Quality Improvement Study

Medicine®

OPEN

Association rule mining for the ordered placement of traditional Chinese medicine containers

An experimental study

Chih-Wen Chen, MS[a,b,c], Chih-Fong Tsai, PhD[d], Yi-Hong Tsai, PhD[a], Yang-Chang Wu, PhD[e,f,*], Fang-Rong Chang, PhD, Dr.h.c.mult.[a,g,h,i,*]

Abstract

In traditional Chinese medicine (TCM) clinics, the pharmacists responsible for dispensing the herbal medicine usually find the desired ingredients based on positions of the shelves (racks; frames; stands). Generally, these containers are arranged in an alphabetical order depending on the herbal medicine they contain. However, certain related ingredients tend to be used together in many prescriptions, even though the containers may be stored far away from each other. This can cause problems, especially when there are many patients and/or the limited number of pharmacists. If the dispensing time takes longer, it is likely to impact the satisfaction of the patients' experience. Moreover, the stamina of the pharmacists will be consumed quickly.

In this study, we investigate on an association rule mining technology to improve efficiency in TCM dispensing based on the frequent pattern growth algorithm and try to identify which 2 or 3 herbal medicines will match together frequently in prescriptions. Furthermore, 3 experimental studies are conducted based on a dataset collected from a traditional Chinese medicine hospital. The dataset includes information for an entire year (2014), including 4 seasons and doctors. Afterward, a questionnaire on the usefulness of the extracted rules was administered to the pharmacists in the case hospital. The responses showed the mining results to be very valuable as a reference for the placement and ordering of the frames in the TCM pharmacies and drug stores.

Abbreviations: FP-growth = frequent pattern growth, TCM = traditional Chinese medicine.

Keywords: association rule mining, frequent pattern growth algorithm, medicinal storage containers, traditional Chinese medicine

1. Introduction

In addition to western medicine, a well-known and popular type of medical treatment for different kinds of diseases is traditional Chinese medicine (TCM) found more than 2500 years. Treatment can include various forms of herbal medicine, acupuncture, dietary therapy, and so on.[1,2]

In TCM clinics/hospitals, pharmacists dispense the herbal medicines required for each patient's prescription. The various kinds of herbal medicines located in shelves usually arranged in an alphabetical order depending on the corresponding containers; however, to find right medicines together in a prescription is the most time-consuming work during the dispensing process.

Editor: Muhammad Shahzad Aslam.

Editor: Patrick Wall

This study was approved by the review board of Municipal Chinese Medical Hospital.

The Apriori and FP-growth algorithm data (Association Rule Mining) used to support the findings of this study are included within the article, Supplementary Information File (Table 6), and are available from the corresponding author upon request.

The research project was supported by grants from the Kaohsiung Municipal Chinese Medical Hospital, Department of Health, Kaohsiung City Government, Kaohsiung, Taiwan (KMCMH-10401). This work was also supported by grants from the Ministry of Science and Technology of Taiwan, awarded to Fang-Rong Chang (Grant Nos. 1 MOST 108-2320-B-037-022-MY3, 108-2811-B-037-511, 109-2927-I-037-502), and awarded to Yang-Chang Wu (MOST 106-2622-B-037-003-CC2, 106-2320-B-037-007-MY3). In addition, this research was funded by the Drug Development and Value Creation Research Center, Kaohsiung Medical University & Department of Medical Research, Kaohsiung Medical University Hospital, awarded to Fang-Rong Chang (grant number: KMU-TC108A03-11). Special thanks to Dr. Hsiu-Chuan Yen, Dr. Chih-Hao Chang and Dr. Hong-Tien Hwang for assistance with the experiments.

Conflict of Interest: The authors of this work have nothing to disclose.

[a] Graduate Institute of Natural Products, College of Pharmacy, Kaohsiung Medical University, [b] Department of Pharmacy, Kaohsiung Municipal Chinese Medical Hospital 801, [c] Department of Nursing, Fooyin University, Kaohsiung, [d] Department of Information Management, National Central University 320, Taoyuan, [e] Graduate Institute of Integrated Medicine, China Medical University, [f] Chinese Medicine Research and Development Center, China Medical University Hospital, Taichung, [g] Drug Development and Value Creation Research Center, Kaohsiung Medical University, Kaohsiung, [h] Department of Marine Biotechnology and Resources, National Sun Yat-sen University, Kaohsiung, [i] Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan.

[*] Correspondence: Yang-Chang Wu, Graduate Institute of Integrated Medicine, China Medical University, Taichung 404, Taiwan (e-mail: yachwu@mail.cmu.edu.tw); Fang-Rong Chang, Graduate Institute of Natural Products, College of Pharmacy, Kaohsiung Medical University, Kaohsiung 807, Taiwan (e-mail: aaronfrc@kmu.edu.tw).

How to cite this article: Chen CW, Tsai CF, Tsai YH, Wu YC, Chang FR. Association rule mining for the ordered placement of traditional Chinese medicine containers: an experimental study. Medicine 2020;99:18(e20090).

Received: 17 July 2019 / Received in final form: 28 November 2019 / Accepted: 28 March 2020

http://dx.doi.org/10.1097/MD.0000000000020090

Problems may arise when the locations of herbal medicines frequently used together are far away from each other in the pharmacy, making them difficult to be accessed. It not only makes more time consuming in dispensing processes but also places a greater burden on the spirit and strength of pharmacists. The speed in dispensing is one of the main reasons that patients have to spend a long time in waiting their medicine.

In the western medicine domain, there are many automated systems designed to reduce prescription error rates and increase the efficiency for both pharmacy and nursing staffs.[3,4] However, very few automatic approaches or systems have been developed specifically and aimed at dealing with the TCM dispensing problems. In TCM pharmacy and drug stores, it becomes an important issue to be concerned.

Recently, Chou et al, 2012, an intelligent system for quality assurance in a TCM dispensing system was introduced using radio frequency identification (RFID) technique to double check the prescriptions and pharmacy dispensing procedures.[5] This system is successful in reducing possibility of human error and ensuring dispensing quality; however, pharmacists did not reduce their loading. The cost of initial investment to develop this kind of system is high for many small or medium sized TCM clinics/hospitals, pharmacy or drug stores. In addition, many pharmacists have insufficient background knowledge in information technology (IT), and thus lack of skills to effectively manage and maintain the system.

Therefore, the aim of this study is to introduce an approach to enhance efficiency in manual dispensing processes, which can find the right Chinese medicines quickly based on the association rule mining technique. This method is also known as market basket analysis and aimed at discovering some interesting relations between frequent items (ie, products) in a transactional dataset.[6,7] For the domain problem, the association rule mining technique is able to identify those herbal medicines frequently used together in various prescriptions (eg, when dispensing the medicine A, the medicine B or medicines B&C will be followed [The drug-paired theory of TCM. For example, bai-hua-she-she-cao, and ban-zhi-lian are always prescribed together, and han-lian-cao and nu-zhen-zi together, etc]).

TCM clinics/hospitals can refer to the discovered rules to relocate some frequently used medicinal containers into nearby positions on shelves. They should be able to reduce the time required to find the right medicines during the dispensing process as well be less taxing on the physical strength of the pharmacist. In addition, the waiting times of patients could also be reduced, which would enhance the satisfaction of the patient experience.

In this study, the well-known and efficient frequent pattern growth (FP-growth) association rule algorithm is used.[8,9] The dataset is collected from a TCM clinics/hospitals in Taiwan. Three experimental studies are conducted, which are based on analyzing data for the whole year, including 4 individual seasons, and different subsets for individual doctors for possible prescription-making preferences. The association rule mining technology including the Apriori and FP-growth algorithms were investigated.

2. Materials and methods

2.1. The dataset of a TCM hospital

The dataset collected from a case TCM hospital in Taiwan, contains 114,708 TCM prescriptions in the year 2014. In total, there are 753 different TCMs used in this dataset. Given the objectives of this study each data sample includes 4 columns related to the prescription used: the identity number of medicines, the name of medicines, the prescription date, and the doctor who made the prescription. No patient related information is included.

2.2. Association rule mining

Association rule mining is a type of machine learning method, which can produce some rules from a given dataset. In particular, the rules describe specific relations between variables in the dataset. This technique was introduced for the purpose of discovering regularities between products in large-scale transaction data in supermarkets.[6] In general, a transactional dataset contains a number of transactions, including the times at which customers make their purchase. Specifically, each transaction is regarded as a set of items.

An association rule is usually represented by a set of symbols, for example, {products A, B} → {product C}, which means that if a customer buys products A and B, then he or she is likely to also buy product C. Furthermore, the criteria of minimum support and confidence should be followed with an association rule.

Support means the frequency of appearance of a specific item set in the dataset. A rule with a high support value means that it appears frequently in the case dataset. Confidence refers to the probability of both A and B item sets appearing, and is used to measure the usefulness of the rule. Therefore, those rules that meet or exceed a pre-defined threshold, that is, the minimum support and confidence values, are found.

The following subsections describe 2 widely used algorithms, the Apriori and FP-growth algorithms.

2.2.1. The apriori algorithm. The Apriori algorithm was the earliest association rule mining algorithm. It was proposed for frequent itemset mining from transactional databases. It is used to identify frequent individual items in a chosen dataset and extend these individual items to larger item sets if they frequently appear in the database.[10]

In general, 2 steps are required to produce the association rules. In the first step, the whole datasets to identify candidate items until the high frequency item sets are identified. Next, related association rules are extracted based on the high frequency item sets.

The equations were detailed in the previous reference.[10] The pseudo code of the Apriori algorithm for generating the high frequency time sets. Note that C_κ means the k candidate item set and F_κ indicates the high frequency set of the κ item set. First of all, all high frequency item sets F_1 are identified after searching the whole dataset. Then, the "apriori-gen" function is used to generate the candidate k item set based on the identified high frequency $(k - 1)$ item sets. The minimum support and confidence values are used as the thresholds for keeping or removing related candidate item sets. The algorithm is stopped when no new high frequency itemsets are generated, that is, $F_k = \varnothing$.

2.2.2. The FP-growth algorithm. The FP-growth algorithm, in which FP stands for frequent pattern, is aimed at developing an extended prefix-tree to store crucial and quantitative information about frequent patterns. To build the FP-tree, each data record is read and its corresponding route on the FP-tree recorded. Since different records can have the same items, some routes can partially overlap. More records produce more overlapping routes, which compresses the structure of the FP-tree.

Specifically, this algorithm uses a partitioning-based and divide-and-conquer method rather than an Apriori-like bottom-up search technique for the generation of frequent itemset combinations. This produces a substantial reduction in the search cost, which allows computers to efficiently extract high frequency itemsets from their main memory without access to hard disc drives, which would require much longer processing times.

The FP-growth algorithm as proposed by Han (2000) and Han et al (2004) is one of the most representative association rule mining algorithms.[8,9] Earlier, Wur and Leu (1999) proposed a Boolean algorithm, related to the FP-growth algorithm, for mining the association rules in large databases.[11] Tsay and Chang-Chien (2004) introduced a cluster-decomposition association rule (CDAR) algorithm for decomposing the transactional databases into related clusters for the generation of association rules.[12] These 3 algorithms have been proven to be more efficient than the Apriori algorithm. To process the FP-growth algorithms, the Weka data mining software is used.[13]

2.2.3. A survey to the response after mining studies. After the studies of extraction of association rules, we created a questionnaire for related doctors who provided the prescriptions in 2014. It was to assess the reliability and usefulness of the association rules from the extracted dataset of 2014 (questions 1–4), 4 different seasons (questions 5–7), and individual doctors' prescriptions (questions 8–10). The questions are listed below:

(1) Do you think that the extracted association rules can be provided as referenceable value for the case hospital? Why?
(2) Do you think that the extracted rules with higher minimum supporting values are more useful than those lower ones?
(3) Can you provide some examples (ie, association rules) to explain their meanings relate to TCMs?
(4) Apart from the optimization of the permuted positions of TCMs on shelves in the pharmacy, what else of the circumstance(s) can also be improved or applied according to the extracted association rules?
(5) Were there any extracted association rules that you never notice before? If yes, please provide some examples.
(6) Did the association rules extracted from 4 different seasons match the clinical medication? If yes, please provide some examples.
(7) Do you think that the association rules extracted from individual doctors' prescriptions can be provided as referenceable value for the case hospital? Why?
(8) The association rules extracted from individual doctors revealed that some doctors have their own preferred prescriptions on some pairs of herbal medicines. Was this situation due to the prescribing habits of some specialists, or any other reason(s)?
(9) What else of the circumstance(s) can be applied according to the association rules extracted from individual doctors' prescriptions?

3. Results and discussion

3.1. Experimental setup

In this study, 3 experimental studies are conducted. The first one is designed to extract the association rules from the whole dataset of the year 2014. The second one is aimed at extracting the association rules individually for the 4 seasons for purposes of comparison, in order to examine whether there is a seasonal correspondence for specific prescriptions. Thus, the whole dataset is divided into 4 independent datasets according to the prescription date, corresponding to the 4 different, January to March, April to June, July to September, and October to December, labeled Q1, Q2, Q3, and Q4, respectively. Finally, the third study is aimed at extracting the association rules from the subsets corresponding to individual doctors in the case hospital. For this purpose, the whole dataset is divided into 29 subsets, because there are 29 doctors. The relatively small subsets, that is, those containing fewer than 1000 data samples, are omitted, leaving 8 subsets (corresponding to 8 doctors) to be used for association rule mining. The association rules extracted from the 8 subsets are compared to be analyzed by the preferences of different doctors when writing prescriptions.

To process the FP-growth algorithms, the related parameters are all follows: first of all, the minimum confidence level is set to 0.1 in order to filter the usefulness rules. In addition, the maximum number of rules is set to 100. The minimum support is set to start at 0.1 and this value is gradually reduced to 0.06 (for study 1) and 0.05 (for studies 2 and 3) at intervals of 0.01.

3.2. Experimental results: the study I

The first study is aimed at extracted association rules from the entire dataset for the whole year. Based on the minimum confidence value of 0.1, the minimum support value is set to begin from 0.1 and then reduced gradually to 0.01. The number of rules starts from 56 and increases to 264, based on the minimum support value of 0.06 (Fig. 1).

To avoid the problems of too many association rules cause information overload for the pharmacists, in this study the number of rules to be extracted is limited to 100. Note that it only takes about 10 seconds to process the 114,708 prescriptions. Tables 1 and 2 show 2 examples of lists of the associated herbal medicines obtained from the extracted rules suing minimum support values of 0.01 and 0.09, respectively. Note that for simplicity we only show the herbal medicine identity number.

According to the questionnaire results for the usefulness of the extracted rules, the doctors in the case clinic think that most of these rules are useful, even when the minimum support and confidence values are very low. For example, with a minimum support value of 0.08, it is reasonable that the associated herbal medicines 11403 Suan-Zao-Ren-Tang (Suan-Zao-Ren-Tang)

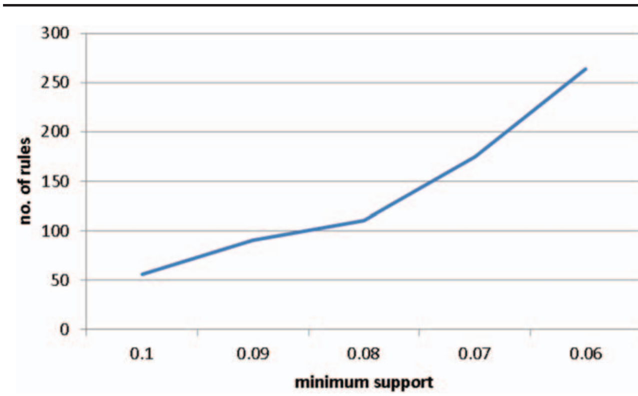

Figure 1. The number of rules obtained with different minimum support values.

Table 1

The associated TCM ingredients listed using a minimum support value of 0.01 (For practical use, the English-translated name of TCMs were applied on this table. The scientific name of TCMs were detailed in the Table 6.[14]).

TCM code/medicine 1	TCM code/ medicine 2	Confidence value
20528 Ban-Zhi-Lian	20505 Bai-Hua-She-She-Cao	0.89
20715 Han-Lian-Cao	20313 Nu-Zhen-Zi	0.88
20602 He-Huan-Pi	20812 Ye-Jiao-Teng	0.8
10713 Xing-Su-Yin (Pediatric)	20915 Qian-Hu	0.74
21115 Zhi-Zi	21205 Hwang-Chyn	0.69
20313 Nu-Zhen-Zi	20715 Han-Lian-Cao	0.68
20505 Bai-Hua-She-She-Cao	20528 Ban-Zhi-Lian	0.67
10527 Zheng-Gu-Zi-Jin-Dan	11303 Dang-Gui-Nian-Tong-Tang	0.66
20306 Shan-Yao	21211 Tu-Si-Zi	0.6
21015 Hai-Piao-Xiao	20703 Bei-Mu	0.58
21211 Tu-Si-Zi	20306 Shan-Yao	0.53
20910 Zhi-Ke	20915 Qian-Hu	0.52
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.52
21015 Hai-Piao-Xiao	10507 Bann-Shiah-Shein-Shin-Tan	0.51
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.48
10410 Zhi-Sou-San	20703 Bei-Mu	0.42
20915 Qian-Hu	20910 Zhi-Ke	0.39
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.4
20915 Qian-Hu	20703 Bei-Mu	0.37
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.35
10704 Shaur-Yaw-Gan-Tsao-Tan	11211 Shu-Ging-Hwa-Shiueh-Tan	0.34
11303 Dang-Gui-Nian-Tong-Tang	10527 Zheng-Gu-Zi-Jin-Dan	0.33
11303 Dang-Gui-Nian-Tong-Tang	11211 Shu-Ging-Hwa-Shiueh-Tan	0.33
10308 San-Huang-Xie-Xin-Tang	23001 LipoCol Forte Capsule	0.32
10507 Bann-Shiah-Shein-Shin-Tan	21015 Hai-Piao-Xiao	0.31
21004 Fu-Shen	11403 Suan-Zao-Ren-Tang	0.3

TCM = traditional Chinese medicine.

and 21004 Fu-Shen (Fu-Shen) be dispensed together. Similarly, the associated herb medicines 11211 Shu-Ging-Hwa-Shiueh-Tang (Shu-Ging-Hwa-Shiueh-Tang) and 11303 Dang-Gui-Nian-Tong-Tang (Dang-Gui-Nian-Tong-Tang) are reasonable dispensed together, even though the confidence value is 0.31. In addition, when the confidence value is 0.3 and the minimum support value is 0.07, the association rule that the herbal medicines 11501 Jia-Wei-Xiao-Yao-San (Jia-Wei-Xiao-Yao-San) and 21028 Yi-Mu-Cao (Yi-Mu-Cao) are dispensed together is deemed a very valuable association rule by doctors at the case clinic.

Another question asked of the doctors is which minimum support value is more useful? The answer is that the association rules obtained using different minimum support values all have highly reference values. These rules act as guidelines for the arrangement of the medicinal containers on the shelves of the TCM pharmacy. Additionally, we make sure that TCMs should place at desired locations on the shelves and those TCMs with higher support values according the algorithm mining results should be located together or nearby in the TCM pharmacy, and pharmacists can access easier.

3.3. Experimental results: the study II

The second study focuses on dividing the original dataset into 4 subsets based on the 4 seasons. Tables 3 and 4 show the results for 2 examples of the association rules extracted for the 4 seasons using minimum support values of 0.1 and 0.09, respectively.

We found that some associated herb medicines appeared more frequently in specific seasons. For example, Ban-Zhi-Lian (20528 Ban-Zhi-Lian) → Bai-Hua-She-She-Cao (20505 Bai-Hua-She-She-Cao) or Tu-Si-Zi (21211 Tu-Si-Zi) → Zuo-Gui-Wan (10525) are prescribed less in Q1. As explained by the doctors, the nature of the season, for example, growth in Spring, development in Summer, fade in Autumn, and storage in Winter, is one of the main references for making prescriptions. In Q1, for example, herbal medicines are used to strengthen the spleen (growth) for liver disease. However, since spring is the season for germination, the dispensing of chilling herbs (In TCM, certain foods and herbs are considered to heat or cool the body) is avoided. On the other hand, Autumn is the chilling season, and prescriptions would not usually contain “germinal” or “growth” medicines.

Therefore, the associated Bai-Hua-She-She-Cao (20505 Bai-Hua-She-She-Cao) and Ban-Zhi-Lian (20528 Ban-Zhi-Lian), which can clear away pathogenic heat and remove toxins (ie, detoxification), are less dispensed in Q1. However, in the fall as the weather becomes cooler and the temperature difference between day and night becomes bigger, asthma, allergy, and decreased immunity related diseases are more likely to be occurred. For these, it is necessary to replenish depleted reserves and improve the natural functions of the human body, thus the prescriptions usually contain herbal medicines for replenishing the kidney and the yin, using herbs such as Zuo-Gui-Wan (10525 Zuo-Gui-Wan), Shan-Yao (20306 Shan-Yao), Tu-Si-Zi (21211 Tu-Si-Zi), Han-Lian-Cao (20715 Han-Lian-Cao), and Nu-Zhen-Zi (20313 Nu-Zhen-Zi).

According to the doctors’ opinions, the association rules based on different seasons allow them to understand the relationship between the epidemiology and the distribution of diseases depending on the seasons.

3.4. Experimental results: the study III

In the final study, the original dataset is divided into 29 subsets based on 29 different doctors, respectively. Since the number of data samples in some subsets is relatively small, for example, less than 1000, it would be impossible to determine the value of the extracted rules. Therefore, we selected subsets containing more than 1000 data samples. Eight subsets met this requirement. In particular, there are 3 top doctors (denoted as A, B, and C) who wrote the greatest number of prescriptions, which are 21089, 16219, and 7489, respectively. Based on a minimum support value of 0.05, the number of extracted rules from these 3 subsets is only 2 to 4 (as shown in Table 5).

As explained by the doctors, doctors in specialized clinics, such as for the elderly, for treating obesity, cancer clinics, and so on, usually use fixed prescriptions. Moreover, based on long term experience with the same types of patients, such as infertility, women’s diseases, acupuncture treatments, and so on, doctors will prescribe similar prescriptions.

There are 42 rules extracted based on a minimum support value of 0.05 from the subset of doctor D who only made 2179 prescriptions. Similarly, for the subset of doctor E who wrote about 5000 prescriptions, the number of extracted rules for a minimum support value of 0.06 is 100. The prescriptions made by doctors D and E showed a less fixed pattern. This indicates that they were visited by a large variety of patients.

The doctors’ responses to the questions regarding the usefulness of the extracted rules for each doctor indicate that

Table 2

The associated TCM ingredients listed using a minimum support value of 0.009.

TCM code/medicine 1	TCM code/medicine 2	TCM code/medicine 3	Confidence value
20528 Ban-Zhi-Lian		20505 Bai-Hua-She-She-Cao	0.89
20715 Han-Lian-Cao		20313 Nu-Zhen-Zi	0.88
20602 He-Huan-Pi		20812 Ye-Jiao-Teng	0.8
10713 Xing-Su-Yin (Pediatric)		20915 Qian-Hu	0.74
11403 Suan-Zao-Ren-Tang	21004 Fu-Shen	20812 Ye-Jiao-Teng	0.74
20703 Bei-Mu	10507 Bann-Shiah-Shein-Shin-Tan	21015 Hai-Piao-Xiao	0.7
21115 Zhi-Zi		21205 Hwang-Chyn	0.69
20313 Nu-Zhen-Zi		20715 Han-Lian-Cao	0.68
20313 Nu-Zhen-Zi		20715 Han-Lian-Cao	0.68
20505 Bai-Hua-She-She-Cao		20528 Ban-Zhi-Lian	0.67
10527 Zheng-Gu-Zi-Jin-Dan		11303 Dang-Gui-Nian-Tong-Tang	0.66
10507 Bann-Shiah-Shein-Shin-Tan	21015 Hai-Piao-Xiao	20703 Bei-Mu	0.62
20306 Shan-Yao		21211 Tu-Si-Zi	0.6
21015 Hai-Piao-Xiao		20703 Bei-Mu	0.58
20703 Bei-Mu	21015 Hai-Piao-Xiao	10507 Bann-Shiah-Shein-Shin-Tan	0.54
21211 Tu-Si-Zi		20306 Shan-Yao	0.53
20910 Zhi-Ke		20915 Qian-Hu	0.52
21004 Fu-Shen		20812 Ye-Jiao-Teng	0.52
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	21004 Fu-Shen	0.52
20713 Chi-Shao		21028 Yi-Mu-Cao	0.51
20812 Ye-Jiao-Teng		21004 Fu-Shen	0.48
22104 Xu-Duan-Xu-Duan		20702 Du-Zhong	0.44
20812 Ye-Jiao-Teng	21004 Fu-Shen	11403 Suan-Zao-Ren-Tang	0.42
10410 Zhi-Sou-San		20703 Bei-Mu	0.42
20812 Ye-Jiao-Teng		11403 Suan-Zao-Ren-Tang	0.4
21028 Yi-Mu-Cao		20713 Chi-Shao	0.38
20915 Qian-Hu		20703 Bei-Mu	0.37
21020 Jie-Geng		21205 Hwang-Chyn	0.36
11403 Suan-Zao-Ren-Tang		20812 Ye-Jiao-Teng	0.35
20915 Qian-Hu		20910 Zhi-Ke	0.35
21020 Jie-Geng		20703 Bei-Mu	0.35
21111 Liao-Qiao		21205 Hwang-Chyn	0.34
10704 Shaur-Yaw-Gan-Tsao-Tan		11211 Shu-Ging-Hwa-Shiueh-Tan	0.34
11303 Dang-Gui-Nian-Tong-Tang		10527 Zheng-Gu-Zi-Jin-Dan	0.33
11303 Dang-Gui-Nian-Tong-Tang		11211 Shu-Ging-Hwa-Shiueh-Tan	0.33
10308 San-Huang-Xie-Xin-Tang		23001 LipoCol Forte Capsule	0.32
21015 Hai-Piao-Xiao	20703 Bei-Mu	10507 Bann-Shiah-Shein-Shin-Tan	0.32
10507 Bann-Shiah-Shein-Shin-Tan		21015 Hai-Piao-Xiao	0.31
21004 Fu-Shen		11403 Suan-Zao-Ren-Tang	0.3

TCM = traditional Chinese medicine.

they are helpful. For instance, after these specific doctors' survey, pharmacists can arrange and improve the accessibility of TCMs in the specific clinics and assist in checking and finding right items of the medicines often dispensed by the doctors. This will also alleviate the problems of having insufficient medicine supply and storage during the dispensing process.

In the past, TCM doctors whose basic knowledge is usually obtained from related historical records and ancient medicinal books. A TCM formula is usually composed by 4 roles of drugs, such as sovereign, minister, assistant and courier. In clinical practice, parts of prescriptions issued with traditional pair theories of Chinese medicines, for example, medicines A and B were usually dispensed together for improving the medical efficacy or reducing some adverse effects. However, there were not any systematic and scientific methodology reported for clarifying and analyzing the influences between the dispensing efficiency and pair theories of Chinese medicines. In this study, we clarify not only the traditional pair theories of Chinese medicines but also patterns of pair (triple) medicines in modern and real TCM practices.

Besides the frequently dispensed pair TCMs found in this study, Chinese medication usually refers to 4 seasons which is related to solar terms for human bodies. For example, chill medicines are fewer used in prescriptions during spring, whereas Yin or kidney related ones are usually prescribed in autumn and winter times. The popular pair prescriptions used for seasons are also recommended.

In addition to the special clinics, TCM doctors always concern patients' personal background characteristics and prescribe TCMs based on the nature and flavor of the dispensed medicines. According to our analytical results, doctors and pharmacists can not only understand the modern pairs (triples) of TCMs but also different types of prescriptions. Through information provided, doctors and pharmacists can have more discussion and lead safer uses in TCM medication on patients.

For the implementation issue, the research findings based on the extracted association rules can bring some guidance for positioning TCMs on shelves in pharmacy. In practice, while setting a pharmacist at the central position, most frequently used medicines and pairs of together targets with higher confidence

Table 3

The associated herbal medicines for 4 seasons using a minimum support value of 0.1.

TCM code/medicine 1	TCM code/medicine 2	Confidence value
Q1 (January–March)		
20910 Zhi-Ke	20915 Qian-Hu	0.54
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.47
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.44
20915 Qian-Hu	20910 Zhi-Ke	0.4
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.36
20915 Qian-Hu	20703 Bei-Mu	0.36
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.3
Q2 (April–June)		
10527 Zheng-Gu-Zi-Jin-Dan	11303 Dang-Gui-Nian-Tong-Tang	0.73
20306 Shan-Yao	21211 Tu-Si-Zi	0.65
21211 Tu-Si-Zi	20306 Shan-Yao	0.63
20910 Zhi-Ke	20915 Qian-Hu	0.48
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.44
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.43
20915 Qian-Hu	20910 Zhi-Ke	0.41
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.33
11303 Dang-Gui-Nian-Tong-Tang	10527 Zheng-Gu-Zi-Jin-Dan	0.33
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.31
Q3 (July–September)		
20715 Han-Lian-Cao	20313 Nu-Zhen-Zi	0.93
20313 Nu-Zhen-Zi	20715 Han-Lian-Cao	0.76
20306 Shan-Yao	21211 Tu-Si-Zi	0.67
21211 Tu-Si-Zi	20306 Shan-Yao	0.67
21015 Hai-Piao-Xiao	20703 Bei-Mu	0.58
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.57
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.46
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.35
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.31
Q4 (October–December)		
20715 Han-Lian-Cao	20313 Nu-Zhen-Zi	0.86
20313 Nu-Zhen-Zi	20715 Han-Lian-Cao	0.72
21015 Hai-Piao-Xiao	20703 Bei-Mu	0.63
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.52
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.5
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.37
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.31

TCM = traditional Chinese medicine.

values should be placed at closer spots on shelves nearby him/her. This confidence-value dependent order would be performed and could be considered as the optimized implementation.

Under this designed system, the operation of dispensing will be certainly affected when any one of the positioned TCM on shelf is adjusted. Pharmacists need time to adopt rules and the system, and then the efficiency of the work can be improved. To improve the new system, "3 checks and 5 rights" which is the principle of final checks after dispensing should also be also conducted.[15] Moreover, the Electric Barcode Check System (EBCS) used in Chinese medication should be correctly operated in order to avoid any incorrect filling to prescriptions.

4. Conclusions

In the present study, we reported an investigation for improving the dispensing efficiency of positioning TCMs on shelves of pharmacy depending on factors of pair theories of TCM, 4 seasons, and doctors' preferences. The statistical technology conducted, Association Rule Mining Algorithms, is a reliable tool of primarily focused on finding frequent co-occurring associations among collections of items. The results presented in this work are the first systematic research in the world with direct evidences about how to enhance the dispensing speed, reduce the fatigue of pharmacists, prevent possible errors when filling prescriptions, and even related to avoid the incidence of medical dispute. These mining data provide a very valuable reference for elevating the healthcare quality of TCM medications.

However, some issues are worthy to be noticed:

Table 4

The associated herbal medicines for 4 seasons using a minimum support value of 0.09.

TCM code/medicine 1	TCM code/medicine 2	Confidence value
Q1 (January–March)		
10713 Xing-Su-Yin (Pediatric)	20915 Qian-Hu	0.64
20910 Zhi-Ke	20915 Qian-Hu	0.54
21015 Hai-Piao-Xiao	20703 Bei-Mu	0.48
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.47
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.44
20915 Qian-Hu	20910 Zhi-Ke	0.4
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.36
10410 Zhi-Sou-San	20703 Bei-Mu	0.36
20915 Qian-Hu	20703 Bei-Mu	0.36
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.3
Q2 (April–June)		
10527 Zheng-Gu-Zi-Jin-Dan	11303 Dang-Gui-Nian-Tong-Tang	0.73
20306 Shan-Yao	21211 Tu-Si-Zi	0.65
21211 Tu-Si-Zi	20306 Shan-Yao	0.63
21015 Hai-Piao-Xiao	20703 Bei-Mu	0.53
20910 Zhi-Ke	20915 Qian-Hu	0.48
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.44
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.43
20915 Qian-Hu	20910 Zhi-Ke	0.41
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.33
11303 Dang-Gui-Nian-Tong-Tang	10527 Zheng-Gu-Zi-Jin-Dan	0.33
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.31
Q3 (July–September)		
20715 Han-Lian-Cao	20313 Nu-Zhen-Zi	0.93
20313 Nu-Zhen-Zi	20715 Han-Lian-Cao	0.76
20306 Shan-Yao	21211 Tu-Si-Zi	0.67
21211 Tu-Si-Zi	20306 Shan-Yao	0.67
21015 Hai-Piao-Xiao	20703 Bei-Mu	0.58
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.57
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.46
21211 Tu-Si-Zi	10525 Zuo-Gui-Wan	0.36
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.35
10525 Zuo-Gui-Wan	21211 Tu-Si-Zi	0.34
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.31
Q4 (October–December)		
20528 Ban-Zhi-Lian	20505 Bai-Hua-She-She-Cao	0.92
20715 Han-Lian-Cao	20313 Nu-Zhen-Zi	0.86
20313 Nu-Zhen-Zi	20715 Han-Lian-Cao	0.72
20505 Bai-Hua-She-She-Cao	20528 Ban-Zhi-Lian	0.7
21015 Hai-Piao-Xiao	20703 Bei-Mu	0.63
21004 Fu-Shen	20812 Ye-Jiao-Teng	0.52
20910 Zhi-Ke	20915 Qian-Hu	0.5
20812 Ye-Jiao-Teng	21004 Fu-Shen	0.5
20812 Ye-Jiao-Teng	11403 Suan-Zao-Ren-Tang	0.37
20915 Qian-Hu	20910 Zhi-Ke	0.32
11403 Suan-Zao-Ren-Tang	20812 Ye-Jiao-Teng	0.31

TCM = traditional Chinese medicine.

Table 5

The extracted rules for the top 3 doctors.

Dr.	Min. support	Confidence	Medicine 1	Medicine 2
A	0.05	0.83	20812 Ye-Jiao-Teng	21004 Fu-Shen
		0.69	20306 Shan-Yao	21211 Tu-Si-Zi
		0.63	21211 Tu-Si-Zi	20306 Shan-Yao
		0.57	21004 Fu-Shen	20812 Ye-Jiao-Teng
B	0.06	0.9	20910 Zhi-Ke	20915 Qian-Hu
		0.62	20915 Qian-Hu	20910 Zhi-Ke
	0.05	0.9	20910 Zhi-Ke	20915 Qian-Hu
		0.62	20915 Qian-Hu	20910 Zhi-Ke
C	0.07	0.49	21205 Hwang-Chyn	20310 Da-Huang
		0.36	20310 Da-Huang	21205 Hwang-Chyn
	0.06	0.49	21205 Hwang-Chyn	20310 Da-Huang
		0.36	20310 Da-Huang	21205 Hwang-Chyn
	0.05	0.58	20708 Mu-Li	21410 Suan-Zao-Ren
		0.49	21205 Hwang-Chyn	20310 Da-Huang
		0.42	21410 Suan-Zao-Ren	20708 Mu-Li
		0.36	20310 Da-Huang	21205 Hwang-Chyn

Table 6

The scientific names of TCMs listed in Table 1.

Compound Chinese medicines

No.	Chinese name	Ingredients and the scientific names
10308	San-Huang-Xie-Xin-Tang	*Rhizoma rhei*, *Rhizoma coptidis*, and *Scutellaria baicalensis*
10410	Zhi-Sou-San	*Platycodon grandiflorum*, herba schizonepetae, *Aster tataricus*, *Stemona sessilifolia*, *Cynanchum stauntonii*, *Glycyrrhiza uralensis* and *Citrus reticulata*
10507	Ban-Xia-Xie-Xin-Tang	*Pinellia ternate*, *Scutellaria baicalnsis*, *Zingiber officinale*, *Panax ginseng*, *Glycyrrhiza uralensis*, *Coptis chinensis* and *Ziziphus jujuba*
10527	Zheng-Gu-Zi-Jin-Dan	*Eugenia caryophyllata*, *Aucklandia lappa*, *Daemonorops draco*, *Acacia catechu*, *Rheum officinale*, *Carthamus tinctorius*, *Angelica sinensis*, *Nelumbo nucifera*, *Poria cocos*, *Paeonia lactiflora*, *Paeonia suffruticosa* and *Glycyrrhiza uralensis*
10704	Shao-Yao-Gan-Cao-Tang	*Paeonia lactiflora* and *Glycyrrhiza uralensis*
10713	Xing-Su-Yin (Pediatric)	*Prunus armeniaca*, *Perilla frutescens*, *Peucedanum praeruptorum*, *Citrus reticulate*, *Glycyrrhiza uralensis* , *Citrus aurantium*, *Platycodon grandiflorum*, *Morus alba*, *Ophiopogon japonicas*, *Scutellaria baicalnsis*, *Fritillaria thunbergii* and *Zingiber officinale*
11211	Shu-Jing-Huo-Sie-Tang	*Angelica sinensis*, *Glycyrrhiza uralensis* , *Paeonia lactiflora*, *Rehmannia glutinosa*, *Atractylodes lancea*, *Achyranthes bidentate*, *Citrus reticulate*, *Prunus persica*, *Clematis chinensis*, *Ligusticum chuanxiong*, *Stephania tetrandra*, *Notopterygium incisum*, *Angelica dahurica*, *Gentiana scabra*, *Poria cocos* and *Zingiber officinale*
11303	Dang-Guei-Nian-Togn-Tang	*Artemisia capillaris*, *Notopterygium incisum*, *Saposhnikovia divaricate*, *Cimicifuga heracleifolia*, *Pueraria lobate*, *Atractylodes lancea*, *Atractylodes macrocephala*, *Glycyrrhiza uralensis*, *Scutellaria baicalnsis*, *Sophora flavescent*, *Anemarrhena asphodeloides*, *Angelica sinensis*, *Polyporus umbellatus* and *Alisma orientalis*
11403	Suan-Zao-Ren-Tang	*Ziziphus jujube*, *Glycyrrhiza uralensis*, *Anemarrhena asphodeloides*, *Poria cocos* and *Ligusticum chuanxiong*
11501	Jia-Wei-Xiao-Yao-San	*Angelica sinensis*, *Poria cocos*, *Gardenia jasminoides*, *Mentha haplocaly*, *Paeonia lactiflora*, *Bupleurum chinensis*, *Glycyrrhiza uralensis*, *Atractylodes macrocephala*, *Paeonia suffruticosa* and *Zingiber officinale*

Single Chinese medicines

No.	Chinese name	Latin name	Scientific name	English name
20306	Shan-Yao	Dioscorea Rhizoma	*Dioscorea opposita*	Chinese Yam
20313	Nu-Zhen-Zi	Ligustri Lucidi Fructus	*Ligustrum lucidum*	Glossy Privet Fruit
20505	Bai-Hua-She-She-Cao	Hedyotidis Diffusae Herba	*Oldenlandia diffusa*	
20528	Ban-Zhi-Lian	Scutellariae Barbatae Herba	*Scutellaria barbata*	Barbed Skullcap Herb
20602	He-Huan-Pi	Albiziae Cortex	*Albizzia julibrissin*	Silktree Albizzia Bark
20702	Du-Zhong	Eucommiae Cortex	*Eucommia ulmoides*	Eucommia Bark
20703	Bei-Mu	Fritillariae Thunbergii Bulbus	*Fritillaria thunbergii*	Thunberg Fritillary Bulb
20713	Chi-Shao	Paeoniae Rubra Radix	*Paeoniae rubra*	Red Peony Root
20715	Han-Lian-Cao	Ecliptae Herba	*Eclipta prostrata*	Yerbadetajo Herb
20812	Ye-Jiao-Teng	Polygoni Multiflori Caulis	*Fallopia multiflora*	Tuber Fleeceflower Stem
20910	Zhi-Ke	Citri Immaturus Fructus	*Citrus aurantium*	Bitter Orange
20915	Qian-Hu	Peucedani Radix	*Peucedanum praeruptorum/ Peucedanum decursivum*	Hogfennel Root
21004	Fu-Shen	Poria	*Poria Cocos*	Indian Bread with Pine
21015	Hai-Piao-Xiao	Sepiae Os	*Sepiella maindroni/ Sepia esculenta*	Cuttlebone
21020	Jie-Geng	Schizonepetae Herba	*Schizonepeta tenuifolia*	Schizonepeta Herb
21028	Yi-Mu-Cao	Leonuri Herba	*Leonurus japonicus*	Motherwort Herb
21111	Liao-Qiao	Forsythiae Fructus	*Forsythia suspensa*	Forsythia Fruit
21115	Zhi-Zi	Gardeniae Fructus	*Gardenia jasminoides*	Capejasmine Fruit
21205	Hwang-Chyn	Scutellariae Radix	*Scutellaria baicalensis*	Scutellaria Root
21211	Tu-Si-Zi	Cuscutae Semen	*Cuscuta australis*	Chinese Dodder Seed
22104	Shiuh-Duann	Dipsaci Radix	*Dipsacus asperoides*	Dipsacus Root
23001	LipoCol Forte Cap.		*Monascus purpureus*	

TCM = traditional Chinese medicine.

(1) Different theories or experiences of TCM might be taught. The concepts of medication may not easy to be practiced in consistent.[16] Additional factors should be concerned for the datasets;
(2) Personal habits, such as the right- or left-handed pharmacists, did not be considered in this work;
(3) Regular treatment and theories of TCM in acupuncture points and meridians were not considered;
(4) The analyses of this study focused on concentrated TCM (or scientific Chinese medicine, extraction processed products).

TCMs from other resources like as decoctions of raw TCM materials and formulas are not included. These items can be considered as limitations for prospects in future.

Author contributions

Conceptualization: Fang-Rong Chang.
Data curation: Chih-Wen Chen.
Formal analysis: Chih-Wen Chen, Chih-Fong Tsai, Yi-Hong Tsai, Yang-Chang Wu.
Funding acquisition: Chih-Wen Chen, Yang-Chang Wu, Fang-Rong Chang.
Investigation: Chih-Wen Chen.
Methodology: Chih-Wen Chen, Yi-Hong Tsai, Yang-Chang Wu.
Project administration: Fang-Rong Chang.
Software: Chih-Wen Chen, Chih-Fong Tsai, Yi-Hong Tsai.
Validation: Chih-Wen Chen.
Writing – original draft: Chih-Wen Chen, Chih-Fong Tsai, Yang-Chang Wu.
Writing – review and editing: Fang-Rong Chang.
Fang-Rong Chang orcid: 0000-0003-2549-4193.

References

[1] Beinfield H, Komgold E. Between Heaven and Earth: A Guide to Chinese Medicine. New York, USA: Random House Publishing Group; 1992.
[2] Kaptchuk TJ. The Web That Has No Weaver: Understanding Chinese Medicine. New York, USA: McGraw-Hill Education; 2000.
[3] Perini VJ, Vermeulen LC. Comparison of automated medication-management systems. Am J Hosp Pharm 1993;51:1883–91.
[4] Murray MJ. Information technology: the infrastructure for improvements to the medication-use process. Am J Health Syst Pharm 2000;57:565–71.
[5] Chou SY, Hwang KY, Shieh SC. Stjepandic J, Rock G, Bil C. Design of an intelligent system to improve traditional Chinese medicine dispensing practice. Boncurrent Engineering Approaches for Sustainable Product Development in a Multi-disciplinary Environment Germany: Springer; 2012;657–66.
[6] Agrawal R, Imielinski T, Swami A. Buneman P, Jajodia S. Mining association rules between sets of items in large databases. Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data New York, USA: ACM Press; 1993;207–16.
[7] Tan PM, Steinbach M, Kumar V. Introduction to Data Mining. Boston, USA: Addision-Wesley; 2014;154–96.
[8] Han J. Dunham M, Naughton J, Chen WD, Koudas N. Mining frequent patterns without candidate generation. Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data New York, USA: ACM Press; 2000;1–2.
[9] Han J, Pei J, Yin Y, et al. Mining frequent patterns without candidate generation: a frequent-pattern tree approach. Data Mining Knowl Discov 2004;8:53–87.
[10] Agrawal R, Srikant R. Bocca JB, Jarke M, Jarke M, Zaniolo CA. Fast algorithms for mining association rules in large databases. Proceedings of the 20th International Conference on Very Large Data Bases, Edited by Bocca JB, Jarke M, Jarke M, Zaniolo CA. San Francisco, USA: Morgan Kaufmann; 1994;487–499.
[11] Wur SY, Leu Y. Chen ALP, Lochovsky FH. An effective Boolean algorithm for mining association rules in large databases. Proceedings of the 6th International Conference on Advanced Systems for Advanced Applications, Edited by Chen ALP and Lochovsky FH, New York, USA: IEEE; 1999;179–186.
[12] Tsay YJ, Chang-Chien YW. An efficient cluster and decomposition algorithm for mining association rules. Inf Sci 2004;160:161–71.
[13] Witten IH, Frank E, Hall MA et al. Algorithms: the basic methods, in Data mining: practical machine learning tools and techniques, 4th ed. Edited by Kent C, Burlington, Massachusetts, USA, 2007.
[14] Ministry of Health and WelfareTaiwan Herbal Pharmacopeia. 2nd ed. Taipei, Taiwan: Health and Welfare; 2013.
[15] Choo J, Hutchinson A, Bucknall T. Nurses' role in medication safety. J Nurs Manag 2010;18:853–61.
[16] Zhu B, Wang H. Basic theories of traditional Chinese medicine. Singing Dragon, London, UK, 2010.