Published in final edited form as:

Urol Oncol. 2017 January ; 35(1): 33.e21–33.e26. doi:10.1016/j.urolonc.2016.07.022.

External validation of the pathological nodal staging score in upper tract urothelial carcinoma: a population-based study

Evangelos Xylinas^{1,2,*}, Luis Kluth^{1,3,*}, Malte Rieken^{1,4}, Morgan Rouprêt⁵, Bashir Al Hussein Al Awamlh¹, Thomas Clozel¹, Maxine Sun⁶, Pierre I. Karakiewicz⁶, Mithat Gonen⁷, and Shahrokh F. Shariat^{1,8}

¹Department of Urology, Weill Cornell Medical College, New York-Presbyterian Hospital, New York, NY, USA ²Department of Urology Cochin Hospital, APHP, Paris Descartes University, Paris, France ³Department of Urology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany ⁴Department of Urology, University Hospital Basel, Basel, Switzerland ⁵Department of Urology, Pitie-Salpetriere Hospital, Assistance PubliqueHôpitaux de Paris, Faculté de Médecine Pierre et Marie Curie, University Paris 6, Paris, France ⁶Department of Urology, University of Montreal, Montreal, QC, Canada ⁷Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY ⁸Department of Urology, Medical University of Vienna, Vienna, Austria

Abstract

Objectives—To externally validate our previously developed pathological nodal staging model (pNSS) that allows quantification of the likelihood that a pathologically node-negative patient has, indeed, no lymph node metastasis (LNM).

Patients and methods—We analyzed data from 2768 patients treated with RNU and LND using the SEER database from 1988–2010. We estimated the sensitivity of pathologic nodal staging using a beta-binomial model and developed a new pNSS. Then, we compared these findings with those of the initial cohort.

Results—The mean and median number of LN removed were 5 and 2, respectively (IQR:5) in the validation cohort, while 66.5% of the patients (n=1814) were pN0. Similarly to the development cohort, the probability of missing a LNM decreased as the number of nodes examined increased in the validation cohort. If only a single node was examined, 35% of patients would be misclassified as pN0 while harboring LNM. Even when 5 nodes were examined, 8% would be misclassified. The proportion of having a positive node increased with advancing pathological T-stage in both cohorts. Patients with pT0-Ta-Tis-T1 disease will have more than a

Corresponding Author: Shahrokh F. Shariat, M.D., Professor and Chairman, Department of Urology, University of Vienna, Währinger Gürtel 18-20, A-1090 Vienna, Austria, Tel: +43 1 40400 2615, Fax: +43 1 40400 2332, sfshariat@gmail.com.

*Both authors contributed equally to the manuscript.

Financial disclosure and conflicts of interest: None conflicts to declare.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

95% chance of a correct pathologic nodal staging with two examined nodes in both cohorts. However, if a patient has pT3–T4 more than 12 examined LN are needed to reach 95% accuracy.

Conclusions—We confirmed that the number of examined nodes needed for adequate staging depends on pT-stage. We externally validated in a population-based database our previous pNSS, which could help the clinical decision-making regarding adjuvant chemotherapy administration.

Keywords

upper tract urothelial carcinoma; lymphadenectomy; validation; pathological nodal staging score; SEER

Introduction

Upper-tract urothelial carcinoma (UTUC) is a relatively rare malignancy, accounting for approximately 5% of all urothelial cancers [1]. Up to 30% of patients with muscle-invasive UTUC have metastasis to the regional lymph nodes (LNs) [2, 3], which represents a powerful prognostic factor of survival [3–5]. While lymphadenectomy improves tumor staging, its therapeutic benefit remains controversial [3–5]. In addition to LN status, the extent of lymphadenectomy, the number of LNs examined, the number of positive LNs detected, and the LN density, are suggested to have both prognostic and potentially therapeutic implications [6]. Knowledge of the true LN status, therefore, is important as it influences patient counseling and, more importantly, clinical decision-making regarding follow-up scheduling and adjuvant chemotherapy [7–10].

We have recently developed a model (ie.; pathological nodal staging score) allowing to determine the probability that a patient with pathologic node negative status at RNU truly has no LN metastasis (LNM) using the number of examined LNs and established pathological features (such as pathological tumor stage) [11]. The aim of the current study was to externally validate our model in a population-based cohort of patients using the Surveillance, Epidemiology and End Results database.

Patients and methods

Patient selection and data collection

The development cohort comprised 814 patients who underwent RNU and lymphadenectomy for UTUC between 1994 to 2007 at 7 centers worldwide [11]. In this cohort, LN dissections were examined grossly, and all lymphoid tissue was submitted for histological examination. The extent of LN dissection was at the surgeon's discretion. None of patients received preoperative chemotherapy or radiotherapy. Adjuvant chemotherapy was administered at the clinicians' discretion based on tumor stage, LN involvement and overall health status as well as patients preference.

For the validation cohort, we used the Surveillance, Epidemiology and End Results (SEER) registry data from 1988 to 2010 for our analyses. This period was chosen because SEER did not collect detailed LN data between 1973 and 1987. By the end of the study period, the registry captured approximately 28% of the United States population, and is considered to

be representative of the general population. Patients who underwent RNU for UTUC (codes ICD-0-2 C65.9 and C66.9) were identified. Inclusion criteria consisted of having a diagnosis of UTUC, and documentation of the number of LN examined as well as the number of pathologically positive LN. Patients were excluded from analyses when the tumor grade and stage were unknown, if there was evidence of metastatic disease at diagnosis, or if patients underwent a partial nephroureterectomy.

Statistical Analysis

Overview—Validation of nodal staging scores cannot be carried out using the standard validation methodology of obtaining predictions in an independent data set and comparing them with the observed outcome. This is because there is no gold standard observation, i.e. it is impossible to know if a LN staged as pN0 has indeed no LNM. For this reason we applied similar methodology as our previous work to the validation cohort in order to build a similar pathological nodal staging score (pNSS) [11] and compare the nodal staging scores across the two data sets. The primary endpoint was the probability of incorrect nodal staging as a function of the examined nodes (n). The true nodal status is unascertainable but we can use the information from LN-positive patients to determine if n examined and negative LNs are sufficient to classify a patient as LN-negative. For example, consider a patient with n large and k small but positive (k =number of positive nodes from patients with node involvement): if less than n LNs had been examined there would be a chance that this patient would have been incorrectly deemed LN negative. Conversely, for a patient with small n and large k , even with fewer examined LNs, it is unlikely that nodal disease would have been missed. Hence the data from LN-positive patients is used to interpret the data for the LN-negative patients. The probability that a LN-negative patient has nodal disease can be computed using the following algorithm: (1) Compute the probability of missing a positive node (sensitivity), (2) compute the prevalence and (3) compute the nodal staging score from sensitivity and prevalence [11–14].

Probability of missing a positive LN—Probability of missing a positive LN (one minus the sensitivity) is inherent to the process of pathological detection and as such depends on the number of examined LNs but not on patient characteristics [11–14]. We used a beta-binomial model for this purpose, allowing for heterogeneity in the intensity of nodal spread across the patients [11–14]. Three key assumptions underlie this step: (i) There are no false positives (if the specimen contains a positive LN, it will be correctly identified by the pathologist); (ii) All LNs are exchangeable, that is they all have an equal probability of being involved; (iii) Sensitivity is the same for LN-positive and LN negative patients. These assumptions may not be completely tenable, but we find them to be sufficient approximations to our biological understanding of nodal spread and clinical practice of nodal staging [11–14].

Estimation of Prevalence of Nodal Disease—The observed prevalence (called apparent prevalence hereafter) is an underestimate and needs to be adjusted for false negatives [11–14]. This was done in two steps; first step invokes Assumption 1 and estimates #FN k as a function of k :

where #TP_k is the number of true positives for a given k [11–14]. Since prevalence is not a function of k, the second step obtains the adjusted prevalence by averaging over k:

Estimation of prevalence is stratified by T stage and LVI for pNSS, but this is not explicitly noted in the above formula to avoid cumbersome notation [11–14].

Nodal Staging Score—Adequate staging was assessed by computing NSS, the probability that a pathologically LN negative patient is indeed free of nodal disease:

Confidence Intervals—Precision of the reported estimates was assessed by creating 400 bootstrap samples from the entire data set and replicating the estimation process [15]. The 2.5th and the 97.5th quartiles were used as the lower and upper 95% confidence bounds for the corresponding estimates.

External validation of the initial model—After elaboration of a new model based on the validation cohort, we compared these findings to the development cohort. First, we compared the clinical characteristics of the two populations using the chi-square test for categorical variables and the Kruskal-Wallis test for continuous variables. Second, we compared the probabilities of missing a LN either based on the number of LN removed/examined or combining the number of LN removed/examined and the pathological stage. These comparisons were performed using a z-test. Since there is no analytic form for the standard error of the predicted probabilities we used 200 bootstrap samples separately for the two populations to obtain prediction standard errors, which were then used in conducting the z-test. All statistical analyses were performed using SAS.

Results

Clinicopathological characteristics and LND characteristics

Table 1 shows the clinicopathological characteristics of the development (n=814) and the validation cohort patients (n=2768). The mean and median number of LN removed were 5 and 2, respectively (range:1–88, IQR:5) in the validation cohort, while 66.5% of the patients (n=1814) were pN0. Both were slightly lower rates than the development cohort (p-values<0.001). The number of examined LN is shown stratified by pT-stage in Table 2. The number of examined LN was slightly lower in the validation cohort independently of the pathological stage category (p-values<0.001).

Probability of missing a positive LN

Using our model, the beta-binomial parameters were estimated to be 1.24 (95% CI, 1.22 to 1.26) and 2.89 (95% CI, 2.83 to 2.94) in the validation cohort. We assessed the probability of LN metastasis (one minus the sensitivity) as a function of the number of LN examined (Figure 1). As expected and similarly to the development cohort, the probability of missing LN metastasis decreased as the number of LN examined increased in the validation cohort. Specifically, if only one single LN was examined, 35% of all LN-positive patients would be misclassified as LN negative. Additionally, when 5 LN were examined (median for our data), 8% of patients would be misclassified. Furthermore, even when 8 LN were examined, 5% of patients would be misclassified as being LN negative. Only when the number of LN

examined was 10, the sensitivity of the method surpassed 95%. When compared to the probabilities of missing a LN in the development cohort, a slightly fewer number of LN was needed in the validation cohort to reach the same level of probability; however these differences were not statistically significant (all p-values>0.05).

Pathological nodal staging score

Pathological nodal staging scores (pNSS) are presented in Figure 2. For the validation cohort, in patients with <pT2 UTUC, a likelihood of 95% to predict the correct pathologic diagnosis of locoregional extent was reached by evaluation of two LNs. The same levels of accuracy required 3 and 12 examined LN in pT2 and pT3–T4 patients, respectively. Bootstrap CIs for all the estimates were all within 1% (in absolute terms) of the estimates (data not shown). These findings were similar to the probabilities of the development cohort (all p values >0.05, Figure 2).

Apparent and corrected prevalence of nodal disease

The apparent and corrected prevalence of nodal metastasis stratified by pT-stage in both cohorts are reported in Table 3. In the validation cohort, the apparent prevalence of nodal disease was 35.6%, but accounting for false negatives, the corrected prevalence was 45.5%. Similarly to the development cohort, underestimation of prevalence due to false negatives was observed for all pT-stages and its extent increased by pT-stage category. As many as 60.3% of pT3–T4 patients were estimated to have LN metastasis, up from an apparent rate of 47.6%. Similarly to the probabilities of missing a LN and to the pathological nodal staging scores, no statistically significant differences were found between the validation and the development cohorts with regards to apparent and corrected prevalences of nodal diseases (all p-values>0.05).

Discussion

LN metastasis is the strongest adverse pathological feature with regards to prognosis in patients treated with RNU for UTUC [3–5]. Various investigators have attempted to determine the adequacy of LND including the assessment of the number of LNs removed [5], the number of positive LNs and the concept of LN density [3–6]. We recently developed a pNSS, whereby nodal staging is subjected to the statistical standards of a diagnostic test by computing the false negative rate and using the negative predictive value in order to define a score that captures the adequacy of node-negative classification [11].

We successfully externally validated our pNSS in a US population-based cohort. This simple probabilistic model accurately calculates the probability of freedom from occult LN metastasis as a function of pathological tumor stage and the number of LN examined. Since outcomes prediction based on a physician's experience alone might be subjectively influenced, postoperative models based on standard pathological features have been introduced for use in daily clinical practice to improve decision-making process [16, 17]. However, to date, neither intuition nor predictive models offer personalized risk/benefit analysis integrating threshold probabilities. Our model is a simple tool that could serve as guidance in the postoperative clinical decision-making regarding follow-up scheduling,

patient counseling and potentially administration of adjuvant systemic chemotherapy [7–10]. Our model allows for a direct estimation of the extent of understaging, which enables physician and the patient to engage in shared-decision making and determine an individualized treatment plan including risks and benefits of adjuvant chemotherapy.

Despite the limitations of this validation study, accurate prediction of the true LN status seems feasible based on the pathologic T-stage and the number of LNs examined using the pNSS. In our study, over 30% of patients harbored LN metastasis and the proportion having a positive LN increased proportionally with advancing pT-stage. In addition, we confirmed that every patient treated with RNU for UTUC needs a LND to ensure accurate nodal staging. This remains true even for those patients with non-muscle invasive (pTa-Tis-T1) UTUC. Moreover, the probability of missing a positive LN decreases with increasing number of LNs examined (1 LN examined: 35%, 10 LNs: 4%, 15 LNs: 3%). According to our results, even an extended LN dissection does not assure 100% accuracy with regards to nodal status.

Our study has several limitations. First and foremost are limitations inherent to the validation cohort. We used the SEER database enabling us to test our model in a large cohort of patients treated in both academic and community centers. Limitations of the SEER database are well known and include the lack of information on patient comorbidities, prior radiation, palpable LN or radiologic LN metastasis. These factors may impact the decision to perform LND at the time of RNU. Second, the database is designed to capture treatment type in the first 6 months after diagnosis, and therefore we may not have captured information on patients who underwent delayed treatment. Finally, the SEER database does not provide information on the processing of the LN (en bloc or as packets) and pathological variability. Although we were able to control for numerous potential confounders, we could not control for surgeon's and pathologist's experience, treatment decisions including patient and surgeon preferences, as well as the anatomical template of the preferred LN dissection. Thus, we could not control for the number of sections per LN and whether it was related or not to the size of the LN. Moreover, the number of LNs examined is not an exact surrogate for the extent of LND. In addition, the number of LNs examined is not only a factor of the extent of LND but is also dependent on the pathological evaluation and inherent differences between patients. Conversely, our data reflect a real world multicenter experience. Due to the selection process for LND, our results might be biased since a significant number of patients who did not undergo LND, but potentially had LN metastasis, were excluded from the analyses. Finally, our model is based on assumptions. Although these might seem debatable, every single mathematical model and theory is built on assumptions. Prospective validation of our model can test whether the assumptions were realistic or not. Moreover, in a prospective fashion one could define by decision curve-analyses a threshold of probability that could define the subgroup of patients who would benefit the most from adjuvant chemotherapy administration.

Conclusions

Despite the limitations of our study, we confirmed using the SEER database that the number of examined LN needed for adequate staging depends on pT-stage. We externally validated

in this population-based cohort our pathological nodal staging score that estimates the likelihood of false-negative LNM, which could help to refine clinical decision-making regarding administration of adjuvant chemotherapy. Our model could be easily incorporated into clinical practice.

References

1. Roupret M, et al. European guidelines on upper tract urothelial carcinomas: 2013 update. *Eur Urol*. 2013; 63(6):1059–71. [PubMed: 23540953]
2. Margulis V, et al. Outcomes of radical nephroureterectomy: a series from the Upper Tract Urothelial Carcinoma Collaboration. *Cancer*. 2009; 115(6):1224–33. [PubMed: 19156917]
3. Roscigno M, et al. Lymphadenectomy at the time of nephroureterectomy for upper tract urothelial cancer. *Eur Urol*. 2011; 60(4):776–83. [PubMed: 21798659]
4. Roscigno M, et al. Impact of lymph node dissection on cancer specific survival in patients with upper tract urothelial carcinoma treated with radical nephroureterectomy. *J Urol*. 2009; 181(6): 2482–9. [PubMed: 19371878]
5. Roscigno M, et al. The extent of lymphadenectomy seems to be associated with better survival in patients with nonmetastatic upper-tract urothelial carcinoma: how many lymph nodes should be removed? *Eur Urol*. 2009; 56(3):512–8. [PubMed: 19559518]
6. Bolenz C, et al. Risk stratification of patients with nodal involvement in upper tract urothelial carcinoma: value of lymph-node density. *BJU Int*. 2009; 103(3):302–6. [PubMed: 18990164]
7. Hellenthal NJ, et al. Adjuvant chemotherapy for high risk upper tract urothelial carcinoma: results from the Upper Tract Urothelial Carcinoma Collaboration. *J Urol*. 2009; 182(3):900–6. [PubMed: 19616245]
8. Vassilakopoulou M, et al. Outcomes after adjuvant chemotherapy in the treatment of high-risk urothelial carcinoma of the upper urinary tract (UUT-UC): results from a large multicenter collaborative study. *Cancer*. 2011; 117(24):5500–8. [PubMed: 21638278]
9. Kawashima A, et al. The result of adjuvant chemotherapy for localized pT3 upper urinary tract carcinoma in a multi-institutional study. *World J Urol*. 2012; 30(5):701–6. [PubMed: 21984472]
10. Yafi FA, et al. Adjuvant chemotherapy for upper-tract urothelial carcinoma treated with nephroureterectomy: Assessment of adequate renal function and influence on outcome. *Urol Oncol*. 2013
11. Xylinas E, et al. Prediction of true nodal status in patients with pathological lymph node negative upper tract urothelial carcinoma at radical nephroureterectomy. *J Urol*. 2013; 189(2):468–73. [PubMed: 23253960]
12. Gonen M, Schrag D, Weiser MR. Nodal staging score: a tool to assess adequate staging of node-negative colon cancer. *J Clin Oncol*. 2009; 27(36):6166–71. [PubMed: 19901106]
13. Shariat SF, et al. Clinical nodal staging scores for bladder cancer: a proposal for preoperative risk assessment. *Eur Urol*. 2012; 61(2):237–42. [PubMed: 22033174]
14. Shariat SF, et al. Pathologic Nodal Staging Score for Bladder Cancer: A Decision Tool for Adjuvant Therapy After Radical Cystectomy. *Eur Urol*. 2012
15. Joseph NE, et al. Accuracy of determining nodal negativity in colorectal cancer on the basis of the number of nodes retrieved on resection. *Ann Surg Oncol*. 2003; 10(3):213–8. [PubMed: 12679304]
16. Cha EK, et al. Predicting clinical outcomes after radical nephroureterectomy for upper tract urothelial carcinoma. *Eur Urol*. 2012; 61(4):818–25. [PubMed: 22284969]
17. Roupret M, et al. Prediction of cancer-specific survival after radical nephroureterectomy for upper tract urothelial carcinoma: development of an optimized post-operative nomogram using decision curve analysis. *J Urol*. 2012

Highlights

- To externally validate our previously developed pathological nodal staging model (pNSS) that allows quantification of the likelihood that a pathologically node-negative patient has, indeed, no lymph node metastasis (LNM).
- Similarly to the development cohort, the probability of missing a LNM decreased as the number of nodes examined increased in the validation cohort.
- The proportion of having a positive node increased with advancing pathological T-stage in both cohorts.
- We externally validated in a population-based database our previous pNSS, which could help the clinical decision-making regarding adjuvant chemotherapy administration.

Figure 1. Probability of missing nodal disease as a function of lymph nodes examined in 2768 patients treated with radical nephroureterectomy and lymph node dissection for upper tract urothelial carcinoma.

Figure 2.
Pathological nodal staging score stratified by pT-stage in 814 patients treated with radical nephroureterectomy and lymph node dissection for upper tract urothelial carcinoma.

Table 1

Characteristics of the development and validation cohorts of patients treated with radical nephroureterectomy and lymph node dissection for upper tract urothelial carcinoma.

	Development cohort	Validation cohort	
Characteristics	n= 814	n=2768	
Median age (Range)	68 (27–97)	71 (30–99)	<0.001
Gender (n, %)			
Male	510 (62.7%)	1640 (59.3%)	<0.001
Female	304 (37.3%)	1128 (40.7%)	
Pathologic T-stage (n, %)			
pT0-Ta-Tis-T1	266 (33%)	554 (20%)	<0.001
pT2	163 (20%)	576 (20.8%)	
pT3–T4	385 (47%)	1638 (59.2%)	
Pathologic N-stage (n, %)			
pN0	593 (73%)	1814 (65.5%)	<0.001
pN+	221 (27%)	954 (34.5%)	
Median number of examined lymph nodes (n, range, IQR)			
All patients	5 (1–46, 8)	2 (1–88, 5)	<0.001
pN0 patients	5 (1–41, 8)	2 (1–88, 4)	
pN+ patients	5 (1–46, 8)	2 (1–54, 5)	
Mean number of examined lymph nodes (n, SD)			
All patients	6.8 (6.5)	5.0 (7.0)	<0.001
pN0 patients	6.9 (6.5)	5.0 (7.0)	
pN+ patients	6.5 (6.7)	5.1 (6.9)	

IQR: Interquartile range

SD: Standard deviation

Table 2

Number of examined lymph nodes in the development and validation cohorts of patients treated with radical nephroureterectomy and lymph node dissection for upper tract urothelial carcinoma stratified by pT-stage.

	pT stage	n	%	Mean	SD	Median	IQR
Development cohort	pT0-Ta-Tis-T1	266	33	7.0	6.7	5.0	7
	pT2	163	20	7.6	6.07	6.0	8
	pT3-T4	385	47	6.3	6.6	4.0	7
Validation cohort	pT0-Ta-Tis-T1	554		5.0	7.4	2.0	5
	pT2	576		4.9	6.4	2.0	4
	pT3-T4	1638		5.0	7.0	2.0	5

SD: Standard Deviation; IQR: Interquartile Range.

Table 3

Apparent and corrected prevalence of nodal disease in the development and validation cohorts of patients treated with radical nephroureterectomy and lymph node dissection for upper tract urothelial carcinoma.

	Development cohort		Validation cohort	
	Apparent Prevalence	Corrected Prevalence	Apparent Prevalence	Corrected Prevalence
Overall cohort	30.8%	38.9%	35.6%	45.5%
pT Stage				
T0-Ta-Tis -T1	13.5%	15.4%	12.3%	16.8%
T2	21.6%	26.5%	21.3%	27.7%
T3–T4	47.3%	61.4%	47.6%	60.3%

Abbreviation: LVI: Lymphovascular invasion

Apparent prevalence is based on the final pathologic stage regardless of the number of negative lymph nodes. Corrected prevalence takes into account the probability of false negative findings based on the number of negative lymph nodes.