

· 数据挖掘 ·

基于数据挖掘技术探析中医药治疗中晚期非小细胞肺癌的 用药规律

邓正亭^{1,2}, 赵凡¹, 赵彤¹, 吕灿³, 林玩福³, 梁淑芳³, 刘涛^{1*}, 程彬彬^{3*}

(1. 南京中医药大学中医学学院·中西医结合学院, 南京 210023;

2. 南京中医药大学附属泰州市中医院, 江苏泰州 225300;

3. 海军军医大学第一附属医院, 上海 200433)

[摘要] 目的: 挖掘中医药治疗非小细胞肺癌(NSCLC)的成方规律, 以为临床遣方用药及新药研发提供一定的理论依据。**方法:** 以中国知网(CNKI)中近40年来中医药治疗中晚期NSCLC的临床研究文献作为数据来源, 通过筛选后建立方剂数据库, 再运用数据挖掘软件进行频数统计、关联规则分析、系统聚类、因子分析及复杂网络构建等。**结果:** 经过纳排标准筛选出307篇临床文献, 包含方剂483首。中晚期NSCLC常见证候以气阴两虚为主, 证候要素以气虚、阴虚、痰、血瘀、热(火)邪、毒邪、湿邪多见。使用的药物类别以补虚药、清热药、化痰止咳平喘药、利水渗湿药和活血化瘀药较为集中。使用高频次药物为黄芪、甘草、麦冬、浙贝母、茯苓等, 其中四气以寒性最为常见, 五味以苦、甘、辛味居多, 归经主要集中在肺、脾、胃经。关联规则分析发现关联强度较高的规则有17条, 因子分析中可提取公因子数为10个, 系统聚类分析得出5大类。复杂网络分析提示核心中药处方为六君子汤合益气养阴解毒方加减。**结论:** 益气养阴、健脾化痰、清热解毒、活血化瘀是中医药治疗中晚期NSCLC的基本治则, 数据挖掘的核心组合和新方具有重要的指导借鉴意义, 但仍需与中医理论相结合, 有待进一步的临床及实验研究给予验证。

[关键词] 非小细胞肺癌; 数据挖掘; 中医药; 用药规律

[中图分类号] R285; R289; R22; R2-031; R33 **[文献标识码]** A **[文章编号]** 1005-9903(2022)03-0171-09

[doi] 10.13422/j.cnki.syfjx.20220114

[网络出版地址] <https://kns.cnki.net/kcms/detail/11.3495.R.20211115.1851.002.html>

[网络出版日期] 2021-11-16 11:19

Analysis on Medication Regularity of Traditional Chinese Medicine in Treatment of Intermediate and Advanced Non-small Cell Lung Cancer Based on Data Mining

DENG Zheng-ting^{1,2}, ZHAO Fan¹, ZHAO Tong¹, LYU Can³, LIN Wan-fu³, LIANG Shu-fang³,
LIU Tao^{1*}, CHENG Bin-bin^{3*}

(1. School of Traditional Chinese Medicine (TCM) · School of Integrated Chinese and Western Medicine, Nanjing University of Chinese Medicine, Nanjing 210023, China;

2. Taizhou Hospital of TCM, Taizhou 225300, China; 3. Changhai Hospital, Shanghai 200433, China)

[Abstract] **Objective:** To explore medication regularity of traditional Chinese medicine (TCM) in the treatment of non-small cell lung cancer (NSCLC) and thereby to lay a theoretical basis for clinical medication and drug development. **Method:** Articles on clinical treatment of intermediate and advanced NSCLC with TCM

[收稿日期] 20210828(011)

[基金项目] 国家自然科学基金项目(82074203);“十三五”军队重点学科专业建设项目(2020SZ21-10);国家“十一五”科技支撑计划“名老中医临床经验、学术思想传承研究”项目(2007BA110B01-033)

[第一作者] 邓正亭, 博士, 主治医师, 从事呼吸系统疾病的中西医结合诊治, E-mail: 202050002@njucm.edu.cn

[通信作者] * 刘涛, 教授, 博士生导师, 从事温病理论与运用研究, E-mail: 13952091103@139.com;

* 程彬彬, 博士, 副教授, 从事中西医结合治疗肝癌、胃癌、肠癌等消化系统恶性肿瘤, E-mail: cbb8202@126.com

in the past 40 years were retrieved from CNKI, which were taken the data source. Then the articles were screened to establish a formula database, followed by frequency statistics, association rule analysis, cluster analysis, factor analysis, and complex network construction. **Result:** A total of 307 eligible articles were screened out, involving 483 formulas. The common syndrome of intermediate and advanced NSCLC was the deficiency of both Qi and Yin, with the common syndrome elements of Qi deficiency, Yin deficiency, phlegm, blood stasis, pathogenic heat (fire), toxin, and pathogenic dampness. The frequently used medicinals mainly had the functions of tonifying deficiency, clearing heat, resolving phlegm and relieving cough and dyspnea, promoting urination and draining dampness, and activating blood and resolving stasis. The high-frequency medicinals were Astragali Radix, Glycyrrhizae Radix et Rhizome, Ophiopogonis Radix, Fritillariae Thunbergii Bulbus, and Poria, which were mainly cold, bitter, sweet, and pungent, with tropism at lung, spleen, and stomach. The association rule analysis yielded 17 rules with strong association. Ten common factors were extracted from the factor analysis, and cluster analysis classified the medicinals into 5 groups. Complex network analysis suggested that the core formula was modified LiuJunzi Tang and Yiqi Yangyin Jiedu prescription. **Conclusion:** The treatment principle for intermediate and advanced NSCLC is replenishing Qi and nourishing Yin, invigorating spleen and resolving phlegm, clearing heat and detoxifying, promoting blood circulation and removing blood stasis. The core combinations new prescription discovered by data mining are of important guiding significance, but they should be further verified in clinical practice and by experiments based on the theory of TCM.

[Keywords] non-small cell lung cancer; data mining; traditional Chinese medicine; medication regularity

肺癌是我国发病率和死亡率均居首位的恶性肿瘤,非小细胞肺癌(NSCLC)是其最常见的病理类型,约占85%^[1-2]。而大多肺癌患者一经确诊已发展至中晚期,5年生存率较低,总体预后较差^[3]。从化疗到分子靶向治疗、抗肿瘤血管生成治疗以及逐渐兴起的肿瘤免疫治疗,中晚期NSCLC的西医治疗理念在不断地发生变化,患者总体生存时间也有所延长。然而这些治疗手段依然面临着费效比高、副作用多、远期疗效欠理想等问题,临床上越来越多的患者对更高的生活质量、更长的生存预期诉求迫切。中医药具有个体化、多靶点、多效应、低毒性等优势,近年来在肺癌防治方面中发挥了重要的作用。大量临床研究表明,中医药结合西医治疗肺癌的模式在减轻不良反应、调节身体机能、改善生存质量等方面显示出独特而明显的临床获益^[4-5]。

中医古籍并无肺癌病名的记载,依据肺癌的临床表现及特点,多将其归属于“咳嗽”“咯血”“肺积”“息贲”“肺萎”“肺岩”等病证范畴。此外,类似肺癌晚期症状及预后的论述也散见于历代医家著作中,如《景岳全书·积聚论治》记录:“劳嗽,声哑,声不能出或喘息气促者,此肺脏积也,必死。”现代医家结合长期临床实践对肺癌的病因病机进行了深入的理论探究,周仲瑛教授倡导癌毒学说,指出肺癌的

基本病机是痰瘀郁肺,毒耗气阴,痰瘀郁毒为其主要的病理因素^[6]。周岱翰教授认为,肺癌为病,亦多由痰,以治痰为核心^[7]。刘嘉湘教授推崇“正气虚则成岩”,首倡扶正治癌学术思想,提出正气亏虚为肺癌发病之本,继而邪毒乘虚而入,致肺气膈郁,宣降失司,气机不利,血行受阻,津液失于输布,聚而为痰,痰凝气滞,瘀阻络脉,痰气毒瘀胶结,日久形成肺部积块^[8]。另有医家从络病、伏毒、阳虚等诸多方面对肺癌病因病机展开论述^[9-10]。虽各家观点不同,但普遍认为肺癌是一种整体为虚、局部属实的疾病,正虚与痰、气、热(火)、毒、瘀等病理因素相互搏结,临床也多以此立论,辨证施治。

由此可见,医家们遣方用药复杂,中医药治疗肺癌的数据繁多,后世学者着实难以借鉴。而这些庞大的中医药数据中蕴藏着巨大的临床价值,因此,高效客观地整合数据、剖析其潜在价值是一项繁重且有意义的工作。鉴于近年来数据挖掘技术在处方配伍规律研究中的应用已比较成熟,包含发现核心药对或药组、方剂效用归纳、发现新处方、量效关系研究等诸多方面^[11]。目前,数据挖掘技术多集中来源于医药典籍和当代名老中医诊治肺癌的用药处方,本研究引入数据挖掘技术着重对中医药治疗中晚期NSCLC的临床研究文献进行系统性分

析,从临床研究的角度归纳分析肺癌的辨证分型,为证候规范化研究及疗效评价体系建立提供一定的参考。此外还探究中药药性以揭示其现代科学意义,最终综合运用多种数据挖掘算法进行统计分析,探索用药规律,旨在为临床应用提供有益的数据支撑和诊疗思路。

1 资料与方法

1.1 数据来源 本研究的文献数据库来源于中国知网(CNKI),检索策略使用主题词检索法,其中主题词“肺癌”OR“肺肿瘤”,副主题词“中医”OR“中药”,两部分主题词之间以“AND”连接。检索年限限定为1981年01月01日至2020年12月31日,检索日期为2021年03月24日,初步获得近40年有关中医药治疗中晚期NSCLC的文献5330篇。

1.2 纳入标准 ①研究文献类型为采用中医药治疗为主、可合并西医常规肺癌治疗方案的临床研究;②研究对象为明确诊断“中晚期NSCLC”的患者;③中医药辨治效果确切:WHO实体瘤疗效评价、卡氏功能状态评分、中医证候临床疗效评价三者中至少有一项优于对照组;④文献所载方剂为具有明确辨证分型且药物组成及剂量完整的中药复方;⑤同一文献中与证型相应的方药只纳入基本方。

1.3 排除标准 ①重复发表或处方相同的文献仅纳入1次;②干预措施为中药注射剂、灌肠、贴敷、针灸等非口服汤剂;③治疗肺癌患者因放化疗、靶向治疗、免疫治疗引起不良反应的文献;④治疗肺癌患者伴发其他严重原发疾病或并发症的文献。

1.4 数据处理 经过纳排标准筛选出307篇临床文献,包含方剂483首。依据《中华人民共和国药典》(2020年版)^[12]、《中药大辞典》(第2版)^[13]和《中药学》^[14]对中药名称、分类、性味归经等进行统一规范,药名如七叶一枝花、白蚤休、蚤休统一为重楼,生天南星、生南星、制南星、蛇六谷统一为天南星;药味如味酸甘拆解为味酸、味甘;归经如归肺脾胃经拆分为归肺经、脾经、胃经;分类如将黄芪、白术、党参归为补气药,麦冬、女贞子、百合归为补阴药等。再参照《中华人民共和国国家标准·中医临床诊疗术语·证候部分》^[15]对文献中辨证分型的证型名称进行规范,未涉及者按《中医诊断学》^[16]和《中医证候鉴别诊断学》^[17]进行规范,如若对于上述标准中未描述而不便归类的证型,遵照原文献予以保留。此外,通过参考朱文锋等^[18]关于证候要素的文献,本研究进一步提取了病位和病性证候要素并对

其进行规范化处理。最终将清洗处理后的数据全部录入Microsoft Excel 2016,建立“中医药治疗NSCLC方剂数据库”。

1.5 统计分析 本研究采用Microsoft Excel 2016, IBM SPSS Modeler 18.0, IBM SPSS Statistics 26.0, Gephi 0.9.2等软件进行频数统计、关联规则分析、聚类及因子分析、复杂网络分析,多种数据挖掘方法联合运用,对方剂数据库中的病机、药物、配伍关系等进行全面探索研究。

2 结果

2.1 证型统计 经证型规范及去重后得到96个证型,其中单证12种,复证84种。以气阴两虚、脾虚痰湿、气滞血瘀、阴虚内热证型为主,共占44.93%。筛选频数≥10次的的数据,见表1。

表1 CNKI数据库NSCLC相关证型分布

Table 1 Frequency of syndromes distribution more than ten times of NSCLC in CNKI database

| 证型 | 频数/次 | 频率/% | 证型 | 频数/次 | 频率/% |
|------|------|-------|-----------|------|------|
| 气阴两虚 | 139 | 28.78 | 气虚血瘀 | 15 | 3.11 |
| 脾虚痰湿 | 36 | 7.45 | 气虚痰湿 | 13 | 2.69 |
| 气滞血瘀 | 22 | 4.55 | 气阴两虚 瘀毒互结 | 12 | 2.48 |
| 阴虚内热 | 20 | 4.14 | 气虚 | 11 | 2.28 |
| 脾肺气虚 | 17 | 3.52 | 肺郁痰瘀 | 10 | 2.07 |
| 阴虚痰热 | 17 | 3.52 | 阴虚热毒 | 10 | 2.07 |

2.2 证素分析 从483个临床样本中共提取病位证素7个,病性证素15个。病位证素中肺占比例最高,为43.81%,其次是脾,占35.71%,体现了本病病位在肺,而与脾关系密切,见表2。病性证素频数分布,出现频率为5%以上的证候要素是气虚、阴虚、痰、血瘀、热(火)邪、毒邪、湿邪,其中又以气虚、阴虚所占比例最大;实性证候要素10个,占比例较高的是痰、血瘀;虚性证候要素5个,以气虚和阴虚占比例最高。结果提示中晚期NSCLC证候主要为虚实夹杂,其中虚又以气虚、阴虚为主,邪实以痰、血瘀为主。见表3。

进一步分析15个病性证候要素中两证素同时出现的频率,提取前20位复合证素,结果显示气虚+阴虚同时出现的频率较高,达38.51%,其次分别为气虚+痰、湿邪+痰、气虚+血瘀、痰+血瘀、热(火)邪+阴虚、气虚+湿邪、毒邪+气虚等,体现了本病气阴两虚,痰瘀湿热毒等互结,正虚与邪实相互交杂的复杂病机,结果见表4。

表2 CNKI数据库治疗NSCLC病位证素分布

Table 2 Essential elements distribution of the location of treat NSCLC in CNKI database

| 病位证素 | 频数/次 | 频率/% |
|------|------|-------|
| 肺 | 400 | 43.81 |
| 脾 | 326 | 35.71 |
| 心 | 120 | 13.14 |
| 肝 | 31 | 3.40 |
| 肾 | 29 | 3.18 |
| 经络 | 4 | 0.44 |
| 胃 | 3 | 0.33 |

表3 CNKI数据库治疗NSCLC病性证素分布

Table 3 Essential elements distribution of nature of treat NSCLC in CNKI database

| 病性证素 | 频数/次 | 频率/% |
|-------|------|-------|
| 气虚 | 316 | 28.21 |
| 阴虚 | 252 | 22.50 |
| 痰 | 152 | 13.57 |
| 血瘀 | 121 | 10.80 |
| 热(火)邪 | 80 | 7.14 |
| 毒邪 | 66 | 5.89 |
| 湿邪 | 66 | 5.89 |
| 气滞 | 35 | 3.13 |
| 阳虚 | 19 | 1.70 |
| 血虚 | 7 | 0.63 |
| 精亏 | 2 | 0.18 |
| 寒邪 | 1 | 0.09 |
| 气逆 | 1 | 0.09 |
| 食积 | 1 | 0.09 |
| 燥邪 | 1 | 0.09 |

注:气逆为尚未形成共识的病性证素,本研究暂将其纳入。

2.3 药物归类及频数分布 数据库共提取278味中药,总频数达6162次,根据传统药物功效归为19类,其中前5位为补虚药、清热药、化痰止咳平喘药、利水渗湿药和活血化瘀药,频率分别为35.10%, 19.15%, 15.63%, 7.06%和6.57%,累计频率达到了83.51%。可见临证治疗以益气养阴、清热解毒、化痰祛瘀为主,此类药物使用较集中。见表5。

2.4 高频药物分析

2.4.1 高频单味药 本研究中,使用频数≥70次的中药共24种,累计频数3344次,占药物总数54.27%,规定其为高频药物^[19]。其中频率居前5位药物依次是黄芪(282次,8.43%),甘草(231次,6.91%),麦冬(226次,6.76%),浙贝母(214次,

表4 CNKI数据库治疗NSCLC复合病性证素分布

Table 4 Simultaneous and essential elements distribution of the nature of treat NSCLC in CNKI database

| 复合病性证素 | 频数/次 | 频率/% | 复合病性证素 | 频数/次 | 频率/% |
|----------|------|-------|----------|------|------|
| 气虚+阴虚 | 186 | 38.51 | 痰+阴虚 | 36 | 7.45 |
| 气虚+痰 | 81 | 16.77 | 气滞+血瘀 | 34 | 7.04 |
| 湿邪+痰 | 65 | 13.46 | 毒邪+血瘀 | 31 | 6.42 |
| 气虚+血瘀 | 56 | 11.59 | 血瘀+阴虚 | 30 | 6.21 |
| 痰+血瘀 | 56 | 11.59 | 毒邪+热(火)邪 | 26 | 5.38 |
| 热(火)邪+阴虚 | 52 | 10.77 | 毒邪+痰 | 16 | 3.31 |
| 气虚+湿邪 | 50 | 10.35 | 气滞+痰 | 11 | 2.28 |
| 毒邪+气虚 | 40 | 8.28 | 热(火)邪+血瘀 | 11 | 2.28 |
| 毒邪+阴虚 | 38 | 7.87 | 气虚+阳虚 | 10 | 2.07 |
| 热(火)邪+痰 | 37 | 7.66 | 气虚+热(火)邪 | 10 | 2.07 |

6.40%)和茯苓(209次,6.25%);其次为白花蛇舌草、白术、半夏、党参、薏苡仁、半枝莲等。使用频数≥70次的单味中药见表6。

2.4.2 性味与归经 对24味高频药物的性味归经进行统计分析,若同一味中药具有不同药味及归经则分别统计,性味中微苦、微辛、微温、微寒分别归入苦、辛、温、寒。结果显示,累计出现3344次药性,寒(1268次,37.92%)最为常见,四气频数统计见图1(A);累计出现5487次药味,以苦、甘、辛味药物居多,其中苦(2029次,36.98%),甘(1989次,36.25%),辛(1079次,19.66%),五味频数统计见图1(B);归经频率前5位分别是肺经(30.13%),脾经(21.76%),胃经(17.60%),心经(11.50%),肝经(6.61%),归经频数统计见增强出版附件材料。

2.4.3 关联规则分析 对高频药物采用SPSS Modeler 18.0软件的Apriori算法进行关联规则分析,进一步探究药物之间的配伍关系。设定支持度≥10%,置信度≥90%,最大前项数为5,得出关联规则17条(按置信度百分比降序排列),见表7;高频药物之间关联网络展示,见增强出版附件材料。

2.4.4 因子分析 为了挖掘高频中药间潜在的结构关系,本研究运用SPSS Statistics 26.0统计软件进行因子分析,结果显示KMO=0.618, Bartlett球形度检验 $P=0$,认为变量间存在较强的依赖性,符合因子分析标准。采用主成分分析法,选取凯撒正态化最大方差法旋转得到旋转成分矩阵成分图,见图1。提取初始特征值>1时,用药累计贡献率为62.21%,设置载荷系数最小绝对值为0.4,可提取公因子数为10个,见表8。

表5 CNKI数据库治疗NSCLC中药类别及频数分布

Table 5 Category and frequency distribution of treat NSCLC traditional Chinese medicine in CNKI database

| 编号 | 总类别(药味数) | 分类别 | | | 编号 | 总类别(药味数) | 分类别 | | | | | | | | | |
|----|-------------|-------|-------|---------|-------|----------|---------|----------|-------|------|------------|-------|-------|------|-------|----|
| | | 名称 | 频数/次 | 频率/% | | | 名称 | 频数/次 | 频率/% | | | | | | | |
| 1 | 补虚药(51) | 补气药 | 1 157 | 2 163 | 8 | 消食药(7) | 凉血止血药 | 13 | 140 | 2.27 | | | | | | |
| | | 补阴药 | 799 | 19.15 | | | 温经止血药 | 3 | | | | | | | | |
| | | 补血药 | 139 | | | | 9 | 9 | | | 攻毒杀虫止痒药(4) | 115 | 1.87 | | | |
| | | 补阳药 | 68 | | | | 10 | 平肝息风药(8) | | | 息风止痉药 | 85 | 112 | 1.82 | | |
| 2 | 清热药(59) | 清热解毒药 | 825 | | 15.63 | 11 | 收敛药(8) | 平抑肝阳药 | 27 | 64 | 95 | 1.54 | | | | |
| | | 清热泻火药 | 157 | 敛肺涩肠药 | | | | 64 | | | | | | | | |
| | | 清热凉血药 | 138 | 固精缩尿止带药 | | | | 30 | | | | | | | | |
| | | 清热燥湿药 | 51 | 固表止汗药 | | | | 1 | | | | | | | | |
| | | 清虚热药 | 9 | 12 | | | | 解表药(16) | 发散风热药 | | | | 49 | 87 | 1.41 | |
| 3 | 化痰止咳平喘药(32) | 清化热痰药 | 464 | 963 | 13 | 化湿药(6) | 发散风寒药 | 38 | 38 | 0.62 | | | | | | |
| | | 温化寒痰药 | 335 | | | | 14 | 安神药(1) | | | 养心安神药 | 27 | 27 | 0.44 | | |
| | | 止咳平喘药 | 164 | | | | 15 | 祛风湿药(11) | | | 祛风寒湿药 | 14 | 24 | 0.39 | | |
| 4 | 利水渗湿药(12) | 利水消肿药 | 408 | 435 | 14 | 祛风湿热药 | 祛风湿强筋骨药 | 1 | 11 | 12 | 0.19 | | | | | |
| | | 利湿退黄药 | 14 | | | | 16 | 泻下药(4) | | | | 攻下药 | 11 | | | |
| | | 利尿通淋药 | 13 | | | | 17 | 温里药(4) | | | | 峻下逐水药 | 1 | | | |
| 5 | 活血化瘀药(27) | 活血调经药 | 159 | 405 | 18 | 开窍药(1) | 开窍药 | 2 | 11 | 0.18 | | | | | | |
| | | 破血消癥药 | 120 | | | | 3.33 | 19 | | | 驱虫药(1) | 驱虫药 | 1 | 0.02 | | |
| | | 活血止痛药 | 111 | | | | | | | | | 2.39 | 化痰止血药 | | 化痰止血药 | 52 |
| | | 活血疗伤药 | 15 | | | | | | | | | | | | | |

表6 治疗NSCLC使用频数≥70的单味中药

Table 6 Single traditional Chinese medicine with frequency more than seventy times of treat NSCLC

| 药物 | 频数/次 | 频率/% | 药物 | 频数/次 | 频率/% |
|-------|------|------|-----|------|------|
| 黄芪 | 282 | 8.43 | 山慈菇 | 110 | 3.29 |
| 甘草 | 231 | 6.91 | 太子参 | 110 | 3.29 |
| 麦冬 | 226 | 6.76 | 桔梗 | 101 | 3.02 |
| 浙贝母 | 214 | 6.40 | 北沙参 | 95 | 2.84 |
| 茯苓 | 209 | 6.25 | 陈皮 | 92 | 2.75 |
| 白花蛇舌草 | 202 | 6.04 | 莪术 | 87 | 2.60 |
| 白术 | 201 | 6.01 | 鱼腥草 | 85 | 2.54 |
| 半夏 | 176 | 5.26 | 女贞子 | 84 | 2.51 |
| 党参 | 160 | 4.78 | 仙鹤草 | 78 | 2.33 |
| 薏苡仁 | 155 | 4.64 | 天冬 | 74 | 2.21 |
| 半枝莲 | 120 | 3.59 | 夏枯草 | 72 | 2.15 |
| 南沙参 | 110 | 3.29 | 瓜蒌 | 70 | 2.09 |

2.4.5 系统聚类 采用SPSS Statistics 26.0软件对上述高频药物进行系统聚类分析,选取组间联接的

聚类方法,分析得出5类。类1:茯苓、白术、党参、半夏、陈皮;类2:生地黄、甘草、当归、桔梗;类3:黄芪、女贞子、白花蛇舌草、半枝莲、莪术、麦冬、北沙参、天冬、南沙参;类4:太子参、丹参;类5:夏枯草、瓜蒌、仙鹤草、壁虎、鱼腥草、浙贝母、薏苡仁、山慈菇。见表9。

2.5 复杂网络分析 采用Gephi 0.9.2软件构建数据库药物-药物及药物-证素的复杂网络,分析范围为483个临床样本涉及的278味中药和22个证素,在置信度>40%,支持度>5%的条件下进行关联分析,并按照特征向量中心度显示节点大小,依据权重设置边的粗细,导出复杂网络图,见图2。复杂网络静态指标见表10。

3 讨论

证型统计分析显示,中晚期NSCLC证候分布复杂,复证为主,以两证组合多见。频数前5依次为气阴两虚、脾虚痰湿、气滞血瘀、阴虚内热、脾肺气

表7 治疗NSCLC高频药物之间的关联规则

Table 7 Association rules analysis between high-frequency drugs of treat NSCLC

| 后项 | 前项 | 支持度百分比/% | 置信度百分比/% | 增益 |
|----|--------------|----------|----------|-------|
| 黄芪 | 女贞子+白术 | 12.008 | 100.0 | 1.719 |
| 黄芪 | 白花蛇舌草+白术+麦冬 | 11.594 | 98.214 | 1.688 |
| 黄芪 | 女贞子+茯苓 | 10.973 | 98.113 | 1.686 |
| 黄芪 | 白花蛇舌草+茯苓+麦冬 | 12.836 | 96.774 | 1.663 |
| 黄芪 | 浙贝母+白术+麦冬 | 11.801 | 96.491 | 1.659 |
| 黄芪 | 女贞子+麦冬 | 10.559 | 96.078 | 1.651 |
| 黄芪 | 女贞子+白花蛇舌草 | 12.008 | 94.828 | 1.630 |
| 黄芪 | 白术+麦冬 | 19.462 | 94.681 | 1.627 |
| 黄芪 | 白术+茯苓+麦冬 | 15.321 | 94.595 | 1.626 |
| 黄芪 | 白术+甘草+麦冬 | 11.387 | 94.545 | 1.625 |
| 黄芪 | 浙贝母+茯苓+麦冬 | 10.973 | 94.34 | 1.622 |
| 黄芪 | 女贞子 | 17.391 | 94.048 | 1.617 |
| 白术 | 浙贝母+茯苓+麦冬+黄芪 | 10.352 | 92.0 | 2.211 |
| 白术 | 薏苡仁+党参+茯苓 | 10.145 | 91.837 | 2.207 |
| 黄芪 | 薏苡仁+白花蛇舌草+白术 | 10.145 | 91.837 | 1.579 |
| 黄芪 | 白花蛇舌草+白术 | 20.497 | 90.909 | 1.563 |
| 黄芪 | 薏苡仁+白花蛇舌草+茯苓 | 10.559 | 90.196 | 1.55 |

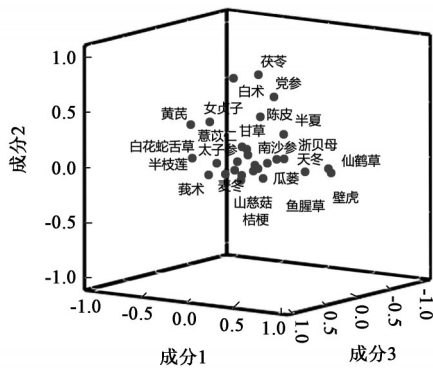


图1 治疗NSCLC高频药物旋转因子空间成分
Fig. 1 Component plot in rotated space of high-frequency drugs of treat NSCLC

虚,该结果与李萍等^[20]研究结论相近。近年来虽有学者在肺癌证候研究等方面做了大量的工作,但目前仍尚无统一的辨证分型标准。学者们大多从肺癌病理类型、临床分期、治疗手段、生活质量、流行病学调查等不同层次探讨肺癌中医证型的分布规律,以期建立统一的证候标准^[21]。笔者认为肺癌首先是一种异质性疾病^[22],中晚期肺癌的病程也是一个极为复杂的动态过程,患者经不同治疗手段后所表现出的临床症状及体征更是存在着一定的差异。因此,若仅总结若干个常见的固定证型来作为分型

表8 治疗NSCLC高频药物公因子信息

Table 8 Common factor information table of high-frequency drugs of treat NSCLC

| 公因子 | 因子贡献率/% | 药物 |
|-----|---------|---------------------|
| F1 | 10.817 | 黄芪、鱼腥草、仙鹤草、壁虎 |
| F2 | 9.198 | 茯苓、白术、党参 |
| F3 | 8.201 | 白花蛇舌草、半夏、半枝莲、莪术、夏枯草 |
| F4 | 6.191 | 浙贝母、薏苡仁、山慈菇 |
| F5 | 5.559 | 北沙参、天冬 |
| F6 | 5.042 | 甘草、陈皮 |
| F7 | 4.972 | 麦冬、南沙参 |
| F8 | 4.383 | 生地黄、当归 |
| F9 | 4.060 | 瓜蒌 |
| F10 | 3.786 | 桔梗、丹参 |

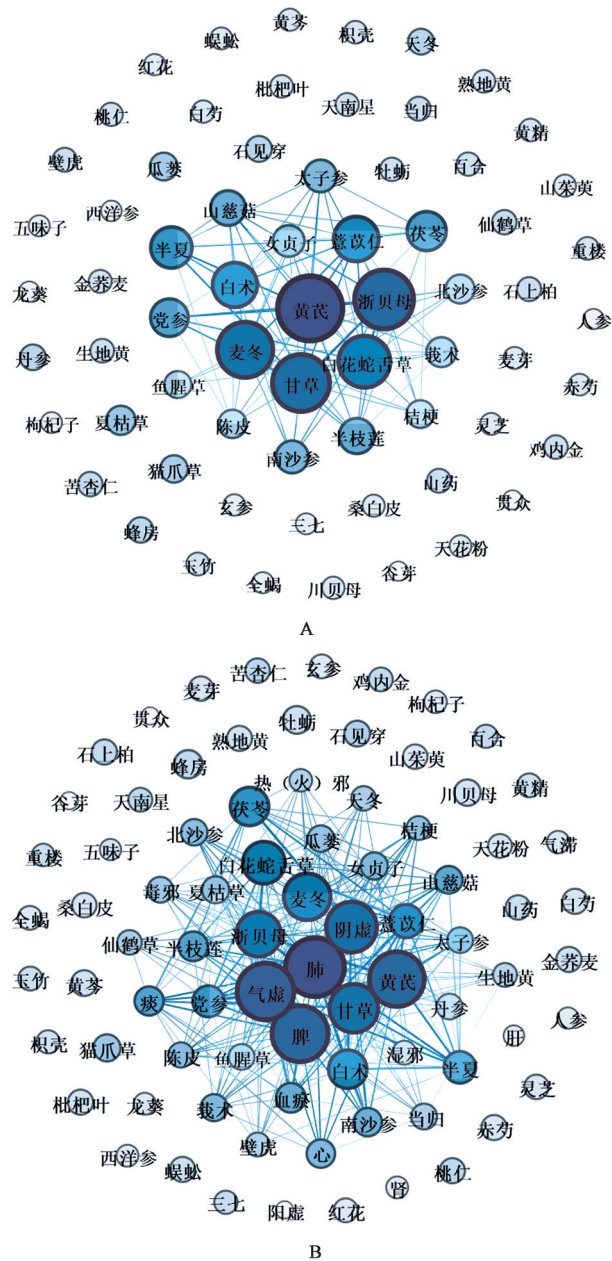
表9 治疗NSCLC高频药物聚类分析树状图药物组成

Table 9 Cluster analysis of high-frequency drugs extract results in dendritic chart of treat NSCLC

| 类别 | 药物 |
|----|-----------------------------------|
| C1 | 茯苓、白术、党参、半夏、陈皮 |
| C2 | 生地黄、甘草、当归、桔梗 |
| C3 | 黄芪、女贞子、白花蛇舌草、半枝莲、莪术、麦冬、北沙参、天冬、南沙参 |
| C4 | 太子参、丹参 |
| C5 | 夏枯草、瓜蒌、仙鹤草、壁虎、鱼腥草、浙贝母、薏苡仁、山慈菇 |

标准既不能为肺癌证候规范化提供足够的客观化证据,也不符合临床和科研的实际需求。部分学者认为证素研究更适应肺癌临证实践,具有可操作性,也便于量化规范肺癌复杂动态多变的证型^[23-24]。笔者据此从文献研究的角度,以证素组合为切入点,结果显示与中晚期NSCLC最密切相关的脏腑是肺和脾,气虚、阴虚为基本的虚性证素,同时夹杂痰、血瘀、热(火)邪、毒邪、湿邪等实性证素。此外,针对中晚期NSCLC中医证型的规范化,今后有必要开展多中心、大样本的临床研究,有益于揭示证候规律及统一分型标准,并为肺癌的临床辨治提供执简驭繁的纲领依据。

药物归类结果显示,数据库处方共涉及中药278味,功效分类主要为补虚药,包括黄芪、甘草、白术、党参等补气药类,麦冬、南沙参、北沙参、女贞子等补阴药类,当归、熟地黄、白芍等补血药类;其次为清热药,如白花蛇舌草、半枝莲、山慈菇等清热解毒药类,夏枯草、天花粉等清热泻火药类等;此外还有化痰止咳平喘药、利水渗湿药及活血化瘀药等。



A. 药物-药物,边权重>80;B. 药物-证素,边权重>60
图2 治疗NSCLC药物-药物及药物-证素复杂网络

Fig. 2 Drug-drug and drug-essential elements complex network diagram of treat NSCLC

表10 智力NSCLC药物-药物及药物-证素复杂网络静态指标
Table 10 Static index of reat NSCLC drug-drug and drug-essential elements complex network

| 编号 | 网络特征 | 药物-药物网络 | 药物-证素网络 |
|----|--------|---------|---------|
| 1 | 节点数 | 67 | 81 |
| 2 | 边数 | 391 | 719 |
| 3 | 网络密度 | 0.177 | 0.222 |
| 4 | 平均路径长度 | 1.901 | 1.845 |
| 5 | 平均聚类系数 | 0.391 | 0.349 |

由此可见,临证遣方常用益气养阴、清热解毒、化痰祛瘀等为主,此类药物使用较集中,这也侧面与肺癌的“虚痰热毒瘀”的病机特点相符。

用药频次分析发现,使用频数最高的前10味药为黄芪、甘草、麦冬、浙贝母、茯苓、白花蛇舌草、白术、半夏、党参、薏苡仁。黄芪作为用药频次最高的单味药物,味甘,性微温,归肺、脾经,具有补气升阳、固表止汗、利水消肿、生津养血、行滞通痹、托毒排脓、敛疮生肌之效^[7]。现代药理学研究表明,黄芪可发挥促进肺癌细胞凋亡、抑制细胞增殖和迁移、调节免疫和逆转耐药等作用^[25-27]。高频药物药性以寒为主,约占38%。张景岳认为“寒热乃阴阳之化也”,寒热是八纲辨证中的纲领,虽中晚期NSCLC患者症状表现各异,但总体均可以寒热分型辨证。吸烟和被动吸烟是目前公认的肺癌最重要的危险因素,而中医理论认为烟为辛热之魁,加之放疗、化疗等治疗手段也具“火热”之性等,故用药方面宜选相对寒性之药^[28]。但临床实际中更应重视寒热转化,随证施治,调和阴阳,方有桴鼓相应之效。高频药物药味以苦、甘、辛使用居多,而中晚期NSCLC病机常表现出升降失调寒热虚实错杂的矛盾性,这也提示取辛开苦降甘补调和方法往往可以缓解临床症状,改善机体状态等^[29]。高频药物归经主要为肺、脾、胃经,符合该病病位在肺,病变常累及脾胃。李东垣曰:“脾胃一虚,肺气先绝。”故临证治疗中也应注重运用调理脾胃之法。

基于关联规则的组方规律分析显示高频组合前5位的是黄芪→女贞子+白术、黄芪→白花蛇舌草+白术+麦冬、黄芪→女贞子+茯苓、黄芪→白花蛇舌草+茯苓+麦冬、黄芪→浙贝母+白术+麦冬,可见黄芪、白术、茯苓、女贞子、麦冬、白花蛇舌草、浙贝母此7味药常常相互配伍出现。该结果侧面体现出多数学者在临证实践中重视益气养阴健脾以肺脾同补、清热解毒化痰以相须为用的基本治法^[30]。通过关联规则网络图可见,20味药物组合为黄芪、党参、麦冬、北沙参、半夏、半枝莲、南沙参、太子参、女贞子、山慈菇、桔梗、浙贝母、甘草、白术、白花蛇舌草、茯苓、莪术、薏苡仁、陈皮、鱼腥草,分析可知此药物组合为六君子汤合益气养阴解毒方加减。临床研究表明,六君子汤加减辅助治疗晚期NSCLC能显著改善生活质量,降低化疗后所致的严重血液学毒性发生率^[31]。益气养阴解毒方及其类方是现代医家肺癌常用的基础方,不仅具有控瘤抑瘤作用,也能改善机体免疫,现代实验研究也在不断揭

示其协同治疗的分子机制及效应物质等,以期延长患者总生存期^[32]。

因子分析结果表明,因子F1为黄芪、鱼腥草、仙鹤草、壁虎(补虚药、清热药、止血药、攻毒杀虫止痒药),因子F2为茯苓、白术、党参(利水渗湿药、补虚药、补虚药),因子F3为白花蛇舌草、半夏、半枝莲、莪术、夏枯草(清热药、化痰止咳平喘药、清热药、活血化瘀药、清热药),因子F4为浙贝母、薏苡仁、山慈菇(化痰止咳平喘药、利水渗湿药),因子F5为北沙参、天冬(补虚药、补虚药),因子F6为甘草、陈皮(补虚药、理气药),因子F7为麦冬、南沙参(补虚药、补虚药),因子F8为生地黄、当归(清热药、补虚药),因子F9为瓜蒌(化痰止咳平喘药),因子F10为桔梗、丹参(化痰止咳平喘药、活血化瘀药)。可见补虚药在中医治疗肺癌方剂中占有重要地位,常与清热药、止咳化痰平喘药等配伍使用。而聚类分析与因子分析的结果在一定程度上存在相似性。类1是F2+F6的基础上配伍半夏与C1基本相同,实为六君子汤组成,提示该类主要以健脾益气、燥湿化痰为主,尤适用于晚期肺脾气虚之证候。类2为生地黄(清热凉血、养阴生津),甘草(补脾益气、清热解毒),当归(补血活血、消肿止痛),桔梗(宣肺祛痰、利咽排脓),与类4太子参(补气健脾、生津润肺),丹参(活血祛瘀、凉血消痈),均提示该两类在补益为主,兼以活血祛瘀,与F8+F10组成相似,多适用于NSCLC气虚血瘀证^[33]。类3为黄芪(补气健脾),女贞子(滋补肝肾),白花蛇舌草(清热解毒),半枝莲(清热解毒、活血化瘀),莪术(破血祛瘀、行气消积),麦冬(养阴润肺),北沙参(养阴清肺、益胃生津),天冬(养阴润燥、清肺生津),南沙参(养阴清肺、益胃生津、补气化痰),提示该类主要以益气养阴解毒药物聚为一类,与临床气阴两虚兼瘀毒互结证相对应。类5为夏枯草(清热泻火),瓜蒌(清热化痰),仙鹤草(收敛止血补虚),壁虎(祛风解毒散结),鱼腥草(清热解毒),浙贝母(清热化痰),薏苡仁(健脾渗湿、清热排脓),山慈菇(清热解毒、消痈散结),提示该类主要以清热解毒为主,同时该结果也佐证了火热毒邪与肺癌的病因有着密切联系,清热需解毒,毒解则热清^[34]。公因子F5,F7以养阴为主,其中沙参、麦冬相结合,为沙参麦冬汤的基础,有甘寒养阴、清热润燥之功。公因子F1,F3,F4,F9以清热解毒化痰为主。综上所述,聚类分析与因子分析结果可相互对照,也提示治疗中晚期NSCLC中药有规可循。

近年来数据挖掘融合复杂网络的交叉性探索

在中医药领域得到了较为广泛的应用,同时也取得了诸多研究进展^[35]。本研究通过构建复杂网络分析研究药物-药物、药物-证素的相关性,将单味药物及单个证素作为一个节点,连接某2种药物及药物与证素的边的权重则表示其在复方中配伍使用的频数,进而可反映出整体药物组方规律及针对特定证素的药物配伍特点。复杂网络分析结果显示,黄芪、麦冬、浙贝母、白术、女贞子、甘草、薏苡仁、白花蛇舌草、茯苓、半枝莲、桔梗、太子参、南沙参、北沙参、半夏、党参、山慈菇、陈皮、莪术、鱼腥草,此20味药为治疗中晚期NSCLC的核心药物组成,核心中药处方为六君子汤合益气养阴解毒方加减,此结果也与高频药物之间的关联规则网络图相符。围绕肺脾两脏,气虚阴虚仍为最主要的病性证素,兼见痰、湿邪、热(火)邪、血瘀、毒邪,其中黄芪、白术、白花蛇舌草、麦冬、浙贝母、甘草、茯苓、党参、半夏、陈皮、薏苡仁、半枝莲、壁虎、当归、仙鹤草、鱼腥草、桔梗、丹参、生地黄、莪术、瓜蒌、南沙参、北沙参、太子参、麦冬、女贞子、夏枯草、山慈菇为七大病机特点的公共用药。六君子汤出自《太平惠民和剂局方》,具有益气健脾、燥湿化痰之效,是培土生金法的经典代表方。由此也可验证中晚期NSCLC患者多正气耗伤,大毒之品攻伐,则脾胃虚弱,故晚期治疗重在扶脾胃,安谷则昌,脾胃健运方使水谷精微变生气血,濡养周身。目前医家也普遍认为气阴两虚是中晚期NSCLC的主要证型,遣方用药时也常从痰、从癌毒、从血瘀等不同角度论治。益气养阴解毒方及其类方遵循益气、养阴、解毒的组方原则,临床上单独或联合使用多能起到增效减毒、改善患者生存质量的作用。此方也反映出现代医家临床遣方用药多西医辨病与中医辨证相结合,同时注重现代中药药理学方面的研究。中医药是一个信息量巨大的宝库,复杂网络分析在研究应用中仍存在一些不足,仍需不断研究^[29]。综上所述,本研究基于数据挖掘探索中医药治疗中晚期NSCLC的用药规律,期待该结果能一定程度上为肺癌中医治疗及成药研发提供一些参考,后期可进一步开展实验研究,临证实践中更应审慎运用。

[利益冲突] 本文不存在任何利益冲突。

[参考文献]

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in

- 185 countries [J]. *CA Cancer J Clin*, 2021, 71 (3) : 209-249.
- [2] ROSELL R, KARACHALIOU N. Large-scale screening for somatic mutations in lung cancer [J]. *Lancet*, 2016, 387(10026) : 1354-1356.
- [3] GETTINGER S, HORN L, JACKMAN D, et al. Five-year follow-up of nivolumab in previously treated advanced non-small-cell lung cancer: results from the CA209-003 study [J]. *J Clin Oncol*, 2018, 36 (17) : 1675-1684.
- [4] 周岱翰, 林丽珠, 田华琴, 等. 益气化痰法为主中医药治疗方案对老年非小细胞肺癌中位生存期的影响: 一项多中心、前瞻性临床队列研究[J]. *世界中医药*, 2014, 9(7) : 833-838, 844.
- [5] 王雯珺, 钦敬茹, 张程程, 等. 养阴解毒颗粒联合吉非替尼治疗表皮生长因子受体突变晚期肺腺癌阴虚型的临床研究[J]. *中华中医药杂志*, 2021, 36(4) : 2415-2418.
- [6] 王珊珊, 郭茗, 朱焱, 等. 国医大师周仲瑛教授辨治肺癌经验[J]. *中华中医药杂志*, 2015, 30(12) : 4332-4335.
- [7] 唐幸林子, 方灿途, 孟金成, 等. 国医大师周岱翰运用星夏涤痰饮治疗肺癌经验[J]. *中医药导报*, 2019, 25(8) : 35-36, 40.
- [8] 张文曦, 刘苓霜, 朱欣佚. 国医大师刘嘉湘从顾护脾胃论治肺癌经验[J]. *南京中医药大学学报*, 2020, 36(4) : 557-560.
- [9] 郑红刚, 花宝金, 朴炳奎. 朴炳奎辨证治疗肺癌的学术思想[J]. *北京中医*, 2007, 26(5) : 273-275.
- [10] 田建辉. “正虚伏毒”为肺癌发病的核心病机[J]. *上海中医药杂志*, 2016, 50(12) : 10-14.
- [11] 陈志奎, 宋鑫, 高静, 等. 基于数据挖掘的中医诊疗研究进展[J]. *中华中医药学刊*, 2020, 38(12) : 1-9.
- [12] 国家药典委员会. 中华人民共和国药典: 一部[M]. 北京: 中国医药科技出版社, 2020.
- [13] 南京中医药大学. 中药大辞典[M]. 2版. 上海: 上海科学技术出版社, 2006.
- [14] 唐德才, 吴庆光. 中药学[M]. 3版. 北京: 人民卫生出版社, 2016.
- [15] 国家技术监督局. GB/T 16751. 2—1997: 中医临床诊疗术语·证候部分[S]. 北京: 国家技术监督局, 1997.
- [16] 李灿东. 中医诊断学[M]. 4版. 北京: 中国中医药出版社, 2016.
- [17] 中国中医研究院. 中医证候鉴别诊断学[M]. 北京: 人民卫生出版社, 1987.
- [18] 朱文锋, 甘慧娟. 证素内容的辨析[J]. *中医药导报*, 2005, 11(1) : 11-13.
- [19] 杨福双, 高法, 谭田慧, 等. 基于数据挖掘的中医药治疗小儿肺炎组方用药规律研究[J]. *中国中药杂志*, 2020, 45(8) : 1942-1947.
- [20] 李萍, 舒琦瑾. 肺癌中医辨证分型的文献分析[J]. *中华中医药学刊*, 2008, 26(12) : 2694-2696.
- [21] 王保芹, 李泽庚. 肺癌中医辨证分型研究进展[J]. *陕西中医药大学学报*, 2016, 39(4) : 99-102.
- [22] LIU Y, ZHANG J, LI L, et al. Genomic heterogeneity of multiple synchronous lung cancer[J]. *Nat Commun*, 2016, 7: 13200.
- [23] 李丛煌, 花宝金. 283例中晚期非小细胞肺癌患者证候分布及证候要素组合特点分析[J]. *北京中医药大学学报*, 2009, 32(10) : 712-715.
- [24] 游捷, 李泰峰, 汪霞, 等. 303例老年原发性支气管肺癌中医证候分型[J]. *中医杂志*, 2012, 53(16) : 1404-1407, 1426.
- [25] 刘小敏, 卢琳琳, 黄艳芳, 等. 黄芪水提物抑制PI3K/Akt通路预防肺癌发生的作用及机制研究[J]. *中草药*, 2019, 50(7) : 1621-1628.
- [26] 李杨, 杨玲玲, 白彦丽, 等. 黄芪多糖对气阴两虚Lewis肺癌荷瘤小鼠肿瘤生长、转移及细胞周期的影响[J]. *肿瘤防治研究*, 2018, 45(3) : 131-137.
- [27] 柳叶, 陈龙云. 黄芪多糖对人肺癌顺铂耐药株A549/DDP细胞耐药性的影响及其机制[J]. *吉林大学学报: 医学版*, 2020, 46(6) : 1162-1168, 1348.
- [28] 孟林凤. 60例晚期非小细胞肺癌患者紫杉醇联合顺铂化疗前后中医寒热证型变化研究[D]. 北京: 北京中医药大学, 2018.
- [29] 谢利, 余德海, 申涛. 辛开苦降甘补调和法治疗老年晚期非小细胞肺癌的临床研究[J]. *成都中医药大学学报*, 2019, 42(4) : 45-48.
- [30] 沈琦. 中药治疗中晚期非小细胞肺癌的系统评价及用药规律分析初探[D]. 天津: 天津中医药大学, 2020.
- [31] 周斌, 单泽松, 方媚媚, 等. 六君子汤加减辅助化疗治疗晚期非小细胞肺癌31例临床观察[J]. *中医杂志*, 2015, 56(3) : 219-222.
- [32] 姜怡, 张朋, 蔡雨晴, 等. 益气养阴解毒方及其联合抗CTLA-4单抗对肺癌原位模型小鼠生存期及瘤体组织Foxp3、CTLA-4蛋白表达的影响[J]. *中医杂志*, 2021, 62(10) : 908-913.
- [33] 李晓娟, 陈滨海, 杨雪静, 等. 气虚血瘀证支气管肺癌患者“瘀毒”病机的临床研究[J]. *中华中医药杂志*, 2019, 34(2) : 773-777.
- [34] 张磊. 国医大师周仲瑛治疗肺癌角药撷菁[J]. *上海中医药杂志*, 2013, 47(11) : 1-2.
- [35] 吕庆莉. 数据挖掘与复杂网络的融合及其在中医药领域应用[J]. *中草药*, 2016, 47(8) : 1430-1436.

[责任编辑 顾雪竹]