

• 数据挖掘与循证医学 •

基于功能导向-算法集成的西洋参双功能保健食品组方规律和设计研究

宋若兰^{1,2}, 马嘉慕^{1,2}, 郭海峰³, 于爽^{1,2}, 姚鉴玲^{1,2}, 董英^{1,2}, 钟祥健^{1,2}, 单东杰^{1,2}, 吕芳^{1,2}, 折改梅^{1,2*}

1. 北京中医药大学中药学院, 北京 100029
2. 中药品质评价北京市重点实验室, 北京 100029
3. 中国营养保健食品协会, 北京 100028

摘要: **目的** 基于西洋参“保健食品原料”与“中药材”的双重身份, 以功能为导向, 融合算法分析, 实现西洋参保健食品组方规律挖掘与组方设计研究。**方法** 依据功能导向-算法集成, 收集西洋参保健食品数据与临床应用数据, 借助 BDP 平台、自然语言处理与信息检索 (natural language processing & information retrieval sharing platform, NLPPIR) 共享平台实现数据处理与信息挖掘, 综合分析具有潜在开发价值的保健食品功能; 以选定功能为导向, 收集西洋参保健食品、中成药、方剂的组方信息并建立数据库, 统计高频药味及类别, 采用 Apriori 关联规则算法、层次分析法 (analytic hierarchy process, AHP)-熵权-自组织特征映射 (self-organizing maps, SOM) 联用的赋权聚类算法 (主观与客观算法联用)、逼近理想解排序 (technique for order preference by similarity to an ideal solution, TOPSIS) 综合评价法等机器学习算法提取样本数据特征并分析总结其内在规律, 实现组方设计。**结果** 综合西洋参保健食品功能与临床疾病和证型双方信息的分析结果, 以“有助于增强免疫力”“对化学性肝损伤有辅助保护功能”双功能为导向, 收集组方信息并建库。数据库中高频保健食品原料药味以补虚药居多, 经赋权聚类筛选得到黄芪、麦冬、茯苓、五味子、枸杞子、三七、葛根、甘草共 8 味候选保健食品原料药味。结合药味关联分析及药对组合情况, 进一步配伍组成备选组方, 经评分优选得到保健食品组方为三七、西洋参、葛根、五味子、枸杞子。**结论** 针对中药保健食品的中医药背景, 依托大数据平台分析挖掘保健食品功能, 借助机器学习算法抽取组方配伍规律, 实现功能导向-算法集成西洋参双功能保健食品组方设计, 或可为中药保健食品组方设计的功能导向化、组方计算化提供新思路。

关键词: 保健食品; 西洋参; 组方规律; 组方设计; 功能; 算法; 数据挖掘

中图分类号: R283.21; TS218; TP312 **文献标志码:** A **文章编号:** 0253-2670(2022)11-3415-10

DOI: 10.7501/j.issn.0253-2670.2022.11.018

Rule and design of dual-function health food formula of *Panaxis Quinquefolii Radix* based on function orientation-algorithm integration

SONG Ruo-lan^{1,2}, MA Jia-mu^{1,2}, GUO Hai-feng³, YU Shuang^{1,2}, YAO Jian-ling^{1,2}, DONG Ying^{1,2}, ZHONG Xiang-jian^{1,2}, SHAN Dong-jie^{1,2}, LYU Fang^{1,2}, SHE Gai-mei^{1,2}

1. School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 100029, China
2. Beijing Key Laboratory for Quality Evaluation of Chinese Materia Medica, Beijing 100029, China
3. China Nutrition and Health Food Association, Beijing 100028, China

Abstract: Objective Based on the dual-identity of Xiyangshen (*Panaxis Quinquefolii Radix*) as health food functional ingredients and traditional Chinese medicine, the functional orientation and algorithm analysis were combined to realize the mining of health food formula rule and formula design of *Panaxis Quinquefolii Radix*. **Methods** According to the function orientation-algorithm integration, the data of health food and clinical application of *Panaxis Quinquefolii Radix* were collected and processed by business data platform (BDP) and natural language processing and information retrieval (NLPPIR) sharing platform to analyze the health food functions with potential development value comprehensively through the information mining. Guided by selected functions, the

收稿日期: 2022-01-04

基金项目: 桂林市技术创新引导计划 (2020010902); 河南省重大科技专项 (211110310100)

作者简介: 宋若兰 (1998—), 女, 博士研究生, 研究方向为中药化学。E-mail: songruolan@126.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

formula information of *Panacis Quinquefolii Radix* health food, Chinese patent medicine and prescription was collected to establish a database. The high-frequency health food functional ingredients and categories were counted. The characteristics of data were extracted and the formula internal rules were analyzed and summarized through machine learning algorithms as Apriori algorithm, analytic hierarchy process (AHP)-entropy-self-organizing maps (SOM) algorithm (subjective and objective algorithms combined) and technique for order preference by similarity to an ideal solution (TOPSIS) to achieve the formula design. **Results** Based on the analysis results of health food functions and clinical disease and syndrome types, the formula information was collected and the database was established with the orientation of “contributing to boost immunity” and “auxiliary protective function on chemical liver injury”. In the database, most of the high-frequency health food ingredients were tonifying deficiency herbs. Eight candidate health food ingredients including Huangqi (*Astragali Radix*), Maidong (*Ophiopogonis Radix*), Fuling (*Poria*), Wuweizi (*Schisandrae Chinensis Fructus*), Gouqizi (*Lycii Fructus*), Sanqi (*Notoginseng Radix et Rhizoma*), Gegen (*Puerariae Lobatae Radix*) and Gancao (*Glycyrrhizae Radix et Rhizome*) were obtained by weighted clustering. Combined with the herbs association and herb pair analysis, the alternative prescriptions were further formed. The health food prescription optimized by scoring was consisted of *Notoginseng Radix et Rhizoma*, *Panacis Quinquefolii Radix*, *Puerariae Lobatae Radix*, *Schisandrae Chinensis Fructus* and *Lycii Fructus*. **Conclusion** According to the traditional Chinese medicine background of Chinese materia medica health food, relying on the big data analysis platform for the health food functions, the prescription rules were summarized and extracted by the machine learning algorithms. The realization of the function orientation-algorithm integration of *Panacis Quinquefolii Radix* health food formula rule mining and dual-function health food formula design may provide a new development idea for the function-oriented and formula-calculated of Chinese materia medica health food formula design.

Key words: health food; *Panacis Quinquefolii Radix*; prescription rule; prescription design; function; algorithm; data mining

中药保健食品作为我国保健食品行业不可或缺的重要组成部分^[1], 是中医药传统养生保健理念与实践相结合的产物^[2]。此类产品所用原料拥有“保健食品原料”和“中药材”的双重身份, 具有明确的保健功能与临床应用基础, 独具我国保健食品特色^[2]。然而, 中药保健食品在研发时却极易陷入忽略或弱化“中药材”属性的困境, 无法充分彰显其中医药背景优势, 难以确立其在大健康产业中的独特定位^[2-4]。究其原因, 一是保健食品功能与中药功效主治存在不对等关系, 较难构建有效结合, 导致功能无法有效引领中药保健食品; 二是在组方设计过程中未能充分体现中药的配伍应用, 存在缺乏中医药理论指导、忽视单味药功效及其在组方配伍中作用的区别、尚未充分利用中药在疾病治疗及预防中的临床应用经验等现象。如何在融合中医药背景的同时凸显保健食品特色, 在研发过程中兼顾其作为中药的特点, 是中药保健食品的破局之题。

基于以上问题, 本研究以组方设计为切入点, 围绕中药保健食品原料收集保健食品批文与临床应用数据, 选用大数据分析平台 BDP (business data platform, <https://me.bdp.cn/home.html>)、自然语言处理与信息检索 (natural language processing & information retrieval sharing platform, NLPir) 共享平台 (<http://www.nlpir.org/wordpress/>) 作为大体量数据及文本型数据的处理途径, 通过关键词提取与

数据挖掘分析完成信息处理工作^[5-6], 综合分析西洋参具有潜在开发价值的保健食品功能。以西洋参应用基础广泛、研究基础扎实的保健食品功能“有助于增强免疫力”, 以及从西洋参配伍和临床应用角度考量、大数据分析处理结果指向的“对化学性肝损伤有辅助保护功能”为导向, 开展后续双功能组方设计。围绕所选保健食品功能, 收集保健食品、中成药、方剂的组方信息。通过 Apriori 关联规则算法、层次分析法 (analytic hierarchy process, AHP)-熵权-自组织特征映射 (self-organizing maps, SOM) 联用的赋权聚类算法、逼近理想解排序 (technique for order preference by similarity to an ideal solution, TOPSIS) 综合评价法等机器学习算法提取样本数据特征并分析总结其内在规律, 设计并优选组方。并在团队前期单一主观赋权算法的基础上, 开展主观与客观赋权算法联用的方法提升, 进一步增强筛选结果的可信度^[7]。本研究将中药保健食品功能方向的筛选、确定, 以及在指明功能方向的前提下进行组方设计称为“功能导向”; 将针对中药保健食品的中医药背景, 融合多种算法总结分析规律性、经验性较强的组方配伍规律, 并实现组方设计的全过程称为“算法集成”。由此, 二者相结合, 以特定保健食品功能为导向, 融合多种机器学习算法, 实现保健食品组方规律挖掘与组方设计。

西洋参 *Panacis Quinquefolii Radix* 作为典型的

外来中药,首载于清代《补图本草备要》,在保健食品开发中应用广泛^[8]。西洋参较其他中药保健食品原料药味而言,在中医药理论指导下的传统应用历史相对不足,但其临床应用研究较为丰富^[9-12]。本研究选择广泛用于保健食品开发且临床研究基础扎实的中药保健食品代表性原料“西洋参”作为研究对象,依据功能导向-算法集成,开展中药保健食品组方设计研究,或可为中药保健食品组方设计的功能导向化、组方计算化提供新的研发思路。

1 方法

1.1 功能数据收集及处理分析

围绕保健食品与临床应用 2 个方面,统计整理西洋参应用概况。保健食品数据来源于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/>)中含西洋参的保健食品批文;临床应用数据来源于中国知网,以“西洋参”和“临床观察”为关键词进行检索(2021 年 8 月 31 日前收录文献)。数据统计及结果可视化处理依托大数据分析平台和 CytoScape 3.9.0 实现,分析平台包括 BDP 平台和 NLPIR 自然语言处理与信息检索共享平台中的分词标注、实体抽取、词频统计、词向量分析模块^[5]。

1.2 组方信息收集

依据分析所得保健食品功能方向,从保健食品、中成药、方剂 3 个方面,开展含西洋参保健食品的组方信息收集工作。保健食品信息来源于国家市场监督管理总局,中成药信息来源于《中国药典》2020 年版^[13]及药智数据(<https://db.yaozh.com/>),方剂信息来源于中国知网、万方数据库、维普数据库和药智数据(均为 2021 年 8 月 31 日前收录)。对于保健食品信息,筛选具有“有助于增强免疫力”或“对化学性肝损伤有辅助保护功能”的含西洋参保健食品批文,完成组方信息收集。借助 Microsoft Excel 2019 实现组方信息录入与建库。

1.3 药味筛选

1.3.1 保健食品原料范围锁定 根据组方数据库,使用 Microsoft Excel 2019,分别整理统计保健食品、中成药、方剂各组方药味种类及频次,并进一步在规范药味名称的基础上(将同物异名者、同种中药但规格相异者、仅经提取等特殊处理但功效未改变者统一规范为《中国药典》名称),依照卫生部《既是食品又是药品的物品名单》《可用于保健食品的物品名单》和《保健食品禁用物品名单》^[14]将非保健食品原

料剔除,筛选得到可作为保健食品原料的药味。

1.3.2 各药味频次统计分析 运用中药系统药理学分析平台(<http://tcmospw.com/tcmosp.php>)的中药数据库,结合《中国药典》2020 年版^[13]、《临床中药学》(第 2 版)^[15]的相关记载,将各药味依照功效主治归类,确定其所属类别。分别统计各药味及各类别在含西洋参保健食品、中成药、方剂组方中的应用频次,筛选得到高频保健食品原料药味及类别。

1.3.3 药味间的关联分析及赋权聚类 通过 SPSS Modeler 18.0 中的 Apriori 关联规则算法,对高频药味在组方中的应用信息进行网络图可视化处理^[16-17],呈现药味关联情况及药对组合情况。

借助 Matlab R2021a,采用 AHP-熵权-SOM 联用算法,完成高频药味的赋权与聚类。采用主观与客观赋权算法相结合的分析方法(AHP-熵权法)^[18],对 2 级指标(药味类别)开展基于评分的主观权重分析,对 3 级指标(药味)分别进行主观和客观权重分析。在赋权的基础上,进一步运用 SOM 聚类算法^[7],得到主观和客观聚类结果,关注二者结果中的最优选药味,取共有者,作为后续组方配伍的候选保健食品原料药味。

1.4 组方配伍

基于候选保健食品原料药味,结合分析所得药味关联情况及药对组合情况,提供“有助于增强免疫力”且“对化学性肝损伤有辅助保护功能”的备选组方。并进一步在 Matlab R2021a 中,运用 TOPSIS 综合评价法计算各备选组方得分并排位^[7],对所设计保健食品组方进行综合评价分析,完成西洋参双功能中药保健食品组方设计。

本研究技术路线如图 1 所示。

2 结果

2.1 双功能选定

2.1.1 西洋参保健食品功能分布 “有助于增强免疫力”居首位 共收集 972 条含西洋参保健食品批文,所涉及保健食品功能种类、频次及分布如图 2 所示(功能名称已参照《保健功能释义(2022 年版)》进行规范^[19])。所涉及功能种类丰富、覆盖范围广,包括有助于增强免疫力、缓解体力疲劳、有助于抗氧化、有助于维持血糖健康水平、耐缺氧、有助于润肠通便、有助于改善黄褐斑、对化学性肝损伤有辅助保护功能、有助于维持血脂健康水平、清咽润喉等共 20 项功能。功能分布较为集中,以“有助于增强免疫力”居多,占比达 37%。

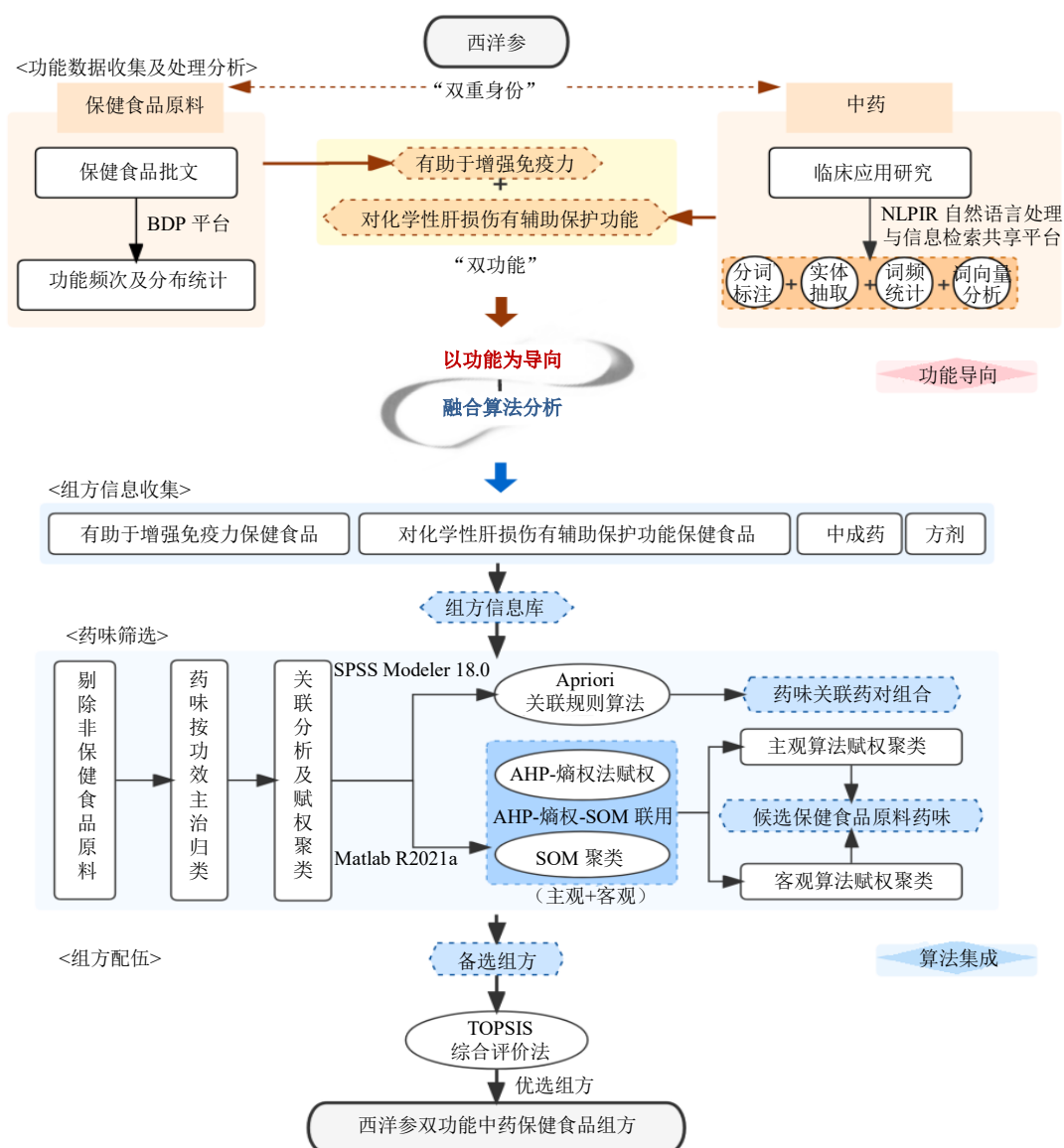


图1 技术路线

Fig. 1 Technical route

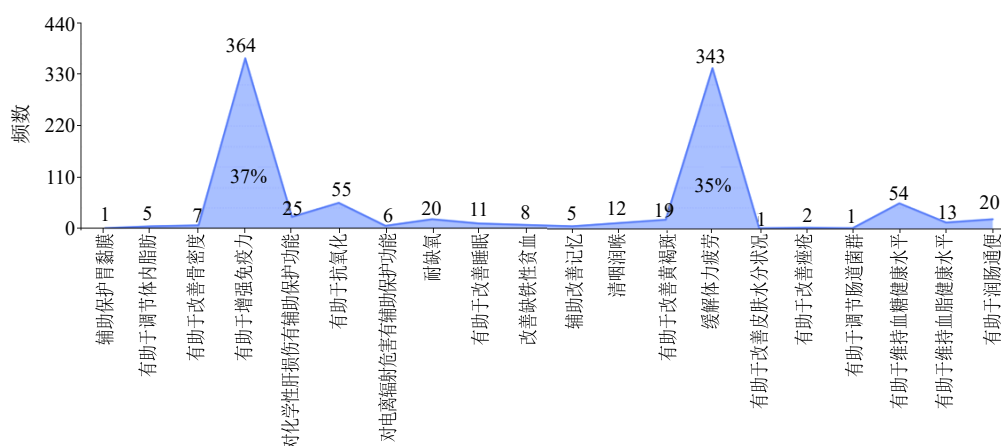


图2 含西洋参保健食品功能频次统计及分布

Fig. 2 Statistic of functional frequency and distribution of health food containing *Panacis Quinquefolii Radix*

2.1.2 西洋参临床应用与“肝”关系密切 提取并梳理 1089 项西洋参临床应用研究中的疾病及证型, 输入 NLPiR 自然语言处理与信息检索共享平台处理框, 处理流程及结果如图 3 所示。使用分词标注模块, 依据《计算所汉语词性标记集 Version 3.0》识别并区分名词、动词、形容词、前后缀、时间词、状态词等词性。在此基础上, 通过实体抽取模块的完美双数组 Trie 树算法抽提得到“气阴两虚”“糖尿病”“乙型肝炎”“心脏病”“脓毒症”共 5 个关键词。进一步借助词频统计模块, 聚焦名词

统计排名前 10 位的词语, 可得西洋参临床应用频率较高的病症为“糖尿病”, 且多用于与心、肺、肝、肾等器官相关的疾病或证型。利用词向量分析模块, 采用基于 CBOW (continuous bag-of-words model) 方法的 Word2vector 改进模型, 综合词性、词的分布特点, 提取本体关联关系, 借助 CytoScape 3.9.0 实现分析结果的可视化^[6,20-21]。结果显示, 西洋参临床应用的疾病及证型中心词为“乙型肝炎”, 而与之相关联的 2 层、3 层关键词多围绕“真菌”“肝炎”等展开。

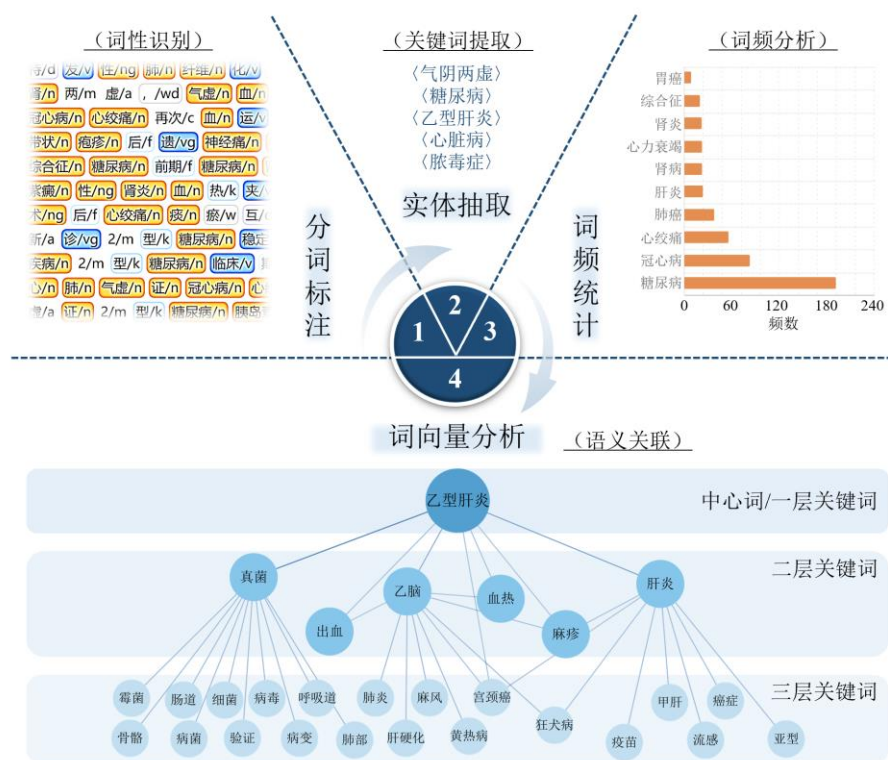


图 3 西洋参临床应用数据 NLPiR 平台处理流程及结果

Fig. 3 Process and results of processing *Panacis Quinquefolii Radix* clinical application data through NLPiR platform

2.2 组方数据库构建

以“有助于增强免疫力”“对化学性肝损伤有辅助保护功能”双功能为导向, 构建西洋参组方信息库, 共收集信息 441 条, 包括“有助于增强免疫力”西洋参保健食品 364 条、“对化学性肝损伤有辅助保护功能”西洋参保健食品 25 条、中成药 33 条以及方剂 19 条。

2.3 候选保健食品原料药味确定

2.3.1 高频保健食品原料药味筛选 规范西洋参组方数据库中各保健食品原料药味名称, 经初步筛选得到可作为保健食品原料的药味共 155 个。按照功

效主治归类, 药味分属于 18 个类别, 以补虚药、清热药、理气药居多 (3 者占比超过半数)。进一步筛选得到在保健食品、中成药、方剂中应用的高频保健食品原料药味及类别, 结果如图 4 所示。

2.3.2 药味关联分析及药对组合情况 采用 Apriori 关联规则算法描绘高频保健食品原料药味在保健食品、中成药、方剂中应用的关联情况及药对组合情况, 为后续组方的药味配伍提供参考依据, 如图 5 和表 1 所示 (图 5 呈现各药味间的关联关系网络; 表 1 展示支持度大于 10%, 置信度大于 40% 的药对组合结果)。

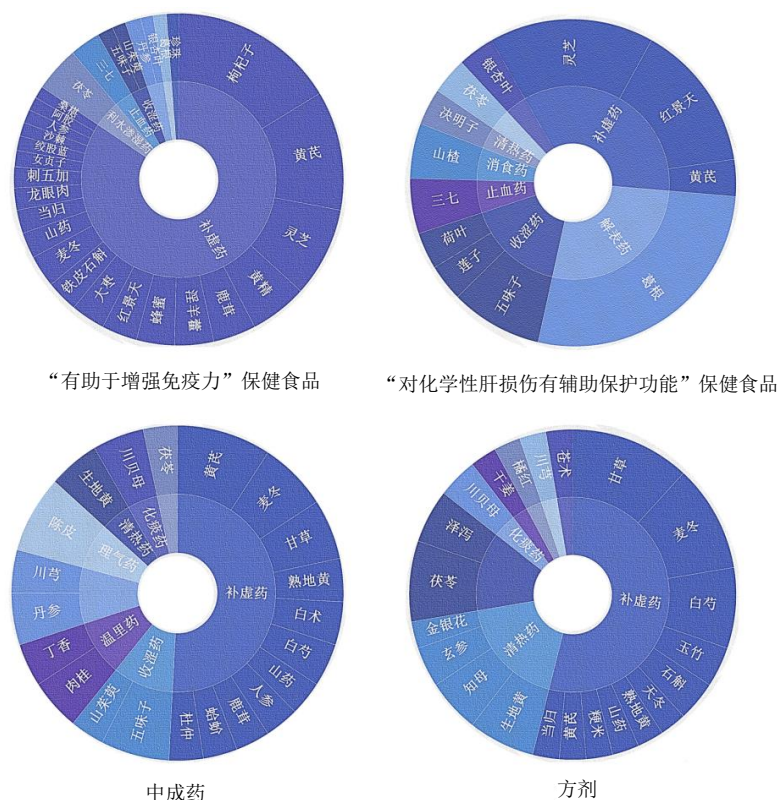


图 4 高频保健食品原料药味及类别

Fig. 4 Functional ingredients and categories of high-frequency health food

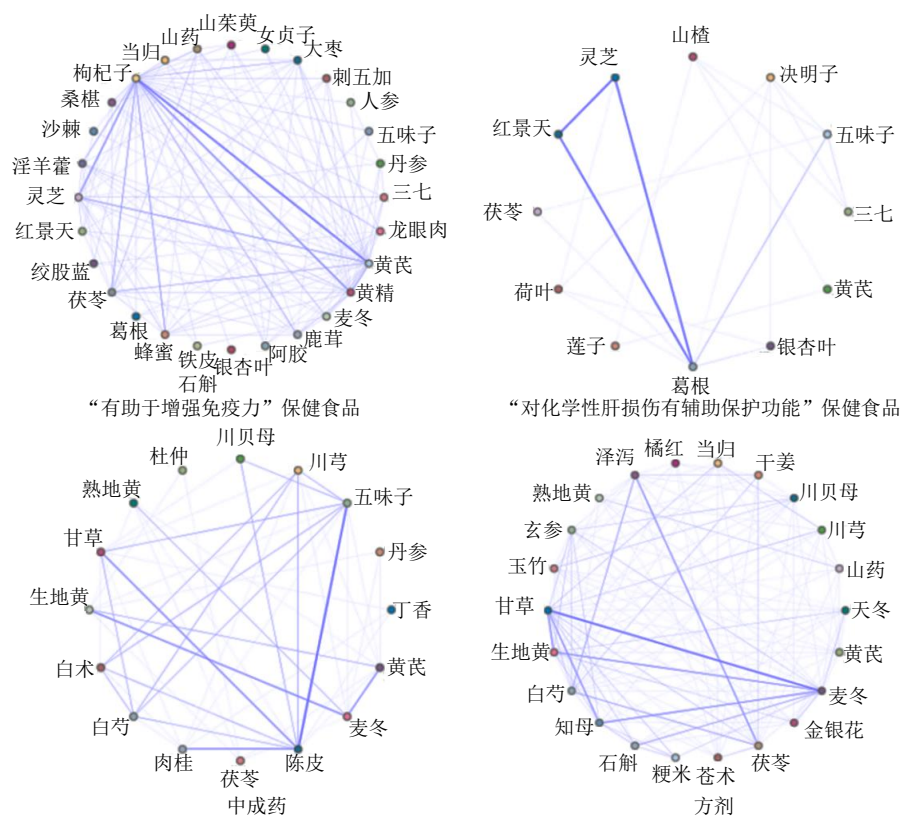


图 5 高频保健食品原料药味关联关系网络

Fig. 5 Association networks of high-frequency health food functional ingredients

表 1 高频保健食品原料药味药对组合情况

Table 1 Combination of herb pairs of high-frequency health food functional ingredients

组方信息来源	药对组合
“有助于增强免疫力”保健食品	黄芪-枸杞子、灵芝-枸杞子、黄精-枸杞子、黄精-黄芪、茯苓-枸杞子、茯苓-黄芪
“对化学性肝损伤有辅助保护功能”保健食品	葛根-红景天、葛根-灵芝、灵芝-红景天、五味子-葛根、五味子-三七
中成药	黄芪-麦冬、陈皮-五味子、陈皮-甘草、陈皮-肉桂、麦冬-生地黄
方剂	甘草-麦冬、甘草-知母、甘草-生地黄、麦冬-知母、麦冬-生地黄

保健食品、中成药、方剂的分析结果具有一定差异。在西洋参保健食品中，“有助于增强免疫力”形成以枸杞子为中心的关联网络，其分别与黄芪、黄精、灵芝、茯苓、蜂蜜组成关联性较强的药对关系；“对化学性肝损伤有辅助保护功能”表现为葛根、灵芝、红景天关联强，葛根-五味子、五味子-三七具有一定关联的网络关系。中成药和方剂的药味关联具有一定相似性，主要药味均包含甘草、麦冬、生地黄，但又分别展现出五味子-陈皮-肉桂和茯苓-泽泻-白芍的关联关系。

2.3.3 赋权与聚类分析 运用 AHP-熵权-SOM 联用算法，对高频保健食品原料药味进行赋权与聚类。

参考文献中的各级指标设置方法，设定“类别”

为 2 级指标，共包含 15 个类别；“保健食品原料药味”为 3 级指标，共包含 64 个药味^[7]，各级指标采用不同算法进行处理。

对于 2 级指标，采用 AHP 法，选择以频次为依据的最优打分赋值方式对各类别及各类别项下药味进行权重计算。最优打分赋值方式是以“有助于增强免疫力”保健食品、“对化学性肝损伤有辅助保护功能”保健食品、中成药、方剂及以上 4 者综合统计所得频次为依据的 5 种打分赋值方式中一致性指标(consistency index, CI)和一致性比例(consistency ratio, CR)绝对值较小的方式。结果如表 2 所示，最优方式为“4 者综合”。

表 2 AHP 法 2 级指标打分赋值方式考察结果

Table 2 Evaluation of rating assignment score for secondary index by AHP

打分赋值方式的频次统计信息来源	CI 值	CR 值
“有助于增强免疫力”保健食品	-8.55×10^{-5}	-5.38×10^{-5}
“对化学性肝损伤有辅助保护功能”保健食品	-1.18×10^{-3}	-7.43×10^{-4}
中成药	-3.38×10^{-4}	-2.13×10^{-4}
方剂	-4.31×10^{-4}	-2.71×10^{-4}
4 者综合	-7.58×10^{-5}	-4.77×10^{-5}

对于 3 级指标，考虑到其包含的药味数量较多，仅选择 AHP 法进行主观打分赋值并开展权重计算可能存在一定的不合理性。采用主观与客观相结合的赋权算法可提高筛选结果的可信度^[18]，故分别采用 AHP 法与熵权法进行计算。在考虑“有助于增强免疫力”保健食品、“对化学性肝损伤有辅助保护功能”保健食品、中成药、方剂及以上 4 者综合统计所得频次的基础上，计算得到各药味在 2 种算法下的权重赋值。

依据 2 级指标及 2 种算法下 3 级指标的加权结果，采用 SOM 聚类分析为各高频保健食品原料药味赋予分值，得到主观和客观 2 种赋权聚类结果，图 6 依次展示了每种分类结果的聚类关联图（图中

灰色正六边形表示神经元节点，正六边形间的色块颜色越偏向于黑色则表示节点间的聚类关联越弱）和神经元分布情况图。根据各药味分值进行分类，分值接近者聚为一类，每 10 分为一类，共 7 类。药味得分低于 10 分者属于第 1 类，即最优选类别。取二者最优选类别的共有药味作为候选保健食品原料药味，结果如图 6 所示（共有药味以红色框标识，0 分表示未存在于最优选类别中）。候选保健食品原料药味为黄芪、麦冬、茯苓、五味子、枸杞子、三七、葛根、甘草共 8 味药材。

2.4 组方评分及优选

基于传统中医药理论和现代研究，针对 8 味候选保健食品原料药味，参考关联分析及药对组合情

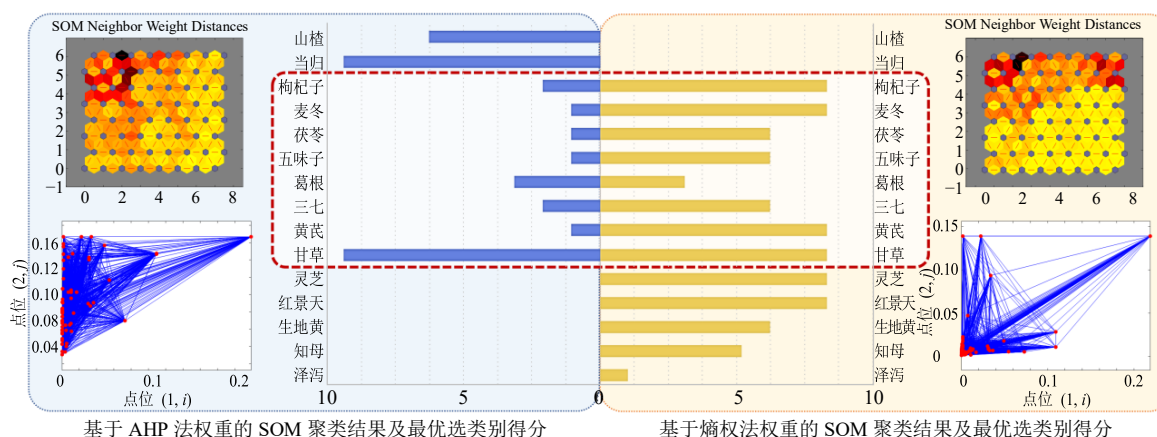


图 6 AHP-熵权-SOM 联用算法的聚类分析结果及最优选类别药味分值

Fig. 6 Cluster results by AHP-entropy-SOM combined algorithm and score of optimal drugs

况的分析结果,生成 10 个备选组方。运用 TOPSIS 综合评价法,以各药味权重(2 级指标及 2 种算法下 3 级指标权重加和)为依据,为各组方评分并排位,结果如表 3 所示。结果表明,排名前 4 位的组方评分较高,最优选的西洋参“有助于增强免疫力”和“对化学性肝损伤有辅助保护功能”双功能中药保健食品组方由三七、西洋参、葛根、五味子、枸杞子组成。

3 讨论

3.1 西洋参双功能保健食品功能选定

“有助于增强免疫力”是西洋参应用最为广泛的保健食品功能,同时也是受现有保健食品市场青睐的“热门”功能,有较为扎实的研究基础^[2,22]。参考《保健功能释义(2022 年版)》^[19]，“免疫力”包括机体对外防御和对内环境维持稳定 2 个方面。在一定程度上,2 个方面可分别与传统中医理论中的“祛

邪”与“扶正”相联系^[4],即在扶助正气、调补脏腑、增强体质的同时通过利水祛湿、清热解毒等方法达到抵御病邪及清除病理产物的效果^[3]。西洋参作为补虚药,长于补气益阴、清热生津,在传统中医药理论的指导下常与其他药味配伍以达到“扶正祛邪”的目的。同时,NLPIR 平台分析结果中的“气阴两虚”“脓毒症”等关键词均与“免疫”关系密切。由此可得,西洋参与“有助于增强免疫力”这一保健功能的关联性较强。

“对化学性肝损伤有辅助保护功能”在功能频次统计中位列第 8,但由于西洋参保健食品功能的分布较为集中,故相较而言,西洋参用于“对化学性肝损伤有辅助保护功能”保健食品的研发背景略显不足。然而,在 NLPIR 平台的临床应用数据分析中,词向量分析结果以“乙型肝炎”为中心词进行展开(词向量分析可依据生语料进行训练产生专业领域的本体关联关系,更适于呈现词语间的指向性与关联性),且实体抽取与词频分析均显示其临床研究与肝脏呈现较强的相关性。因此,西洋参用于肝脏疾病的相关临床研究具有大数据分析背景作为支撑。

此外,“化学性肝损伤”在中医视角下的关键诱因因为湿、热、毒、痰、瘀等互结,致使肝失疏泄、脾失健运、肾阴亏虚^[23],其基本治则符合“扶正祛邪”理论,即扶固正气,补虚养阴,去湿、热、毒、痰、瘀等病邪,这与西洋参的补阴功效相合。加之,考虑到西洋参的“中药”属性,中药通常以药物配伍的方式进行应用^[24],除西洋参自身功效的考量外,西洋参可与其他药味配伍组合,作为中药方剂中的臣药或佐使药,针对具体应用情况,辅佐

表 3 备选组方的 TOPSIS 评分

Table 3 Scores of alternative prescriptions by TOPSIS

排名	备选组方	评分
1	三七、西洋参、葛根、五味子、枸杞子	0.151 9
2	三七、西洋参、黄芪、葛根、甘草	0.124 8
3	三七、西洋参、黄芪、葛根、五味子	0.123 3
4	西洋参、黄芪、葛根、茯苓、枸杞子	0.118 6
5	三七、西洋参、黄芪、茯苓、五味子	0.094 4
6	三七、西洋参、黄芪、麦冬、甘草	0.086 6
7	三七、西洋参、黄芪、麦冬、茯苓	0.081 2
8	西洋参、黄芪、茯苓、五味子、枸杞子	0.078 7
9	西洋参、黄芪、麦冬、五味子、枸杞子	0.074 6
10	西洋参、黄芪、麦冬、茯苓、枸杞子	0.065 8

君药,在其单味药的应用基础上扩大应用范围,达到不同的治疗效果,这也在一定程度上为西洋参单味药的传统功效与基于大数据平台的临床应用分析结果存在差异供了合理解释。西洋参单味药的功效看似与“对化学性肝损伤有辅助保护功能”缺少直接联系,但其通过与其他药味的配伍组合,可用于与肝脏相关疾病的临床治疗。因此,后续的组方配伍设计研究也可进一步为西洋参“对化学性肝损伤有辅助保护功能”保健食品的研发增加可能性。

基于此,从“保健食品原料”及“中药材”2方面考虑,选择应用最为广泛且研究基础扎实的“有助于增强免疫力”保健食品功能,以及与大数据分析处理平台的临床应用分析结果相关的“对化学性肝损伤有辅助保护功能”进行西洋参双功能中药保健食品开发,具有一定可行性。故进一步以此双功能为导向,在此基础上开展组方设计研究。

3.2 AHP-熵权-SOM 赋权聚类算法联用

机器学习是挖掘海量组方信息、提取内在配伍规律的有力抓手^[25-26]。在赋权聚类中,本研究最初采用 AHP 法对高频保健食品原料药味进行赋权。然而,由于本研究 3 级指标所包含样本数较多,仅凭 AHP 主观赋权算法进行赋值打分并依据其权重计算结果开展聚类分析会略有偏颇。

现有赋权算法可分为主观与客观 2 类,结合主观判断与客观指标之间的变异系数或相关关系分别确定权重结果,算法联用,互为补充,可提高筛选结果的可信度^[18]。由此,本研究采用 AHP-熵权法-SOM 联合的赋权聚类方法,结合主观(AHP 法)与客观(熵权法)赋权算法,依据二者的赋权结果,分别进行 SOM 聚类分析,选择最优类别中的共同者,作为候选保健食品原料药味,开展后续组方配伍研究。

3.3 优选组方分析

优选所得西洋参“有助于增强免疫力”和“对化学性肝损伤有辅助保护功能”双功能中药保健食品组方由三七、西洋参、葛根、五味子、枸杞子组成。

在传统中医药理论中,免疫功能属于中医“正气”范畴;而肝损伤则与正虚不御邪,毒邪伤正,耗损肝肾阴精气血关系密切^[4,23]。本组方中三七为补血第一要药,补血活血而不留瘀;西洋参补气养阴、清热生津;葛根解肌退热升阳、生津;五味子益气生津、补肾宁心;枸杞子补益肝肾、益精。5 者

配伍组合可益气补血、养阴生津、补肾宁心,补气血、益精血、祛瘀生新,同时清热、利水、消肿、散瘀以驱除邪气,发挥保健食品双功能^[27-29]。

此外,现代研究表明,本组方中三七和西洋参含有的皂苷类成分、葛根含有的葛根素、五味子含有的五味子醇甲以及枸杞子含有的多糖类成分或在保健食品双功能中发挥主要作用。三七总皂苷能降低促炎因子(肿瘤坏死因子- α 、白细胞介素-6)的表达,升高抑炎因子表达从而减缓炎症反应的发生,与免疫力密切相关^[30]。西洋参皂苷类成分可通过增加中性粒细胞数量、改善巨噬细胞吞噬功能及促进体内 γ -干扰素的分泌来实现对机体免疫功能的调控作用^[31]。西洋参皂苷类成分对 D-半乳糖胺/脂多糖所致急性肝损伤具有一定的保护作用^[32]。其中,含量最高的人参皂苷(人参皂苷 Rb_1)可通过抑制肝实质中性粒细胞浸润,以及下调核因子- κB 通路相关促炎因子,缓解肝脏炎症^[33]。葛根素、五味子醇甲可调节肝细胞白细胞介素-1 β 、环氧合酶-2、前列腺素 E_2 等炎症因子的蛋白表达,减轻肝脏炎性反应,减少肝细胞死亡,恢复肝功能,发挥保护肝脏的作用^[34-35]。枸杞糖肽作为具有免疫活性的糖蛋白,可调节外周和中枢炎症,改善多脏器功能,枸杞子含有的多糖类成分可有效缓解肝功能损伤,并在多种肝脏疾病治疗中发挥积极作用^[28]。

综上所述,本研究围绕中药保健食品原料西洋参的“双重身份”,依托大数据平台,分析挖掘具有潜在开发价值的保健食品功能,以“有助于增强免疫力”“对化学性肝损伤有辅助保护功能”双功能为导向,收集保健食品、中成药、方剂组方信息,联合多种算法完成药味关联分析、赋权聚类、组方配伍及评价,实现功能导向-算法集成西洋参双功能保健食品组方设计。后续拟在此研究思路的基础上,进一步开展在保健食品领域应用广泛且在中医药理论指导下传统应用历史悠久的中药保健食品原料(如枸杞子)的组方设计研究。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 孙蓉,齐晓甜,陈广耀,等. 中药保健食品研发、评价和产业现状及发展策略 [J]. 中国中药杂志, 2019, 44(5): 861-864.
- [2] 李美英,姜雨,余超. 我国保健食品功能与原料管理的一点思考 [J]. 营养学报, 2018, 40(3): 215-221.
- [3] 杨明,胡彦君,王雅琪,等. 基于中医药理论与优势的中药保健食品设计思路 [J]. 中草药, 2017, 48(3): 419-

- 423.
- [4] 齐晓甜, 赵春媛, 张家祥, 等. 中药类保健食品增强免疫力功能评价现状和研究策略 [J]. 中国中药杂志, 2019, 44(5): 875-879.
 - [5] 张华平, 商建云. NLP-Parser: 大数据语义智能分析平台 [J]. 语料库语言学, 2019, 6(01): 87-104.
 - [6] 张华平, 高凯, 黄河燕. 大数据搜索与挖掘 [M]. 北京: 科学出版社, 2014: 281.
 - [7] 马嘉慕, 刘晓云, 任雪阳, 等. 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 [J]. 中草药, 2021, 52(18): 5676-5687.
 - [8] 王林元, 王淳, 赵丹萍, 等. 新外来中药现代研究范式的构建及应用 [J]. 中国中药杂志, 2020, 45(5): 967-977.
 - [9] 李庆, 金润浩, 姜国哲, 等. 2006—2015 年我国已注册保健食品现状分析 [J]. 食品科学, 2017, 38(3): 292-298.
 - [10] 郭洁, 贾伯阳, 张蓉, 等. 国产保健食品原料与功效/标志性成分分析 [J]. 食品研究与开发, 2018, 39(24): 218-224.
 - [11] 于晓娜, 崔波, 任贵兴. 西洋参多糖的研究进展 [J]. 食品科学, 2014, 35(9): 301-305.
 - [12] 林红强, 李平亚, 刘金平. 野生西洋参鉴别、化学成分及药理作用研究进展 [J]. 天然产物研究与开发, 2017, 29(12): 2157-2162.
 - [13] 中国药典 [S]. 一部. 2020.
 - [14] 卫生部关于进一步规范保健食品原料管理的通知 [EB/OL]. (2002-03-11) [2021-07-01]. <http://www.nhcc.gov.cn/zwgk/wtwj/201304/e33435ce0d894051b15490aa3219cdc4.shtm>.
 - [15] 王建, 张冰. 临床中药学 [M]. 第 2 版. 北京: 人民卫生出版社, 2016.
 - [16] 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.
 - [17] 薛宁, 郭桓博, 马瑞雪, 等. 基于数据挖掘的中医药治疗消化性溃疡用药特点研究 [J]. 中草药, 2022, 53(3): 799-805.
 - [18] 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.
 - [19] 市场监管总局关于公开征求《关于发布允许保健食品声称的保健功能目录非营养补充剂 (2022 年版) 及配套文件的公告(征求意见稿)》意见的公告 [EB/OL]. (2022-01-13) [2022-01-28]. https://www.samr.gov.cn/hd/zjdc/202201/t20220113_339092.html.
 - [20] Mikolov T, Zweig G. Context dependent recurrent neural network language model [A] // 2012 IEEE Spoken Language Technology Workshop (SLT) [C]. Miami: IEEE, 2012: 234-239.
 - [21] 王飞, 谭新. 一种基于 Word2Vec 的训练效果优化策略研究 [J]. 计算机应用与软件, 2018, 35(1): 97-102.
 - [22] 高涛, 唐华丽, 孙桂菊, 等. 保健食品产业中存在的问题及对策分析 [J]. 食品与发酵工业, 2021, 47(2): 311-321.
 - [23] 付双楠, 朱平生, 高达. 中医药防治化学性肝损伤的研究进展 [J]. 中医学报, 2017, 32(3): 449-454.
 - [24] 姚文博, 李丰, 石彬彬, 等. 系统药理学III: 在诠释中医药整体作用机制中的应用进展 [J]. 中国实验方剂学杂志, 2020, 26(13): 219-227.
 - [25] 高婉卿, 程宁, 李力松, 等. 机器学习在方剂研究中的应用概述 [J]. 中国中医药信息杂志, 2021, 28(6): 131-137.
 - [26] 林九国, 史梦瑶, 金磊, 等. 基于数据挖掘的中药治疗血液病的用药规律分析 [J]. 中草药, 2021, 52(4): 1073-1082.
 - [27] 刘剑刚. 西洋参三七配伍改善双抗至急性心肌梗死后大鼠胃黏膜损伤的实验研究 [A] // 第十五届中国中西医结合学会微循环专业委员会暨第二届中国微循环学会瘀瘀专业委员会学术会议资料汇编 [C]. 泸州: 中国中西医结合学会微循环专业委员会, 中国中西医结合学会, 2015: 1-5.
 - [28] 卢泽源, 王志斌, 张玉薇, 等. 西洋参五味子口服液对小鼠免疫功能的影响 [J]. 人参研究, 2016, 28(4): 6-9.
 - [29] 张力, 祁伟, 苏国辉. 清肝明目, 理气通心: 枸杞提取物功效开发的现状和展望 [J/OL]. 科学通报, [2022-01-27]. <http://kns.cnki.net/kcms/detail/11.1784.N.20220120.1009.002.html>.
 - [30] 黄家林, 田代雄. 三七总皂苷抗炎免疫药理研究进展 [J]. 中华中医药杂志, 2016, 31(11): 4657-4660.
 - [31] 吕婧, 高燕, 李晨, 等. 基于斑马鱼模式生物的西洋参皂苷类成分增强免疫作用研究 [J]. 中草药, 2020, 51(14): 3728-3733.
 - [32] Yoshikawa M, Murakami T, Yashiro K, et al. Bioactive saponins and glycosides. XI. structures of new dammarane-type triterpene oligoglycosides, quinquenosides I, II, III IV, and V, from American ginseng, the roots of *Panax quinquefolium* L. [J]. Chem Pharm Bull, 1998, 46(4): 647-654.
 - [33] Lai Y Q, Tan Q X, Xv S, et al. Ginsenoside Rb1 alleviates alcohol-induced liver injury by inhibiting steatosis, oxidative stress, and inflammation [J]. Front Pharmacol, 2021, 12: 616409.
 - [34] 赵月蓉, 侯碧玉, 张莉, 等. 葛根素对实验性肝损伤的治疗作用研究进展 [J]. 中国新药杂志, 2017, 26(9): 1005-1010.
 - [35] 任伟光, 张翠英. 五味子的研究进展及质量标志物 (Q-marker) 的预测分析 [J]. 中草药, 2020, 51(11): 3110-3116.

[责任编辑 潘明佳]