

基于隐结构模型的血瘀证症状体征研究^{*}

□王阶^{**} 李海霞^{**} (中国中医科学院广安门医院 北京 100053)
 王焱 张连文 (香港科技大学计算机系 香港 九龙)
 姚魁武 (中国中医科学院西苑医院 北京 100091)

摘要:以血瘀证为例,按照循证医学的原则,开展了多中心、大样本的流行病学调查,获取了2004份血瘀证症状体征调查表,用隐变量分析方法对获得的症状体征数据进行了分析,建立了一个可以用中医理论解释的隐结构模型。研究表明,利用隐结构模型可以探讨中医辨证的客观化、定量化,揭示众多症状间以及症状与证候间复杂的相互关系,为中医辨证的科学发展提供依据。

关键词: 隐结构 中医辨证 规范化

隐变量分析是应用统计学领域中近20年来发展最为迅速的一个分支。它首先应用于心理学、社会学,后来几乎应用到了一切科学。现在我们尝试将其应用于中医学。中医学是以整体观为核心的辨证论治医学体系,重视患者生存质量,无论在临床研究还是基础研究中,均努力为最佳的临床诊治寻求科学依据^[1]。因此,中医证候的量化诊断和数理研究方法在促进学科发展上具有至关重要的作用。隐变量分析是一种很受实际工作者欢迎的统计方法。国外介绍隐变量的图书至少已有50本以上了,但国内还很少有正式出版物。

随着血瘀证标准化、客观化研究的深入,中医证候研究需要新的方法以使研究深入进行,而隐变量分析方法比较适合中医证候的分布规律。本文以血瘀证为例,运用循证医学的原则开展多中心临床随机对照试验研究,用隐变量的方法对数据进行分析,建立隐结构模型,进一步完善临床病证诊断标准化与疗效评价体系的规范化,以构建疗效确切、特色明显的中医临床诊治模式和辨证标准。

一、临床资料

1. 诊断标准

高血压诊断标准:采用世界卫生组织(WHO/ISH)1999年标准。

收稿日期:2006-03-28

修回日期:2006-05-29

* 国家重点基金研究发展规划项目(2003CB517103):证候规范及其与疾病、方剂相关的基础研究—冠心病心绞痛病症结合诊断标准及疗效评价体系研究,负责人:王阶;国家中医药管理局课题(02-03JP47):血瘀证病位分类与诊断研究,负责人:王阶;国家自然科学基金重大项目(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。

** 联系人:王阶,本刊副主编,博导,主要研究方向:中西医结合心血管病防治,病证结合方证对应研究, Tel:010-88001238, E-mail:wangjie0103@yahoo.com.cn; 李海霞,博士,研究方向:中西医结合心血管病防治研究, E-mail:haixiali2008@yahoo.com.cn

糖尿病诊断标准:采用美国糖尿病协会(ADA) 1997年修订的糖尿病诊断标准。

脑梗塞诊断标准:采用中华医学会全国第四届脑血管病学术会议制订(1995年)的诊断要点。

冠心病心绞痛诊断参照 ACC/AHA2002年发布的 ST段抬高心梗治疗指南^[1];冠心病心绞痛症状分级参照1979年中西医结合治疗冠心病心绞痛及心律失常座谈会制订的《冠心病心绞痛及心电图疗效评定标准》^[2]。

功能失调性月经出血:符合《实用妇产科学》有关诊断标准,除外大出血病例、合并其他妇科疾病的病例。

2. 纳入及排除标准

西医诊断:所有入选病例的西医诊断应符合以上疾病诊断标准;年龄、性别、民族不限;可以合并其他疾病,但不能存在本课题所选病种的各种急性并发症,并按要求排除合并其他急性疾病的病例。

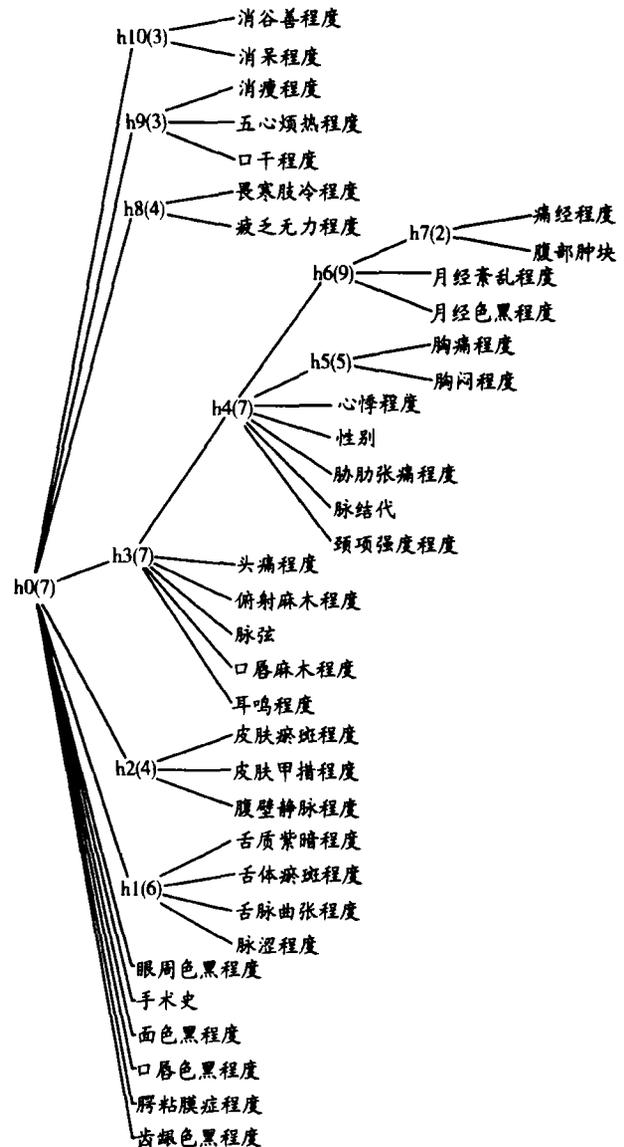
排除标准:进行过系统中医治疗者;存在智力障碍、精神、神志障碍,不能正确回答问卷者;血瘀证在证候诊断中作为次要证候或伴随证候者;正患有急性疾病如外伤、上呼吸道感染等疾病;有上述几种疾病的急性并发症如糖尿病酮症酸中毒。

3. 病历来源

2001年7月至2004年5月期间来自中国中医科学院西苑医院、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,共2004份病例,流行病学调查包括脑梗死、冠心病心绞痛、糖尿病、功能失调性子宫出血、原发性高血压病5种疾病,其中血瘀证1082例,非血瘀证922例。

二、方法

选用血瘀证的36个症状和体征,在统计中未进行分类,不含诊断结果,通过隐结构模型提取总体血瘀证、血瘀证亚型以及不同病位血瘀证的症状体征,观察它们之间的相关性,从而为血瘀证诊断标准的规范化和客观化提供合适的诊断方法。用多层隐类模型分析^[4],最终得出 BIC^[1]分最高的模型 M_{.....1}(结果如



对数似然度:40472.3;BIC分:43414.6

图1 模型 M_{.....1}之结构

图1),称之为模型辨证^[3]。它的对数似然度和 BIC 分别是 -40472.3 和 -43414.6。

三、结果

图2给出了模型的诠释。总体而言,由图2我们可以很容易看出,h0可以看作总体血瘀证,而与h0直接相关的h1、h2、h8、h9、h10和其他血瘀症状体征,如

眼周色黑、手术史、面色黑、口唇色黑、腭黏膜征及齿龈色黑等,都是血瘀证及其亚型的直接表现。它的分支分别为不同病位的血瘀证表现,如 h3 代表头部疾病, h4 代表胸部疾病, h5 代表心脏疾患, h6 代表妇科疾病, h7 代表妇科疾病的进一步发展,而 h3 - h7 从上往下又可以看成分别为头、胸、腹部疾病。

对于 h0 与其他项之间的联系,我们可以看为双向,从上往下可以看成由总体血瘀证衍生出的其他相关血瘀或亚型,从下往上可以看成是由不同的血瘀亚型和相关血瘀疾病归纳而成的总体血瘀证。

1. 隐变量 h3、h4、h5、h6、h7 的诠释

现在对模型 M_{manual} 1 中的隐变量逐一分析。M_{manual} 1 中隐变量 h3 直接影响头痛、偏身麻木、脉弦、口唇麻木、耳鸣症状的出现和程度的轻重,可以看成头部疾病的主要症状体征; h4 直接影响心悸、肋肋胀痛、脉结代、颈项强痛的程度,可以看成胸部疾病的症状体征; h4 通过影响 h5,即胸痛、胸闷程度,间接地影响心脏疾病,同时又影响 h6,月经紊乱、月经色黑程度,可以看成小腹部疾病; h6 直接影响 h7,即直接影响痛经和腹部肿块程度,可以看成妇科疾病。其中 h3、h4、(h5、h6 并列)、h7 之间有线相连,说明它们相关,箭头可以从 h3 指向 h7,从上到下分别可以解释为头、胸、腹部疾病。

2. 隐变量 h1、h2 的诠释

h1 直接影响舌质紫黯、舌体瘀斑、舌脉曲张、脉涩程度这 4 个体征的出现和程度,表示血瘀对舌脉的影响; h2 直接影响皮肤瘀斑、肌肤甲错、腹壁静脉曲张的出现和轻重程度,表示血瘀对全身皮肤的影响。

3. 隐变量 h8、h9、h10 的诠释

隐变量 h8 直接影响畏寒肢冷、疲乏无力的出现和程度。中医认为气虚出现畏寒肢冷、疲乏无力,进而导致血瘀。隐变量 h9 直接影响消瘦、五心烦热、口干出现和程度,从中医角度

讲,阴虚可以导致虚火上炎和血瘀。隐变量 h10 直接导致消谷善饥、纳呆的出现和程度,说明血瘀可以瘀阻胃的功能。

4. 隐变量 h0 的诠释

以上隐变量中有 6 个与隐变量 h0 直接相连,它们分别是 h1、h2、h3、h8、h9、h10。这些变量分别对应血瘀的不同侧面, h1 为血瘀在舌脉的表现, h2 为血瘀在肌肤的表现, h3 为头部血瘀表现, h8 为气虚血瘀, h9 为阴虚血瘀, h10 为血瘀阻滞胃腑。h0 直接影响眼周色黑、手术史、面色黑、口唇色黑、腭黏膜征及齿龈色黑程度,这些症状、体征是血瘀的典型体征和致病因素。换句话说,如果病人出现了 h0 直接影响的症状体征,就可以确诊为血瘀证。

5. 定量部分

模型 M_{manual} 1 有两方面的定量内容:一方面,它明确规定隐变量的取值个数;另一方面,它用条件概率分布来具体刻画变量之间的关系。以隐变量 h9 为例,基于它直接影响消瘦程度、五心烦热、口干的出现和程度这一事实,把它解释为阴虚程度。在定量层面, h9 共有 4 个不同的取值,分别记为 s0、s1、s2、s3;它与消瘦、五心烦热、口干程度之间的定量关系由图 3 所示的两个条件概率分布给出。通过对隐变量与症状变量之间的定量关系进行分析,可以揭示隐变量取值的直接含义。

图 3 中第三副图表示当 h9 = s2 时, 3 个变量的概率分布,症状变量消瘦、五心烦热、口干程度有无、轻、中、重 4 个取值;与之相应,柱体分为不同灰度的 4 段,每段对应一个取值,其高度表示该取值的概率。图 3

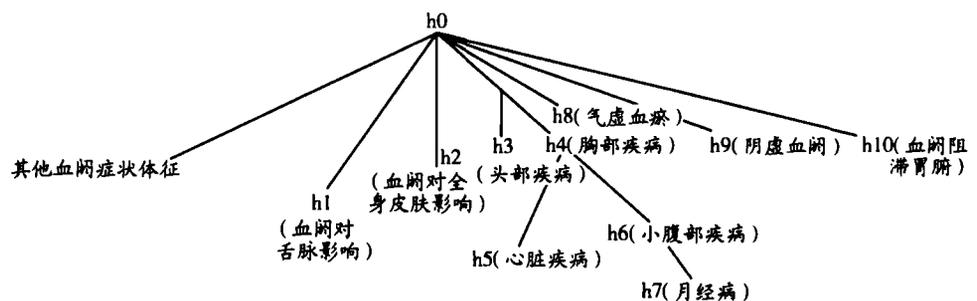


图 2 模型 M_{manual} 的诠释

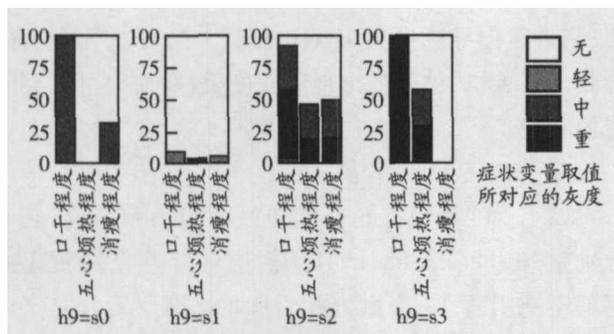


图3 变量 h_9 (阴虚程度) 取值的类概率柱图

同时也给出了 $h_9 = s_0, h_9 = s_1, h_9 = s_3$ 的类概率柱图。类概率柱图让我们对隐变量取值的意义有了一个既直观又具体的认识。 h_9 的取值 s_2 的意义是, s_2 代表一类人, 这类人中消瘦、五心烦热、口干程度的出现情况如图3中第三副柱图所示。

类概率柱图同时也让我们对隐变量不同取值之间的关系有了一个既直观又具体的认识。图2显示: 口干在 $h_9 = s_0, h_9 = s_2, h_9 = s_3$ 这类人身上重度出现, 在 $h_9 = s_1$ 这类人身上几乎不出现; 五心烦热, 在 $h_9 = s_0, h_9 = s_1$ 这类人身上几乎不出现, 在 $h_9 = s_2, h_9 = s_3$ 这类人身上中度出现; 消瘦, 在 $h_9 = s_0$ 这类人身上轻度出现, 在 $h_9 = s_2$ 这类人身上中度出现, 在 $h_9 = s_0$ 和 $h_9 = s_3$ 这类人身上几乎不出现。

四、讨论

用 $M_{\dots\dots\dots 1}$ 来指导辨证的最重要的意义在于, 它能大大提高辨证的客观性^[3]。一是推理过程的客观性, 即贝氏网中的推理完全是按照概率论原则进行, 没有主观性; 二是模型 $M_{\dots\dots\dots 1}$ 内容的客观性。模型 $M_{\dots\dots\dots 1}$ 是计算机按照统计学原则对数据进行分析处理的结果, 这里的数据只包括症状显变量的取值, 而不包括证型隐变量的取值, 在数据收集过程中排除了对证型隐变量的判断, 症状轻重程度判断中的主观性大大减少; 三是结果的直观性: 图2可以直观地显示出不同病位血瘀证和总体血瘀证、血瘀亚型之间的关系; 图表的形式可以直观地展示 h 与一些症状变量的关系, 使得 h 的取值所对应的类的性质一目了然, 同时可以把 h 的取值按轻重程度排序。实际上, 隐变量的诠释过程, 是

把模型 $M_{\dots\dots\dots 1}$ 与中医理论的定性分析进行对照比较的过程。模型 $M_{\dots\dots\dots 1}$ 中所有隐变量均能被诠释, 说明 $M_{\dots\dots\dots 1}$ 的定性内容与中医(血瘀辨证)理论在整体上是吻合的。它一方面显示 $M_{\dots\dots\dots 1}$ 是一个高质量模型, 可以指导辨证, 另一方面, 它为中医辨证理论的科学性提供了依据。我们可以根据不同类病人症状、体征程度的轻重指导临床辨证治疗, 在总体血瘀证辨证治疗的基础上根据病人症状体、征轻重程度的不同对治疗方剂进行化裁。在血瘀证隐结构模型成熟后, 新病人的症状、体征输入微机就判断类型种类, 直接进行电脑处方, 为中医的诊疗方法带来新的进展。

为了使调查结果具有的普遍性, 我们选取代表中国北、中、南三个地区的北京、河南、广西做为调查地域。在对数据未进行分类和不含诊断结果的前提下, 通过对一组没有任何人为特征的数据进行计算机自动分析, 得出了不同病位血瘀证及亚型的结构图, 可以直观地看出总体血瘀证和不同病位血瘀证、总体血瘀证和血瘀亚型之间的相互关系, 以及头、胸、腹部由上往下的血瘀对应关系, 客观地验证了血瘀证与不同病位的关系以及血瘀证与亚型的关系, 为临床辨证和用药提供了客观依据, 减少了主观因素对证候诊断的影响。结果表明, 隐结构模型可以揭示众多症状间以及症状与证候间复杂的相互关系, 有助于加强临床诊断正确性和规范化, 有助于确定证候诊断的标准和规范, 并且可以指导临床辨证及用药。它不仅能提高中医专家诊断的客观性和准确性, 还有可能发现新的医学知识, 为中医客观化的法展提供思路, 带来中医规范化研究的突破。

参考文献

- 1 陈家旭. 中医证候与辨证体系研究现状、展望. 北京中医药大学学报, 2001, 24(4): 3-8.
- 2 孙尚拱. 隐变量分析简介. 数理统计与管理, 2002, 21(1): 52-56.
- 3 张连文. 袁世宏. 隐结构模型与中医辨证. Technical Report HKUST-CS04-12. Department of Computer Science, The Hong Kong University of Science & Technology (Revised: Feb 2006), <http://www.cs.ust.hk/~lzhang/tcm/>.
- 4 Zhang, N. L. (2002). Hierarchical latent class models for cluster analysis. AAAI-02, 230-237.

(Continued on Page 62)

**Effects of Fertilization on Yield, Quality and
Biomass Distribution of *Tussilago farfara***

Peng Rui, Ye Yulan, Qin Songyun, Li Longyun

(Chongqing Academy of Chinese Materia Medica, Chongqing 400065, China)

Wang Jun, Liu Zhihe

(Chongqing Yuanfan Pharmaceutical Plant Company, Wuxi 405800, China)

Liu Yi

(College of Bioinformatics, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

The paper discusses the effect of fertilization on yield, quality and biomass distribution of *Tussilago farfara*, in the context of fertilizer type and application quantity. The following are among the basic methods and processes used in the study: survey of biomass distribution, yield statistics collection, and HPLC and UV for measuring rutin content and total flavones in *T. farfara*. It is found that both fertilizer type and application quantity are of a noticeable effect on biomass distribution, yield, rutin content, and total flavones. Authors conclude that phosphate and potash fertilizer can noticeably increase the number of *T. farfara* buds, with a controlled application of nitrogen. Nitrogen, phosphate, and potash fertilizer and organic manure can, in a reasonable range, increase both the yield and quality of *T. farfara*.

Keywords: *Tussilago farfara*, fertilization, biomass, yield, rutin, total flavones

(责任编辑:张述庆, 责任编辑:丁万隆, 责任译审:邹春申)

(Continued from Page 26)

**A Study of Bloodstasis Symptoms and Physical Signs
Based on Latent Structure Models**

Wang Jie, Li Haixia

(Guanganmen Hospital of China Academy of Chinese Medicine Sciences Beijing 100053)

Wang Yi, Zhang Lianwen

(Department of Computer, Hong Kong University of Science and Technology)

Yao Kuiwu

(Xiyuan Hospital of China Academy of Chinese Medicine Sciences Beijing 100091)

Latent variable analysis is a widely used statistical method. Standardized and objective study of bloodstasis syndrome calls for a more sophisticated approach. Latent variable analysis makes a right method for such study. In the paper, bloodstasis syndrome is selected as the study objective. An extensive epidemiological survey is conducted in line with the principle of evidence-based medicine. Some 2004 questionnaires concerning bloodstasis symptoms and physical signs are collected, and a Latent Structure Model is established to analyze the data. The model results are interpreted using TCM theory. The study suggests that Latent Structure Model is desirable for differentiating symptoms and signs, which creates a scientific ground the further development of TCM.

Keywords: latent structure, symptom assessment, standardization

(责任编辑:张述庆, 责任编辑:李彭涛, 责任译审:邹春申)