



Published in final edited form as:

J Urol. 2016 February ; 195(2): 283–289. doi:10.1016/j.juro.2015.08.093.

External Validation of Bladder Cancer Predictive Nomograms for Recurrence, Cancer-Free Survival and Overall Survival following Radical Cystectomy

Michael Brooks^{*}, Guilherme Godoy^{*}, Maxine Sun, Shahrokh F. Shariat, Gilad E. Amiel, and Seth P. Lerner[†]

Scott Department of Urology, Baylor College of Medicine (MB, GG, SPL), Houston, Texas, University of Montreal Health Center (MS), Montreal, Québec, Canada, Center for Surgery and Public Health, Brigham and Women's Hospital, Harvard School of Medicine (MS), Boston, Massachusetts, Bladder Cancer Research Consortium (SFS, GEA, SPL), Department of Urology, Comprehensive Cancer Center, Vienna General Hospital, Medical University of Vienna (SFS), Vienna, Austria, Weil Medical College Cornell University (SFS), New York, New York, and Department of Urology, Rambam Health Care Campus (GEA), Haifa, Israel

Abstract

Purpose—We externally validated 3 previously published nomograms to predicting recurrence, and cancer specific and overall survival following radical cystectomy and pelvic lymph node dissection for urothelial carcinoma of the bladder.

Materials and Methods—Two surgeons from a single institution performed a total of 197 consecutive radical cystectomies and pelvic lymph node dissections for bladder cancer from January 2003 to September 2009. A total of 23 patients were excluded from analysis. Examined parameters were those used in the original nomograms, including patient age, gender, pathological T stage, N stage, tumor grade, presence of carcinoma in situ and lymphovascular invasion, neoadjuvant chemotherapy, adjuvant chemotherapy and adjuvant radiation therapy. Nomogram predictions were compared to actuarial outcomes and predictive accuracy was quantified using measures of discrimination and calibration.

Results—At the time of analysis 34 patients had experienced recurrence, of whom 28 died of disease and 6 were currently alive with disease. Discrimination at 2, 5 and 8 years was 0.776, 0.809 and 0.794 for recurrence, 0.822, 0.840 and 0.849 for cancer specific survival, and 0.812, 0.820 and 0.825, respectively, for overall survival. Calibration plots revealed nomogram overestimation of all 3 end points.

Conclusions—Nomograms for bladder cancer recurrence, cancer specific survival and overall survival following radical cystectomy and pelvic lymph node dissection performed well in our

[†]Correspondence: Scott Department of Urology, Baylor College of Medicine, 7200 Cambridge St., Suite 10B, Houston, Texas 77030 (telephone: 713-798-6841; FAX: 713-798-5553; slerner@bcm.edu).

^{*}Equal study contribution.

The corresponding author certifies that, when applicable, a statement(s) has been included in the manuscript documenting institutional review board, ethics committee or ethical review board study approval; principles of Helsinki Declaration were followed in lieu of formal ethics committee approval; institutional animal care and use committee approval; all human subjects provided written informed consent with guarantees of confidentiality; IRB approved protocol number; animal approved project number.

series with accuracy comparable to that in the original series. The use of nomogram predictions should be further explored in clinical trials to assess the impact on patient care in clinical practice.

Keywords

urinary bladder neoplasms; cystectomy; nomograms; validation studies; mortality

The natural history of bladder UC varies considerably from noninvasive low grade disease, which has a high probability of recurrence but low mortality following local treatment, to high grade muscle invasive disease, which has rates of recurrence and mortality that depend on a variety of clinical, pathological and treatment factors. Although treatment modalities continue to evolve, RC with PLND remains the primary treatment for muscle invasive and refractory high grade nonmuscle invasive UC of the bladder.¹ Despite treatment with curative intent oncologic outcomes still vary considerably and recurrence-free survival after RC and PLND at 5 years ranges from 58% to 62%.^{2,3}

As defined by AJCC after RC for UC of the bladder prognostic pathological stages represent the gold standard for outcome prediction and are based on pathological staging and lymph node status.⁴ Longitudinal outcome based studies using AJCC stages alone have demonstrated less than ideal recurrence and mortality predictions due to the relatively low predictive accuracy and significant outcome heterogeneity for individuals in AJCC stages.^{3,5} In contrast, nomograms use a multivariable approach that incorporates significant clinical, pathological and treatment factors to improve overall predictive accuracy and provide individualized risk assessments for each patient.⁶

Multivariable nomograms were previously developed to predict recurrence, and overall and cancer specific survival following RC for bladder UC.^{7,8} These nomograms, which are based on the patient outcomes observed in a single multi-institutional series from March 11, 1984 through January 24, 2003, use patient characteristics including age, pathological T stage and N stage, presence of lymphovascular invasion, presence of carcinoma in situ and use of neoadjuvant chemotherapy, adjuvant chemotherapy and adjuvant external beam radiotherapy. In the study group both nomograms demonstrated statistically significant improvement in predictive accuracy over the accuracy of AJCC predictions.

At the time of the current study the recurrence, overall survival and cancer specific survival nomograms had not been externally validated in the United States. This critical step was required before these predictive models could finally be tested in clinical trials and/or be incorporated into clinical practice for patient treatment.⁶

The objective of the current study was external validation of the previously published nomograms. We hypothesized that in our series of patients with UC of the bladder treated with RC and extended PLND by 2 experienced surgeons using a standardized technique the nomogram predictions of recurrence, and overall and cancer specific survival would be accurate and highly correlate with actual oncological outcomes.

MATERIALS AND METHODS

Study Group

Patient consent was obtained during the initial evaluation and patient data were entered into a prospectively updated institutional review board approved database (Caisis, <http://caisis.org/>). The study group included 197 consecutive patients who underwent RC and PLND performed by 2 experienced urological oncologists at a tertiary academic medical center from January 2003 to September 2009. All patients were treated with curative intent and 7 had pathologically positive margins at cystectomy. All patients underwent RC using a standardized approach as well as bilateral extended PLND, including para-aortic and paracaval lymph nodes above the bifurcation of the aorta up to the level of the inferior mesenteric artery as the proximal limit of dissection. PLND also included the presacral nodes and the bilateral presciatic nodes as well as the common iliac, external iliac, obturator and hypogastric nodes by removing all potential node bearing tissue surrounding those vessels.

Indications for RC and PLND included refractory, high grade, nonmuscle invasive disease and muscle invasive UC of the bladder. Neoadjuvant and adjuvant chemotherapy regimens varied but typically included 4 cycles of methotrexate, vinblastine, doxorubicin and cisplatin or gemcitabine and cisplatin. Salvage external beam radiation was performed in only 3 cases and regimen and radiation dose varied depending on treatment facility. After excluding 23 patients from analysis, including 12 because of nonUC histology, 8 because of lack of sufficient data and 3 because they had undergone salvage RC, the final study group included 174 patients.

Pathology Evaluation

All surgical specimens were evaluated by pathologists with expertise in genitourinary cancer. Pathological staging was performed in accordance with the 2002 AJCC staging system to provide consistency with definitions used in the original nomogram development.

Followup

Patients were typically followed at 4-month intervals for the first year, at 6-month intervals for the second year and annually thereafter. The majority of patients were followed by the treating surgeon. Outcome information was obtained from patients, primary physicians, referring urologists and death certificates. As with the original definitions used for nomogram development patients in whom upper urinary tract or urethral UC developed were defined as having second primary urothelial tumors rather than recurrent disease.^{7,8}

Statistical Analysis

Differences between characteristics of patients in the original series and those of patients in the validation series were assessed using the Pearson chi-square and Fisher exact tests. Actuarial survival curves were estimated using the Kaplan-Meier method. Performance of the nomograms was evaluated with measures of predictive accuracy. 1) Discrimination accuracy was quantified using ROC derived AUC analysis.⁹ In this type of analysis 1.0 represents perfect predictions and 0.5 is the equivalent of a coin toss. Recurrence, death from

cancer and death from any cause were coded as events in the recurrence, the cancer specific survival and the overall survival analyses, respectively. Patients who did not have events were censored at the last followup. 2) Actual and predicted probabilities were compared using methods of calibration. The mentioned steps were repeated on each examined end point.

Statistical analyses were performed using R (<https://www.r-project.org/>) with significance considered at $p < 0.05$. The nomograms are available as an online calculator (www.nomogram.org) where one can enter covariate data and retrieve the specific prediction probabilities for recurrence or survival after RC and PLND.¹⁰

RESULTS

The patient characteristics of our series closely resembled those of the series used for nomogram development (table 1). The proportion of male patients was the only statistically significant difference between the validation series and original nomogram series (88.5% vs 82.4%, $p = 0.05$). The majority of patients had muscle invasive disease and high grade tumors, and 24.1% had positive nodes. Lymphovascular invasion and carcinoma in situ were present in 32.2% and 52.9% of patients, respectively.

Of the patients 59.8% underwent RC and PLND as the only treatment for bladder cancer. NACH was done in 9.2% of patients while ACH was performed in 31.6% (table 2). AXRT was administered in only 1.7% of cases. Median followup in patients alive at the end of the study was 24 months (IQR 11, 48) and median followup in all patients was 23 months (IQR 11, 46).

At the time of analysis 34 patients had experienced recurrence with a median time to recurrence of 9 months (IQR 6, 16) after RC with PLND. In 79.4% of these cases disease recurred within 24 months. Of the 34 patients 28 died of disease and 6 were alive with disease at the time of analysis. Additionally 15 patients died of another cause. There were no perioperative deaths but 1 patient died of complications associated with ACH. Figure 1 shows Kaplan-Meier plots with estimates of the 3 oncologic end points.

ROC-AUC values at 2, 5 and 8 years for recurrence, cancer specific survival and overall survival ranged from 0.776 to 0.849 (table 3). Calibration plots revealed overestimation of outcomes for all 3 end points (fig. 2).

DISCUSSION

A new emphasis on the development of multivariable predictive tools has led to an abundance of predictive models for oncologic outcomes but only a minority have been subjected to external validation.¹¹ While the majority undergo some type of internal validation to eliminate overfitting data used in model development, external validation with separate cohorts represents the only true assessment of model predictive accuracy. Consequently external validation is a crucial step toward integrating predictive models into clinical practice.

This study represents a statistically rigorous external validation of the accuracy of all 3 nomograms to predict the respective oncologic end points. The nomograms performed well with predictive accuracy comparable to that of other robust and accepted models for long-term outcome prediction in genitourinary cancer following curative radical surgery.¹²⁻¹⁴

In the original development of the nomograms internal validation was performed using a bootstrap method. Each nomogram was significantly superior to AJCC staging for the 3 end points, including 0.780 vs 0.748, 0.791 vs 0.630 and 0.732 vs 0.615 for recurrence, and cancer specific and overall survival, respectively.^{7,8}

Recently external validation including a total of 246 patients from 2 German academic centers demonstrated the superiority of the nomograms to AJCC staging alone to predict 5-year end points, including 0.84 vs 0.82, 0.78 vs 0.75 and 0.85 vs 0.82 for recurrence, and overall and cancer specific survival, respectively.¹⁵ Subsequently in a larger study including patients from 8 European academic centers mortality predictions were validated in 2,404 patients and recurrence predictions were validated in 2,243.¹⁶ Similarly discrimination accuracy ranged from 0.68 to 0.75. Neither study group included patients who received NACH. In the current series 9.2% of patients received NACH. Our series also provides additional evidence of accuracy in a patient population in the United States and validates 2-year and 8-year outcome predictions.

In this study end points were overestimated in all 3 nomograms, which could be due to several causes. The original nomograms were derived using cystectomy data extending back to March 11, 1984. Given improvements in detection, treatment and supportive care, it is reasonable to expect that the nomogram predictions would overestimate cancer recurrence, cancer related mortality and overall mortality in a more contemporary series.

Additionally perioperative therapy practices have evolved significantly whereas AXRT has largely been abandoned because of lack of efficacy and level 1 evidence supports NACH for muscle invasive disease.¹⁷⁻¹⁹ Therefore, in our more contemporary series NACH as the standard of care is reflected in these changes (9.2% vs 5.2% in the original series), although it is still quite low during the accrual period. The increased use of NACH in this study may have contributed to the increased proportion of low pT stage disease (pT0, pTa and pTis) compared with the proportion in the original series (29.9% vs 23.4%, $p = 0.07$), although differences were not statistically significant.

Finally, our median followup was 24 months, similar to the median 24.9 months in the original cohort. However, with longer followup one would expect that the additional events might potentially contribute to decrease overestimation. Nonetheless it is known that most oncologic outcomes after RC develop within the first 2 years after the procedure with multiple large series demonstrating a median time to recurrence of 10 to 16 months after RC.^{2,3}

Another available nomogram to predict 5-year recurrence risk following RC and PLND was created by IBCNC.¹² One distinguishing feature of the nomograms validated in our study is that NACH, ACH and AXRT are included as predictive variables while patients who received these treatments were excluded from IBCNC nomogram development.¹² Including

these variables allows for risk prediction in the growing subset of patients who receive these treatments and increases overall predictive accuracy. Because the recommendation for NACH has become the standard of care, we are already observing an increase in its use, especially in academic centers.²⁰ The inclusion of this treatment as a predictive variable is the main advantage of the nomogram addressed in our study and also the reason why the validation that we performed is unique to our knowledge. Furthermore, a recent study demonstrated superior accuracy of the Bladder Cancer Research Consortium nomogram in 5-year recurrence risk prediction vs IBCNC predictions and TNM staging alone (ROC-AUC 0.863 vs 0.847 and 0.829, respectively, $p = 0.02$).²¹

Although our current study represents a statistically rigorous validation of the outcome nomograms, limitations exist. Patient data were prospectively entered in a database at the time of evaluation and treatment but followup data were at times obtained in retrospective fashion. Thus, selection bias may have been created when several patients lacked sufficient followup information for inclusion in analysis. In our study group multiple data reviews assured the accuracy of data.

The nomograms are a statistical tool based on correlation of patient characteristics, and pathological and treatment factors with patient outcomes. In the original nomogram study series and the validation series patients were selected for treatments based on physician judgment and available evidence. Therefore, the accuracy of outcome predictions may differ in groups with different treatment practices.

Additionally statistical correlations require cautious interpretation as causation cannot be inferred without an appropriate prospective, controlled study design. The nomograms include a negative correlation of patient outcomes with NACH, ACH and AXRT.^{7,8} A likely explanation for this observed phenomenon is that patients with poor prognostic factors were selected for these treatments and subsequently did worse because of advanced and aggressive disease. Therefore, it is inappropriate to infer a causal relationship based on nomogram predictions alone between adjuvant therapies and an increased risk of poor outcomes. Rather, nomogram predictions should be used to identify certain patient subgroups with particularly poor predicted outcomes for clinical trials that may involve supplementary adjuvant therapies or altered followup. Conversely patients with particularly good outcome predictions could be entered into clinical trials that spare certain routine followup after RC or toxic therapies with the aim of decreasing undue stress, unwanted side effects, unnecessary invasive procedures and financial burden.

CONCLUSIONS

The nomograms for bladder cancer recurrence, cancer specific survival and overall survival following RC and PLND performed well in our series with accuracy comparable to that in the original series. To our knowledge this study is the first validation of these nomograms in an American cohort that received NACH and ACH. Calibration plots demonstrated significant nomogram overestimation of end points as actuarial patient outcomes were better than nomogram predictions for recurrence, and cancer specific and overall survival. Clinical

trials are required to determine the potential impact of nomogram predictions on disease management with further implementation into clinical practice.

Acknowledgments

Supported by Aventis Fellowship in Urologic Oncology—SWOG/Hope Foundation and National Institutes of Health/National Cancer Institute Career Development Award Grant K23CA160664 (GG).

Abbreviations and Acronyms

ACH	adjuvant chemotherapy
AJCC	American Joint Committee on Cancer
AXRT	adjuvant external beam radiotherapy
IBCNC	International Bladder Cancer Nomogram Consortium
NACH	neoadjuvant chemotherapy
PLND	pelvic lymph node dissection
RC	radical cystectomy
UC	urothelial carcinoma

REFERENCES

- Stein JP. Indications for early cystectomy. *Urology*. 2003; 62:591. [PubMed: 14550424]
- Shariat SF, Karakiewicz PI, Palapattu GS, et al. Outcomes of radical cystectomy for transitional cell carcinoma of the bladder: a contemporary series from the Bladder Cancer Research Consortium. *J Urol*. 2006; 176:2414. [PubMed: 17085118]
- Madersbacher S, Hochreiter W, Burkhard F, et al. Radical cystectomy for bladder cancer today—a homogeneous series without neoadjuvant chemotherapy. *J Clin Oncol*. 2003; 21:690. [PubMed: 12586807]
- Greene F, Trotti A, Fritz A. , et al. *AJCC Cancer Staging Handbook*. 7th. Vol. chapt 45. New York: Springer-Verlag; 2010. Urinary bladder.
- Tilki D, Reich O, Karakiewicz P, et al. Validation of the AJCC TNM substaging of pT2 bladder cancer: deep muscle invasion is associated with significantly worse outcome. *Eur Urol*. 2010; 58:112. [PubMed: 20097469]
- Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996; 15:361. [PubMed: 8668867]
- Shariat S, Karakiewicz P, Palapattu G, et al. Nomograms provide improved accuracy for predicting survival after radical cystectomy. *Clin Cancer Res*. 2006; 12:6663. [PubMed: 17121885]
- Karakiewicz P, Shariat S, Palapattu G, et al. Nomogram for predicting disease recurrence after radical cystectomy for transitional cell carcinoma of the bladder. *J Urol*. 2006; 176:1354. [PubMed: 16952631]
- Harrell FE Jr, Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. *JAMA*. 1982; 247:2543. [PubMed: 7069920]
- Take the Nomogram Challenge. [Accessed August 6, 2015] Available at <http://www.nomogram.org>.
- Vickers A. Prediction models in cancer care. *CA Cancer J Clin*. 2011; 61:315. [PubMed: 21732332]

12. International Bladder Cancer Nomogram Consortium: Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer. *J Clin Oncol.* 2006; 24:3967. [PubMed: 16864855]
13. Stephenson AJ, Scardino PT, Eastham JA, et al. Postoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J Clin Oncol.* 2005; 23:7005. [PubMed: 16192588]
14. Kattan MW, Reuter V, Motzer RJ, et al. A postoperative prognostic nomogram for renal cell carcinoma. *J Urol.* 2001; 166:63. [PubMed: 11435824]
15. Zaak D, Burger M, Otto W, et al. Predicting individual outcomes after radical cystectomy: an external validation of current nomograms. *BJU Int.* 2009; 106:342. [PubMed: 20002664]
16. Nuhn P, Matthias M, Sun M, et al. External validation of postoperative nomograms for prediction of all-cause mortality, cancer-specific mortality, and recurrence in patients with urothelial carcinoma of the bladder. *Eur Urol.* 2012; 61:58. [PubMed: 21840642]
17. Sonpavde G, Goldman BH, Lerner SP, et al. Bladder cancer after neoadjuvant chemotherapy. *Cancer.* 2009; 115:4104. [PubMed: 19517476]
18. Grossman HB, Natale RB, Tangen C. Neoadjuvant chemotherapy plus cystectomy compared with cystectomy alone for locally advanced bladder cancer. *N Engl J Med.* 2003; 349:859. [PubMed: 12944571]
19. Advanced Bladder Cancer (ABC). Meta-analysis Collaboration: Neoadjuvant chemotherapy in invasive bladder cancer: update of a systematic review and meta-analysis of individual patient data. *Eur Urol.* 2005; 48:202. [PubMed: 15939524]
20. Bochner B, Sperling D, Feifer A. Multi-institutional quality care initiative to improve the care of patients with muscle invasive bladder cancer. Presented at annual meeting of American Urological Association; May 16–21, 2014; Orlando, Florida. MP 60–12
21. Gakis G, Todenhoefer T, Renninger M, et al. External validation of current nomograms in invasive bladder cancer. *J Clin Oncol.* 2011; 29(suppl.) abstract e15184.

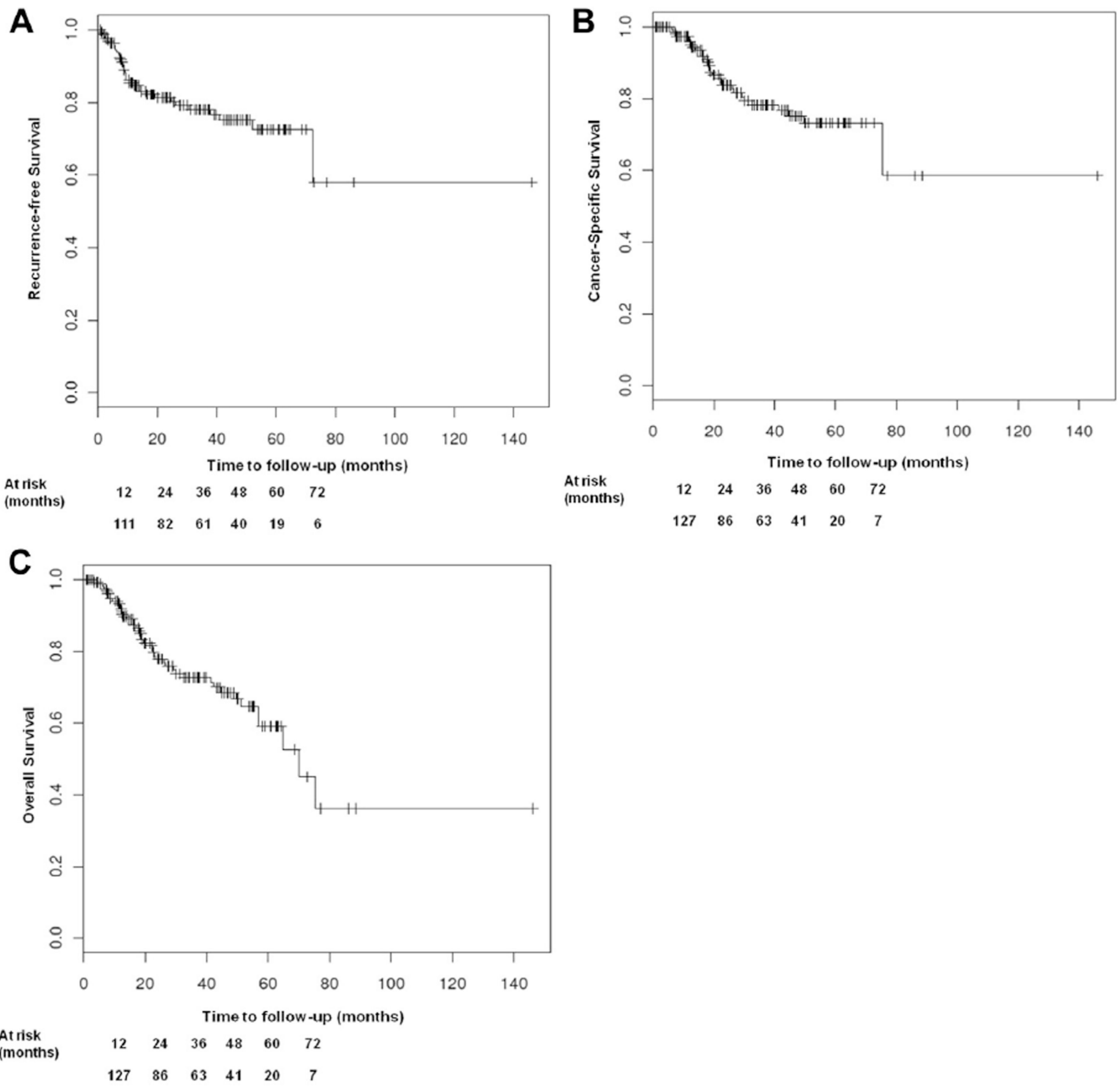


Figure 1. Kaplan-Meier plots of estimated recurrence-free (A), cancer specific (B) and overall (C) survival outcomes

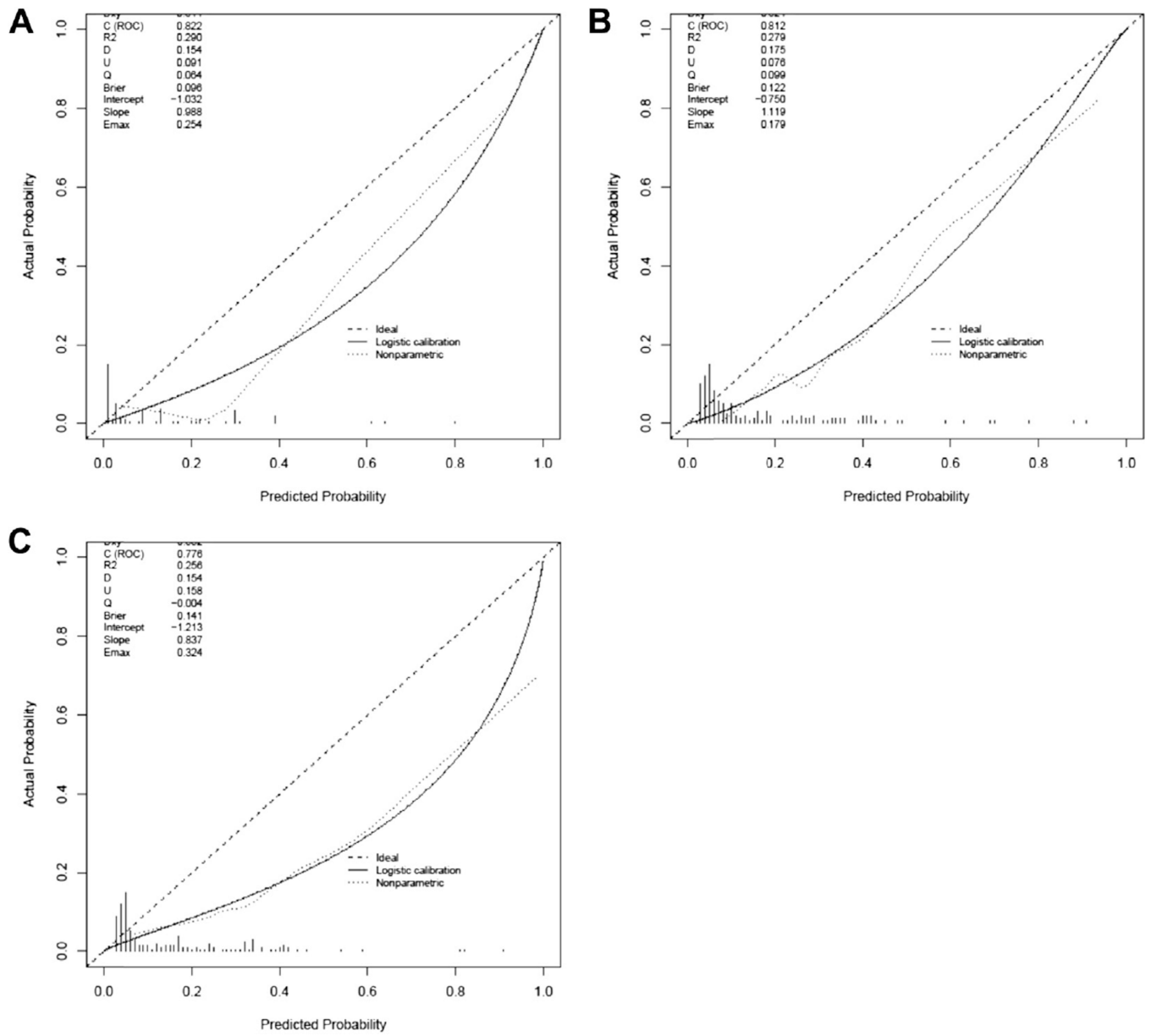


Figure 2. Calibration plots of predicted cancer specific (A), overall (B) and recurrence-free (C) survival outcomes calculated by nomogram.

Table 1

Patient characteristics and pathological staging

Covariates	Validation	Original	p Value
No. pts (%)	174	731	–
Median age	69	65.9	–
No. male (%)	154 (88.5)	602 (82.4)	0.05
No. tumor grade (%): *			0.16
G1, G2	19 (10.9)	56 (7.7)	0.16
G3	137 (78.7)	619 (84.6)	0.06
GX [†]	18 (10.3)	56 (7.7)	0.25
No. tumor stage (%): *			0.30
T0, Ta, Tis	52 (29.9)	171 (23.4)	0.07
T1	23 (13.2)	94 (12.9)	0.89
T2a, T2b	29 (16.7)	164 (22.4)	0.10
T3a, T3b	52 (29.9)	216 (29.5)	0.92
T4a, T4b	18 (10.3)	86 (11.8)	0.60
No. nodal stage (%): *			0.11
N0	132 (75.9)	557 (76.2)	0.92
N1	10 (5.7)	67 (9.2)	0.19
N2	31 (17.8)	93 (12.7)	0.10
N3	1 (0.6)	14 (1.9)	0.33
Lymphovascular invasion pos	56 (32.2)	274 (37.5)	0.19
Ca in situ pos	92 (52.9)	392 (53.6)	0.86

* AJCC 2002, 6th edition definitions.

[†]T0 on final pathological evaluation.

Table 2

Additional treatment and observed end points after RC and PLND

	No. Pts (%)	Median Mos Followup (IQR)
Treatment:		
NACH	16 (9.2)	–
ACH	55 (31.6)	–
AXRT	3 (1.7)	–
Observed end points:		
Recurrence	34	9 (6–16)
Death from disease	28	18 (12–26)
Overall deaths	43	19 (12–33)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Nomogram ROC-AUC analysis results

Predicted Outcome	ROC-AUC		
	2 Yrs	5 Yrs	8 Yrs
Recurrence	0.776	0.809	0.794
Ca specific survival	0.822	0.840	0.849
Overall survival	0.812	0.820	0.825

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript