Received July 28, 2020, accepted August 13, 2020, date of publication August 21, 2020, date of current version September 1, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.3018504*

# Selection of Acupoints Combination Based on Pseudo Mutual Information Maximization

**JINGJING XU[1], YIJIA LIU[1], XIANGNAN XU[2], KIAN-KAI CHENG[3], ZONGBAO YANG[4], AND JIYANG DONG[1]**

[1]Fujian Provincial Key Laboratory for Plasma and Magnetic Resonance, Department of Electronic Science, Xiamen University, Xiamen 361005, China
[2]School of Mathematics and Statistics, The University of Sydney, Sydney, NSW 2006, Australia
[3]Innovation Center in Agritechnology, Universiti Teknologi Malaysia, Johor Bahru 84600, Malaysia
[4]Department of Chinese Medicine, Xiamen University, Xiamen 361005, China

Corresponding author: Jiyang Dong (jydong@xmu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 81871445 and Grant 81801788.

**ABSTRACT** Acupuncture has been used as an alternative and complementary therapy, however the selection of acupoints combination for a specific disease (or symptoms) highly depends on the experience of acupuncturists. The current study aimed to develop a reliable and quantifiable method for selection of acupoints combinations, which may improve treatment efficiency and reproducibility among acupuncturists. In the current study, a number of acupuncture prescriptions for treatment of acute or chronic gastritis had been collected and organized from public databases and literatures. Three types of associations, acupoints-symptoms associations (ASA), acupoints-acupoints associations (AAA), and symptoms-symptoms associations (SSA) were defined to delineate their connections. The network constructed by AAA was used to study the synergistic effects among acupoints combination whereas SSA network was used to examine coexistence of symptoms. On the other hand, ASA characterized the empirical association between acupoints and symptoms in the prescription database which may be considered as the mathematical representation of acupoints selection for a particular prescription. Then a novel method namely mutual information screening (MIS) was proposed for screening acupoints combinations based on pseudo mutual information maximization. The selected acupoints combinations of MIS method were verified to be in accordance with the rules of acupoints compatibility based on the theory of acupuncture. Validation of MIS through conditional entropy and significance test of the clinical records suggested the method to be reliable and robust. Furthermore, a ten-fold cross validation were carried out to evaluate the performance of MIS compared with artificial neural network (ANN)-based method. The F1-score of MIS was found to be 0.67, higher than that of the ANN-based method (0.58). The current results provided a reliable and practical tool for symptom-specific acupoints selection, which may help to develop a scientific basis for clinical acupuncture therapy.

**INDEX TERMS** Acupoints combination, mutual information screening (MIS), gastritis, associations network.

## I. INTRODUCTION

Acupuncture, an important constituent in traditional Chinese medicine (TCM), has been widely used as a complementary and alternative therapy [1], [2] for diseases such as migraines [3], hyperlipemia [4], arthritis [5], and gastritis [6]. According to acupuncture theory, there are more than 360 acupoints associated with 12 principal meridians. Therapeutic effect of clinical acupuncture is rarely achieved through treatment on a specific acupoint. Instead, it is normally obtained through acupuncture on a combination of acupoints.

It was found that combination of different acupoints may trigger either synergistic therapeutic or antagonistic effects that inhibit the effectiveness of acupuncture treatment [7]. The selection of acupoints combinations is carried out based on experience of acupuncturists and classical ancient TCM references [8]. Therefore in practice, there may be large variation in the selection of acupoints combination for a particular disease, which may probably lead to inconsistency in therapeutic effects and excessive acupoints selection. Thus, a reliable and quantifiable evidence-based method for selection

The associate editor coordinating the review of this manuscript and approving it for publication was Jenny Mahoney.

of acupoints combination is essential to achieve a better efficacy and reproducibility of acupuncture treatment among different acupuncturists.

To the best of our knowledge, most of studies on acupoints combinations to focus on assessing the different therapeutic effects of various acupoints combinations for diseases [9], exploring the synergistic mechanism on rules of acupoints compatibility such as ''He-Mu'' combination [10], or demonstrating the occasional antagonistic effects happened in animal experiments [11]. However, little attention has been paid to the selection method of acupoints combinations for confirmed symptoms.

Previously, scientometric analysis including frequency analysis, association rules mining [12] and network analysis [13] have been carried out to reveal the regularity of acupoints combinations in literatures. For example, frequency analysis, a common topic in this field, adopts the frequency of usage in practice to evaluate the roles of acupoints on relieving symptoms of a particular disease. Association rules mining (ARM) [14], another common topic, identifies the significant patterns of acupoints combinations and extract the rules underlying those patterns. In addition, ARM was often used to draw the association rule from acupoints combinations [15]. Network analysis, originated from social sciences to mine the essence of social phenomena [16], has also been introduced to explore the relationship between acupoints and symptoms in TCM. In network analysis, relationship of acupoints are modelled by a graph/network, where one node represents an acupoint and one edge represents the association between the corresponding two nodes. Topological properties of the acupoints network were analyzed and visualized the regularities and mechanisms of a reasonable acupoints combination [17]. All in all, the application of these analytical strategies in acupuncture helps to reveal combinational patterns of acupoints and shed light on whether the application principles are in accordance with the traditional empirical theory.

The present study aimed to develop a reliable and quantifiable method for selection of acupoints combinations for a given symptoms through learning of the compatibility rules behind clinical acupoints prescription. Similar to TCM herbal formula [18], acupoints prescription can be considered as an empirical system of multi-component therapeutics that treats the complex disease in an integrated manner. Recently, an artificial neural network (ANN) model was designed to predict the acupoints combination for a given symptoms set as an input. A total of 232 clinical records were collected to train the ANN model [19]. Since all clinical records were collected from a traditional Korean medical clinic, the ANN model achieved a very high accuracy of 0.865 (precision 0.911 and recall 0.811). For more general cases, however, the clinical records were often not collected from a single clinic, but from public literature databases. This means a more complex relationship between symptoms and acupoints present in the records. In such cases, a larger number of acupoints records are needed to train the model to achieve a robust result, while insufficient training set may lead to model underfitting and unreliable predictions.

Here, we proposed a mutual information-based screening (MIS) method to select an acupoints combination for a specific symptom subset. A total of 220 clinical records (prescriptions) for gastritis treatment were collected. The co-occurrence of symptoms and acupoints in these prescriptions were represented by a joint probability density matrix. Pseudo mutual information was used to evaluate the association between acupoints and symptoms. For a given symptoms subset, the optimal acupoints combination can be screened in accordance with pseudo mutual information maximization. The performance of MIS was compared to ANN through a ten-fold cross validation. In addition, the reliability and robustness of this MI-based acupoints selection were assessed by conditional entropy and significance tests.

## II. DATA COLLECTION

Acupuncture has been widely used as an alternative therapy to gastritis including acute gastritis, chronic superficial gastritis (CSG) and chronic atrophic gastritis (CAG). Databases including the Web of Science, ScienceDirect, EMBASE, Wan Fang database, China National Knowledge Infrastructure (CNKI) and Chinese Biomedical Literature Database (CBM) were searched for published literatures on acupuncture therapy for gastritis from May 2009 until May 2019. The search strings were as follows: (''acupuncture'' or ''electroacupuncture'' or ''moxibustion'' or ''acupoint'') AND (''gastritis'' or ''CAG'' or ''CSG'').

From the search results, we excluded animal studies and systematic reviews from further study. In addition, we have also excluded studies with involved single acupoint prescription, as well as studies with similar acupoints combination prescriptions. Trials that proved the therapeutic efficacy of acupuncture and/or moxibustion for gastritis were used for further analysis.

Based on the criteria, we had selected a total of 220 prescriptions. The IDs for acupoints involved in these 220 prescriptions were standardized according to the ''Standard Acupuncture Nomenclature'' published by the World Health Organization in 1991 [20]. Description about 13 gastritis-related symptoms associated with the 220 prescriptions were standardized as follows: abdominal discomfort, epigastric bloating, regurgitation, appetite loss, nausea, weight loss, bitter mouth, light red tongues and thin white coating, weak pulse, constipation, diarrhea, gastrointestinal bleeding, anemia. Then, the 220 acupuncture prescriptions involving 30 acupoints for treatment of the 13 gastritis-related symptoms were used in the subsequence analysis.

## III. METHODS

### A. ASSOCIATION DEFINITIONS

Let $K$ be the prescriptions number in a prescriptions dataset, and a prescription $k$ with $I_k$ acupoints and $J_k$ symptoms can be denoted by $P^k(A^k, S^k)$, where $A^k = \{a_i^k, i = 1, 2, \ldots, I_k\}$ and $S^k = \{s_j^k, j = 1, 2, \ldots, J_k\}$ be the acupoints set and

symptoms set of the prescription, respectively. Then, the prescription dataset $P = \{P^k | k = 1, 2, \ldots, K\}$ can be denoted by $P(A, S)$, where $A = \bigcup_k A_k$ and $S = \bigcup_k S_k$ be the acupoints set and the symptoms set of the dataset, respectively. In addition, we can denote $A = \{a_i | i = 1, 2, \ldots, I\}$ and $S = \{s_j | j = 1, 2, \ldots, J\}$. In the current study, the total number of acupoints, $I = 30$; the total number of symptoms, $J = 13$; and the total number of prescriptions, $K = 220$.

#### 1) ACUPOINT-SYMPTOM ASSOCIATION

For a given prescription $P^k$, the acupoint-symptom cooccurrence matrix is defined as,

$$w_{ij}^k = \begin{cases} \dfrac{1}{I_k \cdot J_k} & \text{if } i \in A^k \text{ and } j \in S^k \\ 0 & otherwise \end{cases}$$

where $\mathrm{W}^k = (w_{ij}^k)_{I \times J}$ is a matrix of $I$ rows and $J$ columns. Then the empirical cumulative acupoint-symptom cooccurrence matrix of the dataset is defined as,

$$\mathbf{W} = \frac{1}{K} \sum_{k=1}^{K} \mathrm{W}^k$$

where $\frac{1}{K}$ is a normalization coefficient so that $\sum \mathbf{W} = 1$. Note that $\sum_{ij} w_{ij}^k = 1$, which means that all prescriptions equally contributed to the cooccurrence matrix. This empirical cumulative acupoint-symptom cooccurrence matrix reflects the contribution of an acupoint working for a specified symptom.

#### 2) ACUPOINT-ACUPOINT ASSOCIATION

The acupoint-acupoint association in a given prescription $P^k$ can be defined as follows,

$$w_{ij}^{A,k} = \begin{cases} \dfrac{1}{I_k \cdot (I_k - 1)} & for\ \forall i \in A^k, i \neq j \\ 0 & otherwise \end{cases}$$

where $\mathrm{W}^{A,k} = (w_{ij}^{A,k})_{I \times I}$ is a matrix of $I$ rows and $I$ columns. Then the empirical acupoint-acupoint association across the dataset can be defined as,

$$\mathbf{W}^A = \frac{1}{K} \sum_{k=1}^{K} \mathrm{W}^{A,k}$$

$\frac{1}{K}$ is a normalization coefficient so that $\sum \mathbf{W}^A = 1$. $\mathbf{W}^A$ describes the synergistic effect among the acupoints for treatments of a specific disease. For all $k$, $\sum_{ij} w_{ij}^{A,k} = 1$, which means that each prescription has a similar contribution to $\mathbf{W}^A$.

#### 3) SYMPTOM-SYMPTOM ASSOCIATION

The symptom-symptom cooccurrence matrices for a given prescription $P^k$ and the overall dataset can be defined as follows,

$$w_{ij}^{S,k} = \begin{cases} \dfrac{1}{J_k \cdot (J_k - 1)} & for\ \forall j \in S^k, i \neq j \\ 0 & otherwise \end{cases}$$

$$\mathbf{W}^S = \frac{1}{K} \sum_{k=1}^{K} \mathrm{W}^{S,k}$$

where $\mathrm{W}^{S,k} = (w_{ij}^{S,k})_{J \times J}$ is a matrix of $J$ rows and $J$ columns, and $\frac{1}{K}$ is a normalization coefficient so that $\sum \mathbf{W}^S = 1$. $\mathbf{W}^S$ indicates the variety and coexistence pattern of symptoms for a specific disease. Similarly, $\sum_{ij} w_{ij}^{S,k} = 1$, for all $k$, which means that each prescription has a similar contribution to $\mathbf{W}^S$.

### B. JOINT PROBABILITY AND MUTUAL INFORMATION

From statistical point of view, acupoint selection can be regarded as a discrete random variable. For the sake of simplicity, we denoted the random variable of acupoint selection as $A = \{a_i | i = 1, 2, \ldots, I\}$ which has $I$ possible values. Similarly, the emergence of symptom can also be regarded as a discrete random variable, which can be denoted as $S = \{s_j | j = 1, 2, \ldots, J\}$ of $J$ possible values. Next, we used the following probability density functions.

#### 1) JOINT PROBABILITY DENSITY FUNCTION

The joint probability density function between $A$ and $S$ can be defined by the empirical cumulative cooccurrence matrix between them as follows,

$$\mathrm{Prob}\,(A,S) = \mathbf{W}$$

Then, we used the following definitions of probability,

$$p\,(a_i) = \mathrm{Prob}\,(A = a_i) = \sum\nolimits_j \boldsymbol{W}_{ij}$$

$$p\,(s_j) = \mathrm{Prob}\,(S = s_j) = \sum\nolimits_i \boldsymbol{W}_{ij}$$

$$p\,(a_i, s_j) = \mathrm{Prob}\,(A = a_i, S = s_j) = \boldsymbol{W}_{ij}$$

Similarly, the joint probability density function between two acupoints $a_i \in A$ and $a_j \in A$ can be defined as follows,

$$p\,(a_i, a_j) = \mathrm{Prob}\,(A = a_i, A = a_j) = \mathbf{W}_{ij}^A$$

And the joint probability density function between two symptoms $s_i \in S$ and $s_j \in S$ can be defined as,

$$p\,(s_i, s_j) = \mathrm{Prob}\,(S = s_i, S = s_j) = \mathbf{W}_{ij}^S$$

#### 2) MUTUAL INFORMATION

Based on the information theory, the mutual information between $A$ and $S$ is defined as follows [21],

$$I\,(A, S) = \sum_{a_i \epsilon A} \sum_{s_j \epsilon S} p\,(a_i, s_j) \log_2 \frac{p\,(a_i, s_j)}{p\,(a_i)\,p\,(s_j)}$$

The conditional entropy of $A$ when $S$ is given is,

$$I\,(A \mid S) = -\sum_{a_i \epsilon A} \sum_{s_j \epsilon S} p\,(a_i, s_j) \log_2 p\,(a_i \mid s_j)$$

### C. NETWORK MODELING AND VISUALIZATION

Network, or graph, is usually used to model and visualize the associations of interacted entities in many scientific research areas such as social science, computer science, and biomedicine. Here, several networks were constructed to explore the relationship between acupoints and symptoms.

#### 1) ACUPOINT-SYMPTOM ASSOCIATION (ASA) NETWORK

A weighted undirected bipartite network $G = (V, E)$ was used to model the acupoint-symptom association, where the nodes set $V = \{A, S\}$ is consisted of 30 acupoints and 13 symptoms, the adjacent matrix $E$ of the network is,

$$e(a_i, s_j) = e(s_j, a_i) = \begin{cases} \mathbf{W}_{ij} & if\ p\left(a_i, s_j\right) > p\left(a_i\right) \cdot p\left(s_j\right) \\ 0 & otherwise \end{cases}$$

where $\mathbf{W} = (\mathbf{W}_{ij})$ is the joint probability density matrix of $(A, S)$.

#### 2) ACUPOINT-ACUPOINT ASSOCIATION (AAA) NETWORK

A weighted undirected network $G^A = \left(A, E^A\right)$ was used to model the acupoint-acupoint association, where $A$ is the acupoints set, and the edges set $E^A$ is assigned as follows,

$$e(a_i, a_j) = \begin{cases} \mathbf{W}_{ij}^A & if\ \ p\left(a_i, a_j\right) > p\left(a_i\right) \cdot p\left(a_j\right) \\ 0 & otherwise \end{cases}$$

where $\mathbf{W}^A = (\mathbf{W}_{ij}^A)$ is the joint probability density matrix of $A$.

#### 3) SYMPTOM-SYMPTOM ASSOCIATION (SSA) NETWORK

A weighted undirected network $G^S = \left(S, E^S\right)$ was used to model the symptom-symptom association, where $S$ is the symptoms set, and the edges set $E^S$ is assigned as follows,

$$e(s_i, s_j) = \begin{cases} \mathbf{W}_{ij}^S & if\ \ p\left(s_i, s_j\right) > p\left(s_i\right) \cdot p\left(s_j\right) \\ 0 & otherwise \end{cases}$$

where $\mathbf{W}^S = (\mathbf{W}_{ij}^S)$ is the joint probability density matrix of $S$.

Networks were visualized using Cytoscape 3.7.1 (http://www.cytoscape. org) [22]. A circular layout was used to show the network characteristics with degree centrality denoted as the size of nodes and the value of adjacent matrix denoted as linewidth of edges.

### D. ACUPOINTS COMBINATION EVALUATION AND SCREENING

Let $A^k \subseteq A$ be a non-empty acupoints subset of $A$, $S^k \subseteq S$ be a non-empty symptoms subset of $S$. We defined pesudo mutual information between $A^k$ and $S^k$ as,

$$I\left(A^k, S^k\right) = \sum_{a_i \epsilon A^k} \sum_{s_j \epsilon S^k} p\left(a_i, s_j\right) \log_2 \frac{p\left(a_i, s_j\right)}{p\left(a_i\right) p\left(s_j\right)} \quad (1)$$

The pseudo mutual information $I\left(A^k, S^k\right)$ can be used to evaluate the fitness of the combination of $A^k$ for the symptoms pattern $S^k$. The higher $I\left(A^k, S^k\right)$, the better fitness of $A^k$ is for $S^k$. In this study, we use the pseudo mutual information to screen an acupoints combination for a given symptoms set.

#### 1) SIGNIFICANCE TEST OF ACUPOINTS COMBINATION

For a given non-empty symptoms subset $S^k \subseteq S$, the significance of fitness of a non-empty acupoints combination $A^k \subseteq A$ can be tested by the following procedure,

*Step 1:* Calculate the pseudo mutual information $I^k = I\left(A^k, S^k\right)$ using **Eq.1**.

*Step 2:* Randomly select $m = |A^k|$ acupoints from $A$ to make a subset $A^i$, calculate the pseudo mutual information $I^i = I\left(A^i, S^k\right)$ using **Eq.1**.

*Step 3:* Repeat **Step 2** for 1000 times, calculate the nominal $p$ value by,

$$p = 1 - \frac{\#(I^k \geq I^i)}{1000}$$

where $\#(I^k \geq I^i)$ means the number of times that $I^k \geq I^i$. Given a confident level, e.g., $\alpha = 0.05$, it means that the fitness of $A^k$ is significant with respect to $S^k$ if $p \leq \alpha$.

#### 2) ACUPOINT COMBINATION SCREENING

For a given non-empty symptoms subset $S^k \subseteq S$, we define a cost function $J_m\left(A^i\right)$ for an acupoint combination $A^i$ $(m = \left|A^i\right|)$ by the pseudo mutual information between $A^i$ and $S^k$ as follows,

$$J_m\left(A^i\right) = I\left(A^i, S^k\right)$$

where $m$ means there are $m$ acupoints in $A^i$. Then, the optimal acupoints combination $A^o$ $(m = |A^o|)$ is a subset which satisfies with,

$$I\left(A^o, S^k\right) = \max_{A^i \subseteq A} J_m\left(A^i\right) \quad (2)$$

Acupoint combination screening is to find the optimal subset $A^o$ for a given symptoms subset $S^k$.

#### 3) ANN MODEL FOR ACUPOINTS SCREENING

A feed-forward network, with a single hidden layer between an input and an output layer, was implemented using Matlab (R2018b). Similar to Ref. [19], a rectified linear function was used as an activation function for input nodes to convert weighted input into activation output. A sigmoid function was used as the activation function for hidden and output nodes. In our case, 13 different symptoms were assigned to 13 input nodes and 30 acupoints were assigned to 30 output nodes. Confirmed by a tenfold cross validation, the optimal number of hidden nodes was set as 5. Performance of MIS in acupoints selection was compared with ANN through a tenfold cross validation. In the validation process, the parameters including precision, recall and F1 score were used to indicate the performance. Precision represents the ratio of matched acupoints prescriptions to those ones predicted by ANN/MIS and recall is the ratio of matched acupoints prescriptions to actual ones in records. On the other hand, F1 score represented a compromise of precision and recall. It was defined as,

$$\text{F1 score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

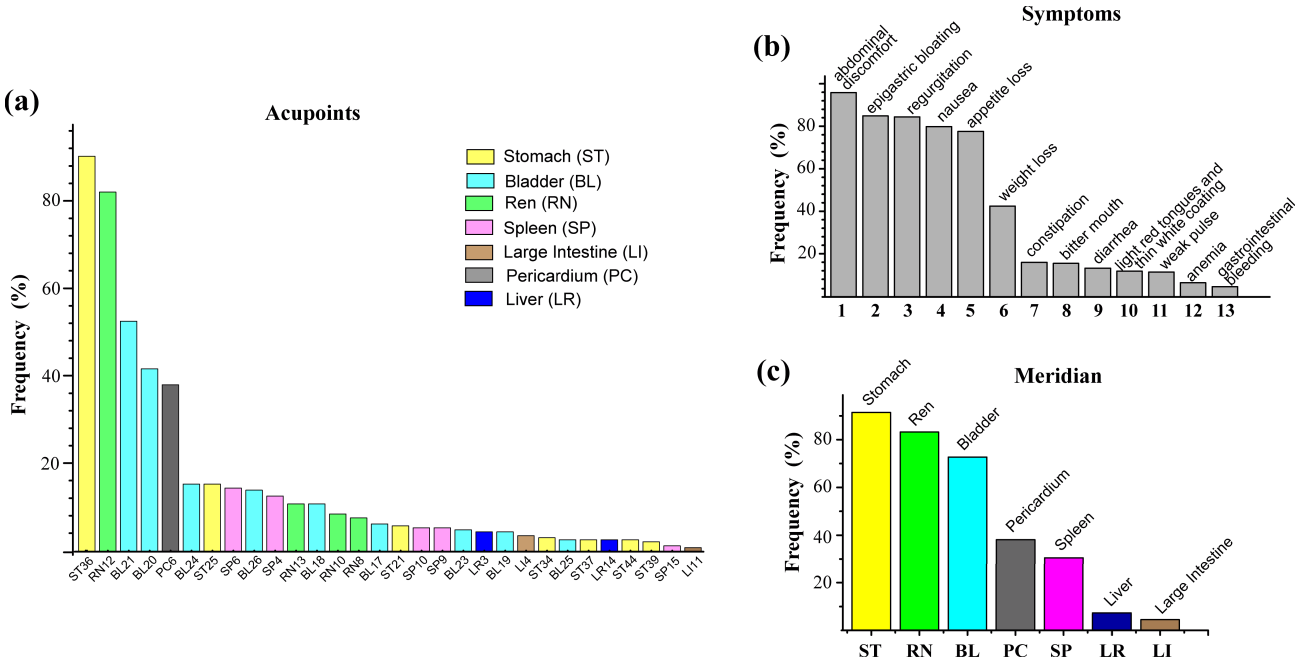


**FIGURE 1.** Frequency of occurrence of acupoints (a), symptoms (b), and meridians (c) in the prescription dataset.

## IV. RESULTS AND DISCUSSION

### A. OVERVIEW OF COLLECTED PRESCRIPTIONS

A total of 30 acupoints on seven meridians were used to alleviate the 13 symptoms of acute and chronic gastritis. Statistics on occurrence frequencies of the collected database were plotted in **Figure 1**.

The three most frequently used acupoints included the ST36 (*Zusanli*) on the Stomach Meridian, the RN12 (*Zhongwan*) on the Ren Meridian and the BL21 (*Weishu*) on the Bladder Meridian, accounted for 90.5%, 82.3% and 52.7% of the prescriptions, respectively, as shown in **Figure 1a**. ST36 has been widely used in clinical acupuncture for treating a variety of medical conditions especially gastrointestinal disorders [23]. In addition, electroacupuncture at RN12 and BL21 acupoints have been previously reported to regulate gastric motility probably through the paraventricular hypothalamic nucleus (PVN) - dorsal vagal complex (DVC) – vagus – gastric neural pathway [24].

**Figure 1b** shows that the top five symptoms included abdominal discomfort (associated with 95.9% prescriptions), epigastric bloating (85.0%), regurgitation (84.5%), nausea (80%) and appetite loss (77.7%).

**Figure 1c** shows that the most frequently used meridians included the Stomach Meridian (90.9%), the Ren meridian (83.2%), and the Bladder Meridian (72.7%). As expected, Stomach meridian is on the top of meridians list since it is the main meridian for gastric disease treatment. Acupoints on the Bladder Meridian are not only useful from a musculoskeletal perspective [25] but also can be used for a range of aliments from eye disorders [26] to gastric motility [27].

### B. ASSOCIATION NETWORKS VISUALIZATION

The acupuncture treatment has been viewed as a comprehensive regulation of multiple functional organs. Here, we provided the visualization and numerical characterization of three association networks to facilitate illustrating the patterns of symptoms, acupoints, and their cooccurrence relation in the prescriptions collection. In consideration of the non-linearity of acupoint-symptom association, mutual information [28] is used to measure the dependency of acupoint-symptom, acupoint-acupoint and symptom-symptom in acupoints combinations.

As shown in **Figure 2a,** ASA network is a weighted bipartite network, it is also a heterogeneous network with two-different kinds of nodes, each node represents an acupoint or a symptom. ASA network may provide a holistic perspective to explore the characteristics of acupoints through their connecting patterns with symptoms. Taking symptom of abdominal discomfort as an example, it was observed to be strongly connected to ST21, PC6, RN12 and RN8, which implied that part or all of these four acupoints had a higher probability of being selected as the candidate prescription to alleviate ''abdominal discomfort'' by acupuncturists. The acupoint of BL21 contributes extensively to most of symptoms in gastritis indicating the general applicable feature of BL21, whereas the acupoint of ST34 showed with only a single edge exclusively applicable on a specific symptom (i.e. nausea) reflecting the acupoint specificity as in the acupuncture theory.

In **Figure 2b**, the AAA network is shown with the nodes representing to acupoints and the edges showing

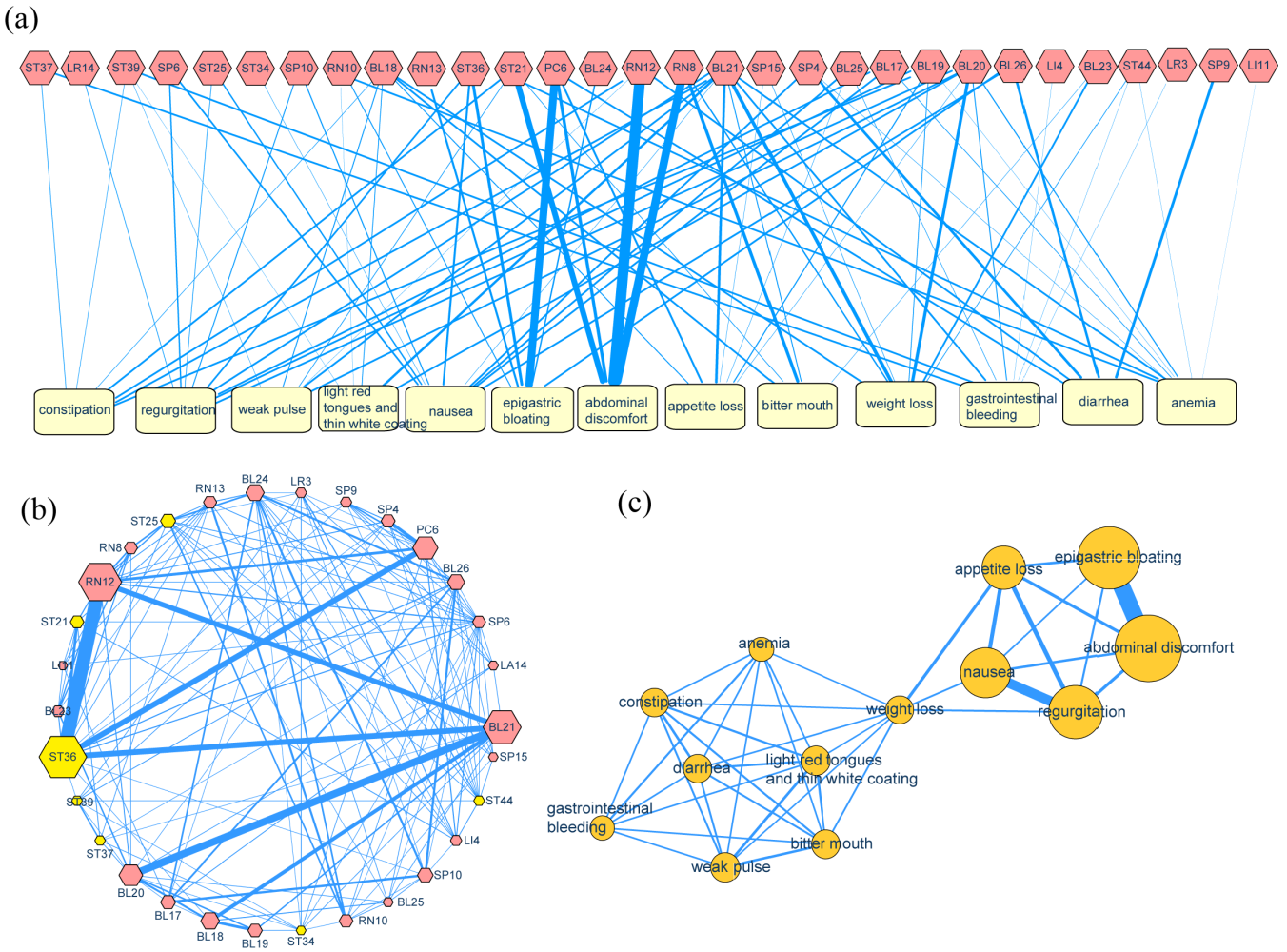


**FIGURE 2.** **Network visualization. (a) The weighted bipartite network of ASA; (b) The weighted undirected network of AAA; (c) the weighted undirected network of SSA.**

the collaboration between different acupoints. Degree of node [29] representing the number of one acupoint connected to others, which highlights the role of one acupoint to be applied cooperatively with other acupoints. The larger size of node means a higher synergistic capacity of that acupoint to be combined with other acupoints. ST36 was observed to be largest node in AAA network which indicated a general cooperativity of ST36 to other acupoints in treating gastritis. It is in accordance with the findings of occurrence frequency in **Figure 1**. Nodes of RN12, BL21, BL 20 and PC6 also showed in high degree, reflecting their roles as significant hubs in the AAA network. Moreover, weight of edge between two acupoints reflects their co-occurrence probability, i.e. a larger weigh of edge showed a higher frequency for the two acupoints to be used together in prescription. The top five weighted edges were the connection of ST36-RN12 (''*ZuSanLi*'' and ''*Zhong Wan*''), BL21-BL20 (''*Wei Shu*'' and ''*Pi Shu*''), ST36-BL21 (''*ZuSanLi*'' and ''*Wei Shu*''), RN12-BL21 (''*Zhong Wan*'' and ''*Wei Shu*'') and ST36-PC6 (''*ZuSanLi*'' and ''*Nei Guan*''). Based on literatures about acupoints selection, the ST36-RN12 combination was identified to be compliance with the rule of ''He-Mu combination'' that are lower sea points of six Fu-organs and Mu points of the affected meridian with a recovering effect on the physiological function of organs [30]. The back-Shu acupoint of the Stomach meridian BL21 combining to the lower sea point ST36, locating at the back and front of the body respectively, followed the rule of ''anterior-posterior combination'' and has been widely used to relieve the nausea and vomiting induced by acute gastritis. The RN12-BL21 combination, known as the ''Shu-Mu combination'' in the acupuncture theory, has been shown to be effective for regulation of gastric motility in clinical practice [24]. PC6, the collateral point of the Pericardium Meridian and also links the Tri-energizer Meridian, thus can function in treating diseases of both the cardiovascular system and the digestive system [31]. With the cooperation of ST36 or RN12, PC6 was reported to treat multiple types of gastrointestinal disorders such as gastroparesis [32], functional dyspepsia [33], achalasia of cardia and gastritis.

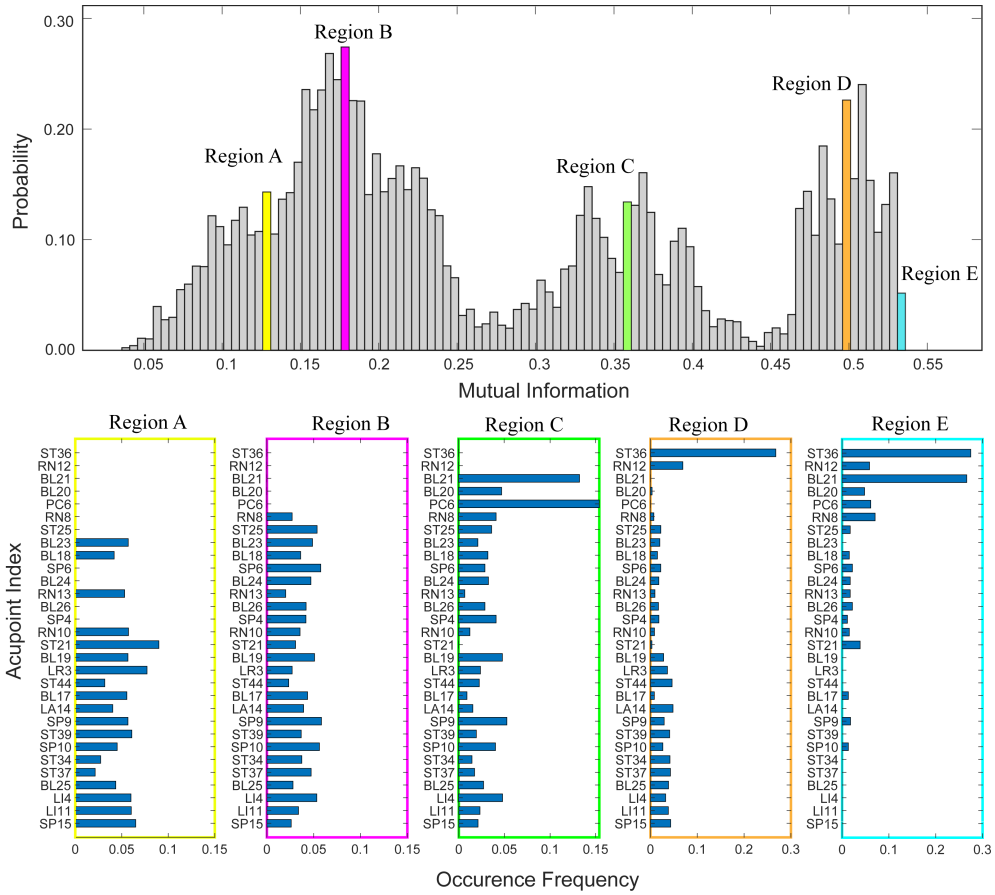


**FIGURE 3.** **The distribution of pseudo MI (the top row) on a given symptoms including abdominal discomfort, epigastric bloating, nausea and bitter mouth with 3-acupoint combination by 100 000 resampling. (Region A-E shows the occurrence frequency of acupoints corresponding to the intervals of mutual information 0.130-0.135, 0.175-0.180, 0.365-0.370, 0.495-0.500, 0.531-0.536 respectively).**

Gastritis was known to arise from various endogenous and exogenous factors thus leading to complicated clinical manifestation. In current database, there are 9.5% of prescriptions recorded with more than 10 symptoms at one time, but 15.9% of prescriptions were recorded with less than three symptoms. In average, the prescription was associated with about 5 symptoms. SSA network (**Figure 2c**) showed that the 13 symptoms can be divided into two communities. One is consisted of five symptoms including abdominal discomfort, epigastric bloating, regurgitation, nausea, and appetite loss; another is consisted of seven symptoms including bitter mouth, diarrhea, constipation, anemia, weak pulse, gastrointestinal, light red tongues and thin white coating. The two communities are connected by ''weight loss''. Symptoms from a same community may be caused by similar pathology of gastritis. This phenomenon suggested that there are at least two different pathologies in the patients involved in the clinical records. Further study is worthy to do on deciphering the difference between the two communities of symptoms.

### C. ACUPOINTS SELECTION FOR GIVEN SYMPTOMS

In the proposed MIS method, the optimal acupoints set for a given symptoms subset $S^k$ is defined as an acupoint subset $A^o$ which maximizes the pseudo mutual information between $S^k$ and $A^o$, which can be screened using **Eq.2**.

To validate the principle of pseudo mutual information maximization, a specific symptoms subset including abdominal discomfort, epigastric bloating, nausea and bitter mouth, was taken as an example. A total of 100 000 acupoints sets (three acupoints in each set) were resampled from the 30 acupoints in a random way, then pseudo MI values between the resampled acupoints sets and the given symptoms set was calculated and plotted in **Figure 3**. Five intervals corresponding to MI values, say [0.130, 0.135], [0.175, 0.180], [0.365, 0.370], [0.495, 0.500], [0.531, 0.536], were picked out manually to display the occurrence frequency of each acupoint. The ST36, RN12, BL21, BL20 and PC6 acupoints were identified to be the five most frequently used acupoints (**Figure 1**). When pseudo MI was lower than 0.2 (Region A and B), none of these top five acupoints were found in any acupoints combination. The occurrence of other acupoints was just about 0.05, suggesting their comparable but relative less importance to be chosen into a prescription. When moderate pseudo MI was obtained (Region C), three of these top five acupoints (BL21, BL20 and PC6) were selected as the candidates which occupied a key place in

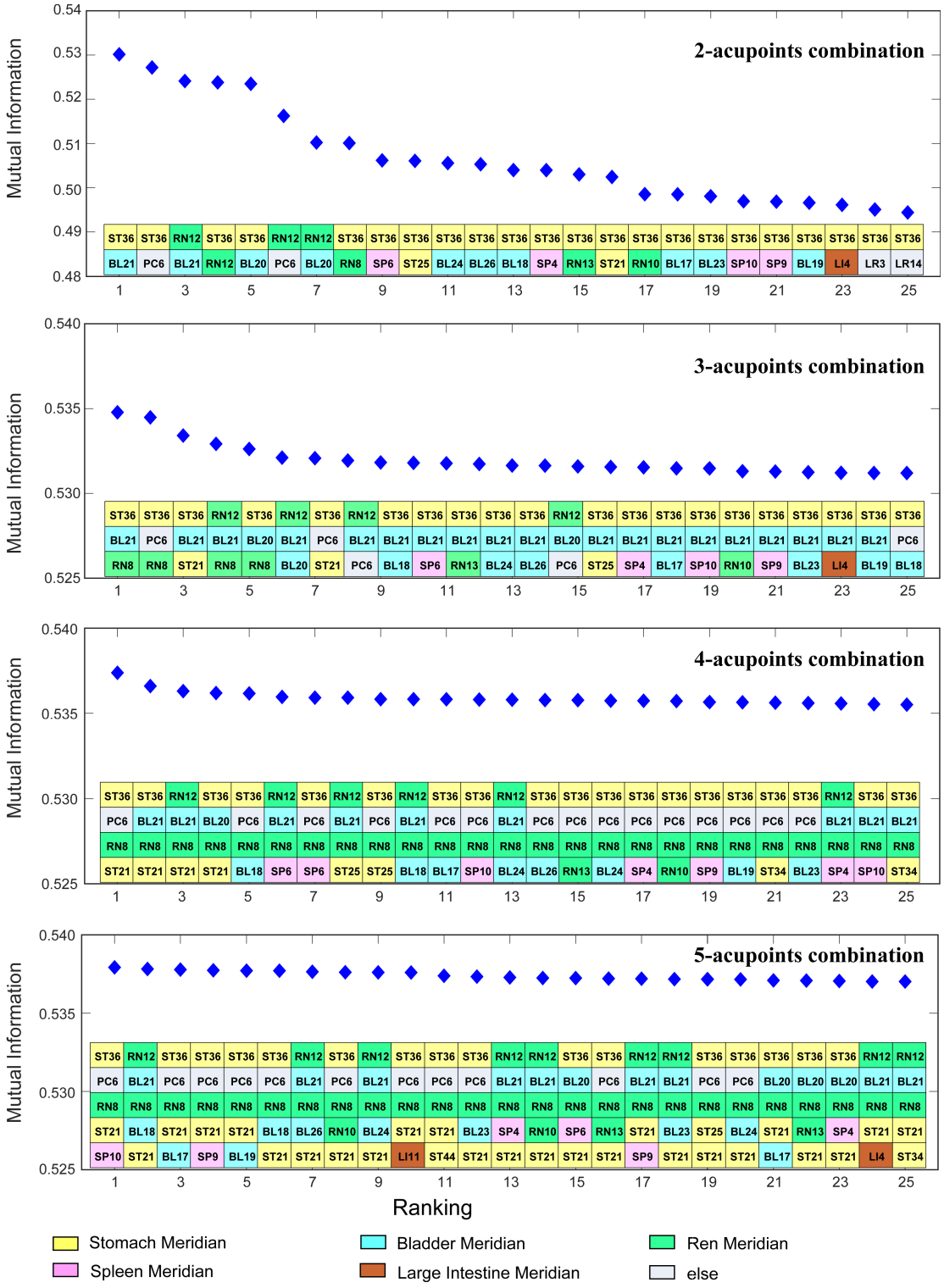


**FIGURE 4.** The top 25 MI from specified number of acupoints combinations ranked in a descending order with the corresponding combinations below.

acupoints combination. When pseudo MI increased to around 0.5 (Region D), ST36 was chosen as part of prescription by a greater probability compared to other acupoints. When pseudo MI reached the maximum (Region E), ST36 and BL21 made a dominant contribution that they were prone to be selected in one prescription. This result is consistent with a previous study on acupuncture treatment on gastric ulcer [34].

To investigate the combination pattern of acupoints with respect to the pseudo MI, the top 25 pseudo MI values were plotted with the acupoints combinations in **Figure 4**.

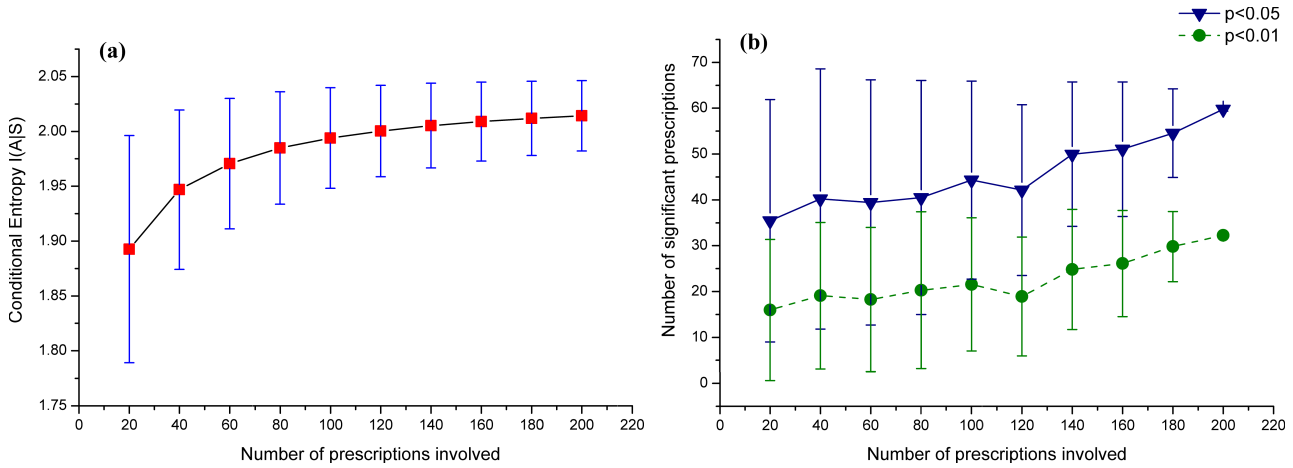


**FIGURE 5.** The conditional entropy of acupoints and symptoms (a) and number of significant prescriptions (b) when different number of prescriptions are involved in the calculation of joint probability density matrix Prob (*A*,*S*).

The maximum pseudo MI slightly increased with the increasing number of selected acupoints. For two-acupoints combination, the first four recommended combinations are ST36-BL21, ST36-PC6, RN12-BL21 and ST36-RN12. These selected acupoints combinations have corresponded to the combination patterns shown in ASA network (**Figure 2a**) which implied an effective learning on the rules of acupoints compatibility hiding in practical experience. Briefly, the ST36 acupoint combined with the BL21 acupoint is in accordance with the rule of ''anterior-posterior combination''. ST36 and PC6, locating at the lower limb and upper limb, respectively, are categorized as the ''upper-lower combination''. RN12-BL21 and ST36-RN12 are known as the rule of ''Shu-Mu combination'' and ''He-Mu combination'' in acupuncture theory. While in the situation of the three-acupoints combination, RN8 (*Shen Que*) was added on the basis of 2-acupoints combination which constituted the first two recommendations. The major functions of RN8 are reinforcing the fundamental and cultivating the vital energy, invigorating the spleen and stomach and regulating gastrointestinal function. It is a commonly used acupoint involved in alleviating abdominal pain [35] on menstrual period that is the analogical symptom in present testing. As in four-acupoints prescription, ST21 is more likely to be introduced as the fourth acupoint because it is believed to cooperate with ST36 for promoting damage repair on gastric mucosa [36]. For five-acupoints prescription, the acupoints combination is more complicated and not straightforward to be explained by using a single rule as in acupuncture theory.

### *D. PERFORMANCE OF MIS METHOD*

The calculation of mutual information, visualization of association network, and screening of acupoints combinations are all based on the joint probability density function (PDF) between acupoints and symptoms Prob($A$, $S$). The estimation of Prob($A$, $S$) plays a crucial role in the reliability and stability of the proposed method. While, the accuracy of joint PDF estimation is mainly determined by the sample size $n$, i.e., the number of prescriptions. The larger sample size is, the small estimation error is for PDF [37]. Two metrics – conditional entropy $I(A \mid S)$ and nominal $p$-value are employed to evaluate the performance of MIS method with respect to the sample size $n$, as shown in **Figure 5**. Also, prediction accuracies of MIS method are calculated and compared with ANN-based method.

The conditional entropy $I(A \mid S)$ can be interpreted as the measure of uncertainty of variable $A$ under the condition of known variable $S$ in the dataset [38]. In other words, $I(A \mid S)$ represents the information content of variable $A$ undetermined by variable $S$. A higher $I(A \mid S)$ indicates a lower entropy of variable $A$ determined by a given variable $S$. Ten-different sample size of prescriptions are resampled and used to estimate the joint PDF, then the conditional entropy $I(A \mid S)$ is calculated, as shown in **Figure 5a**. With the increasing of sample size (numbers of documents), $I(A \mid S)$ rose gradually while the corresponding variance decreased. In the case of screening the acupoints combinations for a given symptoms, high $I(A \mid S)$ with low variance from the joint probability matrix with larger level of involvement suggested the trend of a stable and reliable acupoints combination to be selected for symptoms. It can also be inferred that a joint PDF estimated using more prescriptions will lead to a more reliable MIS method for acupoints combinations.

Validation of clinical records (prescriptions) are carried out to evaluate the significance and efficiency of MIS method. Nominal $p$ value of each prescription is calculated using different joint PDFs estimated under different sample sizes $n$. Nominal $p < 0.05$ (or $p < 0.01$) means that acupoints combination of the prescription is significant in the sense of pseudo mutual information maximization. The more prescriptions identified to be significant, the more significant of the MIS method. With the increase of the sample size $n$, as shown in **Figure 5b**, more prescriptions are of significance ($p < 0.05$ or $p < 0.01$), as well as the obviously

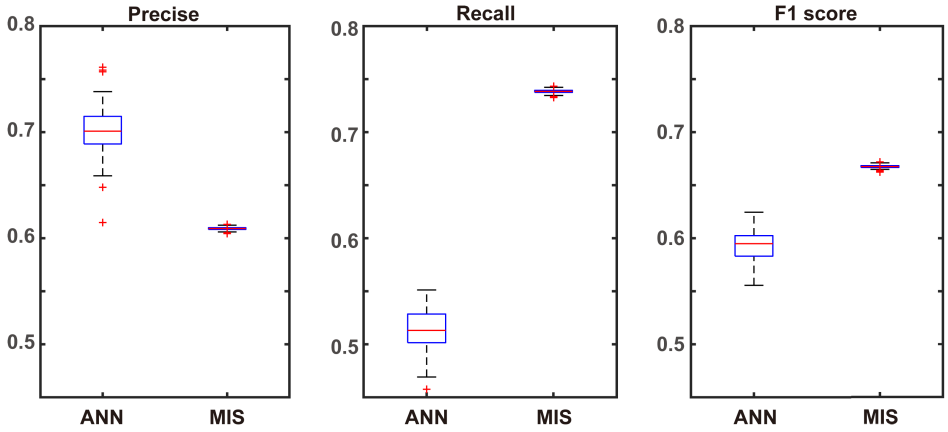


**FIGURE 6.** The comparison of MIS and ANN in performance of acupoints selection through a tenfold cross validation.

decreased variance. The results imply that more significant acupoints combinations would be selected by MIS method when larger prescriptions database is applied. The way to promote efficiency and robustness of MIS is to expand the sample size of clinical records.

Besides, ANN-based method [19] was used to compare with the developed MIS method (**Figure 6**). When the ANN was trained with 5 hidden nodes for 100 epochs, an average F1score $= 0.593 \pm 0.014$ was obtained (precision, $0.702 \pm 0.023$; recall, $0.513 \pm 0.020$). F1 score of MIS was $0.667 \pm 0.002$ and precision and recall were $0.609 \pm 0.001$ and $0.738 \pm 0.002$. The training of ANN model was susceptible to arrangement of training samples and initialization of weights thus leading to larger variance of predicted results than MIS. The current results indicated that MIS provided a better performance and more reliable prescriptions than the ANN model.

## V. CONCLUSION

This study applied network analysis to explore the interconnectedness of acupoints and acupoints, symptoms and symptoms which were present in 220 prescriptions of acupuncture treatment of gastritis. The associations of acupoints and symptoms were comprehended and characterized as an acupoints-symptoms joint probability matrix from which the patterns of acupoints combination can be learned. The acupoints combination for given symptoms was screened with the principle of pseudo mutual information maximization. The optimized acupoints combination was prone to be the most appropriate one for treating current symptoms and recommended for clinical applications. In addition, the significance and robustness of MIS were also assessed by conditional entropy and significance test. The developed MIS method for acupoints combination was found to be a reliable, quantitative and practical tool for symptom-specific acupoint selections, which may provide a scientific basis of clinical acupuncture therapy and help to pave the way for exploring the rules of acupoints compatibility.

## REFERENCES

[1] J. Jia, Y. Yu, J.-H. Deng, N. Robinson, M. Bovey, Y.-H. Cui, H.-R. Liu, W. Ding, H.-G. Wu, and X.-M. Wang, "A review of omics research in acupuncture: The relevance and future prospects for understanding the nature of meridians and acupoints," *J. Ethnopharmacology*, vol. 140, no. 3, pp. 594–603, Apr. 2012.

[2] T. J. Kaptchuk, "Acupuncture: Theory, efficacy, and practice," *Ann. Internal Med.*, vol. 136, no. 5, pp. 374–383, Mar. 2002.

[3] E. Doll, B. Threlkeld, D. Graff, R. Clemons, O. Mittel, M. K. Sowell, and M. McDonald, "Acupuncture in adult and pediatric headache: A narrative review," *Neuropediatrics*, vol. 50, no. 06, pp. 346–352, Dec. 2019.

[4] H. Abdi, B. Zhao, M. Darbandi, M. Ghayour-Mobarhan, S. Tavallaie, A. A. Rahsepar, S. M. R. Parizadeh, M. Safariyan, M. Nemati, M. Mohammadi, P. Abbasi-Parizad, S. Darbandi, S. Akhlaghi, and G. A. A. Ferns, "The effects of body acupuncture on obesity: Anthropometric parameters, lipid profile, and inflammatory and immunologic markers," *Sci. World J.*, vol. 2012, Jan. 2012, Art. no. 603539.

[5] Y. Zhang, J. A. Dennis, F. L. Bishop, H. Cramer, M. Leach, R. Lauche, T. Sundberg, B. Leung, A. L. Zhang, L. Bacon, D. Sibbritt, and J. Adams, "Complementary and alternative medicine use by U.S. adults with self-reported doctor-diagnosed arthritis: Results from the 2012 national health interview survey," *PM&R*, vol. 11, no. 10, pp. 1059–1069, Oct. 2019.

[6] J. Xu, X. Zheng, K.-K. Cheng, X. Chang, G. Shen, M. Liu, Y. Wang, J. Shen, Y. Zhang, Q. He, J. Dong, and Z. Yang, "NMR-based metabolomics reveals alterations of electro-acupuncture stimulations on chronic atrophic gastritis rats," *Sci. Rep.*, vol. 7, no. 1, p. 45580, Mar. 2017.

[7] Y.-R. Chen, Z. Jiang, J.-S. Song, and Y.-F. She, "Discussion on point selection and compatibility of acupuncture formula," *Chin. Acupuncture Moxibustion*, vol. 32, no. 1, pp. 65–68, 2012.

[8] G. Stux and R. Hammerschlag, Eds., "Clinical acupuncture," in *Scientific Basis*. Berlin, Germany: Springer, 2001.

[9] X. Wang, H. Sun, A. Zhang, W. Sun, P. Wang, and Z. Wang, "Potential role of metabolomics apporoaches in the area of traditional Chinese medicine: As pillars of the bridge between Chinese and western medicine," *J. Pharmaceutical Biomed. Anal.*, vol. 55, no. 5, pp. 859–868, Jul. 2011.

[10] S. Yin, Y. Chen, R. R. S. Du Lei, T. T. Ma, P. M. Feng, Z. X. He, X. L. Suo, P. H. Ma, Y. Z. Qu, K. Qiu, and M. M. Jing, "Cerebral mechanism of puncturing at He-Mu point combination for functional dyspepsia: Study protocol for a randomized controlled parallel trial," *Neural Regen. Res.*, vol. 12, no. 5, pp. 831–840, May 2017.

[11] F. Xu and R. Chen, "Reciprocal actions of acupoints on gastrointestinal peristalsis during electroacupuncture in mice," *J. Traditional Chin. Med.*, vol. 19, no. 2, pp. 4–141, Jun. 1999.

[12] C.-C. Yu, L.-H. Kong, C.-Y. Ma, F. Shen, G.-J. Yao, Y. Xiong, and H.-J. Zhou, "The rules of acupoint-selection of acupuncture for polycystic ovary syndrome based on data mining," *World J. Acupuncture-Moxibustion*, vol. 26, no. 3, pp. 73–78, Sep. 2016.

[13] S.-H. Lee, C.-E. Kim, I.-S. Lee, W.-M. Jung, H.-G. Kim, H. Jang, S.-J. Kim, H. Lee, H.-J. Park, and Y. Chae, "Network analysis of acupuncture points used in the treatment of low back pain," *Evidence-Based Complementary Alternative Med.*, vol. 2013, Oct. 2013, Art. no. 402180.
[14] A. Inokuchi, T. Washio, and H. Motoda, "An Apriori-based algorithm for mining frequent substructures from graph data," in *Proc. Eur. Conf. Princ. Data Mining Knowl. Discovery* (Lecture Notes in Computer Science), vol. 1910, 2000, pp. 13–23.
[15] C.-R. Meng, H.-L. Zhang, L.-F. Zeng, Z.-P. Li, J. Huang, and Z. Liang, "Evidence-based decision support for the clinical practice of acupuncture: Data mining approaches," in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, G. Z. Li, Eds. Dec. 2013, pp. 180–181.
[16] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *Science*, vol. 323, no. 5916, pp. 892–895, 2009.
[17] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, Mar. 2009.
[18] S. Li, B. Zhang, D. Jiang, Y. Wei, and N. Zhang, "Herb network construction and co-module analysis for uncovering the combination rule of traditional Chinese herbal formulae," *BMC Bioinf.*, vol. 11, no. S11, Dec. 2010, Art. no. S6.
[19] W.-M. Jung, I.-S. Park, Y.-S. Lee, C.-E. Kim, H. Lee, D.-H. Hahm, H.-J. Park, B.-H. Jang, and Y. Chae, "Characterization of hidden rules linking symptoms and selection of acupoint using an artificial neural network model," *Frontiers Med.*, vol. 13, no. 1, pp. 112–120, Feb. 2019.
[20] O. Akerele, "WHO and the development of acupuncture nomenclature: Overcoming a tower of babel," *Amer. J. Chin. Med.*, vol. 19, no. 01, pp. 89–94, Jan. 1991.
[21] D. Beeferman, A. Berger, and J. Lafferty, "Statistical models for text segmentation," *Mach. Learn.*, vol. 34, nos. 1–3, pp. 177–210, Feb. 1999.
[22] M. E. Smoot, K. Ono, J. Ruscheinski, P.-L. Wang, and T. Ideker, "Cytoscape 2.8: New features for data integration and network visualization," *Bioinformatics*, vol. 27, no. 3, pp. 431–432, Feb. 2011.
[23] Y. Chen, J. Xu, S. Liu, and X. Hou, "Electroacupuncture at ST36 increases contraction of the gastric antrum and improves the SCF/c-kit pathway in diabetic rats," *Amer. J. Chin. Med.*, vol. 41, no. 6, pp. 1233–1249, Jan. 2013.
[24] H. Wang, W. J. Liu, G. M. Shen, M. T. Zhang, S. Huang, and Y. He, "Neural mechanism of gastric motility regulation by electroacupuncture at RN12 and BL21: A paraventricular hypothalamic nucleus-dorsal vagal complex-vagus nerve-gastric channel pathway," *World J. Gastroenterology*, vol. 21, no. 48, pp. 13480–13489, Dec. 2015.
[25] K. Ishida, L. Qin, T. Wang, Y. Lei, W. Hu, G. Zhu, Q. Zhuang, Y. Gong, Q. Wang, G. Arijit, D. Ma, X. Li, and S.-T. Li, "Local mechanisms for acupoint sensitization in gall bladder meridian of foot Shaoyang—A gene chip study," *Acupuncture Electro-Therapeutics Res.*, vol. 44, no. 2, pp. 77–87, Jun. 2019.
[26] M. S. Lee, B.-C. Shin, T.-Y. Choi, and E. Ernst, "Acupuncture for treating dry eye: A systematic review," *Acta Ophthalmologica*, vol. 89, no. 2, pp. 101–106, Mar. 2011.
[27] H. Wang, G.-M. Shen, W.-J. Liu, S. Huang, and M.-T. Zhang, "The neural mechanism by which the dorsal vagal complex mediates the regulation of the gastric motility by Weishu (RN12) and Zhongwan (BL21) stimulation," *Evidence-Based Complementary Alternative Med.*, vol. 2013, pp. 1–7, Jan. 2013.
[28] L. Hussain, A. Ali, S. Rathore, S. Saeed, A. Idris, M. U. Usman, M. A. Iftikhar, and D. Y. Suh, "Applying Bayesian network approach to determine the association between morphological features extracted from prostate cancer images," *IEEE Access*, vol. 7, pp. 1586–1601, 2019.
[29] A. L. Barabasi, "Network science," *Philos. Trans. Roy. Soc. A, Math. Phys. Eng. Sci.*, vol. 371, no. 1987, Mar. 2013, Art. no. 20120375.
[30] F. C. Wang, "Discussion on 'analogical acupoints,"' *World J. Acupuncture-Moxibustion*, vol. 25, no. 1, pp. 24–27, 2015.
[31] S. B. Wang and J. L. Liu, "Clinical application of Neiguan (PC6) in the treatment of disorders of the digestive system-the related research progress," *World J. Acupuncture-Moxibustion*, vol. 14, no. 4, pp. 3–11, 2004.
[32] K. B. Cheong, J.-P. Zhang, and Y. Huang, "The effectiveness of acupuncture in postoperative gastroparesis syndrome—A systematic review and meta-analysis," *Complementary Therapies Med.*, vol. 22, no. 4, pp. 767–786, Aug. 2014.
[33] J. Yin and J. D. Z. Chen, "Gastrointestinal motility disorders and acupuncture," *Autonomic Neurosci.*, vol. 157, nos. 1–2, pp. 31–37, Oct. 2010.
[34] L. S. Wang, Y. X. Liu, and Q. Tu, "Effect of acupuncture of 'Zusanli' and 'Weishu' acupoints on experimental gastric ulcer," *World J. Acupuncture-Moxibustion*, vol. 13, no. 1, pp. 24–29, 2003.
[35] J. Gao, Q. Wang, S. Xian, Y.-M. Feng, W.-X. Cao, J.-Y. Ye, Q.-Q. Zhang, C.-L. Zou, Q.-G. Wu, and S.-J. Liu, "The effect of moxibustion on alleviating menstrual pain in a population of young nursing students: A prospective randomized cross-over pilot study," *Complementary Therapies Med.*, vol. 23, no. 6, pp. 773–781, Dec. 2015.
[36] Y. Zhang, M.-S. Huang, C.-C. Liu, L.-Y. Lian, J.-C. Shen, Q.-D. He, Y.-J. Wang, L.-B. Zhang, M. Liu, and Z.-B. Yang, "Dynamic observation and analysis of metabolic response to moxibustion stimulation on ethanol-induced gastric mucosal lesions (GML) rats," *Chin. Med.*, vol. 14, no. 1, Oct. 2019, Art. no. 44.
[37] B. W. Silverman, *Density Estimation for Statistics and Data Analysis* (Monographs on Statistics and Applied Probability). London, U.K.: Chapman & Hall, 1986.
[38] T. M. Cover and J. A. Thomas, *Entropy, Relative Entropy, and Mutual Information*. Hoboken, NJ, USA: Wiley, 2005.

**JINGJING XU** received the B.S. degree in electronic information science and technology and the Ph.D. degree in radio physics from Xiamen University, Xiamen, China, in 2005 and 2011, respectively. She was an Engineer from 2011 to 2014 and a Senior Engineer from 2015 to 2019 with the Department of Electronic Science, Xiamen University. Her research interests include the applications of metabolomics on therapy mechanisms of acupuncture and moxibustion and development of preprocessing methods on NMR-based metabolomics.

**YIJIA LIU** received the bachelor's degree in communication engineering from Jilin University, Changchun, China, in 2013. She is currently pursuing the master's degree in electronic and communication engineering with Xiamen University, Xiamen, China. From 2016 to 2017, she was a Research Assistant with the Academy of Communication Engineering, Changchun, China. Her interests include application of machine learning in the field of communication.

**XIANGNAN XU** was born in Pingdingshan, Henan, China, in 1992. He received the B.S. degree in information engineering from the Nanjing University of Aeronautics and Astronautics, in 2014, and the M.S. degree in electronic and communication engineering from Xiamen University, Fujian, China, in 2017. He is currently pursuing the Ph.D. degree in bioinformatics with The University of Sydney, Sydney, NSW, Australia. From 2017 to 2018, he was an Algorithmic engineer in Vismarty, Xiamen, Fujian, China. His research interests include the analysis of high-throughput biotechnological data including that from nuclear magnetic resonance, mass spectrometry and next generation sequencing and centered on the development of methods, and the application of statistics to problems in -omics and biomedical research.

**KIAN-KAI CHENG** received the B.S. and M.S. degrees in bioprocess engineering from Universiti Teknologi Malaysia, in 2002 and 2005, respectively, and the Ph.D. degree in biochemistry from the University of Cambridge, in 2010. He is currently a Senior Lecturer with Universiti Teknologi Malaysia. Since 2019, he has been the Director of the Innovation Centre in Agritechnology, Universiti Teknologi Malaysia Pagoh Research Centre. His research interests include metabolomics, systems biology, and multivariate data analysis.

**ZONGBAO YANG** was born in Changde, Hunan, China, in 1973. He received the master's degree in traditional Chinese medicine from the Fujian University of TCM, and the doctor's degree in acupuncture and moxibustion from the Hunan University of TCM, Changsha, China, in 2007. He was an Associate Professor with the Department of Traditional Chinese Medicine, Xiamen University, in 2013. His research interests include the applications of metabolomics on therapy mechanisms of acupuncture and moxibustion on NMR-based metabolomics.

**JIYANG DONG** was born in Quanzhou, Fujian, China, in 1974. He received the B.S. and M.S. degrees in physics education and optical engineering from Fujian Normal University, Fuzhou, China, in 1994 and 1998, respectively, and the Ph.D. degree in condensed matter physics from Xiamen University, China, in 2001. From 2001 to 2004, he was a Postdoctoral Scientist and an Associate Professor with the National Key Laboratory for Radar Signal Processing, Xidian University, Xi'an, China. From 2004 to 2011, he was an Associate Professor with the Department of Physics, Xiamen University. Since 2011, he has been a Professor with the Department of Electronic Science, Xiamen University. He is the author of more than 100 peer review articles. His research interests include network medicine, machine learning with graph, and data analysis in biomedicine.