Neurofuzzy Modeling to Determine Recurrence Risk Following Radical Cystectomy for Nonmetastatic Urothelial Carcinoma of the Bladder

James W.F. Catto,1 Maysam F. Abbod,3 Derek A. Linkens,2 Stéphane Larré,1 Derek J. Rosario,1 and Freddie C. Hamdy4

Abstract
Purpose: Bladder cancer recurrence occurs in 40% of patients following radical cystectomy (RC) and pelvic lymphadenectomy (PLND). Although recurrence can be reduced with adjuvant chemotherapy, the toxicity and low response rates of this treatment restrict its use to patients at highest risk. We developed a neurofuzzy model (NFM) to predict disease recurrence following RC and PLND in patients who are not usually administered adjuvant chemotherapy.

Experimental Design: The study comprised 1,034 patients treated with RC and PLND for bladder urothelial carcinoma. Four hundred twenty-five patients were excluded due to lymph node metastases and/or administration of chemotherapy. For the remaining 609 patients, we obtained complete clinicopathologic data relating to their tumor. We trained, tested, and validated two NFMs that predicted risk (Classifier) and timing (Predictor) of post-RC recurrence. We measured the accuracy of our model at various postoperative time points.

Results: Cancer recurrence occurred in 172 (28%) patients. With a median follow-up of 72.7 months, our Classifier NFM identified recurrence with an accuracy of 0.84 (concordance index 0.92, sensitivity 0.81, and specificity 0.85) and an excellent calibration. This was better than two predictive nomograms (0.72 and 0.74 accuracies). The Predictor NFMs identified the timing of tumor recurrence with a median error of 8.15 months.

Conclusions: We have developed an accurate and well-calibrated model to identify disease recurrence following RC and PLND in patients with nonmetastatic bladder urothelial carcinoma. It seems superior to other available predictive methods and could be used to identify patients who would potentially benefit from adjuvant chemotherapy.

Bladder cancer is the fifth commonest malignancy in the United States with an estimated incidence of 67,160 new cases and 13,750 deaths in 2007 (1). The majority of tumors are urothelial carcinoma and clinical evidence suggests that two distinct varieties exist. Whereas most tumors belong to the nonmuscle invasive pathway, around one third are muscle invasive. The prognosis of muscle invasive urothelial carcinoma is poor and strongly dependent on radical treatment and the stage of the disease. Currently, RC with bilateral pelvic lymphadenectomy (PLND) is the gold standard treatment for muscle invasive bladder cancer (2, 3). Following RC and PLND, tumor recurrence occurs in >40% of patients and the response to chemotherapy is poor (4, 5). Metastatic bladder urothelial carcinoma is a rapidly progressing disease with a median time of 13 months from RC to death (2, 6). Currently, the strongest predictors of recurrence risk following RC are pathologic stage and lymph node metastasis.

For patients with advanced bladder urothelial carcinoma, improvements in survival have been made by the addition of extended PLND to RC (3, 6–9) and the use of chemotherapy in either an adjuvant or neoadjuvant setting (10). Both PLND and chemotherapy seem to have the best response rates in the patients with the lowest tumor burden/metastatic volume (9, 11, 12). Whereas the nodal status is important for predicting tumor recurrence following RC and PLND, ~30% of patients without lymphatic metastasis will die from their urothelial carcinoma (2, 3, 6). Primary tumor stage is a strong predictor of this failure with post-RC recurrence rates for organ-confined (~30% at 5 years) and extravesical (50% at 5 years) nonmetastatic disease differing significantly (2, 3, 6). Although patients at high risk of disease recurrence would benefit from administration of adjuvant chemotherapy or inclusion into clinical trials of adjuvant therapy, pathologic stage is not sufficiently accurate to identify patients at high risk. Therefore, development of a predictive tool that integrates multiple tumor
and patient features would help the identification of patients who are likely to experience disease recurrence and therefore would benefit from intensified therapy.

Various predictive strategies have been used within urological oncology and to date the most popular approach is the nomogram. This is developed using Cox proportional hazards regression to measure the relationship between parameters and outcome. Nomograms are easy to use and provide predictions with an accuracy of ~70% to 80% using traditional clinical parameters (13–16). An alternate predictive approach is with artificial intelligence modeling (17). Artificial intelligence techniques are not dependent on numerical linearity and do not infer relationships by statistical proximity. As such, they function well within refined data sets and can recognize complex nonlinear relationships among contaminating noisy data. Although artificial neural networks have been used by the medical community, their hidden working layer prevents model understanding and has led to skepticism within the scientific community (18–20). We have previously used neurofuzzy modeling (NFM) systems to predict the outcome of bladder cancer (21). NFM systems have a similar structure to neural networks but have a transparent fuzzy logic internal layer. As such, the internal layer of the model can be used to facilitate interpretation and ensure clinical sense.

Here, we develop a new NFM using a large multicenter cohort of patients who have undergone RC with PLND. To produce a homogenous cohort, we select only those bladder urothelial carcinoma patients without lymph node metastasis and who did not receive neoadjuvant or adjuvant chemotherapy. Those patients who develop disease recurrence in this population are indistinguishable postoperatively from those with an excellent prognosis following RC and PLND. As such, they often do not receive adjuvant chemotherapy and may enter less intensive surveillance regimens. An accurate predictor of recurrence risk in patients without advanced disease such as lymph node metastasis following RC would optimize the care of these patients.

### Translational Relevance

Following radical cystectomy (RC) with lymphadenectomy for bladder cancer, up to 40% of patients develop tumor recurrence. Treatment of this recurrence is usually palliative and low response rates are seen. To reduce postoperative recurrence, systemic chemotherapy is used in either the adjuvant or neoadjuvant setting. The toxicity and low response rates seen with adjuvant chemotherapy limit the use to patients at highest recurrence risk. This risk is often hard to define and patients with identical histologic findings frequently have different outcomes. Here, we have used artificial intelligence to analyze a large well-defined cohort of patients without nodal metastases (typically defined as low recurrence risk). These patients comprise the postcytectomy group that shows most heterogeneity in outcome (cure or recurrence) and typically do not receive adjuvant chemotherapy. We produce a model that has excellent predictive parameters for defining recurrence in this population. Our model compares favorably with reported nomograms in this field and can be used to identify those patients who would benefit from adjuvant chemotherapy, despite optimistic-looking histopathologic criteria. We feel that this model significantly improves the care of these patients by allowing better treatment targeting and postoperative counseling.

### Patient population and pathologic data

The study comprised 1,034 patients who had undergone RC and PLND for bladder urothelial carcinoma with curative intent between 1984 and 2003 at either The James Buchanan Brady Urological Institute, The Johns Hopkins Hospital, Baltimore, MD; The Dallas and South Department of Urology, Baylor College of Medicine, Houston, TX; or the Department of Urology, University of Texas Southwestern Medical Center, Dallas, TX (2). No patient had distant metastatic disease at cystectomy. Four hundred twenty-five patients were excluded due to lymph node metastases and/or administration of neoadjuvant or adjuvant chemotherapy. This left 609 patients for analysis.

In each center, pathologic specimen assessment was done by staff genitourinary pathologists according to well-defined protocols (22). The 1997 tumor-node-metastasis classification was used for pathologic staging and 1973 WHO classification for tumor grade. For each patient, data were collected using strict criteria in a retrospective and prospective manner and entered into an institutional review board–approved database (2). To ensure information rigor, numerous internal and external data reviews were done and all pathologic data were scored by two separate clinicians with a measured intraclass correlation coefficient of >0.95. This study was done with the approval and oversight of the institutional review board at each institution.

### Postcytectomy care

All patients underwent postoperative surveillance according to protocols described in detail elsewhere (2). In general, physical examination and biochemical analysis were done every 3 to 6 mo for 2 y and then annually thereafter. Radiological imaging was used when indicated and annually for upper urinary tract surveillance. Recurrence-free survival was measured in months from RC to the first documented tumor recurrence or when the patient was last seen, if lost to follow-up. Patients who died before clinical recurrence were censored at the time of death. Hospital charts and physician records were reviewed to obtain these data and cause of death was determined by the treating physician.

### Artificial intelligence

The development of a NFM involves a number of iterative loops to refine the model parameters and component terms, to simplify its structure to the minimum complexity consistent with the model (most parsimonious model) before validating the results. We developed two NFMs to predict the risk (Classifier: 0–100%) and the timing of post-RC tumor recurrence (Predictor: months from RC). The models were combined in series such that patients with a ≥50% risk of recurrence in the Classifier were automatically evaluated by the Predictor NFM. Both models were produced within Matlab (version 6.5) and using in-house software for predictions (23, 24). For model development, the data were divided into 90% for training (of which 60% was learning and 30% for validation) and 10% for testing. External testing was done using data from a single center when the model was trained on the data from other centers. Ensembling and cross-validation were used to maximize data and because they have least bias and error among resampling methods (25). To train the Predictor model only, cases with actual tumor recurrence were used (n = 172). To evaluate the importance of imbalanced outcome data, bootstrapping was used. Within Matlab, the data set was outcome balanced using bootstrapping to create a data set with an overall 50% recurrence rate. A second Bootstrap Corrected Classifier was then generated and tested as described above. Finally,

---

5 [www.mathworks.com](http://www.mathworks.com)
we compared our model predictions with those made using published nomograms for disease recurrence post-RC (13, 26).

Statistical analysis. Two-tailed statistical analyses were done using SPSS (SPSS, Inc., version 14). Categorical variables were compared using the $\chi^2$ test and continuous variables were analyzed with a Mann-Whitney U test. Recurrence-free survival probability following RC was analyzed using the Kaplan-Meier method and the log-rank test. Multivariate analysis for predictors of tumor recurrence was done using Cox regression analysis. A $P$ value of <0.05 was interpreted as statistically significant. To evaluate model accuracy, a concordance index (C-index) was calculated as described (27). Predictive accuracies at various time intervals following RC were tested across a range of probability risks and illustrated using a calibration plot (28).

Results

Description of patient and tumor parameters. The data from the 609 patients are described in Table 1. The majority (79%) of patients had muscle-invasive disease and high histologic tumor grade (85%). Median age was 70 years (range 40-93 years) and most patients were of male gender (81%). Only one patient had a positive surgical margin at surgery and consequently this parameter was excluded from the NFM design. Concomitant carcinoma in situ was present in 44% of specimens and lymphovascular invasion in 29%. All patients underwent a PLND with a median of 18 nodes reported (IQ range 12-25 nodes). Less than 10 nodes were histologically identified in 102 (17%) patients and this suggested threshold (8) was not associated with a significant increase in recurrence rate (32% versus 27%, log-rank $P = 0.14$). Postoperative tumor recurrence occurred in 172 (28%) patients with a median time to recurrence of 10.3 months (range 1-219). The majority of tumor recurrences occurred within the first 2 years post-RC ($n = 128$; 74%) and by 5 years nearly all events had occurred ($n = 155$; 90%). The median follow-up of patients without disease recurrence was 72.7 months (range 1-253 months). The risk of urothelial carcinoma recurrence varied significantly with the presence of extravesical tumor extension when compared with organ-confined tumors (log-rank $P = 0.0001$; Fig. 1). Univariate and multivariate analyses revealed that tumor stage (odds ratio, 1.65; 95% confidence interval, 1.38-1.99; Cox multivariate analysis, $P < 0.0001$), lymphovascular invasion (odds ratio, 2.66; 95% confidence interval, 1.92-3.67; Cox multivariate analysis $P < 0.0001$), and the number of removed lymph nodes (odds ratio, 0.984; 95% confidence interval, 0.97-0.99; Cox multivariate analysis $P = 0.017$), but not grade or the presence of concomitant carcinoma in situ, were associated with disease recurrence.

Description of NFMs. The Classifier NFM accurately identified the recurrence status of 0.84 cases (C-index 0.92, sensitivity 1-220 months). The risk of urothelial carcinoma recurrence varied significantly with the presence of extravesical tumor extension when compared with organ-confined tumors (log-rank $P = 0.0001$; Fig. 1). Univariate and multivariate analyses revealed that tumor stage (odds ratio, 1.65; 95% confidence interval, 1.38-1.99; Cox multivariate analysis, $P < 0.0001$), lymphovascular invasion (odds ratio, 2.66; 95% confidence interval, 1.92-3.67; Cox multivariate analysis $P < 0.0001$), and the number of removed lymph nodes (odds ratio, 0.984; 95% confidence interval, 0.97-0.99; Cox multivariate analysis $P = 0.017$), but not grade or the presence of concomitant carcinoma in situ, were associated with disease recurrence.

<table>
<thead>
<tr>
<th>Table 1. Description of the clinical and pathologic characteristics of the patient cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Included in NFM model</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Pathologic stage</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Pathologic grade</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Lymph node status</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Carcinoma in situ</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Margin status</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>LVI</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Abbreviations: LVI, lymphovascular invasion; NA, not applicable.

Fig. 1. Kaplan-Meier analysis of bladder cancer recurrence stratified for extravesical tumor extension.

Imaging, Diagnosis, Prognosis

Clin Cancer Res 2009;15(9) May 1, 2009

3152 www.aacrjournals.org

Downloaded from clincancerres.aacrjournals.org on April 10, 2017. © 2009 American Association for Cancer Research.
0.81, and specificity 0.85). A calibration plot (Fig. 2) reveals that the model does well over a wide range of recurrence risk probabilities at various time points. The 2-, 5-, 8-, and 20-year post-RC end points represent 74%, 90%, 96%, and 100% of recurrences, respectively. When input parameters were individually analyzed, the model accuracy ranged between 0.91 and 0.81; females (0.91), males (0.83), pT1 (0.88), pT2 (0.84), pT3 (0.81), pT4 (0.90), grade 2 (0.88), grade 3 (0.84), carcinoma in situ present (0.82) or absent (0.86), and lymphovascular invasion present (0.82) or absent (0.85). As recurrence occurred in only 172 (28%) patients, we used bootstrapping to balance the cohort (to 50% recurrence) and trained a new NFM. This Bootstrap Corrected Classifier NFM accurately identified recurrence in 0.76 of cases (C-index 0.92, sensitivity of 0.93, and specificity 0.68).

The Predictor NFM estimated the timing of disease recurrence in the 172 affected patients with a median error of 8.15 months (range 0.235-69.27 months). In Fig. 3, the predictions of disease recurrence are illustrated against actual events using both a scatter plot (Fig. 3A) and Kaplan-Meier plot (Fig. 3B) survival plots. These reveal the close proximity of predicted and actual events and that predictive accuracy does not vary with duration of follow-up. For example, the median error for predicted recurrence timing was 6.3 and 7.1 months for disease recurrence within the first 2 and 5 years post-RC, respectively.

Comparison of the NFM with reported nomograms. Two nomograms have been reported that estimate recurrence risk following RC (13, 26). Using each, we calculated the 5-year post-RC recurrence risk for the 609 patients. We used a threshold of 50% risk to classify the recurrence status for each case. We compared this with the status of each patient in our cohort at 5 years post-RC (n = 155 recurrences). The accuracy of the nomograms varied between 0.72 and 0.74. These results are similar to the original reported accuracy parameters (0.74/0.78 and 0.75) and include an estimate of time to RC for the International Bladder Cancer Nomogram (this variable was missing from our database). Of note, altering the recurrence risk threshold to 60% did not significantly alter the performance of either nomogram (0.73 and 0.74). Our Classifier NFM (accuracy 0.84, C-index 0.92) compares favorably with these nomogram predictions.

Discussion

Recurrence risk following RC and PLND can be difficult to predict but is important as it is used to stratify chemotherapy use. Bladder cancer can be chemoresistant and the highest response rates are seen with multidrug platinum-based regimens. Adjuvant chemotherapy using these regimens results in a 25% relative reduction in the risk of death and leads to a 5% increase in 5-year survival (4, 11, 29). Whereas response is associated with an increase in survival of around 7 months (11, 29), mortality rates of 1% to 3% from chemotherapy are reported (10). As the benefit from chemotherapy is small if used in all patients post-RC, it is rational to use it only in those at highest risk of disease recurrence. Current criteria to identify this risk are pathologic such as lymph node metastasis and extravesical extension. As shown in this report, tumor recurrence occurs in ~30% of patients without nodal metastasis and in 20% of those with organ-confined disease. On the other hand, patients can be cured by RC and PLND monotherapy despite ominous pathologic features. In our series, 58% of patients with extravesical extension did not
develop tumor recurrence, including 43% of those with lymphovascular invasion.

Two previous nomograms have been developed to estimate disease recurrence following RC (13, 26). Due to the relative low volume of RC in most centers and variations in cystectomy outcome, both nomograms are developed from multicenter collaborative cohorts. The generated data sets are large (n = 9,064 and n = 958) and collected from 12 and 3 academic centers, respectively. The larger series includes all histologic varieties of bladder cancer (26% of tumors are nonurothelial carcinoma) and both nomograms evaluated all patients following RC (including those with nodal metastasis and those receiving chemotherapy). The larger series produced a nomogram with a C-index of 0.75 using seven parameters: patient age, gender, histologic type, pathologic stage and grade, and time from diagnosis to treatment. The latter series (i.e., Bladder Cancer Research Consortium) produced a nomogram with an accuracy of 0.75, increasing to 0.78 with bootstrapping, using pathologic stage, histologic features (the presence of lymphovascular invasion or carcinoma in situ), and treatment received (adjuvant chemotherapy or radiotherapy). In our patient cohort, both nomograms had similar overall predictive accuracies to those reported in the original series.

Here, we used the database of the Bladder Cancer Research Consortium to generate our NFM (2). In each case, patient data have been meticulously collected and checked by several physicians before inclusion into the database. Our goal was to create a model that could be used to identify those patients at high risk of post-RC recurrence from a population often thought to be at low risk for disease recurrence (i.e., with no nodal metastasis) who did not receive neoadjuvant or adjuvant chemotherapy. Our selected cohort seems well representative of this patient group with almost identical pathologic features to those reported by other large centers and a PLND of sufficient contemporary standard (6, 8). The overall recurrence rate (~30%) is analogous to that reported for this population elsewhere (3, 8).

The accuracy of the Classifier NFM seems favorable to those reported by the two nomograms. This may reflect either the carefully defined patient cohort or the different methodology used to generate predictions. In general, NFM does well within data that is from homogenous samples, clean, and of limited size. Thus, our cohort seems ideal with minimal dimension imbalance and input parameters that include interval scores (such as number of excised nodes and patient age) and evenly distributed numerical scores (tumor stage). The long median follow-up also allows the occurrence of most events, reducing the inclusion of false-negative cases in the recurrence-free population. Previous reports support the artificial intelligence approach to predictive modeling (reviewed in refs. 17, 18) and suggest that its independence on reliance of data linearity can improve model development. The most common method of artificial intelligence modeling in medicine has been the artificial neural network. Traditionally, these have been constructed using a feed-forward multilayer Perceptron and seem favorable to statistical regression when cancer predictions are compared (30). Reports of artificial intelligence in medicine have suffered from methodologic flaws such as biased cohorts, small sample sizes, and reporting or interrogation problems with respect to the artificial intelligence model (20, 31). We have addressed these problems by using a large data set and NFM. The NFM approach produces a transparent model that avoids the black box phenomena of other artificial intelligence methods. As such, one can see the meaning of the resultant rule base and the importance of various input parameters. It is also possible to interrogate individual variables (21) to check established predictive and biological knowledge. Although our sample size is modest, when compared with the 9,064 used for one post-RC nomogram (26), we created a clean cohort of homogenous cases for model development and used cross-validation with ensembling to maximize the use of each patient (25). Of note, the sample size is sufficient for events of the number of parameters tested (20, 32). A final benefit of the artificial intelligence approach is its ease of accessibility. This accessibility is both in development and application. Artificial intelligence model development is done in a semiautomatic manner, enabling unbiased processing of data while allowing direction from clinical knowledge.

In conclusion, we have developed an accurate predictive model to identify recurrence risk and timing in patients with nonmetastatic bladder urothelial carcinoma following RC and PLND who did not receive chemotherapy. The model could be used to stratify the use of adjuvant chemotheraphy within these patients and deserves evaluation in further cohorts (i.e., external validation).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank the Bladder Cancer Research Consortium for giving us access to their data and allowing us to produce this model. Members of the consortium who specifically contributed to the collection of these data include Pierre I. Karakiewicz, Shahrokh F. Shariat, Yair Lotan, Arthur I. Sagalowsky, Seth P. Lerner, Ganesh S. Palapattu, Patrick J. Bastian, and Mark Schoenberg.

References

Neurofuzzy Modeling to Determine Recurrence Risk Following Radical Cystectomy for Nonmetastatic Urothelial Carcinoma of the Bladder

James W.F. Catto, Maysam F. Abbod, Derek A. Linkens, et al.