

证的研究 ·

基于关联规则的2型糖尿病方证关系的横断面研究

龚燕冰¹ 罗增刚² 高思华³ 倪青¹ 易丹辉⁴ 谢雁鸣⁵ 王永炎⁵

(1. 中国中医科学院广安门医院, 北京市宣武区北线阁5号, 100053; 2. 中国中医科学院院长办公室;
3. 北京中医药大学; 4. 中国人民大学统计学院; 5. 中国中医科学院中医临床基础医学研究所)

[摘要] **目的** 探索2型糖尿病的方-症-证三者之间的关系。**方法** 以9家医院的2501例回顾性2型糖尿病病例的临床数据为基础, 运用关联规则的方法分析方-症-证之间的关联关系。**结果** 从122首方剂中选取生脉饮、六味地黄汤、四君子汤等3首方剂分析用药患者症状间的关系, 服用生脉饮的患者倦怠乏力、口干口渴、尿频尿多、肢体麻木、视物模糊、口渴欲饮、夜尿频多、气短懒言等症状出现的频率较高, 服用六味地黄汤的患者倦怠乏力、口渴欲饮、视物模糊、夜尿频多、口干口渴、肢体麻木、气短懒言、腰膝酸软等症状出现的频率较高, 且这些症状之间存在多重相关的关联关系。服用四君子汤的患者中, 倦怠乏力、腰膝酸软、夜尿频多、气短懒言、肢体麻木、失眠、视物模糊等症状出现的频率较高, 且这些症状之间存在多重相关的关联关系。**结论** 关联规则可以分析2型糖尿病的方证关系, 能探索症状间的多重相关关系。

[关键词] 2型糖尿病; 证候; 关联规则; 症状; 中医药疗法

Cross-sectional Study on the Relationship between Formulas and Syndromes of Type 2 Diabetes Based on Association Rules

GONG Yanbing¹, LUO Zenggang², GAO Sihua³, NI Qing¹, YI Danhui⁴, XIE Yanming⁵, WANG Yongyan⁵

(1. Guang'anmen Hospital, China Academy of Chinese Medical Sciences; 2. China Academy of Chinese Medical Sciences; 3. Beijing University of Chinese Medicine; 4. Renmin University of China; 5. Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences)

ABSTRACT **Objective** To explore the relationship among the disease - syndrome - formula of type 2 diabetes. **Methods** Based on the clinical data of 2501 cases of type 2 diabetes, to study the relationship among the disease - syndrome - formula of type 2 diabetes by making use of association rules. **Results** Three formulas, Shengmai Yin (Pulse-activating Fluid), Liuwei Dihuang Decoction (Decoction of Six Herbs Including Rehmannia), and Sijunzi Decoction (Decoction of Four Noble Herbs), were selected from 122 ones to analyze the relationship between the formula and symptoms. The occurrence of fatigue with weakness, dry mouth with thirst, frequent urination with increased volume of urine, numbness of limbs, blurred vision, thirst with preference of drinking water, frequent nocturnal urination, and shortness of breath with inactive talking was in high frequency in those patients who were prescribed Shengmai Yin. The occurrence of fatigue with weakness, thirst with preference of drinking water, blurred vision, frequent nocturnal urination, dry mouth with thirst, numbness of limbs, shortness of breath with inactive talking, and soreness and weakness of lumbus and legs was in high frequency in those patients prescribed Liuwei Dihuang Decoction, and there existed a multi-association among these symptoms. The occurrence of fatigue with weakness, soreness and weakness of lumbus and legs, frequent nocturnal urination, shortness of breath with inactive talking, numbness of limbs, insomnia, and blurred vision was in high frequency in those patients who were prescribed Sijunzi Decoction, and there existed a multi-association among these symptoms. **Conclusion** The association rule can be adopted to analyze the relationship between formula and syndrome, and to explore the multi-association relationship among the symptoms of type 2 diabetes.

Key Words Type 2 Diabetes; TCM syndromes; Association rule; Symptoms; TCM therapy

方证关系是近年来证候研究的重要组成部分, 在本研究中我们拟采用横断面研究与纵向研究相结合的办法, 在时空演变中探讨方-症-证三者之间关系。关联规则(association rules)^[1]是数据挖掘中最活跃的研究方法之一, 能发现数据中存在的隐含知识。故本研究采用关联规则研究2型糖尿病方证关系。

1 临床资料

1.1 病例来源

病例选取自2003年6月~2005年2月中国中医科学院广安门医院、北京中医药大学东直门医院、北京中医药大学东方医院、山东中医药大学附属医院、济南市中

医医院、潍坊市中医院、保定市中医院、河南中医学院第一附属医院、新疆中医院等医院内分泌科的2型糖尿病患者2501例。

1.2 诊断标准

西医诊断标准:符合1999年世界卫生组织“糖尿病及其他类型高血糖的诊断标准”;中医诊断标准:符合《中药新药临床研究指导原则》的有关诊断标准^[2]。

2 方法

2.1 病例数据库的建立

将2501份原始病例资料采用双人双机录入Epidata3.0软件,核查无误后导出为Excel表格文件。数据库包括患者的入院信息、体格检查、理化检查、症状、证型、中西药运用、出院诊断等7个模块。数据库中不含有文字,将每个变量对应唯一的编码,并通过填充空缺值、纠正非法值、数据变换等措施进行严格的质量控制。

2.2 关联规则计算

本研究中关联规则应用SAS8.2版软件运算。按照患者服用的不同汤剂,将其分为若干群体,将每种群体的症状做关联分析,得出症状之间、症-证之间、方-症-证之间的关系。本研究取最小支持度(8%),最小置信度(50%),获取症状间的关联关系。

2.3 网络图的构建

网络图是一种描述不同类别事物相互关联程度的图形,设定将出现频数大于10的症状纳入分析。采用Statistica 7.0统计软件运行关联规则得到网络图。在网络图中,每个单元的方块的大小表示该种症状出现的比例,方块越大表示比例越高。同时,若有直线连接两个单元,则说明两种单元有共同出现的概率,线越粗表示共同出现的概率越高。线的颜色越深表示两种症状相互关联的作用越明显。

3 结果

数据库中包含122首中药汤剂,选取中药汤剂中出现频数最多的3种汤剂(生脉饮、六味地黄汤、四君子汤)探讨用药患者的症状间的关联关系。

3.1 生脉饮用药患者症-症关系、方-症-证关系

所得症状间的关联规则结果见表1。提升度表示第一个变量同时伴有第二个变量出现的可能性是第一个变量单独出现的可能性的倍数。如第一条提升度为1.24,表示“尿频尿多”的患者中具有症状“倦怠乏力”的可能性是单独出现“尿频尿多”可能性的1.24倍。

支持度表示两个变量同时出现的样本数占总样本数的百分率。如第一条支持度为18.83%,表示“尿频尿多”和“倦怠乏力”同时出现的患者占全体人群的18.83%。

表1 服用生脉饮患者症状间的关联关系

关联数目	提升度	支持度(%)	置信度(%)	项数频数统计比	结果
2	1.24	18.83	60.42	29	尿频尿多→倦怠乏力
2	1.15	18.83	55.77	29	口干口渴→倦怠乏力
2	1.93	16.88	65	26	口渴欲饮→口干口渴
2	1.93	16.88	50	20	口干口渴→口渴欲饮
2	1.05	14.94	51.11	23	肢体麻木→倦怠乏力
2	1.07	14.94	52.27	23	视物模糊→倦怠乏力
2	1.67	14.29	81.48	22	气短懒言→倦怠乏力
2	1.08	13.64	52.5	21	口渴欲饮→倦怠乏力
2	1.35	12.34	65.52	19	夜尿频多→倦怠乏力
2	1.99	11.69	62.07	18	夜尿频多→尿频尿多
2	1.34	11.04	65.38	17	腰膝酸软→倦怠乏力
2	1.85	9.74	57.69	15	腰膝酸软→尿频尿多
2	1.53	9.74	51.72	15	夜尿频多→口干口渴
2	1.44	9.09	70	14	健忘→倦怠乏力
2	2.85	8.44	50	13	腰膝酸软→气短懒言
2	1.48	8.44	72.22	13	心慌→倦怠乏力
3	1.03	8.44	50	13	口渴欲饮 & 口干口渴→倦怠乏力
3	1.83	8.44	61.9	13	口渴欲饮 & 倦怠乏力→口干口渴

置信度表示在第一个变量中含有第二个变量的百分数。如第一条置信度为60.42%,表示所有“尿频尿多”的患者中有60.42%出现“倦怠乏力”。

项数频数统计比表示同时出现两个变量的样本数。如第一条有29个,表示所有人群中有29个患者同时出现者两个症状。

在服用生脉饮的患者有两个症状的关联和三个症状之间的关联。其中,两个症状的关联中:尿频尿多和倦怠乏力两种症状同时出现的概率是18.83%,有尿频尿多症状的患者,出现倦怠乏力症状的概率是60.42%;口干口渴和倦怠乏力两种症状同时出现的概率是18.83%,有口干口渴症状的患者,出现倦怠乏力症状的概率是55.77%;依次类推。三个症状的关联中:口渴欲饮、口干口渴和倦怠乏力同时出现的概率是8.44%,有口渴欲饮和口干口渴症状的患者出现倦怠乏力的概率是50%;口渴欲饮、倦怠乏力和口干口渴同时出现的概率是8.44%,有口渴欲饮和倦怠乏力症状的患者出现口干口渴的概率是61.9%。提示这些症状间存在较强的相关性。

图1示,服用生脉饮的患者倦怠乏力、口干口渴、尿频尿多、肢体麻木、视物模糊、口渴欲饮、夜尿频多,气短懒言等症状出现的频率较高,且这几个症状之间存在多重相关的关联关系。生脉饮是治疗气阴两虚的主要方剂,在上述以气阴两虚为主的症状群中也有较好的体现。

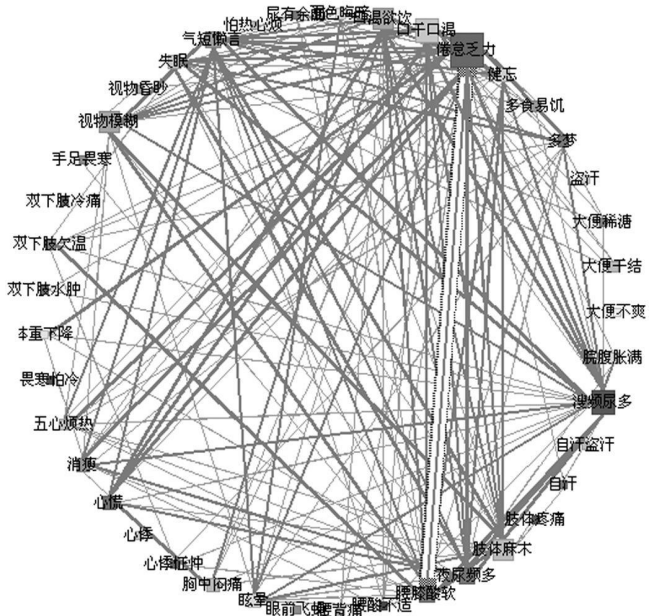


图1 生脉饮用药患者的症状网络图

3.2 六味地黄汤用药患者症状间的关联关系

图2示,服用六味地黄汤的患者倦怠乏力、口渴欲饮、视物模糊、夜尿频多,口干口渴、肢体麻木、气短懒言、腰膝酸软等症出现的频率较高,且这些个症状之间存在多重相关的关联关系。六味地黄汤是治疗肾阴虚的主要方剂,在上述以肾阴虚为主的症状群中也有较好的体现。

表2 服用六味地黄汤患者症状间的关联关系

关联数目	提升度	支持度 (%)	置信度 (%)	项数频数统计比	结果
2	2.56	16.42	61.11	11	口渴欲饮→口干口渴
2	2.56	16.42	68.75	11	口干口渴→口渴欲饮
2	1.87	11.94	61.54	8	气短懒言→倦怠乏力
2	2.25	10.45	53.85	7	气短懒言→夜尿频多
2	2.51	8.96	60	6	失眠→夜尿频多
2	5.74	8.96	60	6	失眠→多梦
2	5.74	8.96	85.71	6	多梦→失眠
3	2.61	8.96	85.71	6	夜尿频多 & 气短懒言→倦怠乏力
3	4.42	8.96	85.71	6	夜尿频多 & 倦怠乏力→气短懒言
3	3.14	8.96	75	6	气短懒言 & 倦怠乏力→夜尿频多

表2示,服用六味地黄汤的患者有两个症状的关联和三个症状之间的关联。其中,两个症状的关联中:口渴欲饮和口干口渴两种症状同时出现的概率是16.42%,有口渴欲饮症状的患者,出现口干口渴症状的概率是61.11%;口干口渴和口渴欲饮两种症状同时出现的概率是16.42%,有口干口渴症状的患者,出现口渴欲饮症状的概率是68.75%;依次类推。三个症状的关联中,夜尿频多、气短懒言和倦怠乏力同时出现的

概率是8.96%,有夜尿频多和气短懒言症状的患者出现倦怠乏力的概率是85.71%;夜尿频多、倦怠乏力和气短懒言同时出现的概率是8.96%,有夜尿频多和倦怠乏力症状的患者出现气短懒言的概率是85.71%,依次类推。

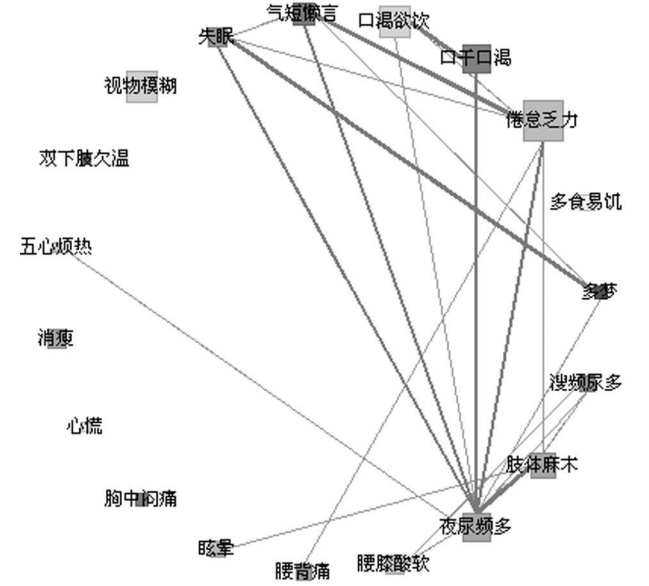


图2 六味地黄汤用药患者的症状网络图

3.3 四君子汤用药患者症状间的关联关系

表3示,服用四君子汤的患者有两个症状的关联。其中,腰膝酸软和倦怠乏力两种症状同时出现的概率是24.62%,有腰膝酸软症状的患者,出现倦怠乏力症状的概率是69.57%;倦怠乏力和腰膝酸软两种症状同时出现的概率是24.62%,有倦怠乏力症状的患者,出现腰膝酸软症状的概率是51.61%;依次类推。

表3 服用四君子汤患者症状间的关联关系

关联数目	提升度	支持度 (%)	置信度 (%)	项数频数统计比	结果
2	1.46	24.62	69.57	16	腰膝酸软→倦怠乏力
2	1.46	24.62	51.61	16	倦怠乏力→腰膝酸软
2	1.54	21.54	73.68	14	夜尿频多→倦怠乏力
2	1.54	21.54	45.16	14	倦怠乏力→夜尿频多
2	1.48	18.46	70.59	12	气短懒言→倦怠乏力
2	1.48	18.46	38.71	12	倦怠乏力→气短懒言
2	1.44	16.92	68.75	11	失眠→倦怠乏力
2	1.44	16.92	35.48	11	倦怠乏力→失眠
2	1.49	15.38	52.63	10	夜尿频多→腰膝酸软
2	1.49	15.38	43.48	10	腰膝酸软→夜尿频多
2	2.01	15.38	52.63	10	夜尿频多→气短懒言
2	2.01	15.38	58.82	10	气短懒言→夜尿频多
2	2.39	15.38	62.50	10	失眠→气短懒言
2	2.39	15.38	58.82	10	气短懒言→失眠

图3示,服用四君子汤的患者倦怠乏力、腰膝酸软、夜尿频多、气短懒言、肢体麻木、失眠、视物模糊等症出现的频率较高,且这些个症状之间存在多重相关的

关联关系。四君子汤是治疗气虚的主要方剂,在上述以气虚为主的症状群中也有较好的体现。

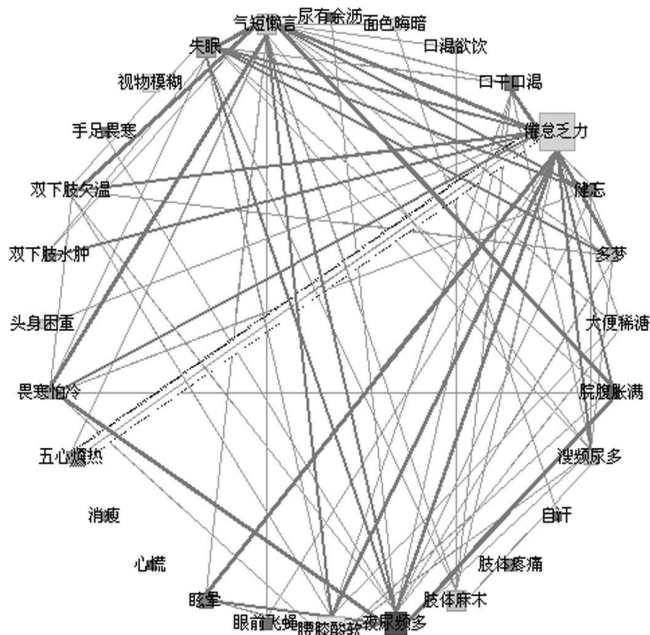


图3 四君子汤用药患者的症状网络图

4 讨论

关联规则能发现数据中存在的隐含知识,能找出多维数据中隐藏的关联规则,有时数据中的关联函数未知,或即使已知也不确定,因此,关联分析生成的规则带有可信度。为了在建模过程中提高模型在实际应用中的准确性,通常我们用最小支持度、最小置信度和兴趣度来衡量关联规则,只有支持度和置信度分别大于指定的最小值的关联规则才是符合要求的关联规则模型。

在关联规则中,我们得出了使用某种方剂的人群

的症状的分布规律,从几种症状同时出现的可能性的 大小,可以判定哪些症状是经常相伴出现的,对于同时出现的可能性较大的几个症状组成的症状群往往提示着某种证型的出现,这对于方-症-证-证之间的关系 的探讨也提供了一种有效的方法。同时对于临床认识也 有一定的帮助。比如在使用生脉饮的患者中,三项关联 的结果显示,出现口渴欲饮、口干口渴症状的患者出现 倦怠乏力可能性达50%,出现口渴欲饮、倦怠乏力症状 的患者出现口干口渴可能性达61.9%。说明这三个症 状在2型糖尿病中经常是结伴出现的,他们所体现出的 气阴两虚的表现,正是生脉饮治疗的主要方向。反之, 生脉饮治疗的气阴两虚的主要症状也正包含口渴欲 饮、倦怠乏力和口干口渴。这说明只有在方证对应的 情况下,才能体现较好的临床疗效。通过统计学方法得出 的结论基本符合中医理论和临床实际,这对于方-症- 证-证关系的探讨也提供了一种有效的方法。本文仅从 横断面研究了方-症-证-证的关系,同时我们也用了线 性混合模型的方法探索了纵向数据所反映的方-症- 证-证关系,两种方法的联合使用起到了取长补短的作用。 也说明了运用单一的统计学方法对数据挖掘具有一定的 局限性,多种统计方法的联用才能做到较为全面准 确。

参考文献

- [1]王华,胡学钢.基于关联规则的数据挖掘在临床上的应用[J].安徽大学学报(自然科学版),2006,30(2):21-25.
- [2]郑筱萸.中药新药临床研究指导原则[M].北京:中国医药科技出版社,2002:215.

(收稿日期:2009-03-23;修回日期:2009-08-30)

[编辑:行 之]

欢迎邮购《中医杂志》“专题笔谈”文萃》

备受广大读者关注的《中医杂志》“专题笔谈”文萃一书已经由人民卫生出版社出版。本书收集了《中医杂志》“专题笔谈”栏目1995~2004年10年间关于单味中药单独应用或以单味中药为主配合其他药物治疗各类疾病的临床治疗经验共900余篇文萃。

“专题笔谈”是《中医杂志》特色栏目之一,汇集了全国各地中医师丰富的临床实践经验,其中不乏对一些疑难病症的独特治法,也包括了一些常用中药的非常规用法,深受广大医务人员和读者的欢迎,对于临床医生临证用药具有较高的参考价值。全书共分为两辑,定价83元,需要者可通过本刊编辑部邮购。为答谢广大读者长期以来对本刊的支持和厚爱,通过编辑部邮购本书者免邮寄费。按照著作权法有关规定,我们将向所刊载文章的作者寄赠样书1册作为全部稿酬的变通支付方式。由于年代间隔较长,不少作者地址变更,为了便于此项工作顺利,请1995~2004年期间在《中医杂志》专题笔谈栏目发表文章的作者及时与编辑部联系,并说明目前的详细联系方式,我们将在收到来信后以挂号方式寄赠样书。

联系地址:北京市东直门南小街16号《中医杂志》编辑部,邮编:100700,电话:(010) 64014411转3040,64013162,联系人:崔京艳、邓媛。

购书者请在汇款单附言栏注明所购书名及联系电话。

欢迎邮购《中医杂志》

1997年2、4、6、8~12期,每册3.50元;1998年3~12期,每册4元,合订本65元;1999年4~12期,每册4元,合订本65元;2000年2~12期,每册5.80元,合订本85元;2001年2~12期,每册6.50元,合订本95元;2002年1~12期,每册6.80元,合订本105元;2003年1~12期,每册6.80元,合订本105元;2004年1、2、4~12期,每册6.80元,合订本115元;2005年1~7、9、11、12期,每册6.80元,合订本125元;2006年1~12期,每册7.30元,合订本135元;2007年2~12期,每册8.80元,合订本140元;2008年2~12期,每册9.80元,合订本140元;2009年每册9.80元,合订本140元。《中医杂志》2006~2008年增刊每册40元。以上免费邮寄,如需挂号每件另加3元。

《中医杂志》光盘合订本:1951~1986年(4张光盘)480元;1995~2004年(1张光盘)280元。光盘合订本为计算机用数据光盘,可对所有文章进行任意字检索,方便实用,欢迎邮购。http://www.cnki.net

欲购者请汇款至北京东直门内南小街16号中医杂志社读者服务部,邮编:100700,电话:(010) 64014411转3036。