

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[△]

潘树茂^{1*}, 章新友^{1#}, 吴地尧¹, 郭永坤², 丁亮², 谈荣珍³ (1. 江西中医药大学药学院, 南昌 330004; 2. 江西中医药大学计算机学院, 南昌 330004; 3. 南昌市洪都中医院骨质疏松科, 南昌 330006)

中图分类号 R2-03 文献标志码 A 文章编号 1001-0408(2020)24-2966-08

DOI 10.6039/j.issn.1001-0408.2020.24.04

摘要 目的:建立“中药数据策略模式分析平台”(TCMDP),并据此对临床治疗原发性肝癌(PHC)的中药处方进行组方规律分析,为临床治疗PHC提供有益参考。方法:根据策略模式思想研发了封装有频次统计、关联、聚类、分类等多个方法及其涵盖的不同算法的TCMDP软件。从南昌市4所省级三级甲等医院信息系统和中国知网、PubMed数据库中,收集治疗PHC的中药有效处方并建立相应的PHC处方数据库(PDOPHC)。将该数据库中的中药处方导入TCMDP,借助CMC-DD改进聚类算法、融合深度和宽度搜索后改进的WD-Get Rules算法和频次统计方法分析上述处方的证型分布,并以其中的痰瘀互结证型肝癌(TYPHC)为例进行药物功效类别、四气五味归经、药性组合、核心药对和核心药物组合的数据挖掘。结果:共收集到907张治疗PHC的中药处方,主要涉及痰瘀互结证、正虚瘀结证、肝脾血瘀证等10个证型。其中,TYPHC相关的中药处方中常用补虚药、化痰药、活血化瘀药等;四气多为寒、温、平,五味多为甘、苦、辛,归经多为脾、肝、胃、肺、心、肾经;药性组合多为温甘脾、温辛脾和寒苦肝等模式;用药频次≥30次的药对共有36对;处方中的强关联规则共31条,核心药物组合8组。WD-Get Rules算法结果显示,临床治疗TYPHC的处方中以白术、党参、茯苓、甘草等4味中药处于核心地位,且常与多种药物配伍出现,这与四者通常组方为四君子汤以达健气补脾之功有关;CMC-DD算法结果显示,临床治疗TYPHC以健补脾胃、补益气血为主,同时又从疏肝理气、化痰、清热解毒、活血化瘀、滋阴等多角度治疗TYPHC。结论:临床上治疗TYPHC的中药处方主要采用健脾益气、疏肝理气、活血化瘀等方法多角度进行治疗,符合医家“扶正祛邪”的用药法则;应用TCMDP能多层次、多方位地实现中药数据的挖掘与分析,可为中药方剂配伍规律的挖掘研究提供一种新的有效途径。

关键词 中药数据策略模式分析平台;策略模式;中药处方;原发性肝癌;痰瘀互结证型;数据挖掘;组方规律

Analysis of Medication Rules of TCM Prescription against Primary Hepatic Carcinoma Based on TCM Data Analysis Platform Based on Strategy Pattern

PAN Shumao¹, ZHANG Xinyou¹, WU Diyao¹, GUO Yongkun², DING Liang², TAN Rongzhen³ (1. School of Pharmacy, Jiangxi University of TCM, Nanchang 330004, China; 2. School of Computing, Jiangxi University of TCM, Nanchang 330004, China; 3. Dept. of Osteoporosis Nanchang Hongdu Hospital of TCM, Nanchang 330006, China)

ABSTRACT OBJECTIVE: To establish “TCM Data Analysis Platform Based on Strategy Pattern” (TCMDP), and to analyze the medication rules of TCM prescription against primary hepatic carcinoma (PHC), so as to provide reference for TCM prescription screening. METHODS: According to the idea of strategy pattern, TCMDP software was developed, and the frequency statistics, association, clustering, classification and other methods and the different algorithms covered were encapsulated. Effective TCM prescriptions against PHC were collected and screened from 4 tertiary-grade class-A hospitals in Nanchang city and CNKI, PubMed. A database of prescriptions for PHC (PDOPHC) was established, and TCM prescriptions in the database were put into TCMDP. Clustering algorithm modified by CMC-DD, modified WD-Get Rules algorithm after merging depth and width search and frequency statistics were used to analyze the distribution of prescription syndrome. Taking phlegm-blood stasis syndrome type (TYPHC) as example, data mining was carried out on drug efficacy category, four properties, five flavors and channel tropism, drug property combination, key couplet medicines and key drug combination. RESULTS: A total of 907 TCM prescriptions against PHC were collected, mainly involving 10 syndrome types, such as phlegm-blood stasis syndrome, deficiency of the vital essential

[△] 基金项目:国家自然科学基金资助项目(No.81660727)

* 硕士研究生。研究方向:中药信息挖掘与应用。E-mail: psm6418@163.com

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。E-mail: xinyouzhang@jxutcm.edu.cn

and blood stasis syndrome and liver and spleen blood stasis syndrome. TCM related to TYPHC commonly used in prescriptions were tonifying deficiency medicine, phlegm-resolving medicine, promoting blood circulation and removing blood stasis medicine, etc. Four properties were mostly cold,

warm and calm; the five flavors were mostly sweet, bitter and pungent; channel tropism were mostly spleen, liver, stomach, lung, heart and kidney; drug combination were mostly warming-sweet-spleen, warming-pungent-spleen and cold-bitter-liver, etc.; there were 36 couplet medicines with frequency ≥ 30 times; there were 31 strong association rules and 8 key drug combinations in the prescriptions. WD-Get Rules results showed that in the TCM prescriptions against PHC, *Atractylodes macrocephala*, *Codonopsis pilosula*, *Poria cocos* and *Glycyrrhiza uralensis* were in the key position. They were often combined with many kinds of drugs, which were related to the effect of Sijunzi decoction on strengthening Qi and tonifying spleen. CMC-DD results showed that therapy for TYPHC focused on strengthening the spleen and stomach, replenishing Qi and blood; and at the same time, it treated TYPHC with phlegm-blood stasis syndrome from multiple angles, such as soothing the liver and regulating Qi, eliminating phlegm, clearing heat and detoxification, promoting blood circulation and nourishing Yin. CONCLUSIONS: Clinical treatment of TYPHC mainly adopts the methods of invigorating the spleen and replenishing Qi, soothing the liver and regulating Qi, promoting blood circulation and removing blood stasis, which conforms to the medication principle of "strengthening the body and eliminating pathogenic factors". The application of TCMDP can realize the data mining and analysis of TCM in multi-level and multi-directional way, which provides a new and effective way for the mining and research of compatibility rules of TCM prescriptions.

KEYWORDS TCM Data Analysis Platform Based on Strategy Pattern; Strategy pattern; TCM prescription; Primary hepatic carcinoma; Phlegm-blood stasis syndrome type; Data mining; Compatibility rules

肝癌是我国常见恶性肿瘤之一,并以原发性肝癌(Primary hepatic carcinoma, PHC)为主^[1]。长期的中医临床实践为肝癌的诊疗积累了丰富的经验,并留有大量的中药处方数据^[2]。如何有效地利用中药处方数据,如何发现隐藏在它们背后的潜在关系并将其合理有效地用于临床,这些都是当今研究的热点问题。

数据挖掘技术是中医药研究的重要手段之一,常用的数据挖掘方法有关联规则、回归分析、决策树、聚类分析、频次分析和人工神经网络等^[3]。李文林等^[4]利用关联规则方法,分析了明清时期中医治疗疫病的药-症关系,为疫病的预防和治疗提供了借鉴。刘耸峰等^[5]利用聚类方法对中药复方治疗特发性肺纤维化的用药规律进行分析,发现该症的临床治疗以益气养阴、活血化瘀通络为主。但是每种数据挖掘方法通常只能解决单一问题,在数据挖掘的过程中存在一定的局限性,无法对问题进行多维度的深入分析。因此,本研究提出一种基于策略模式的中药数据挖掘新方法,其可以采用多个数据挖掘方法解决多个问题或优选出某一方法来解决某一问题,最大可能地挖掘中药处方中的潜在信息,解决单一数据挖掘方法无法解决的难题。

本课题组首先根据策略模式思想研发了“中药数据策略模式分析平台”(Traditional Chinese Medicine Data Analysis Platform Based on Strategy Pattern, TCMDP);其次,收集大量治疗PHC的中药处方,并以PHC的某一疾病证型为例,探讨中医“方-证”之间的用药规律,以期为临床治疗PHC提供有益参考;同时,本课题组所开发的TCMDP软件(目前正在申请软件著作权)也有望为中药方剂配伍的研究提供一套新的方法。

1 资料与方法

1.1 TCMDP的建立

目前,数据挖掘方法广泛应用于中药信息数据挖掘中,但传统的数据挖掘方式多为一对一的方式,即一种数据挖掘方法处理一类中药挖掘问题(如图1所示),但

此种方式无法充分保证最优或最有效地挖掘中药信息,从而导致得出的结果具有一定的局限性^[6]。为此,本研究提出了一种基于策略模式的中药数据挖掘方法,其能够在不同的方法和不同的问题之间实现“多对多”关系的中药数据挖掘(如图2所示)。策略模式是一种行为模式,其核心思想是算法的封装,在封装过程中,算法的职责与算法本身分离,由不同的对象管理^[3]。该模式通常将一系列算法封装在一系列策略类中,使其作为一个抽象策略类的子类,即策略模式定义了一系列算法,封装了每个算法,并使这些算法可以相互替换^[3,7]。基于策略模式建立的该数据挖掘方法可使某个问题被多个算法处理,亦可使单个算法应用到多个问题的解决上,从而可以多层次、全方位地挖掘中药中隐含的关系,不仅大大提高了挖掘效率,而且可以实现方法与问题之间的匹配优选,以及数据挖掘方法的优劣比较。



图1 传统的中药信息数据挖掘方式的“一对一”关系示意图

Fig 1 Schematic diagram of single relationship in traditional data mining mode of TCM

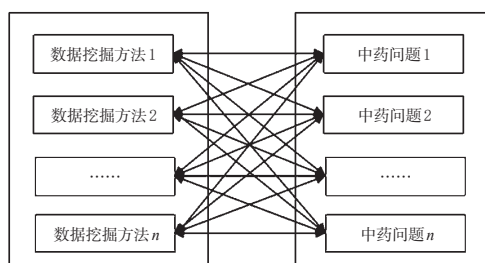


图2 基于策略模式的中药信息数据挖掘模式的“多对多”关系示意图

Fig 2 Schematic diagram of many-to-many relationship in data mining of TCM based strategy mode

本课题组根据策略模式思想,研发了TCMDP这一软件,其中封装了频次统计、关联、聚类、分类等多个方法,在每个不同的方法里面又有多个不同的算法(如图3所示)。以聚类挖掘为例,其功能操作界面如图4所示。

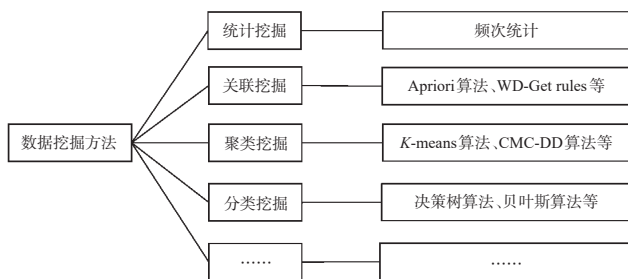


图3 TCMDP的数据挖掘方法示意图

Fig 3 Schematic diagram of data mining method of TCMDP



图4 TCMDP中聚类挖掘功能界面图

Fig 4 Clustering mining function interface diagram of TCMDP

1.2 数据挖掘方法示例

在本研究中,以治疗PHC的中药处方为例,采用TCMDP软件中最常用的CMC-DD算法[Clustering method combining DBSCAN and DPCA,即融合了基于密度噪声应用的空间聚类算法(Density based spatial clustering of applications with noise, DBSCAN)和密度峰值聚类算法(Density peaks clustering algorithm, DPCA)的改进聚类算法]、WD-Get Rules算法(Width and depth-get rules,即融合了深度和宽度搜索的改进关联规则算法)和频次统计方法,对此类处方的配伍规律进行分析。

1.2.1 CMC-DD算法

CMC-DD算法是融合了DBSCAN算法和DPCA算法的改进聚类算法。DPCA算法可以快速地发现数据集的密度峰值点,并能高效地分配样本点和剔除离群点,在社区发现、图像处理、计算机视觉和文本处理等领域得到广泛应用,但是截断距离会对其聚类结果产生影响^[8];DBSCAN算法是基于密度的聚类算法,它不需要预先指定聚类的簇数,能够在含有噪声的数据集中发现任意数量和形状的簇,所以在聚类分析中有着广泛的应用^[9-10]。为了消除截断距离对DPCA算法聚类结果的影响,本研究提出了CMC-DD算法,即主要利用DBSCAN

算法优选出最佳的密度半径,进而消除截断距离对DPCA算法局部密度的影响。

为使DBSCAN算法优选出最佳的密度半径,CMC-DD算法首先根据文献算法^[10]优选出最佳密度参数(ϵ_r)来计算每个数据对象的局部密度(ρ),再计算出每个数据对象的聚类中心距离(δ),计算出 $\rho \cdot \delta$ 的标准差(STD),选出满足条件STD值最大的 k (簇类数)个点作为聚类中心,最后把剩余的点划分到相邻最近的类中。

①CMC-DD算法基本概念如下:以 ρ 表示数据对象的局部密度, δ 表示数据对象的聚类中心距离,设样本数据集: $G = \{x_1, x_2, \dots, x_n\}$ 。

定义1(欧氏距离)——属于常见距离度量中的一种,衡量的是空间中两点 x 和 y 间的直线距离,其计算公式如下:

$$d(x, y) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n (x_i - y_j)^2} \dots \dots \dots (1)$$

定义2(局部密度)——数据对象 x_i 的局部密度 ρ_i 的计算公式如下:

$$\rho_i = \sum_{x_j \in \Omega} e^{-\frac{d(x_i, x_j)}{\epsilon_r}} \dots \dots \dots (2)$$

式(1)(2)中 Ω 为 ϵ_r 的邻域, $|\Omega|$ 为 Ω 内数据点数目, $d(x_i, x_j)$ 为数据点 x_i 与 x_j 间的欧氏距离。

定义3(聚类中心距离)——聚类中心距离 δ_i 的计算公式如下:

$$\delta_i = \begin{cases} \min_j \rho_j < \rho_i [d(x_i, x_j)] & i \geq 1 \\ \max_j \rho_j d(x_i, x_j) & \rho_i = \max \end{cases} \dots \dots \dots (3)$$

定义4(聚类中心选择的约束条件)——聚类中心选择的约束条件的计算公式如下:

$$\Phi = \rho_i \delta_i \dots \dots \dots (4)$$

式(3)(4)中, $\rho_i \delta_i$ 为每个数据点的局部密度 ρ_i 与聚类中心距离 δ_i 的乘积形成的列表数据。

②CMC-DD具体算法流程:CMC-DD聚类算法能够很好地解决截断距离参数的影响,且无需人为干扰算法聚类中心的选择,其具体算法流程见图5。

1.2.2 WD-Get Rules算法

WD-Get Rules算法为本课题组设计的改进关联规则算法,可以更加有效地挖掘出隐藏在数据集中有意义的关系。该算法首先采用宽度优先的策略,找出能作为强关联规则的后件集 H ,该集合中每个后件为1项集,以 H 为基准,采用集合枚举树的关联规则进行深度搜索^[11]。不同于集合枚举树的关联规则,WD-Get Rules算法在进行深度搜索时,规则后件仅源自于 H 中所含有的元素,这样可避免不能作为规则后件的元素;且若项集是频繁 $k(k > 2)$ 项集, H 中的元素个数大于1,才进行深度搜索,否则对下一个频繁项集进行操作,其具体算法流程详见文献[12]。

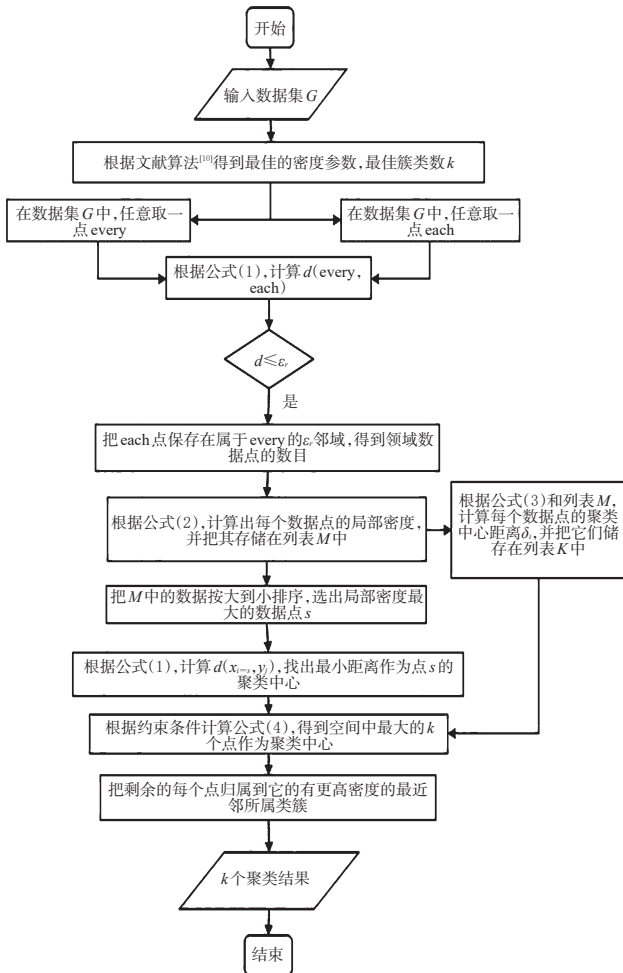


图5 CMC-DD具体算法流程

Fig 5 Flow of CMC-DD algorithm

1.2.3 频次统计方法

频次统计是指根据某些标准对总数据进行分组, 并对每组中的个体数量进行计数。

1.3 中药处方收集与数据库建立

在江西省南昌市4个区(东湖区、青山湖区、西湖区、红谷滩区)的4所三级甲等医院(江西省中医院、江西中医药大学第二附属医院、南昌市第三医院、南昌市洪都中医院), 检索并提取所有临床诊断为“原发性肝癌”的电子中药处方, 检索时限为2016年1月至2019年12月。剔除无效处方(即临床诊断与用药不符的处方), 剔除相同处方(即排除同一医师或同一科室开具的经验处方), 剔除无对应证型的处方, 剔除证型描述不规范的处方, 剔除伴有其他疾病的处方, 最终筛选出临床治疗PHC的有效中药处方。在中国知网(CNKI)进行文献检索, 检索式为: SU=(‘原发性肝癌’+‘肝肿瘤’)*(‘中药’+‘复方’+‘汤’+‘方’)*(‘疗效’+‘变化’+‘对照组’+‘临床’)*‘随机’-‘鼠’-‘动物’-‘meta’; 检索时限为建库起至2020年5月31日。在PubMed中进行文献检索, 检索式为: ((search(‘liver cancer’[MeSH]) OR (‘Primary hepatic carcinoma’[MeSH]) AND (‘TCM’ OR ‘traditional Chinese medi-

cine’)); 同时, 将文献类型设为“Clinical Trial”, 检索时限为建库起至2020年5月31日。

结果, 从CNKI共得到17篇原始文献, 阅读题目和摘要, 剔除13篇与中药治疗PHC无关的文献; 阅读全文, 剔除无对应证型以及随机分组不规范的文献3篇, 最终得到1篇文献^[13]; 从PubMed共得到12篇文献, 阅读题目和摘要, 剔除10篇与中药治疗PHC无关的文献, 最后阅读全文, 得到2篇文献^[14-15]。使用Excel 2019软件提取医院和文献里的中药处方(包括证型和中药), 共得到治疗PHC的有效中药处方907张, 将纳入处方中的中药按照2015版《中国药典》的要求规范化处理(例如将“元参”规范为“玄参”等), 即得到PHC处方数据库(Prescription Database of Primary Hepatic Carcinoma, PDOPHC)

1.4 治疗PHC的中药处方的录入与数据分析

将PDOPHC中的中药处方导入TCMDP中, 采用CMC-DD算法、WD-Get Rules算法和频次统计方法分析处方中的药物功效类别、四气五味归经、药性组合、核心药对和核心药物组合, 进而探讨治疗PHC的中药处方配伍规律。

2 结果与分析

2.1 基于频次统计方法的组方规律分析

2.1.1 治疗PHC的中药处方涉及证型分布

经频次统计显示, 907张治疗PHC中药处方的证型主要涉及痰瘀互结证、正虚瘀结证、肝脾血瘀证、气滞血瘀证等10个证型, 详见图6。

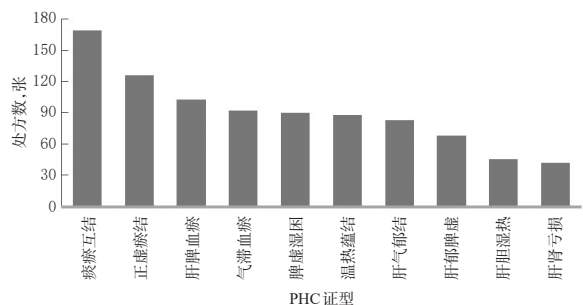


图6 治疗PHC的中药处方涉及证型分布

Fig 6 Distribution of syndrome types involved in TCM prescriptions against PHC

以证型进行配伍规律的分析更符合中医“辨证论治”的思想, 故本研究随机选取了PHC的某一证型, 即痰瘀互结型原发性肝癌(TYPHC)为例, 对其所涉及到的169张中药处方进行配伍规律的分析(其他证型的分析可参考本证型进行类推)。

2.1.2 治疗TYPHC中药处方药物功效分析

经频次统计显示, 治疗TYPHC的中药处方中的药物功效种类主要涉及到补虚、利水渗湿、清热、活血化瘀、理气等15种药物种类, 详见图7。

如图7所示, 在治疗TYPHC时, 医家多使用补虚药、利水渗湿药、清热药、活血化瘀药、理气药、化痰药以达

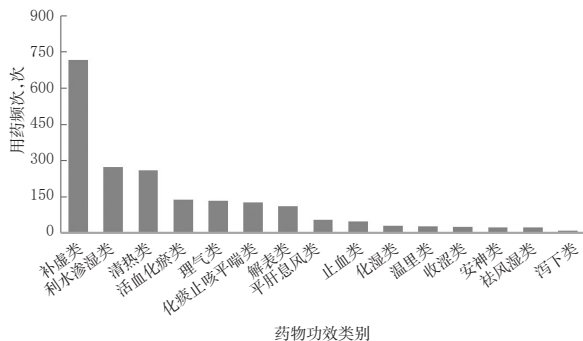


图7 治疗TYPHC的中药处方药物功效类别分布

Fig 7 Distribution of drug efficacy category of TCM prescriptions against TYPHC

到治病之目的。TYPHC患者多为正气亏虚、邪毒入侵，致脏腑失和、气血运行不畅、痰浊内生、瘀阻脉络、气滞血瘀痰凝，日久恶变则为肝癌；且因肝失疏泄，所以常伴有痰、饮、水、湿与瘀血内阻^[6]。中医治疗此证之本应是补益气血、调肝论治，再佐以活血化瘀、理气化痰、利水渗湿等法治标，标本兼治，可共奏奇效^[7]。总体来看，治疗TYPHC中药处方所用药物的功效符合上述医家用药法则。

2.1.3 TYPHC中药处方中药物的性味与归经分布

经频次统计显示，治疗TYPHC的中药处方中药物的四气多为寒、温、平，五味多为甘、苦、辛，归经多为脾、肝、胃、肺、心、肾经，其频次分布如图8所示。四气之中，寒温居多，其次为平，痰瘀互结日久而化热，故会多用寒性药以清热解毒，但一味使用寒性药攻邪亦会导致患者脾胃受损、癥瘕难消，因此又多用温性药，既能防止寒性药攻邪太过而导致的正气受损，又能达到温化痰瘀、暖肝散结等功效；而平性药药性平和，有利于阴阳互补、气血调和^[18-19]。可见，纳入处方的四气用药规律整体上体现了医家注重寒温平调的思想。五味之中，甘、苦、辛味药居多，甘味药具有扶正健脾的功效，可改善TYPHC患者的气血不足之症；痰瘀为阴邪，久则损伤肝络，苦可燥土疏木、泻其壅滞，辛可行气活血、疏通肝络，并使甘味药补而不滞，辛苦合用，为辛开苦降之法，可调畅中焦气机、健补脾胃，以正本清源；另外，辛入肺，又可通过佐金平木之法使肝功能恢复正常^[20]。纳入处方的药物，多入脾胃、肝肾、心肺经。TYPHC病位在肝，同时肝肾同源，故方中多有归肝肾经的药物；又因“肝病实脾”的治疗思想，所以归脾胃经的药物亦常用^[21]；另外，TYPHC以气血不足为本虚，而肺主一身之气、心主血脉、肝主疏泄、肾主纳气，脾胃又为气血生化之源，合理使用归此六经之药，可使患者气血通畅^[20-24]。因此，临床多配以归此六经的药物用于辨证治疗TYPHC，这也体现了中医大家重视“先天之肾，后天之脾”的用药思想^[22,25]。

2.1.4 TYPHC中药处方中核心药对分析

经频次统计显示，在治疗TYPHC的中药处方中用

药频次 ≥ 30 次的核心药对共有36组，并根据2015年版《中国药典》(一部)查询并归纳总结各组药对的功效，结果见表1。



图8 治疗TYPHC的中药处方中药物四气五味归经频次分布

Fig 8 Distribution of frequencies of the four properties, five flavors and channel tropism in TCM prescriptions against TYPHC

表1 治疗TYPHC的中药处方中用药频次 ≥ 30 次的核心药对

Tab 1 Key couplet medicines used more than 30 times in TCM prescriptions against TYPHC

核心药对	频次	功效	核心药对	频次	功效
白术-茯苓	73	健脾除湿	党参-鳖甲	37	健脾益气、软坚散结
白术-党参	68	益气健脾	白术-白芍	37	柔肝安脾
白术-黄芪	61	补气固表、健脾益气	柴胡-鳖甲	37	疏肝解郁、软坚散结
茯苓-党参	55	健脾、益气、除湿	党参-陈皮	36	理气健脾、燥湿化痰
白术-陈皮	52	健脾除湿	白术-茵陈	36	健脾益气、清热除湿
白术-鸡内金	50	补脾和胃、消积化滞	党参-薏苡仁	35	祛湿健脾、补益气血
甘草-黄芪	49	大补元气、清热解毒	陈皮-当归	34	理气健脾、活血调经
黄芪-当归	46	补气养血	甘草-白芍	34	滋阴敛肝、益气补中
党参-柴胡	46	疏肝解郁、补中益气	柴胡-白芍	33	疏肝养肝、调畅气血
白术-当归	46	健脾、补气、益血	柴胡-半边莲	32	疏肝解郁、利尿消肿
白术-半夏	44	健脾益气、化痰散结	陈皮-柴胡	32	疏肝、理气、化痰
黄芪-白花蛇舌草	41	清热解毒、补气固表	半枝莲-柴胡	32	疏肝解郁、清热解毒
白术-鳖甲	39	健脾和胃、软坚散结	茯苓-当归	31	利水渗湿、活血活血
茯苓-半夏	39	燥湿化痰、和胃化饮	白术-莪术	31	健脾胃、破血祛瘀
茯苓-陈皮	38	渗湿健脾、理气和中	半枝莲-白花蛇舌草	30	清热解毒、活血化痰
党参-黄芪	38	补中益气	陈皮-半夏	30	燥湿化痰
茯苓-黄芪	38	健脾益气、利水消肿	黄芪-茵陈	30	健脾益气、清热利湿
茯苓-柴胡	37	疏肝解郁、健脾化湿	黄芪-陈皮	30	补中益气、燥湿化痰

如表1所示，从功效来看，核心药对主要从健脾渗湿、补益气血、软坚散结、清热解毒、疏肝解郁、燥湿化痰等方面治疗TYPHC。“见肝之病，知肝传脾，当先实脾”，脾为气血生化之源，肝癌患者常肝失疏泄，进而导致脾失健运，故水津难运、聚湿成痰；另一方面，“气为血之帅”，脾气虚则导致血行无力出现气虚血瘀，进一步导致肝癌恶化，所以这也是为什么核心药对中健脾益气药对出现频次较高的原因之一^[26-27]。“脾为生痰之源”，脾胃健则痰湿自消；气行则血行，气血得行而瘀血渐消，在源头上治疗痰瘀之症，以达扶正之效^[28]，例如采用“党参-陈皮”药对理气健脾、燥湿化痰。痰、瘀互结日久而化热，故配伍清热解毒之药以祛热毒，例如采用“半枝莲-白花蛇舌草”药对有清热解毒之效。此外，在扶正基础上，医家还常辨证使用活血化瘀、软坚散结的药物，以达到消散结块之目的，例如采用“白术-莪术”药对配伍以健补

脾胃、破血祛瘀,采用“柴胡-鳖甲”药对配伍以疏肝解郁、软坚散结。由此可见,医家对核心药对进行合理应用,可发挥扶正祛邪的效果。

2.2 基于WD-Get Rules算法的组方规律分析

利用WD-Get Rules算法,设定支持度 ≥ 0.2 、置信度 ≥ 0.7 ,对治疗TYPHC的处方进行关联规则分析,结果见表2(表中,“ \Rightarrow ”为关联规则符号,表示前者出现的情况下后者出现的概率)。其中,高于最小支持度和最小置信度的规则即为强关联规则;提升度 > 1 则表示该强关联规则属于有效的强关联规则,提升度 < 1 则是无效的强关联规则^[29]。利用Cytoscape 3.6.1软件将强关联规则的药物进行网络可视化展示,结果见图9。

表2 治疗TYPHC的中药处方药物的关联规则分析结果

Tab 2 Association rule analysis results of TCM in prescriptions against TYPHC

序号	规则	支持度	置信度	提升度	序号	规则	支持度	置信度	提升度
1	党参 \Rightarrow 白术	0.402	0.907	1.459	17	半枝莲 \Rightarrow 甘草	0.308	0.776	1.338
2	茯苓 \Rightarrow 白术	0.432	0.901	1.451	18	白芍 \Rightarrow 白术	0.219	0.771	1.241
3	党参 \Rightarrow 甘草	0.396	0.893	1.541	19	白花蛇舌草 \Rightarrow 黄芪	0.243	0.759	1.734
4	柴胡 \Rightarrow 甘草	0.331	0.875	1.509	20	薏苡仁 \Rightarrow 甘草	0.201	0.756	1.303
5	薏苡仁 \Rightarrow 白术	0.231	0.867	1.395	21	半夏 \Rightarrow 甘草	0.231	0.750	1.293
6	茵陈 \Rightarrow 白术	0.213	0.857	1.380	22	半夏 \Rightarrow 茯苓	0.231	0.750	1.565
7	半夏 \Rightarrow 白术	0.260	0.846	1.362	23	半枝莲 \Rightarrow 白术	0.296	0.746	1.201
8	白花蛇舌草 \Rightarrow 白术	0.266	0.833	1.341	24	白花蛇舌草 \Rightarrow 当归	0.237	0.741	1.987
9	甘草 \Rightarrow 白术	0.479	0.827	1.330	25	鳖甲 \Rightarrow 白术	0.231	0.736	1.184
10	黄芪 \Rightarrow 白术	0.361	0.824	1.327	26	党参 \Rightarrow 茯苓	0.325	0.733	1.530
11	鳖甲 \Rightarrow 甘草	0.254	0.811	1.399	27	当归 \Rightarrow 白术	0.272	0.730	1.175
12	茯苓 \Rightarrow 甘草	0.379	0.790	1.363	28	当归 \Rightarrow 黄芪	0.272	0.730	1.668
13	陈皮 \Rightarrow 甘草	0.308	0.788	1.359	29	柴胡 \Rightarrow 党参	0.272	0.719	1.620
14	陈皮 \Rightarrow 白术	0.308	0.788	1.268	30	白芍 \Rightarrow 甘草	0.201	0.708	1.222
15	柴胡 \Rightarrow 白术	0.296	0.781	1.257	31	白花蛇舌草 \Rightarrow 甘草	0.225	0.704	1.214
16	薏苡仁 \Rightarrow 党参	0.207	0.778	1.753					

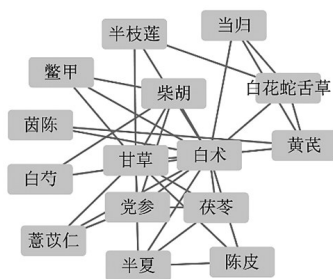


图9 治疗TYPHC的中药处方中强关联药物的网络关系图

Fig 9 Network diagram of strongly related drugs in TCM prescriptions against TYPHC

如图9所示,在治疗TYPHC的中药处方中,白术、党参、茯苓、甘草等4味中药处于核心地位。这4味中药在中医临床上常一起配伍成经典方剂——四君子汤。四君子汤为古今补气健脾的基础方,广泛应用于临床中,具有益气养血、健脾和胃的功效^[30]。在临床上,许多肝癌患者被确诊时已到晚期,又经过手术、介入等手段,常

正气虚弱。《景岳全书》曰:“若积聚渐久,元气日衰,此而攻之,则积气本远,攻不易及,胃气切近,先受其伤,愈攻愈虚,则不死与积而死于攻矣。故凡治虚邪者当以缓治,只宜专培脾胃以固其本”^[24]。四君子汤具有调理脾胃、扶正固本的效果,由此可见,临床上注重以培土固本之法改善患者的正气不足之症,就如《时方歌括》曰:“胃气为生人之本,参、术、苓、草从容和缓,补中宫土气,达于上下四旁,而五脏六腑皆以受气,故一切虚证皆以此方为主。若加陈皮,则有行滞进食之效;再加半夏,即有除痰宽胀之功;再加木香、砂仁,则行气之药多于补守,凡肿满痰饮结聚等证,无不数除,此犹人所易知也”^[30]。

此外,半夏、陈皮与处于核心地位的白术、党参、茯苓、甘草等4味中药关联紧密,合用可成六君子汤,共奏健脾和胃、化痰消痞之功;柴胡、白芍共用可疏肝柔肝,二者一散一收,与四君子汤或六君子汤配伍,可奏疏肝理气、健脾和胃之功;半枝莲、白花蛇舌草常配伍应用,在扶正基础上辨证使用可增强抗肿瘤之功;黄芪、当归联用,是源于当归补血汤,有补气生血之效;黄芪、白术联用,是源于玉屏风散,有补气、利水、止汗之功;柴胡、鳖甲联用,有疏肝解郁、软坚散结之效^[26, 30-35]。此外, TYPHC患者会伴有湿热黄疸的症状,故使用茵陈、薏苡仁来发挥其利水除湿功效^[36]。

2.3 基于CMC-DD算法的组方规律分析

对治疗TYPHC中药处方中的高频药物(用药频次 ≥ 10 次),通过CMC-DD算法得到5组药性组合,依次为寒苦肝、温甘脾、温辛脾、寒咸肝、寒辛肝药性组合,详见表3。另外,通过CMC-DD算法对治疗TYPHC的中药处方进行分析,根据聚类中心距离等原理得到8组核心药物组合,详见表4。

表3 基于CMC-DD算法挖掘的治疗TYPHC中药处方的药性组合

Tab 3 Drug property combination of TCM prescription against TYPHC based on CMC-DD algorithm

序号	组合	中药
1	寒苦肝	白花蛇舌草、白茅根、白芍、北沙参、赤芍、川楝子、垂盆草、大黄、丹参、瓜蒌、虎杖、黄芩、金银花、绿萼梅、麦门冬、牡丹皮、桃仁、王不留行、郁金、枳壳、枳实、重楼、半枝莲
2	温甘脾	当归、茯苓、甘草、桂枝、鸡内金、降香、麦芽、山慈菇、神曲、熟地黄、酸枣仁、五味子
3	温辛脾	白术、半边莲、半夏、冰片、陈皮、大腹皮、党参、莪术、干姜、厚朴、黄芪、姜黄、木香、砂仁、山药、吴茱萸、延胡索、薏苡仁
4	寒咸肝	鳖甲、车前子、枸杞子、龟甲、海浮石、金钱草、猕猴桃藤、牡蛎、山茱萸、生地黄、天龙、土鳖、泽泻
5	寒辛肝	柴胡、夏枯草、茵陈蒿

通过查询文献发现,寒苦肝组合主要功效是清热解毒、凉血、消肿、燥湿,温甘脾组合主要功效是健脾和中、补中益气、消食,温辛脾组合主要功效是温中行气、理气止痛,寒咸肝组合主要功效是平肝明目、软坚散结、滋阴益肾,寒辛肝组合主要功效是清热消肿、解毒化痰^[37]。由表3的药性组合可以看出,温性药(如神曲、白术、砂

表4 基于CMC-DD算法挖掘的治疗TYPHC中药处方的核心药物组合

Tab 4 Key drug combination of TCM prescription against TYPHC based on CMC-DD algorithm

序号	药物组合	功效
1	甘草,白术,茯苓,陈皮,白花蛇舌草,金银花,当归,黄芪	补气养血、健脾除湿化痰、清热解毒
2	甘草,白术,茯苓,党参,麦芽,当归,半枝莲,柴胡	补益气血、疏肝解郁、解毒化痰
3	白术,甘草,党参,茯苓,薏苡仁,柴胡,鸡内金,半边莲	化痰解毒消积、理气健脾除湿
4	党参,甘草,白术,柴胡,半夏,陈皮,黄芪,黄芩	清泄邪热、健脾益气、化痰散结
5	白术,当归,黄芪,枸杞子,猕猴桃藤,麦冬,生地,白花蛇舌草	滋阴养血生津、健脾利湿解毒
6	白术,甘草,茯苓,党参,当归,黄芪,半夏,半枝莲	补气养血、健脾利湿、化痰散结
7	白术,甘草,党参,茯苓,柴胡,白芍,鳖甲,莪术	补中益气、化痰散结、柔肝养血
8	柴胡,半枝莲,牡蛎,鳖甲,天龙,枸杞子,白芍,猕猴桃藤	解毒软坚、理气活血

仁)多以温甘脾、温辛脾组合模式以补益气血、理气和中;寒性药(如柴胡、郁金、夏枯草)多以寒苦肝、寒辛肝的组合模式以清热解毒、消肿、燥湿,又以寒咸肝的组合方式以滋补肝肾、软坚散结,体现了“肝脾同治”的思想。另外,将药物的性味归经作为整体研究,可使药性中不同要素之间的联系更加清晰,更符合中医整体观的理念,也为方剂的配伍提供了参考。

由表4的8组核心药物组合可以看出,其基本以健脾理气、补益气血为主,多在四君子汤基础上加减应用,或佐以清热解毒之药,例如半枝莲、金银花;或佐以利水除湿之药,例如薏苡仁;或佐以化痰散结之药,例如半夏、陈皮;或佐以补气血之药,例如黄芪、当归;或佐以活血化瘀、软坚散结之药,例如鳖甲、莪术。其中,核心药物组合8偏于理气活血,方中柴胡舒肝、调气、解毒,鳖甲、牡蛎软坚散结,天龙活血化瘀,白芍、枸杞子养阴柔肝,半枝莲、猕猴桃藤清热解毒;诸药合用,共奏解毒软坚、理气活血之功。此外,核心药物组合5偏重于滋阴养血,一般阴虚证中多采用滋阴之法,而TYPHC属于实证,不应使用滋阴之法,用于此处似乎矛盾,但是一些TYPHC患者会出现痰瘀互结日久、热极伤阴的情况,故以滋阴之法达到祛除“痰”“瘀”实邪的效果^[38]。中医认为津血同源,《灵枢·邪客》曰:“营气者,泌其津液,注之于脉,化以为血”^[39]。以滋阴之法补其津则血有源,助力血液运行,祛其瘀血;同时,痰生于津液,津液不足或有虚火炼津生痰,则可见难化痰湿之候,此时给予滋阴药使津液充沛,水湿即可从体内以汗液、二便等形式缓缓排出,以达到辅助祛湿化痰之效;此外,在治疗中,难免会使用一些燥热伤阴或苦寒伤阴之药,滋阴亦是为防止其耗散阴津^[38]。

3 讨论

本研究收集了中医临床治疗PHC的907张处方,发现其主要涉及痰瘀互结证、正虚瘀结证、肝脾血瘀证、气滞血瘀证等10个证型。在此基础上,通过TCMDP挖掘了治疗TYPHC的处方用药规律:从用药频次上分析临床常用的中药功效类别,并从性味归经、药性组合、核心药对、关联程度和核心药物组合的角度分析探讨了临床

组方规律。结果发现,临床治疗TYPHC常用补虚药、利水渗湿药、清热药、活血化瘀药、理气药、化痰药;药物四气多为寒、温、平,五味多为甘、苦、辛,归经多为脾、肝、胃、肺、心、肾经。其中,温性药多以温甘脾、温辛脾组合模式以补益气血、理气和中;寒性药多以寒苦肝、寒辛肝的组合模式以清热解毒、消肿、燥湿,又以寒咸肝的组合模式以滋补肝肾、软坚散结。白术、党参、茯苓、甘草等4味药处于核心地位,常与多种药物配伍出现,这与TYPHC患者正气虚弱有关,故以四君子汤加减配伍,以达健气补脾之功效,充分体现了中医“见肝之病,知肝传脾,当先实脾”之思想^[30]。

此外,本研究通过研发TCMDP以研究方剂配伍规律,相对于传统的数据挖掘方法,TCMDP可以多层次、全方位地挖掘中医药中隐含的关系,可以实现多个数据挖掘方法并行,大大提高了数据挖掘效率。本文所采用的WD-Get Rules算法、CMC-DD算法均为改进的新算法:WD-Get Rules算法相对于传统的宽度优先算法和深度优先算法,其挖掘效果得以提升;CMC-DD聚类算法能够很好地解决截断距离参数的影响,且无需人为干扰算法聚类中心的选择,提高了聚类效果。

综上所述,临床治疗TYPHC以健补脾胃,补益气血为主,“见肝之病,知肝传脾,当先实脾”“脾胃乃后天之本”“脾胃是气血生化之源”^[26]。由此可见,中医重视培土益元之法由来已久。同时,《读医随笔》中说:“凡脏腑十二经之气化,皆必藉肝胆之气化以鼓舞之,始能调畅而不病”^[40]。肝主疏泄,调畅气机,当肝失疏泄,气机不畅,或会出现痰瘀互结,所以在治疗时也重视疏肝之法。在此基础上,采用理气化痰、清热解毒、活血化瘀等法辨证治疗;此外,为防耗散阴津,在配伍之时,也会适当配以滋阴之药。由于存在剂量记载不明等原因,故本研究未将剂量纳入分析范畴;同时,考虑到文章的篇幅所限,本研究仅探讨了治疗TYPHC的处方配伍规律,对其他证型及其差异化用药规律研究将是进一步探讨的问题。本研究以TYPHC为例,探讨分析了中药处方治疗PHC的组方规律,可为肝癌的临床用药提供一定的参考,从“方-证”角度探讨中药的配伍规律更符合中医辨证论治的思想。本研究所应用的TCMDP平台可以对处方规律进行多层次的挖掘,得到处方中的隐含关系,相对于传统的单一数据挖掘方法,TCMDP能更加全面地分析中药组方规律,具有一定的实用性。相关研究者也可对算法进行优化,为数据挖掘和中医药研究的融合提供科学的技术手段。

参考文献

- [1] 张金坤,王燕燕.原发性肝癌发生机制及其治疗的研究进展[J].中国药房,2015,26(29):4171-4173.
- [2] 朱元章,张贵彪,朱国福.中药复方抗肿瘤机制研究进展[J].中国实验方剂学杂志,2017,23(16):227-234.
- [3] 罗山水,章新友,张春强,等.基于策略模式的中药数据挖掘

- 掘研究与系统设计[J].世界科学技术:中医药现代化, 2015,17(5):929-933.
- [4] 李文林,屠强,彭丽坤,等.基于关联规则分析明清古籍中疫病文献的药-症关系[J].时珍国医国药,2010,21(4):957-959.
- [5] 刘耸峰,吕晓东,庞立健,等.基于聚类分析探究中药复方治疗肺间质纤维化临床用药规律[J].世界中医药,2016,11(7):1370-1373.
- [6] 罗山水.基于策略模式的中医药信息数据挖掘研究与系统设计[D].南昌:江西中医药大学,2015.
- [7] 吴地尧,章新友.基于策略模式的中药数据挖掘系统的设计和应用:英文[J].中华中医药杂志,2019,34(10):4873-4876.
- [8] 张雨农.密度峰值聚类的优化及其应用研究[D].蚌埠:安徽财经大学,2020.
- [9] 冯振华.基于DBSCAN聚类算法的研究与应用[D].无锡:江南大学,2016.
- [10] 李文杰,闫世强,蒋莹,等.自适应确定DBSCAN算法参数的算法研究[J].计算机工程与应用,2019,55(5):1-7、148.
- [11] ZHANG XY. *Data mining and its application in traditional Chinese medicine*[M]. London: ISCI Publishing Ltd., 2020:251-264.
- [12] 吴地尧,章新友,周小玲,等.运用WD-Get Rules算法对治疗肾虚方剂用药规律的挖掘分析[J].中华中医药学刊,2019,37(4):881-885.
- [13] 王荣茂,张丽君,吴建荣.自拟肝复方联合化疗治疗中晚期原发性肝癌疗效分析[J].中国误诊学杂志,2009,9(25):6084-6085.
- [14] ZHONG C, LI HD, LIU DY, et al. Clinical study of hepatectomy combined with Jianpi Huayu therapy for hepatocellular carcinoma[J]. *Asian Pac J Cancer Prev*, 2014, 15(14):5951-5957.
- [15] HOU EC, LU YX. Primary hepatocarcinoma treated by traditional Chinese medicine combined with transcatheter arterial chemoembolization[J]. *Zhongguo Zhong Xi Yi He Za Zhi*, 2009, 29(3):225-227.
- [16] 吴毓英,钱俊华,连建伟.连建伟教授重视调肝论治痰瘀互结证方法初探[J].浙江中医药大学学报,2019,43(7):657-659.
- [17] 孔怡琳,张海波,张玉佩,等.从痰瘀角度探析肝癌的发病与防治思路[J].中国药物经济学,2012(2):354-356.
- [18] 杜松,胡镜清,卢红蓉.痰瘀互结证现代理论研究进展述评[J].中国中医基础医学杂志,2015,21(4):477-482.
- [19] 吴雄志,彭涛.运用伏邪理论治疗肝癌浅析[J].中西医结合肝病杂志,2020,30(4):289-291.
- [20] 李中玉,曲骞,吴煜.吴煜治疗脾虚痰凝型肝癌用药规律研究[J].新中医,2020,52(9):15-18.
- [21] 薛姣,王雄文,卢燊,等.基于肝-肠轴探讨“实脾法”治疗肝癌的科学内涵[J].中医杂志,2019,60(7):551-555.
- [22] 欧志穗,刘友章,杨汉彬.肝脾相关理论在肝硬化腹水治疗中的应用[J].时珍国医国药,2011,22(9):2272-2273.
- [23] 喻明,王文萍.癌痛“心”论[J].辽宁中医杂志,2010,37(7):1247-1248.
- [24] 王珊珊.周仲瑛教授从痰瘀热毒辨治肺癌的临床经验及益肺解毒汤的抗肿瘤实验研究[D].南京:南京中医药大学,2017.
- [25] 徐德成,马迎民,范吉平.中医“肾精”的现代医学内涵[J].中医杂志,2017,58(22):1891-1897.
- [26] 陈启亮,唐东昕,龙奉玺.“见肝之病,知肝传脾”在肝癌防治中的运用[J].中医学报,2016,31(12):1833-1835.
- [27] 贾文霞,赵红兵,陈桂杰,等.肝癌的中医用药方法[J].中国中医基础医学杂志,2017,23(1):121-123.
- [28] 洪荣健,奚胜艳,王彦晖.基于《丹溪心法·痰》探讨肝癌发病与痰的相关性[J].中华中医药学刊,2014,32(1):164-166.
- [29] 纪文璐,王海龙,苏贵斌,等.基于关联规则算法的推荐方法研究综述[J].计算机工程与应用,2020,56(22):33-41.
- [30] 张广唱,武哲丽,李杰斌.四君子汤治疗肝癌的临床研究进展[J].辽宁中医杂志,2015,42(1):218-220.
- [31] 郭丹丹,黄秋思,张金良,等.张金良治疗肝癌的临床经验总结[J].中医药导报,2019,25(21):102-104.
- [32] 秦桂超.焦中华教授治疗恶性肿瘤遣方用药规律数据挖掘研究[D].济南:山东中医药大学,2015.
- [33] 陈赛里.基于数据挖掘李家庚教授辨治肿瘤临床经验研究[D].武汉:湖北中医药大学,2018.
- [34] 黄月纯,黄水清,刘东辉,等.当归补血汤煎剂中醋酸乙酯部位的主要化学成分分析[J].中国药房,2005,16(16):1275-1277.
- [35] 安海燕,冯思同,林俊豪,等.基于数据挖掘分析鳖甲用药配伍规律[J].中药材,2018,41(8):2021-2025.
- [36] 王利敏,闫军堂,刘晓倩,等.仲景从“痰瘀互阻”辨治黄疸的证治特色与方证解析[J].国医论坛,2018,33(1):4-6.
- [37] 肖斌,王耘,郭维嘉,等.中药药性组合及其与功效的关系研究[J].世界科学技术:中医药现代化,2010,12(6):902-908.
- [38] 田心,高安,张军茹.滋阴药在痰瘀互结病证中的应用[J].中国中医基础医学杂志,2019,25(12):1737-1738.
- [39] 张晨阳,王维,陈文慧.基于气机失调初步探讨非酒精性脂肪肝的病机[J].时珍国医国药,2018,29(5):1148-1150.
- [40] 刘亚琪,胡作为,郑承红.基于“怪病从痰论治”谈肝癌辨治思维[J].中医肿瘤学杂志,2020,2(1):24-27.

(收稿日期:2020-08-11 修回日期:2020-11-20)

(编辑:段思怡)