

# 금원사대가 처방운용패턴의 기하학적 분석 : 한의학 기초이론에 대한 새로운 접근

배효진 · 김창업\*

가천대학교 한의과대학 생리학교실

## Geometrical Analysis of Geum-Won Era Prescriptions : A Novel Approach to Traditional Asian Medicine's Basic Theory

Hyojin Bae, Chang-Eop Kim\*

Department of Physiology, College of Korean Medicine, Gachon University

Throughout history, Traditional Asian Medicine (TAM) has accumulated diverse theories and extensive knowledge from different medical scholars and schools of thought. While numerous qualitative studies have explored the unique theoretical and prescriptive characteristics of each scholar and school of thought, quantitative research with complementary roles remains limited. This study seeks to explore the potential of geometric data structure analysis as a novel quantitative method to enhance understanding of the basic theories of TAM. Specifically, by analyzing the prescriptions of the four noted scholars of Geum-Won era, we evaluated the similarity of representation of herbs and analyzed the topological properties of prescription networks. Our results suggest that there are both subspaces that are significantly shared among scholars and unique subspaces that are not aligned by linear transformation. From the constructed prescription networks, we found a pattern of continuous variation among the prescriptions, which generates a specific directional trajectory on the herbal manifold. Furthermore, we can derive specific directions for future research from the analysis results. Our study highlights that the analysis of TAM theories using mathematical methods can contribute to the modernization of the theories and allow the existing theories to be understood from a novel perspective.

keywords : Geometrical Analysis, Representational Similarity, Network Topology, Modernization of Traditional Medicine, Traditional Medicine Education

### 서론

한의학은 역사적으로 다양한 의가와 학파를 통해 다채로운 이론과 풍부한 지식을 축적해왔다. 각 의가와 학파의 독특한 이론과 처방특성에 대한 정성적인 고찰과 연구는 다수 이루어져 왔으나<sup>1-3)</sup> 상호보완적 관계의 정량적 연구는 비교적 제한적으로 수행되었다<sup>4-6)</sup>. 정성적인 연구방법론은 경험한 것을 설명할 수 있는 이론을 개발하기 위해 실재를 관찰하고 해석하는 것을 목적으로 하며, 여러 실재가 존재하고 이에 대한 개인마다의 여러 해석이 가능하며 모두 똑같이 타당함을 전제로 한다. 이에 반해 정량적인 연구방법론은 개인간의 공통된 객관적 실재를 가정하며, 이론 혹은 가설로부터 시작하여 이를 검증하는 것을 목적으로 한다.<sup>7)</sup>

본 연구에서는 한의학 기초이론의 정량적 연구확대를 위한 새로운 방법론 제시를 목표로 한다. 한의학이 포착하고 있는 시스템 수준의 복잡성은 선형적 분석이 가능한 형태가 아닌 고차원의 변수 공간에 비선형적 매니폴드(manifold, 다양체) 형태로 나타나고 있을 것이라 예상된다. 따라서 고차원 공간상의 매니폴드로 표현된 데이터의 구조 혹은 모양(shape)에 대한 정량적 분석은 한의학의 내재된 비선형적 관계정보를 포착하고 이해할 수 있는 기반을 제공할 것이다. 본 논문에서는 한의학 기초이론에 대한 기하학적 구조 분석의 시험적 사례를 제공하기 위하여 금원사대가의 처방정보를 활용하여 의가 별 본초조합 패턴의 상관관계와 처방 네트워크의 구조적 특성에 대한 정량적 비교분석을 수행하였다.

본 논문에서 소개하는 연구방법론의 중요성은 크게 두가지 측면에서 논의될 수 있다. 첫째는 기초 한의학 이론의 현대화를 위한 노력의 일환으로서 다양한 정량적 분석방법론 도입의 중요성이다. 이는 기초이론에 대한 시험가능한 가설의 생성과 검증을 가능케 하며, 객관적 형태의 지식창출을 통해 한의학의 주관성이라는 오랜 한계를 극복하고 현대화를 가능케하는 발판을 마련할 수 있다. 둘째,

교육적 측면에서 볼 때 이 같은 접근방식은 한의학 기초이론에 대한 수학적, 구조적 사고의 틀을 제공한다. 더불어 연구결과로서 도출된 다양한 시각적 자료를 통한 학습의 효율성 증진 효과를 기대할 수 있다.

### 연구대상 및 방법

1. 금원사대가 본초공간의 표상 유사도(representational similarity) 분석

1) 데이터 수집

본 연구에서는 금원사대가의 대표서적과 4가지 질환군을 다음과 같이 특정한 후 의가와 질환계통별로 사용된 처방과 그 본초 구성을 수집하였다. 이후 각 처방을 행(row)으로, 전체 처방에 포함된 모든 본초를 열(column)으로 하는 데이터 행렬을 구성하고, 해당 행의 처방에 포함된 본초는 1으로, 나머지는 0으로 값을 할당하였다(one-hot-encoding). 분석에 포함된 전체 처방과 본초의 수는 각각 1683, 1211개였다.

- 서적: 유완소(劉完素):황제소문선명론방(黃帝素問宣明論方), 장종정(張從正):유문사친(儒門事親), 이고(李杲):비위론(脾胃論), 의학발명(醫學發明), 주진형(朱震亨):단계심법(丹溪心法).

- 질환계통: 소화기계(Digestive system), 호흡기계(Respiratory system), 부인과(Gynecology), 피부과(Dermatology)

2) 표상 유사도

Svcca(Singular vector canonical correlation analysis)<sup>8)</sup>는 심층신경망 모델에서 각 레이어 간 학습된 표상의 유사도를 비교하기 위하여 제안된 방법론으로, 변수군간 상관관계를 분석하기 위해 사용되는 정준상관분석(canonical correlation analysis, CCA)에 앞서 특이값 분해(singular value decomposition, SVD)를 통해 차원을 축소한다는 것이 특징이다. 특이값 분해는 신호처리와 통계

\* Corresponding author

Chang-Eop Kim, College of Korean Medicine, Gachon University, 1342 Seongnamdaero, Sujeong-gu, Seongnam-si, Republic of Korea

E-mail : eopchang@gachon.ac.kr · Tel : +82-31-750-5493

Received : 2023/09/05 · Revised : 2023/10/24 · Accepted : 2023/10/25

© The Society of Pathology in Korean Medicine, The Physiological Society of Korean Medicine

pISSN 1738-7698 eISSN 2288-2529 http://dx.doi.org/10.15188/kjopp.2023.10.37.5.129

Available online at https://kmpath.jams.or.kr

학 분야 등에서 자주 사용되며 행렬을 특이벡터들로 구성된 행렬과 특이값을 원소로 갖는 행렬의 곱으로 나타내는 방법이다. 특이값의 크기 순으로 일정 수의 벡터만 선택함으로써 데이터의 차원을 줄일 수 있으며, 선택된 특이벡터 세트를 각 변수 집합으로 간주하고 정준상관분석을 통해 두 변수집합 간 상관구조를 가장 잘 설명할 수 있는 선형결합을 찾아서 상관관계를 평가한다.

본 연구에서는 각 의가별로 구성된 본초 x 처방 행렬의 표상 유사도를 비교함으로써 의가별 본초의 조합패턴이 얼마나 유사한지 정량적으로 분석하고자 하였다. 본초 x 처방 행렬의 특이값 분해를 통해 얻어지는 특이벡터는 처방들의 선형결합으로 구성되며, 특이값이 크다는 것은 전체 본초들을 해당 특이벡터로 투영 (projection) 했을 때의 분산이 크다는 것을 의미한다. 본초들은 처방에의 포함유무에 따라 0 또는 1로 인코딩 되었으므로 이 경우 처방 축들의 조합으로 만들어진 특이벡터들은 본초들을 최대한 잘 분리(shattering) 할 수 있는 (즉, 본초들의 포함여부에 대한 설명력이 높은) 축으로서 각 의가들이 처방을 운용하는 주요 패턴(예, 열성 처방, 소화기 처방 등)을 반영한다고 볼 수 있다. 의가별로 추출된 동일한 수의 특이벡터 집단에 계산된 정준상관계수는 의가들의 주요 처방 패턴들이 전반적으로 얼마나 유사한지를 나타낸다. 상관계수의 유의성을 판단하기 위하여 의가 별 처방 당 본초 수의 고유 분포를 보존하면서 처방의 본초 구성을 랜덤화 시킨 귀무 행렬(null-matrix) 간 유사도를 기준 값으로 사용하였다.

## 2. 금원사대가 처방네트워크의 string 구조와 매니폴드 차원 (manifold dimensionality) 분석

### 1) 처방 네트워크

질환계통별 각 의가들의 처방을 노드(node)로, 처방 간 코사인 유사도(cosine similarity)로 노드 간 연결(edge)을 정의하여 처방 네트워크를 구성하였다. 코사인 유사도에 대한 역치는 0.1부터 0.9까지의 구간에서 0.1 단위로 설정되었고, 노드 간 코사인 유사도가 설정된 역치값을 통과한 경우에만 연결을 형성하도록 하였다.

### 2) string 구조

대부분의 처방 네트워크에서 관찰된 ‘처방들의 연속적 변이’ 구조를 string 구조라 칭하였다. 전체 네트워크에서 연결(degree)이 2인 노드에서 시작하여, 좌-우로 부모 노드(parent node)를 제외하고 연결된 이웃 노드(neighbor node)가 하나(leaf node) 거나 3개 이상인 (추가 branch 형성) 노드가 발견되면 멈추는 방식으로 string 구조를 포착하도록 하고 closed loop 형태의 사이클은 제외하였다. 계통과 의가 별 string 구조를 비교할 때 처방의 수나 연결의 밀도(edge density)에 의한 영향을 배제하고자 1000번의 permutation test를 통해 구해진 string 수의 귀무 분포를 이용하여 string의 수를 Z-표준화하였다. Network permutation은 노드와 연결의 수를 보존한 상태에서 노드간 연결을 랜덤화 하는 방식으로 진행하였다.

### 3) 매니폴드 차원

매니폴드 가설(manifold hypothesis)<sup>9)</sup>은 고차원 공간에 주어진 데이터가 실제로는 고차원 공간에 내재한 훨씬 저차원의 매니폴드 인근에 집중되어 있음을 가정하며, 모든 차원축소 및 매니폴드 학습방법은 매니폴드 가설을 전제하고 있다. 의가와 질환계통별로 구성된 처방 x 본초 행렬은 전체 N개의 본초 각각이 하나의 축(차원)을 이루는 N차원의 본초공간에 대응되고, 이때 각 처방(본초 조합)은 본초 공간상의 한점으로 표현된다. 만약 본초들의 조합 원리가 무작위적이라면 처방은 N차원 본초공간에 전체에 걸쳐 균등하게 분포할 것이지만, 본초조합의 유의미한 패턴이 존재한다면 대다수의 처방은 저차원의 부분공간 (이하 본초 매니폴드라 칭함)에 놓여있을 것이다.

앞서 N개의 본초 각각이 개별 본초의 처방에의 포함여부를 나타내는 공간을 N차원 공간이라고 표현할 때의 차원개념은 주변차원(ambient dimensionality)으로 지칭된다. 주변차원은 가능한 모든 본초조합의 공간에 대응되며, 관측된 데이터의 실제 분포를 반영하지 않는다. 우리는 처방들이 N차원 보다 낮은 저차원의 본초 매니폴드에 놓여있다고 가정하여, 두가지 유형의 매니폴드 차원을 계산하였고, 이 둘을 비교함으로써 본초 매니폴드의 기하학적 특성을 도출하고자 하였다. 첫번째는 임베딩 차원(embedding

dimensionality)로, 주변 유클리드 공간(ambient Euclidean space)내에서 매니폴드가 걸쳐 있는 차원의 수를 측정하는 개념이다<sup>10)</sup>. 본 연구에서는 임베딩 차원을 정량화 하는 일반적 전략에 따라 데이터 행렬의 특이값 분해 이후 원 데이터의 분산을 95% 이상 설명하기 위해 필요한 최소한의 특이벡터 수로 계산하였다(Eq. 1). 두번째는 효과적 차원(effective dimensionality)으로 변수집합의 전반적인 상관구조(correlational structure)를 분산이 동일한(등방성) n개의 직교 차원(orthogonal dimension)으로 요약하는 개념이다. 기하학적 관점에서는 같은 부피의 매니폴드라 하더라도 데이터가 모든 축방향으로 동등하게 분포하는 경우 효과적 차원이 가장 높고, 각 축에 대한 분산이 비대칭적 일수록 낮아진다. 효과적 차원은 보편적으로 사용되는 아래 두 가지의 추정식을 이용하였다 (Eq. 2-3)<sup>11)</sup>.

$$EbdD = \min \left\{ j \mid \frac{\sum_{i=1}^j \lambda_i^2}{\sum_{i=1}^k \lambda_i^2} \geq 0.95 \right\} \quad (1)$$

$$EfdD_1 = \prod_{j=1}^k \frac{\lambda_j}{\sum_{i=1}^k \lambda_i} - \frac{\lambda_j}{\sum_{i=1}^k \lambda_i} \quad (2)$$

$$EfdD_2 = \frac{\left( \sum_{i=1}^k \lambda_i \right)^2}{\sum_{i=1}^k \lambda_i^2} \quad (3)$$

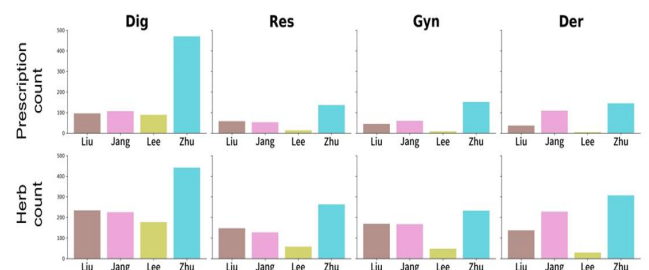
$\lambda$ 는 공분산 행렬의 특이값(eigenvalues)을, k는 데이터의 변수 수 의미한다.

## 결 과

### 1. 금원사대가 처방공간의 표상 유사도

수집된 처방수와 구성 본초의 가짓수는 의가와 질환 계통에 따라 다소간의 차이가 있었다. 의가로서는 주진형이, 질환 계통으로는 소화기계통의 처방 수 및 구성 본초 수가 많았으며, 이고는 소화기계를 제외한 타 질환군에서는 처방의 수가 특이적으로 적음을 알 수 있었다(Fig. 1).

(a)



(b)

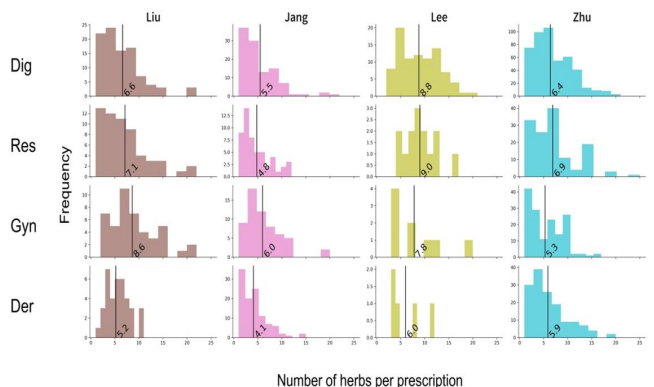


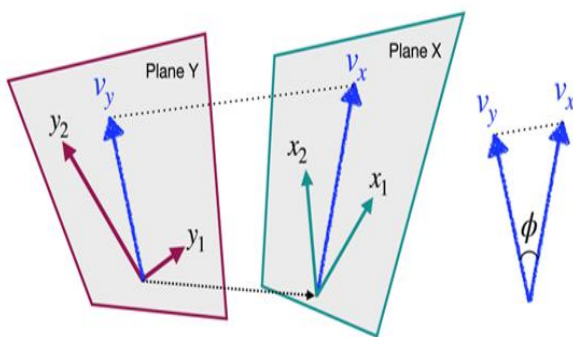
Fig. 1. Basic statistics of dataset. (a) Total number of collected prescriptions (top) and herbs (bottom) for each disease system and each scholars. (b) Histograms of the number of herbs per prescription. Abbreviations: Dig, Digestive system; Res, Respiratory system; Gyn, Gynecology; Der, Dermatology; Liu, 劉完素; Jang, 張從正; Lee, 李杲; Zhu, 朱震亨.

의가 간 처방공간의 표상유사도를 비교하기 위하여 각 의가별 본초 x 처방 데이터 행렬에 대하여 특이값 분해를 시행하고, 특이값이 큰 순서대로 20개의 특이벡터를 추출하여 내재적 부분공간(처방 매니폴드)을 정의하였다(explained variance = 0.37-0.45). 의가들의 추출된 20개 특이벡터의 선형 조합을 통해 서로 가장 유사한 벡터를 만들도록 하고 두 벡터간 상관계수를 계산하였을 때

(Fig. 2a), 귀무 행렬에 비해 의가 간 의미 있는 상관관계가 포착되었다(Fig. 2b). 흥미롭게도, 설명 분산의 기준을 높여 더 많은 수의 특이벡터로 공간을 근사함에 따라 귀무 행렬에서 계산된 상관계수는 증가하였고 실제 관측된 의가 간 상관계수는 귀무 값에 비해 낮게 나타났다. 저차원 공간에서의 유의미한 표상 유사도는 의가마다의 주요 처방 계통별로 본초를 운용하는 패턴(예, 열성계통 처방에는 본초군 A를 주로 활용하고, 소화기 계통 처방에는 본초군 B를 주로 활용하는 등)에서 어느 정도 의미 있게 공유되는 부분공간이 있음을 시사하며, 이후 특이벡터들을 추가함에 따라 표상 유사도가 귀무 값에 비해 낮아지는 것은 의가간 공유되지 않는 고유한 공간(적어도 선형변환으로 정렬(aligned)될 수 없는 공간) 역시 존재함을 알 수 있다.

단 Svcca는 공간 간 상관관계를 평가함에 있어 높은 자유도의 공간 변형을 허용함을 고려할 때, 추후 보다 제약된 방식의 상관관계 평가방식과의 비교가 필요할 것이다. 또한, 매니폴드의 비선형성이 크다면 Svcca방식의 차원축소가 많은 정보를 손실시켜 결과가 왜곡될 가능성이 있으므로 부분공간의 비선형성 정도에 대한 확인이 필요하다.

(a)



(b)

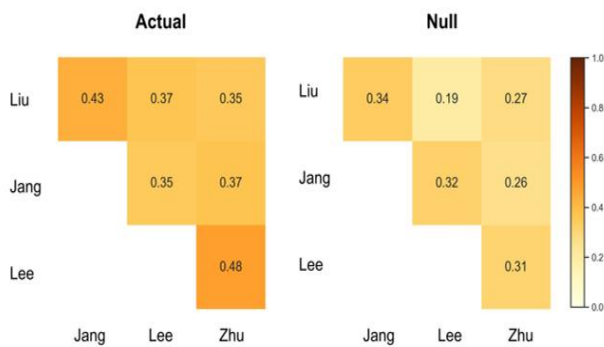
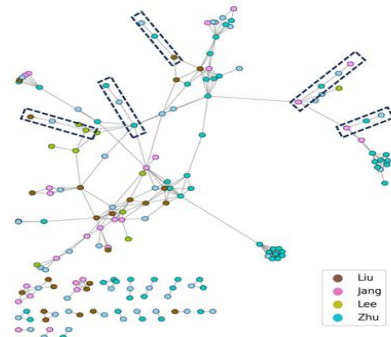


Fig. 2. Representational Similarity. (a) Schematic of canonical correlation analysis; the canonical correlation of two planes X,Y that intersect each other only at the origin is computed as the angle  $\phi$  between the pair of canonical variates  $v_x, v_y$ . Each canonical variable is found as a linear combination of vectors in the corresponding plane ( $x_i, y_i$ ) that minimize the angle between them. (b) Canonical correlation coefficients between the pairs of the scholars (left: value obtained from collected prescription data, right: average of results obtained from repeated random permutations). Abbreviations: Liu, 劉完素; Jang, 張從正; Lee, 李杲; Zhu, 朱震亨.

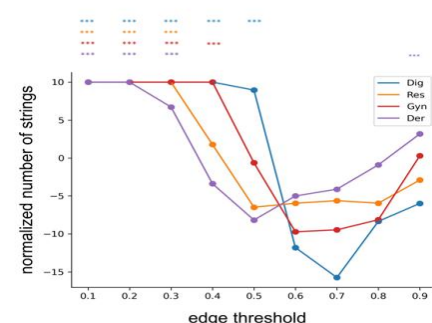
2. 금원사대가 처방네트워크의 기하학적 특징

각 의가와 질환계통별로 구성된 처방네트워크에서 관측된 처방들의 연속적인 변이 패턴을 string 구조라 칭하고(Fig. 3a, dashed box), 네트워크 별 string의 수를 계산하였다. string 구조는 부모 노드에서부터 처음 분기된 이후 분기되거나 수렴하지 않고 연속적으로 이어지는 변이라는 점에서 string 구조를 형성하는 처방들은 다른 처방들과는 구분되는 일정한 본초 조합의 패턴을 갖는다는 것을 시사한다. 모든 처방네트워크에서 코사인 유사도로 정의된 노드 간 연결 역치값이 0.3 이하에서는 유의미한 string 구조를 갖고 있음을 확인할 수 있었다(Fig. 3b). 질병계통에서는 소화기계통(Fig.3b, blue), 의가에서는 주진형의 처방네트워크가 특히 넓은 범위의 연결 역치값에서 일관되게 유의미한 string 구조를 보였고(Fig.3c, cyan), 이고의 경우 string 구조가 유의하지 않은 경우가 많았다(Fig. 3c, olive).

(a)



(b)



(c)

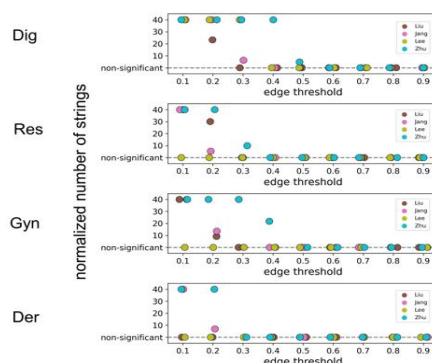


Fig. 3. String structure of prescription networks. (a) The network shown is a network of all scholar's prescriptions for the respiratory system with an edge threshold of 0.5. Some representative examples of detected string structures are marked with a dashed box. (b) Normalized number of strings for each disease system across different edge thresholds. When no strings were observed for the entire permutation network (and hence the z-value was infinite), we denoted the normalized number of strings as 10 and 40 in b and c, respectively, considering the maximum of the overall distribution. Three stars indicate a p-value below 0.001. (c) Normalized number of strings for each disease system and each scholar across different edge thresholds. If the p-value from the null distribution was less than 0.05, it was shown as 'non-significant'. Abbreviations: Dig, Digestive system; Res, Respiratory system; Gyn, Gynecology; Der, Dermatology; Liu, 劉完素; Jang, 張從正; Lee, 李杲; Zhu, 朱震亨.

이어서 우리는 특정방향으로의 처방 변이(variation)을 반영하는 string 구조가 본초 매니폴드의 모양(shape)과 어떠한 연관성을 갖는지 조사하였다. 먼저 우리는 처방데이터의 임베딩 차원과 효과적 차원의 차이를 통해서 본초 매니폴드의 기하학적 형태에 대한 특징을 포착하고자 하였다. 임베딩 차원과 효과적 차원 각각은 데이터 기하학에 대한 직접적인 표현이 아니지만, 이들의 차이가 클수록 매니폴드의 형태가 각 방향으로 고르지 않고 비대칭적으로 우그러져 있음을 시사한다(Fig. 4a). 두가지 차원에 대한 메트릭 모두 주진형에서 가장 높고, 이고에서 가장 낮게 측정되었으며(Fig. 4b), 이는 의가별로 수집된 처방과 본초의 수에 많은 영향을 받은 것으로 판단된다. 의가(4)와 질병계통(4)별로 구성된 16개의 처방 네트워크(edge threshold = 0.1)에서 효과적 차원과 임베딩 차원의 비율과 string 수 간의 피어슨 상관 계수는  $-0.425(EFD_1)$ ,  $-0.453(EFD_2)$ 으로 통계적으로 유의미하지는 않았지만 서로 반대방향의 상관관계를 보였다(Fig.4c). 이는 처방들의 string 구조가 본초 매니폴드의 찌그러진 정도에 유의미하게 기여한다는 것으로, 처방이 변이해 나가는 방향들(string들이 생기는 방향)이 고차원 공간에서 완전한 독립이 아님을 시사한다. 단, 본초 공간이 상당히 고차원임을 고려할 때 string이 변이해 나가는 방향들 간의 상관성이 우연에 의해서도 발생할 수 있는 수준 이상으로 의미 있는 정도인

지는 추가적 검증이 필요하다.

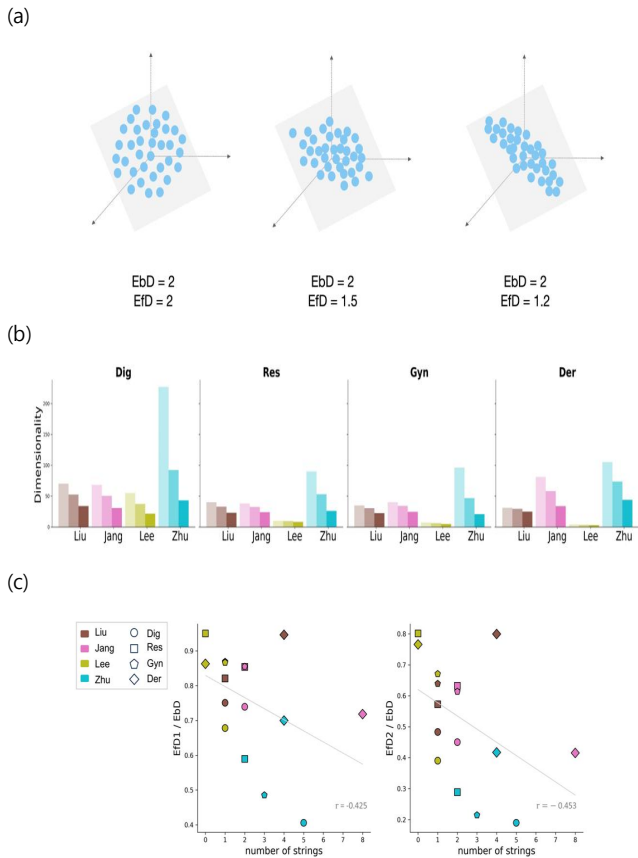


Fig. 4. Correlation between string structure and manifold geometry. (a) Since the points in the three-dimensional plane are actually embedded in a two-dimensional plane, their embedding dimension (EbD) is two. Meanwhile, the more asymmetric the variance of the data on each axis (from left to right), the lower the effective dimension (EfD), and thus the greater the difference between EbD and EfD. (The depicted data points and numerical values are not derived from actual data but are provided to offer a conceptual understanding of embedding and effective dimensionality). (b) EbD, EfD<sub>1</sub> and EfD<sub>2</sub> calculated from prescription data of each scholar and disease system. (In order of decreasing transparency, EbD, EfD<sub>1</sub> and EfD<sub>2</sub> are shown). (c) Correlation between string structure and manifold shape, which can be estimated from the difference between different dimensionality types. The resulting fitted linear regression lines and Pearson's correlation coefficients are shown. The color and shape of the markers represent doctors and disease systems, respectively. Abbreviations: Dig, Digestive system; Res, Respiratory system; Gyn, Gynecology; Der, Dermatology; Liu, 劉完素; Jang, 張從正; Lee, 李杲; Zhu, 朱震亨.

## 고찰

본 연구에서는 데이터의 기하학적 특성 분석을 중심으로 한의학 기초이론에 대한 새로운 접근방안을 제안하였다. 나아가 얻어진 분석결과를 바탕으로 우리는 향후의 구체적 연구방향을 도출할 수 있었다.

금원사대가 처방을 활용한 표상유사도 분석은 다음과 같이 확장될 수 있다. 각 처방(데이터 샘플)을 설명하는 변수를 본 연구에서와 같이 본초단위로 구성할 수도 있지만, 처방에 포함된 성분(compound), 표적 단백질(protein), 표적 경로(pathway), 혹은 한의학 성미 등 다양한 레벨로 설명 변수를 구성할 수 있을 것이다. 설명변수에 따라 동일한 처방데이터를 성분, 단백질, 경로, 성미 등 다양한 변수공간으로 맵핑 할 수 있으며, 처방들이 놓여진 매니폴드는 각각의 공간 특성에 따라 다른 형태를 취할 것이다. 이후 본초공간과 관심변수 공간의 매니폴드 구조간 맵핑관계를 조사함으로써 본초 조합의 패턴과 원리를 관심 변수 수준에서 탐구할 수 있을 것이다(Fig. 5). 이 같은 접근방식을 통해 처방경험에 내재된 풍부한 암묵적 지식으로부터 구체적이고 객관적인 생물/화학적 정보 수준에서 설명가능한 일반적 원리를 도출할 수 있을 것이다.

금원사대가 처방네트워크 분석결과를 처방들이 연속적으로 변이되는 구조(string)가 있으며, 이 변이는 본초매니폴드 상에서 특정한 방향성을 갖는 궤적(trajjectory)을 갖는다는 것을 시사한다. 그리고 이 궤적은 본초 매니폴드를 구성하는 잠재 변수(본초 조합 패턴)들의 조합과 그 변화에 의해 결정된다(Fig. 6). 분석하고자 하는 처방들을 기본 구성요소가 되는 개별 본초들에 대한 포함여부를

기반으로 수치화 할 수도 있지만, 동시에 본초들의 조합 패턴 수준에서 수치화 할 수도 있다. 상대적으로 개별 본초들은 처방데이터를 설명하는 저차원의 설명변수(low-level feature)이고, 개별 본초들의 조합은 고차원의 설명변수(high-level feature)라 할 수 있다. 비유하자면, 사람 얼굴의 이미지를 단순히 선이나 각도와 같은 저차원 변수를 이용하여 표현할 수도 있고, 이들이 특정한 패턴으로 조합된 눈, 코, 입 등의 해석가능한 수준의 고차원 변수로 나타낼 수 있는 것과 같다. 만약 인간의 진화 과정에서 얼굴 생김새가 어떻게 변화해왔는지를 설명하고자 한다면, 각도 등의 저차원 변수 보다는 눈, 코, 입 등의 고차원 변수 수준에서 이들의 조합 패턴이 어떻게 변화해왔는지를 설명하는 것이 보다 만족스러운 수준의 해석을 제공할 수 있을 것이다. 이처럼 처방들의 연속적 변이과정을 본초들의 조합패턴의 변이로 이해함으로써, 본초조합 원리가 시대적, 혹은 지역적으로 변화되어 온 패턴을 명시적으로 설명하고 시각화 할 수 있을 것이다.

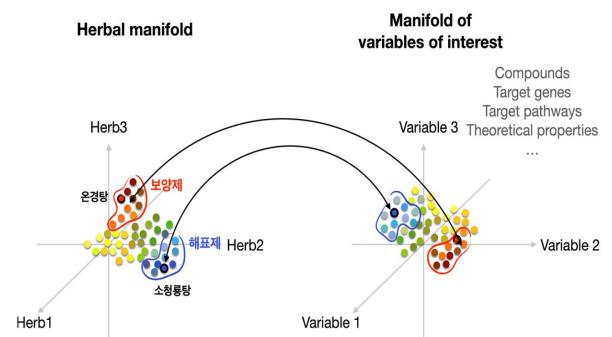


Fig. 5. Investigation of correlational correspondence between different variable spaces. It is possible to investigate how a particular subspace (e.g., 解表劑, 補陽劑) representing specific herbal combination is described in terms of other bioinformation (e.g., a particular combination patterns of compounds, gene target).

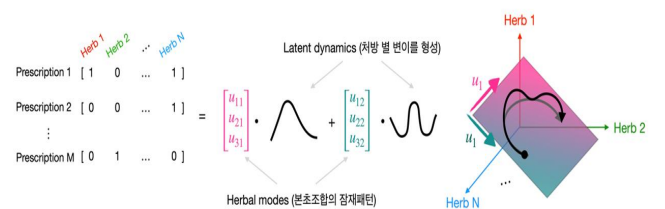


Fig. 6. Investigation of variation of prescriptions in terms of dynamics of herbal modes. M prescriptions can be represented in terms of herbal compositions, which in turn can be represented by a much smaller number of dominant herbal combination patterns (referred to in the figure as herbal modes), which can well describe the overall prescriptions. The successive evolutionary patterns of prescriptions can be described by the latent dynamics of the herbal modes, which form a trajectory on a subspace consisting of herbal modes (i.e., herbal manifold). Continuous variation of prescriptions can be observed and explained more clearly on a low-dimensional herbal manifold than on the high-dimensional original herb space.

## 결론

본 연구에서는 금원사대가 처방들의 본초구성을 기반으로 처방데이터의 기하학적 특성을 분석하였다. 각 의가들의 처방공간을 비교했을 때, 의가 상호간에 유의미한 수준으로 공유되는 부분공간과 선형변환에 의해 정렬되지 않는 고유의 부분공간이 모두 존재함을 확인할 수 있었다. 또한 각 의가의 전체 처방들간 연관관계를 네트워크 수준에서 관찰했을 때 처방들의 연속적 변이 패턴이 포착되었으며, 이는 본초 매니폴드 상에서 특정한 방향의 궤적을 생성함을 시사한다.

본 연구에서 도출된 탐색적 분석결과를 바탕으로, 향후 한의학 처방들에 내재된 본초조합의 원리를 다양한 변수(성미귀경, 성분, 표적단백, 표적경로 등)수준에서 탐구하기 위한 구체적 연구방향을 도출하였고, 처방들의 변이 패턴을 본초 조합 패턴들의 역학(즉 본초 조합 패턴들이 서로 어떻게 조합되고, 그 조합 패턴이 시간에 따라 어떻게 변화하는지)을 통해 설명할 수 있음을 제안하였다.

본 연구에서는 수학적 분석방법론에 기반하여 한의학 기초이론을 새로운 방식으로 해석하고 정량적 형태의 지식을 도출할 수 있음을 보이고자 하였다. 이 같은 접근방식이 기초한의학에서 갖는

가치는 다음 두가지 측면으로 요약될 수 있다. 첫째, 객관적 지식의 도출은 한의학 이론의 주관적 측면을 효과적으로 보완함으로써 현대화에 기여할 수 있다. 둘째, 구조적이고 수학적인 사고의 틀을 제공함으로써 학생들로 하여금 한의학 이론을 더 깊고 체계적으로 이해할 수 있는 기회를 제공할 수 있다.

## 감사의 글

본 연구는 과학기술정보통신부의 재원으로 한국연구재단(No. 2022R1F1A1068841)과 한국한의학연구원 주요사업 AI 한의사 개발을 위한 임상 빅데이터 수집 및 서비스 플랫폼 구축(KSN2021110)의 지원을 받아 수행되었으며, 이에 감사드립니다.

## References

1. Park M, Kim YM. Comparative Study on Etiological Cause, Pathogenesis Mechanism of 'Shanghanlun' and 'Wenbingtiaobian'. *J Physiol & Pathol Korean Med.* 2013;27(1):1-10.
2. Choi HY, Kim KJ. Basic Principles of the 'Spleen-stomach theory' by Li Dong-yuan. *J Physiol & Pathol Korean Med.* 2010;24(6):911-20.
3. Kim K-W, Park H-G. Comparative study on the theory of 'Ministerial-fire' between Zhu Dan-Xi and Jang Gyung-Ak. *J Korean Med Classics.* 2006;19(4):12-26.
4. Jang DY, Oh KC, Jung ES, et al. Diversity of acupuncture point selections according to the acupuncture styles and their relations to theoretical elements in traditional asian medicine: a data-mining-based literature study. *Journal of Clinical Medicine.* 2021;10(10):2059.
5. Oh J. Can Similarities in Medical thought be Quantified? *J Korean Med Classics.* 2018;31(2):71-82.
6. Kim KW, Kim TY, Lee BW. Analysis of Prescriptions from Taepyeonghyeminhwajegukbang, Somunmyungronbang and Nansilbijang based on Herb weight ratio grade. *J Korean Med Classics.* 2014;27(4):73-84
7. Newman I, Benz CR. Qualitative-quantitative research methodology: Exploring the interactive continuum. SIU Press: 1998.
8. Raghu M, Gilmer J, Yosinski J, Sohl-Dickstein J. Svcca: Singular vector canonical correlation analysis for deep learning dynamics and interpretability. *Advances in neural information processing systems.* 2017;30.
9. Fefferman C, Mitter S, Narayanan H. Testing the manifold hypothesis. *Journal of the American Mathematical Society.* 2016;29(4):983-1049.
10. Jazayeri M, Ostojic S. Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. *Current opinion in neurobiology.* 2021;70:113-20.
11. Del Giudice M. Effective dimensionality: A tutorial. *Multivariate behavioral research.* 2021;56(3):527-42.