

基于 AHP-SOM 聚类-TOPSIS 的白及保胃护肝保健食品配方设计与评价

马嘉慕, 刘晓云, 任雪阳, 王宇, 董英, 宋若兰, 于啊香, 魏静, 范琦琦, 折改梅*

北京中医药大学中药学院, 北京 100029

摘要:目的 以白及 *Bletillae Rhizoma* 为例, 通过关联、决策和聚类机器学习算法的联合使用进行数据挖掘研究, 设计具有保胃、护肝双功能的保健食品配方并对其进行评价。方法 对现有具有保护胃黏膜或保护肝功能的保健食品批文, 以及可用于治疗这 2 种疾病的中成药和方剂数据库中信息进行整理, 先后使用关联规则算法 (Apriori)、层次分析法 (analytic hierarchy process, AHP)、SOM (self-organizing map) 聚类、理想解法 (technique for order preference by similarity to solution, TOPSIS) 等关联、决策和聚类机器学习算法, 挖掘其中高频原料药味的配方规律, 结合现代功能活性及营养学研究成果, 组成以白及为核心原料的具有配伍思想的保胃、护肝双功能保健食品配方并进行评价。结果 对各数据库进行原料药味频次统计和相应高频原料药味的关联性分析, 发现高频原料药味之间极易产生较强的关联性。继而对数据库中出现的 64 个可用于保健食品的高频原料药进行 2 个层次共 17 个指标的 AHP 分析和加权后, 得到除白及外加权值排在前 5 位的原料药味为甘草、陈皮、黄芪、茯苓和五味子, 这些原料药可以考虑在组方时进行优先选择。SOM 聚类显示所有高频原料可分为 7 类, 其中在最优选配方原料药味与 AHP 分析结果权重排名前列的原料药味有极高的重叠。对原料进行保胃、护肝功能方面的文献检索, 同时结合之前的数据分析结果, 确定了 15 个原料进行君、臣、佐、使的配伍组合, 最终设计了 10 个可能配方; 并对其进行 TOPSIS 分析评价, 其中排名前 2 位的配方黄芪、白及、五味子和甘草及黄芪、白及、陈皮和五味子的分值相近, 均大于 0.14, 与其他配方评分相差较大。结论 在中医药基本理论的指导下结合各类机器学习算法, 以核心药味白及为例建立了保胃、护肝双功能保健食品的配方设计和评价方法, 为以后保健食品配方研发提供新的思路 and 方向。

关键词: 保健食品; 白及; 配方设计; AHP; SOM 聚类; TOPSIS; 数据挖掘

中图分类号: R285.64

文献标志码: A

文章编号: 0253-2670(2021)18-5676-12

DOI: 10.7501/j.issn.0253-2670.2021.18.022

Formula design and evaluation of health food with protection of gastric mucosa and liver of *Bletillae Rhizoma* based on AHP-SOM-TOPSIS algorithm

MA Jia-mu, LIU Xiao-yun, REN Xue-yang, WANG Yu, DONG Ying, SONG Ruo-lan, YU A-xiang, WEI Jing, FAN Qi-qi, SHE Gai-mei

School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 100029, China

Abstract: Objective This study takes Baiji (*Bletillae Rhizoma*) as an example to design a health food formula with the protection of gastric mucosa and liver, furthermore develop an evaluation method. Some machine learning algorithms are taken into use in this study either, like association, decision, and clustering analysis. **Methods** Four databases from health food approval, Chinese patent medicines, and traditional Chinese medicine prescriptions with functions mentioned above were organized in order. Some machine learning algorithm, such as Apriori association rules algorithm, analytic hierarchy process (AHP), self-organizing map (SOM), and technique for order preference by similarity to solution (TOPSIS) was used in turn to find the regular relation among high-frequency ingredients which rooting from the databases. Combined with the function and nutrition research of all high-frequency ingredients, health food formula with the protection of gastric mucosa and liver, which full of thoughts of compatibility were designed and graded. **Results** Through the statistics of the frequency of all functional ingredients showed up and the correlation analysis of the corresponding high-frequency function ingredients, it was found that the high-frequency function ingredients were easy to have a stronger correlation with each other. Then 64 high-frequency functional ingredients from four databases above were analyzed by the AHP algorithm, including 17 indicators of two levels. The weight value showed

收稿日期: 2021-04-12

作者简介: 马嘉慕 (1996—), 女, 硕士研究生, 研究方向为中药化学。E-mail: majiamu96@163.com

*通信作者: 折改梅 (1976—), 博士生导师, 研究员, 研究方向为中 (民族) 药药效成分和新药创制研究。

Tel: (010)53912129 E-mail: shegaimei@126.com

that Gancao (*Glycyrrhizae Radix et Rhizoma*), Chenpi (*Citri Reticulatae Pericarpium*), Huangqi (*Astragali Radix*), Fuling (*Poria*), and Wuweizi (*Schisandrae Chinensis Fructus*) were in the top rank, which could be favorable evidence for the final formula. All functional ingredients were divided into seven groups in SOM clustering. The results indicated that the best choices group had a great common with the AHP top rank. Through literature search on the functions of protecting stomach and liver of ingredients, 15 functional ingredients were determined to be compatible with king, minister, adjuvant and enthal, combined with the previous data analysis results. Ten possible formula was comprised by 15 functional ingredients, with using TOPSIS analysis to provide the basis of the final formula. The grades showed in tie between formula composed with *Astragali Radix*, *Bletillae Rhizoma*, *Schisandrae Chinensis Fructus*, *Glycyrrhizae Radix et Rhizoma* and formula composed with *Astragali Radix*, *Bletillae Rhizoma*, *Schisandrae Chinensis Fructus*, *Citri Reticulatae Pericarpium*, which grades were all above 0.14 and have obvious difference with other formula. **Conclusion** Under the guidance of the basic theories of traditional Chinese medicine and combined with various machine learning algorithms, this study took the core medicine *Bletillae Rhizoma* as an example to establish the formula design and evaluation method of dual-function health food protecting stomach and liver, also provides a new idea and direction for the research and development of Chinese medicine prescription health food in the future.

Key words: health food; *Bletillae Rhizoma*; formula design; AHP; SOM cluster; TOPSIS; data mining

保健食品中的辅助保护胃黏膜和辅助保护化学性肝损伤 2 项功能即保胃护肝, 分别可以对应中医学理论中的“胃脘痛”和“胁痛”, 是胃、肝处于亚健康时期的一种辅助治疗手段。胃、肝病的同时发生、治疗在中医药典籍及现代医学研究均有论述。医圣张仲景曾提出:“见肝之病, 知肝传脾, 当先实脾。”清代名医叶天士也提到:“肝为起病之源, 胃为传病之所。”说明胃、肝病具有多因素作用及可同时发生的特点。胃、肝病的致病因素中精神因素往往起着先导作用, 中医常将胃、肝病的情志因素归于肝、胃与脾的关系中, 故多疏肝补脾健胃。酒精对于人类健康而言已经成为最大的威胁因素之一, 饮酒人群长期面临着饮酒引起的相关疾病威胁^[1]。酒精的摄入可能会引起诸如心脑血管疾病、中枢神经系统的紊乱, 以及胃黏膜损伤、肝损伤等消化系统疾病等多种不良健康结局同时发生^[2]。在这些疾病中 80% 以上是慢性病, 治愈难度较大。全球范围内的成年人人均酒精消费量在 1990—2017 年已从 5.9 L/年上升至 6.5 L/年, 预计到 2030 年将达到 7.6 L/年。酒精的消费量巨大, 消费人口众多。因此, 在发病之初或未发病的亚健康及健康状态通过服用保健食品、增强锻炼提高自身免疫力等外界辅助措施来控制相关疾病的发生、发展态势是极为重要的^[3-4]。

白及是兰科植物白及 *Bletilla striata* (Thunb.) Reichb. f. 的干燥块茎, 具有收敛止血、消肿生肌的功效^[5], 自古便是白及糖等养生保健药膳的主要原料, 临床也多用于治疗各类消化道溃疡, 也有治疗化学性肝损伤的报道^[6-9]。白及的产品形式主要集中于保护消化道溃疡及止血的药品、美白类的化妆品和生物材料等, 白及同时也可作为保健食品的原料^[10], 在

改善消化系统病症方面有其独特的优势。但以白及为主要原料的保健食品很少, 具有极大的研发潜力。

本研究以白及作为一种具有保胃护肝双功能的保健食品的主要原料, 采用包括 Apriori、层次分析法 (analytic hierarchy process, AHP)、SOM (self-organizing map) 聚类、理想解法 (technique for order preference by similarity to solution, TOPSIS) 在内的多种关联、决策分析 (decision analysis) 和聚类分析 (cluster analysis) 的机器学习算法对注册的保胃和护肝的保健食品批文信息、方剂信息和中成药信息等现有数据进行挖掘, 探究白及保护化学性肝损伤的新功能, 并为以白及为核心药味的具有保护胃黏膜和化学性肝损伤双功能的保健食品开发提供依据。为后续类似双功能保健食品的开发思路提供参考, 为保健食品配方的筛选评价提供方法。

1 方法

1.1 数据来源及纳入标准

药智数据 (<https://db.yaozh.com/>) 是国内最大的医药信息收集网站。含有白及的方剂和中成药数据均来源于药智数据的中药方剂数据库和中成药处方数据库。保健食品信息分别以“胃黏膜”和“肝损伤”为关键词, 在国家市场监督管理总局特殊食品安全监督管理局 (<http://www.samr.gov.cn/tssps/>) 官网上进行检索, 收集截至 2020 年 11 月 30 日所收录的所有保健食品批文信息。

包含白及的方剂在收集整理过程中, 排除部分中药组成、功效或主治相同以及仅含白及单味药的方剂之后, 共收集到 241 首方剂, 其中具有补脾益气功效的有 31 首。包含白及的中成药共收集到 121 个处方, 其中可用于治疗脾气虚弱的成药共 44 个。

保健食品的信息仅纳入原料中包含 2 味及以上中药的中药类保健食品,同时排除没有配方信息的批文,共收集到对胃黏膜有辅助保护作用的批文 46 条,对化学性肝损伤有辅助保护功能的批文 302 条。

1.2 数据库的建立与标准化

1.2.1 原料数据库的建立 将收集得到的数据信息分别导入 Microsoft Excel 2019 建立 4 个数据库。在数据处理过程中参照《中国药典》2020 年版^[5]、《临床中药学》第 2 版^[11]和地方标准对方剂、中成药和保健食品批文的配方药味信息进行标准化处理:(1)将习用名改为药典或地方标准中记载规范名称,如“白芨”统一为“白及”,“黄耆”统一为“黄芪”等。(2)将同一味药的不同炮制规格改为原药名,如“炒白术”统一为“白术”,“蜂胶粉”统一为“蜂胶”等。(3)将药材提取物改为原药材名,如“葛根提取物”统一为“葛根”,“广藿香油”统一为“广藿香”等。(4)合并不同方剂名、中成药名、不同批文,但是药味组成一致且功效一致的数据;另外,在整理中药类保健食品的过程中,将相同药味、不同辅料的批文合并处理。(5)规范化过程中若有改变名称后功效影响显著者,则保留原有记录,分别录入处理,如“梔子炭”和“梔子”,“生地黄”和“熟地黄”等。

1.2.2 配方数据库的建立 用于 AHP 分析的层级和指标分别按照原料数据库中内容和统计项进行录入。参与统计的原料药味包括在具有辅助保护胃黏膜和辅助保护化学性肝损伤的保健食品中及具有补脾益气功能的中成药、方剂中的全部高频药味。一级评价指标为药味类别和现有处方,二级评价指标为现有处方类型,包括具有辅助保护胃黏膜功能的保健食品、具有辅助保护化学性肝损伤的保健食品、包含白及的具有补脾益气功能的中成药和包含白及的具有补脾益气功能的方剂共 4 个。通过统计高频药味在各库中的频次及出现在 4 个库中的总数等信息,按照降序排列形成统计表并赋值 1~9 分。

根据以白及为配方核心的具有保胃护肝双功能保健食品的研发总目标和传统中医理论治疗胃病和肝病补脾益气的核心理论,确定各层级内要素之间的相对重要关系,构建两两比较矩阵,相邻 2 个元素相比的重要性可用 1~9 分赋值。其中 1 表示同等重要,3 表示稍微重要,5 表示明显重要,7 表示强烈重要,9 表示极端重要。2、4、6、8 为上述相邻判断的中值^[12]。

根据 AHP 分析后得到的权重对高频药味进一步分析,按照线性加权公式对各药进行加权评分,并以其最后得分进行排序^[13]。SOM 聚类应用于药味类别和出现数据库种类 2 项分别加权得分后的原料药,可得到在保胃护肝保健食品中这些高频药味的分组,采用聚类方法对评价结果进行分组,其分组结果更具科学性和客观性,较好解决了由于数据微小差别而导致分组界限不清晰的问题。在 TOPSIS 分析中数据依据君、臣、佐、使的顺序来进行排列分析,对拟定配方进行评价以供最终选择。

1.3 分析方法

1.3.1 原料筛选 利用 Microsoft Excel 2019 对数据库中的数据信息进行功效和药味的频数统计并进行分析。通过 SPSS Modeler 18.0 中的 Apriori 关联规则算法和网络图对整理得到的数据进行挖掘与统计分析,得到方剂、中成药以及具有辅助保护胃黏膜和辅助保护化学性肝损伤功能保健食品中的药味之间的二次关联度和高频药对组合。

1.3.2 配方分析 采用 AHP 确定评价体系指标权重,通过 SOM 聚类算法对各药进行分析并通过传统中医理论和现代医学研究来筛选得到可能的配方,最后利用 TOPSIS 法来综合评定各配方优劣,最终得到可与白及配伍使用的原料配伍组合。

本研究的分析及流程如图 1 所示。

2 结果

2.1 保健食品的数据分析

2.1.1 药味频数分析 在 46 条具有辅助保护胃黏膜的中药类保健食品批文中,共涉及到原料药味 70 种,出现频次 ≥ 3 次的共有 28 种原料药味,占总频次的 75.78%。除出现 8 次的白及外,排名前 5 位的原料药味分别为砂仁、蜂胶、白术、黄芪和陈皮,具体情况见表 1。对 70 种原料药物按照类别分类的发现,出现频次 5 次及以上的药味类别分别为补虚药(24 次)、理气药(7 次)、清热药(7 次)、消食药(5 次),共占 61.43%。补虚药和理气药等类别原料药物的使用与中医在治疗胃黏膜损伤时需要补脾益气的治疗原则是一致的^[14]。

在 302 条具有辅助保护化学性肝损伤的中药类保健食品批文中,涉及到 150 种原料药味,其中出现频次 ≥ 10 次的共有 30 种药味,这 30 种原料在所有原料药味出现频次中占比 74.87%,其中频次出现排名前 5 位的分别是葛根、五味子、枸杞子、灵芝和枳椇子。具体信息见表 2。通过对原料的类别进

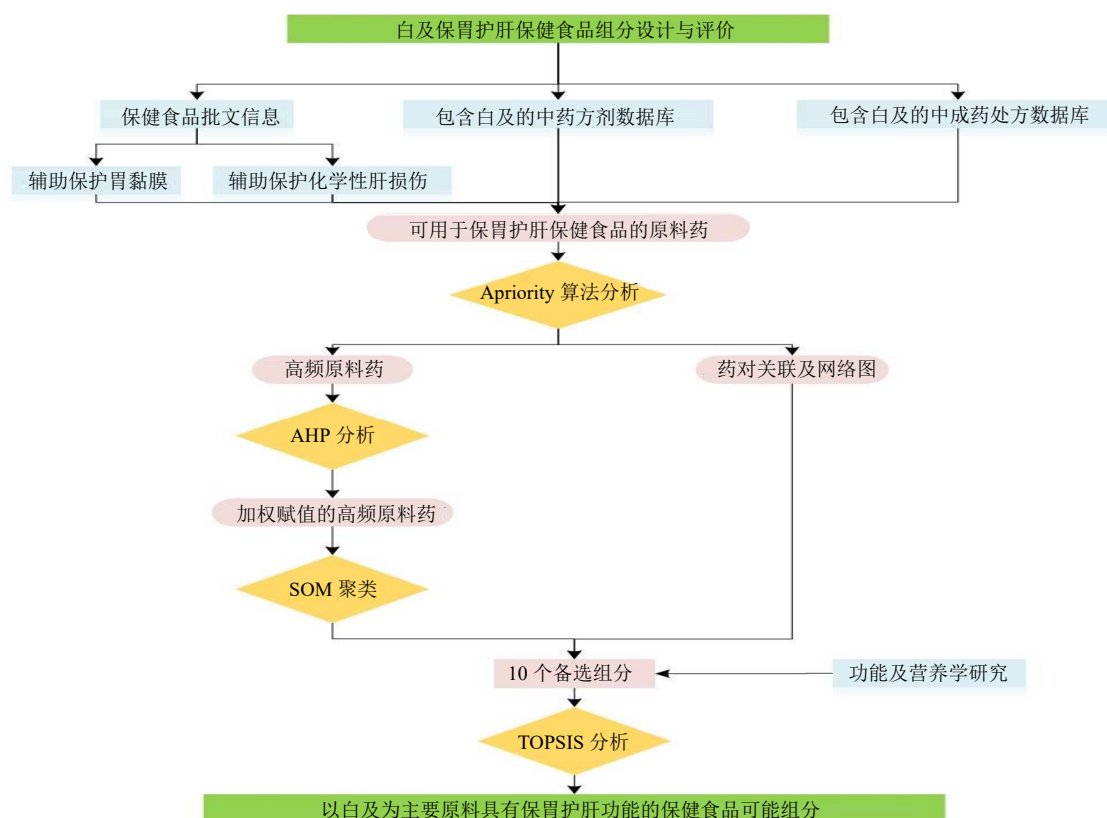


图 1 白及保胃护肝保健食品配方设计分析方法及流程

Fig. 1 Analysis method and process of health food formula with protection of gastric mucosa injury and liver injury of *Bletillae Rhizoma*

表 1 具有辅助保护胃黏膜的保健食品中原料药物频次

Table 1 Frequency of functional ingredients in health food with protection of gastric mucosa injury

序号	名称	频次	序号	名称	频次
1	砂仁	13	15	山药	6
2	蜂胶	11	16	猴头菇	6
3	白术	9	17	白芷	5
4	黄芪	9	18	麦芽	5
5	陈皮	8	19	佛手	4
6	白及	8	20	甘草	4
7	蜂蜡	8	21	木香	4
8	茯苓	8	22	薏苡仁	4
9	白芍	7	23	大枣	3
10	三七	7	24	枸杞子	3
11	吴茱萸	7	25	广藿香	3
12	葛根	6	26	牡丹皮	3
13	党参	6	27	人参	3
14	蒲公英	6	28	山楂	3

表 2 具有辅助保护化学性肝损伤的保健食品中药物频次

Table 2 Frequency of functional ingredients in health food with protection of liver injury

序号	名称	频次	序号	名称	频次
1	葛根	163	16	陈皮	21
2	五味子	84	17	人参	21
3	枸杞子	64	18	菊花	20
4	灵芝	60	19	姜黄	18
5	枳椇子	58	20	西洋参	18
6	丹参	57	21	桑葚	16
7	茯苓	40	22	当归	15
8	甘草	40	23	余甘子	15
9	绞股蓝	38	24	蜂胶	14
10	黄芪	36	25	决明子	14
11	山楂	34	26	银杏叶	14
12	栀子	27	27	大枣	13
13	白芍	25	28	灵芝孢子粉	12
14	泽泻	25	29	蒲公英	12
15	三七	23	30	红景天	10

行统计,发现出现频次前 5 位的类比分别为补虚药(57 次)、清热药(18 次)、理气药(16 次)、利水渗湿药(10 次)和解表药(8 次),共占 72.67%。对于化学性肝损伤,因亚健康状态主要为“胁痛”范畴,多属实热证,易聚湿生痰,因此治疗上多祛湿理气、健脾化痰^[15-16],宜用补虚、理气等类别的药味。保健食品中高频次药物类别的使用契合了这种治则。但是这种根据药味类别分类选择配方原料的方法依然存在不尽完善的地方,例如,在频次统计中排第 2 名的五味子虽为收涩药,其性也与补虚、理气等不尽相同,但是在人参五味子汤等经典方剂中,五味子也可以与其他药物进行配伍发挥补脾益气的功效。故在配方原料的选择上还应结合配伍等其他的因素综合考量^[17]。

综合二者高频药味,出现频次较高的相同药味

包括:蜂胶、黄芪、陈皮、茯苓、白芍、三七、葛根、蒲公英、甘草、大枣、人参、山楂共 12 味原料药。说明在辅助保护胃黏膜和辅助保护化学性肝损伤的原料药味存在较高的重合,且这 12 味药在保护胃黏膜和化学性肝损伤中均可以起到显著的作用。

2.1.2 药味关联度及网络分析 支持度表示在所有数据中关联药味出现的比例;置信度表示关联药味之间联系的强度;提升度主要是用来衡量关联规则的前项与后项间的独立性,如果提升度>1,则可以认为前项和后项之间有良好的相关性,相关性的强弱与数字大小呈正相关^[18-20]。分别对具有辅助保护胃黏膜和辅助保护化学性肝损伤的高频药味进行 Apriori 算法分析和网络图分析,以支持度≥10%、置信度≥50%作为筛选条件,得到高频药味组合和网络图,结果如表 3 和图 2 所示。

表 3 基于 Apriori 算法的保健食品批文中高频药味间的二次关联分析

Table 3 Quadratic combination of high-frequency functional ingredients in health food based on Apriori algorithm

功能	原料中药组合(前项→后项)	频次	支持度/%	置信度/%	提升度
辅助保护胃黏膜	砂仁→蜂胶	11	23.91	54.55	1.93
	蜂胶→蜂蜡	8	17.39	62.50	2.61
	茯苓→陈皮	8	17.39	50.00	2.88
	陈皮→茯苓	8	17.39	50.00	2.88
	白芷→吴茱萸	7	15.22	57.14	5.26
	砂仁→白芍	7	15.22	57.14	2.02
	陈皮→党参	6	13.04	50.00	2.88
	吴茱萸→党参	6	13.04	50.00	3.29
	茯苓→党参	6	13.04	50.00	2.88
	蜂胶→猴头菌	6	13.04	50.00	2.09
	砂仁→猴头菌	6	13.04	50.00	1.77
	麦芽→山药	6	13.04	50.00	4.60
	黄芪→蒲公英	6	13.04	50.00	2.56
	吴茱萸→白芷	5	10.87	80.00	5.26
	黄芪→白芷	5	10.87	60.00	3.07
	山药→麦芽	5	10.87	60.00	4.60
	白术→麦芽	5	10.87	60.00	3.07
辅助保护化学性肝损伤	葛根→五味子	84	27.81	59.52	1.10
	葛根→枳椇子	58	19.21	86.21	1.60
	葛根→丹参	57	18.87	54.39	1.01
	葛根→甘草	40	13.25	62.50	1.16
	葛根→茯苓	40	13.25	52.50	0.97
	葛根→绞股蓝	38	12.58	60.53	1.12
	五味子→丹参、葛根	31	10.26	58.06	2.09

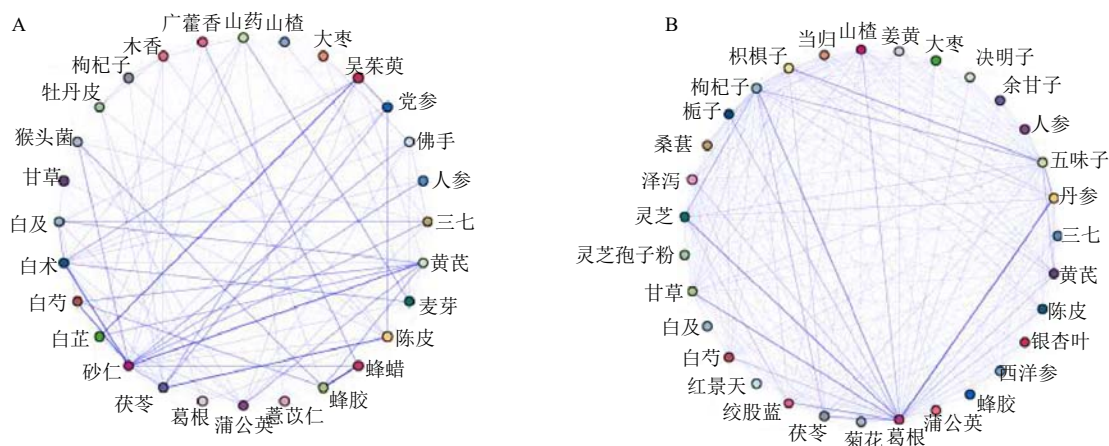


图 2 辅助保护胃黏膜 (A) 和辅助保护化学性肝损伤 (B) 的保健食品中高频药味网络分析图
Fig. 2 Network of high-frequency functional ingredients in health food with protection of gastric mucosa injury (A) and liver injury (B)

通过表 3 可以看出, 在辅助保护胃黏膜和辅助保护化学性肝损伤的原料中药组合中, 关联性最强的分别是白芷和吴茱萸以及五味子和丹参、葛根的配伍。

为进一步揭示高频药味之间的关系, 进行网络图分析。可以看出药物之间的关联强度基本符合二次关联分析的结果。其中在辅助保护胃黏膜功能中, 与白芷显示较强关联的药味有三七、黄芪、蒲公英、甘草、白术等。在辅助保护化学性肝损伤的功能中, 与白芷显示关联的药味主要有白芍和甘草等。

2.2 方剂的数据分析

2.2.1 药味频数分析 对收集得到的方剂药味进行统计分析, 除白及外共有 114 种药味, 累计频次为 238 次。出现频次 ≥ 3 次属于保健食品原料目录内的药味 (除白及外) 统计情况如表 4 所示。通过对出现药味进行类别统计, 出现频率前 5 种类别分别为补虚药 (21 次)、清热药 (17 次)、活血化瘀药 (12 次)、化痰药 (8 次) 和解表药 (6 次), 占比 56.64%。与保健食品的原料类别相比, 方剂中还额外使用了活血化瘀药和化痰药。这类药物的使用可能与在治疗脾气虚弱等病症时存在着活血化瘀和祛湿化痰等疗法有关^[14]。

2.2.2 药味关联度分析 通过对表 4 中高频药味进行二次关联分析得到, 药味间相关性最强的是白术和川芎 (支持度为 12.90%、置信度 75%、提升度为 7.75)。方剂中高频药味的网络分析如图 3 所示, 显示有很多药物之间有强关联, 其中黄芪和甘草之间的关联也比较紧密。

表 4 包含白及的具有补脾益气功效的方剂中高频药味频次
Table 4 High-frequency functional ingredients in prescriptions with effects of tonifying spleen and replenishing qi containing *Bletillae Rhizoma*

序号	名称	频次	序号	名称	频次
1	甘草	14	13	橘红	4
2	当归	9	14	人参	4
3	白芷	8	15	肉桂	4
4	贝母	8	16	白术	3
5	黄芪	8	17	百合	3
6	薏苡仁	8	18	红花	3
7	桔梗	7	19	麦冬	3
8	阿胶	4	20	木香	3
9	陈皮	4	21	生地	3
10	川芎	4	22	吴茱萸	3
11	茯苓	4	23	五味子	3
12	金银花	4			

2.3 中成药的数据分析

2.3.1 药味频数分析 统计包含白及的具有补脾益气功效的中成药中原料频次 (表 5), 并筛选不在保健食品原料目录中的原料药味后, 发现排名前 5 位原料的为甘草 (26 次)、黄芪 (17 次)、白芍 (16 次)、三七 (11 次) 和砂仁 (11 次)。这与保胃护肝功能的保健食品中高频药味有很高的重叠性。

对药味类别进行统计分析, 排在前 5 位的分别是补虚药 (23 次)、清热药 (21 次)、理气药 (12

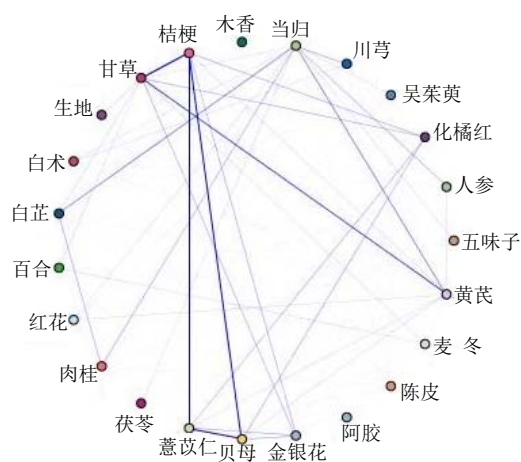


图 3 包含白及的益气健脾方剂中高频次药味网络分析图
Fig. 3 Network of high-frequency functional ingredients in prescriptions with effects of tonifying spleen and replenishing *qi* containing *Bletillae Rhizoma*

表5 包含白及的具有补脾益气功效的中成药高频药味频次
Table 5 High-frequency functional ingredients in Chinese patent drugs with effects of tonifying spleen and replenishing *qi* containing *Bletillae Rhizoma*

序号	名称	频次	序号	名称	频次
1	甘草	26	15	豆蔻	5
2	黄芪	17	16	佛手	4
3	白芍	16	17	珍珠	4
4	三七	11	18	枳壳	4
5	砂仁	11	19	白芷	3
6	白术	10	20	川芎	3
7	木香	10	21	丁香	3
8	党参	8	22	高良姜	3
9	茯苓	8	23	青皮	3
10	陈皮	7	24	山药	3
11	丹参	7	25	山楂	3
12	鸡内金	7	26	吴茱萸	3
13	肉桂	7	27	浙贝母	3
14	香附	6			

次)、活血化瘀药(10次)和解表药(8次),占比53.62%。在中成药的使用原料药物类别方面与方剂中结果类似。

2.3.2 药味关联度分析 通过二次关联分析得到的数据, 药物间配伍关联最强的为丁香和豆蔻(支持度 11.11%、置信度 60%、提升度为 9.0)。在网络图(图 4)中显示甘草与黄芪、陈皮相关性良好, 同时黄芪与陈皮的相关性连接线也比较清晰, 关联性较强。

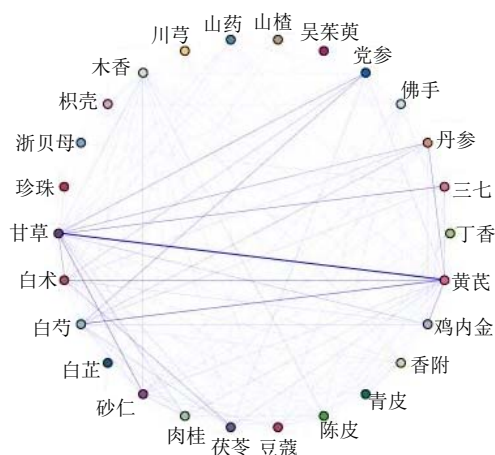


图4 包含白及的益气健脾中成药中高频次药味网络分析图
Fig. 4 Network of high-frequency functional ingredients in Chinese patent drugs with effects of tonifying spleen and replenishing *qi* containing *Bletillae Rhizoma*

2.4 以白及为主要原料具有保胃护肝双功能保健食品配方分析

2.4.1 建模过程 影响保健食品配方选择的因素有很多,包括药物类别和保健食品原料目录等。在选择配方原料时常常把这些因素视为同等重要进行处理,但是由于配方对应功能的不同,因素间的权重也存在着差别,且各因素权重相等也难以判断在该功能下的配方选择中的主导因素。非等权评价方法 AHP 分析是一种可以使权值在定性的基础上得到量化的评价方法。本研究综合现有研究的基础上,根据在保健食品配方设计中原料选择常采用的评价项目建立 AHP 评价指标体系,如表 6 所示。

在 AHP 分析中,选择了多种评分依据来进行打分赋值,以期筛选出更为可靠的评价方式。在二级指标药味类别中采用根据方剂库、中成药库、保健食品库(包括辅助保护胃黏膜损伤和辅助保护化学性肝损伤 2 个库)和 4 个库综合的 4 种评分进行比较。

各高频原料经过 AHP 分析优化权重赋值后得到的数据依然具有一定的人为主观性,且数据结构模糊,除了根据结果进行排序之外,更多的数据信息有待挖掘。SOM 聚类是一种适用于处理不精确的模糊信息的自组织竞争神经网络,对非线性问题求解能力比较强,可以更好地反映数据间的规律。本研究中用于分析各原料药的数据来自于通过 AHP 分析之后经过线性加权、归一化处理的二维数据点,最终得到的聚类结果供后续评价使用。

TOPSIS 是通过判断评价对象与正、负理想解

表 6 AHP 评价指标体系
Table 6 Evaluation index system of AHP

一级指标	二级指标
药物类别 (A1)	安神药 (B1)
	补虚药 (B2)
	化湿药 (B3)
	化痰药 (B4)
	活血化瘀药 (B5)
	解表药 (B6)
	理气药 (B7)
	利水渗湿药 (B8)
	清热药 (B9)
	收涩药 (B10)
	温里药 (B11)
	消食药 (B12)
	止血药 (B13)
现有处方类型 (A2)	辅助保护胃黏膜保健食品 (B14)
	辅助保护化学性肝损伤 (B15)
	包含白及的补脾益气中成药 (B16)
	包含白及的补脾益气方剂 (B17)

的聚类对方案进行评价。若方案最靠近正理想解同时最远离负理想解则为最佳方案，并以此为依据对方案进行优劣排序，是一种关注指标数据本身特性并评价较客观的综合评价方法。但是 TOPSIS 分析无法确定每项评价要素的权重，在应用时需要预先给定权重，所以本研究中 TOPSIS 分析数据来源于经过 AHP 分析赋值和 SOM 聚类后的备选配方中药物的权重和组合权重。

2.4.2 AHP 分析结果 λ 为 AHP 分析结果的最大特征根。CI 值是在 AHP 分析中用于表示权重计算矩阵一致性的一个参数，CR 值是一致性比例，一般认为 CR 值小于 0.1 时矩阵的一致性可以接受^[21]。这 2 个指标值越小，说明矩阵的一致性越好，原始评价赋值越合适。4 库综合评分和 C 层次的原料药的评分中以出现总频次乘以出现库的数目表示。药味类别下不同评价方法中的最大特征根、CI 和 CR 值如表 7 所示。

通过对 64 味高频原料的 AHP 分析后，依据 2 个二级评价指标和经过线性加权之后的总权重进行排序，结果如表 8 所示。

表 7 药味类别下不同 AHP 赋值评价方法结果
Table 7 Results of different evaluation method of AHP in drug category

评价方法	最大特征根	CI 值	CR 值
药味类别-4 库综合排序	13.001 4	$1.140\ 8 \times 10^{-4}$	$7.312\ 7 \times 10^{-5}$
药味类别-中成药频次排序	13.012 3	$1.000\ 0 \times 10^{-3}$	$6.549\ 9 \times 10^{-4}$
药味类别-方剂频次排序	13.015 6	$1.300\ 0 \times 10^{-3}$	$8.325\ 0 \times 10^{-4}$
药味类别-保健食品综合频次排序	13.004 4	$3.647\ 9 \times 10^{-4}$	$2.338\ 4 \times 10^{-4}$

结果显示，综合现有保健食品批文、中成药和方剂数据，在保胃护肝方面具有优先选择可能性的主要原料药，即总权重大于 0 的原料药共有 20 个。其中除白及外排名前 5 的分别为甘草、陈皮、黄芪、茯苓和五味子。这些药味的权重排序为之后配方的建立提供了依据。

2.4.3 SOM 分析结果 将经过 AHP 赋值加权后的 64 味高频药味的 2 个二级指标进行 SOM 聚类，可以得到 7 类结果，分类情况如表 9 所示。通过神经元之间的关系图也可以进一步验证分类结果，如图 5 所示。图 5-A 中灰色的表示神经元节点，红线表示神经元之间有连接，中间的颜色块表示神经元权重向量的远近程度，从黄色到黑色，颜色越深表明越远。图 5-C、D 中表示每个输入向量和竞争层神

经元之间的权值连接情况，权值最小的颜色块为蓝色，权值为 0 的颜色块为黑色，权值最大的颜色块为红色。

2.4.4 TOPSIS 分析结果 通过对每味原料在现在医学研究中对辅助保护胃黏膜和辅助保护化学性肝损伤功能的统计分析，并根据 AHP 分析并线性加权后的总权重排序和 SOM 聚类结果，选择白及、甘草、陈皮、黄芪、茯苓、五味子、丹参、白芍、贝母、白术、党参、人参、木香、山药和当归共 15 种不同药物类别的原料进行配方。依据传统中医理论中补脾益气的治疗原则和配伍理论等，共形成 10 个配方，经过 TOPSIS 分析之后按照打分高低进行排序，结果如表 10 所示。

打分结果显示 10 个配方的差别较大，但排名前

表 8 高频原料药 AHP 分析总权重结果及排名

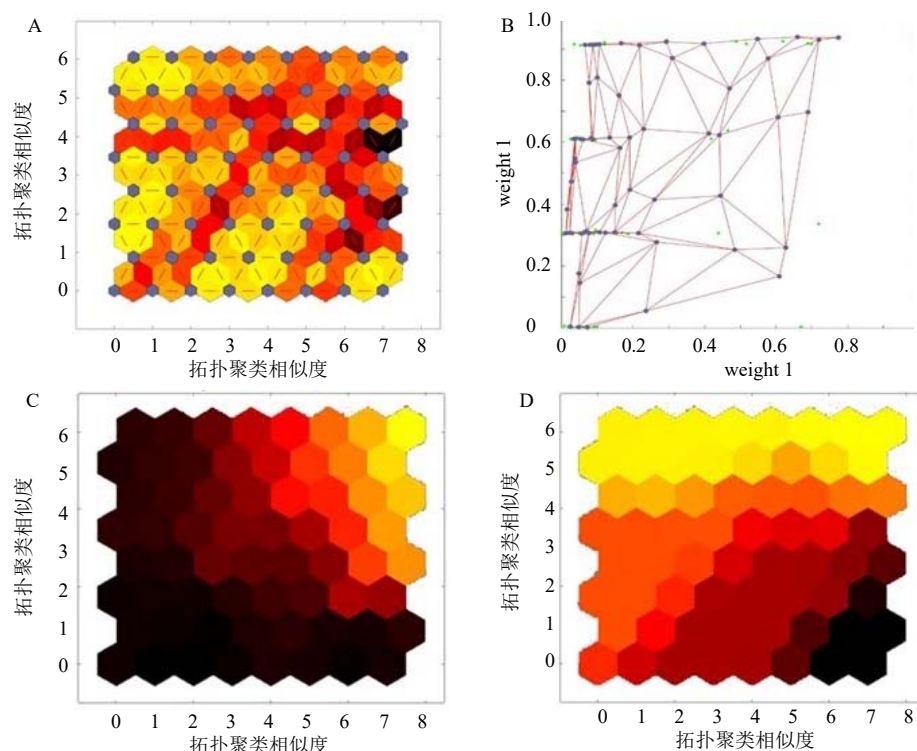
Table 8 Weight value and ranking of high-frequency functional ingredients with AHP analysis

总权重排名	药名	总权重	总权重排名	药名	总权重
1	白及	0.395 70	33	鸡内金	-0.947 65
2	甘草	0.209 35	34	西洋参	-0.947 90
3	陈皮	0.157 28	35	葛根	-1.812 80
4	黄芪	0.147 12	36	蜂胶	-1.921 76
5	茯苓	0.143 20	37	珍珠	-1.929 09
6	五味子	0.138 10	38	白芷	-1.935 07
7	丹参	0.129 84	39	枸杞子	-1.943 63
8	白芍	0.122 20	40	枳椇子	-1.944 51
9	贝母	0.121 93	41	灵芝	-1.947 01
10	白术	0.111 12	42	川芎	-1.947 41
11	党参	0.091 79	43	蜂蜡	-1.947 50
12	人参	0.089 72	44	阿胶	-1.949 39
13	木香	0.089 36	45	猴头菇	-1.951 17
14	山药	0.087 77	46	牡丹皮	-1.955 24
15	当归	0.087 64	47	豆蔻	-1.957 33
16	薏苡仁	0.086 19	48	麦芽	-1.960 73
17	大枣	0.076 62	49	青皮	-1.961 59
18	百合	0.074 94	50	高良姜	-1.961 83
19	泽泻	0.074 21	51	余甘子	-1.963 08
20	丁香	0.072 93	52	红景天	-1.964 27
21	砂仁	-0.819 92	53	广藿香	-1.964 40
22	山楂	-0.912 73	54	银杏叶	-1.965 55
23	吴茱萸	-0.914 43	55	生地	-2.928 17
24	肉桂	-0.921 86	56	桔梗	-2.962 18
25	三七	-0.922 04	57	橘红	-2.967 01
26	佛手	-0.935 65	58	绞股蓝	-2.968 15
27	金银花	-0.936 93	59	栀子	-2.972 13
28	蒲公英	-0.937 82	60	菊花	-2.975 22
29	麦冬	-0.938 31	61	桑葚	-2.976 61
30	香附	-0.938 92	62	灵芝孢子粉	-2.976 61
31	枳壳	-0.941 56	63	决明子	-2.978 65
32	红花	-0.943 40	64	姜黄	-2.978 80

表 9 高频原料药味 SOM 聚类分析结果

Table 9 Clustering results of high-frequency functional ingredients with SOM analysis

类别	涵盖原料药味名称
1 (最优选)	甘草、白及、陈皮、茯苓、贝母、黄芪、白芍
2 (优选)	党参、山药、当归、大枣、木香、百合、五味子、丹参、白术、人参、丁香
3 (较优推荐)	薏苡仁、泽泻、砂仁、山楂、葛根
4 (一般推荐)	吴茱萸、肉桂、金银花、西洋参、麦冬、生地、珍珠、牡丹皮
5 (较不推荐)	枸杞子、枳椇子、灵芝、蜂胶、白芷
6 (不推荐使用)	三七、佛手、香附、枳壳、蒲公英、红花、绞股蓝、栀子、桑葚、灵芝孢子粉、猴头菇、余甘子、红景天、阿胶、蜂蜡
7 (不建议使用)	鸡内金、菊花、桔梗、橘红、决明子、姜黄、高良姜、银杏叶、川芎、青皮、豆蔻、麦芽、广藿香



A-近邻高频原料之间权值的直接距离 B-输入样本的划分情况 C、D-近邻高频原料之间权值的距离
A-direct weight distance between nearby high-frequency functional ingredients B-partition of the input samples C, D-distance of weights between nearby high-frequency functional ingredients

图 5 高频原料 SOM 聚类分析结果神经元关系示意图

Fig. 5 Nerve relationship diagram of high-frequency functional ingredients with SOM analysis

表 10 可能配方 TOPSIS 分析结果

Table 10 TOPSIS analysis results of possible formula

排名	君药	臣药	佐/使药 1	佐/使药 2	TOPSIS 打分结果
1	黄芪	白及	五味子	甘草	0.158 6
2	黄芪	白及	陈皮	五味子	0.148 8
3	白及	白术	黄芪	甘草	0.097 5
4	白及	白术	茯苓	甘草	0.096 8
5	白及	五味子	陈皮	茯苓	0.093 7
6	白及	丹参	五味子	黄芪	0.087 3
7	白及	茯苓	五味子	贝母	0.086 7
8	白及	党参	五味子	陈皮	0.080 9
9	白及	白芍	白术	茯苓	0.078 6
10	白及	白芍	五味子	白术	0.071 0

2 位的配方在分数上差别并不大，在最后的选择上还需要结合其他的研究进一步筛选确定。

3 讨论

保健食品行业快速发展的同时^[22-23]，在研发阶段还存在着申报功能集中，对目录中其他功能研究

挖掘不充分，相同功能下配方原料选择同质化严重，对原料目录中部分原料功能利用不足，中药类保健食品配方原料配伍缺乏依据等诸多问题。针对以上情况，本研究通过对传统药食两用保胃良药白及的相关数据库进行整理，采用 Apriori、AHP、SOM 聚类、TOPSIS 等多种决策和聚类算法对数据库进行分析，得到以白及为主要原料的具有保胃护肝双功能的保健食品配方。

保健食品配方的一般思路即通过对数据库的内容整理，完成对原料的频次、原料所属类别等方面的统计，根据统计结果进行配方分析并根据实际情况确定配方^[17]。本研究在此基础上，增加了对统计结果的多角度分析，并对多个潜在配方进行量化评价。其中包括使用 Apriori 进行药味原料间关联度和网络图分析，作为后续配方的参考。对 AHP 的标准流程做出改进，通过比较多种较为客观、易量化且易重复的评价方法，选用依据所有数据库综合情况进行打分赋值。计算出各高频药味原料各级指标的权重后，采用线性加权的方法对各药进行处理并排

序,赋值缺项为-1,最后形成在保胃护肝双功能下高频药味原料的综合排名^[13]。然后对各原料进行聚类分析,将高频原料分成最优选药味到不建议使用药味共 7 类。根据中医理论中补脾益气的原则和对从最优选药味到一般推荐药味的功能和营养学等现代研究来综合进行选择,形成可与白及配伍使用并具有保胃护肝双功能的保健食品配方。最终根据配方中原料的君、臣、佐、使和总权重共 5 项数据经过 TOPSIS 分析得到的评分对潜在配方进行评价。

通过 Apriori 对各数据库的关联分析发现,保健食品信息、中成药库和方剂库中使用药味类别和高频原料药味重复率高。其中在 4 个数据库中出现频次都较高的药物类别是补虚药、清热药和理气药。4 个数据库中二次关联分析结果均可支持补脾益气的治疗原则,但是数据结果仅为频次或支持度等表现形式,缺乏对各原料本身或配伍的定量分析;另外二次关联分析数据过于分散,对数据库中原料药味的整体评价不充分。基于此,考虑引入 AHP、SOM 聚类分析来对 4 个数据库中的原料药味进行更多层次的分析,挖掘出在过往应用中可能没有被充分应用的原料。AHP 分析和 TOPSIS 分析在药学领域中应用广泛,涉及到提取工艺的优化、质量标记物的识别、保健食品原料评价体系的构建、药材质量的评价和药物利用评价及活性评价等多个方面^[13,24-27]。SOM 聚类之前多用于工程、金融等领域,在药学中应用罕见,但其与 AHP 和 TOPSIS 分析联用研究应用实例多样且比较成熟。AHP-SOM 聚类的联合应用多应用于多种方案的选择或者是对于定性和定量指标并存的情况下建立一个质量评价体系^[28-29]。相较而言,SOM-TOPSIS 联合使用可以大幅降低在最后选择评价时的对象数量^[30]。

本研究的局限性在于所得数据库容量相对较小,所以分析结果也相对集中。同时评分结果主要还是以各项指标的频次统计结果,所以在评分赋值中依然存在缺陷,后期拟采用结合临床资料、专利、科研项目等来综合赋值,为原料排序提供更多的客观参考。另外,还缺乏对最终结果算法的校验和处方可行性的验证环节,在未来的研究中会加强这方面的内容。

利益冲突 所有作者均声明不存在利益冲突

参考文献

[1] Manthey J, Shield K D, Rylett M, et al. Global alcohol

exposure between 1990 and 2017 and forecasts until 2030: A modelling study [J]. *Lancet*, 2019, 393(10190): 2493-2502.

[2] 黄富林,朱晓磊,郭岩,等. 中国酒的兴起、文化、促销及防控策略 [J]. *中国慢性病预防与控制*, 2020, 28(12): 957-960.

[3] Rehm J, Rehn N, Room R, et al. The global distribution of average volume of alcohol consumption and patterns of drinking [J]. *Eur Addict Res*, 2003, 9(4): 147-156.

[4] GBD Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017 [J]. *Lancet*, 2018, 392(10159): 1923-1994.

[5] 中国药典 [S]. 一部. 2020: 106-107.

[6] 巩子汉,王强,段永强,等. 白及多糖对胃溃疡模型大鼠胃组织 TNF- α 、IL-1 β 、IL-6 及 JNK、p38 MAPK 基因蛋白表达水平的影响 [J]. *中药药理与临床*, 2019, 35(4): 90-95.

[7] 高俊,丁兴红,丁志山,等. 白及对阿司匹林致大鼠胃溃疡的治疗作用研究 [J]. *浙江中医药大学学报*, 2019, 43(2): 182-187.

[8] 郑子春,沈洪,朱萱萱,等. 黄柏、地榆、白及对溃疡性结肠炎大鼠组织中 NF- κ B 和细胞因子表达的影响 [J]. *中国中医急症*, 2010, 19(3): 469-472.

[9] 贺国芳,丁伊玲,徐清霞,等. 白及多糖对小鼠急性酒精性肝损伤的保护作用 [J]. *中国医院药学杂志*, 2015, 35(18): 1658-1661.

[10] 中华人民共和国卫生部. 卫生部关于进一步规范保健食品原料管理的通知 (卫法监发[2002]51 号) [EB/OL]. (2002-03-11) [2018-12-05]. <http://www.nhfpc.gov.cn/sps/s3593/200810/bc239ea3d226449b86379f645dfd881d.shtml>.

[11] 王建,张冰. 临床中药学 [M]. 第 2 版. 北京: 人民卫生出版社, 2016: 15-19.

[12] Saaty T L. Decision making with the analytic hierarchy process [J]. *Int J Serv Sci*, 2008, 1(1): 83.

[13] 刘葵,陈紫雯,华桂丰,等. 基于层次分析法 (AHP) 的保健食品原料评价体系构建及分析 [J]. *中草药*, 2020, 51(18): 4829-4836.

[14] 陈春晖,杨军辉. 消化性溃疡的中医证型及证素研究 [J]. *中医学报*, 2020, 35(5): 1077-1080.

[15] 杨柳,薄颖异,于冰莉,等. 中医药防治酒精性肝病概况及相关机制研究进展 [J]. *中成药*, 2020, 42(3): 719-726.

[16] 钟赣生,苗彦妮. 中医药治疗酒精性肝病 [J]. *中国临床医生*, 2009, 37(5): 41-44.

- [17] 陈丰, 陈绍红, 柳海艳, 等. 辅助性保护化学性肝损伤保健食品配方的特点分析 [J]. 中草药, 2018, 49(7): 1703-1709.
- [18] 赵洪英, 蔡乐才, 李先杰. 关联规则挖掘的 Apriori 算法综述 [J]. 四川理工学院学报: 自然科学版, 2011, 24(1): 66-70.
- [19] Han J W, Pei J, Kambe R. *Data Mining Concepts and Techniques* [M]. Second Edition. San Francisco: Morgan Kaufmann, Elsevier, 2006: 234-242.
- [20] 张磐, 丁冷允, 姜宁, 等. 基于支持度-置信度-提升度的配网自动化系统数据挖掘算法及应用 [J]. 电测与仪表, 2019, 56(10): 62-68.
- [21] 邓雪, 李家铭, 曾浩健, 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.
- [22] “健康中国 2030”规划纲要 [EB/OL]. (2016-10-25) [2018-12-05]. <https://baike.baidu.com/item/%E2%80%9C%E5%81%A5%E5%BA%B7%E4%B8%AD%E5%9B%BD2030%E2%80%9D%E8%A7%84%E5%88%92%E7%BA%B2%E8%A6%81/19926077?fr=aladdin>.
- [23] 国务院办公厅关于印发国民营养计划 (2017—2030 年) 的通知 (国办发[2017]60 号) [EB/OL]. (2017-07-13) [2018-12-05]. http://www.gov.cn/zhengce/content/2017-07/13/content_5210134.htm.
- [24] 张娇, 蒋倩倩, 张伯言, 等. 基于 AHP-CRITIC 法正交优选乌甘袋泡茶提取工艺及抗炎作用研究 [J]. 中草药, 2020, 51(8): 2177-2184.
- [25] 冯利梅, 陈艳琰, 乐世俊, 等. 基于层次分析-熵权法的中药质量标志物量化辨识方法研究: 以芍药甘草汤为例 [J]. 药学报, 2021, 56(1): 296-305.
- [26] 陈溪, 毛艳梅, 陈艳平, 等. 基于加权 TOPSIS 法的盐酸左氧氟沙星注射液药物利用评价 [J]. 中国药房, 2015, 26(32): 4471-4475.
- [27] 谈梦霞, 陈佳丽, 邹立思, 等. 不同加工麦冬中多元活性成分的含量测定及多元统计分析 [J]. 药物分析杂志, 2019, 39(6): 992-1002.
- [28] Yu P, Lee J H. A hybrid approach using two-level SOM and combined AHP rating and AHP/DEA-AR method for selecting optimal promising emerging technology [J]. *Expert Syst Appl*, 2013, 40(1): 300-314.
- [29] Wei Z S, Ma X P, Zhan P, *et al*. Flavor quality evaluation system of Xinjiang milk knots by using SOM neural network and the fuzzy AHP [J]. *Food Sci Nutr*, 2020, 8(4): 2088-2093.
- [30] Jędrkiewicz R, Tsakovski S, Lavenu A, *et al*. Simultaneous grouping and ranking with combination of SOM and TOPSIS for selection of preferable analytical procedure for furan determination in food [J]. *Talanta*, 2018, 178: 928-933.

[责任编辑 潘明佳]