

土家医药半结构化数据的 功效挖掘及可视化研究*

余羊羊¹, 彭芳胜², 田华咏², 曾商禹¹, 郭维加¹, 张 艺³, 刘秀华^{1**}

(1. 成都中医药大学民族医药学院 成都 611137; 2. 湘西自治州民族医药研究所 吉首 416000;

3. 成都中医药大学民族医药学术传承创新研究中心 成都 611137)

摘 要:目的 利用数据挖掘技术对半结构化数据的土家族医药进行品种规范和常用功效的可视化发现。方法 收集整理《中国民族药辞典》、《土家医方剂学》等土家族代表性医药文献,建立土家族医药原始数据库,包含药物拉丁学名、药用部位、性味、毒性、功效、主治疾病等。借助 Visual FoxPro 9.0 软件对土家族医药特性进行频次统计和品种规范,主要以法定标准收录;借助 TCMISS (V2.5)、Cytoscape3.6.1、Gephi0.9.2 等软件对土家药“药物-功效-主治”间的语义关系,进行功效挖掘和可视化展示。结果 筛选土家药共 1453 味,方剂共 771 首,其中土家药主要以“寒(凉)、平”性偏多,苦味药频次最高,其次是辛味药、甘味药;清热败毒、赶风除湿等功效出现频次最高;皮肤疮癣、妇科疾病等是关联程度较大的主治病症。结论 土家族医药的品种规范和“药性-药味-功效-主治”隐形关系可视化分析,可为完善土家药信息化研究奠定基础。

关键词:土家医药 半结构化数据 品种规范 功效挖掘 可视化发现

doi: 10.11842/wst.20190223003 中图分类号: R29 文献标识码: A

土家族是我国人口较多的少数民族,排在全国少数民族人口数的第 7 位。据国家统计局网站公布,2010 年全国第 6 次人口普查数据土家族人口为 835 万。土家族古称“巴人”、“乌蛮”、“板盾蛮”,主要居住在武陵山区的湘鄂渝黔四省市交界处。因其独特的生态环境,药材丰富,品种达 2600 余种,较常用及有特色的土家药约有 600 个品种,目前医疗机构常用土家药有 170 余个品种,被称为“华中的天然药库”^[1]。19 世纪 80 年代,田华咏、万定荣等民族医药工作者开展了土家药匠(医生)的用药经验、单方验方、药材标本等土家医药的田野调查和发掘整理工作,并对收集的手抄本、“口述”资料进行总结整理,相继出版了《土家族医药学》^[2]《医疗机构处方常用土家药手册》^[3]等专著。2018 年 7 月 10 日,中国民族医药学会官网发布了 236 条“土家医药名词术语”、8 种土家医疗效评价标准

及诊疗指南等技术标准的通知。为推动土家医医师资格“开考”工作,湘西土家族苗族自治州民族中医院开展了特色常用土家药的炮制规范、土家医药名词术语、土家医医疗技术等医药技术标准与规范研究。赶酒火疗法、滚蛋疗法等土家族特色医疗技术的整理总结,在很大程度上补充了民族地区现代医疗活动中的不足,为患者提供了方便、有效的治疗方法^[4]。据 2019 年《中国民族药辞典》最新民族药品种统计结果,含有确切拉丁名的土家药有 1453 种,排在全国民族药第 2 位。

半结构化数据是介于结构化数据(如关系数据库、面向对象数据库)和非结构化数据(如声音文件、图像文件等)之间的数据形式,让“正常人不必费力获取”人们快速获取所需的信息^[5]。本文针对半结构化土家医药文献数据信息提取,构建土家医药原始数据

收稿日期:2019-07-10

修回日期:2019-08-11

* 国家科技部国家重点研发计划项目(2017YFC1703900):民族医药发掘整理与学术传承研究,负责人:张艺。

** 通讯作者:刘秀华,讲师,博士,主要研究方向:民族医药传承与发展以及《黄帝内经》理论、临床研究。

1932〔Modernization of Traditional Chinese Medicine and Materia Medica—World Science and Technology〕

库,并通过对土家药的品种整理和“药物-功效-主治”隐性关系分析,对土家医药的民间用药经验进行了初步探究。

1 资料与方法

1.1 构建土家医药原始数据库

根据土家医药数据特性确定其字段和编目的提取方法,建立土家医药原始数据库,包含拉丁学名、别名、性味、毒性、功效、主治疾病等数据信息。为确保数据源准确,安排两人校对相同数据并审核。同时,利用VFP9.0(Visual FoxPro 9.0)软件编写程序,将校对的土家医药“标准化”数据录入数据库。

1.2 数据来源

根据《中国民族药辞典》^[6]、《土家医方剂学》^[7]、《土家医治毒药物集》^[8]、《土家族医药学》^[2]、《医疗机构处方常用土家药手册》^[3]等专著,筛选得到土家药共1453味,方剂共771首。

1.3 数据标准化

在校对土家医药数据源的过程中发现,土家医药仍存在药名不规范现象。首先,对已收集的土家医药的药物名称、处方等信息进行整理和数据清洗,然后参照《中国民族药辞典》^[6]和《医疗机构处方常用土家药手册》^[3]中的药材基原记载,对原始数据库中的药名进行规范化处理。对于方剂中无法根据基原记载进行药名规范的药物,仍保留其原有药名。

1.4 数据分析

半结构化数据是相对结构化数据而言的,是伴随着互联网应用产生的一种新的数据形式^[9]。本文从半结构化数据的土家医药文本转化为结构化数据,并建立土家医药原始数据库。同时,利用中医传承辅助系统(Traditional Chinese Medicine Inheritance Support System,TCMISS)(V2.5)的“数据分析”模块中的“方剂分析”功能,对土家医临床常用方剂进行分析;借助Cytoscape3.6.1和Gephi0.9.2软件,对土家药“药物-功效-主治”的隐性关联关系进行可视化和知识发现分析。

2 结果

2.1 土家药品种整理

利用VFP9.0构建土家医药原始数据库,对关键字段“标准”进行识别和提取。筛选得到1454味土家药,其中九头狮子草 *Peristrophe japonica* (Thunb.)Bremek.

六月雪 *Serissa serissoides* (DC.) Druce.等土家药被地方标准收载。《湖北省中药材质量标准》(2009版)、《湖南省中药材质量标准》(2009版)和《贵州省中药材、民族药材质量标准》(2003版)各收载土家药49种、27种及43种(表1)。本文将最具代表性且被标准收录的18种土家药通过Cytoscape 3.6.1软件将其“药性-药味-功效”进行网络化展示(图1)。土家药主要以“寒(凉)、平”性偏多,苦味药频次最高,其次是辛味药、甘味药;清热败毒、赶风除湿等功效出现频次最高。

2.2 常用土家医方剂关联分析

利用TCMISS(V2.5)对《土家医方剂学》中771首方剂的临床常用方剂“表出剂”、“清败剂”、“松祛剂”、“撒开剂”等158首方剂进行关联分析。将“支持度”设为2%，“置信度”设为0.9时,得到核心组合60个,其中前20排名的核心药对(表2,图2)。土家医常用方剂中主要以散热发表、疏风发表的土家药偏多,其次是清热败毒(清热解毒)、赶风胜湿(祛风除湿)等功效。

2.3 土家药主治病症统计

在整理土家医药文献过程中发现,土家医药对“疾病”的认识多为应用经验,一种土家药治疗疾病从一种到十几种不等。对关键词“主治病症”进行识别和提取关键词,得出其主治病症的统计表(表3),可以看出治疗皮肤疮癣、妇科疾病、蛇虫咬伤的土家药较多。

2.4 土家医“药物-主治”关联度分析

土家族地区以植物药为主,草本植物居首位,多为鲜用。土家族医民间用药经验丰富,总结出了许多关于药物防病治病的谚语、口诀和歌谣,如药物功能的口诀有“打得地上爬,离不开四两麻。打在地下困,离不开五虎进”^[7]。本文就“药物”、“主治”字段提取信息,利用Gephi0.9.2软件从土家药的“药物-主治”角度进行关联分析两者之间的关系是十分必要的(图3)。从图中看出,字段节点越大,代表其出现关联程度越大;线条越粗,代表两个关键词节点之间的关联关系越深。以“皮肤湿疹”、“跌打损伤”、“咽喉肿痛”等疾病与药物的关联程度较为密切。

2.5 土家医“功效-主治”关联度分析

在整理土家医药文献时,发现对土家药的功效和应用主治有大量的文字记录。本文就“功效”、“主治”字段提取信息,利用Gephi0.9.2软件对土家药的“功效-主治”之间的关联关系进行分析(图4)。从图中可以看出,以“清热败毒”功效为中心的线条最粗,说明

表 1 最具代表性 18 种土家药统计分析

序号	药名	基原	别名	药性	药味	毒性	功效	标准
1	八月瓜	三叶木通 <i>Akebia trifoliata</i> (Thunb.) Koidz.	八月炸、 八月札	凉	甘、辛		疏肝理气、益肾利水	[10]
2	白三七	大叶三七 <i>Panax japonicus</i> C.A. Mey. [<i>Panax pseudoginseng</i> Wall. var. <i>japonicus</i> (C. A. Mey.) Hoo et Tseng	竹节参、 竹节三七	温	辛、苦	毒	补益强体、活血止血、消肿止痛	[10]
3	大血藤	大血藤 <i>Sargentodoxa cuneate</i> (Oliv.) Rehd. et Wils.	红藤、 活血藤	温	苦、甘		清热败毒、活血止血	[10]
4	七叶一枝花	七叶一枝花 <i>Paris polyphylla</i> Sm. 或狭叶重楼 <i>Paris polyphylla</i> Sm. var. <i>stenophylla</i> Franch.	白蚤休、 铁灯台	寒	苦	小毒	清热败毒、消肿止痛、凉肝定惊	[10]
5	隔山消	耳叶牛皮消 <i>Cynanchum auriculatum</i> Royle ex Wight	白首乌	温	甘、苦	毒	健胃导滞、排毒利湿、补虚生乳	[11-13]
6	糯稻根	糯稻 <i>Oryza sativa</i> L. var. <i>glutinosa</i> Matsum	楞谷根	平	甘		益胃生津、清热解毒	[11-13]
7	九头狮子草	九头狮子草 <i>Peristrophe japonica</i> (Thunb.) Bremek.	辣子七、 青泽兰、	凉	辛、苦		清热败毒、消肿止痛、凉肝定惊	[11-13]
8	仙桃草	蚊母草 <i>Vernicia peririna</i> L.	接骨仙桃草	平	甘、辛		活血止血、消肿止痛	[11-13]
9	龙葵	龙葵 <i>Solanum nigrum</i> L	野辣椒、野 茄子	寒	苦	小毒	清热败毒、利尿消肿。	[11-13]
10	八角枫	瓜木 <i>Alangium platanifolium</i> (Sieb. et Zucc.) Harms 或八角枫 <i>Alangium chinense</i> (Lour.) Harms	白龙须、 九胡子	温	辛	毒	赶风除湿、舒筋活络、消肿止痛	[11-13]
11	满天星	天胡荽 <i>Hydrocotyle sibthorpioides</i> Lam.	落得打、 雨点草	凉	辛、苦		清热败毒、利尿通淋、明目退翳	[11-13]
12	牛血莲	薯蓣 <i>Dioscorea cirrhosa</i> Lour.	朱砂莲、 朱砂七	平	苦、涩	小毒	活血止血、止泻止痢	[11-13]
13	马兰	马兰 <i>Kalimeris indica</i> (L.) Sch.-Bip	马兰草、 鱼鳅串	凉	辛、苦		清热败毒、活血止血、消积利尿	[11-13]
14	白四块瓦	宽叶金粟兰 <i>Chloranthus henryi</i> Hemsl. 或多穗金粟兰 <i>Chloranthus multistachys</i> Pei	四块瓦、 白蛇吐箭	温	辛、苦	毒	赶风除湿、活血散瘀	[11-13]
15	吉祥草	吉祥草 <i>Reineckia carnea</i> (Andr.) Kunth	蛇尾七、 竹叶七	凉	甘		清肺止咳、活血止血	[11-13]
16	白英	白英 <i>Solanum elaeagnifolium</i> Thunb. 或千年不烂心 <i>Solanum chaynwm</i> c. Y. Wu et. S.C.Huang	排风藤、 白毛藤、	寒	苦	小毒	清热败毒、赶风除湿	[11-13]
17	六月雪	白马骨 <i>Serissa serissoides</i> (DC.) Druce	白鸽骨、 路边荆	凉	辛		清热败毒、健脾利湿、疏肝活血	[11-13]
18	八角莲	八角莲 <i>Dysosma versipellis</i> (Hance) M. Cheng ex Ying	叶下花、 八角金盘	凉	辛、苦	小毒	清热败毒、化痰散结、消肿止痛、 活血通络	[11-13]

“清热败毒”与咽喉肿痛、疮痈肿毒等多种主治疾病相关联,这也与土家医对药物和疾病的认识相关。名老土家医彭芳胜在《土家医毒气病学》^[14]书中写道,“土家医认为毒邪为百病之首,万恶之源。把毒邪分为天毒、生毒、蔫毒 3 大类,风毒、寒毒、热毒、湿毒、火毒、脓毒等 18 种。”祛风除湿、利水毒等功效的关联度排名稍偏前。而在主治疾病中,“疮痈肿毒”和“痢疾”与功效之间的关联度最高,说明土家医治疗“疮痈肿毒”和“痢疾”的疾病数量最多。同时痔疮、湿疹、月经不调等疾病也是关联数量较多的疾病。通过对土家药“功效-主治”隐性关系分析,可以加强对土家族地区多发

病的认识,完善相应的医疗技术。

3 讨论

土家医药研究始于 1978 年,我国土家医药专家学者对土家药的历史、资源、用药经验进行发掘整理及开发利用,取得了较好成效^[15]。土家医药工作者就历代文献和“口述”资料进行整理,并对土家医药民间用药经验进行田野调查,总结出土家医药对药物的分类、命名、性味、功效、主治应用等土家药理学理论^[16]。杨娜娜^[17]通过对湘西自治州民族医卫生资源调查,发现

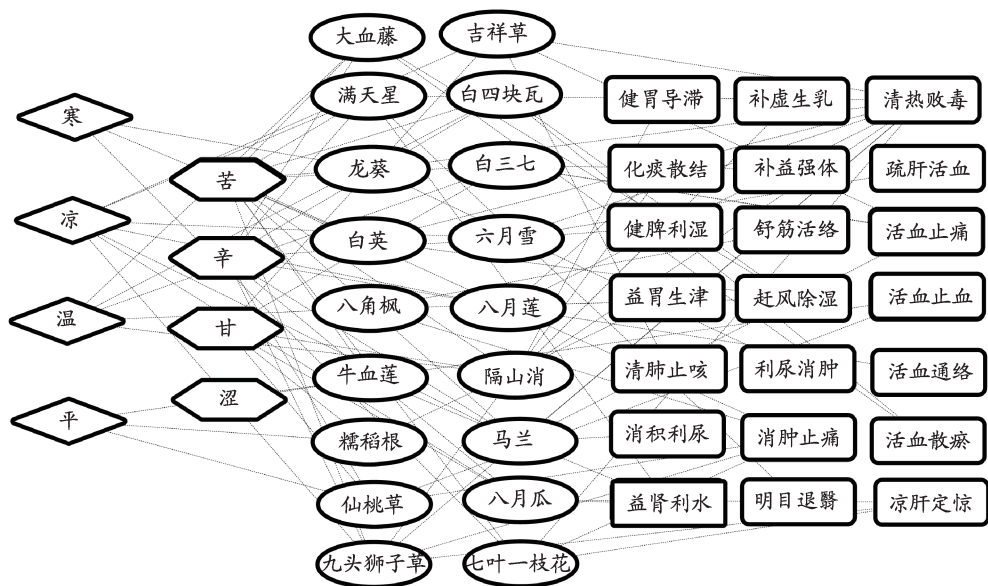


图1 最具代表性18种土家药“药性-药味-功效”网络图展示

注：◇代表“药性”、◇代表“药味”、○代表“药物”、□代表“功效”。

其土家族等民族医的传承主要依靠“口耳相传”，师承和祖传的学习方式占全州民族医学习方式的88.85%，其中84%无传承人，且仅62人存有手抄本资料或者古书典籍。民族医学传承人越来越少、传承人所习得的知识和经验也越来越少。许多民间名医积累的独特医疗经验都随时间而消失，使民族医药文化的流失等问题随之出现，保护和传承民族医药知识迫在眉睫而任重道远。

本文以“半结构化数据”的土家医药资料为基础，参照最新出版且权威的《中国民族药辞典》和《医疗机构处方常用土家药手册》两本专著进行了土家医药数据的规范化处理。同时借助VFP9.0构建土家医药原始数据库，对“关键字段”进行识别和提取，并对其进行土家药的品种规范和功效挖掘。据统计，《湖北省中药材质量标准》(2009版)、《湖南省中药材质量标准》(2009版)和《贵州省中药材、民族药材质量标准》(2003版)各收载土家药49种、27种及43种。利用Cytoscape 3.6.1软件，对地方标准收载最多的18种土家药进行“性-味-药-功效”网络图可视化，表明土家药主要以“寒(凉)、“平”性偏多，苦味药频次最高，其次是辛味药、甘味药；清热败毒、赶风除湿等功效出现频次最高。同时，借助TCMISS (V2.5)对临床常用158首方剂进行关联分析，其中核心药对频次最高是“水灯草-水竹叶”“水灯草-车前草”。水灯草 *Juncus effusus* L.具有走表利窍、清血热的功效，主治鼻塞、流泪等疾

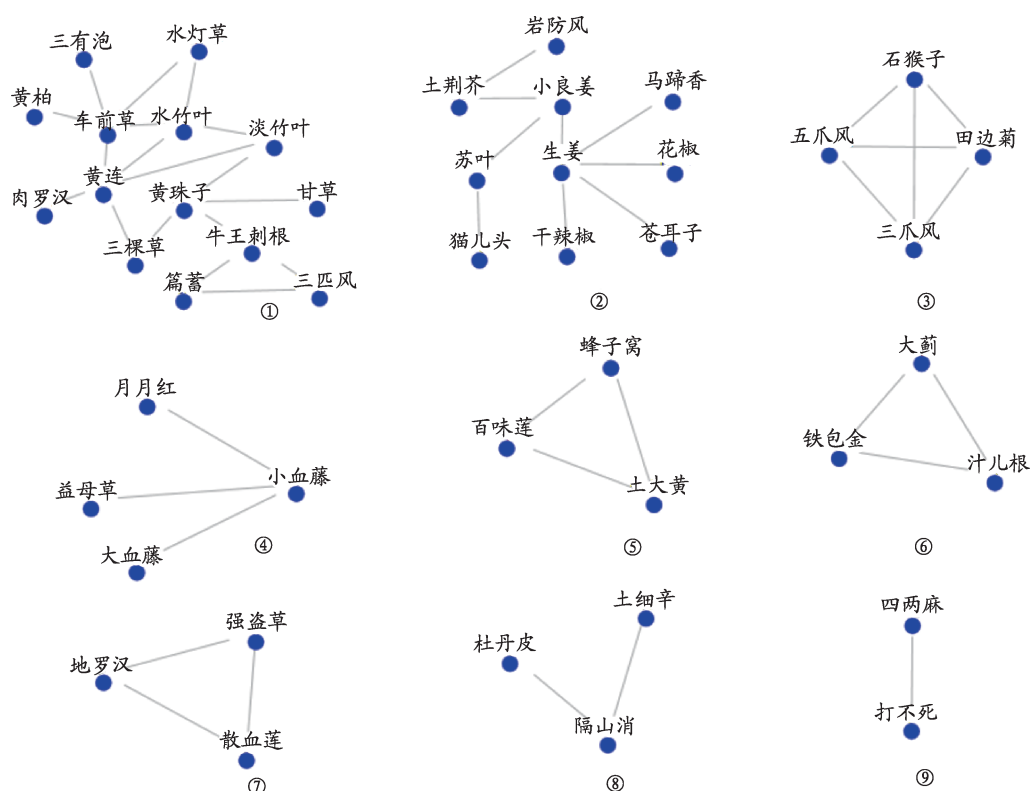
表2 土家医临床方剂中的常用药物配伍
(支持度 ≥ 2%,置信度 ≥ 0.9)

序号	药物配伍	频次	序号	药物配伍	频次
1	水灯草-水竹叶	4	11	四楞草-四块瓦	2
2	水灯草-车前草	4	12	土荆芥-岩防风	2
3	生姜-花椒	3	13	田边菊-三爪风	2
4	小良姜-苏叶	2	14	田边菊-五爪风	2
5	生姜-小良姜	2	15	田边菊-石猴子	2
6	小良姜-土荆芥	2	16	淡竹叶-水竹叶	2
7	苏叶-猫儿头	2	17	黄连-水竹叶	2
8	生姜-马蹄香	2	18	水竹叶-车前草	2
9	生姜-辣椒	2	19	大蓟-鱼腥草	2
10	生姜-苍耳子	2	20	鱼腥草-铁包金	2

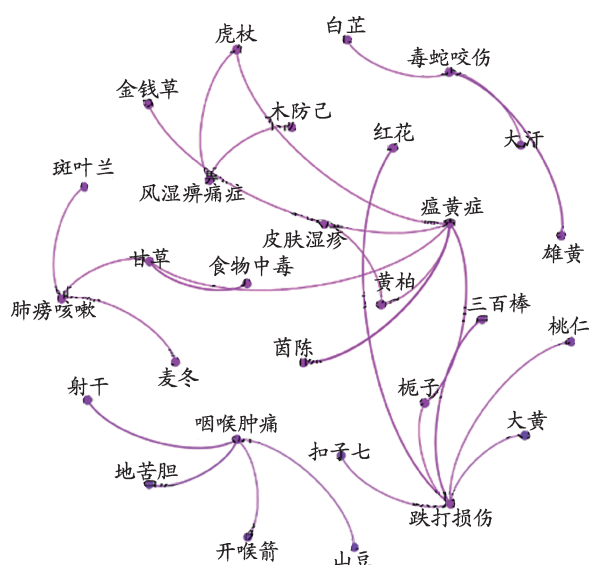
表3 土家医主治病症统计表

序号	主治病症	频次	频率/%	序号	主治病症	频次	频率/%
1	皮肤疮癖	40	10	8	感冒	16	4
2	妇科疾病	36	9	9	跌打损伤	16	4
3	蛇虫咬伤	35	9	10	外伤出血	16	4
4	肛肠	29	7	11	食物中毒	16	4
5	痢肿	22	5	12	风湿	15	4
6	肿痛	20	5	13	寄生虫病	13	3
7	痢疾	19	5	14	热毒	13	3

病^[5]；水竹叶 *Murdannia triquetra* (Wall.) Bruckn. 具有清热败毒、利尿通便的功效，主治咽喉肿痛、肺热咳嗽等疾病^[5]；车前草 *Plantago depressa* Willd. 具有清热明目、利尿的功效，主治咳嗽、小便不利等疾病^[5]。通过“主



注:○代表常用药物配伍网络图的功效字段;①散发表 ②疏风发表 ③清热败毒 ④赶风湿湿 ⑤消痈排脓 ⑥退内腑火 ⑦活血化痰 ⑧温化水湿 ⑨活血散瘀。



治病症”字段检索,获悉土家族人民生活中皮肤疮癣、妇科疾病以及蛇虫咬伤等疾病最为出现,这与土家族生活在武陵山区息息相关。在提取“药物-主治”、“功效-主治”两组数据时,可以知晓土家医在用药和

治疗疾病时,射干 *Belamcanda chinensis* (L.) Redoute、地苦胆 *Tinospora sagittata* (Oliv.) Gagnep. 等清热败毒药常与咽喉肿痛疾病关联程度高;虎杖 *Reynoutria japonica* Houtt 等赶风胜湿(祛风除湿)药常与风湿疾病关联程度高。

本研究在构建土家医药原始数据库时,发现土家药还存在“同物异名”“同名异物”的情况,如三叶青为葡萄科植物三叶崖爬藤 *Tetrastigma hemsleyanum* Diels & Gilg 的干燥块根,别名是石猴子、金钱吊葫芦,但在《医疗机构处方常用土家药手册》^[2]中,白药子为防己科植物金线吊葫芦 *Sphonin arnho* Hayata 的干燥块根。利用半结构化土家医药文献数据构建数据库,虽然可以减少人工录入的工作量,但在快速提取不同版本土家医药文献时,忽视了不同地区土家族对药物使用和疾病治疗的差异性,尚不能较好的反映当地实际医药卫生情况,还仍需向权威的土家族民族医药专家进行咨询。土家族民间用药经验历史悠久,目前的数据挖掘技术对土家族用药经验的探究还处在起步阶段,但为完善土家药信息化研究提供了新的方向,促进了土家医药的传承和发展。

Yu Yangyang¹, Peng Fangsheng², Tian Huayong², Zeng Shangyu¹, Guo Weijia¹, Zhang Yi³, Liu Xiuhua¹

(1. *Ethnic Medicine School, Chengdu University of Traditional Chinese Medicine, Chengdu*

611137, China; 2. *Institute of Ethnic Medicine of Xiangxi Autonomous Prefecture, Jishou*

416000, China; 3. *Ethnic Medicine Academic Inheritance Innovation Research Center,*

Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China)

Abstract: Objective The data mining technology was used to visually discover the variety specification and common efficacy of Tujia medicine with semi-structured data. Methods The representative medical documents of Tujia such as *Chinese National Medicine Dictionary* and *Tujia Medical Formulary* was collected and sorted out to establish the original database of Tujia medicine, including Latin scientific name, medicinal position, sexual taste, toxicity, efficacy, and indication of diseases, etc. With the help of Visual FoxPro 9.0, researching the frequency statistics and variety specifications of Tujia medicine was made. With the help of TCMISS (V2.5), Cytoscape 3.6.1, Gephi0.9.2 and other software, researching the semantic relationship between Tujia medicine "medicine-efficacy-indication" was explored and visually displayed. Results A total of 1453 Tujia medicines with 771 prescriptions were screened, of which Tujia medicines were mainly cold (cool) and mild, bitter medicines had the highest frequency, followed by pungent medicines and sweet medicines. The efficacy of clearing away heat and toxic materials, expelling wind and dehumidifying had the highest frequency. Skin sores and tinea, gynecological diseases and so on were the main indications with a greater degree of correlation. Conclusion The visualization of the hidden relationship between Tujia medicine variety specification and "Property-Taste-Efficacy-Indication" can provide an effective reference. It laid a foundation for the improvement of Tujia medicine information research.

Keywords: Tujia medicine, Semi-structured data, Variety specification, Efficiency mining, Visual discovery

(责任编辑: 周阿剑, 责任译审: 邹建华)