基于数据挖掘探索中药复方治疗围绝经期综合征的用药规律

张素，张天嵩，李秀娟，张伟伟，潘宝峰，杨克敏

上海市静安区中心医院中医科，上海 200040

摘要：目的 探索治疗围绝经期综合征中药复方的用药规律。**方法** 检索公开发表的文献中治疗围绝经期综合征的中药复方，建立相应数据库，采用频数分析、聚类分析、关联规则等方法对复方进行数据挖掘，对主要药物、药对(组)规律进行探讨。**结果** 在治疗围绝经期综合征的 49 首中药复方中共用药 100 种、553 频次；以使用频次≥3 的共 57 种、504 频次，作为主要药物进行聚类分析，共分为滋补药、清热药、温补药、补气药、安神药、平肝药、疏肝药、益智药 8 类药物；经关联规则分析，共得药对规则 39 条，药组规则 17 条。**结论** 应用数据挖掘方法探索治疗围绝经期综合征中药复方用药规律是可行的。

关键词： 围绝经期综合征；中药复方；用药规律；数据挖掘

DOI： 10.3969/j.issn.1005-5304.2013.12.008

中图分类号： R2-05；R271.91 **文献标识码：** A **文章编号：** 1005-5304(2013)12-0020-04

Exploration on Regularity of Compound Herbal Formulae for Perimenopausal Syndrome by Data Mining

ZHANG Su, ZHANG Tian-song, LI Xiu-juan, ZHANG Wei-wei, PAN Bao-feng, YANG Ke-min

(Department of TCM, Jing'an District Central Hospital of Shanghai, Shanghai 200040, China)

Abstract: Objective To study the regularity of compound herbal formulae for perimenopausal syndrome. **Methods** Compound herbal formulae published on journals for the treatment of perimenopausal syndrome was searched and collected. A database was established for data mining using frequency count, cluster analysis and association rule analysis. Regularity of key herbs and pair (group) of herbs was summarized. **Results** There were 49 prescriptions (totaling 100 herbs and 553 counts of frequency) in the compound herbal formulae for perimenopausal syndrome. The herbs which were used for more than 3 times included 57 herbs (totally 504 counts of frequency), and they were cluster analysed and classified into 8 clusters including drugs for invigoration, clearing heat, warming and nourishing, restoring vital energy, tranquilizing, calming the liver, soothing the liver, and promoting mentality. The rules in pair of herbs were 39, and rules in group of herbs were 17, by using association rule analysis. **Conclusion** It was feasible that study the regularity of Chinese compound prescription for the treatment of perimenopausal syndrome by using data mining.

Key words: perimenopausal syndrome; compound herbal formulae; regularity; data mining

围绝经期综合征是指妇女绝经前后由于性激素减少所致的一系列躯体及精神心理症状，如月经改变、潮热、汗出、失眠、易激动、抑郁、疲乏、感觉异常等，一般发生在 40～55 岁，但近年来发病年龄有提前倾向[1]。西医治疗本病主要采用雌激素替代疗法(HRT)，或辅以孕激素、雄激素联合用药。虽疗效肯定，但因其阴道出血、乳房胀痛等不良反应及中远期致子宫内膜癌、乳腺癌等事件无法排除，临床使用受到限制[2]。而中药对于本病的治疗不仅效果明显，而且无明显的毒副作用，

基金项目： 上海市静安区卫生系统学科建设项目(JWXK201206)；上海市静安区卫生系统“十百千”人才培养项目(201006C001)

通讯作者： 张天嵩，E-mail: ztsdoctor@126.com

鉴定、验收和评审等是 3 种中药国家科技成果评价的主要形式，占 56.08%；成果推广和应用数均较低说明，成果的转化率较低[4]。

在国家科技成果数据库(知网版)中，记录 3 种中药国家科技成果的数据不完整，特别是有关成果的课题来源、鉴定结果、推广新式、应用状态等项目数据缺失严重，有待完善。

参考文献：

[1] 蔺海明. 甘肃省中药材产业现状与发展取向[J]. 中国现代中药，2011，13(6)：16-19.

[2] 张尚智，韩黎明，贺莉萍. 等. 中国黄芪专利分析与研究[J]. 中国中医药信息杂志，2012，19(11)：19-21.

[3] 张尚智，杨文玺，刘玲玲. 等. 中国当归专利文献学分析[J]. 中国中医药信息杂志，2013，20(2)：33-36.

[4] 高凡珠，刘保延，李振吉. 等. 中医药科技成果概念及其分类评价与转化的研究[J]. 中医药管理杂志，2012，18(12)：1085-1088.

(收稿日期：2013-03-07，编辑：华强)

因此具有广阔的发展前景。本研究旨在通过对治疗围绝经期综合征的中药复方进行数据挖掘,将众多医家治疗该病的用药经验加以整理和归纳,找出隐藏在数据中的规律,发现对该病治疗上的共识,用以指导临床。

1　资料与方法

1.1　数据来源

以“围绝经期综合征”为关键词检索中国生物医学文献数据库(CBMdisc,1978 年 1 月－2012 年 12 月)、万方数据全文数据库(1998 年 1 月－2012 年 12 月)、中国数字图书馆期刊全文数据库(CHKD,2004 年 1 月－2012 年 12 月),为了避免文献漏失,未以“中医药疗法”、“中医疗法”、“中药疗法”等词限定。通过阅读标题和摘要,初步获得有关中医药治疗围绝经期综合征的文献,再仔细阅读全文,将符合以下 3 项条件的中药复方纳入：①围绝经期综合征中医药临床研究；②组方符合中医理论、药味完整、主治明确、临床病例数 30 例及以上；③单纯中药汤剂治疗,中药剂量在中药学规定用量范围内;④文献表明疗效确切者。共获得中药复方 49 个,按照编号、出处、第一作者、方名、药名等依次输入 Microsoft Excel 中,建立数据库。

1.2　数据清洗

选取中药数据库进行数据清洗,根据 2010 年版《中华人民共和国药典》(一部)及《中药学》[3]教材将药名规范化,如将“山茱萸”统一为“山萸肉”,将“生地”统一为“生地黄”,将“龟板”统一为“龟甲”,将“仙灵脾”统一为“淫羊藿”,将“旱莲草”统一为“墨旱莲”等。组合概念拆分,如药味“味甘辛”拆为“味甘”、“味辛”,归经“归肝肾经”拆分为归“肝经”、“肾经”,功效“补益肝肾”拆分为“补肝”和“补肾”等。在新得到的中药数据库中,按《中药学》输入每味药物的性味、归经和功效,并将其数字化。其中寒、热、温、凉、平等药性按蒋氏方法[4]赋值；对于酸、苦、甘、辛、咸等药味及归经、功效等,某药的描述与其某一项相符则计为 1,无则计为 0。由 2 人分别输入数据,完成交叉核对,遇到不一致处,讨论解决。

1.3　数据挖掘

1.3.1　描述性分析　采用频数分析方法,计算药物的种类及每味药的使用频次。

1.3.2　聚类分析　对使用频次超过 3 次以上的药物,按药物的性味、归经、功效主治以分层聚类法进行聚类分析。

1.3.3　关联规则分析　采用关联规则挖掘药对配伍规律,设置最小支持度 20%,最小置信度 50%。

1.4　数据挖掘工具

描述性分析和关联规则分析由 STATISTICA8.0 软件完成,聚类分析由 SPSS17.0 软件完成。

2　结果

2.1　描述性分析结果

49 个复方中共使用 100 种药物、553 频次。其中出现频次≥3 次的有 57 种,由高到低为：白芍(27)、山萸肉(27)、熟地黄(26)、山药(23)、淫羊藿(22)、茯苓(21)、牡丹皮(19)、当归(19)、甘草(18)、女贞子(17)、生地黄(14)、知母(14)、黄柏(13)、柴胡(12)、墨旱莲(12)、枸杞子(12)、牡蛎(10)、菟丝子(10)、巴戟天(9)、浮小麦(9)、龙骨(9)、白术(8)、酸枣仁(8)、仙茅(8)、合欢皮(7)、黄连(7)、何首乌(7)、泽泻(7)、龟甲(6)、杜仲(5)、黄芩(5)、莲子心(5)、香附(5)、百合(4)、大枣(4)、地骨皮(4)、钩藤(4)、黄芪(4)、麦冬(4)、肉苁蓉(4)、夜交藤(4)、郁金(4)、紫贝齿(4)、栀子(4)、阿胶(3)、补骨脂(3)、赤芍(3)、丹参(3)、党参(3)、鹿角(3)、肉桂(3)、桑寄生(3)、太子参(3)、五味子(3)、玄参(3)、远志(3)、茯神(3),共 504 频次。

2.2　聚类分析结果

57 味常用药物共分为 8 类,根据每类药的主要功效加以归类,并将每味药在治疗围绝经期综合征中所应用的主要作用进行整理。结果见表 1。

表 1　57 味药物聚类结果

类别	归类	频次	所含药物	主要功效
1 类	滋补药	118	女贞子、枸杞子、桑寄生、杜仲、山萸肉	阴中之阳,偏补肾气
			熟地黄、阿胶、墨旱莲、龟甲何首乌	阴中之阴,偏填精补血补精益血
2 类	清热药	106	麦冬、百合、莲子心	偏于清心
			牡丹皮、生地黄、赤芍、丹参	偏于凉血活血
			知母、玄参、地骨皮	偏于退虚热
			黄连、黄芩、黄柏、栀子、钩藤	偏于清实热
3 类	温补药	81	当归、肉桂	偏于温通血脉
			菟丝子、补骨脂	偏补肾阳,补肾气
			淫羊藿、仙茅、巴戟天	偏补肾阳
			肉苁蓉、鹿角	偏补肾阳,益精血
4 类	补气药	63	党参、太子参、大枣、白术、甘草、黄芪、山药	益气健脾
5 类	安神药	63	茯苓、茯神、泽泻	偏于健脾
			酸枣仁、夜交藤、合欢皮、浮小麦、紫贝齿	偏于养心
6 类	平肝药	46	龙骨、牡蛎、白芍	平肝,潜阳,敛汗
7 类	疏肝药	21	柴胡、香附、郁金	疏肝,解郁
8 类	益智药	6	五味子、远志	益智,宁心

2.3　关联规则分析结果

药对共得关联规则 39 条,获得药对 28 对：当归、甘草；当归、白芍；当归、熟地黄；当归、淫羊藿；柴胡、白芍；黄柏、知母；淫羊藿、知母；牡丹皮、山药；熟地黄、山药；山萸肉、山药；茯苓、山萸肉；白芍、山萸肉；女贞子、山萸肉；生地黄、山萸肉；牡丹皮、山萸肉；熟地黄、山萸肉；白芍、淫羊藿；女贞子、淫羊藿；黄柏、淫羊藿；茯苓、熟地黄；甘草、熟地黄；白芍、熟地黄；牡丹皮、熟地黄；茯苓、牡丹皮；女贞子、墨旱莲；女贞子、白芍；茯苓、白芍；甘草、白芍。见表 2。

药组(每组 3 味药以上)共得关联规则 17 条,获得药组 5 个:甘草、当归、白芍;淫羊藿、知母、黄柏;山药、山萸肉、牡丹皮;茯苓、山萸肉、牡丹皮;熟地黄、茯苓、山萸肉。见表 3。

表 2 药对关联规则

序号	关联规则	支持度(%)	置信度(%)	提升度
1	甘草→当归	22.45	61.11	1.58
2	当归→甘草	22.45	57.89	1.58
3	当归→白芍	24.49	63.16	1.15
4	当归→熟地黄	20.41	52.63	1.03
5	当归→淫羊藿	20.41	52.63	1.17
6	柴胡→白芍	20.41	83.33	1.51
7	黄柏→知母	20.41	76.92	2.69
8	知母→黄柏	20.41	71.43	2.69
9	淫羊藿→知母	22.45	50.00	1.75
10	知母→淫羊藿	22.45	78.57	1.75
11	牡丹皮→山药	20.41	52.63	1.12
12	熟地黄→山药	26.53	52.00	1.11
13	山药→熟地黄	26.53	56.52	1.11
14	山萸肉→山药	32.65	59.26	1.26
15	山药→山萸肉	32.65	69.56	1.26
16	茯苓→山萸肉	28.57	66.67	1.21
17	山萸肉→茯苓	28.57	51.85	1.21
18	白芍→山萸肉	28.57	51.85	0.94
19	山萸肉→白芍	28.57	51.85	0.94
20	女贞子→山萸肉	22.45	64.71	1.17
21	牡丹皮→山萸肉	26.53	68.42	1.24
22	熟地黄→山萸肉	28.57	56.00	1.02
23	山萸肉→熟地黄	28.57	51.85	1.02
24	白芍→淫羊藿	28.57	51.85	1.15
25	淫羊藿→白芍	28.57	63.64	1.15
26	黄柏→淫羊藿	22.45	84.61	1.88
27	淫羊藿→黄柏	22.45	50.00	1.88
28	茯苓→熟地黄	22.45	52.38	1.03
29	甘草→熟地黄	20.41	55.56	1.09
30	白芍→熟地黄	28.57	51.85	1.02
31	熟地黄→白芍	28.57	56.00	1.02
32	牡丹皮→熟地黄	22.45	57.89	1.13
33	茯苓→牡丹皮	22.45	52.38	1.35
34	牡丹皮→茯苓	22.45	57.89	1.35
35	女贞子→墨旱莲	24.49	70.59	2.88
36	墨旱莲→女贞子	24.49	100.00	2.88
37	女贞子→白芍	20.41	58.82	1.07
38	茯苓→白芍	22.45	52.38	0.95
39	甘草→白芍	26.53	72.22	1.31

表 3 药组关联规则

序号	关联规则	支持度(%)	置信度(%)	提升度
1	白芍,甘草→当归	18.37	69.23	1.78
2	当归,甘草→白芍	18.37	81.82	1.48
3	当归,白芍→甘草	18.37	75.00	2.04
4	黄柏→知母,淫羊藿	18.37	69.23	3.08
5	淫羊藿,黄柏→知母	18.37	81.82	2.86
6	知母→淫羊藿,黄柏	18.37	64.29	2.86
7	知母,黄柏→淫羊藿	18.37	90.00	2.00
8	知母,淫羊藿 →黄柏	18.37	81.82	3.08
9	山萸肉,牡丹皮→山药	18.37	69.23	1.47
10	山药,牡丹皮→山萸肉	18.37	90.00	1.63
11	山药,山萸肉→牡丹皮	18.37	56.25	1.45
12	牡丹皮,茯苓→山萸肉	18.37	81.82	1.48
13	山萸肉,茯苓→牡丹皮	18.37	64.29	1.66
14	山萸肉,牡丹皮→茯苓	18.37	69.23	1.61
15	熟地黄,茯苓→山萸肉	18.37	81.82	1.48
16	山萸肉,茯苓→熟地黄	18.37	64.29	1.26
17	山萸肉,熟地黄→茯苓	18.37	64.29	1.50

3 讨论

数据挖掘主要用于解决“数据丰富,知识贫乏”的状况,已逐渐成为处理海量数据、并从中获取知识、模式和规律的重要工具。该技术已应用于中医药研究,是促进中医药发展、实现中医药现代化的重要组成部分。但数据挖掘只是人类认识和探索世界的一个工具和方法,它可以发现一些潜在的现象,但不会探索其原因。因此,数据挖掘所获得的结果必须在相关领域专家的指导下进行解释和评价[5-6]。

3.1 主要药物分析

聚类分析是根据事物本身潜在的特性从而对研究对象进行分类的方法。通过聚类把一个数据集合中的个体按照某种相似性归成若干类别,使其“物以类聚”生成一系列有意义的子集[7]。本研究将出现频次≥3 次的 57 味药物按性味、归经、功效主治进行聚类分析,共归为 8 类:滋补药、清热药、温补药、补气药、安神药、平肝药、疏肝药、益智药。其特点如下:①滋补药类使用频次最多,且以平补为主,多为助肾阴、填肾精、补肾血为主要治疗目的,如熟地黄、女贞子、何首乌等;②清热药类使用频次居次,又分为 4 个亚组,分别具有清心(麦冬)、凉血(牡丹皮、生地黄等)、退虚热(知母)、清实热(黄连)的功效;③温补药类,少用大热、大辛之药,大多选用甘温润之品,如淫羊藿、菟丝子、巴戟天、肉苁蓉等,可以补肾益气,兼养肾精;④补气药类中,多数医家喜用可补肺、脾、肾的山药;⑤安神药中,可细分为健脾(茯神)、养心(酸枣仁)两个亚类;⑥其他三类则针对心、肝两经,以平肝、疏肝、益智为治疗目的。综上所述,以现代数理方法分析所得到的药物分类与临床实践相一致。

对主要药物选用频次及性味归经、主治功效的分析结果,体现了大多数医家对围绝经期综合征的认识[8-14]。①病位主要

集中在肾、肝、心、脾；②病机属于本虚标实,本虚为肾、肝、心、脾等脏器功能失调,尤以肾虚为主,标实为郁、火、痰湿、瘀血等；③治则以注重滋肾益阴,佐以扶阳,平调肾中阴阳；④用药多攻补兼施,寒热并用,清热不过于苦寒,祛寒不过于温燥,更不可妄用攻伐,以免犯虚虚之戒。

3.2　药对(组)规律分析

两个或多个变量的取值之间存在某种规律性,如“同时发生”或“从一个对象可以推出另一个对象”,称之为关联。关联规则挖掘就是通过关联分析,找出数据库中隐藏的知识,利用所获得的知识可以根据已知情况对未知问题进行推测[15]；在探索药对或药组规律时,应结合关联规则中支持度、置信度、提升度来考虑,如果某一规则同时满足用户自定义的最小置信度和最小支持度阈值,则可认为它是有趣的,而提升度反映了关联规则中元素关系的密切程度,大于 1 才有意义[5]。通过关联规则挖掘所得的结果来看,有经典的药对(组),如柴胡与白芍、黄柏与知母、女贞子与墨旱莲、熟地黄与山药、熟地黄与山萸肉、山药与山萸肉及牡丹皮等；一些是新近出现的药对(组)如淫羊藿与知母、淫羊藿与知母及黄柏等,表明两方面的的信息：①挖掘所得药组和药对主要是经典名方六味地黄丸、知柏地黄丸、逍遥丸,以及新近出现的名方二仙汤等方剂的主要组成药物；②体现了大多数医家重在治肾的学术思想。

由于中医方剂理论体系极其杂,充满海量的信息,而且方剂的效应评价具有多靶点的特征,所以在方剂研究领域,如果能够充分利用信息技术和数据挖掘技术等方法,以现代计算机技术为工具,可以不断获得新的知识,并对其进一步分析和应用验证更为重要,因此本文作为引玉之砖,质之博雅斧正。

参考文献：

[1] 刘春梅,王永宏.围绝经期综合征的流行病学调查现状[J].中医药导报,2010,16(4)：115-117.

[2] 孙津津,周菲菲.中医治疗围绝经期综合征概述[J].中华中医药学刊,2010,28(4)：860-863.

[3] 凌一揆.中药学[M].上海：上海科学技术出版社,1984.

[4] 蒋永光,李力,李认书,等.中医脾胃方配伍规律的数据挖掘试验[J].世界科学技术－中医药现代化,2003,5(3)：33-37.

[5] 张天嵩,张素,李秀娟,等.治疗肺纤维化中药复方用药规律的数据挖掘[J].中国中医药信息杂志,2011,18(2)：31-34.

[6] 张俊美,王娜娜.数据挖掘技术在方剂文献研究中的应用现状[J].甘肃中医,2008,21(1)：5-7.

[7] 张颖,杨钧.数据挖掘在中医药研究中的应用[J].辽宁中医药大学学报,2008,10(3)：153-154.

[8] 王雪华,曲苗,庄颖梅,等.围绝经期综合征的中医理论探析[J].浙江中医药大学学报,2010,34(2)：140-141.

[9] 吴宏进,周昌乐,许家佗.围绝经期综合征中医证候分布特点文献研究[J].中国中医药信息杂志,2012,19(3)：22-25.

[10] 袁烁,邓高丕,叶慧慈.围绝经期综合征的中医证治规律探讨[J].辽宁中医药大学学报,2011,13(2)：119-120.

[11] 蔡芳英.围绝经期综合征证治规律初探[J].新中医,2010,42(4)：6-7.

[12] 唐晓华.浅析围绝经期综合征的中医药防治思路与方法[J].江苏中医药,2010,42(11)：6-7.

[13] 乔平,何贵翔.中医药治疗围绝经期综合征研究进展[J].中医药信息,2011,28(3)：139-141.

[14] 鲁雅娟,魏睦新.中医治疗围绝经期综合征的分型探讨与名家经验[J].中华中医药学刊,2012,30(3)：610-612.

[15] 李文林,赵国平,陆建峰,等.关联规则在名医临证经验分析挖掘中的应用[J].南京中医药大学学报,2008,24(1)：21-24.

(收稿日期：2013-06-13,编辑：华强)

(上接第 6 页)　陈皮君臣相伍以理气行滞、燥湿化痰；加用丹参、赤芍以活血清热,是针对老年肺炎的热、瘀而设[1]。现代研究表明,老年肺炎患者营养不良率高,营养支持治疗是老年患者治疗中不可缺少的重要内容[4]。参苓白术散可以提高老年肺部感染患者血清瘦素水平及体质指数,改善营养不良状况[5]。二陈汤具有镇咳祛痰的作用[6]。活血化瘀中药能降低毛细血管的通透性,促使感染过程终止,有利于组织的修复；另外,尚具有抑制病原微生物生长、解毒消炎作用。

PCT 是降钙素的前肽,在生理情况下,是由甲状腺 C 细胞产生的含量极低的无激素糖蛋白(<0.1 μg/L)。在细菌感染时,血清 PCT 水平会异常增高,当感染控制后,血清 PCT 水平会明显降低。在感染后 3～6 h 即可检测到 PCT 水平升高,目前,国外学者已经将 PCT 作为细菌感染的标志物及预后判断指标[7-8]。本次 PCT 检测结果从实验室角度表明,参苓术散联合二陈汤可以调整老年肺炎患者的炎性介质。

本研究结果显示：尽管临床疗效总有效率 2 组差异无统计学意义(P>0.05),但加用参苓白术散合二陈汤可使老年肺炎患者 PCT 水平快速下降、胸片炎症吸收显著；同时能明显缩短症状、体征消失时间。表明合用参苓白术散、二陈汤对老年肺炎患者具有明显的疗效,值得进一步研究。

参考文献：

[1] 王守富.中医药治疗老年人肺炎的思路与方法[J].中医研究,2008,21(3)：10-12.

[2] 国家中医药管理局.中医病证诊断疗效标准[S].南京：南京大学出版社,1994：5.

[3] 王东梅,张怡,冷建春.中医扶正疗法在老年肺炎治疗中的应用探讨[J].中国中医急症,2012,21(9)：1447-1448.

[4] Pirlich M, Lochs H. Nutrition in the elderly[J]. Best Pract Res Clin Gastroenterol,2001,15：869-884.

[5] 黄少祥,蒋荣民,曲由,等.参苓白术散对 COPD 稳定期营养不良患者血清瘦素的影响[J].天津医药,2012,40(3)：290-291.

[6] 梁中琴,陈星织,王晓霞,等.二陈汤粗粉与二陈汤提取物镇咳祛痰作用比较[J].苏州医学院学报,2000,20(9)：802-803.

[7] 郭靓,王占科.降钙素原生化特征及其临床应用[J].现代诊断与治疗,2009,20(4)：217-219.

[8] 杨滨,康梅.降钙素原在细菌感染性疾病诊断及治疗中的应用[J].现代预防医学,2009,36(3)：596-597.

(收稿日期：2013-05-01,编辑：陈静)