论著

文章编号:1000-5404(2008)17-1638-03

兔血液凝血功能诊断的逐步判别与分析

罗万春[1,3],刘 宿[2],易 东[3] (第三军医大学:[1]基础医学部数学教研室,重庆 400038,[2]大坪医院野战外科研究所麻醉科,重庆 400042,[3]军事预防医学院卫生统计学教研室,重庆 400038)

摘 要:目的 测定兔胸部不同程度撞击伤后凝血指标的变化,并对传统的活化部分凝血活酶时间(APTT)、凝血酶原时间(PT)、凝血酶时间(TT)和纤维蛋白原(FIB)含量等指标进行筛选。方法 用相关分析做数据预处理剔除指标,再用逐步判别分析(stepwise discriminant analysis, SDA)选入指标,并用多元线性判别分析诊断,最后随机抽样比较剔除指标前后的诊断效果。结果 用逐步判别分析法选入 4 项指标,判别正确率达 87.60%;剔除指标 FIB 后,判别正确率达 96.12%;1 000 次抽样诊断检验,4 项指标和 3 项指标的平均诊断正确率分别为 83.07% 和 85.16%。结论 逐步判别分析可有效判别兔血液出、凝血功能状况,剔除指标 FIB 后诊断正确率略有提高。

关键词:血液;凝血功能;逐步判别分析;诊断
中图法分类号:R311;R331.1;R446.11 文献标识码:A

Application of stepwise discriminatory analysis to blood coagulation function diagnosis

LUO Wan-chun[1,3], LIU Su[2], YI Dong[3] ([1]Department of Mathematics, College of Medicine, [3]Department of Medical Statistics, College of Preventive Medicine, Chongqing 400038, [2]Department of Anesthesiology, Institute of Surgery Research, Daping Hospital, Third Military Medical University, Chongqing 400042, China)

Abstract: Objective To determine the changes of coagulation indexes of rabbits struck at different levels and sieve the traditional coagulation indexes, such as activated partial thromboplastin time (APTT), prothrombin time (PT), thrombin time (TT) and fibrinogen (FIB). Methods We used correlation analysis to precondition the data to reject indexes, and used stepwise discriminatory analysis to choose indexes and made the diagnosis by multiple linear discrimination analysis. Then, by random sampling we compared the correct diagnostic rate by 4 indexes with that by 3 indexes. Results The correct discrimination rate of blood coagulation function diagnosis in rabbits was up to 87.60% by the four indexes and 96.12% with the three indexes (APTT, PT and TT). The average correct discrimination rate of blood coagulation function diagnosis in 1 000-time random sampling was up to 83.07% by the four indexes and 85.16% by the three indexes. Conclusion Stepwise discriminatory analysis can be used to discriminate hemorrhage from thromboxane. The correct diagnostic rate gets higher when FIB was rejected.

Key words: blood; coagulation function; stepwise discrimination analysis; diagnosis

出、凝血功能异常包括高凝和低凝两类,可以是因血小板、凝血因子或纤溶功能异常等原因引起。活化部分凝血活酶时间(activated partial thromboplastin time, APTT)、凝血酶原时间(prothrombin time, PT)、凝血酶时间(thrombin time, TT)和纤维蛋白原(fibrinogen, FIB)含量等指标的测定可以初步判断出凝血功能的异常,但出、凝血功能异常时各项指标并非完全按照某种规律变化[1],临床上根据某项指标的正常范围来判断未必一定准确。因此,以医师的临床经验根据某项指标的变化有时难以准确判定其凝血功能是高凝、正常还是低凝。本研究拟从数学模型的角度研究诊断方法,并剔除某些凝血指标,期望能更科学、准确地进行临床诊断。

判别分析(discriminant analysis, DA)是根据判别对象若干个指标的观测结果判定其应属于哪一类的统计学方法。临床医师主要根据经验对患者的主诉、体征、检查结果等做出诊断,其准确性往往不能保证。判别分析往往能根据几项判别指标做出准确诊断,在医学研究中的应用日益广泛[2-4]。对于多项指标的情

作者简介:罗万春,男,四川省隆昌县人,博士研究生,讲师,主要从事预测和优化建模方面的研究。电话:(023)68752211, E-mail:aluo@tmmu.edu.cn
通信作者:易 东,电话:(023)68752344
收稿日期:2007-12-13;修回日期:2008-03-24

形,需要应用逐步判别分析或逐步回归分析剔除一些对诊断影响不显著的指标,这样不但可以节约检查成本,还可以降低误判率。刘升学等[5]利用逐步判别分析在脑卒中分类诊断中有比较好的判别效果。

1　材料与方法

1.1　材料来源

兔 136 只,用水平式生物撞击机造成兔胸部创伤,再分别对其进行微量泵注内毒素,测得凝血 4 项指标 APTT、PT、TT、FIB258 组,其中凝血功能正常(Ⅰ类)的赋值为 1,共 122 组;高凝状态(Ⅱ类)为 2,共 10 组;低凝状态(Ⅲ类)为 3,共 126 组。在实验获得的原始数据中,抽取Ⅰ类的前 61 组、Ⅱ类的前 5 组、Ⅲ类的前 63 组共 129 组组成训练样本集(即为寻找判别函数的原始数据集),其余 129 组组成检测样本集,以检测所得到的模型的诊断效果。

判别指标分别为 x_1(APTT)、x_2(PT)、x_3(TT)、x_4(FIB)。分类变量 $g=1$(正常),$g=2$(高凝),$g=3$(低凝)。

1.2　研究方法

1.2.1　逐步判别分析的原理　　逐步选取判别指标进入判别函数,每一步选 1 个判别效果最显著的自变量进入判别函数,在每次选入变量之前都要对先前选入的变量进行显著性检验,对于在判别函数中判别效果不显著的变量将予以剔除,直到选完所有变量。

1.2.2　选择统计量　　逐步判别分析选用 Wilks 统计量 Λ 来筛选指标,判别指标的选入或剔除都会导致 Λ 的减小或增大,从而对指标进行筛选。统计 Λ 定义为:

$$\Lambda_n=\frac{|W_n|}{|T_n|} \qquad (1)$$

其中,n 是指判别指标 $x_1,x_2,\cdots,x_n$ 的个数,$|W_n|$是类内离差矩阵的行列式,$|T_n|$是总离差矩阵的行列式。

Λ 与 F 的关系为:

$$F=\frac{1-\Lambda}{\Lambda}\cdot\frac{N-G-r}{G-1}\sim F(G-1,N-G-r) \qquad (2)$$

其中 r 为入选变量个数,G 为类数。

1.2.3　逐步判别分析步骤　　第 1 步:选入变量。由公式(1)计算每个指标的 Λ_i,$i=1,2,\cdots,n$。令 $\Lambda_0=\min\limits_{1\leqslant i\leqslant n}\{\Lambda_i\}$,由公式(2)计算出相应的 F 值。若 $F>F_a$,则选入 Λ_0 所对应的指标,否则无指标选入。以此类推。

第 2 步:剔除变量。假定已选入 $r+1$ 个变量,计算该 $r+1$ 个变量相应的 Λ_{r+1},对其中任一变量,按(2)计算相应的 F 值,若 $F\leqslant F_a$,则剔除该变量,否则进行第 1 步。

第 3 步:重复进行第 1 步和第 2 步,直到入选变量不能被剔除,候选变量不能被选入为止。

第 4 步:建立判别函数,对于选入的 m 个变量,第 i 类的判别函数由下式确定:

$$f_i(X)=a_{i0}+a_{i1}x_1+a_{i2}x_2+\cdots a_{im}x_m\ (i=1,2,\cdots,G) \qquad (3)$$

第 5 步:判别检验,将需要判别的样本 $X=(x_{j1},x_{j2},\cdots,x_{jm})^T$ 代入公式(3),若 $\min\limits_{1\leqslant i\leqslant G}\{f_i(X)\}=f_k(X)$,则将 X 判为第 k 类。

1.3　数据的预处理

以 258 组数据对 4 个变量 x_1、x_2、x_3、x_4 之间作相关分析,得到 x_1 与 x_2、x_1 与 x_3、x_2 与 x_3、x_4 与 x_1、x_4 与 x_2、x_4 与 x_3 之间的相关系数分别为 0.49、0.20、0.51、−0.31、−0.32、−0.33,显著性检验均有 $P<0.01$,x_4 与其余 3 项指标呈显著负相关关系,因此,有必要分析剔除 x_4 后的判别诊断效果。

1.4　统计学分析

运用 SPSS 10.0 统计软件做逐步判别分析及逐步回归分析,用数学软件 Matlab 6.5 编程对检测样本做相关分析及诊断正确率检验并绘图。

2　结果

2.1　逐步判别分析结果

对训练样本集进行逐步判别分析,取 $a=0.05$,选取变量结果。第 1 步:选入 x_2,$\Lambda=0.5358$,$F=54.666$;第 2 步:选入 x_1,$\Lambda=0.319$,$F=48.196$;第 3 步:选入 x_3,$\Lambda=0.242$,$F=42.661$;第 4 步:选入 x_4,$\Lambda=0.222$,$F=34.564$。

2.2　判别函数

先验概率取等概率,得到 4 项指标的判别函数为:

$$\begin{aligned}f_1(X)&=4.144x_1+1.77x_2+2.713x_3+8.938x_4-74.935\\ f_2(X)&=3.631x_1+1.549x_2+2.746x_3+11.002x_4-74.356\\ f_3(X)&=5.315x_1+2.656x_2+3.27x_3+7.736x_4-111.125\end{aligned} \qquad (4)$$

根据 1.3 的结果,先验概率取等概率,用前 3 项指标建立判别函数为:

$$\begin{aligned}g_1(X)&=4.007x_1+1.554x_2+2.447x_3-58.795\\ g_2(X)&=3.462x_1+1.284x_2+2.419x_3-49.902\\ g_3(X)&=5.197x_1+2.47x_2+3.04x_3-99.035\end{aligned} \qquad (5)$$

2.3　判别检验

将检测样本集的样本 129 例代入公式(4)、(5)并做判别,在 Matlab 6.5 中编程,从表 1 可以看出,用 4 项指标判别,在第Ⅰ类中判别正确率为 90.16%;第Ⅱ类判别正确率为 80.00%;第Ⅲ类判别正确率为 85.71%;总判别正确率为(55+4+54)/129=87.60%,判别效果较好。

表 1　4 项指标判别正确率检验[例数(%)]

凝血类别	判别为Ⅰ类	判别为Ⅱ类	判别为Ⅲ类
第Ⅰ类	55(90.16)	6(9.84)	0(0.00)
第Ⅱ类	1(20.00)	4(80.00)	0(0.00)
第Ⅲ类	6(9.52)	3(4.76)	54(85.71)

从表 2 可以看出,用 3 项指标判别,在第Ⅰ类中判别正确率为 96.72%;第Ⅱ类判别正确率为 100.00%;第Ⅲ类判别正确率为 95.24%;总判别正确率为 96.12%,判别效果非常好。

表 2　3 项指标判别正确率检验[例数(%)]

凝血类别	判别为Ⅰ类	判别为Ⅱ类	判别为Ⅲ类
第Ⅰ类	59(96.72)	2(3.28)	0(0.00)
第Ⅱ类	0(0.00)	5(100.00)	0(0.00)
第Ⅲ类	3(4.76)	0(0.00)	60(95.24)

2.4　3 项指标和 4 项指标的验证

为了进一步验证 3 项指标和 4 项指标的诊断效果,在 Mat-

lab 6.5 中产生随机正整数作为各类组数，以此随机抽取Ⅰ类61组、Ⅱ类5组、Ⅲ类63组共129组组成训练样本集，其余129组组成检测样本集，如此随机抽样1 000次得到的诊断正确率为：4项指标最高诊断正确率93.02%，平均正确率83.07%；3项指标最高诊断正确率95.35%，平均正确率85.16%。由图1可知，3项指标和4项指标的诊断差异在0附近，大于0的较多一些，表明3项指标比4项指标的平均诊断正确率略高。

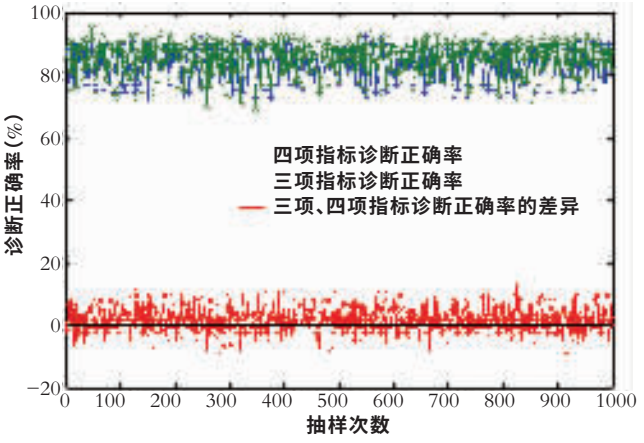


图1 抽样1 000次诊断正确率

3 讨论

在兔血液凝血功能的诊断中，虽然逐步判别分析没有剔除指标FIB，但剔除指标FIB后的判别正确率却有所提高。一方面，添加指标FIB后不会引起FIB的较大变化，有很好的判别效果，正确率达到了87.60%，因此指标FIB被选入；另一方面，由相关分析可以看出，FIB可能与其他指标有相互影响，从而剔除FIB后判别效果会提高。判别函数(5)可以作为兔血液凝血功能诊断的数学模型，并以此建立兔凝血功能的诊断系统，输入APTT、PT、TT的值，即可得到凝血功能的定性结果。

自Macfarlane、Davie和Ratnoff[6,7]分别提出凝血的瀑布学说后，人们对机体的凝血和抗凝机制已做了深入研究。研究表明，许多疾病，包括肺部疾病、肾脏疾病、肝脏疾病、糖尿病、感染、视网膜疾病、系统性红斑狼疮、弥散性血管内凝血、恶性肿瘤等以及器官移植和外科手术均可能出现出、凝血功能异常[8]，因此，凝血功能的确诊有积极的临床意义。长期以来，临床诊断凝血功能时，FIB是一项常规检测的指标[9]。而数学方法诊断有别于人的经验判断，因为数学方法诊断的正确率更高，诊断更可靠，且诊断所需指标可以更少，由实验结果可知，用判别分析诊断凝血功能时，剔除FIB后效果会更好。若人的血液凝血功能诊断也只需3项指标，则可以降低诊断成本，有极大的临床应用推广价值。

本研究获得的仅仅是兔的凝血诊断结果，至于人的血液凝血功能是否有相同结果，尚待进一步实验研究。

参考文献：

[1] 达万明，裴雪涛. 现代血液病学[M]. 北京：人民军医出版社，2003：171.
[2] 赵晓进，刘延霆，胡春红，等. SD大鼠下颌骨的性别判别分析[J]. 解剖学杂志，2006，29(4)：421－423.
[3] 黄小波，李宗信，陈文强，等. 慢性疲劳综合征中医证型判别分析[J]. 中国中医药信息杂志，2006，13(6)：21－22，49.
[4] 张晓东，申 洪. 肺癌细胞核三维结构的体视学定量研究(Ⅱ)——肺癌诊断分型的判别分析[J]. 中国体视学与图像分析，2006，11(3)：177－182.
[5] 刘升学，高 歌，袁 霞. 逐步判别分析在脑中风分类诊断中的应用[J]. 中国卫生统计，2006，23(5)：400－402.
[6] Macfarlane R G. An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier[J]. Nature, 1964, 202：498－499.
[7] Davie E W, Ratnoff O D. Waterfall sequence for intrinsic blood clotting[J]. Science, 1964, 145：1310－1312.
[8] 王振义，李家增，阮长耿，等. 血栓与止血基础理论与临床[M]. 第3版. 上海：上海科学技术出版社，2004：607－768，82－85.
[9] 刘泽霖，贺石林，李家增. 血栓性疾病的诊断与治疗[M]. 第2版. 北京：人民卫生出版社，2006：174.

(编辑 张 维)

文献标识码的标识方法

为了便于文献的统计和期刊评价，确定文献的检索范围，提高检索结果的适用性，每篇文章或资料应标识1个文献标识码。一般设置以下5种标准。

A 理论与应用研究学术论文(包括综述报告)
B 实用性技术成果报告(科技)、理论学习与社会实践总结(社科)
C 业务指导与技术管理性文章(包括领导讲话、特约评论等)
D 一般动态性信息(通讯、报道、会议活动、专访等)
E 文件、资料(包括历史资料、统计资料、机构、人物、书刊、知识介绍等)

不属于上述各类的文章以及文摘、零讯、补白、广告、启事等不加文献标识码。

本刊编辑部