

血瘀证量化诊断流行病学 调查数据的判别分析^{*}

□姚魁武 (中国中医科学院西苑医院 北京 100091)

王阶^{**} (中国中医科学院广安门医院 北京 100053)

朱翠玲 武继涛 (河南中医学院第一附属医院 郑州 450008)

王恰如 (广西中医学院第二附属医院 南宁 530001)

王在意 (广西医科大学第四附属医院 南宁 530021)

摘要:目的:运用统计学方法研究对于血瘀证诊断具有较好区分能力的诊断方法和诊断标准。方法:在现代流行病学原则指导下进行血瘀证诊断的多中心、大样本临床资料收集,对其结果进行判别分析处理。结果:对75个变量进行35步的逐步判别,共有33个项目进入方程,由此建立了血瘀型和非血瘀型的判别函数,并对33个自变量对血瘀证判别的贡献度进行区分。并对2004例患者进行了对判别结果的回顾性检验,总预测正确率为91.6%。结论:逐步判别分析可筛选出对鉴别两类具有不同属性的人群有较大贡献的条目,从而使其结果具有较好的区分度。

关键词:血瘀证 量化诊断 流行病学调查 判别分析

判别分析是根据判别对象若干个指标的观测结果来判定其应属于哪一类的统计学方法,常用于临床辅助鉴别诊断。临床上经常需要根据患者的主诉、体征、检查结果来作出诊断。逐步判别分析可以筛选出对于鉴别两类具有不同属性的人群有较大贡献的条目,从

而使其结果具有较好的区分度。血瘀证在中医证候学研究中因其普遍存在性而广受关注。目前对于血瘀证的诊断,多运用各种统计方法从不同角度进行考虑,但这些研究多是小样本的研究,尚缺乏现代流行病学指导下的多中心、大样本调查以提供更为客观、科学的数据支持。本研究拟通过对血瘀证量化诊断流行病学调查的结果运用判别分析进行处理,以发现对于血瘀证

收稿日期:2007-01-12

修回日期:2007-04-06

* 中国中医研究院创新工程项目(CX-00-02):血瘀证量化诊断、病证相关及证候实质研究,负责人:王阶;国家自然科学基金项目(90209011):基于计算智能的血瘀证量化诊断方法学研究,负责人:王阶;科技部973计划项目(2003CB517103):冠心病心绞痛病证结合的诊断标准及疗效评价体系研究,负责人:王阶。

** 联系人:王阶,主任医师,本刊学术副主编,博导,研究方向:中西医结合心血管病学、中医证候学研究, Tel 010-88001238 E-mail wangjie0103@yahoo.com.cn

和非血瘀证诊断具有较好区分能力的条目,以进一步作为血瘀证量化诊断量表的备选条目。

一、对象与方法

1 调查时间及医院选择

于 2002 年 5 月至 2004 年 1 月选择分别位于中国北部、中部、南部的北京、河南、广西 3 个中心的 4 家医院即中国中医研究院西苑医院、河南中医学院第一附属医院、广西中医学院第二附属医院、广西医科大学第四附属医院作为调查医院,来收集符合入选标准患者的临床病证信息。

2 研究对象

(1) 流行病学调查病种选择。

拟调查冠心病、非胰岛素依赖型糖尿病、高血压病、功能失调性子宫出血、脑梗死血瘀证和非血瘀证患者共 2000 例。其中血瘀证 1000 例,非血瘀证 1000 例。例数计算方法参照文献,依据既往文献记载的血瘀证发病率以及预试验结果采用查表法得出。

(2) 诊断标准。

冠心病采用 WHO 1979 年制订的诊断与分型标准;糖尿病采用 1997 年美国糖尿病协会/世界卫生组织 (ADA/WHO) 修订的糖尿病诊断标准;高血压病采用 WHO/ISH 1999 标准并删除继发性高血压;脑梗死采用中华医学会全国第四届脑血管病学术会议制定 (1995 年) 的诊断要点;功能失调性子宫出血采用《妇产科学》第四版的诊断标准。血瘀证诊断采用中国中西医结合学会活血化瘀专业委员会制订的血瘀证诊断标准 (1986 年,广州)^[1];其他气滞、气虚等证候诊断采用相关诊断标准进行诊断。

(3) 纳入标准。

符合样本选择中疾病诊断与证候诊断标准,无其他急性疾病以及严重并发症患者。

(4) 排除标准。

智力障碍、精神志障碍、不能全面回答问卷者,不愿意接受调查者,以及不符合上述诊断标准和纳入标准者。

3 调查方法

在文献分析、专家调查等基础上制定血瘀证量化

诊断调查表,内容包括患者基本情况、症状、体征等 (详细过程另文发表)。以访谈为调查方法,在安静和自然光线下对每一个病人进行望、闻、问、切的直接数据采集。回答率在 95% 以上的病例为合格病例。

4 运算程序

调用 SPSS 软件中 Discriminant Analysis 程序进行判别分析,引入自变量的方法采用 Use probability of F 方法。运算过程中考察各自变量之间的相关性,并计算其 Box's M 统计量,以检验各类协方差矩阵的齐同性;计算其 Wilks' Lambda 值,以检验各类均数的齐同性。计算其典则判别函数,考察标准化典则判别函数系数,分析各个自变量重要性大小^[2]。最后得出血瘀型和非血瘀型的判别函数。考察该判别函数判别结果的灵敏度、特异度。

二、结果

1 血瘀证流行病学调查基本情况分析

共调查 2100 例,合格病例 2004 例,血瘀证 1082 例,占 54%,非血瘀证 922 例,占 46%。其中完成实验室指标检测 308 例。各个中心病例数分布为:北京 884 例,河南 785 例,广西 335 例。

2 可靠性分析

平均测量类内相关系数 (ICC) = 0.7971, 其 95% CI 0.7852 ~ 0.8179, $P < 0.00005$ 认为平均测量类内相关系数不为 0。由于平均测量类内相关系数较高,故认为调查项目结果可靠性较高。

3 自变量间的两两相关检验

合并类内相关矩阵:

计算得出 75 个自变量的两两相关系数绝对值处于 0.001 ~ 0.428 之间,提示各自变量间相互独立,符合判别分析数据的要求条件。

4 各类协方差矩阵的齐同性检验

计算得 Box's M = 5999.573, 转换近似的 $F = 10.514$, $P = 0.052$ 按 $\alpha = 0.05$ 检验水准,接受 H_0 , 可以认为两类协方差矩阵差异无统计学意义,基本符合判别分析数据的假设条件。

5 引入/删除自变量结果

共进行了 35 步的逐步过程 (step), 共有 33 个项

目进入方程(函数),具体见表 1。

容差是反映自变量之间线性相关的指标,某自变量容差较小表示他几乎是其他自变量的线性组合。而 Wilks' Lambda值也称为 u- 统计量,即类内离均差平方和与总离均差平方和的比值,取值 0~ 1之间,如果该值较小,表示各类均数差异较大。

表 1 各个入选项目的容差、Wilks' Lambda值

入选项目	容差	删除自变量 F 水平的 P 值	Wilks' Lambda
舌质紫暗	0.593	0.000	0.394
舌脉曲张	0.828	0.000	0.380
脉涩	0.812	0.000	0.376
月经血块	0.415	0.001	0.352
头昏	0.625	0.000	0.352
肌肤甲错	0.701	0.000	0.356
行走不稳	0.428	0.000	0.352
耳鸣	0.623	0.000	0.354
齿龈色黑	0.584	0.006	0.351
嗜睡	0.821	0.034	0.351
腹部肿块	0.874	0.000	0.353
关节疼痛	0.884	0.001	0.352
舌体瘀斑	0.697	0.000	0.352
性别	0.774	0.28	0.351
恶心呕吐	0.875	0.000	0.352
皮肤瘀斑	0.694	0.000	0.352
痛经	0.554	0.001	0.352
辛辣饮食	0.823	0.000	0.353
职业	0.910	0.001	0.352
眼周色黑	0.536	0.005	0.351
五心烦热	0.798	0.000	0.353
消瘦	0.753	0.008	0.351
心绞痛	0.750	0.000	0.353
发育情况	0.771	0.010	0.351
眩晕	0.686	0.010	0.351
月经色黑	0.686	0.017	0.351
口唇色黑	0.639	0.15	0.351
语謇失语	0.560	0.011	0.351
月经紊乱	0.694	0.019	0.351
面色黑	0.534	0.051	0.351
外伤史	0.854	0.048	0.351
血管异常	0.931	0.65	0.351
反跳痛	0.959	0.084	0.350

最终方程含有 33个自变量,其容差均大于 0.4 提示有较大的容差,各自变量之间相互独立。他们的 P 值介于 0.000~ 0.084之间,按照删除水准 = 0.11,均拒绝 H_0 ,可认为方程中的 33个自变量均不可被删除。

6 典型判别函数系数

各入选项目的典则判别函数系数见表 2。

7 分类函数系数

入选项目以及其分数函数系数见表 3。

8 血瘀型和非血瘀型判别函数

血瘀型的判别函数:

$$Z1 = -33.158 + 8.038x_1 + 2.836x_2 + 2.001x_3 + 6.413x_4 + 0.48x_5 + 0.646x_6 + 0.487x_7 + 2.757x_8 - 0.651x_9 + 0.439x_{10} + 0.362x_{11} + 2.053x_{12} - 0.004x_{13} + 1.064x_{14} + 1.015x_{15} + 1.371x_{16} - 0.784x_{17} + 1.872x_{18} + 0.895x_{19} + 0.848x_{20} - 2.741x_{21} - 1.979x_{22} + 1.821x_{23} + 0.711x_{24} + 1.097x_{25} + 0.287x_{26} + 0.777x_{27} + 1.281x_{28} + 0.153x_{29} + 0.899x_{30} + 0.074x_{31} + 1.936x_{32} + 1.699x_{33}$$

表 2 典则判别函数系数表

入选项目	典则判别函数系数	入选项目	典则判别函数系数
舌质紫暗	0.140	辛辣饮食	-0.181
舌脉曲张	0.115	职业	0.156
脉涩	-0.293	眼周色黑	0.109
月经血块	-0.163	五心烦热	-0.164
头昏	-0.200	消瘦	-0.156
肌肤甲错	0.155	心绞痛	0.222
行走不稳	0.123	发育情况	-0.150
耳鸣	0.365	眩晕	0.272
齿龈色黑	-0.145	月经色黑	-0.380
嗜睡	-0.107	口唇色黑	0.296
腹部肿块	-0.247	语謇失语	0.289
关节疼痛	-0.136	月经紊乱	-0.267
舌体瘀斑	-0.118	面色黑	0.768
性别	-0.177	外伤史	0.175
恶心呕吐	-0.571	血管异常	0.604
皮肤瘀斑	-10.076	反跳痛	0.845
痛经	-0.243	Constant	-1.170

表 3 分类函数系数表

入选项目	是否血瘀	
	非血瘀型	血瘀型
性别 (x ₁)	7 655	8 038
职业 (x ₂)	7 655	2 836
辛辣饮食 (x ₃)	2 803	2 001
发育情况 (x ₄)	6 858	6 413
外伤史 (x ₅)	0 066	0 480
齿龈色黑 (x ₆)	0 223	0 646
口唇色黑 (x ₇)	0 153	0 487
肌肤甲错 (x ₈)	1 761	2 757
头昏 (x ₉)	- 0 256	- 0 651
眩晕 (x ₁₀)	0 733	0 439
耳鸣 (x ₁₁)	1 038	0 362
嗜睡 (x ₁₂)	2 423	2 053
语塞失语 (x ₁₃)	0 32	- 0 042
心绞痛 (x ₁₄)	1 549	1 064
反跳痛 (x ₁₅)	2 576	1 015
腹部肿块 (x ₁₆)	- 1 571	1 371
恶心呕吐 (x ₁₇)	- 0 118	- 0 784
皮肤瘀斑 (x ₁₈)	1 366	1 872
眼周色黑 (x ₁₉)	0 470	0 895
面色黑 (x ₂₀)	- 1 146	- 0 848
行走不稳 (x ₂₁)	- 2 293	- 2 741
消瘦 (x ₂₂)	- 1 552	- 1 979
五心烦热 (x ₂₃)	1 215	1 821
月经色黑 (x ₂₄)	1 122	0 711
关节疼痛 (x ₂₅)	0 355	1 097
血管异常 (x ₂₆)	1 325	0 287
痛经 (x ₂₇)	- 0 032	0 777
月经血块 (x ₂₈)	0 490	1 281
月经紊乱 (x ₂₉)	0 577	- 0 153
舌质紫暗 (x ₃₀)	- 1 199	0 899
舌体瘀斑 (x ₃₁)	- 0 405	0 074
舌脉曲张 (x ₃₂)	0 285	1 936
脉涩 (x ₃₃)	- 0 612	1 699
Constant	- 30 418	- 33 158

非血瘀型的判别函数:

$$Z_2 = - 30.418 + 7.655x_1 + 7.655x_2 + 2.803x_3 +$$

$$6.858x_4 + 0.066x_5 + 0.223x_6 + 0.153x_7 + 1.761x_8 - 0.256x_9 + 0.733x_{10} + 1.038x_{11} + 2.423x_{12} + 0.32x_{13} + 1.549x_{14} + 2.576x_{15} - 1.571x_{16} - 0.118x_{17} + 1.366x_{18} + 0.47x_{19} - 1.146x_{20} - 2.293x_{21} - 1.552x_{22} + 1.215x_{23} + 1.122x_{24} + 0.355x_{25} + 1.325x_{26} - 0.032x_{27} + 0.490x_{28} + 0.577x_{29} - 1.199x_{30} - 0.405x_{31} + 0.285x_{32} - 0.612x_{33}$$

9 标化典则函数

标化典则判别函数系数用于解释典则判别函数中自变量的相对重要性,系数绝对值大表示该自变量比其他自变量相对重要。表 4 中显示了 33 个自变量对血瘀证判别的贡献度。

10 回顾性检验结果

在对 2004 例患者进行回顾性检验时,判别结果总预测正确率为 91.6%,检验结果见表 5。对该标准的评价,采用文献^[3]的方法运用以下指标进行评价:灵敏度为 90.48%,特异度为 92.84%,真实度为 91.57%,阳性预测值为 93.68%,阴性预测值为 89.26%,阳性似然比为 12.6369,阴性似然比为 0.1025。

表 4 标化典则判别函数系数表

自变量	标化典则判别函数系数	自变量	标化典则判别函数系数
性别 (x ₁)	0 070	皮肤瘀斑 (x ₁₈)	- 0 081
职业 (x ₂)	0 099	眼周色黑 (x ₁₉)	0 108
辛辣饮食 (x ₃)	- 0 133	面色黑 (x ₂₀)	0 075
发育情况 (x ₄)	- 0 082	行走不稳 (x ₂₁)	- 0 154
外伤史 (x ₅)	- 0 060	消瘦 (x ₂₂)	- 0 086
齿龈色黑 (x ₆)	0 101	五心烦热 (x ₂₃)	0 122
口唇色黑 (x ₇)	0 085	月经色黑 (x ₂₄)	- 0 119
肌肤甲错 (x ₈)	0 192	关节疼痛 (x ₂₅)	0 095
头昏 (x ₉)	- 0 126	血管异常 (x ₂₆)	- 0 054
眩晕 (x ₁₀)	- 0 102	痛经 (x ₂₇)	0 128
耳鸣 (x ₁₁)	- 0 175	月经血块 (x ₂₈)	0 143
嗜睡 (x ₁₂)	- 0 065	月经紊乱 (x ₂₉)	- 0 078
语塞失语 (x ₁₃)	- 0 095	舌质紫暗 (x ₃₀)	0 541
心绞痛 (x ₁₄)	- 0 138	舌体瘀斑 (x ₃₁)	0 116
反跳痛 (x ₁₅)	- 0 049	舌脉曲张 (x ₃₂)	0 386
腹部肿块 (x ₁₆)	- 0 133	脉涩 (x ₃₃)	0 361
恶心呕吐 (x ₁₇)	- 0 112		

表 5 诊断试验评价四格表

判别分析结果	传统血瘀证诊断结果		合计
	血瘀证	非血瘀证	
血瘀证	真阳性 (a) 979	假阳性 (b) 66	(a+ b) 1045
非血瘀证	假阴性 (c) 103	真阴性 (d) 856	(c+ d) 959
合计	(a+ c) 1082	(b+ d) 922	N 2004

三、讨 论

统计学家认为医学诊断的推理过程可用数学的方法来精确描述。判别分析常用于临床辅助鉴别诊断, 计量诊断学就是以判别分析为主要基础迅速发展起来的一门科学。判别分析是根据判别对象若干个指标的观测结果来判定其应属于哪一类的统计学方法。临床上经常需要根据患者的主诉、体征、检查结果来作出诊断。逐步判别分析可以筛选出对于鉴别两类具有不同属性的人群有较大贡献的条目, 从而使其结果具有较好的区分度。

血瘀证在中医证候学研究中因其普遍存在性而广

受关注。目前对于血瘀证的标准化研究多集中于量化诊断, 而量化诊断多运用各种统计方法从不同角度进行考虑, 其结果也都从不同侧面反映了血瘀证的实质, 促进了证候研究。但这些研究多是小样本的研究, 尚缺乏现代流行病学指导下的多中心、大样本调查以提供更为客观、科学的数据支持。本研究将血瘀证量化诊断流行病学调查的结果运用判别分析进行处理, 来发现对于血瘀证和非血瘀证诊断具有较好区分能力的条目, 以进一步作为血瘀证量化诊断量表的备选条目。本研究运用判别分析从使诊断方法具有良好区分能力的角度筛选出了 33 个指标, 来作为血瘀证量化诊断量表制定的备选条目。这为下一步综合评价血瘀证诊断的入选项, 进而为血瘀证量化诊断量表的制定, 提供了实证依据。

参考文献

- 1 中西医结合学会活血化瘀研究会. 血瘀证诊断标准. 中西医结合杂志, 1987, 7(3): 129
- 2 刘润幸. SPSS10.0 医学统计方法与应用(上、下). 广州: 广东人民出版社, 2001: 235~ 238.
- 3 林果为. 诊断试验的研究与评价. 实用内科学(第 11 版). 北京: 人民卫生出版社, 2001. 173~ 177.

A Discriminant Analysis of Epidemiological Survey Data on Quantitative Diagnosis of Blood Stasis Syndrome

Yao Kuwui, Wang Jie, Zhu Cuiling

(Xiyuan Hospital, China Academy of Chinese Medical Sciences 100091, Beijing)

Here we report an investigation into the diagnostic approaches that are in a better position to differentiate the blood stasis syndrome (BSS). We used epidemiological survey data on BSS collected from three Chinese areas and made a discriminant analysis of the database. We made 35 stepwise discriminant analysis of 75 items, allowing 33 items into the equation/function. Discriminant functions of BSS and Non-BSS were established. The contributions of 33 items were differentiated. We had a 91.6% prediction accuracy for 2004 cases through a retrospection comparison. It is concluded that the stepwise discriminant analysis can be employed to screen the items that have a better contribution to differentiating the patients of different diseases.

Keywords: blood stasis syndrome, quantitative diagnosis, epidemiological survey, discriminant analysis

(责任编辑: 王 璠, 责任编审: 张志华, 责任译审: 邹春申)