ELSEVIER

Urologic Oncology: Seminars and Original Investigations 37 (2019) 356.e9−356.e18

UROLOGIC ONCOLOGY

Clinical-Kidney cancer

External validation of the updated Leibovich prognostic models for clear cell and papillary renal cell carcinoma in an Asian population

Han Jie Lee*, Alvin Lee, Hong Hong Huang, Weber Kam On Lau

Department of Urology, Singapore General Hospital, Singapore

Received 6 January 2019; received in revised form 17 February 2019; accepted 20 February 2019

Abstract

Purpose: The Leibovich model was updated to prognosticate oncological outcomes in postnephrectomy nonmetastatic renal cell carcinoma (RCC) for each major histological subtype including clear cell (ccRCC), papillary (papRCC), and chromophobe RCC. We evaluated its performance in an independent population of predominantly Asian patients from Singapore.

Materials and Methods: Nine hundred and forty two binephric patients with nonmetastatic unilateral RCC treated with radical/partial nephrectomy from 1990 to 2015 from Singapore were retrospectively reviewed. Based on the Leibovich model, ccRCC patients were scored from 0 to 25 and papRCC patients divided into 3 risk groups. Primary outcomes of progression-free survival (PFS) and cancer-specific survival (CSS) were assessed with the Kaplan−Meier method. Receiver operating characteristic curves and calibration plots were obtained to determine discrimination and calibration respectively.

Results: Eight hundred and twenty nine patients (88%) had ccRCC where 16.2% experienced disease progression while 11.9% died of RCC over a median follow-up of 76 (42−117) months. There was good discrimination (c-index 0.81 for PFS, 0.83 for CSS) and calibration (PFS calibration-in-the-large 0.002 and calibration slope 0.99, CSS calibration-in-the-large 0.005 and calibration slope 0.96). One hundred and thirteen patients (12%) had papRCC, where 18.6% progressed while 14.2% died from RCC over a median follow-up of 69.5 (36.0−112.0) months. Discrimination was slightly weaker (c-index 0.72 for PFS, 0.74 for CSS), and the model was only calibrated for CSS (calibration-in-the-large 0.002, calibration slope 0.98), not for PFS (calibration-in-the-large 0.09, calibration slope 1.93).

Conclusions: The updated Leibovich score is applicable for prognostication of progression and death in both ccRCC and papRCC, even when applied to an independent population of Asian patients. Further validation is required to ensure accuracy in prognosticating PFS for papRCC. © 2019 Elsevier Inc. All rights reserved.

Keywords: Renal cell carcinoma; Pathology; Prognosis; Prognostic model; Survival; External validation

1. Introduction

The incidence of Renal Cell Carcinoma (RCC) remains high, constituting 2% to 3% of adult malignancies [1]. Surgery remains the mainstay of curative therapy for nonmetastatic RCC, but a third of cases that undergo surgery for early-stage localized disease have relapsed after a median of 1.9 years [2], and heterogeneity in outcomes and survival still exists. Numerous postoperative models and nomograms have therefore been designed for prognostication of such outcomes [3−6], to determine surveillance strategies, or selection for adjuvant therapy [7]. Trials for adjuvant therapy after nephrectomy have yielded conflicting results [8−9], and due to high toxicity and costs, accurate risk-stratification of patients for appropriate inclusion into such adjuvant treatment trials is key.

The Leibovich score, one of the commonly evaluated prognostic models developed in 2005, was based on pathological tumor features and sought to predict disease-free survival in patients with nonmetastatic clear cell RCC (ccRCC) [6]. An update was recently published [10], and is of interest as it proposed different scoring systems for major histologic subtypes of RCC including clear cell, papillary (papRCC), and chromophobe RCC (chrRCC). It also sought to incorporate individual components of tumor

*Corresponding author. Tel: (65) 65767678; Fax: (65) 62273787
E-mail address: hanjie.lee@mohh.com.sg (H.J. Lee).

https://doi.org/10.1016/j.urolonc.2019.02.014
1078-1439/© 2019 Elsevier Inc. All rights reserved.

staging instead of using the AJCC TNM stage, in order to account for future changes that might occur with revisions to the staging system. For ccRCC, scores were assigned to each feature as shown in Supplementary Table A, which were then added together to obtain the total score. The scores for progression and death were then used to predict progression-free survival (PFS) and cancer-specific survival (CSS). For papRCC and chrRCC, patients were alternatively categorized into risk groups based on certain features present (Supplementary Table B and C), and these were used to predict PFS and CSS. Risk groups predicting death for chrRCC were not available due to the low number of events observed for death.

Prior to uptake of new prognostic models into routine clinical practice, validation of accuracy and applicability is essential [11–12], and several existing prognostic models and Normogram have been extensively studied and validated externally in independent populations [13–16]. This updated Leibovich model is of great interest given strong predictive ability on internal validation [10], but appropriate external validation has yet to be done. This study seeks to evaluate its performance in an independent population of patients from a single institution in Singapore.

2. Patients and methods

2.1. Study population

The RCC database in Singapore General Hospital was retrospectively reviewed for binephric patients with sporadic, unilateral RCC treated with radical or partial nephrectomy from 1990 to 2015. Ethics approval was obtained from domain-specific review board before commencement of the study (CIRB Reference number: 2009/763/D).

Variables evaluated were based on the original study [10], and included age at time of surgery, gender, race, smoking history, and performance status according to both the Eastern Cooperative Oncology Group criteria and the Charlson Co-morbidity score. Patients were considered to be symptomatic preoperatively if they demonstrated abdominal pain, palpable mass, gross haematuria, or constitutional symptoms (night sweats, loss of appetite/early satiety, and loss of weight). Laboratory findings that were available were included, including pre- and postoperative full blood count findings and estimated glomerular filtration rate. Like the original study, serum calcium levels and liver function tests were excluded from analysis due to significant missing data. Pathologic findings analysed included tumor size (recorded as largest diameter in cm), renal sinus invasion, perinephric fat invasion, extension beyond Gerota's fascia, presence of renal vein invasion/tumor thrombus, lymph node status (if applicable), tumor grade (based on 2016 World Health Organization/International Society of Urological Pathology grade), histologic subtype, presence of sarcomatoid/rhabdoid differentiation, tumor necrosis, and presence of positive surgical margins. Tumor thrombus levels as defined by Neves and Zinke levels were analyzed as no thrombus versus level 0 versus levels I-IV.

Patients with bilateral synchronous tumors, genetic causes of RCC including von Hippel–Lindau syndrome, presence of metastases (regional/nonregional lymph node involvement or distant metastases), or incomplete medical records were excluded from this study. In addition, similar to the original Mayo Clinic study [10], patients with RCC of other histology including collecting duct RCC, acquired cystic disease-associated RCC, mixed clear cell/papRCC and other rare histologic subtypes were excluded as well.

All patients were followed up until death or the end of study on July 1, 2017, whichever occurred first. The main outcomes were CSS and PFS, with progression being defined as either presence of tumor in the ipsilateral kidney/nephrectomy bed, contralateral kidney, or metastatic spread. Cause of death was determined with reference to that registered in the Singapore Death Registry from the Ministry of Home Affairs.

2.2. Statistical analysis

Baseline differences in clinicopathologic features between our local validation population and the Mayo clinic series were evaluated with 2-sample *t*-test, Pearson's chi-square test, or Fisher Exact Test where feasible.

For the purposes of validation, the prognostic index of the original updated Leibovich score for both progression and death was applied to our validation dataset. As presented in the original study, scores were calculated for ccRCC cases without cases being pooled together into risk groups, while papRCC cases were categorized into risk groups 1/2/3. Due to a small number of diagnosed chrRCC in our population ($n = 30$) and lack of events (progression/cancer-related death), analysis was not possible. Survival estimates were performed with the Kaplan–Meier method, and log-rank test was used to detect differences in the survival curves between individual scores and different risk groups. A multivariate Cox proportional hazards model was also used to detect differences between risk scores for ccRCC and risk groups for papRCC.

The performance of the updated Leibovich score was assessed in 2 aspects: discrimination and calibration. Discrimination, the ability of the model to distinguish between a patient who suffered an outcome (progression/death) from a patient who has not, was evaluated with receiver operating characteristic (ROC) curves and the associated concordance index (c-index, equal to area under the curve). An c-index of 0.5 suggests no discriminative ability, while a c-index of 1.0 represents perfect discrimination, with of >0.8 indicating sufficient prognostic accuracy.

Calibration refers to the model's ability to demonstrate agreement between predicted and observed outcomes, and was assessed in a calibration plot between observed outcomes and predicted risks. Calibration-in-the-large,

suggesting mean calibration, was estimated from the intercept of the regression curve, while the calibration slope suggesting magnitude of miscalibration was determined by the regression slope of the linear predictor. A model is perfectly calibrated in the validation dataset if it forms a diagonal line, with an intercept (calibration-in-the-large) of 0 and calibration slope of 1.

All *P* values were 2-sided and considered significant if ≤0.05. All statistical analysis was performed using IBM SPSS v.24 (IBM Corp., Armonk, NY).

3. Results

3.1. Study population

A total of 942 patients with nonmetastatic RCC were included for validation of the updated Leibovich score. Eight hundred and twenty nine (88%) were noted on pathology to have (ccRCC), and 113 (12%) had (papRCC). The clinicopathologic features of the patients included are summarized in Table 1, where it was noted that several variables showed a significant difference when compared to the Mayo clinic dataset used in the updated Leibovich score, including demographic features like median age, race, and smoking history, surgical approach, as well as pathological factors with smaller, more well-differentiated T1 tumors seen in our population.

3.2. Clear cell RCC (ccRCC)

The median length of follow-up for patients with ccRCC after nephrectomy was 76 (42−117) months. 134 (16.2%) patients experienced progression, and a total of 195 (23.5%) patients died, of which 99 (11.9%) died of RCC. As shown in Table 2, estimated PFS rates for all patients at 5, 10, and 15 years were 88% [95% confidence interval (CI) 86.0−90.0], 77% (95% CI 73.1−80.9), and 57% (95% CI 45.2−68.8), and estimated CSS rates were 90% (95% CI 88.0−92.0), 83% (95% CI 79.1−86.9), and 72% (95% CI 62.3−81.8).

Fig. 1 depicts Kaplan−Meier curves for ccRCC risk groups. Calculation of hazard ratios across the risk scores showed that scores of 6 and above were significantly associated with a higher risk of progression in our validation population, and scores of 11 and above are significantly associated with a higher risk of death (Table 3). No patients in our validation dataset had a ccRCC score of 0 for CSS, so survival estimates were not included for this score.

ROC curves are presented in Fig 3A and B. The c-index was 0.81 (SE 0.02, 95% CI 0.77−0.85) for PFS and 0.83 (SE 0.02, 95% CI 0.79−0.87) for CSS, which suggests good discriminative ability. The calibration plot (Fig 4), suggested close to ideal calibration for both PFS and CSS, with calibration-in-the-large of 0.002 and calibration slope of 0.99 in the former, and calibration-in-the-large of 0.005 and calibration slope of 0.96 in the latter.

3.3. Papillary RCC (papRCC)

Median length of follow-up was 69.5 (33−115) months. At the end of follow-up, 21 (18.6%) patients progressed while a total of 33 (29.2%) patients died, of which 16 (14.2%) were from RCC. Estimated PFS rates for all patients at 5, 10, and 15 years were 83% (95% CI 75.2−90.8), 77% (95% CI 67.2−86.8), and 51% (95% CI 29.4−72.6), and estimated CSS rates were 85% (95% CI 77.2−92.8), 79% (95% CI 69.2−88.8), and 79% (95% CI 69.2−88.8), as shown in Table 2.

Kaplan−Meier curves are depicted in Fig. 2. Log-rank tests suggested a significant difference between the risk groups for both PFS and CSS, and calculation of hazard ratios in Table 3 showed that higher risk groups were significantly associated with a higher risk of progression and mortality.

ROC curves for this validation dataset are presented in Fig. 5, and c-index calculated was 0.72 (SE 0.06, 95% CI 0.59−0.85) for PFS and 0.74 (SE 0.07, 95% CI 0.60−0.87) for CSS, suggesting reasonable discriminative ability in our cohort. Calibration plots (Fig. 6) for PFS showed an estimated calibration-in-the-large of −0.09 and calibration slope of 1.93, which deviates from the ideal slope of 1. This deviation was more prominent at higher predicted risks of progression, where observed cancer progression was greater than that predicted by the model. For CSS, the model appears well-calibrated to our population, with calibration-in-the-large being 0.002 and a calibration slope of 0.98.

4. Discussion

A number of prognostic models have been developed and externally validated for prognostication of survival in patients with RCC. The Leibovich score was originally developed in 2003 [6] and externally validated in several different populations [17−19]. An updated prognostic model was recently published [10]. To the best of our knowledge, this is the first external validation based on a relatively large independent dataset of patients with RCC. Our results suggest that the models for nonmetastatic post-nephrectomy ccRCC provide excellent estimates for both progression and death, even in a population of Asian patients with dissimilar characteristics. However, limitations are observed in the prognostic model for papRCC, and should be kept in mind for clinical application.

The model for ccRCC retains good discrimination when applied to our validation cohort compared to the Mayo Clinic derivation cohort [10], and had either a similar or better performance compared to other prediction models including the SSIGN Score [3], UCLA-Integrated Scoring System [5], original Leibovich Score [6], as well as Kattan, Karakiewicz, and Cindolo Nomograms [16]. This improved accuracy could be attributed to its comprehensiveness in including a large number of variables, although this might render it less practical in clinical practice due to the large

Table 1
Comparison of clinicopathologic data for ccRCC and papRCC between the validation study population and the Leibovich series.

	ccRCC			papRCC		
	Present series (*N* = 829)	Leibovich series *(N = 2726)*	*P value*	Present series (*N* = 113)	Leibovich series *(N = 607)*	*P value*
Age at surgery, y						
Median (range)	58 (50−67)	63 (54−71)	<0.001	61 (52−68)	65 (58−73)	<0.001
Gender (%)						
Male	529 (64)	1795 (65)	0.297	79 (70)	493 (81)	0.01
Female	300 (36)	931 (35)		34 (30)	114 (19)	
Race, no. (%)						
Chinese	696 (84)	12 (1)	<0.001	92 (81)	4 (1)	<0.001
Malay	52 (6)	0		10 (9)	0	
Indian	48 (6)	19 (1)		5 (5)	12 (2)	
Others	33 (4)	2389 (98)		6 (5)	527 (97)	
Smoking history, no. (%) (*N = 782:109*)[a]						
Nonsmoker	548 (70)	1116 (42)	<0.001	75 (69)	217 (36)	<0.001
Current	131 (17)	496 (18)		15 (14)	113 (19)	
Former	103 (13)	1073 (40)		19 (17)	267 (45)	
Median BMI, kg/m^2 (range) (*N* = 720:101)	25 (22−27)	28 (25-33)	<0.001	24 (22−26)	28 (25−31)	<0.001
Median Charlson score (range) (*N* = 818:113)	1 (1−2)	1 (0-2)	1.00	2 (1-3)	1 (0-2)	<0.001
ECOG performance status (%) (*N* = 784:111)						
0	677 (86)	2349 (86)	0.009	91 (82)	547 (90)	0.045
1	95 (12)	271 (10)		18 (16)	49 (8)	
2	8 (1)	65 (2)		2 (2)	8 (1)	
3	3 (<1)	38 (1)		0 (0)	3 (<1)	
4	1 (<1)	3 (<1)		0 (0)	0	
Median eGFR, ml/min/1.73m^2 (range) (*N = 779:107*)	74 (55−96)	65 (53-78)	<0.001	61 (40-86)	65 (50-74)	<0.001
Median haemoglobin, g/dl (*N = 816:113*)	13.5 (12.1−14.7)	13.8 (12.4-14.9)	<0.001	12.6 (11.6-14.4)	14.0 (12.9-15.0)	<0.001
Symptoms at presentation, no. (%)	748 (90)	1524 (56)	<0.001	98 (87)	260 (43)	<0.001
Constitutional symptoms	116 (14)	591 (22)	<0.001	14 (12)	72 (12)	0.88
Surgical factors						
Surgical approach (%)						
Open radical	192 (23)	1757 (64)	<0.001	29 (26)	252 (42)	<0.001
Open partial	123 (15)	692 (25)		16 (14)	278 (46)	
Laparoscopic radical	387 (47)	174 (6)		51 (45)	24 (4)	
Laparoscopic partial	127 (15)	103 (4)		17 (15)	53 (9)	
Concomitant organ resection (%)						
Adrenalectomy	321 (39)	959 (35)	0.06	46 (41)	134 (22)	<0.001
Splenectomy	2 (<1)	52 (2)	0.001	0 (0)	3 (<1)	1.00
Lymphadenectomy (%) (*N = 828:113)*	55 (7)	638 (23)	<0.001	9 (8)	57 (9)	0.22
Pathological factors						
Positive surgical margins (%) (*N = 757:108)*	34 (4)	61 (2)	0.001	1 (<1)	16 (3)	0.50
Tumor grade (%) (*N = 828:113)*						
1	87 (10)	257 (9)	<0.001	6 (5)	12 (2)	0.03
2	484 (59)	1255 (46)		58 (51)	357 (59)	
3	205 (25)	1005 (37)		44 (39)	228 (38)	
4	52 (6)	209 (8)		5 (5)	10 (2)	
Tumor necrosis (%)	128 (15)	639 (23)	<0.001	31 (27)	246 (41)	<0.001
Sarcomatoid differentiation (%)	18 (2.2)	70 (3)	0.61	4 (3.5)	7 (1)	0.08
Rhabdoid differentiation (%)	8 (1)	65 (2)	0.01	0		0
2010 pT-classification (%)						
pT1a	374 (45)	937 (35)	<0.001	46 (41)	333 (55)	0.007
pT1b	189 (23)	650 (24)		23 (20)	146 (24)	
pT2a	62 (7)	272 (10)		13 (12)	41 (7)	
pT2b	24 (3)	111 (4)		11 (10)	31 (5)	
pT3a	156 (19)	533 (20)	17 (15)	44 (7)		
pT3b	11 (1)	153 (6)		2 (2)	8 (1)	
pT3c	4 (1)	23 (1)		0 (0)	1 (<1)	
pT4	9 (1)	35 (1)		1 (<1)	1 (<1)	

(*continued*)

Table 1 (***Continued***)

	ccRCC			papRCC		
	Present series (*N* = 829)	Leibovich series *(N = 2726)*	*P value*	Present series (*N* = 113)	Leibovich series *(N = 607)*	*P value*
Pathologic tumor size, cm (%) (*N = 828:113)*						
Median (range)	4.3 (2.9−6.5)	5.4 (3.4−8.5)	<0.001	4.5 (3.0−8.0)	3.8 (2.5-6.0)	<0.001
≤4.0	404 (49)	976 (36)	<0.001	52 (46)	342 (57)	0.004
>4.0 to ≤7.0	259 (31)	820 (30)		27 (24)	167 (28)	
>7.0 to ≤10.0	106 (13)	552 (20)		20 (18)	55 (9)	
>10.0	59 (7)	364 (13)		14 (12)	40 (7)	
Perinephric fat invasion (%) (*N = 827:113*)	138 (17)	541 (20)	0.04	15 (13)	43 (7)	0.04
Tumor thrombus (%)						
No	755 (91)	2248 (83)	<0.001	107 (94)	588 (97)	0.429
Level 0	56 (7)	281 (10)		3 (3)	8 (1)	
≥ Level 1	18 (2)	185 (7)		3 (3)	10 (2)	
Extension beyond the kidney (%)	9 (1)	35 (1)	0.64	1 (<1)	1 (<1)	1
Nodal involvement (%)						
No nodal dissection (pNx)	740 (89)	2088 (77)	<0.001	103 (92)	550 (91)	0.268
No nodal involvement (pN0)	82 (10)	533 (20)		5 (4)	43 (7)	
Nodal involvement (pN1)	7 (1)	105 (4)		5 (4)	14 (2)	

ccRCC = clear cell renal cell carcinoma, papRCC = papillary renal cell carcinoma

BMI = body mass index; ECOG = Eastern Cooperative Oncology Group; eGFR = estimated glomerular filtration rate.

[a] For variables with missing data, the number stated in parentheses ($N = x_1: x_2$) refer to the number of nonmissing patients with ccRCC as the former and papRCC as the latter.

Table 2
Progression-free and cancer-specific survival at 5-, 10- and 15-years after surgery, predicted by the Leibovich scoring/risk group system for clear cell carcinoma (ccRCC) and papillary RCC (papRCC)

	Progression-free survival (SE[a])				Cancer-specific survival (SE[a])			
	N (%)	5 y	10 y	15 y	*N (%)*	5 y	10 y	15 y
ccRCC score								
0	40 (4.8)	100 (-)	100 (0)	82 (12)	-	-	-	-
1	10 (1.2)	100 (-)	100 (0)	100 (0)	29 (3.5)	100 (0)	100 (0)	87 (12)
2	189 (22.7)	99 (1)	97 (2)	97 (2)	18 (2.2)	100 (0)	100 (0)	100 (0)
3	101 (12.2)	94 (2)	83 (7)	83 (7)	104 (12.5)	100 (0)	100 (0)	100 (0)
4	28 (3.4)	88 (7)	88 (7)	44 (31)	113 (13.6)	99 (1)	96 (3)	96 (3)
5	133 (15.9)	91 (4)	87 (5)	87 (5)	72 (8.7)	97 (3)	97 (3)	97 (3)
6	119 (14.4)	75 (5)	60 (8)	45 (14)	51 (6.2)	91 (4)	91 (4)	91 (4)
7	68 (8.2)	69 (7)	51 (10) 51	(10)	54 (6.6)	96 (3)	96 (3)	96 (3)
8	54 (6.5)	62 (8)	49 (10)	25 (18)	90 (10.9)	93 (3)	83 (6)	50 (26)
9	35 (4.2)	59 (10)	28 (11)	28 (11)	81 (9.8)	86 (5)	70 (9)	70 (9)
10	25 (3.0)	56 (12)	33 (19)	16 (18)	60 (7.2)	82 (6)	63 (11)	21 (25)
11	12 (1.4)	27 (14)	27 (14)	0 (0)	53 (6.4)	57 (9)	48 (11)	48 (11)
12	8 (1.0)	25 (15)	25 (15)	0 (0)	41 (4.9)	57 (11)	57 (11)	57 (11)
13	4 (0.5)	67 (27)	0 (0)	0 (0)	18 (2.2)	69 (12)	69 (12)	69 (12)
14	2 (0.2)	0 (0)	0 (0)	0 (0)	13 (1.6)	39 (14)	39 (14)	38 (15)
>15[b]	3 (0.4)	0 (0)	0 (0)	0 (0)	11 (1.3)	36 (16)	36 (16)	37 (16)
16					8 (1.0)	54 (20)	0 (0)	0 (0)
17					6 (0.7)	33 (19)	0 (0)	0 (0)
>18[b]					6 (0.7)	17 (15)	0 (0)	0 (0)
papRCC group								
1	59 (52)	94 (3)	94 (3)	62 (20)	59 (52)	94 (3)	94 (3)	94 (3)
2	31 (27)	71 (12)	55 (17)	55 (17)	31 (27)	75 (11)	58 (17)	58 (17)
3	23 (21)	60 (11)	52 (12)	52 (12)	23 (21)	65 (11)	56 (12)	56 (12)

ccRCC = clear cell renal cell carcinoma, papRCC = papillary renal cell carcinoma

[a] Standard Error (SE)

[b] Similar to the original study, Scores above >15 were collapsed into a single group for PFS, and scores above >18 were collapsed into a single group for CSS.

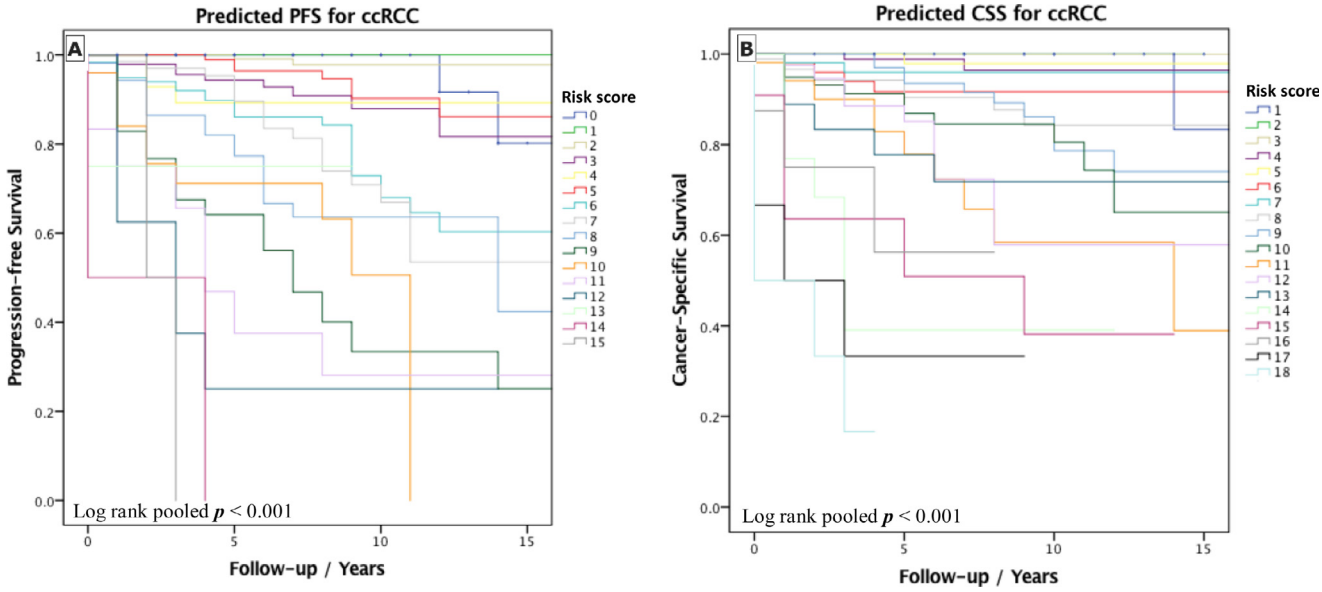

Fig. 1. Kaplan−Meier Curves showing (A) predicted PFS and (B) predicted CSS for ccRCC risk scores. Abbreviations: ccRCC = clear cell renal cell carcinoma; CSS = cancer-specific survival; PFS = progression-free survival.

amount of data required for each patient. Nevertheless, most of the information required for scoring is readily available as routinely obtained clinicopathologic information, and operationalization of the score as an online risk calculator as proposed by the authors would significantly improve ease of use.

We noted suboptimal performance of the papRCC prognostic model−although there was reasonable discrimination (c-index of 0.72 for PFS and 0.74 for CSS), it was not as strong compared to the derivation cohort from Mayo Clinic (c-index 0.77 for PFS and 0.83 for CSS). Additionally, the model prognosticating PFS was not calibrated to our

Table 3
Hazard ratios across different risk scores/groups for predicted by the Mayo Clinic scoring/risk group system for ccRCC and papRCC based on Cox proportional hazards

	Progression			Cancer-specific mortality		
	N (%)	Hazard ratio (95% CI)	*P value*	*N (%)*	Hazard ratio (95% CI)	*P value*
ccRCC score						
0	40 (4.8)		1.00		-	-
1	10 (1.2)	0		0.96	29 (3.5)	1.00
2	189 (22.7)	0.3 (0.1−2.1)	0.22	18 (2.2)	0	0.98
3	101 (12.2)	2.2 (0.5−10.3)	0.31	104 (12.5)	0	0.96
4	28 (3.4)	2.8 (0.5−15.7)	0.23	113 (13.6)	0.7 (0.1−7.7)	0.77
5	133 (15.9)	1.3 (0.3−6.1)	0.76	72 (8.7)	0.4 (0.1−6.8)	0.54
6	119 (14.4)	5.4 (1.3−22.9)	**0.02**	51 (6.2)	2.4 (0.3−21.8)	0.43
7	68 (8.2)	6.0 (1.4−25.9)	**0.02**	54 (6.6)	1.0 (0.1−11.3)	0.99
8	54 (6.5)	8.2 (1.9−35.4)	**<0.01**	90 (10.9)	4.1 (0.5−32.3)	0.18
9	35 (4.2)	16.6 (3.9−71.3)	**<0.01**	81 (9.8)	3.7 (0.5−29.2)	0.21
10	25 (3.0)	13.2 (2.9−60.5)	**<0.01**	60 (7.2)	6.0 (0.8−46.5)	0.09
11	12 (1.4)	23.5 (5.1−108.8)	**<0.01**	53 (6.4)	12.0 (1.6−90.2)	**0.02**
12	8 (1.0)	33.4 (6.7−166.2)	**<0.01**	41 (4.9)	10.5 (1.3−82.1)	**0.03**
13	4 (0.5)	13.1 (1.18−144.9)	**0.04**	18 (2.2)	8.7 (1.0−74.6)	**0.04**
14	2 (0.2)	91.2 (12.7−655.6)	**<0.01**	13 (1.6)	31.8 (3.9−259.3)	**<0.01**
>15	3 (0.4)	76.8 (10.6−556.0)	**<0.01**	11 (1.3)	24.2 (2.9−201.2)	**<0.01**
16				8 (1.0)	23.7 (2.4−229.4)	**<0.01**
17				6 (0.7)	48.8 (5.4−439.7)	**<0.01**
>18				6 (0.7)	114.1 (13.1−991.7)	**<0.01**
papRCC group						
1	58 (51)	1.00		58 (51)	1.00	
2	31 (28)	3.5 (1.1−11.6)	**0.04**	31 (28)	4.4 (1.0−18.4)	**0.04**
3	23 (21)	8.3 (2.7−25.3)	**<0.01**	23 (21)	9.0 (2.3−34.0)	**<0.01**

Abbreviation: ccRCC = clear cell renal cell carcinoma; CI = Confidence Interval; papRCC = papillary renal cell carcinoma

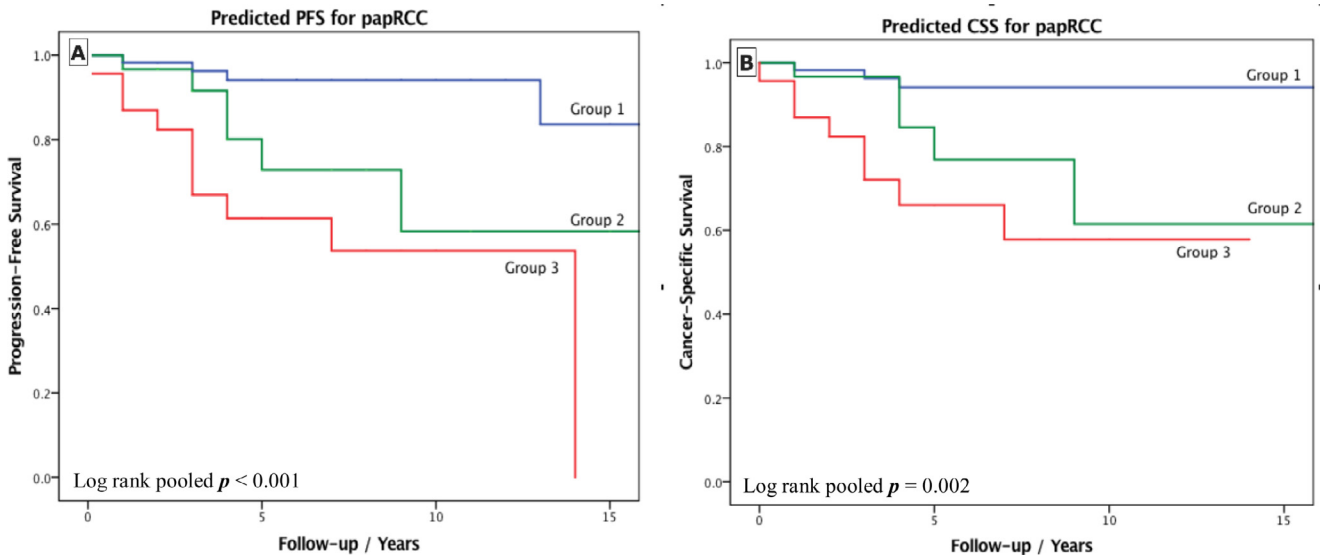

Fig. 2. Kaplan–Meier curves showing (A) predicted PFS and (B) predicted CSS for papRCC risk groups. Abbreviations: CSS = cancer-specific survival; papRCC = papillary renal cell carcinoma; PFS = progression-free survival.

population, with deviation from the ideal line at higher predicted risks suggesting that the number of observed relapses were higher than predicted by papRCC risk group 3. One possible explanation could be the differences in strength of association between certain variables and outcome –papRCC risk group 3 is diverse and includes patients with either Fuhrman grade 4, perinephric/renal sinus fat invasion or any tumor thrombus. Consequently, a greater proportion of our patients (21%) were classified under this group compared to the original study (9%) [10]. When we attempted to investigate the variables in risk group 3, we found that tumor thrombus was no longer significantly associated with both PFS/CSS on multivariate analysis, and unlike the original study, weaker associations were seen between fat invasion and oncologic outcomes (Supplementary Table D). papRCC is widely accepted as a highly heterogeneous disease with inherent variations in both disease progression and outcome [20], and current literature examining associations between prognostic variables and outcomes in papRCC remains mixed–Pichler et al. [21] observed a significantly higher risk of PFS and CSS with male gender, T-stage, tumor grade, tumor necrosis, and vascular invasion; Méjean et al. [22] noted associations of CSS with higher tumor stage and grade, while Ha et al. [23] only found associations between tumor stage and both PFS/CSS. To address differences in association

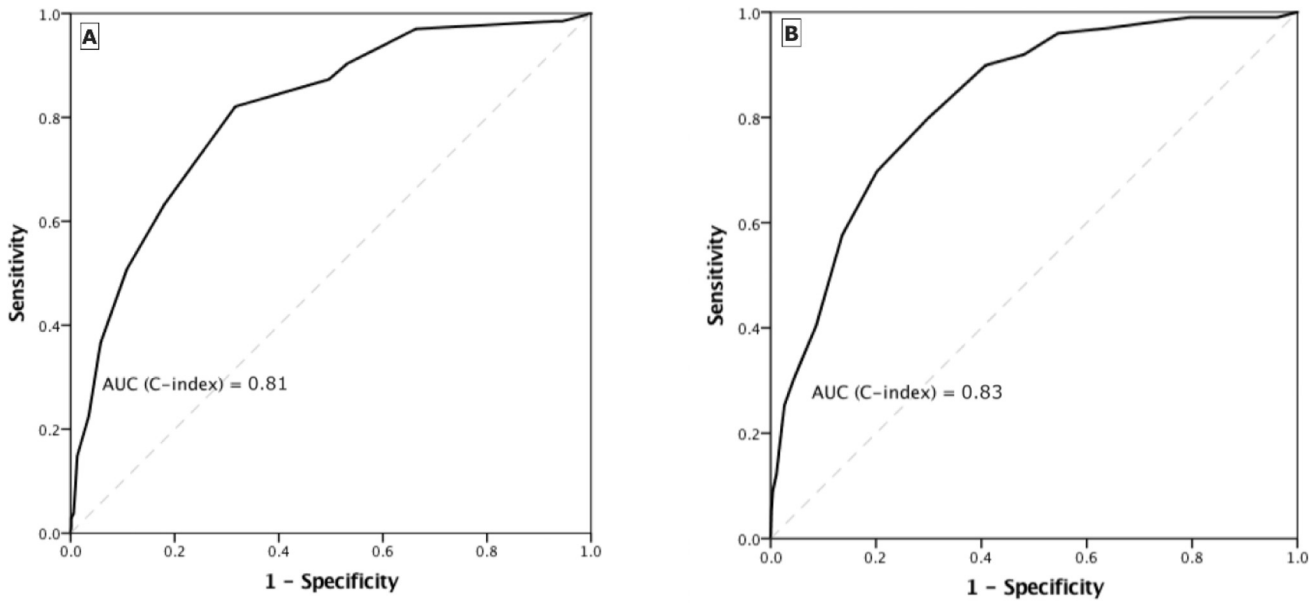

Fig. 3. ROC curves for prognostic models predicting PFS (A) and CSS (B) in ccRCC. Abbreviations: ccRCC = clear cell renal cell carcinoma; CSS = cancer-specific survival; PFS = progression-free survival; ROC = receiver operating characteristic.

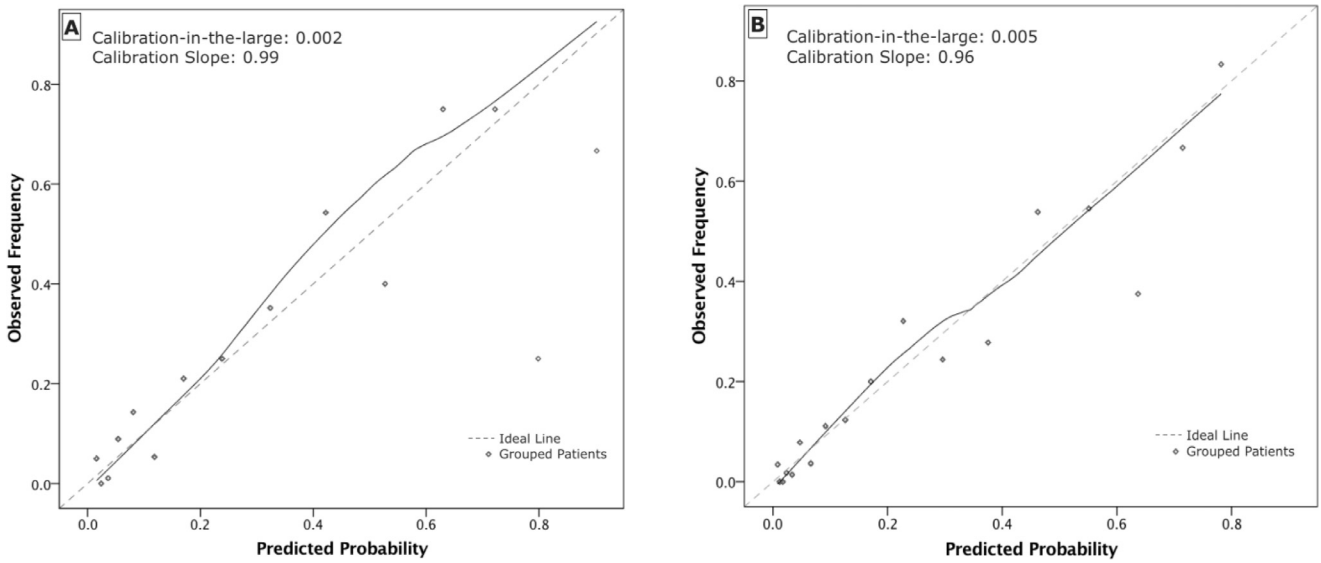

Fig. 4. Calibration plots for prognostic models predicting PFS (A) and CSS (B) in ccRCC. Abbreviations: ccRCC = clear cell renal cell carcinoma; CSS = cancer-specific survival; PFS = progression-free survival; ROC = receiver operating characteristic.

of variables with outcomes, it can be recommended that for papRCC, rather than grouping based on presence/absence of features, risk scores similar to the ccRCC model could be developed instead, with scores designated based on strength of association between the variable and the outcome (either progression or death). Risk scores can subsequently be pooled together into stratified risk groups if applicable, like the original Leibovich score [6].

We were unable to include chrRCC due to the small number diagnosed in our population ($n = 30$), and small number of events (3 cases of progression and 3 cases of cancer-specific death) even for data collected over 25 years, which precludes effective analysis. There is a paucity of studies evaluating oncological outcomes in a significant number of chrRCC patients, and the clinicopathological factors associated with survival appear to vary. Volpe et al. [24] analyzed 291 chrRCC patients and found significant associations between gender, T-stage, N-stage, and sarcomatoid change with PFS and CSS; Przybycin et al. [25] similarly looked at 203 chrRCC patients and found associations between tumor size, necrosis, and sarcomatoid change with recurrence and metastasis, while Xie et al. [26] proposed that tumor size, T-stage, and tumor grade were predictors of disease-free survival. These factors are largely similar to those included by Leibovich into the chrRCC

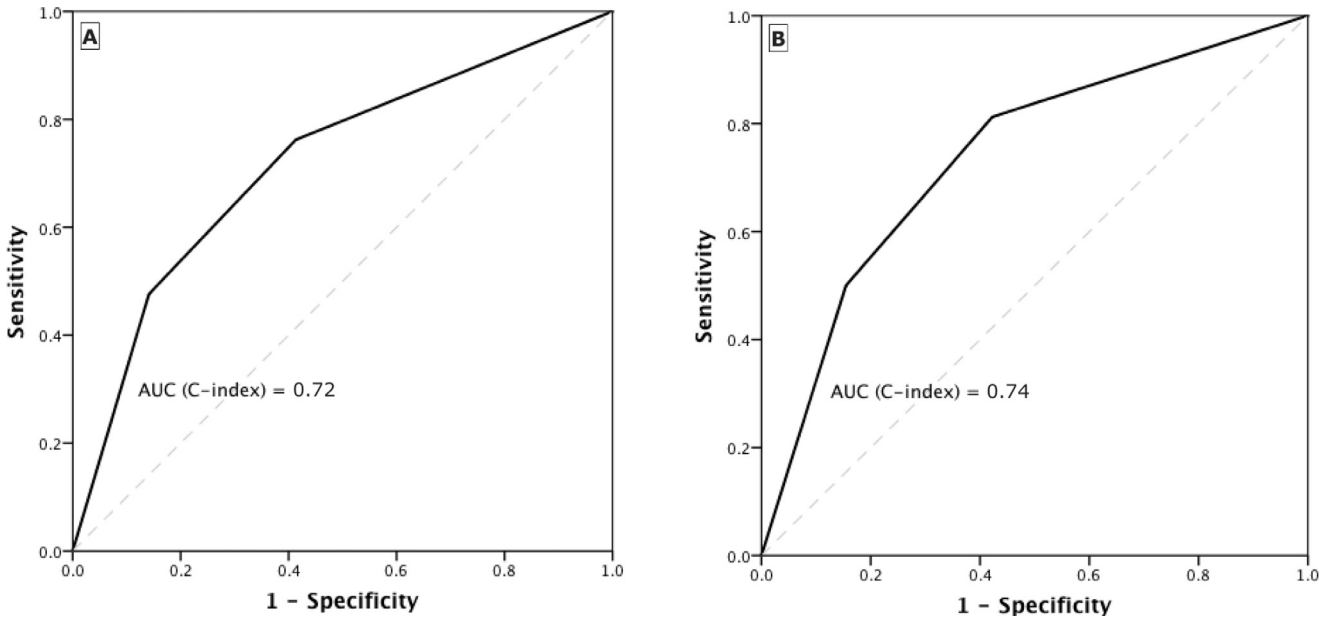

Fig. 5. ROC curves for prognostic models predicting PFS (A) and CSS (B) in papRCC. Abbreviations: CSS = cancer-specific survival; papRCC = papillary renal cell carcinoma; PFS = progression-free survival.

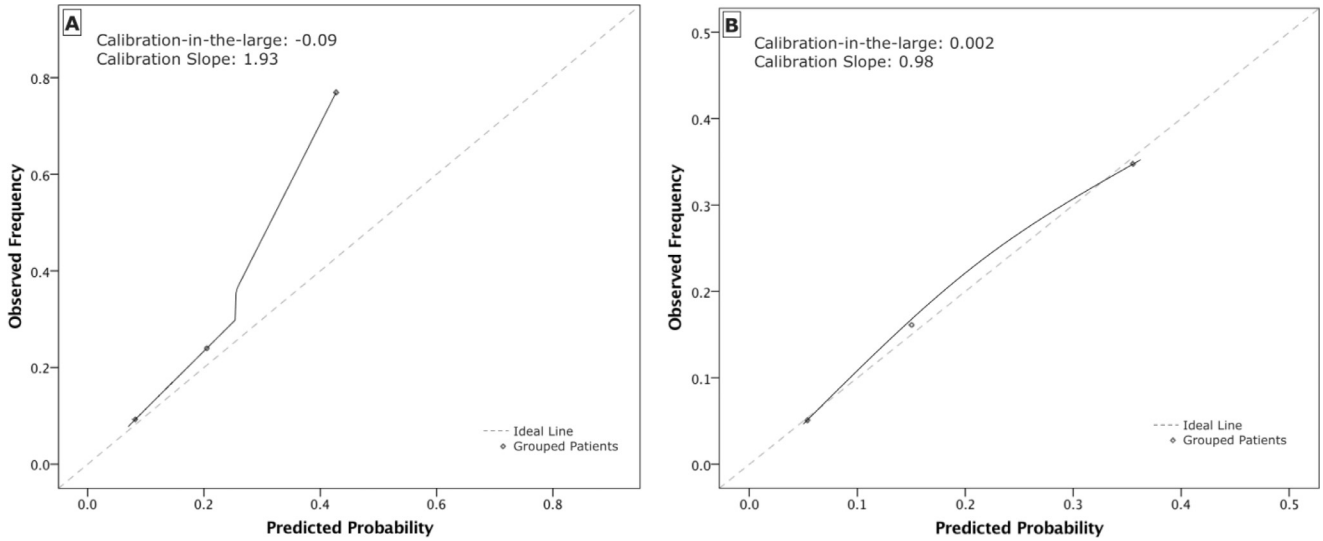

Fig. 6. Calibration plots for prognostic models predicting PFS (A) and CSS (B) in papRCC. Abbreviations: CSS = cancer-specific survival; papRCC = papillary renal cell carcinoma; PFS = progression-free survival.

risk groups [10], which might imply clinical utility. Nevertheless, external validation by other studies with a substantial population of chrRCC patients is still essential prior to implementation.

There are a few notable limitations to our study: First, this external validation is based on retrospective data from a single institution, which is susceptible to bias and usual concerns regarding generalizability to other populations. The updated Leibovich model [10] was developed based on a population of patients in Mayo Clinic, a single institution in the United States well-known for high-quality care, and concerns of limited applicability in other hospitals or patient populations in another geographical area is understandable. Furthermore, as our population was selected in a more contemporary timeframe (1990−2015) compared to the original study (1980−2010), RCC stage migration could be another possible contributing factor. Compared with the derivation cohort, our validation cohort also showed significant differences in multiple clinicopathological factors, including tumor characteristics, and surgical technique with more laparoscopic approaches used. This is unsurprising given the contrasting clinical context with a population of predominantly Asian patients. Despite these differences, the prognostic models proposed by Leibovich generally performed well in our population, which demonstrates its applicability in other centers.

Overall, our results demonstrate that the updated Leibovich score is applicable for prognostication of progression and death even in an independent population of Asian patients, for both ccRCC and papRCC. However, further validation is required to confirm applicability of the papRCC score in determining PFS. In order to ensure that the Leibovich prognostic model is robust, further external validation and evaluation using pooled multi-institutional data is encouraged.

Disclosures

All authors involved in this study declare that they have no conflict of interest. Ethics approval was obtained from domain-specific review board before commencement of the study (CIRB Reference number: 2009/763/D), and written informed consent was obtained from all patients prior to partial/radical nephrectomy. No funding or financial incentives were received for this study.

Supplementary materials

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.urolonc.2019.02.014.

References

[1] Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–86.
[2] Stewart SB, Thompson RH, Psutka SP, et al. Evaluation of the national comprehensive cancer network and American Urological Association renal cell carcinoma surveillance guidelines. J Clin Oncol 2014;32:4059–65.
[3] Frank I, Blute ML, Cheville JC, et al. An outcome prediction model for patients with clear cell renal cell carci- noma treated with radical nephrectomy based on tumor stage, size, grade and necrosis: The SSIGN score. J Urol 2002;168:2395–400.
[4] Karakiewicz PI, Suardi N, Capitanio U, et al. Multi-institutional validation of a new renal cancer-specific survival nomogram. J Clin Oncol 2007;25:1316.
[5] Patard JJ, Kim HL, Lam JS. Use of the University of California Los Angeles integrated staging system to predict survival in renal cell carcinoma: An international multicenter study. J Clin Oncol 2004;22:3316–22.
[6] Leibovich BC, Blute ML, Cheville JC, et al. Prediction of progression after radical nephrectomy for patients with clear cell renal cell

carcinoma: A stratification tool for prospective clinical trials. Cancer 2003;97:1663–71.
[7] Janowitz T, Welsh SJ, Zaki K, et al. Adjuvant therapy in renal cell carcinoma — past, present, and future. Semin Oncol 2013;40:482–91.
[8] Ravaud A, Motzer RJ, Pandha HS, et al. Adjuvant sunitinib in high-risk renal-cell carcinoma after nephrectomy. N Engl J Med 2016;375:2246–54.
[9] Haas NB, Manola J, Uzzo RG, et al. Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): A double-blind, placebo-controlled, randomised, phase 3 trial. Lancet 2016;387:2008–16.
[10] Leibovich BC, Lohse CM, Cheville JC, et al. Predicting oncologic outcomes in renal cell carcinoma after surgery. Eur Urol 2018;73:772–80.
[11] A Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453–73.
[12] Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: A clinical example. J Clin Epidemiol 2003;56:826–32.
[13] Zigeuner R, Hutterer G, Chromecki T, et al. External validation of the Mayo Clinic stage, size, grade, and necrosis (SSIGN) score for clear-cell renal cell carcinoma in a single European centre applying routine pathology. Eur Urol 2010;57:102–9.
[14] Cindolo L, Chiodini P, Gallo C, et al. Validation by calibration of the UCLA integrated staging system prognostic model for nonmetastatic renal cell carcinoma after nephrectomy. Cancer 2008;113:65–71.
[15] Liu Z, Lv J, Ding K, et al. Validation of the current prognostic models for nonmetastatic renal cell carcinoma after nephrectomy in Chinese population: A 15-year single center experience. IJU 2009;16:268–73.
[16] Russo GI, Di Rosa A, Favilla V, et al. Accuracy capabilities comparisons between Karakiewicz, Kattan and Cindolo nomograms in predicting outcomes for renal cancer carcinoma: A systematic review and meta-analysis. Can Urol Assoc J 2015;9:E359–66.
[17] Pichler M, Hutterer GC, Chromecki TF, et al. External validation of the Leibovich prognosis score for nonmetastatic clear cell renal cell carcinoma at a single European center applying routine pathology. J Urol 2011;186:1773–8.
[18] Tan MH, Kanesvaran R, Li H, et al. Comparison of the UCLA integrated staging system and the Leibovich score in survival prediction for patients with nonmetastatic clear cell renal cell carcinoma. Urology 2010;75:1365–70.
[19] Beisland C, Gudbrandsdottir G, Reisæter LAR, et al. Contemporary external validation of the Leibovich model for prediction of progression after radical surgery for clear cell renal cell carcinoma. Scand J Urol 2014;49:205–10.
[20] Linehan WM, Spellman PT, Ricketts CJ, et al. Comprehensive molecular characterization of papillary renal-cell carcinoma. N Engl J Med 2016;374:135–45.
[21] Pichler M, Hutterer GC, Chromecki TF, et al. Histologic tumor necrosis is an independent prognostic indicator for clear cell and papillary renal cell carcinoma. Am J Clin Pathol 2012;137:283–9.
[22] Méjean A, Hopirtean V, Bazin JP, et al. Prognostic factors for the survival of patients with papillary renal cell carcinoma: Meaning of histological typing and multifocality. J Urol 2003;170:764–7.
[23] Ha YS, Chung JW, Choi SH, et al. Clinical significance of subclassification of papillary renal cell carcinoma: Comparison of clinicopathologic parameters and oncologic outcomes between papillary histologic subtypes 1 and 2 using the Korean renal cell carcinoma database. Clin Genitourin Cancer 2017;15:181–6.
[24] Volpe A, Novara G, Antonelli A, et al. Chromophobe renal cell carcinoma (RCC): Oncological outcomes and prognostic factors in a large multicentre series. BJUI 2011;110:76–83.
[25] Przybycin CG, Cronin AM, Darvishian F, et al. Chromophobe renal cell carcinoma: A clinicopathologic study of 203 tumors in 200 patients with primary resection at a single institution. Am J Surg Pathol 2011;35:962–70.
[26] Xie Y, Ma X, Li H, et al. Prognostic value of clinical and pathological features in Chinese patients with chromophobe renal cell carcinoma: A 10-year single-center study. J Cancer 2017;8:3474–9.