

RESEARCH

Open Access



# Predictive models for chronic kidney disease after radical or partial nephrectomy in renal cell cancer using early postoperative serum creatinine levels

Dongwoo Chae<sup>1\*†</sup>, Na Young Kim<sup>2†</sup>, Ki Jun Kim<sup>2</sup>, Kyemyung Park<sup>1</sup>, Chaerim Oh<sup>2</sup> and So Yeon Kim<sup>2\*</sup> 

## Abstract

**Background:** Several predictive factors for chronic kidney disease (CKD) following radical nephrectomy (RN) or partial nephrectomy (PN) have been identified. However, early postoperative laboratory values were infrequently considered as potential predictors. Therefore, this study aimed to develop predictive models for CKD 1 year after RN or PN using early postoperative laboratory values, including serum creatinine (SCr) levels, in addition to preoperative and intraoperative factors. Moreover, the optimal SCr sampling time point for the best prediction of CKD was determined.

**Methods:** Data were retrospectively collected from patients with renal cell cancer who underwent laparoscopic or robotic RN (n = 557) or PN (n = 999). Preoperative, intraoperative, and postoperative factors, including laboratory values, were incorporated during model development. We developed 8 final models using information collected at different time points (preoperative, postoperative day [POD] 0 to 5, and postoperative 1 month). Lastly, we combined all possible subsets of the developed models to generate 120 meta-models. Furthermore, we built a web application to facilitate the implementation of the model.

**Results:** The magnitude of postoperative elevation of SCr and history of CKD were the most important predictors for CKD at 1 year, followed by RN (compared to PN) and older age. Among the final models, the model using features of POD 4 showed the best performance for correctly predicting the stages of CKD at 1 year compared to other models (accuracy: 79% of POD 4 model versus 75% of POD 0 model, 76% of POD 1 model, 77% of POD 2 model, 78% of POD 3 model, 76% of POD 5 model, and 73% in postoperative 1 month model). Therefore, POD 4 may be the optimal sampling time point for postoperative SCr. A web application is hosted at [https://dongy.shinyapps.io/aki\\_ckd](https://dongy.shinyapps.io/aki_ckd).

**Conclusions:** Our predictive model, which incorporated postoperative laboratory values, especially SCr levels, in addition to preoperative and intraoperative factors, effectively predicted the occurrence of CKD 1 year after RN or PN and may be helpful for comprehensive management planning.

**Keywords:** Chronic kidney disease, Creatinine, Nephrectomy, Predictive factors, Renal cell cancer

## Background

Acute kidney injury (AKI) is a common complication after nephrectomy in renal cell cancer (RCC), with radical nephrectomy (RN) associated with a noticeably higher risk than partial nephrectomy (PN) [1, 2]. AKI is defined as an abrupt decrease in kidney function occurring over

\*Correspondence: DONGY@yuhs.ac; KIMSY326@yuhs.ac

<sup>†</sup>Dongwoo Chae and Na Young Kim contributed equally to this work

<sup>1</sup> Department of Pharmacology, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea

<sup>2</sup> Department of Anesthesiology and Pain Medicine, Anesthesia and Pain Research Institute, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

7 days or less, whereas chronic kidney disease (CKD) is defined by the persistence of kidney disease for a period of >90 days [3]. The severity of AKI and recovery time have been implicated as important predictors of CKD progression [3–6]. Surgically induced CKD may be associated with a lower risk of progression and mortality than CKD due to medical causes [7]. However, despite this, an estimated glomerular filtration rate (eGFR) less than 45 mL/min/1.73 m<sup>2</sup> in patients with surgically induced CKD has been associated with an increased risk of mortality [8].

Numerous studies have investigated the predictive factors for CKD following RN or PN [9–16], some of which developed predictive models for CKD [13, 14]. However, these studies analyzed preoperative and intraoperative factors (such as patient characteristics, preoperative laboratory values, and surgical type or technique) as possible predictors, without including postoperative laboratory values [9–14]. Although few studies included the occurrence of AKI or time to nadir eGFR as one of the predictors of CKD [15–17], serial changes of serum creatinine (SCr) was not considered. When considering the variable trajectories following AKI [18], postoperative laboratory values, especially SCr, should be considered for better prediction. Therefore, we hypothesized that SCr levels collected in the first 5 days after nephrectomy would provide important information to predict the SCr levels 1 year after surgery and ultimately, the occurrence of CKD. Thus, this study aimed to develop predictive models for CKD after RN or PN using early postoperative laboratory values, including SCr levels, in addition to preoperative and intraoperative factors, and build a web application to facilitate their implementation. Moreover, we aimed to find optimal SCr sampling time points for accurate CKD prediction.

## Methods

### Patients

The analysis data set included 1,556 patients with RCC who received either laparoscopic or robotic RN (n = 557) or PN (n = 999) between December 2005 and May 2019 and were at least followed up to 1 year after surgery. Patients lost to follow up or died before 1 year were excluded. Data were retrospectively collected from the electronic medical records of a single institution.

### Features used for prediction

Our study aimed to predict the rise in SCr levels relative to preoperative value at 1 year after surgery (between 11 and 13 months after surgery), hereafter denoted as  $\Delta SCr_{1y}$ , given the following features:

- (i) Patient characteristics: age, sex, weight, history of diabetes mellitus, history of hypertension, history of CKD, and tumor size.
- (ii) Preoperative and intraoperative factors: preoperative complete blood count (CBC; hematocrit, neutrophil, lymphocyte, monocyte, and platelet counts), preoperative routine chemistry (SCr, blood urea nitrogen [BUN], uric acid, total protein, albumin, aspartate aminotransferase, alanine aminotransferase [ALT], alkaline phosphatase, total bilirubin, and cholesterol), preoperative serum electrolytes (sodium, potassium, calcium, phosphate), preoperative vital signs (systolic and diastolic blood pressure [BP], and heart rate), size of mass removed, duration of anesthesia, and bleeding amount.
- (iii) Postoperative factors: baseline (preoperative) subtracted SCr, CBC, routine chemistry, serum electrolytes, and vital signs immediately after surgery, postoperative day (POD) 1–5, and postoperative 1 month

Data were also collected on POD 7 and 14 and at 3 and 6 months; however, they were not included in the model development. For brevity, a common notation involving  $\Delta$  has been used throughout to represent the baseline-subtracted level of different variables, with the associated subscript indicating the sampling time point. For example, baseline subtracted SCr on POD 3 has been denoted as  $\Delta SCr_{3d}$ . The letters *d*, *m*, and *y*, used as subscripts, represent the day, month, and year, respectively.

Missing values were imputed using multiple imputation by chained equations (MICE), also known as fully conditional specification or sequential regression multiple imputation. The method operates under the assumption that missing data are Missing at Random (MAR), i.e., the probability of a particular value being missing depends only on the observed values and not the unobserved values [19]. Since missing values in our data showed a clear time-dependency and were likely unrelated to the true value of SCr, we assumed that the condition of MAR was fulfilled, thus allowing the partial deduction of the missing values based on the measurements immediately before and after them. The highest proportions of missing values occurred on POD 4 and 5 in both RN and PN (Additional file 1: Fig. S1). The proportions of patients without any missing value were 14.7% and 16.8% in RN and PN, respectively. The widely validated R package, *mice*, was used to carry out the imputation process [20].

### Model development

First, the features were grouped into eight categories, namely  $F_{pre}$ ,  $F_{0d}$ ,  $F_{1d}$ ,  $F_{2d}$ ,  $F_{3d}$ ,  $F_{4d}$ ,  $F_{5d}$ , and  $F_{1m}$ ,

based on their time of acquisition.  $F_{pre}$  only included factors available prior to completion of the surgery, mentioned above as patient characteristics and preoperative and intraoperative factors. Other feature sets, denoted as  $F_i$  with  $i=0, 1, 2, 3, 4, 5$  days, included  $F_{pre}$  and information collected on the  $i^{th}$  postoperative day. The last feature set,  $F_{1m}$ , included  $F_{pre}$  and features collected 1 month after surgery. Thereafter, Lasso regression models were built on each of the feature sets to predict  $\Delta SCr_{1y}$ , hereafter referred to individually as  $Model_{Lasso,pre}$ ,  $Model_{Lasso,0d}$ ,  $Model_{Lasso,1d}$ ,  $Model_{Lasso,2d}$ ,  $Model_{Lasso,3d}$ ,  $Model_{Lasso,4d}$ ,  $Model_{Lasso,5d}$ , and  $Model_{Lasso,1m}$ , and collectively, as  $Model_{Lasso}$ . The features with non-zero regression coefficients in  $Model_{Lasso}$  were collectively referred to as  $F_{Lasso}$  and individually as  $F_{Lasso,pre}$ ,  $F_{Lasso,0d}$ ,  $F_{Lasso,1d}$ ,  $F_{Lasso,2d}$ ,  $F_{Lasso,3d}$ ,  $F_{Lasso,4d}$ ,  $F_{Lasso,5d}$ , and  $F_{Lasso,1m}$ ; each of these consisted of selected features from  $F_{pre}$ ,  $F_{0d}$ ,  $F_{1d}$ ,  $F_{2d}$ ,  $F_{3d}$ ,  $F_{4d}$ ,  $F_{5d}$ , and  $F_{1m}$ , respectively.

Prior to model development, we split the dataset into training and test datasets in the ratio of 8:2. The features were z-score normalized. A grid search algorithm and four-fold cross-validation were used on the training dataset to tune the shrinkage hyper-parameters.

### Construction of the final model

To retain only the most parsimonious set of features, we calculated the Spearman's partial correlation coefficients [21] of  $F_{Lasso}$  with  $\Delta SCr_{1y}$  and eliminated features with absolute values less than 0.1. This yielded the final selected features, collectively referred to as  $F_{final}$  and individually as  $F_{final,pre}$ ,  $F_{final,0d}$ ,  $F_{final,1d}$ ,  $F_{final,2d}$ ,  $F_{final,3d}$ ,  $F_{final,4d}$ ,  $F_{final,5d}$ , and  $F_{final,1m}$ . We then constructed multivariate linear regression models using  $F_{final}$ . Unlike the Lasso models, the features were not z-score normalized, so the estimated regression coefficients could be readily interpreted. The final models were referred to as  $Model_{final,pre}$ ,  $Model_{final,0d}$ ,  $Model_{final,1d}$ ,  $Model_{final,2d}$ ,  $Model_{final,3d}$ ,  $Model_{final,4d}$ ,  $Model_{final,5d}$ , and  $Model_{final,1m}$ .

The predictive performances of  $Model_{final}$  were then compared with those of  $Model_{Lasso}$ .  $R^2$  and mean squared error (MSE) between the predicted and observed  $\Delta SCr_{1y}$  values calculated using the test dataset were used as the performance metrics. The eGFR was calculated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation [22]. CKD was categorized according to the eGFR: stage 1 ( $\geq 90$  mL/min/1.73 m<sup>2</sup>), stage 2 (60–89 mL/min/1.73 m<sup>2</sup>), and stage 3 and higher ( $< 60$  mL/min/1.73 m<sup>2</sup>) [23]. The overall analysis workflow is schematically shown in Fig. 1.

### Model stacking

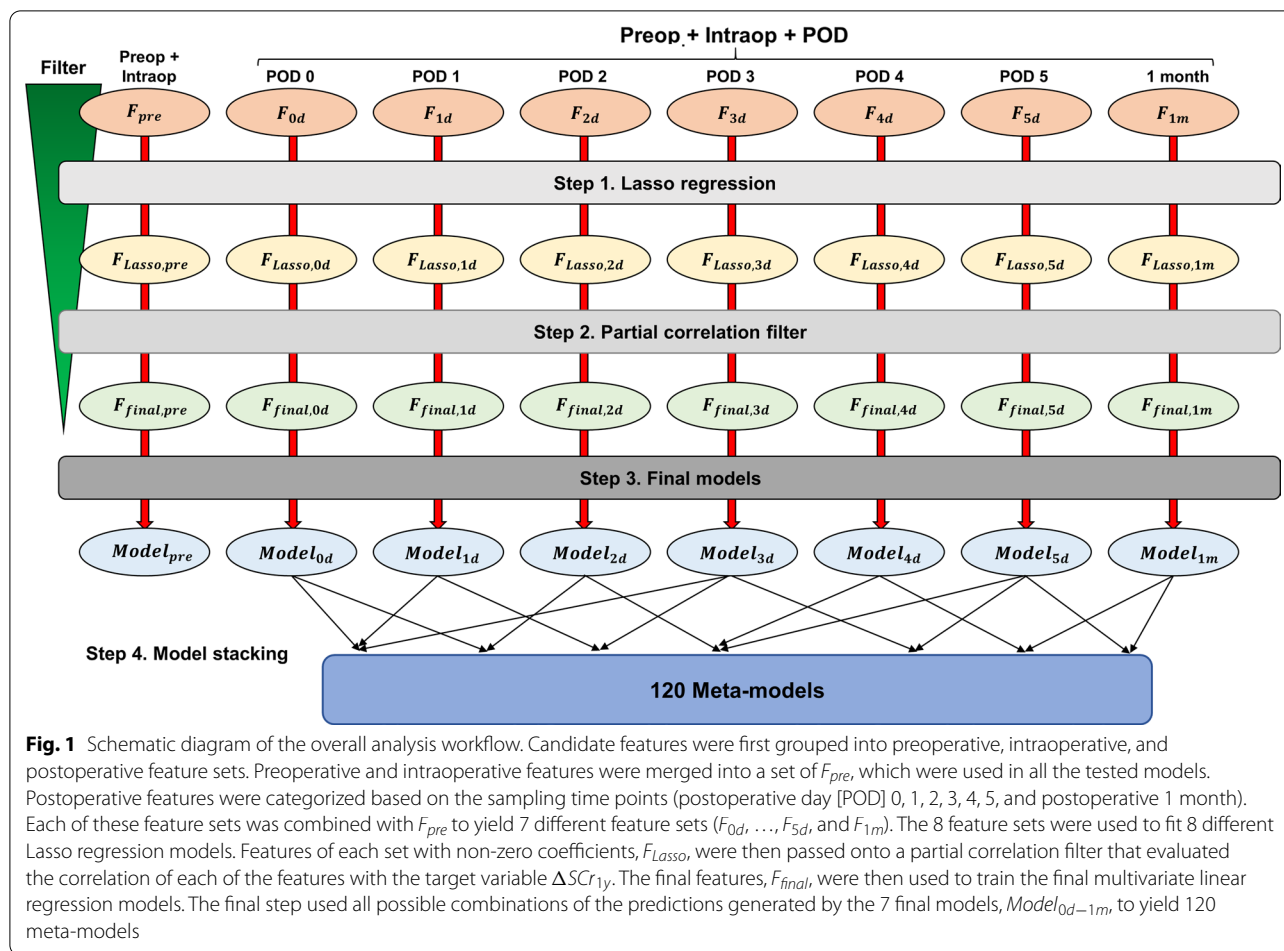
We performed model stacking by developing meta-models using predictions generated from all possible subsets of  $Model_{final,0d}$ ,  $Model_{final,1d}$ ,  $Model_{final,2d}$ ,  $Model_{final,3d}$ ,  $Model_{final,4d}$ ,  $Model_{final,5d}$ , and  $Model_{final,1m}$  as features. Ridge regressions with four-fold cross-validation were used to acquire appropriate weights to be assigned to each of the predictions generated by the component models. A total of 120 ( $=2^7 - 1 - 7$ ) meta-models were thus developed. Meta-models trained using predictions of  $k$  different final models ( $k=2, \dots, 7$ ) are hereafter be referred to as  $Meta\_models_k$ . For example, a meta-model developed using predictions of  $Model_{final,0d}$ ,  $Model_{final,3d}$ , and  $Model_{final,4d}$  constituted one of  $Meta\_models_3$ , and was used when supplied with SCr and laboratory values measured on POD 0, 3, and 4.

### Web application development

To facilitate the automatic selection and implementation of the meta-model, we developed a web application with a user-friendly interface, hosted at [https://dongy.shinyapps.io/aki\\_ckd](https://dongy.shinyapps.io/aki_ckd). The *shiny* package in R (<https://shiny.rstudio.com>) was used for programming the application. The application used 8 basic models ( $Model_{final,pre}$ ,  $Model_{final,0d}$ ,  $Model_{final,1d}$ ,  $Model_{final,2d}$ ,  $Model_{final,3d}$ ,  $Model_{final,4d}$ ,  $Model_{final,5d}$ , and  $Model_{final,1m}$ ) and 120 meta-models built from their predictions. Given only preoperative and intraoperative factors,  $Model_{final,pre}$  is activated to generate the predictions. Following the input of extra postoperative factors, the optimal model is chosen based on the number of samples ( $=k$ ) and the corresponding PODs of their acquisition. Outputs of the model are predicted values of  $\Delta SCr_{1y}$ , eGFR, and CKD stage at 1 year.

### Results

Table 1 summarizes the baseline characteristics of patients that underwent RN and PN. The longitudinal trajectories of postoperative SCr in RN and PN are shown in Additional file 2: Fig. S2. SCr typically increased from POD 0 to POD 3, decreased from POD 4 to 7, showed a secondary surge until POD 15, and gradually declined towards the final level. SCr levels on POD 4 and 5 showed the highest correlation with SCr level at 1 year (Additional file 3: Table S1). In RN and PN, 43.1% (240 of 557) and 7.2% (72 of 999) of the patients, respectively, developed CKD stage 3 and higher 1 year after surgery. All the above exploratory analyses were carried out using the raw data prior to imputation.



### Model development

Lasso regression models— $Model_{Lasso,pre}$ ,  $Model_{Lasso,0d}$ ,  $Model_{Lasso,1d}$ ,  $Model_{Lasso,2d}$ ,  $Model_{Lasso,3d}$ ,  $Model_{Lasso,4d}$ ,  $Model_{Lasso,5d}$ , and  $Model_{Lasso,1m}$ , were developed using 8 different feature sets— $F_{pre}$ ,  $F_{0d}$ ,  $F_{1d}$ ,  $F_{2d}$ ,  $F_{3d}$ ,  $F_{4d}$ ,  $F_{5d}$ , and  $F_{1m}$ , respectively. We then identified subsets of the original feature sets that were associated with non-zero regression coefficients (i.e.,  $F_{Lasso}$ ). The features included in each subset of  $F_{Lasso}$  and their estimated regression coefficients are shown in Additional file 3: Table S2.

The  $R^2$  statistics of  $Model_{Lasso,pre}$ ,  $Model_{Lasso,0d}$ ,  $Model_{Lasso,1d}$ ,  $Model_{Lasso,2d}$ ,  $Model_{Lasso,3d}$ ,  $Model_{Lasso,4d}$ ,  $Model_{Lasso,5d}$ , and  $Model_{Lasso,1m}$  were 0.432, 0.481, 0.507, 0.538, 0.562, 0.603, 0.589, and 0.630; their MSEs were 0.025, 0.023, 0.022, 0.021, 0.02, 0.018, 0.018, and 0.017, respectively. The best performing model was, as expected, the model using factors collected at 1 month ( $Model_{Lasso,1m}$ ). Among the models utilizing single time point information between POD 0 and 5,  $Model_{Lasso,4d}$  showed the best predictive performance for  $\Delta SCR_{1y}$ . Overall, the models utilizing postoperative factors

showed superior predictive performance than those using only preoperative and intraoperative factors (i.e.,  $Model_{Lasso,pre}$ ).

For all features included in each  $F_{Lasso}$  subset, partial Spearman’s correlation coefficients with  $\Delta SCR_{1y}$  were calculated; only features whose absolute values of the coefficients were greater than 0.1 (i.e.,  $F_{final}$ ) were retained. Final regression models (i.e.,  $Model_{final}$ ) were developed on  $F_{final}$  and their estimation results are shown in Table 2.

The  $R^2$  statistics of  $Model_{final,pre}$ ,  $Model_{final,0d}$ ,  $Model_{final,1d}$ ,  $Model_{final,2d}$ ,  $Model_{final,3d}$ ,  $Model_{final,4d}$ ,  $Model_{final,5d}$ , and  $Model_{final,1m}$  were 0.444, 0.530, 0.505, 0.550, 0.575, 0.608, 0.565, and 0.643 and their MSEs were 0.025, 0.021, 0.022, 0.020, 0.019, 0.018, 0.019, and 0.016, respectively. Comparison of the predictive performances of  $Model_{final}$  with  $Model_{Lasso}$  suggested that  $Model_{final}$  generally performed better than  $Model_{Lasso}$ , despite a fewer number of predictive features being included. Similar to  $Model_{Lasso}$ , the best prediction was achieved using  $F_{final,1m}$  followed by  $F_{final,4d}$ . The goodness-of-fit plots of observed vs. predicted values are shown in Fig. 2. The classification performances in terms of predicting the



**Table 1** Patient characteristics

Variable	Radical nephrectomy (n = 557)	Partial nephrectomy (n = 999)
Age (years)	56 (12)	54 (12)
Male sex	364 (67%)	668 (67%)
ASA physical status		
I	116 (21%)	299 (30%)
II	318 (57%)	588 (59%)
III	123 (22%)	112 (11%)
Medical history		
Hypertension	256 (46%)	381 (38%)
Diabetes mellitus	98 (18%)	156 (16%)
Chronic renal disease	6 (< 1%)	8 (< 1%)
Tumor size (cm)	5.2 (2.3)	2.6 (1.4)
Size of mass removed (cm)	13.4 (3.3)	4.4 (2.2)
R.E.N.A.L. nephrectomy score	–	6.0 (1.8)
Warm ischemia time (min)	–	23.4 (12.6)
Duration of anesthesia (min)	207 (58)	224 (53)
Bleeding amount (mL)	171 (380)	232 (279)
Preoperative laboratory values		
Hematocrit (%)	41.4 (4.6)	42.5 (4.3)
Neutrophil count ( $10^3/\mu\text{L}$ )	4.31 (1.74)	3.77 (1.51)
Lymphocyte count ( $10^3/\mu\text{L}$ )	1.99 (0.63)	1.99 (0.61)
Monocyte count ( $10^3/\mu\text{L}$ )	0.44 (0.17)	0.42 (0.15)
Platelet count ( $10^3/\mu\text{L}$ )	264 (85)	252 (67)
Creatinine (mg/dL) <sup>a</sup>	0.88 (0.25)	0.85 (0.21)
Blood urea nitrogen (mg/dL) <sup>a</sup>	14.4 (4.4)	14.2 (3.9)
Uric acid (mg/dL) <sup>a</sup>	5.1 (1.4)	5.2 (1.4)
Total protein (g/dL) <sup>a</sup>	7.3 (0.5)	7.3 (0.5)
Albumin (g/dL) <sup>a</sup>	4.5 (0.4)	4.6 (0.4)
Aspartate aminotransferase (IU/L) <sup>a</sup>	22 (11)	23 (11)
Alanine aminotransferase (IU/L) <sup>a</sup>	23 (16)	25 (16)
Alkaline phosphatase (IU/L) <sup>a</sup>	68 (26)	65 (19)
Total bilirubin (mg/dL) <sup>a</sup>	0.7 (0.3)	0.7 (0.3)
Cholesterol (mg/dL) <sup>a</sup>	179 (42)	186 (36)
Sodium (mEq/L) <sup>a</sup>	141 (2)	141 (2)
Potassium (mEq/L) <sup>a</sup>	4.4 (0.4)	4.4 (0.4)
Calcium (mEq/L) <sup>a</sup>	9.4 (0.5)	9.4 (0.4)
Phosphate (mEq/L) <sup>a</sup>	3.6 (0.5)	3.5 (0.5)
Preoperative vital signs		
Systolic blood pressure (mmHg)	132 (14)	132 (16)
Diastolic blood pressure (mmHg)	81 (10)	83 (11)
Heart rate (beats/min)	74 (11)	74 (11)

Values represent the mean (SD) or number of patients (proportion). ASA American Society of Anesthesiologists, PN partial nephrectomy, R.E.N.A.L. radius, exophytic or endophytic, nearness to collecting system or sinus, anterior or posterior location, and location relative to polar lines

<sup>a</sup> Data were obtained from serum

CKD stage are shown in Table 3.  $Model_{final,Ad}$  was found to confer the best accuracy, weighted averaged precision, and weighted averaged recall.

### Model stacking

Ridge regressions with zero intercept were performed using predictions of all possible subsets of  $Model_{final}$ , yielding 120 meta-models. The average performances of  $Meta\_model(s)_k$  improved with increasing k. For k=2, 3, 4, 5, 6, and 7, the mean  $R^2$  statistics were 0.606, 0.624, 0.637, 0.649, 0.659, and 0.669, respectively.  $Meta\_model_7$ , which utilizes all 7 models (i.e.  $Model_{final,0d}$ ,  $Model_{final,1d}$ ,  $Model_{final,2d}$ ,  $Model_{final,3d}$ ,  $Model_{final,4d}$ ,  $Model_{final,5d}$ , and  $Model_{final,1m}$ ) offered the best predictive performance. In particular, it outperformed  $Model_{final,1m}$  ( $R^2=0.630$ ).  $Meta\_model_7$  showed precisions of 90%, 74%, and 79% and recalls of 78%, 83%, and 76% for classifying CKD stages 1, 2, and 3, respectively (average accuracy, 80%; weighted average precision, 81%; weighted average recall, 80%).

### Web application development

Figure 3 shows the screenshot of the developed web application. The minimum information required to run the application are patient characteristics, such as age, sex, history of CKD, and preoperative and intraoperative factors, including the type of nephrectomy (RN or PN), size of mass removed, and preoperative SCr and BUN levels. The left panel is used for generating predictions. As postoperative measurements of SCr become available, they can be used to update the predictions.

### Discussion

We developed predictive models for CKD 1 year after RN or PN that fully incorporated preoperative, intraoperative, and postoperative factors. Our work can be summarized as follows: 1) We clearly demonstrated the need to incorporate early postoperative information (specifically SCr levels) to accurately predict long-term renal function. 2) Within the first postoperative week we identified POD 4 as the optimal sampling point for SCr (and BUN). 3) We identified the magnitude of early SCr elevation, history of CKD, surgery type (RN or PN), and patient age as the most robust predictors of CKD. 4) We provide a practical framework to predict CKD and offer an easy-to-use web application to implement our models.

Although surgically induced CKD may have a better prognosis than CKD due to medical causes [7], the risk of mortality is known to increase if patients have a reduced eGFR ( $<45$  mL/min/1.73 m<sup>2</sup>) following RCC surgery [8]. Our results showed that 14% (78 out of 557) of RN patients and 2.2% (22 out of 999) of PN patients had an eGFR  $<45$  mL/min/1.73 m<sup>2</sup> 1 year after surgery.

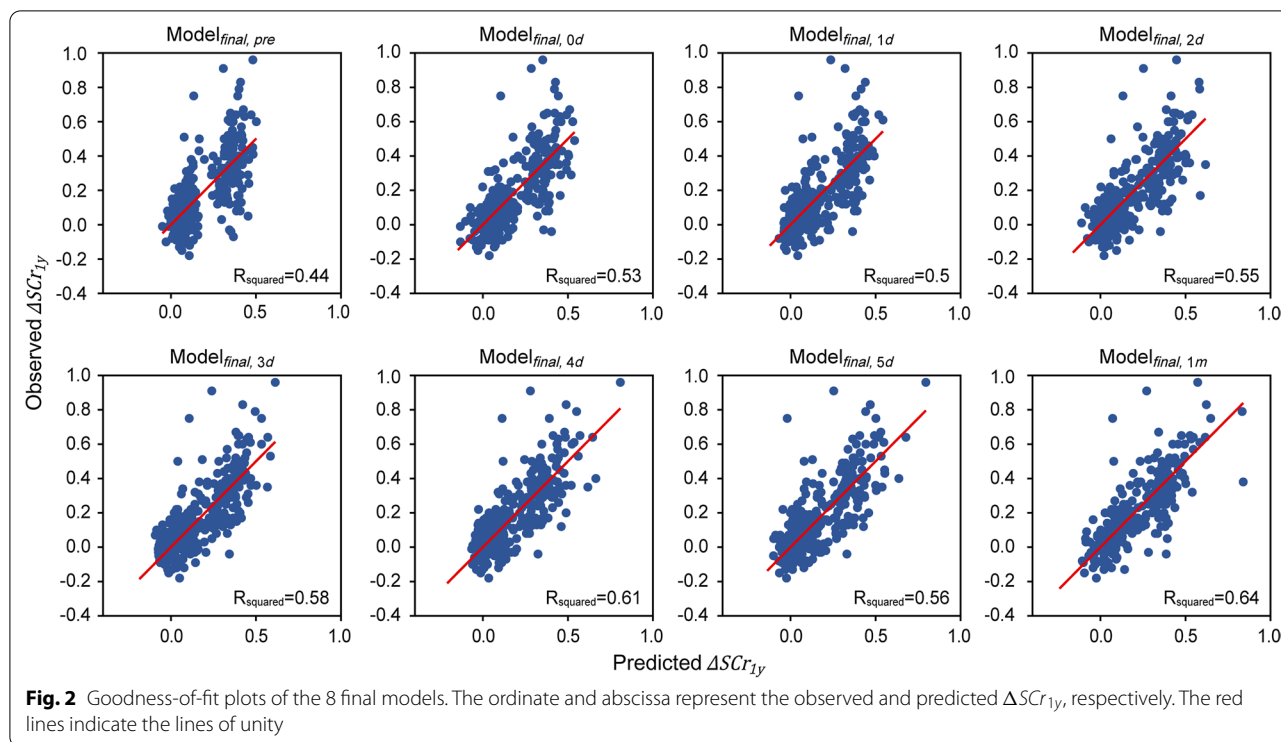
**Table 2** Regression coefficients of the final models sorted in the order of highest statistical significance

Model	Predictive performance	Feature	Estimates	t values	p values
<i>Model<sub>final,pre</sub></i>	$R^2=0.444$ MSE = 0.025	History of CKD	0.67	13.051	< 0.001
		Radical nephrectomy	0.188	8.684	< 0.001
		Male	0.06	5.243	< 0.001
		Size of mass removed (cm)	0.01	4.827	< 0.001
		Age (years)	0.0021	4.808	< 0.001
<i>Model<sub>final,0d</sub></i>	$R^2=0.530$ MSE = 0.021	History of CKD	0.722	14.938	< 0.001
		$\Delta SCr_{0d}$ (mg/dL)	0.540	13.716	< 0.001
		Radical nephrectomy	0.2	9.754	< 0.001
		Age (years)	0.0033	7.771	< 0.001
<i>Model<sub>final,1d</sub></i>	$R^2=0.505$ MSE = 0.022	Size of mass removed (cm)	0.0055	2.769	0.006
		$\Delta SCr_{1d}$ (mg/dL)	0.391	15.443	< 0.001
		History of CKD	0.632	13.193	< 0.001
<i>Model<sub>final,2d</sub></i>	$R^2=0.550$ MSE = 0.02	Radical nephrectomy	0.168	8.343	< 0.001
		Size of mass removed (cm)	0.0029	1.456	0.146
		$\Delta SCr_{2d}$ (mg/dL)	0.414	20.76	< 0.001
		History of CKD	0.471	10.157	< 0.001
<i>Model<sub>final,3d</sub></i>	$R^2=0.575$ MSE = 0.019	Age (years)	0.0022	5.537	< 0.001
		Radical nephrectomy	0.120	6.421	< 0.001
		History of CKD	0.44	9.716	< 0.001
<i>Model<sub>final,4d</sub></i>	$R^2=0.608$ MSE = 0.018	Size of mass removed (cm)	0.0007	0.361	0.718
		$\Delta SCr_{4d}$ (mg/dL)	0.449	17.84	< 0.001
		Radical nephrectomy	0.117	10.031	< 0.001
		History of CKD	0.441	9.735	< 0.001
		$\Delta BUN_{4d}$ (mg/dL)	0.0083	7.062	< 0.001
<i>Model<sub>final,5d</sub></i>	$R^2=0.565$ MSE = 0.019	Preoperative BUN (mg/dL)	0.0070	4.913	< 0.001
		Age (years)	0.0015	3.732	< 0.001
		Radical nephrectomy	0.122	9.918	< 0.001
		History of CKD	0.573	12.658	< 0.001
<i>Model<sub>final,1m</sub></i>	$R^2=0.643$ MSE = 0.016	$\Delta SCr_{5d}$ (mg/dL)	0.472	18.7	< 0.001
		$\Delta BUN_{5d}$ (mg/dL)	0.0059	5.954	< 0.001
		Radical nephrectomy	0.079	6.565	< 0.001
		History of CKD	0.526	12.324	< 0.001
		$\Delta SCr_{1m}$ (mg/dL)	0.638	26.397	< 0.001

CKD chronic kidney disease, SCr serum creatinine, BUN blood urea nitrogen, d day, m month, MSE mean squared error

This percentage was enough to warrant attention. Studies investigating the risk factors for CKD after RN or PN have shown older age, male sex, history of CKD, diabetes mellitus, and RN as independent predictors [9–15]. However, most of these studies did not consider perioperative laboratory values. Another study showed that time to nadir eGFR was one of the predictors of CKD [18]. However, computing the exact time to nadir eGFR requires intensive sampling, and may limit easy clinical implementation. Moreover, the perioperative laboratory values

tested was limited only to eGFR in that study. In contrast to the aforementioned studies, our study comprehensively considered perioperative laboratory values, including CBC, routine chemistry, and serum electrolytes as predictive factors. In all the final models, an increase in SCr levels from the preoperative value and history of CKD were the most important features (Table 2). Among other features, RN (compared to PN) was robustly associated with a higher  $\Delta SCr_{1y}$ . Older age was additionally depicted as a significant risk factor in most final models



( $Model_{final,pre}$ ,  $Model_{final,0d}$ ,  $Model_{final,2d}$ ,  $Model_{final,3d}$ , and  $Model_{final,4d}$ ). Male sex was only significant in  $Model_{final,pre}$ , which was built using just preoperative information; its effect was fully accounted for by other factors once postoperative information was available. The size of mass removed was a significant feature only immediately after surgery ( $Model_{final,0d}$ ), and its contribution disappeared after incorporating the SCr values on POD 1 and onwards. Overall, the important predictors for CKD occurrence were postoperative SCr levels, history of CKD, RN, and older age. Most other factors such as postoperative electrolytes, CBC, routine chemistry, and vital signs that were tested as candidate predictors were insignificant.

In addition to model development, we aimed to identify the optimal time point of postoperative SCr sampling for predicting CKD. To this end, we compared the predictive performances of 6 final models that used the information obtained on POD 0 to 5 ( $Model_{final,0d}$ – $Model_{final,5d}$ ) to that of a reference model that only used preoperative information ( $Model_{final,pre}$ ), and then to that of a model that used information collected at 1 month ( $Model_{final,1m}$ ). The predictive performances of  $Model_{final,0d}$ – $Model_{final,5d}$  were better than those of  $Model_{final,pre}$  but were almost similar to those of  $Model_{final,1m}$ . Hence, postoperative SCr levels measured in the first 5 days after surgery constituted crucial, nearly sufficient information to predict CKD at 1 year. Among

$Model_{final,0d}$ – $Model_{final,5d}$ ,  $Model_{final,4d}$  showed the best performance. In predicting the CKD stage, this model demonstrated classification accuracy of 79%, weighted averaged precision of 80%, and weighted averaged recall of 79% (Table 3). This suggested that POD 4 may be the optimal sampling point for predicting CKD.

To maximize the predictive performance, we adopted model stacking, a technique increasingly used in the medical field [24–26], wherein predictions from each of the 7 models (excluding the  $Model_{final,pre}$ ) were combined in all possible ways to generate 120 feature sets. Ridge regression models were then trained on these sets, yielding 120 meta-models. Our prediction strategy was to select from the meta-models, the one that makes best use of all available information. For example, if we had SCr measurements on POD 3, 4, and 5, we would choose a meta-model built on predictions of  $Model_{final,3d}$ ,  $Model_{final,4d}$ , and  $Model_{final,5d}$ .

This study has a few limitations. First, the data used for model building were retrospectively collected at a single center primarily comprising Korean patients. Hence, for generalization to patients of different ethnic backgrounds or those treated under different hospital environments, external validation is required. Second, this study only included surgeries that used minimally invasive laparoscopic or robotic techniques, and not open techniques. One study reported a lower risk of CKD in minimally invasive approaches [14], whereas others reported

**Table 3** Classification performances of chronic kidney disease stage based on the final models

		Predicted		
		Normal or Stage 1	Stage 2	Stage 3
<i>Model<sub>final,pre</sub></i>				
Observed	Normal or Stage 1	77	32	0
	Stage 2	16	101	19
	Stage 3	0	17	49
Accuracy: 73%				
Weighted average precision: 74%				
Weighted average recall: 73%				
<i>Model<sub>final,0d</sub></i>				
Observed	Normal or Stage 1	74	35	0
	Stage 2	12	108	16
	Stage 3	0	14	52
Accuracy: 75%				
Weighted average precision: 76%				
Weighted average recall: 75%				
<i>Model<sub>final,1d</sub></i>				
Observed	Normal or Stage 1	81	28	0
	Stage 2	15	108	13
	Stage 3	0	18	48
Accuracy: 76%				
Weighted average precision: 77%				
Weighted average recall: 76%				
<i>Model<sub>final,2d</sub></i>				
Observed	Normal or Stage 1	84	25	0
	Stage 2	14	105	17
	Stage 3	1	16	49
Accuracy: 77%				
Weighted average precision: 77%				
Weighted average recall: 77%				
<i>Model<sub>final,3d</sub></i>				
Observed	Normal or Stage 1	83	26	0
	Stage 2	13	109	14
	Stage 3	1	13	52
Accuracy: 78%				
Weighted average precision: 79%				
Weighted average recall: 78%				
<i>Model<sub>final,4d</sub></i>				
Observed	Normal or Stage 1	84	25	0
	Stage 2	16	110	10
	Stage 3	1	12	53
Accuracy: 79%				
Weighted average precision: 80%				
Weighted average recall: 79%				
<i>Model<sub>final,5d</sub></i>				
Observed	Normal or Stage 1	81	28	0
	Stage 2	18	105	13
	Stage 3	1	16	49

**Table 3** (continued)

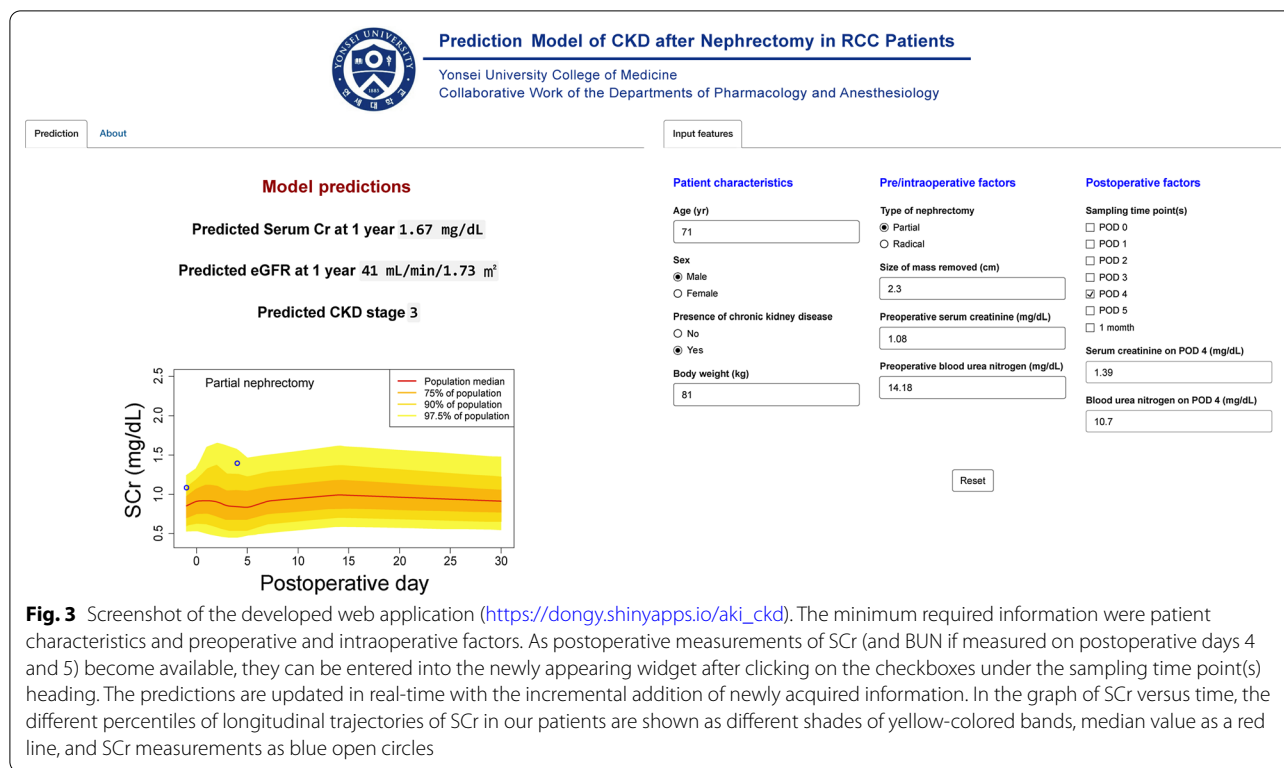
		Predicted		
		Normal or Stage 1	Stage 2	Stage 3
Accuracy: 76%				
Weighted average precision: 76%				
Weighted average recall: 76%				
<i>Model<sub>final,1m</sub></i>				
Observed	Normal or Stage 1	84	24	0
	Stage 2	14	107	15
	Stage 3	1	14	51
Accuracy: 73%				
Weighted average precision: 74%				
Weighted average recall: 73%				
Normal or stage 1, estimated glomerular filtration rate (eGFR) $\geq$ 90 mL/min/1.73 m <sup>2</sup> ; stage 2, eGFR 60–89 mL/min/1.73 m <sup>2</sup> ; stage 3, eGFR < 60 mL/min/1.73 m <sup>2</sup> . <i>d</i> day, <i>m</i> month				

a similar risk [10, 13]. However, minimally invasive approaches are being used more frequently for RN and PN [27]; thus, our model may be appropriate for future studies. Third, SCr was used as the surrogate of postoperative renal function and the target to be predicted, although the definition of CKD is based on eGFR [3]. However, as AKI is defined by changes in SCr, we wanted to examine the longitudinal changes of postoperative SCr with the concept of AKI, and CKD in continuum. Moreover, eGFR can easily be calculated using SCr. Therefore, we displayed eGFR in a web application by converting the predicted SCr to predicted eGFR and then finally classifying the CKD stages of the patients. Fourth, the strong correlation between the early increase in SCr levels and CKD at 1 year after surgery, while being useful for prediction, offers little to modify treatment for improving the clinical outcome. However, our results recommend that further investigations to prevent CKD progression be focused on preventing AKI in the first place, since early SCr elevation is strongly associated with long-term clinical outcomes. Despite these limitations, our model was the first to show the serial trends of SCr during 1 year with the incorporation of preoperative, intraoperative, and postoperative information.

**Conclusions**

We developed a model for predicting CKD after RN or PN, effectively extending the applicability of our prior model for predicting AKI after RN or PN [2]. The main strengths of our study were the active utilization of postoperative SCr and other laboratory values for CKD prediction and a clear demonstration of the importance of SCr measured within the first 5 days after surgery as





a predictor of the 1-year SCr level. Furthermore, our web application may be helpful for patient counseling and comprehensive management planning.

**Abbreviations**

AKI: Acute kidney injury; BUN: Blood urea nitrogen; CBC: Complete blood count; CKD: Chronic kidney disease; eGFR: Estimated glomerular filtration rate; MAR: Missing at Random; PN: Partial nephrectomy; POD: Postoperative day; RCC: Renal cell cancer; RN: Radical nephrectomy; SCr: Serum creatinine.

**Supplementary Information**

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-021-02976-2>.

**Additional file 1: Figure S1.** The proportion of missing values among serial serum creatinine measurements preoperative (CrP), on POD 0 to 5 (Cr0, ..., Cr5) and postoperative 1 month (Cr1M) (Left) and the frequency distribution of different combinations of missing values (Right) in (A) radical and (B) partial nephrectomy.

**Additional file 2: Figure S2.** Longitudinal trajectories of postoperative serum creatinine (SCr) levels in (A) radical and (B) partial nephrectomy, with the red line representing the median and the yellow band, the 95 percentiles.

**Additional file 3: Table S1.** Pairwise correlations of serum creatinine (SCr) at different time points with SCr at 1 year based on raw data prior to imputation. **Table S2.** Selected features and their associated regression coefficients of the Lasso models. Features with zero coefficients are not shown.

**Acknowledgements**

None.

**Authors' contributions**

Conceptualization and design: DC, NYK, KJK, and SYK; acquisition of data: NYK, CO, and SYK; analysis and interpretation of data: DC, NYK, and SYK; writing the original draft: DC, NYK, and SYK; writing, reviewing, and editing: DC, NYK, KJK, KP, CO, and SYK; final approval: DC and SYK; agreement to be accountable for all aspects of the work: DC, NYK, KJK, KP, CO, DC, and SYK. All authors read and approved the final manuscript.

**Funding**

This study was supported by new faculty grant of Yonsei university, medical college, Seoul, Korea (Dongwoo Chae 6-2019-0197 and 6-2020-0239).

**Availability of data and materials**

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Declarations**

**Ethics approval and consent to participate**

This study was approved by the Institutional Review Board and Hospital Research Ethics Committee of Severance Hospital, Yonsei University Health System, Seoul, Korea (IRB number: 4-2020-0221). Due to the retrospective nature of the study, the requirement for informed consent from patients was waived.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

Received: 4 March 2021 Accepted: 4 July 2021

Published online: 16 July 2021

**References**

- Schmid M, Abd-El-Barr AE, Gandaglia G, Sood A, Olugbade K Jr, Ruhotina N, et al. Predictors of 30-day acute kidney injury following radical and partial nephrectomy for renal cell carcinoma. *Urol Oncol*. 2014;32(8):1259–66.
- Kim NY, Chae D, Lee J, Kang B, Park K, Kim SY. Development of a risk scoring system for predicting acute kidney injury after minimally invasive partial and radical nephrectomy: a retrospective study. *Surg Endosc*. 2021;35(4):1626–35.
- Chawla LS, Bellomo R, Bihorac A, Goldstein SL, Siew ED, Bagshaw SM, et al. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 Workgroup. *Nat Rev Nephrol*. 2017;13(4):241–57.
- Heung M, Steffick DE, Zivin K, Gillespie BW, Banerjee T, Hsu CY, et al. Acute kidney injury recovery pattern and subsequent risk of CKD: An analysis of veterans health administration data. *Am J Kidney Dis*. 2016;67(5):742–52.
- Bravi CA, Vertosick E, Benfante N, Tin A, Sjoberg D, Hakimi AA, et al. Impact of acute kidney injury and its duration on long-term renal function after partial nephrectomy. *Eur Urol*. 2019;76(3):398–403.
- Coca SG, Singanamala S, Parikh CR. Chronic kidney disease after acute kidney injury: a systematic review and meta-analysis. *Kidney Int*. 2012;81(5):442–8.
- Lane BR, Campbell SC, Demirjian S, Fergany AF. Surgically induced chronic kidney disease may be associated with a lower risk of progression and mortality than medical chronic kidney disease. *J Urol*. 2013;189(5):1649–55.
- Wu J, Suk-Ouichai C, Dong W, Antonio EC, Derweesh IH, Lane BR, et al. Analysis of survival for patients with chronic kidney disease primarily related to renal cancer surgery. *BJU Int*. 2018;121(1):93–100.
- Jeon HG, Jeong IG, Lee JW, Lee SE, Lee E. Prognostic factors for chronic kidney disease after curative surgery in patients with small renal tumors. *Urology*. 2009;74(5):1064–8.
- Barlow LJ, Korets R, Laudano M, Benson M, McKiernan J. Predicting renal functional outcomes after surgery for renal cortical tumours: a multifactorial analysis. *BJU Int*. 2010;106(4):489–92.
- Ngo TC, Hurley MP, Thong AE, Jeon SH, Leppert JT, Chung BI. Estimating the risk of chronic kidney disease after nephrectomy. *Can J Urol*. 2013;20(6):7035–41.
- Pignot G, Bigot P, Bernhard JC, Bouliere F, Bessedes T, Bensalah K, et al. Nephron-sparing surgery is superior to radical nephrectomy in preserving renal function benefit even when expanding indications beyond the traditional 4-cm cutoff. *Urol Oncol*. 2014;32(7):1024–30.
- Abdel Raheem A, Shin TY, Chang KD, Santok GDR, Alenzi MJ, Yoon YE, et al. Yonsei nomogram: A predictive model of new-onset chronic kidney disease after on-clamp partial nephrectomy in patients with T1 renal tumors. *Int J Urol*. 2018;25(7):690–7.
- Bhindi B, Lohse CM, Schulte PJ, Mason RJ, Chevillie JC, Boorjian SA, et al. Predicting renal function outcomes after partial and radical nephrectomy. *Eur Urol*. 2019;75(5):766–72.
- Cho A, Lee JE, Kwon GY, Huh W, Lee HM, Kim YG, et al. Post-operative acute kidney injury in patients with renal cell carcinoma is a potent risk factor for new-onset chronic kidney disease after radical nephrectomy. *Nephrol Dial Transplant*. 2011;26(11):3496–501.
- Kim WH, Shin KW, Ji SH, Jang YE, Lee JH, Jeong CW, et al. Robust association between acute kidney injury after radical nephrectomy and long-term renal function. *J Clin Med*. 2020;9(3):619.
- Lane BR, Babineau DC, Poggio ED, Weight CJ, Larson BT, Gill IS, et al. Factors predicting renal functional outcome after partial nephrectomy. *J Urol*. 2008;180(6):2363–8.
- Breslow A, Kaufman RM, Lawsky AR. The effect of surgery on the concentration of circulating megakaryocytes and platelets. *Blood*. 1968;32(3):393–401.
- Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods*. 2002;7(2):147–77.
- Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann Transl Med*. 2016;4(2):30.
- Sato T, Yamanishi Y, Horimoto K, Kanehisa M, Toh H. Partial correlation coefficient between distance matrices as a new indicator of protein-protein interactions. *Bioinformatics*. 2006;22(20):2488–92.
- Rule AD. The CKD-EPI equation for estimating GFR from serum creatinine: real improvement or more of the same? *Clin J Am Soc Nephrol*. 2010;5(6):951–3.
- Levin A, Stevens PE. Summary of KDIGO 2012 CKD Guideline: behind the scenes, need for guidance, and a framework for moving forward. *Kidney Int*. 2014;85(1):49–61.
- Zhang Z, Qiu H, Li W, Chen Y. A stacking-based model for predicting 30-day all-cause hospital readmissions of patients with acute myocardial infarction. *BMC Med Inform Decis Mak*. 2020;20(1):335.
- Kurz CF, Maier W, Rink C. A greedy stacking algorithm for model ensemble and domain weighting. *BMC Res Notes*. 2020;13(1):70.
- Kim C, You SC, Reys JM, Cheong JY, Park RW. Machine-learning model to predict the cause of death using a stacking ensemble method for observational data. *J Am Med Inform Assoc*. 2020. <https://doi.org/10.1093/jamia/ocaa277>.
- Poon SA, Silberstein JL, Chen LY, Ehdabe B, Kim PH, Russo P. Trends in partial and radical nephrectomy: an analysis of case logs from certifying urologists. *J Urol*. 2013;190(2):464–9.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

