Validation of 1997 Partin Tables' lymph node invasion predictions in men treated with radical prostatectomy in Montreal Quebec

Pierre I Karakiewicz, MD, Jean-Baptiste Lattouf, MD, Paul Perrotte, MD, Luc Valiquette, MD, François Bénard, MD, Michael McCormack, MD, Catherine Ménard, MD, Thierry Lebeau, MD, Serge Benayoun, MD, Alvaro Ramirez, MD, Simon Ouaknine, MD, Fred Saad, MD

Department of Surgery (Urology), Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada

KARAKIEWICZ PI, LATTOUF J-B, PERROTTE P, VALIQUETTE L, BENARD F, MCCORMACK M, MENARD C, LEBEAU T, BENAYOUN S, RAMIREZ A, OUAKNINE S, SAAD F. Validation of 1997 Partin Tables' lymph node invasion predictions in men treated with radical prostatectomy in Montreal Quebec. The Canadian Journal of Urology. 2005;12(2):2588-2592.

Objective: *The accuracy of 1997 Partin Tables' lymph node invasion (LNI) predictions exhibits important variability in different testing populations. We explored the LNI predictive accuracy in radical prostatectomy (RP) patients from Montreal, Canada. Moreover, we assessed the extent of change in predictive accuracy related to a modification of PSA coding from categorical to continuous.*
Methods: *We used pretreatment serum PSA, clinical stage, and biopsy Gleason sum from 537 men treated with RP to compare predicted and observed rates of LNI. Accuracy was quantified with receiver-operating characteristics curves.*
Results: *Accuracy was 0.760 in 369 evaluable patients, when categorically coded pretreatment PSA (0-4, 4.1-10, 10.1-20, 20.1+) was combined with clinical stage and biopsy Gleason sum. A 2.7% accuracy increase was noted when categorically coded PSA was replaced with continuously coded values.*
Conclusion: *Partin Tables' LNI predictions showed comparable accuracy to a community-based sample from the United States (0.766), and to a recent, multi-institutional sample (0.740). However, accuracy was lower than reported in internal (0.818), and external (0.837) academic, validation cohorts. Accuracy of LNI predictions was appreciably higher, when continuously coded PSA was used.*

Key Words: prostate cancer, lymph node invasion, Partin Tables

Introduction

In recent years, several statistical models have been introduced, which predict either pathological stage or cancer control outcome after radical prostatectomy (RP).[1] Many of them are based on pretreatment tumor characteristics.[1] The Partin Tables predict organ confined disease, capsular invasion, seminal vesicle involvement, and pelvic lymph node invasion (LNI) at radical prostatectomy and rely on clinical stage, pretreatment prostate specific antigen (PSA) and biopsy Gleason grade.[2-4] The validity of the Tables has been tested in patients from the United-States and Europe. These studies suggested that their performance might be affected by case-mix characteristics.[5-9] Recently, Penson et al demonstrated that in a community-based sample the accuracy of the Tables is lower than in tertiary care settings.[7] These findings indicate that population-specific validation of the Tables represents an essential step before

Accepted for publication February 2005

Address correspondence to Dr. Pierre I. Karakiewicz, Prognostics and Outcomes Unit, CHUM, 1058, rue St-Denis, Montreal, Quebec H2X 3J4 Canada

implementing the Tables into routine clinical practice.[10] Thus, we decided to test the predictive accuracy of the Partin Tables' LNI predictions in men from Montreal, Canada. In addition to testing the standard version of the Tables, where categorically coded serum PSA values are used, we decided to explore whether the accuracy may be enhanced, if serum PSA is coded as a continuous variable, as suggested by Cagiannos.[11]

Materials and methods

Between 1985 and 2000, 537 patients with clinically localized prostate cancer underwent RP at a tertiary referral center. Patients with missing clinical stage (36), pretreatment PSA level (41), biopsy Gleason (133) or lymph node invasion (5) were excluded, which left 369 evaluable men. No patient received androgen ablation or radiotherapy. Clinical stage, pretreatment serum PSA, biopsy Gleason sum and LNI were recorded for each patient. Clinical stage was assigned using the 1992 AJCC/TNM guidelines based on digital rectal examination, regardless of the results of ultrasonography or other imaging techniques. PSA was measured using the Hybritech PSA assay (Hybritech Tandem R, San Diego, California). Standard template pelvic lymph node dissection was performed, encompassing all nodal tissues from the medial inferior margin of the external iliac vein down to the internal iliac and obturator vessels. Lymph node dissection specimens were analyzed by staff pathologists.

Statistical tests were performed using SPSS, version 10 (SPSS, Inc., Chicago, Illinois) and S-Plus Professional (MathSoft Inc. Seattle, Washington) software. The association between predictor variables (clinical stage, pretreatment PSA, and biopsy Gleason sum) and the observed rate of LNI was tested in multivariate binary logistic regression analyses. The predictive accuracy of LNI was compared between Partin Tables with categorized PSA and the modified approach, where continuously coded pretreatment PSA was used. Predictive accuracy was quantified with receiver operating characteristics-derived area under the curve. The extent of overestimation or underestimation of the observed LNI rate was explored graphically with nonparametric, local regression (Loess) smoothing technique.

Results

The clinical and LNI characteristics of our patient population are described in Table 1. The lymph node dissection specimens contained from 2 to 11 nodes (mean 4 nodes). Table 2 shows univariate and multivariate logistic regression analyses predicting the probability of LNI, based on clinical stage, continuously coded and categorized PSA, as well as biopsy Gleason sum. In the multivariate model, where PSA was coded as a continuous variable, PSA demonstrated a strong association with LNI (p=0.001). In the multivariate model, where PSA was coded as a categorical variable, PSA demonstrated lack of statistically significant association with LNI (p=0.286). Clinical stage and biopsy Gleason sum demonstrated higher statistical significance, when PSA was coded as a continuous variable, relative to categorically coded PSA.

TABLE 1. **Percentage distribution of clinical and pathological stage variables**

Variable	Percentage distribution of 369 evaluable patients
Clinical stage	
T1c	32.3
T2a-c	66.9
T3	0.8
PSA (ng/ml)	
0 – 4.0	12.3
4.1 – 10.0	52.2
10.1 – 20	25.6
> 20	9.9
Mean	(10.667 ng/ml)
Std. dev	(9.809 ng/ml)
Biopsy Gleason sum	
2 – 4	17.6
5	19.8
6	41.3
7	16.1
8 – 9	5.2
LNI	
Positive	6.0
Negative	94.0

Bootstrap corrected predictive accuracy of LNI predictions using PSA categories, as suggested by Partin, was 0.760 Table 3.[4] Use of continuously coded PSA, as suggested by Cagiannos, resulted in bootstrap corrected predictive accuracy of 0.787 (2.7% accuracy gain).[11]

Figure 1 shows the performance characteristics of LNI predictions, using either categorically coded PSA, as suggested by Partin et al Figure 1a, or continuously coded serum PSA Figure 1b. The apparent (non-bias corrected), and 200 bootstrap corrected (bias-corrected) performance characteristics of LNI

TABLE 2. **Univariate and multivariate logistic regression analyses for prediction of LNI based on PSA, clinical stage, and Gleason sum**

	Univariate model		Multivariate model Categorized PSA - according to 1997 Partin Tables -		Multivariate model Continuously coded PSA - according to Cagiannos et al 2003 -	
	OR	**P value**	**OR**	**P value**	**OR**	**P value**
PSA (ng/ml)	-	0.083	-	0.230		
0 – 4.0					N/A	N/A
4.1 – 10.0	1.204	0.814	0.807	0.801	N/A	N/A
10.1 – 20	2.543	0.238	2.065	0.383	N/A	N/A
Greater than 20	4.116	0.092	2.667	0.286	N/A	N/A
PSA						
Continuously coded	1.032	0.014	N/A	N/A	1.070	0.001
Clinical stage	-	0.060	-	0.093	-	0.049
T1c						
T2a-c	4.303	0.018	5.387	0.030	7.135	0.014
T3	0.039	0.860	0.069	0.917	0.146	0.940
Biopsy Gleason sum	-	0.088	-	0.286	-	0.251
2 – 4						
5	1.789	0.509	1.466	0.670	1.483	0.663
6	1.732	0.494	1.355	0.714	1.350	0.720
7	3.517	0.132	1.774	0.524	1.733	0.545
8 – 9	7.999	0.022	6.584	0.057	7.168	0.050

predictions with categorically coded serum PSA are appreciably farther away from what would be considered an ideal prediction (diagonal line, with a slope of 1), than are predictions made with continuously coded serum PSA. Predictions relying on categorically coded PSA underestimate the observed rate of LNI, for observed LNI rates in excess of 10%. When predictions are made with continuously coded PSA, fewer departures from an ideal prediction are made, as shown Figure 1a.

TABLE 3. **Accuracy of studies addressing LNI predictions**

Tables	**LNI prediction using categorized PSA**	**LNI prediction using continuously coded PSA**
Partin et al	0.818	-
Blute et al	0.837	-
Penson et al	0.766	-
Augustin et al	0.799	-
Cagiannos	0.740	0.760
Present series	0.760	0.787

Discussion

Established presence of LNI represents a pivotal point in prostate cancer treatment. For men who are candidates for RP, it usually implies aborting of the surgery.[11] Alternatively, immediate or delayed hormonal ablation is recommended.[11] Until recently, pelvic lymph node dissections were routinely performed. As of 1997, the Partin Tables allowed many urologic surgeons to decide before the surgery, whether pelvic lymph node dissection is indicated.[4] Although, Partin's LNI predictions have been internally and externally validated, a recent report suggested that their accuracy may vary from cohort to cohort Table 3.[4,5,7,9]

Previous studies that addressed the accuracy of Partin predictions were based on select academic cohorts, and most reported very favorable results.[5,9] Penson reported markedly lower accuracy relative to previous reports.[7] Penson's testing cohort was community-based and differed from academic cohorts with respect to clinical stage and Gleason grade.[4,5]

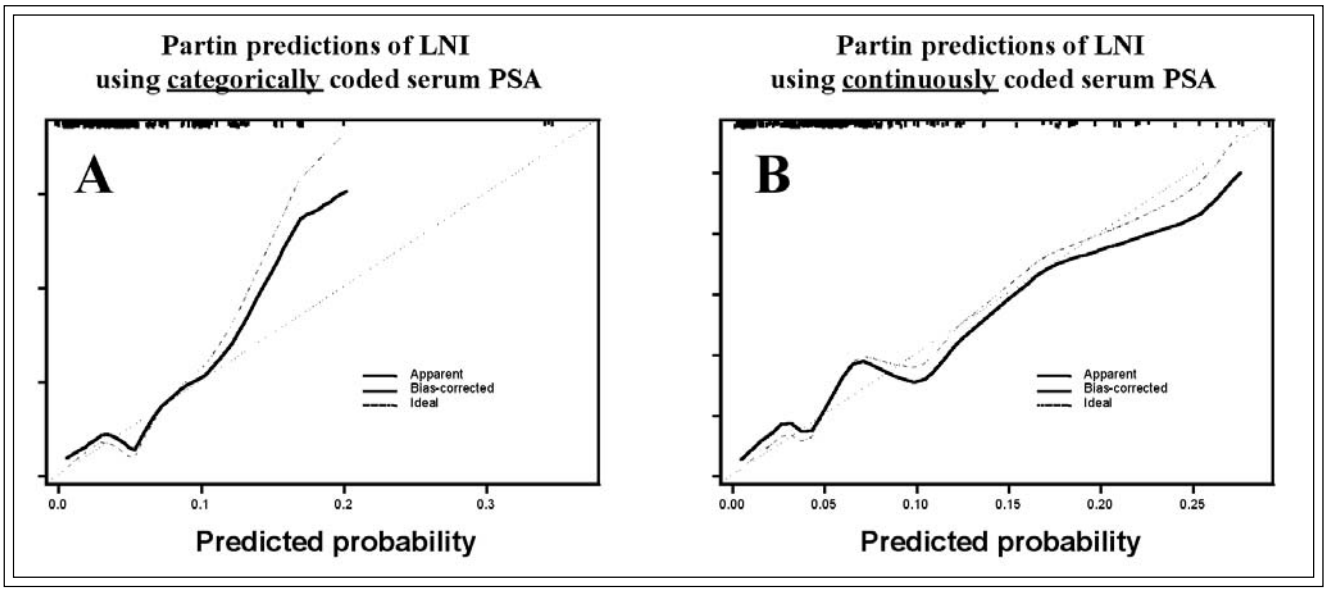

Figure 1a and 1b. Loess local regression non-parametric smoothing plots of the performance of 1997 Partin Tables' LNI predictions relative to observed pathologic rate of LNI. The x-axes represent the Tables' predicted probabilities, whereas the y-axes represent the observed rate of LNI. The plots represent smoothed lines between continuous predictors and binary outcome variables. Within each panel, a perfect prediction is depicted by a diagonal, 45 degree line (Ideal). The non-bootstrap corrected predictions are depicted by the broken line (Apparent), and finally the bootstrap corrected predictions are depicted by the solid line (Bias-corrected). Predictions situated below ideal predictions indicate overestimation by the model. Predictions situated above ideal predictions indicate underestimation by the model. In figure 1A, LNI predictions that are based on categorically coded serum PSA demonstrate a more appreciable departure from what would be considered as ideal predictions, than in figure 1B, which demonstrates predictions based on continuously coded PSA.

Penson's observations suggest that the accuracy of the Tables may be affected by case-mix characteristics. The accuracy of LNI predictions may therefore also differ when applied to men from Montreal, based on their PSA or Gleason score distribution differences.[4,11] The 6% LNI rate seen in our cohort is comparable to the 5% rate reported by Partin.[4] However, it is slightly higher than in most recent radical prostatectomy series (average of 3.7%).[11] Despite this LNI rate similarity, our predictive accuracy is appreciably lower than that reported by Partin (0.760 versus 0.818). This series' LNI predictive accuracy (0.767) closely resembles that reported by Cagiannos (0.740).[11] Therefore, LNI rate does not appear to substantially affect predictive accuracy. Conversely, differences in population characteristics, such as biopsy Gleason sum or serum PSA (for example, in Montreal 35.5% had PSA>10 versus 28.6% in 1997 Partin cohort) appeared to exert a more pronounced effect.

The second goal of our analysis was to test whether a modified version of LNI predictions could yield better predictive accuracy, relative to the 1997 version of LNI Partin predictions.[11] Cagiannos suggested that LNI predictions are more accurate if a continuous PSA scale is used, instead of Partin Tables' PSA categories.[11] This hypothesis is based on the assumption that the introduction of artificial strata in a continuous variable may limit its predictive accuracy.[12]

When PSA was coded as a categorical variable, the predictive accuracy in our cohort was 76.0%. When we used PSA as a continuous variable, accuracy increased to 78.7%. This 2.7% gain in predictive accuracy confirms the findings of Cagiannos et al, where a 2% gain was noted.[11] Moreover, these findings reinforce the generally accepted view, that predictor variables are most informative when introduction of artificial floor and ceiling effects is avoided.12

As shown in Figures 1a and 1b, the gain in predictive accuracy stems from fewer departures from what would be considered as an ideal prediction. The use of continuously coded serum PSA Figure 1b limited the rate of underestimation of the observed LNI rate, which was seen when categorically coded serum PSA was used Figure 1a. Our findings indicate that LNI predictions are reasonably accurate in men treated with radical prostatectomy in Montreal. Moreover, our data suggest that superior predictive

accuracy may be expected when variables are used in the most unaltered fashion, i.e., when introduction of artificial strata is avoided.

Small sample size represents a limitation of our study and likely contributed to lack of significance of biopsy Gleason sum in multivariate models. Moreover, missing values further undermined the significance level of some predictor variables. These phenomena are readily noticeable when overall, or stratum-specific Gleason sum p-values are examined Table 2. It is of interest to note that sample size and missing values have a far lesser impact on nomogram predictive accuracy, where a cumulative linear predictor is defined according to the combined contribution of all predictors. Under such testing conditions, one degree of freedom is necessary, instead of one degree of freedom for each category of stratified variables, or for each continuous predictor.

The observed absence of significant association between Gleason sum and LNI might be related to pathologic Gleason grading of biopsy specimens. Low Gleason grades, corresponding to sums between 2 and 4 were noted in 13.2% of biopsy samples, which is higher than in most institutions, where large numbers of radical prostatectomies are performed.[11] For example, in Hamburg, Gleason sums between 2 and 4 were not assigned to a single patient.[11] At Memorial Sloan-Kettering, 3.7% had Gleason sums 2-4, and 4.1% were diagnosed as such at Cleveland Clinic Foundation.[11]

Finally, we decided to use the 1997 Tables as a benchmark. The rationale for not selecting the 2001 Tables stemmed from fewer validation studies that addressed these updated Tables.[14] Moreover, we have recently shown that the 2001 Tables have the same accuracy and performance characteristics as the original 1997 version.[15]

Conclusion

Despite potential limitations, such as sample size, or lack of central pathologic review, our findings indicate that LNI predictions are 76% accurate in men treated with radical prostatectomy in Quebec, Canada. Moreover, our data suggest that superior predictive accuracy may be expected when predictor variables are used in the most unaltered fashion, i.e., when introduction of artificial strata is avoided. In men treated in Montreal the use of continuously coded PSA, instead of categorically coded PSA, was associated with a predictive accuracy gain of 2.7%. This increment is similar to the one reported by Cagiannos et al (2%), and suggests that predictive models should rely on variables in the most unaltered format.[11,13] □

References

1. Ross LR, Scardino PT, Kattan MW. A catalog of prostate cancer nomograms. *J Urol* 2001;165:1562.
2. Partin AW, Yoo J, Carter HB, Pearson JD, Chan DW, Epstein JI et al. The use of prostate specific antigen, clinical stage and Gleason score to predict pathological stage in men with localized prostate cancer. *J Urol* 1993;150:110.
3. Partin AW, Carter HB. The use of prostate-specific antigen and free/total prostate-specific antigen in the diagnosis of localized prostate cancer. *Urol Clin North Am* 1996;23:531.
4. Partin AW, Kattan MW, Subong ENP, Walsh PC, Wojno KJ, Oesterling JE et al. Combinaison of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. *JAMA* 1997;277:1445.
5. Blute ML, Bergstralh EJ, Partin AW, Walsh PC, Kattan MW, Scardino PT et al. Validation of Partin Tables for predicting pathological stage of clinically localized prostate cancer. *J Urol* 2000;164:1591.
6. Graefen M, Augustin H, Karakiewicz PI, Hammerer PG, Haese A, Palisaar J et al. Can predictive models for prostate cancer patients derived in the United States of America be utilized in European patients? A validation study of the Partin Tables. *Eur Urol* 2003;43:6.
7. Penson DF, Grossfeld GD, Li YP, Henning JM, Lubeck DP, Carroll PR. How well does the Partin nomogram predict pathological stage after radial prostatectomy in a community based population? Results of the cancer of the prostate strategic urological research endeavor. *J Urol* 2002; 167:1653.
8. Graefen M, Karakiewicz PI, Cagiannos I, Quinn DI, Henshall SM, Grygiel JJ, Sutherland RL, Stricker PD, Klein E, Kupelian P, Skinner DG, Lieskovsky G, Bochner B, Huland H, Hammerer PG, Haese A, Erbersdobler A, Eastham JA, De Kernion J, Cangiano T, Schroder FH, Wildhagen MF, Van Der Kwast TH, Scardino PT, Kattan MW. International validation of a preoperative nomogram for prostate cancer recurrence after radical prostatectomy. *J Clin Oncol* 2002;20:3206-3212.
9. Augustin H, Eggert T, Wenske Sven, Karakiewicz PI, Palisaar J, Daghofer F et al. Comparison of accuracy between the Partin Tables of 1997 and 2001 to predict final pathological stage in clinically localized prostate cancer. *J Urol* 2004;171:177.
10. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Inter Med* 1999;130:515.
11. Cagiannos I, Karakiewicz P, Eastham JA, Ohori M, Rabbani F, Gerigk C, Reuter V, Graefen M, Hammerer PG, Erbersdobler A, Huland H, Kupelian P, Klein E, Quinn DI, Henshall SM, Grygiel JJ, Sutherland RL, Stricker PD, Morash CG, Scardino PT, Kattan MW. A preoperative nomogram identifying decreased risk of positive pelvic lymph nodes in patients with prostate cancer. *J Urol* 2003;170:1798.
12. Harrell Fe, Lee KL, Mark DB. Multivariable prognostic models: issues in developping models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-387.
13. Kattan MW, Eastham JA, Stapleton AM, Wheeler TM, Scardino PT. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J Natl Cancer Inst* 1998;90:766-771.
14. Partin AW, Mangold LA, Lamm DM, Walsh PC, Epstein JI, Pearson JD. Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium. *Urology* 2001;58(6):843-848.
15. Augustin H, Eggert T, Wenske S, Karakiewicz PI, Palisaar J, Daghofer F, Huland H, Graefen M. Comparison of accuracy between the Partin tables of 1997 and 2001 to predict final pathological stage in clinically localized prostate cancer. *J Urol* 2004;171(1):177-181.