

探讨 Bayes 判别分析在食管癌淋巴结转移 CT 诊断中的价值

邵华飞, 刘剑芳, 姚丽波, 刘戡, 屈东, 王铸

【摘要】 目的:利用 Bayes 判别分析初步建立诊断食管癌淋巴结转移的各种 CT 征象的联合诊断模型。**方法:**共搜集胸段食管癌 208 例,将其分为实验组即训练样本(166 例,80%)及检验组即验证样本(42 例,20%),实验组用以建立诊断方程,检验组用以验证方程。**结果:**经计算获得的方程为: $Y_0 = -7.499X_1 + 7.957X_2 + 167.761X_3 - 0.087X_4 + 0.459X_5 - 5.528X_6 - 0.711X_7 - 66.080$, $Y_1 = -6.697X_1 + 8.231X_2 + 181.686X_3 - 0.106X_4 + 2.219X_5 - 3.331X_6 - 0.124X_7 - 83.183$,其中 Y_0 为非转移组, Y_1 为转移组, X_1 为淋巴结最大短径, X_2 为淋巴结最大长径, X_3 为横纵比, X_4 为最大截面积, X_5 为是否边缘模糊, X_6 为是否中央低密度, X_7 为是否成簇分布。利用自身检验法所得诊断模型的符合率为 87.7%,误判率为 12.3%,交叉检验法与自身检验法所得结果相近。当利用验证样本数据代入方程,所得模型诊断符合率为 84.7%,误判率为 15.3%。**结论:**通过 Bayes 判别分析法所建立的不同 CT 征象对食管癌淋巴结转移的联合诊断模型具有一定的诊断价值,但诊断模型还有待进一步完善。

【关键词】 体层摄影术, X 线计算机; Bayes 判别分析; 食管癌; 淋巴结转移

【中图分类号】 R814.42; R735.1 **【文献标识码】** A **【文章编号】** 1000-0313(2015)08-0835-03

DOI:10.13609/j.cnki.1000-0313.2015.08.008

The value of Bayes discriminatory analysis in the diagnosis of lymph node metastasis of esophageal cancer SHAO Hua-fei, LIU Jian-fang, YAO Li-bo, et al. Department of Radiology, Cancer Hospital, Chinese Academy of Medical Sciences, Beijing 100021, P. R. China

【Abstract】 Objective: To construct mechanism model for the diagnosis of LNs (Lymph nodes metastasis) of esophageal cancer by Bayes discriminatory analysis. **Methods:** 208 thoracic esophageal cancer cases were divided into two groups, including training set (166 cases) and test set (42 cases). Training set was used to construct mechanism model, test set was used to analyze its value of diagnosis. **Results:** Non-metastatic LNs: $Y_0 = -7.499X_1 + 7.957X_2 + 167.761X_3 - 0.087X_4 + 0.459X_5 - 5.528X_6 - 0.711X_7 - 66.080$; metastatic LNs: $Y_1 = -6.697X_1 + 8.231X_2 + 181.686X_3 - 0.106X_4 + 2.219X_5 - 3.331X_6 - 0.124X_7 - 83.183$. X_1 = short-axis diameter, X_2 = long-axis diameter, X_3 = aspect ratio, X_4 = cross-sectional area, X_5 = fuzzy rim, X_6 = central low density area, X_7 = clustered distribution. The accuracy was 87.7%, the P (the probability of false prediction) was 12.3%, with similar result by cross validation. For the test set, the accuracy of Bayesian analysis was 84.7%, the P was 15.3%. **Conclusion:** It is feasible to use Bayes discriminatory analysis for diagnosis of LNs metastasis of esophageal cancer.

【Key words】 Tomography, X-ray computed; Bayes discriminatory; Esophageal cancer; Lymph node metastasis

食管癌是我国高发的恶性肿瘤之一,其主要症状为吞咽不顺、疼痛等,但早期阶段症状往往不明显,患者就诊时常已处于中晚期,因此预后较差,5 年生存率仅 15%~25%^[1]。淋巴转移是食管癌最常见的转移方式,也是关系到食管癌患者预后的主要影响因素,因此正确判断淋巴结转移是肿瘤分期的重要组成部分,也是选择最优治疗方案的重要前提。

判别分析是一种分类诊断统计方法,是根据已掌握的一批明确分类的样本建立判别函数,再将待判样本的数据代入判别函数中,判别其类型,从而做出诊断。贝叶斯(Bayes)判别准则是根据先验概率并综合

样本所具备的信息,寻求一种判别规则使得属于第 K 类的样品,在第 K 类中取得最大的后验概率,从而判断样品的分类。这种方法在自然科学及国民经济等许多领域应用甚广,近年来也被一些学者应用于疾病诊断,尤其是中医辨证分析中,但这种分类诊断方法在影像诊断中应用很少。

材料与方法

将本院 2011 年 7 月—2013 年 6 月期间收治的 208 例有完整临床资料、并经根治性切除手术病理证实的胸段食管癌患者纳入研究。所有患者术前 1 周均行螺旋 CT 增强扫描,患者取仰卧位,扫描范围自下颈部或锁骨上水平至上腹部。扫描参数:120 kV,80~150 mAs,64i×0.625 mm,螺距 1.375,层厚 5 mm,重建层厚 1.0 mm,层间距 0.8 mm;经肘静脉快速注射非

作者单位:100021 北京,北京协和医学院中国医学科学院肿瘤医院影像诊断科
作者简介:邵华飞(1987—),女,山东青岛人,硕士,住院医师,主要从事食管癌的综合影像诊断工作。
通讯作者:王铸, E-mail: wangzhu@cicams.ac.cn
基金项目:国家自然科学基金资助项目(61372192)

离子型对比剂 60~100 mL,注射流率 2.0~3.0 mL,注药后延迟 30~35 s 开始扫描。

将扫描的原始数据进行图像后处理,采用多平面重组 (multi-planar reformation, MPR),随后由两位影像科医师通过 Carestream PACS Diagnostic Profile 5.3 观察分析图像,测量并记录淋巴结的长径、短径、强化程度(淋巴结与同层肌肉 CT 值之比),观察淋巴结有无中央低密度、边缘模糊、是否成簇出现等 CT 征象,由一位高年资医师进行质控。

在 208 例患者资料中随机抽选 166 例(80%)作为实验组即训练样本,用以建立方程,其余 42 例(20%)作为检验组即验证样本,用以检验方程。

使用 SPSS 17.0 统计分析软件进行 Bayes 逐步判别分析,建议诊断模型,并用自身检验法、交叉验证法以及 42 例患者资料对其进行考核。

结 果

1. 观察指标的测量结果

本组 208 例食管癌患者中共有 1524 枚淋巴结,其中转移淋巴结 270 枚。实验组中共 1217 枚淋巴结,转移淋巴结 214 枚;检验组中 307 枚,转移淋巴结 56 枚,患者的其它观察指标见表 1。

表 1 患者资料描述

指标	临床特征
测量值	
年龄(中位数)	36~76 岁(59 岁)
性别(男/女)	169/39
鳞癌:非鳞癌	197:11(其中有双原发 3 例,均为鳞癌)
淋巴结参数	
短径(mm)	
实验组	6.2(3.5~39.2)
检验组	6.1(3.5~17.5)
长径(mm)	
实验组	10.4(3.5~45.0)
检验组	10.7(3.9~24.5)
横纵比	
实验组	0.6(0.2~1.0)
检验组	0.6(0.2~1.0)
最大截面积(mm ²)	
实验组	69.7(12.3~1764.0)
检验组	70.2(17.5~334.3)
强化程度	
实验组	1.3(-0.2~45.7)
检验组	1.3(0.1~15.9)
是否边缘模糊(枚)	
实验组(是/否)	51/1166
检验组(是/否)	23/284
有无中央低密度(枚)	
实验组(是/否)	71/1146
检验组(是/否)	6/301
是否成簇分布(枚)	
实验组(是/否)	203/1014
检验组(是/否)	33/274

2. 筛选并建立判别函数

将以上各项数据进行逐步判别分析(分类变量赋值化:0 为特征阴性,1 为特征阳性),筛选出淋巴结最大短径、最大长径、横纵比、最大截面积、有无中央低密度、有无边缘模糊、是否成簇分布这 7 项变量纳入方程,先验概率因未获得相关历史资料而取等概率^[2-3]。计算后获得诊断方程的函数系数,并以此建立 Bayes 判别函数:非转移组为 $Y_0 = -7.499X_1 + 7.957X_2 + 167.761X_3 - 0.087X_4 + 0.459X_5 - 5.528X_6 - 0.711X_7 - 66.080$;转移组为 $Y_1 = -6.697X_1 + 8.231X_2 + 181.686X_3 - 0.106X_4 + 2.219X_5 - 3.331X_6 - 0.124X_7 - 83.183$ 。其中, X_1 为淋巴结最大短径, X_2 为淋巴结最大长径, X_3 为横纵比, X_4 为最大截面积, X_5 为是否边缘模糊, X_6 为是否中央低密度, X_7 为是否成簇分布。

3. 判别函数的结果及检验

采用自身检验法及交叉验证法来检验判别函数模型的诊断能力,结果见表 2。

表 2 Bayes 诊断模型检验结果 (枚)

检验方法	预测组非转移	预测组转移	合计
自身检验			
非转移	898(89.5%)	105(10.5%)	1003
转移	45(21.0%)	169(79.0%)	214
交叉验证			
非转移	898(89.5%)	105(10.5%)	1003
转移	46(21.5%)	168(78.5%)	214

再将检验组 42 例(20%)患者共 307 枚淋巴结的数据代入诊断模型以验证模型的诊断能力,结果见表 3。

表 3 检验组数据对诊断模型的验证结果 (枚)

病理结果	预测非转移	预测转移	合计
无转移淋巴结	223	28	251
转移淋巴结	19	37	56
合计	242	65	307

讨 论

传统的影像诊断主要凭借个人知识经验加以综合分析推理,因此准确性很大程度上依赖于诊断医师的知识结构和医学经验,即使面对相同的信息,不同的医生也很有可能得到不同的诊断。而 Bayes 判别分析的优势恰恰弥补了传统影像诊断的不足,将医师的经验半定量化,将检查报告中给出的"可能性大"量化为发生的概率,从而降低医生在诊断过程中的主观性,使诊断结果更加客观准确,文献报道利用 Bayes 判别分析建立的诊断模型有助于提高不同医师乃至不同医疗机构间的诊断一致性^[4],陈伟等^[5]的研究结果亦证实,利用这一方法建立的诊断模型与高年资医师的诊断符合率相当,并明显高于经验欠缺的低年资医师。

本研究通过对以往影像学资料的归纳处理,并综

合多位学者提出的诊断食管癌淋巴结转移的 CT 征象,即淋巴结最大短径、最大长径、横纵比、最大截面积、强化程度、有无中央低密度、有无边缘模糊、是否成簇分布,利用 Bayes 逐步判别分析筛选变量:除强化程度外,上述 7 项变量均可选入用以建立诊断方程。利用 Bayes 判别分析法建立诊断模型,非转移组为 $Y_0 = -7.499X_1 + 7.957X_2 + 167.761X_3 - 0.087X_4 + 0.459X_5 - 5.528X_6 - 0.711X_7 - 66.080$;转移组为 $Y_1 = -6.697X_1 + 8.231X_2 + 181.686X_3 - 0.106X_4 + 2.219X_5 - 3.331X_6 - 0.124X_7 - 83.183$ 。其中, X_1 为淋巴结最大短径, X_2 为淋巴结最大长径, X_3 为横纵比, X_4 为最大截面积, X_5 为是否边缘模糊, X_6 为是否中央低密度, X_7 为是否成簇分布。对上述检验模型进行验证,结果显示全部 1217 枚淋巴结,对 1003 枚非转移淋巴结共判对 898 枚,正确率为 89.5%(即特异度);214 枚转移淋巴结中,判对 169 枚,正确率为 79.0%(即敏感度),诊断模型的诊断符合率为 87.7%,共误判 150 枚,误判率为 12.3%。交叉检验法与自身检验法所得结果相近。

由于自身检验法及交叉验证法常常低估误判率,从而夸大判别效果,因此我们采用验证样本对诊断模型作前瞻性误判概率的估计,这种方法所得的误判率比较客观。非转移淋巴结组共 251 枚淋巴结,判对 223 枚,正确率为 88.8%(即特异度);转移淋巴结组共 56 枚,判对 37 枚,正确率为 66.1%(即敏感度);模型诊断符合率为 84.7%,共误判 47 枚,误判率为 15.3%,与上述两种检验法所得结果相近,且未超过 20%,说明该诊断模型具有临床应用价值^[4]。

本研究的不足之处是采用验证样本所得出的诊断模型的诊断敏感度偏低,分析可能的原因:①验证样本中的转移淋巴结数量较少造成误差较大;②本组病例均为可进行食管癌根治手术的患者,肿瘤处于早中期,

可能造成一定的数据偏倚;③诊断模型中分类变量所占比例较大,而分类变量不同于连续变量,且一定程度上受到判断者的主观因素影响,可能也是造成数据偏倚的原因之一;④分类变量所对应的 CT 征象尽管对食管癌淋巴结转移的诊断有一定价值,并证明具有统计学意义,但在样本中出现概率较小;⑤部分征象,如在癌灶旁的淋巴结常因与病灶重叠分界不清而出现“边缘模糊”征象,再如隆突下淋巴结常常呈现为“成簇分布”,而食管癌患者在这一区域的淋巴结却很少被证实为转移,这些都可能導致误判。

通过 Bayes 判别分析法所建立的不同 CT 征象对食管癌淋巴结转移的联合诊断模型具有一定的诊断价值(可以量化诊断,减少人为及经验误差)。但本文仅仅是对此的初步探索,由于本组病例包含的转移与非转移淋巴结分类欠均衡,因此所得的诊断模型可能尚不具备广泛的代表性,若将其用于其他样本,诊断准确性可能存在一些偏差。目前理论上认为病例数越多,分析指标越细,判别函数的准确性越好,所以随着病例数的增加以及对各种影像学征象的进一步观测、分析,判别诊断模型将会得以进一步的完善。

参考文献:

- [1] Pennathur A, Gibson MK, Jobe BA, et al. Oesophageal carcinoma [J]. Lancet, 2013, 381(9864): 400-412.
- [2] 孙振球, 徐勇勇. 医学统计学[M]. 北京: 人民卫生出版社, 2008: 384-398.
- [3] 黄泽丰, 谷凌雁. 基于 Bayes 公式的医生辅助诊断系统研究[J]. 中国数字医学, 2009, 4(5): 43-45.
- [4] 吴明祥, 马捷, 凌人男, 等. 基于判别分析模型的乳腺 BI-RADS 计算机辅助分类[J]. 放射学实践, 2013, 28(3): 346-351.
- [5] 陈伟, 刘进康, 李文政, 等. Bayes 分析在孤立性肺结节 CT 诊断中的应用[J]. 实用放射学杂志, 2009, 25(5): 646-743.

(收稿日期: 2015-02-22 修回日期: 2015-05-11)

下期要目

低剂量脑 CT 灌注的迭代重建研究

自发性腹腔干夹层的 MSCT 血管成像表现

CTA 对于诊断外伤性四肢动脉损伤的 Meta 分析

DSA 仿真内窥镜在脑动脉瘤介入治疗中的应用

ADC 及 rADC 值对缺血性脑梗死缺血半暗带的诊断价值

3D-ASL 与 DSC 对星形细胞瘤血流灌注评估的相关性研究

冠状动脉痉挛患者冠状动脉 CT 表现特点及临床随访分析

3D-UTE 评价腰椎间盘软骨终板缺损与椎间盘退变的相关性