

四种指标联合检测在肺癌鉴别诊断中的价值

徐 鹏, 胡 冬(四川省绵阳市中心医院检验科 621000)

【摘要】 目的 探讨血清中癌胚抗原(CEA)、糖类抗原(CA)125、细胞角蛋白 19 片段(CYFRA21-1)和胃泌素释放前肽(Pro-GRP)四种指标联合检测在肺癌鉴别诊断中的价值。**方法** 选择 2014 年 8 月至 2015 年 9 月该院收治的 92 例肺癌患者为肺癌组,选择同期 142 例肺部良性疾病患者作为对照组,采用化学发光法和酶联免疫吸附试验(ELISA)测定血清中 CEA、CA125、CYFRA21-1 和 Pro-GRP 水平。运用线性判别分析(LDA)法建立肺癌诊断与预测模型,并与并联试验结果的准确性进行比较。采用内代法,随机选取 2/3 的患者为训练集,1/3 的患者为测试集,考察模型的预测效果。**结果** 肺癌组血清中 CEA、CA125、CYFRA21-1 和 Pro-GRP 水平均显著高于对照组,差异均有统计学意义($P < 0.05$)。LDA 模型能正确诊断 81.5% 的肺癌患者和 83.1% 的肺部良性疾病患者,模型总的诊断准确率为 82.5%。该模型能对 76.5% 的肺癌和 80.0% 的肺部良性疾病进行预测,其预测准确率为 80.8%。并联试验结果表明,CEA、CA125、CYFRA21-1 和 Pro-GRP 联合检测诊断肺癌的灵敏度、特异度和准确率分别为 80.4%、76.1% 和 77.8%。**结论** 联合检测血清中 CEA、CA125、CYFRA21-1 和 Pro-GRP 提高了肺癌诊断的灵敏度。基于血清中 CEA、CA125、CYFRA21-1 和 Pro-GRP 的 LDA 模型能较好地诊断和预测肺癌,该模型为肺癌的鉴别诊断和预测提供了一种全新的思路。

【关键词】 肺癌; 癌胚抗原; 糖类抗原 125; 胃泌素释放前肽; 线性判别分析

DOI: 10.3969/j.issn.1672-9455.2016.10.026 文献标志码: A 文章编号: 1672-9455(2016)10-1367-03

Value of combined detection of 4 indicators in differentiation diagnosis of lung cancer XU Peng, HU Dong (Department of Clinical Laboratory, Mianyang Municipal Central Hospital, Mianyang, Sichuan 621000, China)

【Abstract】 **Objective** To evaluate the value of 4-indicator combined detection of serum carcinoembryonic antigen(CEA), carbohydrate antigen 125 (CA125), cytokeratin 19 fragments (CYFRA21-1) and pro-gastrin-releasing peptide(Pro-GRP) in the differentiation diagnosis of lung cancer. **Methods** Ninety-two cases of lung cancer in our hospital from August 2014 to September 2015 were selected as the lung cancer group and contemporaneous 142 cases of benign lung disease were taken as the control group. The chemiluminescence and ELISA assays were used to determine the levels of serum CEA, CA125, CYFRA 21-1 and Pro-GRP. The linear discriminant analysis(LDA) method was applied to establish the diagnosis and prediction model for lung cancer. Then its accuracy was compared with that in the parallel test. The inner test was adopted to randomly select 2/3 of patients as the training set and 1/3 of patients as the testing set for the investigating the prediction effect of the model. **Results** The levels of serum CEA, CA125, CYFRA21-1, and Pro-GRP in the lung cancer group were significantly higher than those in the control group with statistical difference($P < 0.05$). The LDA model could correctly diagnose 81.5% of patients with lung cancer and 83.1% of patients with benign lung disease, reaching the total accuracy of 82.5%. This model could predicted 76.5% of patients with lung cancer and 80.0% of patients with benign diseases, reaching the predictive accuracy rate of 80.8%. The parallel test results showed that the sensitivity, specificity and accuracy of the combined detection of serum CEA, CA125, CYFRA21-1 and Pro-GRP for diagnosing lung cancer were 80.4%, 76.1% and 77.8% respectively. **Conclusion** The combined detection of serum CEA, CA125, CYFRA21-1 and Pro-GRP increases the sensitivity for the diagnosis of lung cancer. The LDA model based on serum CEA, CA125, CYFRA21-1 and Pro-GRP can better diagnose and predict lung cancer and provides a brand-new thinking for the differentiation diagnosis and prediction of lung cancer.

【Key words】 lung cancer; tumor marker; carbohydrate antigen 125; pro-gastrin-releasing peptide; linear discriminant analysis

肺癌是近年来发病率和病死率最高的恶性肿瘤,为肿瘤导致相关死亡的首要病因^[1]。肺癌发病较隐匿,且 5 年生存率多低于 20%,故早期诊断、及时手术对提高肺癌患者的生存率具有重要意义^[2]。肺癌的诊断主要依赖于影像学和组织病理检查,但都存在一定的局限性。肿瘤标志物是一类用于肿瘤早期诊断、药物疗效及预后判断的辅助诊断指标^[3]。临床上常用的

肿瘤标志物,如癌胚抗原(CEA)、糖类抗原(CA)125、细胞角蛋白 19 片段(CYFRA21-1)和胃泌素释放前肽(Pro-GRP)等,尽管特异性较高但灵敏度较低。联合检测肿瘤标志物可提高肺癌诊断的灵敏度,但特异性也会下降。因此,合理利用临床常用的肿瘤标志物建立诊断和预测模型,有望为肺癌的鉴别诊断提供新思路。线性判别分析(LDA)模型是一种广泛用于语音、

人脸识别和多元分类问题的数据处理技术^[4],但较少有 LDA 模型用于肺癌与肺部良性疾病鉴别诊断的相关报道。本研究以肺癌患者为研究对象,探讨基于 CEA、CA125、CYFRA21-1 和 Pro-GRP 的 LDA 模型对肺癌的鉴别诊断和预测价值。现报道如下。

1 资料与方法

1.1 一般资料 肺癌组为本院 2014 年 8 月至 2015 年 9 月首次入院的 92 例肺癌患者,均经病理组织切片或脱落细胞学确诊。肺癌组中男 69 例,女 23 例;平均年龄(65.49±11.83)岁;其中鳞癌 40 例,腺癌 27 例,大细胞癌 2 例,小细胞肺癌 23 例。随机选取同期 142 例肺部良性疾病患者作为对照组,其中男 107 例,女 35 例;平均年龄(64.96±16.38)岁;肺结核 89 例,慢性支气管炎 23 例,肺炎 14 例,肺气肿 16 例。两组患者的性别、年龄等一般资料比较,差异无统计学意义,具有可比性($P>0.05$)。

1.2 方法 采用美国雅培 SRI2000 化学发光仪和配套试剂测定血清中 CEA、CA125 和 CYFRA21-1 水平。采用酶联免疫吸附试验(ELISA)检测血清中 Pro-GRP 水平,配套试剂为瑞典康奈格生物技术公司 Pro-GRP 试剂盒。参考区间:CEA<6.0 ng/mL,CA125<35.0 U/mL,CYFRA21-1<7.0 ng/mL,Pro-GRP<10.0 ng/L。

1.3 LDA 模型的建立 LDA 法是一种对数据进行降维的处理方法,将多组 P 维数据通过空间投影转换到某个方向,寻找最佳投影平面,实现多组间分离^[5]。基于 Fisher 的 LDA 最常

用,通过计算组间的离差平方和组内离差平方和的比值来寻找最容易实现分类的投影方向,既能使同类样本尽量聚集也能实现不同类别样本尽量区分^[6]。基于 Fisher 的 LDA 模型借助一元方差分析的思想建立 Fisher 线性判别函数,求解方程最大的特征根 λ_1 对应的特征向量 l_1 和判别效率 $\Delta(l_1)=\lambda_1$,依次建立一个或多个 Fisher 线性判别函数对个体属性进行判别和预测。本文以病理组织切片或脱落细胞学确诊的肺癌组为分组变量(因变量),赋予逻辑值“1”,患肺部良性疾病对照组赋予逻辑值“0”。以两组血清 CEA、CA125、CYFRA21-1 和 Pro-GRP 的测定值为自变量(X)。为消除不同肿瘤标志物量纲(单位和测量数值)的影响,采用(测定值-均值)/标准差,将数据归一化处理。采用 Fisher 相关系数和组内离差平方和最小的原则建立 Fisher 诊断和预测模型。

1.4 统计学处理 采用 SPSS18.0 和 SIMCA P12.0 软件对数据进行统计分析。正态分布的计量资料以 $\bar{x}\pm s$ 表示;偏态分布的计量资料以中位数及 25 位百分位数、75 位百分位数 $[M(P_{25}\sim P_{75})]$ 表示,组间比较采用独立样本秩和检验;计数资料以百分率表示,组间比较采用 χ^2 检验。以 $P<0.05$ 为差异有统计学意义。

2 结果

2.1 两组血清中各肿瘤标志物水平比较 肺癌组患者血清中 CEA、CA125、CYFRA21-1 和 Pro-GRP 水平均高于对照组,差异均有统计学意义($P<0.05$)。见表 1。

表 1 两组血清中各肿瘤标志物水平比较 $[M(P_{25}\sim P_{75})]$

组别	n	CEA(ng/mL)	CA125(U/mL)	CYFRA21-1(ng/mL)	Pro-GRP(ng/L)
肺癌组	92	74.60(14.70~139.00)*	88.6(51.90~232.00)*	21.80(13.60~86.90)*	19.60(12.70~41.50)*
对照组	142	2.87(1.57~6.44)	14.9(2.64~41.90)	3.22(2.51~7.03)	4.34(2.36~8.91)

注:与对照组比较,* $P<0.05$ 。

2.2 四种标志物对肺癌的诊断价值 单项指标中 CYFRA21-1 诊断肺癌的灵敏度相对较高而 Pro-GRP 最低。四种标志物诊断肺癌的特异度均中等,其中 CYFRA21-1 的特异度最高,CEA 最低。CEA、CA125、CYFRA21-1 和 Pro-GRP 诊断肺癌的准确率均在 70.0% 以上,CYFRA21-1 诊断肺癌的准确率最高,达到 81.2%。四种指标联合检测的灵敏度为 80.4%,特异度为 76.1%,诊断准确率为 77.8%。见表 2。

表 2 四种标志物诊断肺癌的灵敏度、特异度和准确率 $[\%(n/n)]$

指标	灵敏度	特异度	准确率
CEA	59.8(55/92)	76.8(109/142)	70.1(164/234)
CA125	51.1(47/92)	84.5(120/142)	71.4(167/234)
CYFRA21-1	69.6(64/92)	88.7(126/142)	81.2(190/234)
Pro-GRP	47.8(44/92)	85.2(121/142)	70.5(165/234)
CEA+CA125+			
CYFRA21-1+Pro-GRP	80.4(74/92)	76.1(108/142)	77.8(182/234)

2.3 四种肿瘤标志物的 LDA 模型 肺癌组的 Fisher 线性判别方程: $Y_1 = 0.024X_{CEA} + 0.008X_{CA125} + 0.014X_{CYFRA21-1} + 0.001X_{Pro-GRP} - 5.919$ 。对照组的 Fisher 线性判别方程: $Y_2 = 0.007X_{CEA} + 0.006X_{CA125} + 0.005X_{CYFRA21-1} - 0.002X_{Pro-GRP} -$

2.554。把每例患者的标准化肿瘤标志物测定值带入 Y_1 和 Y_2 两个判别方程,得到 Y_1 和 Y_2 值,两者中较大的表示患者纳入该类的可能性最大,反之则归入另一类。92 例肺癌组患者中 75 例被正确诊断,模型的诊断准确率为 81.5%。142 例对照组患者中 118 例被正确诊断,诊断的准确率为 83.1%。该模型总的诊断准确率为 82.5%(193/234)。两类的模型的判别符合率比较,差异无统计学意义($\chi^2 = 0.096, P = 0.86$)。见图 1。

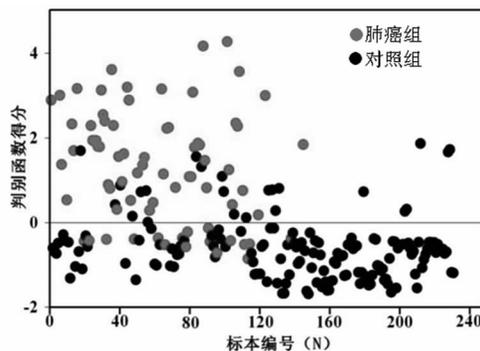


图 1 基于血清中 CEA、CA125、CYFRA21-1 和 Pro-GRP 的 LDA 诊断模型

为考察 LDA 模型的预测能力,采用内代法对模型的有效

性进行验证,随机选取 2/3 的研究对象为训练集,1/3 为测试集,判断模型的预测效果结果表明,测试模型能对 76.5% 的肺癌患者和 80.0% 的肺部良性疾病患者进行预测,预测准确率为 80.8%。见图 2。

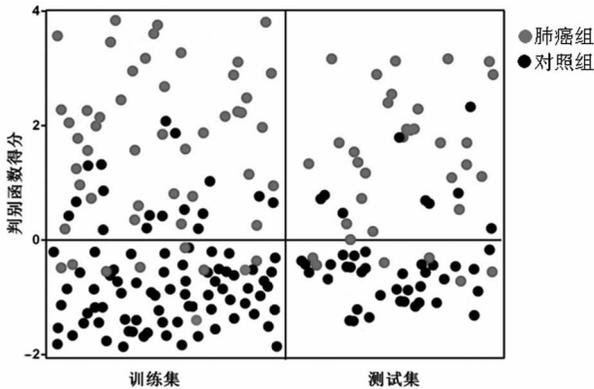


图 2 基于血清中 CEA、CA125、CYFRA21-1 和 Pro-GRP 的 LDA 预测模型

3 讨论

从结果 2.1 可以看出,肺癌患者血清中 CEA、CA125、CYFRA21-1 水平均显著高于肺部良性疾病患者,单项指标的特异性较高但灵敏度较低,这与前期的研究报道一致^[7-8]。四种标志物中,肺癌患者的 Pro-GRP 水平也显著高于对照组。Pro-GRP 是小细胞肺癌的自主生长因子,在健康人和肺部良性疾病患者中血清水平不升高,对小细胞肺癌具有较高的诊断价值^[9]。

目前国内较少有 CEA、CA125、CYFRA21-1 和 Pro-GRP 联合检测对肺癌的诊断价值的相关报道。本研究表明,联合检测血清中 CEA、CA125、CYFRA21-1 和 Pro-GRP 对肺癌的诊断具有较高的灵敏度(80.4%),中等的特异度(76.1%)和准确率(77.8%)。基于血清中四种标志物 CEA、CA125、CYFRA21-1 和 Pro-GRP 的 Fisher-LDA 模型能较好地鉴别诊断肺癌(81.5%)与肺部良性疾病(83.1%),具有较高的预测准确率(80.8%),该模型的诊断准确率优于传统的并联试验(82.5% vs. 77.8%)。肺癌组的 Fisher-LDA 模型中,CEA 的 Fisher 分类系数最大(0.024)而 Pro-GRP 最小(0.001),提示 Pro-GRP 对肺癌诊断模型的作用较小,可能与本研究 75% (69/92) 的肺癌类型为非小细胞肺癌有关。冯阳春等^[10] 利用血清中 Pro-GRP、CEA、CA125、鳞癌抗原(SCC)、CYFRA21-1 和神经元特异性烯醇化酶(NSE)建立 Fisher-LDA 模型并用于小细胞肺癌和非小细胞肺癌的鉴别诊断,该模型中 Pro-GRP 也没有参与非小细胞肺癌模型的建立,这也从侧面证明了 Pro-GRP 是诊断小细胞肺癌较好的指标^[9]。

综上所述,本研究建立了基于血清中 CEA、CA125、CY-

FRA21-1 和 Pro-GRP 的 Fisher-LDA 模型并用于肺癌的鉴别诊断和预测分析,该模型将单一的肿瘤标志物转化成综合的线性判别函数,具有更高的准确性。采用内代法对随机选择的标本属性进行验证取得了较高的预测准确性。目前,肺癌的诊断还缺乏灵敏、特异的标志物,相关的模式识别技术用于其诊断的报道也较少。因此,尝试以 LDA 模型建立肺癌的诊断和预测模型有望挖掘更多有价值的综合信息,实现肺癌的鉴别诊断。

参考文献

- [1] 陈万青,张思维,邹小农. 中国肺癌发病死亡的估计和流行趋势研究[J]. 中国肺癌杂志,2010,13(5):488-493.
- [2] Petrelli NJ, Winer EP, Brahmer J, et al. Clinical cancer Advances 2009: major research advances in cancer treatment, prevention, and screening: a report from the American Society of Clinical Oncology[J]. J Clin Oncol, 2009, 27(35): 6052-6059.
- [3] Schneider J. Tumor markers in detection of lung cancer [J]. Adv Clin Chem, 2006, 42(1): 1-41.
- [4] 张建国,刘新国. 线性判别分析的迭代解法及其应用[J]. 中国海洋大学学报, 2015, 45(11): 119-124.
- [5] 何晓群. 多元统计分析[M]. 3 版. 北京: 中国人民大学出版社, 2012.
- [6] 赵海滨,颜世玉,于清文,等. 采用 Fisher 线性判别分析进行 MEG 信号的分类[J]. 东北大学学报(自然科学版), 2013, 34(12): 1695-1698.
- [7] 朱景倩,朱丹,陈慧,等. 血清肿瘤标志物检测在肺癌诊断中的临床价值[J]. 中国卫生检验杂志, 2010, 20(1): 141-143.
- [8] 冷平,乔凤伶,余蓉,等. 4 种血清肿瘤标志物联合检测在肺癌诊断中的价值[J]. 成都医学院学报, 2015, 10(2): 180-182.
- [9] Shibayama T, Ueoka HK, Kiura K, et al. Complementary roles of pro-gastrin-releasing peptide(Pro-GRP) and neuron specific enolase(NSE) in diagnosis and prognosis of small-cell lung cancer(SCLC)[J]. Lung Cancer, 2001, 32(1): 61-69.
- [10] 冯阳春,徐怡,黄艳春. 聚类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.

(收稿日期:2016-01-05 修回日期:2016-03-10)

医学统计工作的基本内容

按工作性质及其先后顺序,可将医学统计工作分为实验设计、收集资料、整理资料、分析资料。实验设计是开展某项医学研究工作的关键,包括医学专业设计和统计学设计,医学专业设计的内容包括研究对象纳入和排除标准、样本含量、获取样本的方法、分组原则、观察(检测)指标、统计方法等。收集资料的方法包括各种试验、检测或调查,要求资料完整、准确、及时、有足够数量、具有代表性和可比性等。整理资料包括原始资料的检查与核对、对资料进行分组与汇总等。分析资料即对资料进行统计学分析,包括进行统计描述和统计推断。