

基于数据挖掘的中医治疗肝病处方用药规律分析

宋京美^{1,2}, 吴嘉瑞^{1*}, 姜迪², 张丹¹

(1. 北京中医药大学, 北京 100102;

2. 北京市卫生局临床药学研究所, 北京 100035)

[摘要] **目的:**探讨中医治疗肝病的用药规律。**方法:**收集治疗肝病处方,基于中医传承辅助系统软件构建数据库,进而采用关联规则 Apriori 算法和复杂系统熵聚类方法,确定处方中各药物和药物组合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、丹参、茵陈、白术、柴胡等;高频次药物组合包括“丹参、茯苓”,“白术、茯苓”,“丹参、赤芍”等;置信度为 0.6 以上的关联规则包括“黄芪->丹参”,“白术->茯苓”,“黄芪->茯苓”,“党参->茯苓”。**结论:**纳入研究治疗肝病的处方多用健脾渗湿、凉血活血、养阴柔肝、疏肝解郁之品,配伍精巧,符合中医标本兼治的用药原则。

[关键词] 肝病; 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)22-0218-04

[doi] 10.13422/j.cnki.syfjx.2015220218

Medication Rules of Prescriptions of Hepatopathy Based on Data Mining SONG Jing-mei^{1,2}, WU Jia-rui^{1*}, JIANG Di², ZHANG Dan¹ (1. *Beijing University of Chinese Medicine, Beijing 100102, China*; 2. *Institute of Clinical Pharmacy of Beijing Municipal Health Bureau, Beijing 100035, China*)

[Abstract] **Objective:** To investigate the medication rules of traditional Chinese medicines (TCM) for the treatment of hepatopathy by National Medical Masters. **Method:** The prescriptions for hepatopathy were collected to build a database based on TCM inheritance assist system. Apriori algorithm and complex system entropy cluster were adopted to determined the frequency of single medicine and drug combinations that used in prescriptions, and the association rules between drugs e. **Result:** Frequently used herbs included Poria, Salviae Miltiorrhizae Radix et Rhizoma, Artemisiae Scopariae Herba, Atractylodis Macrocephalae Rhizoma and Bupleuri Radix; Frequently used drug combinations included ‘Salviae Miltiorrhizae Radix et Rhizoma and Poria’, ‘Atractylodis Macrocephalae Rhizoma and Poria’ and ‘Salviae Miltiorrhizae Radix et Rhizoma and and Paeoniae Radix Rubra’. The association rules with a confidence of above 0.6 included ‘Astragali Radix-> Salviae Miltiorrhizae Radix et Rhizoma and’, ‘Atractylodis Macrocephalae Rhizoma-> Poria’, ‘Astragali Radix-> Poria’ and ‘Pilosula Codonopsis-> Poria’. **Conclusion:** The prescriptions used for the treatment of hepatopathy mostly have the effects of reinforcing spleen and eliminating dampness, cooling blood and activating blood circulation, nourishing Yin to calm liver and discharging liver and relieving depression, with the refined compatibility and the TCM medication principles for treating symptoms and root causes.

[Key words] hepatopathy; association rules; clustering algorithm

肝病 (hepatopathy) 在现代医学中主要包括病毒性肝病、酒精性肝病、脂肪性肝病、自身免疫性肝病、肝硬化及肝癌等疾病,是临床常见疾病^[1]。肝病初期若不及时治疗会发展

成肝炎、黄疸、脂肪肝、肝硬化腹水、肝囊肿、免疫性肝病、原发性肝癌等。治疗肝病的药物一般按功效可以大致分类为抗病毒药物、护肝和恢复肝功能药物、免疫调节药物、抗肝纤

[收稿日期] 20150817(002)

[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养安堂生物科技有限公司合作横向课题项目(BUCM-YAT-2015001)

[第一作者] 宋京美, 实习研究员, 从事中药学研究, Tel:010-83227445, E-mail:2428083078@qq.com

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

维化药物、中成药及中药各大类^[2]。其中,中医对肝病的病因病机、诊断和辨证论治都有科学的理论体系,《中医辨证学》列举肝气郁结、肝火上炎、肝阴虚、肝阳上亢、肝风内动、肝血虚等 13 类证候,《中医脏腑辨证鉴别诊治手册》分 8 类肝证 30 个证型,各证均有对应方剂;《中医五脏病学》将肝病的常用中药分为疏风、清肝、泻肝、利胆等 15 类,列药物 70 种^[3-4]。

近年来对中医肝病的报道不断增多,但缺乏对此类中药处方用药规律数据挖掘的深入研究,因此,本研究采用数据挖掘的方法对中医治疗肝病的 1 181 首处方进行了用药规律的探讨。纳入本研究 1 181 首治疗肝病的验案良方,选案丰富,按语精详,处方精妙,全面反应了中医治疗肝病独特的辨证治疗思想和用药经验。应用中医传承辅助系统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无监督数据挖掘方法,分析处方中药物的使用频次及药物之间的关联规则等,探讨 1 181 首处方治疗肝病用药规律,以期对中医临床治疗肝病提供理论参考和科学依据。

1 资料与方法

1.1 处方来源 以《中医肝病案例选》^[5],《国医大师验案良方——肝胆肾卷》^[6],《名中医肝病科绝技良方》^[7],《国家级名老中医专科专病丛书——肝病验案良方》^[8],《肝病百草良方》^[9],《肝胆病奇效良方》^[10]为处方遴选来源,纳入研究处方共 1 181 首。

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

1.3 处方的录入和核对 将上述处方录入“中医传承辅助系统(V2.0.1)”。录入完成后,由双人负责数据的审核,以确保数据的准确性。通过“中医传承辅助系统(V2.0.1)”软件中“数据分析”模块中的“处方分析”功能,进行用药规律挖掘。

1.4 数据分析

1.4.1 频次统计分析 将肝病处方中每味药的出现频次从大到小排序,并将“频次统计”结果导出。

1.4.2 处方规律分析 “支持度个数”(表示在所有药物中同时出现的次数)分别设为 120,“置信度”设为 0.4,按药物组合出现频次从大到小的顺序进行排序。

1.4.3 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、复杂系统熵聚类),在聚类分析前,先选择合适的相关度和惩罚度,然后点击“提取组合”按钮,发现新组方。

2 结果

2.1 用药频次 处方中使用频次高于 130 的有 20 味药,使用频次前五位分别是茯苓、丹参、白术、茵陈、柴胡。见表 1。

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“丹参,茯苓”,“白术,茯苓”,“丹参,赤芍”。见表 2。分析所得的用药规则见表 3,关联规则网络图见图 1。

表 1 处方中使用频次 >130 的药物

Table 1 Drugs with usage frequency of above 130 in prescriptions

No.	中药名称	频数/次	No.	中药名称	频数/次
1	茯苓	579	11	当归	281
2	丹参	510	12	陈皮	248
3	白术	480	13	党参	218
4	茵陈	421	14	泽泻	217
5	柴胡	374	15	黄芩	198
6	甘草	372	16	虎杖	190
7	郁金	340	17	鸡内金	169
8	黄芪	338	18	白花蛇舌草	148
9	赤芍	320	19	车前子	142
10	白芍	305	20	猪苓	132

表 2 处方中使用频次 >120 的药物组合

Table 2 Drugs with usage frequency of above 120 in prescriptions

No.	药物组合	频数/次	No.	药物组合	频数/次
1	丹参,茯苓	259	11	丹参,当归	143
2	白术,茯苓	221	12	茯苓,茵陈	138
3	丹参,赤芍	175	13	陈皮,茯苓	137
4	丹参,郁金	174	14	郁金,茵陈	134
5	丹参,茵陈	169	15	郁金,茯苓	134
6	丹参,白术	168	16	当归,茯苓	134
7	丹参,柴胡	165	17	白术,柴胡	130
8	茯苓,柴胡	155	18	郁金,柴胡	130
9	丹参,黄芪	154	19	茯苓,党参	129
10	茯苓,黄芪	145	20	赤芍,茵陈	128

表 3 处方中药物组合关联规则(置信度 >0.4)

Table 3 Association rules of drug combinations in prescriptions (confidence level >0.4)

No.	规则	置信度	No.	规则	置信度
1	黄芪->丹参	0.652 5	12	当归->茯苓	0.505 6
2	白术->茯苓	0.642 4	13	白芍->茯苓	0.500 0
3	黄芪->茯苓	0.614 4	14	白术->丹参	0.488 3
4	党参->茯苓	0.614 2	15	柴胡->丹参	0.483 8
5	赤芍->丹参	0.595 2	16	柴胡->茯苓	0.454 5
6	泽泻->茯苓	0.579 4	17	郁金->茵陈	0.436 4
7	陈皮->茯苓	0.578 0	18	郁金->茯苓	0.436 4
8	郁金->丹参	0.566 7	19	赤芍->茵陈	0.435 3
9	当归->丹参	0.539 6	20	茯苓->白术	0.433 3
10	丹参->茯苓	0.522 1	21	茵陈->丹参	0.430 0
11	茯苓->丹参	0.507 8	22	郁金->柴胡	0.423 4

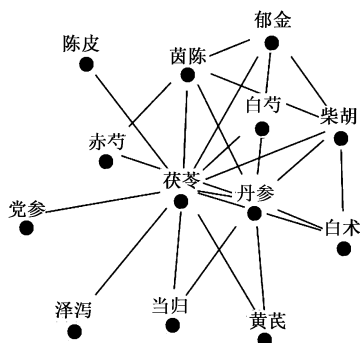


图 1 关联规则网络展示(支持度 120,置信度 0.4)

Fig. 1 Network display of association rules (support level = 120, confidence level = 0.4)

2.3 基于熵聚类的组方规律分析

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 6,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 0.011 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

Table 4 Association analysis between drugs based on improved mutual information method

药对	关联系数	药对	关联系数
石斛,苦参	0.018 8	丹参,茯苓	0.012 0
石斛,仙鹤草	0.018 4	葛根,莪术	0.011 9
红花,茜草	0.018 3	苦参,苍耳草	0.011 6
红花,当归	0.015 7	制大黄,牡丹皮	0.011 4
葛根,桂枝	0.013 5	制大黄,珍珠母	0.011 4
茵陈,大黄	0.012 6	垂盆草,苍耳草	0.011 2
垂盆草,苦参	0.012 6	苦参,虎杖	0.011 1
苦参,鸡骨草	0.012 2	茵陈,垂盆草	0.011 1
茵陈,田基黄	0.012 2	赤芍,地肤子	0.011 1
葛根,牡丹皮	0.012 0	丹参,桃仁	0.011 1

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,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 味药核心组合。见表 5。

2.3.3 基于无监督熵层次聚类的新处方分析 在以上核心组合提取的基础上,运用无监督熵层次聚类算法,得到 10 个新处方。见表 6。

3 讨论

本研究运用关联规则和聚类算法对治疗肝病的用药经验进行了较为深入的分析。研究显示,处方中高频次药物包括茯苓、丹参、茵陈、白术、柴胡等;高频药物组合包括“丹参,茯苓”“白术,茯苓”“丹参,赤芍”等;置信度为 0.4 的关联规则包括:“黄芪->丹参”“白术->茯苓”“黄芪->茯苓”“党参->茯苓”“赤芍->丹参”等;新处方包括“灵芝,枸

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	灵芝,枸杞,徐长卿	11	灵芝,贯众,徐长卿
2	半边莲,半枝莲,胎盘粉	12	半边莲,白芍,牡丹皮
3	黄精,山楂,何首乌	13	山楂,决明子,何首乌
4	葛根,丹参,赤芍	14	白术,茯苓,黄芪
5	生蒲黄,延胡索,小茴香	15	延胡索,牡蛎,川楝子
6	三七,山药,人参	16	白芍,山药,炒党参
7	制附子,炒白术,干姜	17	白术,茯苓,党参
8	沙参,麦冬,地黄	18	枸杞子,麦冬,沙参,地黄
9	生山楂,泽泻,决明子	19	生山楂,泽泻,决明子
10	平地木,田基黄,垂盆草	20	赤芍,垂盆草,鸡骨草

表 6 基于熵层次聚类的新处方

Table 6 New prescriptions based on entropy clustering

No.	新方组合
1	灵芝,枸杞子,徐长卿,贯众
2	半边莲,半枝莲,胎盘粉,白芍,牡丹皮
3	黄精,山楂,何首乌,决明子
4	葛根,丹参,赤芍,白术,茯苓,黄芪
5	生蒲黄,延胡索,小茴香,牡蛎,川楝子
6	三七,山药,人参,白芍,炒党参
7	制附子,炒白术,干姜,白术,茯苓,党参
8	沙参,麦冬,地黄,枸杞子
9	山楂,泽泻,决明子
10	平地木,田基黄,垂盆草,赤芍,鸡骨草

杞,徐长卿,贯众”“半边莲,半枝莲,胎盘粉,白芍,牡丹皮”“葛根,丹参,赤芍,白术,茯苓,黄芪”。组方中频次较高的药物大多具有健脾渗湿、活血通络、凉血清热等功效,这与有关治疗肝病的经典理论的阐释具有较高的吻合度。

同时,本研究在支持度 120,置信度 0.4 条件下所得关联规则网络效果图较好地展示了中药处方治疗肝病的辨证思想。具体来看,网络图中共包括 13 味药物,分别是:茵陈、白芍、丹参、柴胡、郁金、茯苓、白术、黄芪、赤芍、当归、泽泻、党参、陈皮。其中:茵陈味苦、辛,性微寒,归脾、胃、肝、胆经,功效利湿退黄,解毒疗疮,为利胆退黄之要药,亦是中医治疗肝病的常用药;白芍味苦、酸、甘,性微寒,归肝、脾经,功效养血调经,柔肝止痛,敛阴止汗,平抑肝阳;柴胡味苦、辛,性微寒,归肝、胆经,功效疏散退热,疏肝解郁,升阳举陷;郁金味辛、苦,性微寒,归心、肺、肝、胆经,疏肝解郁,行气消胀,祛瘀止痛^[11]。综上,基于关联规则网络图可见,该 1 181 首良方治疗肝病立法以健脾渗湿、凉血活血为主。

同时,本研究基于熵层次聚类方法挖掘出多个新处方,其中多数处方配伍精巧合法。如新处方“葛根,丹参,赤芍,白术,茯苓,黄芪”中,葛根辛凉解肌;丹参活血化瘀以促退黄;赤芍《药品化义》中曰:“味苦能泻,带酸入肝,专泻肝火,养肝藏血”;茯苓、白术、黄芪益气健脾,具有补而不滞,滋而不腻,化湿不伤阴,补中有滋,滋中有消,相得益彰。再如新处方“黄精,山楂,何首乌,决明子”中,黄精补气养阴,健脾益肾;山楂行气散瘀,化浊降脂;何首乌补肝肾、益精血、强筋骨;决明子清肝明目,润肠通便。

综上所述,本研究应用数据挖掘方法对1181首良方治疗肝病用药规律进行了深入研究,获得了既往医案分析较难获得的用药数据和信息。诚然,应用关联规则和聚类研究得到的组方规则、核心组合和新处方等还需进一步临床验证^[12-18],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,希望本研究为中医治疗肝病经验的提炼与升华提供参考。

[参考文献]

[1] 姜喜梅. 肝病中医外治法临床应用综述[J]. 山东中医药杂志, 2015, 34(4): 318-320.

[2] 刘宇, 李东杰. 肝病用药综述[J]. 黑龙江医药, 2007, 20(4): 76-77.

[3] 余世峰, 刘凤斌, 罗世娟. 中医肝病临床疗效评价量表理论结构模型构建的探讨[J]. 中药新药与临床药理, 2010, 21(4): 449-450.

[4] 徐经世. 中医肝病的内涵与外延探析[J]. 中医药学刊, 2005, 23(2): 240-243.

[5] 国家中医药管理局中医肝病重点专科写作组. 中医肝病案例选[M]. 上海: 上海科技教育出版社, 2006: 4.

[6] 吴大真, 李瑶, 杨建宇. 国医大师验案良方肝胆肾卷[M]. 北京: 学苑出版社, 2010: 7.

[7] 吴大真, 李素云, 李奇, 等. 名中医肝病科绝技良方[M]. 北京: 科学技术文献出版社, 2010: 4.

[8] 徐江雁, 刘文礼, 杨建宇, 等. 国家级名老中医肝病验案良方[M]. 郑州: 中原出版传媒集团, 中原农民出版社, 2010: 1.

[9] 王景霞, 吴丽坤, 金磊. 肝病百草良方[M]. 福州: 福建科学技术出版社, 2010: 1.

[10] 谭勇. 肝胆病奇效良方[M]. 北京: 人民军医出版社, 2012: 10.

[11] 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: 116-290.

[12] 高学敏. 中药学[M]. 北京: 中国中医药出版社, 2009: 7-78.

[13] 吴嘉瑞, 张冰, 杨冰. 基于关联规则和复杂系统熵聚类研究颜正华治疗泄泻用药规律[J]. 中华中医药杂志, 2013, 28(8): 2274-2277.

[14] 吴嘉瑞, 张冰, 叶恩培. 基于数据库分析的颜正华消化系统用药总体规律研究[J]. 中国实验方剂学杂志, 2012, 18(2): 242-244.

[15] 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.

[16] 吴嘉瑞, 宋京美, 张冰, 等. 基于数据挖掘的国家级名老中医治疗肝病用药规律研究[J]. 中国中医药信息杂志, 2014, 21(6): 30-33.

[17] 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.

[18] 宋京美, 吴嘉瑞, 姜迪. 基于数据挖掘的国家级名老中医治疗肿瘤用药规律研究[J]. 中国中医药信息杂志, 2015, 22(6): 50-53.

[责任编辑 邹晓翠]