Prediction of post-nephroureterectomy renal function in non-dialysis upper tract urothelial carcinoma using machine learning with PyCaret and SHAP explanations

HSIANG YING LEE[1-3], WEI-MING LI[1-4], HSIN-CHIH YEH[1-3,5], WEN-JENG WU[1-3], CHING-CHIA LI[1-4], HUNG-LUNG KE[1-3,5] and YEN-CHUN WANG[1,6]

[1]Department of Urology, Kaohsiung Medical University Hospital, Kaohsiung 807, Taiwan, R.O.C.; [2]Department of Urology, School of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung 807, Taiwan, R.O.C.; [3]Graduate Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung 807, Taiwan, R.O.C.; [4]Department of Urology, Kaohsiung Medical University Gangshan Hospital, Kaohsiung 820, Taiwan, R.O.C.; [5]Department of Urology, Kaohsiung Municipal Ta-Tung Hospital, Kaohsiung 801, Taiwan, R.O.C.; [6]Department of Public Health, College of Medicine, National Cheng Kung University, Tainan 701, Taiwan, R.O.C.

Received October 31, 2024; Accepted January 28, 2025

DOI: 10.3892/ol.2025.14946

Abstract. Renal function after radical nephroureterectomy (RNU) in patients with advanced upper tract urothelial carcinoma (UTUC) is an indicator of eligibility for postoperative chemotherapy. The present study aimed to utilize machine learning for predicting renal function status after RNU and to investigate the contribution of several features to this prediction. The present retrospective study included 764 medical records of patients with non-metastatic UTUC who received RNU from 2008 to 2022. The records were divided into training (n=534) and testing (n=230) datasets. Several demographic and clinicopathological parameters were collected for analysis. PyCaret was utilized for data processing, model establishment and comparison, while SHapley Additive exPlanations values were adopted to assess the contribution of each variable to the predictive performance of the model. The results demonstrated that the Random Forest Regressor model had improved accuracy in predicting postoperative renal function compared with other algorithms. In both the estimated glomerular filtration rate (eGFR) and chronic kidney disease outcome, preoperative renal function had the most marked contribution to postoperative renal function. Additionally, the Charlson Comorbidity Index (CCI), body mass index (BMI) and tumor size were highly associated with renal function after RNU, and hydronephrosis was another main factor for predicting postoperative eGFR. The retrospective design was a limitation of the present study. The results demonstrated that important predictors, including preoperative eGFR, tumor size, BMI, CCI and hydronephrosis, were associated with postoperative renal function. To conclude, using a machine learning prediction model may be used in the future to determine appropriate therapeutic strategies and inform the timing of perioperative systemic chemotherapy, in particular perioperative systemic chemotherapy, for advanced UTUC.

Correspondence to: Mr. Yen-Chun Wang, Department of Urology, Kaohsiung Medical University Hospital, 100 Shih-Chuan 1st Road, Sanmin, Kaohsiung 807, Taiwan, R.O.C.
E-mail: ashum1009@gmail.com

Key words: prediction of renal function, upper tract urothelial carcinoma, machine learning

Introduction

Upper tract urothelial carcinoma (UTUC) accounts for ~40% of all UC cases, with a higher prevalence among women than men in Taiwan (1,2). The incidence of UTUC has been rising in previous years, predominantly impacting the elderly population (3,4). As UTUC is aggressive and ~60% of UTUCs are invasive at presentation, improving the oncological outcome is imperative, especially in the current comorbid patient population (5). Malnutritional status is also a poor prognostic factor of survival outcome in patients with UTUC receiving radical surgery (6).

Radical nephroureterectomy (RNU) with bladder cuff excision is a curative treatment in UTUC. Although early-stage disease achieves durable long-term control, the 5-year overall survival rate for patients with non-organ-confined UTUC is <50% and it drops to <35% for those with nodal metastases. Given the limited effectiveness of surgery alone in treating locally advanced UTUC, there is a crucial need to develop multimodal treatment strategies to improve the survival rates for patients with poor prognostic UTUC (7).

The optimal timing for perioperative systemic chemotherapy is still controversial. Cisplatin-based treatments are presently administered in both neoadjuvant chemotherapy (NAC) and adjuvant chemotherapy (AC) settings. The importance of AC for patients with the aim of a cure has been confirmed (8). However, patient eligibility must be evaluated, especially concerning renal function in relation to definitive

surgical treatment. The eligibility for cisplatin-based chemotherapy decreases after RNU, as some patients are unable to benefit from the advantage of AC. A previous systematic review and meta-analysis reported that patients with nonmetastatic UTUC who received NAC before RNU experienced improved OS and CSS compared with those who underwent RNU alone (9). NAC has the potential to offer these patients the benefits of chemotherapy. However, NAC presents several disadvantages. Firstly, it may postpone definitive RNU, potentially allowing disease progression, especially in patients who are chemoresistant. Secondly, chemotherapy-related toxicities in NAC recipients can further delay surgery. Lastly, there is a risk of overtreatment in patients who do not have pathologically confirmed muscle-invasive disease (9).

A phase 3 randomized trial from the UK presented compelling evidence supporting the use of AC. These results showed a 55% improvement in disease-free survival at a median follow-up period of 30.3 months. This study showed the use of AC may be superior when compared with RNU alone. The primary disadvantage of AC in the treatment of patients with UTUC is the decline in renal function following RNU (8) In addition, in the clinical trial, a subgroup analysis revealed that patients treated with cisplatin experienced notable benefits, whereas those treated with carboplatin did not (8,9).

The decision to administer NAC is difficult due to challenges in obtaining an accurate diagnosis. Currently, conventional imaging and endoscopic biopsy often underestimate the cancer stage (10). Therefore, developing nomograms that predict an advanced pathological stage at RNU could enhance the ability to identify patients who are most likely to benefit from preoperative chemotherapy. The aim of the present study was to create a model to predict renal function after RNU for improved planning of perioperative chemotherapy.

Materials and methods

Study cohort and clinical characteristics. A total of 764 medical records of patients with non-metastatic UTUC who received RNU from January 2008 to December 2022 at Kaohsiung Medical University Hospital (Kaohsiung, Taiwan) and Kaohsiung Municipal Ta-Tung Hospital (Kaohsiung, Taiwan) were retrospectively included in the present study. The present study was approved by Kaohsiung Medical University Hospital ethics committee [approval no. KMUHIRB-E(I)-20180214]. The inclusion criteria used were as follows: i) Patients with UTUC; and ii) patients who underwent RNU. The exclusion criteria used were as follows: i) Patients undergoing dialysis; and ii) missing eGFR data from medical records. The included data was divided into the training (n=534) and testing (n=230) datasets. The flowchart of the inclusion process is shown in Fig. 1. The training dataset was used to preliminarily train the prediction model, and the model was subsequently tested using the testing dataset. Several demographic and clinicopathological parameters were collected to predict postoperative renal function, including preoperative estimated glomerular filtration rate (eGFR), tumor size, body mass index (BMI), Charlson Comorbidity Index (CCI), pathological tumor stage ≥2 (pT2), pathological lymph node stage ≥Nx/N0 (pN), lymphovascular invasion (LVI), diabetes mellitus (DM), sex, hypertension (HTN), smoking, radiotherapy (RT), surgical margin, grade, carcinoma *in situ* (CIS), perineural invasion (PNI) and time. The Chronic Kidney Disease (CKD)-Epidemiology Collaboration equation was used to calculate the GFR, defined as: GFR=141 x $\min(Scr/\kappa, 1)^{\alpha}$ x $\max(Scr/\kappa, 1)^{-1.209}$ x 0.993^{age} x 1.018 (if female) x 1.159 (if black), where Scr is serum creatinine (mg/dl), κ is 0.7 for women and 0.9 for men, α is-0.329 for female patients and -0.411 for male patients, min indicates the minimum of Scr/κ or 1 and max indicates the maximum of Scr/κ or 1 (11). The definition of CKD according to current The Kidney Disease: Improving Global Outcomes international guidelines (2012) is GFR <60 ml/min per 1.73 m^2 (12,13).

Statistical analysis. The unpaired t-test (normalized data) and Mann-Whitney U test (non-normalized data) were used to assess the association between the dependent variable and continuous variables, respectively. Pearson's chi-square test or Fisher's exact test were used to assess the association between dependent and categorical variables. $P<0.05$ was used to indicate a statistically significant difference.

Linear regression is a foundational algorithm in machine learning that predicts continuous outcomes by modeling the relationship between independent and dependent variables. Logistic regression estimates probabilities using the logistic function and is employed for binary classification tasks. Regularization techniques like Lasso regression and Ridge regression address overfitting by penalizing large coefficients. The Elastic Net combines both Lasso and Ridge penalties for greater flexibility. Huber Regression is robust to outliers, blending squared and absolute loss functions. Support Vector Machines are versatile for classification and regression, leveraging hyperplanes for separation. K-Nearest Neighbors is a simple, non-parametric method relying on proximity for predictions. Naïve Bayes assumes feature independence and is efficient for text classification. Decision Trees provide interpretable models by splitting data hierarchically, while ensemble methods like Gradient Boosting, XGBoost, LightGBM, AdaBoost, CatBoost, Random Forest and Extra Trees combine multiple models to enhance performance. Least-angle regression and its Lasso variant are efficient for high-dimensional data. Orthogonal Matching Pursuit is a greedy algorithm for sparse representation to make locally optimal choices at each step and find a global optimum solution. Dummy Regression serves as a baseline model. Passive-aggressive regression is designed for online learning scenarios (14).

All analyses were performed using Python (version 3.11; Python Software Foundation), and machine learning algorithm was implemented with the PyCaret 3.0.4 package (Moez Ali), on a 64-bit Windows 11 (Microsoft Corporation) computer. PyCaret was used for data preprocessing and training the model with the available algorithms using a 10-fold cross-validation. Hyperparameter tuning was automatically performed via random grid search. The dataset was split into a 70% training set for model development and a 30% test set for validation. The regression was used to establish the prediction model for the continuous target (eGFR) and the classification model was used for the category target (CKD). To assess the accuracy of the machine learning models, R^2 and accuracy (the proportion of all classifications that were correct) were utilized as performance evaluations for predicting eGFR and

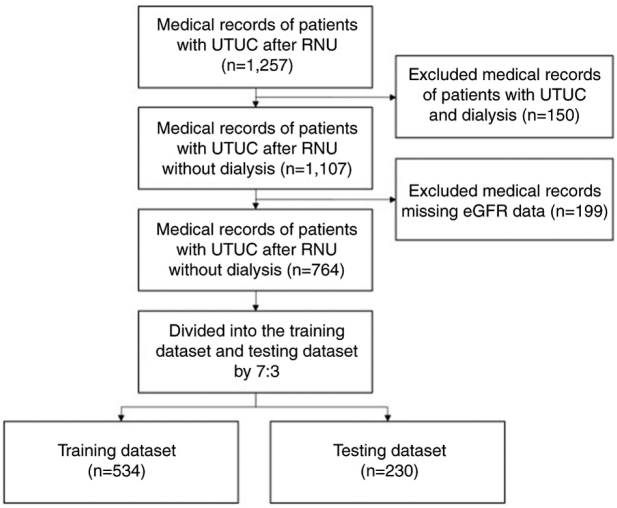


Figure 1. Process flow diagram. UTUC, upper tract urothelial carcinoma; RNU, radical nephroureterectomy; eGFR, estimated glomerular filtration rate.

CKD, respectively. The best model with the highest accuracy would be further explored, and the ROC curve was formulated according to the model. The SHapley Additive exPlanations (SHAP) method was used to identify key feature factors, presenting individual SHAP values in a violin plot, and the ranking of the mean absolute SHAP values were shown in a bar chart to represent the impact mode of features and the importance ranking.

Results

Distribution of patient characteristics. There was no significant difference between the median age of the training [70 (64-76) years] and the testing [69 (63-76) years] datasets (Table I), and 48.9% of the training dataset were male compared with 43.9% of the testing dataset. In addition, there were no significant differences in patient characteristics (BMI, preoperative eGFR, DM, HTN, hydronephrosis, RT, CCI, time, and preoperative CKD) and tumor characteristics (pT2, pN, grade, LVI, PNI, surgical margin, CIS, and tumor size) between the training and test sets.

Prediction of postoperative eGFR. As shown in Table II, the Random Forest Regressor outperformed the other algorithms and achieved the highest average score across multiple indicators. Fig. 2A illustrated the correlation between the predicted and actual values of eGFR when using Random Forest Regressor, showing that the model tended to overestimate when the actual value was low and underestimate when the actual value was high. Both the training and test sets achieved an R^2 >0.6, and there was a larger residual in the test set compared with the training set (Fig. 2B). The summary of the model performance was assessed, including the mean square error (141.279), root mean squared error (11.886), mean absolute error (8.888), mean absolute percentage error (0.287) and R^2 (0.674) (Fig. 2C). The R^2 showed that the model explained more than one-half of the variance.

The SHAP method identified the importance of features for the Random Forest Regressor model, and violin and bar plots were used to evaluate the correlation between features and predicted values, as well as their importance in prediction. In the violin plot (Fig. 3A), the higher feature value was colored red and the lower feature value was colored blue. If the red points are distributed on right side of center axis (SHAP value=0), it means that a higher feature value corresponds to a higher SHAP value. On the contrary, the higher feature value is associated with the lower SHAP value when the red points are scattered on the left. In the bar plots shown in Fig. 3B, all features were sorted by the mean absolute SHAP values (Fig. 3B). Preoperative eGFR, tumor size, BMI, CCI and hydronephrosis were the top five factors attributed to prediction, among which preoperative eGFR had the greatest contribution to the model for prediction. The features in the top five showed a positive correlation with the SHAP value, except for CCI. Other features, including pT stage, LVI, pN stage, DM, sex, HTN, smoking, RT, surgical margin, grade, CIS and PNI also had medium to small contributions to predicting post-operative eGFR. The ability to predict time was minimal.

Prediction of postoperative CKD status. Regarding CKD, the Random Forest Classifier, with the best accuracy, also had the highest accuracy compared with other algorithms regarding the predicted index scores (Table III). The model's performance is illustrated in Fig. 4, with the confusion matrix results (Fig. 4A) indicating that 208 (90.4%) patients were accurately classified, 182 (79.1%) were patients with CKD and 26 (11.3%) were healthy controls. Fig. 4B displays the receiver operating characteristic (ROC) curve for the best model. However, the model's performance in predicting healthy groups was lower compared with patients with CKD (Fig. 4C). The overall model performance is summarized in Fig. 4D, with accuracy (0.904), precision (0.905), recall (0.984) and F1 score (0.943) all >0.9. In Fig. 5, the results from the SHAP method were consistent with the eGFR prediction analysis, with preoperative CKD serving as the most notable role for predicting the post-operative CKD status. The top four factors with the highest SHAP values were preoperative CKD, CCI, tumor size and BMI. Preoperative CKD and higher CCI had above-average SHAP values, while higher BMI and tumor size had below-average SHAP values. Other characteristics, including pT stage, LVI, pN stage, DM, sex, HTM, smoking, RT, surgical margin, grade, CIS, hydronephrosis and PNI, were also of similar importance in predicting postoperative CKD. The SHAP values for time were almost zero.

Discussion

For patients with advanced UTUC, optimum timing of perioperative systemic treatment is important to decrease cancer progression and recurrence. Although NAC can be a good strategy to decrease the stage of the tumor, there is a lack of confidence in the existing staging methods before surgery (15). The issue of over-staging and the subsequent over-treatment is a concern. However, some patients who are eligible for chemotherapy before surgery may become ineligible after surgery. Factors predicting a decline in renal function due to RNU may

Table I. Characteristics and distribution of patients after radical nephroureterectomy in the training and testing datasets.

Characteristic	Training dataset (n=534)	Testing dataset (n=230)	P-value
Age, years [median (Q1-Q3)]	70 (64-76)	69 (63-76)	0.4820
BMI, kg/m^2 [median (Q1-Q3)]	24.50 (22.04-27.03)	24.63 (22.00-27.39)	0.8208
Tumor size, cm [median (Q1-Q3)]	3.0 (2.2-4.5)	3.2 (1.9-5.0)	0.9774
Charlson Comorbidity Index, [median (Q1-Q3)]	4.0 (3.0-5.0)	4.5 (3.0-6.0)	0.6381
Preoperative estimated glomerular filtration rate, ml/min/1.73 m^2 [median (Q1-Q3)]	53.84 (36.74-73.53)	51.52 (36.76-67.75)	0.3489
Carcinoma *in situ*, n (%)			0.8810
No	449.0 (70.3)	190.0 (29.7)	
Yes	43.0 (69.4)	19.0 (30.6)	
Diabetes mellitus, n (%)			0.4503
No	370.0 (70.7)	153.0 (29.3)	
Yes	164.0 (68.1)	77.0 (31.9)	
Sex, n (%)			0.2075
Female	273.0 (67.9)	129.0 (32.1)	
Male	261.0 (72.1)	101.0 (27.9)	
Pathological grade, n (%)			0.5221
Low	65.0 (73.0)	24.0 (27.0)	
High	449.0 (69.7)	195.0 (30.3)	
Hypertension, n (%)			0.9976
No	188.0 (69.9)	81.0 (30.1)	
Yes	346.0 (69.9)	149.0 (30.1)	
Hydronephrosis, n (%)			0.1818
No	168.0 (73.0)	62.0 (27.0)	
Yes	358.0 (68.2)	167.0 (31.8)	
Lymphovascular invasion, n (%)			0.9095
No	405.0 (70.2)	172.0 (29.8)	
Yes	106.0 (70.7)	44.0 (29.3)	
Surgery margin of resection, n (%)			0.9330
No	482.0 (69.9)	208.0 (30.1)	
Yes	31.0 (70.4)	13.0 (29.6)	
Perineural invasion, n (%)			0.9891
No	473.0 (70.3)	200.0 (29.7)	
Yes	38.0 (70.4)	16.0 (29.6)	
Radiotherapy, n (%)			0.6216
No	480.0 (70.2)	204.0 (29.8)	
Yes	54.0 (67.5)	26.0 (32.5)	
Smoking, n (%)			0.5931
No	418.0 (69.4)	184.0 (30.6)	
Yes	116.0 (71.6)	46.0 (28.4)	
Time of postoperative eGFR measurement, months [n (%)]			0.6037
<3	235.0 (69.5)	103.0 (30.5)	
3-12	194.0 (71.9)	76.0 (28.1)	
12-24	105.0 (67.3)	51.0 (32.7)	
Pathological lymph node stage, n (%)			0.3164
N0/Nx	118.0 (73.3)	43.0 (26.7)	
N1/N2	400.0 (69.2)	178.0 (30.8)	

Table I. Continued.

Characteristic	Training dataset (n=534)	Testing dataset (n=230)	P-value
Pathological tumor stage, n (%)			0.5578
<2	241.0 (71.1)	98.0 (28.9)	
≥2	282.0 (69.1)	126.0 (30.9)	
Postoperative CDK status			0.1868
No	208.0 (72.7)	78.0 (27.3)	
Yes	326.0 (68.2)	152.0 (31.8)	

Q, quartile.

Table II. Performance of 20 machine learning algorithms for predicting estimated glomerular filtration rate in the training dataset.

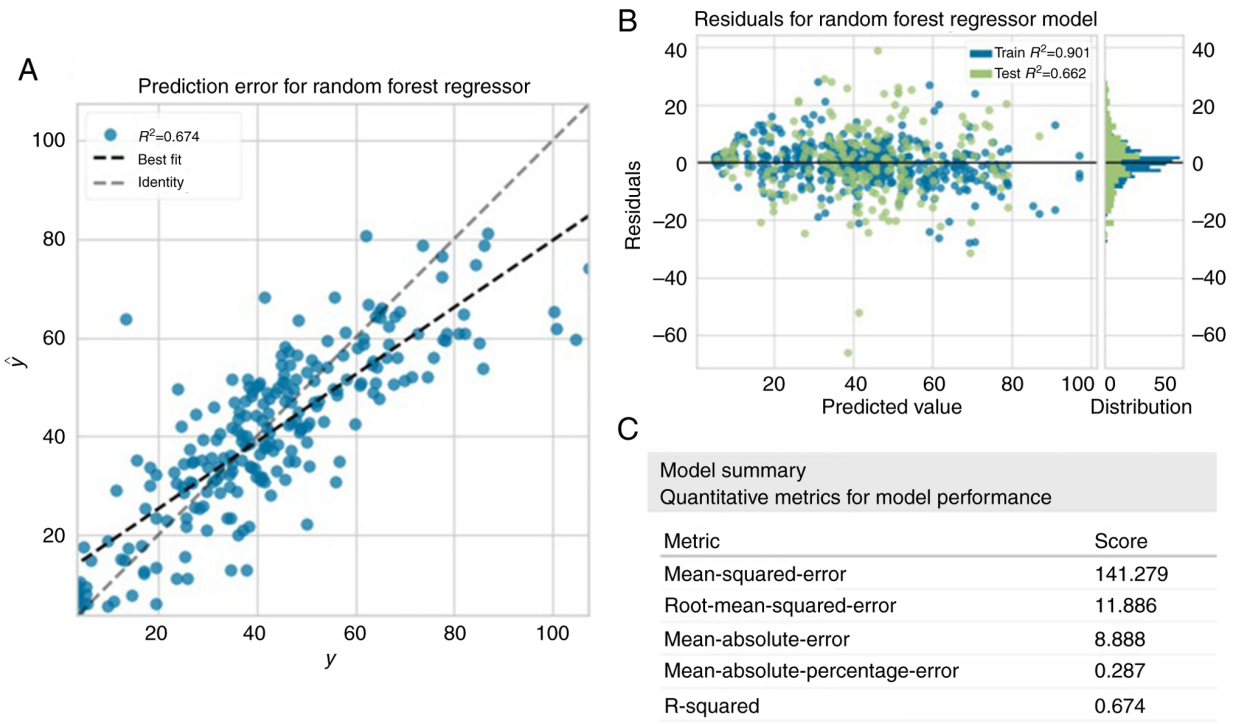
Model	MAE	MSE	RMSE	R^2	RMSLE	MAPE	TT (sec)
Random Forest Regressor	9.1430	172.1723	12.8377	0.6081	0.4496	0.4783	0.0800
Light Gradient Boosting Machine	9.3503	177.7188	13.0598	0.5971	0.4613	0.4893	0.0860
CatBoost Regressor	9.4182	184.8543	13.3495	0.5774	0.4592	0.4759	0.4640
Extreme Gradient Boosting	9.5244	190.4045	13.5610	0.5620	0.4574	0.4498	0.0440
Extra Trees Regressor	9.5420	195.3132	13.7607	0.5512	0.4751	0.4738	0.0680
Gradient Boosting Regressor	10.3896	202.9164	14.0047	0.5343	0.4959	0.5524	0.0510
Bayesian Ridge	11.2124	226.7562	14.9715	0.4680	0.5378	0.6426	0.0340
Elastic Net	11.4288	229.5513	15.0674	0.4612	0.5385	0.6514	0.0310
Huber Regressor	11.0380	230.6556	15.0620	0.4597	0.5358	0.6082	0.0340
Ridge Regression	11.1474	230.5441	15.0782	0.4596	0.5420	0.6415	0.0310
Linear Regression	11.1538	230.8561	15.0878	0.4589	0.5422	0.6419	0.0340
Least Angle Regression	11.1538	230.8561	15.0878	0.4589	0.5422	0.6419	0.0330
Lasso Regression	11.5126	231.7457	15.1408	0.4559	0.5398	0.6550	0.0290
Lasso Least Angle Regression	11.5126	231.7457	15.1408	0.4559	0.5398	0.6550	0.0300
Orthogonal Matching Pursuit	11.8809	242.4413	15.4791	0.4314	0.5439	0.6685	0.0310
AdaBoost Regressor	12.1183	247.7008	15.5625	0.4287	0.5662	0.7149	0.0440
Decision Tree Regressor	10.7265	244.8949	15.3545	0.4228	0.5399	0.4943	0.0310
K Neighbors Regressor	12.1525	263.8043	16.0848	0.3841	0.5367	0.5995	0.0340
Dummy Regressor	16.5575	437.9160	20.7940	-0.0103	0.7276	1.1068	0.0340
Passive Aggressive Regressor	16.3600	497.8276	20.7364	-0.0947	0.5926	0.7304	0.0340

MAE, mean absolute error; MSE, mean square error; RMSE, root mean squared error; RMSLE, root mean squared logarithmic error; MAPE, mean absolute percentage error; TT, training time.

help identify patients at the highest risk of becoming ineligible for postoperative cisplatin treatment.

In the present retrospective study, a machine learning model that focused on preoperative prediction accuracy was adopted and included 18 different feature values for predicting renal function insufficiency postoperatively. The findings indicated that the Random Forest algorithms outperformed other models in both regression and classification tasks. In the regression analysis, a slight difference in residual distribution was observed between the training and testing datasets, which may be attributed to limitations in potential features and sample size. Regarding classification, the prediction of postoperative CKD appears to strongly depend on the preoperative CKD status. Most patients with preoperative CKD remain in a CKD state post-surgery. However, for individuals with normal renal function in the healthy group, postoperative CKD cannot be reliably predicted based on preoperative renal function alone. Furthermore, SHAP was utilized to gain insights into the importance and impact of clinical features on predictions, making the machine learning model more interpretable. Integrating clinical risk factors into a prediction model enhances preoperative personalized risk assessment and helps determine the best timing for perioperative cisplatin-based chemotherapy.

Random Forest provides insights into feature importance, aiding in feature selection and model interpretability. It has robustness and high accuracy and can reduce overfitting. In addition to being computationally intensive, it may struggle with small datasets where individual trees lack sufficient



Metric	Score
Mean-squared-error	141.279
Root-mean-squared-error	11.886
Mean-absolute-error	8.888
Mean-absolute-percentage-error	0.287
R-squared	0.674

Figure 2. Performance of Random Forest Regressor for the testing dataset. (A) Prediction error, (B) residuals plot and (C) Model summary for the Random Forest Regressor.

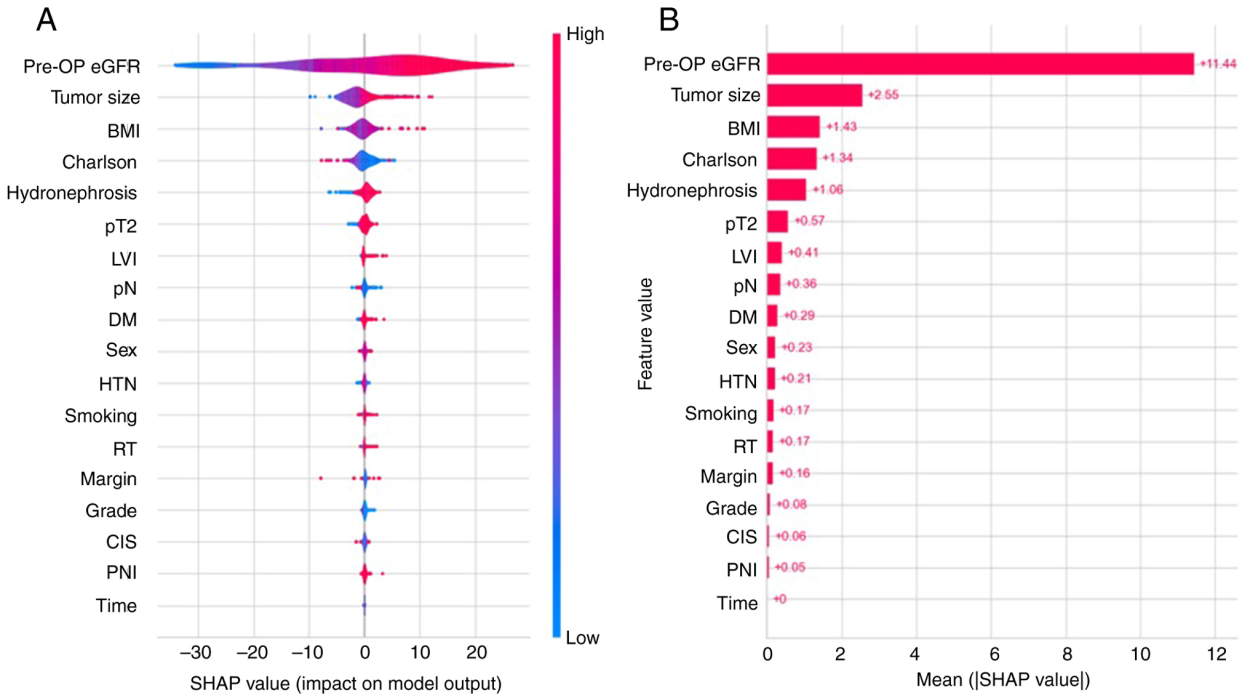


Figure 3. SHAP method for postoperative eGFR regression prediction model. (A) Violin plot and (B) bar chart of the SHAP values. SHAP, SHapley Additive exPlanations; pre-OP, preoperative; eGFR, estimated glomerular filtration rate; BMI, body mass index; Charlson, Charlson Comorbidity Index; pT2, pathological tumor stage ≥2; LVI, lymphovascular invasion; pN, pathological lymph node stage ≥Nx/N02; DM, diabetes mellitus; HTN, hypertension; RT, radiotherapy; CIS, carcinoma *in situ*; PNI, perineural invasion.

diversity. While it provides feature importance, the overall model is less interpretable compared with simpler algorithms like linear regression (16,17). In the present dataset, most patients had CKD before surgery, which may have contributed to the high accuracy of the model. The prediction of postoperative CKD depends largely on the preoperative CKD status. The prediction model was less effective in the preoperative CKD group in predicting the postoperative CKD status compared

Table III. Performance of 16 machine learning algorithms for predicting chronic kidney disease in the training dataset.

Model	Accuracy	AUC	Recall	Precision	F1	k	MCC	TT (sec)
Random Forest Classifier	0.8746	0.8882	0.9559	0.8966	0.9249	0.5434	0.5585	0.0650
CatBoost Classifier	0.8745	0.8744	0.9489	0.9022	0.9243	0.5525	0.5683	0.7260
Extreme Gradient Boosting	0.8707	0.8757	0.9373	0.9070	0.9214	0.5534	0.5616	0.1450
Gradient Boosting Classifier	0.8690	0.8710	0.9605	0.8871	0.9220	0.5098	0.5316	0.0490
Light Gradient Boosting Machine	0.8689	0.8740	0.9373	0.9051	0.9202	0.5455	0.5559	0.1040
Extra Trees Classifier	0.8558	0.8465	0.9187	0.9048	0.9112	0.5239	0.5291	0.0670
Decision Tree Classifier	0.8502	0.8007	0.8932	0.9198	0.9055	0.5390	0.5450	0.0300
K Neighbors Classifier	0.8390	0.7759	0.9490	0.8654	0.9049	0.3786	0.4056	0.3760
Linear Discriminant Analysis	0.8351	0.8327	0.9374	0.8692	0.9013	0.3964	0.4173	0.0300
Ridge Classifier	0.8277	0.0000	0.9722	0.8401	0.9011	0.2528	0.3107	0.0340
Ada Boost Classifier	0.8238	0.7899	0.9280	0.8643	0.8947	0.3546	0.3665	0.0630
Logistic Regression	0.8201	0.8342	0.9513	0.8456	0.8949	0.2770	0.3149	0.6260
Dummy Classifier	0.8072	0.5000	1.0000	0.8072	0.8933	0.0000	0.0000	0.0340
Naive Bayes	0.8032	0.7821	0.8978	0.8646	0.8803	0.3211	0.3264	0.0350
SVM-Linear Kernel	0.7920	0.0000	0.8980	0.8610	0.8722	0.2229	0.2339	0.0320
Quadratic Discriminant Analysis	0.6571	0.5337	0.7374	0.8424	0.7313	0.0355	0.0451	0.0340

AUC, area under the curve; F1, F1-score; k, Cohen's kappa coefficient; MCC, Matthews Correlation Coefficient; TT, training time.

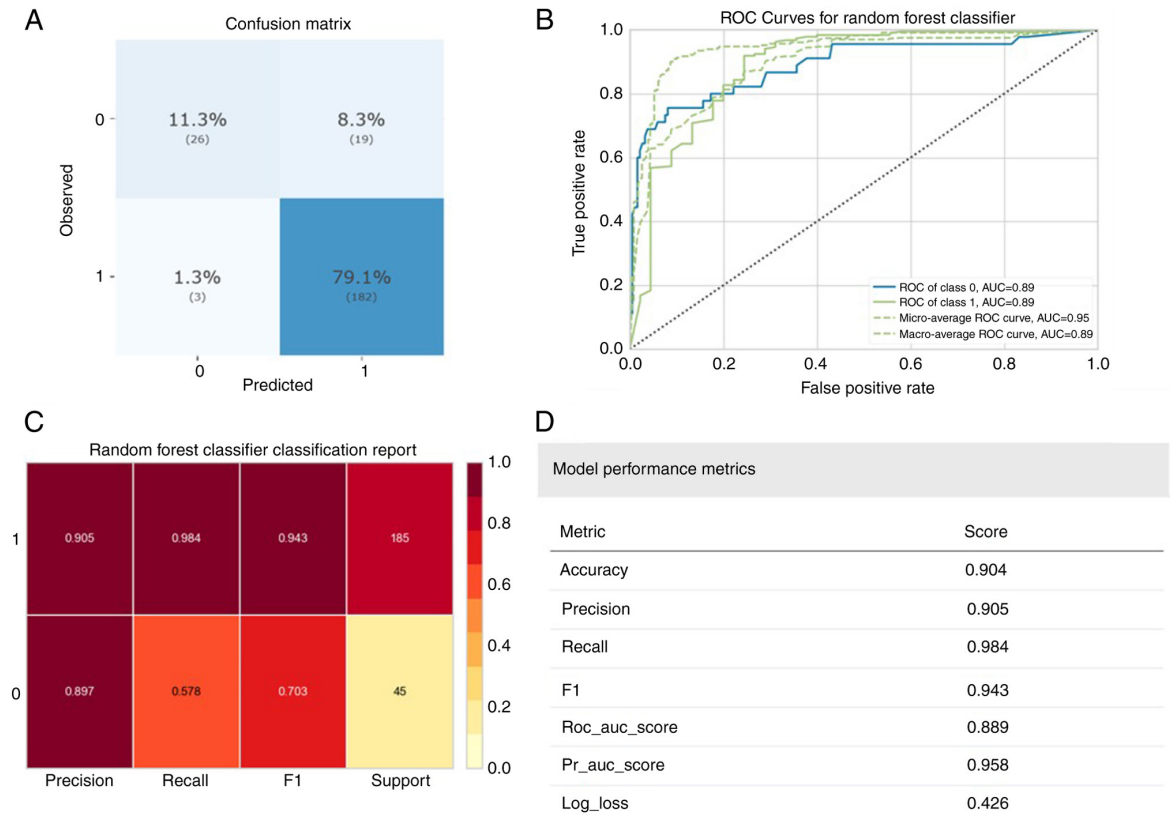


Figure 4. Performance of Random Forest classifier for the testing dataset. (A) Confusion for Random Forest classifier. (B) ROC curves for Random Forest classifier. (C) Classification report for Random Forest classifier. (D) Summary of model performance for random forest classifier. ROC, receiver operating characteristic; AUC, area under the curve.

with the healthy control group. Therefore, it is necessary to establish different prediction models based on the presence or absence of CKD before surgery.

In the present study, from the analysis of eGFR and CKD outcomes, preoperative eGFR, tumor size, BMI and CCI were the main predictors of postoperative renal function. It has

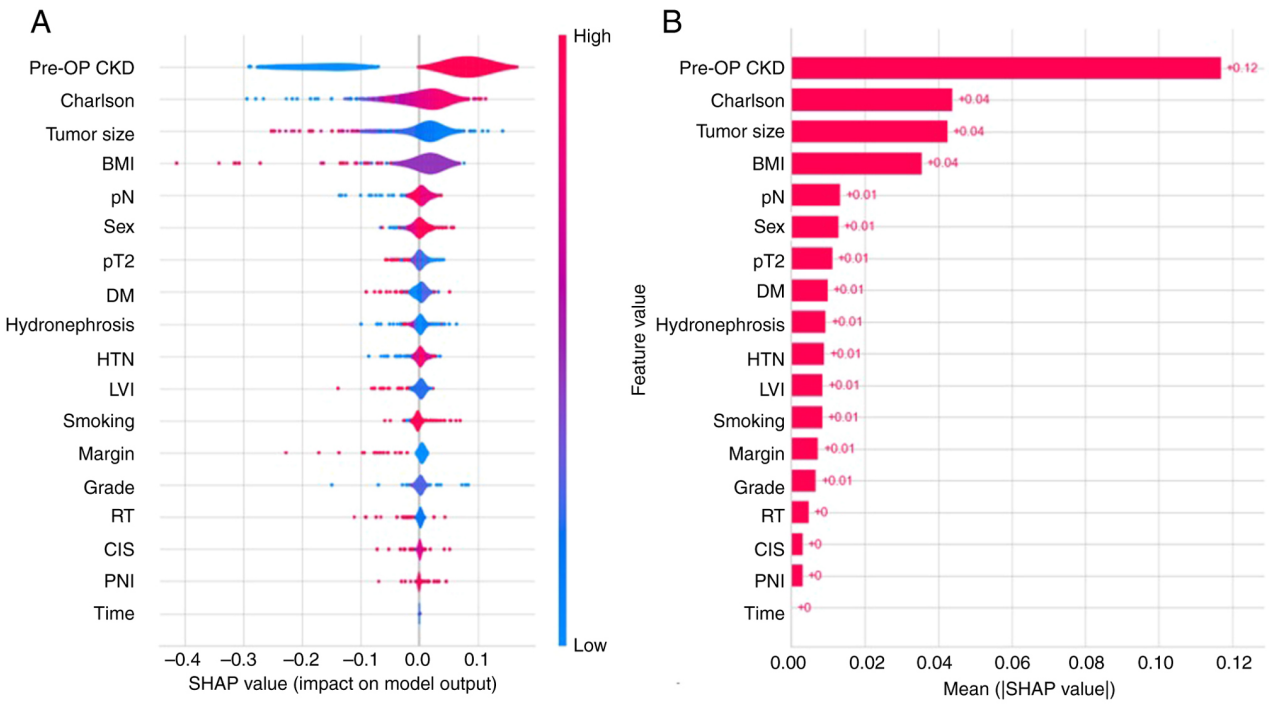


Figure 5. SHAP method for postoperative CKD status classification prediction model. (A) Violin plot and (B) bar chart for the SHAP values. SHAP, SHapley Additive exPlanations; pre-OP, preoperative; eGFR, estimated glomerular filtration rate; BMI, body mass index; Charlson, Charlson Comorbidity Index; pT2, pathological tumor stage ≥2; pN, pathological lymph node stage ≥Nx/N02; DM, diabetes mellitus; HTN, hypertension; LVI, lymphovascular invasion; RT, radiotherapy; CIS, carcinoma *in situ*; PNI, perineural invasion.

been previously shown that numerous clinical factors such as preoperative eGFR <60 ml/min, increased age, coronary artery disease and acute kidney injury on postoperative day 1 have been identified as being associated with poorer renal function outcomes after RNU, highlighting that preoperative renal function serves a crucial role in preserving postoperative function (18). Kaag *et al* (19) reported that there is a postoperative decline in renal function that does not improve over time observed between the early (1-5 months) and late (>5 months) postoperative time points. In addition, the authors discovered that older patients with lower preoperative eGFR tend to have a poorer renal function decline after surgery (19). Rodríguez Faba *et al* (20) also identified that some factors can predict postoperative impaired renal function including preoperative eGFR and hydronephrosis. Song *et al* (21) reported that preoperative renal function and contralateral kidney volume can predict new-onset CKD after surgery based on a multivariate logistic regression analysis, and that a small contralateral kidney volume is associated with a risk of postoperative CKD. The results of the present study indicated that BMI is an independent risk factor for renal function decline following RNU. Obesity is associated with a significantly increased risk for CKD progression. However, malnutrition is a prevalent issue among patients with cancer who often progress to cachexia, which can result in a diminished response to treatment, a poorer prognosis and a reduced quality of life. While the link between obesity and CKD progression is intricate, its adverse impact has been shown in several previous studies (21-23).

In the present study, hydronephrosis was demonstrated to be a predictive factor for postoperative renal function decline. Kanno *et al* (24) also discovered that the severity of hydronephrosis has a notable impact on postoperative renal function decline. The aforementioned study also reported that moderate to severe hydronephrosis, but not mild hydronephrosis, was associated with a slight decrease in renal function after surgery. The difference between preoperative and postoperative renal function is larger in patients without hydronephrosis or mild hydronephrosis compared with severe hydronephrosis (24). Similarly, a study by Singla *et al* (25) revealed that patients with hydronephrosis receiving surgery experience a smaller decrease in renal function compared with those without hydronephrosis. A possible explanation for the findings of the present study is that hydronephrosis was consistently associated with a thinner renal cortex and lower eGFR (25). The overall renal function would be compensated by the opposite kidney, so removing the impaired kidney would not lead to a marked drop in total eGFR. A smaller tumor size predicted impaired postoperative overall renal function as this would be compensated by the opposite kidney, which is in accordance with a previous study (26), suggesting that a larger tumor size would be associated with decreased function in the affected kidney; therefore, it is a protective factor for postoperative eGFR after accounting for preoperative renal function. Based on this concept, split renal function studies using nuclear renal scans may offer preoperative contralateral kidney function information, which could serve a key role in predicting postoperative renal function following surgery.

Compared with previous studies (27-29), the present model demonstrates greater flexibility, effectively accommodating both linear and non-linear outcome distributions. Furthermore, the performance of various models was evaluated to ensure the final selection achieved optimal accuracy. However, the present study does have some limitations. Firstly, although 18 features were included in the prediction model, there may be other important features such as tumor variants, cardiovascular

disease, impact of medication and patient age, that were not added, as these may not have been collected or the data were missing. Secondly, this is a retrospective design study from a single institution and the number of cases is limited. Therefore, the original data may limit the reliability and accuracy of the model. Additionally, some patients were excluded from the present study due to the lack of complete preoperative data. Nevertheless, the present research provides important predictive factors that are associated with postoperative renal function decline. The predictors identified in the present study may be used for treatment planning strategies especially perioperative systemic chemotherapy.

To conclude, preoperative eGFR, tumor size, BMI, CCI and hydronephrosis were found to be independent predictors of postoperative eGFR in patients with UTUC. The correlation between the predicted and observed postoperative eGFR and CKD demonstrated a strong relationship. While further research is required to validate these findings, the machine learning model based on these factors could be valuable for identifying patients with impaired renal function after RNU and those who may be suitable for NAC.

Acknowledgements

Not applicable.

Funding

The present study was partially supported by the Ministry of Science and Technology (grant nos. MOST-111-2314-B-037-100-MY2 and NSTC 113-2314-B-037-027), Kaohsiung Medical University Hospital (grant nos. KMUH-111-1R55 and KMUH-112-2R57) and Regenerative Medicine and Cell Therapy Research Center (grant no. KMU-TC112A02).

Availability of data and materials

The data generated in the present study are included in the figures and tables of this article.

Authors' contributions

Conception and design was performed by HYL, HCY, WML, WJW and HLK. Acquisition of data and analysis was performed by YCW and CCL. YCW and HYL analyzed and interpreted the data. HYL drafted the manuscript. HYL and YCW confirm the authenticity of all the raw data. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

The present study was approved by the institutional review board of Kaohsiung Medical University Hospital (Kaohsiung, Taiwan; approval no. KMUHIRB-E(I)-20180214). A waiver for consent to participate was granted by the IRB/Ethics Committee for this purpose.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Han J, Xian Z, Zhang Y, Liu J and Liang A: Systematic overview of aristolochic acids: Nephrotoxicity, carcinogenicity, and underlying mechanisms. Front Pharmacol 10: 648, 2019.
2. Chang YH, Hsu WL, Lee YK, Chiang CJ, Yang YW, You SL, Chen YC and Lai TS: Trends and sex-specific incidence of upper urinary tract cancer in Taiwan: A birth cohort study. Cancer Med 12: 15350-15357, 2023.
3. Yeh HC, Chang CH, Fang JK, Chen IHA, Lin JT, Hong JH, Huang CY, Wang SS, Chen CS, Lo CW, *et al*: The value of preoperative local symptoms in prognosis of upper tract urothelial carcinoma after radical nephroureterectomy: A retrospective, multicenter cohort study. Front Oncol 12: 872849, 2022.
4. Chen IA, Chang CH, Huang CP, Wu WJ, Li CC, Chen CH, Huang CY, Lo CW, Yu CC, Tsai CY, *et al*: Factors predicting oncological outcomes of radical nephroureterectomy for upper tract urothelial carcinoma in Taiwan. Front Oncol 11: 766576, 2022.
5. Margulis V, Shariat SF, Matin SF, Kamat AM, Zigeuner R, Kikuchi E, Lotan Y, Weizer A, Raman JD and Wood CG; Upper Tract Urothelial Carcinoma CollaborationThe Upper Tract Urothelial Carcinoma Collaboration: Outcomes of radical nephroureterectomy: A series from the upper tract urothelial carcinoma collaboration. Cancer 115: 1224-1233, 2009.
6. Chang LW, Hung SC, Chen CS, Li JR, Chiu KY, Wang SS, Yang CK, Lu K, Chen CC, Wang SC, *et al*: Geriatric nutritional risk index as a prognostic marker for patients with upper tract urothelial carcinoma receiving radical nephroureterectomy. Sci Rep 13: 4554, 2023.
7. Abouassaly R, Alibhai SMH, Shah N, Timilshina N, Fleshner N and Finelli A: Troubling outcomes from population-level analysis of surgery for upper tract urothelial carcinoma. Urology 76: 895-901, 2010.
8. Birtle A, Johnson M, Chester J, Jones R, Dolling D, Bryan RT, Harris C, Winterbottom A, Blacker A, Catto JWF, *et al*: Adjuvant chemotherapy in upper tract urothelial carcinoma (the POUT trial): A phase 3, open-label, randomised controlled trial. Lancet 395: 1268-1277, 2020.
9. Leow JJ, Chong YL, Chang SL, Valderrama BP, Powles T and Bellmunt J: Neoadjuvant and adjuvant chemotherapy for upper tract urothelial carcinoma: A 2020 systematic review and meta-analysis, and future perspectives on systemic therapy. Eur Urol 79: 635-654, 2021.
10. Petros FG, Qiao W, Singla N, Clinton TN, Robyak H, Raman JD, Margulis V and Matin SF: Preoperative multiplex nomogram for prediction of high-risk nonorgan-confined upper-tract urothelial carcinoma. Urol Oncol 37: 292.e1-292.e9, 2019.
11. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF III, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, *et al*: A new equation to estimate glomerular filtration rate. Ann Intern Med 150: 604-612, 2009.
12. Webster AC, Nagler EV, Morton RL and Masson P: Chronic kidney disease. Lancet 389: 1238-1252, 2017.
13. Kidney Disease: Improving Global Outcomes (KDIGO) CKD-MBD Work Group: KDIGO clinical practice guideline for the diagnosis, evaluation, prevention, and treatment of chronic kidney disease-mineral and bone disorder (CKD-MBD). Kidney Int Suppl 76 (Suppl 113): S1-S130, 2009.
14. Ray S: 'A quick review of machine learning algorithms', 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, pp35-39, 2019.
15. Favaretto RL, Shariat SF, Savage C, Godoy G, Chade DC, Kaag M, Bochner BH, Coleman J and Dalbagni G: Combining imaging and ureteroscopy variables in a preoperative multivariable model for prediction of muscle-invasive and non-organ confined disease in patients with upper tract urothelial carcinoma. BJU Int 109: 77-82, 2012.
16. Qi Y: Random forest for bioinformatics. In: Zhang C and Ma Y (eds) Ensemble Machine Learning. Springer, New York, NY, pp 307-323, 2012.

17. Smith PF, Ganesh S and Liu P: A comparison of random forest regression and multiple linear regression for prediction in neuroscience. J Neurosci Methods 220: 85-91, 2013.
18. Tafuri A, Marchioni M, Cerrato C, Mari A, Tellini R, Odorizzi K, Veccia A, Amparore D, Shakir A, Carbonara U, *et al*: Changes in renal function after nephroureterectomy for upper urinary tract carcinoma: Analysis of a large multicenter cohort (radical nephroureterectomy outcomes (RaNeO) research consortium). World J Urol 40: 2771-2779, 2022.
19. Kaag M, Trost L, Thompson RH, Favaretto R, Elliott V, Shariat SF, Maschino A, Vertosick E, Raman JD and Dalbagni G: Preoperative predictors of renal function decline after radical nephroureterectomy for upper tract urothelial carcinoma. BJU Int 114: 674-679, 2014.
20. Rodríguez Faba O, Palou J, Breda A, Maroto P, Fernández Gómez JM, Wong A and Villavicencio H: Predictive factors for impaired renal function following nephroureterectomy in upper urinary tract urothelial cell carcinoma. Urol Int 92: 169-173, 2014.
21. Song W, Sung HH, Han DH, Jeong BC, Seo SI, Jeon SS, Lee HM, Choi HY and Jeon HG: The effect of contralateral kidney volume on renal function after radical nephroureterectomy: Implications for eligibility for neoadjuvant chemotherapy for upper tract urothelial cancer. Urol Oncol 35: 114.e1-114.e7, 2017.
22. Yun HR, Kim H, Park JT, Chang TI, Yoo TH, Kang SW, Choi KH, Sung S, Kim SW, Lee J, *et al*: Obesity, metabolic abnormality, and progression of CKD. Am J Kidney Dis 72: 400-410, 2018.
23. Madero M, Katz R, Murphy R, Newman A, Patel K, Ix J, Peralta C, Satterfield S, Fried L, Shlipak M and Sarnak M: Comparison between different measures of body fat with kidney function decline and incident CKD. Clin J Am Soc Nephrol 12: 893-903, 2017.
24. Kanno T, Kobori G, Saito R, Ito K, Nakagawa H, Takahashi T, Koterazawa S, Takaoka N, Somiya S, Haitani T, *et al*: Hydronephrosis severity as a predictor of postoperative renal function decline following laparoscopic radical nephroureterectomy. Int J Clin Oncol 29: 464-472, 2024.
25. Singla N, Hutchinson R, Haddad A, Sagalowsky A, Lotan Y and Margulis V: Preoperative hydronephrosis is associated with less decline in renal function after radical nephroureterectomy for upper tract urothelial carcinoma. Can J Urol 23: 8334-8341, 2016.
26. Fang D, Zhang Q, Li X, Qian C, Xiong G, Zhang L, Chen X, Zhang X, Yu W, He Z and Zhou L: Nomogram predicting renal insufficiency after nephroureterectomy for upper tract urothelial carcinoma in the Chinese population: Exclusion of ineligible candidates for adjuvant chemotherapy. Biomed Res Int 2014: 529186, 2014.
27. Hashimoto T, Ohno Y, Nakashima J, Gondo T, Nakagami Y, Namiki K, Horiguchi Y, Yoshioka K, Ohori M and Tachibana M: Prediction of renal function after nephroureterectomy in patients with upper tract urothelial carcinoma. Jpn J Clin Oncol 45: 1064-1068, 2015.
28. Hensley PJ, Labbate C, Zganjar A, Howard J, Huelster H, Durdin T, Pham J, Xiao L, Pallauf M, Lombardo K, *et al*: Development and validation of a multivariable nomogram predictive of post-nephroureterectomy renal function. Eur Urol Oncol 7: 1313-1319, 2024.
29. Liu J, Wu P, Lai S, Wang J, Hou H and Zhang Y: Prognostic models for upper urinary tract urothelial carcinoma patients after radical nephroureterectomy based on a novel systemic immune-inflammation score with machine learning. BMC Cancer 23: 574, 2023.