

本文引用:刘安平,刘吉善,沈绍武,于琦,陈刚.基于名老中医刘吉善门诊病案的数据挖掘分析[J].湖南中医药大学学报,2020,40(1):75-79.

基于名老中医刘吉善门诊病案的数据挖掘分析

刘安平¹,刘吉善¹,沈绍武²,于琦³,陈刚^{4*}

(1.十堰市中医医院,湖北 十堰 442012;2.湖北中医药大学,湖北 武汉 430065;3.中国中医科学院中医药信息研究所,北京 100700;4.湖北中医药高等专科学校,湖北 荆州 434020)

〔摘要〕目的 基于国家级知名中医刘吉善4年的门诊病案,分析刘吉善主任医师擅长治疗的病种、遣方用药规律及其学术思想。方法 本研究以2008年至2012年间刘吉善专家门诊病案为样本,采用内容分析方法,确定分析单元,设定分析类目,建立病案数据库,对病案各类目进行频数分析,对单味中药进行聚类分析,对中医病症诊断与方药进行关联分析。结果 共纳入合格病案7873例,通过频数分析统计了刘吉善专家擅长治疗的疾病:不寐(523/6.64)、眩晕(465/5.91)、感冒(380/4.83)等,常用方剂:丹温方(2232/28.35)、藿陈方(1506/19.13)等及常用药:罗汉果(5818/73.90)、炒麦芽(4868/61.83)、炒神曲(4833/61.39)、炒山楂(4833/61.37)、酸枣仁(4690/59.57)、延胡索(4690/59.57)等;聚类分析发现了常用对药:C2<6,7>:酸枣仁、延胡索;C3<11,12>:莱菔子、砂仁;C3<9,16>:葶苈子、王不留行;C3<10,13>:大枣、全蝎;C4<26,39>:凌霄花、木槿花;C4<15,20>:天麻、炒杜仲;C4<31,37>:蝉蜕、僵蚕;C4<27,25>:百合、生地黄;关联分析发现:方剂清暑益气汤与伤暑病、荆防方与粉刺病、藿陈方与咳嗽病、感冒病呈强关联(L>3);丹温方与乳癖、不寐、眩晕、月经过少、痛经、郁证关联强度较高(L>1),对药天麻、杜仲与眩晕呈强关联(L>3),酸枣仁、延胡索与不寐关联强度较高(L>1)。结论 刘吉善专家擅长治疗不寐、眩晕、咳嗽、感冒等疾病,擅长运用清暑益气汤加减治疗伤暑,藿陈方加减治疗感冒病及咳嗽,荆防方治疗粉刺病;擅长运用对药天麻、杜仲治疗眩晕,酸枣仁、延胡索治疗不寐;学术思想主要体现在擅长从肝论治诸多疾病及遣方用药重视顾护脾胃。

〔关键词〕 医案;不寐;眩晕;丹温方;清暑益气汤;天麻;杜仲;酸枣仁;刘吉善

〔中图分类号〕 R249

〔文献标志码〕 A

〔文章编号〕 doi:10.3969/j.issn.1674-070X.2020.01.017

Data Mining Analysis Based on Outpatient Medical Records of Famous and Experienced TCM Doctor LIU Jishan

LIU Anping¹, LIU Jishan¹, SHEN Shaowu², YU Qi³, CHEN Gang^{4*}

(1. Shiyan City Traditional Chinese Medicine Hospital, Shiyan, Hubei 442012, China; 2. Hubei University of Chinese Medicine, Wuhan, Hubei 430065, China; 3 Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China; 4. Hubei College of Chinese Medicine, Jingzhou, Hubei 434020, China)

〔Abstract〕 Objective Based on the 4 years' outpatient medical record of LIU Jishan, to analyze the chief physician LIU Jishan's specialized treatment of diseases, medication rules and academic thoughts. Methods In this study, the clinical records of LIU Jishan's expert outpatients from 2008 to 2012 were used as samples. The content analysis method was used to determine the analysis unit. The analysis categories were set. The medical record database was established. Frequency analysis was performed on various types of medical records. Cluster analysis was performed for single Chinese materia medica, and correlation analysis was conducted to TCM diagnosis and prescription. Results A total of 7873 qualified medical records were included. Through frequency analysis, the diseases that LIU Jishan experts were good at treating (number / frequency%) were: insomnia (523/6.64), dizziness (465/5.91), cold (380/4.83), etc. Common prescriptions were Danwen Recipe (2232/28.35), Huochen Recipe (1506/19.13), etc. and commonly used herbs were *Fructus Momordicae* (5818/73.90), fried *Fructus Hordei*

〔收稿日期〕 2019-06-14

〔基金项目〕 湖北省卫生计生委中医药中西医结合科研项目(LY2017-19);全国名老中医药专家传承工作室建设项目(国中医药人教函[2012]149号);十堰市科学技术研究与开发项目(16Y79)。

〔作者简介〕 刘安平,女,博士,副主任医师,研究方向:中医内科。

〔通讯作者〕 *陈刚,男,博士研究生导师,教授,E-mail:wofavo@qq.com。

Germinatus (4868/61.83), fried *Massa Medicata Fermentata* (4833/61.39), fried *Fructus Crataegi* (4833/61.37), *Semen Ziziphi Spinosae* (4690/59.57), *Rhizoma Corydalis* (4690/59.57), etc.; cluster analysis found common drug pairs: C2 <6,7>: *Semen Ziziphi Spinosae*, *Rhizoma Corydalis*; C3 <11,12>: *Semen Raphani*, *Fructus Amomi Villosi*; C3 <9,16>: *Semen Descurainiae*, *Semen Vaccariae*; C3 <10,13>: *Fructus Jujubae*, *Scorpio*; C4 <26,39>: *Flos Campsis*, *Flos Hibisci L*; C4 <15,20>: *Rhizoma Gastrodiae*, fried *Cortex Eucommiae*; C4 <31,37>: *Periostracum Cicadae*, *Bombyx Batryticatus*; C4 <27,25>: *Bulbus Lili*, *Radix Rehmanniae Recens*; Correlation analysis found that prescription Qingshu Yiqi Decoction and summerheat damage disease, Jingfang Formula and acne disease, Huochen Formula and cough disease, and cold disease were strongly associated ($L>3$); Danwen Formula was strongly associated with breast Bi, insomnia, dizziness, less menstruation, dysmenorrhea, and depression syndrome ($L>1$). Drug pair of *Rhizoma Gastrodiae* and *Cortex Eucommiae* was strongly associated with dizziness ($L>3$). *Semen Ziziphi Spinosae* and *Rhizoma Corydalis* had a stronger association with insomnia ($L>1$). **Conclusion** Expert LIU Jishan is good at treating diseases such as insomnia, dizziness, cough, and cold. He is good at using modified Qingshu Yiqi Decoction to treat summerheat damage, modified Huochen Formula to treat cold and cough, and Jingfang Fang to treat acne disease. He is good at treating dizziness with *Rhizoma Gastrodiae* and *Cortex Eucommiae*, and insomnia with *Semen Ziziphi Spinosae* and *Rhizoma Corydalis*. His academic thoughts are mainly reflected in his ability to treat many diseases from the liver and to take care of the spleen and stomach when giving prescription and medication.

[**Keywords**] medical cases; insomnia; dizziness; Danwen Prescription; Qingshu Yiqi Decoction; *Rhizoma Gastrodiae*; *Cortex Eucommiae*; *Semen Ziziphi Spinosae*; LIU Jishan

名老中医诊疗疾病的临床实录大多存在于门诊病案中,名老中医擅长治疗的疾病及遣方用药规律也蕴含于其中,故门诊病案是研究名老中医临床诊疗经验的重要文献资料。刘吉善主任医师系湖北省中医名师,2012年被国家中医药管理局遴选为第五批全国老中医药专家学术经验继承工作指导老师。刘老师自幼家传习医,从医已有40余年,日诊患者百余例,临床积累了大量的门诊病案。为了传承刘老师临床经验,探析其学术思想,本研究选取了刘吉善专家门诊4年间记录相对连续、完整的门诊病案作为样本,采用内容分析结合数据挖掘技术进行研究。内容分析法是一种主要以文献为对象的研究方法,该方法将非定量的文献资料转化为定量的数据,并依据这些数据对文献内容做出定量分析和做出关于事实的判断和推论^[1]。内容分析法在对门诊病案的内容进行分类中具有独特的优势,数据挖掘技术为大样本文献资料的内容分析提供了方便与快捷,并且拓展了内容分析的深度与广度^[2-5]。本研究运用内容分析法结合数据挖掘技术旨在力求相对准确、客观地分析刘老师擅长治疗的疾病、遣方用药规律及辨证论治思路,同时为名老中医临床经验的传承与研究方法提供更多的实践经验。

1 资料

1.1 样本资料

样本资料系笔者于2008年至2012年跟师期间笔录的刘吉善专家4年的门诊医案,总计收集到符合纳入标准与排除标准医案7 873例。

1.2 纳入标准

(1)有明确的就诊日期、姓名、性别、年龄、症状、证型、方药、剂量;(2)有初诊、复诊记录,且复诊记录有症好转等显示治疗有效的医案;(3)满足前两项要求的初诊记录。

1.3 排除标准

(1)就诊日期、姓名、性别、年龄、症状、证型、方药、剂量任意一项缺失者;(2)无复诊记录或无初诊记录的医案;(3)复诊记录提示无效的医案。

2 方法

2.1 确立分析单元、细化分析粒度

分析样本资料,根据研究目的确立分析单元,选择合适的分析粒度。分析粒度是指对分析单元的细化程度,亦即最小分析单元。如中医诊断类目下的分析单元细化到每一个独立的中医疾病名称;中医证候类目下的分析单元则依据《国家标准·中医临床诊疗术语》^[6]证候部分中每一个独立证候名称为最小分析单元;方剂的分析粒度细化到每一个独立方剂名;中药则以一个单味中药名加剂量及剂量单位组成最小分析单元,如“甘草 6 g”。

2.2 划分分类目录

根据刘吉善专家门诊医案的构成特点,将每则医案划分为:就诊日期、姓名、性别、年龄、中医诊断、西医诊断、中医证候、四诊资料、方剂、单味药物组成及其剂量,并组织专家对分析单元所属类目进行信度测评。

2.3 建立专家医案数据库

以分析类目为字段名,将分析单元归入相应类

目建立专家医案数据库。步骤:首先对医案录入人员进行培训,包括对各类标准的培训及操作培训,然后以“古今医案云平 V1.4”软件(中国中医科学院中医药信息研究所开发)中的医案管理模块为建库工具,新建专家医案数据库,录入专家医案,实行双人双录入,多次校对后锁定备用。

2.4 数据预处理

2.4.1 规范中医名词术语 中医病证名称的规范参照《国家标准·中医临床诊疗术语》疾病部分、证候部分,中医病证名代码参照国家标准《中医病证分类与代码》^[7];方剂名的规范参照第 7 版《方剂学》教材及《中华人民共和国药典》中成药部分,专家经验方保留原方剂名称;中药名的规范参照《中华人民共和国药典》(2015 版)^[8]。中药剂量统一规范为“阿拉伯数字”,剂量单位规范为“g”。

2.4.2 数据清洗 去除重复;清除不构成最小分析单元的其他数据;去除本字段统一分隔符以外的所有符号,包括不可见但实际存在空格、段落符号等。

2.5 数据分析

运用 SPSS 19.0 进行统计分析,对医案各要素进行频数分析,对中医病症诊断与方、药进行关联分析,对单味中药进行聚类分析。关联分析采用 Apriori 算法,设置最小置信度为 0.5,最小支持度为 0.05,同时满足最小置信度与最小支持度的规则为强关联规则。为了检验强关联规则是否具有实际意义,用提升度检验(提升度能够反映关联规则中的前项与后项的相关性,提升度(LIFT)用 L 表示,提升度大于 1 且越高表明正相关性越高,强关联规则具有实际意义。通常情况下前项与后项构成强关联规则,一般要求 $L>3$);聚类方法取 Ward 方法,距离类型设置为 Euclidean 距离。

3 结果

本研究收集了符合纳入标准的病案共计 7 873 例,根据研究的目的,统计分析得出如下数据。

3.1 频数分析结果

3.1.1 中医病症诊断分布情况 刘吉善治疗的疾病中,病名达 816 种,病种涉及内科、妇科、皮外科、儿科、五官科。7 873 例病例资料中:内科病有 3 665 例,占总病例 47%;妇科病有 1 455 例,占 18%;其他如皮外科、儿科、五官科等共 2 753 例,共占总病例的 35%。其中以内科疾病最多见,其次为妇科疾病。内科疾病中又以不寐、眩晕、感冒等为多见;妇科病则以月经失调类疾病多见,如月经过少、月经过

多、崩漏、闭经、痛经、月经先期、月经后期、月经先后不定期共 634 例,共占医案总病例数的 8.1%。频次位居前十位的疾病分布情况,详见表 1。

表 1 频次位居前十的中医病症诊断分布表

中医诊断(TCD*)	频次/例	频率/%
眩晕(BNG070)	523	6.64
不寐(BNX040)	465	5.91
感冒(BNW060)	380	4.83
粉刺(BWP200)	302	3.84
咽菌病(BRA011)	273	3.47
咳嗽(BWF010)	207	2.63
月经过少(BFY050)	188	2.39
虚病(BNV020)	174	2.21
头痛(BNG060)	154	1.96
胃痛(BNP010)	147	1.87

注:*TCD:中医病证分类代码[国家标准]

3.1.2 方药分布情况 纳入研究范围的 7 873 例病例中总用方达 89 首,运用频次较高的有:丹温方、藿陈方、贯煎方、荆防方等。其中以丹温方为主方的病案例数最多,藿陈方次之(详见表 2)。加减用单味药中以罗汉果、炒麦芽、炒神曲、炒山楂、酸枣仁、延胡索等为高频次用药(详见表 3)。

表 2 频次位居前十的方剂分布表

方剂名	频次/例	频率/%
丹温方*	2 232	28.35
藿陈方*	1 506	19.13
贯煎方*	709	9.01
荆防方*	651	8.27
三仁汤	629	7.99
清暑益气汤	453	5.75
甘露饮	377	4.79
加味逍遥散	305	3.87
瓜石汤	158	2.01
香橼汤	147	1.87

注:*经验方

表 3 频次位居前十的单味中药分布表

中药名	频次/例	频率/%
罗汉果	5 818	73.90
炒麦芽	4 868	61.83
炒神曲	4 833	61.39
炒山楂	4 832	61.37
酸枣仁	4 690	59.57
延胡索	4 365	55.44
川贝母	3 973	50.46
铁皮石斛	3 393	43.10
葶苈子	1 974	25.07
大枣	1 899	24.12

3.2 聚类分析结果

对7 873例医案中的单味中药进行聚类分析,发现有价值的对药120余组,但限于篇幅,仅显示频次大于500的前40味中药聚类分析的结果,详见图1。

以分组距离>5为界,可将药物聚类分为4组,具体分组情况如下。

C1:炒山楂、炒麦芽、炒建曲、罗汉果;

C2:酸枣仁、延胡索、川贝母、铁皮石斛;

C3:莱菔子、砂仁、葶苈子、王不留行、大枣、全蝎、枇杷叶、三七;

C4:凌霄花、木槿花、天麻、炒杜仲、玄参、蚕茧、蝉蜕、僵蚕、百合、生地黄、穿山甲、生甘草、黄芩、黄柏、水牛角、白花蛇舌草、生何首乌、白及、马齿苋、鱼腥草、太子参、竹茹、小蓟、茺菢。

以上各类又可以距离>1为界分为若干亚类,如图1所示,其中以两味药为一类的有:C2<6,7>:酸枣

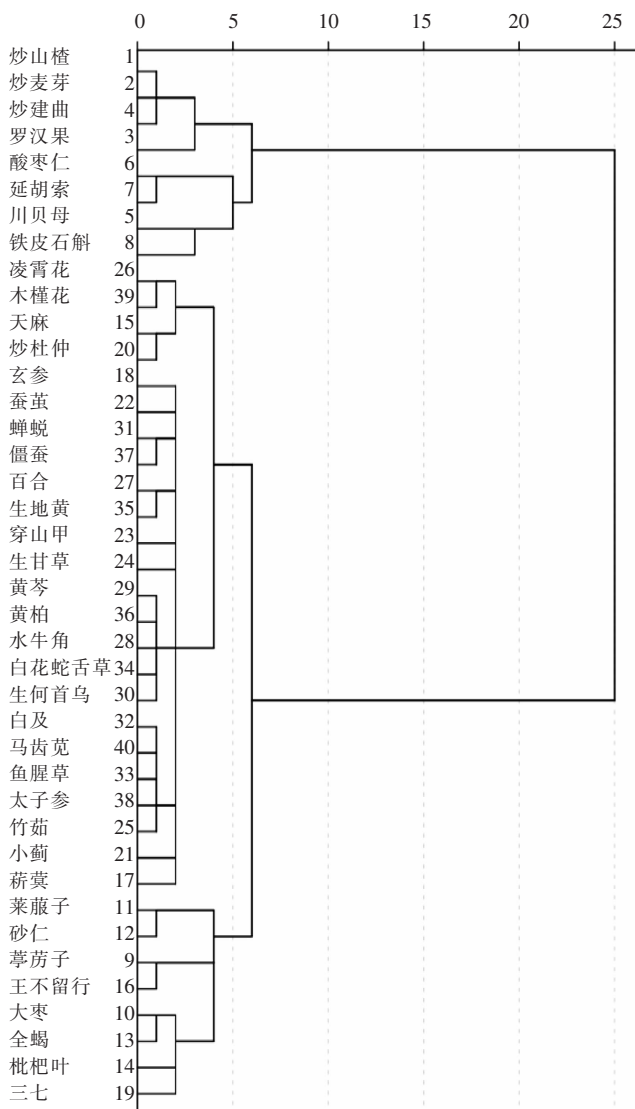


图1 药物聚类树状图

仁、延胡索;C3<11,12>:莱菔子、砂仁;C3<9,16>:葶苈子、王不留行;C3<10,13>:大枣、全蝎;C4<26,39>:凌霄花、木槿花;C4<15,20>:天麻、炒杜仲;C4<31,37>:蝉蜕、僵蚕;C4<27,25>:百合、生地黄。

3.3 关联分析结果

3.3.1 中医病症诊断与方剂 关联分析对中医病症诊断与方剂进行关联分析,发现伤暑与清暑益气汤,粉刺病与荆防方,咳嗽、感冒与藿陈方呈强关联(L>3);乳癖、不寐、眩晕、月经过少、痛经、郁证与丹温方关联强度较高(L>1),详见表4。

表4 中医病症诊断与相关方剂关联分析表 (最小置信度:0.5,最小支持度0.05,L>1)

中医病症诊断(TCD)	方剂名	置信度	支持度	提升度(L)
伤暑(BNW060)	清暑益气汤	1.00	0.05	16.13
粉刺(BWP200)	荆防方	0.73	0.06	10.74
感冒(BNW010)	藿陈方	0.88	0.06	4.49
咳嗽(BNF101)	藿陈方	0.64	0.06	3.27
泄泻(BNP10)	藿陈方	0.56	0.05	2.86
乳癖(BNR050)	丹温方	0.75	0.05	2.06
不寐(BNX040)	丹温方	0.71	0.09	1.95
眩晕(BNG070)	丹温方	0.66	0.05	1.81
月经过少(BFY050)	丹温方	0.63	0.05	1.73
痛经(BFY070)	丹温方	0.53	0.05	1.46
郁证(BNG110)	丹温方	0.50	0.05	1.37

注:TCD为中医证证分类代码[国家标准]

3.3.2 中医病症诊断与单味中药关联分析 对中医病症诊断与单味中药进行关联分析,发现不寐与酸枣仁、延胡索、茺菢关联强度较高(L>1);眩晕与木槿花、凌霄花关联强度较高(L>1),与天麻、杜仲呈强关联(L>3),详见表5。

表5 中医病症诊断与单味中药关联分析表 (最小置信度:0.5,最小支持度0.05,L>1)

中医诊断(TCD)	中药名	置信度	支持度	提升度(L)
眩晕(BNG070)	杜仲	0.66	0.05	5.08
眩晕(BNG070)	天麻	0.69	0.05	4.31
不寐(BNX040)	茺菢	0.54	0.05	1.93
不寐(BNX040)	延胡索	0.90	0.09	1.82
不寐(BNX040)	酸枣仁	0.92	0.09	1.65
眩晕(BNG070)	木槿花	0.69	0.05	1.40
眩晕(BNG070)	凌霄花	0.77	0.05	1.38

4 讨论

4.1 频数分析

从频数分析的结果可以看出,刘吉善专家治疗的病种较为广泛,涉及内科、外科、妇科、儿科等,其

中以内科疾病、妇科疾病为多见,擅长治疗的疾病有不寐、眩晕、咳嗽、月经不调类疾病等。擅长使用的方剂有丹温方、藿陈方、荆防方、贯煎方、三仁汤、清暑益气汤等。

4.2 聚类分析

通过聚类分析频次大于500的前40味中药,分类出4类药物组合(组间距离大于5),其中C1类药物是刘老师用于促进汤药吸收及矫正汤药味道,改善汤药口感的辅助用药。炒麦芽、炒神曲、炒山楂是小方“焦三仙”的药物组成,焦三仙有健胃消食作用,旨在促进汤药吸收。汤药的给药途径主要是口服,经胃肠吸收。此三味药物的使用既保护了脾胃,促进汤药吸收,又增进了患者食欲,体现了刘老师用药顾护脾胃的学术思想。罗汉果的运用除外用其功效,重在用其“味”,罗汉果味甘,却不是普通的甘,其甘甜程度是蔗糖的300倍左右^[9],用之旨在调整汤药的味,使患者服用汤药后不仅疗效好,口感也好,提高了服药的依从性。C2-C4类药物组合主要体现了刘老师擅长治疗的疾病分别是:不寐、咳嗽、眩晕,这在频数分析中已经体现出来,故不赘述。由于该研究运用聚类分析目的在于发现刘老师擅长运用的对药,故对此4类药物再次进行分类,以距离>1为界分为若干亚类,其中以两味药为一类的有:酸枣仁、延胡索;莱菔子、砂仁;葶苈子、王不留行;大枣、全蝎;凌霄花、木槿花;天麻、炒杜仲;蝉蜕、僵蚕;百合、生地黄。酸枣仁与延胡索是刘吉善专家治疗不寐的常用对药;天麻与杜仲、木槿花与凌霄花是治疗眩晕的常用对药;全蝎、大枣是相制配伍,大枣可制约全蝎之毒性;砂仁、莱菔子是治疗腹胀的常用对药;而葶苈子、王不留行是利水通经的常用对药;百合、生地黄常用于阴虚内热精神不定的情志病;蝉蜕、僵蚕则是利咽化痰的常用对药。

4.3 关联分析

对中医病症诊断与单味中药进行关联分析,发现眩晕与天麻、杜仲构成强关联规则,这说明刘老师治疗眩晕的经验对药是天麻、杜仲,这与聚类分析结果相吻合。

对中医病症诊断与方剂进行关联分析,发现伤暑与清暑益气汤,粉刺与荆防方,咳嗽、感冒与藿陈方呈强关联($L>3$),而乳癖、不寐、眩晕、月经过少、痛经、郁证则与丹温方关联强度较高($L>1$)。清暑

益气汤主要用于治疗伤暑,藿陈方主要用于治疗感冒、咳嗽,荆防方主要用于治疗肺风粉刺,故呈强关联。但丹温方的用途较广,常用于证属肝郁化火兼胆郁痰扰的诸多疾病,例如,乳癖、不寐、眩晕、月经过少、痛经、郁证等,故关联强度相对略低,但仍与之呈正相关。“清暑益气汤”出自《脾胃论》,系李杲所创,该方具有清暑益气健脾祛湿之功效,每于夏至以后立秋之前,使用该方频次最高,凡病伤暑,刘老师必用之^[10]。荆防方具有祛风解毒、养阴清热等功效,是刘老师家传经验方,常用于治疗肺风粉刺、风疹等;藿陈方亦是刘老师家传方,主要功效为祛风散寒兼清肺热,常用于治疗咳嗽、感冒等疾病。丹温方由加味道遥散合温胆汤加减化裁而成,加味道遥散适用于肝郁化火证,温胆汤则适用于胆郁痰扰证。刘老师认为肝胆有调节情志及气血津液代谢的作用,“一有拂郁,百病生焉”。运用丹温方重在调节肝胆,从肝论治,使得肝之疏泄与藏血的功能和调,不郁不亢,从而达到“疏其气血,令其调畅,以平为期”的目的,故该方刘老师临床运用较多,常“有是证用是方”,一方广用治疗多种疾病。

参考文献

- [1] 刘燕,邹小娟,陈家旭,等.基于内容分析法的不孕症中医证候文献研究[J].中华中医药杂志,2015,7(30):2345-2348.
- [2] 岳明明,比丽克孜·艾克木,沙塔娜提·穆罕默德.名医扁鹊治规律认识的文献研究的内容分析及聚类分析[J].中华医学图书情报杂志,2017,26(6):26-29.
- [3] 李尚昊,郝琦.内容分析与文本挖掘在信息分析应用中的比较研究[J].图书馆学研究,2015(23):37-42.
- [4] 梁科.数据挖掘技术在中医医案研究中的应用[J].信息与电脑(理论版),2017(24):155-156.
- [5] 滕文静,孙长岗,李雁.浅谈不同中医医案研究方法对临床思维建立的重要性[J].中华中医药杂志,2018,33(3):811-815.
- [6] 国家技术监督局.中华人民共和国国家标准·中医临床诊疗术语[S].北京:中国标准出版社,1997.
- [7] 国家技术监督局.GB/T15657-1995 中医病证分类与代码[S].北京:中国标准出版社,1995.
- [8] 国家药典委员会.中华人民共和国药典(2015年版)[M].北京:中国医药科技出版社,2015.
- [9] 秦莹莹,李啸红.罗汉果甜苷的药理毒理研究进展[J].现代医药卫生,2012,28(18):2794-2796.
- [10] 刘安平,刘吉善.刘吉善主任医师运用清暑益气汤治疗暑病经验[J].西部中医药,2013,26(1):39-40.

(本文编辑 李路丹)