Nomograms for predicting long-term overall survival and disease-specific survival of patients with clear cell renal cell carcinoma

Guanghao Zhang1,*
Yun Wu2,*
Jiashu Zhang3
Zhiqing Fang4
Zhaoxu Liu4
Zhonghua Xu4
Yidong Fan4

1Department of Financial Mathematics and Financial Engineering, School of Economics, Shandong University, Jinan, Shandong, People’s Republic of China
2Department of Medical Oncology, National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, People’s Republic of China
3Department of Breast Surgery, Qilu Hospital of Shandong University, Jinan, Shandong, People’s Republic of China
4Department of Urology, Qilu Hospital of Shandong University, Jinan, Shandong, People’s Republic of China

*These authors contributed equally to this work

Objectives: The aim of this study was to establish comprehensive and practical nomograms, based on significant clinicopathological parameters, for predicting the overall survival (OS) and the disease-specific survival (DSS) of patients with clear cell renal cell carcinoma (ccRCC).

Patients and methods: The data of 35,151 ccRCC patients, diagnosed between 2004 and 2014, were obtained from the database of the Surveillance, Epidemiology, and End Results (SEER) program. The Kaplan–Meier method and Cox proportional hazards regression model were used to evaluate the prognostic effects of multiple clinicopathological variables on survival. Based on Cox models, a nomogram was constructed to predict the probabilities of OS and DSS for an individual patient. The predictive performance of nomograms was evaluated using the concordance index (C-index) and calibration curves.

Results: According to univariate and multivariate analyses, age at diagnosis, sex, race, marital status, surgical approach, tumor node metastasis (TNM) stage, and Fuhrman grade significantly correlated with the survival outcomes. These characteristics were used to establish nomograms. The nomograms showed good accuracy in predicting 3-, 5-, and 10-year OS and DSS, with a C-index of 0.79 (95% CI, 0.79–0.80) for OS and 0.87 (95% CI, 0.86–0.88) for DSS. All calibration curves revealed excellent consistency between predicted and actual survival.

Conclusion: Nomograms were developed to predict death from ccRCC treated with nephrectomy. These new prognostic tools could aid in improving the predictive accuracy of survival outcomes, thus leading to reasonable individualized treatment.

Keywords: clear cell renal cell carcinoma, prognosis, nomogram, survival analysis

Introduction

Renal cell carcinoma (RCC) is one of the most common tumors worldwide, which accounted for >140,000 deaths per year.1 The incidence of renal cancer and the overall survival (OS) have continued to increase over the past 3 decades in the world.2 Patients with RCC usually have unfavorable prognosis, which makes RCC a serious problem for public health.3,4

It has been well established that RCC is an extremely heterogeneous type of disease with different genetic and molecular alterations, resulting in many histological subtypes such as clear-cell, papillary (types 1 and 2), and chromophobe RCC.5 The most common type of RCC, characterized by compact nests of tumor cells with clear cytoplasm, is known as clear cell RCC (ccRCC) and represents ~70% of all cases.6 In a recent study, Leibovich et al7 found that ccRCC subtype was an independent predictor of cancer-specific death and metastasis compared with papillary and chromophobe carcinoma after adjusting for other clinicopathological features. Considering the higher
incidence rate and worse prognosis of ccRCC than other kinds of RCC, we focused on ccRCC in this study.\textsuperscript{8–11}

The tumor node metastasis (TNM) classification system derived from the American Joint Committee on Cancer (AJCC) represents the gold standard staging scheme after nephrectomy for RCC.\textsuperscript{12} This system is most widely used for prognostic estimates and clinical treatments in patients with cancer. However, some researchers have reported that other clinicopathological factors such as race, sex, age, tumor size, and mode of presentation may also influence the RCC patients’ outcomes.\textsuperscript{13–15} No single factor can account for the wide variability of outcomes observed in individual patients. Neglecting these significantly prognostic parameters may reduce the accuracy of survival predictions. Therefore, a comprehensive prognostic evaluation system that includes clinicopathological and demographic variables is needed in clinical practice.

A nomogram is a convenient graphical representation of a mathematical model, in which various important factors are combined to predict a specific endpoint.\textsuperscript{11} By integrating those clinical and pathological factors, a nomogram can provide individualized estimates of the probability of an event over time, such as a patient’s individual probability of disease recurrence and death. Therefore, these algorithms might be used as a reliable tool for predicting the clinical outcomes and guiding decisions regarding surgery, surveillance, and adjuvant therapies.

To date, several postoperative nomograms have been generated for predicting the probability of RCC recurrence and survival.\textsuperscript{16} Kattan et al.\textsuperscript{11} published the first nomogram in 2001, which was designed to calculate the likelihood of recurrence after surgery for RCC. The Kattan nomogram is currently known to be the most accurate nomogram, although there are other kinds of nomograms that have been used, including University of California, Los Angeles, Integrated Staging System (UISS) from the University of Los Angeles and Stage, Size, Grade and Necrosis (SSIGN) score from the Mayo Clinic.\textsuperscript{17,18} Based on the data obtained from long follow-ups, nomogram could present a graphical predictive model to integrate multiple variables, which make nomogram outstanding from many prognostic models. However, there is no prediction model to estimate the OS for ccRCC patients. The purpose of this study was to develop a prognostic nomogram for estimating the survival of ccRCC patients.

Patients and methods

Patient selection

In this study, all data were acquired from the National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) cancer database. From 20 cancer registries that cover ~28% of the population of the United States, the SEER program collects and publishes data including cancer incidence and mortality. Only patients diagnosed between 2004 and 2014 with ccRCC were considered. Patients diagnosed before 2004 were excluded since TNM stage information was not recorded in the SEER database until 2004. Additionally, to ensure adequate follow-up time, patients diagnosed after 2014 were not included. From the SEER database, we identified a cohort of 35,151 patients according to the following inclusion criteria: age at diagnosis between 18 and 80 years, known Fuhrman grade, positive histology, active follow-up, known race, known marital status at diagnosis, and surgical treatment with either partial or radical nephrectomy (RN). Patients whose race was recorded as American Indian/Alaskan Native or Asian/Pacific Islander in SEER were assigned to an “others” race category for analysis. Moreover, patients with T0/TX/NX/MX stage were excluded.

Construction of the nomograms

The variables analyzed in this study were race, sex, Fuhrman grade, pathological stage (T/N/M, derived AJCC, sixth edition), surgical treatment, age at diagnosis, and marital status. Some of the variables were grouped in the analysis: the surgical treatment variable was grouped into partial nephrectomy (RX Summ-Surg Prim Site code: 10-30) and RN (RX Summ-Surg Prim Site code: 40-90), while AJCC T stage was grouped into T1 (T1/T1a/T1b/T1 not otherwise specified), T2, T3 (T3/T3a/T3b/T3c/T3NOS), and T4. Frequency and proportion were reported for each variable analyzed in this study.

The univariable and multivariable Cox regression analyses were used to calculate the effect of variables on OS and disease-specific survival (DSS). The measure of the effect of each variable on OS and DSS was presented as a hazard ratio (HR). Nomograms in this study were created using information obtained from the Cox regression analysis.

Validation of the nomograms

Validation of this nomogram was performed with repeated sampling (bootstrapping) of our data. Marginal estimates and model-average prediction probabilities were used to create calibration curves. To decrease the overfit bias, the nomograms were subjected to 1,000 bootstrap resamples for internal validation in the validation cohort. In a perfectly calibrated model, the predictions should fall on the diagonal 45° line of the calibration plot. Predictive performance was assessed using the concordance index (C-index), which
remakes the area under the curve (AUC) but appears to be better suited for censored data.\textsuperscript{19} The value of the C-index statistic ranged from 0.5 (no discrimination) to 1 (perfect discrimination), and higher C-index values indicated a better prognostic model.

**Statistical analyses**

The Cox regression analysis was performed by the IBM SPSS Statistics for statistical computing and graphics (Version 24.0.0), and the nomogram was graphed by the R software (Version 3.3.0). All statistical tests were considered statistically significant at $P<0.05$, and estimates were considered significant at a CI of 95%.

**Results**

**Clinicopathological characteristics of the patients**

According to the inclusion criteria, a total of 35,151 patients were analyzed in this study. Descriptive analysis of the clinicopathological characteristics is shown in Table 1. Generally, most patients were male (21,692; 61.71%), older than 50 years (27,264; 77.56%), married (23,433; 66.66%), and Caucasian (30,226; 85.99%). Moreover, most patients underwent RN (24,280; 69.07%).

**Independent prognostic factors in the cohort**

The univariable and multivariable Cox regression analyses were used for the analysis of variables that influence OS and DSS, and we use HR to quantify the effect of each variable on OS and DSS. The following variables are selected to be analyzed in Cox regression analysis: race, sex, Fuhrman grade, pathological stage (T/N/M, AJCC, sixth edition), surgical treatment, age at diagnosis, and marital status; the results are shown in Table 2. In the univariable Cox regression analysis for DSS, all variables except race reached statistical significance, while all variables but race and sex reached statistical significance in the multivariable Cox regression analysis for DSS. All variables have statistical significance both in the univariable and multivariable Cox regression analyses for OS.

Based on results in Table 2, it is possible to evaluate the risk of death faced by patients. For example, older patients are more likely to experience worse DSS outcome. In addition, single patients are more likely to have poor prognosis than married patients. With higher Fuhrman grade and higher T/N/M stage, patients have lower probability of survival due to cancer progression. Moreover, patients with RN have higher HR than those with partial nephrectomy.

**Prognostic nomograms for OS and DSS**

Considering outcomes of multivariable Cox regression analysis for OS and DSS, all of the variables are included to create the nomogram for OS and DSS. The prognostic nomogram for 3-, 5-, and 10-year OS is shown in Figure 1. The prognostic nomogram for 3-, 5-, and 10-year DSS is shown in Figure 2. By adding up the scores associated with each variable, and projecting total scores to the bottom scale, probabilities can be estimated for 3-, 5-, and 10-year OS and DSS.

To clarify, we can take two ccRCC patients for examples. A 55-year-old single male patient with Fuhrman grade II,
### Table 2
Univariable and multivariable Cox regression model analyses of overall survival and disease-specific survival in nomogram cohort

<table>
<thead>
<tr>
<th>Variables</th>
<th>Disease-specific survival</th>
<th>Overall survival</th>
<th>Disease-specific survival</th>
<th>Overall survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariable Cox</td>
<td></td>
<td>Multivariable Cox</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P-value</td>
<td>HR (95% CI)</td>
<td>P-value</td>
</tr>
<tr>
<td>Race</td>
<td>0.416</td>
<td>0.006</td>
<td></td>
<td>0.269</td>
</tr>
<tr>
<td>Black</td>
<td>1 (Reference)</td>
<td></td>
<td>1 (Reference)</td>
<td>0.872 (0.731–1.041)</td>
</tr>
<tr>
<td>Others</td>
<td>1.043 (0.875–1.243)</td>
<td>0.001</td>
<td>0.723 (0.630–0.830)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>White</td>
<td>1.085 (0.950–1.240)</td>
<td>0.027</td>
<td>0.839 (0.759–0.926)</td>
<td>0.001</td>
</tr>
<tr>
<td>Sex</td>
<td>&lt;0.001</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1 (Reference)</td>
<td></td>
<td>1 (Reference)</td>
<td>0.955 (0.834–1.092)</td>
</tr>
<tr>
<td>Male</td>
<td>1.359 (1.271–1.454)</td>
<td>&lt;0.001</td>
<td>1.198 (1.133–1.266)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age at diagnosis (years)</td>
<td>&lt;0.001</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>1 (Reference)</td>
<td></td>
<td>1 (Reference)</td>
<td>1.050 (0.979–1.126)</td>
</tr>
<tr>
<td>40–49</td>
<td>2.295 (1.798–2.931)</td>
<td>&lt;0.001</td>
<td>1.609 (1.311–1.974)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>50–59</td>
<td>3.391 (2.685–4.284)</td>
<td>&lt;0.001</td>
<td>2.066 (1.698–2.513)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>60–69</td>
<td>3.968 (3.144–5.008)</td>
<td>&lt;0.001</td>
<td>2.780 (2.288–3.378)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>70–80</td>
<td>4.440 (3.508–5.619)</td>
<td>&lt;0.001</td>
<td>4.199 (3.454–5.105)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Marital status</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>1 (Reference)</td>
<td></td>
<td>1 (Reference)</td>
<td>1.204 (1.125–1.289)</td>
</tr>
<tr>
<td>Single</td>
<td>1.087 (1.018–1.162)</td>
<td>0.013</td>
<td>1.385 (1.312–1.461)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Grade</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1 (Reference)</td>
<td></td>
<td>1 (Reference)</td>
<td>1.339 (1.048–1.711)</td>
</tr>
<tr>
<td>II</td>
<td>1.644 (1.393–1.941)</td>
<td>&lt;0.001</td>
<td>1.609 (1.311–1.974)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>III</td>
<td>5.837 (4.967–6.858)</td>
<td>&lt;0.001</td>
<td>2.066 (1.698–2.513)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IV</td>
<td>19.026 (16.110–22.469)</td>
<td>&lt;0.001</td>
<td>4.199 (3.454–5.105)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>T</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>5.211 (4.704–5.774)</td>
<td>&lt;0.001</td>
<td>2.395 (2.132–2.691)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>T2</td>
<td>10.283 (9.483–11.151)</td>
<td>&lt;0.001</td>
<td>1.410 (1.271–1.563)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>T3</td>
<td>12.417 (10.769–14.316)</td>
<td>&lt;0.001</td>
<td>1.410 (1.271–1.563)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>T4</td>
<td>2.728 (6.941–8.605)</td>
<td>&lt;0.001</td>
<td>2.780 (2.288–3.378)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>N</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>10.850 (9.741–12.086)</td>
<td>&lt;0.001</td>
<td>1.797 (1.612–2.004)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>N1</td>
<td>14.948 (13.199–16.928)</td>
<td>&lt;0.001</td>
<td>2.384 (2.103–2.703)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>M</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>3.649 (3.354–3.970)</td>
<td>&lt;0.001</td>
<td>1.790 (1.636–1.958)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>M1</td>
<td>15.854 (14.850–16.925)</td>
<td>&lt;0.001</td>
<td>4.031 (3.765–4.316)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Nephrectomy</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial</td>
<td>7.873 (6.807–9.106)</td>
<td>&lt;0.001</td>
<td>2.399 (2.058–2.796)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviation:** HR, hazard ratio.
Figure 1 Nomograms for predicting the 3-, 5-, and 10-year OS of ccRCC patients.

Note: Instructions for using the nomogram: first, assign the points of each characteristic of the patient by drawing a vertical line from that variable to the points’ scale, then, sum all the points, and draw a vertical line from the total points scale to the 3-, 5-, and 10-year OS to obtain the probability of death.

Abbreviations: ccRCC, clear cell renal cell carcinoma; OS, overall survival.

Figure 2 Nomograms for predicting the 3-, 5-, and 10-year DSS of ccRCC patients.

Note: Instructions for using the nomogram: first, assign the points of each characteristic of the patient by drawing a vertical line from that variable to the points scale, then, sum all the points, and draw a vertical line from the total points scale to the 3-, 5-, and 10-year DSS to obtain the probability of death.

Abbreviations: ccRCC, clear cell renal cell carcinoma; DSS, disease-specific survival.
T2N1M1 stage who underwent RN would score $>300$ points, which means that the patient has $\sim 50\%$ probability of survival at the third year. However, for a 76-year-old single male patient with Fuhrman grade IV, T4N2M1 stage who underwent RN, the total score would be 450 points; thus, the patient’s probability of survival in the third year would be far $<10\%$.

Validation of the nomograms

In the nomogram, each variable will be cleared based on the status of patient; thus, each variable could provide a point value, which will be found on the top line of the nomogram. Then, the points of each variable should be added up to get a total point, by which we could find the probability of survival at 3, 5, or 10 years. Validation of the nomogram was processed internally. The $C$-indexes for the nomograms to predict OS and DSS were 0.79 (95% CI, 0.79–0.80) and 0.87 (95% CI, 0.86–0.88), respectively. This finding implied that these models were reasonably accurate. The internal calibration plots of the OS and DSS are presented in Figure 3, revealing an excellent correlation in OS and DSS between the nomogram and observed outcome.

Discussion

RCC is the most common kidney malignancy in adults, accounting for 2%-3% of all malignant diseases in adults. RCC is a heterogeneous group of malignancies that vary histologically, genetically, and molecularly. Clear cell subtype represents $>70\%$ of all RCC patients and is the most aggressive subtype.20-22 The remaining 30% of RCC subtypes are generally indolent such as chromophobe and papillary RCC.23,24 In light of the high incidence and the relative unfavorable prognosis of ccRCC, a brief nomogram based on long-term follow-up of a population-based cohort that predicts long-term OS and DSS should be quite useful and practical for clinicians.

The nomogram is a graphic depiction of a prediction model that combines multiple biological and clinical variables to determine the probabilities of clinical events.12,25 Given that many prognostic factors of patients with ccRCC are available for physicians in practice, no single factor can account for the wide variability in outcomes observed in individual patients. And compared to the current tumor staging system (TNM), the nomogram showed better prediction accuracy and prognostic value.26,27 Hence, a more comprehensive and precise prognostic model is needed.

To date, there is no comprehensive nomogram containing the demographic and clinicopathological variables for ccRCC. Using the SEER database, we developed novel nomograms based on 35,151 cases predicting the 3-, 5-, and 10-year OS and DSS rates of ccRCC patients based on univariable and multivariable Cox regression analyses. The predictive performance of the nomograms was evaluated using an internal bootstrap resampling method. The $C$-index demonstrated the accuracy of nomograms to predict the 3-, 5-, and 10-year OS and DSS rates of ccRCC patients. With these nomograms, clinicians can classify patients into different risk groups, thus rendering individualized treatment and surveillance possible.

The TNM classification is the most widely used system for prognostic estimates of patients with cancer. Notably, our nomogram also showed that the T/N/M categories are making the largest contribution to prognosis. As the enlargement of tumor size, the increasing number of metastatic lymph nodes and the presence of distant organ metastasis, the risk of mortality increased remarkably.28 Improvements in imaging techniques and biologic makers may help identify patients with minimal metastatic disease, thus improving the prognosis of ccRCC patients apparently. In this study, we found that patients who received partial nephrectomy showed a better OS and DSS than those with RN. Many retrospective reports have suggested a benefit in OS for patients treated with nephron-sparing surgery (NSS) relative to RN.29,30 On one hand, current guidelines recommend NSS as the standard surgical treatment for clinical T1a (tumor size $\leq 4$ cm) renal tumors.31 As for RN, it is utilized for many cT1b (4–7 cm) tumors and it is the recommended treatment for T2a (7–10 cm) and larger tumors.32,33 Mir et al’s34 study confirmed that patients undergoing PN were younger and had smaller masses, which might partially account for better OS for NSS. On the other hand, patients with solitary RCC that has been surgically treated usually have long-term cancer survival time (85%-96% cancer-specific survival 10 years after surgery), so renal functional outcomes are important.35 A decrease in renal function is associated with higher risk of severe cardiovascular disease and all-cause mortality.36 In a meta-analysis of 34 comparative studies, Kim et al37 calculated a cumulative 61% reduction in the risk of severe chronic kidney disease, as well as a 19% risk reduction for all-cause mortality for those undergoing NSS. Thereby, NSS has been used increasingly in patients with small renal mass due to the additional benefit of renal function preservation and the reduced risk for cardiovascular events and other adverse effects (eg, new-onset hypertension, diabetes, and cerebrovascular disease).38
As shown in the nomograms that we have presented, age at diagnosis and Fuhrman grade had a great prognostic value on OS and DSS. Several studies have found that younger patients with ccRCC have better prognosis than older patients, which might be related to a lower grade and stage at diagnosis. Jung et al\textsuperscript{14} proved that young age was associated with lower grade and stage and was an independent predictor of survival among patients with ccRCC. A study of over 1,000 cases demonstrated that young patients defined as age <45 years were more likely to have lower stage and grade and had a higher 5-year cancer-specific survival than older patients.\textsuperscript{39} Fuhrman grade has been demonstrated to be a significant prognostic factor for clear cell tumors in multiple studies as well.\textsuperscript{40,41} Fuhrman system distinguishes grade 1–4 based on nuclear size, morphology, and nucleolar prominence and correlates greatly with tumor size, stage, and the presence of metastasis.\textsuperscript{42,43}

There are several prognostic models that have been developed to predict disease recurrence and survival after nephrectomy for RCC, using different variables, tools (nomograms or prognostic categories), and endpoints previously. In 2001, Kattan et al\textsuperscript{11} reported a nomogram that can...
be used to predict the 5-year probability of treatment failure among patients with newly diagnosed RCC. The four factors included in this nomogram were the presence of symptoms, histological subtype, tumor size, and standard TNM stage according to the 1997 version. The bootstrapping estimate of the internally predictive accuracy of the nomogram was 0.74. When applied to external populations in Europe, the Kattan nomogram has shown variable prognostic accuracy ranging from 61 to 71%. In the same year, Zisman et al reported the UISS for patients with RCC, based on 661 patients undergoing nephrectomy at the University of California, Los Angeles, between 1989 and 1999. The UISS integrated stage, grade, and Eastern Cooperative Oncology Group (ECOG) performance status to predict OS following partial or RN. This system has been validated in a cohort of >4,000 external patients, and the predictive accuracy was found to be 81% for patients with localized RCC. In the next year, researchers in the Mayo Clinic reported SSIGN for patients with ccRCC treated with RN, based on 1,801 patients who underwent RN between 1970 and 1988. The C-index from an SSIGN model that contained the ccRCC score was 0.84. Subsequently, European and Japanese studies have confirmed the prognostic accuracy of the SSIGN algorithm ranging from 81 to 88%. Finally, in 2007, Karakiewicz et al developed the model included >2,500 patients with RCC incorporated TNM stage, tumor size, histological subtype, age, sex, and symptoms at presentation in order to predict cancer-specific mortality among patients with RCC. The internally validated accuracy of this nomogram was 86%. In the same study, an external validation of this nomogram yielded a predictive accuracy of ~86.7%. From the above, our results indicate that the discriminant properties of our model (79% for OS and 87% for DSS) are comparable to other models that addressed similar end points. Nevertheless, external validation on large population and multicenter datasets is needed to compare the discriminating accuracy of those prognostic models with our models more precisely.

However, all of these models have defects more or less. Nomogram created by Kattan et al did not include tumor grade, which is a strong predictor that has been shown to be significant by other investigators. Histological subtype is not illustrated in the UISS. As for SSIGN, this model is based on limited scale of data and its applicability for nationwide patients remains to be confirmed. And the model created by Karakiewicz was devoid of demographic variables, which may influence the predictive accuracy. To the best of our knowledge, this is the first large-population-based study to construct a nomogram to predict the OS and CSS for patients with ccRCC. One merit of our nomogram over that of previous studies resides in our large sample size and long follow-up. With a median follow-up of 65.5 months, our nomogram cohort (n=35,151) is the largest and covered almost thousands of hospitals, allowing the nomograms to be used widely for decision making in clinical practice. Besides, our study first took some demographic variables into account, such as race, sex, and marital status, which are likely to improve nomogram accuracy. In this study, White married females have a better OS and DSS than Black single males. Racial disparities could be attributed to lower socioeconomic status, poor health coverage, no or less access to medical care for undergoing ultrasound and CT scans, and later stage at cancer diagnosis among African American patients. Many findings suggest that marriage may have a protective effect on RCC patients, which related to the support from spouses, such as receiving medical assistance, assisting in activities of daily living, and medication reminders. Women with RCC had significantly better OS rates than men, and its sex-specific differences have previously been reported. The reason may be associated with genetic and social environmental factors. Sex-specific mutation of genes such as BAP1, is evident in ccRCC, which may affect the OS of the patient. Furthermore, women had a significantly higher proportion of incidentally detected RCC and lower grade and smaller size of the tumors at presentation, which may be related to more extensive use of the health care system and more likely to have incidental detection of tumors during imaging for abdominal or pelvic pathology than men. Finally, our cohorts are the most updated version compared with those previous models. Accurate outcome prediction for patients with ccRCC treated with nephrectomy is critically important for postoperative management.

Moreover, our study has several limitations. First, the SEER database is devoid of variables such as ECOG prognostic scores, mode of presentation, and detailed histological information such as tumor necrosis, which have proven to be predictive of survival. Additionally, it has been widely accepted that treatment details regarding surgery and other adjuvant therapy such as chemotherapy, radiation therapy, and targeted therapy were important prognostic factors for ccRCC. However, that information except for surgical treatment was not available in the SEER database and cannot be incorporated in our analysis. Furthermore, due to the retrospective nature of our study, these nomograms must be further validated in a prospective cohort or a clinical trial before being applied to clinical use. Finally, despite having achieved accuracy that exceeded that of other existing
models, our nomogram is not perfect. Indeed, 21% of predictions will be made incorrectly. This flaw is shared with virtually all predictive models, given that 100% correct predictions virtually are never achieved.

The present study showed that age, sex, race, marital status, surgical approach, TNM stage, and Fuhrman grade were independent risk factors for survival in patients with ccRCC. Nomograms were developed to accurately predict the 3-, 5-, and 10-year OS and DSS rates of these patients, based on patient-specific characteristics. These predictive tools could help clinicians identify high-risk patients, tailor therapeutic and follow-up strategies as well as make individualized predictions of patient survival.

**Data sharing statement**
All data are fully available upon request to the corresponding author, Zhiqing Fang.

**Acknowledgments**
The authors would like to thank the SEER program for providing open access to the database. This study was supported by the Key Project of Research and Development Plan in Shandong Province (no 2016GSF201172) and the Scientific Research Foundation of Shandong Province for Outstanding Young Scientist Award (no BS2014YY036). None of the funding bodies had any part in the design of the study and collection, analysis, interpretation of data, and manuscript writing.

**Author contributions**
ZF contributed to the conception and design. GZ and YW contributed to the collection of patient information and drafting of the article. ZF interpreted the data. JZ, GZ, ZL, ZX, and YF helped revising the manuscript. All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

**Disclosure**
The authors report no conflicts of interest in this work.

**References**


