

체질진단설문도구의 체질별 진단정확률 차이에 대한 고찰

김상혁^{1,*}

¹한국한의학연구원 한의약데이터부 책임연구원

Abstract

Differences in the Diagnostic Accuracy of the Questionnaire for Sasang Constitution Diagnosis Across Constitutional Types

Sang-Hyuk Kim^{1,*}

¹KM Data Division, Korea Institute of Oriental Medicine

Objectives

This study aimed to compare the diagnostic accuracy of the Questionnaire for Sasang Constitution Diagnosis(QSCD) across constitutional types by recalculating sensitivity and specificity, rather than relying solely on Percent Correctly Predicted. It also sought to explore potential reasons for variations in diagnostic performance based on prior research.

Methods

Studies reporting the diagnostic accuracy of the QSCD were reviewed. For each constitutional type, 2×2 contingency tables of true positives, false positives, true negatives, and false negatives were reconstructed to compute sensitivity and specificity. When multiple studies were available for a single QSCD version, pooled estimates with 95% confidence intervals were derived using a diagnostic meta-analysis approach.

Results & Conclusions

Considering the QSCD's frequent use as a screening tool in clinical settings, the SDQ_S showed superior accuracy in diagnosing Taeum(TE) and Soyang(SY) types, whereas the KS-15 demonstrated better performance for the Soeum(SE) type. However, the SY type consistently exhibited relatively low sensitivity across questionnaires. This may result from the insufficient development of scales specific to the SY type, methodological limitations of discriminant analysis commonly applied in QSCD studies, and the absence of a consistent SY type diagnostic image among experts.

Key Words : Questionnaire, Sensitivity and specificity, Sasang constitution

Received 23, October 2025 Revised 31, October 2025 Accepted 11, December 2025

*Corresponding author : Sang-Hyuk Kim

KM Data Division, Korea Institute of Oriental Medicine,

1672 Yuseongdae-ro, Yuseong-gu, Daejeon 34054, Republic of Korea

Tel : +82-42-868-9238, Fax : +82-42-868-9480, E-mail : kataz@kiom.re.kr

© The Society of Sasang Constitution and Immune Medicine. All rights reserved. This is an open access article distributed under the terms of the Creative Commons attribution Non-commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>)

I. 서론

체질진단은 환자의 평가와 진단, 치료 및 건강관리 등 사상체질의료의 모든 과정에 필수적으로 전제된다¹. 이러한 체질진단의 객관성 확보를 위해 정량적 체질진단 도구 개발을 위한 연구들이 진행되어 왔으며 체질진단설문도구는 오랜 기간 연구가 지속된 분야 중 하나이다².

사상체질 진단을 위한 설문도구 연구들은 도구의 진단정확률을 평가하는데 주로 정확예측률(Percent Correctly Predicted, PCP)을 사용해 왔다. 정확예측률은 체질이 올바르게 진단된 경우의 수를 진단에 사용한 연구대상자의 총수로 나눈 백분율을 말한다. 그러나 정확예측률은 진단검사의 유용성을 숫자 하나로 간단하게 표시하는 이점이 있는 반면 그 외의 타당성 정보들을 표현하지 못하는 단점 또한 지니고 있다³. 이에 기존 연구들에서 제시하는 진단정확률을 체질유형별 민감도, 특이도로 수량화하여 평가할 필요가 제시되었으며³, 다빈도로 사용되는 사상체질분류검사Ⅱ(Questionnaire for Sasang Constitutional Classification Ⅱ, QSCCⅡ)에 대해서는 체질별 진단 정확률의 체계적 문헌고찰 결과가 보고되기도 하였다. 해당 연구는 QSCCⅡ의 전반적인 진단정확률을 보통(moderate)으로 평가하고 서로 다른 체질별 진단정확률을 보고하면서 상대적으로 소양인 진단정확률이 낮다고 하였다⁴.

본 연구는 이러한 선행연구를 참고하여 여러 체질진단설문도구의 진단정확률을 체질유형별 민감도, 특이도로 다시 수량화하여 비교함으로써 체질별 진단정확률 차이를 확인하고 평가하고자 하였다. 아울러 체질별 진단정확률이 차이가 나게 된 원인에 대해서 체질진단 관련 기존 연구 내용들을 검토하여 추론해보고자 하였다.

II. 연구방법

1. 체질진단설문도구 선정

대상 체질설문도구는 선행연구⁵를 참고하여 사상체질변증 기초설문지(1)⁶, 사상체질분류검사(Questionnaire for Sasang Constitutional Classification, QSCC)⁷, QSCCⅡ⁸, QSCCⅢ⁹, QSCCⅡ⁺¹⁰, 의사결정나무법에 근거한 16문항 사상체질진단검사지(QSCCⅡ Decision tree)¹¹, 의사용 및 환자용 설문지(Sasang Constitution Questionnaire for Doctors and Patients, SSCQ-D and SSCQ-P)¹², 사상체질진단을 위한 2단계 설문지(Two Step Questionnaire for the Sasang Constitution Diagnosis, TS-QSCD)¹³, 단축형 사상체질 진단설문지(Korea Sasang Constitutional Diagnostic Questionnaire, KS-15)¹⁴, 전문가용 사상체질진단지(Sasangin Diagnostic Questionnaire for Specialist, SDQ-S)¹⁵로 하였다. 의사용 진료기록지에 가까운 형태인 체질판별표¹⁶는 진단정확률을 보고한 예가 없어 제외하였다.

2. 진단정확률 연구 선정 기준

선정된 체질진단설문도구를 사용하여 태음인, 소음인, 소양인의 진단정확률을 보고한 연구 중 한국 성인을 대상으로 하고, 각 체질별 참양성(true positive), 거짓양성(false positive), 참음성(true negative), 거짓음성(falsenegative)을 추출할 수 있으며 검증표본(test group)을 사용하여 외적타당도(external validity)를 제시한 연구들을 선정하였다. 단, 해당 체질진단설문도구에 검증표본을 사용한 연구가 한 건도 없을 경우에는 분석표본(training group)만을 사용한 연구를 포함하였다. 유소아나 청소년, 외국인 등 한국 성인을 대상으로 하지 않은 연구는 제외하였다.

3. 연구 검색

연구 검색은 한의학문헌 특화 데이터베이스인 OASIS 전통의학정보포털(<https://oasis.kiom.re.kr>)을 사용하였다. 위에서 대상으로 선정할 10종의 체질진단 설문도구를 주제로 검색을 수행하였으며, 이렇게 검색된 연구들이 선정 기준에 맞는지 제목과 초록 검토를 통해 선별하였다. 이후 선별된 연구들의 참고문헌에 포함된 연구들을 다시 검토하는 스노우볼링(snowballing) 방식을 사용하였다. 아울러 선행연구들^{1,4,5,17}에서 제시한 연구들 역시 포함하였다. 이어 중복된 연구들을 제거한 후 연구들의 검증표본 사용 여부를 검토하였다.

4. 자료 추출

선정된 연구들에서 연구의 저자, 출판연도, 사용된 체질진단설문도구의 종류, 연구대상자 수, 사용한 표본이 분석표본인지 검증표본인지의 여부, 참조표준(reference standard) 대비 사용된 체질진단설문도구의 예측빈도수를 기술한 태음인·소음인·소양인의 체질별 3X3 표, 각 체질별 참양성·거짓양성·참음성·거짓음성 수를 정리하였다. 자료의 정리는 Microsoft Office Excel 2021을 이용하여 진행하였다.

5. 통계방법

참조표준 대비 사용된 체질진단설문도구의 예측빈도수를 기술한 태음인·소음인·소양인의 체질별 3X3 표는 Microsoft Office Excel 2021을 사용하여 각 연구의 정확예측률을 계산하는데 사용하였다. 선정된 진단정확률 연구들의 각 체질별 참양성·거짓양성·참음성·거짓음성 수는 Meta-DiSc 2.0¹⁸을 사용하여 체질별 민감도와 특이도 값을 구하는데 사용하였다. Meta-DiSc 2.0은 진단검사 정확도 연구의 메타분석에 사용되는 소프트웨어로, 분석 과정에서 각 개별 연구의 민감도와 특이도 값을 95% 신뢰구간과 같이 표시해 주기 때문에 이를 활용하였다. 한 종류의 체질진단

설문도구에 두 건 이상의 진단정확률 보고 연구가 있을 경우에는 Meta-DiSc 2.0이 제공하는 진단 메타분석(diagnostic meta-analysis)을 사용하여 95% 신뢰구간과 같이 표시하였다. 소프트웨어의 권고에 따라, 통합된 체질별 민감도와 특이도 값을 구하는데 있어 네 건 이상의 진단정확률 보고가 있을 경우에는 bivariate random effects model이 적용되었으며 네 건 미만의 보고가 있을 경우에는 univariate model이 적용되었다.

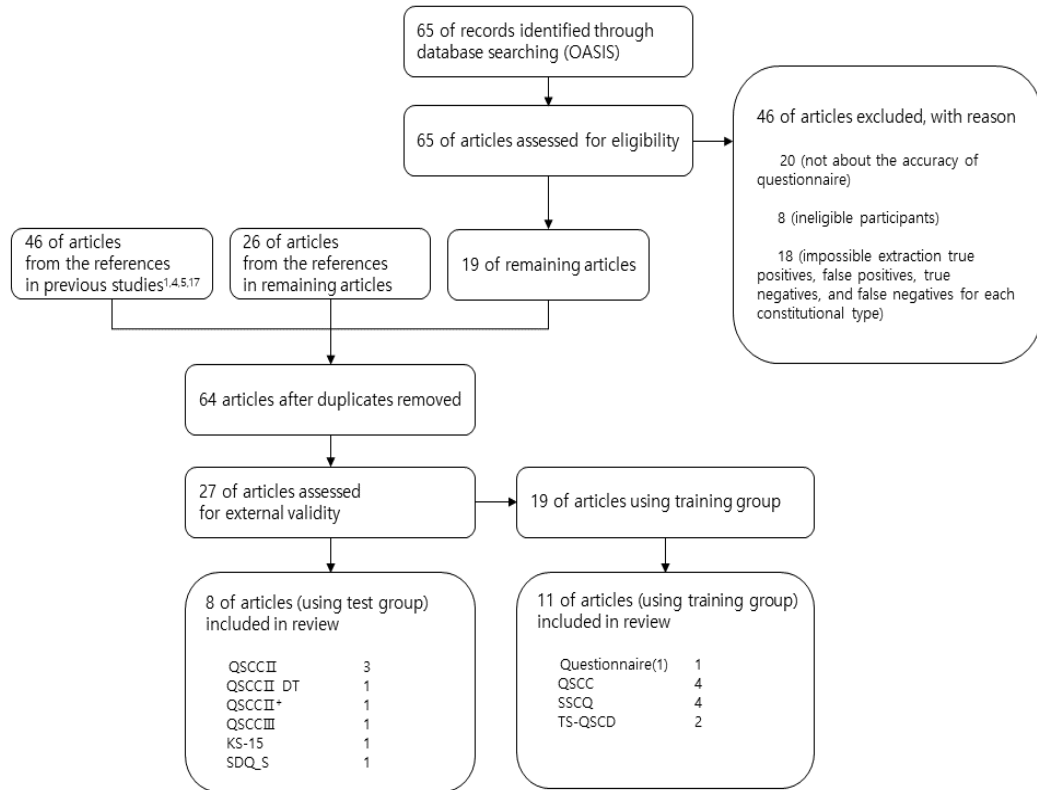
III. 결과

연구 검색 및 선정 과정은 Figure 1과 같으며 최종적으로 19건의 연구가 선정되었다. QSCCⅡ, 의사결정나무법에 근거한 16문항 사상체질진단검사지, QSCCⅡ*, QSCCⅢ, KS-15, SDQ-S 등 6종의 설문지는 검증표본을 사용한 진단정확률이 보고되었으나, 사상체질변증 기초설문지(1), QSCC, SSCQ, TS-QSCD 등 4종의 설문지는 분석표본만을 사용한 진단정확률이 보고되었다.

1. 각 체질진단설문도구의 체질별 진단정확률

사상체질변증 기초설문지(1)는 한 건의 연구⁶가 포함되었다. 사상체질변증 기초설문지(1)는 원래 점수 계산법으로 개발되었으나 해당 연구는 편별분석을 적용한 결과를 보고하고 있었으며 정확예측률은 57.9%로 계산되었다. 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.56(0.47-0.65), 소음인 0.63(0.53-0.72), 소양인 0.55(0.45-0.65)로 소음인 민감도가 가장 높았고 소양인 민감도가 가장 낮았다. 특이도에서는 태음인 0.80(0.73-0.85), 소음인 0.79(0.73-0.84), 소양인 0.79(0.73-0.84)로 비슷한 수치를 보였다.

QSCC는 네 건의 연구^{7,19,20,21}가 포함되었으며 정확예측률은 연구에 따라 54.9~100%로 계산되었다. 통합한 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.52(0.47-0.57), 소음인 0.61(0.55-0.67), 소양인 0.62(0.53-0.70)로 소양인 민감도가 가장 높았고 태음



DT, Decision tree; KS-15, Korea Sasang Constitutional Diagnostic Questionnaire; QSCC, Questionnaire for the Sasang Constitution Classification; SDQ_S, Sasangin Diagnostic Questionnaire for Specialist

Figure 1. Study flow diagram: result of the search

인 민감도가 가장 낮았다. 특이도에서는 태음인 0.80 (0.76-0.83), 소음인 0.79(0.76-0.83), 소양인 0.81(0.78-0.84)로 소음인에서 상대적으로 낮은 수치를 보였다.

QSCCII는 세 건의 연구^{8,22,23}가 포함되었으며 정확 예측률은 연구에 따라 50.9~66.0%로 계산되었다. 통합한 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.52(0.46-0.57), 소음인 0.57(0.50-0.63), 소양인 0.62(0.55-0.67)로 소양인 민감도가 가장 높았고 태음인 민감도가 가장 낮았다. 특이도에서는 태음인 0.90 (0.87-0.93), 소음인 0.83(0.80-0.86), 소양인 0.70(0.67-0.74)로 태음인 특이도가 가장 높은 반면 소양인 특이도가 가장 낮았다.

QSCCII 문항 중 16문항 만을 선택하여 의사결정나무법을 사용한 의사결정나무법에 근거한 16문항 사상 체질진단검사지는 한 건의 연구¹¹가 포함되었으며 정확예측률은 55.4%로 계산되었다. 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.67(0.56-0.76), 소음인 0.63(0.54-0.72), 소양인 0.31(0.21-0.42)로 소양인 민감도가 현저히 낮았다. 특이도에서는 태음인 0.72 (0.66-0.79), 소음인 0.76(0.69-0.82), 소양인 0.84(0.78-0.89)로 소양인 특이도가 가장 높은 반면 태음인 특이도가 가장 낮았다.

QSCCII의 문항을 축소한 QSCCII+는 한 건의 연구¹⁰가 포함되었으며 정확예측률은 58.4%로 계산되었다.

체질별 진단정확률을 살펴보면 민감도에서 태음인 0.68(0.61-0.73), 소음인 0.77(0.68-0.84), 소양인 0.41(0.35-0.48)로 소양인 민감도가 현저히 낮았다. 특이도에서는 태음인 0.77(0.72-0.81), 소음인 0.76(0.71-0.79), 소양인 0.87(0.83-0.90)로 소양인 특이도가 가장 높은 반면 소음인 특이도가 가장 낮았다.

QSCCⅢ는 한 건의 연구⁹가 포함되었으며 정확예측률은 73.3%로 계산되었다. 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.67(0.45-0.84), 소음인 0.82(0.67-0.93), 소양인 0.64(0.31-0.89)로 소음인 민감도가 가장 높았고 소양인 민감도가 가장 낮았다. 특이도에서는 태음인 0.90(0.79-0.97), 소음인 0.78(0.61-0.90), 소양인 0.89(0.79-0.96)로 태음인 특이도가 가장 높은 반면 소음인 특이도가 가장 낮았다.

SSCQ는 네 건의 연구^{12,24,25,26}가 포함되었으며 정확예측률은 연구에 따라 65.7~89.2%로 계산되었다. 통합한 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.69(0.67-0.71), 소음인 0.71(0.69-0.74), 소양인 0.63(0.61-0.66)로 소음인 민감도가 가장 높았고 소양인 민감도가 가장 낮았다. 특이도에서는 태음인 0.87(0.86-0.88), 소음인 0.87(0.86-0.88), 소양인 0.85(0.84-0.86)로 비슷한 수준이었지만 소양인 특이도가 가장 낮았다.

TS-QSCD는 두 건의 연구^{13,27}가 포함되었다. TS-QSCD는 점수계산법으로 개발되었으나 한 건의 연구는 관별분석을 적용한 결과를 보고하고 있었으며, 정확예측률은 연구에 따라 각각 76.2%, 74.9%로 계산되었다. 통합한 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.81(0.76-0.86), 소음인 0.71(0.61-0.79), 소양인 0.68(0.60-0.76)로 태음인 민감도가 가장 높았고 소양인 민감도가 가장 낮았다. 특이도에서는 태음인 0.76(0.70-0.82), 소음인 0.93(0.90-0.96), 소양인 0.90(0.86-0.93)로 소음인 특이도가 가장 높은 반면 태음인 특이도가 가장 낮았다.

KS-15는 한 건의 연구¹⁴가 포함되었으며 정확예측률은 63.1%로 계산되었다. 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.63(0.54-0.71), 소음인 0.83

(0.65-0.94), 소양인 0.52(0.38-0.66)로 소음인 민감도가 가장 높았고 소양인 민감도가 가장 낮았다. 특이도에서는 태음인 0.85(0.75-0.91), 소음인 0.81(0.75-0.87), 소양인 0.80(0.74-0.86)로 태음인 특이도가 가장 높은 반면 소양인 특이도가 가장 낮았다.

SDQ_S는 한 건의 연구¹⁵가 포함되었으며 정확예측률은 74.3%로 계산되었다. 체질별 진단정확률을 살펴보면 민감도에서 태음인 0.77(0.61-0.88), 소음인 0.76(0.56-0.90), 소양인 0.70(0.51-0.84)로 소양인 민감도가 가장 낮았다. 특이도에서는 태음인 0.90(0.80-0.96), 소음인 0.82(0.71-0.90), 소양인 0.90(0.81-0.96)로 소음인 특이도가 가장 낮았다.

그리고 추가적으로 분석표본을 사용한 연구들의 각 체질간 민감도, 특이도 값의 차이는 검증표본을 사용한 연구들에 비해 그 차이가 상대적으로 작게 나타났다. (Table 1)

2. 체질진단설문도구간 체질별 진단정확률 비교

전술하였듯 사상체질변증 기초설문지(1), QSCC, SSCQ, TS-QSCD 등 4종의 설문지는 분석표본을 사용한 진단정확률만이 보고되었다. 분석표본을 사용한 진단정확률은 해당 연구집단에 최적화되어 있고 외적타당도를 담보하기 어려워 검증표본을 사용한 연구들과 직접 비교하기에는 어려움이 따른다. 이에 체질진단설문도구간 비교에서는 위 4종 설문지는 제외하였다.

설문지간 진단정확률을 비교해보았을 때, 민감도 측면에서는 태음인의 경우 SDQ_S가 가장 높은 값을 보인 반면 QSCCⅡ가 가장 낮은 값을 보였고, 소음인의 경우 KS-15가 가장 높은 값을 보인 반면 QSCCⅡ가 가장 낮은 값을 보였으며, 소양인의 경우 SDQ_S가 가장 높은 값을 보인 반면 의사결정나무법에 근거한 16문항 사상체질진단검사지가 가장 낮은 값을 보였다. 특이도 측면에서는 태음인의 경우 QSCCⅡ, QSCCⅢ, SDQ_S가 비슷하게 높은 값을 보인 반면 의사결정

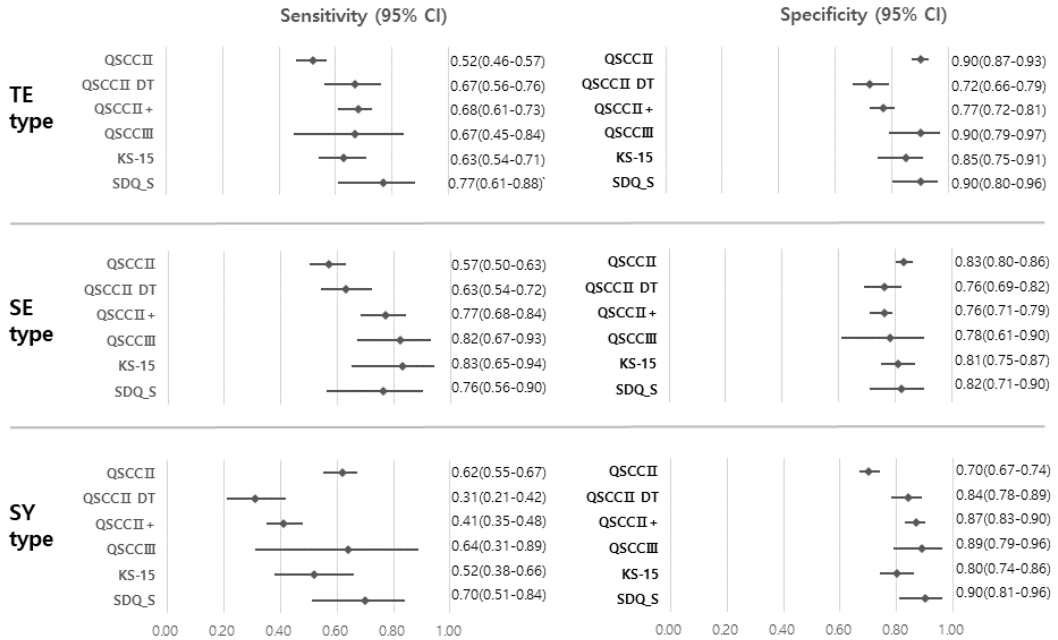
Table 1. Summary of Diagnostic Accuracy (Sensitivity and Specificity) in Each Questionnaire for Sasang Constitution Diagnosis

Type of Questionnaire	No. of studies (participants)	Used sample	PCP(%)	Target constitution	Sensitivity (95% CI)	Specificity (95% CI)
Questionnaire(1)	1 (n=328) ⁶	Training group	57.9	TE type SE type SY type	0.56(0.47-0.65) 0.63(0.53-0.72) 0.55(0.45-0.65)	0.80(0.73-0.85) 0.79(0.73-0.84) 0.79(0.73-0.84)
QSCC	4 (n=810) ^{7,19,20,21}	Training group	55.6 ⁷ 58.2 ¹⁹ 100 ²⁰ 54.9 ²¹	TE type SE type SY type	0.52(0.47-0.57) 0.61(0.55-0.67) 0.62(0.53-0.70)	0.80(0.76-0.83) 0.79(0.76-0.83) 0.81(0.78-0.84)
QSCC II	3 (n=898) ^{8,22,23}	Test group	66.0 ⁸ 55.8 ²² 50.9 ²³	TE type SE type SY type	0.52(0.46-0.57) 0.57(0.50-0.63) 0.62(0.55-0.67)	0.90(0.87-0.93) 0.83(0.80-0.86) 0.70(0.67-0.74)
QSCC II (Decision tree)	1 (n=285) ¹¹	Test group	55.4	TE type SE type SY type	0.67(0.56-0.76) 0.63(0.54-0.72) 0.31(0.21-0.42)	0.72(0.66-0.79) 0.76(0.69-0.82) 0.84(0.78-0.89)
QSCC II ⁺	1 (n=587) ¹⁰	Test group	58.4	TE type SE type SY type	0.68(0.61-0.73) 0.77(0.68-0.84) 0.41(0.35-0.48)	0.77(0.72-0.81) 0.76(0.71-0.79) 0.87(0.83-0.90)
QSCC III	1 (n=75) ⁹	Test group	73.3	TE type SE type SY type	0.67(0.45-0.84) 0.82(0.67-0.93) 0.64(0.31-0.89)	0.90(0.79-0.97) 0.78(0.61-0.90) 0.89(0.79-0.96)
SSCQ	4 (n=5412) ^{12,24,25,26}	Training group	77.3 ¹² 65.7 ²⁴ 66.9 ²⁵ 89.2 ²⁶	TE type SE type SY type	0.69(0.67-0.71) 0.71(0.69-0.74) 0.63(0.61-0.66)	0.87(0.86-0.88) 0.87(0.86-0.88) 0.85(0.84-0.86)
TS-QSCD	2 (n=494) ^{13,27}	Training group	76.2 ¹³ 74.9 ²⁷	TE type SE type SY type	0.81(0.76-0.86) 0.71(0.61-0.79) 0.68(0.60-0.76)	0.76(0.70-0.82) 0.93(0.90-0.96) 0.90(0.86-0.93)
KS-15	1 (n=222) ¹⁴	Test group	63.1	TE type SE type SY type	0.63(0.54-0.71) 0.83(0.65-0.94) 0.52(0.38-0.66)	0.85(0.75-0.91) 0.81(0.75-0.87) 0.80(0.74-0.86)
SDQ_S	1 (n=105) ¹⁵	Test group	74.3	TE type SE type SY type	0.77(0.61-0.88) 0.76(0.56-0.90) 0.70(0.51-0.84)	0.90(0.80-0.96) 0.82(0.71-0.90) 0.90(0.81-0.96)

KS-15, Korea Sasang Constitutional Diagnostic Questionnaire; PCP, Percent Correctly Predicted; QSCC, Questionnaire for the Sasang Constitution Classification; SDQ_S, Sasangin Diagnostic Questionnaire for Specialist; SE, Soeum; SSCQ, Sasang Constitution Questionnaire for Doctors and Patients; SY, Soyang; TE, Taeum; TS-QSCD, Two Step Questionnaire for the Sasang Constitution Diagnosis

나무법에 근거한 16문항 사상체질진단검사가 가장 낮은 값을 보였고, 소음인의 경우 QSCC II가 가장 높은 값을 보인 반면 의사결정나무법에 근거한 16문항 사상체질진단검사가, QSCC II⁺가 비슷하게 가장 낮은 값을 보였으며, 소양인의 경우 SDQ_S가 가장 높은 값을 보인 반면 QSCC II가 가장 낮은 값을 보였다.

(Figure 2)



DT, Decision tree; KS-15, Korea Sasang Constitutional Diagnostic Questionnaire; QSCC, Questionnaire for the Sasang Constitution Classification; SDQ_S, Sasangin Diagnostic Questionnaire for Specialist; SE, Soeum; SY, Soyang; TE, Taeum

Figure 2. Comparison of sensitivity and specificity in each questionnaire for Sasang constitution diagnosis

IV. 고찰 및 결론

정확예측률은 체질진단설문도구 연구들에서 진단 정확률을 제시하기 위한 지표로 많이 사용되어 왔다. 그러나 정확예측률은 진단정확률을 숫자 하나로 표시하는 간편함이 있지만 그 외의 정보들을 표현하지 못하는 한계점 또한 존재한다. 예컨대 모든 체질에서 고른 정확률을 보이는 설문지와, 한 체질에서는 극단적으로 높은 정확률을 보이는 반면 다른 체질에서는 극단적으로 낮은 정확률을 보이는 설문지가 동일한 정확예측률 값을 가질 때 두 설문지가 동일한 타당도를 가진다고 평가할 수 있는나는 문제가 발생할 수 있는 것이다. 또한 그 동안 체질진단 정확률 연구들은 진단정확률을 제시하는데 있어 정확예측률 이외에도 일치율, 정확률, 감별정확률, 평균진단정확률, 정판별력, 오분류율 등의 용어 및 개념을 혼용하고 있어 일관된 지표로 진단정확률을 비교·평가하기 어렵기도 하

였다.

이에 본 연구는 체질진단설문도구의 진단정확률을 진단검사의 유효성을 평가하는 일반적 평가 지표인 민감도와 특이도를 이용한 체질유형별 민감도, 특이도로 재수량화하여 비교함으로써 체질별 진단정확률 차이를 확인하고 평가하고자 하였다. 이를 통해서 각 체질진단설문도구의 장단점과 활용시 주의점을 확인할 수 있을 것이기 때문이다.

먼저 사상체질변증 기초설문지(1), QSCC, SSCQ, TS-QSCD 등 4종의 설문지는 분석표본을 사용한 진단 정확률만이 보고되었다. 이는 분석표본을 통해 도출된 판별식을 다시 동일한 집단에 적용하여 나온 진단 정확률로, 모델 자체가 해당 집단에 최적화되어 있어 검증표본을 사용한 연구에 비해 일반적으로 높은 값을 보이며 외적타당도를 담보하기 어렵다는 문제점이 있다. 이에 이후 논의에서는 위 4종 설문지는 제외하였다.

1. 각 체질진단설문도구의 장단점

검증표본을 사용한 QSCC II, 의사결정나무법에 근거한 16문항 사상체질진단검사지, QSCC II*, QSCC III, KS-15, SDQ_S 등 6종의 설문지는 서로 다른 체질별 민감도, 특이도 값을 보여 각각 다른 장단점이 있음을 알 수 있었다.

태음인에서는 SDQ_S가 가장 높은 민감도를 보였으며 QSCC II, QSCC III, SDQ_S가 비슷하게 높은 특이도를 보였다. 소음인에서는 KS-15가 가장 높은 민감도를 보였으며 QSCC II가 가장 높은 특이도를 보였다. 소양인에서는 SDQ_S가 가장 높은 민감도와 특이도를 보였다. 즉 태음인의 경우 SDQ_S를 선별검사(screening test)로 사용하고 QSCC II, QSCC III, SDQ_S 등을 확진검사(diagnostic test)로 사용할 경우 보다 나은 진단율을 보일 수 있을 것이다. 마찬가지로 소음인의 경우 KS-15를 선별검사로 사용하고 QSCC II를 확진검사로 사용하여 보다 높은 진단율을 보일 수 있을 것이며, 소양인의 경우 SDQ_S가 가장 좋은 선택지가 될 것이다.

반대로 태음인 민감도가 낮은 QSCC II를 태음인 선별검사에 사용하거나 태음인 특이도가 낮은 의사결정나무법에 근거한 16문항 사상체질진단검사지를 태음인 확진검사에 사용하는 것은 좋지 않은 선택이 될 것이다. 소음인 민감도가 낮은 QSCC II를 소음인 선별검사에 사용하거나 소음인 특이도가 낮은 의사결정나무법에 근거한 16문항 사상체질진단검사지, QSCC II*를 사용하는 것 역시 좋지 않은 선택이 될 것이다. 마찬가지로 소양인 민감도가 낮은 의사결정나무법에 근거한 16문항 사상체질진단검사지를 소양인 선별검사에 사용하거나 소양인 특이도가 낮은 QSCC II를 소양인 확진검사에 사용하는 것은 좋지 않은 선택일 것이다.

이처럼 동일한 설문도구라 하더라도 진단하고자 하는 체질에 따라 장단점이 다름을 확인할 수 있었다. 그러나 임상 현장에서 동일한 환자에게 두 종류 이상의 설문지를 사용하는 것은 현실적으로 어려우며, 실

제적으로 체질진단설문도구는 전문가 진료 전 선별검사의 의미로 시행되는 점을 고려하면 민감도 측면의 고려가 더 중요하다고 할 것이다.

2. 체질별 민감도 차이

검증표본을 사용한 연구들의 체질별 민감도 값에 초점을 맞추면, 태음인과 소양인 진단에 있어서는 SDQ_S가 우선적으로 추천되며 소음인 진단에 있어서는 KS-15가 우선적으로 추천된다는 것을 확인할 수 있었다.

그리고 동시에 QSCC II를 제외한 체질진단설문도구 모두에서 소양인 민감도가 다른 두 체질의 민감도에 비해 낮음을 확인할 수 있었다. 선행 체계적 문헌고찰 연구⁴에서는 QSCC II의 소양인 민감도도 낮게 평가하고 있는데, 이는 QSCC II*를 QSCC II의 범주 안에 같이 병합하여 도출된 결과였다. QSCC II*는, 의사결정나무법에 근거한 16문항 사상체질진단검사지와 같이, 특히 낮은 소양인 민감도를 보이고 있었기 때문이다. 이처럼 낮은 소양인 민감도는 『東醫壽世保元』 『四象人辨證論』의 소양인 진단이 가장 쉽다¹⁾는 서술과 차이가 있다.

이처럼 낮은 소양인 진단률과 관련하여 타 체질에 비해 소양인의 척도가 안정적이지 않고 검도가 필요하다고 지적한 선행 연구들^{28,29,30}이 있었다. 체질진단을 위한 척도 중 소양인 척도가 적어 추가 발굴이 필요하고, 물리적인 변수량으로 표현하기 어려운 소양인 특징이 있다고도 하였다³¹. 하지만 소양인의 민감도가 전반적으로 낮게 나타나는 것은 개별 소양인 척도가 가지는 문제점 외에 다른 부분에서도 그 원인을 찾아볼 수 있었다.

먼저, 체질진단설문도구에서 주로 적용하고 있는 통계분석 방법에 대해 고려해볼 필요가 있다. 의사결정나무법에 근거한 16문항 사상체질진단검사지와 다중 로지스틱 회귀분석을 사용한 KS-15를 제외한 설문

1) 少陽人 體形 上盛下虛 胸實足輕 剽銳好勇 而人數亦多 四象人中 最爲易辨

도구들은 모두 판별분석을 사용하고 있다. 판별분석은 새로운 사례가 이미 정의된 집단 분류 어디에 속하는지 구분하는데 사용될 수 있으며 의학적 진단 분야에서도 활용되고 있다³². 그러나 특정 집단이 절대적 다수를 차지하는 상황(집단 불균형, class imbalance)에서는 전체 정확도를 높이기 위해 주로 다수 집단으로 과도하게 예측하는 편향(다수 집단 편향, majority class bias)이 나타날 수 있다³³. 즉, 인구집단에서 태음인이 다수인 것이 예상되고³⁴ 연구대상 집단에서도 태음인이 다수인 상황에서 정확예측률만으로 성능을 평가한다면 다른 체질에 비해 태음인을 더 많이 예측하게 되는 것이다. 그리고 이 조건은 소음인과 소양인 예측에 있어 다르게 작용한다. 체형이나 식사량, 식욕, 땀 등 다수의 진단 지표에서 소음인은 태음인과 통계적 차이가 큰 정반대 상에 위치하는 반면 소양인은 통계적 차이가 없거나 적은, 태음인과 소음인의 중간에 위치하는 경우가 많다^{5,17,35}. 태음인의 정반대 상으로서 소음인의 예측률은 동반 상승할 여지가 있지만, 중간 상으로서 비태음, 비소음의 배제적 진단²⁾이 이루어지는 소양인의 경우 도리어 예측률이 하락하게 되는 것이다. QSCC II가 개발된 이래 QSCC II보다 더 나은 정확예측률을 목적으로 하는 연구들이 반복되면서 이러한 경향이 강화되어 소양인 민감도가 전반적으로 낮아진 것으로 추정된다.

다음으로는, 설문도구 외적으로 다른 체질과 비교하여 소양인의 전문가 진단 일치도가 상대적으로 낮다는 것, 다시 말해 전문가 사이에 공통된 소양인의 상이 적다는 점이다. 전문가간 체질 진단 일치도에 대해서는 33.5%에서 90% 이상까지 다양한 보고가 있지만³⁶, 체질별 진단 일치도를 보고한 연구들에서 소양인의 진단 일치도가 더 낮게 나타나고 있음을 확인할 수 있었다. 장의 연구³⁷는 전문가 6인을 3인씩 두 그룹으로 나누고 동일한 체질 증례 자료를 주어 체질 진단을 수행하게 하였을 때, 결과가 일치한 증례의

체질 일치율은 태음인, 소음인이 각각 38%, 37%였던 반면 소양인은 23%로 낮았으며 반대로 불일치율은 소양인이 36%로 가장 높았다고 보고하며 전문가 사이의 견해가 체질별로 차이가 난다고 해석하였다. 백의 연구³⁸에서도 전문가 3인을 대상으로 동일한 체질 증례 자료를 주어 독립적으로 체질진단을 수행하였을 때, 전문가 3인 모두 소양인에서 가장 낮은 양성예측도를 보였으며 다수 전문가의 체질진단이 일치하는 경우에도 소양인의 양성예측도가 낮다고 보고하면서 전문가들의 소양인 진단경향이 매우 다르다는 것을 의미한다고 해석하고 있다. 이처럼 전문가 사이에 공통된 소양인의 상이 적다는 점은 통계적으로 유의미한 소양인 척도의 발굴을 어렵게 하는 요인이 되며, 전문가간 소양인 진단 일치도가 낮다는 점은 체질진단 설문도구의 타당성 연구에 있어 소양인의 참조표준(reference standard) 자체가 안정적이지 않게 되는 요인이 된다.

그리고 적절한 소양인 척도의 부족과 체질진단 설문도구의 낮은 소양인 민감도는 후에 전문가 집단에서 소양인 상 구축을 위한 재학습 과정 자체를 어렵게 함으로써 전문가 집단 공통의 소양인 상 형성을 더욱 어렵게 한다. 이를 도식화하면 Figure 3과 같다.

객관적 체질진단을 위한 도구로서 다양한 체질진단 설문도구들이 연구 및 개발되어 왔다. 그러나 대부분 정확예측률만을 사용해 도구의 타당도를 평가해온 것이 사실이다. 이에 본 연구는 체질진단 설문도구의 체질유형별 민감도, 특이도를 분석하여 각 도구의 타당성 정보를 다시 비교, 평가하였다. 실제 임상 현장에서 체질진단 설문도구가 예진 과정의 선별검사로 시행되는 경우가 많다는 점을 고려했을 때, 태음인과 소양인 진단에서는 SDQ-S의 사용이 더 나은 성능을 기대할 수 있었으며 소음인 진단에서는 KS-15의 사용이 더 나은 성능을 기대할 수 있었다. 그리고 체질진단 설문도구의 체질유형별 민감도를 비교하는 과정에서 비교적 일관된 체질간 민감도 차이, 즉 소양인 진단의 낮은 민감도를 확인할 수 있었다. 그 원인으로는, 기존의 연구들에서 지적한 적절한 소양인 척도의 부족 외

2) 실례로 한 연구¹⁵에서 제시한 체질판별함수를 살펴보면 연구에서 상정한 소양척도는 사용되지 않고 태음척도와 소음척도만이 사용되고 있음을 확인할 수 있었다.

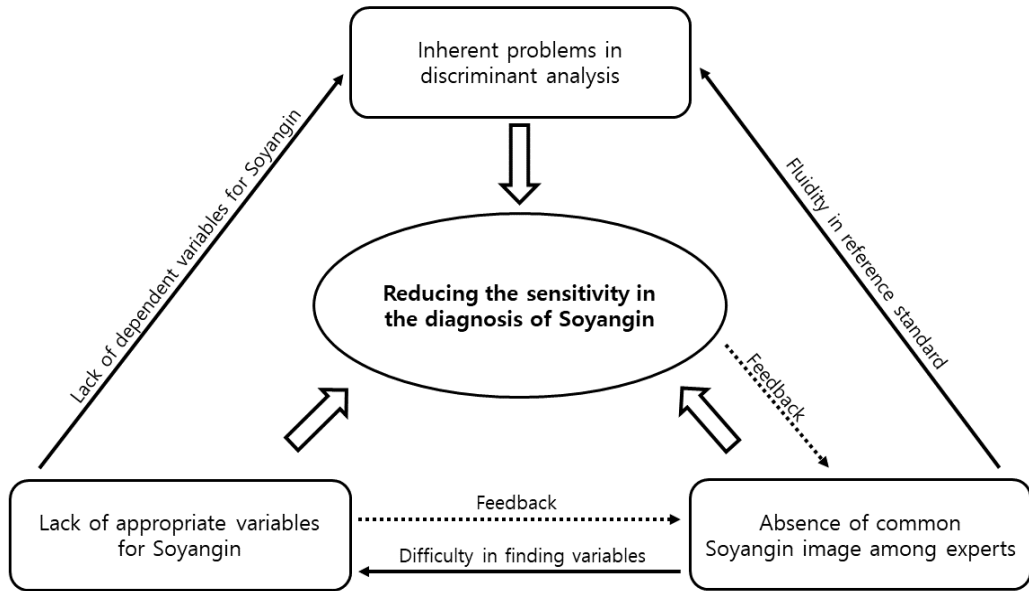


Figure 3. Pathways reducing the sensitivity in the diagnosis of Soyangin

에도, 체질진단설문도구 연구에서 주로 사용하고 있는 판별분석의 내재적 문제점과 전문가간 공통의 소양인 상이 적다는 점을 추론할 수 있었다.

그러나 본 연구는 몇 가지 측면에서 한계점을 지닌다. 첫째, 본 연구는 체계적 문헌고찰(Systematic Review)로 설계되지 않았다. 스노우볼링 방식을 통해 사전 선정한 체질진단설문도구가 사용된 연구들을 최대한 검색하고자 하였으나 일부 누락된 연구가 있을 수 있다. 또한 각 연구들의 질 평가(quality assessment)가 수행되지 않았으며 연구대상자, 참조표준으로서 사용한 체질진단방법, 연구의 질 수준 등 이질성(heterogeneity)을 고려하고 있지 않아 결과의 해석에는 제한이 따른다. 둘째, 체질진단설문도구간 진단정확률을 비교한 6종의 설문지는 QSCC II를 제외하고는 모두 한 건의 진단정확률 연구만을 포함하고 있었다. 검증표본을 사용하고 있어 어느정도의 외적타당도를 기대할 수 있으나 실제 임상현장에서의 성능을 반영하고 있는지는 확인하기 어렵다. 셋째, 분석을 위해 태음인, 소음인, 소양인 세 체질로 구성된 체질별 3X3 표를 특정

체질과 비특정 체질의 2X2 표로 재구축하는 과정에서 참양성 값에 비해 참음성 값이 증가하게 됨으로써 전반적으로 민감도에 비해 높은 특이도 값을 얻게 되었다. 이는 기존 데이터를 참양성·거짓양성·참음성·거짓음성으로 재구축하는 과정에서 발생한 결과이기 때문에 체질진단설문도구 전반이 실제로 민감도에 비해 특이도가 높다고 결론 내리기에는 제한이 있다.

V. 감사의 말씀

이 연구는 2025년도 한국한의학연구원의 ‘빅데이터 기반 한의 예방 치료 원천기술 개발(KSN1739121) 과제’의 지원을 받아 수행되었습니다.

VI. 참고문헌

1. Kim SH, Jang ES, Koh BH. The Methodological Review on the Accuracy Study of Questionnaire for Sasang Constitution Diagnosis. *J Sasang Constitut Med.* 2012;24(3):1-16. (Korean) DOI: 10.7730/JSCM.2012.24.3.1
2. Lee SW, Jang ES, Lee J, Kim JY. Current researches on the methods of diagnosing sasang constitution: an overview. *eCAM* 2009;6(S1):43-9. DOI: 10.1093/ecam/nep092
3. SJ Lee, MG Kim, H Chae. Multi-facet Analysis on Validity of Sasang Type Diagnostic Test. *J Korean Oriental Med.* 2008;29(1):7-14. (Korean)
4. SH Kim, Y Lee, BH Koh, E Jang. Assessing the diagnostic accuracy of the Questionnaire for Sasang Constitutional Classification II (QSCC II): A systematic review. *EuJIM* 2013;5(5):393-398. DOI: 10.1016/j.eujim.2013.05.003
5. Kim SH. Revisit of the Digestion-related Items for Diagnosing Soeumin. *J Sasang Constitut Med.* 2022;34(2):15-35. (Korean) DOI: 10.7730/JSCM.2022.34.2.15
6. Lee EJ, Koh BH, Song IB. The Study on the Validity of "Questionnaire of Sasang Constitution Classification(1)". *J Sasang Constitut Med.* 1995; 7(2):89-100. (Korean)
7. Kim SH, Koh BH, Song IB. A Validation Study of Questionnaire for Sasang Constitutional Classification(QSCC). *J Sasang Constitut Med.* 1993; 5(1):67-85. (Korean)
8. Jung HJ, Chou LS, Yoo JY, Shin DE, Han CH. Comparison study on Constitution Classification between QSCC II and Clinical Research. *J Korean Oriental Med.* 2000;21(4):73-83. (Korean)
9. Kim YW, Kim JW. The Study in Objectification of the diagnosis of Sasang Constitution(According to Analysis of the Past Questionnaires). *J Sasang Constitut Med.* 1999;11(2):151-184. (Korean)
10. Choi KJ, Choi YS, Cha JH, Hwang MW, Lee SK, Koh BH, et al. A Study on the Reliability and Validity test of the QSCC II* (Revised Questionnaire for the Sasang Constitution Classification). *J Sasang Constitut Med.* 2006;18(1):62-74. (Korean)
11. Park SS, Choi JY. The Characteristics of Questionnaire Response Using Decision Tree Method. *J Sasang Constitut Med.* 2003;15(3):177-186. (Korean)
12. Seul YK. The study on sasang constitutional discriminant function through integration study of Questionnaire for doctors and patients. Doctor thesis. Dongeui univ. 2006. (Korean)
13. Kim YW, Shin DY, Kim JH, Choi DS, Lim MK, Lee KL, et al. A Development of the Two Step Questionnaire for the Sasang Constitution Diagnosis (TS-QSCD). *J Sasang Constitut Med.* 2006;18(1): 75-90. (Korean)
14. Baek YH, Jang ES, Park KH, Yoo JH, Jin HJ, Lee S. Development and Validation of Brief KS-15 (Korea Sasang Constitutional Diagnostic Questionnaire) Based on Body Shape, Temperament and Symptoms. *J Sasang Constitut Med.* 2015; 27(2):211-221. (Korean) DOI: 10.7730/JSCM.2015.27.2.211
15. Pak YS, Kim JS, Park BJ, Yoo JH, Lee JH, Koh BH, Lee EJ. The Reliability and Validity Test of Sasangin Diagnostic Questionnaire for Specialist (SDQ_S). *J Sasang Constitut Med.* 2016;28(2): 110-122. (Korean) DOI: 10.7730/JSCM.2016.28.2.110
16. Kim JY, Kim KY, Song JM. A Study on the Objectification of Sasang Constitution Diagnosis Using Constitutional Discrimination Table. *J Sasang Constitut Med.* 1998;10(1):181-214. (Korean)

17. Kim SH. Reconsideration of Body Mass Index for Diagnosing Sasang Constitution. *J Physiol & Pathol Korean Med.* 2024;38(4):81-88. (Korean) DOI: 10.15188/kjopp.2024.08.38.4.81
18. Plana, M.N., Arevalo-Rodriguez, I., Fernández-García, S. et al. Meta-DiSc 2.0: a web application for meta-analysis of diagnostic test accuracy data. *BMC Med Res Methodol* 2022;22:306. DOI: 10.1186/s12874-022-01788-2
19. Koh BH. The Study of Analysis with Sixteen Personality Factor in Sasang Constitution. *J Korean Oriental Med.* 1993;14(2):64-76. (Korean)
20. Lee JC, Koh BH, Song IB. The Validation Study of the Questionnaire of Sasang Constitution Classification: Comparative Analysis with Sixteen Personality Factor Questionnaire (16PF). *J Sasang Constitut Med.* 1993;5(1):87-104. (Korean)
21. Ro SH, Lee EJ, Hong SC, Koh BH, Song IB. The Study on Validity of "Questionnaire of Sasang Constitution Classification". *J Sasang Constitut Med.* 1996;8(2):239-253. (Korean)
22. Kim KK, Kim JW, Lee EJ, Kim JY, Choi SM. Study on Classification Function into Sasang Constitution Using Data Mining Techniques. *Korean J Oriental Physiol Pathol* 2004;18(6):1938-44. (Korean)
23. Park HS, Ju JC, Kim JH, Kim KY. A Study on Clinical Application of the QSCC II (Questionnaire for the Sasang Constitution Classification II). *J Sasang Constitut Med.* 2002;14(2):35-44. (Korean)
24. Beum SG. Study on development of classification model and implementation for diagnosis system of Sasang constitution. Doctor thesis. Donggeui univ. 2008. (Korean)
25. Jeon SH, Jeong SI, Kwon SD, Park SJ, Kim KK, Kim JW. A Study on the Diagnostic Accuracy Rate of the Sasang Constitution Questionnaire for Doctors Used in the On-line System. *J Sasang Constitut Med.* 2008;20(3):82-93. (Korean)
26. Kim JW, Jeong SI, Kim KG, Lee EJ, Kim JY, Lee YT. Validity Study of the Questionnaire for Doctors for Sasang Constitution Diagnosis. *Korean J Oriental Physiol Pathol.* 2006;20(1):209-214. (Korean)
27. Shin DY, Song JM. A Study on the Standardization of TS-QSCD. *J Sasang Constitut Med.* 2009; 21(1):99-126. (Korean)
28. Kim SB, Lee JH, Park GS, Jung YJ, Lee SK, Song IB. A Study on responses to the questionnaire based on of revised Sasang Constitution Classification II (QSCC II⁺). *J Sasang Constitut Med.* 2001; 13(3):15-23. (Korean)
29. Jeong YJ, Kim SB, Lee SK, Lee EJ, Choi SM, Koh BH, Song IB. Revision of the Questionnaire for Sasang Constitution Classification. *J Sasang Constitut Med.* 2003;15(3):100-117. (Korean)
30. Bae KH, Kim SH, Go HY, Park KH, Lee SW, Lee SJ. One year test-retest reliability of the Korea Sasang constitutional diagnostic questionnaire (KS-15) in university students. *J Sasang Constitut Med.* 2019;31(2):12-21. (Korean) DOI: 10.7730/JSCM.2019.31.2.12
31. Jang ES, Park KH, Baek YH, Lee SW, Kim SH, Joo JH. A Development of Sasang Constitution Diagnosis Program Based on Characteristics of Body Shape, Face, Physiological Symptom and Personality. *J Sasang Constitut Med.* 2012;24(1): 21-31. (Korean)
32. Shariatnia, S., Ziaratban, M., Rajabi, A. et al. Modeling the diagnosis of coronary artery disease by discriminant analysis and logistic regression: a cross-sectional study. *BMC Med Inform Decis Mak.* 2022;22:85. DOI: 10.1186/s12911-022-01

- 823-8
33. Xie J, Qiu Z. The effect of imbalanced data sets on LDA: A theoretical and empirical analysis. *Pattern Recognit.* 2007;40(2):557-562. DOI: 10.1016/j.patcog.2006.01.009
34. Jin HJ, Baek Y, Kim HS, Ryu J, Lee S. Constitutional multicenter bank linked to Sasang constitutional phenotypic data. *BMC Complement Altern Med.* 2015;15:46. DOI: 10.1186/s12906-015-0553-3
35. Jang ES, Kim HS, Lee HJ, Baek YH, Lee SW. The Clinical Study on the Ordinary and Pathological Symptoms according to Sasang Constitution. *J Sasang Constitut Med.* 2007;19(3):144-155. (Korean)
36. Han EK, Kwon YK. Inter-expert Agreement and Diagnostic Accuracy of Sasang Constitution Medicine. *J Physiol & Pathol Korean Med* 2018; 32(4):185-196. (Korean) DOI: 10.15188/kjopp.2018.08.32.4.185
37. Jang ES, Kim HS, Yoo JH, Kim SH, Baek YH, Lee SW. The Verification of Concordance Coefficient in Sasang Constitutional Factors by Expert. *J Sasang Constitut Med.* 2009;21(2):79-86. (Korean)
38. Baek YH, Kim HS, Lee SW, Jang ES. The Concordance and Validity Assessment of Diagnosis for the Expert in Sasang Constitution. *J Sasang Constitut Med.* 2014;26(3):295-303. (Korean) DOI: 10.7730/JSCM.2014.26.3.295

<저자들의 소속 및 지위>

김상혁/한국한의학연구원 한의약데이터부/
책임연구원