

利用文本挖掘技术探索中西药治疗慢性乙型肝炎的用药规律

谭勇¹, 杨静¹, 赵宁¹, 郑光^{1,2}, 蔡锋¹, 姜春燕³, 郭洪涛^{1,4}, 姜淼¹, 吕爱平^{1*}

(1. 中国中医科学院中医临床基础医学研究所, 北京 100700;

2. 兰州大学应用数学与统计学院, 兰州 730000; 3. 首都医科大学附属北京
中医医院皮肤科, 北京 100010; 4. 上海中医药大学, 上海 201203)

[摘要] **目的:**利用文本挖掘探索中西药治疗慢性乙型肝炎的核心药物及其组合关系。**方法:**采集治疗慢性乙型肝炎的相关文献,建立 Access 数据库,运用 SQL 对数据进行处理,结合人工降噪后,挖掘中西药系统治疗慢性乙型肝炎的用药规律。**结果:**丹参、黄芪、柴胡、虎杖、甘草、苦参、天冬、茵陈、栀子 是中医复方配伍治疗乙肝的核心药物;干扰素和丹参注射液分别是治疗乙肝的最常用西药和中成药;干扰素、拉米夫定、胸腺肽、谷胱甘肽、维生素 C 等常与丹参注射液联合使用,黄芪注射液、益肝灵、护肝片、肝苏颗粒等常与拉米夫定联合使用。**结论:**文本挖掘获得的中西药治疗慢性乙型肝炎的用药规律不仅反映了临床用药实际,而且拓展了临床用药的范畴与思路。

[关键词] 慢性乙型肝炎;中西药使用规律;文本挖掘

[中图分类号] R287 **[文献标识码]** A **[文章编号]** 1005-9903(2011)24-0232-04

Regularity of Chinese and Western Medicine Application for Chronic Hepatitis B with Text Mining Technique

TAN Yong¹, YANG Jing¹, ZHAO Ning¹, ZHENG Guang^{1,2}, CAI Feng¹,
JIANG Chun-yan³, GUO Hong-tao^{1,4}, JIANG Miao¹, LV Ai-ping^{1*}

(1. Institute of Basic Research In Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China; 2. School of Information Science and Engineering Technology, Lanzhou University, Lanzhou 730000, China; 3. Beijing Chinese Medicine Hospital, Capital Medical University, Beijing 100010, China; 4. Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China)

[Abstract] **Objective:** To explore primary Chinese and western medicines treating chronic hepatitis B and their combination regularity with text mining technique. **Method:** Literatures on Chinese and western medicines treating chronic hepatitis B were collected from Chinese biomedical literature database (CBM) and access database was established. Above data were processed by SQL, meanwhile, data noise was removed with artificial means. And then, primary Chinese and western medicines treating chronic hepatitis B and their combination regularity were received by text mining technique. **Result:** Chinese herbs, such as salvia, astragalus, bupleurum, polygonum, licorice, matrine, aspartame, capillaries and gardenia, were main components of Chinese herb compound treating chronic hepatitis B. Danshen injection and interferon were the most commonly used Chinese patent medicine and western medicine for chronic hepatitis B. Some western medicine, for example, Interferon, lamivudine, thymosin,

[收稿日期] 20110526(002)

[基金项目] 国家自然科学基金杰出青年项目(30825047);国家自然科学基金面上项目(30902003);第 47 批博士后基金项目(20100470522)

[第一作者] 谭勇,医学博士,在站博士后,从事中医病证关联研究和中药毒性评价研究,Tel: 13718069135

[通讯作者] *吕爱平,医学博士,首席研究员,博士生导师,E-mail: lap64067611@126.com

glutathione and vitamin C were often used in combination with Danshen injection. In addition, lamivudine was applied with astragalus injection, yiganling, hugan slices or gansu particles. **Conclusion:** Primary application of Chinese and western medicines treating chronic hepatitis B and their combination regularity not only reflect clinical practical state, but also expanded the scopes and thinking of clinical application of Chinese and western medicine.

[**Key words**] chronic hepatitis B; regularity of Chinese and western medicine application; text mining

慢性乙型肝炎是我国常见的慢性传染病之一,严重危害人民健康。目前认为对慢性乙肝应采取综合治疗,西医治疗主要包括抗病毒、免疫调节、抗炎保肝、抗纤维化和对症治疗,其中抗病毒治疗是关键。中医治疗采用辨证施治、整体治疗,除了中草药配伍个性化治疗之外,也采用中成药治疗,尤其中草药配伍治疗更体现了中医的特色和优势。中西医结合,优势互补、扬长避短,是目前治疗慢性乙型肝炎的重要措施。本文利用文本挖掘技术探索中西药治疗慢性乙型肝炎的用药规律,为临床用药和药物研究提供依据。

1 材料

登录中国生物医学文献数据库(Chinese biomedical literature database, CBM)在主题检索下检索关键词“慢性乙型肝炎”,共得到文献 17 834 篇(检索日期 2011 年 3 月 6 日)。

2 方法

2.1 文本数据处理 将采集的数据按照先后顺序整合到一个平面文件(后缀 TXT)中,以 ANSI 编码格式保存。然后,利用专有的文本提取工具(软件著作权,登字第 0261882 号,登记号 2010SR073409)对下载的非结构化文本数据进行信息提取,保存成格式化的、便于数据库(Access 和 Microsoft SQL Server)处理的格式。

2.2 文本挖掘以及分析 在同一篇文章中出现的关键词,部分反映整篇文章的信息,相关的关键词之间存在着“共同出现”的特点。这种“共同出现”不是随机的,而是蕴含有一定的意义^[1],尤其对于高频共同出现的关键词对,在一定程度上反映了科研工作者的重视程度。这些关键词对可作为文本挖掘的主要素材^[1-3]。在本研究中,首先采用表 1 的算法构造针对每一篇文献共同出现的关键词对,得到 DN_pairs 数据表。然后采用表 2 的算法,将上述数据表中相同的关键词对进行合并处理,仅保留它们出现的频数,得到 DN_pairs_frqcy 数据表。

表 1 构建关键词对程序算法

```
USE Table Initial
FOR each PMID
    k = Number_of_DescriptorName(PMID)
    j = 1
    FOR DescriptorNames(i) (i = 1, 2, ..., k)
        DO while j ≤ k
            DescriptorNames_Pair = DescriptorNames(i) +
                DescriptorNames(j)
            j = j + 1
            OUTPUT DescriptorName_Pair INTO
                table DN_pairs
        ENDDO
    j = 1
    ENDFOR
ENDFOR
```

表 2 合并筛检关键词对程序算法

```
USE table DN_pairs
k = max_line_number
DO while k ≥ 1
GO top
FOR DescriptorName_Pair(1) // The 1st pairs in CHD_RA
    COUNT its Frequency
EndFor
OUTPUT DescriptorName_Pair, Frequency INTO table
    DN_pairs_frqcy
DELETE all DescriptorName_Pair(1) from table
    DN_pairs
k = max_line_number
ENDDO
```

2.3 数据的可视化 从 DN_pairs_frqcy 数据表中抽出不同频数的关键词对,用 Cytoscape 2.7 进行可视化处理。

3 结果

3.1 治疗慢性乙型肝炎的中草药 通过对有关“慢性乙型肝炎”文献的分析,得到用药关系网络图 1, 2。单味药之间连线上的数字是 2 个单味药共同出现的频数。频数大小体现了 2 个单味药在疾病治疗中共同出现的几率大小,反映其配伍关系的重要程度。与某个单味药连接的线条多少体现其与其他单味药配伍的几率高低,反映这个单味药在治疗用药中的重要性。

图 1 展示了 2 个单味药同时出现频次在 17 次以上的用药情况。丹参与 15 个单味药有配伍关系,配伍几率由大到小排列前 10 味药依次为:甘草、苦参、黄芪、虎杖、柴胡、茵陈、赤芍、白花蛇舌草、郁金、白芍,黄芪与 8 个单味药有配伍关系,甘草和柴胡各自与 6 个单味药有配伍关系,茵陈与 5 个单味药有配伍关系,苦参、虎杖、茯苓、白术各自与 3 个单味药有配伍关系,郁金、天冬、栀子、黄芩、白芍、白花蛇舌草、当归与 2 个单味药有配伍关系,赤芍、鳖甲、佛手、党参各自与 1 个单味药有配伍关系。图 1 中所有具有配伍关系的单味药是组成治疗慢性乙型肝炎复方的常用中药。图 2 展示了 2 个单味药同时出现频次在 34 次以上的用药情况,是对图 1 用药的进一步筛选。丹参与 5 个单味药有配伍关系,其次甘草与 3 个单味药有配伍关系,而苦参与 2 个单味药有配伍关系。图 2 中展示的所有单味药在治疗银屑病的复方中更常用。以丹参为中心,依次甘草、苦参、黄芪、虎杖和柴胡与之配伍,体现了治疗慢性乙型肝炎复方的配伍核心。

3.2 治疗慢性乙型肝炎的西药和中成药 文本挖掘显示,干扰素和丹参注射液分别是治疗慢性乙型肝炎使用最频繁

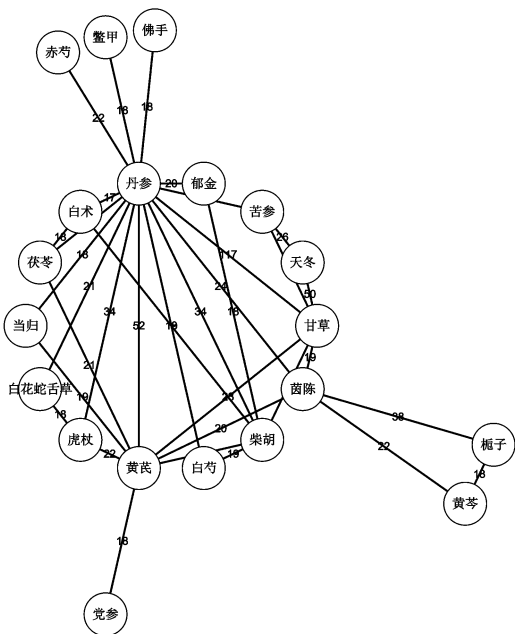


图 1 治疗乙肝中药网络第 1 层

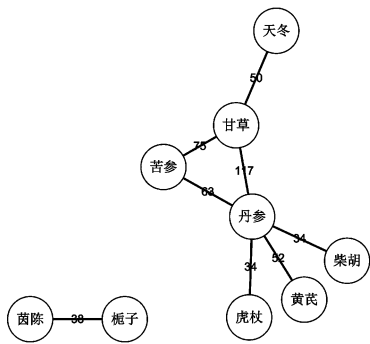


图 2 治疗乙肝中药网络第 2 层

的西药和中成药。使用频率由高到低:拉米夫定、阿德福韦、胸腺肽、白细胞介素、恩替卡韦、谷胱甘肽、阿糖腺苷等是治疗慢性乙型肝炎的常用西药;益肝灵、护肝片、黄芪注射液、丹参片、肝炎灵注射液、茵栀黄注射液、肝苏颗粒等是治疗慢性乙型肝炎的常用中成药(图 3,4)。

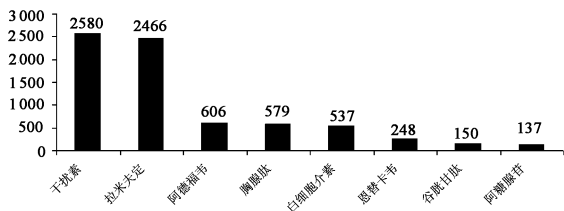


图 3 治疗慢性乙型肝炎的西药使用频率

3.3 治疗慢性乙型肝炎西药和中成药联合使用情况 文本挖掘结果显示,干扰素和拉米夫定是中西药联用时最常用的西药,干扰素最多与丹参注射液和黄芪注射液联合使用,也常与益肝灵、丹参片或护肝片联合使用,拉米夫定最多与黄芪注射液或益肝灵联合使用,也常与护肝片、丹参注射液或

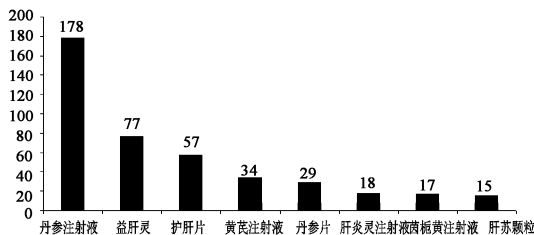


图 4 治疗慢性乙型肝炎的中成药使用频率

肝苏颗粒联合使用。丹参注射液是中西药联用时最常用的中成药,除了常与干扰素和拉米夫定联合使用外,也与胸腺肽、谷胱甘肽、维生素 C、门冬氨酸钾镁、硫普罗宁等一起使用。益肝灵是继丹参注射液之后,有 5 种西药与之组合使用,黄芪注射液、护肝片和当飞利肝宁分别与 3 种西药有组合关系(图 5)。

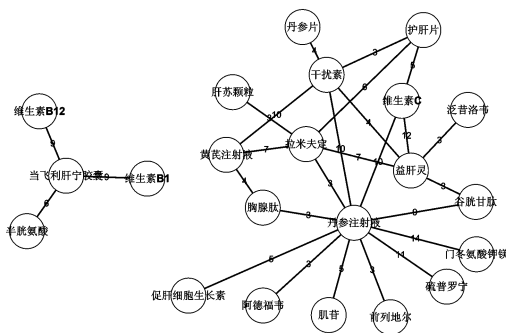


图 5 治疗慢性乙型肝炎中西药联用情况

4 讨论

在生物学领域,由于生物学数据和生物医学文献数量的急骤增长,通过数据挖掘寻找规律和新知成了生物学研究的一个新热点和重要分支^[4]。文本挖掘是从文本集中发现潜在、隐藏的归纳性知识的一门技术^[5],它能对海量数据进行整合、分析,获得的结果更具有代表性、可信度更高^[6]。既往的研究表明,文本挖掘具有探索治疗疾病用药规律的技术特征和条件,可操作性强^[7]。本研究从大量文献(17 834 篇)中挖掘中西药治疗慢性乙型肝炎的用药规律,文献信息全面,挖掘过程恰当,质量控制严格,所得结果具有指导临床用药价值。

中医认为,慢性乙型肝炎病机特点为湿热羁留、毒痰瘀交阻,肝胆不疏、脾肾受损。治疗以“攻补兼施”为原则,清热、祛湿、凉血、解毒、化痰、化瘀、疏肝利胆、补脾益肾诸法择情选用。本研究发现,丹参是辨证论治乙肝的核心配伍药物。肝体阴而用阳,乙肝者肝血易虚、易瘀,肝疏泄情志失常、心神受扰。丹参苦、微寒,归心、肝经,具有养血活血、祛瘀止痛、清心安神的功效,是针对乙肝病机的要药。丹参除了与甘缓和中的甘草配伍之外,还与清热利湿解毒药如苦参、虎杖、茵陈、白花蛇舌草等联合使用,也与和解少阳的柴胡、养阴柔肝的白芍、凉血化瘀的赤芍、疏肝活血的郁金及补气升阳的黄芪等联合使用。黄芪是另一个在复方中经常使

用的核心配伍药物。黄芪甘、温,归肺、脾经,具有补气固表、健脾除痹的功效。“肝为罢极之本”,乙肝病人常见乏力症状,是肝气虚、不能升发的表现,用黄芪切中病机,通过益肺气、补脾气达到补肝、疏肝的目标。现代诸多研究均发现黄芪对各种肝损伤均具有保护作用,能够保护肝细胞膜,降低血清ALT、AST水平等,是治疗乙肝的常用中草药^[8-10]。柴胡苦、微寒,归肝、胆经,和解表里、疏肝升阳,治疗肝胆病的著名方剂如小柴胡汤、柴胡疏肝散、四逆散等均以其为君药。丹参、黄芪、柴胡分别从和血、补气、和解少阳三方面调治,三者构成了治疗乙肝复方的核心配伍药物。综观本研究发掘的所有中草药,均为临床上的常用药物,其中的配伍关系,体现了中医对证治疗主旨。

西医认为,治疗乙肝的关键是抗病毒,因此临床上干扰素和拉米夫定被经常使用,配合保肝、调节免疫药物如胸腺肽、白细胞介素、谷胱甘肽等,共同达到治疗目标。与挖掘的中草药使用规律结果相近,临床上经常使用的中成药是丹参、黄芪制剂等,亦体现了中医审因对证论治的思想。益肝灵、护肝片、肝炎灵注射液、茵栀黄注射液、肝苏颗粒等均为治疗乙肝的经典中草药按照现代制药工艺制成的现代中药,具有质量稳定、使用方便、药力集中等特点,在保肝降酶、利胆退黄方面具有一定的优势。

慢性乙型肝炎是临床疑难病证,中西药联合使用治疗是目前的主要手段。与西药单一成分、单靶点的作用不同,中成药具有多成分、多途径、多靶点的作用特点,因此其药力持久、副作用较少。50多年的中西医结合实践证明,恰当的中西药联用具有增效减毒的特点^[11]。本研究整合、挖掘大量临床专家的经验,发现干扰素和丹参注射液分别是中西药联用时最常用的西药和中成药。针对不同的临床情况,干扰素还常与益肝灵、丹参片、护肝片联合使用;丹参注射液还常与胸腺肽、谷胱甘肽、维生素C、门冬氨酸钾镁、硫普罗宁等一起使用。长期、反复的临床实践验证了上述中西药联用的安全性和有效性,为了进一步国内推广应用和走向国际化,优化中西药联用配伍剂量、找准中西药联用适应症以及深入探索中西药联用的作用机制必不可少。另外,根据中医学“以方测证”的方证研究思路,通过中西药联用的研究,逆向发掘慢性乙型肝炎发生、发展、转归的新机制,将有助于慢性乙型肝炎诊治理论和实践水平的提高。

综上,文本挖掘是探索中西药治疗慢性乙型肝炎用药规律的有效方法,获得的结果不仅真实反映了临床用药实际,而且其新的发现拓展了临床用药的范畴和思路,为科学合理地使用治疗慢性乙型肝炎的中西药提供了客观依据。本研究也为进一步探索治疗慢性乙型肝炎中西药联用机制奠定

了基础,并且对治疗其他疾病用药规律的探索有一定的启迪和示范作用。

[参考文献]

- [1] Andrea Campagna, Rasmus Pagh (2009) Finding associations and computing similarity via biased pair sampling [C]. Miami, Florida: 2009 Ninth IEEE International Conference on Data Mining, 2009:61.
- [2] Jeffrey W Seifert. Data mining: An overview[R]. CRS Report RL31798, 2004.
- [3] Brigitte Mathiak, Silke Eckstein. Five steps to text mining in biomedical literature [C]. Pisa: The second european workshop on data mining and text mining for bioinformatics, held in conjunction with ECML/PKDD. 2004: 47.
- [4] Tari L, Anwar S, Liang S, et al. Discovering drug-drug interactions: a text-mining and reasoning approach based on properties of drug metabolism [J]. Bioinformatics, 2010, 26(18): 1547.
- [5] Al-Mubaid H, Singh R K. A text-mining technique for extracting gene-disease associations from the biomedical literature [J]. Int J Bioinform Res Appl, 2010, 6(3): 270.
- [6] Rodriguez-Esteban R. Biomedical text mining and its applications [J]. PLoS Comput Biol, 2009, 5(12): e1000597.
- [7] 谭勇,郭洪涛,郑光,等.利用文本挖掘技术探索中医药治疗疾病的用药规律[J].世界科学技术——中医药现代化,2010,12(5):823.
- [8] 张银娣,沈建平.黄芪皂苷抗实验性肝损伤作用[J].药学学报,1992,27(6):401.
- [9] 马红,王宝恩.黄芪对免疫损伤性肝纤维化大鼠的治疗作用[J].中西医结合肝病杂志,1997,7(1):32.
- [10] 张毅,李筠,赵景民.复方黄芪颗粒对肝纤维化大鼠肝脏基质金属蛋白酶-2、金属蛋白酶组织抑制因子-2蛋白表达的影响[J].中西医结合肝病杂志,2008,18(4):242.
- [11] 周顺,滕树忠.中西药联用对临床药效的影响[J].中国现代医生,2008,46(8):100.

[责任编辑 邹晓翠]