International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 5441-5451

https://doi.org/10.70517/ijhsa464383

Methods for constructing a knowledge map of the relationship between acupuncture points and diseases

Yue Liu[1,*] and Yunzhi Zhang[2]
[1] Hainan Vocational University of Science and Technology, Haikou, Hainan, 571101, China
[2] The 928th Hospital of People's Liberation Army Joint Logistic Support Force, Haikou, Hainan, 571101, China
Corresponding authors: (e-mail: hkd2022138011@126.com).

Abstract Traditional Chinese medicine acupuncture, as a unique medical system in China, has formed a rich clinical experience and theoretical system after thousands of years of development, and modern medical research has confirmed that acupuncture has significant efficacy in the treatment of many diseases. In this paper, the association law between acupuncture points and diseases was deeply analyzed through the construction of knowledge mapping technology, which provides a scientific basis for clinical acupuncture treatment. The study used a systematic search method to retrieve relevant literature from China Knowledge Network, Wanfang Knowledge Service Platform, and Wipu Chinese Science and Technology Journal Database from January 1, 2018 to August 15, 2023, and used data mining techniques such as systematic cluster analysis and association rule mining, combined with CiteSpace software to visualize and analyze the knowledge graph. Finally, 354 literatures were included, involving 462 acupoints and 22 meridians. Frequency statistics showed that the frequency of foot Sanli was used up to 61 times, and the frequency of foot solar bladder meridian was used up to 173 times. The association rule analysis found that the confidence level of the association between Fengchi and Hegu was as high as 91.54%, and the cluster analysis divided the high-frequency acupoints into 5 major groups.CiteSpace analysis constructed a keyword co-occurrence network graph containing 570 nodes, and generated 7 major clusters. The study demonstrated that the knowledge graph technology can effectively reveal the association patterns between acupuncture points and diseases, provide data support for clinical selection of acupuncture points, and promote the modernization and development of acupuncture and evidence-based medical practice.

Index Terms Knowledge graph, acupuncture points, association rules, cluster analysis, data mining, visualization analysis

I. Introduction

Acupuncture is an important part of traditional Chinese medicine therapy, which regulates the operation of qi and blood and promotes the body's self-healing ability by stimulating specific acupoints in the human body [1]. There exists a close relationship between acupuncture points and diseases, and the common acupuncture points and their association with diseases mainly include Quchi point, Hegu point, Shusanli point and Taiyuan point [2], [3].

Quchi point is located on the medial side of the arm, on the transverse elbow stripe, in the depression below Quchi [4]. Acupuncture and moxibustion at the Quchi point can relieve respiratory disorders such as chest pain, cough, asthma, etc. It is effective in unblocking qi and blood and relieving respiratory difficulties [5]. Located at the junction of the thumb and index finger on the back of the hand, Hegu point is one of the most commonly used acupuncture points [6]. Acupuncture Hegu point can relieve headache, insomnia, dry mouth, and other disorders related to the nervous system and digestive system [7]. The Foot Sanli point is located on the lateral side of the lower leg, three inches from the anterior edge of the tibia [8]. Acupuncture and moxibustion of the foot Sanli point can enhance the body's immunity, regulate the function of the digestive system, and improve gastrointestinal problems, such as stomach pain and diarrhea [9]. Taiyuan point is located at the front of the sole of the foot, in the depression between the first and second metatarsal bones [10].

Acupuncture and moxibustion at the Taiyuan point can improve the function of the cardiovascular system and regulate heartbeat and blood pressure [11]. At the same time, Taiyuan acupoints also help to alleviate anxiety, insomnia, depression, and other problems related to mental health [12]. There is a strong link between acupuncture points and diseases. By correctly selecting and stimulating the relevant points, acupuncture can help to regulate the balance of yin and yang in the human body, and promote the circulation of qi and blood, which in turn can play a role in treating and preventing diseases [13]-[15]. Despite the therapeutic efficacy of acupuncture in treatment, it

needs to be combined with specific symptoms and individual characteristics in practical application and operated by a professional TCM acupuncturist to ensure safety and effectiveness [16]-[18].

In this study, we designed a research program on the association between acupuncture points and diseases based on the construction of knowledge graph. First, a standardized literature search strategy was established to systematically collect clinical research literature on acupuncture for diseases in recent years from major Chinese databases. Then, the literature was screened using strict inclusion and exclusion criteria, and the acupuncture prescription information was extracted to construct a structured dataset. Next, systematic cluster analysis was applied to identify the acupuncture point dispensing patterns, the association relationship between acupuncture points was discovered by association rule mining technique, and the knowledge graph was constructed by using CiteSpace software for visualization and analysis. Finally, the results of multiple analyses were synthesized to summarize the association patterns between acupuncture points and diseases, and to provide a scientific basis for clinical acupuncture point selection.

II. Materials and methods

II. A. Data preparation

II. A. 1) Literature sources

Retrieval year: according to the preliminary literature preliminary inspection, due to the large number of literature, the retrieval year is appropriately shortened, and the year is set from January 1, 2018 to August 15, 2023.

Search database: China Knowledge Network, Wanfang Knowledge Service Platform, Wipu Chinese Science and Technology Journal Database.

II. A. 2) Literature search strategy

A combination of computer retrieval and manual inspection is adopted. Computer search: The three major databases were searched, mainly using subject headings, and the search terms were "disease treatment", "disease occurrence", "disease prevention", "lung disease", "cardiovascular disease", "common disease", "infectious disease", "related disease", "immune disease", "cardiovascular and cerebrovascular disease", "infectious disease", "acupuncture", "acupuncture", "needle", "warm acupuncture", "warm needle", "electroacupuncture", etc.

II. A. 3) Literature inclusion criteria

(1) Literature type: clinical research-based literature related to acupuncture treatment of diseases published publicly within the period of 2018.1.1-2023.8.15 with a randomized controlled trial approach.

(2) Interventions: the treatment group used acupuncture alone or acupuncture combined with other therapies, of which acupuncture included acupuncture, warm acupuncture, and electroacupuncture; the control group had no limit to the interventions.

(3) There must be a clear prescription of acupuncture points.

II. A. 4) Literature exclusion criteria

(1) Animal experiments, experience sharing, theoretical discussions, literature reviews, and non-randomized controlled trials.

(2) Treatment group interventions did not include acupuncture (warm or electroacupuncture) treatment.

(3) More than 2 experimental groups.

(4) No explicit acupuncture prescription

(5) Comparative study of the same acupuncture point prescription, different needling methods, and depth of needling.

(6) Repeatedly examined or published literature, taking the latest published one.

II. A. 5) Literature screening methods

(1) Primary screening: export the first retrieved literature to NoteExpress 3.6 software and use its “weight check” function to remove duplicates.

(2) Secondary screening: through the title of the article, abstract, keywords and other contents of the rapid browsing, eliminating the literature that is obviously not in line with the study, such as: non-randomized controlled trials, review literature, animal experiments and so on.

(3) Re-screening: Retrospective searching of the literature after the second screening, downloading the original articles and reading them carefully. Literature that did not meet the inclusion and exclusion criteria established in this study was eliminated strictly.

II. A. 6) Acupoint screening methods

(1) The acupoints selected in the literature belong to the twelve main meridians, the Ren and Du channels and the extra-meridian odd points.

(2) If the literature involves bilateral acupoints, one side of the acupoints will be recorded.

(3) When two or more groups of acupoints in the literature meet the inclusion criteria, a group of acupoints with better therapeutic effect will be selected for prescription statistics.

II. A. 7) Literature search results

According to the specified search strategy, the literature with the years from January 1, 2018 to August 15, 2023 was searched from the three major databases of Zhi.com, Wanfang, and Wipu, of which 705 documents were first checked in Zhi.com, and after three screenings according to the inclusion and exclusion criteria, 267 documents were finally obtained; 679 documents were first checked in Wanfang database, and after three screenings according to the inclusion and exclusion criteria, 75 documents were finally obtained; 812 documents were first checked in Wipu (duplicates in the three major databases were counted among the documents included in Zhi.com). There were 75 documents in Wanfang database; 812 documents in Wipu database were firstly examined, and 12 documents were finally obtained after three times of screening according to the inclusion and exclusion criteria (duplicates in the three databases were counted in the documents included in Zhi.com). In total, 354 documents were included.

II. B.Research methodology

II. B. 1) Systematic cluster analysis

Systematic clustering method [19] in addition to defining the affinity index between things, but also to define the affinity index between classes and classes and to derive the recursive formula for the value of the affinity index between classes. Classes and classes of different affinity indicators, the recursive formula for the value of the affinity indicators between classes is also different.

Various algorithms are described below:

If n_r is the number of objects in class r, n_s is the number of objects in class s, and x_{ri} is the i th object in class r, then different clustering algorithms are defined as follows:

Algorithm 1: Shortest distance method

The shortest distance method is also known as single connection or nearest neighbor connection. The distance between two classes if defined as the smallest distance between the elements in the two classes, and accordingly choose the closest class aggregation method one by one is called the shortest distance method.

The formula indicates:

$$d(r,s) = \min\left(dist\left(x_{ri}, x_{sj}\right)\right), i \in (1,\cdots,n_r), j \in (1,\cdots,n_s) \tag{1}$$

Algorithm 2: Longest Distance Method

The longest distance method is also known as complete linkage or farthest immediate neighbor linkage. The longest distance method clusters in the same way as the shortest distance method, with the difference that the distance between classes is defined as being the one with the greatest distance between the elements of the two classes.

The formula indicates:

$$d(r,s) = \max\left(dist\left(x_{ri}, x_{sj}\right)\right), i \in (1,\cdots,n_r), j \in (1,\cdots,n_s) \tag{2}$$

Algorithm 3: Intermediate Distance Method

The class-to-class distance is called the intermediate distance method if instead of taking the shortest distance between the elements of the two classes or the longest distance, some intermediate distance is taken. For example, suppose that in the process of clustering two classes G_1 and G_2 are merged into a new class $G_N = (G_1, G_2)$, then the distance between G_N and any other class G_3 is defined as the square of the median of the triangle in the figure below.

Formula representation:

$$d_{G_3G_N}^2 = d^2 = \frac{1}{2}\left(d_{G_3G_1}^2 + d_{G_3G_2}^2 - \frac{1}{2}d_{G_1G_2}^2\right) \tag{3}$$

Algorithm 4: Average Distance Method

The distance between two classes is defined as being the average distance between two and two of the elements in those two classes.

The formula indicates:

$$d(r,s)=\frac{1}{n_r n_s}\sum_{i=1}^{n_r}\sum_{j=1}^{n_s} dist\left(x_{ri},x_{sj}\right) \tag{4}$$

Algorithm 5: Center of Gravity Method

The distance between two classes is defined as the distance between the centers of gravity of these two classes.

The formula indicates:

$$d(r,s)=d\left(\overline{x}_r,\overline{x}_s\right) \tag{5}$$

where $\overline{x}_r=\frac{1}{n_r}\sum_{i=1}^{n_r}x_{ri}$, and $\overline{x}_s=\frac{1}{n_s}\sum_{i=1}^{n_s}x_{si}$

Algorithm 6: Sum of Squares of Departures

The idea of its classification is similar to that of ANOVA. In the classification process, the sum of squares of variances between elements within a class is made as small as possible, while the sum of squares of variances between classes is made as large as possible.

The formula indicates:

$$d(r,s)=n_r n_s d_{rs}^2/(n_r+n_s) \tag{6}$$

where d_{rs}^2 is the center of gravity distance between classes r and s.

By introducing the above six methods, we have understood the idea of cluster analysis. Among these six methods, the clustering process of both the shortest distance method and the longest distance method are monotonic, i.e., the distance at each step of clustering is greater than that of the previous step, and thus the clustering tree will be readily apparent if it is illustrated graphically. The intermediate distance method is non-monotonic, i.e., sometimes the distance of clustering may instead be smaller than the distance at the previous step of clustering, so the tree diagram is sometimes not easy for people to understand. But the intermediate distance method is spatially conserved, i.e. the distance between two classes is basically taken to be in the middle, not the shortest or the longest, which is an advantage. The center of gravity method is spatially conserved, but it is also non-monotonic.

Comparing these six methods, it is found that although the hierarchical clustering method is more objective, the choice of distances as well as the choice of various aggregation methods is still somewhat subjective. Therefore, in carrying out the clustering analysis, this paper tries a variety of distances, a variety of aggregation methods, and finally determines a more appropriate clustering result according to the nature of the actual problem. In view of the actual mandatory selection, this paper uses the average distance method.

Since this paper uses the clustering statistics tool in the general-purpose programming software MATLAB to categorize the flood samples, the following is combined with the relevant content in MATLAB. Using d_{ij} to denote the distance between object i and object j, $C_1,C_2,\cdots C_n$ to denote the classes, and D_{pq} to denote the distance between class C_p and class C_q, the steps of the systematic clustering method are as follows:

(1) Classification pre-processing - data standardization

When the sample data between the order of magnitude or magnitude difference is very large, will interfere with the system clustering, affecting the clustering results. Therefore, the usual way to deal with such problems in scientific research is to standardize all the data, standardization formula: $Z=(V-\text{mean}(V))./\text{std}(V)$, where Z - standardized data; V - column vector; $mean(V)$ - mean of column vector; $std(V)$ - standard deviation of column vector. standard deviation of the column vector.

(2) Specify the distance between objects and calculate the sample distance

A variety of distance calculation formulas are provided in the systematic clustering, and the most commonly used Euclidean distance is used in this paper. Start each sample into a class of its own, this time $D_{pq}=d_{pq}$, all the distance between the samples can be listed as a pair of angular arrays, the two two distances form a symmetric array, notated as $D_{(0)}$.

(3) Specify the distance between classes and calculate the distance between classes

In this paper, the average distance method is used. Select the smallest element in $D_{(0)}$, set as D_{pq}, and merge C_p and C_q into a new class, denoted as $C_r=\{C_p,C_q\}$.

(4) Calculate the distance between the new class and the other classes

$D_{rk} = \min_{\substack{i\in G_r \\ j\in G_k}} d_{ij} = \min\left\{\min_{\substack{i\in G_p \\ j\in G_k}} d_{ij}, \min_{\substack{i\in G_q \\ j\in G_k}} d_{ij}\right\} = \min\{D_{pk}, D_{qk}\}$, p, q rows, p, q columns in $D_{(0)}$ are merged into a new row and column, the new row and column correspond to C_r, and the resulting new matrix is denoted as $D_{(1)}$.

(5) Repeat step (4) for $D_{(1)}$, and then obtain $D_{(2)}$, and so on until all elements become one class. If there is more than one minimal element in $D_{(k)}$ in a step, the classes corresponding to these minimal elements can be merged simultaneously.

(6) In MATLAB, up to this point, a clustering tree can be formed and a clustering tree diagram is drawn.

(7) Determine the location of the classification on the clustering tree to form the clustering results.MATLAB statistical tools can be detected through the cluster() function on the clustering tree of the natural classification of the situation and take the threshold or by truncating the clustering tree of any node for classification.

II. B. 2) Association Rule Mining

Association rules [20], as one of the commonly used algorithms in data mining, are used to mine valuable, hidden, and meaningful information from massive data, and are an important technique to reflect the correlation between things. Categorized according to the processed values, they can be divided into quantitative and Boolean type association rules. Boolean association rules study the occurrence or non-occurrence of data items in a transaction. While quantitative association rules are mainly for discrete data, in which the data items are quantitative, in the association rule mining of quantitative data, it is often necessary to discretize the quantitative data with the help of methods such as statistics and discretization, into several discrete intervals, and then transformed into Boolean association rule mining. The most common association rule algorithms include Apriori algorithm, FP-growth algorithm and Eclat algorithm.

The basic concepts in association rules are:

(1) Item set

An itemset is a collection of elements in a dataset. An itemset can contain one or more items. In association rule mining, itemsets are often used to describe combinations of items that frequently occur together in data.

(2) Transaction

A transaction represents a record or an event in a data set. Each transaction contains information about the itemset, which describes the simultaneous occurrence of certain items.

(3) Support

Support is a metric used to measure how often an item set occurs in all data. For the itemset X, support is defined as the ratio of the number of transactions containing X to the total number of transactions.

$$Support(X) = \frac{Transaction\ containing\ X}{Total\ transactions} \tag{7}$$

(4) Confidence level

Confidence level is a metric used to measure the reliability of a rule $A \Rightarrow B$. For the association rule $A \Rightarrow B$, the confidence level is the ratio of the number of transactions in which both A and B occur to the number of transactions in which only A occurs.

$$Confidence(A \Rightarrow B) = \frac{Support(A \cup B)}{Support(B)} \tag{8}$$

(5) Lift

The degree of lift is a metric used to measure whether rules $A \Rightarrow B$ are more likely to occur simultaneously compared to their support. The lifting degree is calculated as:

$$Lift(A \Rightarrow B) = \frac{Confidence(A \Rightarrow B)}{Support(B)} \tag{9}$$

(6) Frequent itemset

Frequent itemset refers to the set of items in the data set whose frequency of occurrence reaches a predetermined threshold. Frequent itemsets are the basis of association rule mining, and association rules can be further generated by discovering frequent itemsets.

(7) Association rule

Assuming that $I=\{I_1,I_2,I_3,\ldots,I_k\}$ is a set consisting of all items, and the set I is a $K-$ itemset, the association rule is an expression in the form of $A \Rightarrow B$, in which A is the premise of the rule, and B is the result, and A, B have to satisfy both $\{A,B \mid A \subset I, B \subset I, A \cap B=\varnothing\}$.

(8) Strong association rules

A rule is a strong association rule if it satisfies both the minimum support threshold and the minimum confidence threshold.

Association rules are mainly mined by the support and confidence of the item set, and their efficiency mainly lies in whether they can effectively mine frequent item sets, and the mining steps are as follows:

(1) Generate candidate itemsets: use Apriori algorithm or other frequent itemset mining algorithms to generate all possible itemsets. These itemsets are usually candidate itemsets.

(2) Calculate support: Count the frequency of occurrence of each candidate itemset in the data set to get the support.

(3) Filtering frequent itemsets: By setting the minimum support threshold, the frequent itemsets are obtained by keeping the itemsets with support greater than the threshold.

(4) Generate association rules: For each frequent itemset, generate all its possible association rules. Consider the confidence level of the rules and keep the rules with confidence level greater than the set threshold.

(5) Evaluate the quality of rules: Evaluate the rules obtained from mining by other metrics, such as lift and full confidence.

(6) Visualization and interpretation: association rules can be interpreted and understood by visualization tools or analysis tools to discover hidden relationships in the data.

II. C.Statistical tools

(1) The collected 354 acupuncture prescriptions in the literature were organized into an Excel (Microsoft Excel 2019) sheet, and the Excel sheet was uploaded into the Traditional Chinese Medicine Inheritance Computing Platform (V3.0) software for statistical analysis, association rule analysis, and cluster analysis.

(2) Literature with years from January 1, 2018 to August 15, 2023 that met the inclusion criteria was searched in the three major databases, entered into Citespace 6.3.R1 software, and keyword co-occurrence, keyword clustering, keyword emergence, and keyword timeline mapping were analyzed.

Table 1: The acupuncture treatment of the mild acupoint frequency of the first 20 cases

Code	Acupuncture points	Frequency	Frequency(%)
S1	Tsusanli	61	13.78
S2	Yang bai	44	4.87
S3	Picking bamboo	40	4.84
S4	San yin Jiao	35	4.33
S5	Ho ku	33	4.25
S6	Si zhu kong	26	3.83
S7	Bai hui	24	3.71
S8	Sun	22	3.54
S9	Yu yao	21	3.47
S10	The kidney heals	19	3.42
S11	Quchi	18	3.37
S12	Fengchi	18	3.24
S13	Bl 20	16	2.99
S14	Three in hand	15	2.76
S15	Yang ling quan	14	2.62
S16	Internal closure	13	2.34
S17	White	12	2.31
S18	Chung	11	2.27
S19	Sosky	10	2.36
S20	Closing yuan	10	2.41

III. Results and analysis

III. A. Frequency count results

III. A. 1) Statistical Analysis of Frequency of Acupoints

The top 20 frequency of acupuncture points for myasthenia gravis are shown in Table 1. The frequency of acupuncture points selected was ranked from high to low as 462 acupoints, including Foot Sanli (61), Yangbai (44), Zanzhu (40), Sanyinjiao (35), Hegu (33), Sizhukong (26), Baihui (24), Taiyang (22), Yubai (21), Renyu (19), and Quchi (18).

III. A. 2) Frequency analysis of acupoint attribution to meridians

The top 20 frequencies of acupuncture for myasthenia gravis attributed to the meridians are shown in Table 2. The frequency of using the selected acupoints totaled 828 times, involving 22 meridians, of which the highest frequency was the foot solar bladder meridian (173), followed in order by the foot yangming stomach meridian (113), the foot Shaoyang gallbladder meridian (101), the hand yangming large intestinal meridian (77), and the directing meridian (74).

Table 2: The acupuncture treatment of the syndrome was unable to return to the top 20 times

DAIMA	Meridian	Ffrequency	Frequency	Cavities
1	Urocystis	173	21.47	31
2	Pingming gastric meridian	113	14.23	22
3	The bladder is less than the bile	101	13.41	19
4	The hand of the hand Yang Ming large intestine	77	9.87	6
5	Governor	74	9.43	19
6	The meridian of the foot	66	8.37	9
7	External cavity	62	8.15	15
8	Vein	56	6.67	5
9	Oligo trifocal	41	6.94	4
10	The heart of the hand	19	5.38	4
11	Hepatic meridian	13	2.13	3
12	Renal renal meridian	10	1.81	3
13	The heart of the hand	9	1.53	2
14	The small intestine of the hand	8	1.21	2
15	The heart of the hand is little	6	0.97	2

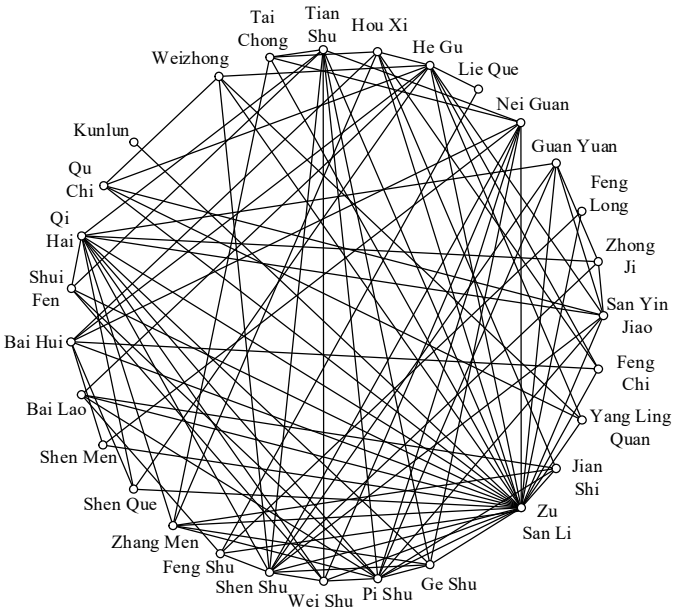

Figure 1: The rules of the association of acupuncture and acupuncture in the treatment of diseases

III. B. Analysis of acupuncture point and disease association rules

Using Apriori algorithm in Clementine 12.0, 31 high-frequency acupoints with a frequency of use ⩾5 were analyzed for association rules. The mesh diagram of the association rules is shown in Figure 1, and the thicker the line between them indicates the higher their correlation, in which the results of the two acupoints association are shown in Table 3.

Table 3: The analysis of the correlation between acupuncture and acupuncture treatment

Sort	Item set	Support(%)	Confidence(%)
1	Wind pool→Closed valley	22	91.54
2	Meander→Closed valley	18.21	100
3	Meander→Triathlon	18.21	100
4	Interlock→Triathlon	18.21	100
5	Armature→Triathlon	18.21	100
6	Middle pole→Air sea	15.24	100
7	Middle pole→Closing yuan	15.24	100
8	Triyin→Closing yuan	15.24	100
9	Yang ling quan→Weizhong	15.24	100
10	Weishu→Armature	15.24	100
11	Weishu→Pishu	15.24	100
12	Weishu→Triathlon	15.24	100
13	God door→Internal closure	12.57	100

III. C. Cluster analysis of acupoints

Combined with statistics, the top 20 acupoints in terms of frequency were clustered and analyzed, as shown in Figure 2. As can be seen from Figure 2, it can be divided into 5 major groups, with Group 1 being Euming/ Guanyuan/ Guanzhong/ Sibai/ Zhongguang/ Neiguan/ Sanli/ Yanglingquan; Group 2 being Quchi/ Fengchi/ Spleen Yu/ Silk Bamboo Hollow/ Baihui/ Fish Waist/ Kidney Yu/ Sun; Group 3 being Yangbai/ Zanzhu; Group 4 being Sanyinjiao/ Hegu; and Group 5 being Shusanli.

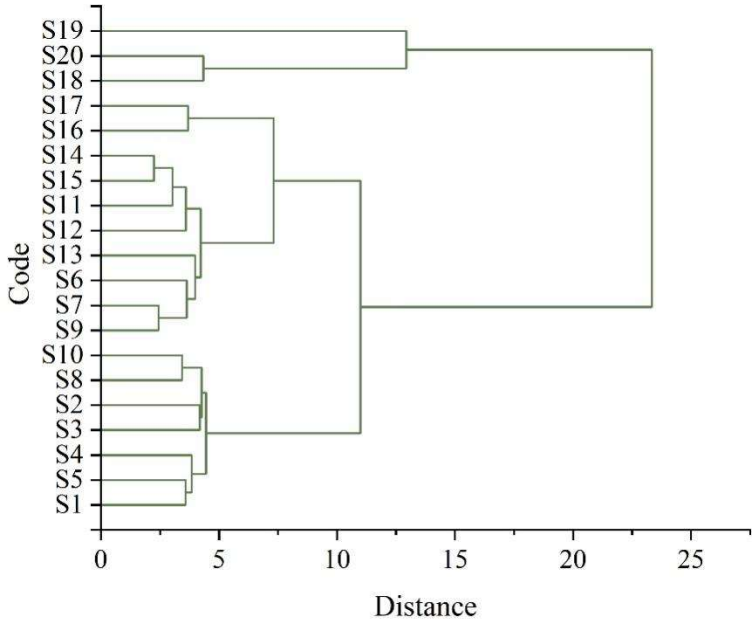

Figure 2: The results of the analysis of the treatment of myasthenia

III. D. CiteSpace Knowledge Graph Visualization and Analysis

III. D. 1) Keyword co-occurrence analysis

CiteSpace 6.1.R6 software was used to construct a network graph of literature keyword co-occurrence in the field of acupuncture data mining. The time range was set to 2008-05-2022-05, and the data of each year were statistically

analyzed by taking one year as the division unit, and the node type was selected as "Keyword", and a network graph containing 570 nodes, 11741 edges, and a density of 0.0111 was constructed.

The keyword co-occurrence network mapping was analyzed as shown in Table 4. 14 keywords appeared 18 times or more, and the keywords with higher frequency were “data mining” 753 times and “acupuncture” 693 times, “acupuncture point selection law” appeared 382 times. The high-frequency keywords mainly include data mining algorithms, such as association rules, cluster analysis, etc.; syndromes research, such as depression, stroke, etc.; data sources, such as literature research, ancient books, acupuncture prescriptions, etc.; and data mining objectives, such as the law of selecting acupoints, acupoints matching, etc. The above mentioned high-frequency keywords indicate that the data mining of acupuncture and moxibustion is a very important part of the research. The above high-frequency keywords indicate that the research hotspot in the field of acupuncture and moxibustion data mining is the mining of point selection laws and acupoint prescription laws from a large number of acupuncture and moxibustion prescriptions recorded in the modern literature and ancient books by using data mining algorithms such as clustering analysis and association rule analysis, etc. for the conditions such as depression, stroke and so on.

Table 4: The number of keywords in the mining area of acupuncture data is 20 and above

Serial number	Key words	Frequency	Serial number	Key words	Frequency
1	Data mining	753	8	Cluster analysis	33
2	Acupuncture	693	9	Regularities	25
3	Cavitation pattern	382	10	Electric needle	22
4	Acupuncture points	71	11	Correlation analysis	21
5	Literature study	65	12	moxibustion	19
6	Association rule	45	13	Ancient book	19
7	The acupuncture points match	37	14	Migraine	18

III. D. 2) Keyword clustering analysis

The keywords were clustered and analyzed using CiteSpace 6.1.R6 software with the aim of understanding the historical span and research frontiers in the field of acupuncture data mining. The keywords were clustered and seven major clusters were generated with a module value (Q-value) of 0.4323, which is greater than 0.3, indicating a significant association structure in the graph, and an average silhouette value (S-value) of 0.8559, which is greater than 0.7, indicating that the keyword clustering results have a high degree of confidence.

The cluster labels, number of nodes and contour values of the seven main clusters are shown in Table 5; the cluster labels show the research hotspots and research trends in the field; the number of nodes indicates the research heat, and the larger its value means the higher the heat, and the contour value indicates the concentration of the research topic, and the larger its value means the higher the concentration of the topic. #0, #1, #2 clustering node number is higher, and the research theme is to excavate the law of selecting acupuncture points in the prescription of acupuncture treatment for various diseases. The research theme of #3 clustering is to mine the core acupoints in acupuncture prescriptions using association rules. The #4 clustering research theme is literature research, including ancient books, famous doctors' cases and modern literature. #5 The clustering research theme is the rules of meridian identification and taking acupoints with symptoms. #6 Clustering research theme for data mining of electroacupuncture therapy, moxibustion therapy, pestle and mortar therapy, etc.

Table 5: The result of keywords in acupuncture data mining

Cluster number	Cluster label	Node number	Contour number
0	Acupuncture	113	0.686
1	Data mining	87	0.923
2	Cavitation pattern	71	0.719
3	Association rule	54	0.821
4	Literature study	44	0.803
5	regularities	35	0.808
6	Electric needle	31	0.859

III. D. 3) Keyword emergence analysis

Keyword emergence refers to the rapid increase in the frequency of a keyword in a certain time period, also known as emergent words, which to some extent reflects the development path and trend of the research field. Keyword emergence was analyzed, and since 2008, migraine has been the main research symptom in the field of

acupuncture data mining, which consistently lasted until 2012. Database technology, as the main technical means of data management, received extensive attention during 2009-2015. As the intersection of Chinese medicine and acupuncture and information science, moxibustion, tongkong acupoints, and acupuncture have gradually received more and more attention, and have successively become research hotspots since 2013. The keyword emergence mapping in the field of acupuncture and moxibustion data mining is shown in Table 6. From the table, it can be seen that in recent years, the law, acupuncture point preference, acupuncture point matching, and prescription law have become the research hotspots. It can be seen that the research hotspots in the field of acupuncture and moxibustion data mining are numerous, in which the research directions of acu-point dispensing and prescription law will still be used as cutting-edge hotspots to promote the development and innovation of the field.

Table 6: Key words in acupuncture data mining

Key words	Sudden beginning	The end of the sudden	Sudden strength
migraine	2008	2012	5.57
Acupuncture prescription	2009	2016	5.53
database	2009	2015	3.61
moxibustion	2013	2015	3.95
Homological cavity	2015	2017	3.91
regularities	2016	2020	8.3
Acupoint preference	2016	2021	6.74
Acupuncture	2017	2020	2.87
The acupuncture points match	2017	2022	5.23
Prescription law	2017	2022	2.91

IV. Conclusion

In this study, the association between acupuncture points and diseases was deeply analyzed by knowledge mapping technology, and important regular features were found. Frequency statistics showed that Foot Sanli ranked first with 61 times of use, and Yangbai ranked second with 44 times, indicating the core position of these points in clinical treatment. Meridian distribution analysis showed that the foot-sun bladder meridian was used 173 times, accounting for 21.47% of the total frequency, reflecting the importance of the bladder meridian in acupuncture treatment. Association rule mining revealed the deep connection between acupoints, in which the association confidence level of Fengchi and Hegu was as high as 91.54%, indicating that these two acupoints have a strong tendency to match in clinical practice. Cluster analysis scientifically classified the high-frequency acupoints into five major groups, which provided a systematic reference framework for the clinical selection of acupoints.

The knowledge graph visualization analysis constructed a keyword co-occurrence network containing 570 nodes, and the seven major clusters generated reflected the research hotspots and development trends in the field of acupuncture data mining. This big data-based analysis method breaks through the limitations of traditional experience summarization and provides scientific support for the modernization and development of the acupuncture discipline. The research results not only help to improve the standardization of clinical acupuncture treatment, but also lay a theoretical foundation for building an intelligent acupuncture diagnosis and treatment system. In the future, the data sample size can be further expanded, more dimensional clinical information can be integrated, and the knowledge map of the association between acupuncture points and diseases can be continuously improved, so as to promote the development of acupuncture medicine in the direction of precision and individualization.

Funding

Project Source (Name of Fund Granting Unit): Hainan Vocational University of Science and Technology School level Industry Education Integration Characteristic Professional Group Fund Funding Project Name: Health Management and Rehabilitation Services Professional Group (NO: HKZYQ2024-05).

References

[1] Chon, T. Y., & Lee, M. C. (2013, October). Acupuncture. In Mayo Clinic Proceedings (Vol. 88, No. 10, pp. 1141-1146). Elsevier.
[2] Smith, C. A., Armour, M., Zhu, X., Li, X., Lu, Z. Y., & Song, J. (2016). Acupuncture for dysmenorrhoea. Cochrane Database of Systematic Reviews, (4).
[3] Kelly, R. B., & Willis, J. (2019). Acupuncture for pain. American family physician, 100(2), 89-96.

[4] Li, C. R., Lin, Y., Guan, H. Y., Liang, Z. R., Zhang, Z. X., Kim, A., ... & Chen, Y. G. (2013). Different surface electromyography of propagated sensation along meridians produced by acupuncturing Quchi acupoint (LI11) or control points. Evidence‐Based Complementary and Alternative Medicine, 2013(1), 198451.
[5] Li, X., Chen, F., Wang, W., Liu, Y., Han, J. Q., Ke, Z., & Zhu, H. H. (2024). Visual analysis of acupuncture point selection patterns and related mechanisms in acupuncture for hypertension. Technology and Health Care, 32(1), 397-410.
[6] Guangjun, W., Yuying, T., Shuyong, J., Tao, H., & Weibo, Z. (2012). Change of blood perfusion in Hegu acupoint after contralateral Hegu acupoint was stimulated. The Journal of Alternative and Complementary Medicine, 18(8), 784-788.
[7] Hsiu, H., Hsu, W. C., Hsu, C. L., & Huang, S. M. (2011). Assessing the effects of acupuncture by comparing needling the hegu acupoint and needling nearby nonacupoints by spectral analysis of microcirculatory laser Doppler signals. Evidence‐Based Complementary and Alternative Medicine, 2011(1), 435928.
[8] Wang, L. Q., Chen, Z., Zhang, K., Liang, N., Yang, G. Y., Lai, L., & Liu, J. P. (2018). Zusanli (ST36) acupoint injection for diabetic peripheral neuropathy: a systematic review of randomized controlled trials. The Journal of Alternative and Complementary Medicine, 24(12), 1138-1149.
[9] Chen, Z., Wang, X., Du, S., Yao, K., Guo, Y., & Lin, X. (2024). Acupuncture at the Zusanli acupoint can reduce the inflammatory response in AIA mice by regulating the arachidonic acid and pentose phosphate pathways. Journal of Chromatography B, 1247, 124307.
[10] Wei, H. N., Jiang, L. P., Xiong, B., Zhou, S., Yu, L., Huang, Y. M., ... & Zhao, H. L. (2017). Characteristic patterns of normal meridian acupoint temperature. Journal of the Chinese Medical Association, 80(7), 419-426.
[11] Wang, M. (2019). Acupuncture styles in current practice. Translational acupuncture research, 75-163.
[12] Godson, D. R., & Wardle, J. L. (2019). Accuracy and precision in acupuncture point location: a critical systematic review. Journal of acupuncture and meridian studies, 12(2), 52-66.
[13] Zhang, Y. Q., Wen, J., Li, Y. N., Liu, R. R., Zhang, X. P., Zhang, X., & Jia, C. S. (2022). Visual analysis of the acupoint prescription characteristics of acupuncture and moxibustion in treatment of rheumatoid arthritis. Zhongguo Zhen jiu= Chinese Acupuncture & Moxibustion, 42(11), 1311-1316.
[14] Kwon, S., Seo, B. K., & Kim, S. (2016). Acupuncture points for treating Parkinson's disease based on animal studies. Chinese journal of integrative medicine, 22(10), 723-727.
[15] de Lima Pimentel, R., Duque, A. P., Moreira, B. R., & Junior, L. F. R. (2019). Acupuncture for the treatment of cardiovascular diseases: a systematic review. Journal of Acupuncture and Meridian Studies, 12(2), 43-51.
[16] Hwang, Y. C., Lee, I. S., Ryu, Y., Lee, M. S., & Chae, Y. (2021). Exploring traditional acupuncture point selection patterns for pain control: data mining of randomised controlled clinical trials. Acupuncture in Medicine, 39(3), 184-191.
[17] Li, X., Liu, S., Liu, H., & Zhu, J. J. (2023). Acupuncture for gastrointestinal diseases. The Anatomical Record, 306(12), 2997-3005.
[18] Lee, Y. S., Ryu, Y., Yoon, D. E., Kim, C. H., Hong, G., Hwang, Y. C., & Chae, Y. (2020). Commonality and specificity of acupuncture point selections. Evidence‐Based Complementary and Alternative Medicine, 2020(1), 2948292.
[19] Yu Bin,Zheng Zijian & Dai Jianhua. (2023). K-DGHC: A hierarchical clustering method based on K-dominance granularity. Information Sciences,632,232-251.
[20] Kamyar Sabri Laghaie,Farid Momayezi,Negar Ghaleshakhani & Leila Maroufi. (2025). Association rule mining based approach to consider users' preferences in the energy management of smart homes. Journal of Building Engineering,104,112361-112361.