

# 基于数据挖掘的中医治疗糖尿病处方用药规律分析

吴嘉瑞<sup>1\*</sup>, 王凯欢<sup>1</sup>, 纪凯<sup>2</sup>, 张丹<sup>1</sup>, 张晓朦<sup>1</sup>, 杨冰<sup>1</sup>, 俸润彬<sup>1</sup>

(1. 北京中医药大学, 北京 100102; 2. 北京灸道堂中医研究院, 北京 100024)

**[摘要]** **目的:**探讨中医治疗糖尿病的用药规律。**方法:**广泛收集整理中医治疗糖尿病的处方,在构建处方数据库的基础上,采用关联规则算法和熵聚类方法,确定处方中各药物和药物组合的使用频次、药物间关联规则、核心药物组合等。**结果:**高频次药物包括黄芪、茯苓、地黄、山药、丹参等;高频次药物组合包括“丹参、黄芪”,“黄芪、地黄”,“茯苓、黄芪”等;置信度大于0.85的关联规则包括“山茱萸,牡丹皮,茯苓->山药”,“牡丹皮,泽泻,茯苓->山药”,“山药,牡丹皮,泽泻->茯苓”等。**结论:**中医治疗糖尿病多用益气养阴、补脾益肾、清热生津、活血化瘀之品,配伍精巧合法。

**[关键词]** 糖尿病; 关联规则; 聚类算法

**[中图分类号]** R287 **[文献标识码]** A **[文章编号]** 1005-9903(2015)22-0214-04

**[doi]** 10.13422/j.cnki.syfjx.2015220214

**Medication Rules for Treating Diabetes Based on Data Mining** WU Jia-rui<sup>1\*</sup>, WANG Kai-huan<sup>1</sup>, JI Kai<sup>2</sup>, ZHANG Dan<sup>1</sup>, ZHANG Xiao-meng<sup>1</sup>, YANG Bing<sup>1</sup>, FENG Run-bin<sup>1</sup> (1. *Beijing University of Chinese Medicine, Beijing 100102, China*; 2. *Beijing Jiudaotang Academy of Traditional Chinese Medicine, Beijing 100024, China*)

**[Abstract]** **Objective:** To explore the medication rules of traditional Chinese medicine for treating diabetes. **Method:** The traditional Chinese medicine (TCM) prescriptions for treating diabetes were collected to build a database based on TCM inheritance assist system. Association rules algorithm and entropy clustering methods were used to determine the frequency of single drug and drug combinations, the association rules between drugs, core drug combinations and so on. **Result:** Based on the analysis of prescriptions for treating diabetes, the frequently adopted drugs were Astragali Radix, Poria, Rehmanniae Radix, Dioscoreae Rhizoma, and Salviae Miltiorrhizae Radix et Rhizoma. The frequently adopted drug combinations were ‘Salviae Miltiorrhizae Radix et Rhizoma and Astragali Radix’, ‘Astragali Radix and Rehmanniae Radix’, and ‘Poria and Astragali Radix’. In the association rules with a confidence of over 0.85, the drug combinations included ‘Corni Fructus, Moutan Cortex, Poria -> Dioscoreae Rhizoma’, ‘Moutan Cortex, Alismatis Rhizoma, Poria->Dioscoreae Rhizoma’, and ‘Dioscoreae Rhizoma, Moutan Cortex, Alismatis Rhizoma -> Poria’. **Conclusion:** In TCM, the drugs used for treating diabetes mostly have the effects of tonifying Qi and Yin, strengthening spleen and kidney, clearing heat and promoting fluid, and activating circulation to remove blood stasis, which reflects a refined and rational compatibility.

**[Key words]** diabetes; association rules; clustering algorithm

糖尿病(diabetes)一词是描述多病因的代谢疾病,特点是慢性高血糖,伴随因胰岛素分泌或作用缺陷引起的糖、脂肪和蛋白质代谢紊乱,以多饮、多尿、多食及消瘦为主要临床特征,可导致各种组织,特别是眼、肾、心脏、血管、神经的慢性损害或功能障碍<sup>[1-2]</sup>。我国目前将糖尿病共分4大类,即1型糖尿病,2型糖尿病,妊娠糖尿病和特殊类型的糖尿病,其中以2型糖尿病最为常见<sup>[3]</sup>。中国糖尿病流行状况严峻、

未诊断人群比较高、糖尿病的高危人群广,相比较于西医治疗糖尿病的药物常伴不同程度的副作用而言<sup>[4]</sup>,中医药防治糖尿病具有简便验廉的特点<sup>[5]</sup>。复方中药以多环节、多靶点、整体调节的方式在改善糖尿病临床症状和控制其并发症方面发挥着较好的优势<sup>[6]</sup>。中医认为,本病属“消渴”范畴,其中,先天禀赋不足是糖尿病发生的基本原因,后天饮食不节、情志失调、过服温燥药物等是糖尿病发生的重要条

**[收稿日期]** 20150908(026)

**[基金项目]** 国家自然科学基金项目(81473547);北京市中医药科技发展基金项目(JJ-2010-70);北京中医药大学与北京灸道堂中医研究院合作项目(BCUM-JDT2014001);北京市与中央在京高校共建项目(BJGJ1420)

**[通讯作者]** \* 吴嘉瑞, 博士, 副教授, 硕士生导师, 从事临床中药学研究, Tel:010-84738662, E-mail:exogamy@163.com

件<sup>[7]</sup>。中医学将糖尿病分为上、中、下三消,主要认为糖尿病的基本病理是肺燥、肾虚、胃热<sup>[8]</sup>。

面对糖尿病防治这一世界医学难题,中医治疗糖尿病获得了高度重视,而且中医对其病因病机、诊断和辨证论治等方面均有科学的理论体系,但目前仍缺乏对此类处方的深度分析和数据挖掘研究。对此,本研究在广泛收集中医治疗糖尿病处方的基础上,应用中医传承辅助平台软件建立处方数据库,进而采用关联规则 Apriori 算法和复杂系统熵聚类等方法,分析处方中药物的使用频次、药物之间的关联规则、核心组合等,探讨中医治疗糖尿病的用药规律,以期中医临床治疗糖尿病提供理论参考和科学依据。

### 1 资料与方法

**1.1 处方来源与规范化整理** 以《国家级名老中医专科专病丛书——糖尿病验案良方(2版)》<sup>[9]</sup>(书中国家级名老中医均为国家中医药管理局遴选的第一~四批全国名老中医专家学术经验继承工作指导老师)和 CNKI 相关文献中收录的处方为依据(全文检索“糖尿病”及“处方”)。纳入研究的处方共 1 330 首,依据 2010 年版《中国药典》<sup>[10]</sup>和全国中医药高等教育规划教材《临床中药学》<sup>[11]</sup>,对处方中药物名称进行规范整理。

**1.2 分析软件** “中医传承辅助平台系统(V2.0.1)”软件,由中国中医科学院中药研究所提供。

**1.3 数据分析** 应用“中医传承辅助平台(V2.0.1)”软件中“组方规律”模块和“新方分析”模块中嵌入的关联规则和熵聚类算法,设置支持度为 15%,置信度为 0.85,惩罚度为 2,相关度为 8,开展研究<sup>[12-15]</sup>。

### 2 结果

**2.1 用药频次** 处方中使用频次前 30 位的中药具体见表 1,其中使用频次前 5 位的分别是黄芪、茯苓、地黄、山药、丹参。

表 1 处方中使用频次居前 30 位的药物

No.	药名	频数/次	No.	药名	频数/次
1	黄芪	661	16	泽泻	247
2	茯苓	546	17	玄参	239
3	地黄	489	18	知母	236
4	山药	454	19	党参	224
5	丹参	451	20	赤芍	219
6	山茱萸	333	21	桂枝	218
7	当归	331	22	川芎	211
8	黄连	315	23	五味子	190
9	麦冬	311	24	苍术	183
10	白术	306	25	炙甘草	179
11	葛根	299	26	太子参	177
12	天花粉	298	27	牡丹皮	177
13	白芍	271	28	柴胡	176
14	枸杞子	265	29	陈皮	169
15	甘草	260	30	熟地黄	168

**2.2 基于关联规则分析的组方规律** 药物组合出现频次前 3 位分别是“丹参、黄芪”,“山药、丹参”,“丹参、茯苓”,频次前 10 位的药物组合详见表 2。置信度  $\geq 0.85$  的关联规则项见表 3,关联规则网络见图 1。

表 2 处方中频次前 10 位的药物组合情况

Table 2 Top 10 frequently adopted drug combinations prescriptions

No.	药物组合	频数/次	No.	药物组合	频数/次
1	丹参, 黄芪	281	6	白术, 茯苓	218
2	黄芪, 地黄	253	7	山茱萸, 山药	218
3	茯苓, 黄芪	252	8	山药, 地黄	214
4	山药, 茯苓	251	9	当归, 黄芪	213
5	山药, 黄芪	250	10	丹参, 地黄	203

表 3 处方药物关联规则(支持度为 5%,置信度为 0.85)

Table 3 Association rules of drug combinations in prescriptions (Support level = 5%, confidence level = 0.85)

No.	规则	置信度
1	山茱萸, 牡丹皮, 茯苓->山药	0.943 2
2	牡丹皮, 泽泻, 茯苓->山药	0.935 9
3	山药, 牡丹皮, 泽泻->茯苓	0.912 5
4	牡丹皮, 泽泻->山药	0.909 1
5	山茱萸, 泽泻, 茯苓->山药	0.908 2
6	山茱萸, 牡丹皮, 泽泻->山药	0.906 7
7	山茱萸, 牡丹皮, 泽泻->茯苓	0.906 7
8	山茱萸, 山药, 牡丹皮->茯苓	0.902 2
9	白术, 泽泻->茯苓	0.893 3
10	山茱萸, 牡丹皮->山药	0.893 2
11	牡丹皮, 泽泻->茯苓	0.886 4
12	山茱萸, 山药, 泽泻->茯苓	0.881 2
13	牡丹皮, 泽泻, 茯苓->山茱萸	0.871 8
14	玄参, 麦冬->地黄	0.870 7
15	山药, 牡丹皮, 茯苓->山茱萸	0.864 6
16	山茱萸, 牡丹皮->茯苓	0.854 4
17	牡丹皮, 泽泻->山茱萸	0.852 3

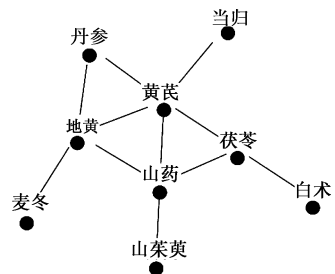


图 1 关联规则网络展示(支持度 15%,置信度 0.85)

Fig.1 Network display of association rules (support level = 15%, confidence level = 0.85)

### 2.3 基于熵聚类的方剂组方规律研究

**2.3.1 基于改进的互信息法的药物间关联度分析** 在惩罚度为 2, 相关度为 8 的研究条件下, 开展聚类分析, 得到处方中两两药物间的关联系数, 关联系数前 30 位的药对见表 4。

表 4 处方中关联系数前 30 位的药对

Table 4 Top 30 herbal pairs with high correlation coefficient in the prescriptions

No.	药对	关联系数	No.	药对	关联系数
1	桂枝, 麦冬	0.018 3	16	知母, 陈皮	0.012 1
2	桂枝, 鸡血藤	0.017 2	17	茯苓, 知母	0.012 1
3	天花粉, 黄连	0.016 5	18	天花粉, 玉竹	0.012 1
4	麦冬, 制附子	0.016 2	19	山药, 天花粉	0.011 9
5	麦冬, 太子参	0.015 4	20	地黄, 山药	0.011 8
6	桂枝, 知母	0.014 9	21	白术, 猪苓	0.011 8
7	地黄, 制附子	0.014 4	22	地黄, 炙甘草	0.011 5
8	桂枝, 山药	0.014 1	23	麦冬, 干姜	0.011 5
9	地黄, 半夏	0.014 0	24	白芍, 天花粉	0.011 4
10	桂枝, 玄参	0.014 0	25	白术, 法半夏	0.011 4
11	地黄, 生姜	0.013 5	26	地黄, 大枣	0.011 3
12	山药, 白扁豆	0.013 5	27	桂枝, 当归	0.011 0
13	茯苓, 法半夏	0.013 5	28	地黄, 陈皮	0.010 9
14	茯苓, 地黄	0.013 2	29	知母, 黄连	0.010 8
15	知母, 麦冬	0.012 7	30	茯苓, 玄参	0.010 8

**2.3.2 基于复杂系统熵聚类的药物核心组合分析** 以改进的互信息法的药物间关联度分析结果为基础, 用相关度和惩罚度为约束, 基于复杂系统熵聚类, 演化出的药物核心组合, 见表 5。

### 3 讨论

中医认为, 糖尿病为本虚标实之证, 燥热为标, 证见口干口苦、多食喜饥、面红目赤、心烦恶热、舌红苔黄等, 宜清热生津; 阴虚为本, 证见五心烦热、腰膝酸软、头晕耳鸣、心悸失眠、舌红少津等, 故治疗多采用益气养阴、调补脾肾之法。本研究应用关联规则和聚类算法分析了中医治疗糖尿病处方的用药规律。结果显示, 高频次药物多具有补气养阴、补益脾肾、活血等功效。如居用药频次首位的黄芪, 性微温, 味甘, 归肺、脾经, 功能补气升阳、固表止汗、利水消肿、托毒生肌, 可有效治疗由于本虚标实所致的气虚津亏等证。另外, 已有药理实验证明黄芪中的黄芪甲苷具有促进糖尿病大鼠血浆胰岛素分泌的作用<sup>[16]</sup>。因此, 黄芪通常被用作治疗糖尿病的首选药, 蜜炙后补气效果更佳。又如, 高频药物中的丹参, 味苦微寒, 归心、肝经, 善于活血化瘀。自古就有“一味丹参, 功同四物”之说, 因其性偏寒, 尤宜于血热郁滞者。从配伍而言, 高频次药对多为高频次中药两两相伍而成, 如山药配伍山茱萸, 山药味甘性平, 入脾、肺、肾经, 有益气养阴、补肺脾肾、固精止带之功; 山茱萸酸涩微温, 入肝、肾经,

表 5 基于复杂系统熵聚类的药物核心组合

Table 5 Core drug combinations based on complex system entropy clustering

No.	核心组合	No.	核心组合
1	瓜蒌, 降香, 薤白	34	瓜蒌, 红曲, 山楂, 清半夏
2	厚朴, 法半夏, 陈皮	35	厚朴, 半夏, 陈皮
3	茜草, 石决明, 三七	36	三七, 谷精草, 木贼
4	白术, 茯苓, 陈皮	37	泽泻, 牡丹皮, 茯苓, 山药
5	白术, 知母, 地黄	38	知母, 地黄, 天花粉
6	地骨皮, 枸杞子, 玉竹	39	女贞子, 枸杞子, 菊花
7	火麻仁, 锁阳, 肉苁蓉	40	火麻仁, 阿胶, 秦皮
8	川芎, 红花, 丹参	41	丹参, 生姜, 大枣
9	柴胡, 黄芩, 黄芪	42	柴胡, 枳壳, 黄芪
10	麻黄, 炮姜, 白芥子	43	白薇, 珍珠母, 炒蒲黄, 炮姜
11	木瓜, 吴茱萸, 没药	44	木瓜, 焦神曲, 焦麦芽
12	桂枝, 地黄, 干姜	45	玄参, 知母, 地黄
13	石榴皮, 炒薏苡仁, 乌梅炭	46	石榴皮, 炒薏苡仁, 炒知母
14	桑葚, 天麻, 荔枝核	47	钩藤, 天麻, 夜交藤
15	苍术, 薏苡仁, 白豆蔻	48	鹿角胶, 紫苏, 白豆蔻
16	制何首乌, 制黄精, 决明子	49	制何首乌, 决明子, 谷精草, 木贼
17	秦艽, 炒杜仲, 羌活	50	独活, 羌活, 防风, 升麻
18	秦艽, 炒杜仲, 威灵仙	51	络石藤, 威灵仙, 熟附子
19	杜仲, 续断, 桑寄生	52	杜仲, 续断, 狗脊
20	焦神曲, 焦山楂, 焦麦芽	53	焦神曲, 焦山楂, 乌梅炭
21	牛蒡子, 土茯苓, 蝉蜕	54	小茴香, 土茯苓, 地榆
22	锁阳, 肉苁蓉, 何首乌	55	沙参, 乌药, 百合, 何首乌
23	豨莶草, 远志, 琥珀	56	远志, 琥珀, 青礞石
24	蛇床子, 覆盆子, 菟丝子	57	萆薢, 蛇床子, 磁石, 覆盆子
25	白芍, 山药, 生姜	58	桂枝, 白芍, 炙甘草, 生姜, 大枣
26	白芍, 当归, 鸡血藤	59	川芎, 红花, 当归, 赤芍, 桃仁
27	侧柏叶, 槐米, 小蓟	60	侧柏叶, 小蓟, 大蓟
28	知母, 当归, 天花粉	61	麦冬, 天花粉, 石膏
29	竹茹, 法半夏, 陈皮	62	竹茹, 枳实, 半夏, 陈皮
30	白扁豆, 莲子, 乌梅炭	63	桑螵蛸, 心木, 白花蛇舌草, 莲子, 沙苑子
31	金银花, 蒲公英, 紫花地丁	64	金银花, 连翘, 穿山甲
32	柏子仁, 炒枣仁, 龙骨	65	炒枣仁, 龙骨, 炒苍术
33	茜草, 珍珠母, 石决明, 墨旱莲	66	白茅根, 小蓟, 墨旱莲, 大蓟

功能补益肝肾、收敛固涩。两药合用, 共奏养阴、平补脾肾之功。关联规则分析显示, 置信度较高的关联规则均为经典名方六味地黄丸中的药物组合, 如置信度为 0.943 2 的规则“山茱萸, 牡丹皮, 茯苓-> 山药”, 又如置信度为 0.935 9 的规则“牡丹皮, 泽泻, 茯苓-> 山药”, 再如置信度为 0.912 5 的规则“山药, 牡丹皮, 泽泻-> 茯苓”等, 六味地黄丸是治疗肾阴虚的经典方剂, 由此可见中医治疗糖尿病以补肾滋阴为

首要治法。

基于复杂系统熵聚类的药物核心组合中分析发现,在治疗糖尿病时,药物组合除益气活血、清热生津类外,还有化湿理气类药物组合,如厚朴-清半夏-陈皮。究其原因,《医学衷中参西录》中曾提到:脾为脾之副脏,脾即为胰。由此,中医认为,胰腺分泌胰岛素的功能可能与脾有关,脾虚则胰腺分泌胰岛素不足导致血糖升高<sup>[17]</sup>。脾喜燥恶湿,应运用化湿理气之品以健脾益气。另外,有一核心组合与经典名方桂枝汤组成一致,即桂枝-白芍-炙甘草-生姜-大枣,其中以桂枝之辛温来发汗解表,以芍药之酸寒和营敛阳,炙甘草合桂枝辛甘化阳以鼓舞卫气、合白芍酸甘化阴以和营益阴,加之生姜温胃止呕,大枣益气补中,诸药合用,通过健运脾胃,调和营卫,滋阴和营以奏功效<sup>[18]</sup>。

#### [参考文献]

[1] 钱荣立. 关于糖尿病的新诊断标准与分型[J]. 中国糖尿病杂志, 2000, 8(1): 62-65.

[2] 司富春, 李亚楠. 糖尿病的中医证型分布和方药规律分析[J]. 中医研究, 2009, 22(8): 13-16.

[3] 中华医学会糖尿病学分会. 中国2型糖尿病防治指南(2010年版)[J]. 中国糖尿病杂志, 2012, 20(1): S1-S37.

[4] 徐斯盛, 张惠斌, 周金培, 等. 新型抗糖尿病药物的研究进展[J]. 中国药科大学学报, 2009, 42(2): 97-106.

[5] 李学军, 杨叔禹. 预防糖尿病的问题与展望[J]. 中华中医药杂志, 2009, 24(8): 981-983.

[6] 姚蓝, 辛小红, 姚雪梅, 等. 古今中医治疗糖尿病用药异同探析[J]. 中国实验方剂学杂志, 2013, 19(18): 342-345.

[7] 栗明, 丁常宏, 方芳. 糖尿病中医治疗进展[J]. 中

药信息, 2012, 29(6): 112-115.

[8] 罗毅, 付玉山. 中医治疗糖尿病用药规律及效果分析[J]. 中医药信息, 2015, 31(1): 118-119.

[9] 徐江雁, 毋莹玲, 杨建宇, 等. 国家级名老中医糖尿病验案良方[M]. 2版. 郑州: 中原农民出版社, 2013.

[10] 国家药典委员会. 中华人民共和国药典. 一部[S]. 北京: 中国医药科技出版社, 2010: 1-286.

[11] 王建, 张冰. 临床中医学[M]. 北京: 人民卫生出版社, 2012: 1-237.

[12] 吴嘉瑞, 郭位先, 张晓朦, 等. 基于数据挖掘的国医大师颜正华含首乌藤处方用药规律研究[J]. 中国中药杂志, 2014, 39(22): 4464-4469.

[13] 郭位先, 吴嘉瑞, 张晓朦, 等. 基于数据挖掘的国医大师颜正华含炒酸枣仁处方用药规律研究[J]. 中华中医药杂志, 2015, 29(1): 62-65.

[14] 吴嘉瑞, 童有健, 张晓朦, 等. 基于关联规则和熵聚类的邓星伯治疗肺系病证用药规律研究[J]. 中国实验方剂学杂志, 2014, 20(7): 223-226.

[15] 郭位先, 吴嘉瑞, 张冰, 等. 基于关联规则和复杂系统熵聚类的颜正华教授治疗血瘀证用药规律研究[J]. 中国实验方剂学杂志, 2014, 20(5): 218-221.

[16] 孙建宁. 中药药理学[M]. 北京: 中国中医药出版社, 2013: 294.

[17] 张永顺. 中医药调治糖尿病[M]. 北京: 人民军医出版社, 2009: 26.

[18] 李静华, 赵玉堂, 郭玉成, 等. 桂枝汤对四氧嘧啶致糖尿病小鼠模型的降糖作用[J]. 承德医学院学报, 2006(2): 115-116.

[责任编辑 邹晓翠]

热烈祝贺中国中医科学院

中药研究所成立60周年!