Basic treatment principles for bladder cancer with Chinese herbal medicine: an application of text mining

Cheng Xiao
China-Japan Friendship Hospital
Yinghuayuan Dongjie Chaoyang District, Beijing, China
+86 10 84205442, 100029
xiaocheng2002812@163.com

Shenglong Jing
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
+86 10 64014411, 100700
1538736048@qq.com

Miao Jiang
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
+86 10 64014411, 100700
miao_jm@126.com

Cheng Lu
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
+86 10 64014411, 100700
lv_cheng0816@163.com

Xiaojuan He
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
+86 10 64014411, 100700
hxj19@yahoo.com.cn

Peng Xu
Jiangxi University of Traditional Chinese Medicine, Nanchang, China
+86 791 7118826, 330004
xp0420@163.com

Dan Luo
Beijing University of Chinese Medicine, Beisanhuang Donglu 11# Chaoyang District, Beijing, China
+86 10 64213841, 100030
330996648@qq.com

Guang Zheng
School of Information Science-Engineering, Lanzhou University, Lanzhou, China
+86 931 8910864, 730000
forzhengguang@163.com

Aiping Lu*
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
+86 10 64067611, 100700
lap64067611@163.com

ABSTRACT

Bladder cancer occurs in the majority of cases in males with a male/female sex ratio of 3:1. Of the three main histological variants of epithelial malignancies arising from the urothelium of the urinary bladder, transitional cell carcinoma (TCC) is the most prevalent in Japan, North America, and other developed countries, while squamous cell carcinoma and adenocarcinoma are diagnosed less frequently, and it is also the most expensive cancer to treat. Antineoplastic resistance is one of the most common problems in bladder cancer which can cause treatment failure and serious complications. Antineoplastic-resistance has been increasing in the past decades worldwide. Chinese herbal medicine (CHM) has been an integral part of Traditional Chinese Medicine (TCM) for thousands of years. Some Chinese herbs in the study were proved to have antineoplastic activity, which prompted their compound prescription use in the management of bladder cancer. Many herbal formulations have been developed and used in the treatment of Bladder cancer and were proved to be effective and safe. Yet the principles of treating bladder cancer with CHMs are hard to manage due to the complexity of TCM theory. In this study, a novel text mining method was development based on a comprehensive collection of literatures in order to explore the treatment principles more intuitively. Networks of TCM patterns and CHMs which are most frequently used in bladder cancer treatment were built-up and analyzed, two major principles are explored in treating bladder cancer from 14097 records of literature: Clearing the damp-heat with strengthening healthy qi. These findings might guide the clinicians in treatment of bladder cancer.

Categories and Subject Descriptors

Algorithms: Language Constructs and Features – abstract data types, polymorphism, control structures. In the process of data mining, we construct an data slicing algorithm called discrete derivatives.

General Terms


Keywords

Bladder cancer, Chinese herbal medicine, Pattern, Traditional Chinese medicine, Text mining.

1. INTRODUCTION

Bladder cancer is the 10th most common cancer worldwide, with the highest rates reported in Europe, North America and Australia.
and accounting for an estimated 261,000 new cases diagnosed and 115,000 deaths each year; by comparison, relatively low rates are found in the Far Eastern countries[1, 2]. It affects men more frequently than women[3]. Typical of solid tumours, bladder cancer incidence increases with age. Tumours of the bladder rarely occur before the age of 40–50, arising most commonly in the seventh decade of life[4, 5]. The median ages at diagnosis are 69 years for men and 71 for women[6].

Gross hematuria (presence of blood in urine) is the most common symptom in BC which was noted in 63 to 88% of the cases[7–9]. Dysuria (painful voiding) has been reported as the second most common symptom. Urinary obstruction, abdominal pain, urinary tract infection and weight loss have been reported occasionally[7, 10]. Rare cases of paraneoplastic syndromes such as ectopic ACTH secretion and hypercalcaemia were also reported[11].

There are two major biological pathways in human bladder tumor development, leading to two major subtypes of bladder cancer: superficial/non-muscle-invasive (NMIBC) and advanced/muscle-invasive (MIBC)[12].

An article by Bassi et al. [13] confirmed that delays in diagnosis and initiation of therapy have adverse effects on stage and survival. The median age at diagnosis is 65 years with as many as 10% of patients >85 years of age. Studies in this group are essential to define the optimal approach to care because a substantial number of people with advanced BC are elderly or have other poor-risk features (characterized by visceral metastasis, poor performance status (PS), and >5% weight loss).

Currently, cisplatin-based combination chemotherapy is considered to be the standard therapy for this disease. Regimens such as MVAC (methotrexate, vinblinaste, adriamycin, cisplatin), CMV (cisplatin, methotrexate, vinblinaste) and GC (gemcitabine, cisplatin) have been employed with relative risks (RRs) reported in up to 70% [14, 15]. Despite these high RRs, toxicity and survival outcomes remain suboptimal. For example, MVAC therapy results in a median survival time of approximately 13 months and is associated with severe toxicities including myelosuppression, nephrotoxicity, stomatitis, and emesis[16]. In addition, treatment-related death rates have been up to 3%. In an attempt to minimize these toxicities, studies are under way to determine which patients would best benefit from therapy [17, 18].

Chinese herbal medicine (CHM) has been an integral part of Traditional Chinese Medicine (TCM) for thousands of years. Some Chinese herbs in the study were proved to have antineoplastic activity [19–21]. These properties had prompted their compound prescription maybe use in the management of BC.

However, due to the complexity of TCM theory, the treatment principles of BC are complicated and mysterious. In order to explore the treatment principles more intuitively, a novel text mining method was development based on a comprehensive collection of 14,097 records of literatures[22]. The study would provide an accessible way for understanding the treatment principles for BC with CHMs.

2. MATERIAL AND METHODS

2.1 Data Collection

The dataset were downloaded from SinoMed (http://sinomed.cintcm.ac.cn/index.jsp) with the query term of “Bladder cancer” on Feb. 23, 2012. This dataset contains 14,097 records of literatures on clinical practices or theoretical research on bladder cancer. In this dataset, each record/paper is tagged with a unique ID. These records contain the title, keywords, and abstract of published papers[22].

2.2 Data Filtering

Please use a 9-point Times Roman font, or other Roman font with serifs, as close as possible in appearance to Times Roman in which these guidelines have been set. The goal is to have a 9-point text, as you see here. Please use sans-serif or non-proportional fonts only for special purposes, such as distinguishing source code text. If Times Roman is not available, try the font named Computer Modern Roman. On a Macintosh, use the font named Times. Right margins should be justified, not ragged.

2.2.1 TCM Patter

Pattern (also called as Syndrome, or Zheng) differentiation is regarded as the key role in the clinical practise of TCM traditional Chinese medicine[23]. Usually, pattern identification is the basis of the prescription of herb formulae, CHMs, or other TCM therapies. Thus it is natural and intuitive to filter out the pattern and then try to find the associate rules between pattern and CHMs. The top TCM patterns in BC are: qi deficiency (Qi xu), blood heat (Xue re), blood stasis (Xue yu) and damp-heat stagnation (Shi re yun jie).

2.2.2 Chinese herbal medicine

Based on the keyword list of CHMs (both legal names and other popular names are included for calculation), we filtered the CHMs in the plain text format, and then converted all popular names into legal names. All the CHMs were tagged with their unique paper ID. Based on the unique paper ID, we could construct the pairs of co-existed CHMs as they coexisted in literature. For example, in one paper, CHMs of Huangqi (Radix Astragali seu Hedyssari), Renshen (Radix Ginseng), and Shengdihuang (Radix Rehmanniae Recens) are mentioned. Then, the pairs of co-existed CHMs of “Huangqi-Renshen”, “Huangqi-Shengdihuang”, and “Renshen-Shengdihuang” are constructed.

3. RESULTS

In this paper, focused on BC, we explored the principles of pattern differentiation and CHMs prescription and the association between the two aspects under the framework of TCM theory from 14,097 literatures. The network construction is based on the analysis of networks of pattern and CHM correlated with BC in literature. The connections among these networks are built-up under the professional knowledge of TCM.

3.1 Major TCM Patterns in BC

Pattern identification is regarded as the first step during TCM clinical practice procedure. After the pattern is approved, the treatment principle can be determined. For example, when the pattern of blood stasis is approved, then the treatment principle of active blood and resolve stasis is...
determined. In our results, 4 TCM patterns are detected to be related with BC, and the TCM patterns in BC are presented in Fig. 1. Remarkably, the patterns of BC less amount due to lack of literatures.

3.2 Most frequently prescribed CHMs in BC treatment
P- Altogether 64 CHMs are mined from the literature in treatment of BC. As herbal formulae are composed by the CHMs, the list of most frequently used CHMs can certainly provide the information of TCM treatment principles more effectively due to the stability and uniqueness of each CHMs rather than formulæ which can be renamed easily after slight regulation. The top 10 frequently prescribed CHMs are shown in Fig. 2. It is demonstrated that most CHMs prescribed in BC management are those with functions of clearing away heat evil and removing carbuncle, only 3 CHMs stand for tonification, which can help strengthen the principal curative action of clearing damp-heat and invigorating qi.

3.3 Networks of the Pattern and CHMs in BC
The networks of patterns and CHMs in BC treatment can be constructed based on the co-existence frequency among patterns or CHMs, respectively. By checking these two networks, the correlation between TCM patterns and CHMs can be analyzed and explored. In order to achieve better visualization, the CHM network is simplified to preserve 13 CHMs which are the most frequently used in combination in treating BC. The networks of patterns and CHMs which with their correlation on BC is demonstrated in Fig. 3. The major correlation between TCM patterns identification and CHMs are demonstrated with arrows.

4. CONCLUSION AND DISCUSSION
Based on the analysis described in previous section, it is naturally come to the point that TCM treatment principles of a disease can be reasonably mined out and presented from dataset downloaded from SinoMed. Compared with the knowledge of BC in text book, the most knowledge is covered by the simple and succinct networks demonstrated in Fig.3 which can be summarized with following points and their internal connections.
4.1 TCM Networks of Patterns and CHMs can be Constructed and Analyzed

In this study, through mass calculation on dataset on BC, the main aspect of TCM networks were built-up. The pathogenesis related with BC includes Dampness-heat, blood heat, Deficiency of qi in spleen and kidney. To follow the matter of course, CHMs most frequently prescribed in BC treatment can be grouped in to 2 major classes, one group is responsible for clearing damp-heat, the other for help the principle action and reinforcing the healthy qi. These major principles might guide the clinicians in treatment of BC.

4.2 Internal Connections among Networks

Through directed text mining, the internal connections among TCM networks were also found. These internal connections can be grouped into two major hierarchical clusters. Each cluster is associated with one major kind of patterns. The major treatment principles of TCM treatment of BC can be explored by text mining method and summarized in a succinct figure.

4.3 TCM Network might be Useful in both TCM Clinical Practices and Scientific Researches

The network demonstrated in Fig. 3 can be taken as a high level of abstraction on the treatment of BC out of dataset contains 14,097 records. From the viewpoint of clinicians, it can be taken as a kind of reference. From the viewpoint of basic researchers, this result might be useful to illuminate some further studies in BC.

5. ACKNOWLEDGMENTS

This work was partially supported by National Science Foundation of China (No. 30902003 and 81072982). 2012’ Traditional Chinese medicine Professional project (No. 201207012).

6. REFERENCES


