Original Article

Kidney Res Clin Pract [Epub ahead of print]
pISSN: 2211-9132 • eISSN: 2211-9140
https://doi.org/10.23876/j.krcp.25.052

KIDNEY RESEARCH AND CLINICAL PRACTICE

Machine learning-based utilization of lipid panels in predicting kidney dysfunction

Soie Kwon[1,*], Donghwan Yoon[2,*], Changhwa Park[3,*], Sehoon Park[2], Yong Chul Kim[2], Dong Ki Kim[2], Kook-Hwan Oh[2], Kwon Wook Joo[2], Yon Su Kim[2], Sungroh Yoon[3], Seung Seok Han[2]

[1]*Department of Internal Medicine, Chung-Ang University, Seoul, Republic of Korea*
[2]*Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea*
[3]*Department of Electrical and Computer Engineering, Seoul National University, Seoul, Republic of Korea*

Background: Dyslipidemia is a risk factor for atherosclerotic cardiovascular disease, but its relationship with kidney dysfunction varies depending on the patient and kidney status. Herein, we used a machine learning model to select lipid panels to improve the predictability of kidney dysfunction.

Methods: A total of 9,403 patients whose lipid panels and low-density lipoprotein subfraction scores were measured were enrolled. The primary outcome was kidney outcome, which was defined as the development of end-stage kidney disease or a 50% decline in kidney function. The secondary outcome was a composite outcome, defined as the occurrence of either the kidney outcome or all-cause death. Five machine learning models were utilized to predict 1- and 3-year outcomes. Feature ranking was employed to identify key factors contributing to model performance.

Results: At 3 years, 117 patients (1.2%) experienced kidney outcomes, and 691 (7.4%) experienced composite outcomes. All of the models that used all the features demonstrated high predictive power, with the area under the receiver operating characteristic curve (AUROC) exceeding 0.85. Although the overall performance of the model that selectively uses lipid panels and creatinine was lower than that of models using all features, its predictive ability for kidney outcomes remained compatible, with AUROC values exceeding 0.78. Feature ranking analysis indicated that apolipoprotein A1, apolipoprotein B, and low-density lipoprotein cholesterol were significant contributors to model performance.

Conclusion: Machine learning models incorporating lipid panels successfully predict the risk of kidney dysfunction, which can help clinicians in precisely identifying patients who are at risk of kidney dysfunction through lipid panel assessments.

Keywords: Apolipoproteins, Cholesterol, Chronic kidney disease, Lipids, Machine learning

Received: February 17, 2025; **Revised:** July 15, 2025; **Accepted:** July 31, 2025

Correspondence: Seung Seok Han
Department of Internal Medicine, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea.
E-mail: hansway7@snu.ac.kr
ORCID: https://orcid.org/0000-0003-0137-5261

Sungroh Yoon
Department of Electrical and Computer Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea.
E-mail: sryoon@snu.ac.kr
ORCID: https://orcid.org/0000-0002-2367-197X

*Soie Kwon, Donghwan Yoon, and Changhwa Park contributed equally to this study as co-first authors.

© 2025 by The Korean Society of Nephrology
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial and No Derivatives License (https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits unrestricted non-commercial use, distribution of the material without any modifications, and reproduction in any medium, provided the original works properly cited.

Introduction

Dyslipidemia is a well-established risk factor for atherosclerotic cardiovascular disease (ASCVD) [1,2], with several studies demonstrating an association between altered levels of serum lipid panels, such as total cholesterol, low-density lipoprotein (LDL) cholesterol, triglyceride (TG), and high-density lipoprotein (HDL) cholesterol, and the risk of ASCVD [3–5]. Among the various forms of LDL cholesterol, small dense particles pose a greater atherogenic risk than large buoyant particles do [6–8]. Patients with chronic kidney disease (CKD) exhibit hypertriglyceridemia and low HDL cholesterol levels [9,10]. However, total and LDL cholesterol levels can vary depending on the underlying cause of CKD. Patients with nephrotic syndrome and diabetic kidney disease commonly experience elevated total and LDL cholesterol levels, whereas those with non-nephrotic CKD may have lower cholesterol levels [11]. Despite fluctuations in total LDL cholesterol levels, these patients generally have a greater proportion of small, dense LDL cholesterol [9,11].

Because ASCVD and CKD share common risk factors, previous studies have explored the associations between various lipid panels and kidney dysfunction. While some studies have demonstrated a significant association [12,13], others have not confirmed this link [14,15]. Establishing the association between lipid panels and kidney dysfunction via conventional analysis is challenging, particularly in patients with advanced CKD, because comorbid conditions in these patients may confound the results [16,17]. The continuous spectrum of lipoproteins, varying in size, composition, and apolipoprotein content, and the intercorrelations between lipid panels further complicate simultaneous analysis. Therefore, a comprehensive approach that includes lipid panels and subfractions, comorbidities, other laboratory findings, and medications is necessary to better understand these complex associations.

Machine learning models have recently been used to predict clinical outcomes by comprehensively considering a large number of features without the issue of intercorrelation [18]. Feature ranking in these models can provide clinical insights by indicating the extent to which a feature contributes to model performance [19]. There is a growing trend of utilizing machine learning models in nephrology, particularly for predicting kidney dysfunction, including conditions such as acute kidney injury and diabetic kidney disease [20–24]. Herein, we addressed whether machine learning models incorporating lipid panels, including both conventional and subfraction parameters, could predict kidney dysfunction. We also applied a machine learning-based feature ranking method to identify lipid panels that were primarily associated with kidney dysfunction.

Methods

Patients and study design

A cohort from a single tertiary referral center (Seoul National University Hospital) was retrospectively evaluated. A total of 14,123 adult patients (aged ≥18 years) who underwent serum LDL subfraction testing along with conventional lipid panels were reviewed between 2009 and 2016. Patients with end-stage kidney disease (ESKD; e.g., hemodialysis, peritoneal dialysis, and kidney transplantation; n = 591) or those who had received organ transplants other than kidneys (n = 15) at the time of enrollment were excluded. Patients with missing values for serum creatinine and lipid parameters were excluded (n = 4,128). Finally, 9,389 patients were included in the analysis (Supplementary Fig. 1, available online). All the patients were randomly assigned to training, validation, or test groups at a ratio of 7:1:2.

Features for model development

Demographic information, such as age, sex, body mass index, and smoking status, was collected at the time of enrollment. Comorbidities, such as diabetes mellitus, hypertension, CKD, and coronary artery disease, were defined on the basis of the International Classification of Diseases, 10th Revision (ICD-10) codes (Supplementary Table 1, available online). The serum LDL subfraction was measured once at enrollment, while other lipid panels (total cholesterol, LDL cholesterol, HDL cholesterol, TG, apolipoprotein A1 [ApoA1], and apolipoprotein B [ApoB]) were repeatedly tested. Additional blood parameters, such as creatinine, white blood cell count, hemoglobin, glucose, and cystatin C, were also collected. The estimated glomerular filtration rate (eGFR) was calculated via the CKD-EPI (Chronic Kidney Disease Epidemiology Collab-

oration) equation [25]. Prescription information, such as antidyslipidemia drugs (statin, ezetimibe, fibrate, omega-3), antiplatelets (aspirin, clopidogrel, prasugrel, ticagrelor), anticoagulants (warfarin, direct oral anticoagulant, low-molecular-weight heparin), nitrates, oral antidiabetic drugs, insulin, protein pump inhibitors, H_2 receptor antagonists, renin-angiotensin system blockers, beta blockers, calcium channel blockers, minoxidil, loop diuretics, thiazide, spironolactone, xanthine oxidase inhibitors, and erythropoietin-stimulating agents, was summarized at the time of enrollment.

Study outcomes

Kidney dysfunction, as the primary outcome, was defined by a >50% decline in eGFR, a doubling of serum creatinine from baseline, or the occurrence of ESKD (i.e., the need for dialysis or transplantation). The secondary outcome was a composite outcome, defined as the occurrence of either kidney dysfunction or all-cause death. Outcome data were collected as of March 31, 2020. Information on ESKD was obtained from the Korean Society of Nephrology database, and mortality data were retrieved from the Korean Statistical Information Service. Considering that the latest enrollment occurred in 2016, a minimum follow-up period of 3 years was available. Therefore, outcomes were assessed at discrete time points of 1, 2, and 3 years after enrollment, with censoring applied at each time point. Both primary and secondary outcomes were predicted based on these intervals.

Statistical analysis and model development

Statistical analyses were performed via R (version 3.6.2, R Foundation for Statistical Computing) and Python (version 3.9.16, Python Software Foundation). The features are expressed as frequencies and percentages for categorical parameters, means ± standard deviations for normally distributed continuous parameters, and medians and interquartile ranges (IQRs) for nonnormally distributed continuous parameters. Baseline characteristics were compared using the chi-square test for categorical parameters, the Student t test for normally distributed continuous parameters, and the Mann-Whitney U test for nonnormally distributed continuous parameters.

Prediction models were developed using a training set with machine learning algorithms, such as logistic regression (LR), random forest (RF), and light gradient boosting (LGBM), as well as deep learning algorithms, such as multilayer perceptron (MLP) and recurrent neural network (RNN). The models were trained using the training set, and the optimal hyperparameters for LR, RF, and LGBM, along with the training epochs for early stopping in MLP and RNN, were determined based on validation set results. A list of hyperparameters for the LR, RF, and LGBM, as well as detailed methods for the MLP and RNN, is provided in the Supplementary Table 2 (available online) and Supplementary Methods (available online). For LR, we performed a grid search over two regularization penalties (L1 and L2) and four inverse regularization strengths (C = 0.01, 0.1, 1, 10); elastic-net regularization was deliberately excluded to avoid an additional tuning parameter and preserve model interpretability. The Python code used for the machine learning analyses has been made publicly available at the following repository (https://github.com/dactylogram/lipid_panel).

Model performance was assessed using the area under the receiver operating characteristic curve (AUROC) and area under the precision–recall curve (AUPRC) in the testing dataset. Feature importance was assessed in two ways: individually for each feature and grouped by category (demographics, comorbidities, creatinine, lipid panels, other laboratory findings, and medications). Lipid panels that ranked highly in feature importance were further validated using a multivariable Cox regression model, and hazard ratios (HRs) and 95% confidence intervals (CIs) were estimated. A two-sided p-value of less than 0.05 was considered statistically significant.

Ethics statement

The study was approved by the International Review Board (IRB) of Seoul National University Hospital (No. H-1807-013-955) and was conducted in accordance with the principles of the Declaration of Helsinki. In terms of ethics approval, the requirement for informed consent was waived by the IRB.

Results

Cohort characteristics

The average age of the patients was 64.8 ± 10.9 years, and 64.7% were male (Table 1). The patients who developed kidney or composite outcomes tended to be older and had a higher prevalence of diabetes mellitus and CKD. The proportion of patients diagnosed with coronary artery disease and the prescription rates of statin, aspirin, and nitrate

Table 1. Baseline characteristics of the study subjects according to kidney and composite outcomes

Characteristic	Total	Kidney outcome			Composite outcome		
		No	Yes	p-value	No	Yes	p-value
No. of patients	9,389	9,202	187		7,892	1,497	
Age (yr)	64.8 ± 10.9	64.7 ± 10.9	67.0 ± 10.6	0.004	63.5 ± 10.6	71.5 ± 9.8	<0.001
Male sex	6,071 (64.7)	5,964 (64.8)	107 (57.2)	0.04	5,101 (64.6)	970 (64.8)	0.93
Body mass index (kg/m^2)	24.2 ± 3.3	24.2 ± 3.3	23.4 ± 4.2	0.02	24.5 ± 3.1	22.4 ± 3.7	<0.001
Current smoker	1,361 (14.5)	1,337 (14.5)	24 (12.8)	0.58	1,164 (14.7)	197 (13.2)	0.12
Comorbidities							
Diabetes mellitus	994 (10.6)	947 (10.3)	47 (25.1)	<0.001	774 (9.8)	220 (14.7)	<0.001
Hypertension	1,430 (15.2)	1,393 (15.1)	37 (19.8)	0.10	1,187 (15.0)	243 (16.2)	0.26
Chronic kidney disease	1,309 (13.9)	1,185 (12.9)	124 (66.3)	<0.001	753 (9.5)	556 (37.1)	<0.001
Coronary artery disease	6,014 (64.1)	5,910 (64.2)	104 (55.6)	0.02	5,136 (65.1)	878 (58.7)	<0.001
Laboratory findings							
WBC counts (×10^3/L)	6.9 ± 2.6	6.9 ± 2.5	7.5 ± 3.9	0.06	6.8 ± 2.5	7.3 ± 3.1	<0.001
Hemoglobin (g/dL)	13.3 ± 1.8	13.3 ± 1.8	11.6 ± 2.0	<0.001	13.5 ± 1.7	12.2 ± 2.0	<0.001
Glucose (mg/dL)	125.3 ± 51.2	125.1 ± 50.9	139.1 ± 63.9	0.003	124.1 ± 49.2	131.8 ± 60.2	<0.001
Albumin (g/dL)	4.1 ± 0.5	4.1 ± 0.5	3.8 ± 0.6	<0.001	4.2 ± 0.4	3.8 ± 0.6	<0.001
Creatinine (mg/dL)	0.9 ± 0.4	0.9 ± 0.3	1.8 ± 1.1	<0.001	0.9 ± 0.3	1.2 ± 0.7	<0.001
eGFR (mL/min/1.73 m^2)	80.6 ± 19.1	81.3 ± 18.4	49.8 ± 26.6	<0.001	83.3 ± 16.9	66.5 ± 23.3	<0.001
Cystatin C (mg/L)	0.9 ± 0.3	0.9 ± 0.3	1.5 ± 0.8	<0.001	0.8 ± 0.2	1.2 ± 0.6	<0.001
Lipid panel							
Total (mg/dL)	152 (130–178)	152 (130–178)	152 (131–181)	0.66	153 (131–179)	147 (125–174)	<0.001
LDL (mg/dL)	86 (68–109)	86 (68–109)	86 (67–113)	0.93	86 (68–109)	83 (66–106)	0.003
HDL (mg/dL)	37 (31–44)	37 (31–44)	32 (27–39)	<0.001	37 (31–44)	35 (28–42)	<0.001
Triglyceride (mg/dL)	105 (78–146)	105 (78–146)	124 (92–166)	<0.001	106 (79–148)	101 (74–136)	<0.001
Apolipoprotein A1 (mg/dL)	109 (98–121)	109 (98–121)	104 (91–117)	<0.001	110 (99–121)	104 (92–118)	<0.001
Apolipoprotein B (mg/dL)	77 (65–93)	77 (65–92)	81 (70–99)	0.003	78 (65–93)	75 (63–91)	0.001
LDL subfraction (mg/dL)	1.6 (1.4–1.8)	1.6 (1.4–1.8)	1.6 (1.4–1.9)	0.13	1.6 (1.4–1.9)	1.5 (1.3–1.8)	<0.001
Medications (%)							
Statin	1,797 (19.1)	1,739 (18.9)	58 (31.0)	<0.001	1,425 (18.1)	372 (24.8)	<0.001
Omega-3	565 (6.0)	548 (6.0)	17 (9.1)	0.10	497 (6.3)	68 (4.5)	0.01
Aspirin	1,934 (20.6)	1,870 (20.3)	64 (34.2)	<0.001	1,481 (18.8)	453 (30.3)	<0.001
Other antiplatelets	482 (5.1)	469 (5.1)	13 (7.0)	0.33	367 (4.7)	115 (7.7)	<0.001
Nitrate	1,974 (21.0)	1,918 (20.8)	56 (29.9)	0.003	1,564 (19.8)	410 (27.4)	<0.001
Proton pump inhibitor	514 (5.5)	505 (5.5)	9 (4.8)	0.81	384 (4.9)	130 (8.7)	<0.001
RAS blocker	1,019 (10.9)	974 (10.6)	45 (24.1)	<0.001	731 (9.3)	288 (19.2)	<0.001
Beta blocker	1,175 (12.5)	1,133 (12.3)	42 (22.5)	<0.001	872 (11.0)	303 (20.2)	<0.001
Calcium channel blocker	957 (10.2)	919 (10.0)	38 (20.3)	<0.001	742 (9.4)	215 (14.4)	<0.001

Data are expressed as number only, number (%), mean ± standard deviation, or median (interquartile range).
eGFR, estimated glomerular filtration rate; HDL, high-density lipoprotein; LDL, low-density lipoprotein; RAS, renin-angiotensin system; WBC, white blood cell.

were notably high (Table 1). This was largely because lipid subfraction scoring was primarily performed in patients with coronary artery disease to assess their prognosis. Serum cystatin C results were available for the majority of patients (92.8%). Diseases with a prevalence rate of less than 10% and drugs with a prescription rate of less than 5% are presented in Supplementary Table 3 (available online). Baseline characteristics of the study subjects by study groups (training, validation, and test sets) are presented in Supplementary Table 4 (available online).

Model performance

The median follow-up duration was 6 years (IQR, 4.3–8.6 years). During this period, kidney dysfunction occurred in 187 patients (2.0%), and the composite outcome was observed in 1,497 patients (15.9%). Within the first year after enrollment, kidney dysfunction occurred in 54 patients (0.6%), followed by 87 patients (0.9%) in the second year, and 117 patients (1.2%) in the third year. A total of 281 patients (3.0%) experienced the composite outcome in the first year, 489 patients (5.2%) in the second year, and 691 patients (7.4%) in the third year.

When the performance of each model was compared in terms of AUROC, the short-term (1-year) predictive power tended to be greater than the long-term (3-year) predictive power for kidney outcomes (Fig. 1, Table 2; Supplementary Table 5, available online). However, for the composite outcome, the short-term and long-term predictive powers were relatively similar (Fig. 2, Table 2; Supplementary Table 5, available online). Additionally, the trend in AUPRC for the composite outcome tended to increase more in long-term predictions than in short-term predictions. The minimum and maximum values of AUPRC with all features ranged from 0.182 to 0.250 at 1 year and from 0.285 to 0.327 at 3 years; when lipid panels and creatinine were used, they ranged from 0.158 to 0.243 at 1 year and from 0.256 to 0.294 at 3 years (Table 2; Supplementary Table 5, available online; Supplementary Figs. 2 and 3, available online).

Overall, model performance was better when all features

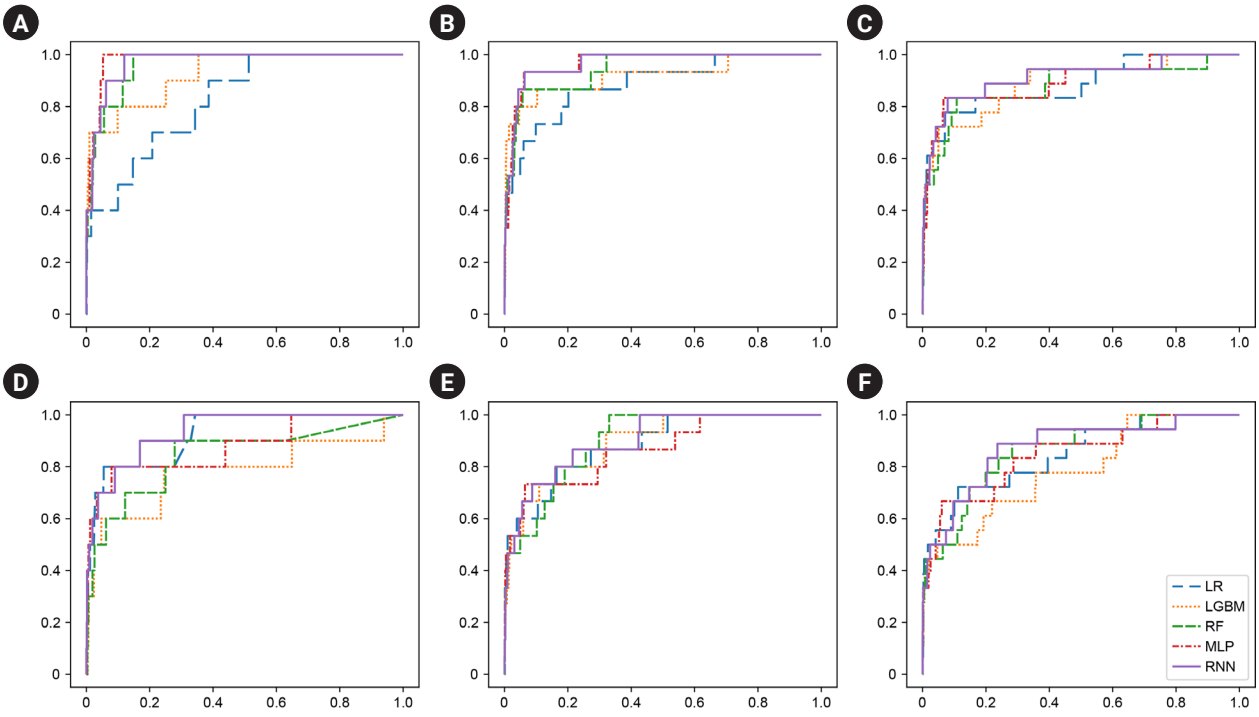

Figure 1. Area under the receiver operating characteristic curve for predicting kidney dysfunction. (A–C) Model with all features for 1-year (A), 2-year (B), and 3-year (C) prediction. (D–F) Model with lipid plus serum creatinine for 1-year (D), 2-year (E), and 3-year (F) prediction. The x-axis represents the false positive rate (1 − specificity), and the y-axis represents the true positive rate (sensitivity). LR, logistic regression; LGBM, light gradient boosting machine; RF, random forest; MLP, multilayer perceptron; RNN, recurrent neural network.

Table 2. Model performance for predicting kidney and composite outcomes based on 10-fold cross-validation by mean AUROCs and mean AUPRCs

Model	Time (yr)	Kidney outcome				Composite outcome			
		Model 1		Model 2		Model 1		Model 2	
		Mean AUROC	Mean AUPRC	Mean AUROC	Mean AUPRC	Mean AUROC	Mean AUPRC	Mean AUROC	Mean AUPRC
Logistic regression	1	0.886	0.250	0.845	0.243	0.886	0.321	0.726	0.178
	2	0.893	0.280	0.811	0.253	0.869	0.413	0.735	0.250
	3	0.831	0.302	0.792	0.283	0.865	0.472	0.727	0.298
LGBM	1	0.870	0.202	0.786	0.158	0.892	0.322	0.703	0.136
	2	0.866	0.302	0.826	0.240	0.861	0.376	0.721	0.220
	3	0.830	0.324	0.784	0.256	0.859	0.452	0.701	0.262
Random forest	1	0.918	0.215	0.828	0.217	0.886	0.309	0.739	0.176
	2	0.927	0.314	0.833	0.289	0.875	0.405	0.742	0.231
	3	0.883	0.327	0.807	0.279	0.869	0.461	0.727	0.287
RNN	1	0.910	0.182	0.871	0.228	0.888	0.316	0.748	0.185
	2	0.910	0.268	0.828	0.252	0.872	0.388	0.758	0.254
	3	0.880	0.305	0.801	0.261	0.864	0.443	0.764	0.304
MLP	1	0.894	0.182	0.840	0.179	0.876	0.283	0.732	0.164
	2	0.888	0.253	0.865	0.240	0.872	0.393	0.741	0.241
	3	0.859	0.285	0.822	0.294	0.865	0.446	0.734	0.300

Model 1: all the features are used. Model 2: serum creatinine and lipid panels are used.
AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve; LGBM, light gradient boosting; RNN, recurrent neural network; MLP, multilayer perceptron.

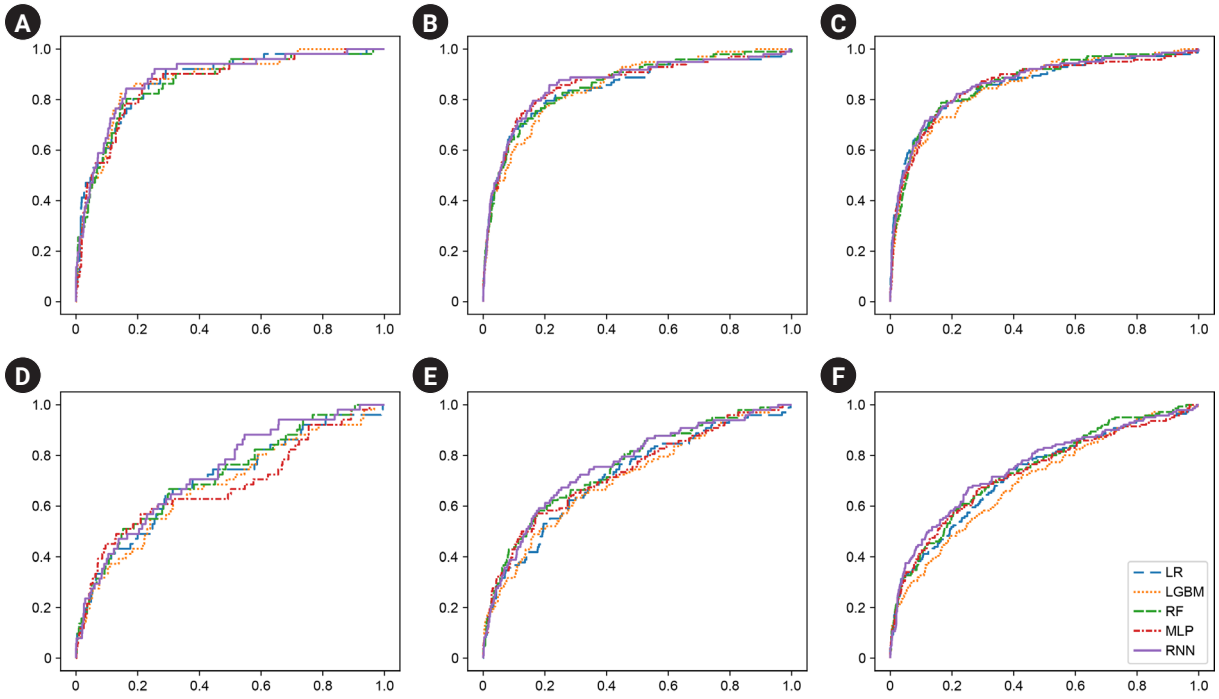

Figure 2. Area under the receiver operating characteristic curve for predicting composite outcomes. (A–C) Model with all features for 1-year (A), 2-year (B), and 3-year (C) prediction. (D–F) Model with lipid plus serum creatinine for 1-year (D), 2-year (E), and 3-year (F) prediction. The x-axis represents the false positive rate (1 − specificity), and the y-axis represents the true positive rate (sensitivity).
LR, logistic regression; LGBM, light gradient boosting machine; RF, random forest; MLP, multilayer perceptron; RNN, recurrent neural network.

were used than when only lipid panels and creatinine were used. For the 1-year prediction of kidney outcomes, the AUROC ranged from 0.870 to 0.918, and the AUPRC ranged from 0.182 to 0.250 with all features, whereas the use of only lipid panels and creatinine resulted in AUROCs ranging from 0.786 to 0.871 and AUPRCs ranging from 0.158 to 0.243. For the 1-year prediction of composite outcome, the AUROCs ranged from 0.876 to 0.892, and the AUPRCs ranged from 0.283 to 0.322 with all features, whereas using lipid panels and creatinine yielded AUROCs of 0.703–0.748 and AUPRCs of 0.136–0.185 (Table 2). However, when predicting kidney outcomes using the LR model, the performance was comparable when all features were used and when only lipid panels and creatinine were used. Detailed performance metrics of the models, including accuracy, sensitivity, specificity, precision, and F1 score, are presented in Supplementary Table 6 (available online) for kidney outcomes and Supplementary Table 7 (available online) for composite outcomes.

Identifying important lipid panels

To identify features that significantly contributed to predicting kidney and composite outcomes, with a particular focus on lipid panels, feature ranking was performed. Among the five algorithms, the MLP demonstrated the most favorable blend of AUPRC, precision, and F1-score and was therefore selected for a detailed feature-importance analysis (Supplementary Tables 6 and 7, available online). When examined in a grouped manner, the lipid panel emerged as the most important feature in the MLP model for predicting 3-year kidney and composite outcomes, highlighting its

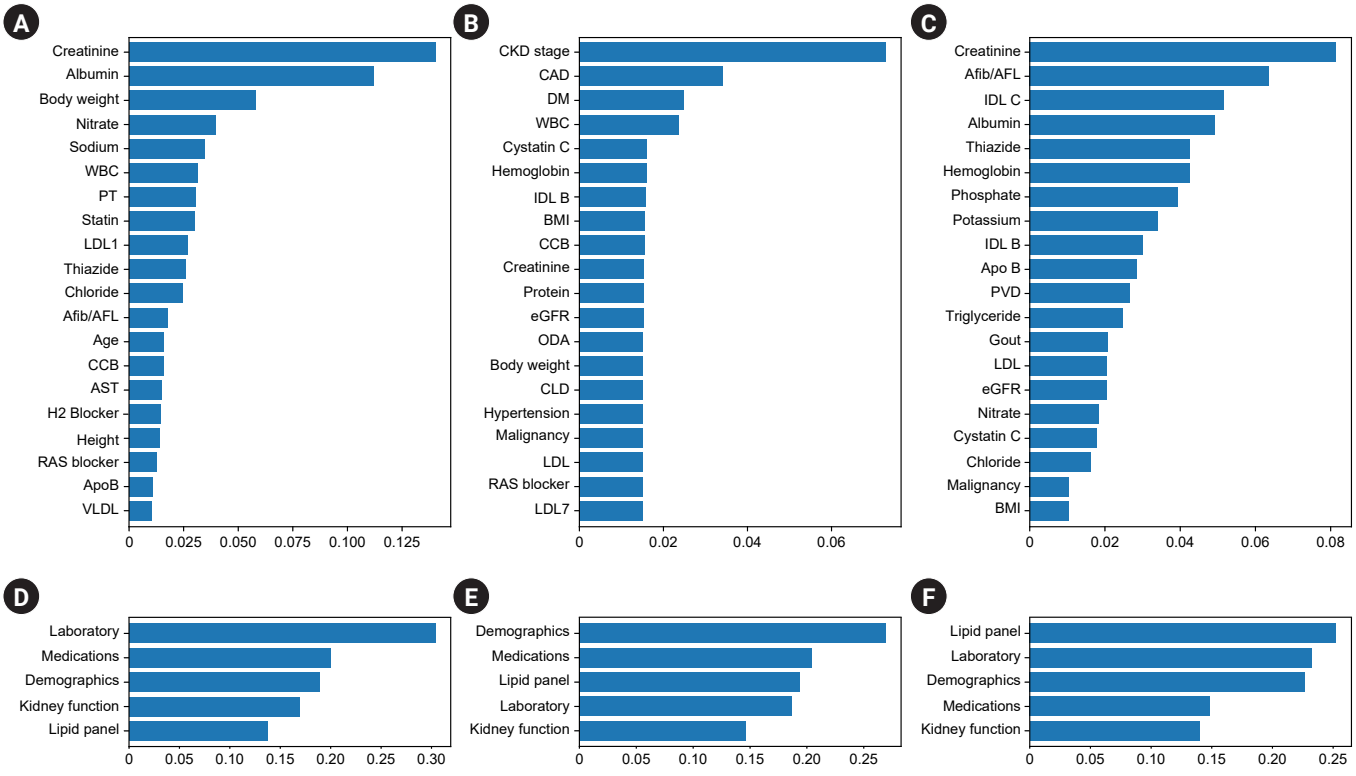

Figure 3. Feature importance for predicting kidney dysfunction in the multilayer perceptron model. (A–C) Importance of each feature for 1-year (A), 2-year (B), and 3-year (C) prediction. (D–F) Importance of grouped features for 1-year (D), 2-year (E), and 3-year (F) prediction. The x-axis indicates the mean absolute SHAP (SHapley Additive exPlanations) value, representing the average impact of each feature on the model output, and the y-axis shows the input features ranked by importance.
AFL, atrial flutter; Afib, atrial fibrillation; ApoB, apolipoprotein B; AST, aspartate aminotransferase; BMI, body mass index; CAD, coronary artery disease; CCB, calcium channel blocker; CKD, chronic kidney disease; CLD, chronic liver disease; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; IDL, intermediate-density lipoprotein; LDL, low-density lipoprotein; ODA, oral anti-diabetic agent; PT, prothrombin time; PVD, peripheral vascular disease; RAS, renin-angiotensin system; VLDL, very low-density lipoprotein; WBC, white blood cell.

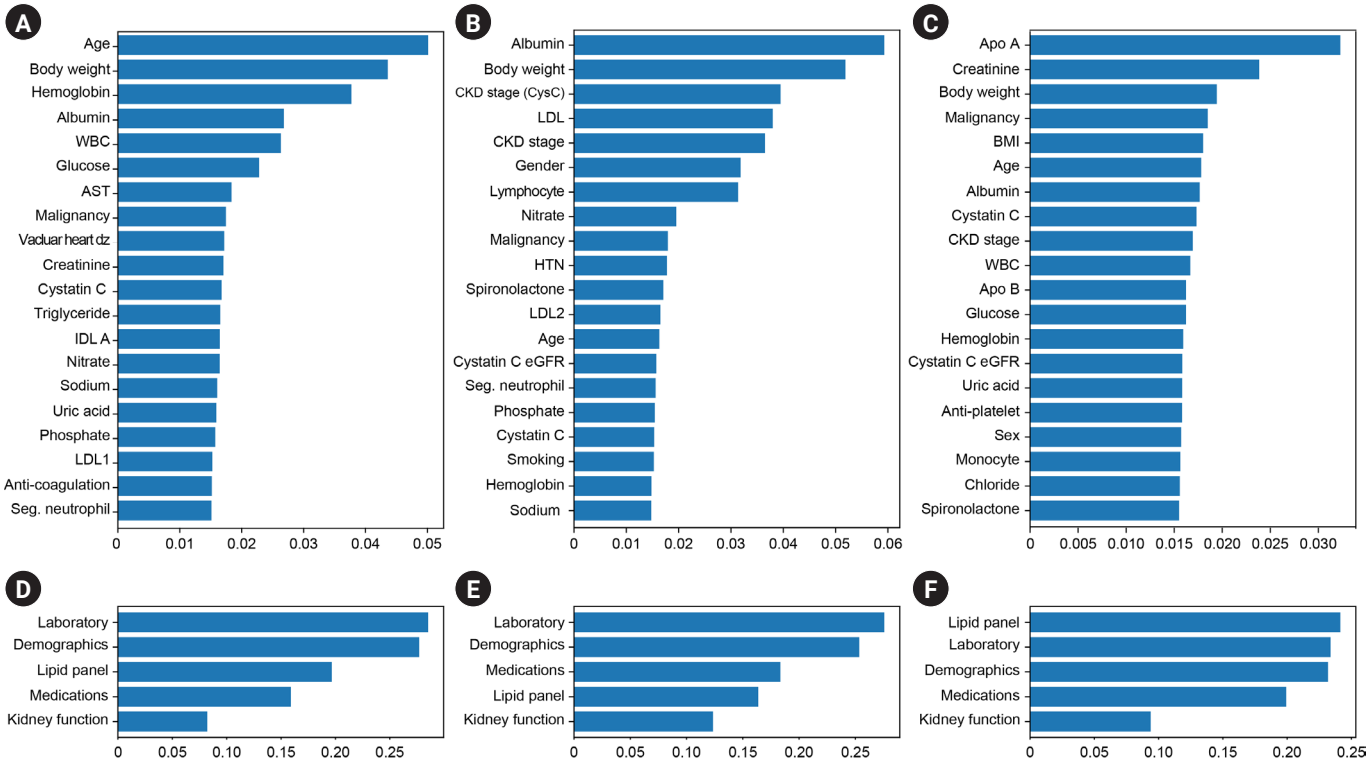

Figure 4. Feature importance for predicting composite outcomes in the multilayer perceptron model. (A–C) Importance of each feature for 1-year (A), 2-year (B), and 3-year (C) prediction. (D–F) Importance of grouped features for 1-year (D), 2-year (E), and 3-year (F) prediction. The x-axis indicates the mean absolute SHAP (SHapley Additive exPlanations) value, representing the average impact of each feature on the model output, and the y-axis shows the input features ranked by importance.
Apo A, apolipoprotein A; AST, aspartate aminotransferase; BMI, body mass index; CKD, chronic kidney disease; dz, disease; eGFR, estimated glomerular filtration rate; HTN, hypertension; IDL, intermediate-density lipoprotein; LDL, low-density lipoprotein; WBC, white blood cell.

significance (Figs. 3 and 4). In the LGBM model, it consistently ranked as the second most important feature across all prediction tasks (Supplementary Figs. 4 and 5, available online). Similarly, in the RF model, the lipid panel ranked second to third in importance, reaffirming its relevance (Supplementary Figs. 6 and 7, available online). Although the RNN model showed relatively lower rankings for composite outcome prediction, the third to fourth, the actual SHAP (SHapley Additive exPlanations) values were comparable to those of the top-ranked features (Supplementary Figs. 8 and 9, available online).

When the influence of individual features was separately estimated to identify the top 20 features, more than one lipid parameter appeared in all of the models. Notably, in the MLP model, where the lipid parameter group had the greatest influence, LDL and ApoB were frequently identified. For kidney outcomes, the top lipid parameters included LDL1 and ApoB at 1 year; intermediate-density lipoprotein (IDL)-B, LDL, and LDL7 at 2 years; and IDL cholesterol, ApoB, IDL-B, and LDL at 3 years. For composite outcomes, the top lipid parameters were TG, IDL-A, and LDL1 at 1 year; LDL and LDL2 at 2 years; and ApoA1 and ApoB at 3 years (Figs. 3 and 4). ApoA1, ApoB, and LDL cholesterol were consistently identified as important lipid parameters, as they frequently appeared in the feature rankings across all machine learning and deep learning algorithms (Figs. 3 and 4; Supplementary Figs. 4–9, available online).

Validation of feature importance

A multivariate Cox analysis was conducted to validate three lipid parameters identified by the algorithms (ApoA1, ApoB, and LDL cholesterol). When analyzed as continuous

variables in the fully adjusted model, high ApoB levels were associated with an elevated risk of kidney outcomes (HR, 1.01; 95% CI, 1.00–1.02; p = 0.004) (Table 3), whereas high ApoA1 levels were associated with a low risk of composite outcomes (HR, 1.00; 95% CI, 0.99–1.00; p = 0.02) (Table 4). However, LDL cholesterol did not show a significant linear relationship with either kidney or composite outcomes.

Discussion

Owing to the fluctuating correlation between lipid parameters and various underlying conditions that influence dyslipidemia in CKD patients, the relationship between lipid parameters and kidney dysfunction remains controversial. To address the challenge of intercorrelations, we employed various artificial intelligence models to explore the associations between lipid parameters and kidney outcomes. These models incorporated all relevant features, with a specific focus on lipid parameters and serum creatinine levels, achieving high predictive power for both kidney and composite outcomes, particularly with the MLP and LGBM models. The feature ranking results consistently placed the group of lipid parameters among the top or second most important features in predicting both kidney and composite outcomes. These results demonstrate that machine learning models can accurately predict the associations between kidney dysfunction-related outcomes and changes in lipid panels measured in CKD patients.

Previous studies have utilized various conventional analytical methods, such as analyzing the distribution [26] or ratio [27,28] of lipid parameters and apolipoproteins [29–32], to address the intercorrelations and explore the relationships between dyslipidemia and kidney outcomes. Despite these efforts, no specific lipid factor has been definitively identified as being strongly associated with the progression of kidney disease. Additionally, a Cochrane review reported that while statin usage in CKD patients consistently lowers ASCVD risk, its effect on kidney function remains uncertain [15]. The ongoing debate in studies using traditional analyses highlights the need for new and more advanced analytical approaches to better understand the complex relationship between lipid profiles and kidney dysfunction. To address this challenge, this study utilized machine learning models, which were able to effectively explore the association with kidney dysfunction.

Traditionally, machine learning algorithms have been used to predict short-term outcomes, such as the occurrence of in-hospital acute kidney injury [21] or mortality in intensive care units [22]. However, recent efforts have expanded the application of machine learning algorithms

Table 3. Multivariate Cox model for predicting kidney outcome according to the level of lipid parameters

Variable	Model 1		Model 2		Model 3	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Apolipoprotein A1	0.98 (0.98–0.99)	<0.001	0.99 (0.95–1.00)	0.03	1.00 (0.99–1.01)	0.55
Q1	1 (Reference)		1 (Reference)		1 (Reference)	
Q2	0.58 (0.40–0.85)	0.005	0.71 (0.48–1.04)	0.08	0.80 (0.54–1.19)	0.27
Q3	0.60 (0.41–0.87)	0.007	0.75 (0.51–1.10)	0.14	0.96 (0.65–1.42)	0.82
Q4	0.41 (0.27–0.63)	<0.001	0.57 (0.37–0.89)	0.01	0.72 (0.46–1.14)	0.16
Apolipoprotein B	1.01 (1.00–1.02)	0.001	1.01 (1.01–1.02)	<0.001	1.01 (1.00–1.02)	0.004
Q1	1 (Reference)		1 (Reference)		1 (Reference)	
Q2	1.15 (0.74–1.79)	0.53	1.30 (0.83–2.02)	0.25	1.30 (0.83–2.05)	0.25
Q3	1.16 (0.75–1.81)	0.50	1.21 (0.78–1.89)	0.39	1.24 (0.79–1.95)	0.35
Q4	1.74 (1.16–2.60)	0.008	1.91 (1.26–2.89)	0.002	1.68 (1.10–2.58)	0.02
LDL cholesterol	1.00 (1.00–1.01)	0.91	1.00 (1.00–1.01)	0.19	1.00 (1.00–1.01)	0.12
Q1	1 (Reference)		1 (Reference)		1 (Reference)	
Q2	0.84 (0.56–1.27)	0.42	0.91 (0.60–1.37)	0.64	1.03 (0.67–1.57)	0.91
Q3	0.87 (0.58–1.30)	0.50	0.98 (0.65–1.48)	0.94	1.18 (0.77–1.81)	0.44
Q4	0.95 (0.64–1.42)	0.82	1.17 (0.78–1.76)	0.44	1.27 (0.83–1.94)	0.27

Model 1: unadjusted. Model 2: adjusted for age, sex, body mass index, smoking, and comorbidities. Model 3: adjusted for variables in Model 2 plus laboratory findings and medications.
CI, confidence interval; HR, hazard ratio; LDL, low-density lipoprotein; Q, quartile.

Table 4. Multivariate Cox model for predicting composite outcome according to the level of lipid parameters

Variable	Model 1		Model 2		Model 3	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Apolipoprotein A1	0.98 (0.98–0.99)	<0.001	0.99 (0.99–0.99)	<0.001	1.00 (0.99–1.00)	0.02
Q1	1 (Reference)		1 (Reference)		1 (Reference)	
Q2	0.71 (0.48–1.04)	0.08	0.75 (0.66–0.86)	<0.001	0.89 (0.78–1.02)	0.10
Q3	0.75 (0.51–1.10)	0.14	0.67 (0.58–0.77)	<0.001	0.81 (0.70–0.93)	0.004
Q4	0.57 (0.37–0.89)	0.01	0.69 (0.60–0.80)	<0.001	0.87 (0.75–1.01)	0.07
Apolipoprotein B	1.00 (1.00–1.00)	0.02	1.00 (1.00–1.01)	0.06	1.00 (1.00–1.00)	0.52
Q1	1 (Reference)		1 (Reference)		1 (Reference)	
Q2	0.93 (0.81–1.06)	0.28	1.06 (0.93–1.22)	0.39	1.02 (0.88–1.17)	0.83
Q3	0.83 (0.72–0.96)	0.01	0.98 (0.85–1.13)	0.78	0.96 (0.83–1.21)	0.59
Q4	0.85 (0.74–0.98)	0.03	1.14 (0.98–1.32)	0.08	1.04 (0.90–1.61)	0.58
LDL cholesterol	1.00 (1.00–1.00)	<0.001	1.00 (1.00–1.00)	0.45	1.00 (1.00–1.00)	0.54
Q1	1 (Reference)		1 (Reference)		1 (Reference)	
Q2	0.89 (0.77–1.02)	0.09	1.00 (0.87–1.16)	0.96	1.02 (0.89–1.18)	0.76
Q3	0.84 (0.73–0.97)	0.02	1.04 (0.90–1.20)	0.59	1.06 (0.92–1.22)	0.44
Q4	0.76 (0.66–0.88)	<0.001	1.04 (0.90–1.20)	0.62	1.04 (0.90–1.21)	0.59

Model 1: unadjusted. Model 2: adjusted for age, sex, body mass index, smoking, and comorbidities. Model 3: adjusted for variables in Model 2 plus laboratory findings and medications.
CI, confidence interval; HR, hazard ratio; LDL, low-density lipoprotein; Q, quartile.

to long-term outcome prediction in conditions such as the development of diabetic kidney disease [23,24,33], and microvascular complications in diabetic patients [34]. A previous study that compared various machine learning models for predicting 3-year outcomes revealed that LGBM outperformed other models, demonstrating promising predictive power for diabetic kidney disease progression [33]. This study also identified several key factors contributing to the prediction, including old age, high homocysteine levels, poor glycemic control, hypoalbuminemia, low eGFR, and high bicarbonate levels [33]. Additionally, a deep learning model that simultaneously incorporated structural data (serum/urine laboratory findings and medications) and unstructured data (ICD-10 codes and medical text) successfully predicted the progression of diabetic kidney disease 6 months in advance, achieving an accuracy of 71% [24].

To our knowledge, while several studies have explored the relationship between lipid parameters and kidney outcomes, our study is the first to investigate this relationship using machine learning algorithms. Our comprehensive model, which incorporated all relevant features, demonstrated strong predictive power for both kidney and composite outcomes, with AUROC values of 0.92 and 0.87, respectively. Notably, even when limited to lipid parameters and serum creatinine, the model maintained significant predictive accuracy, achieving AUROC values of 0.87 for kidney outcomes and 0.72 for composite outcomes. Among the lipid parameters identified through feature ranking, ApoA1 and ApoB consistently showed linear associations with the study outcomes. However, although LDL cholesterol was highlighted in the feature ranking and is known for its variability in CKD patients, a linear relationship was not observed in conventional analysis. This suggests that the machine learning model may have uncovered associations that traditional analysis methods could not detect, highlighting the potential for applying advanced explainable artificial intelligence techniques in future research.

However, our study has certain limitations. Although the overall cohort size was large, the proportion of outcome events was relatively low (kidney outcomes: 0.6%–1.2% and composite outcomes: 3.0%–7.4% across 1 to 3 years). Despite our efforts, this introduces a potential risk of overfitting, highlighting the need for future studies with larger numbers of outcome events to ensure the validity and generalizability of our findings. Second, external validation could not be performed due to the lack of an independent cohort. We attempted to mitigate this limitation by performing validation through conventional survival analysis. Considering the limitations of artificial intelligence in

long-term outcome prediction, particularly its inability to account for time lags, this approach was appropriate. Additionally, we were unable to account for the time-varying characteristics of lipid parameters, as it was challenging to assess follow-up results for various lipid parameters simultaneously due to high rates of missing data.

In conclusion, this study demonstrates the utility of lipid parameters for predicting kidney and composite outcomes via diverse machine learning models. Notably, our models identified ApoB, ApoA1, and LDL cholesterol as potential prognostic factors, and their linear relationships were further supported by survival analysis. However, external validation of our machine learning models is essential to establish their robustness and generalizability.

Conflicts of interest

All authors have no conflicts of interest to declare.

Data sharing statement

The data presented in this study are available from the corresponding author upon reasonable request.

Authors' contributions

Conceptualization, Supervision: SY, SSH
Data curation: SK, SP, YCK, DKK, KHO, KWJ, YSK
Formal analysis: DY, CP
Investigation: SP, YCK, DKK, KHO
Methodology, Visualization: DY
Project administration: SY, SSH
Validation: SK
Writing–original draft: SK, DY
Writing–review & editing: All authors
All authors read and approved the final manuscript.

ORCID

Soie Kwon, https://orcid.org/0000-0002-0878-5469
Donghwan Yoon, https://orcid.org/0000-0001-6566-5183
Changhwa Park, https://orcid.org/0000-0002-6685-8072
Sehoon Park, https://orcid.org/0000-0002-4221-2453
Yong Chul Kim, https://orcid.org/0000-0003-3215-8681
Dong Ki Kim, https://orcid.org/0000-0002-5195-7852
Kook-Hwan Oh, https://orcid.org/0000-0001-9525-2179
Kwon Wook Joo, https://orcid.org/0000-0001-9941-7858
Yon Su Kim, https://orcid.org/0000-0003-3091-2388
Sungroh Yoon, https://orcid.org/0000-0002-2367-197X
Seung Seok Han, https://orcid.org/0000-0003-0137-5261

References

1. Dayimu A, Wang C, Li J, et al. Trajectories of lipids profile and incident cardiovascular disease risk: a longitudinal cohort study. *J Am Heart Assoc* 2019;8:e013479.
2. Zhao X, Wang D, Qin L. Lipid profile and prognosis in patients with coronary heart disease: a meta-analysis of prospective cohort studies. *BMC Cardiovasc Disord* 2021;21:69.
3. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97:1837–1847.
4. Boullart AC, de Graaf J, Stalenhoef AF. Serum triglycerides and risk of cardiovascular disease. *Biochim Biophys Acta* 2012;1821:867–875.
5. Michos ED, McEvoy JW, Blumenthal RS. Lipid management for the prevention of atherosclerotic cardiovascular disease. *N Engl J Med* 2019;381:1557–1567.
6. Ip S, Lichtenstein AH, Chung M, Lau J, Balk EM. Systematic review: association of low-density lipoprotein subfractions with cardiovascular outcomes. *Ann Intern Med* 2009;150:474–484.
7. Ikezaki H, Lim E, Cupples LA, Liu CT, Asztalos BF, Schaefer EJ. Small dense low-density lipoprotein cholesterol is the most atherogenic lipoprotein parameter in the prospective framingham offspring study. *J Am Heart Assoc* 2021;10:e019140.
8. Krauss RM. Dietary and genetic effects on low-density lipoprotein heterogeneity. *Annu Rev Nutr* 2001;21:283–295.
9. Theofilis P, Vordoni A, Koukoulaki M, Vlachopanos G, Kalaitzidis RG. Dyslipidemia in chronic kidney disease: contemporary concepts and future therapeutic perspectives. *Am J Nephrol* 2021;52:693–701.
10. Wanner C, Tonelli M. KDIGO clinical practice guideline for lipid management in CKD: summary of recommendation statements and clinical approach to the patient. *Kidney Int* 2014;85:1303–1309.
11. Barbagallo CM, Cefalù AB, Giammanco A, et al. Lipoprotein abnormalities in chronic kidney disease and renal transplantation. *Life (Basel)* 2021;11:315.
12. Kosugi T, Eriguchi M, Yoshida H, et al. Association between chronic kidney disease and new-onset dyslipidemia: the Ja-

pan Specific Health Checkups (J-SHC) study. *Atherosclerosis* 2021;332:24–32.
13. Ferro CJ, Mark PB, Kanbay M, et al. Lipid management in patients with chronic kidney disease. *Nat Rev Nephrol* 2018;14:727–749.
14. Rahman M, Yang W, Akkina S, et al. Relation of serum lipids and lipoproteins with progression of CKD: the CRIC study. *Clin J Am Soc Nephrol* 2014;9:1190–1198.
15. Palmer SC, Navaneethan SD, Craig JC, et al. HMG CoA reductase inhibitors (statins) for people with chronic kidney disease not requiring dialysis. *Cochrane Database Syst Rev* 2014;(5):CD007784.
16. Mikolasevic I, Žutelija M, Mavrinac V, Orlic L. Dyslipidemia in patients with chronic kidney disease: etiology and management. *Int J Nephrol Renovasc Dis* 2017;10:35–45.
17. Bulbul MC, Dagel T, Afsar B, et al. Disorders of lipid metabolism in chronic kidney disease. *Blood Purif* 2018;46:144–152.
18. Sidey-Gibbons JA, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol* 2019;19:64.
19. Remeseiro B, Bolon-Canedo V. A review of feature selection methods in medical applications. *Comput Biol Med* 2019;112:103375.
20. Lee Y, Ryu J, Kang MW, et al. Machine learning-based prediction of acute kidney injury after nephrectomy in patients with renal cell carcinoma. *Sci Rep* 2021;11:15704.
21. Vagliano I, Chesnaye NC, Leopold JH, Jager KJ, Abu-Hanna A, Schut MC. Machine learning models for predicting acute kidney injury: a systematic review and critical appraisal. *Clin Kidney J* 2022;15:2266–2280.
22. Hyland SL, Faltys M, Hüser M, et al. Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat Med* 2020;26:364–373.
23. Allen A, Iqbal Z, Green-Saxena A, et al. Prediction of diabetic kidney disease with machine learning algorithms, upon the initial diagnosis of type 2 diabetes mellitus. *BMJ Open Diabetes Res Care* 2022;10:e002560.
24. Makino M, Yoshimoto R, Ono M, et al. Artificial intelligence predicts the progression of diabetic kidney disease using big data machine learning. *Sci Rep* 2019;9:11862.
25. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604–612.
26. Chen SC, Hung CC, Kuo MC, et al. Association of dyslipidemia with renal outcomes in chronic kidney disease. *PLoS One* 2013;8:e55643.
27. Wang X, Chen H, Shao X, et al. Association of lipid parameters with the risk of chronic kidney disease: a longitudinal study based on populations in Southern China. *Diabetes Metab Syndr Obes* 2020;13:663–670.
28. Liao S, Lin D, Feng Q, et al. Lipid parameters and the development of chronic kidney disease: a prospective cohort study in middle-aged and elderly Chinese individuals. *Nutrients* 2022;15:112.
29. Hsu CC, Kao WH, Coresh J, et al. Apolipoprotein E and progression of chronic kidney disease. *JAMA* 2005;293:2892–2899.
30. Kronenberg F. Apolipoprotein L1 and apolipoprotein A-IV and their association with kidney function. *Curr Opin Lipidol* 2017;28:39–45.
31. Boes E, Fliser D, Ritz E, et al. Apolipoprotein A-IV predicts progression of chronic kidney disease: the mild to moderate kidney disease study. *J Am Soc Nephrol* 2006;17:528–536.
32. Kwon S, Kim DK, Oh KH, et al. Apolipoprotein B is a risk factor for end-stage renal disease. *Clin Kidney J* 2021;14:617–623.
33. Dong Z, Wang Q, Ke Y, et al. Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records. *J Transl Med* 2022;20:143.
34. Al-Sari N, Kutuzova S, Suvitaival T, et al. Precision diagnostic approach to predict 5-year risk for microvascular complications in type 1 diabetes. *EBioMedicine* 2022;80:104032.