

本文引用:黄亮,余娜,晏峻峰.基于关联规则和复杂系统熵聚类的国医大师熊继柏治疗脾胃病证用药规律研究[J].湖南中医药大学学报,2018,38(12):1412-1415.

基于关联规则和复杂系统熵聚类的国医大师 熊继柏治疗脾胃病证用药规律研究

黄亮,余娜*,晏峻峰*

(湖南中医药大学,湖南长沙 410208)

〔摘要〕目的 研究熊继柏教授治疗脾胃病证的用药规律及组方规律。方法 收集录入熊继柏教授治疗脾胃病证的医案处方数据,采用 Apriori 关联规则算法,改进的复杂系统熵聚类算法等无监督数据挖掘算法,确定处方中各种药物的使用频次以及药物之间的关联规则,分析熊继柏教授治疗脾胃病证的用药经验以及组方规律。结果 对脾胃病证筛选出的 29 个医案,共 71 首处方、108 味中药进行分析,使用频次较高的中药是甘草、陈皮、砂仁、茯苓、木香、法半夏、厚朴、枳实;使用频次较高的药对是“陈皮-甘草”“砂仁-甘草”“茯苓-甘草”;挖掘出熊继柏教授治疗脾胃病的 4 个核心组合及 1 首新处方。结论 熊继柏教授治疗脾胃病证多用醒脾开胃,行气化湿之药品。

〔关键词〕 脾胃病证;熊继柏;关联规则;复杂系统熵聚类

〔中图分类号〕 R256.3

〔文献标志码〕 A

〔文章编号〕 doi:10.3969/j.issn.1674-070X.2018.12.013

A Study of Chinese Medicine Master Xiong Jibo's Medication Rule in Prescriptions for Spleen and Stomach Diseases Based on Association Rules and Entropy Clustering for Complex Systems

HUANG Liang, YU Na*, YAN Junfeng*

(Hunan University of Chinese Medicine, Changsha, Hunan 410208, China)

〔Abstract〕 Objective To study Professor Xiong Jibo's medication and prescription rules for spleen and stomach diseases. **Methods** The medical records and prescription data of Professor Xiong Jibo in the treatment of spleen and stomach diseases were collected and enrolled in this study. Two unsupervised data mining algorithms, Apriori association rule algorithm and improved entropy clustering algorithm for complex systems, were used to determine the usage frequency of medicines and association rules between medicines in prescriptions. The medication experiences and prescription rules of Professor Xiong Jibo were analyzed in the treatment of spleen and stomach diseases. **Results** A total of 29 medical records, including 71 prescriptions and 108 traditional Chinese medicines, for spleen and stomach diseases were analyzed. The most frequently used medicines were Radix Glycyrrhizae, tangerine peel, *Amomum villosum*, *Poria cocos*, *Saussurea costus*, Rhizoma Pinelliae Praeparatum, *Magnolia officinalis*, and Fructus Aurantii Immaturus. The most frequently used medicine pairs were tangerine peel-Radix Glycyrrhizae, *Amomum villosum*-Radix Glycyrrhizae, and *Poria cocos*-Radix Glycyrrhizae. Four core combinations and one new prescription were identified in Professor Xiong Jibo's medical records for spleen and stomach diseases. **Conclusion** To treat spleen and stomach diseases, Professor Xiong Jibo often uses medicines that arouse the spleen, open the stomach, move Qi, and remove dampness.

〔Keywords〕 spleen and stomach diseases; Xiong Jibo; association rule; entropy clustering for complex systems

〔收稿日期〕 2018-04-19

〔基金项目〕 国家中医药管理局“国医大师熊继柏传承工作室建设项目”资助。

〔作者简介〕 黄亮,男,学士,主要从事中医药信息学研究。

〔通讯作者〕 *余娜,女,讲师,E-mail:49488173@qq.com;晏峻峰,女,教授,博士研究生导师,E-mail:381155346@qq.com。

国医大师熊继柏擅长内科病证诊疗。本研究收集、整理熊继柏教授治疗脾胃病证处方数据,采用数据挖掘方法探析熊继柏教授治疗脾胃病证的用药经验。脾胃病证是中医内科病证的重要组成部分,目前国内已有学者采用统计分析和数据挖掘等技术对名医流派、经典古籍治疗脾胃病的用药规律开展相关研究工作^[1-8]。刘文琛等^[9]对熊继柏教授治疗失眠的 254 例病案进行整理和分析,挖掘熊继柏教授治疗失眠常用方剂组合和药物。但是,目前尚未有采用数据挖掘技术分析熊继柏教授治疗脾胃病证用药规律的研究。本研究采用数据挖掘中的关联分析和复杂系统熵聚类等方法对熊继柏教授治疗脾胃病证处方数据进行分析 and 挖掘,探寻熊继柏教授的用药规律和诊疗经验,以期更好地推广和传承熊继柏教授的用药特点和学术思想。

1 资料与方法

1.1 处方来源

本研究以 2009 年熊继柏学术思想与临证经验研究小组整理出的《一名真正的名中医——熊继柏临证医案实录 1》^[10]中的脾胃病证处方作为数据来源并作出适当的筛选。

1.2 处方的录入与核对

将上述筛选出的处方数据录入 Excel 表格中,并对数据进行审核,以确保数据的准确性。然后将数据导入数据库中以便研究。

1.3 处方数据预处理

为了保证处方中药物数据用词的标准化以及准确性,对处方数据进行了术语与数据名词的统一,将药名进行标准化处理,如药物名称的规范,同一味药,名字如果有区别,则将药名统一化;或者是炮制品的规范,同一味药,但是炮制方法不同,也将其看作是同一味药,并统一药名。使用数据库中的 SQL 语句进行统计。药名规范标准按照统编教材《中药学》^[11]进行规范。

1.4 数据分析

1.4.1 频次统计分析 将脾胃病证处方中每味药的出现频次统计并从大到小排序,将结果导出。

1.4.2 关联规则分析 关联规则算法也被称为 Apriori 算法,是数据挖掘中的一类重要算法。该算

法将发现关联规则的过程分为两个步骤,第一步通过“支持度”的设置,迭代检索出事务数据库中的所有频繁项集;第二步利用频繁项集构造出满足用户最小置信度的规则。

设置“支持度”(表示在所有药物中同时出现的次数)为 25 和“置信度”(表示药物 A 在包含药物 B 的事物中出现的频繁程度)为 0.8,按照药物组合出现频次从大到小的顺序进行排序,作为高频药对;按照在支持度范围内,大于或者等于置信度的关联规则从大到小排序。

1.4.3 复杂系统熵聚类分析 首先在复杂系统 S 中^[12-13],设置 n_a 为单个药物在处方中出现的次数, q 为处方数,计算单个药物在处方中出现的概率,对每个药物进行单个熵的计算(如公式 1 所示)。

$$H(X_i) = - \sum_{a=1}^k (n_a/q) \lg(n_a/q) \quad (\text{公式 1})$$

再对药对计算联合熵(如公式 2 所示)。

$$H(X_i, X_j) = - \sum_a \sum_b (n_{ab}/q) \lg(n_{ab}/q) \quad (\text{公式 2})$$

根据联合熵与单个熵计算关联度(如公式 3 所示)。

$$\mu(X_i, X_j) = H(X_i) + H(X_j) - H(X_i, X_j) \quad (\text{公式 3})$$

在计算关联度系数时(如公式 4 所示),由于公式 4 中关联度系数 $\mu(i, j) \neq \mu(j, i)$ (变量 X_i 和 X_j 之间相互依存的程度不等于变量 X_i 和 X_j 之间相互依存的程度)不利于计算变量之间的相关系数和对复杂系统进行分类。所以根据赵亚丽^[14]对该算法的分析,在关联度系数公式的基础上进行改进,改进后如公式 5 所示。

$$\mu(i, j) = \frac{\mu(X_i, X_j)}{H(X_j)} \quad (\text{公式 4})$$

$$\mu(i, j) = \frac{\mu(X_i, X_j)}{\sqrt{H(X_i)} \sqrt{H(X_j)}} \quad (\text{公式 5})$$

通过计算关联度系数,将每味药所对应的关联药物在关联度系数中排名前十的药物进行统计并导出,再使用聚类算法得到新的药物处方组合。

2 结果

2.1 熊继柏教授治疗脾胃病证的用药频次统计

对脾胃病证 29 个有效医案,71 首处方中的数据使用 Java 以及 Oracle 软件,通过整理以及经过相关的统计分析,得出在脾胃病证中共使用 108 味中药,频次总数为 787 次,平均每张处方只用 11.1 味

中药,使用频次排名前十位分别是甘草、陈皮、砂仁、茯苓、木香、法半夏、厚朴、枳实、白术、白芍。见表 1。

表 1 熊继柏教授治疗脾胃病证频次 ≥ 10 的高频单味药

序号	药名	频次
1	甘草	62
2	陈皮	48
3	砂仁	37
4	茯苓	34
5	木香	31
6	法半夏	31
7	厚朴	27
8	枳实	26
9	白术	23
10	白芍	19

2.2 熊继柏教授治疗脾胃病证药对频次统计

对脾胃总处方中药对的出现频次进行排序。前三位分别为“陈皮、甘草”“砂仁、甘草”“茯苓、甘草”。设置频次大于等于 25,见表 2。

表 2 脾胃病证处方中高频药对情况表(频次 ≥ 25)

序号	药名	频次
1	陈皮、甘草	42
2	砂仁、甘草	34
3	茯苓、甘草	34
4	甘草、法半夏	31
5	木香、甘草	30
6	砂仁、陈皮	30
7	法半夏、茯苓	27
8	茯苓、陈皮	25
9	法半夏、陈皮	25

2.3 基于关联规则的用药规律分析

对脾胃病证中二项关联规则,设置支持度为 25,置信度为 0.8,得出高频药物二项关联规则,前三位分别为“法半夏、甘草”“茯苓、甘草”“木香、甘草”,见表 3。相关药物关联规则网络图见图 1。

设置最小支持度阈值为 25,最小置信度阈值为 0.8,获取脾胃病证处方中药物的三项关联规则,前三位分别为“法半夏、茯苓、甘草”“茯苓、陈皮、甘草”“法半夏、陈皮、甘草”,如表 4 所示。

2.4 基于复杂系统熵聚类的组方规律分析

通过对脾胃病证进行关联度系数的计算,现设置关联度系数不小于 0.6,前三位分别为“厚朴、木香”“苍术、厚朴”“厚朴、苍术”,药对列表排名如表 5 所示。

表 3 脾胃病证处方中高频药物二项关联表
(支持度 25,置信度 0.8)

序号	关联规则	置信度
1	法半夏→甘草	1
2	茯苓→甘草	1
3	木香→甘草	0.96
4	砂仁→甘草	0.91
5	陈皮→甘草	0.88
6	法半夏→茯苓	0.87
7	砂仁→陈皮	0.81
8	法半夏→陈皮	0.81

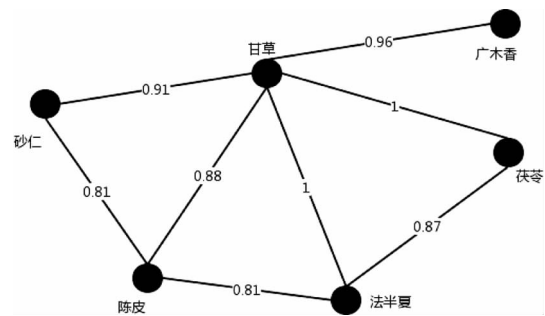


图 1 脾胃病证药物二项关联规则网络图
(支持度 25,置信度 0.8)

表 4 脾胃病证处方中高频药物三项关联表(支持度 25,置信度 0.8)

序号	关联规则	置信度
1	法半夏、茯苓→甘草	1
2	茯苓、陈皮→甘草	1
3	法半夏、陈皮→甘草	1
4	砂仁、陈皮→甘草	0.9
5	法半夏→甘草、茯苓	0.87
6	甘草、法半夏→茯苓	0.87
7	法半夏→陈皮、甘草	0.81
8	法半夏、甘草→陈皮	0.81

表 5 脾胃处方基于改进的关联度系数药物排名表
(关联度系数 ≥ 0.6)

中药	关联的中药	关联度系数
厚朴	木香	0.90
苍术	厚朴	0.89
厚朴	苍术	0.89
苍术	木香	0.84
木香	厚朴	0.79
厚朴	砂仁	0.75
砂仁	厚朴	0.75
砂仁	木香	0.74
陈皮	砂仁	0.72
木香	苍术	0.72
苍术	砂仁	0.68
砂仁	苍术	0.68
陈皮	厚朴	0.61
厚朴	陈皮	0.61

以药物之间的关联度分析结果为基础,基于复杂系统熵聚类的方法,求出 3~8 味药的核心组合,挖掘出治疗脾胃病证新处方:苍术-陈皮-厚朴-砂仁。见表 6。

表 6 脾胃处方基于复杂系统熵聚类的核心药物组合表

编号	候选处方			
1	苍术	陈皮	厚朴	砂仁
2	木香	苍术		
3	厚朴	木香		
4	砂仁	木香		

3 讨论

本研究使用 Java 开发平台以及 Oracle 数据库软件,运用关联规则算法以及改进后的复杂系统熵聚类算法分析熊继柏教授治疗脾胃病证的用药经验,经过关联规则算法分析,熊继柏教授治疗脾胃病证常用的药物包括甘草、陈皮、砂仁、茯苓、木香、法半夏、厚朴、枳实、白术、白芍等。出现频次较高的药对有“陈皮、甘草”“砂仁、甘草”“茯苓、甘草”“甘草、法半夏”等。基于关联规则的高频药物二项关联规则中,使用法半夏就一定会使用甘草,使用茯苓就一定会使用甘草。基于关联规则的高频药物三项关联规则中结果显示,使用法半夏和茯苓的前提下一定会使用甘草,使用茯苓、陈皮的前提下一定使用甘草,使用法半夏、陈皮的前提下一定使用甘草等。基于改进的互信息法的关联度较大的药物组合有:厚朴-木香,厚朴-苍术,苍术-厚朴,苍术-木香等。基于熵层次聚类的治疗脾胃病新处方有:苍术-陈皮-厚朴-砂仁。

甘草是使用频次最高的药物,甘草的作用是能够调和诸药,入脾胃,益中气,并有解毒作用,和其他药配伍以后可渐进性缓解比较急重的症状^[3];陈皮归脾、肺经,功能理气健脾,燥湿化痰;排名第三的砂仁性温,味辛,可以用来行气调中、和胃醒脾,治疗湿浊中阻,腹痛痞胀,胃呆食滞,呕吐泻泄等症。陈皮和法半夏合用有燥湿化痰、健脾和胃、理气止呕的作

用;陈皮与砂仁合用能够行气导滞、开胃醒脾;茯苓与法半夏配伍具有燥湿化痰降逆之功效。厚朴和苍术都属于化湿药,具有燥湿的功效,可以治疗湿阻中焦之证;厚朴和木香都可以治疗胃病;陈皮和砂仁合用可以醒脾开胃行气,治疗脾胃气滞等。法半夏、茯苓和陈皮共有燥湿化痰、健脾和胃的作用。由此可见,熊继柏教授在治疗脾胃病证时,对症下药,注重全身调治,找出根本原因,祛除湿痰,注重治疗脾胃湿气,治疗胃痛等症。

参考文献:

- [1] 梁子钰.易水学派代表医家治疗脾胃病方药规律及学术思想嬗变研究[D].北京:北京中医药大学,2016.
- [2] 李 熙,刘树春.《丹溪心法》治疗脾胃病方剂用药特点分析[J].中国中医药图书情报杂志,2016,40(1):30-33.
- [3] 崔国宁,刘喜平,董军刚,等.基于数据挖掘分析脾胃病治疗中甘草用药规律[J].辽宁中医杂志,2017,44(11):2278-2279.
- [4] 信 楠.基于数据挖掘的《临证指南医案》脾胃病证治规律研究[D].广州:广州中医药大学,2016.
- [5] 刘新发,李廷保,井小会.基于数据挖掘金代名医李东垣治疗脾胃病用药思路探讨[J].中医研究,2013,26(10):75-78.
- [6] 聂多锐,刘兰林,郭锦晨,等.201 例新安王氏内科脾胃病医案组方用药规律及数据挖掘研究[J].西南医科大学学报,2017,40(5):484-487.
- [7] 蔡盛捷.李学麟教授治疗脾胃病处方用药规律研究及学术经验探讨[D].福州:福建中医药大学,2016.
- [8] 宋虎杰,李芳芳,王 辉,张文焕先生 150 张小儿脾胃病处方整理及用药特点研究[J].中国中西医结合儿科学,2016,8(2):206-209.
- [9] 刘文琛,龚后武,文 毅,等.基于数据挖掘的熊继柏教授诊疗失眠模型构建研究[J].湖南中医药大学学报,2016,36(7):75-78.
- [10] 何清湖,袁振仪.一名真正的名中医——熊继柏临证医案实录 1 [M].北京:中国中医药出版社,2009:65-99.
- [11] 高学敏.中药学[M].北京:中国中医药出版社,2007:59-454.
- [12] 西广成.复杂系统方法学与中医证候建模[M].北京:科学出版社,2010:68-82.
- [13] 唐仕欢,陈建新,杨洪军,等.基于复杂系统熵聚类方法的中药新药处方发现研究思路六[J].世界科学技术-中医药现代化,2009,11(2):225-228.
- [14] 赵亚丽.一类复杂系统的熵方法研究[D].北京:中国科学院自动化研究所,2005.

(本文编辑 贺慧娥)