Comparison of accuracy among three generations of Partin tables in a Chinese cohort

Wen-Jun Xiao, MD,[1,3] Ding-Wei Ye, MD,[1,3] Xu-Dong Yao, MD,[1,3] Shi-Lin Zhang, MD,[1,3] Bo Dai, MD,[1,3] Chao-Fu Wang, MD,[2,3] Jian Wang, MD,[2,3] Hai-Liang Zhang, MD,[1,3] Yi-Jun Shen, MD,[1,3] Yao Zhu, MD,[1,3] Yi-Ping Zhu, MD,[1,3] Guo-Hai Shi, MD,[1,3] Chun-Guang Ma, MD,[1,3] Xiao-Jian Qin, MD,[1,3] Guo-Wen Lin, MD[1,3]

[1]Department of Urology, Fudan University Shanghai Cancer Centre, Shanghai, China
[2]Department of Pathology, Fudan University Shanghai Cancer Centre, Shanghai, China
[3]Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China

XIAO W-J, YE D-W, YAO X-D, ZHANG S-L, DAI B, WANG C-F, WANG J, ZHANG H-L, SHEN Y-J, ZHU Y, ZHU Y-P, SHI G-H, MA C-G, QIN X-J, LIN G-W. Comparison of accuracy among three generations of Partin tables in a Chinese cohort. The Canadian Journal of Urology. 2011;18(2):5619-5624.

Introduction: *To perform a head to head comparison among three generations of Partin tables, namely from 1997, 2001 and the last updated version of 2007, in a Chinese cohort of prostate cancer.*

Material and methods: *Clinical and pathological data of 198 consecutive Chinese patients were retrospectively analyzed, who underwent radical prostatectomy for clinically localized prostate cancer between January 2005 and May 2010. Three versions of the Partin tables were compared for their accuracy and performance to predict final pathological stage using receiver operating characteristic (ROC) curve.*

Results: *Of the whole cohort 58.6% were presented with organ-confined disease (OCD), 10.1% had lymph node involvement (LNI), and 31.3% had locally advanced disease (LAD), while 21.2% had extraprostatic extension (ECE) and 10.1% showed seminal vesicle involvement (SVI). The area under the ROC curve (AUC) of the Partin Tables 1997, 2001 and 2007 was 0.732, 0.722 and 0.695 for OCD; 0.647, 0.594 and 0.577 for LAD; 0.856, 0.872 and 0.829 for LNI, respectively.*

Conclusion: *All three generations of the Partin tables showed a good accuracy to predict OCD, and LNI. However, the predictive accuracy for LAD was more limited. Overall, the newer versions of the Partin tables could not exceed the version of 1997 in their predictive accuracy for the present Chinese cohort. Our results suggest caution when using newly introduced predictive tools that are not supported by population-specific accuracy tests.*

Key Words: predictive tests, prostate cancer, prostatectomy, staging

Introduction

Because radical prostatectomy (RP) is most effective for organ-confined prostate cancer, the key question is how to predict the pathologic stage using the preoperative information. Basically, counseling and proper treatment selection requires precise clinical staging and accurate prediction of final pathological stage. Traditionally, a physician's judgment was based on knowledge and experience about preceding similar courses and clinical judgment may be biased. So Partin and colleagues pioneered the field of predictive tools with the development of their probability tables. The 1997 Partin tables, based upon data from 4133 American patients with prostate cancer, uses three preoperative variables, serum prostate-specific antigen (PSA) level, biopsy Gleason score, and clinical stage, to predict four pathologic stages: organ-confined disease (OCD), established capsular penetration (ECP), seminal vesicle involvement (SVI), and lymph node involvement (LNI).[1] This version demonstrated a good performance and reliability not only with other United States (US) patients[2] but also with European patients.[3] In 2001, this predictive tool was updated to reflect a more contemporary condition

Accepted for publication December 2010

Address correspondence to Dr. Ding-Wei Ye, 86-21-64175590-2808, No. 270, Dong'an Road, Shanghai 200032 China

of prostate cancer stage at diagnosis,[4] and their reliability was tested in different clinical scenarios in Europe[5-7] and Asian.[8,9] In 2007, this predictive tool was updated[10] again, their reliability was also tested in different clinical scenarios in North America[11,12] and Europe[13-16] but not in China. Because of the different biologic features and the detection procedure of prostate cancer used in China (i.e., the obviously lower incidence and no PSA screening for the public), such US-derived predictive tools should be tested in China. Moreover, it is also necessary to perform a direct head to head comparison of the three versions of the Partin tables to test the superiority of the newer versions in a Chinese cohort. Therefore, we determined the overall predictive accuracies and estimated the performance characteristics with the same Chinese sample and compared those directly.

Materials and methods

From January 2005 to May 2010, 212 patients underwent RP for localized prostate cancer (confirmed by 10-core biopsy) at the Department of Urology, Shanghai Cancer Center, an academic teaching hospital of Fudan University. These patients had not received neoadjuvant hormonal therapy that could affect the stage or grade of prostate cancer at RP. Patients who had missing clinical stage (n = 6), pretreatment PSA level (n = 5), or Gleason score information (n = 4) were excluded, resulting in a cohort of 198 consecutive patients available for validation. However, patients with clinical stage T1a-T1b (n = 2) and T3a (n = 8) were also excluded from the test of the 2001 and 2007 Partin tables, because predictive values were no longer provided for patients with these disease stages in the 2001 and 2007 version. Besides, when LNI were predicted using the 2001 and 2007 Partin tables, the nomograms could not provide a probability for another four patients.

All pretreatment PSA levels were measured before prostatic manipulation. The clinical stage mainly based upon digital rectal examination (DRE) was assigned by attending urologists using the 2002 American Joint Commission on Cancer (AJCC) TNM staging system. Dedicated genitourinary pathologists examined all biopsy and RP specimens and assigned the pathologic staging into OCD, ECP, SVI and LNI categories and was determined using the criteria described by Partin et al.[1,4,10] In the analysis of the present study locally advanced disease was defined as evidence of cancer outside the prostatic capsule or the seminal vesicles invasion without lymph node involvement, including ECP and SVI.

Statistical calculations were performed using the Statistical Package for Social Sciences software, version 11.0 (SPSS, Chicago, IL, USA). Two-sided tests with significance at $p = 0.05$ were used. The predictive accuracies of the three generation Partin tables were quantified using receiver operating characteristics (ROC) analysis and area under the ROC curve (AUC) assessment, where a value of 100 represents perfect predictions and 50 indicates a chance phenomenon.

Results

All patients in our cohort were Chinese, and the median age was 66.95 (ranged 48-79). The comparison of the distribution of clinical and pathologic variables for the three generations of Partin tables and the present study is presented in Table 1. The present cohort contained the lowest percentage of patients with non-palpable T1a-c disease (27.3%) and the lowest percentage of patients with a PSA value less than 4.1 ng/mL (1%). Correspondingly, the present cohort displayed a high percentage of patients with a PSA value greater than 20 ng/mL (51.5%). The Partin cohorts showed lower percentages of biopsy Gleason scores greater than 6. As the pathological outcomes after RP were concerned, the present study had the highest rates of SVI and LNI (both 10.1%).

To analyze the discriminative ability of the three generations of Partin tables, ROC analyses for OCD, LAD, and LNI were performed. Figure 1 graphically shows ROC derived AUCs of the predictive accuracy of the Partin tables' pathological stage predictions, relative to the observed stage. Table 2 lists the AUC separately for each stage and each version of the Partin tables. For OCD, the Partin tables presented a relatively similar AUC (1997: 0.732, 2001: 0.722 and 2007: 0.695). For LAD, the AUCs of all three versions only revealed a modest or low overall accuracy AUC (1997: 0.647, 2001: 0.594 and 2007: 0.577). For LNI, almost similar AUCs were presented (1997: 0.856, 2001: 0.872 and 2007: 0.829). These statistical data also showed that the 1997 Partin tables had reasonably better predictive accuracy in the present cohort.

Discussion

The prediction of the pathologic prostate cancer stage is of great importance in treatment decision making. The likelihood of an individual patient having OCD versus non-OCD influences the selection of an appropriate therapeutic intervention. Because imaging studies are not accurate for staging prostate cancer, preoperative clinical and pathologic parameters are often used

TABLE 1. **Comparative analysis of two cohorts**

	Present cohort	1997 Partin cohort[1]	2001 Partin cohort[4]	2007 Partin cohort[10]
Patients, n	198	4133	5079	5730
Age, yrs				
Mean	66.95	n/a	57.9	57.4
Range	48-79	n/a	42-74	34-75
PSA (ng/mL), n (%)				
0-2.5	4 (2.0)	943 (23) (0-4.0)	355 (7)	452 (8)
2.6-4.0	3 (1.5)		508 (10)	946 (17)
4.1-6.0	11 (5.6)	2006 (48) (4.1-10.0)	1371 (27)	1994 (35)
6.1-8.0	18 (9.1)		1778 (35) (6.1-10.0)	1093 (19)
8.1-10.0	14 (7.1)			578 (10)
10.1-20.0	46 (23.2)	856 (21)	1067 (21) (> 10)	667 (12) (> 10)
> 20	102 (51.5)	328 (8)		
Gleason score, n (%)				
< 5	4 (2)	222 (5)	31 (0.6)	n/a
5	6 (3)	688 (17)	4012 (79) (5-6)	4402 (77) (5-6)
6	62 (31.3)	2095 (51)		
3 + 4 = 7	46 (23.2)	906 (22) (both 3+4 and 4+3)	660 (13)	816 (14)
4 + 3 = 7	25 (12.6)		224 (4.4)	348 (6)
> 7	55 (27.8)	222 (5)	152 (3)	164 (3)
Clinical stage, n (%)				
T1a-b	2 (1)	223 (4)	n/a	n/a
T1c	52 (26.3)	1358 (33)	3200 (63)	4419 (77)
T2a	32 (16.2)	1186 (29)	1168 (23)	998 (17)
T2b	54 (27.3)	852 (21)	559 (11)	279 (5)
T2c	50 (25.3)	398 (10)	152 (3)	34 (1)
T3a	8 (4)	116 (3)	n/a	n/a
Pathologic stage, n (%)				
OCD	116 (58.6)	1957 (48)	3251 (64)	4204 (73)
LAD	62 (31.3)	1964 (47)	1727 (34)	1456 (25)
ECP	42 (21.2)	1661 (40)	1524 (30)	1276 (22)
SVI	20 (10.1)	303 (7)	203 (4)	180 (3)
LNI	20 (10.1)	212 (5)	101 (2)	70 (1)

PSA = prostate-specific antigen; OCD = organ-confined disease; LAD = locally advanced disease; ECP = extracapsular extension; SVI = seminal vesicle invasion; LNI = lymph node involvement

to predict the pathologic stage and thus identify the patients most likely to benefit from the RP. This poses a problem for urologists, as each of the three traditional prostate cancer diagnostic parameters (DRE, PSA level, and needle biopsy Gleason score) pose distinct dilemmas: intra- and inter-observer variability of DRE;[17] low specificity and sensitivity of PSA, which also has inconsistency over time and poor discrimination of indolent cancers from aggressive cancers;[18] unreliable of needle biopsy Gleason scores which frequently differ from the true Gleason score assigned at RP.[19] With such variation in diagnostic parameters, the formulation of reliable and transferable predictive models for prostate cancer prognosis is immensely challenging.

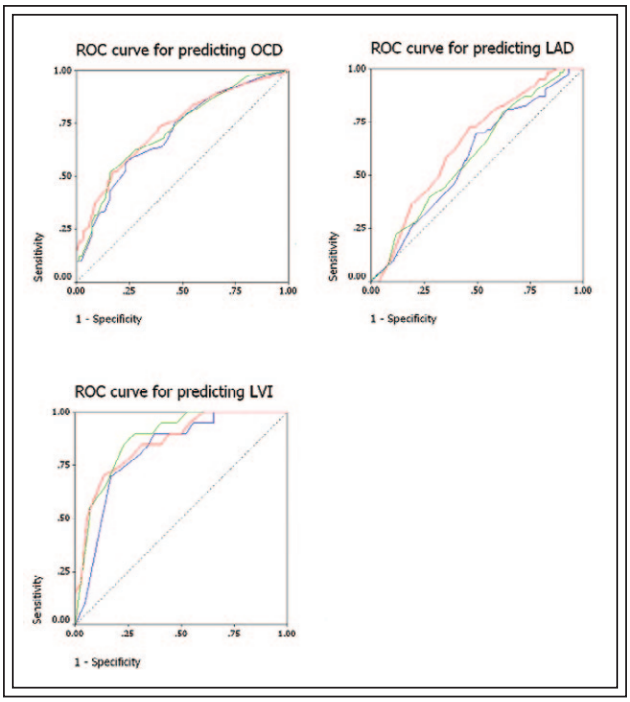

Figure 1. Receiver operating characteristic (ROC) curves of the 1997 (red), 2001 (green) and 2007 (blue) Partin tables for predicting organ-confined disease (OCD, upper left), localized advanced disease (LAD, upper right), lymph node involvement (LNI, lower), respectively.

Many predictive models have been published to predict the pathologic outcomes of patients with clinically localized prostate cancer. These predictive tools aid urologists, medical oncologists and radiation oncologists in clinical decision-making. They can also help counsel individuals in choosing the most appropriate management option. Among these predictive tools, the most widely used is the Partin tables. Although prostate cancer is not a common malignancy in China, several reports have shown a trend toward an increasing incidence of prostate cancer,[20] and the wide spread of the PSA test, which promotes the early diagnosis of prostate cancer, has resulted in RP being performed more widely in China than before.[21] Therefore, Chinese urologists urgently need to adopt a predictive pathologic outcome tool in their clinical practice, such as the Partin tables.

The Partin tables has gained acceptance as a useful guide in clinical practice in the US and other countries since 1997. Even so, there is a real need for thorough population-specific external validation, prior to adoption into routine clinical practice in varying geographical locations. In the PSA era, more prostate cancer are detected with a lower serum PSA level and biopsy Gleason score, especially in the US and European countries. In accordance with the tremendous change, the 1997 Partin tables were updated in 2001 and 2007 with narrower intervals when PSA < 10 ng/mL, expecting to provide a more accurate prediction of the pathologic stage. The new tables have also redesigned Gleason score categories to accommodate the prognostic difference based on observations suggesting different clinical behavior in men with biopsy Gleason sum 7 due to the predominance of pattern 4 or 3. Whether these newer version are more applicable than the older version when applied in Chinese clinical practice remains unknown.

In the presented study, we performed a head to head comparison of the three versions of the Partin tables. Generally, all three versions showed a good accuracy to predict OCD and LNI, but their predictive accuracy for LAD (including ECE and SVI) was only moderate or poor. However, the predictive accuracies of the 1997 version for almost all categories were still superior to those of the versions from 2001 and 2007,

TABLE 2. **Areas under the curve (AUC)**

Versions of the Partin tables	Pathological stage variables AUC (95% confidence interval (CI)) OCD	LAD	LNI
1997	0.732 (p < 0.001)	0.647 (p = 0.001)	0.856 (p < 0.001)
	0.663-0.802	0.569-0.726	0.775-0.937
2001	0.722 (p < 0.001)	0.594 (p = 0.036)	0.872 (p < 0.001)
	0.649-0.795	0.511-0.677	0.792-0.952
2007	0.695 (p < 0.001)	0.577 (p = 0.090)	0.829 (p < 0.001)
	0.618-0.772	0.492-0.661	0.733-0.925

OCD = organ-confined disease; LAD = locally advanced disease, including extracapsular extension and seminal vesicle invasion; LNI = lymph node involvement

Table 2. Interestingly, the AUC of almost all categories even decreased slightly from versions 1997 and 2001 through to version 2007, Table 2. Thus, we could not observe an increase in the predictive accuracy by the 2001 and 2007 modified versions. Based on the results of our test we also identified a lowest accuracy for the 2007 version of the Partin tables.

There should be some reasons contributing to the decreased accuracy of the modified Partin tables. Population characteristics and the basic structural changes of the 2001 and 2007 Partin tables might reduce their predictive accuracy for the present cohort.

Population characteristics among the three versions of the Partin cohort varied greatly. Similarly higher clinical stages were seen in the 1997 Partin cohort and the present Chinese cohort compared with both to the 2001 and 2007 Partin cohort. For example, in the present cohort only 26.3% of men presented with T1c tumor and up to 68.8% presented with T2 tumor, while 33% and 60%, respectively, in the 1997 Partin cohort. But a significantly greater ratio of 63% and 77%, who presented with T1c tumor, were seen in the training set of the 2001 and 2007 tables. Differing population characteristics between the validation cohort and the development cohort are quite common, and to a certain extent, these differences could be compensated for by the transportability of a well-performed predictive tool. However, as the differences show a trend toward more significance, the accuracy of a predictive tool will be compromised, even to the point of no longer qualifying for clinical practice.[9] Compared to the 2007 Partin cohort, the population characteristics of our validation study were so different that it might partly account for the lower accuracy of the 2007 Partin tables.

During the past decades, a tremendous prostate cancer stage migration has occurred in western countries. In accordance with this change, Partin et al revised their former nomogram with a more contemporary cohort in 2001[4] and 2007.[10] The PSA category spanning values 0 to 10 ng/mL were divided into four and five intervals in the 2001 and 2007 tables, respectively, instead of two in the 1997 tables. In China, the PSA screening test has not been widely used and most prostate cancer cases were revealed by urinary symptoms (75.9%) or bone pain (12.8%).[21] So it is not surprising that more than 50% of the present Chinese cohort had a high level of PSA (more than 20 ng/mL). Narrower divisions in pretreatment PSA related to a theoretical advantage to yield an appreciable gain in predictive accuracy, which would have been expected from decreasing the restriction of this continuous variable. However, the new tables put those patients with a PSA of > 10 ng/mL into the same group. This might be more suitable to a validation setting with a low PSA level. Nonetheless, when applied to the present cohort, in which nearly 75% of the patients had a PSA level of > 10 ng/mL, the theoretical advantage related to narrower divisions in the pretreatment PSA level below 10 ng/mL has not yielded an appreciable gain in predictive accuracy.

Decreased restriction of pretreatment PSA was paralleled by a modification of biopsy Gleason score coding in the 2001 and 2007 version of the tables: the group with a Gleason sum 7 was separated into two groups, with Gleason score of 3 + 4 and 4 + 3, respectively. Previous studies have showed that Gleason score 7 tumors in radical prostatectomy specimens were heterogeneous in their biologic behavior and a predominance of pattern 4 in Gleason 7 adversely affected prognosis.[22] But the agreement between biopsy and pathological specimen grades may be limited.[23] Under this premise the identification of the pathological predominance of Gleason pattern 3 may be more difficult if needle biopsy specimens are used. So the potential benefit of the modified coding of Gleason grade in the 2001 and 2007 tables failed to be associated with appreciably higher predictive accuracy in present cohort.

Our study also has some limitations. Although the patients receiving RP in our center may be the most in China during recent 5 years, the sample size of this study was relatively small. Nonetheless, considering the relatively low incidence of prostate cancer in China, these patients consecutively enrolled by the present cohort could still serve as a representative sample for validation. Another limitation of our study was that it was a single center study. But the use of a single center meant that all surgery was performed exclusively by one experienced urologist and central pathologic review could be easily achieved.

Overall, Partin tables showed a good accuracy to predict OCD and LNI. However, the predictive accuracy for LAD was more limited, while the newer versions of the Partin tables could not exceed the version of 1997 in their predictive accuracy for almost all pathological stages. Our results suggest caution when using newly introduced predictive tools that are not supported by several predictive accuracy tests. There is also a clear need for population-specific predictive tools in China. □

References

1. Partin AW, Kattan MW, Subong EN et al. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. *JAMA* 1997;277(18):1445-1451.

2. Blute ML, Bergstralh EJ, Partin AW et al. Validation of Partin tables for predicting pathological stage of clinically localized prostate cancer. *J Urol* 2000;164(5):1591-1595.
3. Graefen M, Augustin H, Karakiewicz PI 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(1):6-10; discussion 11.
4. 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.
5. Ayyathurai R, Ananthakrishnan K, Rajasundaram R, Knight RJ, Toussi H, Srinivasan V. Predictive ability of partin tables 2001 in a Welsh population. *Urol Int* 2006;76(3):217-222.
6. Augustin H, Eggert T, Wenske S 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(1):177-181.
7. Eskicorapci SY, Karabulut E, Turkeri L et al. Validation of 2001 Partin tables in Turkey: a multicenter study. *Eur Urol* 2005;47(2):185-189.
8. Naito S, Kuroiwa K, Kinukawa N et al. Validation of Partin tables and development of a preoperative nomogram for Japanese patients with clinically localized prostate cancer using 2005 International Society of Urological Pathology consensus on Gleason grading: data from the Clinicopathological Research Group for Localized Prostate Cancer. *J Urol* 2008;180(3):904-909; discussion 909-910.
9. Gao X, Ren S, Lu X, Xu C, Sun Y. The newer the better? Comparison of the 1997 and 2001 partin tables for pathologic stage prediction of prostate cancer in China. *Urology* 2008;72(5):1096-1101.
10. Makarov DV, Trock BJ, Humphreys EB et al. Updated nomogram to predict pathologic stage of prostate cancer given prostate-specific antigen level, clinical stage, and biopsy Gleason score (Partin tables) based on cases from 2000 to 2005. *Urology* 2007;69(6):1095-1101.
11. Yu JB, Makarov DV, Sharma R, Peschel RE, Partin AW, Gross CP. Validation of the partin nomogram for prostate cancer in a national sample. *J Urol* 2010;183(1):105-111.
12. Karakiewicz PI, Bhojani N, Capitanio U et al. External validation of the updated Partin tables in a cohort of North American men. *J Urol* 2008;180(3):898-902; discussion 902-893.
13. Bhojani N, Salomon L, Capitanio U et al. External validation of the updated partin tables in a cohort of French and Italian men. *Int J Radiat Oncol Biol Phys* 2009;73(2):347-352.
14. Bhojani N, Ahyai S, Graefen M et al. Partin Tables cannot accurately predict the pathological stage at radical prostatectomy. *Eur J Surg Oncol* 2009;35(2):123-128.
15. Fanning DM, Fan Y, Fitzpatrick JM, Watson RW. External validation of the 2007 and 2001 partin tables in irish prostate cancer patients. *Urol Int* 2010;84(2):174-179.
16. Augustin H, Isbarn H, Auprich M et al. Head to head comparison of three generations of Partin tables to predict final pathological stage in clinically localised prostate cancer. *Eur J Cancer* 2010;46(12):2235-2241.
17. Gosselaar C, Kranse R, Roobol MJ, Roemeling S, Schroder FH. The interobserver variability of digital rectal examination in a large randomized trial for the screening of prostate cancer. *Prostate* 2008;68(9):985-993.
18. Fitzpatrick JM, Banu E, Oudard S. Prostate-specific antigen kinetics in localized and advanced prostate cancer. *BJU Int* 2009;103(5):578-587.
19. King CR. Patterns of prostate cancer biopsy grading: trends and clinical implications. *Int J Cancer* 2000;90(6):305-311.
20. Sim HG, Cheng CW. Changing demography of prostate cancer in Asia. *Eur J Cancer* 2005;41(6):834-845.
21. Peyromaure M, Debre B, Mao K et al. Management of prostate cancer in China: a multicenter report of 6 institutions. *J Urol* 2005;174(5):1794-1797.
22. Chan TY, Partin AW, Walsh PC, Epstein JI. Prognostic significance of Gleason score 3+4 versus Gleason score 4+3 tumor at radical prostatectomy. *Urology* 2000;56(5):823-827.
23. Cookson MS, Fleshner NE, Soloway SM, Fair WR. Correlation between Gleason score of needle biopsy and radical prostatectomy specimen: accuracy and clinical implications. *J Urol* 1997;157(2):559-562.