ISSN 2307-8960 (online)

World Journal of *Clinical Cases*

World J Clin Cases 2023 October 16; 11(29): 6974-7260

Published by Baishideng Publishing Group Inc

World Journal of Clinical Cases

Contents

Thrice Monthly Volume 11 Number 29 October 16, 2023

MINIREVIEWS

ORIGINAL ARTICLE

Retrospective Cohort Study

Retrospective Study

Observational Study

World Journal of Clinical Cases

Thrice Monthly Volume 11 Number 29 October 16, 2023

Contents

World Journal of Clinical Cases

Thrice Monthly Volume 11 Number 29 October 16, 2023

Contents

LETTER TO THE EDITOR

Contents

Thrice Monthly Volume 11 Number 29 October 16, 2023

ABOUT COVER

AIMS AND SCOPE

The primary aim of *World Journal of Clinical Cases* (*WJCC*, *World J Clin Cases*) is to provide scholars and readers from various fields of clinical medicine with a platform to publish high-quality clinical research articles and communicate their research findings online.

WJCC mainly publishes articles reporting research results and findings obtained in the field of clinical medicine and covering a wide range of topics, including case control studies, retrospective cohort studies, retrospective studies, clinical trials studies, observational studies, prospective studies, randomized controlled trials, randomized clinical trials, systematic reviews, meta-analysis, and case reports.

INDEXING/ABSTRACTING

The *WJCC* is now abstracted and indexed in Science Citation Index Expanded (SCIE, also known as SciSearch®), Journal Citation Reports/Science Edition, Current Contents®/Clinical Medicine, PubMed, PubMed Central, Reference Citation Analysis, China National Knowledge Infrastructure, China Science and Technology Journal Database, and Superstar Journals Database. The 2023 Edition of Journal Citation Reports® cites the 2022 impact factor (IF) for *WJCC* as 1.1; IF without journal self cites: 1.1; 5-year IF: 1.3; Journal Citation Indicator: 0.26; Ranking: 133 among 167 journals in medicine, general and internal; and Quartile category: Q4.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Hua-Ge Yu*; **Production Department Director:** *Xiang Li*; **Editorial Office Director:** *Jin-Lei Wang*.

NAME OF JOURNAL
World Journal of Clinical Cases

ISSN
ISSN 2307-8960 (online)

LAUNCH DATE
April 16, 2013

FREQUENCY
Thrice Monthly

PUBLICATION DATE
October 16, 2023

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

World Journal of Clinical Cases

Submit a Manuscript: https://www.f6publishing.com

World J Clin Cases 2023 October 16; 11(29): 7004-7016

DOI: 10.12998/wjcc.v11.i29.7004

ISSN 2307-8960 (online)

ORIGINAL ARTICLE

Retrospective Study

Roles of biochemistry data, lifestyle, and inflammation in identifying abnormal renal function in old Chinese

Chao-Hung Chen, Chun-Kai Wang, Chen-Yu Wang, Chun-Feng Chang, Ta-Wei Chu

Specialty type: Urology and nephrology

Provenance and peer review: Unsolicited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): D, D
Grade E (Poor): 0

P-Reviewer: Lin HH, China; Patel J, United States

Received: July 5, 2023
Peer-review started: July 5, 2023
First decision: July 18, 2023
Revised: August 1, 2023
Accepted: September 11, 2023
Article in press: September 11, 2023
Published online: October 16, 2023

Chao-Hung Chen, Chun-Feng Chang, Division of Urology, Department of Surgery, Kaohsiung Armed Forces General Hospital, Kaohsiung 802, Taiwan

Chao-Hung Chen, Division of Urology, Department of Surgery, Chang Gung Memorial Hospital, Keelung 204, Taiwan

Chun-Kai Wang, Department of Obstetrics and Gynecology, Zuoying Branch of Kaohsiung Armed Forces General Hospital, Kaohsiung 813, Taiwan

Chen-Yu Wang, Ta-Wei Chu, Department of Obstetrics and Gynecology, Tri-Service General Hospital, National Defense Medical Center, Taipei 114, Taiwan

Chun-Feng Chang, Division of Urology, Department of Surgery, Tri-Service General Hospital, National Defense Medical Center, Taipei 114, Taiwan

Ta-Wei Chu, Chief Executive Officer's Office, MJ Health Research Foundation, Taipei 114, Taiwan

Corresponding author: Ta-Wei Chu, MD, PhD, CEO, Chief Physician, Doctor, Department of Obstetrics and Gynecology, Tri-Service General Hospital, National Defense Medical Center, No. 325 Sec. 2, Chenggong Road, Neihu District, Taipei 114, Taiwan. david_chu@mjlife.com

Abstract

BACKGROUND

The incidence of chronic kidney disease (CKD) has dramatically increased in recent years, with significant impacts on patient mortality rates. Previous studies have identified multiple risk factors for CKD, but they mostly relied on the use of traditional statistical methods such as logistic regression and only focused on a few risk factors.

AIM

To determine factors that can be used to identify subjects with a low estimated glomerular filtration rate (L-eGFR < 60 mL/min per 1.73 m^2) in a cohort of 1236 Chinese people aged over 65.

METHODS

Twenty risk factors were divided into three models. Model 1 consisted of demographic and biochemistry data. Model 2 added lifestyle data to Model 1, and Model 3 added inflammatory markers to Model 2. Five machine learning methods

were used: Multivariate adaptive regression splines, eXtreme Gradient Boosting, stochastic gradient boosting, Light Gradient Boosting Machine, and Categorical Features + Gradient Boosting. Evaluation criteria included accuracy, sensitivity, specificity, area under the receiver operating characteristic curve (AUC), F-1 score, and balanced accuracy.

RESULTS

A trend of increasing AUC of each was observed from Model 1 to Model 3 and reached statistical significance. Model 3 selected uric acid as the most important risk factor, followed by age, hemoglobin (Hb), body mass index (BMI), sport hours, and systolic blood pressure (SBP).

CONCLUSION

Among all the risk factors including demographic, biochemistry, and lifestyle risk factors, along with inflammation markers, UA is the most important risk factor to identify L-eGFR, followed by age, Hb, BMI, sport hours, and SBP in a cohort of elderly Chinese people.

Key Words: Biochemistry data; Lifestyle; Machine learning; Renal function

Core Tip: This is a retrospective study that used five machine learning methods to evaluate the impact of lifestyle and chronic inflammation in identifying subjects with abnormal estimated glomerular rates among elderly Chinese subjects. Our results showed that uric acid is the most important risk factor (inflammatory marker), followed by age, hemoglobin, body mass index, sport hours, and systolic blood pressure.

Citation: Chen CH, Wang CK, Wang CY, Chang CF, Chu TW. Roles of biochemistry data, lifestyle, and inflammation in identifying abnormal renal function in old Chinese. *World J Clin Cases* 2023; 11(29): 7004-7016

URL: https://www.wjgnet.com/2307-8960/full/v11/i29/7004.htm

DOI: https://dx.doi.org/10.12998/wjcc.v11.i29.7004

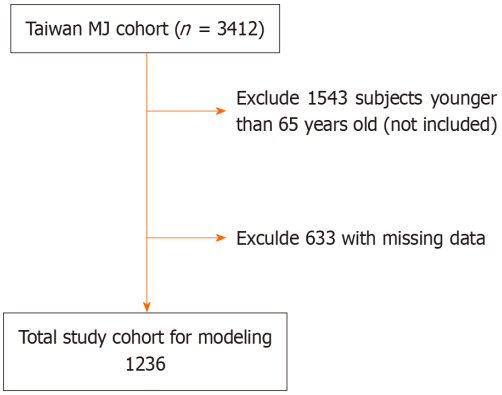
INTRODUCTION

The number of global people suffering chronic kidney disease (CKD) and acute kidney injury is approaching 850 million. CKD is expected to emerge as the 5th cause of death by the year of 2040, and 2nd by 2100 as the global population continues to age. CKD progresses through five stages based on the estimated glomerular filtration rate (eGFR). A decrease in eGFR value to 15 mL/min per 1.73 m^2 is defined as end stage renal disease according to Kidney Disease: Improving Global Outcome. In the United States, there are approximately 80000 patients with end stage renal disease, 71% of whom are presently on dialysis[1]. A similar trend is found in Taiwan. Data from Taiwan's National Health Insurance agency indicates that the prevalence of CKD increased from 1.3 million to 2.2 million from 2005 to 2014[2], while Tsai *et al*[3] found a 15.45% prevalence in a study cohort of 106094 subjects, of which 9.06% were in CKD stages[4]. The determinants were found to be diabetes, hypertension, and metabolic syndrome[4].

Subjects with CKD have a significantly higher chance to have cardiovascular diseases and cerebrovascular disease (stroke, transient ischemic attack, *etc*.), along with associated cognitive dysfunction. Even in early stage CKD, the appearance of albuminuria could be regarded as a representative systemic vascular injury[5].

Many studies have examined the risk factors for CKD. Hannan *et al*[6] found that lifestyle factors such as smoking cessation and exercise significantly retard the onset of CKD. They also reported that increased waking during was associated with a higher risk for CKD. Imig *et al*[7] found that inflammation and immune system activation are common underlying mechanisms for CKD. However, it should be noted that these previous studies have not been subject to meta-analysis and used traditional statistical analysis methods.

In recent years, machine learning (Mach-L) techniques have been widely applied in the field of medicine. Mach-L uses the current computing power to achieve our goal automatically through a computer algorithm[8]. Mach-L can capture nonlinear relationships in the data and complex interactions among multiple predictors, allowing it to potentially outperform conventional multiple logistic regression for diseases[9]. However, to date, no study has applied Mach-L to identifying the risk factors for CKD. The present study, we 1236 healthy elderly Chinese subjects. Five different Mach-L methods were applied to predict high or low eGFR levels (H-eGFR: ≥ 60, L-eGFR < 60 mL/min/1.73 m^2, dependent variable). The independent variables were divided into three models: Model 1: Demographic and biochemistry data; Model 2: Model 1 + lifestyle factors (income, education level, smoking, drinking, sleeping hour, and sport hours); Model 3: Model 2 + inflammatory markers (IM). This study sought to determine whether adding lifestyle and/or IM to Model 1 would increase the prediction accuracy for L-eGFR in elderly Chinese by applying state-of-the-art Mach-L methods.



DOI: 10.12998/wjcc.v11.i29.7004

Figure 1 Flowchart of sample selection from the MJ chronic kidney disease study cohort.

MATERIALS AND METHODS

Patient selection

Data for this study were sourced from the Taiwan MJ cohort, an ongoing prospective cohort of people undergoing health examinations conducted by the MJ Health Screening Centers in Taiwan[10]. These examinations cover more than 100 important biological indicators, including anthropometric measurements, blood tests, imaging tests, *etc.* Each participant completed a self-administered questionnaire to collect personal information and family medical history, current health status, lifestyle, physical exercise, sleep habits, and dietary habits[11]. The MJ Health Database only includes participants who provided informed consent. All or part of the data used in this research were authorized by and received from MJ Health Research Foundation (Authorization Code: MJHRF2020022A). Any interpretations or conclusions described in this paper do not represent the views of MJ Health Research[12]. The study protocol was approved by the Institutional Review Board of the Tri-Service General Hospital, National Defense Medical Center (IRB No.: KAFGHIRB 109-46). A total of 3412 healthy participants were enrolled. After excluding subjects for various causes, a total of 1236 subjects remained for analysis, as shown in Figure 1.

On the day of the study, senior nursing staff recorded the subject's medical history, including information on any current medications, and a physical examination was performed. The waist circumference was measured horizontally at the level of the natural waist. The body mass index (BMI) was calculated as the participant's body weight (kg) divided by the square of the participant's height (m). The systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured using standard mercury sphygmomanometers on the right arm of each subject while seated.

Following previously published protocols, the procedures for collecting demographic and biochemical data are as follows[13]. After fasting for 10 h, blood samples were collected for biochemical analyses. Plasma was separated from the blood within 1 h of collection and stored at 30 °C until the analysis of fasting plasma glucose (FPG) and lipid profiles. FPG was measured using the glucose oxidase method (YSI 203 glucose analyzer; Yellow Springs Instruments, Yellow Springs, OH, United States). Total cholesterol and triglyceride (TG) levels were measured using the dry multilayer analytical slide method with a Fuji Dri-Chem 3000 analyzer (Fuji Photo Film, Tokyo, Japan). Serum high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol concentrations were analyzed using an enzymatic cholesterol assay, following dextran sulfate precipitation. A Beckman Coulter AU 5800 biochemical analyzer was used to determine the urine albumin/creatinine ratio by turbidimetry.

Table 1 defines the 19 baseline clinical variables, categorized into three models (Table 2). Model 1 included sex, age, BMI, blood pressure, FPG, aspartate aminotransferase (AST), alanine aminotransferase (ALT), uric acid (UA), HDL-C, TG, and eGFR; Mode 2 added drinking, daily sleeping and sport hours; Model 3 added white blood cell (WBC) count, hemoglobin (Hb), alkaline phosphatase (ALP), γ-glutamyl transferase (γ-GT), and high sensitivity c-reactive protein (hsCRP). All these variables were regarded as independent variables. At the same time, the dependent variable was categorical and subjects with H-eGFR were defined as 0 while those with L-eGFR were defined as 1 (L-eGFR < 60 mL/min/1.73 m^2).

Traditional statistics

Data are represented as the mean ± SD. Student's *t* test was used to evaluate the differences of continuous data between H-eGFR and L-eGFR subjects. All statistical tests were two-sided, and $P < 0.05$ was considered statistically significant. Statistical analyses were performed using SPSS 10.0 for Windows (SPSS, Chicago, IL, United States).

Proposed machine learning scheme

Models to predict H- or L-eGFR and rank risk factors were constructed using five different Mach-L methods: Multivariate adaptive regression splines (MARS), eXtreme Gradient Boosting (XGBoost), stochastic gradient boosting (SGB), Light Gradient Boosting Machine (LightGBM), and Categorical Features + Gradient Boosting (CATboost) to construct models for predicting whether to have H- or L-eGFR and to identify the importance of the aforementioned risk factors. These

Table 1 Demographic, biochemistry, and lifestyle information of the participants

	Low eGFR	High eGFR
Number	180	1056
Age (yr)	72.1 ± 5.9	69.5 ± 4.6[c]
Sleep time (h)	5.89 ± 1.10	6.1 ± 1.15[a]
Drinking duration	4.76 ± 4.37	5.25 ± 5.74
Sport hours	205.6 ± 36.2	204.4 ± 36.3
Body mass index (kg/m^2)	23.9 ± 3.4	23.7 ± 3.2
White blood cell count (10^3/μL)	5.94 ± 1.75	5.58 ± 1.40[b]
Hemoglobin (g/dL)	13.8 ± 1.5	14.0 ± 1.3
Fasting plasma glucose (mg/dL)	108.2 ± 19.3	108.2 ± 21.5
Alkaline phosphatase (IU/L)	67.8 ± 19.2	66.6 ± 22.2
Serum glutamic oxaloacetic transaminase (IU/L)	27.3 ± 11.3	26.0 ± 12.8
Serum glutamic pyruvic transaminase (IU/L)	25.4 ± 14.7	25.8 ± 19.2
γ-glutamyltransferase (IU/L)	28.5 ± 28.7	27.1 ± 35.5
Systolic blood pressure (mmHg)	131.2 ± 18.9	127.4 ± 18.1[b]
Diastolic blood pressure (mmHg)	75.6 ± 11.2	74.6 ± 10.6
Triglyceride (mg/dL)	121.4 ± 71.4	114.5 ± 67.3
High density lipoprotein cholesterol (mg/dL)	57.4 ± 14.3	59.6 ± 16.0
Uric acid (mg/dL)	6.53 ± 1.48	5.56 ± 1.3
High sensitivity C-reactive protein (mg/L)	2.52 ± 5.08	2.11 ± 4.35[c]
eGFR	74.8 ± 10.1	53.02 ± 6.8

[a]$P < 0.05$.
[b]$P < 0.01$.
[c]$P < 0.005$.
eGFR: Estimated glomerular filtration rate.

Mach-L methods have been used in various healthcare applications and do not have prior assumptions regarding data distribution[14-23].

MARS is a nonparametric and nonlinear statistical method in which several linear segments with different gradients are used to automatically examine the nonlinearity and dependency between multidimensional input and output variables, and then generate the final optimum nonlinear prediction model[24].

XGBoost is a gradient boosting technology based on an SGB optimized extension[25]. It trains and assembles many weak models sequentially using the gradient boosting method of outputs, which achieves a better prediction performance. In XGBoost, the Taylor binomial expansion is used to approximate the objective function and arbitrary differentiable loss functions to accelerate model construction and convergence process[26]. XGBoost then applies a regularized boosting technique to penalize model complexity and correct overfitting, thus increasing model accuracy[25].

SGB is a tree-based gradient boosting learning algorithm that combines both bagging and boosting techniques to minimize the loss function to solve the overfitting problem of traditional decision trees[23]. In SGB, many stochastic weak learners of trees are sequentially generated through multiple iterations, in which each tree concentrates on correcting or explaining errors of the tree generated in the previous iteration. That is, the residual of the previous iteration tree is used as the input for the newly generated tree. This iterative process is repeated until the convergence condition or a stopping criterion is reached for the maximum number of iterations. Finally, the cumulative results of many trees are used to determine the final robust model.

LightGBM is a decision tree-based distributed gradient boosting framework that uses advanced histograms. In each iteration, it learns the approximate value of the decision tree residuals based on one-side sampling and negative gradient fitting[27].

CatBoost is a gradient-boosting decision tree technique in which sequential boosting methods are combined with gradient boosting and multiple categorical features[28]. In CatBoost, the tree combinations and categorical features generated through gradient boosting are aggregated into a sequence to generate the final model.

Figure 2 presents the proposed prediction and important variable identification scheme that combines the five Mach-L methods. First, patient data were collected to prepare the dataset. The dataset was then randomly divided into an 80%

Table 2 Variables contained in three models

	Model 1	Model 2	Model 3
Age	√	√	√
Body mass index	√	√	√
Systolic blood pressure	√	√	√
Diastolic blood pressure	√	√	√
Fasting plasma glucose	√	√	√
Serum glutamic oxaloacetic transaminase	√	√	√
Serum glutamic pyruvic transaminase	√	√	√
Uric acid	√	√	√
High density lipoprotein cholesterol	√	√	√
High sensitivity C-reactive protein	√	√	√
Triglyceride	√	√	√
Estimated glomerular filtration rate	√	√	√
Sleep time		√	√
Drinking duration		√	√
Sport hours		√	√
White blood cell count			√
Hemoglobin			√
Alkaline phosphatase			√
γ-glutamyltransferase			√
High sensitivity C-reactive protein			√

training dataset for model building and a 20% testing dataset for model testing. In the training process, the hyperparameters of each Mach-L method must be tuned to construct an effective model. In this study, a 10-fold cross-validation technique was used for hyperparameter tuning.

The training dataset was further randomly divided into a training dataset to rebuild the model with a different set of hyperparameters and a validation dataset for model validation. All possible hyperparameter combinations were investigated using a grid search. The best performing model in terms of accuracy, sensitivity, specificity, area under the receiver operating characteristic (AUC) curve, F-1 score, and balanced accuracy (Table 3) for the validation dataset was taken as the more accurate one. In the present study, AUC obtained from each Mach-L method was averaged and used as the comparator for the accuracy of the three models. We also ranked the corresponding variable importance. Using different Mach-L methods produces different risk rankings because of the different modeling characteristics. Therefore, we integrated the risk importance ranking to enhance the stability and integrity. After averaging, rank 1 is the most critical factor for L-eGFR.

All methods were performed using R software version 4.0.5 and R-Studio version 1.1.453 with the required packages installed (http://www.R-project.org; https://www.rstudio.com/products/rstudio/).

RESULTS

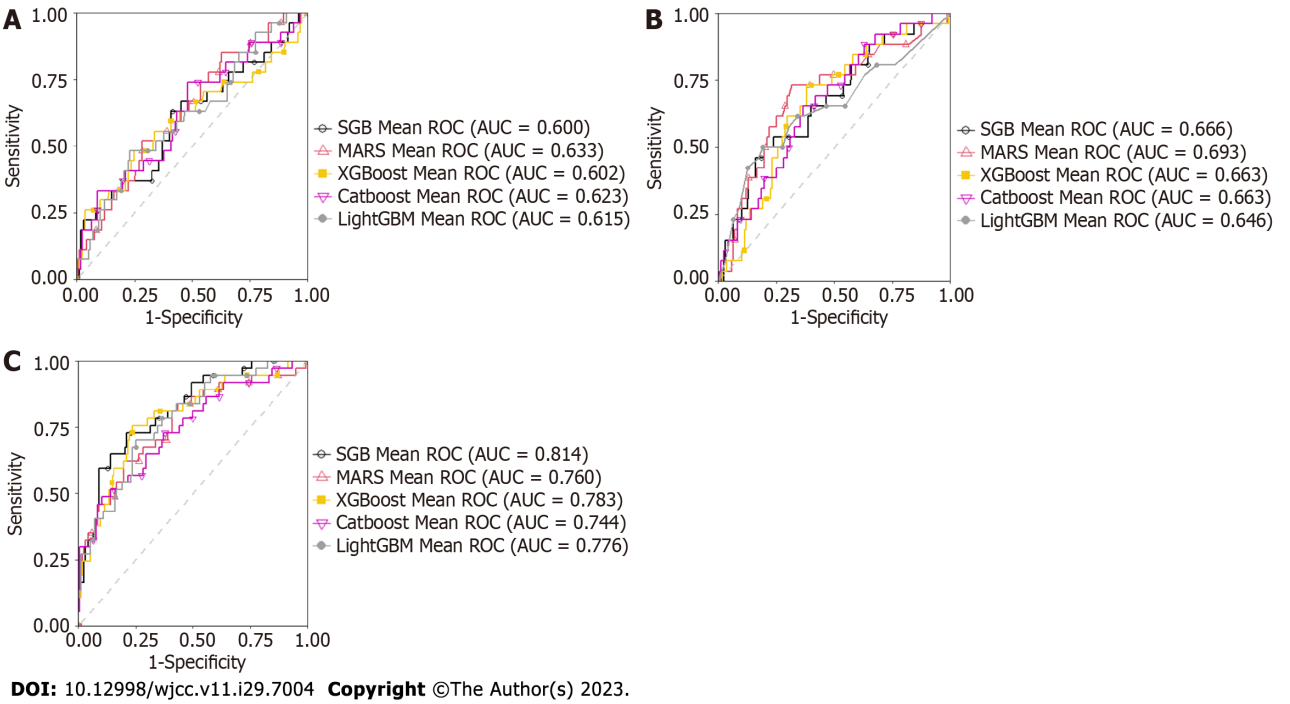
Table 1 summarizes the demographic data of the 1236 participants (mean ± SD). The mean age was significantly higher in subjects with low eGFR (72.1 ± 5.9 *vs* 69.5 ± 4.6 years old). Alcohol consumption was expressed as the multiple of the drinking frequency, alcohol percentage, and drinking duration. Exercise habits were expressed as the multiple of the intensity of the exercise, frequency, and the whole duration. Lifestyle results were consistent across both groups. Interestingly, the high eGFR group was found to have significantly higher sleep hours (6.1 ± 1.15 *vs* 5.89 ± 1.10 h). SBP was significantly higher in the low eGFR group (131.2 ± 18.9 *vs* 127.4 ± 18.1 mmHg), but not DBP. For the laboratory data, only WBC count and hsCRP were higher in the low eGFR group (5.94 ± 1.75 *vs* 5.58 ± 1.40 × 103/μL for WBC count and 2.52 ± 5.08 *vs* 2.11 ± 4.35 mg/L for hsCRP).

Table 3 summarizes the results for accuracy, sensitivity, specificity, AUC, F-1 score and BA derived from each model. Each value was found to increase from Model 1 to Model 3. Since the AUC represents the most important accuracy indicator for a given model, it is listed as the most important one in Table 4, which shows the average AUC values. The mean increased from 0.6144 for Model 1 to 0.776 for Model 3, indicating that, as risk factors were added, the mean AUC

Table 3 Results of five different machine learning methods in three different models

Model	Methods	Accuracy	Sensitivity	Specificity	AUC	F1-score	BA
Model 1	MARS	0.689	0.717	0.519	0.633	0.799	0.618
	XGboost	0.715	0.753	0.482	0.602	0.820	0.617
	SGB	0.596	0.590	0.630	0.599	0.715	0.610
	LightGBM	0.731	0.771	0.482	0.615	0.831	0.626
	Catboost	0.549	0.518	0.741	0.623	0.664	0.629
Model 2	MARS	0.689	0.683	0.731	0.693	0.792	0.707
	XGboost	0.632	0.617	0.731	0.663	0.744	0.674
	SGB	0.762	0.802	0.500	0.666	0.854	0.651
	LightGBM	0.767	0.808	0.500	0.646	0.857	0.654
	Catboost	0.637	0.635	0.654	0.663	0.752	0.644
Model 3	MARS	0.767	0.801	0.622	0.760	0.848	0.711
	XGboost	0.762	0.763	0.757	0.786	0.838	0.760
	SGB	0.777	0.789	0.730	0.814	0.851	0.759
	LightGBM	0.741	0.750	0.703	0.776	0.824	0.726
	Catboost	0.819	0.897	0.487	0.744	0.889	0.692

AUC: Area under receiver operating characteristic curve; BA: Balanced accuracy; MARS: Multivariate adaptive regression splines; XGBoost: eXtreme Gradient Boosting; SGB: Stochastic gradient boosting; LightGBM: Light Gradient Boosting Machine; CATboost: Categorical Features + Gradient Boosting. Model 1: Demographic and biochemistry data; Model 2: Model 1 + lifestyle factors; Model 3: Model 2 + inflammation factors.



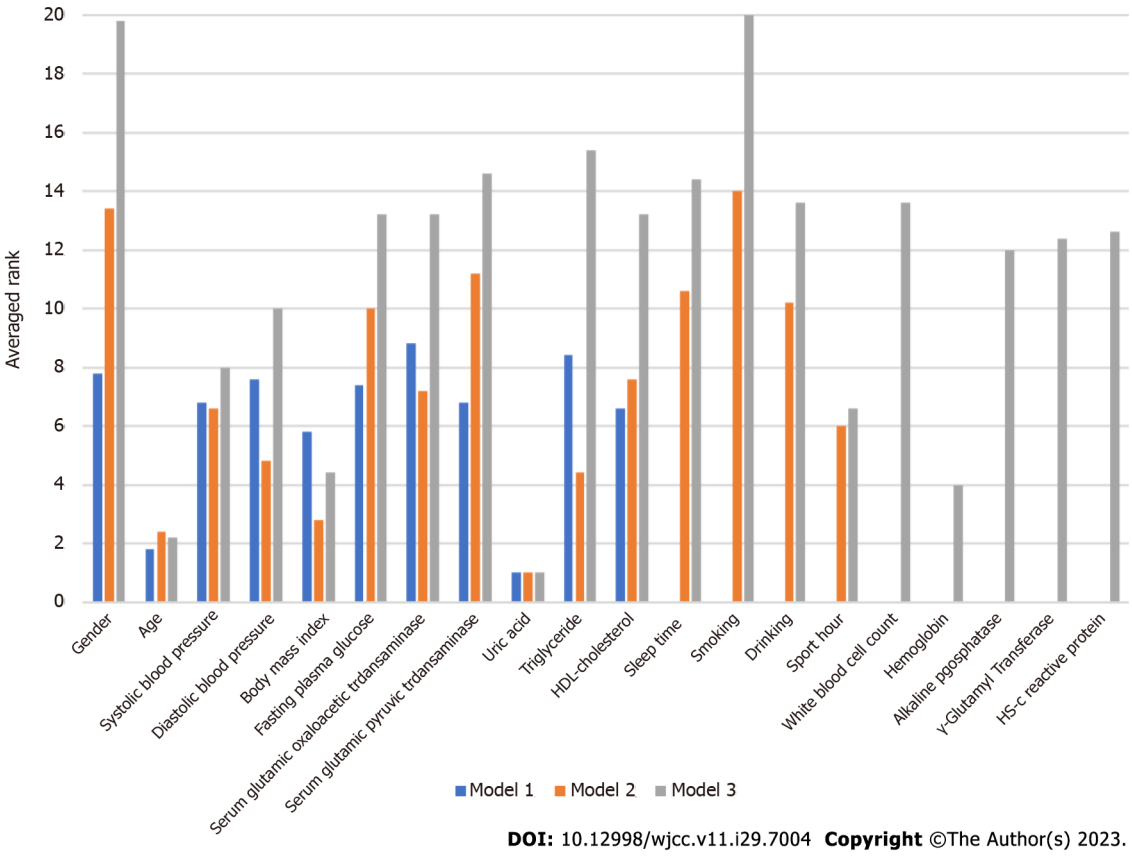
DOI: 10.12998/wjcc.v11.i29.7004

Figure 2 Area under receiver operating characteristic curve derived from five different machine learning methods in different models. A: Model 1; B: Model 2; C: Model 3. MARS: Multivariate adaptive regression splines; XGBoost: eXtreme Gradient Boosting; SGB: Stochastic gradient boosting; LightGBM: Light Gradient Boosting Machine; CATboost: Categorical Features + Gradient Boosting. Model 1: Demographic and biochemistry data; Model 2: Model 1 + lifestyle factors; Model 3: Model 2 + inflammation factors.

Table 4 Area under receiver operating characteristic curve derived from five different machine learning methods of the three different models

Model/AUC	Model 1	Model 2	Model 3
MARS	0.633	0.693	0.760
XGboost	0.602	0.663	0.786
SGB	0.599	0.666	0.814
LightGBM	0.615	0.646	0.776
Catboost	0.623	0.663	0.744
Mean	0.6144	0.6662	0.776

AUC: Area under receiver operating characteristic curve, MARS: Multivariate adaptive regression splines; XGBoost: eXtreme Gradient Boosting; SGB: Stochastic gradient boosting; LightGBM: Light Gradient Boosting Machine; CATboost: Categorical Features + Gradient Boosting. Model 1: Demographic and biochemistry data; Model 2: Model 1 + lifestyle factors; Model 3: Model 2 + inflammation factors.



DOI: 10.12998/wjcc.v11.i29.7004

Figure 3 Averaged ranks of importance for each risk factor in three different models. Model 1: Demographic and biochemistry data; Model 2: Model 1 + lifestyle factors; Model 3: Model 2 + inflammation factors.

increased in each for different Mach-L methods. Not surprisingly, Model 3 had the best AUC. Finally, the importance rankings for the three models are, respectively, shown in Tables 5-7. In Model 1, the most important risk factor was UA, followed by age, BMI, HDL-C, SBP, and GPT. When lifestyle factors were added, the ranking changed to UA, age, BMI, TG, DBP, and sport hours. Finally, integrating inflammation factors, the most important risk factor was UA, followed by age, Hb, BMI, sport hours, and SBP. The AUC of each model is, respectively, shown in Figures 2 to 3, while Table 3 presents the numerical values of the changes to each model. As shown in Figure 2, Model 3 had the highest AUC value. Figure 3 first compares the relative importance of each variable in the models, with color coded in blue, orange, and grey, respectively, for Models 1-3. The figure shows that gender was of greater importance in Model 1 than in Model 3, where a lower value indicated greater importance. Next, comparing columns of the same color allows for a clear observation of the relative importance of the various factors in each model. For example, UA was the most important variable in Model 3, followed by age and BMI.

Table 5 Rank importance of risk factors in model 1, from the most important to the least

Variable	MARS	XGboost	SGB	LightGBM	Catboost	AVG
Uric acid	1	1	1	1	1	1
Age	1	2	2	2	2	1.8
Body mass index	11	3	8	4	3	5.8
HDL-cholesterol	11	4	4	9	5	6.6
Systolic blood pressure	11	6	5	5	7	6.8
Serum glutamic pyruvic transaminase	4	9	7	10	4	6.8
Fasting plasma glucose	5	10	10	3	9	7.4
Diastolic blood pressure	11	5	6	8	8	7.6
Gender	3	8	11	7	10	7.8
Triglyceride	11	11	3	6	11	8.4
Serum glutamic oxaloacetic transaminase	11	7	9	11	6	8.8

AVG: Average of the rank of importance; MARS: Multivariate adaptive regression splines; XGBoost: eXtreme Gradient Boosting; HDL: High-density lipoprotein; SGB: Stochastic gradient boosting; LightGBM: Light Gradient Boosting Machine; CATboost: Categorical Features + Gradient Boosting. Model 1: Demographic and biochemistry data; Model 2: Model 1 + lifestyle factors; Model 3: Model 2 + inflammation factors.

Table 6 Rank importance of risk factors in model 2, from the most important to the least

Variable	MARS	XGboost	SGB	LightGBM	Catboos	AVG
Uric acid	1	1	1	1	1	1
Age	2	2	2	4	2	2.4
Body mass index	2	5	2	2	3	2.8
Triglyceride	7	4	5	1	5	4.4
Diastolic blood pressure	3	7	3	9	2	4.8
Sport hours	4	6	6	10	4	6
Systolic blood pressure	14	1	4	3	11	6.6
Serum glutamic oxaloacetic transaminase	6	8	8	5	9	7.2
HDL-cholesterol	14	3	7	8	6	7.6
Fasting plasma glucose	14	10	10	6	10	10
Drinking	5	12	12	14	8	10.2
Sleep time	14	9	9	14	7	10.6
Serum glutamic pyruvic transaminase	14	11	11	7	13	11.2
Gender	14	13	14	14	12	13.4
Smoking	14	14	14	14	14	14

AVG: Average of the rank of importance; MARS: Multivariate adaptive regression splines; XGBoost: eXtreme Gradient Boosting; HDL: High-density lipoprotein; SGB: Stochastic gradient boosting; LightGBM: Light Gradient Boosting Machine; CATboost: Categorical Features + Gradient Boosting. Model 2: Model 1 + lifestyle factors.

Table 7 Rank importance of risk factors in model 3, from the most important to the least

Variable	MARS	XGboost	SGB	LightGBM	Catboost	AVG
Uric acid	1	1	1	1	1	1
Age	1	2	2	2	4	2.2
Hemoglobin	4	4	6	3	3	4
Body mass index	7	3	3	7	2	4.4
Sport hours	3	8	4	5	13	6.6
Systolic blood pressure	5	6	20	4	5	8
Diastolic blood pressure	20	5	5	6	14	10
Alkaline phosphatase	20	7	11	12	10	12
γ-glutamyl transferase	20	13	8	9	12	12.4
Hs-C reactive protein	9	12	10	15	17	12.6
Fasting plasma glucose	20	9	20	10	7	13.2
Serum glutamic oxaloacetic transaminase	20	18	7	13	8	13.2
HDL-cholesterol	20	10	13	8	15	13.2
Drinking	20	14	12	16	6	13.6
White blood cell count	8	11	20	11	18	13.6
Sleep time	6	17	20	18	11	14.4
Serum glutamic pyruvic transaminase	20	16	14	14	9	14.6
Triglyceride	20	15	9	17	16	15.4
Gender	20	20	20	20	19	19.8
Smoking	20	20	20	20	20	20

AVG: Average of the rank of importance; MARS: Multivariate adaptive regression splines; XGBoost: eXtreme Gradient Boosting; HDL: High-density lipoprotein; SGB: Stochastic gradient boosting; LightGBM: Light Gradient Boosting Machine; CATboost: Categorical Features + Gradient Boosting. Model 3: Model 2 + inflammation factors.

DISCUSSION

The present study evaluated the effects of lifestyle and inflammation factors on eGFR changes in an elderly Chinese cohort. Our data show that, even though lifestyle and inflammation factors did have some predictive impacts for L-eGFR, the main determinants are still traditional factors that had been discussed extensively in previous work, including UA, age, Hb, BMI, sport hours, and SBP.

For all three models, the various Mach-L methods all selected UA as the key factor for determining L-eGFR, a finding supported by previous work. In a four-year longitudinal study, Liu *et al*[29] showed that compared to the highest quartile of UA, subjects with lower UA (quartile 1) are at lower risk for having reduced renal function [hazard ratio = 0.64, 95% confidence interval (0.49–0.85)]. Zhang *et al*[30] reviewed ten randomized controlled trials, finding that, following febuxostate treatment, eGFR was consistently and significantly lower than that in the non-treatment group. From such evidence, it could be concluded that through different mechanisms, hyperuricemia can lead to vascular obstruction and renal hypoperfusion[31].

Age is well-known to be associated with decreased adaptive capacity which leads to morbidity and mortality[32]. In the present study, it is not surprising that age is the 2nd most important factor related to L-eGFR, and this result is consistent with most previous findings. The underlying pathophysiology for this phenomenon has been studied extensively, and loss of renal mass, hyalinization of the afferent capillary, sclerotic glomerular and tubulointerstitial fibrosis are the main causes, leading to reduced blood flow and ultrafiltration of the glomerular capillary along with reduced afferent arteriolar resistance, thus resulting in reduced eGFR[33].

In the present study, Hb level was the third most important risk factor for abnormal eGFR. It is well-known that CKD can cause anemia, and this correlation is strongly supported by the cornerstone study published in 2002 by Coresh *et al* [34] that found that, once the eGFR falls below 60 mL/min per 1.73 m^2, lower renal function is associated with a higher incidence of anemia[35]. On the other hand, anemia might also contribute to the deterioration of renal function. Subjects with anemia have lower exercise tolerance[36], poor left ventricular growth[37], and even higher risk of heart failure[38]. This suggests that even before the CKD, anemia might begin to damage renal function.

It is well-known that BMI is also related to eGFR. Few previous studies have examined this link. Chang *et al*[39] conducted a longitudinal study from 2008 to 2013 with 7357 CKD subjects, finding that subjects with a BMI < 18.5 kg/m^2 had lower eGFR declines compared to other BMI groups. Similar findings were also found in Japanese and Malaysians [40-42]. It is not surprising that higher body weight leads to poor renal function since obesity is related to various sequelae such as hyperglycemia, hypertension, dyslipidemia, and metabolic syndrome[43,44]. It must be stressed here that this is not a causal association, and further longitudinal studies are needed to elucidate our result.

Surprisingly, when sport hours was included in the model, it emerged as the 5th most important factor. An increasing number of publications have suggested that exercise is beneficial for many aspects of CKD. Regular exercise is recommended by the Renal Association Clinical Practice Guidelines to improve renal function[45]. The impact of exercise on renal function could be explained by reduced inflammation, nitric oxide, angiotensin II accumulation, and improved anabolic response in skeletal muscles[46].

High blood pressure is a well-known independent risk factor for decreased renal function[47-49]. The present study is the first to use Mach-L to identify SBP as the 6th most important factor. Our finding was not alone, and in a 7-year longitudinal study, Wang *et al*[50] followed 2383 rural Chinese between the ages of 40 and 60 years old, finding a dose-dependent relationship between blood pressure and eGFR. The highest rate of eGFR decline was observed among subjects with SBP over 140 mmHg (odds ratio 2.9, 95% confidence interval 1.6–5.1) or DBP over 90 mmHg (odds ratio 2.7, 95% confidence interval 1.6–4.6)[50]. Interestingly, both SBP and DBP were important for identifying H-eGFR and L-eGFR. This indicates that SBP and DBP have different and independent effects.

CONCLUSION

In conclusion, we have applied Mach-L techniques to identify and rank risk factors from demographic, biochemistry, and lifestyle factors along with inflammation markers for L-eGFR among elderly Chinese, finding that the most important factors are UA, age, Hb, BMI, sport hours, and SBP.

ARTICLE HIGHLIGHTS

Research background

The incidence of chronic kidney disease (CKD) has significantly increased in recent years, leading to substantial impacts on patient mortality rates.

Research motivation

Previous studies have identified various risk factors for CKD, but they mostly relied on traditional statistical methods, such as logistic regression, and focused only on a limited number of risk factors.

Research objectives

To evaluate the impact of lifestyle and chronic inflammation in identifying subjects with abnormal estimated glomerular rates among elderly Chinese elderly subjects.

Research methods

The main focus of this study is to utilize five machine learning methods (Mach-L) for identifying factors.

Research results

Our results showed that uric acid is the most important risk factor (inflammatory marker), followed by age, hemoglobin, body mass index, sport hours, and systolic blood pressure.

Research conclusions

The study highlights that among demographic, biochemistry, lifestyle risk factors, and inflammation markers, UA is the most crucial risk factor for identifying low estimated glomerular filtration rate in elderly Chinese individuals, followed by age, hemoglobin, body mass index, sport hours, and systolic blood pressure.

Research perspectives

Further longitudinal studies are warranted to validate and clarify the causal relationships between these factors and estimated glomerular filtration rate changes.

FOOTNOTES

Author contributions: Chen CH participated in the design and oversight of the study, and was involved in data collection; Wang CK participated in the design of the study and was involved in data collection; Wang CY was involved in data collection and assisted with data analysis; Chang CF drafted the manuscript and assisted with data analysis; Chu TW drafted the manuscript and assisted with data

analysis; all authors read and approved the final manuscript.

Supported by the Kaohsiung Armed Forces General Hospital.

Institutional review board statement: The study protocol was approved by the Institutional Review Board of the Tri-Service General Hospital, National Defense Medical Center (IRB No.: KAFGHIRB 109-46).

Informed consent statement: All study participants, or their legal guardian, provided written consent prior to study enrollment.

Conflict-of-interest statement: All the authors of this manuscript have no conflicts of interest to disclose.

Data sharing statement: There is no additional data available.

Country/Territory of origin: Taiwan

ORCID number: Chao-Hung Chen 0009-0005-3207-9627; Chun-Kai Wang 0009-0002-4491-240X; Chen-Yu Wang 0000-0002-4985-2488; Chun-Feng Chang 0009-0007-5629-4203; Ta-Wei Chu 0000-0002-7629-7854.

S-Editor: Liu JH
L-Editor: Wang TQ
P-Editor: Zhao S

REFERENCES

1 **Copur S**, Tanriover C, Yavuz F, Soler MJ, Ortiz A, Covic A, Kanbay M. Novel strategies in nephrology: what to expect from the future? *Clin Kidney J* 2023; **16**: 230-244 [PMID: 36755838 DOI: 10.1093/ckj/sfac212]
2 **Sheen YJ**, Hsu CC, Jiang YD, Huang CN, Liu JS, Sheu WH. Trends in prevalence and incidence of diabetes mellitus from 2005 to 2014 in Taiwan. *J Formos Med Assoc* 2019; **118** Suppl 2: S66-S73 [PMID: 31300322 DOI: 10.1016/j.jfma.2019.06.016]
3 **Tsai MH**, Hsu CY, Lin MY, Yen MF, Chen HH, Chiu YH, Hwang SJ. Incidence, Prevalence, and Duration of Chronic Kidney Disease in Taiwan: Results from a Community-Based Screening Program of 106,094 Individuals. *Nephron* 2018; **140**: 175-184 [PMID: 30138926 DOI: 10.1159/000491708]
4 **Chadban SJ**, Briganti EM, Kerr PG, Dunstan DW, Welborn TA, Zimmet PZ, Atkins RC. Prevalence of kidney damage in Australian adults: The AusDiab kidney study. *J Am Soc Nephrol* 2003; **14**: S131-S138 [PMID: 12819318 DOI: 10.1097/01.ASN.0000070152.11927.4A]
5 **Drew DA**, Weiner DE, Sarnak MJ. Cognitive Impairment in CKD: Pathophysiology, Management, and Prevention. *Am J Kidney Dis* 2019; **74**: 782-790 [PMID: 31378643 DOI: 10.1053/j.ajkd.2019.05.017]
6 **Hannan M**, Ansari S, Meza N, Anderson AH, Srivastava A, Waikar S, Charleston J, Weir MR, Taliercio J, Horwitz E, Saunders MR, Wolfrum K, Feldman HI, Lash JP, Ricardo AC; CRIC Study Investigators; Chronic Renal Insufficiency Cohort (CRIC) Study Investigators. Risk Factors for CKD Progression: Overview of Findings from the CRIC Study. *Clin J Am Soc Nephrol* 2021; **16**: 648-659 [PMID: 33177074 DOI: 10.2215/CJN.07830520]
7 **Imig JD**, Ryan MJ. Immune and inflammatory role in renal disease. *Compr Physiol* 2013; **3**: 957-976 [PMID: 23720336 DOI: 10.1002/cphy.c120028]
8 **Mitchell TM**. Machine Learning. 1997; McGraw-Hill Science/Engineering/Math. Available from: https://www.cin.ufpe.br/~cavmj/Machine%20-%20Learning%20-20Tom%20Mitchell.pdf
9 **Nusinovici S**, Tham YC, Chak Yan MY, Wei Ting DS, Li J, Sabanayagam C, Wong TY, Cheng CY. Logistic regression was as good as machine learning for predicting major chronic diseases. *J Clin Epidemiol* 2020; **122**: 56-69 [PMID: 32169597 DOI: 10.1016/j.jclinepi.2020.03.002]
10 **Wu X**, Tsai SP, Tsao CK, Chiu ML, Tsai MK, Lu PJ, Lee JH, Chen CH, Wen C, Chang SS, Hsu CY, Wen CP. Cohort Profile: The Taiwan MJ Cohort: half a million Chinese with repeated health surveillance data. *Int J Epidemiol* 2017; **46**: 1744-1744g [PMID: 28204597 DOI: 10.1093/ije/dyw282]
11 **MJ Health Research Foundation**, MJ Health Resource Center. The introduction of MJ Health Database. Technical Report No. MJHRF-TR-01. August, 2016
12 **MJ Health Research Foundation**. (2014). MJ Health Survey Database, MJ BioData [Data file], MJ BioBank [Biological specimen]. Available from: http://www.mjhrf.org
13 **Lu CH**, Pei D, Wu CZ, Kua HC, Liang YJ, Chen YL, Lin JD. Predictors of abnormality in thallium myocardial perfusion scans for type 2 diabetes. *Heart Vessels* 2021; **36**: 180-188 [PMID: 32816060 DOI: 10.1007/s00380-020-01681-2]
14 **Tseng CJ**, Lu CJ, Chang CC, Chen GD, Cheewakriangkrai C. Integration of data mining classification techniques and ensemble learning to identify risk factors and diagnose ovarian cancer recurrence. *Artif Intell Med* 2017; **78**: 47-54 [PMID: 28764872 DOI: 10.1016/j.artmed.2017.06.003]
15 **Chang CC**, Chen SH. Developing a Novel Machine Learning-Based Classification Scheme for Predicting SPCs in Breast Cancer Survivors. *Front Genet* 2019; **10**: 848 [PMID: 31620166 DOI: 10.3389/fgene.2019.00848]
16 **Shih CC**, Lu CJ, Chen GD, Chang CC. Risk Prediction for Early Chronic Kidney Disease: Results from an Adult Health Examination Program of 19,270 Individuals. *Int J Environ Res Public Health* 2020; **17** [PMID: 32664271 DOI: 10.3390/ijerph17144973]

17 **Lee TS**, Chen IF, Chang TJ, Lu CJ. Forecasting Weekly Influenza Outpatient Visits Using a Two-Dimensional Hierarchical Decision Tree Scheme. *Int J Environ Res Public Health* 2020; **17** [PMID: 32630311 DOI: 10.3390/ijerph17134743]

18 **Chang CC**, Yeh JH, Chen YM, Jhou MJ, Lu CJ. Clinical Predictors of Prolonged Hospital Stay in Patients with Myasthenia Gravis: A Study Using Machine Learning Algorithms. *J Clin Med* 2021; **10** [PMID: 34640412 DOI: 10.3390/jcm10194393]

19 **Chang CC**, Huang TH, Shueng PW, Chen SH, Chen CC, Lu CJ, Tseng YJ. Developing a Stacked Ensemble-Based Classification Scheme to Predict Second Primary Cancers in Head and Neck Cancer Survivors. *Int J Environ Res Public Health* 2021; **18** [PMID: 34886225 DOI: 10.3390/ijerph182312499]

20 **Chiu YL**, Jhou MJ, Lee TS, Lu CJ, Chen MS. Health Data-Driven Machine Learning Algorithms Applied to Risk Indicators Assessment for Chronic Kidney Disease. *Risk Manag Healthc Policy* 2021; **14**: 4401-4412 [PMID: 34737657 DOI: 10.2147/RMHP.S319405]

21 **Wu TE**, Chen HA, Jhou MJ, Chen YN, Chang TJ, Lu CJ. Evaluating the Effect of Topical Atropine Use for Myopia Control on Intraocular Pressure by Using Machine Learning. *J Clin Med* 2020; **10** [PMID: 33396943 DOI: 10.3390/jcm10010111]

22 **Wu CW**, Shen HL, Lu CJ, Chen SH, Chen HY. Comparison of Different Machine Learning Classifiers for Glaucoma Diagnosis Based on Spectralis OCT. *Diagnostics (Basel)* 2021; **11** [PMID: 34574059 DOI: 10.3390/diagnostics11091718]

23 **Chang CC**, Yeh JH, Chiu HC, Chen YM, Jhou MJ, Liu TC, Lu CJ. Utilization of Decision Tree Algorithms for Supporting the Prediction of Intensive Care Unit Admission of Myasthenia Gravis: A Machine Learning-Based Approach. *J Pers Med* 2022; **12** [PMID: 35055347 DOI: 10.3390/jpm12010032]

24 **Friedman JH**. Greedy function approximation: a gradient boosting machine. *Annals of statistics* 2001; 1189-1232 [DOI: 10.1214/aos/1013203451]

25 **Tierney NJ**, Harden FA, Harden MJ, Mengersen KL. Using decision trees to understand structure in missing data. *BMJ Open* 2015; **5**: e007450 [PMID: 26124509 DOI: 10.1136/bmjopen-2014-007450]

26 **Breiman L**. Random Forests. *Machine Learning* 2001; **45**: 5-32 [DOI: 10.1023/A:1010933404324]

27 **Ke G**, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu TY. Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems, 2017; 30. Available from: https://papers.nips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

28 **Dorogush AV**, Ershov V, Gulin A. CatBoost: gradient boosting with categorical features support. arXiv 2018 [DOI: 10.48550/arXiv.1810.11363]

29 **Liu L**, You L, Sun K, Li F, Qi Y, Chen C, Wang C, Lao G, Xue S, Tang J, Li N, Feng W, Yang C, Xu M, Li Y, Yan L, Ren M, Lin D. Association between uric acid lowering and renal function progression: a longitudinal study. *PeerJ* 2021; **9**: e11073 [PMID: 33828916 DOI: 10.7717/peerj.11073]

30 **Zhang S**, Xu T, Shi Q, Li S, Wang L, An Z, Su N. Cardiovascular Safety of Febuxostat and Allopurinol in Hyperuricemic Patients With or Without Gout: A Network Meta-Analysis. *Front Med (Lausanne)* 2021; **8**: 698437 [PMID: 34211992 DOI: 10.3389/fmed.2021.698437]

31 **Sánchez-Lozada LG**, Tapia E, Santamaría J, Avila-Casado C, Soto V, Nepomuceno T, Rodríguez-Iturbe B, Johnson RJ, Herrera-Acosta J. Mild hyperuricemia induces vasoconstriction and maintains glomerular hypertension in normal and remnant kidney rats. *Kidney Int* 2005; **67**: 237-247 [PMID: 15610247 DOI: 10.1111/j.1523-1755.2005.00074.x]

32 **Yin D**, Chen K. The essential mechanisms of aging: Irreparable damage accumulation of biochemical side-reactions. *Exp Gerontol* 2005; **40**: 455-465 [PMID: 15935593 DOI: 10.1016/j.exger.2005.03.012]

33 **Weinstein JR**, Anderson S. The aging kidney: physiological changes. *Adv Chronic Kidney Dis* 2010; **17**: 302-307 [PMID: 20610357 DOI: 10.1053/j.ackd.2010.05.002]

34 **Coresh J**, Astor BC, McQuillan G, Kusek J, Greene T, Van Lente F, Levey AS. Calibration and random variation of the serum creatinine assay as critical elements of using equations to estimate glomerular filtration rate. *Am J Kidney Dis* 2002; **39**: 920-929 [PMID: 11979335 DOI: 10.1053/ajkd.2002.32765]

35 **National Center for Health Statistics (US)**. Plan and operation of the third National Health and Nutrition Examination Survey, 1994; 1988-1994 (No. 32). National Ctr for Health Statistics

36 Association between recombinant human erythropoietin and quality of life and exercise capacity of patients receiving haemodialysis. Canadian Erythropoietin Study Group. *BMJ* 1990; **300**: 573-578 [PMID: 2108751 DOI: 10.1136/bmj.300.6724.573]

37 **Foley RN**, Parfrey PS, Kent GM, Harnett JD, Murray DC, Barre PE. Long-term evolution of cardiomyopathy in dialysis patients. *Kidney Int* 1998; **54**: 1720-1725 [PMID: 9844150 DOI: 10.1046/j.1523-1755.1998.00154.x]

38 **Harnett JD**, Foley RN, Kent GM, Barre PE, Murray D, Parfrey PS. Congestive heart failure in dialysis patients: prevalence, incidence, prognosis and risk factors. *Kidney Int* 1995; **47**: 884-890 [PMID: 7752588 DOI: 10.1038/ki.1995.132]

39 **Chang TJ**, Zheng CM, Wu MY, Chen TT, Wu YC, Wu YL, Lin HT, Zheng JQ, Chu NF, Lin YM, Su SL, Lu KC, Chen JS, Sung FC, Lee CT, Yang Y, Hwang SJ, Wang MC, Hsu YH, Chiou HY, Kao S, Lin YF. Relationship between body mass index and renal function deterioration among the Taiwanese chronic kidney disease population. *Sci Rep* 2018; **8**: 6908 [PMID: 29720598 DOI: 10.1038/s41598-018-24757-6]

40 **Horber FF**, Gruber B, Thomi F, Jensen EX, Jaeger P. Effect of sex and age on bone mass, body composition and fuel metabolism in humans. *Nutrition* 1997; **13**: 524-534 [PMID: 9263233 DOI: 10.1016/S0899-9007(97)00031-2]

41 **Kuk JL**, Lee S, Heymsfield SB, Ross R. Waist circumference and abdominal adipose tissue distribution: influence of age and sex. *Am J Clin Nutr* 2005; **81**: 1330-1334 [PMID: 15941883 DOI: 10.1093/ajcn/81.6.1330]

42 **Hopper J Jr**, Trew PA, Biava CG. Membranous nephropathy: its relative benignity in women. *Nephron* 1981; **29**: 18-24 [PMID: 7329469 DOI: 10.1159/000182232]

43 **Shin HY**, Linton JA, Shim JY, Kang HT. Cancer survivors aged 40 years or elder are associated with high risk of chronic kidney disease: the 2010-2012 Korean National Health and Nutrition Examination Survey. *Asian Pac J Cancer Prev* 2015; **16**: 1355-1360 [PMID: 25743798 DOI: 10.7314/APJCP.2015.16.4.1355]

44 **Wong G**, Hayen A, Chapman JR, Webster AC, Wang JJ, Mitchell P, Craig JC. Association of CKD and cancer risk in older people. *J Am Soc Nephrol* 2009; **20**: 1341-1350 [PMID: 19406977 DOI: 10.1681/ASN.2008090998]

45 **Baker LA**, March DS, Wilkinson TJ, Billany RE, Bishop NC, Castle EM, Chilcot J, Davies MD, Graham-Brown MPM, Greenwood SA, Junglee NA, Kanavaki AM, Lightfoot CJ, Macdonald JH, Rossetti GMK, Smith AC, Burton JO. Clinical practice guideline exercise and lifestyle in chronic kidney disease. *BMC Nephrol* 2022; **23**: 75 [PMID: 35193515 DOI: 10.1186/s12882-021-02618-1]

46 **Bishop NC**, Burton JO, Graham-Brown MPM, Stensel DJ, Viana JL, Watson EL. Exercise and chronic kidney disease: potential mechanisms underlying the physiological benefits. *Nat Rev Nephrol* 2023; **19**: 244-256 [PMID: 36650232 DOI: 10.1038/s41581-022-00675-9]

47 **Haroun MK**, Jaar BG, Hoffman SC, Comstock GW, Klag MJ, Coresh J. Risk factors for chronic kidney disease: a prospective study of 23,534

men and women in Washington County, Maryland. *J Am Soc Nephrol* 2003; **14**: 2934-2941 [PMID: 14569104 DOI: 10.1097/01.ASN.0000095249.99803.85]

48 **Tozawa M**, Iseki K, Iseki C, Kinjo K, Ikemiya Y, Takishita S. Blood pressure predicts risk of developing end-stage renal disease in men and women. *Hypertension* 2003; **41**: 1341-1345 [PMID: 12707291 DOI: 10.1161/01.HYP.0000069699.92349.8C]

49 **Fox CS**, Larson MG, Leip EP, Culleton B, Wilson PW, Levy D. Predictors of new-onset kidney disease in a community-based population. *JAMA* 2004; **291**: 844-850 [PMID: 14970063 DOI: 10.1001/jama.291.7.844]

50 **Wang Q**, Xie D, Xu X, Qin X, Tang G, Wang B, Wang Y, Hou F, Wang X. Blood pressure and renal function decline: a 7-year prospective cohort study in middle-aged rural Chinese men and women. *J Hypertens* 2015; **33**: 136-143 [PMID: 25255396 DOI: 10.1097/HJH.0000000000000360]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com