KOREAN JOURNAL OF ACUPUNCTURE
pISSN 2287-3368 / eISSN 2287-3376

ORIGINAL ARTICLE
Korean Journal of Acupuncture
Vol.33, No.3, pp.102-113, 2016
http://dx.doi.org/10.14406/acu.2016.011

# 기계학습을 적용한 자기보고 증상 기반의 어혈 변증 모델 구축

김현호[1] · 양승범[2] · 강연석[3] · 박영배[1] · 김재효[4]

[1]경희대학교 한의과대학 진단생기능의학과학교실, [2]원광보건대학교 의무부사관과,
[3]원광대학교 한의과대학 의사학교실, [4]원광대학교 한의과대학 경혈학교실

# Machine Learning Approach to Blood Stasis Pattern Identification Based on Self-reported Symptoms

Hyunho Kim[1], Seung-Bum Yang[2], Yeonseok Kang[3], Young-Bae Park[1], Jae-Hyo Kim[4]

[1]Department of Biofunctional Medicine & Diagnostics, College of Korean Medicine, Kyung Hee University,
[2]Department of Medical Non-commissioned Officer, Wonkwang Health Science,
[3]Department of Medical History, College of Korean Medicine, Wonkwang University,
[4]Department of Meridian & Acupoint, College of Korean Medicine, Wonkwang University

**Objectives :** This study is aimed at developing and discussing the prediction model of blood stasis pattern of traditional Korean medicine(TKM) using machine learning algorithms: multiple logistic regression and decision tree model. **Methods :** First, we reviewed the blood stasis(BS) questionnaires of Korean, Chinese, and Japanese version to make a integrated BS questionnaire of patient-reported outcomes. Through a human subject research, patients-reported BS symptoms data were acquired. Next, experts decisions of 5 Korean medicine doctor were also acquired, and supervised learning models were developed using multiple logistic regression and decision tree. **Results :** Integrated BS questionnaire with 24 items was developed. Multiple logistic regression models with accuracy of 0.92(male) and 0.95(female) validated by 10-folds cross-validation were constructed. By decision tree modeling methods, male model with 8 decision node and female model with 6 decision node were made. In the both models, symptoms of 'recent physical trauma', 'chest pain', 'numbness', and 'menstrual disorder(female only)' were considered as important factors. **Conclusions :** Because machine learning, especially supervised learning, can reveal and suggest important or essential factors among the very various symptoms making up a pattern identification, it can be a very useful tool in researching diagnostics of TKM. With a proper patient-reported outcomes or well-structured database, it can also be applied to a pre-screening solutions of healthcare system in Mibyoung stage.

**Key words :** blood stasis, pattern identification, machine learning, logistic regression, decision tree

## 서 론

변증이란 증상과 증후에 대한 분석과 판단으로서, 한의학에서 질병을 진단하는 가장 중요한 근본이 되지만[1], 증상과 증후를 측정하고 통합하는 과정에서 환자의 주관과 한의사의 주관이 개입될 가능성이 많다. 따라서 이러한 증상 정보들을 정량적으로 측정하고 평가할 수 있는 도구가 필요하며, 나아가 측정 및 평가 도구를 이용하여 진단에 활용하거나 측정 자료의 패턴을 분석할 수 있는 표준

Received June 7, 2016, Revised June 19, 2016, Accepted June 25, 2016

Corresponding author: **Jae-Hyo Kim**
Department of Meridian & Acupoint, College of Korean Medicine, Wonkwang University, 514, Iksan-daero, Iksan 54538, Korea
Tel: +82-63-850-6446, Fax: +82-63-857-6458, E-mail: medicdog@wku.ac.kr

This work (Grants No. C0272971) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2015.

화된 연구 방법이 필요하다. 이렇듯 중요한 측정과 진단의 정량화, 표준화를 위하여 그동안 한의계에서는 병인변증 설문지[2-6], 체질판별 설문지[7,8], 특정 질병 및 그에 대한 변증에 관한 설문지[9-11] 등 매우 다양한 설문지들이 연구 개발되었다. 상기 연구들은 전문가 집단의 총의를 형성하는 데에 그치지 않고, 대부분 사회조사방법론과 척도개발방법론에 입각, 실제 인간대상연구를 바탕으로 구인타당도(construct validity), 준거타당도(criterion validity), 신뢰도(reliability)등의 검증을 거쳐 발표된 내용들이다.

또한, 측정 및 평가용의 척도개발 뿐 아니라 사용된 척도를 이용, 측정된 정보를 통합하여 진단에 활용하는 방법에 대해서도 다양한 접근이 시도되었다. 전통적인 진단 통계는 주로 단수 지표와 질병과의 상관성을 민감도, 특이도, 양성예측도, 음성예측도등의 통계량으로 표현하여 해당 지표의 질병 예측도를 확률적인 측면에서 접근하는 방법이었으나, 한의학의 변증은 상술한 듯 다양한 증상과 증후의 결합으로 이루어진 다차원적 접근이기 때문에 수학적인 사칙연산이 불가능하여 단수 지표화하기 힘든 것이 사실이다. 따라서 변증연구는 주로 전통적인 진단 통계를 이용한 방법 외에도, 기계학습 등의 데이터마이닝 기법[12]을 사용하는 경우가 많았으며, 주된 대상은 사상체질 판별을 위한 모델[13-15]과 팔체질 판별을 위한 모델[16,17]이었다. 데이터마이닝의 범위를 명확히 정의하는 것은 쉽지 않으나, 대용량의 자료로부터 정보의 요약과 예측을 목표로 하며 자료에 존재하는 관계, 패턴, 규칙 등을 탐색하고 이를 통계적으로 모형화 하는 일련의 과정으로 정의하는 것이 일반적[18]이기 때문에, 한의학의 변증 모형을 해석하고 판단 모델을 만드는데 가장 적합한 방법이라 할 수 있다. 특히 기존 데이터로부터 특정 값을 예측하는 모델을 높은 적합도로 학습시키는 기계학습(machine learning)의 지도학습(supervised learning)을 기반으로 하고, 새롭게 입력되는 데이터로부터 해당 모델을 다시 학습하게 하여 모델의 정확성을 성장시켜나갈 수 있는 구조화된 데이터베이스와 솔루션이 확립된다면 한의학의 변증 진단의 타당성과 신뢰성이 더욱 향상될 것이다.

따라서 본 연구에서는, 기존에 기계학습 모델링 된 바 없는 어혈 변증을 소재로 하여, 한중일 3개국의 어혈 설문지를 자기보고방식으로 통합하고, 이를 기반으로 2건의 인간대상연구를 수행, 증상 데이터를 수집하였다. 그리고 지도학습 기법 중 가장 일반적인 다중 로지스틱 회귀 모델을 통해 전문가의 변증 진단 행위를 모방한 예측 모델이 어느 정도의 높은 적합성을 가지는지 연구하고, 추가로 모델에 대한 설명력이 높은 의사결정나무 모델을 이용하여 해당 모델의 임상적 의미와 전문가의 변증 진단 행위의 일면을 이해하고자 한다.

## 연구방법

본 연구에서는 통합설문지 개발, 증상데이터 수집을 위한 1차 인간대상연구, 전문가 판단값 수집을 위한 2차 인간대상연구, 그리고 데이터를 기반으로 한 기계학습 알고리듬 개발 작업을 순차적으로 거쳤으며, 전체적인 흐름은 Fig. 1에 나타내었다.

### 1. 한 · 중 · 일 3국 통합 자기 보고형 어혈 변증 설문지 개발

한국[6], 중국[19], 일본[20]에서 각각 개발된 어혈 평가 설문지를 분석하여 정리하였다. 한중일 각국의 한의학(중의학) 체계 차이로 인하여 약간씩 유사하거나 상이한 개념들이 혼재되어 있기에, 전문가

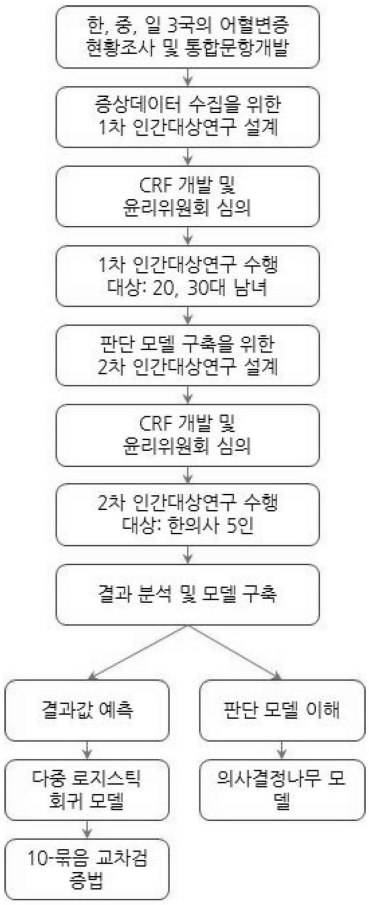


Fig. 1. Whole study procedure.

의 토론을 근거로 항목들을 통합하거나 분리하였으며, 중복된 항목들은 삭제하여 정리하였다. 중국에서 사용하는 평가항목은 상당수가 의료기기 및 병리검사 결과를 이용해야 하는 것으로서 의료기기를 비교적 자유롭게 사용할 수 있는 중의학의 환경을 반영하고 있으나, 자기 보고형 항목에 대한 연구가 본 연구의 주된 목적이므로 환자가 느끼고 직접 기록할 수 있는 자각증상에 한하여 항목을 추출하였다. 또한 맥진, 복진 등 한의사의 직접 진찰이 필요하다고 판단되는 내용은 삭제하였다. 통합형 설문지 항목 추출 과정에 참여한 패널은 한의사 면허를 가지고 있는 한의대 교수 3인, 임상경력 10년 이상인 현직 한의사 5인으로 구성되었다.

## 2. 증상데이터 수집을 위한 1차 인간대상연구

상술한 개발 과정을 통해 개발된 설문지의 항목은 총 23개이며, 여성의 경우는 월경관련 항목이 1개 추가되어 총 24개로 구성되어 있다. 따라서 증상의 유무를 단순히 이분화(있음 또는 없음)하여 응답한다 하더라도 증상들이 모두 서로 독립(independent)하다고 가정한다면 산술적으로 나타날 수 있는 경우는 총 $2^{24}$ 가지, 약 1600만 가지의 경우를 가정하여야 한다. 그러나 상기 증상들이 정상군, 미병군, 그리고 환자군에서 모두 독립적으로 발현된다고 볼 수 없기 때문에, 산술적인 모든 경우를 가정하고 알고리듬을 개발하는 것보다는 실제 조사되는 호소 증상의 조합으로 알고리듬을 개발하는 것이 합리적이라 볼 수 있다.

따라서 본 연구에서는 실제 증상 데이터 수집을 수행하는 인간대상연구를 디자인하고 수행하였다. 일반적으로 설문조사 후 데이터를 탐색하기 위하여 진행하는 탐색적 요인분석(exploratory factor analysis)의 경우 필요한 표본의 수가 확고히 정해져있지는 않으나, 일반적으로 문항수의 10~20배의 표본수를 권장하고 있다[21]. 본 연구에서 개발된 통합 어혈 설문지의 경우 항목이 24개이므로, 전체 표본의 수는 약 500건 이상을 계획하고 진행하였다. 연구 방법은 단면적 설문조사이며, 설문 데이터 수집은 경희대학교와 원광대학교 캠퍼스 내에서, 본인의 건강상태를 판단할 수 있고 설문지를 직접 작성할 수 있는 만 20세 이상 만 40세 이하의 남녀를 대상으로 자기보고형 무기명 방식으로 진행되었으며, 응답 척도는 리커르트 5점 척도를 이용하였다. 본 1차 인간대상연구는 경희대학교 한의과대학의 생명윤리위원회의 면제승인(KHSIRB-15-050(EA))을 획득한 후 진행되었으며, 응답의 성실도를 보장하기 위하여 적절한 금액의 상품이 참여 보상으로 지급되었다.

## 3. 어혈 판단값 수집을 위한 2차 인간대상연구

상술한 1차 인간대상연구를 통해 얻은 실제 증상 데이터를 기반으로, 한의학 전문가의 판단에 기준하여 어혈 변증으로의 판단 여부를 결정하는 2차 인간대상연구를 수행하였다.

본 2차 연구는 선행연구에서 획득한 성별, 나이, 그리고 증상정보를 전문가 5인에게 배포하고, 상기 정보들의 조합으로부터 한방의료 행위 시 어혈증으로 판단을 고려할 것인지의 여부를 물어보았다. 다만, 1차 인간대상연구에서 얻어진 설문 문항의 응답 척도가 리커르트 5점 척도이므로 증상 유무에 대한 정보만을 수록하기 위하여 이분화 작업을 수행하였다. 이분화 작업은 리커르트 5점 척도 중 '그렇다' 와 '매우 그렇다'를 긍정의 응답으로, '전혀 그렇지 않다'와 '그렇지 않다'를 부정의 응답으로 하였으며, 중립 문항인 '보통이다'는 긍정응답으로 한번, 부정응답으로 한번 간주하고 수행하였다. 따라서 2차 인간대상연구의 데이터는 1차 인간대상연구 데이터의 두 배가 된다.

전문가 패널은 모두 임상경력 10년차 이상의 한의사이며, 한국의 의료법 하에서 현재 진료를 하고 있는 5인으로 구성되었다. 추후 산출될 알고리듬의 외적 타당도 및 일반화 가능성, 그리고 한두명의 전문가의 판단에 대한 과적합성을 방지하기 위하여 전문가 5인의 결과상 다수의 의견이 일치하는 것을 최종 판단으로 하였다. 또한 본 연구의 주결과는 아니지만, 같은 증상을 가진 환자에 대하여 한의사의 판단이 어떻게 달라질 수 있는지 분산의 정도를 확인하기 위하여 탐색적인 목적으로 검사자간 신뢰도를 평가하였다. 검사자간 신뢰도를 평가하기 위하여 사용한 통계량은 Fleiss' kappa이다.

본 2차 인간대상연구는 경희대학교 한의과대학의 생명윤리위원회의 승인(KHSIRB-16-011)을 획득한 후 진행되었으며, 응답의 성실도를 보장하기 위하여 적절한 금액의 전문가 자문료가 참여 보상으로 지급되었다.

## 4. 데이터마이닝 기계학습 알고리듬과 모델의 타당성 분석

결과값에 대한 예측을 통계적으로 모형화 하는 지도학습(supervised learning)에서 사용되는 모델링 기법은 회귀분석(regresstion), 로지스틱회귀분석(logistic regression), 판별분석(classification), 의사결정나무(decision tree), 신경망모델(neural network), 서포트벡터머신(support vector machine), 딥 러닝(deep learning)등 매우 다양하지만, 그 중 가장 일반적이며 특별한 가정이 필요하지 않은 다중 로지스틱 회귀분석 기법과, 해석이 비교적 난해하지 않아 사용자가 쉽게 분석결과를 이해할 수 있는 의사결정나무법을 선택하여 분석을 수행하였다.

다중 로지스틱 회귀분석은 통계학에서 오랫동안 사용되어왔던 기법으로, 원래는 독립변수들의 조합으로 종속변수의 범주화된 결과값을 모델링 하는데 사용된다. 다중 로지스틱 회귀분석 기법을

Table 1. Process of Developing Integrated Patient-Reported Outcomes of Blood Stasis Questionnaires of Korea, China, and Japan

| 개발 국가 | 문항(중국, 일본의 경우 번역) | 비고 |
|---|---|---|
| 한국 | 1. 발목이나 손목, 허리가 삐끗한 일로 증상이 있다. | |
| | 2. 최근 넘어지거나 교통사고 등 심하게 부딪친 일로 증상이 있다. | |
| | 3. 일정 부위의 저림 증상이 오래도록 낫지 않는다. | |
| | 4. 몸이 쑤시듯이 아프다. | 한7과 유사 |
| | 5. 아랫배가 아프다. | |
| | 6. 옆구리가 아프다. | |
| | 7. 야간에 쑤시고 아파서 잠자기 힘들다. | |
| | 8. 복부에 덩어리가 느껴진다. | |
| | 9. 멍이 잘 든다. | |
| | 10. 얼굴색이 검다. | 한의사 진찰 필요 |
| | 11. 입술이나 혀, 잇몸의 색이 푸르거나 자주색을 띄며 어둡다. | |
| | 12. 눈밑이 푸르거나 자주색을 띄며 어둡다. | |
| | 13. 대변색이 검다. | |
| | 14. 수술횟수(0회:1점, 1,2회:2점, 3,4회:3점, 5,6회:4점, 7회이상:5점) | |
| 중국 | 1. 혀의 색이 자주색이다. | 한11 중복 |
| | 2. 아랫배가 아프다. | 한5 중복 |
| | 3. 삽맥이 관찰된다. | 한의사 진찰 필요 |
| | 4. 대변색이 검다. | 한13 중복 |
| | 5. 종괴가 느껴진다. | 한8 중복 |
| | 6. 혀 아래 정맥이 두드러져 있다. | 혈관노창으로 통합필요 |
| | 7. 결대맥이 관찰된다. | 한의사 진찰 필요 |
| | 8. 맥이 느껴지지 않는다. | 한의사 진찰 필요 |
| | 9. 배의 정맥이 두드러져 있다. | 혈관노창으로 통합필요 |
| | 10. 피부 아래 멍이 있다. | 한9 중복 |
| | 11. 월경색이 검고 덩어리진다. | 여성만 해당 |
| | 12. 지속적으로 가슴에 통증이 있다. | |
| | 13. 특정 부위에 고정된 통증이 있다. | 국소 지역 압통 및 저항감으로 통합필요 |
| | 14. 잇몸과 이 사이가 암홍색이다. | 한11 중복 |
| | 15. 혈관이 두드러져 있다. | 혈관노창으로 통합필요 |
| | 16. 손발에 내 살 같지 않은 부분이 있다. | |
| | 17. 수술병력 | 한14 중복 |
| | 18. 입천장점막의 혈관이 두드러져 있다. | 혈관노창으로 통합필요 |
| | 19. 반신마비 증상이 있다. | |
| | 20. 정신이상 증상이 있다. | |
| | 21. 피부가 끈적끈적하다. | |
| | 22. 전혈점도가 증가되어 있다. | 병리검사 필요 |
| | 23. 혈장점도가 증가되어 있다. | 병리검사 필요 |
| | 24. 체외혈전건중이 증가되어 있다. | 병리검사 필요 |
| | 25. 체외혈전습중이 증가되어 있다. | 병리검사 필요 |
| | 26. 혈소판 응고성이 증가되어 있다. | 병리검사 필요 |
| | 27. 혈전탄력성이 증가되어 있다. | 병리검사 필요 |
| | 28. 미세순환장애가 관찰된다. | 병리검사 필요 |
| | 29. 혈액동력학적 장애가 있다. | 병리검사 필요 |
| | 30. 섬유소의 용활성이 저하되어 있다. | 병리검사 필요 |
| | 31. 혈소판의 방출이 항진되어 있다. | 병리검사 필요 |
| | 32. 병리절편 검사시 혈어증이 관찰된다. | 병리검사 필요 |
| | 33. 영상검사시 혈관이 막혀있다. | 영상검사 필요 |
| 일본 | 1. 눈 아래가 착색되어 있다. | 한12 중복 |
| | 2. 얼굴이 검다. | 한10 중복 |
| | 3. 피부가 비늘처럼 거칠다. | |
| | 4. 입술이 암홍색이다. | 한11 중복 |
| | 5. 잇몸이 암홍색이다. | 한11 중복 |
| | 6. 혀의 색이 어두운 자주색이다. | 한11 중복 |
| | 7. 혈관이 두드러져 있다. | 혈관노창으로 통합필요 |
| | 8. 피부 아래 멍이 있다. | 한9 중복 |
| | 9. 손바닥에 홍반이 있다. | |
| | 10. 배꼽 왼쪽에 압통 및 저항감이 있다. | 한의사 진찰 필요(복진에 해당) |
| | 11. 배꼽 오른쪽에 압통 및 저항감이 있다. | 한의사 진찰 필요(복진에 해당) |
| | 12. 회맹장부에 압통 및 저항감이 있다. | 한의사 진찰 필요(복진에 해당) |
| | 13. S자 결장 부위에 압통 및 저항감이 있다. | 한의사 진찰 필요(복진에 해당) |
| | 14. 갈비뼈 부위에 압통 및 저항감이 있다. | 한의사 진찰 필요(복진에 해당) |
| | 15. 치질이 있다. | |
| | 16. 월경장애가 있다. | 여성만 해당 |

사용하여 구축된 모델의 엄격한 설명과 인과관계의 크기에 관한 통찰을 얻기 위해서는 회귀분석의 특성상 다중공선성(multicollinearity) 검토를 비롯한 몇 가지 기본가정에 관한 사전 분석이 반드시 진행되어야 한다. 그러나 본 연구는 독립변수 각각이 종속변수에 미치는 영향에 대한 정량적 설명을 위한 연구라기보다는, 어혈변증의 유무, 즉 종속변수에 대한 예측 가능성(predictability)과 그에 기여하는 독립변수들의 전반적인 분포를 탐색하는 기계학습 모델링이 목적이기 때문에 사전 가정 분석들이 중요하게 작용하지 않는다. 오히려 기계학습 모델의 타당성은 그러한 수학적 사전 가정보다는 모델의 성능을 평가할 수 있는 민감도, 특이도, 정확도 등의 통계량으로 판단한다.

기계학습 모델의 목표는 예측이므로, 모델의 타당성은 전통적으로 민감도와 특이도에 기반한 ROC 곡선(receiver operating characteristic curve)를 이용하여 평가하는 경우가 많다. 그러나 해당 방법을 사용하기 위해서는 데이터를 학습용(training set)과 검증용(test set)으로 분리해야 하며, 이 경우 과적합을 방지하기 위하여 검증용 데이터는 학습 목적으로 사용할 수 없기 때문에, 데이터의 양이 절대적으로 많지 않은 경우에는 좋지 않은 방법이다. 이러한 경우에 사용하는 모델의 타당성 평가 방법이 k-묶음 교차검증법(k-folds cross-validation)이며, 특히 데이터 전체를 임의적으

**한중일 통합 어혈 설문 문항**

아래 문항들은 평소(오늘을 포함하여 최근 2주일 이내) 자신이 느끼는 몸의 상태에 대한 질문입니다.
해당항목에 체크(√)를 해주십시오.

| 1점 | 2점 | 3점 | 4점 | 5점 |
|---|---|---|---|---|
| 전혀 그렇지 않다 | 그렇지 않다 | 보통이다 | 그렇다 | 매우 그렇다 |

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1. 관절에 통증이 있다. | | | | | |
| 2. 야간에 쑤시고 아파서 잠자기 힘들다. | | | | | |
| 3. 옆구리가 아프다. | | | | | |
| 4. 아랫배가 아프다. | | | | | |
| 5. 일정 부위의 저림 증상이 오래도록 낫지 않는다. | | | | | |
| 6. 멍이 잘 든다. | | | | | |
| 7. 복부에 덩어리가 느껴진다. | | | | | |
| 8. 눈 밑이 푸르거나 자주색을 띄며 어둡다. | | | | | |
| 9. 입술이나 혀, 잇몸의 색이 푸르거나 자주색을 띄며 어둡다. | | | | | |
| 10. 대변색이 검다. | | | | | |
| 11. 최근 넘어지거나 교통사고 등 심하게 부딪친 후 증상이 있다. | | | | | |
| 12. 발목이나 손목, 허리가 삐끗한 일로 증상이 있다. | | | | | |
| | 1 | 2 | 3 | 4 | 5 |
| 13. 수술횟수(0회:1점, 1,2회:2점, 3,4회:3점, 5,6회:4점, 7회이상:5점) | | | | | |
| 14. 혀 아래, 입천장, 복부, 팔다리등의 혈관이 두드러져 있다. | | | | | |
| 15. 지속적으로 가슴에 통증이 있다. | | | | | |
| 16. 손발에 내 살 같지 않은 부분이 있다. | | | | | |
| 17. 반신마비 증상이 있다. | | | | | |
| 18. 정신이상 증상이 있다. | | | | | |
| 19. 피부가 끈적끈적하다. | | | | | |
| 20. 얼굴이 검다. | | | | | |
| 21. 피부가 비늘처럼 거칠다. | | | | | |
| 22. 손바닥에 붉은 영역이 있다. | | | | | |
| 23. 치질이 있다. | | | | | |
| 24. 월경색이 검고 덩어리지며, 월경통이 있다. | | | | | |

Fig. 2. Integrated blood stasis questionnaire with 5-point Likert scale.

로 10등분하여 교대로 학습-검증하는 10-묶음 교차검증법(10-folds cross-validation)을 주로 사용한다. 이 방법은 전체 데이터를 임의로 묶은 10개 묶음 중 하나의 묶음을 선정, 이를 제외한 나머지 9개의 묶음으로 모델을 학습시킨 후 선택한 원래 묶음으로 해당 모델을 검증하여 적합도 혹은 타당성을 평가한다. 이러한 과정을 나머지 9개의 묶음에 대해서도 모두 수행하며, 산출된 10개의 적합도를 평균하여 모델의 최종 적합성을 산정하는 타당성 평가 기법이다.

의사결정나무법은 1984년 Breiman에 의해 기발된 기법[22]으로, 각 변수를 이분화하는 과정을 반복하여 판단 알고리듬을 나무 형태로 형성하는 것이 특징이다. 이분화 하는 과정이 각 node의 조건문으로 표현되며, 이 조건문은 단순 수치 비교이므로 추후 연구자가 판단의 과정을 따라가면서 해석하고 이해하기가 비교적 편리하다. 본 연구에서는 전문가 판단에 대한 해석을 시도해보기 위하여 검증용 데이터 없이 전체 데이터에 대하여 의사결정나무법을 적용해보았다.

데이터의 전처리, 기초 통계 분석, 검사자간 신뢰도, 다중 로지스틱 회귀분석, 10-묶음 교차검증법, 의사결정나무등 모든 작업과 결과물 산출은 R 3.2.3.(R Core Team, Vienna, Austria)을 사용하였으며, 검사자간 신뢰도는 "irr" 패키지를, 다중 로지스틱 회귀분석은 "stats" 패키지를, 10-묶음 교차검증법은 "DAAG" 패키지를 사용하였고, 의사결정나무기법은 과적합성에 대해서 통계적 유의성을 기준으로 가지치기를 수행해주는 "party" 패키지를 사용하였다.

# 결 과

## 1. 한 · 중 · 일 3국 통합 자기 보고형 어혈 변증 설문지

임상의를 포함한 전문가 패널의 토론 과정(Table 1)을 통하여 중복되는 문항은 제거되거나, 추가 작업이 필요한 문항은 통한 혹은 분리되었다. 그 결과 총 24문항이 통합 개발되었으며, 그 중 24번 문항은 여성에만 해당하는 문항이다. 자기 보고형 응답의 척도는 리커르트 5점 척도를 사용하도록 하였다(Fig. 2).

## 2. 인간대상연구 기초 분석

개발된 설문지를 이용하여 1차 인간대상연구인 단면적 설문조사를 실시한 결과 총 513건의 데이터가 수집되었으며, 기초 빈도분석 결과는 Table 2와 같다. 24번 문항은 여성에게만 해당하기 때문에, 성별을 나누어 분석하였다.

또한 2차 인간대상연구는 2절에서 설명한대로 1차 인간대상연구 데이터의 두배를 이용하여 실시하였는데, 이 중 모든 항목에서 '증상 없음'으로 표시되는 케이스는 삭제하였다. 그 결과 남성 데이터는 368건, 여성 데이터는 419건으로 전처리되었다. 이를 통해 얻은 전문가 5인의 검사자간 일치도를 분석하여, 같은 증상을 가진 환자에 대하여 한의사의 판단이 어떻게 달라질 수 있는지 분산의 정도를 탐색해 보았다(Table 3).

## 3. 다중 로지스틱 회귀분석 결과와 타당성 평가

2차 인간대상연구에 참여한 5인의 전문가 판단 중, 다수의 판단 결과를 최종 판단으로 가정하고 다중 로지스틱 회귀분석 모델을 구하였다. 성별 특이성이 있는 24번 문항의 존재로 인하여, 남성 응답자와 여성 응답자 모델을 각각 산출하였다(Table 4).

상기 산출된 다중 로지스틱 회귀분석 예측 모델의 성능을 평가하기 위하여 10-묶음 교차검증법을 각각 5회씩 수행하였으며, 모두 90% 이상의 성능을 나타내었다(Table 5).

## 4. 의사결정나무법을 이용한 판단 알고리듬

다중 로지스틱 회귀분석에 의한 예측 모델과는 별도로, 전문가 합의에 따른 최종 판단이 어떤 알고리듬으로 구성되어있는지 탐색하기 위하여 의사결정나무기법을 활용, 남성과 여성에 대한 판별 알고리듬을 각각 Fig. 3과 Fig. 4에 도시하였다. 남성의 경우에는 8개의 판단 노드가 존재하며, 여성의 경우에는 6개의 판단 노드가 존재하는 것으로 모델링 되었다.

**Table 2. Basic Characteristic Results of 1st Human Subject Research**

| Gender | Male | Female |
|---|---|---|
| Number of respondents | 262 | 251 |
| Mean of age | 22.63 | 22.94 |
| Standard Deviation of age | 3.00 | 3.19 |
| Missing values in age | 1 | 2 |

**Table 3. Inter-rater Reliability of Blood Stasis Decision among 5 Experts**

| | # of subjects | # of raters | Fleiss' kappa | z | *p*-value |
|---|---|---|---|---|---|
| Male | 368 | 5 | 0.498 | 30.2 | <0.001 |
| Female | 419 | 5 | 0.525 | 34 | <0.001 |

Table 4. Logistic Regression Model based on the Final Decision

| | Male | | Female | |
|---|---|---|---|---|
| | Coefficient | SE | Coefficient | SE |
| (Intercept) | −17.94*** | 3.78 | −20.07*** | 4.90 |
| item01 | 2.74** | 0.95 | 4.49** | 1.66 |
| item02 | 10.45*** | 2.84 | 8.79** | 2.97 |
| item03 | 6.45 | 3.96 | −0.61 | 4.71 |
| item04 | 1.23 | 1.14 | 2.36 | 1.36 |
| item05 | 8.38*** | 1.92 | 7.18** | 2.24 |
| item06 | 6.51*** | 1.90 | 8.19*** | 2.34 |
| item07 | 2.00 | 1.28 | 1.80 | 2.00 |
| item08 | 6.32*** | 1.63 | 5.76** | 1.86 |
| item09 | 9.56 | 11.77 | 8.47*** | 2.34 |
| item10 | 5.47*** | 1.44 | 2.03 | 2.04 |
| item11 | 31.49 | 17.26E+02 | 38.24 | 66.98E+02 |
| item12 | 5.15*** | 1.24 | 5.38*** | 1.55 |
| item13 | 14.72*** | 3.62 | 18.96*** | 4.88 |
| item14 | 8.21*** | 2.03 | 7.71** | 2.60 |
| item15 | 8.69*** | 2.09 | 6.00 | 3.51 |
| item16 | 5.71*** | 1.38 | 4.41 | 4.11 |
| item17 | 8.86 | 32.86 | −20.21 | 47.45E+06 |
| item18 | 1.83 | 3.75 | 13.55** | 4.32 |
| item19 | 5.11** | 1.71 | 4.65 | 3.43 |
| item20 | 1.39 | 1.03 | −3.14 | 2.52 |
| item21 | 5.02** | 1.53 | −5.88 | 3.06 |
| item22 | 1.40 | 0.89 | −1.23 | 1.58 |
| item23 | 2.79* | 1.28 | 5.93* | 2.34 |
| item24(female) | - | - | 46.25 | 29.97E+02 |

SE : Standard Error.
Statistical significance : ***$p$-value<0.001, **$p$-value<0.01, *$p$-value<0.05.

Table 5. 10-folds Cross-validation(5 times repeat) of Multiple Logistic Regression Model for the Final Decision

| | | Accuracy | |
|---|---|---|---|
| | | Male model | Female model |
| 10-folds cross-validation repeat | 1 | 0.93 | 0.94 |
| | 2 | 0.91 | 0.95 |
| | 3 | 0.91 | 0.95 |
| | 4 | 0.92 | 0.95 |
| | 5 | 0.93 | 0.94 |
| Average accuracy | | 0.92 | 0.95 |

## 고 찰

### 1. 한 · 중 · 일 어혈 설문지와 통합 설문지의 특징

본 연구의 기초자료로 사용한 한국, 중국, 일본의 기 개발된 어혈 설문지를 분석함으로써 각각의 특징을 분명히 지니고 있음을 확인하였다. 한국의 어혈 설문지는 증상의 발현 동기, 주로 통증과 관련된 자각증상, 그리고 혈관 관련 망진 정보 등으로 구성되어 있는데 반해, 중국의 어혈 설문지는 전통적인 의서의 기록에서 나타나는 증상정보와 더불어, 의료기기를 사용한 진단검사, 영상검사 등의 결과값을 어혈 평가에 사용하여 그 정확성을 높이고자 한 것이 특징이다. 특히 진단검사와 영상검사 항목(12개)이 전체 항목(33개)의 36%가량을 차지하고 있다는 것은 중국의 의료 환경에서 중의사가 의료기기를 사용하는 데 큰 제약조건이 없는 현실을 그대로 반영하는 것으로 생각된다. 특히 중국의 어혈 설문지가 소개된 시기를 고려한다면, 30년 전부터 객관적이고 정량적인 검사 항목들을 적극적으로 중의학에 접목하여 발전을 시도하였던 것은 다양한 측면으로 시사하는 바가 크다고 하겠다. 일본의 어혈 설문지 역시 전통적인 의서에 기록에서 나타나는 증상정보들은 한국의 그것과 유사하나, 특징적인 것은 절진의 하나인 복진 항목(5개)이 전체 문항(16개)의 30% 정도를 차지할 정도로 강조된다는 점이다(Table 1).

그리고 한국의 어혈 설문지는 델파이기법, 요인분석, 그리고 절단값 산출을 위한 임상시험 등 통계적인 기법과 실제 데이터를 이

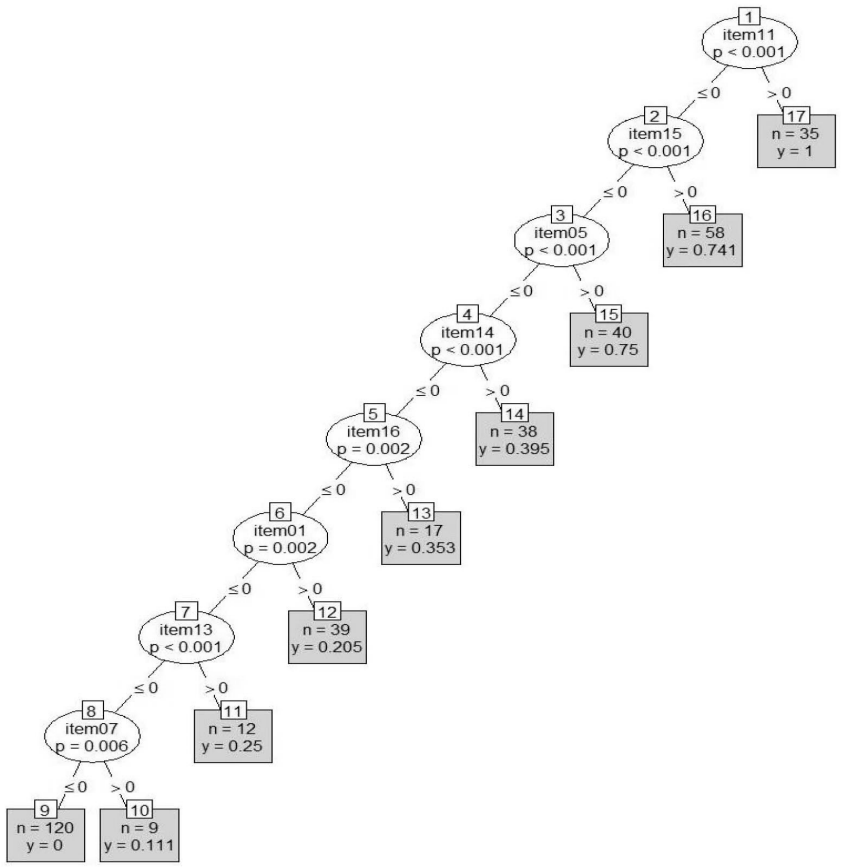


Fig. 3. Decision tree based on the final decision for male.

용하여 신뢰성, 타당성의 검증 절차를 거쳐 개발되었으나, 중국과 일본의 경우에는 정확한 개발 방법과 총의 형성 방법에 대하여 보고되지 않았다[6].

본 연구에서는 한국, 중국, 일본의 어혈 설문지를 통합하여 환자의 자기보고형 설문 항목들을 추출하는 것이 목적이었으므로, 중국 어혈 설문지의 진단검사, 영상검사 등의 항목과 함께 일본 어혈 설문지의 복진 항목들은 제외되었다. 그 외 증상들은 대부분 전통적인 의서에 기록되어 있는 한의학 이론에 따른 증상들이며, 이러한 증상들은 한국, 중국, 일본 어혈 설문지에 공통적으로 나타나는 경우가 많았다(Table 1).

## 2. 검사자간 일치도

2차 인간대상연구를 수행한 5인의 한의사 간의 일치도를 평가하였다. 이 일치도 연구는 환자의 자기보고형 설문 항목이 같았을 경우, 한의사의 판단이 얼마나 일치하는 것인가를 수치적으로 평가하기 위한 목적이다. 2인 이상의 검사자간 신뢰도(inter-rater reliability)를 판단할 때 사용하는 통계량인 Fleiss' kappa를 산출한 결과, 남성에 대한 판단은 0.498의 일치도가, 여성에 대한 판단은 0.525의 일치도가 산출되었다. kappa 통계량의 절대적인 기준은 마련되어 있지 않으나, 전통적으로는 0.41 이상을 중간(moderate)정도의 일치도, 0.61 이상을 상당한(substantial) 일치도로 판단하였으며[23], 최근에 보건의료계에서는 0.40 이상을 약한(weak) 일치도, 0.60 이상을 중간(moderate)정도의 일치도로 해석해야 한다는 의견[24]이 있다. 본 연구에서 산출된 일치도는 전통적인 의견에 따르면 중간 정도, 그리고 최신 의견에 따르면 약한 일치도로 해석할 수 있는데, 이는 5인 검사자간 일치성이 다소 강하지 못하다고 생각할 수 있다. 환자의 자기보고형 증상 정보만을 제공한다는 것은 사진(四診) 중 문진(問診)의 일부만을 가지고 의학적 판단을 시도하는 것이라는 측면에서 볼 때, 본 연구에서 진행한 전문가의 판단 과정은 의학적으로 많은 정보들이 부족한 상태에서 이루어진 것이다. 따라서 충분히 어혈 변증으로 확진할 수 있는 정보의 부족때문에 위와 같이 강하지 못한 일치도가 산출된 것으로 생각된다.

## 3. 다중 로지스틱 회귀 모델과 의사결정나무 모델

남성에 해당하는 23개의 문항과 여성에 해당하는 24개의 문항의 로지스틱 회귀 모델을 구한 결과(Table 4)에 대해 교차검증법을

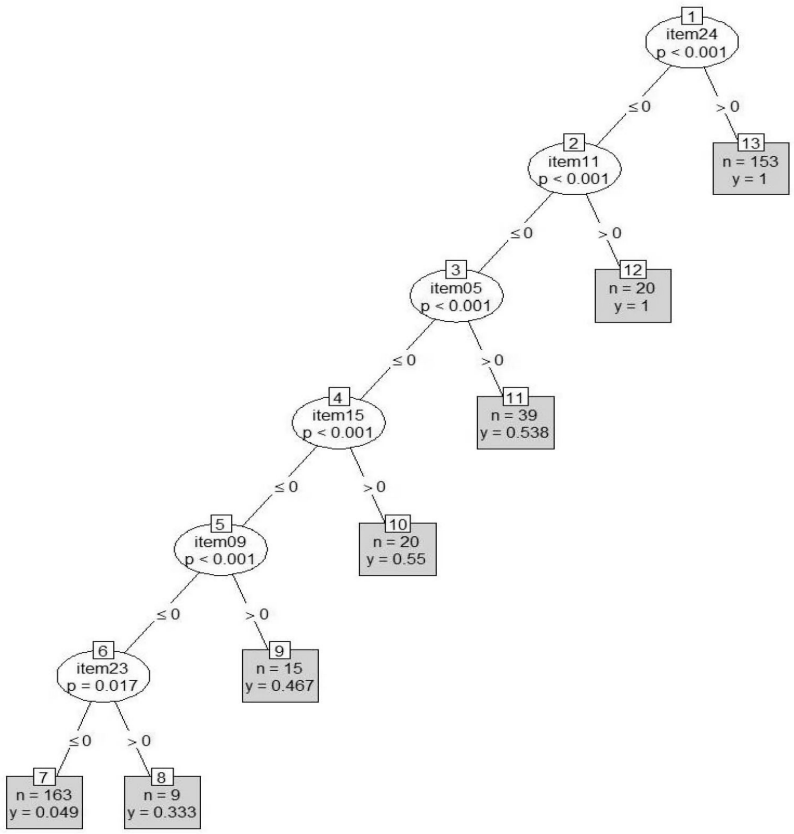


Fig. 4. Decision tree based on the final decision for female.

통해 타당성을 산출한 결과, 90% 이상(남성 92%, 여성 95%)의 높은 정확도를 보였다. 90%의 타당성이란, 학습용 데이터를 가지고 기계학습된 모델에 새로운 열 개의 데이터가 적용되었을 때 아홉 개의 데이터에 대한 결과값을 정확히 예측한다는 의미이다. 따라서 본 회귀 모델은 전문가 5인의 최종 판단을 높은 정확도로 재현하고 있는 모델이라고 할 수 있다.

남성 모델의 경우에는 11번(최근 물리적충격), 13번(수술횟수), 2번(야간통증) 항목이 비교적 큰 계수를 가지고 있으므로 종속변수에 영향을 크게 미치고 있으며, 여성 모델의 경우에는 24번(월경상태), 11번(최근 물리적충격), 13번(수술횟수), 17번(반신마비, negative의 영향), 18번(정신이상) 항목이 비교적 큰 계수를 가지고 있다. 계수의 크기는 크지만 통계적으로 유의하지 않은 항목에 대해서는 표본 집단 내에서는 높은 영향력이 있으나 모집단으로 확장하는 데에는 표본이 부족하거나 실질적으로 유의하지 않을 수 있다고 판단할 수 있다. 실질적으로 본 연구에서 사용된 기초 증상 데이터는 대학캠퍼스에서 활동하는 20대와 30대 남녀에게서 수집한 것이기 때문에 13번(수술횟수), 17번(반신마비), 18번(정신이상), 12번(치질)과 같은 증상들의 긍정 응답 빈도가 다른 항목에 비하여 현저히 낮았다. 따라서 추후 모집단의 특성을 충분히 반영할 수 있도록 문항을 재선정 하거나 혹은 해당하는 케이스를 충분히 더 수집할 필요가 있다.

또한 2.4.절에서 상술한 바와 같이, 기계학습을 이용한 예측도구로서 사용되는 로지스틱 회귀분석은 인과관계의 크기 산출 및 증명에서 사용되는 회귀분석과는 달리 회귀 계수 각각을 고려하여 각 변수에 확정적인 의미를 부여하기가 힘들다. 특히 다중 회귀분석의 경우에는 다중공선성이 충분히 제어되어야만 각 변수의 독립적인 증감이 종속변수에 미치는 영향을 정량적으로 판단할 수 있으나, 기계학습에서의 다중 회귀분석은 다중공선성에 대한 제어 없이 모델의 적합도를 높이는 데 목적으로 두고 있기 때문에 더욱 그러하다. 따라서 단순히 계수의 크기만을 가지고 독립변수들의 영향력을 판단하거나 비교하는 것은 불가능하다.

위와 같은 이유로 인하여 예측도는 높지만 모델의 설명이 난해하다는 특징을 가진 다중 로지스틱 회귀분석 모델과는 달리, 의사결정나무 모델은 연속적인 수치의 비교를 통해 다중 분기문으로 이루어져 있기 때문에 비교적 쉽게 모델을 이해할 수 있으며, 우선적으로 어떤 항목들이 종속변수에 영향을 주는 것인지 판단할 수 있다.

Fig. 3의 결과를 보면, 남성을 대상으로 어혈 판단을 할 때 가장

먼저 고려하는 항목은 11번(최근 물리적충격)으로 나타났으며, 11번 항목에 긍정 응답한 35명의 케이스에 대하여 모두 어혈증으로 판단하였다. 그 후 고려 항목은 15번(가슴 통증)으로, 긍정 응답의 74%를 어혈증으로 판단하였다. 그 다음 나타나는 5번(저림증상)항목에 긍정응답한 40 케이스에 대해 75%를 어혈증으로 판단하였고, 그 다음 나타나는 14번(혈관노창), 16번(사지감각이상)의 긍정응답의 경우는 각각 어혈군이 39.5%와 35.3%로, 분류 결과의 분포가 모호하였다. 그 다음 판단 노드인 1번(관절통증), 13번(수술횟수), 7번(복부종괴)의 경우는 해당 증상에 대해 긍정응답을 하였다 하더라도 어혈증이 아닌 경우가 더 많았기 때문에, 일괄적으로 비어혈증으로 분류할 수 있으며, 어혈증으로 판단할 하위 판단 항목이 통계적으로 유의하게 존재하지 않는다. 따라서 남성의 경우에는 11번(최근 물리적충격), 15번(가슴통증), 5번(저림) 항목에 하나라도 긍정응답을 한 경우에는 대부분 어혈증으로 분류가 되었으며, 그렇지 않은 경우 중 14번(혈관노창)과 16번(사지감각이상)에 긍정 응답한 경우에는 판단이 불확실하며, 그 외의 경우에는 어혈증으로 분류할 수 있는 근거가 존재하지 않았다. 같은 방법으로 여성 모델인 Fig. 4를 분석하면, 가장 우선적이고 중요한 변수로 24번(월경상태)이 추출되었다. 24번(월경상태) 문항에 긍정응답을 한 153명 모두 어혈증으로 판단되었다. 그리고 월경증상은 없으나 11번(최근 물리적충격)에 긍정응답을 한 20명이 모두 어혈증으로 판단되었다. 그 이후에 나타나는 5번(저림), 15번(가슴통증), 9번(구순,설색암자) 항목의 경우는 모두 모호한 판단으로 생각되며, 그 외의 경우에는 어혈증으로 분류할 수 있는 유의한 항목이 존재하지 않았다. 따라서 의사결정나무 모델을 적용하여 모델을 설명하였을 경우, 실질적으로 남녀 모두 어혈증 판단에 있어서 초기 2개 증상의 존재여부가 어혈증 판단에 절대적으로 영향을 미치는 것으로 추출되었다. 그 이후부터는 3~5가지의 증상을 고려하나, 그 판단 결과는 모호한 분포를 가지며, 그 이후에도 증상이 추가되지 않을 경우에는 비어혈증으로 판단하는 양상을 보인다. 모델에 나타난 모호한 판단 결과는, 추가적인 독립변수가 추가된다면 새로운 하위 노드의 등장과 함께 분류 정확도가 상승할 것으로 생각된다.

### 4. 연구의 한계와 추가 연구

본 연구 디자인이 가지는 제한점은 크게 세 가지로 나누어 논의할 수 있다.

첫째, 본 연구에서 가정한 참값은 설문 대상자가 가지는 실제 어혈증의 참값이 아니라 한의사 5인의 판단 여부로 결정된다. 즉, 참값은 전적으로 개인의 임상 경험과 한의학 지식에 기반하였으며, 실제 어혈증 여부는 다른 형태의 연구 디자인을 통하여 얻을 수 밖에 없다. 따라서, 본 연구에서 구축된 모델은 실제 어혈증 여부에 관한 예측이라기보다는, 임상경력 10년 이상의 한의사 5인의 총의를 예측하는 모델로 해석하여야 한다. 한의학의 변증 진단은 특정 화학 물질의 관찰이나 조직학적 검사 등에 의해서 결정되는 것이 아니라 환자의 증상과 증후를 통하여 이루어지는 만큼, 그 참값을 산출하기에는 많은 어려움이 따르며, 다른 변증명들과 마찬가지로 어혈의 개념이 넓은 범위를 아우르기 때문에 그 여부를 쉽게 판단하기가 힘든 것이 사실이다. 추후 치료 전후의 상태를 비교 분석하는 연구를 설계하여 실제로 환자의 증상이 사라졌는가에 대한 임상적 효과를 관찰한다면, 전적으로 전문가 판단에만 의존하기에는 부족한 부분을 일정 부분 보충해 줄 수 있을 것으로 생각된다.

둘째, 본 연구에서 사용된 기계학습을 위한 독립변수, 즉 특징(feature)들이 모두 자기보고결과물이라는 한계가 있다. 의학적 진단을 하기 위하여 필요한 정보가 일부 문진(問診)에만 국한될 수 없다는 것은 매우 분명한 사실이므로, 이 부분은 추후 한의사의 진찰에 의한 결과물, 진단검사 혹은 영상검사의 결과물이 새로운 독립변수들로 입력되고, 어혈 변증에서 나타날 수 있는 유사 징후들에 대한 연구들이 추가로 진행된다면, 보다 더 현실에 가까운 변증 진단 툴이 될 수 있을 것으로 생각된다. 그러나 최근에는 전통적인 의료 개념에 더하여, 질병단계가 아닌 미병단계의 환자가 의료 시설을 방문하지 않고 자신의 건강을 지속적으로 평가하고 관리하는 건강관리 개념이 급부상하고 있으므로, 이러한 형태의 건강관리 솔루션으로서 이 자기보고 증상를 이용한 판단 모델이 사전검사(pre-screening)의 일부분을 담당할 수 있을 것으로 생각한다.

셋째, 1차 인간대상연구의 모집단이 20대와 30대에 국한되었기 때문에 통합 설문지를 개발할 당시에 선정된 독립변수들에 대한 증상데이터들이 모든 항목에 골고루 모이지 못했다는 한계가 있다. 이런 경우 기계학습모델은 해당 변수에 대한 충분한 정보를 얻지 못하여 모집단으로의 일반화 가능성이 낮아진다. 실제로 본 연구에서 선정한 반신마비 혹은 정신이상 증상은 표본 집단에서 거의 관찰되지 않았으며, 이로 인하여 판단 모델에서 해당 변수의 역할이 거의 없거나 해석이 난해한 결과가 도출된 부분도 존재한다.

## 결 론

본 연구에서는 한의학 혹은 중의학의 중요한 변증 개념중 하나인 어혈변증에 대하여 기계학습 중 지도학습 알고리듬을 사용하여 예측 모델을 구축, 평가하고, 그 과정을 분석하였다. 다중 로지스틱 회귀분석을 사용하여 높은 정확도의 판단 모델을 구축하였으며, 추

가로 의사결정나무 기법을 이용하여 어혈 변증에 관하여 전문가의 판단 과정에 대한 이해를 시도해보았다. 한의학에서 변증 참값의 기준, 자기보고 증상의 임상적 제한등 비교적 본질적인 한계를 가지고 있으나, 추후 구조화된 데이터베이스가 개발되고, 다양한 임상시험을 통하여 임상적인 참값들이 기록된다면 이러한 기계학습 알고리듬은 한의사의 진료에 많은 도움을 줄 수 있는 임상의사결정 지원시스템(Clinical Decision Support System)으로서 중요한 역할을 할 수 있을 것이다.

## 감사의 글

This work(Grants No. C0272971) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2015.

## References

1. Yim HJ, Kim SH, Lee SR, Jung IN. Study to develop the instrument of pattern identification for Hwa-byung. Korean J Oriental Physiology & Pathology. 2008 ; 22(5) : 1071-7.
2. Park YJ, Park JS, Kim MY, Park YB. Development of a valid and reliable phlegm pattern questionnaire. J Altern Complement Med. 2011 ; 17(9) : 851-8.
3. Park YJ, Lim JS, Park YB. Development of a valid and reliable food retention questionnaire. European Journal of Integrative Medicine. 2013 ; 5(5) : 432-7.
4. Yoon KJ, Park YB, Park YJ, Kim MY. Development and validation of a Lao Juan questionnaire. Chin J Intergr Med. 2015 ; 21(7) : 500-6.
5. Ryu HH, Lee H, Kim H, Kim JY. Reliability and validity of a cold-heat pattern questionnaire for traditional Chinese medicine. J Altern Complement Med. 2010 ; 16(6) : 663-7.
6. Park YJ, Yang DH, Lee JM, Park YB. Development of a valid and reliable blood stasis questionnaire and its relationship to heart rate variability. Complement Ther Med. 2013 ; 21(6) : 633-40.
7. Kim SH, Ko BH, Song IB. A study on the standardization of QSCC II. J of Sasang Constitutional Medicine. 1996 ; 8(1) : 187-246.
8. Lee JH, Ko BH, Song IB. A study on the validation of QSCC II. J of Sasang Constitutional Medicine. 1996 ; 8(1) : 247-94.
9. Park DM, Lee SR, Kang WC, Jung IC. Preliminary Study to Develop the Instrument of Pattern Identification for Jing Ji and Zheng Chong. Journal of Oriental Neuropsychiatry. 2010 ; 21(2) : 1-15.
10. Lee IS, Bae GM. A Clinical Study on Differentiation of Syndromes of Amenorrhea or Oligomenorrhea with DSOM. The Journal of Oriental Obstetrics and Gynecology. 2009 ; 22(2) : 189-208.
11. Shin JH, Jung WY, Moon YK, Nam HJ, Kim YB, Lee JH, et al. An Expert Survey for Developing the Pattern Diagnosis Instrument of Acne. 2015 ; 28(2) : 23-32.
12. Kim SK, Jang HC, Kim JH, Kim C, Yea SJ, Song MY. Computational methods for traditional Korean medicine : A survey. Korean J Oriental Physiology & Pathology. 2011 ; 25(5) : 894-9.
13. Kim KK, Kim JW, Lee EJ, Kim JY, Choi SM. Study on classification function into Sasang constitution using data mining techniques. Korean J Oriental Physiology & Pathology. 2004 ; 18(6) : 1938-44.
14. Hong JW, Kim YI, Park SJ, Kim BC, Eom IK, Hwang MW, et al. Data mining algorithms for the development of Sasang type diagnosis. Korean J Oriental Physiology & Pathology. 2009 ; 23(6) : 1234-40.
15. Chae H, Hwang SM, Eom IK, Kim BC, Kim YI, Kim BJ, et al. Development of Sasang type diagnostic test with neural network. Korean J Oriental Physiology & Pathology. 2009 ; 23(4) : 765-71.
16. Shin YS, Park YB, Park YJ, Kim MY, Lee SC, Oh HS. A fundamental study for 8 constitution medicine diagnosis expert system development. J Korean Institude of Oriental Medical Diagnostics. 2007 ; 11(1) : 25-47.
17. Shin YS, Park YB, Park YJ, Kim MY, Lee SC, Oh HS. A study for 8 constitution medicine diagnosis expert system development. J Korean Institude of Oriental Medical Diagnostics. 2008 ; 12(1) : 142-84.
18. Hand D, Mannila H, Smyth P. Principles of Data Mining. Cambridge, MA : MIT Press. 2001.
19. Yang DH, Park YJ, Park YB, Lee SC. Development of ques-

tionnaires for blood stasis pattern. J Korean Institude of Oriental Medical Diagnostics. 2006 ; 10(1) : 141-52.
20. Kim JB, Choi SH, Ahn KS. Study on the effects of Taorencheng-qitang and its components on blood stasis model. Korean J Oriental Medical Pathology. 1997 ; 11(1) : 65-76.
21. Gorsuch. Factor Analysis, 2nd edition. NJ : Lawrence Erlbaum. 1983.
22. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Monterey, CA : Wadsworth & Brooks/Cole Advanced Books & Software. 1984.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 ; 33 : 159-74.
24. Mary LM. Interrater reliability: the kappa statistic. Biochem Med. 2012 ; 22(3) : 276-82.