 (1.30, 2.62)	10.946	<0.001*
P-CEA (ng/mL)	125.90 (13.55, 623.90)	1.21 (0.68, 1.78)	11.486	<0.001*
P-FRT (ng/mL)	1410.00 (883.65, 2000.00)	1138.00 (597.53, 2000.00)	2.414	0.016*
P-NSE (ng/mL)	10.67 (5.51, 30.69)	5.73 (2.99, 10.74)	5.136	<0.001*
P-SCC (ng/mL)	0.82 (0.57, 1.81)	1.00 (0.65, 2.03)	1.428	0.153
P-CA50 (U/mL)	8.39 (3.58, 42.95)	3.62 (2.11, 5.93)	6.638	<0.001*
P-CA125 (U/mL)	965.40 (451.85, 1697.50)	397.25 (177.90, 804.38)	6.056	<0.001*
P-CY21-1 (ng/mL)	87.28 (30.16, 227.45)	10.23 (5.63, 22.34)	10.126	<0.001*

* $P < 0.05$ 。

2.2 各单项肿瘤标志物指标的 diagnostic 性能评价

在初步筛选出 2 组间差异有统计学意义的 12 项指标后, 本研究利用二元 Logistic 回归方法对该 12 项指标逐个进行单因素回归分析以进一步明确各指标的鉴别诊断效能。12 项指标中, P-CEA 拥有最好的诊断性能, 灵敏度、特异度、准确率 (π)、约登指数(YI)和曲线下面积(AUC)分别为 84.96%、98.31%、91.24%、0.833 和 0.920; 灵敏度、特异度、准确率、YI 和 AUC 最低的指标分别为 P-NSE(47.37%)、S-FRT(42.37%)、S-CA125 (56.57%)、S-CA125(0.134)和 P-FRT(0.587); 同时具备了 YI 大于 0.5 且 AUC 大于 0.7 的指标有 S-CEA、S-CY21-1、P-CEA 和 P-CY21-1, 见表 3。

2.3 构建 Logistic 回归诊断模型

由于各单项肿瘤标志物指标的 diagnostic 性能参差不齐, 见表 3, 为达到灵敏度和特异度的完美统一, 进一步提高诊断性能。本研究对该 12 项指标进行二元 Logistic 多因素回归分析, 为达到指标数量和诊断性能的良好结合, 尽量避免各指标间的相互影响, 在建模过程中采用“逐步法”。结果显示自变量“S-NSE、S-CY21-1、P-CEA、P-NSE”在 Logistic 诊断模型建立过程中的作用差异有统计学意义 ($P < 0.05$), 其余指标差异无统计学

意义 ($P > 0.05$), 见表 4。将 S-NSE、S-CY21-1、P-CEA、P-NSE 做为协变量, 胸腔积液类型做为因变量, 建立的 Logistic 诊断模型对 251 例病例进行预测, 结果显示: 灵敏度、特异度、YI 和 AUC 分别为 93.23%、97.46%、0.907 和 0.992, 见表 5。

2.4 构建 ANN 诊断模型及 diagnostic 性能分析

本研究同时将通过二元 Logistic 多因素回归分析筛选出的 S-NSE、S-CY21-1、P-CEA、P-NSE 作为协变量输入端, 胸腔积液类型作为输出端, 随机分出的 172 例用来训练神经网络, 其余的 79 例作为测试集, 建立 ANN 诊断模型。采用 BP(back propagation)神经网络, 其拓扑结构包含 3 层, 分别为输入层、隐含层和输出层。输入层神经节点为 4 个, 包括 S-NSE、S-CY21-1、P-CEA、P-NSE, 采用双曲正切传递函数; 输出层神经节点 2 个, 分别对应 BPE 和 LC-MPE, 采用 Softmax 传递函数。应用该模型对测试集样本进行预测, 灵敏度、特异度、YI 和 AUC 分别为 95.35%、97.22%、0.926 和 0.990, 显示出 ANN 模型对 LC-MPE 和 BPE 有较好的预测概率, 见表 5 和图 1。与此同时通过 ANN 模型的敏感性分析显示了 P-CEA 和 S-CY21-1 是构成 ANN 预测模型的最重要

表 3 各单项肿瘤标志物的诊断性能分析

Tab. 3 Analysis of the diagnostic performance of individual tumor markers

指标	灵敏度(%)	特异度(%)	阳性 预测值(%)	阴性 预测值(%)	准确率(%)	约登指数	AUC	<i>P</i>	95%CI
S-CEA	75.19	94.92	94.34	77.24	84.46	0.701	0.871	<0.001*	0.826 ~ 0.917
S-FRT	81.95	42.37	61.58	67.57	63.35	0.243	0.663	<0.001*	0.595 ~ 0.730
S-NSE	57.89	66.10	65.81	58.21	61.75	0.240	0.663	<0.001*	0.596 ~ 0.729
S-CA50	51.88	76.27	71.13	58.44	63.35	0.282	0.677	<0.001*	0.612 ~ 0.743
S-CA125	54.89	58.47	59.84	53.49	56.57	0.134	0.598	0.007*	0.528 ~ 0.668
S-CY21-1	72.93	86.44	85.84	73.91	79.28	0.594	0.900	<0.001*	0.864 ~ 0.937
P-CEA	84.96	98.31	98.26	85.29	91.24	0.833	0.920	<0.001*	0.880 ~ 0.961
P-FRT	66.17	47.46	58.67	55.45	57.37	0.136	0.587	0.017*	0.516 ~ 0.659
P-NSE	47.37	77.97	70.79	56.79	61.75	0.253	0.688	<0.001*	0.623 ~ 0.753
P-CA50	56.39	84.75	80.65	63.29	69.72	0.411	0.743	<0.001*	0.682 ~ 0.803
P-CA125	58.65	68.64	67.83	59.56	63.35	0.273	0.722	<0.001*	0.660 ~ 0.783
P-CY21-1	63.91	89.83	87.63	68.83	76.10	0.537	0.870	<0.001*	0.828 ~ 0.913

**P* < 0.05。

表 4 肿瘤标志物的二元 Logistic 多因素分析

Tab. 4 Binary Logistic Multivariate Analysis of Tumor Markers

指标	β	<i>SE</i>	<i>Wald</i>	<i>P</i>	<i>OR</i>	95%CI
S-NSE	0.099	0.042	5.627	0.018*	1.104	(1.017 ~ 1.197)
S-CY211	0.606	0.214	8.048	0.005*	1.833	(1.206 ~ 2.787)
P-CEA	0.829	0.225	13.577	<0.001*	2.291	(1.474 ~ 3.562)
P-NSE	0.037	0.017	4.738	0.030*	1.038	(1.004 ~ 1.073)
常数	-7.900	1.503	27.645	<0.001*	0.000	-

**P* < 0.05。“-”表示无数据。

表 5 2 种诊断模型的诊断性能比较

Tab. 5 Comparison of diagnostic performance of two diagnostic models

诊断模型	灵敏度(%)	特异度(%)	阳性预测值(%)	阴性预测值(%)	约登指数	AUC	<i>P</i>	95%CI
Logistic	93.23	97.46	97.64	92.74	0.907	0.992	<0.001*	0.986 ~ 0.999
ANN-测试集	90.63	98.37	98.31	90.98	0.890	0.992	<0.001*	0.985 ~ 0.999
ANN-验证集	95.35	97.22	97.62	94.59	0.926	0.990	<0.001*	0.982 ~ 0.998

**P* < 0.05。

因素。

3 讨论

胸腔积液作为 1 种肺癌远处转移的常见并发症，它不仅是独立的危险因素，同时也预示着患者的病情恶化和极低生存率^[8]。因此，通过胸腔穿刺，胸膜细胞学或组织学诊断活检和胸腔镜检查鉴别 LC-MPE 与 BPE 是临床干预的第一步，也是关键的一步，以便早期干预和适当治疗。但是，许多 LC-MPE 患者总是有病理及细胞学阴性的不明确诊断。事实上，检测血清和胸腔积液中的肿

瘤标志物，有助于对高度怀疑 LC-MPE 的患者进行正确诊断^[9]。然而，由于诊断准确率较低，多在 30% 至 60% 徘徊，限制了它们的临床应用价值^[10]。因此，选择多项生物标志物并对其联合进而建立诊断模型对 LC-MPE 的早期检测和预测肺癌的预后至关重要。

本研究共纳入的肿瘤标志物共有 7 种，对血清和胸腔积液分别检测，共获取 14 项指标。首先对 14 项指标进行了 2 组间的结果比较，发现仅有血清和胸腔积液中的 SCC 在 2 组之间无差异。作为抗原糖蛋白，SCC 是由鳞癌细胞产生并分泌，多见于子宫鳞癌和肺鳞癌，用于诊断和预后^[11]。

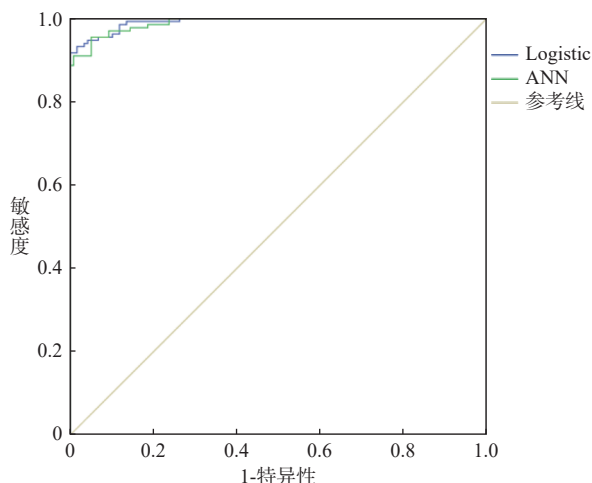


图 1 Logistic 和 ANN 模型诊断 LC-MPE 和 BPE 的 ROC 曲线

Fig. 1 The ROC curves of Logistic and ANN model for differential diagnosis of LC-MPE and BPE

但 Ma Q^[12] 报道 SCC 不仅可在鳞癌中升高, 而且在一些特殊炎症时同样可以升高, 这与本研究结果相符。在对其它 12 项指标进行逐项诊断性能时发现 P-CEA 有最高的诊断价值, 灵敏度、特异度、准确率、YI 和 AUC 分别为 84.96%、98.31%、91.24%、0.833 和 0.920, 原因为癌肿侵及胸膜时, 会大量分泌 CEA 并释放入胸腔积液, 因其分子量较大, 极难经由胸膜入血, 不能像血液中的 CEA 进入肝脏代谢降解, 致使 LC-MPE 中的 CEA 远高于血清^[13]。CEA 在本研究中显示出较好的诊断性能, 这与 Son^[14] 的报道相符。与 CEA 相似的指标还有 CY21-1, LC-MPE 组中的 S-CY21-1 和 P-CY21-1 中位数分别为 5.64 ng/mL 和 87.28 ng/mL, 远高于 BPE 组(1.83 ng/mL 和 10.23 ng/mL)。并且无论是 LC-MPE 组还是 BPE, 胸腔积液中 CY21-1 含量远远高于血清。Shin 等^[15] 认为目前还没有其他肿瘤标志物能够代替 P-CEA 进行 LC-MPE 诊断, P-CEA 升高可预测 LC-MPE 的特异性可大于 90.00%, 但敏感性相对较低。也有学者对其它指标如 CA15-3, CA-125, CA-199, CY21-1 进行了评估, 但结果均不理想^[16]。

近年来, Logistic 回归和 ANN 已在医学领域中得到了广泛应用。Logistic 回归是 1 种经典的统计方法, 以输出结果是概率值的形式将因变量与自变量之间的非线性关系转换为线性关系。它不仅用于疾病危险因素的筛选, 还具有诊断和预测功能, 对定性、半定量、定量的指标均可适用, 具有限制条件少, 样本量相对低的优点^[17-18]。ANN 是 1 种模仿大脑神经元的结构和功能的计算模型,

在没有给定的明确编程的情况下学习任务, 对输入和输出间复杂的关系进行探索, 自动从数据集中抽取函数关系, 具有自适应性、高度容错性和非线性综合推理能力, 可高效、客观地建立准确的疾病诊断模型^[19-20]。Chen^[21] 利用 ANN 模型基于基因表达数据对非小细胞肺癌辅助化疗进行分类。Ligor T^[22] 利用 ANN 模型对肺癌潜在生物标志物进行选择。本研究将二元 Logistic 回归和 ANN 技术进一步应用于血清和胸腔积液肿瘤标志物与胸腔积液类型关系的研究, 在预分析的基础上选择可能影响胸腔积液鉴别诊断的 S-CEA、S-FRT 和 P-CEA 等 12 个肿瘤标志物指标, 通过二元 Logistic 回归多因素分析方法筛选出了 S-NSE、S-CY21-1、P-CEA 和 P-NSE 共 4 项指标, 以此 4 项指标为自变量分别建立了 Logistic 诊断模型和 ANN 诊断模型。笔者建立此 2 种模型的灵敏度、特异度、阳性预测值、阴性预测值、YI 和 AUC 分别为 93.23% VS 95.35%、97.46% VS 97.22%、97.64% VS 97.62%、92.74% VS 94.59%、0.907 VS 0.920 和 0.992 VS 0.990, 显示出良好的诊断性能。此外通过 Logistic 模型的 OR 值和 ANN 模型的敏感性分析均显示了 4 项指标中 P-CEA 和 S-CY21-1 是构成模型的最重要因素。原因可能为: (1) 肺腺癌是肺癌中最易产生胸腔积液且发病率最高的病理类型, 而 CEA 作为肺腺癌相对特异的肿瘤标志物, 加上 CEA 在胸腔积液中相对血清中难降解的特性, 使得 P-CEA 在鉴别诊断 LC-MPE 和 BPE 中具有最强的优势; (2) 肺鳞癌是产生胸腔积液和发病率仅次于肺腺癌的病理类型, 而 CY21-1 作为肺鳞癌相对特异的肿瘤标志物不具有像 CEA 在胸腔积液和血清中浓度差别非常明显的特点, 并且 S-CY21-1 的鉴别诊断能力高于 P-CY21-1。

综上所述, 在鉴别诊断 LC-MPE 和 BPE 中, 利用 Logistic 模型和 ANN 模型联合多项肿瘤标志物指标的方法所具备的诊断效能远高于单项指标, 要在尽可能的情况下进行多指标联合诊断, 但同时应该看到并不是纳入的指标越多越好, 通过 Logistic 回归从 12 项指标筛选出 S-NSE、S-CY21-1、P-CEA 和 P-NSE, 这 4 项指标的联合建模具有最高的诊断性能, 可辅助临床医生提高诊断准确率。本研究的创新之处在于通过建立 Logistic 模型和 ANN 模型, 用较少的检验指标实现最大鉴别诊断性能, 既节约了医疗成本、减轻了病人经济负担还达到了较高的鉴别诊断准确性。但同时也有一些局限性, 如临床中肺腺癌引起的胸腔积液居多, 导致本研究中纳入的鳞癌和小细胞癌的

样本量有限,在后续研究中,笔者将进一步扩大肺鳞癌和小细胞癌病例校正和完善该模型。

[参考文献]

- [1] Kim N Y, Jang B, Gu K M, et al. Differential diagnosis of pleural effusion using machine learning[J]. *Ann Am Thorac Soc*, 2024, 21(2): 211–217.
- [2] Gonnelli F, Hassan W, Bonifazi M, et al. Malignant pleural effusion: current understanding and therapeutic approach[J]. *Respir Res*, 2024, 25(1): 47.
- [3] Han Y Q, Yan L, Li P, et al. A study investigating markers in pleural effusion (SIMPLE): A prospective and double-blind diagnostic study[J]. *BMJ Open*, 2019, 9(8): e27287.
- [4] Brun C, Gay P, Cottier M, et al. Comment from the authors: the tests combination in patients with lung cancer and malignant pleural effusion[J]. *J Thorac Dis*, 2019, 11(5): E74–E75.
- [5] Liu Q, Yu Y X, Wang X J, et al. Diagnostic accuracy of interleukin-27 between tuberculous pleural effusion and malignant pleural effusion: A meta-analysis[J]. *Respiration*, 2018, 95(6): 469–477.
- [6] Da C L V, Ribeiro-Alves M, Da S C R, et al. Predominance of Th1 immune response in pleural effusion of patients with tuberculosis among other exudative etiologies[J]. *J Clin Microbiol*, 2019, 58(1): e919–e927.
- [7] 中华医学会呼吸病学分会胸膜与纵隔疾病学组. 胸腔积液诊断的中国专家共识 [J]. *中华结核和呼吸杂志*, 2022, 45(11): 1080–1096.
- [8] Gonnelli F, Hassan W, Bonifazi M, et al. Malignant pleural effusion: Current understanding and therapeutic approach[J]. *Respir Res*, 2024, 25(1): 1–11.
- [9] Wu Q, Li M, Zhang S, et al. Clinical diagnostic utility of CA 15-3 for the diagnosis of malignant pleural effusion: A meta-analysis[J]. *Exp Ther Med*, 2015, 9(1): 232–238.
- [10] Nguyen A H, Miller E J, Wichman C S, et al. Diagnostic value of tumor antigens in malignant pleural effusion: A meta-analysis[J]. *Transl Res*, 2015, 166(5): 432–439.
- [11] Charakorn C, Thadanipon K, Chaijindaratana S, et al. The association between serum squamous cell carcinoma antigen and recurrence and survival of patients with cervical squamous cell carcinoma: A systematic review and meta-analysis[J]. *Gynecol Oncol*, 2018, 150(1): 190–200.
- [12] Ma Q, Liu W, Jia R, et al. Inflammation-based prognostic system predicts postoperative survival of esophageal carcinoma patients with normal preoperative serum carcinoembryonic antigen and squamous cell carcinoma antigen levels[J]. *World J Surg Oncol*, 2016, 14(14): 1–6.
- [13] 李锐成, 郜赵伟, 董轲, 等. 胸腔积液与血清中的癌胚抗原及其比值对结核性与肺癌性胸腔积液的诊断价值[J]. *南方医科大学学报*, 2019, 39(2): 175–180.
- [14] Son S M, Han H S, An J Y, et al. Diagnostic performance of CD66c in lung adenocarcinoma-associated malignant pleural effusion: Comparison with CEA, CA 19-9, and CYFRA 21-1[J]. *Pathology*, 2015, 47(2): 123–129.
- [15] Shin Y M, Yun J, Lee O J, et al. Diagnostic value of circulating extracellular miR-134, miR-185, and miR-22 Levels in lung adenocarcinoma-associated malignant pleural effusion[J]. *Cancer Res Treat*, 2014, 46(2): 178–185.
- [16] Abbas M, Kassim S A, Habib M, et al. Clinical evaluation of serum tumor markers in patients with advanced-stage non-small cell lung cancer treated with palliative chemotherapy in China[J]. *Front Oncol*, 2020, 10(6): 1–12.
- [17] Heijnen B J, Bohringer S, Speyer R. Prediction of aspiration in dysphagia using logistic regression: oral intake and self-evaluation[J]. *Eur Arch Otorhinolaryngol*, 2020, 277(1): 197–205.
- [18] Chiou S H, Betensky R A, Balasubramanian R. The missing indicator approach for censored covariates subject to limit of detection in logistic regression models[J]. *Ann Epidemiol*, 2019, 38(10): 57–64.
- [19] Teglia C M, Guinez M, Goicoechea H C, et al. Enhancement of multianalyte mass spectrometry detection through response surface optimization by least squares and artificial neural network modelling[J]. *J Chromatogr A*, 2020, 1611(40): 460613.
- [20] Yogeswari M K, Dharmalingam K, Mullai P. Implementation of artificial neural network model for continuous hydrogen production using confectionery wastewater[J]. *J Environ Manage*, 2019, 252(20): 109684.
- [21] Chen Y C, Chang Y C, Ke W C, et al. Cancer adjuvant chemotherapy strategic classification by artificial neural network with gene expression data: An example for non-small cell lung cancer[J]. *J Biomed Inform*, 2015, 56(4): 1–7.
- [22] Ligor T, Pater L, Buszewski B. Application of an artificial neural network model for selection of potential lung cancer biomarkers[J]. *J Breath Res*, 2015, 9(2): 27106.