

DOI: 10.13288/j.11-2166/r.2020.01.008

基于药物剂量及作用的方剂核心用药配伍规律复杂网络算法的构建

马宁¹, 郝秀霞^{2*}, 白金牛¹, 邢俊凤¹

1. 内蒙古科技大学包头医学院, 内蒙古自治区包头市建设路 31 号, 014040; 2. 内蒙古科技大学包头医学院第一附属医院

摘要 利用复杂网络相关算法, 结合方剂中药物“君、臣、佐、使”的药物作用及剂量, 构建一种基于药物剂量及作用的方剂核心用药配伍规律的新算法。通过构建方剂子网, 根据方剂中药物作用及剂量确定核心药物及其权重; 基于药物核心度权重构建中药方剂复杂网络模型; 改进复杂网络标签传递算法对相关方剂的核心用药配伍规律进行挖掘。以《中医方剂大辞典》中治疗产后中风方剂及用药为例, 收集相关方剂 143 首, 频次 > 20 的药物 16 味, 支持度 > 10、置信度 > 80 的关联规则组合 88 个, 构建包含 74 个节点、1245 条边的产后中风治疗用药复杂网络模型, 利用复杂网络算法发现药物聚类组合结果 51 个。通过分析研究发现, 中医治疗产后中风用药以祛风除湿、散寒止痛、补益气血为主, 同时对改进的标签传递算法与传统的关联规则等方法进行比较, 认为改进复杂网络标签传递算法可得出一些围绕核心药物的新颖且更有针对性的药物组合结果。

关键词 复杂网络; 方剂配伍规律; 药物剂量; SLPA 算法; 标签传递算法; 产后中风

中医学经过多年的发展, 在证候诊断、临床诊疗、中医用药等方面积累了宝贵的海量数据资源。挖掘中药复方用药的潜在规律已在现代中医药学中广泛开展。相比于关联规则、聚类等传统数据挖掘算法偏向于发掘用药的表层统计规律, 复杂网络算法通过网络角度描述系统结构, 重在发掘系统内在机制及内涵, 更有助于揭示中医理论体系的复杂性和内隐性^[1], 同时也更易于延伸和扩展。

中药方剂复杂网络方法是在构建中药方剂复杂网络模型基础上采用基于图论的相关算法进行药物聚类的过程, 与因子分析、复杂系统熵聚类等中药配伍规律常用聚类方法相比, 利用复杂网络的重叠社团发现算法, 即寻找聚类社团, 可使同一药物出现在多个聚类配伍结果中, 从而提供不同角度的数据结果。研究人员基于相关算法针对中药配伍规律做了大量的研究, 如李茹等^[2]建立药物属性及药物之间的复杂网络模型, 基于治疗肺痿中药复方使用派系过滤算法进行分析和验证; 韩楠等^[3]定义了药物的重要性并进行组网, 根据局部密度定义社团的适应度来进行社团的划分; 孙道平等^[4]根据

方剂药物贡献度进行组网, 同时对标签传递算法 (COPRA) 进行改进, 在传递过程中根据聚集系统来选择社团标签, 挖掘中药配伍规律; 吕庆莉^[5]使用复杂网络构建中医用药数据拓展网络, 针对脑血管疾病改进 COPRA 重叠社区发现算法, 探讨方剂配伍规律。

本文采用复杂网络重叠社区发现算法——SPLA (speaker-listener label propagation algorithm) 算法进行药物配伍规律的挖掘, 同时利用方剂中药物的君、臣、佐、使特性及药物用量关系来确定核心药物度, 并将核心药物度值作为参数, 改进 SPLA 算法, 提升聚类结果的目标性和针对性。通过选取治疗“产后中风”经典方剂 145 首, 分别利用药物出现频率、关联规则挖掘和本文提出的方法进行计算分析, 通过对比, 分析证实该方法较传统方法可挖掘一些新颖、可信的药物配伍, 从而对于阐明复方配伍的科学内涵、完善中药药性理论、指导中药复方开发产生积极的促进作用^[6-8]。

1 建模方法

1.1 定义方剂子网

由于在一首方剂中药物存在君、臣、佐、使作用之分, 如果抛开君药、臣药去研究佐药、使药的

基金项目: 包头医学院科学研究基金 (BYJJ-QM-2018011)

* 通讯作者: haoxx86@163.com

配伍关系则意义不大,因此,针对药物配伍规律的研究应以君药、臣药为核心。但要注意的是,在选取核心药物时若仅以出现频率较高或复杂网络模型中度值/聚类系数较高的药物作为核心,那么如甘草这类在方剂中经常以调和诸药的“佐使药”作用出现的药物就会对计算结果产生干扰。本研究以每个方剂为对象,选择方剂中的药物作为网络节点,将方剂中药物分类为“君臣药”和“佐使药”,将君药、臣药作为网络节点分别与其他药物建立连接边,佐药、使药只与君药、臣药建立连接边而彼此间不连接,这样君药、臣药的节点度会更大,其核心药物的地位会更突出。例如,方剂小独活汤中独活为君药,葛根为臣药,生姜和甘草分别为佐药、使药,建立的方剂子网如图 1 所示。从图中可知,君药、臣药独活与葛根的度值均为 3,佐药、使药生姜与甘草的度值均为 2,君臣药的核心地位凸显。

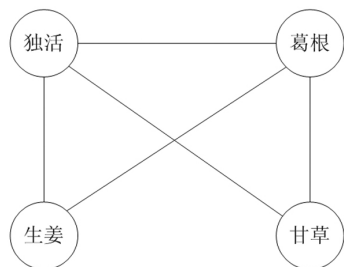


图 1 小独活汤方剂子网图

1.2 定义药物权重

在方剂用药中,一般情况下药物的功效决定药量的大小,主药的用量相比辅药要大,这进一步说明方剂中的核心药物与用药剂量存在正比关系,因此,定义方剂中核心药物度权重值为药物所连接边数(药物的度)与用药剂量比例的乘积,具体公式如下。

设方剂中药物数量为 N ,第 i 味药物的度为 d_i ,用药量为 l_i ,在方剂子网中与 i 相连的药物节点的集合为 $N(i)$,与 i 相连药物节点的个数为 $|N(i)|$,则该药物的权重 w_i 可如下定义:

$$w_i = \frac{d_i}{N} \times \frac{\sum_{j \in N(i)} l_j}{|N(i)|} \quad (1)$$

由公式(1)可以得出,方剂中每一味药物的核心度权重与方剂中药物数量无直接关系,主要与药物在方剂中的作用及与其他药物剂量比例有关,所以该权重可有效体现药物在某一方剂中的重要性。

1.3 建立中药复杂网络模型

通过复杂网络算法进行药物配伍规律的挖掘,

前提条件是要建立相关病症方剂的中药复杂网络模型。本研究以药物为节点,药物间的连接为边,建立复杂网络模型,在构建过程中,所有出现频率大于阈值 v 的药物作为复杂网络模型的节点,如果药物之间在方剂子网中存在连接边,则在复杂网络模型中也建立该连接边。经验证通过上述方法建立的中药方剂复杂网络模型节点度符合幂指数分布特点,可进行重叠社团的发现。

在构建的复杂网络模型中定义节点核心药物度权重 W 为该节点方剂子网中核心度权重的平均值,具体公式如下。

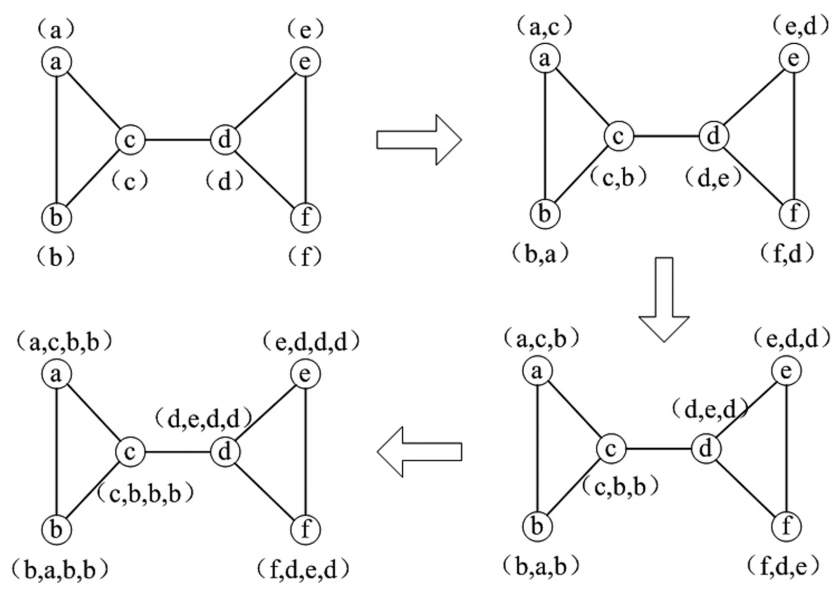
$$w_i = \frac{\sum_{j \in S(i)} w_j}{|S(i)|} \quad (2)$$

其中 $S(i)$ 代表所有包含节点 i 的方剂子网的集合, w_j 代表方剂子网中节点 i 的核心度权重, $|S(i)|$ 为包含节点 i 的方剂子网个数。

1.4 改进 SLPA 算法发现重叠社团

SLPA 算法是由 Xie 等^[7]提出的一种保留历史标签的标签传递算法,该算法首先给网络中每一个节点标记为自身标签,然后开始进行迭代传递标签,在迭代过程中每个节点(speaker)向邻居节点发送自身标签,接收节点(listener)根据邻居节点发送过来的标签并随机选择一个存入本标签的存储空间内,这样经过 t 次迭代,节点的标签存储空间中共包含 t 的标签,此时设定阈值 r ,如果标签存储空间内某一标签出现的比例低于 r 则被删除,剩余的标签即该节点所属的社团标签,如果节点包含至少 2 个标签,说明该节点为重叠节点,即该节点会出现在不同的聚类结果中。SLPA 算法标签传递的过程如图 2 所示。

SLPA 算法性能及稳定性整体要优于 LFM、COPRA 重叠社团发现算法,且不需要提前预估发现社团的数量,但是在迭代过程选择归属标签时存在随机性的问题,导致结果的不稳定性。本研究根据核心药物节点特性,对传递标签的选择方法进行改进,以提升社团发掘的有效度。具体改进方法为:1) 针对接收节点选择邻居发送标签随机性的问题,采用上一节提出的核心药物度权重进行标签的选择,即接收节点在接受各个邻居节点发送过来的标签时,优先选择核心药物度权重最大的节点存入,如果权重相同,则随机选取一个。2) 发送节点包含多个标签时,优先选择核心药物度权重最大的标签进行发送传递。3) 通过上述方法,随着迭代次数的增加,会使核心节点标签有更大的几率被



注: (a)(b)(c)(d)代表节点名称; (a) (b) (c) (d) 代表要传递的标签名称

图 2 SLPA 算法标签传递过程图

发送和接收,这使得每个节点中携带的核心节点度大的标签越来越多,从而使聚类结果围绕核心度高的节点有效聚集。

此外,由于 SLPA 算法在迭代过程中易产生规模过大的社团,这样的结果结构不清,分析意义不大,因此借鉴文献 [6] 中定义的 Max_c (max community) 参数对传播过程中社团包含节点数目及迭代次数进行控制,以 Max_c 参数来记录当前最大社团中节点数目,在迭代过程中如果社团中节点个数超过给定阈值时,则抑制该社团标签的传播,重新选取新的传播标签,当前后两次迭代社团个数不发生变化,则算法迭代终止。通过该方法可有效控制社团节点数目和迭代次数。

2 结果及分析

收集《中医方剂大辞典》[8] 中治疗产后中风的相关方剂共计 143 首。

2.1 词频及药物核心度结果

通过数据整理、清洗、规约构建数据源,利用 SPSS Modeler 18.0 软件计算出词频统计结果、核心药物度权重及关联规则分析结果。计算方剂中包含的 185 种药物中出现频率大于 20 的所有药物,同时按照公式 (2) 计算出对应药物的核心药物权重 (见表 1)。根据统计,185 种药物的平均核心药物权重为 1.69,可见表 1 中大部分药物的权重值高于该值,而甘草、人参等药物的权重较低,与这些药物在产后中风方剂中多以佐使药角色出现有关。

2.2 关联规则 Apriori 算法挖掘结果

设置最小支持度为 10,置信度 > 80,最大前项数为 3,使用 Apriori 算法共发现关联规则 88 条,其中置信度较高的 10 条规则如表 2 所示。

表 1 《中医方剂大辞典》治疗产后中风用药词频及药物核心度统计

序号	药名	词频	权重	序号	药名	词频	权重
1	防风	82	4.21	9	麻黄	40	3.66
2	肉桂	80	3.08	10	附子	36	3.28
3	当归	73	2.82	11	羌活	33	4.17
4	川芎	64	3.10	12	羚羊角	25	2.80
5	独活	57	5.68	13	生地黄	25	2.96
6	甘草	51	1.60	14	天麻	24	7.01
7	白芍	48	2.77	15	秦艽	22	1.71
8	人参	41	1.50	16	白术	20	2.04

表 2 《中医方剂大辞典》治疗产后中风用药关联规则 Apriori 算法挖掘结果

序号	药物组合规则	支持度	置信度
1	羚羊角, 肉桂 == > 防风	13.287	94.737
2	附子, 川芎 == > 肉桂	13.287	94.737
3	羚羊角, 川芎 == > 防风	11.888	94.118
4	人参, 甘草, 肉桂 == > 川芎	11.189	93.750
5	附子, 川芎, 当归 == > 肉桂	10.490	93.333
6	甘草, 川芎, 防风 == > 人参	16.084	91.304
7	人参, 川芎, 肉桂 == > 防风	13.287	89.474
8	人参, 肉桂, 防风 == > 川芎	13.287	89.474
9	人参, 白芍 == > 防风	12.587	88.889
10	甘草, 川芎, 当归 == > 人参	12.587	88.889

2.3 复杂网络改进 SLPA 算法挖掘结果

按照本文中的方法构建方剂复杂网络,设置 v

值为 4, 即删除出现频率小于 4 的药物节点及其关联边, 得到关于治疗产后中风中药复杂网络模型, 该复杂网络模型共包含药物节点 74 个, 药物连接边 1245 条, 然后利用 Java 语言自编程实现本文提出的改进 SLPA 算法, 控制发现社团中药物数量上限为 4 (Maxc 参数), r 值经过实验设置为 0.2, 可得到比较理想的聚类效果。共产生聚类结果 51 个, 部分有代表性 (即药物组合中药物数量及搭配具有合理性) 计算结果如表 3 所示。

表 3 《中医方剂大辞典》治疗产后中风用药复杂网络改进 SLPA 算法挖掘结果

序号	药物组合
1	麻黄、肉桂、甘草、人参
2	独活、当归、细辛、牛膝
3	川芎、荆芥、羌活、甘草
4	天麻、当归、秦艽、川芎
5	葛根、生姜、肉桂、甘草
6	独活、桑寄生、当归、白芍
7	防风、吴茱萸、天雄、生姜
8	白鲜皮、独活、人参、茯苓

产后中风是产后感受外邪而引起的疾病, 患者常出现关节肌肉疼痛、恶风畏寒等症状, 古籍认为其源于产伤动气血, 脏腑劳损, 风邪乘虚伤之^[1], 而临床上也认为本病多因百脉空虚、气血俱去、筋脉失养而致, 在药物治疗中应以补益气血、扶正祛邪、祛风胜湿、表里兼顾为主^[1]。通过表 1 数据结果可知, 防风、肉桂、当归、川芎、独活、甘草、白芍、人参、麻黄、附子、羌活等是治疗产后中风使用频率较高的药物, 其中防风、麻黄、羌胡、独活等祛风散邪, 兼以当归、川芎、肉桂、白芍等补血助阳, 是为治疗产后中风之主药。表 2 的药物组合规则体现以辛温解表类、补血补气类、熄风止痉类和祛风止痛类药物为主的用药配伍规律。

通过分析关联规则挖掘的结果, 发现存在两点不足: 一是由于设置支持度的影响, 关联规则中挖掘的结果基本为出现频率较高的药物组合规则, 一些出现频率较低的药物被过滤掉; 二是会出现一些佐使药的关联规则组合, 这对于分析功效意义不大。通过表 3 的药物组合可以看出, 其结果与关联规则的挖掘结果不尽相同, 该组合是围绕 1 味或多味核心药物产生的聚类结果。例如麻黄、肉桂、甘草、人参的组合, 其中麻黄可祛表邪、散风寒, 配伍肉桂可增强发汗解表之功^[12], 辅以甘草、人参补气正阳。此组合可见于《圣济总录》中小续命汤方中, 治疗妇人产后失血中风, 冒昧不知痛处,

拘急不得转侧, 四肢缓急遗失便利。又如独活、当归、细辛、牛膝的组合, 其中独活散寒解表, 配伍细辛、牛膝既可祛除风、寒、湿、邪又可通痹止痛^[13], 辅以当归补血活血, 肝肾强健, 气血充足。此组合可见于《圣济总录》独活饮方中, 治疗产后中风偏枯, 半身不收, 麻痹不仁。根据统计使用本文提出方法得到的药物组合有 80.4% 在关联规则中没有出现, 进一步说明该方法在挖掘结果上具有一定的差异性。

3 小结

本研究利用方剂中药物的作用及剂量设计了一种基于核心药物度的方剂用药规律挖掘方法, 其计算结果围绕君臣药及剂量大的药物产生了众多新颖的药物组合规律, 虽然本法还存在诸多不足之处, 例如缺乏与证候、性味、归经等的有效结合, 但研究结果可以为总结用药规律、指导中药复方开发提供一定参考。此外, 本研究方法也可作为一种新的分析中医用药规律的工具和手段, 为揭示配伍的科学内涵和用药特点提供有力的支撑。

参考文献

- [1] 吕庆莉. 数据挖掘与复杂网络的融合及其在中医药领域的应用 [J]. 中草药, 2016, 47(8): 1430-1435.
- [2] 李茹, 孙正, 王崇骏, 等. 中药方剂药物属性的组网模型 [J]. 智能系统学报, 2014, 9(2): 148-153.
- [3] 韩楠, 乔少杰, 李天瑞, 等. 面向复杂网络的中药方剂配伍规律挖掘算法 [J]. 计算机科学与探索, 2017, 11(7): 1159-1164.
- [4] 孙道平, 高原, 谢隽, 等. 一种用于中药方剂网络重叠社区发现的改进 COPRA 算法 [J]. 南京大学学报(自然科学版), 2013, 49(4): 483-490.
- [5] 满姗姗, 王立新. 基于数据挖掘及辨证论治探讨中药治疗肝病的用药特点 [J]. 中草药, 2016, 47(15): 2775-2778.
- [6] 李泮霖, 苏薇薇. 网络药理学在中药研究中的最新应用进展 [J]. 中草药, 2016, 47(16): 2938-2942.
- [7] XIE JR, KELLEY S, SZYMANSKI BK. Overlapping community detection in networks: the state-of-the-art and comparative study [J]. ACM Computing Surveys, 2012, 45(4): 115-123.
- [8] 赵雨露, 张曦煌. 一种优化标签传播过程的重叠社区发现算法 [J]. 计算机应用研究, 2018, 35(3): 765-768.
- [9] 彭仁怀. 中医方剂大辞典 [M]. 北京: 人民卫生出版社, 1993.
- [10] 陈自明. 妇人大全良方 [M]. 北京: 人民卫生出版社, 2004: 1838.

- 〔1〕李杰,韩涛,宋立家,等. 基于中医传承辅助系统的治
疗产后中风方剂的组方规律分析〔J〕. 中国药房,
2016,27(2):171-173.
- 〔2〕林红菊,牟艳林,南奕阳,等. 经方中麻黄相关药对的

- 研究进展〔J〕. 亚太传统医药,2018,14(11):110-111.
- 〔3〕曹雯,李崇超. 独活的本草考证及常用药对〔J〕. 陕西
中医药大学学报,2018,41(1):127-130.

Construction of Complex Network Algorithm for the Core Drug Compatibility Law of Prescriptions Based on Drug Dose and Effect

MA Ning¹, HAO Xiuxia², BAI Jinniu¹, XING Junfeng¹

1. Baotou Medical College, Inner Mongolia University of Science and Technology, Baotou, 014040; 2. The First Affiliated Hospital of Baotou Medical College, Inner Mongolia University of Science and Technology

ABSTRACT Using a complex network-related algorithm, combined with the drug action and dose of “monarch, minister, assistant and guide” drugs in the prescription, a new algorithm based on the drug dose and effect of the core drug compatibility rule of the prescription was constructed. By constructing the prescription subnet, the core drugs and their weights were determined according to the drug action and dose in the prescriptions. The complex network model of the traditional Chinese medicine prescriptions was constructed based on the core weights of the drugs. The complex network label transfer algorithm was improved to mine the core drug compatibility rules of related prescriptions. Taking prescriptions and medications for treating postpartum affection by cold in *Dictionary of Traditional Chinese Medicine Prescriptions* (《中医方剂大辞典》) as an example, 143 related prescriptions and 16 flavors of drugs with a frequency greater than 20, 88 association rules with a support greater than 10, and a confidence greater than 80 were collected to construct a complex network model of postpartum stroke medications with 74 nodes and 1245 edges. Fifty-one complex clustering results were found using a complex network algorithm. Through analysis and research, it was found that the traditional Chinese medicine for postpartum affection by cold is mainly through removing wind and dampness, dispersing cold and analgesic, and nourishing qi and blood. At the same time, the improved label delivery algorithm is compared with traditional association rules and other methods. It is considered that improved complex network label delivery algorithm can obtain some new and more targeted drug combination results around the core drugs.

Keywords complex network; prescription compatibility law; monarch, minister, assistant and guide; drug dose; SLPA algorithm; postpartum affection by cold

(收稿日期: 2019-03-28; 修回日期: 2019-08-13)

编辑: 柴倩云

欢迎订阅 2020 年《中医杂志》

《中医杂志》(ISSN 1001-1668, CN 11-2166/R) 是由中华中医药学会和中国中医科学院主办的全国性中医综合性学术期刊。1955 年创刊以来始终坚持“以提高为主, 兼顾普及”的办刊方针, 是我国中医药界创刊早、发行量大、具有较高权威性和学术影响力的国家级中医药期刊之一, 是中文核心期刊和科技核心期刊、中国精品科技期刊、首届国家期刊奖获得者和中国期刊方阵双奖期刊、新中国 60 年有影响力的期刊、中国百强科技期刊、中国百种杰出学术期刊, 荣获第二届、第三届、第四届新闻出版政府奖期刊奖提名奖, 中医药科技期刊分级目录 T1 级期刊。

《中医杂志》主要栏目中当代名医和临证心得分别介绍名老中医和临床医生辨证用药治疗疑难病的经验; 临床研究介绍中医药治疗的新方法、新成果; 学术探讨、思路与方法、专家论坛、病例讨论、综述、百家园等栏目, 提供最新学术观点、研究成果与治疗方法, 成为学习中医药、研究中医药, 不断提高临床及研究水平的良师益友。

《中医杂志》为半月刊, 每月 2 日和 17 日出版, 每期定价 20.00 元, 全年 480 元。读者可以到全国各地邮局办理订阅手续(邮发代号: 2-698), 也可以与本刊读者服务部联系邮购, 邮购免邮费。电话: 010-64035632, 010-64089195。国外发行: 中国国际图书贸易总公司(北京 399 信箱, 邮编: 100044, 代号: M140)。本社地址: 北京市东城区东直门内南小街 16 号, 邮政编码: 100700。网址: <http://www.jtcm.net.cn>。

扫一扫 关注

