

한약재 사용량 데이터 분석을 통한 한국, 중국, 일본 전통의학의 이론적 특성 비교연구

박무순 · 이충열 · 이태희¹ · 김연섭^{2*} · 김창업*

가천대학교 한의과대학 생리학교실, 1: 방제학교실, 2: 해부경혈학교실

Identifying Theoretical Characteristics of Traditional Medicines in Korea, China, and Japan through the Herb Usage Data

Mu Sun Park, Choong Yeol Lee, Tae Hee Lee¹, Youn Sub Kim^{2*}, Chang Eop Kim*

Department of Physiology, 1: Department of Formulae Pharmacology,
2: Department of Anatomy-Pointology, College of Korean Medicine, Gachon University

Traditional medicines (TM) in Korea, China, and Japan share most of the theories and therapeutic tools, but there are also differences due to their unique histories and cultures. Here, we aim to identify the differences in the utilization of TM theory between three countries by analyzing herb usage data in terms of the related traditional theories. Herb usage data of each country was collected from "Investigation of Korean medicine use and herbal medicine consumption survey" (Korea), "Analytical report on circulation of key Chinese medicinal materials" (China), and "Survey report on raw material crude drug usage" (Japan). Fifty five herbs with sixty features belonging to five theoretical categories (four properties, five tastes, targeting meridians, treatment strategies, and herbal parts) were selected and analyzed. Weight Sum Model (WSM) and Network-Based Group Features (NBGF) were used to compare the theoretical characteristics of TM between three countries. For the statistical evaluation, we developed and applied Herb Set Enrichment Analysis (HSEA) for WSM and NBGF results. HSEA for WSM results revealed the kidney meridian were targeted more in Korea than Japan, while the spleen meridian were targeted more in Japan than Korea. Herbs with sour taste were used more in Japan than China. HSEA for NBGF results found that NBGF including warm, neutral, sweet, and tonifying features were more dominant in Korea and than Japan, while NBGF including cold, bitter, heat-clearing features were more dominant in Japan than the others. These results suggest that TM in Korea, China, and Japan have unique aspects of practice patterns and theoretical utilization.

keywords : Traditional medicine, Herb usage, Theoretical utilization, Data-driven

서 론

동아시아 3국인 한국, 중국, 일본의 전통의학은 모두 내경, 상한의학의 전통에 기초하고 있고, 대부분의 의학이론과 임상 치료 수단을 공유하고 있다. 그러나 세 나라의 지리와 기후, 문화적 환경과 전통의학의 역사적 발전 과정, 의료제도가 달라 각국의 전통 의학 임상현상은 차이가 존재할 것으로 기대된다¹⁾. 특히 전통의학의 임상적용이 단순히 과거의 경험에만 의존하는 것이 아닌, 나름의 고유한 이론체계를 기반으로 하고 있는 것이라면 세 나라에서 나타나는 임상적 특성의 차이는 이를 뒷받침하는 이론적 운용의 차

이로서 설명될 수 있을 것이다.

이와 관련하여 조²⁾는 전통의학에서 나타나는 사고의 차이, 치료체계, 병태인식, 진단체계 및 변증체계, 병리관 등을 포함하여 다방면에 걸쳐 한국, 중국, 일본의 전통 의학을 비교하는 연구를 진행한 바가 있다. 조는 세 나라의 전통의학이 내경의학과 상한의학 중 어떤 체계를 취하고 있는지에 따라 사고와 음양, 허실 등 용어 개념에서 차이가 있으며, 나아가 한국과 중국은 주로 병인병기에 기초한 변증논치를, 일본은 증과 처방을 직접적으로 연결하는 '方劑辨證'을 중시하는 특징이 있다고 분석하였다. 또 한국은 후세방과 사상의학, 중국은 온병학과 변증체계에 대한 인식이 독특하고, 일

* Corresponding author

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

E-mail : ysk5708@hanmail.net Tel : +82-31-750-5420

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 : 2018/04/30 Revised : 2018/06/27 Accepted : 2018/06/29

© 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.2018.06.32.3.149>

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

본에서는 체력변증과 복진법이 있는 것이 독특한 점이라고 하였다. 그러나 조의 연구결과는 문헌들의 정성적 분석을 바탕으로 이루어졌으며, 실제 임상현장을 반영하는 데이터의 정량적 분석에 근거하고 있지는 않다는 한계점이 있다. 우리가 아는 한, 지금까지 한, 중, 일 전통의학의 이론적 측면의 차이를 임상현장의 정량적 데이터를 기반으로 분석한 연구는 없었다.

본 연구에서 저자들은 데이터 주도형 접근법(data-driven approach)을 적용하여 한, 중, 일 전통의학의 임상적 활용을 뒷받침하는 이론적 특성의 차이를 객관적으로 확인하고자 하였다. 세 나라의 임상현장을 반영하는 데이터로서 각 국가의 한약재 사용량 데이터를 확보하였으며, 이로부터 전통의학의 이론적 특성의 차이를 추론하기 위하여 약물별로 부여된 이론적 특성들(features) - 시기, 오미, 귀경, 치료 속성 등을 추출하였다. 세 나라의 이론적 차이를 분석하기 위하여 한약재 사용량을 가중치로 설정하여 이론적 특성의 가중치 합 모델(Weighted Sum Model, WSM)을 적용하였으며, 보다 통합적인 이론적 의미를 추론하기 위하여 네트워크 이론을 기반으로 한약재의 특성들을 연관성에 따라 그룹으로 묶은 네트워크 기반 특성 그룹(Network-Based Group Features, NBGF)을 적용하였다. 또한 생물정보학 분야에서 사용하는 GSEA (Gene Set Enrichment Analysis) 방법론을 변형한 HSEA (Herb Set Enrichment Analysis)를 수행하여 통계적 유의성을 확인하였다 (Fig. 1).

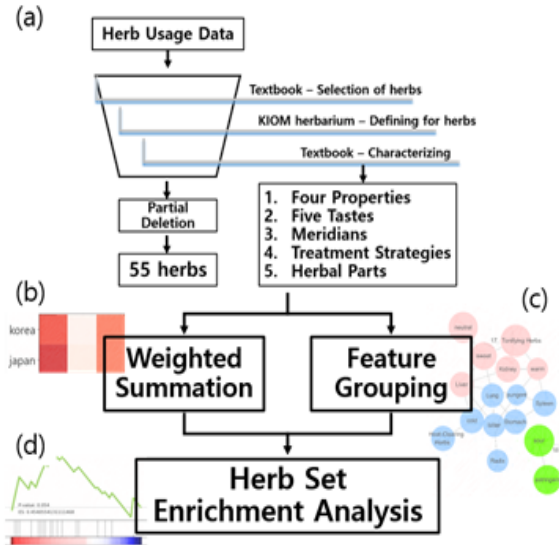


Fig. 1. Workflow of this study. Herb usage data of each country was collected and preprocessed. Fifty five herbs were finally selected for our analysis. Features characterizing the herbs were also collected (a). Weighted sum model and network-based group features were used to compare the theoretical characteristics of TM between three countries (b and c). Finally, herb set enrichment analysis was applied to the results to evaluate the statistical significance (d).

연구 방법

1. 한약재 선정 및 한약재 사용량 데이터 전처리

한국, 중국, 일본의 임상 특징 분석을 위하여 각 국가 기관에

서 발표한 한약재 사용량 데이터를 수집하였다. 한국은 2014년에 보건산업진흥원에서 발표한 <한방의료이용 및 한약소비실태조사>³⁾, 일본은 일본한방생약제협회에서 발간한 <原料生藥 使用量等調査報告書>⁴⁾, 중국은 중화인민공화국 상무부 시장질서사에서 발간한 <中藥材 重點品種 流通分析報告>⁵⁾의 자료를 기초로 하여 일본 285종, 한국 59종, 중국 35종의 한약재 사용량 데이터를 확보하였다.

일본은 2009년부터 2015년, 한국은 2011년부터 2013년, 중국은 2011년부터 2015년까지의 데이터를 포함하고 있었으며, 중국의 경우 2014년까지는 29종의 한약재 사용량 정보를 포함하고 있었으나 2015년에 6종의 한약재 정보가 추가되었다. 본 연구에서는 한, 중, 일 세 나라에서 공통으로 조사된 35종의 한약재를 분석 대상으로 하였으며, 2015년 추가된 한약재의 경우 2015년의 사용량으로 2011-2014년의 결측값으로 대체하였다. 한약재 별 사용량은 각 나라별 총 사용량을 기준으로 하여 상대적 사용량으로 변환되었다.

한약재 중 한국 한의과대학에서 공통 교재로 사용되는 본초학 교과서(한의과대학 본초학 편찬위원회, 영림사, 2008)에 기재되어 있지 않은 한약재는 분석에서 제외하였다. 예를 들어, 은행엽은 한국의 한약재 사용량 데이터에는 포함되어 있으나 본초학 교과서에는 수록되어 있지 않은 약물이므로 제외하였다. 포제 방법이 다르나 기원과 약용부위가 같은 경우 하나의 한약재로 취급하였다. 예를 들어 생지황과 숙지황은 *Rehmanniae Radix Preparata*로 기원과 약용부위가 같다. 포제법이 다르지만, 본 연구에서는 같은 한약재로 취급하였다. 마지막으로, 당귀의 경우 한국, 중국, 일본에서 사용하는 당귀의 기원이 다르지만, 과(family)가 동일하며, 세 나라가 같은 이론적 배경을 가지고 진료에서 사용하기 때문에 같은 한약재로 취급하였다.

최종적으로 두 나라 이상에서 공통으로 포함되어 있는 총 55종의 한약재를 분석 대상으로 선정하였다(Table 1, Fig. 2). 한약재의 평균 상대적 사용량 데이터가 두 나라에만 있는 경우에는 평균 대치법을 사용하여 두 나라 데이터의 평균값을 사용량이 누락된 나라의 사용량으로 삼아, 결측값으로 인해 생길 수 있는 통계적 편향을 보정하였다⁶⁾.



Fig. 2. Selection of herbs. Only the herbs were chosen whose data is available in at least two of three countries. The shaded parts represent the number of selecting herbs in each country.

2. 특성(feature) 선정

한약재의 특성은 본초학 교과서에 명시된 시기, 오미, 귀경, 치

료 속성, 약용부위에서 3종 이상의 한약재에 포함되는 것을 선정하였다. 그 결과 사기 5종, 오미 7종, 귀경 11종, 치료 속성 20종, 약용부위 17종이 특성으로 선정되었다(Table 1).

Table 1. Selected herbs and features

선정된 한약재 목록	선정된 특성 목록
감초, 백작약, 복령, 육계, 당귀, 반하, 인삼, 마황, 시호, 천궁, 갈근, 백출, 대황, 지황, 택사, 의인, 황금, 석고, 황기, 목단피, 길경, 맥문동, 후박, 방풍, 산약, 산수유, 연교, 부자, 오미자, 우슬, 자소엽, 활련, 현호색, 향부자, 백지, 사인, 용안유, 홍화, 독활, 강활, 구기자, 두충, 애엽, 천마, 위령선, 금은화, 단삼, 하수오, 당삼, 산사, 녹용, 계지, 사삼, 우황, 오가피	사기 寒性, 凉性, 平性, 温性, 熱性 오미 酸味, 苦味, 甘味, 辛味, 鹹味, 滋味, 淡味 귀경 肺經, 大腸經, 胃經, 脾經, 心經, 小腸經, 膀胱經, 腎經, 心包經, 膽經, 肝經 치료 속성 解表藥, 清熱藥, 사하藥, 거풍습藥, 방항 화습藥, 이수삼습藥, 온리약, 이기약, 소식약, 구충약, 지혈약, 활혈거어약, 화담 지해명천약, 안신약, 평간약, 계규약, 보익약, 수삼약, 용토약, 외용약 약용부위 동물성약재, 구근식물, 경부, 피질, 꽃, 잎, 과실, 초부, 목질, 광물성약재, 과피, 뿌리, 가지, 수지, 뿌리줄기, 종자, 암술

Fifty five herbs were selected after preprocessing. Only herbs were chosen whose data is available in at least two of three countries. Sixty features were selected from 5 categories.

3. 가중치 합 모델(WSM)

각 나라의 특성 별 한약재 사용량은 WSM을 사용하여 비교하였다. WSM은 한약재 특성과 한약재 사용량 정보를 이용하여 구하였다. WSM은 각각의 한약재마다 한약재가 가지고 있는 특성을 1, 가지고 있지 않은 특성을 0으로 처리한 이진 데이터(binary data)에 전처리 과정을 거친 한약재의 평균 상대적 사용량을 곱하여 특성 별 가중치 값을 구한다. 위 과정을 통해 구한 모든 한약재의 특성 별 가중치를 모두 합하여 한약재의 특성 별 WSM으로 정의했다. 세 나라의 WSM은 python의 matplotlib library를 이용하여 히트맵으로 시각화하였다.

$$WSM\ score_i = \sum_{j=1}^n w_j a_{ij}$$

w_j = 한약재 j의 평균 사용량,
 a_{ij} = 한약재 j에서의 특성 i의 유무(1 또는 0)

4. 한약재 특성 네트워크 구성(feature network)과 네트워크 기반 특성 그룹(NBGF) 정의

한약재 특성 네트워크는 한약재의 단일 특성을 노드로, 특성 간의 유사 정도를 연결로 정의하여 구성하였다. 노드 선정 시 분석에 사용된 한약재에 한 번도 등장하지 않는 특성은 노드에서 제외하였다. 특성 간의 유사 정도는 본초학 교과서를 기반으로 계산한 Jaccard 계수(Jaccard similarity coefficient)⁷⁾의 값으로 정의하였다. 네트워크의 최종 연결은 모든 Jaccard 계수를 가진 연결 중에서 Jaccard 계수 값이 상위 2% 안에 드는 것을 선정하였고 네트워크 구성 후 연결이 없는 분리된 노드는 제거하였다.

$$Jaccard\ similarity\ coefficient = \frac{A \cap B}{A \cup B}$$

A, B = 각 특성을 포함하는 본초들의 집합

한국, 중국, 일본의 한약재 상대적 사용량을 네트워크로 시각화하여 비교했다. 상대적 사용량은 한약재 특성의 WSM을 세 나라의 특성 WSM의 평균으로 나누어 세 나라의 평균 사용량과 비교한 상대적 사용 정도를 계산하고, 로그화하여 계산값을 보정하였다.

네트워크의 시각화는 Cytoscape 프로그램⁸⁾을 이용하였고, 보정한 상대적 사용량은 네트워크 노드의 크기에 반영하여 하여 국가 간의 상대적 사용량 차이를 볼 수 있게 하였다.

네트워크의 하위 모듈은 unweighted Louvain algorithm⁹⁾를 이용해 구성하였고, python community library를 사용하여 알고리즘을 구현하였다. 각각의 모듈에 속한 노드들은 NBGF로 새롭게 정의하였다.

5. Herb Set Enrichment Analysis (HSEA)

WSM과 네트워크를 이용한 탐색적 분석의 결과는 기존의 GSEA 방법론¹⁰⁾을 본 연구에 맞게 변형한 HSEA로 분석하여 평가하였다. GSEA는 특정 질병에서의 유전자 발현량과 실험을 통해 이미 알려진 유전자 집합의 차이를 비교, 분석하는 방법이다. GSEA에서는 질병군과 대조군의 microarray 발현량 차이를 실험을 통해 확인하고, 발현량의 차이를 유전자 주석 데이터베이스(KEGG pathway¹¹⁾, Gene Ontology¹²⁾ 등)의 정보를 이용하여 분석한다.

GSEA는 질병군과 대조군의 microarray의 발현량 상관계수를 가중치로 하여 분석한다. 본 연구에서는 발현량의 상관계수를 HSEA 분석을 진행할 두 나라의 한약재 사용량 차이로 수정하여 분석을 진행하였다.

또한, 유전자 주석 데이터베이스의 유전자 주석 집합은 본 연구에 맞게 한약재 집합으로 수정하였다. WSM 분석에서는 비교하고자 하는 특성을 가지고 있는 한약재를 집합으로 구성하였다. 네트워크 분석에서는 NBGF로 정의된 특성들의 집합과 단일 한약재가 가지고 있는 특성 간의 Jaccard 계수를 계산하고, 계수의 값이 순열 검정(permutation test)을 통과하는 한약재로 집합을 구성했다. 너무 적은 한약재 개수를 포함하는 한약재 집합은 편향된 Enrichment Score (ES) 값을 가질 수 있기 때문에 한약재 집합에 3개 이상의 한약재가 포함된 것만 분석을 진행하였다.

시각화는 python 프로그램으로 진행하였으며, 그림의 상단에는 ES의 변동 누적 그래프를, 2번째 단계는 한약재 집합에 선정된 한약재를 실선으로 표시하였다. 3번째 단계는 두 나라의 사용량 차이를 색상 지도로 나타내었고, 마지막 단계는 사용량 차이를 히스토그램 그래프로 나타내었다.

6. 통계

본 연구의 통계적 분석은 순열 검정을 이용하였다. WSM 분석에서는 ES의 값이 통계적으로 유의함이 있는 지 알아보았다. 귀무분포(null distribution)는 전체 한약재에서 한약재 집합의 개수만큼 무작위로 한약재를 선정한 후, 무작위 한약재로 ES를 계산하여 구성하였다. 한약재의 ES가 양수인 경우는 10000번의 검정 중 양수의 ES값을 가지는 것만, 반대로 ES가 음수인 경우에는 음수인 값만 모아 one-tailed test를 진행하였다.

NBGF의 한약재 집합에 포함되는 한약재는 NBGF의 특성과 한약재의 특성 간의 Jaccard 계수를 계산하고, 한약재 특성을 무작위로 선정하여 만든 10000번의 순열 검정을 통해 p-value<0.005인 것을 선정했다. 또한, NBGF의 한약재 집합은, WSM 분석과 같은 방법으로 귀무분포를 구성하고 one-tailed test를 진행하여

p-value를 계산하였다.

결 과

1. WSM을 이용한 세 나라의 임상 특성 조사

Fig. 3은 한국, 일본, 중국의 특성 별 한약재 사용량을 WSM을 이용하여 분석한 결과이다. 각 특성은 사기(Fig. 3a), 오미(Fig. 3b), 귀경(Fig. 3c), 약용부위(Fig. 3d), 치료 속성(Fig. 3e)을 WSM으로 분석했으며, 가중치 합이 높을수록 진한 색으로 시각화했다.

사기 그룹의 분석에서는 세 나라가 모두 溫性의 WSM이 높게 나타났으나, 일본의 경우 한국, 중국에 비해 溫性 약물의 사용량이 적었다. 반대로 寒性 약물은 일본의 사용량이 다른 두 나라보다 높은 것을 확인할 수 있었다.

오미 그룹 분석에서는 세 나라 모두 甘味의 WSM이 높은 것을 보여주고 있으며, 그 중에서도 중국의 WSM 값이 제일 높았다. 반면, 苦味 한약재와 辛味 한약재의 WSM은 한국과 일본이 중국보다 높은 것으로 나타났다.

귀경의 분석 결과, 세 나라 모두 脾經의 WSM이 가장 높았으며, 그 중에서도 일본의 WSM이 제일 높은 것으로 나타났다. 한국의 경우 腎經과 肝經의 한약재 WSM이 다른 두 나라에 비해 높았으며, 일본에서는 胃經과 脾經, 心經, 膀胱經 한약재의 WSM이, 중국에서는 大腸經 한약재 WSM이 높은 것을 보여줘 다른 나라보다 사용량이 많음을 알 수 있었다.

약용부위 분석에서는 세 나라가 모두 뿌리(Radix)를 압도적으로 많이 사용하였으며, 특히 중국에서는 뿌리의 WSM이 중국 내 전체 사용량의 절반 이상을 차지하였다. 다른 나라와 비교하여 한국에서는 뿌리줄기(Rhizoma)의 WSM이 많았으며, 일본에서는 피

질(Cortex)의 WSM이 제일 많았다.

마지막으로 치료 속성은 세 나라가 모두 보익하는 약물의 WSM이 높은 것으로 나타났으며, 특히 중국의 보익약 WSM이 다른 나라보다 높은 것으로 나타났다. 중국에서 해표약 WSM은 다른 나라보다 낮았으나, 청열약의 사용량은 한, 일보다 높은 것으로 나타났다. 또한, 중국에서는 활혈거어약 WSM이, 일본에서는 이수삼습약 WSM과 온리약 WSM이 다른 나라와 비교했을 때 제일 높은 사용량을 가지는 것으로 확인되었다.

2. WSM 분석에서의 HSEA 분석 결과

단일 특성에 대한 한국, 일본, 중국의 WSM 분석 결과가 유의한 차이가 있는 지 알아보기 위해 HSEA 분석을 시행하였다(Fig. 4).

단일 특성의 HSEA 분석은 기준으로 정한 나라와 비교할 나라의 한약재 사용량 차이를 이용한다. 한약재 사용량 차이가 가장 큰 것부터 시작하여 사용량 차이가 점점 적은 것으로 한 단계씩 내려가면서 분석하고자 하는 사기, 오미 등의 특성을 가지고 있는 한약재가 등장하면 해당 한약재의 사용량 차이의 절댓값을 ES의 점수에 더한다. 반대로 등장한 한약재에 분석하고자 하는 특성이 나타나지 않으면 기존에 설정해 놓았던 값만큼 ES의 점수를 줄여 나간다. 위 과정에서 가장 높게 올라가거나 가장 낮게 내려간 값들 중에서 절댓값의 크기가 큰 것을 최종 ES로 정의한다. 최종 ES가 양수이면 기준으로 정한 나라의 특성 사용량이 많은 것이고, 음수이면 기준으로 정한 나라의 사용량이 적은 것으로 해석할 수 있다.

분석 결과, 한국과 일본 간의 HSEA 분석에서 腎經 특성의 한약재는 한국에서 많이 사용하는 것(p=0.047)으로, 脾經의 한약재는 일본에서 많이 사용하는 것(p=0.031)으로 조사되었다(Fig. 4a). 한

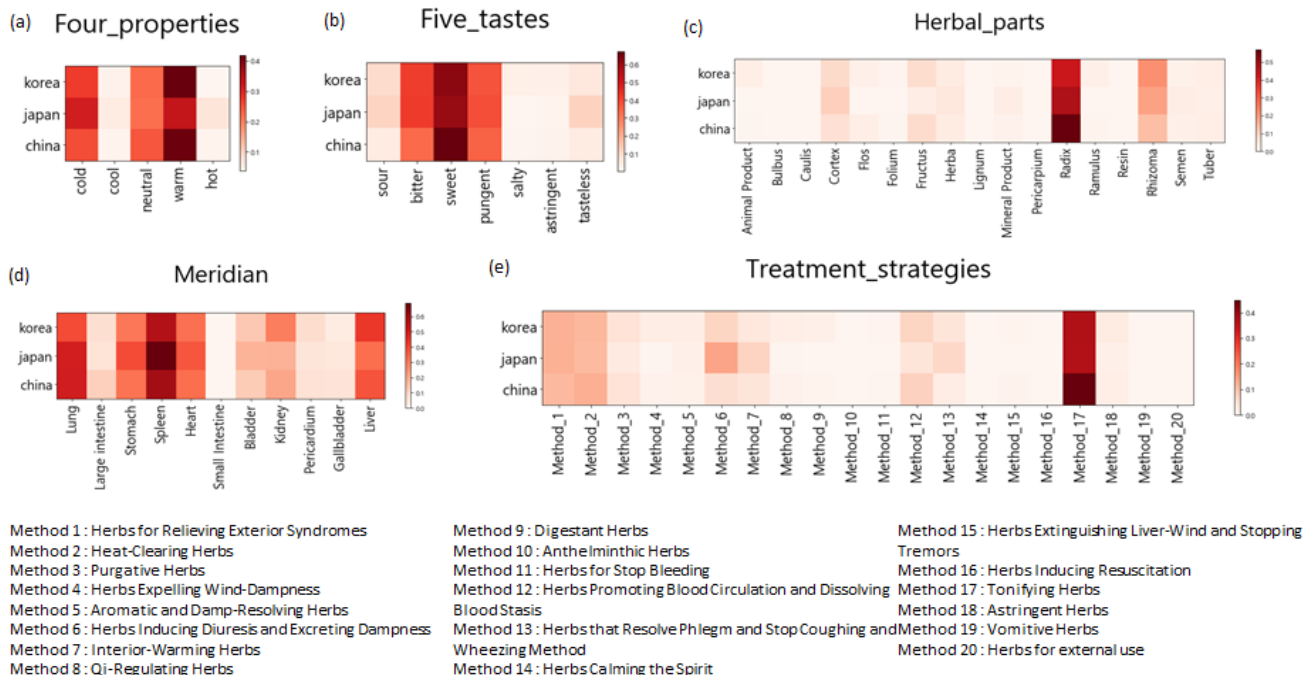


Fig. 3. Clinical characteristics of three countries using WSM. Features characterizing herbal medicine of each country were evaluated from the herbal usage data and compared each other (a, b, c, d, and e). Analysis of patterns of herbal features. The usage of characteristics including warm property (a), sweet taste (b), radix part (c), spleen meridian (d), and tonifying herb (c) is the highest in each country.

국과 중국 간의 HSEA 분석에서(Fig.4b) 뿌리 약용부위는 중국의 사용량($p=0.023$)이, 반대로 뿌리줄기 약용부위는 한국의 사용량이 많은 것($p=0.003$)을 확인할 수 있었다. 마지막으로, 일본과 중국 간의 HSEA 분석(Fig. 4c)에서는 酸味 특성의 한약재가 중국보다 일본에서 더 많이 사용하는 것을 발견하였다($p=0.013$).

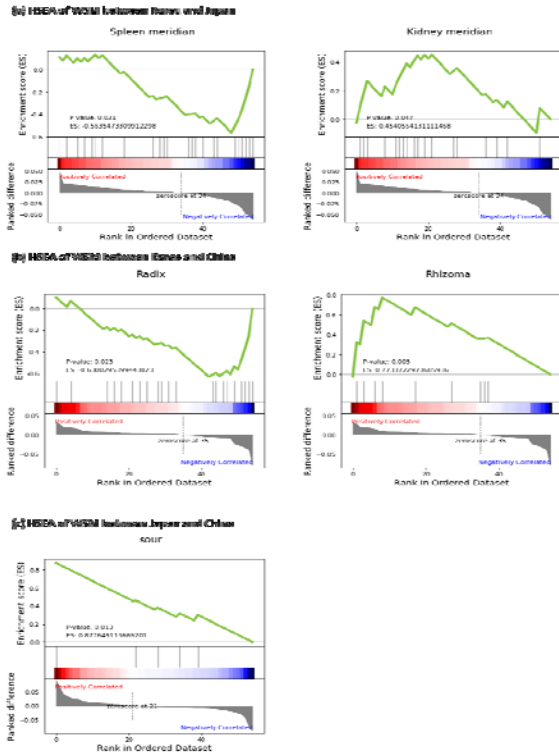


Fig. 4. HSEA result of WSM analysis. HSEA was performed for WSM results. Only the features with significant trends ($p < 0.05$) were shown. Positive (Negative) ES indicated the dominance of the feature in prior (posterior) country.

3. Network-Based group features (NBGF)를 이용한 세 나라의 임상 특성 분석

사기, 오미, 귀경, 치료 속성, 약용부위 범주에 속하는 한약재의 각 특성들은 독립적이지 않으며 한의학의 한열, 장부, 경락 등 이론에 의해 특정 특성들이 서로 연관성을 가질 것으로 추정된다. 그러므로 본 연구에서는 네트워크 방법을 사용하여 한약재의 특성들을 그룹화하여 이를 이용하여 임상 특성을 분석해 보았다. 그룹 분석은 연관성이 높은 특성들을 하나의 변수로 묶어 보여줄 수 있으므로 개별 특성으로 분석할 때와는 또 다른 차원의 임상 특성에 대한 분석을 제공할 수 있을 것으로 생각했다.

본 연구에서는 NBGF의 개념을 도입, 연관 정도가 높은 한약재의 특성들을 그룹화하여 분석하였다. 특성과 특성 간의 연관 정도를 이용하여 네트워크를 구축한 결과(Fig. 5) 총 6개 그룹으로 나누어졌다(Table 2).

한국은 鹹味-동물성 약재 그룹(NBGF 5)의 크기가 다른 나라에 비해 크며, 溫性-平性-甘味-보익약 그룹(NBGF 2)과 酸味-澁味-수삼약 그룹(NBGF 4)이 일본보다 사용량이 많은 것으로 분석되었다. 일본에서는 寒性-苦味-辛味-청열약 그룹(NBGF 1), 熱性-온리

약 그룹(NBGF 3), 淡味-이수삼습약 그룹(NBGF 6)의 사용량이 타 국가에 비해 많은 것을 확인할 수 있었다. 반면, 중국에서는 가장 많이 사용되는 NBGF 그룹이 없었다. 다만, 중국에서의 溫性-平性-甘味-보익약 그룹(NBGF 2)의 사용량이 한국과 비슷했으며 일본보다는 많은 것을 확인할 수 있었다(Fig. 5).

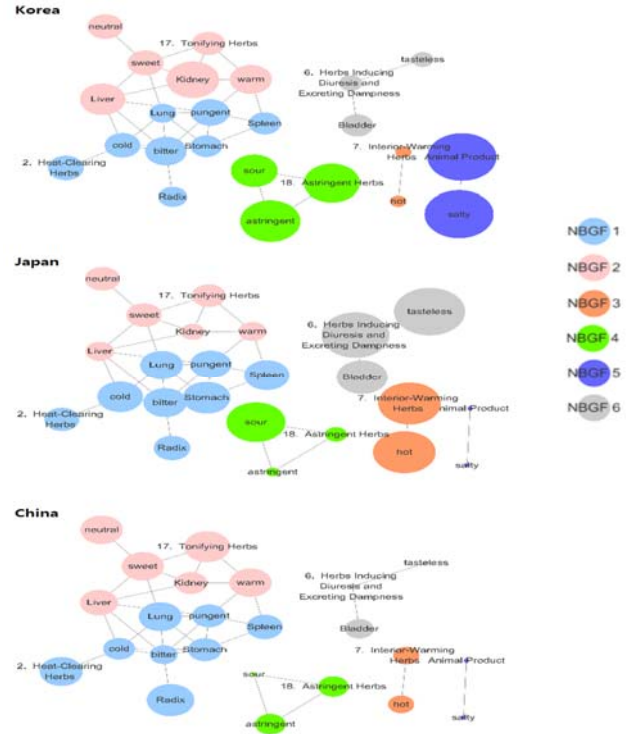


Fig. 5. Clinical practice pattern analysis of three countries using NBGF. Networks of features constructed from the co-occurrence patterns in the herb description. Feature modules were divided into six groups. The color of each node indicates the module (NBGF) they belong to. The size of the each node shows the relative usage of associated herbs with the features compared to the average of the three countries.

Table 2. Grouping features through Jaccard similarity coefficient analysis

NBGF 1	Cold property, Bitter taste, Pungent taste, Lung meridian, Stomach Meridian, Spleen Meridian, Heat-Clearing herbs, Radix
NBGF 2	Warm property, Neutral property, Sweet taste, Kidney meridian, Liver meridian, Tonifying herbs
NBGF 3	Hot property, Interior-warming herbs
NBGF 4	Sour taste, Astringent taste, Astringent herbs
NBGF 5	Salty taste, Animal product
NBGF 6	Tasteless, Bladder meridian, Herbs inducing diuresis and excreting dampness

Features are divided in six groups on the basis of Jaccard similarity coefficient analysis of herbal medicine textbook. Their connections define the two percent of edge density, which is the upper two percent of all possible connected Jaccard similarity scores. The first group which is the biggest contains 8 features, and the second group has six features. The other groups consist of 2 or 3 features.

4. 네트워크 분석에서의 HSEA 분석 결과

네트워크 분석에서 나타난 한국, 일본, 중국의 한약재 사용량 차이가 유의성이 있는지 확인하기 위해 HSEA 분석을 실시하였다(Fig. 6).

네트워크 분석의 HSEA 분석은 기준이 되는 나라와 비교할 나라의 한약재 사용량 차이를 사용하였으며, 한약재 집합에 들어가는

한약재는 Jaccard 계수를 이용하여 선정하였다. WSM의 HSEA 분석과 같은 방법으로 ES를 계산하며, ES가 가장 높게 올라가거나 가장 낮게 내려간 값 중 절댓값의 크기가 큰 값을 최종 ES로 정하며, 최종 ES가 양수이면 비교의 기준이 된 나라의 특성 사용량이 많은 것이고, 음수이면 비교한 나라의 특성 사용량이 많은 것으로 해석할 수 있다.

한국과 일본의 분석 결과에서는(Fig. 6a) NBGF 1은 일본의 사용량이 한국의 사용량보다 유의하게 많은 것을 통계적 검정을 통해 확인하였다($p=0.049$). 반대로 NBGF 2는 한국의 사용량이 유의하게 더 많은 것을 보여주었다($p=0.006$). 일본과 중국의 분석에서(Fig. 6c) NBGF 2는 중국의 사용량이 일본보다 유의하게 더 많은 것을 보여주었다($p=0.001$).

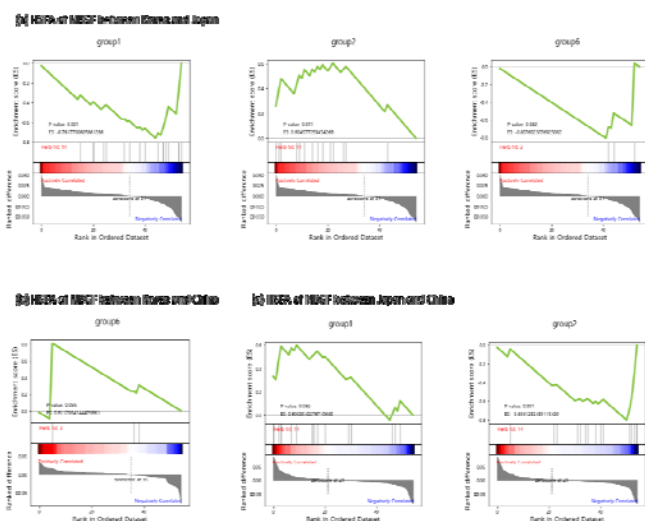


Fig. 6. HSEA result of network analysis. HSEA was performed for NBGF. Only the groups with significant trends ($p < 0.1$) were shown. Positive (Negative) ES indicates the dominance of the feature in prior (posterior) country.

고찰

본 연구에서는 한국, 중국, 일본의 전통의학 임상적 특성을 데이터를 기반으로 비교하였으며, 분석 결과, 세 나라의 전통의학이 가지고 있는 임상적 특성의 차이를 관찰할 수 있었다.

WSM 분석 결과, 일본에서는 이수삼습약을 많이 사용하며 중국에서는 활혈거어약을 많이 사용하는 것으로 나타났다. 이러한 결과는 중국과 일본의 전통의학 차이를 연구했던 기존 연구 결과와 일치한다. 조²⁾는 중국 전통의학은 기와 혈의 이원론으로 병리 개념을 설명하고, 일본 전통의학은 중국에서 사용하는 혈의 개념을 혈과 수로 나누어 기, 혈, 수의 불균형으로 병리를 설명한다고 했다. 조의 연구에 따르면, 같은 병증이라도 중국에서는 혈의 문제로, 일본에서는 수의 문제로 해석하려는 경향이 있을 것이라고 유추할 수 있다.

또 한국과 일본에서는 해표약을, 중국에서는 청열약을 많이 사용하는 것을 확인할 수 있었다. 조¹⁾는 한국에서는 동의보감이 임상에서 큰 비중을 차지하고 있고 일본에서는 상한론을 중심으로 진료하고 있어 고방이 많이 활용되고 있으며, 중국에서는 온병학의 처

방을 자주 사용하는 것이 특징이라고 하였다. 그러므로 이런 분석 결과는 한, 중, 일 세 나라가 서로 다른 의학이론에 근거하여 임상에서 이루어지고 있다는 사실을 객관적, 정량적으로 확인시켜 주는 것이라고 하겠다.

본 연구에서는 네트워크와 NBGF를 적용한 분석도 진행하였다. 한약재의 특성들을 연관성이 높은 것끼리 그룹화한 NBGF를 적용했을 때 개별 특성으로 분석했을 때와는 다른 차원의 임상 특성을 관찰할 수 있다.

Fig. 5에서 산미 특성을 가진 한약재는 일본에서 한국보다 많이 사용되고 있으나 酸味-澁味-수삼약 그룹(NBGF 4)은 한국의 사용량이 일본보다 높은 것을 확인할 수 있다. 산미는 수렴하는 성질을 가지고 있어 개별 특성으로 분석할 경우에는 일본에서 수렴하는 성질을 가지는 한약재를 많이 사용한다고 해석할 여지가 있다. 그러나 NBGF 분석을 통해 일본에서는 한국에 비해 산미를 수렴의 목적으로 많이 사용하지 않는 것을 발견하였다. 이는 NBGF를 적용하여 그룹화한 특성이 치료 방법에 기반한 임상 특성을 더 잘 반영하고 있음을 시사한다.

본 연구에서는 WSM과 NBGF에서 발견한 차이가 통계적으로 유의한 지 평가하기 위해 생물정보학 분야에서 사용하는 GSEA를 차용한 HSEA를 사용하였다. GSEA 방법론은 독립이 아닌 여러 요소들(유전자 발현)의 차이를 다중비교 하는 상황에서 '해석 가능한 묶음'(유전자의 주석)으로서 차이를 비교, 통계적으로 분석할 수 있는 수단을 제공한다. GSEA의 결과에서 차이의 정도는 ES로, 차이의 통계적 유의성은 p-value로 나타나며 그래프로 ES의 값이 표현되기 때문에 직관적으로 결과를 이해할 수 있는 장점이 있다. 본 연구에서 GSEA를 변형하여 적용한 HSEA에서는, 개별 약물들의 사용량과 약물에 속한 특성, 혹은 NBGF가 각각 GSEA의 유전자 발현량과 유전자의 주석에 해당되었으며, 이를 통해 한의학적으로 해석 가능한 이론적 차이를 통계적으로 검정할 수 있었다. 향후 한의학 연구에서 여러 요소들을 동시적으로 분석하는 경우, GSEA 방법론을 차용하는 분석방법이 유용한 전략 중 하나가 될 수 있을 것으로 생각된다.

본 연구는 몇 가지 한계점을 가지고 있다. 우선, 분석에 사용한 한약재 사용량 데이터가 충분하지 못했다. WSM 분석에서 사기와 오미, 귀경의 분석 결과는 한국과 중국의 한약재 사용 특성이 비슷한 것으로 나타났으며, 일본의 한약재 사용 특성은 두 나라에 비해 이질적인 것으로 나타났다. 위의 결과는, 결측값을 처리하는 과정에서 결측값이 없는 두 나라의 사용량 정보를 이용했기 때문에, 데이터의 결측값이 많은 한국과 중국의 한약재 사용 특성이 비슷하게 나온 것일 가능성이 있다. NBGF의 HSEA 분석에서도 조사된 한약재의 개수가 적어 충분한 분석을 할 수 없었다. 3개의 NBGF를 제외하면 한약재 집합에 들어갈 수 있는 한약재가 3개 미만이었기 때문에 편향된 결과를 보여주어 정확한 분석이 불가능했다.

또한, 한국, 일본의 데이터는 의료 현장을 반영하는 데이터였으나, 중국의 데이터는 명확한 사용처에 대한 기준을 찾을 수 없어 데이터의 전수가 의료 현장을 반영한다고 보기 어려웠다. 중국의 데이터도 대부분 의료 현장을 반영하는 데이터일 것으로 가정하여 분석을 진행하였으나 이는 데이터의 오염 가능성을 배제할 수 없

며, 검정력을 낮추는 요소로 작용했을 가능성이 존재한다.

약물들에 의해 공유되는 이론적 특성을 바탕으로 구성된 NBGF는 관찰된 이론 특성들의 차이를 보다 의미론적으로 이해할 수 있게 했지만, 몇몇 특성들의 다면적 속성 때문에 명확한 해석이 어려운 경우가 있었다. 구성된 NBGF에서 溫性, 辛味, 肝經, 脾經, 腎經 등은 네트워크 모듈의 경계에 위치하여 명확한 의미 단위 구성을 어렵게 하였는데, 이는 이 특성들이 임상에서 다면적 의미로 이용되고 있기 때문인 것으로 해석된다. 그러나 본 연구의 NBGF 구성에서는 다면적 속성을 충분히 반영하지 못하였으며, 이는 분석 결과의 검정력을 낮추는 요소로 작용하였을 것으로 생각된다.

이상의 한계점들이 전반적으로 분석 결과의 검정력을 낮추는 요소로 작용했음에도 불구하고, 본 연구는 한, 중, 일 삼국 전통의학의 임상적 적용과 그 이론적 특성에서 의미 있는 차이를 확인할 수 있었다. 특히, 이용된 데이터의 한계에도 불구하고 실제 데이터를 바탕으로 삼국 전통의학의 차이를 분석할 수 있는 새로운 방법론을 제안하였다는 의의를 가지고 있다. 향후 보강된 양질의 데이터를 기반으로 위 논문에서 시행한 정량적인 분석방법을 적용한다면, 더욱 명확한 이론적 특성을 확인할 수 있을 것으로 기대된다.

결 론

본 연구에서는 한국, 중국, 일본의 한약재 사용량 데이터를 기반으로 세 나라 전통의학 임상적용의 차이와 이를 뒷받침하는 이론적 특성의 차이를 분석하고자 하였다. 이를 위하여 한약재 사용량을 가중치로 하여 사기, 오미, 귀경, 치료 속성, 약용부위 등 약물의 특성에 대한 WSM 모델을 적용하고, 이론적 차이의 의미를 추론하기 위하여 네트워크 이론 기반의 NBGF를 정의하고 분석하였다. 또한 HSEA 방법론을 사용하여 WSM, NBGF로 분석한 결과를 통계적으로 검정하였다. 그 결과 다음과 같은 결론을 얻었다.

WSM 분석을 통해 세 나라는 모두 溫性, 甘味, 脾經, 뿌리(Radix), 보익하는 성질의 한약재를 많이 사용하고 있으며, 특히 한국에서는 腎經, 肝經을 귀경으로 하는 한약재와 뿌리줄기(Rhizoma)를, 중국은 大腸經 한약재와 청열약, 일본은 寒性 약물과 이수삼습약, 온리약을 각각 많이 사용하는 것으로 조사되었다.

NBGF 분석을 통해 한국은 鹹味-동물성 약재 그룹(NBGF 5)를, 일본에서는 寒性-苦味-辛味-청열약 그룹(NBGF 1), 熱性-온리약 그룹(NBGF 3), 淡味-이수삼습약 그룹(NBGF 6)을 가장 많이 사용하는 것으로 나타났다. 중국에서는 溫性-平性-甘味-보익약 그룹(NBGF 2)의 사용량이 한국과 비슷했으며 일본보다는 많은 것을 확인할 수 있었다.

WSM 분석을 HSEA로 검정한 결과 몇 가지 유의한 차이를 발견할 수 있었다. 한국에서는 일본보다 腎經을 귀경으로 하는 한약재를, 일본에서는 한국보다 脾經을 귀경으로 하는 한약재를 많이 사용하는 것으로 조사되었다. 한국은 중국보다 뿌리 약용부위를, 중국은 한국보다 뿌리줄기 약용부위를 많이 사용하였으며 일본은 중국과 비교하여 酸味를 가진 한약재를 많이 사용하고 있었다.

NBGF 분석에서는 한국에서는 일본에 비해 溫性-平性-甘味-보익약 그룹(NBGF 2)을 많이 사용하였으며, 일본에서는 한국과 중

국에 비해 寒性-苦味-辛味-청열약 그룹(NBGF 1)의 한약재를 많이 사용하는 것으로 조사되었다.

감사의 글

이 논문은 2016년도 가천대학교 교내연구비 지원에 의한 결과임(GCU-2016-0492).

References

1. Park HL, Lee HS, Shin BC, Liu JP, Shang Q, Yamashita H, et al. Traditional Medicine in China, Korea, and Japan: A Brief Introduction and Comparison. *Evid Based Complement Alternat Med*. 2012;2012:429103.
2. Cho KH. The Comparative Study of Oriental Medicine in Korea, Japan and China. *Journal of Korean Oriental Medicine*. 1998;19(1):271-98.
3. Lim DO, Jung MJ, Park JS, Seo GS, Hwang JW, Kim EY, et al. Investigation of Korean medicine use and herbal medicine consumption survey. *Korea Health Industry Development Institute*; 2014.
4. Survey report on raw material crude drug usage. Japan: Japan Kampo Medicines Manufacturers Association; 2011, 2013, 2015, 2016.
5. Analytical Report on Circulation of Key Chinese Medicinal Materials. China: Ministry of Commerce of the People's Republic of China Department of Market Supervision; 2011, 2012, 2013, 2014, 2015.
6. van der Heijden GJMG, T. Donders AR, Stijnen T, Moons KGM. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: A clinical example. *Journal of Clinical Epidemiology*. 2006;59(10):1102-9.
7. Real R, Vargas JM. The Probabilistic Basis of Jaccard's Index of Similarity. *Systematic Biology*. 1996;45(3):380-5.
8. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-504. Available from: <http://www.cytoscape.org>
9. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech-Theory E*. 2008;2008(Issue 10).
10. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U*

- S A. 2005;102(43):15545-50.
11. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27-30. Available from: <http://www.genome.ad.jp/kegg/>
 12. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000; 25(1):25-9. Available from: <http://www.geneontology.org/>