

ROC 曲线下面积估计的参数法与非参数法的应用研究

宋花玲¹, 贺佳^{2*}, 黄品贤¹, 李素云³

(1. 上海中医药大学预防教研室, 上海 201203; 2. 第二军医大学卫生勤务学系卫生统计学教研室, 上海 200433; 3. 上海中医药大学病理学教研室, 上海 201203)

[摘要] **目的:** 阐明 ROC 曲线下面积估计的参数法和非参数法并进行比较, 为其在诊断试验评价中的应用提供依据。 **方法:** 用双正态模型的参数法和 Mann-Witney 统计量的非参数法估计 ROC 曲线下面积, 并以其在肺癌诊断试验准确度评价中的应用来具体说明。 **结果:** 非参数法估计的肺癌两个标志物 Cyfra21-1 和 CEA 的 ROC 曲线下面积分别为 0.77、0.87, 参数法估计的面积分别为 0.78、0.87, 表明在样本量较大时参数法和非参数法估计的 ROC 曲线下面积近似相等。 **结论:** 样本量较小时可选择非参数法估计 ROC 曲线下面积, 样本量较大时可根据实际情况选择参数法或非参数法。

[关键词] ROC 曲线下面积; 参数法; 非参数法**[中图分类号]** R 195.1 **[文献标识码]** A **[文章编号]** 0258-879X(2006)07-0726-03

Application of parametric method and non-parametric method in estimation of area under ROC curve

SONG Hua-ling¹, HE Jia^{2*}, HUANG Pin-xian¹, LI Su-yun³ (1. Department of Preventive Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China; 2. Department of Health Statistics, Faculty of Health Services, Second Military Medical University, Shanghai 200433; 3. Department of Pathology, Shanghai University of Traditional Chinese Medicine, Shanghai 201203)

[ABSTRACT] **Objective:** To elucidate and compare the parametric method and non-parametric method in estimation of the area under ROC curve, so as to provide a basis for their application in diagnosis assessment. **Methods:** The areas under ROC curves were estimated by parametric method of fitting binomial model and by non-parametric method of Mann-Witney statistics. The method was employed in the diagnostic tests of lung cancer. **Results:** By non-parametric methods, the areas under ROC curves of Cyfra21-1 and CEA were respectively 0.77 and 0.87 in the lung cancer diagnostic tests; by parametric methods, they were 0.78 and 0.87, respectively. It was indicated that when the sample size was large, the values of areas under ROC Curves were similar between parametric method and non-parametric method. **Conclusion:** Non-parametric method should be used to evaluate the area under ROC curve if the sample size is small, and for large sample size, the parametric method or nonparametric method should be chosen according to the actual situation.

[KEY WORDS] area under ROC curve; parametric method; non-parametric method

[Acad J Sec Mil Med Univ, 2006, 27(7): 726-728]

一项新的诊断试验的诊断性能如何, 它能否替代旧的诊断试验, 这在很大程度上依赖于新的诊断试验的准确度大小。诊断试验的准确度评价指标有灵敏度、特异度、一致率、优势比、Youden 指数、ROC 曲线下面积、似然比等等, 而其中的 ROC 曲线下面积指标因其不受患病率和诊断界值的影响, 以及对两个诊断试验的准确度进行综合比较, 因而成为目前公认的最佳评价指标^[1]。

1 方法介绍

ROC 曲线^[2]是以试验结果的每一个值作为可能的诊断界值, 由此计算得到相应的灵敏度和特异度, 以假阳性率即(1-特异度)为横坐标, 以真阳性率即灵敏度为纵坐标绘制而成的曲线即为 ROC 曲

线, 其曲线下面积的大小可作为诊断试验准确度的衡量指标, 其取值范围为 0.5~1。完全无诊断价值的 ROC 曲线(称为参照线或机会线, reference line) 下面积为 0.5, 理想的 ROC 曲线下面积为 1, 而一般认为: 对于一个诊断试验, ROC 曲线下面积在 0.5~0.7 之间时诊断价值较低, 在 0.7~0.9 之间时诊断价值中等, 在 0.9 以上时诊断价值较高^[3]。ROC 曲线作为诊断试验准确度的评价指标可理解为: 所有可能特异度值的平均灵敏度; 所有可能灵敏度值的平均特异度; 随机选择的病例试验结果比随机选择的对照试验结果更有可能怀疑有病的概率。

[作者简介] 宋花玲, 讲师, 硕士。

* Corresponding author. E-mail: hejia@smmu.com.cn

ROC 曲线下面积的估计有参数法和非参数法两种,均适用于结果为连续性资料或等级资料的诊断试验准确度的评价。

非参数法^[4]是根据实验结果直接计算绘制 ROC 曲线所需的工作点(假阳性率,真阳性率)即(1-特异度,灵敏度),由此绘制的 ROC 曲线称为经验 ROC 曲线,其曲线下面积可由梯形规则计算得到, Bamber 于 1975 年发现:经验 ROC 曲线下面积等价于患者和非患者实验结果秩和检验的 Mann-Whitney 检验统计量,因而可由 Mann-Whitney 统计量估计曲线下面积的大小,但在诊断试验检测结果相同值较多(例如结果为有序资料的影像学检查)及时样本量较小时,其结果常小于真实的面积值^[2]。然而非参数法因其没有限制条件,所以适用于任何诊断试验 ROC 曲线下面积的估计。非参数法估计 ROC 曲线下面积可通过 SPSS、SAS 等统计软件实现。

曲线下面积估计的参数法是假设患者和非患者的实验结果均符合正态分布,根据试验结果拟合双正态模型,由模型拟合的 ROC 曲线称为拟合 ROC 曲线或光滑 ROC 曲线,该曲线可用 2 个参数表示,由 2 个参数可得到绘制光滑 ROC 曲线所需的工作点及曲线下面积的估计值,而且得到的面积估计值一般是无偏的。双正态模型的两个参数可由最大似然估计法得到^[5]。参数法的应用条件为:患者与非患者的试验结果服从双正态分布,但这是指 ROC 曲线的函数形式,而不是指试验结果的基本分布,因为变量变换几乎可使任何试验结果转换为双正态分布,而且在样本量较大时参数法与非参数法估计的 ROC 曲线下面积常常近似相等。参数法估计 ROC 曲线下面积的公式为:

$$A = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right)$$

其中 A 为 ROC 曲线下的面积, a 、 b 分别为双正态模型的 2 个参数, Φ 表示标准正态分布函数。参数法估计 ROC 曲线下面积可通过软件 ROCKIT0.9 β 来实现。参数法估计 ROC 曲线下面积的缺点是当样本量较小时,或试验检测值及其变量变换值均不符合参数法的应用条件时,结果可能会严重偏离真实值,此时应增大样本量或改用非参数法估计曲线下面积。

2 实例应用

以杭州、武汉的某两个医院为研究单位,随机从门诊和住院人群及体检人群中选择 334 例研究对象(男性 215 人,女性 119 人),包括正常人、良性疾病患者和肺癌患者,其中的良性疾病包括肺炎及其他良性疾病。正常人为 95 人,平均年龄(49.69 \pm 15.91)岁;其他研究对象由临床病理检查结果分为肺癌患者和良性疾病患者,其中肺癌患者组 163 人,平均年龄(60.31 \pm 12.65)岁,良性疾病组 76 人,平均年龄(52.87 \pm 16.57)岁。由经过培训的检验医师用上海某公司新开发的肺癌诊断试剂,在严格遵循操作规程的条件下对各研究对象血清中的肿瘤标志物 Cyfra21-1 和 CEA 进行定量检测。用参数法和非参数法估计各标志物的 ROC 曲线下面积,从而评价该诊断试剂的准确度,并对两种方法进行比较。参数法和非参数法的分析结果分别由软件 ROCKIT0.9 β 、SPSS 来实现。

结果见图 1 和表 1、2。

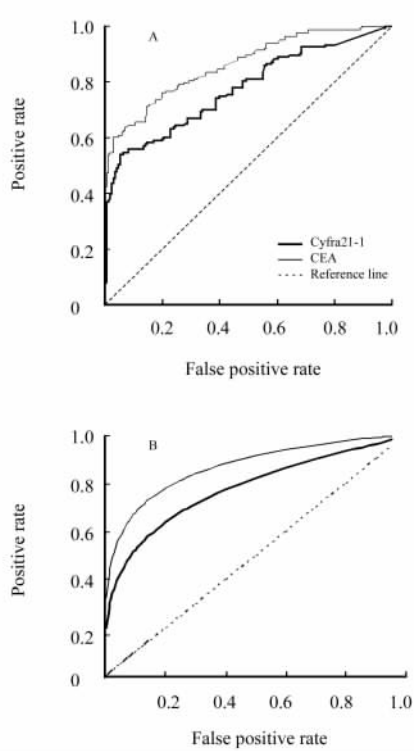


图 1 肺癌肿瘤标志物 Cyfra21-1 和 CEA 的 ROC 曲线

Fig 1 ROC curves of tumor markers of lung cancer: Cyfra21-1 and CEA

A: Experiential ROC curve; B: Fitted ROC curve

表 1 非参数法估计肺癌标志物 ROC 曲线下面积的结果

Tab 1 Areas under ROC curves of lung cancer markers estimated by non-parametric methods

Marker	Areas under ROC curves	Standard error	95%CI	Z	P
Cyfra21-1	0.77	0.03	0.72-0.83	10.69	<0.000 1
CEA	0.87	0.02	0.82-0.90	18.66	<0.000 1

Lung cancer sample size:163; non-lung cancer sample size:171

表 2 参数法估计肺癌标志物 ROC 曲线下面积的结果

Tab 2 Areas under ROC curves of lung cancer markers estimated by parametric methods

Markers	a	b	Areas under ROC curves	Standard error	95%CI	Z	P
Cyfra21-1	0.93	0.68	0.78	0.03	0.73-0.83	10.91	<0.000 1
CEA	1.36	0.71	0.87	0.02	0.82-0.90	18.48	<0.000 1

Lung cancer sample size:163; non-lung cancer sample size:171. a and b are 2 parameters of fitted ROC curve; Z is statistic when the area under ROC curves is compared with 0.5; and P is corresponding normal probability value

由图 1 可知,两标志物的 ROC 曲线均位于机会线以上,表明 2 个标志物在肺癌的诊断中均有一定的诊断价值,而且 CEA 的诊断价值要高于 Cyfra21-1。由表 1 和表 2 可知,参数法估计 2 个标志物的 ROC 曲线下面积分别为 0.78 和 0.87,非参数法估计 2 个标志物的 ROC 曲线下面积分别为 0.77 和 0.87,与 0.5 相比均有统计学意义,说明 2 个标志物对肺癌的诊断均有中等的诊断价值;其次,本研究中的肺癌患者和非患者的样本量分别为 163、171,标志物 Cyfra21-1 的 ROC 曲线下面积用参数法和非参数法估计的结果分别为 0.78 和 0.77,CEA 的 ROC 曲线下面积两种方法估计的结果均为 0.87,表明在样本量较大时参数法和非参数法估计的 ROC 曲线下面积非常接近甚至相等。

3 讨论

ROC 曲线作为诊断试验准确度的评价指标,因其诸多优点而有着广泛应用的前景。本文通过对其曲线下面积估计方法的阐述,说明了参数法和非参数法的应用及异同。ROC 曲线下面积估计的非参数法因其没有条件限制,所以可以应用于所有诊断试验的准确度评价,但计算的 ROC 曲线下面积往往

偏小;而拟合双正态模型的参数法可较准确地估计 ROC 曲线下面积,但在样本量较小时,估计的 ROC 曲线下面积可能会出现位于机会线以下的情况,或者当资料远远偏离其应用条件时,估计的结果可能会严重偏离其真实值;在样本量较大且相同值较少时,参数法和非参数法估计的 ROC 曲线下面积常常近似相等(如本文中的实例应用)。在实践中,样本量大时可根据实际情况选择参数法或非参数法估计 ROC 曲线下面积,样本量较小时可选择非参数法。

[参考文献]

- [1] 傅 华 主编. 预防医学[M]. 4 版. 北京:人民卫生出版社, 2004:318-323.
- [2] Zhou X, Obuchowski NA, Mcclish DK. Statistical methods in diagnostic medicine[M]. New York: Wiley, 2002:111-136, 180-187.
- [3] 宇传华. 诊断试验评价[M]//余松林主编. 医学统计学. 北京:人民卫生出版社,2002:164-178.
- [4] Kaufmann J, Werner C, Brunner E. Nonparametric methods for analyzing the accuracy of diagnostic tests with multiple readers[J]. Statist Methods Medical Res, 2005,14: 129-146.
- [5] 邹莉玲,沈其军,陈 峰,等. ROC 曲线下面积的 ML 估计与假设检验[J]. 中国公共卫生,2003 19:127-128.

[收稿日期] 2006-02-17

[修回日期] 2006-06-14

[本文编辑] 尹 茶