

· 数据挖掘 ·

《中医方剂大辞典》中治疗肺痿方剂的用药规律分析

李健¹, 卢朋^{1,2}, 张瑞贤¹, 信富荣³, 陈建新⁴, 唐仕欢¹, 杨洪军^{1*}

(1. 中国中医科学院中药研究所, 北京 100700; 2. 中国科学院自动化研究所, 北京 100290;
3. 中国中医科学院中医临床基础研究所, 北京 100700; 4. 北京中医药大学, 北京 100029)

[摘要] 目的: 分析《中医方剂大辞典》中治疗肺痿方剂的组方规律。方法: 应用中医传承辅助系统 (V1.1), 收集、整理治疗肺痿疾病的方剂, 录入系统, 利用熵方法与 apriori 算法进行组方规律分析。结果: 对筛选出的 181 个肺痿处方进行分析, 使用频次在 10 以上的药物有 23 味。apriori 算法组方规律分析得到 42 个常用药对 (支持度个数 ≤ 13), 35 个 3 味药物核心组合 (支持度个数 ≤ 8), 8 个 4 味药物核心组合 (支持度个数 ≤ 6), 对肺痿方剂中药物关联规则进行分析, 支持度个数 ≤ 10 , 置信度 ≤ 0.6 。基于熵方法分析得到 3 味药核心组合 32 个, 熵层次聚类得到 9 对组成新方的核心组合以及 9 个候选新方。结论: 中医传承辅助系统 (V1.1) 是中医方剂用药规律分析的良好工具。

[关键词] 中医方剂大辞典; 肺痿; 中医传承辅助系统; 用药规律

[中图分类号] R283.6 **[文献标识码]** A **[文章编号]** 1005-9903(2012)10-0001-05

Analysis on Composition Principles of Prescriptions for Atrophic Lung Disease in the Book of *〈The Prescription of Traditional Chinese Medicine Dictionary〉*

LI Jian¹, LU Peng^{1,2}, ZHANG Rui-xian¹, XING Fu-rong³, CHEN Jian-xin⁴,
TANG Shi-huan¹, YANG Hong-jun^{1*}

(1. Institute of Chinese Material Medica, Academy of Chinese Medical Sciences, Beijing 100700, China;

2. Institute of Automation, Chinese Academy of Sciences, Beijing 100290, China;

3. Institute of Basic Research In Clinical Medicine China Academy of Chinese Medical Sciences, Beijing 100700, China; 4. Beijing University of Traditional Chinese Medicine (TCM), Beijing 100029, China)

[Abstract] **Objective:** To analyze composition principles of the prescriptions for the treatment of atrophic lung disease. **Method:** prescriptions for atrophic lung disease are collected, sorted and entered into the TCM inheritance system (V1.1) to analyze the composition principles by the methods of entropy and apriori. **Result:** Based on the analysis of 181 cases of prescriptions, twenty-three medications with more than 10 frequencies are mined from the database. Composition principles are obtained with apriori method: forty-two frequently used herbal pair (support ≤ 13); thirty-five core combinations with 3 medications (support ≤ 8); 8 core combinations with 4 medications (support ≤ 6); also the association principles for those medications in the prescriptions are analyzed (support ≤ 10 , confidence ≤ 0.6). Principles are obtained with entropy: thirty-two core combinations with implicit rules of 3 medications; nine new core combinations 9 new prescriptions are found by hierarchical clustering method. **Conclusion:** TCM inheritance system (V1.1) is a useful tool to analyze the composition principles of prescriptions.

[收稿日期] 20120130(135)

[基金项目] 中医药行业科研专项(200907001-5); 中国中医科学院中药研究所基本科研业务费自主选题项目(ZZ20090205)

[第一作者] 李健, 硕士, 实习研究员, 从事本草文献研究, Tel: 010-84035184, E-mail: li10jian@yahoo.com.cn

[通讯作者] * 杨洪军, 博士, 研究员, 从事中药新药设计的方法与技术研究, Tel: 010-84035184, E-mail: hongjun0420@vip.sina.com

[Key words] The Prescription of Traditional Chinese Medicine Dictionary; atrophic lung disease; TCM inheritance system; composition principle

历代医家都对肺痿研究、论述颇多,积累下丰富的诊疗经验。肺痿病名首见于《金匱要略·肺痿肺癰咳嗽上气病脉证治》:“热在上焦者,因咳为肺痿。……寸口脉数,其人咳,口中反有浊唾涎沫者何?师曰:为肺痿之病。”^[1]此后,唐代著名医家孙思邈《备急千金要方》:“病咳唾脓血,其脉数,实者属肺癰,虚者属肺痿”^[2],并载有疗肺痿方。后世医书《外台秘要》、《太平圣惠方》、《圣济总录》等都有相关论述及其诊疗方剂。此类方剂是历代中医治疗肺痿疾病经验的有效载体。本文采用改进的互信息法^[3]、复杂系统熵聚类^[4-5]、无监督的熵层次聚类^[6]等无监督数据挖掘方法和 apriori 算法,对古代肺痿方剂的常用药对、核心组合、组方规律等进行分析,以期对肺痿的中医临床治疗提供参考,并为方剂的用药规律研究提供新的途径。

中医传承辅助系统软件由中国中医科学院中药研究所与中国科学院自动化研究所联合开发。该软件集成了上述算法,具有“数据录入-数据管理-数据查询-数据分析-网络可视化展示”功能。同时,实现了对临床病案信息、疾病信息、证候信息、中药信息、方剂信息等数据的管理、检索和分析,可用于名老中医经验继承、临床医师个体经验总结、方剂组方规律分析及新药处方发现等领域。本文利用中医传承辅助系统(V1.1)对《中医方剂大辞典》中治疗肺痿方剂用药规律进行分析。

1 资料与方法

1.1 处方来源 《中医方剂大辞典》中收录的方剂。

1.2 处方筛选 《金匱要略》中已经较为详细地记载了肺痿的主要症状,包含咳嗽,吐稠痰白沫,或伴有寒热,形体消瘦,精神萎靡,心悸气喘,口唇干燥,脉象虚数等。肺痿一病常续发于其他疾病或经误治之后,因津液一再耗损,阴虚内热,肺受熏灼而致。若病久伤气或肺中虚寒而致者,则表现为阳虚,患者多涎唾,常吐出涎沫而无咳嗽。考虑中医其他疾病可能兼见肺痿疾病的临床症状,因此,本次在上述来源方剂中筛选“主治”或“功用”项中明确记载“肺痿”的方剂。例如:《中医方剂大辞典》^[7]所载:“七宝散,【方源】《直指》卷八。【组成】人参、款冬花、钟乳石、鹅管石(并生研)、明矾(煨)各二钱、辣桂、甘草各一钱。【用法】上为细末。临卧以少许咽下。

【主治】肺痿劳嗽、久嗽”。其中“主治”项中明确记载“肺痿”,故收录。通过筛选,共收集明确治疗肺痿方剂共 181 首。

1.3 处方的录入与核对 将上述筛选的方剂由专人录入中医传承辅助系统。考虑录入过程中可能出现的纰漏,在完成录入后,再由双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。通过“中医传承辅助系统”中“数据分析”模块中“方剂分析”功能,进行分析。

1.4 数据分析

1.4.1 方剂提取 在中医疾病查询项中查询“肺痿”,将治疗肺痿疾病的方剂提取出来。

1.4.2 药物频次统计 将方剂中每味药物的出现频次按照由大到小排序,可导出 Excel 文件。

1.4.3 组方规律分析 进入组方规律分析界面后,首先选择合适的“支持度个数”和“置信度”,然后点击“用药模式”按钮,即可提取组合的常用药对,常用核心组合等,点击“规则分析”按钮,可计算出药物之间的关联关系,上述分析结果可实现网络可视化展示以及 Excel 格式导出。

1.4.4 新方分析 首先进行聚类分析(核心算法包括改进的互信息法^[3]、复杂系统熵聚类^[4-5]),在聚类分析前,选择合适的关系系数和惩罚系数,然后点击提取组合按钮,发现新方(核心算法是无监督的熵层次聚类^[6])。上述分析结果可以实现网络可视化展示及 Excel 格式导出。

2 结果

2.1 用药频次 对《中医方剂大辞典》中治疗肺痿的 181 个方剂包含的 273 种药物,进行“频次统计”,可得到治疗肺痿药物频次从高到低的排序。使用频次在 10 以上的药物有 23 味,见表 1。

表 1 肺痿方剂中频次 > 10 的药物

序号	中药名	频次	序号	中药名	频次	序号	中药名	频次
1	苦杏仁	53	9	桔梗	16	17	石膏	12
2	半夏	46	10	紫苏子	15	18	天南星	11
3	甘草	41	11	雄黄	15	19	五味子	11
4	麻黄	35	12	川贝母	14	20	白术	11
5	陈皮	25	13	细辛	14	21	炙甘草	10
6	茯苓	19	14	桑白皮	13	22	胆南星	10
7	黄芩	19	15	白矾	12	23	沉香	10
8	款冬花	17	16	枳壳	12			

2.2 基于关联规则 apriori 算法组方规律分析 根

据本次治疗肺痿方剂数量和用到的中药个数,结合经验判断以及不同参数提取数据的预读,选择组方中常用药对支持度个数为13,提取出42个常用药对,见表2;选择提取组方中3味药物核心组合支持

度个数为8,提取出35个组合,结果见表3;选择提取4味药物核心组合支持度个数为6,见表4;选择支持度个数 ≤ 10 ,置信度 ≤ 0.6 ,对肺痿方剂进行关联规则分析,见表5。

表2 治疗肺痿疾病组方中的常用药对

药对	频次	药对	频次	药对	频次	药对	频次
炙甘草,生姜	22	天冬,麦冬	18	人参,桔梗	15	五味子,生姜	14
地黄,麦冬	21	大枣,生姜	18	五味子,桔梗	15	炙甘草,紫菀	13
桔梗,贝母	20	人参,炙甘草	17	甘草,麦冬	15	甘草,贝母	13
麦冬,桔梗	20	炙甘草,桑白皮	17	麦冬,桑白皮	15	人参,知母	13
生姜,麦冬	20	苦杏仁,桔梗	17	人参,甘草	14	知母,贝母	13
人参,茯苓	19	生姜,桑白皮	17	桔梗,柴胡	14	苦杏仁,桑白皮	13
桔梗,紫菀	19	人参,桑白皮	16	桔梗,桑白皮	14	苦杏仁,贝母	13
甘草,桔梗	19	炙甘草,桔梗	16	人参,苦杏仁	14	天冬,地黄	13
人参,生姜	19	知母,桔梗	16	人参,麦冬	14	生姜,桔梗	13
炙甘草,麦冬	18	人参,五味子	15	麦冬,款冬花			
五味子,麦冬	18	五味子,桑白皮	15	天冬,桔梗			

表3 治疗肺痿组方中3个药物核心组合

3个药物的核心组合	频次	3个药物的核心组合	频次
天冬,地黄,麦冬	12	炙甘草,生姜,桑白皮	8
炙甘草,生姜,麦冬	11	炙甘草,生姜,柴胡	8
甘草,麦冬,桔梗	11	人参,生姜,茯苓	8
人参,大枣,生姜	10	炙甘草,生姜,赤茯苓	8
大枣,炙甘草,生姜	10	天冬,麦冬,桔梗	8
知母,桔梗,贝母	10	天冬,麦冬,款冬花	8
五味子,生姜,麦冬	9	地黄,麦冬,款冬花	8
人参,生姜,麦冬	9	五味子,麦冬,桔梗	8
人参,炙甘草,生姜	9	甘草,麦冬,贝母	8
地黄,麦冬,桔梗	9	五味子,麦冬,桑白皮	8
麦冬,桔梗,贝母	9	炙甘草,麦冬,桑白皮	8
五味子,桔梗,贝母	9	炙甘草,麦冬,赤茯苓	8
五味子,甘草,桔梗	9	苦杏仁,桔梗,款冬花	8
炙甘草,桔梗,紫菀	9	苦杏仁,桔梗,贝母	8
生姜,麦冬,桑白皮	8	甘草,桔梗,贝母	8
生姜,麦冬,黄芩	8	桔梗,紫菀,贝母	8
大枣,生姜,麦冬	8	人参,茯苓,桔梗	8
人参,五味子,生姜	8		

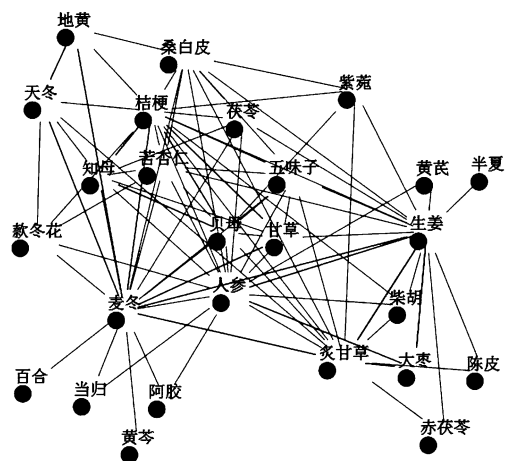
表4 治疗肺痿组方中4个药物核心组合

4个药物核心组合	频次	4个药物核心组合	频次
大枣,炙甘草,生姜,麦冬	7	生姜,麦冬,桑白皮,黄芩	6
炙甘草,生姜,麦冬,赤茯苓	7	天冬,地黄,麦冬,桔梗	6
天冬,地黄,麦冬,款冬花	7	五味子,甘草,麦冬,桔梗	6
生姜,白芍,当归,黄芪	6	五味子,甘草,知母,桔梗	6

关联规则 apriori 算法组方规律分析的上述各项分析结果均可实现网络可视化展示,现仅选取上述肺痿方剂关联规则分析中支持度个数为10,置信度为0.6的分析结果的网络可视化展示,见图1。

表5 肺痿方剂中药物关联规则(支持度个数 ≤ 10 ,置信度 ≤ 0.6)

序号	规则	置信度	序号	规则	置信度
1	大枣->生姜	0.72	10	炙甘草,麦冬->生姜	0.61
2	赤茯苓->生姜	0.67	11	人参,大枣->生姜	0.83
3	半夏->生姜	0.80	12	大枣,炙甘草->生姜	0.91
4	地黄->麦冬	0.72	13	天冬,麦冬->地黄	0.67
5	百合->麦冬	0.65	14	天冬,地黄->麦冬	0.92
6	贝母->桔梗	0.63	15	甘草,麦冬->桔梗	0.73
7	柴胡->桔梗	0.64	16	知母,贝母->桔梗	0.77
8	茯苓->人参	0.61	17	知母,桔梗->贝母	0.63
9	赤茯苓->炙甘草	0.80			

图1 关联规则分析结果的网络化展示
(支持度个数 ≤ 10 ,置信度 ≤ 0.6)

2.3 基于熵方法的新方分析 根据本次治疗肺痿方剂数目,同理结合经验判断和不同参数提取出数据的预读,选择相关系数(correlation)为8,惩罚系数(penalty)为5,进行熵方法聚类分析。

2.3.1 基于熵方法聚类的核心组合分析 按照上

述相关系数与惩罚系数的约束,基于熵方法聚类,挖掘出可能为隐性规律的 3 味药核心组合,共 32 组,见表 6。

表 6 基于复杂系统熵聚类的治疗肺痿的核心组合

序号	核心组合	序号	核心组合
1	麦冬,熟地黄,天冬	17	柴胡,知母,乌梅
2	麦冬,天冬,地黄	18	柴胡,知母,羚羊角
3	麦冬,天冬,百合	19	柴胡,知母,桔梗
4	苦杏仁,贝母,阿胶珠	20	黄连,大黄,香附
5	苦杏仁,蜂蜜,京墨	21	川贝母,诃子,冰片
6	紫菀,旋覆花,白前	22	川贝母,诃子,炙麻黄
7	紫菀,柴胡,桔梗	23	乌梅,诃子,炙麻黄
8	紫菀,赤茯苓,白前	24	蜂蜜,五味子,桑白皮
9	贝母,知母,桔梗	25	生姜,炙甘草,赤茯苓
10	贝母,知母,鹿角胶	26	天冬,款冬花,京墨
11	贝母,桔梗,阿胶珠	27	天冬,款冬花,百合
12	贝母,桔梗,甘草	28	款冬花,薄荷,阿胶珠
13	当归,黄柏,沉香	29	麻黄,五味子,桑白皮
14	当归,黄柏,巴戟天	30	麻黄,白前,厚朴
15	黄芪,人参,茯苓	31	炙甘草,赤茯苓,甘草
16	薏苡仁,百合,北沙参	32	赤茯苓,白前,甘草

2.3.2 基于熵方法层次聚类的新处方分析 在以

上核心组合提取的基础上,点击系统中“提取组合”按钮,通过无监督的熵方法层次聚类算法,可计算得到用于组成新方的 9 对核心组合(表 4)。进一步组合后,可得到 9 个候选新处方(表 7)。利用软件的“网络展示”功能,可以采取不同的网络可视化方式,直观展示出药物不同组合之间的关系,图 2 是用于新方聚类的核心组合的网络可视化的不同展示方式。

表 7 用于新方聚类的核心组合

序号	核心组合 1	核心组合 2
1	苦杏仁,蜂蜜,京墨	蜂蜜,五味子,桑白皮
2	紫菀,柴胡,桔梗	贝母,桔梗,甘草
3	贝母,知母,桔梗	贝母,知母,鹿角胶
4	贝母,桔梗,阿胶珠	柴胡,知母,桔梗
5	柴胡,知母,乌梅	柴胡,知母,羚羊角
6	川贝母,诃子,冰片	川贝母,诃子,炙麻黄
7	生姜,炙甘草,赤茯苓	炙甘草,赤茯苓,甘草
8	天冬,款冬花,京墨	天冬,款冬花,百合
9	麻黄,五味子,桑白皮	麻黄,白前,厚朴

基于熵方法层次聚类,表 7 中的核心组合进一步组合,形成治疗肺痿的新方,结果见表 8,网络可视化展示见图 3。

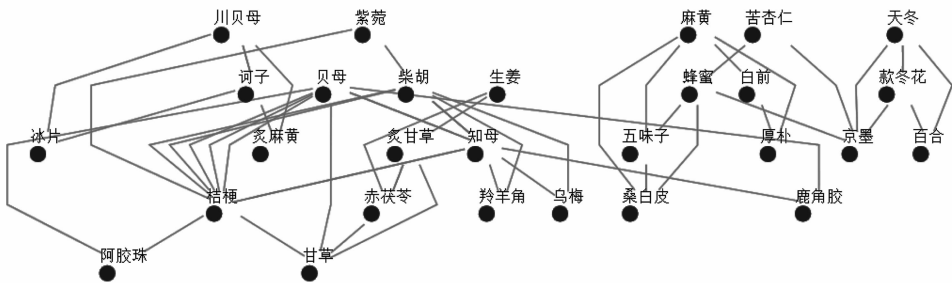


图 2 用于新方聚类核心组合的网络化展示

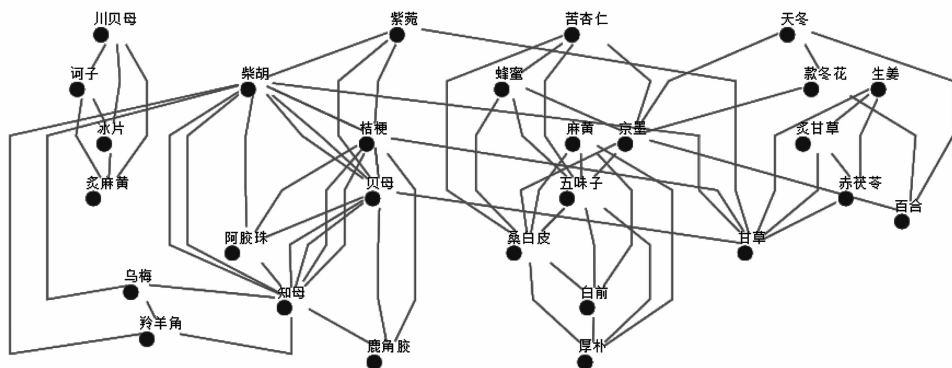


图 3 基于熵层次聚类的治疗肺痿新处方网络化展示

表8 基于熵层次聚类的治疗肺痿新处方

序号	候选新处方
1	苦杏仁,蜂蜜,京墨,五味子,桑白皮
2	紫菀,柴胡,桔梗,贝母,甘草
3	贝母,知母,桔梗,鹿角胶
4	贝母,桔梗,阿胶珠,柴胡,知母
5	柴胡,知母,乌梅,羚羊角
6	川贝母,诃子,冰片,炙麻黄
7	生姜,炙甘草,赤茯苓,甘草
8	天冬,款冬花,京墨,百合
9	麻黄,五味子,桑白皮,白前,厚朴

3 讨论

本文采用“中医传承辅助系统”软件,录入《中医方剂大辞典》中收录的治疗肺痿的方剂,分析得到治疗肺痿方剂中药物的使用频次,利用集成于系统中的 apriori 算法,选取适合的支持度,提取出治疗肺痿疾病的常用药对 42 个、3 味药物的核心组合 35 组、4 味药物的核心组合 6 组,并且进一步揭示了肺痿组方中药物关联规则。

上述基于 apriori 算法的分析结果,是对肺痿疾病组方固有规律、显性规律的总结。肺痿组方用药中一些隐性知识,隐性组合,可能因为其用药频次较少的原因而不能被有效发现。“中医辅助传承系统”软件中“新方分析”模块集成的无监督数据挖掘方法熵方法(改进的互信息法^[3]、复杂系统熵聚类^[4,5]、无监督的熵层次聚类^[6]),即是针对现有统计算法分析的不足,提取低频次高相关的组合,实现对组方中隐性知识的发现,包括隐性药对、隐性核心组合药物之间关联性的定量描述、核心组合演化与提取、新处方发现等功能。“新方分析”模块中,选取适合的相关系数与惩罚系数,基于熵方法聚类,挖掘出可能为隐性规律的 3 味药核心组合,共 32 组,通过熵方法层次聚类,得到用于组成新方的 9 对核心组合,进一步组合后,可得到 9 个候选新方。

上述熵方法层次聚类得到的 9 个候选新方,其在一定程度上亦符合治疗肺痿疾病常用药物规律。例如:表 7 用于新方聚类的核心组合“紫菀,柴胡,桔梗”,通过软件的溯源查询可知,其在 8 个收集的肺痿方中出现。而紫菀、桔梗更是治疗肺痿疾病的常用药对,参看表 2 可知紫菀、桔梗作为药对出现的频次为 19;同理,通过对其他用于新方聚类的核心组合的溯源可以看出,其应用情况在一定程度上反应肺痿疾病的用药规律,这也说明熵层次聚类得到的候选新方有一定的临床应用依据。虽然熵方法聚类挖掘出的隐性知识及候选新处方的临床价值,尚

需要通过溯源、临床医生的判读、临床试验的研究等进一步评判,但此法为临床或基础研究提供了有益的线索。

本文在药物频次、常用药对、核心组合等分析中发现,有炙甘草与甘草两味药物分别出现或同时出现的情况。这是由于从《中医方剂大辞典》中筛选的肺痿方剂,其药物组成中有部分药物附有炮制方法,也有方剂并未记载炮制相关信息,例如入选方剂:“《圣济总录》卷九十三,茯苓汤,【组成】白茯苓(去黑皮),人参,麦门冬(去心,焙),独活(去芦头),槟榔各三分,桂(去粗皮),防风(去叉),防己各一两一分,桔梗(锉,炒),甘草(炙),防葵,枳壳(去瓢,麸炒)各四两,地骨皮十两”和“《宣明论》卷九,人参散,【组成】石膏,甘草各一两,滑石四两,寒水石二两,人参半两”。茯苓汤中甘草录入中医传承辅助系统时,根据其明确记载炮制情况,录为炙甘草;而人参散中甘草与石膏、滑石、寒水石为伍,并不能判断甘草属于炙甘草还是生甘草,从忠实于原著的角度出发,此类未注明炮制情况的均照实录为甘草。所以,前文分析结果有炙甘草和甘草同时出现的情况,由于此类结果属于个别情况,通过进一步的判读可以排除干扰,故并不影响前文分析结果的可靠性。

[参考文献]

- [1] 张仲景著.林亿校正.杨鹏举,侯仙明,杨延巍注释.金匱要略[M].北京:学苑出版社,2007:39.
- [2] 唐·孙思邈著.刘更生,张瑞贤点校.千金方[M].北京:华夏出版社,1993:247.
- [3] 杨洪军,赵亚丽,唐仕欢,等.基于熵方法分析中风病方剂中药物之间的关联度[J].中国中医基础医学杂志,2005,11(9):706.
- [4] 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.
- [5] 陈建新.中医证候的复杂系统建模及其与疾病的相关性研究[D].北京:中国科学院研究生院,2008.
- [6] Yang Hongjun, Chen Jianxin, Tang Shihuan, et al. New drug R&D of traditional Chinese medicine-role of data mining approaches[J]. JBS, 2009,17(3):1.
- [7] 彭怀仁.中医方剂大辞典.第一册.[M].北京:人民卫生出版社,1993:248.

[责任编辑 古云侠]