

基于关联规则的中医药知识发现

——以肺病医案为例

马捷^{1,2} 王馨梓^{1*} 何美慧¹

(1. 吉林大学商学与管理学院, 吉林 长春 130022; 2. 吉林大学信息资源研究中心, 吉林 长春 130022)

摘要: [目的/意义] 中医医案作为重要的医疗资源含有大量的数据与信息, 深度挖掘中医医案, 围绕医案内容进行知识发现, 有利于隐性知识的发掘, 提高中医临床诊疗的效率与医疗服务的精准性。[方法/过程] 以实地肺病医案为数据源, 基于统计分析与关联规则的方法对中医医案内容进行挖掘, 以咳嗽为主证为出发点, 挖掘相关中医用药规律与临床症状以及症—药规律, 探究关联规则分析在中医药知识发现方面的创新应用。[结果/结论] 本地治疗咳嗽的高频药物包括炙甘草、杏仁、半夏、川贝母、茯苓、陈皮等, 舌诊症状以苔薄白、苔薄黄、舌红、舌暗红、有齿印等为常见。基于 Apriori 算法的关联规则分析显示常见的药物配伍为丹参, 桔梗—炙甘草、柴胡, 桔梗—炙甘草、前胡, 枳壳—桔梗等; 舌诊症状以薄淡黄—舌暗红、有齿印—舌暗红、苔薄腻淡黄—有齿印等组合为强关联组合, 苔薄淡黄—舌暗红症状组合的强关联药物有桔梗、炙甘草、丹参、白前等。咳嗽患者的类型有外感咳嗽与内伤咳嗽, 涉及风热证、痰湿证、阴虚证、肺气虚证、痰热壅肺证等证型。

关键词: 关联规则; 知识发现; Apriori 算法; 中医医案; 用药规律

DOI: 10.3969/j.issn.1008-0821.2024.03.004

[中图分类号] G250.7 [文献标识码] A [文章编号] 1008-0821 (2024) 03-0034-13

Knowledge Discovery Based on Association Rules

——Taking Pulmonary Disease Medical Cases as an Example

Ma Jie^{1,2} Wang Xinzi^{1*} He Meihui¹

(1. School of Business and Management, Jilin University, Changchun 130022, China;

2. Information Resource Research Center, Jilin University, Changchun 130022, China)

Abstract: [Purpose/Significance] As an important medical resource, TCM (Traditional Chinese Medicine) medical records contain a large amount of data and information. Deep mining of TCM medical records and knowledge discovery around the contents of medical records are conducive to the mining of tacit knowledge, improving the efficiency of TCM clinical diagnosis and treatment and the accuracy of medical services. [Method/Process] The study took the medical records of local lung disease medical records as the data source, the content of traditional Chinese medicine medical records was mined based on statistical analysis and association rules. Starting from cough as the main symptoms, the relevant traditional Chinese medicine medical patterns, clinical symptoms, and symptom drug patterns were explored, attempted to explore the innovative application of association rule analysis in discovering traditional Chinese medicine knowledge discovery. [Result/Conclusion] The local high-frequency drugs for cough include Radix Glycyrrhizae Preparata, almond, Pinellia ternata, Fritillaria cirrhosa, Poria cocos, dried tangerine peel, etc. The common symptoms of tongue diagnosis are thin white coat-

收稿日期: 2023-03-22

基金项目: 吉林大学“中国式现代化道路”与“人类文明新形态”哲学社会科学研究创新团队项目“推进人类数字生态文明: 数字化转型视域下的数据价值与数据创新”(项目编号: 2022CXTD20)。

作者简介: 马捷 (1973-), 女, 教授, 博士生导师, 研究方向: 知识组织与管理、多系统信息协同。何美慧 (2000-), 女, 硕士研究生, 研究方向: 数字信息资源管理。

通讯作者: 王馨梓 (1999-), 女, 硕士研究生, 研究方向: 数据挖掘与知识发现。

ing, thin yellow coating, red tongue, dark red tongue, tooth printed tongue, etc. Association rule analysis based on the Apriori algorithm shows that common drug combinations include *Salvia miltiorrhiza*, *Platycodon grandiflorum*-*Radix Glycyrrhizae* Preparata, *Bupleurum chinense*, *Platycodon grandiflorum*-*Radix Glycyrrhizae* Preparata, *peucedanum* root, *Fructus aurantii*-*Platycodon grandiflorum*. The symptoms of tongue diagnosis are dark red tongue with thin yellow coating, dark red tongue with tooth marks, thin greasy yellow fur coating with tooth marks and other combinations as the strongly correlated combination. The strongly correlated drugs of thin yellow fur dark red tongue symptom combination are *Platycodon grandiflorum*, *Radix Glycyrrhizae* Preparata, *Salvia miltiorrhiza*, and *Cynanchum glaucescens*. The types of cough patients include the type of external cough and internal injury cough, which involve wind-heat syndrome, phlegm-dampness syndrome, yin deficiency syndrome, lung qi deficiency syndrome, phlegm heat obstructing lung syndrome, etc.

Key words: association rules; knowledge discovery; apriori algorithm; medical records of traditional Chinese medicine; medication pattern

中医医案作为重要的医疗资源包含海量的数据与信息,其中蕴含的隐性知识亟待被挖掘与发现^[1]。提升中医医案的挖掘与利用水平在一定程度上可以提高中医临床诊疗的效率与医疗服务的精准性^[2],因此围绕中医医案内容进行知识发现有着重要的意义。习近平总书记在考察东汉医学家张仲景的墓祠纪念地时强调,中医药学包含着中华民族几千年的健康养生理念及实践经验,是中华民族的伟大创造和中国古代科学的瑰宝,要做好守正创新、传承发展工作,注重用现代科学解读中医药学原理,推动传统中医药和现代科学相结合^[3]。关联规则是数据挖掘的一项重要技术,也是一种发现事物之间关联关系的分析方法^[4]。知识发现(Knowledge Discovery in Database, KDD)是“数据挖掘”的一种更广义的说法,即从繁杂的信息中根据不同的需求获得知识,其最终目的是为使用者提炼出有意义、有价值的规律和知识以便创新利用^[5]。基于关联规则挖掘中医知识,对中医医案数据进行知识新发现,是体现中医药学与现代科学相结合的一种方法。现有基于关联规则的方法对中医知识进行发现的研究中,大多数研究以中医古籍、名老中医医案、已出版中医著作等作为数据挖掘源,使用临床医案作为数据源直接进行知识发现的研究不够充分,而临床医案中蕴含着大量的医生诊疗经验,基于此数据源进行挖掘可以提高研究实用性。基于上述,本文将运用关联规则的技术方法,以实地临床医案为例,挖掘规律、提炼出特定的知识并进行可视化,从而增强知识发现结果的实用性与可利用性。本研究拟解决两个基本问题:①如何运用关联规则方法基于临床医案发现治疗以咳嗽为主证的用药规律、舌诊

症状特征、症—药规律?②以实地肺病医案为例,临床诊疗的用药规律、舌诊症状特征以及症—药规律是什么?

1 文献回顾

1.1 中医药知识发现研究

在中医药领域的知识发现研究中,学者们采用的理论方法丰富多样,利用计算机技术与中医药学相结合分别运用不同方法,从不同角度对中医知识进行了新发现。刘超男等^[6]利用复杂概念网络生成方法对《伤寒论》方药数据进行处理,建立相关知识库,以方剂、药物和方—药关系为基础生成与优化形式背景,实现相关关系结构的可视化,并在此基础上做整体的知识发现,验证了一种新的知识发现方法的可行性。刘鸿儒等^[7]从知识发现的角度分析石学敏院士针灸治疗脑梗死偏瘫过程中症状与治疗原则之间的关联,利用属性偏序图方法对“现症—治疗原则”的形式背景进行知识结构发现,并获得结果关系描述,发现以“醒脑开窍”的主要治疗原则。张楠^[8]收集李赛美教授治疗糖尿病患者以失眠为主诉症状的有效门诊病历,运用基于属性偏序结构理论的知识发现新方法,对诊治病例的证型分布和症状特征进行提取,挖掘李赛美教授治疗的用药规律,总结出糖尿病患者失眠症的主要症状及治疗关键,最终发现其医案中蕴藏的隐性知识。徐笋晶等^[9]提出基于形式概念分析的数学属性偏序表示原理挖掘糖尿病临床经验的方法,通过建立形式背景、生成属性偏序结构图成功进行知识发现。赵翔凤等^[10]检索《中国药典》《中华医典》等数据源,收集与海马相关的方剂,以频数分析、关联规则挖掘方法等分析用药规律以及配伍,挖掘并发现

了历代医家中海马的临床应用情况、配伍规律与潜在应用价值。综上所述,对中医药数据规律挖掘的本质是通过数据统计与分析得出理法方药之间的一系列关系。医学数据中包含着多维关系,涉及药—药、药—症、症—证等联系,而关联规则方法可以在整体中直接挖掘事物与其他事物的关联与相存依性,可以清晰地挖掘非线性数据以及中医药知识,是研究用药以及诊疗规律知识的较好选择,与本研究目的相匹配。

1.2 基于关联规则的中医药知识发现研究

数据挖掘是知识发现过程中的一个关键步骤,通常以技术方法实现数据模式的发现。关联规则是数据挖掘的一项技术,关联规则的挖掘过程主要包括两个阶段:第一阶段是从原始资料集合中找出高频项目组(频繁项集);第二阶段是从高频项目组中产生关联规则^[11]。近几年在国内的应用研究已有许多成果,现有基于关联规则的挖掘技术在中医药知识发现方面的研究主要涉及以下几个方面:对名老中医医案隐性知识进行发掘^[12],分析中医用药规律特点或常用配伍,挖掘症状间的关系发现证候规律,从某病症出发探索一系列方证关系等。

1.2.1 基于关联规则的用药规律研究

目前,基于关联规则挖掘用药规律的研究往往从某一具体疾病出发,通过关联规则等方法分析疾病所用药物,运用中医知识分析药物性味、归经等,从而得出治疗疾病的药物组配规律以及治疗大法。高雅婷等^[13]以多层次数据挖掘研究中医临床治疗咳嗽变异性哮喘的用药特点及组方规律,分析得到二项关联规则 15 项,药物组合 14 项,总结出以祛风解表、祛痰平喘、补虚、祛湿等为主的治疗大法;寒温结合、宣降相依的用药规律。周世琴等^[14]收集名老中医治疗干眼症经验方进行用药规律挖掘,发现药类频次较高的为补虚药、清热药,体现了以补益肝肾、益气养阴等为主的治法。刘丽萍等^[15]运用关联规则技术分析王邦才教授治疗郁证的用药组方规律,发现其形神同调、宣畅气机的核心主张以及寒温并用、气血同调等治疗方法。赵达等^[16]通过古今医案云平台探讨名中医治疗冠心病的用药规律,发现以心气阴亏虚的主要病机以及活血化瘀、燥湿化痰,补益心气、滋阴养血为主的

治疗方法,为临床经验提供了可靠的参考。相丽玲等^[17]以中医配方书籍为来源筛选治疗肺结核的方剂 96 首进行关联规则、频次等分析,挖掘出出现频率最高的 3 味中药以及 3 组药对,为中医治疗肺结核用药及配伍提供依据。

1.2.2 基于关联规则的证候规律相关研究

中医证候的概念包含了望闻问切所诊断出的所有反映状态,表现为所有可被观察到的症状。因此在基于关联规则挖掘中医证候要素的研究中所涉及的内容较为复杂,研究往往基于某疾病前提下,在证型分类的基础上分别构成症状与症状的关联。例如钟震等^[18]通过关联规则分析阵发性房颤中医证候要素,得到 6 个最常见的中医证型以及各个证型所对应的常见主要症状,从而发现阵发性房颤中医证候间的内在规律,对于临床应用以及尽早识别该疾病风险贡献出一定价值。王昆阳^[19]采用关联规则等数据挖掘方法探索焦虑抑郁共病的中医证候要素间的关联规律,提取出 14 项证候要素,并最终归纳成五大中医证型及其相对应的主症与次症。

1.2.3 基于关联规则的中医多维关联研究

中医知识中不仅存在着药—药、症—症等简单的一维关联关系,还存在方—症—证—药等不同性质事物间的关联关系。例如,从某方剂入手挖掘相关症、证、药间的规律,形成完整的多维关联链条。刘梦柳^[20]运用关联规则、聚类分析等方法,从方剂、中药、症状等方面分析梳理《伤寒杂病论》中包含麻黄、石膏经典药对的有关内容,探究出与麻黄—石膏相关的证—方—药间的关系规律,实现了中医经典方剂的数据化。方坤炎等^[21]以关联规则技术分析探究小柴胡汤在治疗热入血室中的应用,从中医医案入手分析相关方证关系,发现“经血夹块—经色紫黯”“红花—桃仁”“生半夏—牡丹皮”“生姜—恶心”等的强关联规则,揭示了症状与药物的关联。

通过对上述文献的梳理可知,数据挖掘的技术方法在发现有价值的知识上发挥着重要作用。目前有关中医知识发现的研究大多集中于对古籍或名老中医的医案分析,以某地医案进行实证研究的数量相对来说较少。本文将基于关联规则的方法挖掘分析实地肺病医案,对中医用药、患者症状特征、症—

药关联等规律进行研究分析,发现医案中潜在、有价值的知识。

2 研究设计

本研究意在基于数据挖掘技术对中医医案内容进行分析挖掘与知识发现,以实地肺病医案为数据源,发现用药知识、患者临床症状特征、症—药二维关联规律与治疗特色,以期提高中医临床诊疗效率、支撑决策,探索关联规则分析在中医药知识发现方面的创新应用。

2.1 研究思路

首先,收集与整理实地肺病医案,将筛选过符合研究要求的医案数据导入到 Excel 表格中,为接下来的数据分析步骤作准备。其次,用 R 语言对所需数据进行分析,分别统计出临床治疗实践中运用频率靠前的高频药物、舌诊特征后再利用 Apriori 算法对数据进行关联规则分析,得到相应的频繁项集与强关联规则。同时将高频舌诊特征作为固定项进行症—药关联规则的挖掘,并加入可视化分析来展示数据挖掘的结果。最后,在上述频数分析与关联规则分析的基础上对所得数据信息进行知识发现,挖掘提炼出隐性的中医知识。

2.2 研究方法

本文在收集、清洗数据后,利用 R 语言软件作为辅助工具依次对医案数据进行统计分析、关联规则分析与可视化分析,并在上述分析结果的基础上对中医药知识进行归纳整理,从而形成有价值的新知识。对医案数据进行统计分析可以达到直观迅速地了解中医数据集中较为重要项目的目的;关联规则是发现事物之间关联关系的一种方法,一般表现的规则形式为 $X \rightarrow Y$ ^[22]。其典型的例子就是超市购物篮分析,通过分析顾客购买不同商品的行为习惯从而得出商品之间的关系。而医学数据与购物篮数据存在许多共同点,根据关联规则所呈现的支持度、置信度等数据观察到事物间所对应的关系与规律,因此关联规则可以合理地在医学知识发现中被应用^[23],利用关联规则的方法对中医医案进行分析,可以发现药物间、症状间、药物与症状间的隐含关系。同时,运用 R 语言中 ggplot2 包的功能对关联规则进行可视化呈现,以实现对知识结果更好的展示。

2.3 数据来源与处理

首先,分别整理从长春中医药大学附属医院吉

林省中医院与长春中医药大学附属第三医院门诊处所获的肺病相关医案^[24],选择数据中所有患者主证为“咳嗽”的医案。本文筛选医案过程中所纳入的医案需满足以下要求:医案需明确诊断患者主证为“咳嗽”;药方明确并以口服为治疗途径。排除标准如下:无主要症状或主要症状不符合的医案,医案中明确药物作用不明或为无效的医案;药物组方不完整、不清晰的医案;不同病患同一药物处方的医案;患者症状恶化成为其他病症的医案^[25]。医案筛选完毕后根据所需分析项目依次录入 Excel 中,药物名称统一按照 2020 年版《中华人民共和国药典》与《中华本草》进行规范校正^[26],例如,“乌贼骨”“墨鱼骨”统一为“海螵蛸”、“胡麻仁”统一为“芝麻”、“玉蝴蝶”统一为“马兜铃”;舌诊名称按照舌色、舌体、苔色、苔质进行拆分并规范,例如,舌色分为“舌红”“舌淡红”“舌白”“舌暗”“舌暗红”等,苔色分为“苔白”“苔黄”“苔淡黄”等,经初步筛选共获得医案 663 例。整理所有医案数据并录入 Excel 中以建立肺病数据库。数据预处理步骤如下:首先在 Excel 中手动进行清洗,删除所有药方不完整的医案以及同病治疗的相同处方,将无效医案过滤掉。将上述医案一并整理、保存为 csv 文件后导入 R 语言待进一步处理。在 R 语言中使用 read() 函数读取数据,输入 format = c(“basket”), cols = 1, header = TRUE 命令,定义数据类型同时指定表格中第一列(编号)为事物标识,并用 rm.duplicates = TRUE 命令删除重复值,将数据转化成事务性数据后完成数据预处理。

3 数据分析

3.1 中医治疗咳嗽的高频用药分析

以治疗咳嗽的医案为例,收集并整理数据。首先针对治疗药物进行频次分析:第一步在 Excel 中将整理好的医案数据保存为.csv 格式以便数据软件间兼容利用,在 R 语言中使用 read.csv() 函数读取此 csv 文件。由于整理好的医案数据为购物篮型数据,因此读取数据时编辑 cols = 1,指定以文件中的第一列为事务标识,并使用 rm.duplicates = TRUE 命令删除重复值。再利用 itemFrequency() 函数对药物频率进行统计,并使用 decreasing = TRUE 命令对统计结果进行降序排列,最后再利用 itemFrequency

cyPlot()包将数据结果呈现为条形统计图。如图 1 所示,为治疗咳嗽的使用频率排名前十高频用药统计图,横坐标为中药名称,纵坐标为药物使用频率。统计分析结果显示,本研究所用基础中药总共 428 种,炙甘草、杏仁、半夏、川贝母、茯苓、陈皮等为高频被使用的药物,其中炙甘草使用频率最高,达到 30%以上,表明这味药物经常在治疗咳嗽当中被使用,可能起着重要的作用或重要辅助调和作用,极具治疗价值。另外,杏仁、半夏、川贝母、茯苓、陈皮的使用频率也极高,这几味药材在临床医案中经常出现,概率均达到 20%以上,单独或组配对治疗咳嗽起着重要作用。

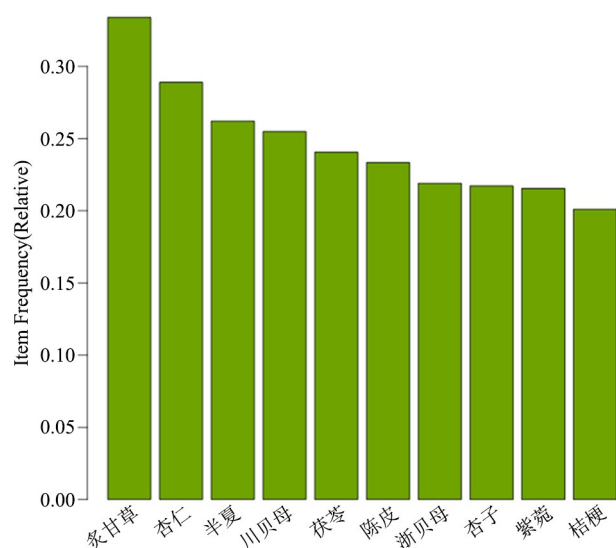


图 1 中医治疗咳嗽主证的高频中药统计图

Fig. 1 High Frequency Drug Statistical Chart for Treating Cough as Main Symptoms in Traditional Chinese Medicine

3.2 基于关联规则的中医咳嗽用药规律分析

关联规则涉及几个基本概念。 $I = \{i_1, i_2, \dots, i_n\}$ 是 n 个不同项目的集合, $i_k | k = 1, 2, \dots, n$ 称为一个项目, 项目的集合简称为项集。支持度是两个事物在总的项集中出现的概率, 其大小能反映出规则是否普遍, 最小支持度表示规则需满足的最低重要性; 置信度是出现前项事物时后项同时出现的概率, 是一种条件概率; 最小置信度表示规则需满足的最低可靠性。关联规则的发现过程分为两个步骤^[27], 首先, 根据所设定最小支持度找出所有满足此条件的频繁项集, 包含所有频繁项集的非空子集; 其次, 在上一步基础上设定最小置信度, 统计出所有满足条件的规则, 即为强规则, 强规则即代表强关联。最小支持度与最小置信度的值需要数据挖掘人员结

合实际需要设定。

首先, 对规则支持度的设定意义在于过滤掉没有意义的规则, 数据挖掘的意义在于通过挖掘结果做出决策而产生价值, 若规则的支持度过低则会失去意义。由于本研究所采用的数据量庞大, 经筛选共得到有效医案 556 条, 药物种类达 428 种, 药物总使用频次达 5 611 次, 为了不遗漏掉一些可能存在价值的数据, 则需要折中设定支持度。在 R 语言关联规则挖掘中, 以超市购物篮数据为示例, 官方设定的最小支持度与最小置信度为 $\text{Support} = 0.06$, $\text{Confidence} = 0.5$, 因为医学数据与购物篮数据有许多共同特点, 所以尝试设置与之相同的阈值以观察关联规则的结果是否理想。在 R 语言中执行 Apriori 算法对药物数据进行分析, 在执行算法之前首先设定最小支持度为 6%, 以生成频繁项集 (Frequent Itemsets), 表 1 为部分频繁项集。

表 1 基于 Apriori 算法的药物频繁项集挖掘

Tab. 1 Mining Frequent Itemsets of Drugs Based on the Apriori Algorithm

项 集	支持度	频次
{炙甘草, 紫菀}	0.14003591	78
{炙甘草, 陈皮}	0.14003591	78
{半夏, 陈皮}	0.12926391	72
{杏仁, 浙贝母}	0.11669659	65
{桔梗, 炙甘草}	0.11490126	64
{半夏, 炙甘草}	0.11490126	64
{杏仁, 桔梗}	0.10592460	59
{南沙参, 川贝母}	0.10412926	58
{前胡, 炙甘草}	0.10412926	58
{款冬花, 紫菀}	0.10053860	56
{半夏, 茯苓}	0.09874327	55
{款冬花, 炙甘草}	0.09694794	54
{紫菀, 陈皮}	0.09694794	54
{前胡, 半夏}	0.09515260	53
{杏仁, 紫菀}	0.09515260	53
{炙甘草, 黄芩}	0.09335727	52
{前胡, 桔梗}	0.09335727	52
{杏仁, 炙甘草}	0.09156194	51

其次, 在生成频繁项集的基础上挖掘关联规则, 置信度越大则说明规则越可靠。设定药物间最

小支持度为 6%，最小置信度为 50%，进而筛选出关联规则(rules)共 97 条。首先需要通过提升度来判断所挖掘关联规则的有效性，提升度是含有 X 的条件下同时含有 Y 的概率，与 Y 总体发生的概率之比，即可理解为 X 发生对 Y 产生的影响程度，如式 (1)：

$$Lift(X \Rightarrow Y) = P(Y|X) / P(Y) \tag{1}$$

提升度的值可以正确揭示两种事物的相关性，当提升度大于 1 时，值越高，则代表两种药物直接

的关联性越强，此时提升度越高则可以证明关联规则越具有实用价值。若提升度等于 1，则代表 X 与 Y 之间相互独立，此时即便置信度相当高也是一条无效的关联规则。当提升度小于 1 时，绝对值越高则代表两种事物间负相关性越强^[28]。由此，根据提升度大小排除无效的关联规则，在验证 X 与 Y 存在正向关联性的基础上，再利用 sort 函数对有效关联规则按照置信度大小进行排序，得到符合标准的规则如表 2 所示。

表 2 基于 Apriori 算法的部分药物关联规则
Tab. 2 Partial Drug Association Rules Based on Apriori Algorithm

序号	关联规则	支持度	置信度	提升度
1	{丹参, 桔梗}⇒{炙甘草}	0.06104129	1	2.994624
2	{丹参, 炙甘草}⇒{桔梗}	0.06104129	1	4.973214
3	{柴胡, 桔梗}⇒{炙甘草}	0.06283662	1	2.994624
4	{柴胡, 炙甘草}⇒{桔梗}	0.06283662	1	4.973214
5	{前胡, 枳壳}⇒{桔梗}	0.06283662	1	4.973214
6	{炙甘草, 郁金}⇒{桔梗}	0.06104129	0.9714286	4.831122
7	{丹参}⇒{桔梗}	0.06104129	0.9189189	4.569981
8	{丹参}⇒{炙甘草}	0.06104129	0.9189189	2.751816
9	{桔梗, 郁金}⇒{炙甘草}	0.06104129	0.9189189	2.751816
10	{半夏, 枳壳}⇒{桔梗}	0.06283662	0.8974359	4.463141
11	{白前}⇒{炙甘草}	0.07181329	0.8888889	2.661888
12	{枳壳, 炙甘草}⇒{桔梗}	0.06463196	0.8780488	4.366725
13	{前胡, 紫菀}⇒{炙甘草}	0.07540395	0.875	2.620296
14	{半夏, 桔梗}⇒{炙甘草}	0.07181329	0.8510638	2.548616
15	{紫菀, 黄芩}⇒{炙甘草}	0.06822262	0.8444444	2.528793
16	{桔梗, 紫菀}⇒{炙甘草}	0.06463196	0.8372093	2.507127
17	{紫菀, 陈皮}⇒{炙甘草}	0.08078995	0.8333333	2.49552

由表 2 可知，符合设定标准的规则有 {丹参, 桔梗}⇒{炙甘草}、{丹参, 炙甘草}⇒{桔梗}、{柴胡, 桔梗}⇒{炙甘草}、{柴胡, 炙甘草}⇒{桔梗}、{前胡, 枳壳}⇒{桔梗}、{炙甘草, 郁金}⇒{桔梗}、{丹参}⇒{桔梗}、{丹参}⇒{炙甘草}、{桔梗, 郁金}⇒{炙甘草}等，序号 1~5 的关联规则均兼具 100% 的置信度，表明出现前项药对或药组时有 100% 的概率后项药物会同时出现在同一处方当中，这说明所挖掘出的规则药物间具有相当紧密的联系。且 {柴胡, 炙甘草}⇒{桔梗}、{前胡, 枳壳}⇒{桔梗}、{炙甘草, 郁金}⇒{桔梗}、{丹参}⇒{桔

梗}、{半夏, 枳壳}⇒{桔梗}等药组具有相当高的提升度，证明本研究所挖掘出的关联规则中，前项药物、药对的出现大大提升了后项药物出现的程度。

本文采用 R 语言中具有可视化功能的 plot() 函数对上述数据结果进一步精炼，设计了基于关联规则的用药规律可视化方式，分别制作点图、平行坐标图与网状图。由点图可以显示关联规则按照支持度与置信度的大小分布情况，如图 2(a) 所示，点的颜色代表提升度大小，颜色越深则提升度越高；如图 2(b) 所示，平行坐标图每一组由左向右的折线代表一组关联规则，折线两侧分别代表关联规则

炙甘草、陈皮一半夏、南沙参—川贝母、浙贝母—杏子等药物组合的共现次数较高，彼此之间具有较强的关联性。

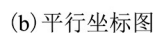


Fig. 2 Scatter Plot and Parallel Coordinate Plot of Drug Association Rules Based on Apriori Algorithm

头折线间的节点大小与深浅则代表两味药物间的支持度与提升度大小。

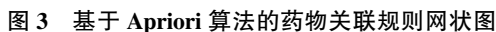


Fig. 3 Network Graph of Drug Association Rules Based on Apriori Algorithm

3.3 咳嗽主证相关舌诊症状分析

在中医理论中脏腑问题通常可以通过舌质、舌苔等特征被反映出，因此基于舌诊信息能够判断疾病与体内的一些相关问题。以治疗咳嗽的医案为例，针对拆分、规范化后的舌诊数据行频次分析：具体步骤如上述 3.1，在 R 语言软件中使用 read() 函数读取 csv 文件，再利用 itemFrequency() 函数对舌诊频率进行统计，图 4 为主证咳嗽的舌诊高频统计图，横坐标为舌诊症状，纵坐标为出现频率。由统计分析结果显示，苔薄白、苔薄腻、苔薄黄，以及舌红、舌暗红、舌淡红、有齿印等为高频出现的舌诊症状。

3.4 基于关联规则的咳嗽舌诊症状规律分析

利用 Apriori 算法对舌诊数据进行分析，首先需根据数据集大小与特点调整最小支持度与置信度的阈值。在调试过程中，首先以 Support = 0.06, Confidence = 0.5 的阈值执行算法，结果只生成一条符合的规则，为了不遗漏有价值的关联关系，不断尝试在调低最小支持度与最小置信度的同时确保二

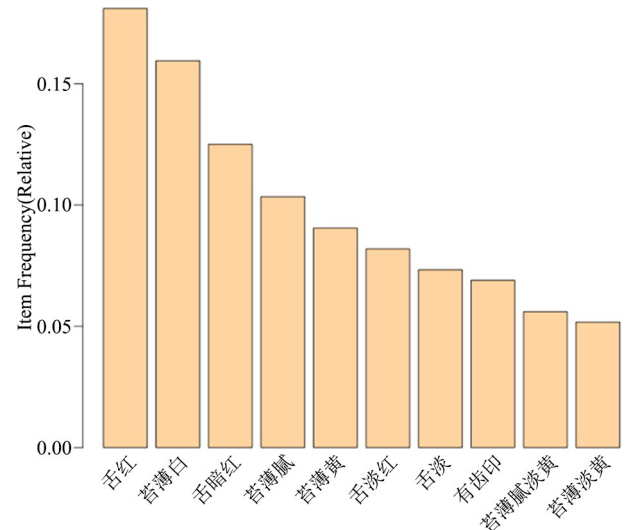


图 4 咳嗽主证的高频舌诊统计图
Fig. 4 High Frequency Tongue Diagnosis Statistical Chart of Cough as Main Symptoms

者的值可以反映出所挖掘规则的一定意义并挖掘出相应数量的结果。最终设定药物间最小支持度为 3%，最小置信度为 10%，进而筛选出关联规则 (rules) 共 16 条，表 3 为所有有效关联规则。

表 3 基于 Apriori 算法的舌诊频繁项集挖掘
Tab. 3 Mining Frequent Itemsets of Tongue Diagnosis Based on Apriori Algorithm

序号	关联规则	支持度	置信度	提升度
1	{ 苔薄淡黄 } ⇒ { 舌暗红 }	0.04741379	0.9166667	7.333333
2	{ 舌暗红 } ⇒ { 苔薄淡黄 }	0.04741379	0.3793103	7.333333
3	{ 苔薄腻淡黄 } ⇒ { 有齿印 }	0.03017241	0.5384615	7.807692
4	{ 有齿印 } ⇒ { 苔薄腻淡黄 }	0.03017241	0.4375	7.807692
5	{ 苔薄腻淡黄 } ⇒ { 舌暗红 }	0.0387931	0.6923077	5.538462
6	{ 舌暗红 } ⇒ { 苔薄腻淡黄 }	0.0387931	0.3103448	5.538462
7	{ 舌淡红 } ⇒ { 苔薄白 }	0.04741379	0.5789474	3.630156
8	{ 苔薄白 } ⇒ { 舌淡红 }	0.04741379	0.2972973	3.630156
9	{ 有齿印 } ⇒ { 舌暗红 }	0.06034483	0.875	7
10	{ 舌暗红 } ⇒ { 有齿印 }	0.06034483	0.4827586	7
11	{ 苔薄白 } ⇒ { 舌红 }	0.05172414	0.3243243	1.791506
12	{ 舌红 } ⇒ { 苔薄白 }	0.05172414	0.2857143	1.791506

由表 3 可知，符合设定标准的舌诊二元组合有：苔薄淡黄—舌暗红、舌淡红—苔薄白、有齿印—舌暗红、苔薄白—舌红、苔薄腻淡黄—有齿印、苔薄腻淡黄—舌暗红；且苔薄淡黄—舌暗红、有齿印—舌暗红、苔薄腻淡黄—有齿印的组合具有相当高的

提升度，证明在临床咳嗽医案中，患者舌质暗红常伴苔薄且淡黄或有齿印的症状一起出现，有齿印的症状常伴有暗红舌色或薄腻淡黄的舌苔。而对于同一组舌诊数据 (如 { 有齿印 } ⇒ { 舌暗红 }、{ 舌暗红 } ⇒ { 有齿印 }) 具有相同的支持度和提升度，置

信度存在差异的原因在于项集顺序的不同，这种前后顺序导致的差异在临床中不具有太大的研究意义。最后再次采用 `plot()` 函数，选择用网状图对上述数据结果进行更好的展示，如图 5 所示，从图中可直观解读出，苔薄淡黄—舌暗红、有齿印—舌暗红组合间的圆点的颜色代表的提升度较高。

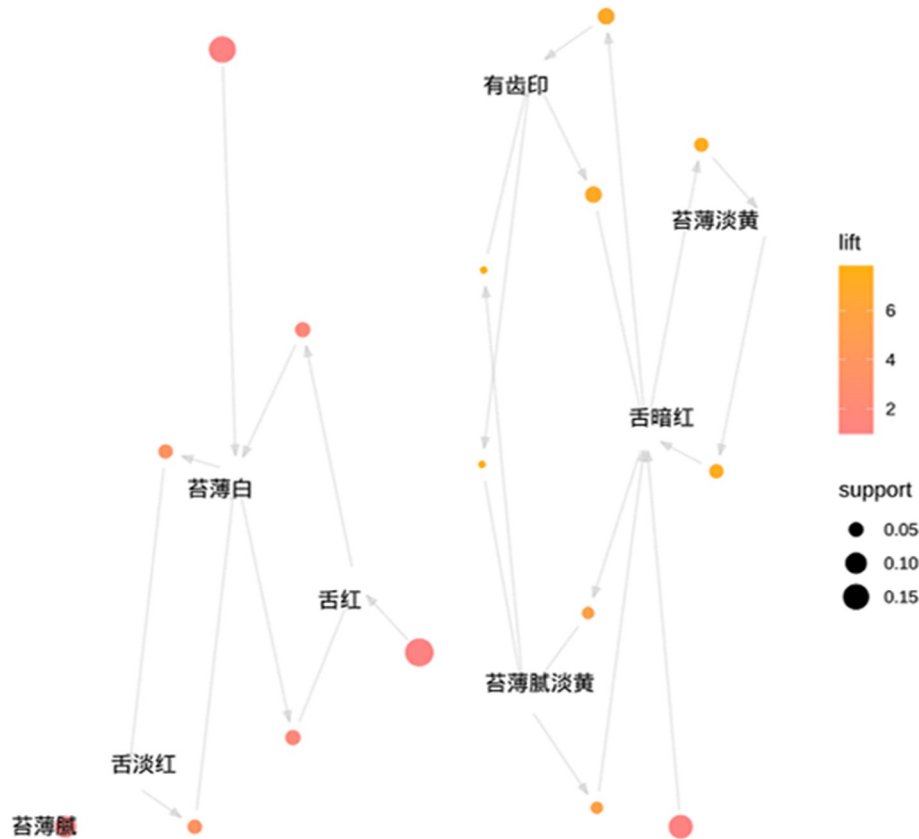


图 5 基于 Apriori 算法的舌诊症状关联规则网状图

Fig. 5 Mesh Graph of Tongue Diagnosis Symptom Association Rules Based on Apriori Algorithm

3.5 基于关联规则的咳嗽相关症—药规律分析

在全部医案中筛选舌诊、中药信息齐全的医案数据，整理保存成所需 csv 格式。首先利用 3.3 中所得高频舌诊数据作为关联规则中的固定项，在 R 语言中将高频舌诊依次放入后项，根据新数据的特点调整最小支持度与置信度的阈值。在调试过程中，既最大限度地使设定的阈值提高，又保证关联规则结果输出的数量，以筛选到有价值的知识。最终设

定条件最小支持度为 5%，最小置信度为 10%，对舌诊和药物间的关联性进行运算，得到舌诊症状—药物关联规则，分别选取舌色、苔色、苔质最高频项的对应关联规则以展示，按置信度排序的关联规则如表 4 所示(序号 1~5 是以“舌红”为后项的关联规则、序号 6~10 是以“苔薄白”为后项的关联规则、序号 11~15 是以“舌暗红”为后项的关联规则)，按支持度排序的关联规则如表 5 所示。

表 4 基于 Apriori 算法的症—药关联规则挖掘(按置信度排序)

Tab. 4 Mining of Symptom Drug Association Rules Based on Apriori Algorithm (Ranked By Confidence)

序号	前 项	后 项	置信度	支持度	提升度
1	{平地木}	{舌红}	0.667	0.052	3.683
2	{陈皮, 青皮}	{舌红}	0.583	0.06	3.222
3	{青皮}	{舌红}	0.56	0.06	3.093
4	{芦根}	{舌红}	0.364	0.052	2.009
5	{黄芩}	{舌红}	0.339	0.086	1.872

表 4 (续)

序号	前 项	后 项	置信度	支持度	提升度
6	{ 防风 }	{ 苔薄白 }	0.542	0.056	3.396
7	{ 舌红 }	{ 苔薄白 }	0.286	0.052	1.792
8	{ 款冬花, 紫菀 }	{ 苔薄白 }	0.286	0.052	1.792
9	{ 射干 }	{ 苔薄白 }	0.267	0.052	1.672
10	{ 炙甘草, 紫菀 }	{ 苔薄白 }	0.264	0.06	1.656
11	{ 有齿印, 麻黄根 }	{ 舌暗红 }	1	0.052	8
12	{ 有齿印, 莱菔子 }	{ 舌暗红 }	1	0.052	8
13	{ 有齿印, 莱菔子, 麻黄根 }	{ 舌暗红 }	1	0.052	8
14	{ 有齿印, 白前, 麻黄根 }	{ 舌暗红 }	1	0.052	8
15	{ 丹参, 有齿印, 麻黄根 }	{ 舌暗红 }	1	0.052	8

表 5 基于 Apriori 算法的症—药关联规则挖掘(按支持度排序)

Tab. 5 Mining of Symptom Drug Association Rules Based on Apriori Algorithm (Ranked By Support)

序号	前 项	后 项	支持度	置信度	提升度
1	{ 炙甘草 }	{ 舌红 }	0.116	0.235	1.297
2	{ 黄芩 }	{ 舌红 }	0.086	0.339	1.872
3	{ 陈皮 }	{ 舌红 }	0.078	0.247	1.362
4	{ 前胡 }	{ 舌红 }	0.069	0.235	1.3
5	{ 紫菀 }	{ 舌红 }	0.065	0.205	1.135
6	{ 炙甘草 }	{ 苔薄白 }	0.108	0.217	1.363
7	{ 紫菀 }	{ 苔薄白 }	0.078	0.247	1.546
8	{ 黄芩 }	{ 苔薄白 }	0.065	0.254	1.594
9	{ 炙甘草, 紫菀 }	{ 苔薄白 }	0.06	0.264	1.656
10	{ 款冬花 }	{ 苔薄白 }	0.06	0.25	1.568
11	{ 炙甘草 }	{ 舌暗红 }	0.125	0.252	2.017
12	{ 桔梗, 炙甘草 }	{ 舌暗红 }	0.121	0.5	4
13	{ 桔梗 }	{ 舌暗红 }	0.121	0.378	3.027
14	{ 丹参 }	{ 舌暗红 }	0.116	0.844	6.75
15	{ 丹参, 桔梗 }	{ 舌暗红 }	0.116	0.844	6.75

通过综合观察关联规则对应的支持度、置信度、提升度大小可以发现,舌红与苔薄白的症状、舌暗红与有齿印具有很强的关联性;“舌红”作为后项时平地木、陈皮—青皮、青皮分别作为前项与其有相当高的提升度,表明这几味药或药物组合对于治疗舌红症状对应证具有相当的作用;“苔薄白”与防风也具有很高的提升度,桔梗、炙甘草、丹参药物单独或组合与“舌暗红”都有很强的关联度,表明这几味药物是治疗舌暗红症状对应证的重要药物;麻黄根、莱菔子分别或组合与“有齿印”“舌

暗红”同时出现时关联规则的置信度达到 100%,说明了“有齿印”“舌暗红”高度地同时出现在同一医案当中,临床中患者有相当大的概率在出现舌质暗红时同时有齿印,麻黄根、莱菔子是治疗其对应证的重要药物。

4 分析结果与讨论

4.1 基于关联规则的中医药知识发现过程

本研究基于实地肺病医案数据,首先,根据上文结果所显示的关联规则筛选出具有关联关系的药对,在对药物各自进行四气五味分析的基础上挖掘

出具有强关联的药对配伍规律，初步分析咳嗽患者的大致类型；其次，根据舌诊关联规则进行归类挖掘咳嗽的病位以及病因；最后，根据症—药关联规则通过临床症状结合药物作用进一步挖掘以咳嗽为主证患者的几类证状，形成症—药—证的逻辑链，同时发现治疗基本原则，最终达到整理与细化医案知识与知识发现的目的。

4.2 基于关联规则的用药规律知识发现

中药的性味可以反映药物的作用与功效，是药物性能的一个重要方面^[29]。基于关联规则从药物四气五味的角度可以发现临床治疗中的常用治法组合，中药的四气是从药物作用于机体所发生的反应概括的，即寒、热、温、凉，此外还有一些药物药性较为平和被称为平性；五味是指酸、苦、甘、辛、咸，有些药物为一气多味。对关联规则中的药物四气五味进行总结发现，在药物四气配伍方面，主要分为 3 个组合方式：第一类为“平+温”组合的配伍，以炙甘草—紫菀、炙甘草—陈皮、杏仁—桔梗、半夏—茯苓药物组合为代表；第二类为“温+温”组合的同气配伍，以半夏—陈皮、款冬花—紫菀、紫菀—陈皮等组合为代表；第三类为“平+寒”组合的配伍。温热药物常用于治疗阴盛寒证，寒凉药多具有解热、镇咳的作用，具体有清热燥湿药、清虚热药的药效多用于治疗阳盛热证，由此咳嗽的患者首先可大致分为两大类，一类为热证，一类为寒证。药物五味方面，第一涉及较多“苦辛”搭配，其中多味药物兼具苦、辛之味。苦能降能泄，具有清泄火热、泄降气逆的功效，辛味先入肺经，可润燥、行气；而苦辛并施可以针对表证、热湿证。第二主要为甘+辛、苦的多味搭配，甘味可调和诸药，先入脾经，中医中脾主运化，脾运化功能失常则会影响其余脏腑之功能。而又由于甘味用于虚证，甘+辛、苦的药物搭配可起到从内调和脏腑的功效，针对里证治疗，由此咳嗽患者可再次被分为表证与里证。

4.3 基于关联规则的舌诊特征知识发现

舌诊是中医望诊中的重要内容之一，通过观察舌头的形态、色泽等特征来辅助诊断。基于关联规则从舌诊特征的角度进行知识发现，可以挖掘临床

治疗中的患者舌诊特征组合。从舌苔颜色角度来看，薄白苔一般为正常舌苔，也可能由外感风邪引起，而薄黄苔大多由白苔转化而成，表示病已由寒化热、由表及里，黄苔则大多由痰热湿交织而成；以舌质来说，淡红色为正常，红舌和暗红舌大多代表里热火旺；从舌形角度看，有齿印大多因舌体胖大，多见于虚证或湿热痰浊壅滞；因此若从病位和病情发展阶段来看，可大致分为两类：第一种以苔薄白—舌淡红为代表，反映出患者可能处在表证初期或存在里证轻病；第二种包括苔薄淡黄—舌暗红、苔薄腻淡黄—有齿印、苔薄腻淡黄—舌暗红、有齿印—舌暗红组合，以苔色变化伴随舌质的变化反映出患者病位已入里，脏腑功能已失调。若从病位结合病因来看，也可分为两类：苔薄白—舌淡红多属外感风寒症，患者应为外感咳嗽，病位在表病情较为轻微；苔薄白—舌红症状所对应患者属外感风热或外感风寒化热。齿痕舌多因脾脏运化水液功能失调而湿阻导致，所以伴有齿印—舌暗红、苔薄腻淡黄—有齿印的患者病为里证虚证；苔薄淡黄—舌暗红、苔薄腻淡黄—舌暗红组关联规则的患者体内湿热积滞病情较为复杂，除肺腑以外肠胃等脏腑可能存在功能障碍，因此患者属于内伤咳嗽，舌形舌质、舌苔较正常舌苔相比有较明显的变化。综上角度看舌诊反映的信息，导致患者咳嗽的主要病因大致有风、寒、热、湿几种外邪，病情的病变部位包括体表与内在脏腑，即所描述“外感咳嗽”与“内伤咳嗽”之分，而具体的证型仍需进一步结合中药进行分析。

4.4 基于关联规则的咳嗽症—药规律知识发现

基于关联规则从舌诊症状与药物相结合的角度进行知识发现，可以更加精确的挖掘患者咳嗽的主因与治疗思路。从咳嗽病位以及主因角度进行知识分类分为以下几组，第一组为苔薄白—舌淡红组合，所涉及的强关联药物有防风、紫菀、款冬花、射干、黄芩等，药物主要类型为清热药、止咳平喘药与解表药，结合中药功效可见苔薄白—舌淡红组合所对应的患者为外寒内热证或风热证；第二组为苔薄腻淡黄—有齿印组合，所涉及的强关联药物有炙甘草、丹参、枳壳、白前、郁金等，药物主要类型为补益

药、行气药、活血化淤药，药物归经涉及心、肝、肺、脾、胃、大肠经，可见苔薄腻淡黄—有齿印组合，所对应患者病情较为复杂，机体内部功能严重失调，结合药物功效与症状看，可能为肺气虚证、痰热壅肺证、肝火犯肺证等；第三组为有齿印—舌暗红组合，所涉及的强关联药物有丹参、枳壳、白前、莱菔子等，药物主要类型有行气药、止咳平喘药，药物归经涉及脾、胃、肺经，结合舌诊信息来看，患者大多因气机郁滞导致痰湿咳嗽，属痰湿证；第四组为苔薄淡黄—舌暗红组合，所涉及的强关联药物有桔梗、麻黄根、莱菔子、丹参、白前；苔薄腻淡黄—舌暗红组合所涉及的强关联药物有麻黄根、莱菔子、丹参、白前，药物主要作用为理气与止咳化痰，表明这两组舌诊症状对应的患者大多气机不通畅且体内湿热属痰热证或阴虚证。综上所述，对舌诊症状与药物的关联规则进行知识发现，一方面可以通过咳嗽患者的舌诊症状与所对应药物作用大致判断其病位深浅与复杂程度；另一方面可以通过研究临床医案发现咳嗽时不同状况下所对应的治疗方法与原则，如病情较复杂时，用药不仅需针对肺腑情况，还需兼顾其余脏腑以达到恢复五脏六腑平和的目的。

5 结 语

对中医药规律进行知识发现，可将中医的治疗经验、用药思想、症状规律等进行清晰化的表述，有利于思想的传承与临床治疗的经验借鉴。本文系统地介绍了基于关联规则对中医药知识进行挖掘的过程，以实地肺病医案为例，从“咳嗽”为主证的医案入手，首先通过统计分析、关联规则算法找出常用药物、舌诊特征以及强关联的药物、舌诊特征、症—药关系组合，最后一步一步对实地治疗咳嗽医案的用药规律、舌诊特征、症—药关联规律进行隐性知识发现，为中医药知识发现提供新角度。本研究选取了实地肺病医案进行研究，后可将此关联规则分析方法运用到其他中医药知识发现的研究中。本研究结果显示，本地治疗咳嗽的药物以炙甘草为首，杏仁、半夏、川贝母、茯苓、陈皮等被高频使用，基于 Apriori 算法的关联规则分析显示，常见的药物配伍为丹参，桔梗—炙甘草、柴胡，桔

梗—炙甘草、前胡，枳壳—桔梗、丹参—桔梗等；本地咳嗽舌诊症状以苔薄白、苔薄黄、舌红、舌暗红、有齿印等最为常见，基于 Apriori 算法的关联规则分析显示，舌诊症状以苔薄淡黄—舌暗红、有齿印—舌暗红、苔薄腻淡黄—有齿印等组合为强关联组合，苔薄淡黄—舌暗红舌诊症状组合的强关联药物有桔梗、炙甘草、丹参、白前等。研究发现，咳嗽患者病变部位包括体表与体内脏腑，导致咳嗽的外邪主要有风、寒、热、湿，证型涉及风热证、痰湿证、阴虚证、肺气虚证、痰热壅肺证等；医治里证时以肺与其他脏腑同时调理以平衡人体五脏六腑最终达到治疗效果。

参 考 文 献

- [1] 范铁兵, 宁秋萍, 杨志旭, 等. 基于数据挖掘技术的中医学隐性知识显性化分析 [J]. 中国实验方剂学杂志, 2017, 23 (10): 221-226.
- [2] 吴信东, 叶明全, 胡东辉, 等. 普适医疗信息管理与服务的关键技术与挑战 [J]. 计算机学报, 2012, 35 (5): 827-845.
- [3] 习近平总书记在南阳市考察调研引发热烈反响 [EB/OL]. <https://baijiahao.baidu.com/s?id=1699979280046580483&wfr=spider&for=pc>, 2023-12-04.
- [4] 胡侃, 夏绍玮. 基于大型数据库的数据挖掘: 研究综述 [J]. 软件学报, 1998, (1): 54-64.
- [5] 唐晓萍. 数据挖掘与知识发现综述 [J]. 电脑开发与应用, 2002, (4): 31-32, 35.
- [6] 刘超男, 徐笋晶, 李赛美, 等. 基于多层次复杂概念网络表示方法的《伤寒论》方药按治法分类的知识发现 [J]. 北京中医药大学学报, 2014, 37 (7): 452-457.
- [7] 刘鸿儒, 刘丹, 张涛. 石学敏院士脑梗死偏瘫的针灸治疗规律分析 [J]. 辽宁中医杂志, 2022, 49 (10): 174-177, 224.
- [8] 张楠. 基于属性偏序结构理论的李赛美六经辨治糖尿病患者失眠知识发现 [D]. 广州: 广州中医药大学, 2017.
- [9] 徐笋晶, 李赛美, 洪文学, 等. 基于数学属性偏序表示原理的李赛美教授治疗 II 型糖尿病处方用药分析 [J]. 中国实验方剂学杂志, 2014, 20 (2): 207-211.
- [10] 赵翔凤, 相光鑫, 石丛薇, 等. 基于方剂构成信息的知识发现和关联网络构建挖掘海洋中药海马历代用药经验 [J]. 世界科学技术-中医药现代化, 2021, 23 (2): 335-343.
- [11] 陆丽娜, 陈亚萍, 魏恒义, 等. 挖掘关联规则中 Apriori 算法的研究 [J]. 小型微型计算机系统, 2000, (9): 940-943.
- [12] 吴嘉瑞, 唐仕欢, 郭位先, 等. 基于数据挖掘的名老中医经验传承研究述评 [J]. 中国中药杂志, 2014, 39 (4): 614-617.

- [13] 高雅婷, 王心恒, 杨勤军, 等. 中医药治疗咳嗽变异性哮喘用药规律多层次数据挖掘研究 [J]. 世界中医药, 2022, 17 (9): 1337-1342, 1349.
- [14] 周世琴, 谢鹏飞, 李青青, 等. 基于中医传承辅助系统的中医治疗干眼症用药规律研究 [J]. 中国民间疗法, 2022, 30 (9): 76-81.
- [15] 刘丽萍, 包科颖, 伍竹君, 等. 基于数据挖掘探讨王邦才治疗郁证用药规律 [J]. 新中医, 2022, 54 (8): 10-15.
- [16] 赵达, 曹晖, 胡志希. 基于关联规则和聚类分析研究全国名中医治疗冠心病用药规律 [J]. 湖南中医药大学学报, 2021, 41 (11): 1726-1730.
- [17] 相丽玲, 王凰. 基于关联规则的肺结核中医治疗方法研究 [J]. 中国民族民间医药, 2022, 31 (4): 20-24.
- [18] 钟霞, 焦华琛, 杨梦琦, 等. 基于数据挖掘技术探讨心房颤动的中医证候分布规律 [J]. 辽宁中医药大学学报, 2022, 24 (6): 164-169.
- [19] 王昆阳. 基于关联规则的焦虑抑郁共病中医证候规律研究 [D]. 北京: 北京中医药大学, 2017.
- [20] 刘梦柳. 基于 Apriori 算法的《伤寒杂病论》“麻黄—石膏”的证—方—药关联分析 [D]. 太原: 山西医科大学, 2021.
- [21] 方坤炎, 吕丽容, 郭超峰. 基于数据挖掘的小柴胡汤加减治疗热入血室方证关系研究 [J]. 中国中医药信息杂志, 2021, 28 (4): 47-51.
- [22] 蔡伟杰, 张晓辉, 朱建秋, 等. 关联规则挖掘综述 [J]. 计算机工程, 2001, (5): 31-33, 49.
- [23] 杨柳, 徐晴文, 黄汝佳, 等. 关联规则在中医药数据挖掘中的应用 [J]. 中医药信息, 2022, 39 (12): 35-40.
- [24] 高鹤萌. 面向智慧医疗的诊疗术语集构建研究 [D]. 长春: 吉林大学, 2022.
- [25] 赵艳青, 项敏泓, 王雪菲, 等. 基于因子分析与关联规则挖掘干眼症的用药规律 [J]. 中华中医药杂志, 2017, 32 (10): 4708-4712.
- [26] 国家药典委员会. 中华人民共和国药典 [M]. 北京: 中国医药科技出版社, 2020.
- [27] 刘红岩, 陈剑, 陈国青. 数据挖掘中的数据分类算法综述 [J]. 清华大学学报 (自然科学版), 2002, (6): 727-730.
- [28] 关联分析中的支持度、置信度和提升度 [EB/OL]. <https://blog.csdn.net/sanqima/article/details/42746419>, 2023-12-04.
- [29] 卢训丛. 中药性味归经理论的研究思路与方法 [J]. 中国中医基础医学杂志, 2006, (11): 803-804.

(责任编辑: 杨丰侨)

(上接第 33 页)

- [42] 刘佩, 林如鹏. 网络问答社区“知乎”的知分享与传播行为研究 [J]. 图书情报知识, 2015, (6): 109-119.
- [43] 陈明红, 漆贤军. 社会资本视角下的学术虚拟社区知识共享研究 [J]. 情报理论与实践, 2014, 37 (9): 101-105.
- [44] 吴琼, 邓胜利. 虚拟社区中知识创新影响因素的实证研究 [J]. 图书情报知识, 2011, (5): 115-121.
- [45] Balog A. Testing a Multidimensional and Hierarchical Quality Assessment Model for Digital Libraries [J]. Stud Inform Control, 2011, 20 (3): 233-246.
- [46] 吴雅威, 张向先, 陶兴, 等. 基于用户感知的学术问答社区答案质量评价指标构建 [J]. 情报科学, 2020, 38 (10): 141-147.
- [47] 姚伟, 康世伟, 柯平, 等. 虚拟网络社区中知识共创的多元衍生研究 [J]. 情报理论与实践, 2021, 44 (7): 58-64.
- [48] 刘宝瑞, 张双双. 虚拟学习社区知识构件的交流机理研究 [J]. 情报科学, 2012, 30 (11): 1603-1608.
- [49] Armstrong A G, Hagel J. The Real Value of Online Communities [J]. Knowledge and Communities, 2000, 74 (3): 85-95.
- [50] 李林晓. 用户创新社区评价指标体系研究 [D]. 大连: 东北财经大学, 2017.
- [51] 张馨遥, 沈涵, 张健, 等. 基于模糊综合评价的在线健康信息服务用户满意度分析 [J]. 情报科学, 2018, 36 (6): 73-78, 99.
- [52] 邓雪, 李家铭, 浩健, 等. 层次分析法权重计算方法分析及其应用研究 [J]. 数学的实践与认识, 2012, 42 (7): 93-100.
- [53] 赵宇翔, 刘周颖, 刘炜, 等. 创意类开放数据竞赛作品评价指标体系构建与测定——以数字人文项目为例 [J]. 中国图书馆学报, 2020, 46 (2): 75-95.
- [54] 小木虫. 关于小木虫 [EB/OL]. <http://muchong.com/aboutus.php>, 2023-12-18.
- [55] 甘春梅, 王伟军. 在线科研社区中知识交流与共享: MOA 视角 [J]. 图书情报工作, 2014, 58 (2): 53-58.
- [56] 刘冰, 张耀辉. 基于网络用户体验与感知的信息质量影响因素模型实证研究 [J]. 情报学报, 2013, 32 (6): 663-672.
- [57] 徐美凤, 叶继元. 学术虚拟社区知识共享主体特征分析 [J]. 图书情报工作, 2010, 54 (22): 111-114, 48.
- [58] 徐美凤, 叶继元. 学术虚拟社区知识共享研究综述 [J]. 图书情报工作, 2011, 55 (13): 67-71, 125.
- [59] Davis F D. User Acceptance of Information Technology: System Characteristics, User Perceptions and Behavioral Impacts [J]. International Journal of Man-Machine Studies, 1993, 38 (3): 475-487.
- [60] 程豪杰, 张宁, 袁勤俭. 学术虚拟社区知识交流的研究进展 [J]. 现代情报, 2018, 38 (5): 158-163.
- [61] 刘蕊. 虚拟社区知识共享影响因素及激励机制探析 [J]. 情报理论与实践, 2012, 35 (8): 39-43.

(责任编辑: 杨丰侨)