

Analyze the Drug Regularity in the Prescription to the Treatment of Primary Hepatocellular Carcinoma By R Language

Yexing Wang^{1,a}, Jingyi Liu^{2,b}, Zhilei Jiang^{2,c} and Jian Gong^{1,d}

¹Qiqihar Medical College

²State Grid Heilongjiang Electric Power Company Limited

^a249137624@qq.com, ^bliujingyi2002@163.com, ^c32096100@qq.com, ^dgongjian_0303@sina.com

Keywords: R language; Liver cancer; Drug pairs; Cluster; Associate

Abstract. Object Application of R language to analyze the drug regularity in the treatment of primary hepatocellular carcinoma by traditional Chinese Medicine. Method Retrieval of the traditional Chinese medicine knowledge database in Chinese Wan-fang Database and extract the prescription for the treatment of primary hepatocellular carcinoma. Drug and symptom data set was built by Excel, and applied R language for frequency analyses, cluster analyses and association analyses, then draw the cluster diagram and associated graph. Results The common drugs used in the treatment of liver cancer are Hedyotis diffusa, Glycyrrhiza, Atractylodes, Scutellariae barbata, Lanceolata, Salvia, Curcuma Longa, Bupleurum, Poria and Peony root in turn. The common drug pairs were Astragalus-Glycyrrhiza, Poria-Atractylodes, Lanceolata-Atractylodes, Lanceolata-Glycyrrhiza, Scutellariae barbata-Salvia. Hedyotis diffusa often composition of the drug pair with Glycyrrhiza, Glycyrrhiza and Lanceolata. Single drug of Gardenia, Bupleurum, Astragalus, Scutellariae barbata and Pinellia were often used for the pain of liver, while the drug pairs were Pinellia-Bupleurum, Hedyotis diffusa-Astragalus, Bupleurum-Poria, Scutellariae barbata-Bupleurum and Bupleurum-Curcuma Longa. Conclusion Application of the R language for data mining of traditional Chinese medicine database can effectively explore the rule of drug treatment for liver cancer, and has a certain value for further elucidating the pharmacological mechanism of traditional Chinese medicine prescriptions and guiding the clinical use of drugs.

China is a high incidence of primary hepatocellular carcinoma country, about 383,000 people died of liver cancer each year, and mortality ranks second in malignant tumors[1]. Because the onset of liver cancer is occult, as well as the early lack of typical symptoms. So it is often found to have progressed to an advanced stage and lost the opportunity for surgery. For the patients with advanced liver cancer with poor therapeutic effect of western medicine, reasonable application of traditional Chinese medicine can often obtain certain curative effect, which is helpful to improve the symptoms, alleviate the pain and improve the life quality of the patients.[2,3] However, it is difficult to sum up the role of traditional Chinese medicine due to the application of syndrome differentiation and the complex compatibility, and makes the scientific nature of traditional Chinese medicine questioned.

Drug pair is the basic form of traditional Chinese medicine compatibility, which constitutes by 2 relatively fixed drugs. It is a bridge for the research between single drug and prescription, because of its simple structure, clear curative effect, etc. This study starts with the drug pair to construct liver cancer treatment prescription data set. The modern statistical method and R language were used to mine the data, which were helpful to further understand the role of compatibility of traditional Chinese medicine. Then reveal the relationship between the core medication and compatibility of traditional Chinese medicine in the treatment of primary hepatocellular carcinoma, and provide a theoretical basis for the treatment of liver cancer by traditional Chinese medicine.

Data and Methods

Data Sources. The data comes from the traditional Chinese medicine knowledge database of Chinese Wan-fang Database.

Data Extraction Standards. Retrieval by keywords, which were liver cancer or malignant

tumor of liver or liver and lump or liver and tumor. Search category was traditional Chinese medicine prescription.

Data Exclusion Standards. Non-liver cancer prescription.

Incomplete information.

No specific symptom description.

Duplicate record.

Data Pretreatment. Data extraction including prescription name, source, drug composition, drug dosage and corresponding symptoms. All the drug name in the knowledge database unified naming according to the 2010 edition of the Chinese Pharmacopoeia.

Statistical Methods. Drug and symptom data set was built by excel, and frequency analyses by R3.3.2, cluster analyses by complete, then the transactions transform the data into the shopping basket format for association analyses, and the results visualization by arulesViz.

Results

Distribution of Drug Frequency in Liver Cancer Treatment Prescriptions. A total of 32 doses of prescriptions, 426 flavours of drugs, and 279 species of symptoms. After degrading and standardize drugs and symptoms, such as panax notoginseng powder classify to panax notoginseng, plantain classify to plantain seed, aversion to wind and cold classify to aversion to cold, dry and bitter mouth classify to dry and bitter. 170 flavours of drugs and 95 species of symptoms were obtained in the finally. The specific frequency distribution of drugs and symptoms were shown in table 1 and 2.

Table 1 Distribution of drug frequency in liver cancer treatment prescriptions (Top 20)

Drugs	Freq uency	Frequenc y rate	Drugs	Freq uency	Frequenc y rate
Hedyotis diffusa	16	50.00%	Zedoary turmeric	8	25.00%
Glycyrrhiza	15	46.88%	Angelica sinensis	7	21.88%
Atractylodes Scutellariae	13	40.63%	Astragalus Red peony	7	21.88%
barbata	12	37.50%	root	6	18.75%
Lanceolata	12	37.50%	Rhubarb	6	18.75%
Salvia	11	34.38%	Scutellaria	6	18.75%
Curcuma	11	34.38%	Coix Seed	6	18.75%
Longa			Pinellia	5	15.63%
Bupleurum	10	31.25%	Turtle shell	5	15.63%
Poria	10	31.25%	Tangerine	5	15.63%
Peony root	8	25.00%	peel		

Table 2 Frequency distribution of liver cancer symptoms (Top 20)

Symptoms	Freq uency	Frequenc y rate	Symptoms	Freq uency	Frequenc y rate
Amorexia	23	71.88%	Jaundice	6	18.75%
Abdominal distension	19	59.38%	Pulse string small	6	18.75%
Liver pain	19	59.38%	Ascites	5	15.63%
Emaciation	13	40.63%	Abdominal pain	5	15.63%
Hypodynami a	8	25.00%	Limb swelling	5	15.63%
Hepatic cirrhosis	8	25.00%	Fever	4	12.50%
Dry mouth	8	25.00%	Abdominal tenderss	4	12.50%
Nausea	7	21.88%	Dark side	4	12.50%
Spiritual collapse	7	21.88%	Dark purple tongue	4	12.50%
Liver mass	6	18.75%	Radiating pain	4	12.50%

Cluster Analyses. Cluster analyses of drugs by complete, the optimal clustering data was judged by BIC index. The results were no kurtosis appeared, and then adjusted to 20 categories manually, shown in Fig. 1.

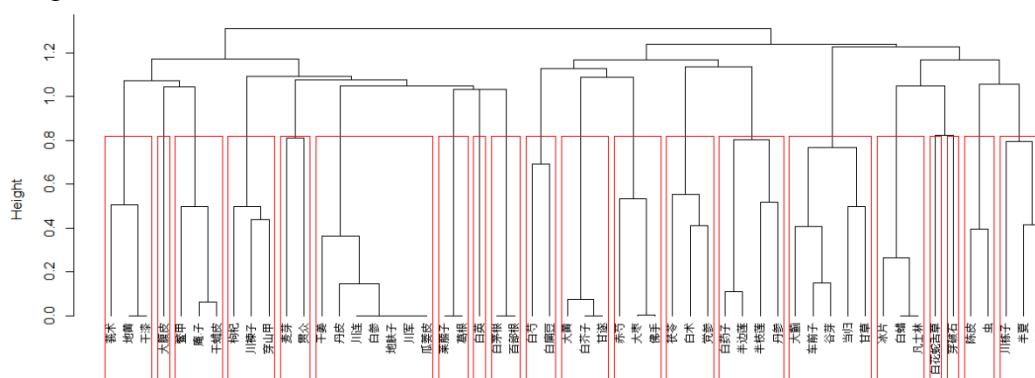


Figure 1. Drug cluster map.

Zedoary turmeric, radix rehmanniae, achyranthes root, areca peel, Turtle shell, anzi, dry skin of toad, medlar, chinaberry fruit, pangolin, malt, male fern rhizome, dried ginger, cortex moutan, sichuan coptis root, white ginseng, broom cypress fruits, Chuanjun, raphnus sativus, gegen, solanum dulcamara, cogongrass rhizome, radix stemonae, Peony root, white hyacinth bean, rhubarb, seeds of brassica alba, kansui root, red peony root, jujube, bergamot, Poria, Atractylodes, Lanceolata, radix stephaniae cepharanthae, lobelia chinensis lour, Scutellariae barbata, Salvia, cirsium japonicum, plantain seed, millet sprout, angelica sinensis, Glycyrrhiza, borneol, white wax, vaseline, Hedyotis diffusa, radix cudraniae, tangerine peel, insect, umbellate pot, Pinellia, Bupleurum

Drug Association Analyses. Set the support degree above 0.1, the confidence degree above 0.25. A single drug is associated with other single drug, ranked according to the lifting degree. The results are shown in table 3.

Table 3 Correlation analyses of common drugs used in liver cancer prescriptions (Top 10)

D rug pair s	Left side	Right side	Support degree	Confidence degree	Lifting degree	Frequency
1	Astragalus	Glycyrrhiza	0.2	1.00	2.	7
	Glycyrrhiza	Astragalus	0.2	0.47	2.	7
2	Poria	Atractylodes	0.2	0.80	1.	8
	Atractylodes	Poria	0.2	0.62	1.	8
3	Lanceolata	Atractylodes	0.2	0.75	1.	9
	Atractylodes	Lanceolata	0.2	0.69	1.	9
4	Lanceolata	Glycyrrhiza	0.3	0.83	1.	10
	Glycyrrhiza	Lanceolata	0.3	0.67	1.	10
5	Scutellariae barbata	Salvia	0.2	0.58	1.	7
	Salvia	Scutellariae barbata	0.2	0.64	1.	7

Specific Drug Association Analyses. Analysis of association between drugs with high frequency and other single drugs, takes the Hedyotis diffusa as an example, the results shown in table 4, visualization results shown in Fig. 2.

Table 4 Analyses of association between Hedyotis diffusa and other single drugs

D rug pairs	Left side	Right side	Support degree	Confidence degree	Lifting degree	Frequency
1	Glycyrrhiza	Hedyotis diffusa	0.34	0.73	1.4	11
	Hedyotis diffusa	Glycyrrhiza	0.34	0.69	1.4	11
2	Curcuma Longa	Hedyotis diffusa	0.25	0.73	1.4	8
	Hedyotis diffusa	Curcuma Longa	0.25	0.50	1.4	8
3	Lanceolata	Hedyotis diffusa	0.22	0.58	1.1	7
	Hedyotis diffusa	Lanceolata	0.22	0.44	1.1	7

Glycyrrhiza.
Lanceolata
Curcuma Longa
Hedyotis diffusa

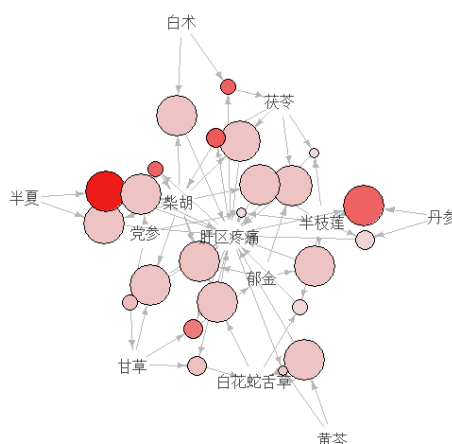
Fig. 2 Analyses of association between Hedyotis diffusa and other single drugs. The arrow indicates that the left side points to the right side, the darker the color, the higher the lift degree, the larger of the cycle, the higher the confidence degree.

Analyses of the Association between Specific Symptoms and Drugs. Analysis of association between symptoms with high frequency and drug pair, the drug is a single drug or drug pairs, sort by lift degree. Take the liver pain as an example, the results shown in table 4, visualization results shown in Fig. 3

Table 5 Analyses of association between liver pain and drugs.

Nu mber	Left side	Right side	Sup port degree	Conf idence degree	Liftin g degree	Frequ ency
------------	-----------	---------------	-----------------------	--------------------------	--------------------	---------------

Sin gle drug	Gardenia	Liver pain	0.13	1.00	1.68	4
	Bupleurum	Liver pain	0.31	1.00	1.68	10
	Scutellaria	Liver pain	0.16	0.83	1.40	5
	Scutellariae barbata	Liver pain	0.31	0.83	1.40	10
	Pinellia ternate	Liver pain	0.13	0.80	1.35	4
Dru g pairs	Pinellia ternate-Bupleurum	Liver pain	0.13	1.00	1.68	4
	Hedyotis diffusa-Scutellaria	Liver pain	0.13	1.00	1.68	4
	Bupleurum-Poria	Liver pain	0.19	1.00	1.68	6
	Scutellariae barbata-Bupleurum	Liver pain	0.13	1.00	1.68	4
	Bupleurum-Curcuma Longa	Liver pain	0.16	1.00	1.68	5



Salvia, Atractylodes, Poria, Pinellia, Glycyrrhiza, Hedyotis diffusa, Curcuma Longa, Lanceolata, Bupleurum, Scutellariae barbata, Scutellaria, Liver pain

Fig. 3 Analyses of association between liver pain and drugs. The arrow indicates that the left side points to the right side, the darker the color, the higher the lift degree, the larger of the cycle, the higher the confidence degree.

Discussion

The R language is a common tool for data mining in modern statistics. In this study, R language was used as a tool to analyze the regularity of prescriptions for the treatment of liver cancer, and to find out the distribution rule of drug pairs and its correlation with the symptoms of liver cancer. The results showed that the common drugs used in the treatment of liver cancer are Hedyotis diffusa, Glycyrrhiza, Atractylodes, Scutellariae barbata, Lanceolata, Salvia, Curcuma Longa, Bupleurum, Poria and Peony root in turn. Cluster analyses divides drugs into 20 categories and find the Hedyotis diffusa was relatively independent, while the Poria, Atractylodes, Lanceolata and other drugs were often appeared together. Further used the correlation method to analyze the drug pairs, found that the common drug pairs used in liver cancer were Astragalus-Glycyrrhiza, Poria-Atractylodes, Lanceolata-Atractylodes, Lanceolata-Glycyrrhiza, Scutellariae barbata-Salvia, etc. Hedyotis diffusa with the function of clearing away heat, dispersing blood stasis, detoxification, relieved pain, and even anti-tumour[4]. Take the Hedyotis diffusa for association analyses, the results showed that it often constitutes drug pairs with Glycyrrhiza, Curcuma Longa and Lanceolata. Using liver pain as an example to represent the common symptoms of liver cancer for association analyses, the results showed that single drug of Gardenia, Bupleurum, Astragalus, Scutellariae barbata and Pinellia were often used for the pain of liver, while the drug pairs were Pinellia-Bupleurum, Hedyotis diffusa-Astragalus, Bupleurum-Poria, Scutellariae barbata-Bupleurum and Bupleurum-Curcuma Longa.

Liver cancer belongs to traditional Chinese medicine liver mass, spleen mass, zheng ji and tympanites[5,6]. The principle of soothing liver and strengthening spleen, qu shi hua yu and heat-clearing and detoxifying should be followed in most treatments[7]. Through data mining, the frequency statistics of commonly used drugs for liver cancer were carried out in this study. The results showed that most of the drugs were belong to complement deficiency and regulating vital energy, heat-clearing and detoxifying, blood-activating and stasis-dissolving, and clearing damp and promoting diuresis, which is consistent with the rules discovered through traditional Chinese medicine heritage assistant system[8,9,10]. Common drug pairs through clustering and association analyses, take the Hedyotis diffusa-Curcuma Longa as an example, the pharmacological displayed that Hedyotis diffusa with the function of heat-clearing and detoxifying, Xiaotong Sanjie, clearing damp and promoting diuresis. And the Curcuma Longa with the function of releasing the pain in the chest and hypochondrium, release jaundice and blood urination. Combine the Hedyotis diffusa with Curcuma Longa to the purpose of detoxicating and resolving a mass, remove pain and release

jaundice. Bupleurum is a good medicine for the treatment of stagnation of liver-qi and thoracic and rib distending pain. Pinellia with the function of dry dampness and disperse phlegm, drop inverse and stop vomit. The results show that both Bupleurum and Pinellia are common drugs for the treatment of liver pain, and they are compatible in drug pairs, which are consistent with clinical experience.

In conclusion, this study uses R language for data mining, focusing on drug pairs, validating the rationality of clinical experience and understanding the law of compatibility of traditional Chinese medicine, which provides a theoretical basis for further treatment of liver cancer and subsequent pharmacological research.

Acknowledgement

This project was from Scientific research projects for basic scientific research in Heilongjiang Provincial Universities in 2017, multidimensional analysis by drug pair-syndrome-symptom and disease and data mining of genuine regional drug in northeast china based on data warehouse.

Reference

- [1] Wang FS, Fan JG, Zhang Z, et al. The global burden of liver disease: The major impact of China[J]. *Hepatology*, 2014,60(6):2099-2108.
- [2] Chai KQ. Discussion on the clinical application of the method of reinforcing the vital energy and consolidating the constitution, removing phlegm and relieving depression in malignant tumor[J]. *Zhejiang Journal of Traditional Chinese Medicine*, 2013,48(11):785-787.
- [3] Zhang XY, Chai KQ, Chen JB, etc. Study on Chai Kequn's medication empirical in the treatment of Primary Hepatocellular carcinoma based on data mining[J]. *Zhejiang Journal of Traditional Chinese Medicine*, 2017,52(4):239-240.
- [4] Wang LL, Lin J. Inhibitory effect of volatile oil of *Hedyotis diffusa* on the proliferation of human hepatoma HepG2 cells[J]. *Fujian Journal of Traditional Chinese Medicine*, 2013, 44(3):60.
- [5] Song YY, Jiang J, Li AQ, etc. Literature analyses of traditional Chinese medicine syndrome differentiation in middle and advanced stage of primary hepatocellular carcinoma[J]. *Heilongjiang Journal of Traditional Chinese Medicine*, 2013(6):2.
- [6] Ren J, Di ZG. Analysis of the recipe rule of prescriptions for the treatment of primary hepatocellular carcinoma based on Inheritance auxiliary system of traditional Chinese Medicine[J]. *Chian Journal of Traditonal Chinese Medicine and Pharmacy*, 2015, 13(12):7-10.
- [7] Zhang T, Kang MR, Zhang RM. The experience of professor Liangshi Kang combine diseases with symptoms for the treatment of primary hepatocellular carcinoma[J]. *Chian Journal of Traditonal Chinese Medicine and Pharmacy*, 2012,27(12):3147.
- [8] Liu JH, Deng JF, He JR, etc. Literature analysis of the traditional Chinese medicine pharmacy rule in the treatment of primary hepatocellular carcinoma[J]. *Chian Journal of Traditonal Chinese Medicine and Pharmacy*, 2014, 55(18):1592.
- [9] Cao XJ, Sun CG, Li YJ, etc. Analysis of the recipe rule of prescriptions for the treatment of primary hepatocellular carcinoma by traditional Chinese medicine based on inheritance auxiliary system of traditional Chinese medicine[J]. *Chinese Archives of Traditional Chinese Medicine*, 2016, 34(1):69-72.
- [10] Liu JH, Wei ZH, Lv DY, etc. Study on the pharmacy rule by famous veteran teran doctors of TCM in the treatment of primary hepatocellular carcinoma based on data mining[J]. *Chian Journal of Traditonal Chinese Medicine and Pharmacy*, 2016, 31(1):58-61.