

# Use of nomograms for predictions of outcome in patients with advanced bladder cancer

Shahrokh F. Shariat, Pierre I. Karakiewicz, Guilherme Godoy and Seth P. Lerner

*Ther Adv Urol*  
[2009] 1(1) 13–26

DOI: 10.1177/  
1756287209103923

© The Author(s), 2009.  
Reprints and permissions:  
[http://www.sagepub.co.uk/  
journalsPermissions.nav](http://www.sagepub.co.uk/journalsPermissions.nav)

## Abstract:

**Introduction:** Accurate estimates of risk are essential for physicians if they are to recommend a specific management to patients with bladder cancer. In this review, we discuss the criteria for the evaluation of nomograms and review current available nomograms for advanced bladder cancer.

**Methods:** A retrospective review of the Pubmed database between 2002 and 2008 was performed using the keywords ‘nomogram’ and ‘bladder’. We limited the articles to advanced bladder cancer. We recorded input variables, prediction form, number of patients used to develop the prediction tools, the outcome being predicted, prediction tool-specific features, predictive accuracy, and whether validation was performed.

**Results:** We discuss the characteristics needed to evaluate nomograms such as predictive accuracy, calibration, generalizability, level of complexity, effect of competing risks, conditional probabilities, and head-to-head comparison with other prediction methods. The predictive accuracies of the pre-cystectomy tools ( $n=2$ ) range from ~65–75% and that of the post-cystectomy tools ( $n=5$ ) range from ~75–80%. While some of these nomograms are well-calibrated and outperform AJCC staging, none has been externally validated. To date, four studies demonstrated a statistically significant improvement in predictive accuracy of nomograms by including biomarkers.

**Conclusions:** Nomograms provide accurate individualized estimates of outcomes. They currently represent the most accurate and discriminatory decision-making aids tools for predicting outcomes in patients with bladder cancer. Use of current nomograms could improve current selection of patients for standard therapy and investigational trial design by ensuring homogeneous groups. The addition of biological markers to the currently available nomograms using clinical and pathologic data holds the promise of improving prediction and refining management of patients with bladder cancer.

**Keywords:** bladder cancer, nomogram, prediction, prognosis, risk

## Introduction

Carcinoma of the urinary bladder, the fourth most common cancer in men and the eighth most common cause of death by cancer in men, results in significant morbidity and mortality [Jemal *et al.* 2008]. There will be an estimated 68,810 new cases and 14,100 deaths in US men and women in 2008, the vast majority of which are urothelial carcinoma (UC). The majority of patients present with non muscle-invasive tumors, while 20–25% of patients have a muscle-invasive cancer as the first manifestation of their disease. Radical cystectomy (RC) is the treatment of choice for patients with muscle-invasive cancers and for select patients with non

muscle-invasive cancers that have failed intravesical therapy or who are considered high-risk for progression to muscle-invasive disease [Witjes and Hendricksen, 2007]. Despite advances in surgical techniques, deeper understanding of the role of lymphadenectomy and progress in postoperative care, 5-year disease-specific survival after RC remains 50–60% [Shariat *et al.* 2006a; Stein *et al.* 2001].

Accurate estimates of the likelihood of treatment success, complications and long-term morbidity are essential for patient counseling and informed medical decision-making (MDM) regarding treatment options. Knowing the probability

Correspondence to:  
**Seth P. Lerner, MD, FACS**  
Professor of Urology, Beth and Dave Swalm Chair of Urologic Oncology, Scott Department of Urology, Baylor College of Medicine, 6620 Main Suite 1325, Houston, Texas 77030, USA  
[slerner@bcm.edu](mailto:slerner@bcm.edu)

**Shahrokh F. Shariat, MD, PhD**  
Division of Urology; Sidney Kimmel Center for Prostate and Urologic Cancer, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, Box 27, New York, NY 10065, USA

**Pierre I. Karakiewicz**  
Cancer Prognostics and Health Outcomes Unit, University of Montreal, Health Center, (CHUM) Montreal, QC, Canada

**Guilherme Godoy**  
Division of Urology; Sidney Kimmel Center for Prostate and Urologic Cancer, Memorial Sloan-Kettering Cancer Center, New York, USA

of progression after RC is essential for selection of patients, who may benefit from adjunctive chemotherapy, particularly with the availability of effective systemic therapies and accurate prediction of the presence of systemic micrometastatic disease [Amiel and Lerner, 2006]. A well-informed patient will be more likely to be compliant with treatment suggestions and studies suggest that a lack of patient involvement is a major risk factor for regret of treatment choice [Miles *et al.* 1999], particularly when complications arise [Clark *et al.* 2001]. Finally, accurate risk-estimates may help identifying homogeneous high-risk patient groups to assist with clinical trial design and accrual.

Traditionally, physician judgment has formed the basis for risk estimation, patient counseling and MDM, however, clinicians' estimates might be biased due to subjective and objective confounders that exist at all stages of the prediction process [Hogarth and Karelaia, 2007; Vlaev and Chater, 2006; Kattan, 2001; Elstein, 1999]. Clinicians do not recall all cases equally; certain cases can stand out and exert an unsuitably large influence when predicting future outcomes. Clinicians might be inconsistent when processing their memory and tend to resort to heuristics when processing becomes difficult [Kattan, 2002]. When it is time to make a prediction, they tend to predict the preferred outcome rather than the outcome with the highest probability [Kattan, 2001]. Finally, without resorting to the use of computers, clinicians might find it difficult to integrate the growing numbers of predictor variables that are important in MDM [Ross *et al.* 2002; Rabbani *et al.* 2000]. To circumvent these limitations and to obtain the most accurate and reliable predictions, researchers have developed predictive/prognostic tools based on statistical models. These tools have been shown to perform better than clinical judgment, when predicting probabilities of outcome [Specht *et al.* 2005; Ross *et al.* 2002]. However, physician input is obviously still essential for the measurement of variables that are used in the prediction process, as well as in the interpretation and application of model-derived outcome predictions in clinical practice.

Among the available decision aids, nomograms currently represent the most accurate and discriminatory tools for predicting outcomes in patients with cancer [Capitanio *et al.* 2008; Shariat *et al.* 2009; 2008a; 2008b; 2008c;

2008d; 2008f; 2008g; 2008h; 2008i; 2008j]. Various studies have documented the superior performance of nomograms compared to risk-grouping [Shariat *et al.* 2009; 2006b; Chun *et al.* 2007b; Kattan *et al.* 2003a; 2003b; 2002; Kattan, 2003a; 2003b], look-up tables [Chun *et al.* 2007b; Gallina *et al.* 2007; Briganti *et al.* 2006; Kattan, 2003a; 2003b], tree analysis [Chun *et al.* 2007b; Steuber *et al.* 2006; Kattan, 2003a], and artificial neural networks (ANN) [Chun *et al.* 2007a; 2007b; Kattan, 2003a; Terrin *et al.* 2003; Schwarzer and Schumacher, 2002; Sargent, 2001]. In this review, we discuss the criteria for the evaluation of nomograms and discuss current available nomograms for advanced bladder cancer. We describe the patient populations to which they apply and the outcomes predicted, and report their individual characteristics.

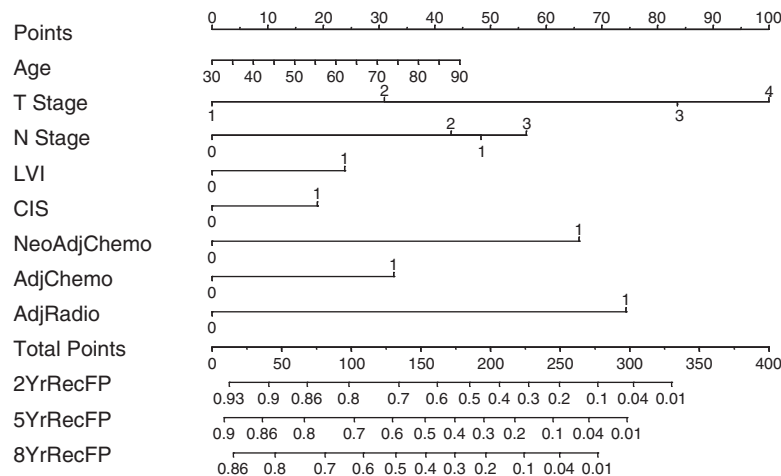
#### *Nomogram: definition*

Various distinct statistical methodologies have broadly been described as 'nomograms.' According to the strict definition, a nomogram represents a graphical calculation instrument, that can be based on any type of function, such as logistic regression or Cox hazards ratio regression models [Kattan, 2002; Kattan *et al.* 1998]. The nomogram can usually incorporate continuous or categorical variables. The effect of the variables on the specific outcome is represented in the format of axes, and risk points are attributed according to the prognostic/predictive importance of the variable of interest. For example, the nomogram in Figure 1 [Karakiewicz *et al.* 2006a] assigns to each pathologic stage a unique point value that represents its prognostic significance.

The nomogram format is unique as it allows numerous combinations of the inputs of many continuously and/or categorically coded variables. This format distinguishes nomograms from look-up tables or decision trees, where continuously coded variables cannot be processed, and where data availability limits the degree of stratification to avoid empty cells or dead-end branches. Nomograms are designed to extract the maximum amount of information from data with the goal of providing the most accurate predictions.

#### **Evaluation predictive tools**

Decision aids can be compared, based on several characteristics: predictive accuracy, calibration



## Legend :

T stage :	N stage :	0 : No	2 YrRecFP: 2 year recurrence-free probability
1 : T <sub>0</sub> , T <sub>a</sub> , T <sub>is</sub> , T <sub>1</sub>	0 : N <sub>0</sub>	1 : Yes	5 YrRecFP: 5 year recurrence-free probability
2 : T <sub>2</sub>	1 : N <sub>1</sub>		8 YrRecFP: 8 year recurrence-free probability
3 : T <sub>3</sub>	2 : N <sub>2</sub>		
4 : T <sub>4</sub>	3 : N <sub>3</sub>		

To obtain nomogram predicted probability of recurrence, locate patient values at each axis, Draw a vertical line to the "Point" axis to determine how many points are attributed for each variable value. Sum the points for all variables. Locate the sum on the "Total Points" line. Draw a vertical line towards the the 2YrRecFP, 5YrRecFP and 8YrRecFP axes to respectively determine the 2 year recurrence-free probability, the 5 year recurrence-free probability and the 8 year recurrence-free probability.

**Figure 1.** Bladder cancer recurrence in 731 patients treated with radical cystectomy and bilateral lymphadenectomy for urothelial carcinoma of the bladder ([Karakiewicz *et al.* 2006a]; reprinted with permission).

(correlation between predicted and observed risk throughout the entire range of predictions), generalizability, level of complexity, adjustment for the effect of competing risks, and use of conditional probabilities.

### Predictive accuracy

Accuracy quantifies the model's ability to discriminate between patients with or without the outcome of interest. The quantification uses the area under the curve (AUC) for binary outcomes (i.e., binary logistic regression models), and the c-index for censored data (i.e., time-to-event data) [Begg *et al.* 2000]. The AUC measures the ability of the nomogram to discriminate between those with or without the outcome of interest, while the c-index measures the ability of the nomogram to randomly chose first, patients who have the event of choice. As for the AUC, a c-index of 0.5 represents no

discriminating ability, whereas a value of 1.0 represents perfect discrimination.

A particular model's accuracy represents the most important consideration for comparison of different models. A valid determination of the model's accuracy would require the application of the same model under novel testing conditions, different from the development cohort. Thus the accuracy should be ideally tested in an independent cohort, however, in the absence of an external cohort, models are usually subjected to internal validation. In this situation, bootstrapping represents the ideal internal validation format, where the development dataset is used to simulate model testing under novel conditions [Steyerberg *et al.* 2003; Kattan, 2003a; 2002; Steyerberg *et al.* 2001; Steyerberg *et al.* 1998; Bradley, 1993]. Split-sample and cross-validation (leave-one-out validation) also represent valid alternatives [Steyerberg *et al.* 2001].

### Calibration

The model's accuracy (discrimination) indicates the overall ability to predict the outcome of interest. However, overall accuracy does not indicate the ability of the model to predict the outcome of interest in specific patient groups or according to risk level. For example, a model that is 80% accurate may predict virtually perfectly in high-risk patients, but may demonstrate dismal performance in low-risk ones. The relationship between predicted risk and observed rate of the outcome of interest should be provided for each new model, along with its overall accuracy. Calibration plots provide this type of information and can be obtained for internal, as well as external data [Steyerberg *et al.* 2007; Steyerberg *et al.* 2003; Steyerberg *et al.* 2001; Steyerberg *et al.* 1998; Bradley 1993]. These plots graphically illustrate the relationship between predicted and observed rates of the outcome of interest. Ideally, a model with perfect ability to predict the outcome of interest should exhibit a 1:1 relationship between predicted and observed rates, which results in a 45° slope.

### Generalizability

Several considerations may undermine the generalizability of a model under specific conditions or in a specific population. Differences in disease and in population characteristics may undermine the accuracy of predictive and prognostic models, when applied to a different population. Specific model criteria, such as inclusion and exclusion criteria, do not allow the use of models for patients with different characteristics or who have been exposed to different treatment modalities. For example, a model that is specific to UC cannot be applied to patients with squamous cell variants. Moreover, models that were developed using high-volume single center databases may not be invariably applicable to community practice. Therefore, it is imperative that the clinician knows whether a specific model is indeed generalizable and applicable to the population they intend to apply it to [Shariat *et al.* 2008b; 2008c; 2008d; 2008g; Shariat *et al.* 2005; Steyerberg *et al.* 2003; 2001; 1998; Bradley, 1993].

A model's ability to predict a specific outcome may be affected by population characteristics that change over time. In general, more contemporary cancer patients are diagnosed with more favorable stage and grade. Therefore, this kind of tool requires periodic reappraisals to assess the effect of stage and grade migration. One may

find that predictions devised on historic cohorts no longer apply to contemporary series of patients. However, models may also show stable accuracy and performance characteristics. External validation in contemporary cohorts is necessary to ensure temporal validity.

### Level of complexity

The level of complexity of a predictive or prognostic model represents an important practical consideration. Excessively complex models, which rely on multiple variables, are clearly impractical in busy clinical practice. Similarly, models that rely on variables that are not routinely available are impractical. These include models that rely on novel biomarkers, which in turn require non-commercially available and/or non-standardized assays, making the instrument unsuitable for the great majority of the clinicians.

### Adjustment for competing risks

Because of the protracted course of cancer, competing causes of mortality are extremely important variables to be taken into consideration, when building predictive/prognostic models [Shariat *et al.* 2007a; 2007b; 2006c; Lotan *et al.* 2005]. Competing risk modeling is able to predict cancer control rates, after accounting for the effects of competing risks. There is a need for more competing risk based modeling to better understand the risk of bladder cancer in the context of other cause mortality. Such predictions are important to clinicians, as well as to patients, especially when over-treatment or sub-optimal treatment considerations are addressed. Indeed, since the morbidity and mortality of bladder cancer treatment are not trivial, clinicians must be able to better risk-stratify bladder cancer patients to assure that those who are to benefit the most from the intervention, derive such benefit [Nielsen *et al.* 2007]. To date, the only modeling tools that allow adjustment for competing risks are nomograms [Shariat *et al.* 2008a; Chun *et al.* 2006].

### Conditional probabilities

The updated versions of the pre- and post-operative Kattan nomograms for prediction of biochemical recurrence after radical prostatectomy provide the opportunity to adjust for disease-free interval from surgery [Stephenson *et al.* 2005; 2006]. Absence of adjustment for disease-free interval presents the clinician with an excessively somber estimate of cancer control over time. Expectedly, the latter improves with



increasing disease-free interval. While the preliminary format of one such conditional probability nomogram has been presented, [Chun *et al.* 2006] no such model has been published, to date, in the bladder cancer literature.

#### *Head-to-head comparison*

When judging a new tool, one should examine its predictive accuracy, validity, and calibration relative to established models, with the intent of determining whether the new model offers advantages relative to available alternatives [Margulis *et al.* 2008a; Shariat *et al.* 2008c; 2008d; 2008j; 2006b; Kattan, 2003a; 2003b, Kattan *et al.* 2003b; Steyerberg *et al.* 2001; 1998]. Head-to-head comparisons represent the most direct and unbiased analysis of objective attributes (accuracy and calibration) of various models. Subsequently, complexity, generalizability and other considerations may also be compared. With this approach, the alternatives are compared directly, without having to judge the concordance index in isolation or against a possibly arbitrary threshold.

The main steps required in a head-to-head comparison consist of the application of the original model to a common external dataset that will serve for testing of all models that will be compared to one another. The original model is then applied to each individual observation to derive the probability of the outcome of interest. The predictions are then compared against observed rates of the outcome of interest, and accuracy (discrimination) is calculated using the receiver operating characteristics curve or another measure of discrimination, such as the Brier index. These steps are repeated for each of the tested models. A common mistake consists of refitting a new model that relies on the same variables as the original model and calling it the validation model.

#### **Currently available prediction tools**

The above discussion is meant to provide guidelines in the process of analyzing and utilizing decision aid selection instruments, such as nomograms. Herein, we provide an overview of nomograms used for advanced bladder cancer (Table 1). We recorded predictor variables, the outcome of interest, the number of patients utilized to develop the tools, tool-specific features, predictive accuracy estimates, and whether

internal and/or external validation has been performed.

#### *Pre-operative nomogram for prediction of pathologic features at RC*

Inaccuracy of the pre-RC clinical staging system is well documented, but continues to be a major determinant, governing MDM [Shariat *et al.* 2007b]. Consequently, development of accurate preoperative risk-stratification models would allow prediction of advanced disease and enable better selection of patients, who would benefit from neo-adjuvant systemic chemotherapy. To this end, pre-RC nomograms for the prediction of advanced pathologic stage (pT3–4) and presence of lymph node metastases were developed from a multicenter cohort of 731 patients with available clinical and pathologic staging data [Karakiewicz *et al.* 2006b]. This cohort was derived from a collaboration of Baylor College of Medicine, UT Southwestern, and Johns Hopkins who pooled their data on 958 patients treated with radical cystectomy and bilateral pelvic lymphadenectomy, and University of Montreal (Bladder Cancer Research Consortium–BCRC). When patient age, transurethral resection (TUR) stage, grade, and presence of carcinoma in situ (CIS) were integrated within the nomogram, 75.7% accuracy was recorded in predicting advanced pathologic stage (pT3–4) *versus* 71.4% when TUR stage alone was used. The nomogram was 63.1% accurate in predicting lymph node metastases, when TUR stage and grade were used *vs.* 61% using TUR stage alone. The various components of these nomograms can be accessed at [www.nomogram.org](http://www.nomogram.org).

The pre-RC nomograms provide only a modest increase in accuracy. Nonetheless, they demonstrate that the combined use of clinical and pathologic variables, which cannot always be integrated within look-up tables, results in more accurate predictions than the use of a single variable. Several variables may have contributed to the suboptimal accuracy of these nomograms. These may include differences in TUR technique, non-standardized use of re-staging biopsies and variability in the pathologic evaluation. Possibly, the integration of other pathologic prognostic markers, such as lymphovascular invasion (LVI), in addition to molecular markers of disease, might enhance predictive accuracy of pre-RC nomograms [Shariat *et al.* 2008e]. However, the limited ability to predict nodal

**Table 1.** Available predictive models for advanced bladder cancer.

Reference	Prediction form	Patient population	Outcome	No. of patients	Variables	Accuracy	Validation
[Shariat <i>et al.</i> 2008a]	PN	pT1a-3NDM0	2-, 5-, and 8-year recurrence free survival 2-, 5-, and 8-year cancer-specific survival	191	p53, p21, pRB, p27, cyclin E1, gender, age, pathologic stage, grade, lymphovascular invasion, and concomitant carcinoma in situ	83.4% 86.9%	Internal
[Qureshi <i>et al.</i> 2000]	ANN	Ta/T1 Ta/T1 T2-T4	Recurrence within 6 months Progression-free survival 1-year cancer-specific survival	56 105 40	EGFR, c-erbB2, p53, stage, grade, tumor size, number of tumors, gender, smoking status, histology of mucosal biopsies, carcinoma in situ, metaplasia, architecture, location	75% 80% 82%	Internal
[Catto <i>et al.</i> 2003]	NFM	Ta-T4	Recurrence-free survival	109	p53, mismatch repair proteins, stage, grade, age, smoking status, previous cancer	88–95%	Internal
[Karakiewicz <i>et al.</i> 2006a]	PN	RC	Cystectomy T and N stage	731	Age, TUR stage, TUR grade, carcinoma in situ	76% for T stage 63% for N stage	Internal
[Karakiewicz <i>et al.</i> 2006a]	PN	RC	2-, 5-, and 8-year recurrence free survival	731	Age, T stage, N stage, grade, lymphovascular invasion, carcinoma in situ, adjuvant radiotherapy, adjuvant chemotherapy, neo-adjuvant chemotherapy	78%	Internal
[Shariat <i>et al.</i> 2006b]	PN	RC	2-, 5-, 8-year all cause and cancer-specific survival	731	Age, T stage, N stage, grade, lymphovascular invasion, adjuvant radiotherapy, adjuvant chemotherapy, time from diagnosis to surgery	79% for all cause survival 73% for cancer-specific survival	Internal
[Bochner <i>et al.</i> 2006]	PN	RC	5-year recurrence free survival	9,064	Age, gender, T stage, N stage, grade, histology, time from diagnosis to surgery	75%	Internal
[Bassi <i>et al.</i> 2007]	ANN	RC	5-year all cause survival	369	Age, gender, T stage, N stage, lymphovascular invasion, grade, concomitant CaP, history of upper tract UC	76%	Internal

FNM = Fuzzy neuro-modeling  
 PN = Probability nomogram  
 RC = Radical cystectomy  
 UC = Urothelial carcinoma

metastases indicates that the accurate pretreatment prediction of this outcome represents one of the challenges in urologic oncology.

#### *Post-operative nomogram for prediction of disease recurrence and survival after RC*

Several post-RC nomograms have been developed to predict the natural history of surgically treated bladder cancer and to assist in the decision process regarding the use of adjuvant therapy after RC [Bochner *et al.* 2006; Karakiewicz *et al.* 2006a; Shariat *et al.* 2006b]. The BCRC used the dataset described above to determine the probabilities of recurrence, cancer-specific and all-cause mortality at 2, 5 and 8 years after surgery (Figures 1 and 2; available at [www.nomogram.org](http://www.nomogram.org)) [Karakiewicz *et al.* 2006a; Shariat *et al.* 2006b]. To address each of the three outcomes, three separate nomograms were developed and internally validated. All three exceeded the accuracy of the American Joint Committee on Cancer (AJCC) stage groupings, as well as of the individual predictors. All showed excellent performance characteristics, which virtually corresponded to ideal predictions. The recurrence nomogram (accuracy: 78%) relied on pathological T and N stages, pathologic grade, presence of LVI and CIS at RC, as well as the delivery of chemotherapy (either neo-adjuvant, adjuvant or both), and/or radiation. The cause-specific mortality nomogram accurately predicted in 78% of cases and the all-cause mortality nomogram in 73% of cases.

The authors demonstrated that their nomograms were significantly more accurate/discriminatory than the AJCC-staging risk grouping resolving some of the heterogeneity of outcome prediction within each AJCC-staging risk group. Moreover, the nomogram predictions were tailored to the risk posed by the characteristics of an individual's cancer, which is more relevant to the patient than are group-level probabilities.

In the same year, the International Bladder Cancer Nomogram Consortium (IBCNC) published a post-operative nomogram predicting the risk of recurrence at 5 years following RC and pelvic lymph node dissection (available at [www.nomograms.org](http://www.nomograms.org)) [Bochner *et al.* 2006]. The dataset developed for this study included over 9,000 patients from 12 centers including the BCRC data. Age, gender, grade, pathological stage, histological type, lymph node status, time from diagnosis to surgery were significant

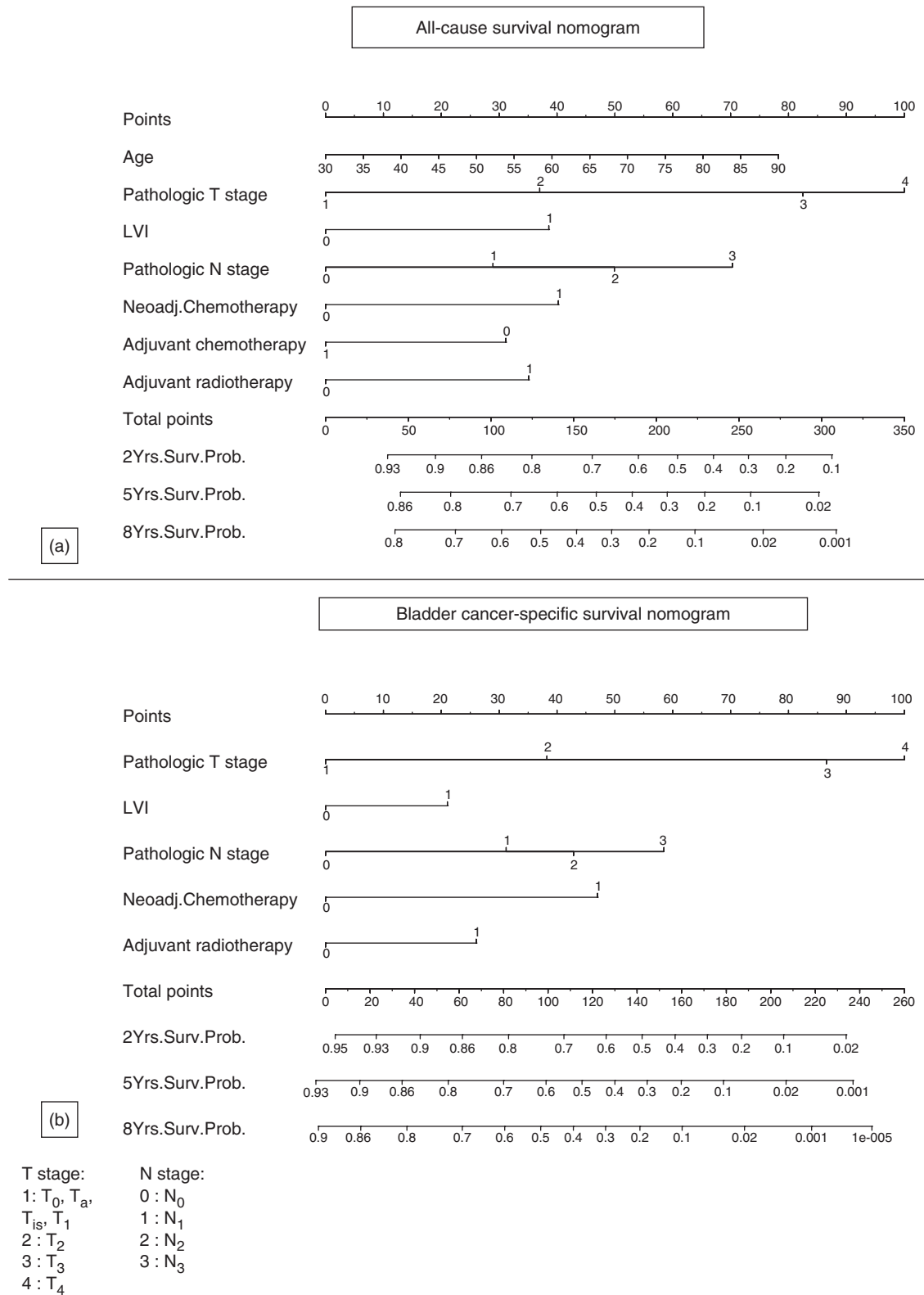
contributing factors in the nomogram. The predictive accuracy of the nomogram (75%) was statistically superior to either AJCC TNM staging (68%) or standard pathologic grouping models (62%).

The recurrence nomogram of the IBCNC is more generalizable than that of the BCRC and can be applied to patients with histological variants other than UC. The BCRC nomogram is best suited for patients from the United States and may be exclusively applied to patients with UC. Despite these limitations, the BCRC nomogram offers a 4% accuracy advantage (78% *versus* 74%) compared to the IBCNC nomogram. This may be due to a higher homogeneity of the population in the BCRC study. In addition, the BCRC offers a suite of nomograms predicting all cause and cancer-specific survival in addition to tumor recurrence. Furthermore, their models provide 2-, 5- and 8-year predictions while the IBCNC nomogram provides only 5-year prediction. This feature allows better identification of early bladder cancer recurrences, which represent the most aggressive type of recurrences.

Several limitations of nomogram approaches to patient risk stratification should be noted. First, and foremost, all currently available predictive tools in bladder cancer are not perfectly accurate. Because of difficulty with uniform data collection, inherent in multi-institutional collaborative efforts, important surgical parameters, such as timing of RC, margin status, extent of lymph node dissection and lymph node density are not reflected in currently available predictive models. The retrospective nature of the data collection imposes the limitations associated with bias in data acquisition and prospective validation and adjustment is imperative. Furthermore, all available nomograms were derived and are applicable to centers of excellence for bladder surgery. Their applicability in different clinical settings must be viewed with caution and certainly requires additional validation. Finally, any further improvement in predictive accuracy of nomograms may require the integration of molecular prognosticators, as clinical and pathologic prognosticators alone appear to have limited prognostic ability.

#### *Other post-operative tools for prediction of disease recurrence and survival after RC*

As an alternative to nomogram-based modeling, Bassi *et al.* developed an ANN utilizing gender, age at surgery, LVI, pathological T and N stage,



**Figure 2.** (a) All cause and (b) bladder cancer-specific survival nomogram in 731 patients treated with radical cystectomy and bilateral lymphadenectomy for urothelial carcinoma of the bladder [(Shariat *et al.* 2006b)]; reprinted with permission.



grade, presence of concomitant prostatic adenocarcinoma and history of upper tract UC as input variables for prediction of 5 year all-cause survival after RC [Bassi *et al.* 2007]. In a single institution cohort of 369 patients, the prognostic accuracy of the ANN (76%; based on 12 variables) was slightly superior to the logistic regression model that was based on only two statistically significant variables (75%; stage and grade) [Bassi *et al.* 2007]. Unfortunately, the comparison of the accuracy of both models was performed on the same population that served for model development, which undermines the validity of such comparison.

#### *Nomograms including novel biomarkers*

There is ample room for improvement of the accuracy of current predictive tools. Better modeling of the data, use of larger datasets, and more systematic and focused data collection (e.g., tighten the definition of symptom status) may help to improve the nomogram accuracy. The limited accuracy of current models is partially related to the heterogeneous biologic behavior of tumors with the same clinical and/or pathologic features. The integration of novel biomarkers and/or data derived from imaging tools, that are associated with the biologic behavior of bladder cancer may help improving the accuracy of nomogram predictions. In addition, the emergence of new therapeutic approaches for bladder cancer cannot flourish without a set of markers to serve as prognosticators and/or therapeutic targets. However, despite numerous reports of promising new biomarkers have been reported in the urological literature [Margulis *et al.* 2008a; 2008b; 2006; Shariat *et al.* 2008e; 2007c; 2006a; Karam *et al.* 2007], only four studies to date have demonstrated a statistically significant improvement in predictive accuracy, when biomarkers were added to established predictors in the tool setting [Shariat *et al.* 2008; Karam *et al.* 2007; Catto *et al.* 2003; Qureshi *et al.* 2000]. One of these studies, for example, demonstrated that the addition of a panel of five well-established cell cycle regulatory bio-markers (p53, pRB, p21, p27, and cyclin E1) improved the predictive accuracy of competing-risk nomograms for prediction of bladder cancer recurrence and survival following RC for patients with pT<sub>a</sub>-pT<sub>3</sub>, node negative tumors by a clinically significant margin (Figure 3) [Shariat *et al.* 2008a]. Two smaller studies have added biomarkers to standard clinico-pathologic features using ANN and neuro-fuzzy modeling [Catto *et al.* 2003;

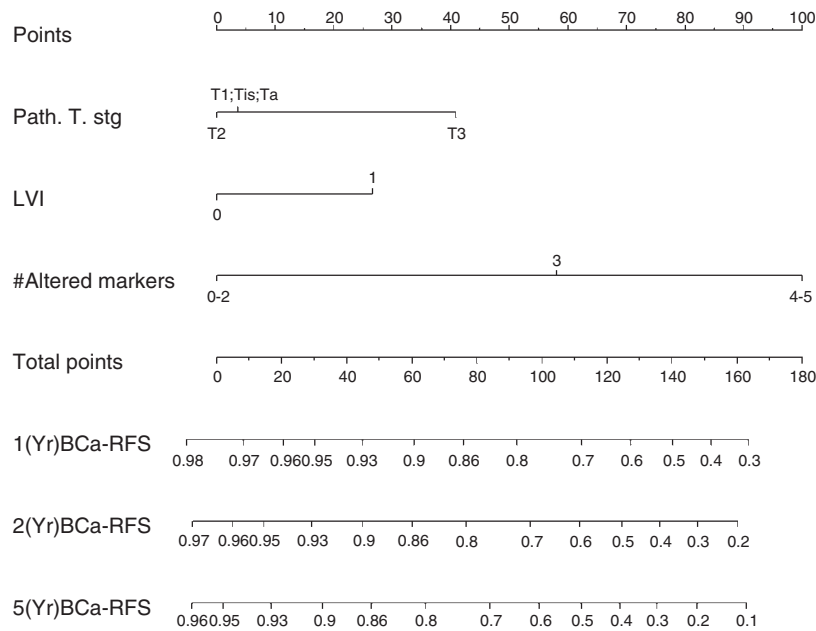
Qureshi *et al.* 2000]. Prediction tools such as these, that incorporate pathological and molecular information could form the basis for counseling patients regarding their risk of disease recurrence following surgery and for designing clinical trials to test adjuvant treatment strategies in high-risk patients.

#### **Final considerations**

Bladder cancer patients need to be involved in the decision-making process, regarding the management of their disease. They should know what their options are and what the consequences can be for each option. Ideally, patients would make their own treatment decisions. At the core of any patient involvement would be accurate prediction of consequences and, essentially, a spreadsheet of these predictions tailored to the individual. This spreadsheet would represent informed consent for any medical decision. Providing this should reduce the likelihood of regret of treatment choice, particularly when complications arise.

Continuous multivariable models such as nomograms are a highly appealing means of calculating accurate predictions with or without the use of a computer. Many nomograms have been constructed for patients with bladder cancer. Nomograms currently represent the most accurate and discriminating tools for predicting outcomes in patients with bladder cancer. When faced with the difficult decision of choosing among the treatment options for each clinical stage of bladder cancer, the nomograms provide patients with accurate estimates of outcomes. Equipped with this information, the patient is more likely to be confident in their treatment decision and less likely to experience regret in the future. However, it should be emphasized that nomogram predictions must be interpreted as such; they do not make treatment recommendations or act as a surrogate for physician-patient interactions, nor do they provide definitive information on symptomatic disease progression or complications associated with treatments.

The fundamental issue raised by opponents of predictive tools, such as nomograms is regarding their utility. Indeed, very limited data exist with respect to the impact of nomograms on MDM. There are no prospective randomized studies that clearly demonstrate that the use of nomograms improves patient care. Despite studies showing



**Figure 3.** Post-operative nomogram that integrates the immunohistochemical status of five established cell cycle regulatory bio-markers (p53, pRB, p21, p27, and cyclin E1) with standard histopathologic variables for predicting 1-, 2- and 5-year risk of disease recurrence. Data analyzed in 191 patients with pTa-3 N0 M0 urothelial carcinoma of the bladder treated with radical cystectomy and bilateral lymphadenectomy [Shariat *et al.* 2008a]; reprinted with permission).

that decision aids improve patient knowledge and affect MDM behavior [O'Connor, 1999], the role of nomograms has yet to be proven. A clinical trial would be very valuable for establishing the effect of nomograms on patient MDM; however, whereas informing patients with predictions regarding the impact of a medical procedure seems ethical, withholding accurate outcome predictions from patients to achieve equipoise in a randomized trial, where the control arm lacks this information, does not. Presently, patients are using very limited information when making their decisions, and direct outcome predictions are the simplest factors for them to consider; MDM is facilitated when patients can see tailored predictions of their outcomes with various alternatives.

Besides improved MDM, accurate risk estimates are also required for evaluation of novel markers and clinical trial design, to ensure homogeneous high-risk patient groups for whom new cancer therapeutics will be investigated. Prediction models have the potential of improving the ability of Phase II trials to discriminate between ineffective and potentially effective therapies. As in most Phase II trials, both the expected response rate

from standard care and the actual response rate of novel therapy are group level estimates: the researchers simply take the number of patients who respond and divide by the total treated to get a proportion. This takes no account of individual patient characteristics and therefore implicitly assumes similarity of trial patients and historical controls. However, trial patients may differ from controls in terms of performance status, prior experience of chemotherapy, and extent of disease, or other prognostic factors. Apparently promising or disappointing results from a Phase II treatment trial can be merely the result of such differences [Bajorin, 2004]. For Phase III trials, prediction models can help to ensure that eligible patients are at sufficient levels of high-risk, thereby increasing event rates and reducing sample size requirements. Risk prediction models define high-risk patients more accurately than risk-grouping strategies. Use of risk predictions for individual patients, therefore, decreases the proportion of low-risk patients enrolled, avoiding unethical inclusion, as well as increasing statistical power. Finally, future designs of Phase III trials should include prediction models to increase the clinical utility of their findings. We, therefore, recommend the wider

adoption of risk prediction models in the design, analysis, and implementation of clinical trials.

## Conclusions

In conclusion, despite the limitations of data, nomograms provide optimum accuracy for individualized evidence-based MDM of bladder cancer. Patients, administrators, peers, and third party payers increasingly demand to have an objective justification for almost all clinical decisions. Evidence-based nomograms can provide individualized estimates for a number of endpoints in urologic oncology. Moreover, when determining the usefulness of nomograms, many patients want to know their likely outcomes, and most clinicians would like to provide accurate estimates of those outcomes. When a nomogram is available, little else is able to make more accurate predictions of outcome, which confirms the usefulness of nomograms.

## Conflict of interest statement

SFS and SPL are co-owners of the patent entitled “Methods to determine prognosis after therapy for bladder cancer.” U.S. Patent Application Serial Number: Docket 675.003US1. Filed June 1, 2001.

## References

- Amiel, G.E. and Lerner, S.P. (2006) Combining surgery and chemotherapy for invasive bladder cancer: current and future directions. *Expert Rev Anticancer Ther* 6: 281–291.
- Bajorin, D. (2004) The phase III candidate: can we improve the science of selection? *J Clin Oncol* 22: 211–213.
- Bassi, P., Sacco, E., De Marco, V., Aragona, M. and Volpe, A. (2007) Prognostic accuracy of an artificial neural network in patients undergoing radical cystectomy for bladder cancer: a comparison with logistic regression analysis. *Br J Urol Int* 99: 1007–1012.
- Begg, C.B., Cramer, L.D., Venkatraman, E.S. and Rosai, J. (2000) Comparing tumour staging and grading systems: a case study and a review of the issues, using thymoma as a model. *Stat Med* 19: 1997–2014.
- Bochner, B.H., Kattan, M.W. and Vora, K.C. (2006) Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer. *J Clin Oncol* 24: 3967–3972.
- Bradley, E.A. and Tibshirani, R.J. (1993) *Monographs on Statistics and Applied Probability: An Introduction to the Bootstrap*. Chapman and Hall/CRC.
- Briganti, A., Chun, F.K., Salonia, A., Gallina, A., Farina, E., Da Pozzo, L.F. *et al.* (2006) validation of a nomogram predicting the probability of lymph node invasion based on the extent of pelvic lymphadenectomy in patients with clinically localized prostate cancer. *Br J Urol Int* 98: 788–793.
- Capitanio, U., Jeldres, C., Shariat, S. F. and Karakiewicz, P. (2008) Clinicians are most familiar with nomograms and rate their clinical usefulness highest, look-up tables are second Best. *Eur Urol* 54: 958–959.
- Catto, J.W., Linkens, D.A., Abbod, M.F., Chen, M., Burton, J.L., Feeley, K.M. *et al.* (2003) Artificial intelligence in predicting bladder cancer outcome: a comparison of neuro-fuzzy modeling and artificial neural networks. *Clin Cancer Res* 9: 4172–4177.
- Chun, F.K., Graefen, M., Briganti, A., Gallina, A., Hopp, J., Kattan, M.W. *et al.* (2007a) Initial biopsy outcome prediction—head-to-head comparison of a logistic regression-based nomogram versus artificial neural network. *Eur Urol* 51: 1236–1240; discussion 1241–1233.
- Chun, F.K., Karakiewicz, P.I., Briganti, A., Walz, J., Kattan, M.W., Huland, H. *et al.* (2007b) A critical appraisal of logistic regression-based nomograms, artificial neural networks, classification and regression-tree models, look-up tables and risk-group stratification models for prostate cancer. *Br J Urol Int* 99: 794–800.
- Chun, K.H., Briganti, A., Shariat, S.F., Palapattu, G.S., Lotan, Y., Rogers, C.G. *et al.* (2006) Competing risks nomogram predicting the probability of bladder cancer recurrence after cystectomy can reliably distinguish between tumours destined to recur and patients who may die of other causes before recurrence. *European Urology Supplements* 5: 38.
- Clark, J.A., Wray, N.P. and Ashton, C.M. (2001) Living with treatment decisions: regrets and quality of life among men treated for metastatic prostate cancer. *J Clin Oncol* 19: 72–80.
- Elstein, A.S. (1999) Heuristics and biases: selected errors in clinical reasoning. *Acad Med* 74: 791–794.
- Gallina, A., Chun, F.K., Briganti, A., Shariat, S.F., Montorsi, F., Salonia, A. *et al.* (2007) Development and split-sample validation of a nomogram predicting the probability of seminal vesicle invasion at radical prostatectomy. *Eur Urol* 52: 98–105.
- Hogarth, R.M. and Karelaia, N. (2007) Heuristic and linear models of judgment: matching rules and environments. *Psychol Rev* 114: 733–758.
- Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T. *et al.* (2008) Cancer statistics, 2008. *CA Cancer J Clin* 58: 71–96.

- Karakiewicz, P.I., Shariat, S.F., Palapattu, G.S., Gilad, A.E., Lotan, Y., Rogers, C.G. *et al.* (2006a) Nomogram for predicting disease recurrence after radical cystectomy for transitional cell carcinoma of the bladder. *J Urol* 176: 1354–1361; discussion 1361–1352.
- Karakiewicz, P.I., Shariat, S.F., Palapattu, G.S., Perrotte, P., Lotan, Y., Rogers, C.G. *et al.* (2006b) Precystectomy nomogram for prediction of advanced bladder cancer stage. *Eur Urol* 50: 1254–1260; discussion 1261–1252.
- Karam, J.A., Lotan, Y., Karakiewicz, P.I., Ashfaq, R., Sagalowsky, A.I., Roehrborn, C.G. *et al.* (2007) Use of combined apoptosis biomarkers for prediction of bladder cancer recurrence and mortality after radical cystectomy. *Lancet Oncol* 8: 128–136.
- Kattan, M. (2001) Expert systems in medicine. In Smelser, N.J. and Baltes, P.B. (eds) *International encyclopedia of the social and behavioral sciences*. Oxford: Pergamon, 5135–5139.
- Kattan, M.W. (2002) Nomograms. introduction. *Semin Urol Oncol* 20: 79–81.
- Kattan, M.W. (2003a) Comparison of cox regression with other methods for determining prediction models and nomograms. *J Urol* 170: S6–9; discussion S10.
- Kattan, M.W. (2003b) Nomograms are superior to staging and risk grouping systems for identifying high-risk patients: preoperative application in prostate cancer. *Curr Opin Urol* 13: 111–116.
- Kattan, M.W., Eastham, J.A., Stapleton, A.M., Wheeler, T.M. and Scardino, P.T. (1998) A Preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J Natl Cancer Inst* 90: 766–771.
- Kattan, M.W., Karpeh, M.S., Mazumdar, M. and Brennan, M.F. (2003a) Postoperative nomogram for disease-specific survival after an R0 resection for gastric carcinoma. *J Clin Oncol* 21: 3647–3650.
- Kattan, M.W., Leung, D.H. and Brennan, M.F. (2002) Postoperative nomogram for 12-year sarcoma-specific death. *J Clin Oncol* 20: 791–796.
- Kattan, M.W., Shariat, S.F., Andrews, B., Zhu, K., Canto, E., Matsumoto, K. *et al.* (2003b) The addition of interleukin-6 soluble receptor and transforming growth factor beta1 improves a preoperative nomogram for predicting biochemical progression in patients with clinically localized prostate cancer. *J Clin Oncol* 21: 3573–3579.
- Lotan, Y., Gupta, A., Shariat, S.F., Palapattu, G.S., Vazina, A., Karakiewicz, P.I. *et al.* (2005) Lymphovascular invasion is independently associated with overall survival, cause-specific survival, and local and distant recurrence in patients with negative lymph nodes at radical cystectomy. *J Clin Oncol* 23: 6533–6539.
- Margulis, V., Lotan, Y., Montorsi, F. and Shariat, S.F. (2008a) Predicting survival after radical cystectomy for bladder cancer. *Br J Urol Int* 102: 15–22.
- Margulis, V., Lotan, Y. and Shariat, S.F. (2008b) Survivin: a promising biomarker for detection and prognosis of bladder cancer. *World J Urol* 26: 59–65.
- Margulis, V., Shariat, S.F., Ashfaq, R., Sagalowsky, A.I. and Lotan, Y. (2006) Ki-67 is an independent predictor of bladder cancer outcome in patients treated with radical cystectomy for organ-confined disease. *Clin Cancer Res* 12: 7369–7373.
- Miles, B.J., Giesler, B. and Kattan, M.W. (1999) Recall and attitudes in patients with prostate cancer. *Urology* 53: 169–174.
- Nielsen, M.E., Shariat, S.F., Karakiewicz, P.I., Lotan, Y., Rogers, C.G., Amiel, G.E. *et al.* (2007) Advanced age is associated with poorer bladder cancer-specific survival in patients treated with radical cystectomy. *Eur Urol* 51: 699–706; discussion 706–698.
- O'Connor, A.M. (1999) Do shared decision making programs work? a systematic overview. *Med Decis Making* 11: 523.
- Qureshi, K.N., Naguib, R.N., Hamdy, F.C., Neal, D.E. and Mellon, J.K. (2000) Neural network analysis of clinicopathological and molecular markers in bladder cancer. *J Urol* 163: 630–633.
- Rabbani, F., Stapleton, A.M., Kattan, M.W., Wheeler, T.M. and Scardino, P.T. (2000) Factors predicting recovery of erections after radical prostatectomy. *J Urol* 164: 1929–1934.
- Ross, P.L., Gerigk, C., Gonen, M., Yossepowitch, O., Cagiannos, I., Sogani, P.C. *et al.* (2002) Comparisons of nomograms and urologists' predictions in prostate cancer. *Semin Urol Oncol* 20: 82–88.
- Sargent, D.J. (2001) Comparison of artificial neural networks with other statistical approaches: results from medical data sets. *Cancer* 91: 1636–1642.
- Schwarzer, G. and Schumacher, M. (2002) Artificial neural networks for diagnosis and prognosis in prostate cancer. *Semin Urol Oncol* 20: 89–95.
- Shariat, S. F., Ashfaq, R., Sagalowsky, A. I. and Lotan, Y. (2006a) Correlation of cyclin D1 and E1 expression with bladder cancer presence, invasion, progression, and metastasis. *Hum Pathol* 37: 1568–1576.
- Shariat, S. F., Capitanio, U., Jeldres, C. and Karakiewicz, P. I. (2009) Can nomograms be superior to other prediction tools? *Br J Urol Int* 103: 492–495; discussion 495–497.
- Shariat, S.F., Karakiewicz, P.I., Ashfaq, R., Lerner, S.P., Palapattu, G.S., Cote, R.J. *et al.* (2008a) Multiple biomarkers improve prediction of bladder cancer recurrence and mortality in patients undergoing cystectomy. *Cancer* 112: 315–325.



- Shariat, S.F., Karakiewicz, P.I., Margulis, V. and Kattan, M.W. (2008b) Inventory of prostate cancer predictive tools. *Curr Opin Urol* 18: 279–296.
- Shariat, S.F., Karakiewicz, P.I., Palapattu, G.S., Amiel, G.E., Lotan, Y., Rogers, C.G. *et al.* (2006b) Nomograms provide improved accuracy for predicting survival after radical cystectomy. *Clin Cancer Res* 12: 6663–6676.
- Shariat, S.F., Karakiewicz, P.I., Palapattu, G.S., Lotan, Y., Rogers, C.G., Amiel, G.E. *et al.* (2006c) Outcomes of radical cystectomy for transitional cell carcinoma of the bladder: a contemporary series from the bladder cancer research consortium. *J Urol* 176: 2414–2422.
- Shariat, S.F., Karakiewicz, P.I., Roehrborn, C.G. and Kattan, M.W. (2008c) An updated catalog of prostate cancer predictive tools. *Cancer* 113: 3075–3099.
- Shariat, S.F., Karakiewicz, P.I., Suardi, N. and Kattan, M.W. (2008d) Comparison of nomograms with other methods for predicting outcomes in prostate cancer: a critical analysis of the literature. *Clin Cancer Res* 14: 4400–4407.
- Shariat, S.F., Karam, J.A. and Lerner, S.P. (2008e) Molecular markers in bladder cancer. *Curr Opin Urol* 18: 1–8.
- Shariat, S.F., Karam, J.A., Walz, J., Roehrborn, C.G., Montorsi, F., Margulis, V. *et al.* (2008f) Improved prediction of disease relapse after radical prostatectomy through a panel of preoperative blood-based biomarkers. *Clin Cancer Res* 14: 3785–3791.
- Shariat, S.F., Margulis, V. and Karakiewicz, P.I. (2008g) Rebuttal from authors re: james w.f. catto. more nomograms or better evidence of efficacy: what do we need in urologic oncology? *Eur Urol* 2008; 54: 11–12. *Eur Urol* 54: 13–15.
- Shariat, S.F., Margulis, V., Lotan, Y., Montorsi, F. and Karakiewicz, P.I. (2008) Nomograms for bladder cancer. *Eur Urol* 54: 41–53.
- Shariat, S.F., Palapattu, G.S., Karakiewicz, P.I., Rogers, C.G., Vazina, A., Bastian, P.J. *et al.* (2007a) Concomitant carcinoma in situ is a feature of aggressive disease in patients with organ-confined Tcc at radical cystectomy. *Eur Urol* 51: 152–160.
- Shariat, S.F., Palapattu, G.S., Karakiewicz, P.I., Rogers, C.G., Vazina, A., Bastian, P.J. *et al.* (2007b) Discrepancy between clinical and pathologic stage: impact on prognosis after radical cystectomy. *Eur Urol* 51: 137–151.
- Shariat, S.F., Trinh, Q.D., Morey, A.F., Stage, K.H., Roehrborn, C.G., Valiquette, L. *et al.* (2008h) Development of a Highly Accurate Nomogram for Prediction of the Need for Exploration in Patients with Renal Trauma. *J Trauma* 64: 1451–1458.
- Shariat, S.F., Walz, J., Roehrborn, C.G., Montorsi, F., Jeldres, C., Saad, F. *et al.* (2008i) Early Postoperative Plasma Transforming Growth Factor-Beta1 Is a Strong Predictor of Biochemical Progression after Radical Prostatectomy. *J Urol* 179: 1593–1597.
- Shariat, S.F., Walz, J., Roehrborn, C.G., Zlotta, A.R., Perrotte, P., Suardi, N. *et al.* (2008j) External Validation of a Biomarker-Based Preoperative Nomogram Predicts Biochemical Recurrence after Radical Prostatectomy. *J Clin Oncol* 26: 1526–1531.
- Shariat, S.F., Zippe, C., Ludecke, G., Boman, H., Sanchez-Carbayo, M., Casella, R. *et al.* (2005) Nomograms including nuclear matrix protein 22 for prediction of disease recurrence and progression in patients with Ta, T1 or cis transitional cell carcinoma of the bladder. *J Urol* 173: 1518–1525.
- Shariat, S.F., Zlotta, A.R., Ashfaq, R., Sagalowsky, A.I. and Lotan, Y. (2007c) Cooperative effect of cell-cycle regulators expression on bladder cancer development and biologic aggressiveness. *Mod Pathol* 20: 445–459.
- Specht, M.C., Kattan, M.W., Gonen, M., Fey, J. and Van Zee, K.J. (2005) Predicting nonsentinel node status after positive sentinel lymph biopsy for breast cancer: clinicians versus nomogram. *Ann Surg Oncol* 12: 654–659.
- Stein, J.P., Lieskovsky, G., Cote, R., Groshen, S., Feng, A.C., Boyd, S. *et al.* (2001) Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. *J Clin Oncol* 19: 666–675.
- Stephenson, A.J., Scardino, P.T., Eastham, J.A., Bianco Jr, F.J., Dotan, Z.A., Diblasio, C.J. *et al.* (2005) Postoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J Clin Oncol* 23: 7005–7012.
- Stephenson, A.J., Scardino, P.T., Eastham, J.A., Bianco Jr, F.J., Dotan, Z.A., Fearn, P.A. *et al.* (2006) Preoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J Natl Cancer Inst* 98: 715–717.
- Steuber, T., Graefen, M., Haese, A., Erbersdobler, A., Chun, F.K., Schlom, T. *et al.* (2006) Validation of a nomogram for prediction of side specific extracapsular extension at radical prostatectomy. *J Urol* 175: 939–944; discussion 944.
- Steyerberg, E.W., Bleeker, S.E., Moll, H.A., Grobbee, D.E. and Moons, K.G. (2003) Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol* 56: 441–447.
- Steyerberg, E.W., Harrell Jr, F.E., Borsboom, G.J., Eijkemans, M.J., Vergouwe, Y. and Habbema, J.D. (2001) Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 54: 774–781.
- Steyerberg, E.W., Harrell, F.E., Jr. and Goodman, P.H. (1998) Neural Networks, Logistic



Regression, and Calibration. *Med Decis Making* 18: 349–350.

Steyerberg, E.W., Roobol, M.J., Kattan, M.W., Van Der Kwast, T.H., De Koning, H.J. and Schroder, F.H. (2007) Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. *J Urol* 177: 107–112; discussion 112.

Terrin, N., Schmid, C.H., Griffith, J.L., D'agostino, R.B. and Selker, H.P. (2003) External validity of predictive models: a comparison of logistic

regression, classification trees, and neural networks. *J Clin Epidemiol* 56: 721–729.

Vlaev, I. and Chater, N. (2006) Game relativity: how context influences strategic decision making. *J Exp Psychol Learn Mem Cogn* 32: 131–149.

Witjes, J. A. and Hendricksen, K. (2008) Intravesical pharmacotherapy for non-muscle-invasive bladder cancer: a critical analysis of currently available drugs, treatment schedules, and long-term results. *Eur Urol* 53: 45–52.

Visit SAGE journals online  
<http://tau.sagepub.com>

SAGE JOURNALS  
Online