

# 应用匹配矩阵、因子分析和聚类分析法 研究冠心病痰瘀证候特征

李 林<sup>1</sup>, 李晓冬<sup>2</sup>, 宋剑南<sup>3</sup>, 房祥忠<sup>2</sup>, 苗 阳<sup>4</sup>, 林 谦<sup>5</sup>

(1. 美国埃默里大学生物统计学部, 美国 亚特兰大 30322; 2. 北京大学数学学院概率统计系, 北京 100871; 3. 中国中医科学院基础理论研究所生化室, 北京 100700; 4. 中国中医科学院西苑医院心内科, 北京 100091; 5. 北京中医药大学东方医院心内科, 北京 100078)

[摘要] 目的: 研究冠心病痰瘀证候与临床四诊信息的相互关系, 为其分类及规范化诊断提供参考。方法: 采用匹配矩阵、因子分析和聚类分析法, 对 200 例冠心病患者痰证、瘀证、痰瘀互阻证及非痰非瘀证 46 项临床症状进行相关性分析。结果: 冠心病不同痰瘀证候舌、脉象与非痰非瘀证之间存在明显差异; 中医病机脏腑定位于心和肾, 以心气虚为主, 兼肾气虚, 痰证患者尚多伴脾气虚; 在 46 项临床症状中, 仅 15 项症状在痰证、瘀证和痰瘀互阻证中相伴出现频率较高。除共有的本虚证外, 痰证的主要临床症状为白腻苔和滑脉, 尚见胸脘痞满和腹胀等; 瘀证则以紫舌、舌生瘀斑为主, 尚见痛有定处; 痰瘀互阻证则兼上述两种证候的主要症状。结论: 综合应用匹配矩阵、因子分析和聚类分析法, 可以获得有关冠心病不同痰瘀证候主要临床特征及脏腑定位的信息, 对冠心病不同痰瘀证候辨证及其分类研究具有一定的意义。

[关键词] 因子分析; 聚类分析; 匹配矩阵; 冠心病; 痰瘀证候; 辨证

[中图分类号] R2-03 [文献标识码] A [文章编号] 1672-1977(2006)04-0343-05

## Applying matching matrix, factor analysis and clustering analysis to investigation of characteristics of syndromes of phlegm and blood stasis in patients with coronary heart disease

Lin LI<sup>1</sup>, Xiao-Dong LI<sup>2</sup>, Jian-Nan SONG<sup>3</sup>, Xiang-Zhong FANG<sup>2</sup>, Yang MIAO<sup>4</sup>, Qian LIN<sup>5</sup>

(1. Department of Biostatistics, Emory University, Atlanta, GA 30322, USA; 2. Academy of Mathematic Science, Beijing University, Beijing 100871, China; 3. Department of Biochemistry, Institute of Basic Theory, China Academy of Chinese Medical Sciences, Beijing 100700, China; 4. Department of Cardiology, Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing 100091, China; 5. Department of Cardiology, Dongfang Hospital, Beijing University of Traditional Chinese Medicine, Beijing 100078, China)

**ABSTRACT** Objective: To investigate the characteristics of syndromes of phlegm and blood stasis in patients with coronary heart disease by multiple statistical methods of matching matrix, factor analysis and clustering analysis, and to provide some references for classification and normalization of diagnosis of syndromes of phlegm and blood stasis of coronary heart disease. Methods: The correlations among 46 kinds of symptoms in syndrome of non-phlegm and non-blood stasis, syndrome of blood stasis, syndrome of phlegm and syndrome of phlegm-blood stasis blocking in 200 patients with coronary heart disease were analyzed by matching matrix, factor analysis and clustering analysis. Results: The manifestations of tongue and pulse in syndromes of phlegm and blood stasis were significantly different from those in syndrome of non-phlegm and non-blood stasis. The pathogenesis of viscera in syndromes of phlegm and blood stasis lied in the heart and kidney, and

[基金项目] 国家自然科学基金资助项目 (No. 90209001, 30171128); 国家重点基础研究发展计划 (973 计划) 资助项目 (No. 2003CB517105)

Correspondence to: Prof. Jian-Nan SONG. E-mail: sjn2003@sina.com

the syndrome of deficiency of heart qi was the most common one while the syndrome of deficiency of kidney qi took the secondary place. The syndrome of phlegm was often accompanied by syndrome of deficiency of spleen qi. Only 15 ones of 46 clinical symptoms showed high frequency in concomitant appearance in syndrome of blood stasis, syndrome of phlegm and syndrome of phlegm-blood stasis blocking. Apart from having the common symptoms in syndrome of deficiency in origin, the syndrome of phlegm especially showed white and greasy fur and slippery pulse as well as distention and fullness of chest and abdominal distension; the syndrome of blood stasis showed purplish tongue and ecchymosis on tongue as well as fixed pain; and the syndrome of phlegm-blood stasis blocking showed the main symptoms of both syndrome of phlegm and syndrome of blood stasis. Conclusion: The statistical methods of matching matrix, factor analysis and clustering analysis are convenient, and can definitely indicate the clinical characteristics and syndrome differentiation of viscera of different syndromes of phlegm and blood stasis, which are beneficial to further research of diagnosis and differentiation of such syndromes of coronary heart disease.

**KEY WORDS** factor analysis; clustering analysis; matching matrix; coronary heart disease; syndromes of phlegm and blood stasis; syndrome differentiation

Zhong Xi Yi Jie He Xue Bao / J Chin Integr Med, 2006, 4(4):343-347

www.jcimjournal.com

证候是对通过四诊(即望、闻、问、切)手段获取的机体在某一时空条件下对各种内外因素(包括机体生理功能及生物、化学、环境、精神、气候等各种致病因子)产生反应而呈现出的生理、病理状态信息进行综合判断的结果表述。因此,临床四诊信息是判定证候类型的重要依据,不同证候的临床表现呈现多样化和复杂性。在本研究中,我们采用匹配矩阵、因子分析、聚类分析等数学方法对冠心病痰瘀证候与症状之间的关系进行研究和分析,现报道如下。

## 1 资料与方法

1.1 一般资料 选择在北京西苑医院和北京东方医院就诊的心肌梗死患者以及经冠状动脉造影或介入治疗诊断为冠心病的患者共 200 例,其中男 105 例,女 95 例;平均年龄( $63.6 \pm 5.8$ )岁;分为痰证(20 例)、瘀证(79 例)、痰瘀互阻证(82 例)和非痰非瘀证(19 例)4 种证候类型<sup>[1]</sup>。上述患者均符合 1999 年美国心脏病协会公布的冠心病分型和诊断标准<sup>[2]</sup>。中医辨证标准参考国家技术监督局发布的《中医临床诊疗术语·证候部分》(国家标准 GB/T16751.2-1997)。

1.2 研究方法 本研究采用匹配矩阵、因子分析及聚类分析等方法<sup>[3,4]</sup>研究冠心病患者中医临床症状与痰瘀证候的相关性。

1.2.1 匹配矩阵 为了衡量临床症状在某一证候中伴随出现趋势的度量,我们提出匹配度的概念。将一组患者中两项症状变量的匹配度定义为在该组中同时出现这两项症状的患者数与至少出现其中一项症状患者数的比值。由此可见,匹配度的最小值

为 0,最大值为 1。匹配度越大,表示该组患者中若出现其中一项症状时将更倾向于同时出现另一项症状。由于我们的关注点在于研究症状伴随出现的趋势,因此并不关心两项症状均未出现的情形,且在匹配度的定义中也未将其考虑在内。为了研究具有同一证候的一组患者的若干症状,我们利用匹配度概念构造了匹配矩阵,矩阵的元素为相应两项症状间的匹配度。

1.2.2 聚类分析 聚类分析是一种比较标准的统计学方法。聚类分析的作用是根据症状将患者进行分类,在分类过程中并未涉及患者的证候信息。最后将分类结果与患者证候分类情况进行比较。如果两者分类结果一致,我们就可以确定不同证候主要的诊断症状。本研究采用聚类分析法中的一种,即离差平方和法。该方法的基本思想基于方差分析,即如果分类正确,则同类样品之间的离差平方和应较小,不同类样品之间的离差平方和应较大。SAS 统计软件既可提供谱系聚类图和最后的聚类情况分析,又可提供用于确定分类个数的阈值,其中应用比较广泛的是 R-square 和伪  $t^2$  统计量。通过分析 R-square 来确定分类个数时,主要视 R-square 值的变化而定,若由第 k 类到第(k+1)类时 R-square 值的变化很大,则可以分为 k 类。通过分析伪  $t^2$  统计量来确定分类个数时,若分 k 类时伪  $t^2$  统计量很大,则应分为(k+1)类。

1.2.3 因子分析 在进行聚类分析时,相对于患者人数来说,所收集的总共 46 项临床症状所包含的伪信息太多,因此要求我们必须加以浓缩提炼以得到有效信息。因子分析提供了这种可能,可以将多项

症状综合为相对少数的几个公因子。这些公因子是原来症状的线性组合,能够承载原来 46 项症状的大部分信息。因子分析法也是一个比较标准的统计学方法,通过分析得到这些公因子后就可以利用他们作进一步的聚类分析,如此得到的结果将更为可靠。

## 2 结果

2.1 匹配矩阵分析结果 我们收集到在冠心病患者中可能出现的症状共 46 项,包括舌象、脉象和主诉。首先分析 46 项症状之间的关系,计算匹配度以找出其中较大的值,即对应于倾向于相伴出现的症状。将所得结果按舌、脉象和主诉分别进行归纳,用图示法表示各项症状之间的相互关系(某项症状与其他症状的连线越多,表明其在相互关系中的作用越大,即伴随出现的机率越大)。冠心病非痰非瘀证候患者中医临床症状沉脉、薄苔、淡红舌、短气、乏力、心悸的集合比较符合心气虚及阳虚证候;而食少、腹胀则归属于脾气虚证候。冠心病瘀证患者舌象以薄苔、舌生瘀斑为临床特征性表现,此外伴随出现的主诉则以胸痛、心悸、乏力、短气(属于心气虚)、眩晕、腰酸(属于肾气虚)多见。冠心病痰证患者多有白腻苔、滑脉,这与中医临床经验判断的结果基本一致。除此之外,心悸、乏力、短气、胸脘痞满和腰酸亦是痰证患者的主要临床表现,同时该类证候患者也常出现腹胀、眩晕等与心气虚有关的症状。冠心病痰瘀互阻证患者常见腻苔和紫舌,同时心悸、胸痛、乏力、短气、胸脘痞满和腹胀等属于痰证或瘀证的症状亦相伴出现。见表 1~4,图 1。

表 1 冠心病非痰非瘀证候临床主要症状之间的匹配度(>70%)  
Table 1 Matching degree (>70%) among main clinical symptoms in syndrome of non-phlegm and non-blood stasis

Symptom	Symptom	Matching degree (%)
White fur	Deep pulse	90 .909
Weakness	Shortness of breath	80 .000
Weakness	Palpitation	80 .000
Weakness	Inappetence	80 .000
Pale tongue	Deep pulse	83 .333
Palpitation	Abdominal distention	75 .000
Palpitation	Inappetence	100
Abdominal distention	Inappetence	75 .000
Thin fur	Deep pulse	76 .923
Pink tongue	Stringy pulse	73 .333
Pale tongue	Pink tongue	76 .923
Pale tongue	White fur	76 .293
Pale tongue	Thin fur	75 .571
Thin fur	White fur	71 .429
Dizziness	Soreness of waist	100

表 2 冠心病瘀证临床主要症状之间的匹配度(>70%)  
Table 2 Matching degree (>70%) among main clinical symptoms in syndrome of blood stasis

Symptom	Symptom	Matching degree (%)
Pain of chest	Fixed pain	85 .714
Pain of chest	Palpitation	85 .714
Weakness	Palpitation	85 .714
Shortness of breath	Palpitation	85 .714
Palpitation	Dizziness	83 .333
Palpitation	Soreness of waist	83 .333
Weakness	Pain of chest	75 .000
Pain of chest	Shortness of breath	75 .000
Pain of chest	Dizziness	71 .429
Pain of chest	Soreness of waist	71 .429
Fixed pain	Palpitation	71 .429
Weakness	Dizziness	71 .429
Weakness	Soreness of waist	71 .429
Shortness of breath	Dizziness	71 .429
Shortness of breath	Soreness of waist	71 .429
Weakness	Shortness of breath	100
White fur	Thin fur	70 .000
Ecchymosis on tongue	Thin fur	70 .588

表 3 冠心病痰证临床主要症状之间的匹配度(>70%)  
Table 3 Matching degree (>70%) among main clinical symptoms in syndrome of phlegm

Symptom	Symptom	Matching degree (%)
Slippery pulse	Greasy fur	80 .000
Floating pulse	Moderate pulse	100
Distention and fullness in chest and hypochondrium	Weakness	100
Shortness of breath	Palpitation	100
Distention and fullness in chest and hypochondrium	Shortness of breath	100
Shortness of breath	Weakness	100
Weakness	Soreness of waist	100
Abdominal distention	Dizziness	100
White fur	Slippery pulse	75 .000
Inappetence	Red tongue	100
Weakness	Palpitation	100
Palpitation	Soreness of waist	100
Distention and fullness in chest and hypochondrium	Palpitation	100
Shortness of breath	Soreness of waist	100
Distention and fullness in chest and hypochondrium	Soreness of waist	100

2.2 冠心病痰瘀证候四诊信息因子分析和聚类分析结果 46 项症状变量按 0.8 累计贡献率,选取与各因子关系较为显著的症状(相关系数 > 0.5)共 18 项,此时累计贡献率为 0.8149。经方差最大化正交旋转后,选取与各公因子关系较为显著的症状(相关系数 > 0.5),获得如下结果:第 1 公因子主要相关症状有:乏力、短气、心悸;第 2 公因子主要相关症状有:胸脘痞满、白腻苔、滑脉;第 3 公因子主要相关症

状有:紫舌、舌生瘀斑、弦脉、胸痛;第 4 公因子主要相关症状有:淡舌、胖大舌或嫩舌、齿痕舌、沉脉、细脉;第 5 公因子主要相关症状有:眩晕、腰酸;第 6 公因子主要相关症状有:腹胀、便溏、小便清长。经方差最大化正交旋转后结果显示,第 1、4、5、6 公因子主要相关症状符合本虚证相关症状,第 2 公因子主要相关症状符合痰证相关症状,第 3 公因子主要相关症状符合瘀证相关症状。

在上述因子分析的基础上,通过计算各个样本各主要因子的得分,进行聚类分析。在本研究中,当选取前 5 个公因子参与聚类时,可将样本分为 4 类。第 1 类与痰瘀互阻证相关,第 2 类与瘀证相关,第 3 类与痰证相关,第 4 类与非痰非瘀证相关。计算聚类后各公因子的算术平均值,即各公因子与各类之间的关联度,选取其中关联度 > 0.5 的公因子。结果显示,第 1 类含 79 例患者,公因子所含症状主要包括:乏力、短气、心悸、胸脘痞满、白腻苔、滑脉、胸痛、痛有定处、紫舌、舌生瘀斑、弦脉,与痰瘀互阻

表 4 冠心病痰瘀互阻证临床主要症状之间的匹配度 (> 70%)  
Table 4 Matching degree (> 70%) among main clinical symptoms in syndrome of phlegm-blood stasis blocking (%)

Symptom	Symptom	Matching degree (%)
Weakness	Pain of chest	94.444
Palpitation	Pain of chest	94.444
Weakness	Shortness of breath	88.889
Weakness	Palpitation	88.889
Distention and fullness in chest and hypochondrium	Shortness of breath	84.211
Shortness of breath	Pain of chest	84.211
Palpitation	Shortness of breath	88.889
Palpitation	Distention and fullness in chest and hypochondrium	84.211
Distention and fullness in chest and hypochondrium	Pain of chest	89.474
Distention and fullness in chest and hypochondrium	Weakness	84.211
Abdominal distention	Weakness	70.588
Abdominal distention	Shortness of breath	70.588
Shortness of breath	Dizziness	76.471
Palpitation	Abdominal distention	70.588

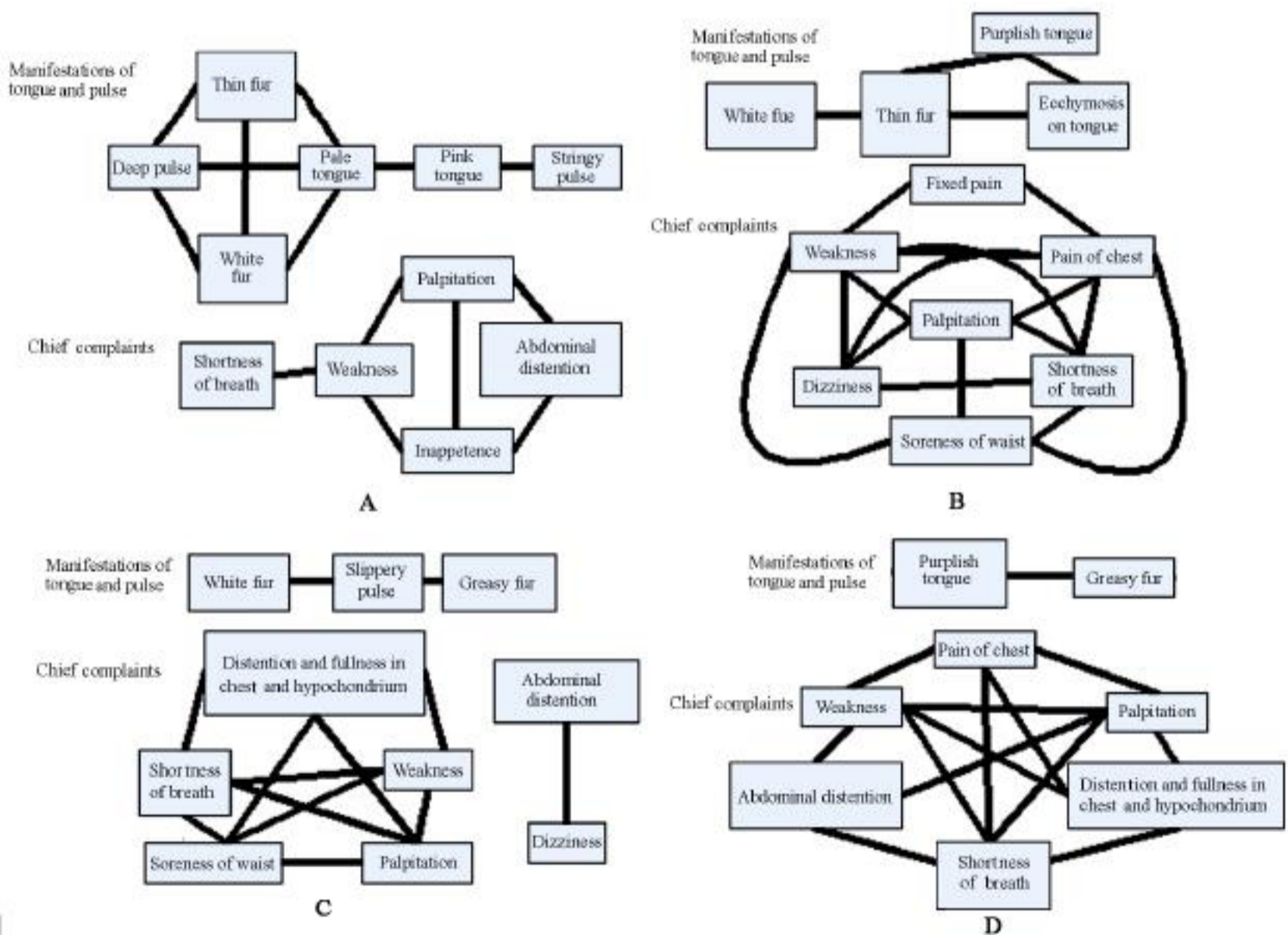


图 1 冠心病不同痰瘀证候与主要临床症状关系的匹配图

Figure 1 Matching results of main clinical symptoms in different syndromes of phlegm and blood stasis in patients with coronary heart disease

A: Syndrome of non-phlegm and non-blood stasis; B: Syndrome of blood stasis; C: Syndrome of phlegm; D: Syndrome of phlegm-blood stasis blocking

兼气虚型(邪实本虚)相关。第 2 类含 64 例患者,公因子所含症状主要包括:紫舌、舌生瘀斑、弦脉、胸痛,与瘀血内阻型(邪实)相关。第 3 类含 27 例患者,公因子所含症状主要包括:胸脘痞满、白腻苔、滑脉、乏力、短气、心悸、胸痛,与痰浊兼心气虚型(邪实本虚)相关。第 4 类含 15 例患者,公因子所含症状主

要包括:乏力、短气、心悸、淡舌、胖大舌或嫩舌、齿痕舌、沉脉、眩晕、腰酸、腹胀、便溏、小便清长,与气虚型(本虚)相关,属非痰非瘀证。另有 15 例患者未被纳入上述 4 类之中,这些患者主要症状包括:剥苔、口干、肢凉、身寒、沉脉、迟脉、烦热、便秘、小便短赤、夜尿多等。这些症状多与阴虚或阳虚相关。见表 5。

表 5 聚类分析结果与痰瘀证候的对应关系

Table 5 Corresponding relationship between clustering results and different syndromes of phlegm and blood stasis

Cluster					(Cases)
	Syndrome of non-phlegm and non-blood stasis	Syndrome of blood stasis	Syndrome of phlegm-blood stasis blocking	Syndrome of phlegm	Total
Cluster 1	5	14	54	6	79
Cluster 2	3	44	15	2	64
Cluster 3	1	11	4	11	27
Cluster 4	10	2	2	1	15
Total	19	71	75	20	185

### 3 讨论

矩阵中每一个元素代表两项症状同时出现的次数与出现总次数的商,若两项症状同时出现的次数越多则越匹配,即这两项症状在这组患者中倾向于相伴出现。本研究结果显示,冠心病痰证、瘀证、痰瘀互阻证主要表现为在舌、脉象上与非痰非瘀证之间存在明显差异。在诸多的临床四诊信息中,白腻苔、滑脉、紫舌是临床判定痰瘀证候的主要因子,其中前两项主要为痰证,后者主要为瘀证的诊断依据。冠心病标实证主要为痰浊和血瘀,本虚证则以心气虚为主,患者多有心悸、短气、乏力等症状,除此之外,多兼有肾气虚(眩晕、腰酸)和脾气虚(腹胀、食少)等。非痰非瘀证和痰证患者除舌象外,脉象也是重要的伴随指标,但瘀证和痰瘀互阻证患者脉象的伴随情况似乎不如舌象明显。

应用因子分析和聚类分析法对 46 项主要临床症状进行研究,证实这些症状之间存在一定的相关性,且各项症状所携带的信息量并不相同,如果直接进行分析,则所得结果一方面不易解释,另一方面也可能由于纳入过多的混杂信息而使典型的症状弱化从而影响结果分析。因此在本研究中,我们考虑将诸多症状中具有共性且相关性较强的症状归入同一类,代表群体信息中的 1 个“共同因子”,即所要寻找的证候的症状组合结构和规律,如此亦不会丢失主要的信息。公因子分析就是一种数据简化分析的常用方法。每 1 个公因子代表 1 个证候群。通过方差最大化正交旋转法以明确这些公因子是否具有统计

学意义。应用这种方法一般能对因子进行明确的解释。本研究结果显示,经过旋转后各公因子更能符合中医临床,可以合理地予以解释。得到公因子后,可以利用这些公因子对病例样本进行聚类分析。当聚类 R-Square 由 0.714(聚为 4 类)骤然下降到 0.520(聚为 3 类)时,说明聚为 4 类是比较合适的;同时,在聚为 3 类时,伪  $t^2$  统计量高达 219,明显高于其他分类情况时的伪  $t^2$  统计量,因此也说明聚为 (3+1)类是较为合适的。根据上述分析结果,我们可以初步认为每一类所包含的主要症状信息可以基本代表冠心病不同痰瘀证候的主要诊断指标。

综上所述,应用匹配矩阵、因子分析和聚类分析法,我们初步获得了冠心病不同痰瘀证候及其兼夹证典型症状主要的脏腑定位,这对于冠心病中医临床辨证分型及诊断规范化研究具有一定的意义。

#### [参考文献]

- 1 陈国伟, 郑宗铎. 现代心脏内科学. 长沙: 湖南科学技术出版社, 2002. 1051-1053.
- 2 欧阳涛, 宋剑南, 苗阳, 等. 冠心病痰瘀证与载脂蛋白 E 基因多态性关系的研究. 中西医结合学报, 2005, 3(6): 438-442.
- 3 高惠璇. 应用多元统计分析. 北京: 北京大学出版社, 2001. 185-209, 277-300.
- 4 高惠璇. SAS 系统与基础统计分析. 北京: 北京大学出版社, 2005. 216-259, 293-317.

[收稿日期] 2005-11-28