

• 思路与方法 •

基于熵的复杂系统分划方法与冠心病
心绞痛证候量化诊断标准研究贾振华¹ 吴以岭¹ 高怀林¹ 西广成² 陈 静²
谷春华¹ 袁国强¹ 吴相春¹ 魏 聪¹

(1. 河北医科大学附属以岭医院, 石家庄市天山大街 238 号, 050091; 2. 中国科学院自动化研究所)

[关键词] 冠心病心绞痛; 基于熵的复杂系统分划方法; ROC 曲线分析; 证候量化诊断

证候规范研究是中医现代化、科学化和标准化研究的热点领域, 建立量化诊断标准是证候规范研究的重要内容。由于构成证候的症状之间存在着非线性的复杂关系, 采用处理线性关系的方法分析症状多变量对证候的非线性关系, 难以充分反映症状间相互作用的复杂关系^[1], 引入非线性复杂科学方法已成为证候规范研究发展的内在需求。

冠心病心绞痛属中医“胸痹”、“心痛”等病证范畴, 历代文献对其证候表现有丰富记载, 有关学术团体及专家亦提出关于冠心病心绞痛的辨证分型标准, 但尚缺乏冠心病心绞痛量化诊断标准的系统研究。本研究综合运用文献研究、专家咨询、临床流行病学调查等方法获取冠心病心绞痛四诊信息, 以基于熵的复杂系统分划及诊断性试验方法, 从原始症状关联度的无监督分析着手, 确立冠心病心绞痛基本证型的症状信息、基本证候之间关联度、症状贡献度和诊断阈值, 建立起量化的诊断标准, 并分析了其基本证型之间组合规律。

1 对象与方法

1.1 研究对象

1.1.1 入选标准 符合 WHO《缺血性心脏病命名及诊断标准》中《冠心病心绞痛诊断标准》; 心电图检查有缺血性改变; 年龄 35~70 岁。

1.1.2 排除标准 急性心肌梗死及其他心脏疾病; 合并脑血管病、周围血管病及糖尿病血管并发症; 重度神经官能症、更年期症候群、颈椎病、胆及食道源性胸痛者; 重度心肺功能不全、重度心律失常者或肝、肾、造血系统等严重原发性疾病、精神病患者。

1.1.3 研究对象与分组 以北京、上海、广州、济南、沈阳、长春等地 8 家三级甲等医院为调查中心, 收集 2006 年 2 月~2006 年 9 月期间就诊的符合入选与排除标准的冠心病心绞痛患者 403 例, 随机分为运算组 299 例, 考核组 104 例, 运算组用于建立证候量化诊断标准, 考核组用于量化诊断标准的前瞻性检验。运算组男 150 例, 女 149 例, 平均年龄 63.6 岁, 平均病程 4.2 年; 考核组男 52 例, 女 52 例, 平均年龄 61.2 岁, 平均病程 4.6 年, 两组患者年龄、病程等基线情况差异无统计学意义($P>0.05$)。

1.2 研究方法

1.2.1 中医四诊信息的筛选与分级 (1) 文献研究检索《中华医典》、《中国生物医学文献数据库》(1978~2005 年), 以及专著、教材、学术团体及国家和行业制定的有关诊断标准, 采集其中涉及中医“胸痹”、“心痛”、“厥心痛”、“卒心痛”或西医冠心病心绞痛中医辨证分型及其相关的四诊信息, 对症状进行修改、删除等初步整理, 对实际含义相同的症状选定和使用统一规范症名, 并明确每一症状所代表的实际含义。(2) 专家咨询根据文献研究筛选的四诊信息条目池, 制定专家问卷, 选择北京、上海、广州等地三级甲等医院中医、中西医结合专业副主任医师以上专家 30 名, 进行二轮专家问卷调查, 采用德尔菲分析法, 综合问卷调查结果确定进行临床调查的中医四诊信息。(3) 对入选的中医四诊信息, 参考文献报道分级方法^[3,4]分为无、轻、中、重四级, 对不能分级的舌脉分为无、有两级。

1.2.2 制定临床调查表 采用流行病学横断面调查方法, 制定临床调查表, 内容包括患者性别、出生年月、婚姻状况、职业、经济收入等基本情况和疾病史及相关危险因素, 分级的中医四诊信息及与本病相关的临床常规客观检测项目。

1.2.3 现场调查 正式调查时调查人员根据调查表

基金项目: 国家重点基础研究发展计划(“973”计划)项目
(2005CB523301)

通讯作者: 吴以岭, jiatcm@163.com, (0311)85901553

内容详细询问受试者,将数据及时、完整、正确、清晰地载入调查表,所有调查表填写正确完整,并与原始资料一致,如有错误和遗漏,及时要求研究者改正。

1.2.4 质量控制 承担调查任务的研究单位指定专人负责调查,参与调查人员为中医、中西医结合专业主治医师以上临床医生;制定临床调查研究人员手册,记录研究中采用的诊断标准、纳入标准和排除标准,并对有关中医名词术语及其他易产生歧义或需定义的调查名词做出统一解释;调查前对参与调查人员进行统一培训,熟悉调查表内容;正式调查前进行小范围预调查,对调查项目及调查可行性作进一步考察;随机抽取 10% 的病例在第一次调查的同一天(分上、下午)对部分代表性中医四诊指标进行重复调查,比较两次调查的一致性,本次重复调查四诊指标的一致率为 92%~95%,具有较好一致性。

1.3 数据处理与分析

1.3.1 对回收的调查表用 Epidata3.0 建立数据库,数据由双人双录入,将二次录入的数据逐项核对,二者不统一的项目与原问卷进行核对纠正;用 EXCEL 对核对无误的数据进行整理、归类,并量化赋值,二分类指标变量记 0(无)、1(有);等级指标记 0(无)、1(轻)、2(中)、3(重)。

1.3.2 将运算组 299 例调查资料输入 Matlab 数据分析软件,采用基于熵的复杂系统分划方法对数据进行处理分析,提取其症状信息,归纳基本证型,分析基本证型之间关联度,计算基本证型中症状贡献度。根据症状赋分对 299 例患者症状进行打分,计算基本证型中症状总积分,将数据输入 SPSS 统计软件,采用诊断性试验中的受试者工作特性(ROC)曲线分析方法,参考不同诊断界点的灵敏度和特异度,结合医理建立各基本证型的诊断阈值。

1.3.3 将运算组建立的量化诊断标准用于考核组 104 例冠心病心绞痛患者诊断,参照专家辨证结果,计算其灵敏度、特异度和准确度,进行诊断标准的前瞻性检验。专家辨证由领域内三位副主任医师以上职称中医师一致性诊断确定。

2 结果

2.1 冠心病基本证型症状提取

对于冠心病心绞痛数据,首先应用熵分划方法在不同的参数下得到大量的症状集合群。以乏力、神疲为例,首先根据乏力在患者中的分布计算其熵值,然后根据乏力和神疲在患者中的分布计算其联合熵值,进而得到乏力、神疲的关联度数值,同理应用算法可对其他症状之间进行自由聚类。症状集合群即以症状之间的

内在关联度为基础,临床患者症状表现不完全与算法所聚症状吻合,可根据每个症状集合中符合中医理论内涵的症状集合,从医理角度进行合并和归结,得到既有算法支持又符合中医理论的症状集合。结果通过算法获得 8 个症状集合,经分析判定为与病位相关反映疾病特异性表现的胸闷、胸痛、心悸,与病情相关反映疾病全身性非特异性表现的气滞、气虚、痰湿、血瘀、阴虚、阳虚、痰热、郁热(相关表格略)。

2.2 各基本证型的关联度分析

将各基本证型视为广义子系统,进行基于熵的子系统关联度分析,方法同上,分别计算各子系统熵值、各子系统间联合熵值、各子系统间关联度。结果显示,气滞可与痰湿、血瘀、郁热、阴虚、痰热相关联,气虚可与痰湿、血瘀、阴虚、阳虚、痰热相关联,痰湿与阳虚、痰热亦有关联关系(相关表格略)。

2.3 各症状对证候诊断贡献度

将各症状与所在症状集合的关联数值作为症状对证型的贡献度,反映症状的重要程度。以气虚证神疲症状为例,其贡献值由神疲和气虚证中其他症状之间的关联度系数整合得到,同时可分别计算出其他症状的贡献度,并转化为整数值。由于病位症状可能在任何单一证型中出现,因此,分别计算其对不同证型的贡献度,通过取均值作为最终贡献度,结果胸痛、胸闷、心悸分别为 0.0257、0.0279、0.0209;转化后的贡献值为 3、3、2(相关表格略)。

2.4 各基本证型 ROC 曲线分析及诊断阈值的建立

根据各基本证型包含症状及其贡献分对 299 例患者按基本证型进行打分,以每个基本证型得分作检验变量,专家辨证作状态变量,作 ROC 曲线分析,计算 ROC 曲线下面积,结果表明,各基本证型 ROC 曲线面积均大于 0.5(P 均 <0.01),其 95% 可信区间也均大于 0.5,表明该证型积分对证候诊断具有一定价值。根据 ROC 曲线移动诊断界点,参考诊断灵敏度和特异度,确定诊断阈值(相关表格略)。

2.5 前瞻性检验结果

对 104 例冠心病心绞痛患者依据各基本证型诊断阈值辨证,同时以三位领域内副主任医师以上职称中医师对该部分患者进行专家辨证,以三者的一致性意见为专家辨证结果作为金标准,依据诊断性试验原理,计算标准辨证相对于专家辨证的灵敏度和特异度,各证型阈值辨证结果均有较好的灵敏度和特异度(相关表格略)。

3 讨论

目前科学研究正呈现出注重非线性和复杂系统研

究的趋向,对于中医证候这一复杂系统也需从研究方法上实现转变^[5]。基于熵的复杂系统分划方法是信息论中熵方法和熵语言在非线形相关模式识别领域的具体应用,该方法遵照数据的内在联系,不对数据作刚性先行分割,依据数据内在关联进行自主聚类,可以无监督地处理多变量、多层次上的复杂数据,对于提取中医证候相关症状并分析症状之间的非线性相关关系具有重要应用价值。

在中医证候相关研究中,熵分划方法可以针对不同对象的数据运算得到不同研究结果。对于本研究给出的冠心病心绞痛病例数据,应用熵分划方法在不同的参数调整下对症状之间进行自由聚类,通过症状熵值、症状间联合熵值推算出症状间关联度数值,进而得到大量的症状集合群,结合医理分析,提取出具有中医理论意义的症状集合(基本证型)。如将基本证候视为广义子系统,通过熵的运算还可进一步推算基本证型之间的关联度,反映出证型之间的组合规律,对于探讨疾病中复合证型及病机演变规律具有重要意义。如以该方法分析基本证型中各个症状与所在症状集合的关联度,则可获得症状对于所在症状集合(证候)的贡献度,反映症状的重要程度。通过结果分析,基于熵的复杂系统分划方法可实现无监督的中医证候相关症状提取及症状贡献度运算,其分析结果较之处理线性相关的多元统计方法更具说服力。

以基于熵的复杂系统分划方法对临床数据分析结果显示,其基本证候包括气滞、气虚、痰湿、血瘀、阴虚、阳虚、痰热、郁热等,与目前文献学研究结果基本一致^[6]。以基本证候之间的关联度分析可见,气滞与痰湿、血瘀、郁热、痰热等具有较强的关联度,气虚与痰湿、血瘀、痰热、阳虚、阴虚具有较强关联度。这些证候与证候之间的关联度不仅显示了临床实际可能出现的复合证型及冠心病心绞痛本虚标实的病机特点,而且揭示了冠心病心绞痛可分为以气滞和气虚为基础证型的两大类病机变化,即气滞心络瘀阻和气虚心络瘀阻,随病变发展,可出现兼夹痰湿、血瘀、郁热、痰热、阴虚、阳虚等病机转趋。同时证候之间关联度分析还显示,痰湿与阳虚、痰湿与痰热也存在关联度,也反映了阳虚生(寒)痰、痰郁化热的病机变化规律。

证候是不同症状依其内在规律有机组合而成,不同症状对证候具有不同的特异性贡献,这也是产生症状之间非线性作用及复杂关系的原因。为充分反映症状之间复杂的相互作用,基于熵的复杂系统分划方法以症状对症状集合(证候)的关联度,反映症状对症状集合的贡献度,如气滞各症状按其原始贡献度排序为

善太息、心胸胀痛、情志抑郁、急躁易怒、脉弦,气虚各症状原始贡献度排序为乏力、神疲、气短、懒言、心胸隐痛、脉弱、舌淡,对气虚具有较高诊断价值的乏力、神疲、气短、懒言等症状的贡献度均显著高于其他症状。其他证型症状贡献度也均基本符合临床实际。在实际临床运用中,可对原始症状贡献度分别乘以 10 的整数倍并取个位整数,既保持了症状之间贡献度的原始比例,又便于临床操作。

利用阈值诊断可最大程度避免诊断中的主观因素,为证候量化诊断提供更为科学的解决方法。建立阈值的方法除早期日本寺泽捷年建立血瘀证诊断阈值所用的逐步线性回归分析^[7]外,近年又有最大似然判别法^[8]和参考值估计法^[9]。客观评价诊断方法是否合理的指标为诊断性试验中的灵敏度和特异度,诊断性试验中 ROC 曲线的建立需参考不同移动诊断界点的灵敏度和特异度。

本研究中各基本证型诊断阈值的建立即利用 ROC 曲线分析,考察不同诊断界点灵敏度和特异度指标,参考文献^[10]并结合医理,以两者取值最大时的诊断界点作为该证型的诊断阈值,经 104 例患者的前瞻性检验结果,所取阈值均具有良好的灵敏度和特异度。

参考文献

- [1]袁世宏,张连文,王米渠,等.数理统计思想及方法在辨证规范化研究中的应用及思考.中国中医基础医学杂志,2003,9(4):15~21.
- [2]梁茂新,洪治平.中医症状量化的方法初探:附虚证 30 症的量化法.中国医药学报,1994,9(3):37~39.
- [3]严石林,李炜弘,王米渠,等.寒证辨证因子等级量化操作标准的研究.中国中医药信息杂志,2002,9(8):64~66.
- [4]李梢,王永炎,季梁,等.复杂系统意义下的中医学及其案例研究.系统仿真学报,2002,14(11):1429~1442.
- [5]袁敬柏,董绍英,王阶,等.2689 例冠心病心绞痛证候要素的文献统计分析.中国中医药信息杂志,2006,13(5):100~101.
- [6]寺泽捷年(陈可冀等译).瘀血证的证候分析与诊断标准刍议.国外医学中医中药分册,1984,6(1):1.
- [7]邱向红.脾虚证计量诊断的探讨.广州中医学院学报,1990,7(1):24~27.
- [8]潘毅,徐志伟,严灿,等.频数优势法对心理应激人群中四个中医常见证型的计量诊断.江西中医学院学报,2006,18(4):63~65.
- [9]金丕焕.医用统计方法.2 版.上海:复旦大学出版社,2004:428~431.
- [10]郭蕾.关于证候概念的诠释.北京中医药大学学报,2003,26(2):5~8.

(收稿日期:2007-03-08)

[编辑:黄健]