

数据挖掘算法在中药方剂研究中的应用现状[△]

李蕙质*,周小玲,杨玉杰,章新友*(江西中医药大学计算机学院,南昌 330004)

中图分类号 R289;R2-03 文献标志码 A 文章编号 1001-0408(2024)01-0112-07

DOI 10.6039/j.issn.1001-0408.2024.01.20



摘要 近年来,数据挖掘算法在中药领域的科研中得到了广泛应用。采用数据挖掘算法可处理和分析中药方剂中的多层次数据,并对其作用机制提供合理解释。这一方法现已较好地应用于中医药的配伍规律和高频药组的挖掘中,提高了临床诊断、靶点筛选和新药研究的可靠性和准确性。本文对147篇中药方剂研究中运用数据挖掘算法的文献进行了整理与分析,结果表明,数据挖掘算法在中药方剂作用机制研究、中药方剂量效研究、挖掘核心药对/药组、挖掘“方-药-证”间的关系、发现新方剂和挖掘配伍规律这6个子领域中发挥了独特优势,尤以关联规则和聚类分析算法最具有代表性。

关键词 数据挖掘算法;中药方剂;文献计量法;应用

Application of data mining algorithms in research on traditional Chinese medicine formula

LI Huizhi, ZHOU Xiaoling, YANG Yujie, ZHANG Xinyou (School of Computer Science, Jiangxi University of Chinese Medicine, Nanchang 330004, China)

ABSTRACT In recent years, data mining algorithms have been widely employed in scientific research within the field of traditional Chinese medicine (TCM). The data mining algorithms are used to effectively handle and analyze the complex data in TCM formulas, providing a rational explanation for the mechanism of action. This method has proven particularly useful in uncovering patterns of compatibility and frequent combinations of herbs in TCM, thereby enhancing the reliability and accuracy of clinical diagnosis, target screening, and the study of new drugs. This paper reviews and analyzes 147 papers on TCM formula research that utilize data mining algorithms. The results indicate that data mining algorithms play a unique advantage in six sub-areas, including the study on the mechanism of action in TCM formula, the dose-efficacy of TCM formulas, the identification of core drugs pairs/groups, mining the relationships among “formulas-drug-symptom”, the discovery of new formulas, and mining the compatibility law. Notably, association rules and clustering algorithms are the most representative.

KEYWORDS data mining algorithms; traditional Chinese medicine formula; bibliometrics analysis; application

中药方剂是中药学中的一个重要研究领域,其基于中医整体观的思想,通过对疾病的辨证施治来构建适用于不同疾病或病情的中药处方,以实现治疗和预防疾病的目的。中药方剂的形成不仅遵循“君臣佐使”的组方原则,还取决于医生个人的诊断和实践经验,因此方剂的有效性很大程度上受到医生经验的影响,这使得方剂研究较为困难。此外,由于病因、病机、中药药味及其组合方式的复杂性,使得单一分析难以准确地获得中药方剂与疾病之间的关系。目前的中药方剂研究主要从传

统经验角度和现代科学角度两个方面展开,其中传统经验角度注重探究草药的性味归经、功效特点等方面,现代科学角度则注重从分子水平理解方剂的药效成分和作用机制^[1]。

近年来,数据挖掘算法在中药领域的科研中得到了广泛应用。采用数据挖掘算法可处理和分析中药方剂中的多层次数据,并对其作用机制提供合理解释^[2]。这一方法现已较好地应用于中医药的配伍规律和高频药组的挖掘中^[3],提高了临床诊断、靶点筛选和新药研究的可靠性和准确性,为临床精准用药提供了科学参考。本文中,笔者较为全面地论述了数据挖掘算法在中药方剂中的应用现状,旨在为该方法在中药方剂研究中的进一步应用提供参考。

1 文献检索与整理

本研究首先以“方剂”“数据挖掘”“Apriori”“层次聚类”“决策树”“支持向量机”等为关键词,组合查询中国

△基金项目 国家自然科学基金项目(No.82360992);江西省中医药管理局科技计划重点项目(No.2022Z007);江西省中医药管理局癌病方证信息数据挖掘重点实验室项目(No.科研字[2022]16号);江西中医药大学校级科技创新团队立项(No.CXTD22015)

* 第一作者 硕士研究生。研究方向:中医药信息挖掘与应用。电话:0791-87118863。E-mail:lihuizhi@jxutcm.edu.cn

通信作者 教授,博士生导师,博士。研究方向:中医药信息挖掘与应用。电话:0791-87118863。E-mail:xinyouzhang@jxutcm.edu.cn

知网、万方数据、维普网和PubMed数据库中2015年1月—2023年6月收录的相关文献,共得到341篇文献。为了保证文献的质量,本研究对检索到的341篇文献进行了筛选整理,选择核心期刊文献以及硕士、博士学位论文,去除无效文献和重复文献后,共纳入有效文献147篇。然后,对纳入文献按照中药方剂研究子领域进行分类,将子领域内文献总数不少于10篇的纳入分析,如子领域里文献数量较少则不做单独讨论,合并归类在“其他应用”项里。对得到的有效文献进行归纳总结后,将纳入文献按照子领域分类并建立了6个文献库,方法流程如图1所示。最后,分析数据挖掘算法在中药方剂研究子领域中的应用现状,并探讨分类任务、聚类任务和关联规则分析中的代表性算法。

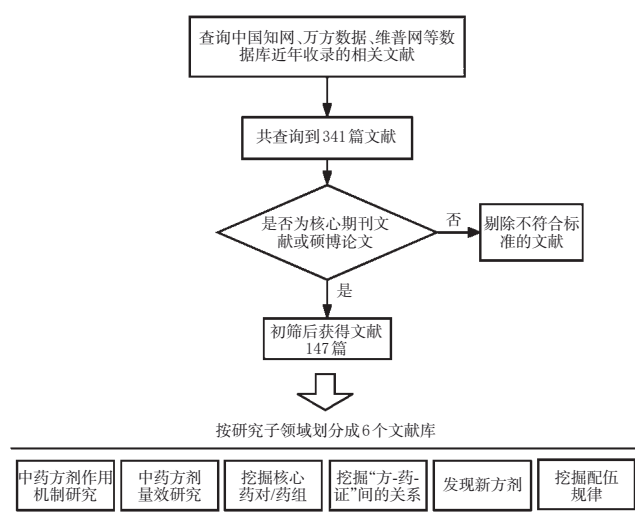


图1 文献收集与归类分析流程

2 数据挖掘算法在中药方剂研究子领域的应用与分析

由检索到的147篇有效文献可知,数据挖掘算法是中药方剂研究的有力辅助工具。不同类型的数据挖掘算法在中药方剂研究子领域的应用如表1所示。

表1 数据挖掘算法在中药方剂研究子领域的应用情况

| 研究子领域 | 文献数量/篇 | 数据挖掘算法 |
|---------------|--------|---------------------------------|
| 中药方剂作用机制研究 | 43 | 复杂网络分析、人工神经网络等 |
| 中药方剂量效研究 | 98 | Apriori、K均值聚类、聚类分析、支持向量机、朴素贝叶斯等 |
| 挖掘核心药对/药组 | 112 | 关联规则、熵聚类、层次聚类等 |
| 挖掘“方-药-证”间的关系 | 14 | Apriori、熵聚类、层次聚类等 |
| 发现新方剂 | 61 | 关联规则、聚类分析、支持向量机、决策树等 |
| 挖掘配伍规律 | 42 | 关联规则、聚类分析、决策树、回归分析等 |
| 其他应用 | 8 | 隐含狄利克雷分布主题模型、马尔可夫链、禁忌搜索算法等 |

2.1 中药方剂作用机制研究

中药具有多靶点、多环节、多途径、整合调节的特点,因而详细阐明其作用机制较为困难。目前多成分干预治疗对于复杂疾病更有优势,如肿瘤的药物治疗已从单一用药向联合用药方向转变。陈恒巍^[4]建立了一种协

同抗肿瘤多药组合深度学习预测模型,在阐明中药协同抗肿瘤的作用机制方面提供了新的思路和方法。

随着医学与生命科学研究进入大数据时代,网络药理学将网络科学和系统生物学方法应用于药物研究,可解释中医药作用机制。研究者利用筛选的药物成分和疾病靶点信息构建药物-靶点-通路生物网络模型,通过找出网络关键节点可揭示目标药物对疾病的作用机制和生物学基础。如复方一枝蒿在治疗流感方面具有确切的效果,但其物质基础及作用机制尚不明确,许律捷等^[5]基于贝叶斯模型的组合机器学习方法探索复方一枝蒿的作用机制,得到其通过多成分、多靶点和多通路发挥抗流感作用。

2.2 中药方剂量效研究

中药药量是决定方剂药效的关键因素,挖掘药物在治疗不同疾病方剂中的常用剂型、剂量和用法,有助于发现潜在组方规律,提高中药组方的准确性,确保用药的有效性和安全性。研究者收集含有目标药物的名家医案或经典古方,整理方剂对应的证型、症状和药物的药量、性味、归经等数据,通过关联规则和聚类分析等数据挖掘算法,分析目标药物在不同功效下的常用剂量,可为中药在临床的精准使用提供科学参考。王妍等^[6]采用关联规则的算法开展酸枣仁用量与主治病证关系的研究,得出了5种常见病证下酸枣仁的常用剂量。骆宾妃等^[7]采用数据挖掘方法分析了《中华医典》中含竹沥的方剂和主治病证,指出针对不同疾病的证型特点不仅应调整药物配伍,还应调整竹沥的剂型和用量。侯王君等^[8]通过中医传承辅助平台分析,得出了冠心病心绞痛方剂中使用频次最高的5味中药——川芎、丹参、黄芪、当归和红花的最佳用量。王焯燃等^[9]采用Apriori算法对治疗胃火炽盛型消渴病方剂中的药物用量、性味、归经和频次进行了分析,得到了方剂中天花粉、知母、黄连、麦冬的药量,印证了“清泻胃火为主,兼顾滋补肺、肾等脏腑之阴”的治疗原则。

2.3 挖掘核心药对/药组

药对是方剂最小的组成单位。作为一种相对固定的中药配伍形式,药对像桥梁一样将单味中药和方剂联系在一起,其不仅具备方剂的基本主治功能,往往还体现了方剂组成的核心。研究人员常采用Apriori算法挖掘高频药对/药组并分析潜在配伍规律,通过调整置信度和支持度过滤关联性弱的药对/药组,绘制出现频次高的核心药对/药组的网络图,从而直观地展示药对/药组间关联性的强弱。孙世光等^[10]采用复杂系统熵聚类算法整理防治失眠的方剂,得到了分别由2味、3味和4味药组成的核心药对/药组。

一方面,挖掘核心药对/药组能够表达方剂的处方思路或首选处方结构,反映方剂的整体功效特点;另一方面,关键药对/药组对方剂的组成和疗效具有重要影响。孙志新等^[11]通过挖掘治疗心悸的方剂,对比方剂中所有药物和常用药对中的高频药物,将11味重合药物命名为核心用药,发现其多分布在心经和脾经,印证了治疗心悸的方剂具有养心、健脾的功效。鞠康等^[12]挖掘《中国方剂数据库》包含“疫”病方剂的药物频次、性味、归经,并通过聚类分析算法对高频药物聚类,得到5类关键药物组合,该结果与方剂的主成分分析结果一致。

2.4 挖掘“方-药-证”间的关系

“方-药-证”的理念体现了中医疗法的整体观念和个体化治疗原则,医生运用辨证施治的方法,根据患者特定的证候特征进行针对性的治疗,在此基础上选择适当的方剂,再确定最终药物组合。研究者可以采用数据挖掘算法,对“方-药-证”之间的关联性进行研究,从更客观的角度阐释中药方剂科学配伍的合理性,为中医药的临床诊治提供参考。

在解释目标方剂或药物与中医证型之间的关系方面,胡慧明等^[13]采用关联规则及聚类分析算法发现,针对不同病证,山楂皆常配伍理气药、消食药、补虚药进行治疗;陈建国等^[14]采用改进的遗传算法提取方证决策表的决策规则,阐释了冠心病方剂与证型间的关联及用药原则;崔一然等^[15]构建了方剂中除丹参-红花以外的药物关系网络图,揭示了该药对常与活血、行气之品联用,解释了该药对与中风、血瘀痛证之间的关系。相关研究还关注于药物与主治疾病之间的联系,如赵萱等^[16]采用频次统计和关联规则的方法,分别阐明了人参-附子在治疗5种不同疾病时的配伍区别,为解释该药对的核心功效提供了重要参考和科学依据;赵艳青等^[17]通过中医传承平台挖掘抑郁症方剂的证型和高频药味,发现常见证型多含“肝郁”,高频药味也多具有疏肝解郁、疏肝理气等功效,揭示了抑郁症与情志失调之间的关联性。

2.5 发现新方剂

发现新的中药方剂具有双重意义:其一,依据现有方剂的高频药物组合生成新处方,丰富了经典古方的内涵;其二,研究评价新处方的方法,判断新处方能否达到更好的疗效。关联规则算法和聚类分析算法在发现新方剂方面的使用频次较高,前者用于在大量药物中挖掘关联度较高的药物组合,然后采用后者将核心或高频药物聚类得到新处方。

王嘉欣等^[18]依据聚类和排序的思想提出 TOPSIS 算法,得到的最优新处方与中医传承辅助平台生成的新处方一致,为生成新处方提供了新的思路和方法,使新处

方的评价更具有客观性和科学性。朱月^[19]通过挖掘治疗泄泻病寒湿内盛证的用药规律,得到支持度最高的药对是甘草-白术,且理气药和补虚药2类药物使用较多,在此基础上自拟了中药处方并通过临床试验证明了其疗效优于蒙脱石散。程立前^[20]收集整理了治疗先兆流产的药物及出现频率、归经和常用药组,经聚类分析算法获得了7个能够改善患者临床症状的新组方。鲁可等^[21]采用无监督的熵聚类分析算法对两类核心组方进一步聚类,得到了6个治疗咳血的新药物组合,并结合中药理论知识分析了新方的合理性。沈莹等^[22]对2015年版《中国药典》中含甘草的止咳方剂进行了挖掘分析,结合中医理论在高频药物的基础上归纳出1组新方剂,并从现代药理学角度阐释了新方剂止咳祛痰的机制,为研发止咳祛痰类药物提供了科学依据。

2.6 挖掘配伍规律

中药配伍规律遵循“七情和合”“四气五味”“君臣佐使”等原则^[23]。同一方剂中的不同药物通过配伍,可起到增效、减毒等作用。研究中药方剂的配伍机制不仅对阐释中医理论内涵具有重要意义,还对中药方剂的发展具有推动作用。

通过分析目标药物的配伍药物可为研究其配伍规律提供参考,如赵文龙等^[24]通过挖掘《中医方剂大辞典》中含有羌活治疗疼痛的方剂,发现羌活配伍的高频药物有防风、甘草、川芎。潘树茂等^[25]采用数据挖掘方法发现治疗原发性肝癌的处方中健脾益气药的使用频次最高,该类药物与白术-莪术配伍可健气补脾,加以柴胡-鳖甲可疏肝解郁。切尼项毛等^[26]将 Apriori 算法与“味性化味”矢量模型结合,解释了石榴籽适合与温性药材配伍的原因。郝庆勋等^[27]采用数据挖掘算法发现不同结缔组织原发病的治疗药物具有不同的配伍特点,如治疗系统性硬化症选用鸡血藤和大血藤可促进血液循环,而治疗原发性干燥综合征相关间质性肺疾病应该使用当归和甘草这2种具有滋阴健脾、清热养血作用的药物。郑丽等^[28]采用聚类分析算法研究治疗肾性贫血的方剂,发现黄芪和当归始终分在同一簇并且处于中心位置,说明黄芪和当归为治疗肾性贫血方剂的关键药物,并且常与白术、茯苓和党参配合使用。

2.7 其他应用

方剂相似度可用于衡量方剂之间的相似性,而成分相似的方剂在功效上具有高相关性。李新龙等^[29]基于方剂相似度构建核心方药及其有效人群特征的复杂网络,采用子网挖掘算法和聚类分析算法对目标人群进行模块划分,发现对应方剂可按功效划分为2个模块。朱志鹏等^[30]基于隐含狄利克雷分布主题模型将方剂和其

组成成分间的关系转换成“方剂-证型”和“证型-组成成分”2个概率模型,表明该方法能够很好地计算方剂间的相似度。郭文龙等^[31]采用词频-逆文档频率算法建立了方剂数据挖掘系统,可用于推荐与目标方剂功能和主治相似度高的方剂。

对于联合用药产生的不良反应,王淳等^[32]采用关联规则算法分析了联合用药对心血管疾病的影响,提出应谨慎使用丹参注射液、参附注射液等。蒋先仲等^[33]采用数据挖掘算法并引入“丰度”的概念用于衡量研究中纳入的药品占比,提出“目标调节+整体调节+肠道微生态环境调节”的个体化组方策略,分别对早、中、晚期3个阶段的胃癌患者给出了组方设计建议。

3 算法应用效果比较

3.1 分类任务中常用算法的比较

分类算法是预测性任务的一种,主要目标是将数据划分为不同的类别。中药方剂领域研究中常采用的分类算法有决策树算法和支持向量机。决策树算法因其易于理解和解释的特点而被用于研究药物与主治病证之间的关系,以及构建配伍决策模型,如构建药物和主治病证关系决策树、方剂配伍决策模型等^[34]。支持向量机常用于建立预测模型、分析方剂配伍的可能性等,如金滋力等^[35]把方剂看作从药物出发的组合决策问题,采用支持向量机达到了90%的剂剂分类准确率。

3.2 聚类任务中常用算法的比较

聚类分析算法将未被标记的数据分为具有相似特征的簇,从而发现数据集中的内在结构和规律,在方剂研究中可实现筛选核心药物的目的,以K均值聚类算法、层次聚类算法、熵聚类算法为代表。K均值聚类算法适用于大规模方剂数据分析,其关键点在于选择合适的聚类个数K。赵凌霄等^[36]在3次聚类中分别比较了不同K值下的药物聚类结果,得到了不同K值下的药物关系图以及核心组合。层次聚类算法适合小规模数据集,如数据量有限的自身免疫性肝炎的研究^[37],可依据药物性味特征进行聚类,用以挖掘药物的核心组合或生成新处方。熵聚类算法在处理噪声和异常值方面表现较佳,如肖红等^[38]采用复杂系统熵聚类算法分析药物关联规则,并对聚类结果进行再次提取,通过演化得到了新的处方。

3.3 关联规则分析中常用算法的比较

关联规则分析揭示的是数据之间的潜在关系^[2],通常包括2个步骤:先查找所有频繁项集,然后基于高频项生成频繁关联规则。如果其满足最小置信度并基于第一步找到的频繁项集,则该规则被视为关联规则。Apriori算法和FP-growth算法都是关联规则分析的经典

算法,前者基于先验知识,后者基于树结构。上述2种算法均有助于识别多种属性之间的关系并生成关联规则,例如挖掘高频药物的配方规律等,从而有助于分析用药原则、验证配伍规律^[39]。如李德琳等^[40]通过FP-growth算法对靶点-通路网络进行了数据挖掘分析,得到34组靶点-通路频繁项集,验证了方剂具有多成分和多靶点的特征。

4 总结与讨论

本文对147篇中药方剂研究中运用数据挖掘算法的文献进行了整理与分析,结果表明,数据挖掘算法已在中药方剂的研究中广泛应用,主要应用的子领域为:中药方剂作用机制研究、中药方剂量效研究、挖掘核心药对/药组、发现新方剂和挖掘配伍规律。常用的数据挖掘算法包括:关联规则、聚类分析、分类分析、复杂网络分析等。

虽然数据挖掘算法具有较广泛的应用范围,但其对中药方剂的研究还有待进一步改进。首先,数据集规模是制约数据挖掘算法在中药方剂研究中应用的重要因素。现有的中药方剂领域数据集规模较小,这对于数据挖掘算法的应用和效果都带来了一定程度的限制。因此,需要采集更多的规范化数据,包括临床试验数据和基础实验数据,以扩大数据集规模,从而提高数据挖掘算法的精度和效果。其次,中药方剂研究中使用的算法比较单一,缺乏针对性的算法选择。尽管部分数据挖掘算法有规范的流程,例如关联规则和聚类分析,但其对于中药方剂研究存在一定的局限性和适用范围,需要根据具体问题和数据特征的选择进行调整。第三,目前数据挖掘算法在中药方剂研究中的应用主要集中在方剂用药规律的分析中,而对于疗效预测、“毒-效”评价、药物禁忌、方剂中药物的炮制等方面的研究相对较少,需要更加关注并大力开展相关领域的研究工作。

随着中医病历数字化的不断推进,必将产生大量复杂、多维度的数据。相关研究人员可专注于构建结构合理的中医临床数据库,以应对研究中提出的问题。为了实现传统中药方剂多层次、全方位的数据挖掘,除了进一步拓宽方剂研究新领域以外,研究者还应灵活地组合使用不同算法以提高结果的准确性,避免不必要的数据处理内耗。中药方剂研究若能更进一步结合数据挖掘等现代科学技术,必将进一步提高临床疗效,促进中医的传承和发展。

参考文献

- [1] 刘志华,孙晓波.网络药理学:中医药现代化的新机遇[J].药学学报,2012,47(6):696-703.
LIU Z H, SUN X B. Network pharmacology: new opportunity for the modernization of traditional Chinese medi-

- cine[J]. *Acta Pharm Sin*, 2012, 47(6): 696-703.
- [2] HAN J W, KAMBER M. 数据挖掘概念与技术[M]. 范明, 孟小峰, 译. 北京: 机械工业出版社, 2012: 4-23.
HAN J W, KAMBER M. Concepts and technologies of data mining [M]. Translated by FAN M and MENG X F. Beijing: Machinery Industry Press, 2012: 4-23.
- [3] 吴地尧, 章新友, 甘宇汾, 等. 数据挖掘算法在中药研究中的应用[J]. *中国药房*, 2018, 29(19): 2717-2722.
WU D Y, ZHANG X Y, GAN Y F, et al. Application of data mining algorithm in the research of traditional Chinese medicine[J]. *China Pharm*, 2018, 29(19): 2717-2722.
- [4] 陈恒巍. 基于深度学习的协同抗肿瘤多药组合预测模型研究[D]. 镇江: 江苏大学, 2020.
CHEN H W. The study of deep learning based multi-drug synergy prediction model[D]. Zhenjiang: Jiangsu University, 2020.
- [5] 许律捷, 姜雯, 庞晓丛, 等. 复方一枝蒿抗流感有效成分的网络药理学研究[J]. *药学报*, 2017, 52(5): 745-752.
XU L J, JIANG W, PANG X C, et al. Network pharmacology study of the effective constituents in the Compound Yizhihao against influenza disease[J]. *Acta Pharm Sin*, 2017, 52(5): 745-752.
- [6] 王妍, 宋亚南, 陈丽名, 等. 基于数据挖掘的酸枣仁现代临床量效关系初探[J]. *中华中医药杂志*, 2016, 31(9): 3729-3732.
WANG Y, SONG Y N, CHEN L M, et al. Discussion on modern clinical dose-effect relationship of Ziziphi Spinosae Semen based on data mining[J]. *China J Tradit Chin Med Pharm*, 2016, 31(9): 3729-3732.
- [7] 骆宾妃, 董佳威, 刘红宁, 等. 基于“病-药-量”探讨《中华医典》含竹沥中药方剂相应规律的数据挖掘[J]. *中草药*, 2023, 54(8): 2536-2545.
LUO B F, DONG J W, LIU H N, et al. Data mining on corresponding patterns of traditional Chinese medicine formulas containing Bambusae Succus in *Chinese Medical Classics* based on “disease-medicine-quantity” [J]. *Chin Tradit Herb Drugs*, 2023, 54(8): 2536-2545.
- [8] 侯王君, 庄贺, 薛一涛, 等. 应用中医传承辅助平台分析冠心病不稳定型心绞痛方剂的组方用药规律[J]. *中国药房*, 2017, 28(32): 4478-4482.
HOU W J, ZHUANG H, XUE Y T, et al. Analysis of the regularity of drug use in coronary heart disease unstable angina prescriptions by TCM inheritance auxiliary platform[J]. *China Pharm*, 2017, 28(32): 4478-4482.
- [9] 王烨燃, 张洋, 秦琦冰, 等. 基于数据挖掘的胃火炽盛型消渴病方剂的用药规律分析[J]. *中国中药杂志*, 2020, 45(1): 196-201.
WANG Y R, ZHANG Y, QIN Q B, et al. Analysis of medication regularity of traditional Chinese medicine prescriptions for gastropyloric excessiveness diabetes based on data mining[J]. *China J Chin Mater Med*, 2020, 45(1): 196-201.
- [10] 孙世光, 井静, 孙蓉. 基于中医传承辅助系统的失眠防治方剂组方配伍规律研究[J]. *中国实验方剂学杂志*, 2015, 21(9): 208-211.
SUN S G, JING J, SUN R. Analysis on composition principles of prescriptions for insomnia disease by traditional Chinese medicine inheritance support system[J]. *Chin J Exp Tradit Med Formulae*, 2015, 21(9): 208-211.
- [11] 孙志新, 张盼盼, 高武霖, 等. 基于中医传承辅助平台的现代中医药治疗心悸的用药规律分析[J]. *中国中药杂志*, 2017, 42(2): 385-389.
SUN Z X, ZHANG P P, GAO W L, et al. Analysis on medication rules of modern traditional Chinese medicines in treating palpitations based on traditional Chinese medicine inheritance support system[J]. *China J Chin Mater Med*, 2017, 42(2): 385-389.
- [12] 鞠康, 张盛, 胡云飞, 等. 基于数据挖掘“疫”病方剂用药规律对新型冠状病毒肺炎防治的用药启发[J]. *中药药理与临床*, 2021, 37(1): 9-16.
JU K, ZHANG S, HU Y F, et al. Medication rules study of prescription for yi (pestilence) disease on prevention and treatment of COVID-19 based on data mining[J]. *Pharmacol Clin Chin Mater Med*, 2021, 37(1): 9-16.
- [13] 胡慧明, 翁家俊, 朱彦陈, 等. 基于数据挖掘的《中医方剂大辞典》含山楂组方用药规律研究[J]. *中国现代应用药学*, 2021, 38(21): 2713-2720.
HU H M, WENG J J, ZHU Y C, et al. Analysis on composition principles of formulae containing Crataegi Fructus in *Dictionary of Traditional Chinese Medicine Prescriptions* based on data mining[J]. *Chin J Mod Appl Pharm*, 2021, 38(21): 2713-2720.
- [14] 陈建国, 李四海, 赵磊. 基于粗糙集和遗传算法的中医方证相关性研究[J]. *计算机应用与软件*, 2018, 35(7): 211-215.
CHEN J G, LI S H, ZHAO L. Study on the relativity of prescription and syndrome in TCM based on rough set and genetic algorithm[J]. *Comput Appl Softw*, 2018, 35(7): 211-215.
- [15] 崔一然, 刘欣, 申丹, 等. 《中医方剂大辞典》含丹参、红花药对组方规律数据挖掘分析[J]. *中国中药杂志*, 2016, 41(3): 528-531.
CUI Y R, LIU X, SHEN D, et al. Data mining analysis of regularity of formulas containing Salviae Miltiorrhizae Ra-

- dix et Rhizoma-Carthami Flos medicin pair in *Dictionary of Chinese Medicine Prescription*[J]. *China J Chin Mater Med*, 2016, 41(3):528-531.
- [16] 赵萱,陈云慧,郑明月,等. 基于数据挖掘的含人参-附子药对方剂的组方规律分析[J]. *中草药*, 2021, 52(4):1083-1091.
ZHAO X, CHEN Y H, ZHENG M Y, et al. Analysis on composition rules of Ginseng Radix et Rhizoma-Aconiti Lateralis Radix Praeparata pairs based on data mining method[J]. *Chin Tradit Herb Drugs*, 2021, 52(4):1083-1091.
- [17] 赵艳青,滕晶,杨洪军. 基于数据挖掘的现代中医药治疗抑郁症用药规律分析[J]. *中国中药杂志*, 2015, 40(10):2042-2046.
ZHAO Y Q, TENG J, YANG H J. Analysis on medication regularity of modern traditional Chinese medicines in treating melancholia based on data mining technology[J]. *China J Chin Mater Med*, 2015, 40(10):2042-2046.
- [18] 王嘉欣,姚鉴玲,马嘉慕,等. 基于AHP-SOM聚类-TOPSIS和中医传承辅助平台研究中医药治疗围绝经期抑郁症组方规律[J]. *中草药*, 2022, 53(22):7153-7163.
WANG J X, YAO J L, MA J M, et al. Analysis on prescription rules of traditional Chinese medicine in treatment of perimenopausal depression based on AHP-SOM-TOPSIS algorithm and traditional Chinese medicine inheritance platform system[J]. *Chin Tradit Herb Drugs*, 2022, 53(22):7153-7163.
- [19] 朱月. 基于数据挖掘的治疗泄泻病寒湿内盛证用药规律的探讨及临床研究[D]. 济南:山东中医药大学, 2022.
ZHU Y. Analysis of medication regularity and clinical study of the treatment of diarrhea diseases with excessive cold and dampness type based on data mining[D]. Jinan: Shandong University of Traditional Chinese Medicine, 2022.
- [20] 程立前. 基于数据挖掘现代中医药治疗早期先兆流产用药规律及新方应用临床研究[D]. 济南:山东中医药大学, 2021.
CHENG L Q. Based on data mining traditional Chinese medicine treatment of early threatened-abortion of the herbal medicine rule and clinical study of the new prescription[D]. Jinan: Shandong University of Traditional Chinese Medicine, 2021.
- [21] 鲁可,陈腾宇,黄慈辉,等. 基于中医传承辅助平台软件的治疗咳血方剂的组方规律分析[J]. *中国药房*, 2018, 29(6):769-773.
LU K, CHEN T Y, HUANG C H, et al. Analysis of composition regularity of prescriptions for hemoptysis based on TCM inheritance support platform software[J]. *China Pharm*, 2018, 29(6):769-773.
- [22] 沈莹,樊建,孟祥才. 基于2015年版《中国药典》(一部)分析含甘草止咳方剂的用药规律[J]. *中国药房*, 2020, 31(3):281-286.
SHEN Y, FAN J, MENG X C. Analysis of the medication rule of *Glycyrrhiza uralensis*-containing antitussive prescriptions based on 2015 edition of *Chinese Pharmacopoeia* (part I)[J]. *China Pharm*, 2020, 31(3):281-286.
- [23] 字磊,李艳娟,李艳芹,等. 中药复方配伍机制研究方法/策略进展[J]. *中国药房*, 2023, 34(11):1393-1398.
ZI L, LI Y J, LI Y Q, et al. Progress in research methods or strategies of compatibility mechanism of traditional Chinese medicine prescriptions[J]. *China Pharm*, 2023, 34(11):1393-1398.
- [24] 赵文龙,其乐木格,张晶晶,等. 基于3种疼痛病的羌活组方用药规律多维度分析[J]. *中草药*, 2023, 54(2):608-619.
ZHAO W L, Qilemuge, ZHANG J J, et al. Multidimensional analysis on medication rule of prescriptions containing *Notopterygii Rhizoma et Radix* based on three kinds of pain diseases[J]. *Chin Tradit Herb Drugs*, 2023, 54(2):608-619.
- [25] 潘树茂,章新友,吴地尧,等. 基于中药数据策略模式分析平台分析治疗原发性肝癌中药处方的用药规律[J]. *中国药房*, 2020, 31(24):2966-2973.
PAN S M, ZHANG X Y, WU D Y, et al. Analysis of medication rules of TCM prescription against primary hepatic carcinoma based on TCM data analysis platform based on strategy pattern[J]. *China Pharm*, 2020, 31(24):2966-2973.
- [26] 切尼项毛,才让吉,文成当智,等. 基于Apriori算法和“味性化味”矢量模型的含石榴籽藏药方剂的用药规律及药性研究[J]. *中国药房*, 2020, 31(8):906-912.
Qienixiangmao, CAI R J, WEN C, et al. Study on medication regulation and pharmaceutical property of Tibetan medicine prescription of *Punica granatum* seeds based on apriori algorithm and “Ro Nus Zhu rJes” vector model[J]. *China Pharm*, 2020, 31(8):906-912.
- [27] 郝庆勋,焦扬,曹芳,等. 基于数据挖掘分析周平安治疗结缔组织病相关间质性肺病的中药配伍规律[J]. *中医杂志*, 2021, 62(15):1320-1325.
HAO Q X, JIAO Y, CAO F, et al. The compatibility rule of Chinese herbal medicine used by professor ZHOU Ping'an in treating connective tissue disease interstitial lung diseases: based on data mining[J]. *J Tradit Chin Med*,

2021,62(15):1320-1325.

- [28] 郑丽,刘明,孙雪林. 基于数据挖掘的中药治疗肾性贫血的用药规律研究[J]. 中国药房, 2023, 34(5): 591-594,619.
ZHENG L, LIU M, SUN X L. Analysis of medication rules of traditional Chinese medicine in treatment of renal anemia based on data mining[J]. China Pharm, 2023, 34(5):591-594,619.
- [29] 李新龙,刘岩,周莉,等. 基于方剂相似度的核心方药及其适应证挖掘方法研究:以失眠症为例[J]. 中医杂志, 2021, 62(2):118-124.
LI X L, LIU Y, ZHOU L, et al. Prescription similarity-based analysis of core formulas and medicinals and related indications through data mining: taking insomnia as an example[J]. J Tradit Chin Med, 2021, 62(2):118-124.
- [30] 朱志鹏,杜建强,刘英锋,等. 基于LDA主题模型的中药方剂相似度计算[J]. 计算机应用研究, 2017, 34(6): 1668-1670,1676.
ZHU Z P, DU J Q, LIU Y F, et al. Similarity calculation of traditional Chinese medicine prescriptions based on LDA topic model[J]. Appl Res Comput, 2017, 34(6): 1668-1670,1676.
- [31] 郭文龙,罗熊,姜惠娟,等. 基于TF-IDF算法的方剂构成相似度可视化研究[J]. 中国中医药信息杂志, 2019, 26(7):104-108.
GUO W L, LUO X, JIANG H J, et al. Visualization study on approximate degree analysis of prescription properties based on TF-IDF algorithm[J]. Chin J Inf Tradit Chin Med, 2019, 26(7):104-108.
- [32] 王淳,刘振丽,宋志前,等. 基于关联规则探索临床治疗心血管疾病方剂与其不良反应关系的研究[J]. 中国中医基础医学杂志, 2018, 24(10):1478-1483.
WANG C, LIU Z L, SONG Z Q, et al. Exploration of the relation between Chinese medicines for cardiovascular disease and clinical adverse effect based on association rules[J]. Chin J Basic Med Tradit Chin Med, 2018, 24(10):1478-1483.
- [33] 蒋先仲,郭红莲. 基于数据挖掘的抗胃癌植物类中药组方设计[J]. 中国药房, 2018, 29(24):3407-3411.
JIANG X Z, GUO H L. Study on the design of plant-based TCM formula for gastric cancer treatment based on data mining[J]. China Pharm, 2018, 29(24):3407-3411.
- [34] 图雅,张春生,白翠兰. 基于决策树的蒙医热证方剂组方规律数据挖掘[J]. 中国药理学杂志, 2022, 57(10):817-822.
Tuya, ZHANG C S, BAI C L. Data mining of prescription rules of Mongolian medicine heat syndrome based on decision tree[J]. Chin Pharm J, 2022, 57(10):817-822.
- [35] 金滋力,胡建星,金宏威,等. 基于支持向量机与层次分析法的中药方剂配伍分析[J]. 中国中药杂志, 2018, 43(13):2817-2823.
JIN Z L, HU J X, JIN H W, et al. Analysis of traditional Chinese medicine prescriptions based on support vector machine and analytic hierarchy process[J]. China J Chin Mater Med, 2018, 43(13):2817-2823.
- [36] 赵凌霄,吕红,呼兴华,等. 基于数据挖掘探讨黄连在糖尿病及其并发症复方中的应用规律[J]. 中国实验方剂学杂志, 2022, 28(14):158-164.
ZHAO L X, LYU H, HU X H, et al. Application law of Coptidis Rhizoma to diabetes and its complications: based on data mining[J]. Chin J Exp Tradit Med Formulae, 2022, 28(14):158-164.
- [37] 黎胜,施梅姐,萧焕明,等. 自身免疫性肝炎中医医案诊治规律数据挖掘[J]. 中医杂志, 2017, 58(14):1237-1240.
LI S, SHI M J, XIAO H M, et al. Data mining for diagnosis and treatment law of autoimmune hepatitis based on Chinese medical records[J]. J Tradit Chin Med, 2017, 58(14):1237-1240.
- [38] 肖红,姜泉,唐晓颇,等. 姜泉治疗类风湿关节炎组方用药规律研究[J]. 中国中药杂志, 2019, 44(2):381-387.
XIAO H, JIANG Q, TANG X P, et al. Regularity of prescription and medication of Jiang Quan in treating rheumatoid arthritis[J]. China J Chin Mater Med, 2019, 44(2): 381-387.
- [39] 何浩强,高嘉良,陈光,等. 治疗气虚血瘀证中成药的组方规律分析[J]. 中国药房, 2017, 28(35):4910-4914.
HE H Q, GAO J L, CHEN G, et al. Analysis of the regularity of Chinese patent medicine prescriptions for the Qi deficiency and blood stasis syndrome[J]. China Pharm, 2017, 28(35):4910-4914.
- [40] 李德琳,魏本征,张诏,等. 基于FP-growth算法的中医抗病毒方剂配伍规律探索[J]. 中华中医药学刊, 2018, 36(3):663-668.
LI D L, WEI B Z, ZHANG Z, et al. FP-growth algorithm-based exploratory study on compatibility law of Chinese traditional anti-virus medicine prescription[J]. Chin Arch Tradit Chin Med, 2018, 36(3):663-668.

(收稿日期:2023-06-13 修回日期:2023-12-07)

(编辑:胡晓霖)