

• 临床检验研究论著 •

最优生化指标检测模型在 4 种疾病中的应用价值初探

朱喜丹, 田刚, 蔡美珠

(泸州医学院附属医院检验科, 四川 646000)

摘要:目的 建立 4 种疾病最优生化指标的检测模型, 探讨不同模型在疾病诊断中的应用价值。**方法** 以肺癌、乳腺癌、直肠癌和淋巴瘤患者为研究对象(113 例), 采用受试者工作特征(ROC)曲线和人工神经网络(ANN)建立诊断模型。**结果** ANN 模型中 31 项生化检测项目在肺癌、乳腺癌、直肠癌和淋巴瘤中的 ROC 曲线下面积(AUC)分别为 0.777, 0.848, 0.827 和 0.733, 而优化检测项目后 4 种疾病的 AUC 分别为 0.869, 0.949, 0.859, 0.947。**结论** 不同疾病最优生化检测模型的建立提高了疾病诊断的准确, 有助于减轻患者医疗费用。

关键词:肺癌; 乳腺癌; 直肠癌; 淋巴瘤; ROC 曲线**DOI:**10.3969/j.issn.1673-4130.2013.04.022**文献标识码:**A**文章编号:**1673-4130(2013)04-0437-02

Primary exploration the application value of optimal biochemical items detection model in four diseases

Zhu Xidan, Tian Gang, Cai Meizhu

(Department of Clinical Laboratory, Affiliated Hospital of Luzhou Medical College, Luzhou, Sichuan 646000, China)

Abstract: Objective A supervised biochemical detection model for four different diseases has been established to evaluate its possible application value for diagnosis of disease. **Methods** The model of biochemical items of 113 samples of carcinoma of the lungs, breast cancer, carcinoma of the rectum and lymphoma were established by receiver operating characteristic(ROC) curve and artificial neural network(ANN). **Results** The area under the ROC(AUC) of 31 detection items in carcinoma of the lungs, breast cancer, carcinoma of the rectum and lymphoma were 0.777, 0.848, 0.827 and 0.733, respectively. The AUC of optimal model of biochemical items in four diseases were 0.869, 0.949, 0.859 and 0.947, respectively. **Conclusion** The established optimal models of biochemical items in different diseases contribute to the enhancement of veracity and alleviation of hospitalization cost of patients.

Key words:carcinoma of the lungs; breast cancer; carcinoma of the rectum; lymphoma; ROC curve

生化指标是临幊上揭示人体代谢变化的常规检测项目, 在疾病的辅助诊断、病情判断、药物的疗效观察及其毒副作用的评估中提供信息支撑和决策依据。目前, 不同疾病尚无指定的生化检测指标, 人为地选取生化全套, 肝功能或肾功能检测可能会增加不必要的检测项目, 从而加重患者的经济负担。随着个体化治疗的开展^[1], 建立合理的个体化检测方案在辅助诊断疾病的同时将有助于减少患者的医疗费用^[2]。本文以 4 种肿瘤为例, 采用受试者工作特征(ROC)曲线和人工神经网络(ANN)初步探讨不同疾病状态下相关生化指标的重要性, 以期建立不同疾病个体化的生化检测项目。

1 资料与方法

1.1 一般资料 收集 2010 年 8 月 1 日至 2011 年 5 月 31 日泸州医学院附属医院检验科 LIS 系统中肺癌(28 例)、乳腺癌(51 例)、直肠癌(19 例)和淋巴瘤(15 例)患者共 113 例, 其中男 39 例, 女 74 例, 平均年龄为(51.1±12.6)岁。(1)纳入标准: ①临床诊断明确; ②标本信息完整; ③入院第一次检查。(2)排除标准: ①临床诊断不明确或信息不全; ②复检病例。

1.2 统计学处理 ROC 曲线分析和基于人工神经网络(ANN)模型的 ROC 曲线分析均采用 SPSS 17.0 软件完成。

2 结果与讨论

受试者工作特征曲线(ROC)是一种全面、准确评价诊断试验的工具, 能够客观地揭示诊断方法对疾病的诊断能力^[3]。ROC 曲线下面积(AUC)是反映诊断试验准确性的关键指标, 通常认为其取值在 0.7~0.9 时具有中等的诊断准确性, 而大于 0.9 时诊断的准确性较高^[4]。4 种疾病的 ROC 曲线分析见

表 1。从表中可知, 4 种疾病有价值的生化指标不完全相同, 肺癌中肌酐(Crea)、血浆同型半胱氨酸(Hcy)、球蛋白(GLO), 总胆固醇(TC)和低密度脂蛋白胆固醇(LDL-C)的 AUC 均有助于揭示患者的肝肾功能, 其中 Crea 和 Hcy 的 AUC 降低具有统计学意义($P<0.05$), 而 TC, GLO 和 LDL-C 的 AUC 升高则具有统计学意义($P<0.05$)。乳腺癌与肺癌比较, 除 GLO 外, 脱抑素 C(CysC)、碱性磷酸酶(ALP)的 AUC 也具有统计学意义($P<0.05$); 直肠癌除以上 5 个指标外, CysC、ALP 和直接胆红素(DBIL)的 AUC 也具有统计学意义($P<0.05$)。淋巴瘤有价值的生化指标最少, 仅有 Crea、GLO、TC 和 LDL-C($P<0.05$)。4 种疾病中共同的生化指标为 Crea, TC 和 LDL-C, 其中 Crea 在 4 种疾病中 AUC 最小而 LDL-C 最大, 除乳腺癌以外, 其余 3 种指标的 AUC 均在 0.70 以上, 说明该指标能够较好地反映患者的肝肾功能状态。

表 1 4 种疾病的 ROC 曲线分析

项目	AUC			
	肺癌	乳腺癌	直肠癌	淋巴瘤
Crea	0.360 5*	0.235 9*	0.294 0*	0.337 1*
Hcy	0.390 3*	0.244 1*	0.296 7*	0.358 4
CysC	0.410 1	0.334 5*	0.305 0*	0.366 7
ALP	0.410 5	0.389 6*	0.322 1*	0.368 0
DBIL	0.415 6	0.403 7	0.348 1*	0.374 9
Urea	0.427 3	0.409 0	0.410 2	0.397 9

续表1 4种疾病的ROC曲线分析

项目	AUC			
	肺癌	乳腺癌	直肠癌	淋巴瘤
UA	0.4286	0.4280	0.4125	0.4103
A/G	0.4315	0.4317	0.4131	0.4162
TBIL	0.4471	0.4814	0.4397	0.4186
AST	0.4496	0.4851	0.4595	0.4313
TG	0.4500	0.4912	0.4734	0.4536
GGT	0.4651	0.5018	0.4846	0.4794
AST/ALT	0.4698	0.5066	0.4970	0.4811
TBA	0.4698	0.5146	0.4973	0.4928
GLU	0.4726	0.5193	0.4985	0.4935
AFU	0.4734	0.5243	0.5000	0.5182
IBIL	0.4766	0.5259	0.5009	0.5213
ALT	0.4887	0.5270	0.5044	0.5340
ALB	0.5034	0.5273	0.5068	0.5375
HDL-C	0.5083	0.5368	0.5139	0.5412
LP(a)	0.5289	0.5384	0.5242	0.5529
LDH	0.5510	0.5395	0.5260	0.5567
CHE	0.5525	0.5545	0.5328	0.5595
5-NT	0.5553	0.5551	0.5603	0.5691
TP	0.5572	0.5789	0.5674	0.5852
GSP	0.5646	0.5837	0.6108	0.6062
APOA	0.5942	0.5884	0.6350	0.6162
APOB	0.5944	0.6035	0.6362	0.6244
GLO	0.6280*	0.6080	0.6773*	0.6711*
TC	0.7674*	0.6176*	0.6868*	0.6914*
LDL-C	0.7736*	0.6191*	0.7054*	0.7378*

*:与其他项目比较, $P < 0.05$ 。

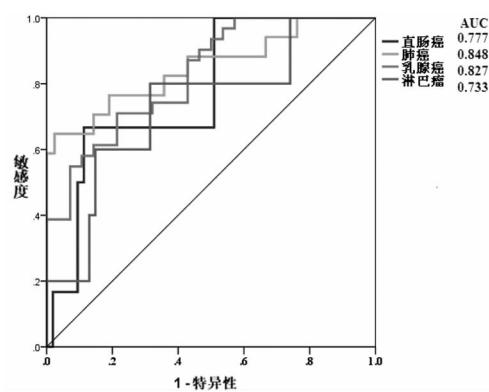


图1 4种疾病的ANN-ROC曲线

人工神经网络模型(ANN),即通过模拟大脑的某些机制对复杂问题进行神经元网络模式化处理,从而实现了复杂问题的简单化^[5]。将表1中31项生化指标进行ANN-ROC分析(图1)。由图1可知,4种疾病综合变量的AUC均大于单一的测定指标,肺癌、乳腺癌、直肠癌和淋巴瘤中的AUC分别为0.777、0.848、0.827和0.733,说明综合的生化指标对4种疾病引起的肝肾功能紊乱具有中等的诊断准确性。运用ANN-ROC模型将113例患者随机分为训练集和测试集,其模型的

平均预测准确性仅有24.0%。

以4种疾病所对应的生化指标($P < 0.05$)建立ANN-ROC曲线模型(图2)。肺癌、乳腺癌、直肠癌和淋巴瘤的AUC分别为0.869、0.949、0.859、0.947,其中乳腺癌和淋巴瘤的AUC均在0.90以上,说明优化后的生化指标模型能够提高疾病的诊断准确性。肺癌、乳腺癌、直肠癌和淋巴瘤所对应的最优生化指标模型的平均预测准确性也明显提高,分别为76.5%、67.9%、35.7%和42.3%。

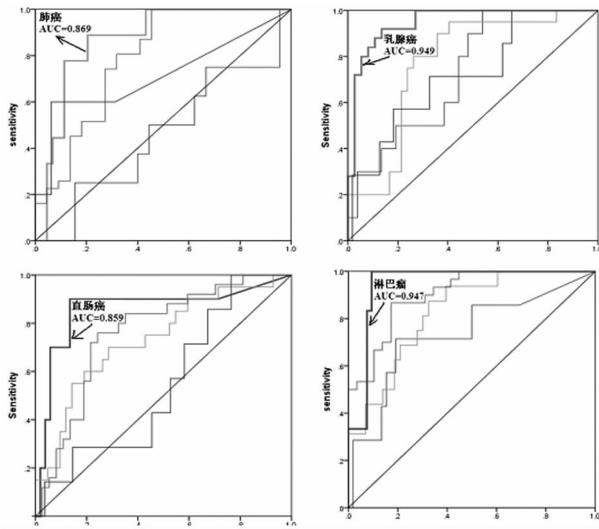


图2 4种疾病筛选指标的ANN-ROC曲线分析

ANN模型是一种较好的非线性数据处理工具^[6],已广泛用于疾病的诊断^[7]和药物分析^[8]等。本文借助ANN-ROC曲线模型,以4种肿瘤为研究对象,用较少的生化指标建立模型获得较高的诊断价值。由于不同疾病的发病机制和治疗的手段的差异,相关的生化指标可能完全不同。以ANN-ROC曲线模型为基础建立大样本的多种疾病的最优生化检测模型将有助于减轻医疗费用,为缓解看病贵的问题提供了一条新思路。

参考文献

- [1] Loscalzo J, Kohane I, Barabasi AL. Human disease classification in the postgenomic era: A complex systems approach to human pathobiology [J]. Mol Syst, 2007, 24(1):1-3.
- [2] 周大勇.阑尾周围肿胀166例个体化治疗分析[J].实用心脑肺血管病杂志,2010,18(12):1869-1870.
- [3] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic curve [J]. Radiology, 1982, 143(1):29-36.
- [4] 张克坚,杨振华.应用ROC曲线图评价检验项目的临床准确性[J].江西医学检验,1999,17(2):66.
- [5] Haykin S. Neural networks: a comprehensive foundation [M]. New York: Macmillan College Publishing Company, 1994.
- [6] Ripley BD. Pattern recognition and neural networks [M]. Cambridge: Cambridge University Press, 1996.
- [7] 姜维,王开正,白克镇等.精神分裂患者血清蛋白标志物的临床应用研究[J].中华神经精神杂志,2008,34(1):27-30.
- [8] Zupan J, Gasteiger J. Neural networks in chemistry and drug design [M]. 2nd ed, Weinheim: Wiley-VCH, 1999.

(收稿日期:2012-09-29)