

설진 영상지표를 활용한 경도 신기능 저하의 예측 가능성

이지윤* · 한경선* · 박정환[†] · 이상훈[†] · 김호준*

동국대학교 일산불교한방병원 한방재활의학과*, 한국한의학연구원 한의약데이터부[†]

Predicting Mild Renal Function Decline Using Tongue Diagnosis Indicators: A Logistic and Machine-Learning Analysis

Jiyeon Lee, K.M.D.*, Kyungsun Han, K.M.D., Ph.D.*, Jeong Hwan Park, K.M.D.[†], Sanghun Lee, K.M.D.[†], Hojun Kim, K.M.D., Ph.D.*

Department of Korean Medicine Rehabilitation, College of Korean Medicine, Dongguk University*, KM Data Division, Korea Institute of Oriental Medicine[†]

This study was supported by the Korea Institute of Oriental Medicine (KIOM) major research project, "Development of an LLM-based integrated clinical decision support artificial intelligence service for Korean medicine big data (KSN2511022)," and the Korean Medicine Innovation Technology Development Program of the Korea Health Industry Development Institute (KHIDI) (RS-2020-KH087713).

RECEIVED December 16, 2025

ACCEPTED December 25, 2025

CORRESPONDING TO

Hojun Kim, Department of Korean Medicine Rehabilitation, College of Korean Medicine, Dongguk University, 27 Dongguk-ro, Ilsandong-gu, Goyang 10326, Korea

TEL (031) 961-9009

FAX (031) 961-9009

E-mail kimklar@dongguk.ac.kr

Copyright © 2026 The Society of Korean Medicine Rehabilitation

Objectives To evaluate whether quantitative tongue diagnosis indicators can predict mild renal function decline and to compare their predictive ability using logistic regression and machine-learning models.

Methods In 2,009 adults who underwent health screening, tongue-image indicators (coating, light-red index, blue-purple index, tooth-mark index), body composition, and metabolic markers were analyzed. estimated glomerular filtration rate (eGFR) was dichotomized (≥ 90 vs. < 90). Logistic regression, random forest, and XGBoost models were used to evaluate predictive performance.

Results The blue-purple index was independently associated with reduced eGFR (odds ratio=3.3). Age and body composition (fat mass, fat percentage) were also significant predictors. Machine-learning models showed modest discrimination (area under the receiver operating characteristic curve 0.64 to 0.70), with the logistic model performing best. Coating and light-red indices showed limited predictive power.

Conclusions Tongue color—particularly the blue-purple index—may serve as a non-invasive indicator related to early renal function decline, although overall predictive power remains moderate. (*J Korean Med Rehabil* 2026;36(1):11-25)

Key words Tongue diagnosis, Kidney function test, Glomerular filtration rate, Machine learning, Logistic regression, Chronic kidney disease

서론»»»»

만성 신장 질환(chronic kidney disease, CKD)은 전 세계적으로 유병률이 증가하고 있으며, 사망률과 심혈관계

합병증 위험을 높이는 주요 만성 질환이다¹⁻³⁾. CKD의 평가는 일반적으로 추정사구체여과율(estimated glomerular filtration rate, eGFR)에 기반하며, 최근 Kidney Disease: Improving Global Outcomes (KDIGO) 2024 가

이드라인에서는 eGFR의 절대 수치보다 G1 (≥ 90 mL/min/1.73m²)에서 G2 (60~89 mL/min/1.73m²)로의 전환과 같은 카테고리 변화 자체가 임상적으로 더욱 중요하다고 강조하고 있다⁴⁾. 이는 정상 범위 내에서도 신기능 감소를 조기에 인지하는 것이 중요함을 의미한다.

한의학의 설진(舌診)은 인체의 전신 상태를 반영하는 주요 진단 도구로, 기혈 · 음양 · 장부 기능의 이상을 간접적으로 파악할 수 있다. 최근에는 디지털 영상 기술과 CIE Lab 색공간 기반 분석을 통해 설태면적, 설질색, 치흔 패턴 등 다양한 설진 지표를 객관적으로 측정할 수 있게 되었으며, 이는 전통 진단의 정량화 및 표준화에 중요한 전환점을 제공한다. 특히 이러한 정량적 설진 지표는 인공지능 기반 예측 시스템과 결합되어, 한의학적 진단의 객관성과 임상 적용 가능성을 높일 수 있다는 점에서 주목받고 있다⁵⁻⁷⁾.

선행연구들에서는 비만⁸⁾, 대사질환^{9,10)}, 순환기능^{11,12)}과의 연관성을 중심으로 설진의 진단적 의의를 제시해왔으며, 예컨대 비만도가 높을수록 설태가 두꺼워지고 설질 색이 변화한다는 연구결과나⁸⁾, 특정 설질 색 패턴이 혈관 기능 저하 또는 대사적 위험도와 관련된다는 보고가 있다¹¹⁾. 특히 일부 연구에서는 설태지수의 증가가 eGFR 감소와 유의하게 관련된다고 제시하면서^{13,14)}, 설진 지표가 신기능 상태를 반영할 수 있다는 가능성이 제기되었다. 그러나 대부분의 기존 연구는 단일 지표에 초점을 맞추거나, 표본 수가 적고, 비만과 신기능을 동시에 고려하지 않아 해석에 제한이 있다.

비만은 CKD의 중요한 위험 인자이나, 비만도와 신기능 및 설진 지표 간의 복합적 상호작용은 아직 충분히 규명되지 않았다. 또한 기존 연구들은 주로 eGFR을 연속형 변수로 모델링하였으나, 실제 임상에서는 KDIGO의 제안처럼 eGFR을 정상군(G1)과 경도 감소 이하군(G2~G5)으로 이분화하여 분류하는 방식이 진단적 실용성 측면에서 더 적합할 수 있다⁴⁾.

최근에는 설진을 단순한 주관적 관찰에 의존하는 전통적 방식에서 벗어나, 디지털 이미지 분석과 인공지능을 기반으로 혀의 설태 · 색상 · 모양 등을 자동화된 시스템으로 정량화하는 연구가 활발히 이루어지고 있다. 예를 들어, 딥러닝 기반 설진 자동화 시스템은 혀 이미지의 분할, 색상 보정, 특징 추출 등을 통해 설질의 변화를 정밀하게 분석하며, 일부 연구에서는 90% 이상의

질병 분류 정확도를 기록하기도 하였다^{5,6)}. 이는 한의학적 설진의 객관화뿐 아니라, AI 기반 예측 시스템의 임상적 활용 가능성을 뒷받침하는 사례이다.

본 연구는 2020~2024년의 건강검진 자료를 활용한 2,000명 이상의 대규모 데이터를 기반으로, 설진 지표가 체질량지수(body mass index, BMI) 및 eGFR과 어떤 연관성을 갖는지를 분석하고, 설진 지표를 이용한 신기능 저하(G1 vs. G2 이하) 예측 가능성을 평가하였다. 또한, 연속형 eGFR 모델과 범주형 eGFR 모델의 성능을 비교함으로써, 설진 기반 예측의 임상적 활용 가능성을 탐색하고자 한다.

대상 및 방법»»»»

1. 연구대상

본 연구는 2020년 1월부터 2024년 8월까지 동국대학교 일산한방한방병원을 방문한 만 19세 이상의 성인을 대상으로 수행되었다. 총 2,648명의 한의건강검진 및 일반건강검진 데이터를 수집하였으며, 이 중 결측치를 제거한 2,009명의 자료를 최종 분석에 포함하였다. 본 연구는 동국대학교 일산한방병원 기관생명윤리위원회(institutional review board, IRB)의 승인을 받아 진행되었다(IRB no. DUIOH 2023-03-001-002, DUIOH 2024-01-001-001). 연구대상자의 선정기준 및 제외기준은 다음과 같다.

1) 선정 기준

- (1) 만 19세 이상 성인 남녀
- (2) 인지기능에 문제가 없고, 본 연구에 대한 자세한 설명을 듣고 완전히 이해한 후 자의로 참여를 결정하고 서면동의를 통해 연구 참여에 동의한 자
- (3) 대상자의 임상연구 참여 종료 시점을 기준으로 6개월 이내에 본인 부담으로 실시한 양방 일반건강검진 데이터를 제공하거나 제공할 수 있는 자

2) 제외 기준

- (1) 거동이 불편하거나 기기 측정이 불가능한 자
- (2) 의료기관에서 다음 질병을 진단받고 유병 중인 자: 심혈관계 질환(심근경색, 울혈성심부전, 협심증, 부정맥 등), 뇌혈관질환(뇌졸중, 중풍 등), 악성 신 생물(암), 정신질환(우울증, 불안장애 등), 관절염 (류마티스), 갑상선질환(갑상선기능항진증, 갑상선 기능저하증) 유병자
- 단, 고혈압, 당뇨, 이상지질혈증, 지방간, 고요산혈증 유병자는 연구 대상으로 포함
- (3) 임신 중이거나, 임신가능성 있거나, 월경 중인 여성
- (4) 본인 명의의 휴대폰이 없는 자
- (5) 기타 연구책임자가 연구 참여에 적절하지 못하다고 판단되는 자
- (6) 당해 연도에 실시 중인 본 한의건강검진 임상 연구에 참여한 이력이 존재하는 자
- (7) 동의 취득일 기준 30일 이내에 다른 임상 연구에 참여한 이력이 존재하는 자

2. 연구방법

1) 건강검진 자료 수집 및 설진 · 체성분 · 혈액검사 측정

(1) 설진 영상 데이터 획득
설진은 한국한의약연구원에서 개발한 설 영상 측정 기기를 이용하여 표준화된 촬영 환경에서 수행하였다. 획득된 RGB 이미지는 CIE Lab 색공간으로 변환하였으며, 자동화된 영상분할 알고리즘을 통해 혀 전체 영역과 설태 영역을 분리하였다. 담홍지수(LIGHTRED)는 평균 Lab 색상값을 기반으로 밝은 홍색 영역의 비율을 계산하였고, 설태지수(COATED)는 전체 혀 면적 대비 설태가 덮인 픽셀 비율로, 청자지수(BLUEPURPLE)는 평균 Lab 색상값을 이용해 청자색 성분의 비율을 정량화하고, 치흔지수(TEETHMARK)는 설변 영역의 밝기 패턴을 푸리에 변환하여 산출하였다. 음영 대비가 클수록 치흔지수가 증가한다.

(2) 체성분 및 혈액학적 지표

체성분 분석은 direct segmental multi-frequency bio-electrical impedance analysis 방식(InBody Co., Korea)을 사용하여 키, 몸무게, BMI, 체지방량, 체지방률을 측정

하였다. 혈액검사는 자동 생화학 분석기를 이용하여 공복혈당과 혈청 크레아티닌을 측정하였다. 추정 사구체 여과율(eGFR)은 Chronic Kidney Disease Epidemiology Collaboration 공식으로 산출하였다.

(3) eGFR 분류(KDIGO 기준)

KDIGO 2024 기준에 따라 eGFR은 G1 (≥ 90), G2 (60~89), G3a (45~59), G3b (30~44), G4 (15~29), G5 (<15)로 분류한다. 본 연구에서는 G1을 ‘정상군’, G2 이하(G2~G5)를 ‘경도 감소 이하군’으로 이분화하여 분석하였다.

2) 자료 전처리

수집된 데이터는 R (version 4.3.2; R Foundation for Statistical Computing, Vienna, Austria)을 사용하여 전처리하였다. 우선 모든 연도별 데이터를 통합한 뒤 주요 변수(나이, 성별, 키, 몸무게, BMI, 설진 지표, 체성분, 혈당, eGFR)에서 결측값이 존재하는 사례를 제거하였다. 이후 BMI를 기반으로 4단계로 비만도 그룹을 구성하였다. BMI <25 는 비비만군(non-obese, NO), $25 \leq \text{BMI} < 30$ 은 1단계 비만(obesity category 1, OC1), $30 \leq \text{BMI} < 35$ 는 2단계 비만(obesity category 2, OC2), BMI ≥ 35 는 3단계 비만(obesity category 3, OC3)으로 정의하였다. 또한 GFR 값을 KDIGO 기준에 따라 정상군(≥ 90)과 경도감소 이하군(<90)으로 이분화하여 로지스틱 회귀 및 머신러닝 예측 모델의 종속변수로 사용하였다.

정규성 검정은 Shapiro-Wilk 검정과 quantile-quantile (Q-Q) plot으로 평가하였으며, 모든 연속형 변수에서 $p < 0.05$ 로 정규성이 기각되어 비모수 분석을 적용하였다.

연속형 변수는 중앙값과 사분위 범위(median [interquartile range, IQR])로, 범주형 변수는 빈도와 비율(n, %)로 요약하였다.

3) 데이터 분석

모든 통계 분석과 시각화는 R 환경에서 수행하였으며, 주요 패키지는 dplyr, rstatix, ggplot2, agricolae, corplot, tidyr, random forest, XGBoost 등을 사용하였다.

먼저 연구대상자의 기본 특성을 기술통계로 요약하였고, 연속형 변수는 Kruskal-Wallis 검정을, 범주형 변수는 χ^2 검정을 사용하여 집단 간 차이를 확인하였다.

그 다음, 상관 분석 및 연속형 eGFR 기반 회귀모형 평

가를 시행하였다(step 1). 설진 지표와 체성분, 신기능 변수 간의 상관관계는 비정규성 분포를 고려하여 Spearman 상관분석으로 평가하였으며, 설진 지표를 독립변수로 하고 연속형 eGFR을 종속변수로 한 다중 선형회귀모형을 평가하였다.

이후 연속형 모델의 성능을 보완하기 위해 GFR을 90 기준으로 정상군과 경도감소 이하군으로 이분화하여 다음 분석을 수행하였다(step 2). 다중 로지스틱 회귀분석은 GFR 감소 여부(<90)를 종속변수로, 설진 지표, BMI, 체성분, 연령 등을 독립변수로 시행하였으며 다중 로지스틱 회귀분석을 시행하였고, 오즈비(odds ratio, OR), 95% 신뢰구간(confidence interval), p-value를 산출하였다. 머신러닝 기반 비교분석을 위해 random forest와 XGBoost 모델을 구축하였다. 두 모델은 5-fold 교차검증을 통해 과적합을 최소화하였으며, 변수 중요도는 random forest에서는 mean decrease accuracy (MDA) 및 mean decrease gini (MDG), XGBoost에서는 gain 값으로 산출하였다. 모델의 판별력은 receiver operating characteristic (ROC) 곡선과 area under the curve (AUC)를 통해 평가하였다. 모든 분석의 유의수준은 양측검정에서 $p < 0.05$ 로 설정하

였다. 성별은 모델의 보정변수로 포함하여 분석하였다.

결과»»»»

초기 분석(step 1)에서는 BMI를 기준으로 4개의 비만도 그룹(obesity category: NO, OC1, OC2, OC3)으로 구분하여, 연속형 eGFR을 종속 변수로 하는 선형 회귀모형을 구축하였다. 설태지수, 청가지수, 치흔지수, BMI 등을 독립 변수로 포함했으나, 모든 모델에서 결정계수(R^2)가 최대 0.017에 불과하여 예측력이 매우 낮은 것으로 나타났다. 이에 따라 본 연구는 KDIGO 기준에 따라 eGFR을 정상군(≥ 90 mL/min/1.73m²)과 경도 감소 이하군(<90)으로 이분화한 뒤, 보다 임상 적용성이 높은 범주형 분석(step 2)을 수행하였다.

Table 1. General Characteristics by GFR Group (Normal vs. Mildly Decreased Group)

Variable	G1	G2-G5	p-value	χ^2
Sex (male/female)	270/1069	242/428	$1.56e^{-14}$	59.0
Age	44 (35-52)	50 (41-57)	$2.73e^{-23}$	98.8
Height	162.2 (157.9-167.8)	163.9 (158.9-170.7)	$1.53e^{-5}$	18.7
Weight	58.7 (53.1-67.8)	63.0 (55.5-73.0)	$2.8e^{-10}$	39.8
LIGHTRED	0.4 (0.3-0.5)	0.4 (0.3-0.5)	0.108	2.58
COATED	0.6 (0.4-0.9)	0.6 (0.4-0.9)	0.101	2.69
BLUEPURPLE	0.4 (0.3-0.4)	0.4 (0.3-0.5)	$1.4e^{-5}$	18.9
TOOTHMARK	11 (8-14)	10 (8-13)	0.267	1.23
eGFR	106 (97-117)	81 (74-85)	$3.27e^{-293}$	1339
FAT	18.1 (14.8-21.5)	18.1 (15.0-22.5)	0.326	0.97
PBF	30.7 (25.5-34.9)	29.3 (23.5-34.4)	$6.48e^{-4}$	11.6
GLUC (mg/dL)	92 (86-98)	93 (87-101)	$3.96e^{-5}$	16.9

Data are presented as median (IQR).

GFR: glomerular filtration rate, LIGHTRED: light red index of the tongue body (representing the brightness and paleness of tongue color; higher values indicate a stronger reddish color), COATED: coating index of the tongue surface (indicating the thickness of tongue coating; higher values reflect a thicker coating), BLUEPURPLE: bluish-purple index of the tongue body (reflecting dark or cyanotic tongue discoloration; higher values indicate a more bluish-purple appearance), TOOTHMARK: tooth mark index (measuring indentation patterns on the tongue edge; higher values suggest more pronounced tooth marks), eGFR: estimated glomerular filtration rate, FAT: body fat mass, PBF: body fat percentage, GLUC: fasting blood glucose level, IQR: interquartile range (25th-75th percentile).

Table II Baseline Characteristics by Obesity Category

Variable	Non-obese	Obesity category 1	Obesity category 2	Obesity category 3	p-value
Sex					
Female	1210 (81)	232 (55)	40 (56.3)	12 (63.2)	2.79e ⁻²⁸
Male	81 (19)	190 (45)	31 (43.7)	7 (36.8)	
Total	1,291	422	71	199	
Age (yr)	45 (37-54)	48 (39-56)	39 (33-49)	43 (38.5-56.5)	1.40e ⁻⁰⁴
Height	162.1 (158-167.175)	165.5 (158.13-172.48)	167.6 (161.75-174.5)	166 (155.5-170.55)	8.44e ⁻¹¹
Weight	56.6 (52.2-62.3)	73.4 (66.3-80.0)	90.6 (83.85-96.7)	100 (88.5-105.45)	3.02e ⁻¹⁸³
BMI	21.79 (20.29-23.30)	26.50 (25.67-27.79)	31.72 (31.05-32.46)	36.63 (35.73-37.08)	9.64e ⁻²⁵¹
CREAT (mg/dL)	0.7 (0.6-0.9)	0.8 (0.7-1)	0.8 (0.68-1)	0.84 (0.65-1)	5.56e ⁻⁰⁹
eGFR	98 (86-111.825)	93 (82.385-107.75)	98 (87-118)	98 (82-108.05)	2.94e ⁻⁰³
FAT	16.6 (13.8-19.3)	23.8 (20.3-26.8)	33.3 (29.8-36.25)	41.8 (37.5-45.8)	8.62e ⁻¹⁶⁴
PBF	29.4 (24.4-33.275)	33.8 (26.625-37.9)	38.8 (32.95-42.3)	45.9 (37.85-47.85)	6.11e ⁻⁴⁷

Data are presented as median (IQR).

Non-obese: BMI<25, Obesity category 1: 25≤BMI<30, Obesity category 2: 30≤BMI<35, Obesity category 3: BMI≥35, BMI: body mass index, CREAT: serum creatinine level, eGFR: estimated glomerular filtration rate, FAT: body fat mass, PBF: body fat percentage, IQR: interquartile range (25th-75th percentile).

1. 기본 정보

결측값이 제거된 총 2,009명이 분석에 포함되었으며, 정상군(1,339명)과 경도 감소 이하군(670명) 간의 인구학적 · 신체계측 · 설진 · 대사 지표의 차이를 비교하였다. 연속형 변수는 중앙값(IQR)으로 제시하였고, Kruskal-Wallis 검정을 이용하여 분석하였다. 범주형 변수는 χ^2 검정을 적용하였다(Table I). 또한 obesity category에 따른 baseline 비교 결과는 Table II에 제시하였다.

1) 인구학적 변수

성별 분포는 두 군 간 유의한 차이를 보였다($p<0.001$). 정상군에서는 남성 270명, 여성 1,069명으로 여성 비율이 높았던 반면, 경도 감소 이하 군에서는 남성 242명, 여성 428명으로 남성 비율이 상대적으로 더 높았다. 나이는 경도 감소 이하군에서 유의하게 더 높게 나타났다($p<0.001$) 신기능 저하가 성별과 나이와 밀접한 관련이 있음을 시사하였다.

2) 신체계측 및 체성분

키와 체중은 두 군 간 유의한 차이를 보였다. 정상군의 키 중앙값은 162.2 cm (사분위 범위: 157.9~167.8), 경

도 감소 이하 군은 163.9 cm (158.9~170.7)로, 경도 감소 이하군에서 더 큰 값을 보였으며 그 차이는 통계적으로 유의하였다($p<0.001$). 체중 또한 정상군군은 58.7 kg (53.1~67.8), 경도 감소 이하 군은 63.0 kg (55.5~73.0)으로 경도 감소 이하 군에서 유의하게 높은 값을 나타냈다($p<0.001$). 체성분 지표 중 체지방률은 정상군 30.7% (25.5-34.9), 경도 감소 이하 군 29.3% (23.5~34.4)로 통계적으로 유의한 차이를 보였으며($p<0.001$), 경도 감소 이하 군에서 더 낮은 값을 나타냈다. 체지방량은 두 군 간 유의한 차이를 보이지 않았다.

3) 설진 지표

설진 지표 중 청자지수는 경도 감소 이하군에서 유의하게 높았으며($p<0.001$), 이는 말초순환 저하 · 율혈 경향과 일치하는 소견이다. 담홍지수, 설태지수와 치혼지수는 두 군 간 차이가 유의하지 않았다.

4) 대사 및 신기능 지표

공복혈당은 정상군에서 92 mg/dL (86~98), 경도 감소 이하군에서 93 mg/dL (87~101)로, 경도 감소 이하군에서 통계적으로 유의하게 높은 값을 보였다($p<0.001$). eGFR은 정상군 106 mL/min/1.73m² (97~117), 경도 감소 이하군 81 mL/min/1.73m² (74~85)로, 두 군 간 명확

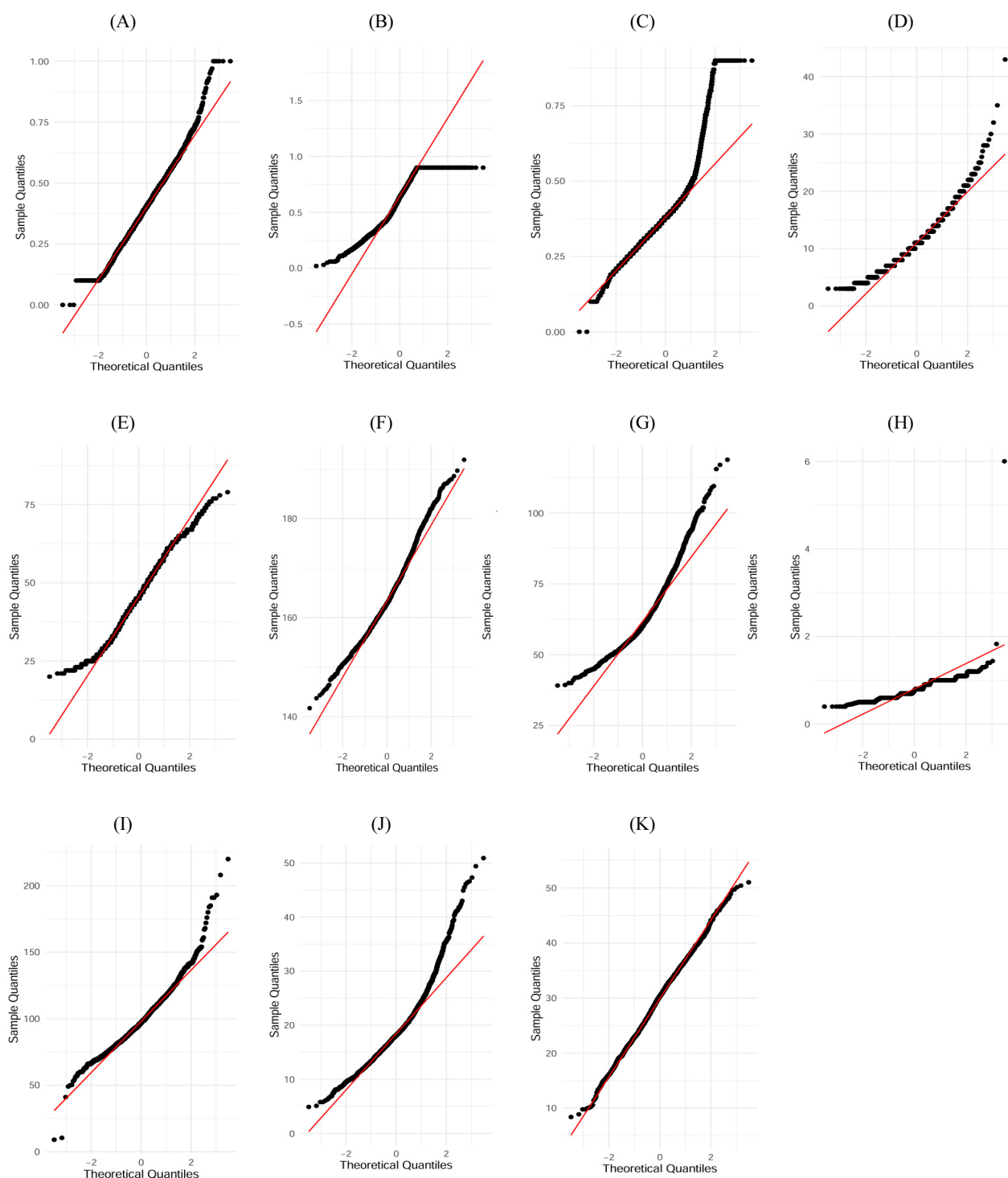


Fig. 1. Q-Q plots of major continuous variables. (A) LIGHTRED index, (B) COATED index, (C) BLUEPURPLE index, (D) TOOTHMARK index, (E) age, (F) height, (G) weight, (H) serum creatinine level, (I) eGFR, (J) body fat mass, (K) body fat percentage. LIGHTRED: light red index of the tongue body, COATED: coating index of the tongue surface, BLUEPURPLE: bluish-purple index of the tongue body, TOOTHMARK: tooth mark index, eGFR: estimated glomerular filtration rate.

한 차이를 나타냈으며($p<0.001$), 경도 감소 이하군에서 유의하게 낮았다.

2. 데이터의 정규성 검정

Shapiro-Wilk 검정과 결과, 모든 연속형 변수에서 $p<0.05$ 로 확인되어 정규성을 만족하지 않았다. Q-Q plot에서도 관측값이 이론적 직선으로부터 벗어나는 분포를 보여(Fig. 1), 시각적으로도 정규성이 부정되었다. 이에 따라 이후 분석은 Spearman 상관분석, Kruskal-Wallis 검정 등 비모수 방법을 사용하였다.

3. 비만도에 따른 설진 지표 변화

비만도가 증가할수록 설태지수는 증가하고(Fig. 2A, C), 담홍지수는 감소하는 경향을 보였다(Fig. 2B, D).

4. 비만도와 신기능 간의 상관관계

연속형 eGFR과 BMI간 상관성을 분석한 결과, BMI는 eGFR과 유의한 음의 상관관계($r<0$, $p<0.001$)를 보였다. 4개 BMI 단계로 구분한 후 비교한 결과, 군 간 eGFR의 중앙값 차이가 통계적으로 유의했다($p<0.001$). 즉, BMI가 높을수록 eGFR이 낮아지는 경향이 뚜렷하

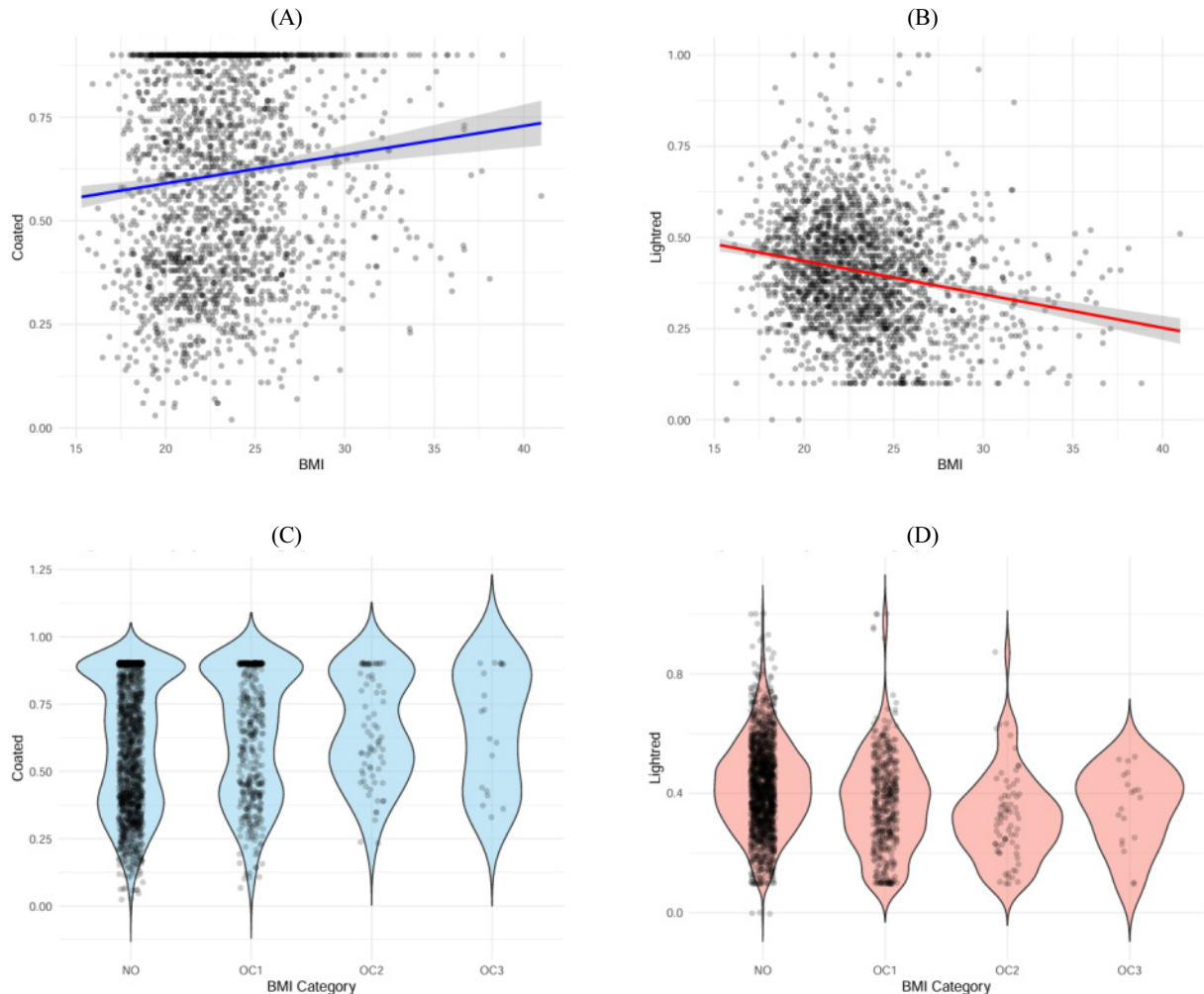


Fig. 2. Associations between BMI and tongue indices. (A) Scatterplot of BMI vs. COATED index. (B) Scatterplot of BMI vs. LIGHTRED index. (C) Violin plot of COATED index by BMI category. (D) Violin plot of LIGHTRED index by BMI category. BMI: body mass index, COATED: coating index of the tongue surface, LIGHTRED: light red index of the tongue body.

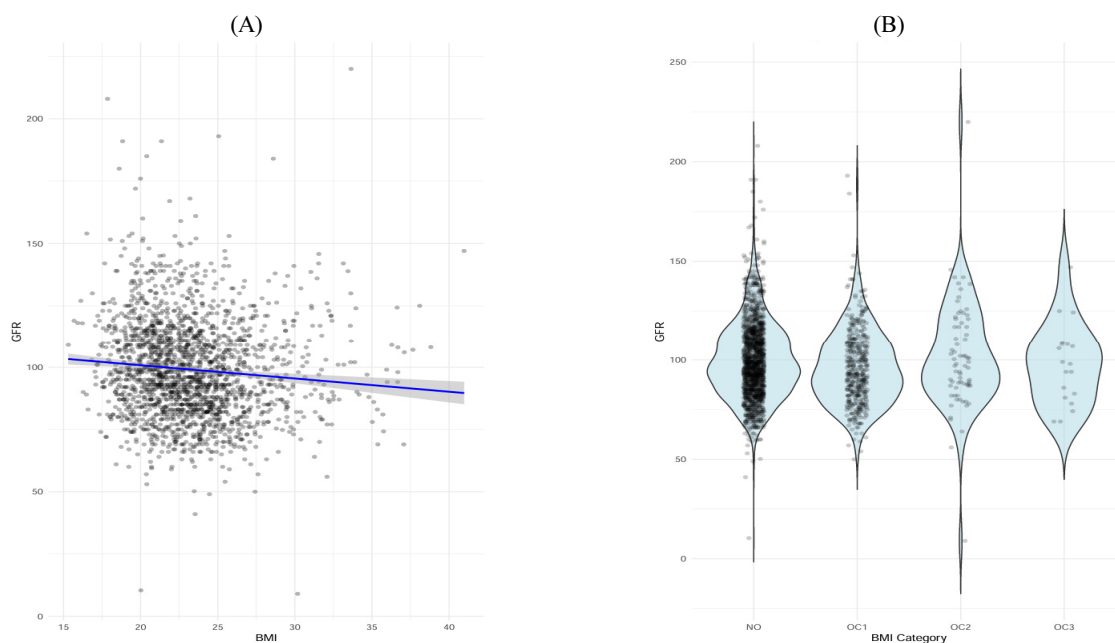


Fig. 3. Associations between BMI and eGFR. (A) Scatterplot of BMI vs. eGFR. (B) Violin plot of eGFR by BMI category. BMI: body mass index, eGFR: estimated glomerular filtration rate.

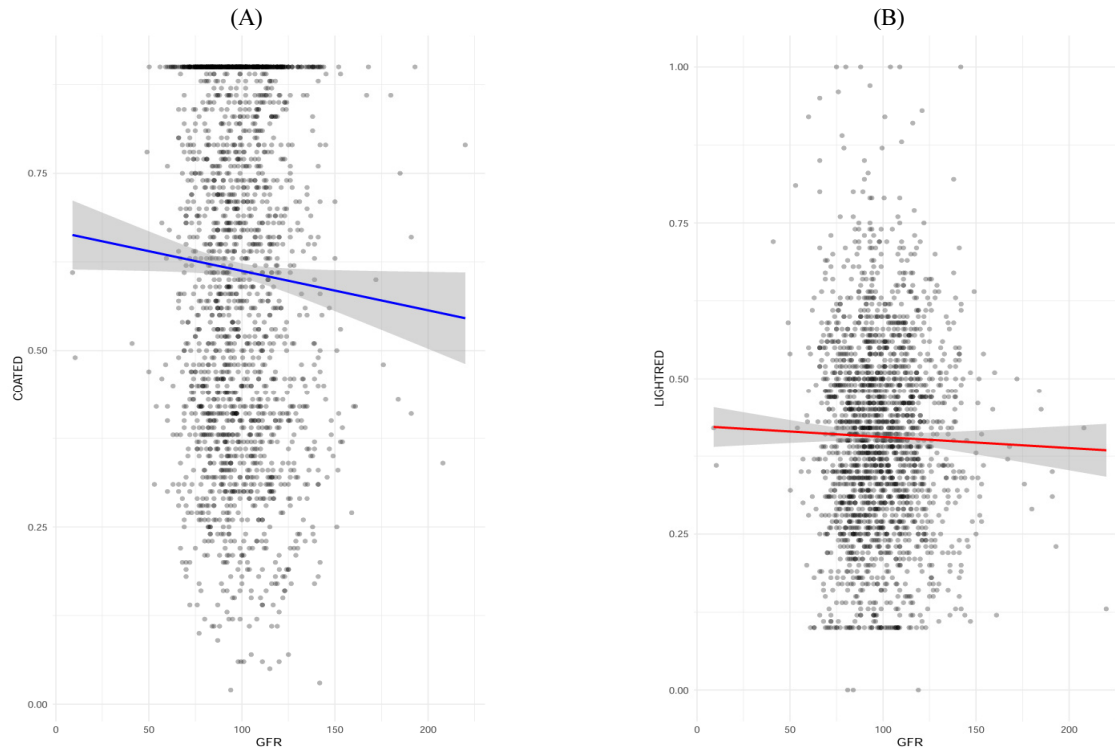


Fig. 4. Scatterplots of eGFR vs. tongue indices. (A) COATED index. (B) LIGHTRED index. eGFR: estimated glomerular filtration rate, COATED: coating index of the tongue surface; indicating the thickness of tongue coating, with higher values reflecting a thicker coating, LIGHTRED: light red index of the tongue body; representing the brightness and paleness of tongue color, with higher values indicating a stronger reddish color.

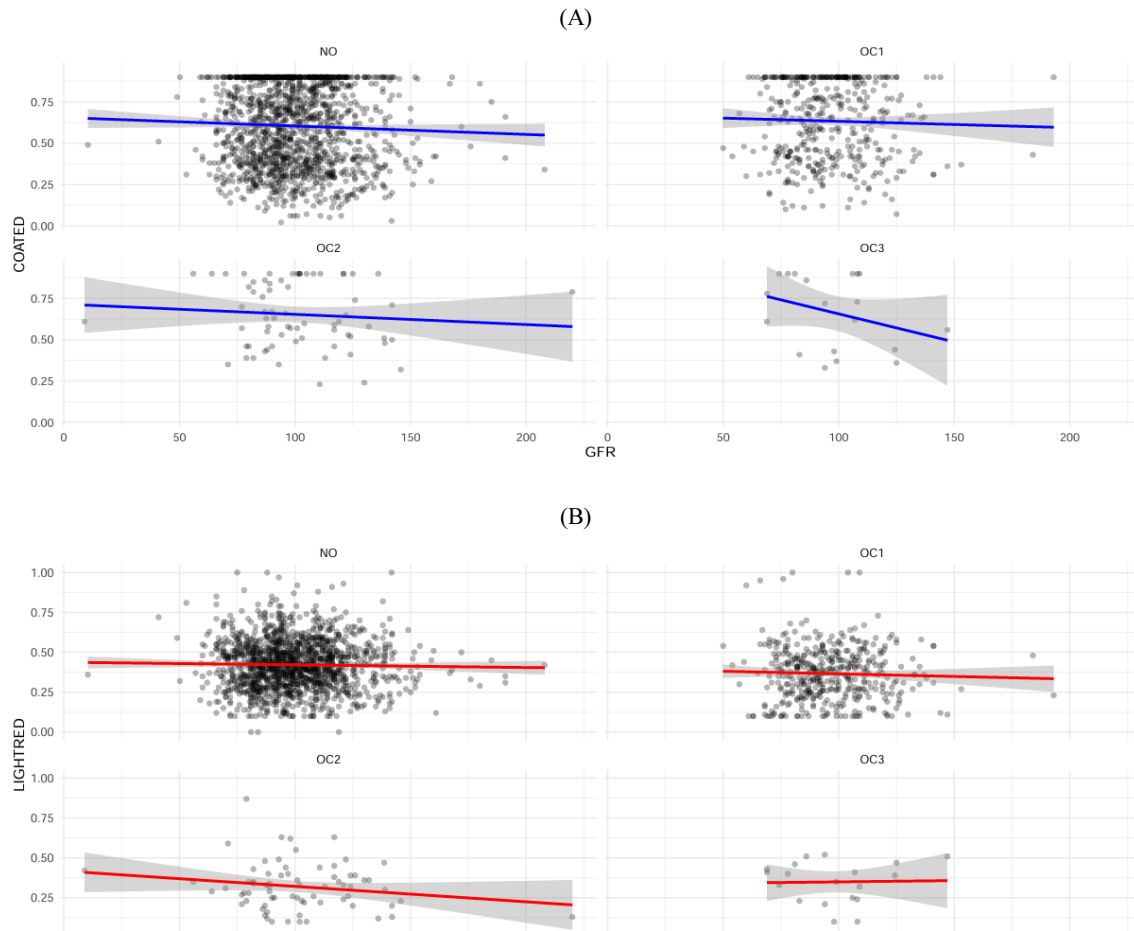


Fig. 5. Scatterplots of eGFR vs. tongue indices stratified by BMI category. (A) COATED index. (B) LIGHTRED index. eGFR: estimated glomerular filtration rate, BMI: body mass index, COATED: coating index of the tongue surface; indicating the thickness of tongue coating, with higher values reflecting a thicker coating, LIGHTRED: light red index of the tongue body; representing the brightness and paleness of tongue color, with higher values indicating a stronger reddish color.

여 비만 자체가 신기능 저하와 독립적으로 연관됨을 시사했다(Fig. 3).

5. 신기능과 설진 지표 간의 상관관계

eGFR은 설태지수 및 담홍지수와 음의 경향을 보여 eGFR 낮을수록 설태지수와 담홍지수가 증가하였고(Fig. 4) 청자지수 및 치혼지수와는 유의한 상관관계를 보이지 않았다.

BMI 카테고리별로 eGFR과 설진 지표 간의 관계를 분석한 결과, 일부 그룹에서는 설태 및 담홍지수가 eGFR과 음의 경향을 보였으나 전체적으로 통계적 유의성은 확보되지 않았다(Fig. 5). 특히 OC3 그룹에서만 설태와

eGFR의 음의 경향이 비교적 유의미했으나($p>0.05$), 경향성 수준에 머물렀다. 담홍지수 또한 모든 BMI 그룹에서 eGFR과 유의한 상관성을 보이지 않았다.

6. 연속형 eGFR 수치 예측을 위한 다중 회귀모델 구축 및 성능 평가

eGFR을 연속형으로 두고 설진 지표로 예측 가능한지 확인하기 위해 회귀모델을 구축해본 결과, 1) 설태와 설질, 2) 설태, 설질, BMI, 3) 모든 설진 지표와 BMI를 활용한 모델의 성능이 가장 높았으나, 3개의 모델 모두 R^2 는 0.0039~0.017 수준으로 낮아 연속형 eGFR을 설진 지표만으로 예측하기 어렵다는 결론에 도달하였다.

즉, step 1에서는 BMI가 eGFR과 음의 상관관계를 보여 비만할수록 신기능이 낮고, 설태 · 담홍지수가 eGFR과 유의한 상관관계를 보여 설진이 신기능을 반영할 수 있음을 확인했으나, 연속형 eGFR 예측 모델의 설명력이 매우 제한적으로 임상적 유용성이 낮다고 판단되었다. 한편 CKD 진단에서는 eGFR의 절대치보다 ‘카테고리 변화(G1 to G2)’가 더욱 중요하므로⁴⁾, 본 연구는 실질적 임상 적용성을 평가하기 위해 KDIGO 기준에 따라 eGFR을 정상군과 경도 감소 이하군으로 이분화하여 로지스틱 회귀 및 머신러닝 모델을 재구축하였다.

7. 로지스틱 회귀에 의한 설진 및 체성분 변수 기반 eGFR 저하 여부 예측

다중 로지스틱 회귀분석 결과, 청자지수(OR=3.26, $p=0.001$), 나이(OR=1.06, $p<0.001$), 체지방량(OR=1.11, $p=0.026$), 체지방률(OR=0.875, $p<0.001$)이 eGFR 감소와 유의하게 관련되었다. 반면 설태지수, 담홍지수, 치흔지

수 및 키는 유의미한 관련성을 보이지 않았다. 청자지수의 OR 값이 3.26이라는 것은 설진이 청자색에 가까울수록 신기능 저하 위험이 약 3.3배 증가함을 의미하며, 이는 이전 CKD 관련 설진 연구에서 보고된 설진 패턴과 일치한다¹⁴⁾.

8. 머신러닝 기반 eGFR 저하 여부 예측

본 연구에서는 설진 지표와 신체계측, 대사 지표를 포함한 다변량 정보를 이용하여 GFR 감소($<90 \text{ mL/min/1.73m}^2$)를 분류하는 머신러닝 모델을 구축하였다. Random forest와 XGBoost 알고리즘을 적용하였으며, 각 모델의 변수 중요도와 ROC 및 AUC를 산출하여 예측 성능을 평가하였다.

1) Random forest 모델

Random forest의 변수 중요도는 크게 두 가지 기준인 MDA와 MDG를 통해 산출된다. MDA는 특정 변수를

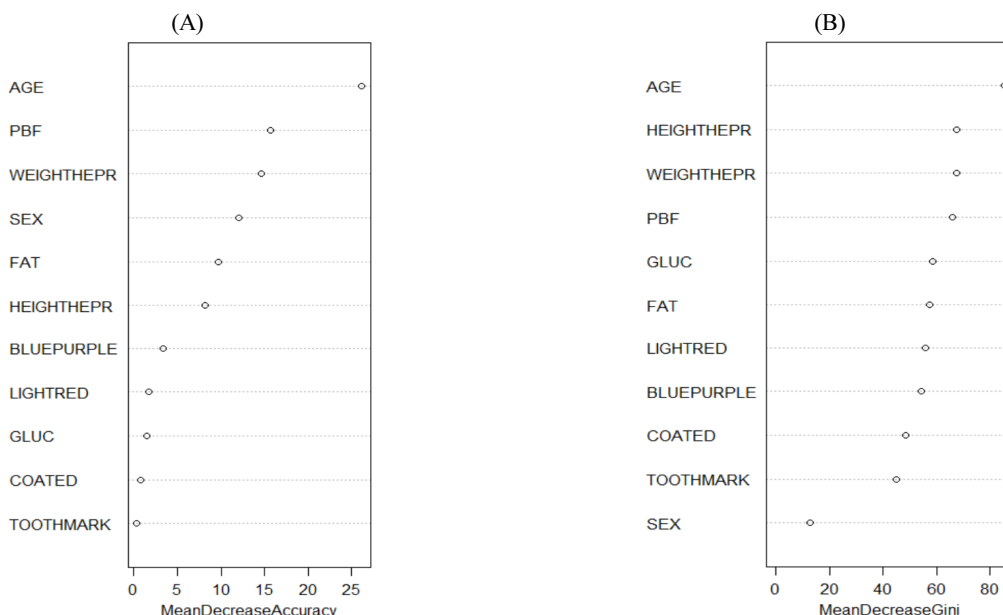


Fig. 6. Variable importance in the random forest model. (A) MDA. (B) MDG. MDA: mean decrease accuracy, MDG: mean decrease gini, PBF: body fat percentage, WEIGHTPR: body weight, FAT: body fat mass, HEIGHTPR: standing height, GLUC: fasting blood glucose level (mg/dL), BLUEPURPLE: bluish-purple index of the tongue body; reflecting dark or cyanotic tongue discoloration, with higher values indicating a more bluish-purple appearance, LIGHTRED: light red index of the tongue body; representing the brightness and paleness of tongue color, with higher values indicating a stronger reddish color, COATED: coating index of the tongue surface; indicating the thickness of tongue coating, with higher values reflecting a thicker coating, TOOTHMARK: tooth mark index; measuring indentation patterns on the tongue edge, with higher values suggesting more pronounced tooth marks.

제거했을 때 전체 모델 정확도가 얼마나 감소하는지를 기반으로 하며, MDG는 분류를 위한 노드 분할에서 해당 변수가 얼마나 자주 사용되었는지를 측정한다.

(1) MDA

MDA 기준 변수 중요도는 나이가 가장 높은 기여도를 보였으며, 체지방률, 몸무게, 성별, 체지방량, 키가 뒤를 이었다. 설진 지표 중에서는 청자지수가 중간 정도의 기여도를 보였으며, 담홍지수, 치혼지수, 설태지수는 모든 변수 중 하위권에 위치하였다. 즉, random forest에서는 신기능 예측에 있어 설진 지표의 기여는 제한적이며, 설진 지표 중에서는 청자지수가 가장 의미 있는 변수였다(Fig. 6A).

(2) MDG

MDG 기준으로 나이가 가장 중요한 변수였으며, 키와 몸무게가 그 뒤를 이었다. 체지방률과 공복혈당,

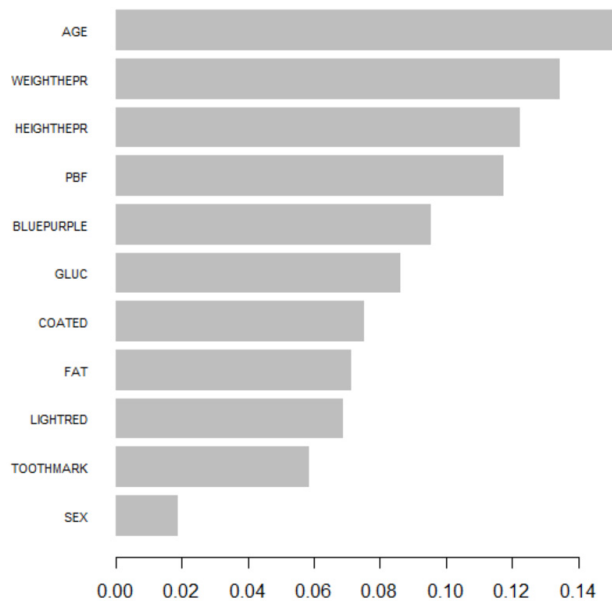


Fig. 7. Feature importance in the XGBoost model. WEIGHTPR: body weight, HEIGHTPR: standing height, PBF: body fat percentage, FAT: body fat mass, GLUC: fasting blood glucose level (mg/dL), BLUEPURPLE: bluish-purple index of the tongue body; reflecting dark or cyanotic tongue discoloration, with higher values indicating a more bluish-purple appearance, LIGHTRED: light red index of the tongue body; representing the brightness and paleness of tongue color, with higher values indicating a stronger reddish color, COATED: coating index of the tongue surface; indicating the thickness of tongue coating, with higher values reflecting a thicker coating, TOOTHMARK: tooth mark index; measuring indentation patterns on the tongue edge, with higher values suggesting more pronounced tooth marks.

체지방량 또한 높은 중요도를 보였다. 설진 지표의 경우 청자지수와 담홍지수가 중·하위 수준을 유지하였고, 설태지수와 치혼지수는 가장 낮은 변수로 나타났다(Fig. 6B).

(3) Random forest 모델 해석

MDA, MDG 두 지표 모두에서 나이, 체성분, 신체계측 변수가 가장 중요한 예측 변수였고 청자지수만이 설진 지표 중 일정한 기여도를 보였으며 다른 설진 지표의 중요도는 매우 낮았다. 이는 신기능 저하가 전신 상태, 특히 나이·체성분·대사 기능에 의해 주로 설명되며, 설진 지표의 독립적 기여는 제한적임을 시사한다.

2) XGBoost 모델

XGBoost 모델의 결과에서도 나이의 중요도가 압도적으로 높았으며, 키 및 몸무게, 체지방률, 혈당이 높은 수준의 기여도를 나타냈다. 설진 변수 중에서는 청자지수가 중간 정도의 중요도를 보였으나, 설태지수, 치혼지수, 담홍지수는 하위권에 머물렀다(Fig. 7).

3) 모델 간 성능 비교

분류 모델의 성능 비교를 위해 ROC 곡선을 분석하고 AUC를 산출하였다. ROC 곡선은 민감도(sensitivity)와 1-특이도(false positive rate) 간의 관계를 시각화한 그래프로, 모델의 분류 임계값 변화에 따른 예측 성능

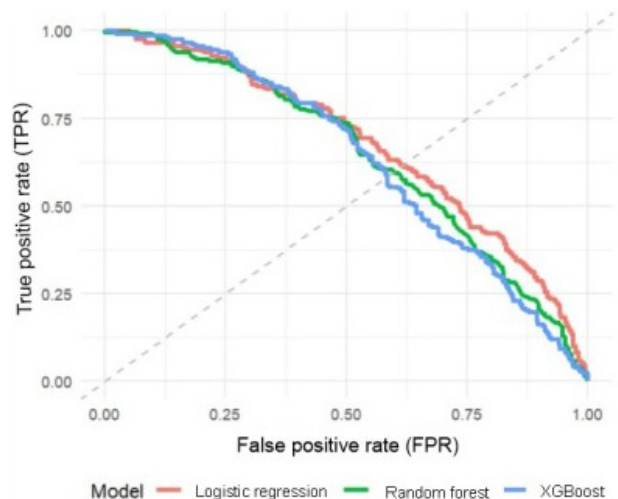


Fig. 8. Receiver operating characteristic curves of logistic regression, random forest, and XGBoost models. XGBoost: extreme gradient boosting.

을 평가할 수 있다. AUC는 이 곡선 아래 면적으로, 1에 가까울수록 분류 정확도가 높음을 의미한다.

본 분석에서 로지스틱 회귀 모델이 가장 높은 AUC 값(0.699)을 보여, 비선형 모델인 random forest (AUC=0.644) 및 XGBoost (AUC=0.652)보다 더 우수한 성능을 나타냈다(Fig. 8).

4) 머신러닝 분석 요약

나이, 체성분(체지방량, 체지방률), 신체계측치(키, 몸무게), 공복혈당이 GFR 감소를 예측하는 데 가장 중요한 변수였다. 설진지표 중 청자지수는 중간 수준의 예측력을 갖는 유일한 변수였다. 설태지수, 담홍지수, 치혼지수는 머신러닝 모델에서 일관되게 낮은 기여도를 보여 예측 변수로서의 영향이 제한적이었다. BMI는 체중·체지방률·체지방량 등 체성분 변수의 정보에 의해 설명력이 대체되는 양상을 보여, 주요 중요 변수로는 부각되지 않았다. 모델의 AUC가 0.64~0.70 범위로 보통 수준에 머무른 점은 설진 지표만으로 신기능 이상을 강하게 예측하기 어렵다는 점을 시사한다. 그러나 청자지수를 중심으로 한 설질 색 변화는 유의한 임상적 의미를 가질 수 있다.

고찰»»»

본 연구는 2020년부터 2024년까지 추적된 대규모 건강검진 자료를 활용하여 설진 지표와 신기능의 연관성을 평가하였으며, 비만도가 이 관계를 어떻게 조절하는지를 검토하였다. 기존 선행연구들은 설진 지표가 간기능, 심혈관계, 대사질환 등 여러 전신 질환의 상태를 반영할 수 있다는 점을 보고해왔으며^{6,9-12)}, 국내 연구에서도 비만이 설태 및 설질과 밀접하게 연관된다는 결과가 제시된 바 있다⁸⁾. 본 연구는 이러한 맥락에서 한의학적 진단 지표가 신기능과도 관련될 수 있는지 검증하고, 나아가 설진을 활용한 신기능 예측 가능성을 탐색하는 확장된 연구라 할 수 있다.

본 연구에서 확인된 바와 같이 BMI는 eGFR과 통계적으로 유의한 음의 상관관계를 보여, 비만할수록 신기능이 저하되는 경향이 명확하였다. 이는 기존 메타분석

연구에서 비만이 사구체고혈압, 사구체비대, 세뇨관 간 질섬유화 등을 통해 CKD 발생위험을 증가시킨다고 보고한 결과와 일치한다¹⁵⁾. 특히 본 연구에서는 BMI가 신기능 변화의 독립적인 관련 인자로 작용했으며, 이는 체중조절이 신기능 보호에 있어 중요한 중재 요인이 될 수 있음을 시사한다. 한편 eGFR 저하 여부를 분류하는 다변량 예측모형에서는 BMI 자체보다 연령 및 체성분(체지방량·체지방률) 변수가 더 큰 설명력을 보였으며, 이는 비만의 영향이 신체구성 및 대사 상태와 밀접하게 중첩되어 나타날 가능성을 시사한다.

연속형 eGFR을 기준으로 한 1차 분석에서는 설태지수(COATED)와 담홍지수(LIGHTRED)가 모두 eGFR과 통계적으로 유의한 음의 상관관계를 보였다. 즉, eGFR이 낮을수록 설태는 두꺼워졌고, 설질은 오히려 더 붉어지는 양상이 나타났다. 설태가 신기능 저하와 함께 증가한다는 결과는 기존 CKD 환자의 설태 변화와 유사하며, 전신 미세순환 장애와 구강 대사환경 변화가 혀 표면에 반영된 것으로 해석할 수 있다^{13,14)}.

반면, 설질 밝기가 eGFR 감소와 함께 증가한 결과는 기존 한의학적 해석과는 상반된 부분이 있다. 전통적으로 담홍지수의 감소는 기허·양허 상태에서 나타나는 창백한 설질과 연관되지만, 본 연구에서는 eGFR이 감소할수록 설질이 홍색을 띠는 방향으로 나타났다. 이러한 결과는 비만도, 연령, 대사 상태, 혈류역학적 변화 등 교란 요인의 영향을 반영한 복합적 결과일 가능성이 높으며, 한의학적 병리 개념과 일대일 대응으로 해석하는 데에는 주의가 필요하다. 특히 본 연구는 건강인을 대상으로 수행되었으며, 신기능에 영향을 줄 수 있는 명확한 신장질환 또는 만성 염증 상태를 가진 대상자는 포함되지 않았기 때문에, 설질의 붉어짐은 중증 염증성 변화보다는 대사적 또는 순환계적 변화의 반영으로 해석해야 한다. 또한 본 연구는 실제 건강검진 참여자를 대상으로 하였기 때문에, 당뇨병이나 고혈압 등의 만성질환자 및 해당 약물 복용자가 일부 포함되었을 가능성이 있다. 향후 연구에서는 이러한 교란 요인을 배제하거나, 약물 복용 이력에 따른 하위분석을 실시함으로써, 설진 지표의 예측력을 보다 명확히 평가할 필요가 있다. 특히 BMI, 연령, 성별 등을 공변량으로 통제하여 보다 엄격한 환자군 선별 및 병력 기반 분석이 병행된다면 임상적 활용 가능성을 더욱 높일 수 있을 것이다.

설진 지표만을 이용한 다중 선형회귀모델에서는 설명력이 0.39~1.7%에 불과하여 임상적 예측 도구로 활용하기에는 현실적 제약이 컸다. 이는 설진 지표가 신기능의 변화를 직접적으로 정량 예측하기보다는, 전신 상태를 간접적으로 반영하는 보조지표로 활용될 가능성이 크다는 점을 시사한다⁶⁾.

연속형 예측의 한계를 극복하기 위해 KDIGO 기준을 적용하여 정상(G1)과 경도 감소(G2~G5)로 범주화한 후 로지스틱 회귀를 시행한 결과, 청자지수, 나이, 체지방량, 체지방률이 신기능 감소와 유의한 관련성을 나타냈다. 이 중 청자지수의 OR=3.26은 설질이 더욱 청자색에 가까울수록 GFR 감소 위험이 증가함을 의미하며, 이는 CKD 환자에서 혀 색깔의 어둡고 자색에 가까운 변화가 보고된 기존 연구들과 일관된다^{13,14)}. 청자색 설질은 말초순환 저하와 울혈을 반영하는 한의학적 병리와도 부합하며, 신기능 저하의 임상적 징후를 설상에서 포착할 수 있다는 점을 시사한다.

한편 체지방률은 OR<1로 나타나 통계적으로는 보호적 방향을 보였으나, 이는 실제로 체지방이 신기능을 보호한다는 의미라기보다는, 고령층에서의 근감소·노쇠와 같은 체성분 변화가 GFR 저하에 더 직접적인 영향을 미치는 역설적 역학 현상으로 해석할 수 있다. 즉, 동일한 BMI 범주 내에서도 근육량이 낮고 체지방률도 상대적으로 낮은 노쇠 집단에서 GFR 감소가 더 빈번하게 나타나는 경향이 반영된 결과로, 체지방률의 감소가 곧 건강함을 의미하지는 않는다는 점을 보여준다. 이러한 패턴은 사르코페니아·근감소성 비만 연구에서도 보고된 바 있으며, 신기능 저하와 체성분 간의 복합적 상호작용을 시사한다.

Random forest¹⁶⁾와 XGBoost¹⁷⁾ 알고리즘 분석에서도 설진 지표보다 나이, 체성분, 공복혈당 등이 GFR 감소 예측에서 훨씬 높은 기여도를 나타냈다. 설진 변수 중에서는 청자지수만이 중간 수준의 중요도를 유지했다. 이는 전신 상태를 반영하는 대사·신체계측 변수가 신기능 저하에 더 직접적으로 영향을 미치며, 설진의 예측력은 제한적임을 보여준다. 다만 머신러닝의 AUC 자체가 0.64~0.70 수준에 머문 것은 본 연구에서 다룬 변수들 자체가 신기능 예측의 한계를 갖고 있다는 의미이기도 하다. 또한 비선형 모델이 선형모델보다 성능이 크게 향상되지 않았다는 점은 신기능 예측에서 설진 변

화의 기여가 비교적 단순하고 제한적이라는 점을 시사한다.

최근 몇 년간 설진의 객관화 및 자동화에 대한 연구가 활발히 진행되고 있다. 실제로 설진 영상을 딥러닝 기반으로 분석하여 색상, 질감, 설태, 치흔 등의 지표를 정량화하고, 이를 다양한 질환의 진단 또는 예측에 활용하는 시도가 증가하고 있다. 예컨대, 자동화 혀 영상 진단 시스템을 이용한 만성 신질환 환자 연구에서는 설태 두께, 혀 색, 치흔, 출혈반 등이 CKD의 중증도와 유의하게 연관되었으며, 특히 출혈반은 AUC 0.757로 비교적 양호한 판별력을 보였다¹⁸⁾. 또한 설질의 색 범주를 픽셀 단위로 분류하는 딥러닝 기반 연구에서는 약 92%에 달하는 색상 분류 정확도를 보고하였고, 이는 기존의 육안 관찰에 비해 훨씬 높은 재현성과 신뢰도를 가진다¹⁹⁾. 나아가, 설진 영상에 구강 미생물 정보를 결합하여 대사성 간질환 환자의 한의학적 변증을 분류한 연구에서는 AUC=0.939, 정확도 85%라는 매우 우수한 성능을 확인하였다²⁰⁾. 이러한 연구들은 설진이 AI를 기반으로 주관성을 극복하고 객관적인 진단 도구로 전환될 수 있음을 보여준다.

본 연구 역시 설진 지표를 중심으로 신기능 저하와의 연관성을 평가하고, 머신러닝 기반 예측모델을 통해 임상 적용 가능성을 탐색했다는 점에서 이 같은 연구 흐름과 맞닿아 있다. 다만 본 연구에서 구축한 머신러닝 모델의 AUC가 0.64~0.70 수준에 머문 점은, 설진 지표 뿐 아니라 체성분, 신체계측, 혈당 등 다양한 지표를 포함했음에도 불구하고 신기능 예측에서 설진 지표의 독립적 기여도가 비교적 낮았음을 보여준다. 또한, 비선형 모델이 선형 모델보다 크게 우수하지 않았다는 점은, 설진 변화가 신기능 저하에 대한 단순한 반영일 가능성이 높고, 복잡한 상호작용보다는 제한적인 지표로 작용함을 시사한다. 그럼에도 불구하고, 청자지수는 모든 모델에서 일관되게 중간 이상의 중요도를 보여, 전통적으로 어혈이나 한증과 연관된 설질 색상이 실제로 신기능 저하와 유의하게 관련될 수 있음을 시사하였다. 이는 한의학적 병리 해석과 현대 생리지표 간의 접점을 제공하며, 향후 설진 지표의 통합적 활용 가능성을 높인다. 향후 연구에서는 설진 영상 기반 색상·질감 분석 뿐 아니라, 구강 미생물, 염증지표, 장-신장 축(gut-kidney axis) 등을 포함한 다중 생체지표 융합 모델을 활용

하여 신기능 평가의 정밀도를 높일 필요가 있다^{10,21,22)}. 특히 최근 대규모 설진 영상 데이터셋과 자동화 설진 시스템이 구축되고 있는 흐름은, 한의학 설진의 현대적 활용 가능성을 더욱 확장할 수 있는 기반이 될 것이다.

결론»»»»

본 연구는 설진 지표가 비만도 및 신기능과 어떠한 연관성을 갖는지 탐색하고, 설진 정보를 기반으로 신기능 저하 여부를 예측할 수 있는 가능성을 평가하였다. 연속형 eGFR 분석에서는 설태지수와 담홍지수가 eGFR과 유의한 음의 상관관계를 보여, 사구체여과율이 낮을수록 설태가 증가하고 설질이 홍색을 띠는 경향이 확인되었다. 그러나 설진 지표만을 이용한 회귀모델의 설명력은 매우 제한적이었다. 또한 BMI는 eGFR과 유의한 음의 상관관계를 보여, 비만도가 증가할수록 신기능 저하 경향이 뚜렷하였다.

KDIGO 기준에 따라 eGFR을 정상군과 경도 감소 이하군으로 이분화하여 분석한 결과, 청자지수가 신기능 저하와 가장 강한 독립적 관련성을 보였으며(OR=3.26), random forest 및 XGBoost 분석에서도 유일하게 일관된 변수 중요도를 나타냈다. 이는 청자색 설질이 신기능 저하를 반영하는 유의미한 지표일 가능성을 시사한다. 반면 설태지수, 담홍지수, 치혼지수 등 다른 설진 지표들의 예측력은 상대적으로 낮았다.

요약하면, 비만 지표는 신기능 저하와 일관된 관련성을 보였으며, 설진 지표만으로 신기능의 절대값을 예측하는 데에는 한계가 있으나, 청자지수는 신기능 저하 위험을 조기에 식별할 수 있는 보조적 진단 지표로 활용될 수 있다는 점에서 임상적 의의가 있다.

References»»»»

1. Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, Saran R, Wang AY, Yang CW. Chronic kidney disease: global dimension and perspectives. *Lancet*. 2013;382(9888):260-72.

2. Xie Y, Bowe B, Mokdad AH, Xian H, Yan Y, Li T, Maddukuri G, Tsai CY, Floyd T, Al-Aly Z. Analysis of the Global Burden of Disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016. *Kidney International*. 2018;94(3):567-81.
3. Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. *Kidney International Supplements*. 2022;12(1):7-11.
4. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney International*. 2024;105(4S):S117-314.
5. Chen T, Chen Y, Zhou Z, Zhu Y, He L, Zhang J. Deep learning-based automated tongue analysis system for assisted Chinese medicine diagnosis. *Frontiers in Physiology*. 2025;16:1559389.
6. Yuan L, Yang L, Zhang S, Xu Z, Qin J, Shi Y, Yu P, Wang Y, Bao Z, Xia Y, Sun J, He W, Chen T, Chen X, Hu C, Zhang Y, Dong C, Zhao P, Wang Y, Jiang N, Lv B, Xue Y, Jiao B, Gao H, Chai K, Li J, Wang H, Wang X, Guan X, Liu X, Zhao G, Zheng Z, Yan J, Yu H, Chen L, Ye Z, You H, Bao Y, Cheng X, Zhao P, Wang L, Zeng W, Tian Y, Chen M, You Y, Yuan G, Ruan H, Gao X, Xu J, Xu H, Du L, Zhang S, Fu H, Cheng X. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. *EClinicalMedicine*. 2023;57:101834.
7. Lin CH, Yang SH, Lee JD. Tongue image segmentation and constitution identification with deep learning. *Electronics*. 2025;14(4):733.
8. Lee J, Han K, Nam DH, Kim H. Analysis of differences in pulse and tongue diagnosis indicators and their correlations with clinical markers based on obesity levels. *Journal of Korean Medicine for Obesity Research*. 2024;24(2):144-56.
9. Son JY, Kim JS. Diagnostic values of tongue coating thickness and sterno-costal angle in functional dyspepsia. *The Journal of Internal Korean Medicine*. 2014;35(2):157-74.
10. Jiang B, Liang X, Chen Y, Ma T, Liu L, Li J, Jiang R, Chen T, Zhang X, Li S. Integrating next-generation sequencing and traditional tongue diagnosis to determine tongue coating microbiome. *Scientific Reports*. 2012;2:936.
11. Li J, Xiong D, Hong L, Lim J, Xu X, Xiao X, Guo R, Xu Z. Tongue color parameters in predicting the degree of coronary stenosis: a retrospective cohort study of 282 patients with coronary angiography. *Frontiers in*

- Cardiovascular Medicine. 2024;11:1436278.
12. Duan M, Mao B, Li Z, Wang C, Hu Z, Guan J, Li F. Feasibility of tongue image detection for coronary artery disease: based on deep learning. *Frontiers in Cardiovascular Medicine*. 2024;11:1384977.
 13. Chen JM, Chiu PF, Wu FM, Hsu PC, Deng LJ, Chang CC, Chiang JY, Lo LC. The tongue features associated with chronic kidney disease. *Medicine*. 2021;100(9):e25037.
 14. Chung CJ, Wu CH, Hu WL, Shih CH, Liao YN, Hung YC. Tongue diagnosis index of chronic kidney disease. *Biomedical Journal*. 2023;46(1):170-8.
 15. Wang Y, Chen X, Song Y, Caballero B, Cheskin LJ. Association between obesity and kidney disease: a systematic review and meta-analysis. *Kidney International*. 2008;73(1):19-33.
 16. Breiman L. Random forests. *Machine Learning*. 2001;45(1):5-32.
 17. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016:785-94.
 18. Hsu PC, Chen JM, Chang CC, Chang YJ, Chiu PF, Chiang JY, Lo LC. Exploring the pivotal variables of tongue diagnosis between patients with chronic kidney disease and health participants. *Frontiers in Big Data*. 2025;7:1443646.
 19. Liu B, Wang Z, Yu K, Wang Y, Zhang H, Song T, Yang H. Research on tongue image segmentation and classification methods based on deep learning and machine learning. *Information*. 2025;16(5):357.
 20. Zhang J, Feng S, Xue J, Zhou C, Fu S, Liu P, He Z, Gong M, Feng H, Zhou S, Zhang N. AI-driven multi-modal fusion of tongue images and clinical indicators for identifying MAFLD patients at risk of coronary artery disease: an exploratory study. *iLIVER*. 2025;4(3):100181.
 21. Shim H, Kwon O, Kim MJ, Song EJ, Moon SY, Nam YD, Nam DH, Lee JH, Koo BS, Kim H. A comparative study on the quality of sleep, tongue diagnosis, and oral microbiome in accordance to the Korean medicine pattern differentiation of insomnia. *Journal of Korean Medicine for Obesity Research*. 2020;20(1):40-51.
 22. Guo S, Wu G, Liu W, Fan Y, Song W, Wu J, Gao D, Gu X, Jing S, Shen Q, Ren L, Dong Y, Liu Z. Characteristics of human oral microbiome and its non-invasive diagnostic value in chronic kidney disease. *Bioscience Reports*. 2022;42(5):BSR20210694.