

思路与方法

基于关联规则与熵聚类的健脾类中成药组方规律研究

金燕萍 吴嘉瑞 张 冰 杨 冰 周 唯 张晓滕

(北京中医药大学, 北京, 100102)

摘要 目的:探讨常用健脾类中成药组方规律。方法:收录《新编国家中成药》中的健脾类中成药处方,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:高频次药物包括茯苓、白术、甘草、党参、陈皮等;高频次药物组合包括“白术、茯苓”“甘草、茯苓”“甘草、白术”等;置信度较高的关联规则包括“陈皮->白扁豆”“陈皮->半夏”等。结论:处方用药中除常见的健脾类中药外,尚包括具有健脾作用的部分理气药、消食药及其他类药物。

关键词 健脾药;关联规则;熵聚类

Analysis on Composition Rules of Chinese Patent Drugs with Tonifying Spleen Based on Association Rules and Clustering Algorithm

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

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

Abstract Objective: To investigate composition rules of Chinese patent drugs with tonifying spleen. **Methods:** The prescriptions of Chinese patent drugs with tonifying spleen in “The New National Medicine” were collected to build a database. The methods of association rules with apriori algorithm and complex system entropy cluster were used to achieve the frequency of medicines and association rules between drugs. **Results:** The data-mining results indicated that in the prescriptions of Chinese patent drugs with tonifying spleen, the most frequency used drugs were Poria Cocos Wolff, Rhizoma Atractylodis Macrocephalae, Radix Glycyrrhizae, Radix Codonopsis, Pericarpium Citri Reticulatae. The most common drug combinations were “Rhizoma Atractylodis Macrocephalae, Poria Cocos Wolff”, “Radix Glycyrrhizae, Poria Cocos Wolff”, “Radix Glycyrrhizae, Rhizoma Atractylodis Macrocephalae”. The drugs with a high degree confidence coefficient of association rules include “Pericarpium Citri Reticulatae-> Semen Dolichoris”, “Pericarpium Citri Reticulatae-> Pinellia ternata”. **Conclusion:** There are not only the common drugs tonifying spleen, but also drugs regulating the flow of vital energy, removing obstruction to it, and helping digest.

Key Words Chinese patent drugs; Tonifying spleen; Association Rules; Clustering Algorithm

中图分类号:R289 文献标识码:A doi:10.3969/j.issn.1673-7202.2015.10.039

《脾胃论·脾胃胜衰论》中提出:“百病皆由脾胃衰而生也,内伤脾胃,百病由生。”^[1]中医理论认为“脾胃为后天之本”“气血生化之源”。临床上脾病证候分虚实两种,其证候包括脾气虚证、脾虚气陷证、脾不统血证、脾阳虚证等4个虚证及寒湿困脾证、湿热蕴脾证2个实证^[2-3],所以健脾药又分理气健脾、化湿健脾、消食健脾等。理气健脾者如陈皮、枳实等,多用于脾胃气滞之证,多与消导药合用;化湿健脾者如苍术、厚朴、砂仁等,多用于湿浊内阻,脾胃湿困,运化失常所致的脘腹痞满、食少体倦等症;

消食健脾者如山楂、神曲、麦芽等,具消食化积、健脾开胃,和中之功。健脾类中成药种类较多,除上述常见的理气健脾、化湿健脾以及消食健脾外,还有诸多其他类别药物,如健脾安神、健脾补肾、健脾止泻等,而其配伍规律值得深入探讨。本研究收集《新编国家中成药》中的健脾类中成药处方,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^[4-5]。

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

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

作者简介:金燕萍(1988.12—),女,在读研究生,从事临床中药学研究,E-mail:qingpingzhumeng@sina.com

通信作者:吴嘉瑞(1977.1—),男,博士,副教授,硕士研究生导师,从事临床中药学研究,Tel:(010)84738662,E-mail:exogamy@163.com;张

冰(1959.8—),女,博士,教授,博士研究生导师,从事临床中药学研究,Tel:(010)84738662,E-mail:zhangbing@263.net

药》(第二版)^[6]为来源,共筛选出健脾药处方 393 首。

1.2 分析软件 “中医传承辅助系统(V2.0)”软件,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^[7-8]。

1.3 方剂的录入与核对 将筛选后的处方录入“中医传承辅助系统(V2.0)”数据管理模块中的“方剂管理”,由双人负责数据的审核,以确保数据的准确性,建立关于补气类中成药的方剂数据库。然后通过中医传承辅助系统中的数据分析模块中的“方剂分析”进行统计分析,包括药物频次统计、组方规律分析、新方分析。对于方剂中所出现的生(炒)白术、炙甘草等药名,为避免药物出现频次降低从而影响关联性,参照 2010 版《中国药典》规范统一录入为白术、甘草等^[9]。

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 394 首健脾方剂。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,前三位分别是茯苓、白术和甘草,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	茯苓	178	16	鸡内金	41
2	白术	177	17	六神曲	41
3	甘草	129	18	苍术	38
4	党参	126	19	柴胡	37
5	陈皮	123	20	丹参	37
6	黄芪	109	21	厚朴	36
7	山药	98	22	白扁豆	35
8	白芍	86	23	枸杞子	33
9	当归	82	24	大枣	32
10	山楂	81	25	熟地黄	32
11	木香	61	26	香附	30
12	麦芽	56	27	薏苡仁	29
13	砂仁	55	28	麦冬	28
14	人参	50	29	枳壳	27
15	泽泻	46	30	五味子	27

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“白术、茯苓”“甘草、茯苓”“甘草、白术”,出现频次 50 次以上的药物组合见表 2。在支持度个数为 50,置信度≥0.6 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。

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	白术,茯苓	119	14	白术,白芍	58
2	甘草,茯苓	90	15	山药,白术	56
3	甘草,白术	87	16	白术,黄芪	56
4	陈皮,茯苓	85	17	茯苓,黄芪	55
5	白术,党参	80	18	白术,当归	54
6	白术,陈皮	78	19	党参,黄芪	54
7	茯苓,党参	74	20	陈皮,党参	53
8	甘草,白术,茯苓	70	21	白芍,茯苓	52
9	甘草,陈皮	69	22	当归,茯苓	52
10	山药,茯苓	68	23	甘草,陈皮,茯苓	51
11	白术,陈皮,茯苓	63	24	山楂,陈皮	50
12	甘草,党参	59	25	甘草,白术,陈皮	50
13	白术,茯苓,党参	59			

表 3 处方中药物组合关联规则(置信度≥0.6)

序号	规则	置信度	序号	规则	置信度
1	白术,陈皮->茯苓	0.807 7	12	白芍->白术	0.674 5
2	甘草,白术->茯苓	0.804 6	13	甘草->白术	0.674 4
3	茯苓,党参->白术	0.797 3	14	白术->茯苓	0.672 3
4	甘草,茯苓->白术	0.777 8	15	茯苓->白术	0.668 5
5	陈皮,茯苓->白术	0.741 2	16	当归->白术	0.658 5
6	甘草,陈皮->茯苓	0.739 1	17	白术,陈皮->甘草	0.641 0
7	白术,党参->茯苓	0.737 5	18	党参->白术	0.634 9
8	甘草,陈皮->白术	0.724 6	19	陈皮->白术	0.634 1
9	甘草->茯苓	0.697 7	20	当归->茯苓	0.634 1
10	山药->茯苓	0.693 9	21	山楂->陈皮	0.617 3
11	陈皮->茯苓	0.691 1	22	白芍->茯苓	0.604 7

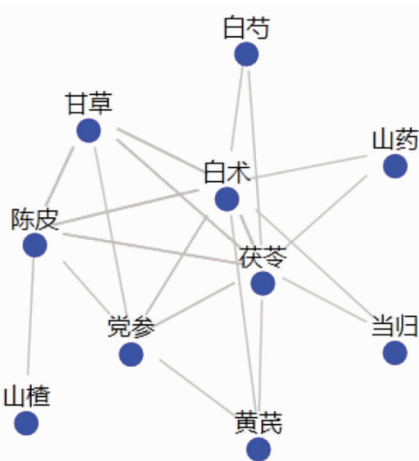


图 1 支持度个数为 50,置信度为 0.6 条件下的关联规则网络展示图

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

依据处方数量,结合经验判断和不同参数提取数据的预读,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 0.02 以上的药对列表,见表 4。

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	陈皮,白扁豆	0.0335	10	木香,陈皮	0.0238
2	陈皮,半夏	0.0314	11	木香,甘草	0.0236
3	白扁豆,山楂	0.0275	12	厚朴,白术	0.0214
4	厚朴,山楂	0.0260	13	木香,肉豆蔻	0.0212
5	陈皮,山药	0.0258	14	木香,丁香	0.0212
6	木香,白扁豆	0.0258	15	厚朴,半夏	0.0210
7	厚朴,六神曲	0.0248	16	麦芽,广藿香	0.0205
8	陈皮,荆芥	0.0241	17	茯苓,芡实	0.0204
9	苍术,滑石	0.0239			

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 13 个新处方,具体见表 6。

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	何首乌,黄芪,当归	黄芪,枸杞子,山茱萸
2	生地黄,知母,天冬	苦参,知母,马齿苋
3	木香,槟榔,滑石	厚朴,木香,砂仁,广藿香
4	甘草,白术,补骨脂	甘草,白术,炮姜
5	川芎,熟地黄,当归	杜仲,续断,熟地黄,当归
6	延胡索,川楝子,颠茄草	延胡索,川楝子,三棱
7	茯苓,陈皮,甘草	郁金,丹参,茵陈,柴胡
8	益母草,蝉蜕,墨旱莲	蝉蜕,石菖蒲,仙鹤草
9	肉桂,人参,附子	肉桂,肉豆蔻,丁香
10	虎杖,白花蛇舌草,蒲公英	虎杖,蒲公英,夏枯草
11	丹参,川楝子,三棱	郁金,丹参,板蓝根,茵陈
12	熟地黄,党参,山药	六神曲,陈皮,麦芽,山楂

表 6 基于熵层次聚类的新处方

序号	新处方
1	何首乌,黄芪,当归,枸杞子,山茱萸
2	生地黄,知母,天冬,苦参,马齿苋
3	木香,槟榔,滑石,厚朴,砂仁,广藿香
4	甘草,白术,补骨脂,炮姜
5	川芎,熟地黄,当归,杜仲,续断
6	延胡索,川楝子,颠茄草,三棱
7	茯苓,陈皮,甘草,郁金,丹参,茵陈,柴胡
8	益母草,蝉蜕,墨旱莲,石菖蒲,仙鹤草
9	肉桂,人参,附子,肉豆蔻,丁香
10	虎杖,白花蛇舌草,蒲公英,夏枯草
11	丹参,川楝子,三棱,郁金,板蓝根,茵陈
12	熟地黄,党参,山药,六神曲,陈皮,麦芽,山楂

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》中健脾类中成药的用药规律。以下结合研究结果,对处方中高频次药物进行分析。茯苓是出现频次最高的单味中药,其味甘淡,性平,归心脾肾经,功能利水渗湿,健脾宁心。常用治疗心脾两虚,气血不足之心悸,失眠,健忘等症。白术出现频次仅次于茯苓,其味甘苦,性温,归脾胃经。功能益气健脾,燥湿利水,安胎。其主归脾胃经,以健脾、燥湿为主要作用,被誉为“脾脏补气健脾第一要药”,常用治脾气虚证,气虚自汗以及脾虚所致的胎动不安等症。甘草其味甘性平,归心肺脾胃经,其善入中焦,具有补气健脾之功,可配伍人参、白术等用于脾气虚弱之证;且有缓急止痛、调和诸药的功效。党参味甘,性平,归脾肺经。其不燥不腻,药力缓和,长于补脾气,益肺气,功似人参而力弱,又善益气生津养血,脾肺气虚及气血不足者宜之。陈皮辛行温通,有行气止痛,健脾和中之功,又苦温而燥,故多用于治疗中焦寒湿脾胃气滞,脘腹胀痛等症。

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“白术,陈皮->茯苓”(置信度为 0.807 7)含义为,处方中有白术、陈皮时,有茯苓的概率为 80.77%;“甘草,白术->茯苓”(置信度为 0.804 6)含义为,处方中有甘草、白术时,有茯苓的概率为 80.46%;“茯苓,党参->白术”(置信度为 0.797 3)含义为,处方中有茯苓、党参时,有白术的概率为 79.73%。又如关联规则网络图中共 10 味药物,这些药物以四君子汤为底方,益气健脾为主,并针对不同的脾虚证进行加减。如加入莲子肉、砂仁、薏苡仁等而成参苓白术散治疗脾虚湿盛证;又如配伍车前子、苍术、柴胡等组成完带汤,治疗肝郁脾虚证等。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,《景岳全书》所说:“或先伤于气,气伤必及于精;或先伤于精,精伤必及于气。”“脾虚及肾”是临床上常见的证型^[10]。如“何首乌、黄芪、当归、枸杞子、山茱萸”的新处方,在补气健脾的基础上,平补肾精肝血。脾喜燥而恶湿,如“木香、槟榔、滑石、厚朴、砂仁、广藿香”的配伍处方,木香健脾消食,槟榔杀虫消积,滑石清热利尿,厚朴下气燥湿,砂仁化湿行气,藿香滑石解暑,故诸药合用既能

(下接第 1610 页)

虚无力推动血行,瘀血内生,或阴虚内热,耗伤津液,血脉为之虚涩而成血瘀。“气为血之帅,血为气之母”,瘀血内阻,则气行受阻,气机郁滞。而郁证的基本病机为肝郁气滞,故消渴患者易生郁证。临床也有调查显示^[14]:糖尿病合并抑郁症患者与非抑郁症患者相比,存在更明显的肝郁气滞,血瘀内生的病机特点。

综上所述,2型糖尿病患者合并抑郁症或抑郁状态的临床危险因素为存在糖尿病并发症、缺乏运动锻炼及高糖化血红蛋白。糖尿病病程较长虽未列入危险因素,但对合并抑郁症仍有负面影响。临床医师在对2型糖尿病患者的诊治过程中应强调血糖控制和运动锻炼的重要性,不仅可预防糖尿病慢性并发症,对预防抑郁症亦有重要意义。对存在上述危险因素,且有情绪障碍的2型糖尿病患者,要早诊断、早干预,积极对其进行心理疏导,加强沟通,同时提高患者对糖尿病的认识,帮助患者建立正确的保健观,减轻或延缓并发症的发生,减少患者抑郁症的发病率。

参考文献

- [1]中华医学会糖尿病分会.中国2型糖尿病防治指南(2013版)[S].中华糖尿病杂志,2014,6(7):447-498.
- [2]Nouwen A, Winkley K, Twisk J, et al. Type 2 diabetes mellitus as a risk factor for the onset of depression: a systematic review and meta-analysis[J]. Diabetologia, 2010, 53(12): 2480-2486.
- [3]Siddiqui S, Jha S, Waghdhare S, et al. Prevalence of depression in pa-

tients with type 2 diabetes attending an outpatient clinic in India[J]. Postgrad Med J, 2014, 90(1068): 552-556.

- [4]Aikens JE, Perkins DW, Piette JD, et al. Association between deression and concurrent Type 2 diabetes outcomes varies by diabetes regimen[J]. Diabet Med, 2008, 25(11): 1324-1329.
- [5]胡随瑜,王素娥,张春虎.抑郁症临床症状分层与中医辨证分型的关系[J].中西医结合学报,2011,9(9):933-936.
- [6]黄洁云.抑郁症的发病机制与治疗进展[J].中国疗养医学,2013,22(3):233-235.
- [7]Fisher L, Skiff MM, Mullah JT, et al. Alouitudinal study of affective and anxiety disorders, depressive affect and diabetes distress in adults with type 2 diabetes[J]. Diabet Med, 2008, 25(9): 1096.
- [8]Bruehl H, Rueger M, Dziobek I, et al. Hypothalamic-pituitary-adrenal, axis dysregulation and memory impairments in type 2 diabetes[J]. Clin Endocrinol Metab, 2007, 92: 2439-2445.
- [9]彭云丽,王雯英,蒋春雷,等.应激诱发抑郁症的细胞因子机制研究进展[J].生理学报,2013,65(2):229-236.
- [10]邱宁岩,李殿富,张静馨,等.2型糖尿病患者垂体-肾上腺轴的变化以胰岛素抵抗[J].放射免疫学杂志,2006,19(6):517-517.
- [11]Lyoo IK, Yoon S, Jacobson AM, et al. Prefrontal cortical deficits in type 1 diabetes mellitus: brain correlates of comorbid depression[J]. Arch Gen Psychiatry, 2012, 69(12): 1267-1276.
- [12]李学岩.大学生身体锻炼对心理健康的影响[J].中国临床康复,2002,6(13):1968.
- [13]钱湘毅,王海琴.社区2型糖尿病患者抑郁和焦虑情绪危险因素流行病学调查[J].中国现代药物应用,2011,5(22):132-134.
- [14]杨晓辉,孙宏峰,吴浩,等.社区2型糖尿病合并抑郁患者中医证候调查[J].北京中医药,2013,32(9):687-689.

(2015-04-15 收稿 责任编辑:王明)

(上接第1606页)

化湿解暑,健脾行气之功。

综上,本研究应用数据挖掘方法对《新编国家中成药》中的健脾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,为健脾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^[14-16]。

参考文献

- [1]刘燕池,雷顺群.中医基础理论[M].北京:学苑出版社,2005:61-62.
- [2]闫丽芳.中医脾病证候的特征症状群研究[D].济南:山东中医药大学,2009:2-3.
- [3]彭颖,李晓波.脾虚证与肠道微生态[J].世界华人消化杂志,2012,20(34):3287-3291.
- [4]唐仕欢,陈建新,杨洪军,等.证-熵-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.

[5]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.

[6]宋民宪,杨明.新编国家中成药[M].2版.北京:人民卫生出版社,2011.

[7]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[J].世界科学技术-中医药现代化,2009,11(2):225-228.

[8]盛倩,刘绍燕,庄曾渊,等.基于中医传承辅助系统的李东垣眼病方剂用药规律研究[J].中国中医眼科杂志,2014,24(3):177-180.

[9]宗鑫,季旭明,魏风琴,等.基于中医传承辅助系统的治疗老年痴呆组方规律分析[J].中国中药杂志,2014,39(4):640-642.

[10]侯丽颖,季幸姝,曾常春.“脾虚及肾”研究思路与方法探索[J].中医杂志,2011,52(20):1736-1738.

[11]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.

(2014-12-17 收稿 责任编辑:徐颖)