Journal of the Korea Academia-Industrial cooperation Society
Vol. 25, No. 7 pp. 320-328, 2024

https://doi.org/10.5762/KAIS.2024.25.7.320
ISSN 1975-4701 / eISSN 2288-4688

Predicting the association between metabolic syndrome and chronic kidney disease using the XGBoost model

Yoonjin Park
Department of Nursing, Joongbu University

XGBoost 모델을 이용한 대사증후군과 만성콩팥질환의 연관성 예측

박윤진
중부대학교 간호학과

Abstract This study aimed to estimate the accuracy of the XGBoost regression model to investigate the predictors of decreases in glomerular filtration rate. This study used data of 29,386 individuals aged ≥20 years who participated in the Korea National Health and Nutrition Examination Survey (KNHANES) 2016–2020 and in whom serum creatinine levels were analyzed. Prediction models were created with neural network machine learning and techniques using the XGBoost model. The areas under the curve (AUC) of the multivariable logistic prediction model, the logistic model score-based XGBoost model, and the XGBoost model for the prediction of important predictors were 0.929, 0.922 and 0.945 respectively. Dyslipidemia, hypertension, age, waist circumference, and fasting serum glucose levels were the main predictors of decreased glomerular filtration rate. The XGBoost regression model showed the highest predictive power. It can thus be employed in real-time health management for estimating the risk of major diseases and predicting possible complications.

요 약 본 연구는 XGBoost 회귀 모델을 이용한 분석을 통해 사구체여과율 감소의 영향 요인을 살펴보고 이를 통하여 machine learning의 기법의 적용 가능성과 정확도를 추정하는 것을 목적으로 하였다. 본 연구의 참여자는 2016년부터 2020년까지 국민건강영양조사 참여자를 대상으로 추출한 연구 모집단으로 하였으며 총 29,386명을 최종 분석하였다. 그중 Training cohort는 23,688명, validation cohort는 5,689명이었다. 본 연구의 결과 테스트 데이터 세트에 대한 다변량 로지스틱 예측 모델의 AUC는 0.929이며 logistic 모델 score를 바탕으로 XGBoost 분석 결과 AUC가 0.922로 최종 분석되었다. 또한, Important predictor를 예측하기 위한 XGBoost 모델 개발의 경우는 AUC가 0.945로 XGBoost의 예측력이 가장 높게 나타났다. 이 밖에 이상지질혈증, 고혈압, 나이, 허리둘레, 공복혈당 수치가 사구체 여과율 감소의 주요 예측 변수로 의미 있게 분석되었다. 본 연구의 결과를 통하여 신기능의 위험도를 예측하는 새로운 예측 모델을 개발하고 실시간 건강관리에 활용함으로써 향후 대사증후군을 가진 대상자의 신장기능의 위험도를 효과적으로 예측하기 위한 서비스프로그램을 개발을 제언한다.

Keywords : Glomerular Filtration Rate, Machine Learning, Metabolic Syndrome, Renal Insufficiency, Health Management

본 논문은 한국연구재단 연구과제로 수행되었음(과제번호: 2021R1G1A1092286)
*Corresponding Author : Yoonjin Park(Joongbu Univ.)
email: pyj2272@joongbu.ac.kr
Received May 29, 2024 Revised June 24, 2024
Accepted July 5, 2024 Published July 31, 2024

1. Introduction

According to the Health Insurance Review and Assessment Service in South Korea, the number of patients with chronic kidney disease (CKD) or chronic renal failure is increasing every year, with 203,978 patients in 2017; 226,877 in 2018; 249,283 in 2019; 259,116 in 2020; and, 277,252 in 2021 [1]. CKD is a condition in which the glomerular filtration rate is less than 60 mL /min/1.73㎡ or one in which markers of kidney damage persist for more than 3 months. The prevalence of CKD is more than 10% worldwide, with over 800 million people with CKD, and it is predicted to increase by 41.5% in 2040 compared to that in 1990 [2]. CKD is closely associated with the increasing prevalence of comorbidities such as hypertension, cardiovascular disease, diabetes mellitus, and obesity, which are chronic diseases in modern society [3]. In particular, there is a growing interest in the association between CKD and metabolic syndrome, a risk factor for cardiovascular disease [4,5]. Metabolic syndrome is an independent risk factor for CKD, and metabolic syndrome was associated with a 2.3-fold increased risk of proteinuria [6] and increased prevalence of CKD [7,8].

According to the American Heart Association/ National Heart, Lung, and Blood Institute (AHA/NHLBI), metabolic syndrome is defined as having at least 3 of 5 factors including abdominal obesity, hypertriglyceridemia, reduced high-density lipoprotein (HDL)-cholesterol level, elevated blood pressure, and elevated blood glucose level [9]. Several studies have revealed the relationship between metabolic syndrome and CKD. Patients with end-stage renal disease (ESRD) accompanied by diabetes mellitus accounted for more than half of all patients with CKD, and the burden of their medical expenses due to drug use and hospitalization was 1.6 times higher than that in those without diabetes mellitus [10]. In addition, diabetes and hypertension accounted for 45–50% and 25% of the major causes of ESRD, respectively, and patients with obesity and metabolic abnormalities among patients with CKD had a 1.4-fold increased risk of progression to ESRD, indicating the presence of an association between CKD and chronic diseases [11]. The prevalence of metabolic syndrome is increasing worldwide, and its prevalence in South Korea is approximately 1 in 4 persons, although there is an age difference in the prevalence [12] This metabolic syndrome is characterized by insulin resistance, which increases uric acid reabsorption in the kidney and blood uric acid concentrations; inhibits nitric oxide activity; and, promotes insulin resistance and the transport of free fatty acids from visceral obesity to the liver. Therefore, CKD is closely associated with metabolic syndrome and is both a cause and a consequence of metabolic syndrome [13,14].

Therefore, various specialized studies on the association between metabolic syndrome and decreased GFR are essential. Previous studies have often used traditional statistical analysis methods such as regression analysis and structural models. However, in recent years, various approaches through machine learning (ML) have been attempted, and the predictive performance is superior to that of conventional statistical modeling [15]. ML is a representative core technology of big data analysis and is a general term for a method of creating optimal learning models from a given training dataset, analyzing large amounts of data through the respective model, and deriving results such as inductive reasoning [16]. Among these models, XGBoost model is a model that uses the boosting technique of ensemble models and is commonly used as a method of making predictions by averaging learning and the prediction results of multiple models [17]. In particular, with this, it is easy to reduce errors when biomedical data are diverse and there are errors depending on patients' situation. In fact, the prediction accuracy was higher than 71–74

(%) in a study using Support Vector Machine, Random Forest, etc., as a classifier for patients with diabetic [18], and an analysis for predicting diabetic nephropathy also showed a high accuracy of 75% [19]. However, studies analyzing the association between metabolic syndrome and CKD using ML are scarce.

Therefore, this study aims to identify factors affecting decreased GFR through ML techniques; to examine the applicability, problems, and solutions of ML techniques; and, to provide basic data for the development of programs that can be used for the prevention of CKD in the health prediction field.

2. Measure

2.1 Data screening

This study used data of 29,386 individuals aged ≥20 years who participated in the 2016–2020 Korea National Health and Nutrition Examination Survey (KNHANES) and underwent serum creatinine level analyses (Fig. 1). General characteristics of the participants included sex and age, and the participants were classified by smoking status, which included cigarette and electronic cigarette smoking. According to the National Cholesterol Education Program Adult Treatment Panel III (NCEP-ATP III) and the Korean Society for the Study of Obesity, metabolic syndrome was defined by the presence of at least 3 of 5 risk factors from waist circumference ≥ 90 cm for men or ≥85 cm for women, fasting blood glucose ≥100 mg/dL, triglycerides 〉150 mg, HDL cholesterol 〈40 mg/dL for men or 〈50 mg/dL for women, and blood pressure≥130/85 mmHg [9]. This study included waist circumference, total cholesterol level, fasting blood glucose level, hypertension, and dyslipidemia in the analysis. In addition, stroke, cardiovascular disease, joint disease, and thyroid disease were added as past history of diseases affecting metabolic syndrome to perform data preprocessing, model learning, and prediction. Kidney function was calculated using the CKD-EPI formula based on creatinine levels using the estimated glomerular filtration rate (eGFR) and the patient's age, sex, and race.

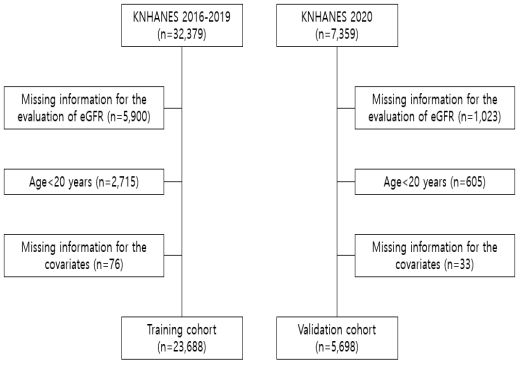


Fig. 1. Flow diagram for the inclusion of study participants

According to the Kidney Disease Improving Global Outcomes (KDIGO) practice guidelines by the National Kidney Foundation in the US, kidney function was classified into 5 groups: normal (eGFR ≥90 mL/min/1.73㎡); mildly decreased (60 mL/min/1.73㎡ ≤ eGFR〉 90 mL/min); mildly to moderately decreased (45 mL/min/1.73㎡ ≤ eGFR〉 60 mL/min); moderately to severely decreased (30 mL/min/1.73㎡ ≤ eGFR〉 45 mL/min/1.73㎡); and, severely decreased (eGFR 〈30 mL/min/1.73㎡) groups [20-22]. In this study, mildly to moderately decreased group (eGFR 〈60 mL/min/1.73㎡) or lower as a dependent variable and excluded some samples with missing information. Based on a total of 14 variables, prediction models were created with neural network ML and techniques using XGBoost model (eXtreme gradient boosting machine; the XGBoost model originates from a Research Project at the University of Washington, Seattle, WA, USA), and the course of learning was monitored to prevent overfitting. This study was approved by the Institutional Review Board of J University.

2.2 Dataset application

In this study, we selected sociodemographic characteristics (age, sex), metabolic risk factors (waist circumference, fasting serum glucose level, total cholesterol level), lifestyle behaviors (physical activity, alcohol consumption, smoking), comorbidities (hypertension, dyslipidemia), and medical history (stroke), coronary heart disease, arthritis, and thyroid disease; identified the important predictors; and, evaluated the predictive power of the XGBoost model after developing the logistic model. The analysis data for ML was classified into training data and test data. In addition, a training set was built with data from KNHANES 2016–2019, a test set was built with data from KNHANES 2020, and the machine learning classifier was trained with a supervised ML methodology that simultaneously received and learned the input variables of the trained set and the presence or absence of kidney disease (eGFR 〈60 mL/min/1.73㎡). At the time of reasoning, the presence or absence of disease was predicted for the test set, and performance was measured by comparing it with a true value, and XGBoost model was used as a classification model. XGBoost model is an algorithm that shows excellent performance in predicting structured data by combining a decision tree model with prediction models that allow for simple classification and applying boosting techniques that can create stronger prediction models. The XGBoost model can be used in regression models or classification models [23]. In this study, the area under the curve (AUC) method was employed in the training dataset and the validation dataset to evaluate the accuracy of the model.

3. Result

3.1 Descriptive characteristics

The study population extracted from the KNHANES 2016–2020 comprised 29,386 participants, of which 23,688 and 5,689 formed the training and validation cohorts, respectively (Table 1). The proportions of men and women in the training cohort were 44.2% and 55.8%, respectively, whereas those in the validation cohort were 44.9%, and 55.1%, respectively, indicating that the proportion of women was high. The mean age of those included in the analysis, except for those with missing information, was 52 and 54 years in the training cohort and the validation cohort, respectively. Waist circumference, a subfactor of metabolic syndrome, in the training cohort and the validation cohort was 82.8 cm and 85.0 cm, respectively; the fasting serum glucose level in the training cohort and the validation cohort was 96 mg/dL and 97 mg/dL, respectively; and, the total cholesterol level in the training cohort and the validation cohort was 191 mg/dL and 187 mg/dL, respectively. The prevalence of hypertension and dyslipidemia was 32.7% and 23.5% in the training cohort and 33.7% and 26.4% in the validation cohort, respectively (Table 1).

Table 1. Descriptive characteristics of the study population

Category		Training cohort (n=23,688)	Validation cohort (n=5,698)
Age, years		52 (39-65)	54 (39-66)
Sex, n (%)	Men	10,472 (44.2)	2,559(44.9)
	Women	13,216 (55.8)	3,139(55.1)
Waist circumference, cm		82.8 (75.5-89.7)	85.0 (77.3-91.6)
Fasting serum glucose, mg/dL		96 (89-105)	97 (90-106)
Total cholesterol, mg/dL		191 (167-216)	187(163-213)
Hypertension, n (%)	Yes	7,743 (32.7)	1,918 (33.7)
	No	15,945 (67.3)	3,780 (66.3)
Dyslipidemia, n (%)	Yes	5,567 (23.5)	1,505 (26.4)
No		18,121 (76.5)	4,193 (73.6)
Physical activity, n (%)	Yes	9,714 (41.0)	2,201 (38.6)
	No	12,813 (54.1)	3,024 (53.1)
	NA	1,161 (4.9)	473 (8.3)
Alcohol	Yes	20,871 (88.1)	4,975 (87.3)

consumption, n (%)	No	2,574 (10.9)	676 (11.9)
	NA	243 (1.0)	47 (0.8)
Smoking, n (%)	〈100 cigarettes	451 (1.9)	139 (2.4)
	≥100 cigarettes	9,004 (38.0)	2,116 (37.1)
	Never	13,967 (59.0)	3,392 (59.5)
Stroke		526 (2.2)	121 (2.1)
Coronary heart disease		668 (2.8)	177 (3.1)
Arthritis		3,000 (12.7)	694 (12.2)
Thyroid disease		886 (3.7)	239 (4.2)

* Data are median (interquartile range) unless indicated otherwise. *Acronyms: NA, not applicable.

3.2 Multivariable logistic regression to identify important predictors for renal failure

The multivariable logistic regression analysis revealed that age, waist circumference, fasting serum glucose level, total cholesterol level, hypertension, dyslipidemia, and drinking status were important predictors of CKD (Table 2). Those with eGFR 〈60 mL/min/1.73㎡ had an average 1.11-fold increased risk of CKD with increasing age, and those with elevated waist circumference had a 1.03-fold increased risk of CKD; elevated fasting glucose, 1.04-fold; and, elevated total cholesterol, 0.99-fold; elevated blood pressure, 1.38-fold; dyslipidemia, 0.81-fold; and, alcohol consumption, 0.81-fold increased risk of CKD (Table 2).

Table 2. Multivariable logistic regression to identify important predictors for renal failure

Category		Unit/ reference	OR (95% CI)	*P* value
Age		Years	1.11 (1.10-1.11)	〈.001
Sex, women		Men	0.89 (0.74-1.07)	.200
Waist circumference		cm	1.03 (1.02-1.04)	〈.001
Fasting serum glucose		mg/dL	1.04 (1.04-1.04)	〈.001
Total cholesterol		mg/dL	0.99 (0.99-0.99)	〈.001
Hypertension, yes		No	1.38 (1.20-1.57)	〈.001
Dyslipidemia, yes		No	1.55 (1.37-1.76)	〈.001
Physical activity	Yes	No	0.97 (0.85-1.11)	.646
	NA	No	0.63 (0.32-1.23)	.174
Alcohol consumption	Yes	No	0.81 (0.69-0.94)	.007
	NA	No	0.31 (0.11-0.94)	.038
Smoking	≥100 cigarettes	〈100 cigarettes	2.01 (0.98-4.12)	.058
	Never	〈100 cigarettes	1.92 (0.93-3.97)	.079
Stroke	Yes	No	0.99 (0.77-1.28)	.947
	NA	No	1.17 (0.13-10.64)	.887
Coronary heart disease		No	1.13 (0.91-1.41)	.276
Arthritis		No	1.08 (0.93-1.25)	.340
Thyroid disease		No	0.98 (0.72-1.35)	.914

* Note. OR = Odds Ratio; CI = Confidence Interval; NA = Not Applicable

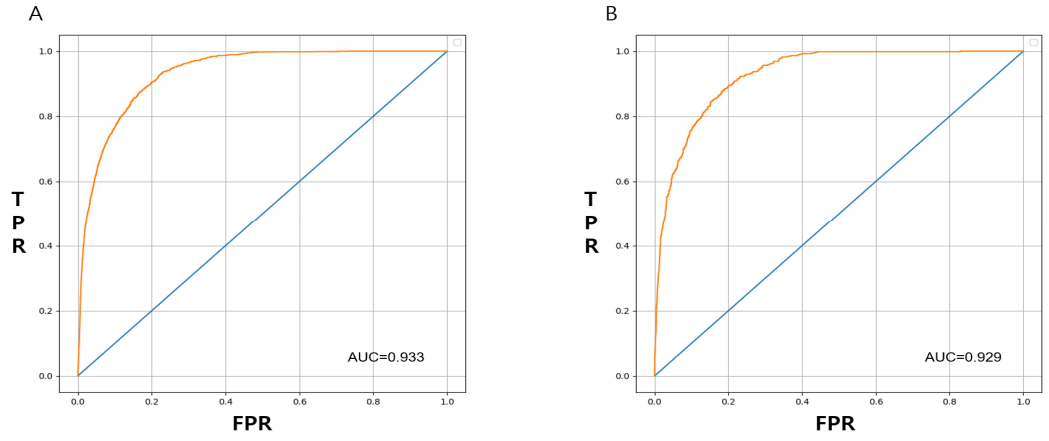
3.3 Multivariable logistic regression of important predictors to predict renal failure

The results of testing to identify important predictors of CKD according to the model

Table 3. Multivariable logistic regression of important predictors to predict renal failure

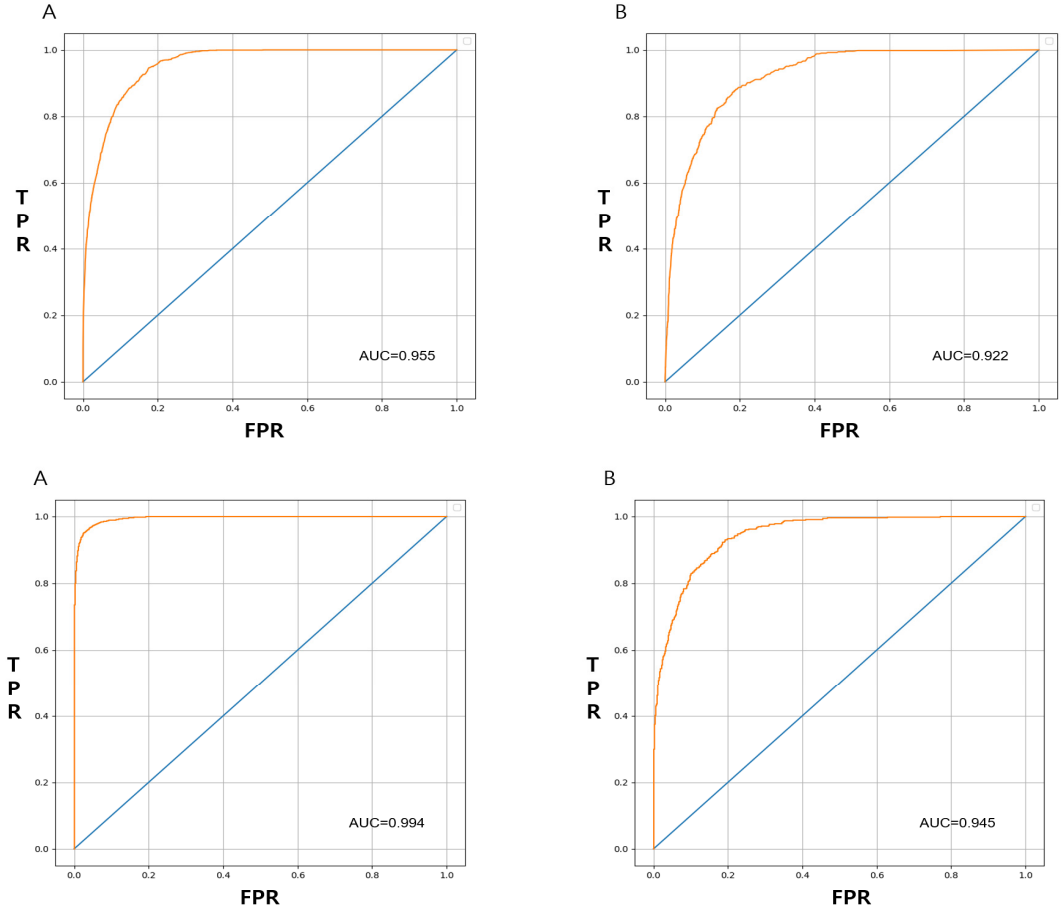
Category		Unit/reference	Estimate	OR (95% CI)	*P* value
Intercept			-14.516		
Age		Years	0.102	1.11 (1.10-1.12)	〈.001
Waist circumference		cm	0.032	1.03 (1.03-1.04)	〈.001
Fasting serum glucose		mg/dL	0.038	1.04 (1.04-1.04)	〈.001
Total cholesterol		mg/dL	-0.012	0.99 (0.99-0.99)	〈.001
Hypertension, yes		No	0.315	1.37 (1.20-1.57)	〈.001
Dyslipidemia, yes		No	0.434	1.54 (1.37-1.75)	〈.001
Alcohol consumption	Yes	No	-0.210	0.81 (0.69-0.95)	.008
	NA	No	-0.972	0.38 (0.13-1.10)	.075
Smoking	≥100 cigarettes	〈100 cigarettes	0.715	2.04 (1.00-4.19)	.051
	Never	〈100 cigarettes	0.596	1.82 (0.89-3.72)	.104

* Note. OR = Odds Ratio; CI = Confidence Interval; NA = Not Applicable



* Note: TPR = True Positive Rate; FPR = False Positive Rate. (A) Training cohort. (B) Validation cohort.

Fig. 1. Receiver operating characteristic curves of the important predictors-based multivariable logistic regression model for the prediction of renal failure



* Note: TPR = True Positive Tate; FPR = False Positive Rate. (A) Training cohort. (B) Validation cohort.

Fig. 2. Receiver operating characteristic curves of the logistic model score-based XGBoost model and receiver operating characteristic curves of the important predictors-based XGBoost model for the prediction of renal failure.

revealed that the odds ratio of CKD was the highest for dyslipidemia (OR: 1.54 CI: 1.37–1.75), followed by hypertension (OR: 1.37, CI: 1.20–1.57), age (OR: 1.11, CI: 1.10–1.12), waist circumference (OR: 1.11, CI: 1.10–1.12), and fasting serum glucose level (OR: 1.04, CI: 1.04–1.04) (Table 3). OR calculated using logistic regression. The AUC of the multivariable logistic prediction model for the test dataset was 0.929(Fig. 1). The AUC of the logistic model score-based XGBoost model was 0.922 and the AUC was 0.945 for the important predictor-based XGBoost model for the prediction of important predictors for CKD (Fig. 2).

4. Discussion

This study aimed to estimate the accuracy of the XGBoost model Regression model to investigate the predictors of decreased GFR. The AUC of the multivariable logistic prediction model was 0.929, the AUC of the logistic model score-based XGBoost model was 0.922, and the AUC for the development of the XGBoost model for the prediction of important predictors was 0.945, indicating that all models had high predictive power. Among these models, the XGBoost model regression model showed the highest predictive power; it also had the highest predictive power in a study using five classification algorithms: kNN, decision tree, LGBM, Voting, and XGBoost model for the quantitative prediction of decreased eGFR in patients with diabetes [19].

In addition, the results are similar to a previous study on diabetes prediction using algorithms in which the average accuracy was 86% [23] and to a study in which the average accuracy for the metabolic syndrome prediction algorithm was more than 85% [24].

Both the multivariable logistic prediction model and the XGBoost model regression model had more than 90% predictive power.

Among them, dyslipidemia, hypertension, age, alcohol consumption waist circumference, and fasting serum glucose level were the main variables for the feature importance of decreased GFR. This is similar to the results of a previous study showing that the risk factors for proteinuria were blood pressure, hyperglycemia, and hypertriglyceridemia, and that the comparative risk of proteinuria for metabolic syndrome was 2.3 times higher [6].

In particular, dyslipidemia was found to be the most important factor for CKD. This is not unrelated to obesity and can be explained by the results of previous studies showing that the prevalence of metabolic syndrome in those with severe obesity was 30.0 times higher; with obesity, 10.0 times higher; and, with overweight, 3.0 times higher compared to those with normal weights.

Among the underlying diseases in subjects with metabolic syndrome, hypertriglyceridemia was the most common at 38.9%, followed by reduced HDL cholesterol at 35.9%, hypertension at 35.4%, hyperglycemia at 34.8%, and abdominal obesity at 27.5%, indicating that dyslipidemia was the most important cause of metabolic syndrome [25]. Considering that the ages of these study participants were diverse, the results of this study may differ from those of other studies in which participants were classified by age. However, considering that dyslipidemia has a strong influence on CKD in the entire adult group, education on the prevention of dyslipidemia including obesity should be included in health examination items and target education. In particular, dyslipidemia is closely related to diabetes and is known to be a major cause of CKD [26]. Diabetic nephropathy may lead to progression to renal hypertrophy, decreased glomerular filtration rate, severe proteinuria, and end-stage renal disease [27]. In particular, since patients with type 2 diabetes may not have albuminuria, and GFR can be decreased by

various factors including hyperglycemia, insulin resistance, proteinuria, advanced glycation end products, and oxidative stress, various approaches for early diagnosis are needed [28]. In this study, fasting serum glucose was also found to be a major predictor for CKD, indicating its importance.

However, this study found that the estimates of alcohol consumption and total cholesterol were -0.210 and -0.012, respectively ($p<0.05$). The mean age of the participants in this study corresponds to the age at which atherosclerosis -induced dyslipidemia accompanied by high LDL cholesterol level, hypertriglyceridemia, and low HDL cholesterol level are common [29]. Considering this, further studies are needed. However, It is very meaningful to confirm that XGBoost is suitable for medical big data.

However, based on the results of this study, a new model for predicting the risk of CKD was built, and the ML model was provided as a web service and used for real-time health management, thereby estimating the risk of major diseases and predicting possible complications. Hence, the development of service programs in the future is suggested to effectively predict the risk of kidney function in individuals with metabolic syndrome.

References

[1] Health insurance review & assessment service. Statistical indicators of medical expenses (2021) [Internet]. [cited January 21, 2022], Available From: https://www.hira.or.kr//bbsDummy.do?pgmid=HIRAA020045020000 (Accessed January 21, 2022).

[2] Kovesdy CP. "Epidemiology of chronic kidney disease: an update 2022", *Kidney International Supplements,* Vol.12, No.1, pp.7-11, Apr, 2022. DOI: https://doi.org/10.1016/j.kisu.2021.11.003

[3] Provenzano M, Coppolino G, Faga T, Garofalo C, Serra R, Andreucci M. "Epidemiology of cardiovascular risk in chronic kidney disease patients: the real silent killer", *Reviews in cardiovascular medicine,* Vol. 20, No.4, pp.209-220, Dec, 2019. DOI: https://doi.org/10.31083/j.rcm.2019.04.548

[4] Obunai K, Jani S, Dangas GD. "Cardiovascular morbidity and mortality of the metabolic syndrome", *Medical Clinics of North America,* Vol.91, No.6, pp.1169-1184, Nov, 2007. DOI: https://doi.org/10.1016/j.mcna.2007.06.003

[5] Wang H, Zhang H, Sun L, Guo W. "Roles of hyperuricemia in metabolic syndrome and cardiac-kidney-vascular system diseases", *American Journal of Translational Research.* Vol.10, No.9, pp.2749, Sep, 2018.

[6] Oh MJ, Lee MJ, Lee KH, Park MS, Ko JA. "Association of Proteinuria with Metabolic Syndrome in Korean Adults", *Korean Journal of Family Medicine*, Vol.31, No.6, pp.437-443, Jun, 2010. DOI: https://doi.org/10.4082/kjfm.2010.31.6.437

[7] Singh AK, Kari JA. "Metabolic syndrome and chronic kidney disease", *Current opinion in nephrology and hypertension*. Vol.22, No.2, pp.198-203, Mar, 2013. DOI: https://doi.org/10.1097/MNH.0b013e32835dda78

[8] Mok CC. "Metabolic syndrome and systemic lupus erythematosus: the connection", *Expert Review of Clinical Immunology,* Vol.12, No.7, pp.765-775, May, 2019. DOI: https://doi.org/10.1080/1744666X.2019.1620601

[9] Ha K, Song Y. "Associations of meal timing and frequency with obesity and metabolic syndrome among Korean adults", *Nutrients*, Vol.11, No.10, PP.2347, Oct, 2019. DOI: https://doi.org/10.3390/nu11102437

[10] Seo HJ. "Effects for Comorbidities of Chronic Kidney Disease on the Progression to End-stage Renal Disease", *Journal of Health Informatics and Statistics*, Vol.45, No.4, pp.356-364, Nov, 2020. DOI: DOI: https://doi.org/10.21032/jhis.2020.45.4.356

[11] Ahn C. 10-year longitudinal study on the patient survival and renal survival by specific causes of the chronic kidney diseases, Seoul: Seoul National University Hospital, Korea Disease Control and Prevention. 2016.

[12] Park E, Kim J. "Gender-and age-specific prevalence of metabolic syndrome among Korean adults: analysis of the fifth Korean National Health and Nutrition Examination Survey", *Journal of Cardiovascular Nursing*, Vol.30, No.3, pp.256-266, May, 2015. DOI: https://doi.org/10.1097/JCN.0000000000000142

[13] Nakagawa T, Hu H, Zharikov S, Tuttle KR, Short RA, Glushakova O, et al. "A causal role for uric acid in fructose-induced metabolic syndrome", *American Journal of Physiology-Renal Physiology*, Vol.290, No.3, pp.625-631, Mar, 2006. DOI: https://doi.org/10.1152/ajprenal.00140.2005

[14] King C, Lanaspa MA, Jensen T, Tolan DR, Sánchez-Lozada LG, Johnson RJ. "Uric acid as a cause of the metabolic syndrome", *Uric Acid in Chronic Kidney Diseas,* Vol.192, pp. 88-102, 2018. DOI: https://doi.org/10.1159/000484283

[15] Yu C-S, Lin Y-J, Lin C-H, Wang S-T, Lin S-Y, Lin SH, et al. "Predicting metabolic syndrome with machine learning models using a decision tree algorithm: retrospective cohort study", *JMIR medical informatics*, Vol.20, No.3, pp.e17110, Mar, 2020.
DOI: https://doi.org/10.2196/17110

[16] Lee JG, Lee TH, Yun SR. "Machine Learning for Big Data Analysis", I*nformation and Communications Magazine*, Vol.31, No.11, pp.14-26, 2014.

[17] Chen T, Guestrin C, editors. XGBoost model: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016.

[18] Sarwar MA, Kamal N, Hamid W, Shah MA. "Prediction of diabetes using machine learning algorithms in healthcare", *2018 24th international conference on automation and computing (ICAC); 2018: IEEE.*
DOI: https://doi.org/10.23919/IConAC.2018.8748992

[19] Park, Y.J, Kang, H.K. "Performance comparison of various classification algorithms of machine learning applications for predicing debetic nephropathy:, *Journal of the Korea Academia-Industrial Cooperation Society*, Vol.23, No.7, pp.184-191. Jul, 2022.
DOI: https://doi.org/110.5762/KAIS.2022.23.7.184

[20] Stevens LA, Coresh J, Greene T, Levey AS. "Assessing kidney function—measured and estimated glomerular filtration rate", *New England Journal of Medicine*, Vol.654, No.23, pp.2473-2483, Jun, 2006.
DOI: https://doi.org/10.1056/NEJMra054415

[21] National Kidney Foundation. Estimated Glomerular Filtration Rate(eGFR)(2023) [Internet] [cited May, 21, 2023], Available From: https://www.kidney.org/atoz/content/gfr [Accessed May 21, 2023].

[22 The Korean society of nephrology. Estimated Glomerular Filtration Rate(eGFR)(2023) [Internet]. [cited May, 21, 2023], Available from: https://ksn.or.kr/general/about/check.php [Accessed May 21, 2023].

[23] Li W, Yin Y, Quan X, Zhang H. "Gene expression value prediction based on XGBoost model algorithm", *Frontiers in genetics*, Vol.10, pp.1077, Nov, 2019.
DOI: https://doi.org/10.3389/fgene.2019.01077

[24] Kim J, Mun S, Lee S, Jeong K, Baek Y. "Prediction of metabolic and pre-metabolic syndromes using machine learning models with anthropometric, lifestyle, and biochemical factors from a middle-aged population in Korea", *BMC Public Health*, Vol.22, No.1, pp.1-10, Apr, 2022.
DOI: https://doi.org/10.1186/s12889-022-13131-x

[25] Park E. "The obesity and the metabolic syndrome among Korean adults: based on National Representative Survey", *Korean Journal of Health Promotion*. Vol.20, No.4, pp.144-150, Dec, 2020.
DOI: https://doi.org/10.15384/kjhp.2020.20.4.144

[26] Gurruchaga AM. "Consecuencias patológicas de la obesidad: hipertensión arterial, diabetes mellitus y dislipidemia", *ARS MEDICA Revista de Ciencias Médicas*, Vo..26, 1997.

[27] Association AD. "11. Microvascular complications and foot care: standards of medical care in diabetes—2019", *Diabetes Care*, Vol.42(Supplement_1), pp.S124-S38, Dec, 2019.
DOI: https://doi.org/10.2337/dc20-S011

[28] Yamazaki T, Mimura I, Tanaka T, Nangaku M. "Treatment of diabetic kidney disease: current and future", *Diabetes & Metabolism Journal*, Vol.45, No.1, pp.11-26, Jan, 2021.
DOI: https://doi.org/10.4093/dmj.2020.0217

[29] Kim Y, Han AL. "The correlation between triglyceride to HDL cholesterol ratio and metabolic syndrome, nutrition intake in Korean adults: Korean National Health and Nutrition Examination Survey 2016", *Journal of Nutrition and Health*, Vol.52, No.3, pp.268-276, Jun, 2019.
DOI: https://doi.org/10.4163/jnh.2019.52.3.268

Yoonjin Park [Regular member]

- Feb. 1997 : Korea armed forces nursing academy, Bachelor's degree
- Sep. 2017 : Catholic Univ. PhD
- Sep. 2017~ current : Joonbu Univ., Dept. of Nursing, Professor

〈Research Interests〉
Rehabilitation nursing, Hospice nursing