

基于数据挖掘的郑钦安用药规律研究

龙晓华 王乐鹏 单体亮 马淑然

【摘要】 目的 利用现代统计学和数据挖掘方法,挖掘郑钦安的用药规律,为深入研究郑钦安的临床用药规律和学术思想提供统计学依据。**方法** 从《医理真传》与《医法圆通》中收集组成明确的方剂。利用 Excel 建立方药数据库,采用 SPSS 17.0 和 clementine 12.0 软件对数据进行描述统计分析、双变量相关性分析、关联规则分析和聚类分析。**结果** 方剂来源医著不少于 3 次者 15 种,来源于《伤寒论》的方剂最多,为 42 首。使用频次不少于 10 次的中药 32 味,排在前 10 味的中药主要是桂枝汤和四君子汤的药物组成。相关系数大于 0.2 的药对 30 对。关联规则分析提取出 84 条规则。聚类分析得出 6 组药和 7 对核心药对。**结论** 发现郑钦安在用药上具有如下几个特点:(1)法宗伤寒,推崇仲景;(2)辨分二纲,药有阴阳;(3)善补脾胃,协和营卫;(4)配伍灵活,用药圆通。

【关键词】 数据挖掘; 郑钦安; 用药规律

【中图分类号】 R249 **【文献标识码】** A doi:10.3969/j.issn.1674-1749.2015.05.007

基金项目:北京中医药大学科研创新团队项目(2011-CXTD-01);国家中医药管理局刘燕池名医传承工作室建设项目(01/010002607);北京市中医药管理局刘燕池名医传承工作站建设项目(2009-SZ-C-24)

作者单位:100029 北京中医药大学基础医学院[龙晓华(硕士研究生)、王乐鹏(博士研究生)、单体亮(硕士研究生)、马淑然]

作者简介:龙晓华(1990-),2013 级在读硕士研究生。研究方向:天人相应理论与临床。E-mail:1090292569@qq.com

通讯作者:马淑然(1964-),女,博士,教授,博士生导师。研究方向:天人相应理论与临床。E-mail:mashuran64@sina.com

Study on the medication regularities of ZHENG Qin-an based on data mining LONG Xiao-hua, WANG Le-peng, SHAN Ti-liang, et, al. School of Basic Medical Sciences, Beijing University of Chinese Medicine, Beijing, 100029, China

Corresponding author: MA Shu-ran, E-mail: mashuran64@ sina. com

【Abstract】 Objective To excavate the medication regularities of ZHENG Qin-an by means of modern statistics and data mining methods, and provide a statistical basis for in-depth study of ZHENG's clinical medication regularities and academic thoughts. **Methods** Prescriptions with specific composition were collected from the *True Principles of Medical Science* (Yi Li Zhen Zhuan) and *Flexibility of Medical Method* (Yi Fa Yuan Tong). Excel is used to set up the prescriptions database, SPSS17.0 and clementine12.0 softwares were used to do descriptive statistical analysis, bivariate correlation analysis, association rules analysis, and cluster analysis. **Results** 15 prescriptions come from at least three medical books, prescriptions from *Treatise on Cold Damages* rank most, totalling up to 42. 32 kinds of Chinese medicinals are used not less than 10 times. The Chinese medicinals that rank top 10 are mainly the composition of Guizhi decoction and Sijunzi decoction. 30 herb pairs' correlation coefficient is greater than 0.2. Association rules analysis has extracted out 84 rules. 6 groups of medicinals and 7 core herb pairs have been extracted by cluster analysis. **Conclusion** It is found that ZHENG's drug application has the following characteristics: (1) his medication regularities follow Cold Damages, and respect Zhong Jing; (2) pattern diagnosis and medicines both should be distinguished by yin and yang; (3) good at invigorating the spleen and stomach, coordinating *nutrient qi* and *defense qi*; (4) flexible compatibility and drug application.

【Key words】 Data mining; ZHENG Qin-an; Medication regularity

郑寿全(1824 - 1911), 字钦安, 四川邛崃人。清末伤寒学家, “火神派”创始人。学医于刘止唐, 郑钦安著有《医理真传》、《医法圆通》和《伤寒恒论》三书。唐步祺^[1]归纳出郑钦安对仲景学术的发挥主要有三: 一是阴阳为纲, 尤重心肾阳气; 二是百病不离六经气化; 三是联系实际, 阐释条文精义。并指出郑钦安对仲景阴阳学说和三阴症发挥颇多, 是近代不可多得的一位杰出的伤寒学家。郭子光在该书序言中指出: “其于阳虚辨治, 所积累之独到经验, 确是祖国医学中一份珍贵宝藏。”

目前关于郑钦安学术的研究多偏重于文献学诠释或临床经验的总结, 缺乏现代统计学研究的报道, 基于此, 本文拟利用现代统计学和数据挖掘方法, 从数据库中挖掘出郑钦安的用药规律, 从而为深入研究郑钦安的临床用药规律和学术思想提供统计学依据。

1 研究对象与方法

1.1 研究对象

查阅唐步祺所著《郑钦安医书阐释》中的《医理真传》与《医法圆通》全文, 找出药物组成明确的方剂。经过筛选, 找出符合选择要求的方剂。将以下字段输入 Excel 建立方药数据库: 方剂名称、方剂来源、药物组成。

1.2 研究方法

1.2.1 方剂选择 收集《医理真传》和《医法圆通·方剂索引》中的组成明确的方剂, 组成不明但审核后明确的方剂也一并纳入研究。二书中重复的方剂仅取一次。一共收集到 174 首方。

1.2.2 数据的标准化 建立规范的数据库前, 参考《中药学》^[2]《方剂学》^[3], 对方剂名称、方剂来源、药物组成和药物名称进行核对确认和规范化处理。

1.2.3 数据库的建立 用 Excel 表建立郑钦安方药数据库, 将方剂名称、方剂来源以及 174 首方所涉及的药物的名称输入 Excel 表第一行作为变量名; 第二行、第三行……一直到第 175 行, 每行代表一首方剂, 根据每首方剂使用中药的有、无, 分别在第一行中药名称下面对应的单元格中输入 1、0。输入完毕后, 再进行检查与核实。

1.2.4 统计分析 将建立的方药数据库分别导入 SPSS 17.0 和 clementine 12.0。在 windows 7 操作平台上, 应用 SPSS 17.0 进行描述统计、双变量相关及聚类分析; 应用 clementine 12.0 软件进行关联规则分析。

2 研究结果

2.1 描述统计分析

2.1.1 方剂来源分析 对选定的 174 首药物组成明

确的方剂的来源医著进行频数分析,涉及医籍共 46 种。其中频次不少于 3 次者 15 种。结果见表 1。

从上表的结果可以看出,来源于《伤寒论》的方剂最多,为 42 首。并且频次不少于 10 次的是张仲景的《伤寒论》与《金匱要略》,以及宋代官修的《太平惠民和剂局方》和郑钦安所著的《医理真传》,并且上表所列 15 种医籍,上至汉唐,下至宋金元明。

2.1.2 药物频次分析 对 174 首方剂中的药物进行频数分析,共涉及药物 182 种,药物使用总频次 1123 次。其中,使用频次不少于 10 次的药物共 32 味。结果见表 2,表中所列频率为单味药的频次与药物使用总频次的百分比。

从下表的结果中可以看出,在排在前 10 位的中药里,7 味味甘,6 味性温。桂枝、白芍、甘草、生姜和大枣是《伤寒论》第一方桂枝汤的组成药物;而人参、茯苓、白术、甘草则是《太平惠民和剂局方》四君子汤的组成药物。从中还可看出在以上 32 味药中,有补虚药 8 味,清热药 7 味,解表药 6 味。温热类药 17 味,苦寒类药 12 味。郑钦安除了常用麻黄、桂枝、附子、干姜等辛温之品外,还常用黄芩、黄连、大黄、黄柏等苦寒药。

2.2 双变量相关与关联规则分析

2.2.1 双变量相关性分析 对使用药物进行双变量相关性分析,在分析中将用药次数不少于 15 次的前 22 种药物作为分析对象,其中相关系数大于 0.2

表 1 方剂来源医籍频次表

医籍名称	频次	医籍名称	频次
伤寒论	42	丹溪心法	4
太平惠民和剂局方	24	宣明论方	3
医理真传	14	小儿药证直诀	3
金匱要略	10	瘟疫论	3
东垣十书	9	卫生宝鉴	3
景岳全书	6	外台秘要	3
证治准绳	4	类证活人书	3
寿世保元	4		

表 2 中药使用频率表

药物名称	频次	频率%	药物名称	频次	频率%
甘草	97	8.64	半夏	19	1.69
茯苓	38	3.38	川芎	18	1.60
当归	37	3.29	柴胡	16	1.42
人参	36	3.21	黄柏	15	1.34
白芍	31	2.76	栀子	15	1.34
白术	31	2.76	熟地黄	15	1.34
生姜	30	2.67	桔梗	14	1.25
黄芩	29	2.58	厚朴	14	1.25
大枣	27	2.40	升麻	14	1.25
桂枝	26	2.32	泽泻	14	1.25
生地黄	25	2.23	黄芪	12	1.07
附子	24	2.14	牡丹皮	11	0.98
黄连	23	2.05	杏仁	10	0.89
陈皮	21	1.87	知母	10	0.89
干姜	20	1.78	细辛	10	0.89
大黄	20	1.78	麻黄	10	0.89

的 30 个药对结果见表 3。

从表 3 可以看出,在相关系数大于 0.4 的前 8 个药对中,有 5 个药对为桂枝汤的组成药物。从上表所列 30 个药对中还可发现,郑钦安用药配伍是非常灵活的,例如生地黄配黄连、黄芩以清热,配当归、川芎以凉血活血。半夏配陈皮、茯苓燥湿化痰,配干姜温化寒饮。

2.2.2 关联规则分析 关联分析的目的就是要寻找事物之间的联系和规律,发现它们之间的关联关系。Clementine 中提供了 Apriori, GRI, Carma 等算法。其中,Apriori 算法最早由 Agrawal 和 Srikant 于 1994 年提出,现已成为数据挖掘中简单关联规则技术的核心算法^[4]。

在本次分析中,以用药次数不少于 15 次的前 22 种药物为分析对象,将前项最小支持度设置为 6%,规则的最小置信度设置为 60%,前项包含的最大项目数设置为 3。最终 Clementine 软件共生成规则 84 条,结果见表 4。

从以上的关联规则分析结果可以看出,药物组合规律主要为桂枝汤、四君子汤(六君子汤)的药物组成,有 50 条规则的药物组成存在于桂枝汤方中,25 条规则的药物组成存在于六君子汤方中。另外,尚有凉血活血的“川芎-当归-生地黄”和清热的“黄连-黄柏”“黄芩-栀子”等。

2.3 聚类分析

聚类分析能够将样本数据,在没有“先验知识”

的前提下,根据数据的诸多特征,按照其在性质上的“亲疏程度”进行自动分组,且使组内部个体的结构特征具有较大相似性,组之间个体的特征相似性较小^[4]。

对使用频次不少于 15 次的前 22 味药采用 SPSS 软件中的系统聚类法进行聚类分析,同时进行核心药对的分析和提取,聚类分析结果见图 1。

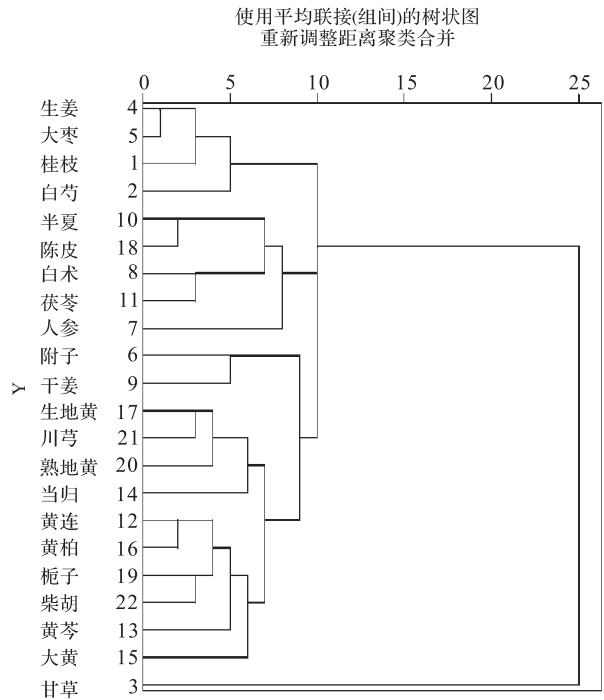


图 1 系统聚类分析树状图

表 3 中药双变量相关性分析

中药		相关系数	中药		相关系数
生姜	大枣	0.65	人参	茯苓	0.28
白术	茯苓	0.55	干姜	半夏	0.28
桂枝	生姜	0.45	黄连	生地黄	0.28
桂枝	大枣	0.44	茯苓	陈皮	0.27
当归	生地黄	0.43	半夏	茯苓	0.26
白芍	生姜	0.42	人参	陈皮	0.25
黄连	黄柏	0.42	白术	陈皮	0.24
白芍	大枣	0.42	当归	熟地黄	0.24
生地黄	川芎	0.40	黄芩	柴胡	0.23
半夏	陈皮	0.38	甘草	陈皮	0.22
当归	川芎	0.38	甘草	大枣	0.22
黄连	黄芩	0.37	附子	干姜	0.22
黄芩	栀子	0.36	黄芩	生地黄	0.21
桂枝	白芍	0.31	桂枝	甘草	0.21
人参	白术	0.28	大黄	栀子	0.21

表 4 药物关联规则分析

后项	前项	支持度	置信度	后项	前项	支持度	置信度
甘草	陈皮-茯苓	6.32	100	茯苓	白术-甘草	13.22	73.91
甘草	桂枝-大枣	8.05	100	桂枝	大枣-白芍	8.62	73.33
甘草	桂枝-白芍	6.9	100	桂枝	生姜-白芍-甘草	8.05	71.43
甘草	白术-人参	8.05	100	白术	人参-茯苓-甘草	8.05	71.43
甘草	桂枝-大枣-生姜	6.9	100	茯苓	白术-人参	8.05	71.43
甘草	桂枝-大枣-白芍	6.32	100	茯苓	白术-人参-甘草	8.05	71.43
甘草	大枣-白芍	8.62	93.33	茯苓	白术	17.82	70.97
大枣	桂枝-生姜-甘草	7.47	92.31	白术	茯苓-甘草	13.79	70.83
甘草	大枣-生姜-白芍	7.47	92.31	桂枝	大枣-生姜-甘草	9.77	70.59
大枣	桂枝-白芍	6.9	91.67	白芍	大枣-生姜-甘草	9.77	70.59
大枣	桂枝-白芍-甘草	6.9	91.67	甘草	生姜	17.24	70
生姜	桂枝-大枣-白芍	6.32	90.91	甘草	柴胡	9.2	68.75
甘草	生姜-白芍	9.2	87.5	生地黄	川芎-当归	6.9	66.67
甘草	人参-茯苓	9.2	87.5	大枣	生姜	17.24	66.67
生姜	大枣-白芍	8.62	86.67	大枣	桂枝-甘草	12.07	66.67
甘草	桂枝-生姜	8.62	86.67	白芍	桂枝-生姜	8.62	66.67
大枣	生姜-白芍-甘草	8.05	85.71	白芍	生姜-甘草	12.07	66.67
生姜	桂枝-大枣	8.05	85.71	当归	川芎	10.34	66.67
生姜	桂枝-大枣-甘草	8.05	85.71	甘草	黄芩	16.67	65.52
生姜	大枣-白芍-甘草	8.05	85.71	白芍	大枣-生姜	11.49	65
甘草	陈皮	12.07	85.71	当归	生地黄	14.37	64
甘草	大枣-生姜	11.49	85	半夏	陈皮-茯苓	6.32	63.64
生姜	桂枝-白芍	6.9	83.33	半夏	陈皮-茯苓-甘草	6.32	63.64
生姜	桂枝-白芍-甘草	6.9	83.33	桂枝	大枣-甘草	12.64	63.64
白芍	桂枝-大枣-生姜	6.9	83.33	白术	陈皮-茯苓	6.32	63.64
甘草	大枣	15.52	81.48	白芍	大枣-甘草	12.64	63.64
大枣	生姜-白芍	9.2	81.25	白术	陈皮-茯苓-甘草	6.32	63.64
大枣	生姜-甘草	12.07	80.95	人参	陈皮-茯苓	6.32	63.64
甘草	桂枝	14.94	80.77	人参	陈皮-茯苓-甘草	6.32	63.64
大枣	桂枝-生姜	8.62	80	甘草	茯苓	21.84	63.16
甘草	半夏	10.92	78.95	桂枝	生姜-白芍	9.2	62.5
桂枝	大枣-白芍-甘草	8.05	78.57	白术	人参-茯苓	9.2	62.5
白芍	桂枝-大枣	8.05	78.57	桂枝	生姜-甘草	12.07	61.9
白芍	桂枝-大枣-甘草	8.05	78.57	生姜	桂枝-甘草	12.07	61.9
生姜	大枣-甘草	12.64	77.27	茯苓	陈皮-甘草	10.34	61.11
甘草	白术-茯苓	12.64	77.27	大枣	白芍-甘草	13.22	60.87
桂枝	大枣-生姜-白芍	7.47	76.92	生姜	白芍-甘草	13.22	60.87
白芍	桂枝-生姜-甘草	7.47	76.92	人参	白术-甘草	13.22	60.87
甘草	人参	20.69	75	陈皮	半夏-甘草	8.62	60
甘草	白芍	17.82	74.19	黄连	黄柏	8.62	60
甘草	白术	17.82	74.19	桂枝	大枣-生姜	11.49	60
生姜	大枣	15.52	74.07	黄芩	栀子	8.62	60.00

通过聚类分析,所选 22 味药被分为了 6 组:生姜、大枣、桂枝、白芍为一组,若加甘草则为桂枝汤的药物组成;半夏、陈皮、白术、茯苓、人参为一组,若加甘草则为六君子汤的药物组成;附子和干姜为一组,若加甘草则为四逆汤的药物组成;生地黄、川芎、熟地黄、当归为一组,若易生地黄为白芍则为四物汤的药物组成;黄连、黄柏、栀子、柴胡、黄芩、大黄为一组;甘草为一组。可见,郑钦安常以桂枝汤和营卫、调脾胃;六君子汤益气健脾化痰;四逆汤“回阳祛阴”;四物汤补血调血;黄芩、黄连、黄柏、栀子、柴胡与大黄泻火救阴。另外,得到核心药对 7 对:栀子-柴胡、黄连-黄柏、生姜-大枣、白术-茯苓、川芎-生地黄、半夏-陈皮、附子-干姜。

3 讨论

数据挖掘,又称为数据库知识发现,可以将其描述为按既定决策目标,对大量的数据进行探索和分析,揭示隐藏的、未知的或验证已知的规律性,并进一步将其模型化的先进有效的方法。它是在没有明确假设的前提下去挖掘信息、发现知识。同时,数据挖掘所得到的信息具有先前未知、有效和可实用 3 个特征^[5]。

本研究采用数据挖掘技术,对郑钦安医著《医理真传》与《医法圆通》中组成明确的方剂的用药规律进行了分析,发现了郑钦安的部分“隐藏着的”用药规律。通过对上述统计结果的分析,发现郑钦安在用药方面具有如下几个特点:(1)法宗伤寒,推崇仲景。通过上文对方剂来源的分析,发现郑钦安临床用药是以仲景方为主的,来源于《伤寒论》的方剂 42 首,源于《金匱要略》的方剂 10 首,充分说明了他对仲景学术的推崇与运用,这正如郑钦安所说:“余实推诚相与,愿与后世医生同入仲景之门,共用仲景之法。”^[6](2)辨分二纲,药有阴阳。通过上述药物频次分析、关联规则分析、聚类分析,发现作为“火神派”创始人的郑钦安既善于运用温热类方药,也不忽视对寒凉药的运用,这充分体现了郑钦安以阴阳二纲辨治疾病的思想,正如郑钦安所说:“凡三阴症,以温补为要,是阴盛阳必衰,故救阳为急;三阳症,以解散清凉为主,是阳盛阴必亏,故救阴为先。”^[6](3)善

补脾胃,协和营卫。通过用药频次、双变量相关、关联规则以及聚类分析,发现排在用药频次前 10 位中的 5 味、相关系数大于 0.4 的前 8 个药对中的 5 个、84 条关联规则中的 50 条规则以及聚类分析的结果均有桂枝汤方的组成药物。在用药频次排在前 10 位的中药里还有四君子汤方的组成药物。关联规则分析、聚类分析还发现有六君子汤方的组成药物。郑钦安谓桂枝汤是“调和阴阳第一法”、“能宣营卫,协和阴阳”^[6],桂枝汤不仅能外调营卫,而且内和脾胃,滋阴和阳。外证得之,解肌祛邪;内证得之,调脾胃,和阴阳^[7]。四君子汤益气健脾,六君子汤兼化痰。(4)配伍灵活,用药圆通。通过方剂来源分析,发现郑钦安用方之来源上至汉唐,下至宋金元明,足见其虽推崇仲景而不泥于仲景方,兼收并蓄、灵活多变的用方风格。通过聚类分析,发现郑钦安常以栀子-柴胡清热散热,黄连-黄柏清热燥湿;生姜-大枣、白术-茯苓健脾和胃;川芎-生地黄凉血活血;半夏-陈皮燥湿化痰;附子-干姜温阳、回阳,充分体现了一代伤寒大家、火神派开山者在用药上的辨证论治精神。

本次研究旨在为进一步探讨郑钦安的临床用药经验和学术思想提供一定的统计学依据,同时也说明利用数据挖掘等现代科学技术研究中医药学不仅是一种科学有效的方法,而且也为中医药现代化发展指明了方向。

参 考 文 献

- [1] 唐步祺. 郑钦安医书阐释[M]. 四川:巴蜀书社,2006:1-9.
- [2] 高学敏. 中药学[M]. 北京:中国中医药出版社,2007:1-524.
- [3] 许济群. 方剂学[M]. 上海:上海科学技术出版社,1985:1-236.
- [4] 薛薇,陈欢歌. 基于 Clementine 的数据挖掘[M]. 北京:中国人民大学出版社,2012:435-450,398.
- [5] 熊平. 数据挖掘算法与 Clementine 实践[M]. 北京:清华大学出版社,2011:2.
- [6] 郑钦安. 中医火神派三书[M]. 周鸿飞点校. 北京:学苑出版社,2007:124,30,237-238.
- [7] 熊曼琪. 伤寒学[M]. 北京:中国中医药出版社,2007:23.

(收稿日期:2014-10-08)

(本文编辑:蒲晓田)