

基于关联规则与熵聚类的补气类中成药组方规律研究

吴嘉瑞 金燕萍 张冰 张晓朦 周唯 杨冰

(北京中医药大学,北京,100029)

摘要 目的:探讨常用补气类中成药组方规律。方法:收录《新编国家中成药》中的补气类中成药处方,采用关联规则apriori算法和复杂系统熵聚类等方法,确定处方中药物的使用频次及药物之间的关联规则等。结果:高频次药物包括黄芪、当归、熟地黄、茯苓、白术等;高频次药物组合包括“当归、黄芪”“熟地黄、当归”“白芍、当归”等;置信度较高的关联规则包括“川芎、熟地黄->当归”“白芍、熟地黄->当归”等。结论:处方用药中除常见的补气类中药外,尚包括具有养阴、生津、补血作用的部分补益药及其他类药物。

关键词 补气药;关联规则;熵聚类

Analysis on Composition Rules of Qi Tonic of Chinese Patent Medicine Based on Association Rules and Clustering Algorithm

Wu Jiarui, Jin Yanping, Zhang Bing, Zhang Xiaomeng, Zhou Wei, Yang Bing

(Beijing University of Chinese Medicine, Beijing 100102, China)

Abstract Objective: To explore the composition rules of qi tonic of Chinese patent medicine. **Methods:** The formula of qi tonic of Chinese patent medicine in The New National Medicine were collected to achieve the frequency of herbs and association rules between herbs by the methods of association rules with apriori algorithm and complex system clustering algorithm. **Results:** The most frequently used herbs were Milkvetch Root, Chinese Angelica, Prepared Rehmannia Root, Indian Bread, White Atractylodes Rhizome etc. The most frequently used herb combinations were Chinese Angelica and Milkvetch Root, Prepared Rehmannia Root and Milkvetch Root, as well as Debarb Peony Root and Prepared Rehmannia Root etc. The herbs with a high confidence degree of association rules include “Sichuan Lovage Rhizome, Prepared Rehmannia Root-> Chinese Angelica”, “Debarb Peony Root, Prepared Rehmannia Root-> Chinese Angelica” etc. **Conclusion:** Besides qi-tonifying herbs, the Chinese patent medicine prescriptions of the type also include some tonics with yin nourishing, fluid production and blood tonifying herbs etc.

Key Words Qi tonic; Association rules; Clustering algorithm

中图分类号:R311;R385.6 文献标识码:A doi:10.3969/j.issn.1673-7202.2016.01.045

中医认为,元气、宗气、肾气等是维持机体生命活动的主要物质,气虚可导致机体的脏腑功能减退^[1]。《素问·天元纪大论》指出:“在天为气,在地为形,形气相感而化生万物矣。”《素问·宝命全形论》说:“人以天地之气生,四时之法成。”“天地合气,命之曰人。”^[2]人也是天地之交的产物,所以人的形体构成,也是以气为最基本物质的。补气药能补益脏气以纠正人体脏气虚衰的病理偏向。其包括补脾气、补肺气、补心气、补元气等,故而使用本类药物可以治疗各种虚证,也可结

合其他类药综合治疗多种病症,如脾虚食滞证,可配伍消食药同用,以清除消化功能减弱而停滞的宿食;用于脾虚中气下陷证,多配伍能升阳的药物,以升举下陷的清阳之气等。本研究收集《新编国家中成药》中的补气类中成药处方,在构建数据库的基础上,应用关联规则和熵聚类算法进行处方用药规律的研究^[3-4]。

1 资料与方法

1.1 处方来源与筛选 本研究以《新编国家中成药》(第二版)^[5]为来源,共筛选出补气药处方 167

基金项目:国家科技支撑计划课题(编号:2007BAI10B01);北京市中医药科技发展基金课题(编号:JJ-2010-70);北京中医药大学科研创新团队项目资助(编号:2011-CXTD-14);北京中医药大学“重点学科”开放课题(编号:2013-ZDXKKF-19);北京市与中央在京高校共建项目(编号:BJGJ1420);北京中医药大学与北京灸道堂中医研究院横向合作课题(编号:BUCCM-JDT-001)

通信作者:吴嘉瑞,副教授,博士,硕士研究生导师,研究方向:从事临床中药学研究,E-mail:exogamy@163.com;张冰,教授,博士,博士研究生导师,研究方向:从事临床中药学研究,E-mail:zhangbing@263.net

首。

1.2 分析软件 “中医传承辅助系统 (V2.0)” 软件^[6], 中国中医科学院中药研究所提供。该系统围绕中医药继承、发展、传播和创新 4 个核心问题, 集数据收集、管理、分析等为一体, 有效实现了疾病、证候、中药、方剂、医案及其相关信息的管理、检索、分析等功能, 在中医临床经验传承与学习、新药研发等领域具有重要应用价值^[7]。

1.3 处方的录入与核对 将上述筛选的方剂录入“中医传承辅助系统”, 并对药物名称进行规范统一。在完成录入后, 由专人负责录入数据源的审核, 以确保数据源的准确, 为数据挖掘结果的可靠性提供保障^[6]。

1.4 数据分析 通过“中医传承辅助系统”中“数据分析”模块中“方剂分析”功能进行组方规律分析。通过系统“数据查询”功能, 提取出录入的 167 首补气方剂。然后点击相应功能按钮进行数据分析, 包括药物频次统计、组方规律分析、新方分析。

2 结果

2.1 用药频次 将药物按使用频次从高到低进行排序, 前三位分别是黄芪, 当归和熟地黄, 频次前 30 位的药物情况见表 1。

表 1 处方中使用频次前 30 位的药物情况表

序号	中药名称	频次	序号	中药名称	频次
1	黄芪	97	16	山药	25
2	当归	85	17	麦冬	24
3	熟地黄	67	18	生地黄	22
4	甘草	65	19	香附	21
5	茯苓	62	20	鹿茸	21
6	白术	57	21	菟丝子	20
7	白芍	54	22	怀牛膝	20
8	党参	51	23	杜仲	19
9	川芎	50	24	大枣	19
10	人参	50	25	丹参	18
11	五味子	38	26	红花	18
12	枸杞子	35	27	益母草	17
13	陈皮	29	28	太子参	17
14	肉桂	28	29	砂仁	16
15	阿胶	25	30	远志	16

2.2 基于关联规则的组方规律分析 按照药物组合出现频次由高到低排序, 前 3 位分别是“当归, 黄芪”“熟地黄, 当归”“白芍, 当归”, 出现频次 35 次以上的药物组合见表 2。在支持度为 20%, 置信度 ≥ 0.75 条件下, 分析所得药对的用药规则, 结果见表 3, 关联规则网络图见图 1。

2.3 基于熵聚类的方剂组方规律分析

2.3.1 基于改进的互信息法的药物间关联度分析 依据不同参数提取的数据并结合样本数和经验判断, 设置相关度为 8, 惩罚度为 4, 进行聚类分析, 得到处方中两两药物间的关联度, 将关联系数 0.035 以上的药对列表, 见表 4。

表 2 处方中高频次药物组合情况表

序号	药物组合	频次	序号	药物组合	频次
1	当归, 黄芪	60	12	白芍, 黄芪	39
2	熟地黄, 当归	50	13	甘草, 当归	37
3	白芍, 当归	49	14	党参, 黄芪	36
4	川芎, 当归	46	15	川芎, 白芍	36
5	白术, 当归	46	16	熟地黄, 当归, 黄芪	36
6	当归, 茯苓	44	17	川芎, 白芍, 当归	36
7	熟地黄, 茯苓	41	18	川芎, 熟地黄	35
8	熟地黄, 黄芪	40	19	白术, 白芍	35
9	白芍, 熟地黄	40	20	川芎, 熟地黄, 当归	35
10	白术, 茯苓	40	21	白芍, 当归, 黄芪	35
11	白芍, 熟地黄, 当归	40			

表 3 处方中药物组合关联规则 (置信度 ≥ 0.75)

序号	规则	置信度	序号	规则	置信度
1	川芎, 熟地黄->当归	1	12	熟地黄, 茯苓->当归	0.8293
2	白芍, 熟地黄->当归	1	13	白芍, 当归->熟地黄	0.8163
3	白芍, 茯苓->当归	1	14	白术->当归	0.8070
4	川芎, 白芍->当归	1	15	熟地黄, 当归->白芍	0.8
5	白术, 白芍->当归	0.9714	16	川芎, 当归->白芍	0.7826
6	白术, 熟地黄->当归	0.9706	17	当归, 茯苓->熟地黄	0.7727
7	川芎->当归	0.92	18	当归, 茯苓->白术	0.7723
8	白芍->当归	0.9074	19	当归, 茯苓->白芍	0.7727
9	熟地黄, 黄芪->当归	0.9	20	甘草->当归	0.7708
10	白芍, 黄芪->当归	0.8974	21	川芎, 当归->熟地黄	0.7609
11	白术, 茯苓->当归	0.85			

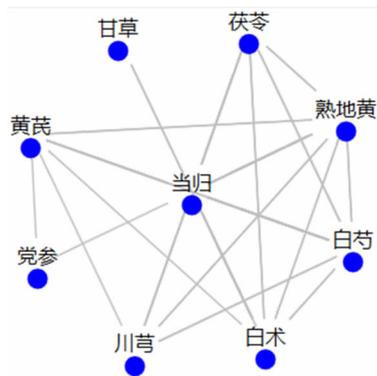


图 1 支持度为 20%, 置信度为 0.75 条件下的关联规则网络展示图

2.3.2 基于复杂系统熵聚类的药物核心组合与新处方分析 以改进的互信息法的药物间关联度分析结果为基础, 按照相关度与惩罚度约束, 基于复杂系统熵聚类, 演化出 3~4 味药物核心组合, 具体

见表5。在核心组合提取的基础上,运用无监督熵层次聚类算法,得到4个新处方,具体见表6。

表4 基于改进的互信息法的药物间关联度分析

序号	药对	关联系数	序号	药对	关联系数
1	白术,香附	0.0538	14	川芎,延胡索	0.0400
2	延胡索,益母草	0.0461	15	茯苓,肉桂	0.0395
3	白术,黄芩	0.0437	16	延胡索,海螵蛸	0.0389
4	茯苓,远志	0.0435	17	香附,鳖甲	0.0384
5	茯苓,砂仁	0.0435	18	白术,续断	0.0382
6	白芍,乌鸡	0.0420	19	阿胶,砂仁	0.0376
7	川芎,肉桂	0.0418	20	白术,砂仁	0.0372
8	肉桂,人参	0.0418	21	杜仲,鹿茸	0.0371
9	延胡索,阿胶	0.0412	22	川芎,艾叶	0.0366
10	白术,熟地黄	0.0409	23	香附,柴胡	0.0357
11	延胡索,杜仲	0.0409	24	川芎,红花	0.0352
12	川芎,益母草	0.0403	25	延胡索,白芍	0.0351
13	黄芩,续断	0.0401			

表5 基于复杂系统熵聚类的药物核心组合

序号	核心组合
1	牡丹皮,杜仲,续断
2	沉香,琥珀,化橘红
3	远志,酸枣仁,苍耳子
4	女贞子,枸杞子,菟丝子
5	杜仲,砂仁,怀牛膝
6	沉香,没药,僵蚕
7	远志,酸枣仁,柏子仁
8	红参,枸杞子,鹿茸,龟甲

表6 基于熵层次聚类的新处方

序号	新处方
1	牡丹皮,杜仲,续断,砂仁,怀牛膝
2	沉香,琥珀,化橘红,没药,僵蚕
3	远志,酸枣仁,苍耳子,柏子仁
4	女贞子,枸杞子,菟丝子,红参,鹿茸,龟甲

3 讨论

本研究应用关联规则和聚类算法系统分析了《新编国家中成药》中补气类中成药的用药规律。以下结合研究结果,对方剂中高频次药物进行分析。黄芪是出现频次最高的单味中药,其味甘微温,归脾肺经,功能补气健脾,升阳举陷,常用于脾肺气虚证以及气虚自汗、气血双亏等证,多与人参、白术等配伍使用。当归出现频次仅次于黄芪,能补血调经、活血止痛、润肠通便;其甘温质润,长于补血,为补血之圣药,多用于血虚诸证,多与滋阴、补气药合用,如当归补血汤、四物汤等。熟地黄甘温,具有补血滋阴、填精益髓之功效,补阴益精以生血,为养血补虚之要药,常用于血虚诸证,多配伍当归、白芍、川芎等;且质润入肾,善

滋补肾阴,为补肾阴之要药,可用于肝肾阴虚诸证,常与山药、山茱萸合用,如六味地黄丸。茯苓味甘淡,能利水渗湿,健脾宁心,甘能补,淡能泄,药性平和,即可祛邪,又可扶正,利水而不伤正气,兼能健脾,故能健脾渗湿而止泻,尤宜于脾虚湿盛泄泻,可与山药、白术、薏苡仁同用,如参苓白术散;且味甘,善入脾经,能健脾补中,常配以人参、白术等,治疗脾胃虚弱等,如四君子汤。白术味苦甘温,善于补气健脾,燥湿利水,为补气健脾燥湿之要药,脾虚湿盛的大便溏泻尤宜^[8],有“脾脏补气健脾第一要药”之称。

本研究所得的药物间关联规则有助于分析药物间的关联程度,如“川芎,熟地黄->当归”(置信度为1)含义为,处方中有川芎、熟地黄时,有当归的概率为100%;“白芍,熟地黄->当归”(置信度为1)含义为,处方中有白芍、熟地黄时,有当归的概率为100%。又如关联规则网络图中共9味药物,这些药物是以八珍汤为主方加减而成的,如当归、白术、茯苓、白芍、川芎、熟地黄等,而党参功似人参而力弱,能补脾肺气。可见中成药处方中高关联度药物组合与中医经典名方的一致性,这也在一定程度上体现了中医理论的传承与延续。

再者本研究应用复杂系统熵聚类方法得出的新处方不乏启迪意义,人体气的生成来源可分为三个方面,一是自然界的清气,二是水谷精微之气,三是肾中之精化生的精气。由此可知气的生成与肺、脾胃、肾等脏器的功能关系密切。临床上尤以脾气虚最为常见,脾虚证本质是以脾气虚为本,挟湿挟痰挟热、气滞为标,临床上根据脾虚兼挟症的轻重不同而治疗上则以健脾为主,辅以渗湿、化痰、清热、行气^[9]。如“牡丹皮、杜仲、续断、砂仁、牛膝”的配伍处方,牡丹皮清热凉血,以清气中实热;杜仲补肝肾,肾藏精,化生元气;续断既能补肝肾,又能行血脉;砂仁化湿行气,理脾胃气滞;怀牛膝补肝肾兼能活血祛风除湿,故诸药合用既能滋补肝肾以固先天之本,又能健脾燥湿,祛邪的同时扶正,兼能凉血活血,营行脉内,化生血液,补气时养血。

综上,本研究应用数据挖掘方法对《新编国家中成药》中的补气类处方用药规律进行研究,获得了传统统计学研究未获得的新知识、新信息,为补气类处方的深入挖掘和组方规律探寻提供了参考。当然,应用关联规则和熵聚类研究得的组方规则等还需进一步临床验证,数据挖掘方法亦有其局限

性,需结合中医药理论与实践综合分析、评价^[10]。

参考文献

- [1] 马文芳,郭新荣,余秀,等.青蒿组分的体外抗氧化活性研究[J].中国医药导报,2014,11(14):4-8.
- [2] 刘渡舟.白话中医四部经典[M].天津:天津科技翻译出版公司,1993:54.
- [3] 唐仕欢,陈建新,杨洪军,等.证-病-方——中医药研究的新领域[J].中国中医基础医学杂志,2010,16(4):267-270.
- [4] 唐仕欢,杨洪军.中医组方用药规律研究进展述评[J].中国实验方剂学杂志,2013,19(5):359-363.
- [5] 宋民宪,杨明.新编国家中成药(第二版)[M].北京:人民卫生出版社,2011.

- [6] 卢朋,李健,唐仕欢,等.中医传承辅助系统软件开发与应用[J].中国实验方剂学杂志,2012,18(9):1-4.
- [7] 盛倩,刘绍燕,庄曾渊,等.基于中医传承辅助系统的李东垣病方剂用药规律研究[J].中国中医眼科杂志,2014,24(3):177-180.
- [8] 钟赣生.常用中药功用鉴别--补气药[J].中国临床医生,2000,28(5):50-52.
- [9] 罗琦,纪云西,陶双友,等.脾虚证补气药运用时机选择[J].中国民族民间医药,2010(15):128,131.
- [10] 吴嘉瑞,郭位先,张冰,等.基于数据挖掘的国医大师颜正华含陈皮处方用药规律研究[J].中国中药杂志,2014,39(4):618-622.
- (2015-01-07 收稿 责任编辑:张文婷)

2015年世界中医药十大新闻揭晓

2016年1月10日,世界中医药学会联合会在北京举行新闻发布会,由世界中医药网和世界中联新媒体专业委员会邀请有关媒体及专家评出的《2015年世界中医药十大新闻》榜单揭晓。

世界中联主席余靖,创会副主席兼秘书长李振吉,副秘书长姜再增,世界中医药网总编辑陈贵廷,国家中医药管理局新闻办公室主任欧阳波,世界中联新媒体专业委员会副会长、以岭药业总经理吴相君等出席发布会。

这十大新闻分别为:1.中国科学家屠呦呦获诺贝尔生理学或医学奖。2.中国中医药振兴发展迎来

天时地利人和大好时机。3.匈牙利《中医立法实施细则》10月19日生效。4.ISO/TC249(即国际标准化组织/中医药技术委员会)正式命名,新发布4项中医药国际标准。5.中捷中医中心、中法中医药中心成立。6.康莱特注射液、莲花清瘟胶囊获批进入美国FDA临床研究。7.世界中医药学会联合会(WFCMS)与世界卫生组织(WHO)建立官方正式关系,《世界中医药杂志》英文刊创刊。8.首家中医针灸国际传承基地落户加拿大多伦多。9.世界卫生组织传统医药合作中心落户澳门。10.国际制药巨头及农场业者瞄准中药市场商机。